7817916

STEFFEN, ANN ELIZABETH
FACILITY LOCATION PROBLEMS AMONG RECTANGULAR REGIONS.

THE UNIVERSITY OF OKLAHOMA, PH.D., 1978

University Microfilms International 300 N. ZEEB ROAD, ANN ARBOR, MI 48106

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

FACILITY LOCATION PROBLEMS AMONG RECTANGULAR REGIONS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

By

ANN ELIZABETH STEFFEN

Norman, Oklahoma

1978

FACILITY LOCATION PROBLEMS AMONG

RECTANGULAR REGIONS

APPROVED BY:

Adel A. Aly

Adel A. Aly, Chairman

Alan L. Dorris

Alan L. Dorris

B L Foote

Bobbie L. Foote

David C. Kay

David C. Kay

Hillel Kumin

Hillel J. Kumin

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

I would like to express my appreciation to all who have helped me through this research effort. I would especially like to thank my advisor, Dr. Adel Aly, for the time, guidance, encouragement, and expertise he provided throughout the course of my dissertation.

I am also grateful to the other members of my committee, Dr. Alan Dorris, Dr. Bob Foote, Dr. Hillel Kumin, and Dr. David Kay for their encouragement throughout my doctoral programs. Their suggestions on improving the technicalities and readability of this dissertation were invaluable.

I would also like to thank my office partner, Yoshi Ikeda, for his help and patience in overcoming some of the computer problems I encountered. Thanks also goes to John Marucheck for all the encouragement he gave me through my graduate career and for leaving town when he did so that I could hasten this research effort.

Finally, I would like to thank the National Science Foundation for their support throughout most of my graduate career.

ABSTRACT

The facility location problem among rectangular regions involves location problems in the urban setting. Here the existing facilities or customers are assumed to be uniformly distributed over rectangular regions. The problem is a special case of the probabilistic formulation or can be interpreted as a limiting case of the location problem in discrete space.

The sensitivity of this location problem with the rectilinear metric to a deterministic solution technique is discussed and contrasted to the relative insensitivities of the same problem with the Euclidean metric to the deterministic solution. Properties of the problem are developed, and algorithms are developed for both the single and multi-facility location problems with rectilinear distances. Computational experience is reported.

The location-allocation problem is then discussed in terms of its properties and its relation to the above facility location problem. The limitations of present solution techniques are discussed. A branch and bound solution technique is then developed. The algorithm is verified, and computational results are given. These results are used to compare the algorithm to other algorithms and to adapt the algorithm to other versions of the location-allocation problem.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Overview of Location Analysis

Location analysis is a class of problems that has intrigued the human mind for centuries. Although applications of location problems have been with man through all ages, the problem did not receive scholarly treatment until the seventeenth century.

Wesolowsky (1973) credits Fermat with proposing the first location problem--determining the location of a fourth point in relation to three given points in the plane so that the sum of the distances from that point to each of the three given points is minimized. This question was aptly named the Fermat problem. Toricelli was one of the first to solve this problem when he presented a geometrical solution in 1640.

In the early 1900's Weber (1929) generalized the Fermat problem with an economic significance by assuming that two of the given points were sites of raw materials; the third was a market; and that the new point represented a factory that was to be located. He also assumed the weights on each of the three given points were unequal; thus, the objective became the minimization of the sum of the weighted distances. Since then, the basic location model has been referred to as the generalized Weber problem.

During the last twenty years, location analysis has received a significant amount of attention in the literature. This has been due to the use of high-speed computers which made solution of intractable location problems possible and a growing awareness of the application of the location problem to many different fields. Researchers in the areas of industrial engineering, operations research, transportation, management science, mathematics, computer science, geography, regional planning, urban development, economics, marketing, and political science have all studied the location problem, although their approach to the problem will be tempered by their particular discipline.

One result of this explosion of literature has been that certain variations of the location problem received most of the attention while other variations were left virtually untreated. For example, up until several years ago, most of the work considered location among deterministic points. Probabilistic formulations and location among areas instead of points had not been studied. On the other hand, in considering new problems there is always a danger of merely reworking existing problems or obtaining a result not significantly different from current methods. Thus, anytime a new solution is proposed it should be superior in either the quality of the solution or computational time required to existing solution techniques.

The location of new facilities among rectangular regions is one area that has received very little attention. In many different settings it is proper to treat the customers to be served as a region, such as in the location of a public service facility to serve a neighborhood or a densely populated urban area. Another example would be the case of locating an emergency services facility relative to an urban area where an incident could arise anywhere in the area with a certain probability distribution. Finally, even when the customers or users of the new facility are discrete, the number of customers may become so large that it may be infeasible in terms of data collection and computational efficiency to consider them discrete. In this situation it may be necessary to make a regional assumption in order to solve the problem.

The traditional solution method has been to represent the regions by their centroids and to solve the problem using deterministic methods. However, the validity of this assumption has not been examined fully. One of the purposes of this research will be to examine cases where the assumption is appropriate, and where it is inappropriate. In the cases where the assumption is not viable, new solution techniques will be developed, tested, and compared to existing techniques.

1.2 Location Problems to be Considered

Two types of location problems will be considered in this research. Both are variations on the generalized Weber problem. The problem involves the location of one or more points relative to several given points in order to minimize the sum of the weighted distances among the new points and the old points.

P1.2.1 minimize $\sum_{j=1}^{n} \sum_{i=1}^{m} w_{ji} \quad |X_j - P_i|_{\ell_p} + \sum_{1 \le j < k \le n} v_{jk} |X_j - X_k|_{\ell_p}$

where: n = number of new facilities to be located.

m = number of existing facilities.

$P_i = (a_i, b_i)$ coordinate location of existing facility i.

$X_j = (x_j, y_j)$ coordinate location of new facility j.

ℓ_p = the type of norm used.

$|X_j - P_i|_{\ell_p}$ = distance between the locations of new facility j and existing facility i.

$|X_j - X_k|_{\ell_p}$ = distance between the locations of new facilities j and k.

v_{jk} = cost per unit time per unit distance between new facilities j and k.

w_{ji} = cost per unit time per unit distance between existing facility i and new facility j.

Here the X_j's are the decision variables, and all other terms are known parameters. A complete taxonomy of classifying location problems on the basis of the parameters can be found in Francis and White (1974).

Notice that the distance metric is expressed as an ℓ_p norm. The ℓ_p norm is defined as follows:

$$|X_j - P_i|_{\ell_p} = (|x_j - a_i|^p + |y_j - b_i|^p)^{1/p}.$$

Thus, when p = 1, the distance metric becomes rectilinear, rectangular, or Manhattan distance.

When p = 2, the distance metric becomes Euclidean or straight-line distance. Finally, if p = ∞, the distance metric becomes the less frequently used Chebyshev distance.

One variation on the generalized Weber problem to be considered is the n-center problem among rectangular regions. It is basically the same as the generalized Weber problem with existing facilities spread over rectangular regions with a particular probability distribution.

P1.2.2 minimize $\sum_{j=1}^{n} \sum_{i=1}^{m} \iint_{R_i} w_{ji} \, |X_j - R_i|_{\ell_p} \, \theta(R_i) dR_i$

$$+ \sum_{1 \le j < k \le n} v_{jk} |X_j - X_k|_{\ell_p}$$

All variables are the same as in the generalized Weber problem except:

$R_i = [a_{i_1}, b_{i_1}] \times [a_{i_2}, b_{i_2}]$ coordinate locations encompassed by region.

$\theta(R_i)$ = the joint probability density function of (a_i, b_i) defined on region R_i.

A further generalization of the Weber problem is the location-allocation (L-A) problem. This problem assumes that the allocation scheme of existing facilities to new facilities is not known a priori. Thus, the v_{jk}'s are assumed to be zero. The solution of the problem is twofold. First, the optimal allocation scheme must be determined and secondly, optimal locations for the new facilities must be determined.

P1.2.3 minimize $\sum_{j=1}^{n} \sum_{i=1}^{m} \iint_{R_i} z_{ij} w_{ij} |X_j - R_i|_{\ell_p} \theta(R_i) dR_i$

subject to: $\sum_{j=1}^{n} z_{ij} = 1 \qquad i = 1, \ldots, m$

$z_{ij} = 0, 1$ for all i and j

where: $z_{ij} = \begin{cases} 1, & \text{if existing facility } i \text{ is allocated to new facility } j \\ 0, & \text{otherwise} \end{cases}$

Notice the implicit assumption that all new facilities have an infinite capacity to serve the existing facilities. If capacity restrictions were made, the problem would become a transportation problem.

The term "facility" has been used in its loose sense. The existing facility may represent the locations of people in need of a service or people that will be potential customers. It may represent the location of a machine that needs to be maintained. Similarly, the "new facility" may assume the same realizations that the "existing facility" does. It is not necessary for the new facility to maintain a relationship of purely service to the existing facilities either. For example, a new warehouse may have to be located relative to both the plants that supply it and the other warehouses it may interact with. Thus, the terms "facility" and "serve" or "interact" should be considered in their broadest senses.

1.3 Application of the Research

With the areas of possible application for location research expanding rapidly, an application exists for this research anytime a service or good is distributed among a group of consumers. Applications would range from the problem of a chain of grocery stores seeking to locate several stores in a city to locating power-generating stations in an area containing both urban and rural regions with different types of needs for electricity.

A specific example of a possible application of this research was presented by Abernathy and Hershey (1972). Several health clinics were to be located within a medical-service area containing a population of 50,000. There were three small cities in the area whose total population was 39,000 and surrounding rural areas comprising the balance of the population.

The planner was interested in maximizing the utilization of the clinics and realized that distance from the health facility was the most important factor in utilization from a set of variables that also included demographic, socioeconomic, and health status factors.

Clearly, a centroid approach would fail to account for the different population densities in the urban and rural areas. Hence, the entire area was decomposed into subareas. Each subarea was given a utilization weight based upon the characteristics of the inhabitants. The weight was used to determine the likelihood of a visit to the clinic from someone living in the area. Thus, the criterion became to allocate subareas to clinics and subsequently locate the new clinics so that the distance per visit to a clinic is minimized.

Numerous other examples could be found in both the private and public sectors. Revelle et al. (1970) discuss such examples and their analyses relative to location models.

1.4 Scope and Limitations

The research effort will concentrate on both the multifacility location problem and the location-allocation problem with respect to rectangular regions. The tradeoffs between using the regions or an "areal" approach versus using their centroids and a deterministic approach will be analyzed. In the cases where the regional approach is warranted, properties and techniques for solution will be developed.

The probabilistic aspects of the areal approach will also be considered using the expected value criterion. Although several distributions will be discussed, the uniform distribution will be emphasized for its practicality.

The computational aspects of all solution procedures are emphasized and will be compared to existing results.

1.5 Order of Presentation

Because the literature dealing with the generalized Weber problem is extremely voluminous, the review of the literature in Chapter II will be a brief discussion of the basic types of approaches to location analysis. Each subsequent chapter will then contain a more detailed description of the works relating to the subject of that chapter.

Chapter III will treat the multifacility location problem among rectangular regions. A model will be formulated and both the Euclidean and rectilinear metrics will be discussed. A heuristic algorithm will be developed and tested.

Chapter IV will include a development of the location-allocation model for rectangular regions. The emphasis will be on the model with a rectilinear metric. It will be shown that this development has an impact for both the corresponding deterministic problem and the problem with Euclidean metric on the basis of extensive computational results. Chapter V will summarize the research effort and include recommendations for further research.

CHAPTER II

STATE OF THE ART

2.1 General Treatments of Location Theory

The literature dealing with location analysis is so extensive that a complete review of all works is impossible. Instead, a general overview of solution techniques will be presented here. Discussions of work relating to the research effort will be deferred to subsequent chapters.

An extensive bibliography of over two hundred works in location theory was compiled by Francis and Goldstein (1974). The reader is also advised to consult Francis and White (1974) for a thorough treatment of the state of the art regarding the formulation and solution of various solution methods.

Other extensive treatments include Elshafei and Haley (1974) who provided a complete literature search in addition to the formulations and solution aspects of different location models. ReVelle et al. (1970) presented a brief overview of the different types of location problems.

In the area of location-allocation problems Scott (1970) presented a review of the different formulations of the general location-allocation model. Lea (1973) compiled an annotated bibliography of all papers relating to location-allocation systems.

In light of these extensive treatments, it should be kept in mind that the location system may be conceptualized in two ways:

1) Location on a network

2) Location on a plane

Location on a network assumes that the existing facilities are nodes and that the new facility(s) may be located anywhere on the network--either on a node or on a link between nodes--in order to minimize the arc flow in the network. Although there is a large body of literature that considers location on a network, this type of problem won't be considered in this research effort.

The formulations in Chapter I are examples of location in the plane. The new facility(s) to be located are represented as points in the plane, and similarly, the existing facilities are represented as points or areas on the plane. Even this broad category may be further subdivided in the following manner:

1) Finite solution space

2) Infinite solution space

The finite solution space means, as its name implies, that the new location site(s) must be chosen from a finite set of points in the plane. This corresponds to the situation where one or more sites must be chosen from a set of alternative feasible sites. The problem is usually referred to in the literature as the "plant location problem." There has been a large amount of work devoted to this latter problem.

On the other hand, problems with the infinite solution space have not been resolved quite as well. The infinite solution space corresponds to the situations when either no set of feasible solution sites has been chosen or the number of feasible solution sites becomes large. This is the focus of the work in this effort and all discussions of the literature will be restricted to this theme.

2.2 Solution Techniques

In this section an overview of solution techniques for the location problem with infinite solution space will be considered. It will become obvious that certain techniques are superior to others; yet the superior techniques are the culmination of the years of work on the location problem which produced the other techniques. The organization of the techniques will be similar to that of Eilon et al. (1971).

2.2.1 Analog Techniques

Analog techniques may be used to solve the location problem by considering the forces at work in the system. One of the simplest analog models is the "string-and-weights" approach to location. Here the existing facilities are represented as weights proportional to the interaction term between the new and the existing facility. String is used to represent distances. Kuhn and Kuenne (1962) discuss several methods used during the early twentieth century to locate new facilities based on a physical representation of the forces. Such analog systems served as a visual representation of variously weighted existing facilities, where the more heavily weighted ones tended to pull the site of the new facility toward them.

Hitchings (1967) built an electrical analog to the same situation. Here the weights of different existing facilities were represented by either different materials with different resistivities or by different cross-sectional areas of materials with similar resistivity. Distance was represented by the length of a particular resistor.

A plotter was used in conjunction with the analog to plot the contour lines or "isocost" lines of location sites with the same weighted distance cost. Francis (1963) also described a method of drawing contour lines for a location system based on a weighted geometrical model. Although using contour lines is not the most efficient solution method, its value lies in the representation of the total surface of the cost function.

2.2.2 Simulation

Simulation is probably the best approach to solving the location system when heuristic or optimal algorithms are hard to solve. Massam (1974) used simulation to solve a facilities location problem with existing facilities uniformly distributed over a rectangular region. Similarly, simulation could be used for irregularly shaped regions or other distributions. Another advantage to the simulation approach is the feeling for the shape of the cost function produced without explicitly having to draw contour lines.

2.2.3 Analytic Methods

Analytic methods are appealing in the respect that they sometimes provide optimal numerical solutions. Needless to say, most of the literature seems to address itself toward this end. It is noteworthy that the first successful developments of analytic methods involved gradient descent approaches. The gradient descent method has been widely used in attempts to develop solutions for some of the then unsolved problems. Analytic techniques for problems considered in this research effort will be discussed in the relevant chapters. Here the solutions to the single facility location problem will be presented.

The single facility location model is

P2.2.1 minimize $\sum_{i=1}^{m} w_i \, |X-A_i|_{\ell_p}$

where: $X = (x,y)$ the coordinate location of the new facility

$A_i = (a_i, b_i)$ the coordinate location of existing facility i

w_i = cost per unit distance to serve facility i

m = number of existing facilities.

When p = 1, Francis (1963) showed through gradient reduction that the optimum is found by satisfying a median condition.

The solution of p = 2 was developed independently by Cooper (1963) and Kuhn and Kuenne (1962). It was an iterative procedure based on a modified gradient condition.

Other innovative techniques used in solving location problems will be discussed later.

2.2.4 Heuristic Methods

When a location problem involves combinatorials, an analytic solution may be possible, but the number of combinations to be considered is so large that the problem becomes computationally infeasible. In such a case a "good" or near-optimal solution is all that can be hoped for. A heuristic never guarantees optimality but may guarantee a solution within a specified percentage of optimal. Sometimes a heuristic can guarantee only a local optimum.

Most location problems employing heuristic solutions involve the making of allocations or assignments. One example of such a problem is the previously formulated location-allocation model. Cooper (1963, 1964 and 1967) has developed several heuristic algorithms to solve the model. These will be discussed in Chapter IV.

Another example is the partial covering problem which is related to the location-allocation problem. It is formulated as:

$$\text{P2.2.2} \quad \text{minimize } Z = \sum_{i=1}^{m} \min_{j} \left\{ z_j w_i \; |X_j - A_i|_{\ell_p} \right\}$$

$$\text{subject to:} \quad \sum_{j=1}^{n} z_j \leq k$$

$$z_j = 0,1$$

where: m = number of existing facilities

n = number of sites for new facilities

k = maximum number of new facilities allowed

w_i = weight on existing facility i

X_j = coordinate location of new facility j

A_i = coordinate location of existing facility i

$$z_j = \begin{cases} 1, & \text{if a facility is located on site } j \\ 0, & \text{otherwise} \end{cases}$$

White and Case (1974) present a treatment of heuristic solutions to this problem. Note this problem corresponds to the location-allocation problem in discrete space by its attempt to assign existing facilities to new facilities.

2.3 Probabilistic Location Problems

Since there are probabilistic aspects to the location problems considered, a brief review of the literature concerning probabilistic formulations will be useful. The works to be reviewed consider either the existing facilities or the interaction weights (w_i) to be random variables. Most of the works have focused on the normal distribution. Works dealing with the uniform distribution will be discussed in the next two chapters.

Katz and Cooper (1974) considered the single facility location problem with random existing destinations and a Euclidean norm under a minimization of expected value criterion. They developed an iterative descent alogrithm to obtain the global minimum and proved that order of convergence of the algorithm is linear. The results were then applied to the bivariate normal density function.

Katz and Cooper (1975) continued their work to develop sufficient conditions to assure that the location of the new facility will be within the convex hull of the means of the existing facilities. The exponential and the symmetric exponential distributions were also considered as further examples of their iterative algorithm. Cooper (1974) presented computational experience for the results of the first paper.

Seppäla (1975) discussed the various objective function criteria relevant to probabilistic formulations. He developed a location problem where the interaction weights are normally distributed and used the Euclidean metric. The problem is defined as follows:

P2.3.1 minimize Z

$$\text{subject to: } P_r \{ \sum_{j=1}^{n} \sum_{i=1}^{m} w_{ij} |X_j - A_i|_{\ell_2} + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} v_{jk} |X_j - X_k|_{\ell_2} \leq Z \} \geq \alpha$$

where: α = a specified probability level

Z = decision variable, and

w_{ij}, X_j, A_i, and v_{jk} are as defined before.

This problem was solved with chance-constrained programming methods. The chance constraints were converted into deterministic equivalents by introducing new variables. Linear approximations were substituted for the nonlinear expressions, and the new linearized problem was solved as a linear programming problem.

Wesolowsky (1977) considered the single facility location problem where the existing facilities were deterministic and collinear, but the interaction weights had a multivariate normal distribution. Also, the rectilinear metric was used. He verified that the solution could be found by an application of the median condition on the means of the interaction weights.

This particular work was preceded by Aly (1975) who considered both random existing facilities and interaction weights. He considered both the Euclidean norm and the rectilinear norm, which had received little attention up to that point. Both single facility and multifacility location problems were considered with the expected value criterion. Both unconstrained and constrained versions of the problem were considered. The constraints consisted of upper bounds on the weighted distance costs and chance constraints representing an aspired service level. The multivariate normal distribution was generally used in the treatment of the random variables.

Wesolowsky (1977) considered the different distributions a random existing facility may assume when the metric is rectilinear. He discussed the bivariate normal, bivariate uniform, and bivariate symmetric exponential distributions. He developed solution techniques that were rudimentary in light of previously developed techniques.

Thus, this review illustrates, not surprisingly, that normally distributed random variables are assumed in the consideration of probabilistic location problem formulations. The expected value criterion is nearly universally assumed as the objective function criterion, and until Aly's work the norm was always considered to be Euclidean.

CHAPTER III

FACILITY LOCATION PROBLEMS AMONG RECTANGULAR REGIONS

3.1 Introduction

In Chapter I the deterministic version of the generalized Weber problem was presented. It was noted in Chapter II that single facility version of this problem has received considerable attention in the literature. In this chapter the different works devoted to the multifacility versions of the problems will also be reviewed.

This chapter will consider a version of the problem which has received relatively little attention--that is, locating new facilities among rectangular regions. There are two different interpretations that may be made from this problem. The first is that the existing facilities are points, and each point is a random variable with a uniform distribution over some rectangular area. The second interpretation is that the existing facilities are areas or rectangular regions. These areas may be deterministic or probabilistic. The deterministic case will receive most of the emphasis because its solution will be the same as the situation posed by the first interpretation.

3.2 Problem Formulation

The generalized formulation to be considered in this chapter is as follows.

P3.2.1 minimize $$\sum_{j=1}^{n} \sum_{i=1}^{m} \iint_{R_i} w_{ji} |X_j - R_i|_{\ell_p} \theta(R_i) dR_i + \sum_{1 \le j < k \le n} |X_j - X_k|_{\ell_p}$$

where: n = number of new facilities

m = number of existing facilities

$X_j = (x_j, y_j)$ coordinate location of new facility j

$R_i = [a_{i_1}, b_{i_1}] \times [a_{i_2}, b_{i_2}]$ coordinate locations encompassed by the existing facility--region i. In general, a particular coordinate location is (a_i, b_i).

$\theta(R_i)$ = the joint probability density function of (a_i, b_i) defined on region i.

w_{ji} = cost per unit distance per unit time between region i and new facility j.

v_{jk} = cost per unit distance per unit time between new facilities j and k.

When $n = 1$, this is the single facility location problem. Otherwise, it is a multifacility location problem.

Now consider the case of many existing facilities or "customers" distributed over rectangular regions. The objective is to minimize the sum of the distances each customer must travel, but since the number of existing facilities is large, it would be infeasible to consider a deterministic point model. Instead the cost of serving each of these customers is to be considered.

P3.2.2 minimize $\sum_{j=1}^{n} \sum_{i=1}^{m} \int_{a_i}\int_{b_i} w_{ji}|X_j-R_i|_{\ell_p}\, da_i db_i$

$$+ \sum_{1 \le j < k \le n} |X_j - X_k|_{\ell_p}$$

where: n = number of new facilities

m = number of regions where existing facilities are located

X_j = (x_j, y_j) coordinate location of new facility j

(a_i, b_i)= general coordinate representation for any existing facility in region i

R_i = $[a_{i_1}, b_{i_1}] \times [a_{i_2}, b_{i_2}]$ rectangular region i

w_{ji} = cost per unit distance per unit time between region i and new facility j

v_{jk} = cost per unit distance per unit time between new facilities j and k.

An example of this model is the location of libraries or other public service facilities in a densely-populated urban area in order to minimize the sum of the distances its patrons must travel to use it.

The probabilistic version of this problem would be to consider the expected distance a customer in region i must travel. In this case, the customers will be considered to be uniformly distributed over the region.

P3.2.3 minimize $\sum_{j=1}^{n} \sum_{i=1}^{m} \int_{a_i}\int_{b_i} w_{ji} |X_j - R_i|_{\ell_p} f(a_i) f(b_i)\, da_i db_i$

$$+ \sum_{1 \le j < k \le n} v_{jk} |X_j - X_k|_{\ell_p}$$

where: $f(a_i) = \frac{1}{a_{i_2} - a_{i_1}}$

$f(b_i) = \frac{1}{b_{i_2} - b_{i_1}}$

and all other parameters are as defined in P3.2.2.

Note that $f(a_i) \cdot f(b_i) = \frac{1}{A_i}$

where A_i is the area of rectangular region i.

This term would be a constant in relation to the integration taking place.

3.3 Related Work

This section will treat the development of the location problem with rectangular regions. Since the multi-facility location problem will be stressed, a discussion of works dealing with that problem will be provided.

3.3.1 Work Relating to Deterministic Formulations

Since the number of works relating to location among areas is few, the literature concerning the deterministic multifacility problem will be briefly reviewed. The solution methods reviewed will serve as background to those that will be later developed.

There are two papers devoted to developing necessary and sufficient conditions for optimal solutions to location problems. Francis and Cabot (1972) considered the multi-facility problem for Euclidean distances. They explored properties of the objective function and its solution in addition to developing the properties of the dual. Wendell and Hurter (1973) developed similar properties for a location problem with general distance metrics. Kay, et al. (1978) extended these properties to the non-Euclidean space.

Surprisingly, there has been little work dealing with only the Euclidean distance metric although solution techniques are available. Eyster et al. (1973) developed an iterative procedure to solve both the rectilinear and the Euclidean distance multifacility location problem by using hyperboloids to approximate the cones formed by the original objective functions. This procedure was necessary since the partial derivatives of the objective function are not defined at all points in the solution space. Love and Morris (1975) used the same idea of using a hyperbolic distance function to approximate nondifferentiable objective functions under general distance metrics. Convex programming methods were then used to solve the problem. Ostresh (1977) observed that if an iterative procedure were used in solving the multifacility problem with a Euclidean metric that the convergence should be linear. He couldn't use this result to improve any of the existing solution techniques.

More solution techniques have been developed for the rectilinear distance problem. The hyperbolic approximation procedures described above could be used since they were developed for general distance metrics. Wesolowsky and Love (1972) applied this procedure.

Pritsker and Ghare (1970) developed a method of locating new facilities by considering movements of each of the new facilities relative to existing facilities such that those movements show a potential for improving the objective function. Rao (1973) showed that this method produced an optimal solution only when no two new facilities had the same location. He suggested that a dual approach would be more efficient in general. However, Pritsker (1973) later corrected the algorithm in light of Rao's counter-example.

Dual results have been very useful in developing new solution techniques for problems with the rectilinear metric. Cabot et al. (1970) decomposed the problem into two subproblems--each of which could be expressed as a linear programming problem. The dual of each linear programming problem was formulated as a minimal cost network flow problem and solved using the out-of-kilter algorithm.

Wesolowsky and Love (1971) used the same approach of solving the dual, but they used linear programming techniques. Morris (1975) essentially used the same technique, but he extended it to cover the constrained problem also.

3.3.2 Works Relating to Rectangular Regions

There have been only two papers done in this area. Love (1972) developed a solution to the single facility location problem among rectangular regions with a Euclidean distance metric. The solution required the evaluation of a complex expression involving seven initial substitutions, four logarithms, and eight integrations. He was able to extend this result to constrained location problems by using SUMT.

Wesolowsky and Love (1971) considered both the corresponding single and multifacility location problems with the rectilinear distance metric. A gradient reduction technique is used to solve the single facility problem. A solution for a two facility problem was developed by examining the entire projection of the hyperplane formed by the objective function onto the two-dimensional plane.

3.4 The Centroid Approach

The motivation for the use of rectangular regions in location problems was the search for a better solution to that problem than the deterministic solution techniques could provide. As the literature illustrates, the "better" solution involved several tradeoffs. The costs involved in this tradeoff were:

1) the evaluation of complex expressions. (As Love (1972) indicates, the expressions that were to be evaluated were far more complicated than anything that had been encountered in the deterministic case.)

2) increased computational time. (This increase in computational time is a function of both the complex expressions to be calculated described above and the number of rectangles to be considered. For example, in a real life situation an irregular shaped region may be approximated by decomposing it into rectangular regions, thus increasing the number of regions to be considered.)

The question is then "is the improved solution worth these costs?"

Bennett and Mirakhor (1974) addressed the work of Love (1972) with that intention. The centroid of each rectangular area was used to represent the region in each of the example problems. A deterministic solution was then obtained.

Aly (1975) applied his probabilistic formulations to deterministic problems worked by Francis and White (1974) in an attempt to make a comparison.

In Table 3.1 the various works will be compared. The deterministic formulations include the centroid approach. The probabilistic formulations include both normal and uniformly distributed variables.

Note that objective function values were not available in Problem 2. Love did give a value for his problem; Bennett and Mirakhor did not. The author was unable to replicate Love's value.

TABLE 3.1

A COMPARISON BETWEEN PROBABILISTIC AND DETERMINISTIC FORMULATIONS IN LOCATION PROBLEMS

Problem	Description	Source of Results	Optimal Location	Objective Function Value (Prob.)	% Deviation	
					Objective Function	Coordi-nate Location**
A.1*	Single facility Euclidean metric constrained	Love (prob.)	(9.15,2.47)	502.81	2%	5%
		Bennett & Mirakhor (deter.)	(9.25, 2.25)	513.23		
A.1*	Single facility Euclidean metric unconstrained	Love (prob.)	(12.4,3.39)	-	-	3.5%
		Bennett & Mirakhor (deter.)	(12.37,3.16)	-		
A.1*	Single facility rectilinear metric	(prob.)	(11.375,3.22)	72.96	6%	17.8%
		(deter.)	(12,2.25)	77.29		
A.2*	Multifacility Euclidean metric	Aly (prob.)	(4.65,5.12) (4.65,5.12)	84.4	5%	54%
		Francis & White (deter.)	(8,7) (8,7)	88.76		
A.2*	Multifacility rectilinear metric	Aly (prob.)	(6,6) (6,5)	136.9	17.6%	27%
		Francis & White (deter.)	(8,7) (8,7)	161.043		

*Found in Appendix A.

**The values are averaged over both coordinates.

Table 3.1 indicates that location problems with Euclidean distance metric are relatively insensitive to a relaxation of the probabilistic assumptions. In other words, using the centroid approach for probabilistic location problems with Euclidean distance metric amounts to using heuristic solution. In these problems, the centroid approach produced a solution within five per cent of optimal. Although there is no guaranteed percentage within optimum, empirical results seem to indicate that there is a large region around the optimum whose objective function values are near optimal.

In considering the deterministic Euclidean distance location problem, the same result holds true. Cooper (1967) was one of the first to notice from computational experience that surface of the objective function was relatively flat around the global optimum. Larson and Stevenson (1972) also discovered the insensitivities of optimal facility location in designing urban service systems. They discovered that for their problem an optimal location reduced the mean service time by 25 per cent over a random location.

In considering the probabilistic formulations of this problem, Massam (1974) discovered through simulation that the objective function value with an expected value criterion is very insensitive to locations or solutions in an area around the global optimum but is characterized by sharp ridges around this area. Seppäla (1974) also noted that under an expected value criterion the probabilistic location problem with Euclidean distance metric was not significantly different from the deterministic formulation.

Thus, when the Euclidean distance metric is used in a probabilistic location problem under an expected value criterion, it may be more efficient to solve the deterministic version of the problem by applying the centroid approach.

On the other hand, Table 3.1 indicates that the tradeoffs in considering the deterministic version of the rectilinear distance location problem are greater. This problem seems to be more sensitive to shifts from the optimal location. There has been no work to document the nature of the surface of the objective function with this metric. Consequently, in considering probabilistic formulations of this location problem it is necessary to develop solution techniques other than the deterministic methods.

3.5 The Single Facility Location Problem with Rectilinear Norm

The problem to be considered is the single facility location problem with expected value criterion.

P3.5.1 minimize $\sum_{i=1}^{m} \frac{w_i}{A_i} \iint_{R_i} \left(|x-a_i| + |y-b_i| \right) da_i db_i$

where: (x,y) = coordinate location of new facility

(a_i,b_i) = general coordinate of any point in region R_i

w_i = cost per unit distance per unit time between the new facility and region R_i

A_i = area of region R_i

Notice that if the A_i term is eliminated from the formulation this formulation corresponds to the single facility version of P3.2.2. That problem does not use the expected value criterion, but the solution technique will be essentially the same.

3.5.1 Assumptions

The assumptions for this model will be stated below. Parts of Assumption 2 are modifications of assumptions presented by Wesolowsky and Love (1971).

Assumption 1: The cost of interaction between the new facility and any of the regions may be deterministic or probabilistic. If it is probabilistic, it will be assumed that its expected value is known. Aly (1975) showed that under the expected value criterion, the expected weighted distance between a new and existing facility is:

$$E[w_i]E[|x-A_i|_{\ell_p}] \qquad (3.5.1)$$

which is the product of the expected value of the interaction term and the expected distance between the two. The implications for this model is that $E[w_i]$ may be substituted for w_i in P3.5.1 without changing either the formulation or solution technique. Thus, this model accommodates both probabilistic or deterministic interaction weights.

Assumption 2: The regions are all distinct and rectangular such that the following conditions hold:

i) None of the rectangular regions overlaps another.

ii) No region contains any barriers that would affect interaction with the new facility.

iii) The region should contain uniform density as far as interaction with the new facility is concerned.

This assumption states that any nonrectangular region must be decomposed into smaller rectangular areas. The decomposition of an irregularly shaped region is illustrated in Figure 3.1.

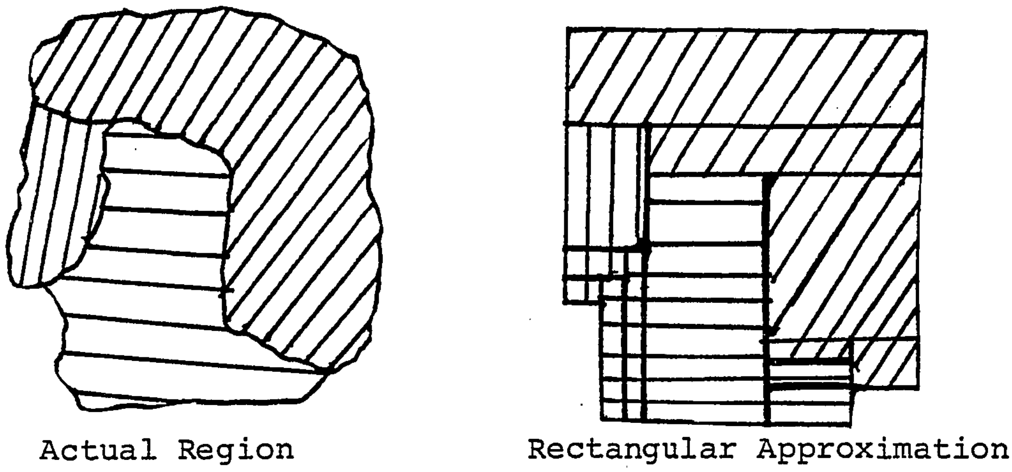


Figure 3.1. A rectangular approximation to an irregularly shaped region.

The first assumption states that the rectangular regions must not overlap. Any situation where this occurs can be rectified by a rectangular decomposition. For example, consider the situation given in Figure 3.2. Rectangle 1 is [2,5] x [1,4] with weight 2. Rectangle 2 is [4,8] x [1,4] with weight 3.

If we consider the problem without the expected value criterion, the overlaps would be removed in the following manner:

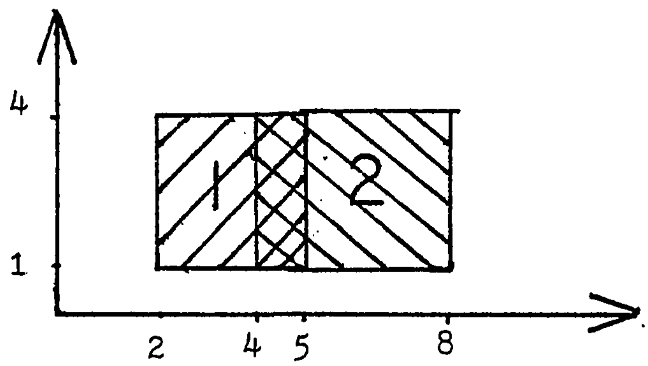


Figure 3.2. Two overlapping rectangles.

i) Decompose the area into nonoverlapping rectangles.

In this case we would have three decomposed rectangles.

1) $[2,4] \times [1,4]$

2) $[4,5] \times [1,4]$

3) $[5,8] \times [1,4]$

ii) Accumulate the weights from the old rectangles for new rectangle j by the following expression:

$$w_j' = \sum_{k \in I_j} w_k \tag{3.5.2}$$

where I_j is the set of indices of the original rectangles whose intersection with new rectangle j is not empty.

Here the weights are computed:

j=1. $I_1 = \{1\}$ $w_1' = 2$

j=2. $I_2 = \{1,2\}$ $w_2' = 2+3 = 5$

j=3. $I_3 = \{2\}$ $w_3' = 3$

This method of decomposition could also be used to rectify any situation violating condition 2. Discussions on removing overlaps will also be included as each of the algorithms is developed.

3.5.2 Development of a Solution Procedure

Notice that P3.5.1 can be decomposed into two subproblems. If $R_i = [a_{i_1}, a_{i_2}] \times [b_{i_1}, b_{i_2}]$, then one subproblem will be used to solve for x, and the other to solve for y. The subproblems are:

$$\text{P3.5.2} \quad \text{minimize } f(x) = \sum_{i=1}^{m} \frac{w_i}{(a_{i_2} - a_{i_1})} \int_{a_{i_1}}^{a_{i_2}} |x - a_i| \, da_i$$

and

$$\text{P3.5.3} \quad \text{minimize } f(y) = \sum_{i=1}^{m} \frac{w_i}{(b_{i_2} - b_{i_1})} \int_{b_{i_1}}^{b_{i_2}} |y - b_i| \, db_i$$

Hence, min f(X) = min f(x) + min f(y).

Thus, a solution technique need only be formulated for solving one coordinate. For the purposes of this discussion, only P3.5.2 will be considered. The same results will also apply to P3.5.3. The objective function in P3.5.2 will henceforth be referred to as f(x).

<u>Theorem 3.5.1</u>: The function f(x) is a convex function for all real values of x and is strictly convex when $i \varepsilon [1,\ldots,m]$ such that $x \varepsilon [a_{i_1}, a_{i_2}]$.

<u>Proof</u>: For this function, in order to prove that f(x) is convex, it is sufficient to show that at least one of the functions under the summation is convex. A sufficient condition for convexity is that the second derivative is always nonnegative. The sufficient condition for strict

convexity is that the second derivative is positive.

Since $\frac{w_i}{a_{i_2}-a_{i_1}}$ is a positive constant, it will be sufficient to examine the convexity of $\int_{a_{i_1}}^{a_{i_2}} |x-a_i|da_i$.

Integration produces the following results:

$$\int_{a_{i_1}}^{a_{i_2}} |x-a_i|da_i = \frac{(a_{i_2}-x)^2 - (a_{i_1}-x)^2}{2} \quad \text{if } x \leq a_{i_1}$$

$$= \frac{(x-a_{i_1})^2 + (a_{i_2}-x)^2}{2} \quad \text{if } a_{i_1} \leq x \leq a_{i_2}$$

$$= \frac{(x-a_{i_1})^2 - (x-a_{i_2})^2}{2} \quad \text{if } a_{i_2} \leq x \qquad (3.5.3)$$

The first derivatives of these expressions are, respectively:

$$a_{i_1} - a_{i_2} \quad , \text{ if } x < a_i$$

$$2x - (a_{i_1}+a_{i_2}), \text{ if } a_{i_1} \leq x \leq a_{i_2}$$

$$a_{i_2} - a_{i_1} \quad , \text{ if } a_{i_2} < x \qquad (3.5.4)$$

The second derivatives are 0, 2, and 0, respectively. Thus, the individual functions are convex so f(x), the sum of convex functions, is also convex.

Notice that if $x \varepsilon [a_{i_1}, a_{i_2}]$ for some i, then the second derivative of f(x) will be positive because one of the functions under the summation has a positive second derivative.

Notice that the first derivatives as calculated in this proof are continuous. Thus, the observation can also be made that f(x) is continuously differentiable over all values of x.

Corollary 3.5.1: If a value of x can be found such that $x \varepsilon [a_{i_1}, a_{i_2}]$ and $\frac{\partial f(x)}{\partial x} = 0$, then that x is a global minimum.

Hence, if we let w_i' represent the weight on region i and assume the i's are labeled such that the a_i's are in ascending order, then at point x = a the derivative is:

$$\sum_{i=1}^{j-1} w_i'(a_{i_2}-a_{i_1}) + 2aw_j' - w_j'(a_{j_1}+a_{j_2}) - \sum_{i=j+1}^{m} w_i'(a_{i_2}-a_{i_1})$$

$$= \quad -\sum_{i=1}^{m} w_i'(a_{i_2}-a_{i_1}) + 2\sum_{i=1}^{j-1} w_i'(a_{i_2}-a_{i_1}) + 2w_j'(a-a_{j_1})$$

where $a \varepsilon [a_{j_1}, a_{j_2}]$ (3.5.5)

This expression must be set equal to zero to solve for a.

3.5.3 The Algorithm

This algorithm is based on the gradient reduction method of Wesolowsky and Love (1971). It will be more general than their algorithm since their algorithm was developed only in the framework of working one example problem. This algorithm is based on the developed theory.

Algorithm for Solving the Single Facility Location Problem

Initialization:

1. Sort the intervals $[a_{i_1}, a_{i_2}]$ i = 1,2,...,m, on the basis of ascending a_{i_1}.

2. Compute $w_i' = \dfrac{w_i}{a_{i_2} - a_{i_1}}$ for each i.

3. Decompose the intervals $[a_{i_1}, a_{i_2}]$ into nonoverlapping intervals $[r_j, s_j]$ $j = 1,\ldots,p$. The weight on each interval is:

$$w_j' = \sum_{i \varepsilon I_j} w_i'$$

where I_j is set of indices of the original intervals i such that $[r_j, s_j] \cap [a_{i_1}, a_{i_2}] \neq \phi$.

Gradient Reduction:

4. Compute $M = \sum_{j=1}^{p} w_j' \, (s_j - r_j) = \sum_{i=1}^{m} w_i$

5. Let $k = 1$

6. Compute $t_k = w_k' \, (s_k - r_k)$.

   The derivative at $x = s_k$ is $d(s_k) = -M + 2 \sum_{j=1}^{k} t_j$

7. If $d(s_k) < 0$, set $k = k+1$ and go to 6.

8. If $d(s_k) = 0$, then $s_k \le x^* \le r_{k+1}$. Stop.

9. If $d(s_k) > 0$, then $x^* = r_k - d(s_{k-1}) \cdot \dfrac{b_k - a_k}{2t_k}$. Stop.

Concisely, the weight $2t_k$ is added to $d(s_{k-1})$ until the sum becomes nonnegative. If the sum becomes zero, there will be an interval solution to the location problem. This means that the new facility may be optimally located at any given point on the interval. If the sum becomes positive, there will be a point solution.

The justification for this result comes from (3.5.5). Obviously,

$$d(s_k) = d(r_{k+1}) \qquad \text{for all } k \qquad (3.5.6)$$

Also, $d(x) = d(s_k)$ for $x \varepsilon (s_k, r_{k+1})$ (3.5.7)

This follows from the fact the intervals are non-overlapping. Thus, if $d(s_k) = 0$, then $d(x) = 0$ for $x \varepsilon [s_k, r_{k+1}]$ and all x on this interval are global optima:

The result for the point solution comes from a study of Figure 3.3.

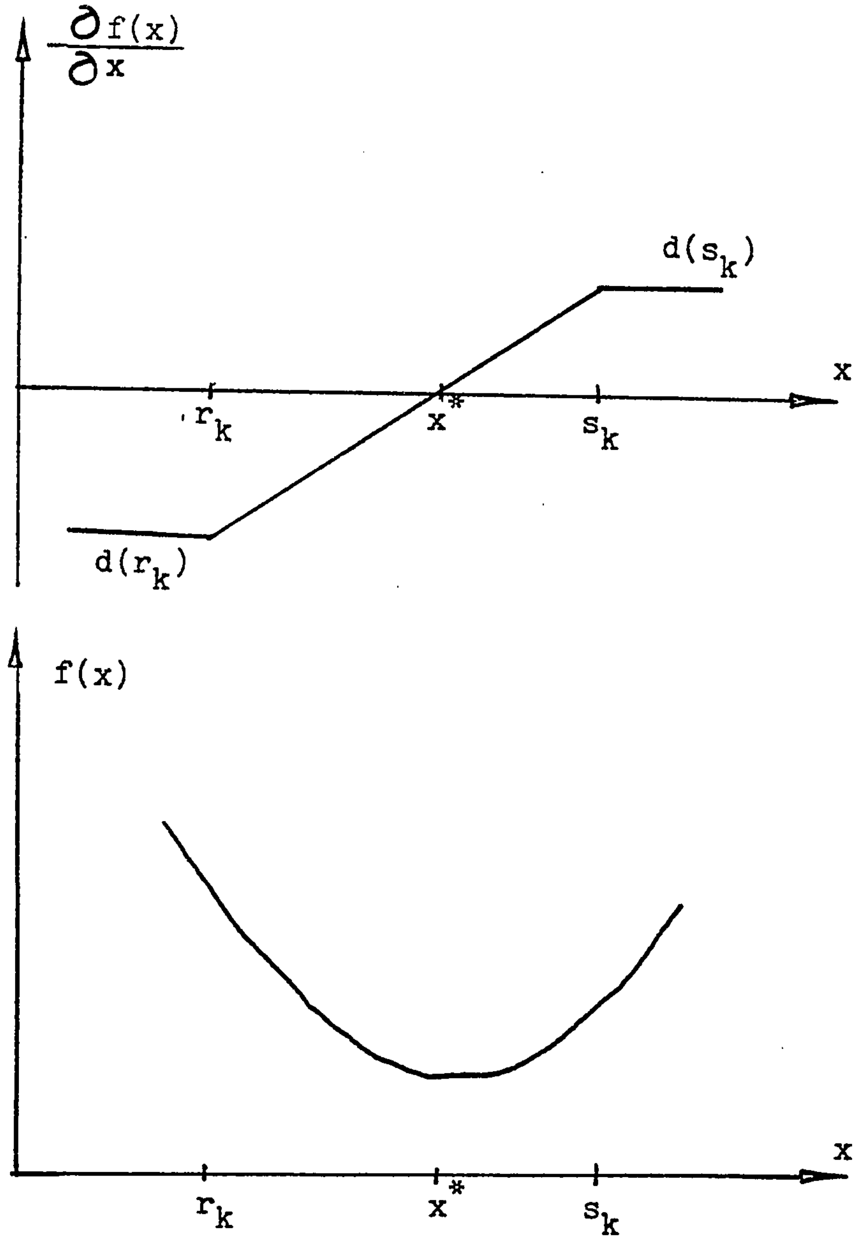


Figure 3.3. A graph of f(x) and its derivative.

x^* is solved by considering the line between $(r_k, d(r_k))$ and $(x^*, 0)$. This result was presented in Step 9 of the algorithm.

3.5.4 A Numerical Example

A simple numerical result will be worked to illustrate the algorithm.

Let $R_1 = [1,3] \times [1,3] \quad w_1 = 2$

$R_2 = [2,3] \times [2,4] \quad w_2 = 1$

$R_3 = [4,5] \times [2,3] \quad w_3 = 3$

x Subproblem:

$j=1 \; [r_j, s_j] = [1,2] \quad w_1' = 1$

$2 \quad [2,3] \quad w_2' = 1 + 1 = 2$

$3 \quad [4,5] \quad w_3' = 3$

$M = 2 + 1 + 3 = 6$

$d(s_1) = -6 + 2 = -4$

$d(s_2) = -6 + 2 + 4 = 0$

Thus, there is an interval solution: $3 \leq x^* \leq 4$.

y Subproblem:

$j=1 \; [r_j, s_j] = [1,2] \quad w_1' = 1$

$2 \quad [2,3] \quad w_2' = 1 + \frac{1}{2} + 3 = 4\frac{1}{2}$

$3 \quad [3,4] \quad w_3' = \frac{1}{2}$

$M = 6$

$d(s_1) = -6 + 2 = -4$

$d(s_2) = -6 + 2 + 9 = 5$

There is a point solution: $y^* = 2 + \frac{4 \cdot 1}{9} = 2\,\frac{4}{9}$.

3.6 The Multifacility Location Problem with Rectilinear Norm

The problem to be considered in this section is the multifacility version of the problem in Section 3.5. It is also formulated with the expected value criterion.

P3.6.1 minimize $\sum_{j=1}^{n} \sum_{i=1}^{m} \frac{w_{ij}}{A_i} \iint_{R_i} \left(|x_j-a_i| + |y_j-b_i|\right) da_i db_i$

$$+ \sum_{1 \le j<k \le n} v_{jk}(|x_j-x_k| + |y_j-y_k|)$$

where: n = number of new facilities to be located

m = number of existing facilities

(x_j,y_j) = coordinates of new facility j

(a_i,b_i) = general coordinates of any point in region R_i. The area of R_i is A_i.

w_{ij} = interaction cost per unit distance per unit time between new facility j and existing facility i.

v_{jk} = interaction cost per unit distance per unit time between new facilities j and k.

The method used by Wesolowsky and Love (1971) involved consideration of the projection of the polyhedral surface formed by the objective function onto the planes formed by the coordinate axes, i.e., the x_1x_2 plane. The projection of each edge of the polyhedral surface is considered in turn. The minimum of the objective function on the considered projection is then computed by solving a single facility location problem. This method continues until all projections have been checked.

There are several reasons why this algorithm can be considered too inefficient,

1. There is no criterion for optimality. The only way optimality may be ascertained is to consider all the projections. Thus, optimality may be achieved at the first iteration; however, one would never know if it were a local or global minimum until the last projection were checked.
2. The procedure becomes very complex for values of n larger than two, as Wesolowsky and Love stated. They gave no evidence that they had been able to program the algorithm for $n > 2$.
3. No computational experience had been developed. Although Wesolowsky and Love essentially presented their procedure for the two facility location problem, they did not even present the computational results for this algorithm. Thus, the algorithm will produce the optimal solution, but it is untested in a computational sense.

The algorithm to be presented will be a heuristic algorithm although it often produces the optimal solution. Computational experience with the algorithm will be stressed.

3.6.1 Assumptions for the Multifacility Problem

The assumptions for this algorithm will be the same as those described in Section 3.5.1. Assumption 1 concerning the w_{ij}'s can also be directly applied to the v_{jk}'s. Thus,

v_{jk} may reflect either a deterministic value or the expected value $E[v_{jk}]$ when v_{jk} is a random variable.

It will be further assumed that there exists a k, k = 1, ..., n such that

$$v_{jk} > 0 \quad \text{for every } j.$$

This means that every new facility must interact with at least one other new facility. Basically, this is a formulational assumption because if there were some j such that $v_{jk} = 0$ for each k, then that new facility j could be found by just solving a single facility location problem. Thus, the v_{jk}'s prevent the problem from degenerating into n single facility problems.

On the other hand, it will be assumed that at least one $w_{ij} = 0$. If all $w_{ij} > 0$, then the problem would become relatively uninteresting. It would only be necessary to solve the single facility location problem and then to locate all n of the new facilities on that solution point so that the interaction terms between new facilities would make no contribution to the objective function.

3.6.2 Properties of the Objective Function

It follows readily from the development of the single facility algorithm that the multifacility problem may also be decomposed into an x-coordinate subproblem and a y-coordinate subproblem. The difference is that the solution to each of the subproblems will be an n-dimensional vector. Once more, the x-coordinate subproblem will be addressed with the assumption that the results may be applied to the y-coordinate subproblem without loss of generality.

3.6.2.1 Convexity

Corollary 3.6.1: The function $f(x_1, x_2, \ldots, x_n)$ is convex for all real values of x_j, $j = 1, \ldots, n$. It is strictly convex when there is an i for some x_j such that $x_j \varepsilon [a_{i_1}, a_{i_2}]$.

Proof: The results of Theorem 3.5.1 will be used. For any value of x_j, $j = 1, \ldots, n$, the general function

$$\sum_{i=1}^{m} \frac{w_{ij}}{a_{i_2} - a_{i_1}} \int_{a_{i_1}}^{a_{i_2}} |x_j - a_i| \, da_i$$

is convex. It is strictly convex if x_j is located in some interval $[a_{i_1}, a_{i_2}]$. The function $v_{jk}|x_j - x_k|$ is also convex since it has a second derivative of zero no matter what the relative values of x_j and x_k are. Thus, since the summation of convex functions is also convex, $f(x_1, \ldots, x_n)$ is convex and strictly convex when there is a $j \varepsilon \{1, \ldots, n\}$ such that $x_j \varepsilon [a_{i_1}, a_{i_2}]$ for some i.

3.6.2.2 Non-Differentiability

Although a partial derivative exists for all real values of x_j, these partial derivatives are not necessarily continuous. This follows from the formula for a partial derivative:

$$\frac{\partial f(x_1,\ldots,x_n)}{\partial x_j} = \sum_{i\varepsilon s_j^+} w_{ij} - \sum_{i\varepsilon s_j^-} w_{ij} + \sum_{i=1}^{m} z_i w_{ij} \frac{(2x_j-(a_{i_1}+a_{i_2}))}{a_{i_2}-a_{i_1}}$$

$$+ \sum_{k\varepsilon t_j^+} v_{jk} - \sum_{k\varepsilon t_j^-} v_{jk} \qquad (3.6.1)$$

where: $z_i = \begin{cases} 1, \text{ if } x_j \varepsilon [a_{i_1}, a_{i_2}] \\ 0, \text{ otherwise} \end{cases}$

s_j^+ = the set of indices i such that $x_j > a_{i_2}$.

$i = \{1, \ldots, m\}$

s_j^- = the set of indices i such that $x_j < a_{i_1}$.

t_j^+ = the set of indices k such that $x_k < x_j$.

$k = \{1, \ldots, n\}$

t_j^- = the set of indices k such that $x_j < x_k$.

Thus, to illustrate a discontinuity, consider the following case. Let $x_1 = a$ be such that $a\varepsilon(a_{i_2}, a_{(i+1)_1})$ for some i. In other words, x_1 is a point located between two nonoverlapping intervals. Assume $v_{12} = 2$, and let the partial derivative with respect to x_2 be -1 at the point (a, a_{i_2}). Thus, for all points (a, x_2) where $x_2 < a$, the partial derivative is -1. When $x_2 > a$, the v_{12} term will change from having a negative sign in the formula to having a positive sign for a net increase of 4.

Thus,

$$\frac{\partial f(a,x_2)}{\partial x_2} x_2 = a^- = -1$$

$$\frac{\partial f(a,x_2)}{\partial x_2} x_2 = a^+ = 3$$

Hence, the partial derivative is not continuous.

The implications from this discussion are obvious. It will not be possible to use a gradient as a solution technique because of discontinuities in the partial derivatives. On the other hand, the convexity results indicate that if a local minimum can be found such that one of the x_j's is located on one of the nonoverlapping intervals, then this solution is a global minimum.

3.6.3 Development of a Solution Technique

The results of the previous discussion indicate that a gradient-free nonlinear search technique is recommended to solve the multifacility problem. Furthermore, if the search could be restricted to the area formed by the projection of the polyhedral edges of the objective function onto the plane, this would be the same as requiring one of the variables x_j to be located within an interval $[a_{i_1}, a_{i_2}]$. Hence, one could be assured that the function being subjected to search was strictly convex.

3.6.3.1 A Direct Search Technique

An appropriate search method would be the Hooke and Jeeves (1961) direct search technique. Direct search techniques as well as other nonlinear optimization procedures are discussed by Himmelblau (1972), Geoffrion (1972), and Box et al. (1969). The Hooke and Jeeves algorithm is gradient free and is adapted to convex nonlinear functions containing no cross products of variables--all characteristics of the multifacility location problem. This is a heuristic algorithm.

3.6.3.2 A Starting Point and Search Direction

With the Hooke and Jeeves algorithm as the backbone to the solution technique, the remaining tasks for total development of the algorithm are to obtain an initial point and a search direction. When the multifacility location problem is considered, there are two interesting points, excluding the optimum, because they represent two extremes. The first point denoted w is the point with a minimum interaction cost between new and existing facilities. This point can be found by disregarding the terms involving interactions between new facilities (v_{jk}'s) and just solving the remaining function as n single facility location problems. The second point is the point with a minimal interaction cost between new facilities and will be denoted v. Obviously an interaction cost of zero occurs when all facilities are located on the same point. This point will be located on the line $x_1 = x_2 = \ldots = x_n$. The location of this point is easily found by disregarding the interaction terms between new facilities (v_{jk}'s) and treating the remaining problem as one single facility location problem.

These two points represent the tradeoffs that are made in the multifacility location problem. The point w represents minimal interaction costs between new and existing facilities at a sacrifice to the interaction costs between existing facilities. On the other hand, the point v minimizes the interaction costs between new facilities while increasing the same costs between new and existing facilities. The optimal solution will strike a balance between these two costs.

Thus, when the v_{jk}'s are small relative to the w_{ji}'s in a location problem, the optimum should be found in a neighborhood of w. On the other hand, when the v_{jk}'s become large relative to the w_{ji}'s, the optimum will be forced toward the line $x_1 = x_2 = \ldots = x_n$ and may actually be located at v. The problem is that there is no way to measure the relative magnitudes of these interaction costs. The following two examples illustrate this concept for the x-coordinate subproblem:

Example 3.6.1: (Two New Facilities--Two Existing Facilities)

$$a_1 = [1,2] \qquad w_{ji} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$
$$a_2 = [5,6]$$

When $v_{12} = 0$ $\quad x_1 = 1\,\frac{3}{4},\ x_2 = 5\,\frac{1}{3}$

When $v_{12} = 1$ $\quad 2 \le x_1 \le 5,\ x_2 = 5\,\frac{1}{6}$

When $v_{12} = 2$ $\quad x_1 = 5\,\frac{1}{8},\ x_2 = 5\,\frac{1}{8}$

A graph of this example is presented in Figure 3.4.

Example 3.6.2: (Four Existing Facilities--Two New Facilities)

$$a_1 = [1,2] \qquad w_{ji} = \begin{bmatrix} 0 & 0 & 2 & 1 \\ 1 & 3 & 0 & 0 \end{bmatrix}$$

$$a_2 = [5,6]$$

$$a_3 = [7,8]$$

$$a_4 = [11,12]$$

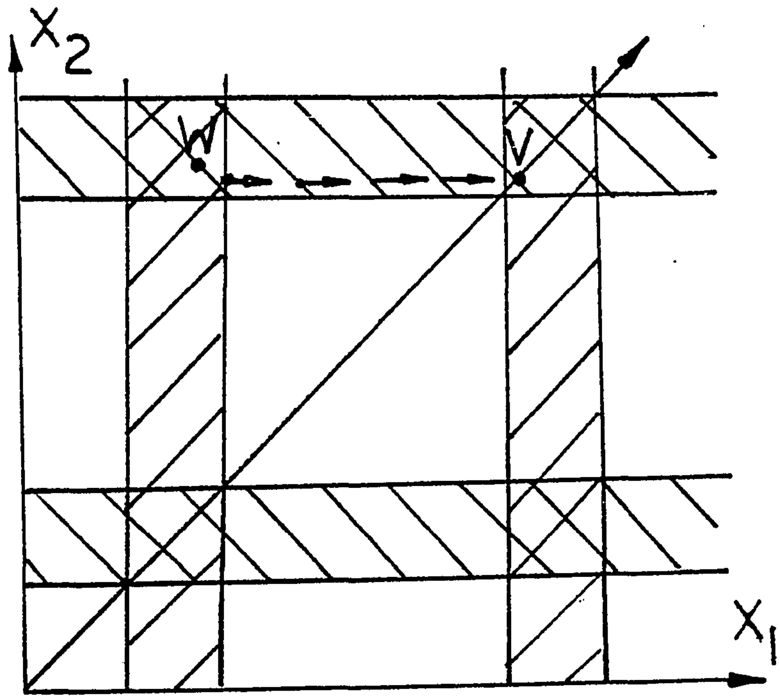


Figure 3.4. A graph of Example 3.6.1.

When $v_{12} = 0$ $\quad x_1 = 7\frac{3}{4},\ x_2 = 5\frac{1}{3}$

When $v_{12} = 1$ $\quad x_1 = 7\frac{1}{2},\ x_2 = 5\frac{1}{2}$

When $v_{12} = 2$ $\quad x_1 = 7\frac{1}{4},\ x_2 = 5\frac{2}{3}$

When $v_{12} = 3$ $\quad 5\frac{5}{6} \le x_1 \le 7,\ x_2 = 5\frac{5}{6}$

When $v_{14} = 4$ $\quad x_1 = 5\frac{5}{6},\ x_2 = 5\frac{5}{6}$

A graph of this example is presented in Figure 3.5.

In both Figures 3.4 and 3.5, the shaded region represents the projections of the edges of the polyhedron onto the x_1x_2 plane. The solution to any of the multifacility location problems is found in the shaded area which was expected because of previously discussed properties of the problem.

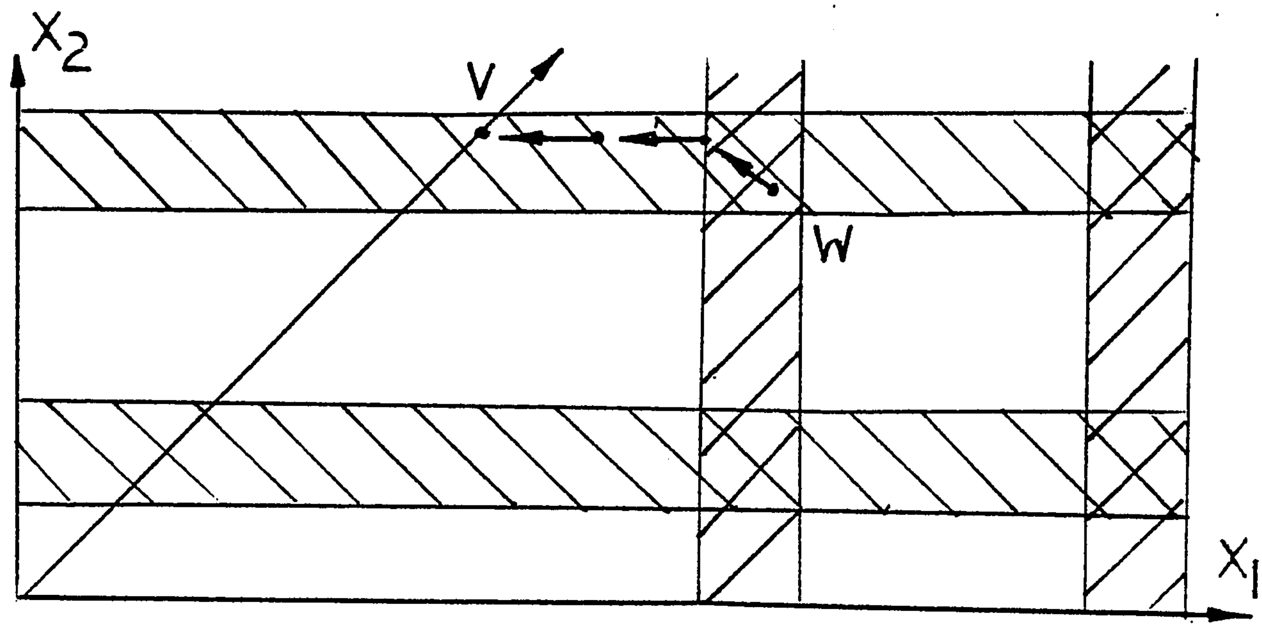


Figure 3.5. A graph of Example 3.6.2.

Empirical results from these two examples indicate that the point w discussed before is a good initial point for a direct search technique. In these examples, some solutions were collinear on a line segment emanating from point w. This point becomes x^o in the algorithm.

Intuitively, a good search direction would search the shaded area between x_o and the line $x_1 = x_2 = \ldots = x_n$. It is obvious that this region reflects the previously discussed tradeoffs between the w_{ji}'s and the v_{jk}'s. Hence, the optimum should be found by searching this region.

The search direction to be developed will be a vector from point x^o toward the line $x_1 = x_2 = \ldots = x_n$, which will be denoted line ℓ. The point x^o also defines a vector originating at the origin and passing through x^o. Let x^ℓ represent a point on ℓ; x^ℓ also defines a vector.

Consider the direction angles of these vectors. The direction angles, denoted $\theta_1, \theta_2, \ldots, \theta_n$, of a vector are just the angles between the vector and the positive x_1-, x_2-, ..., x_n- axes, respectively. For an arbitrary vector $(x_1, x_2, \ldots, x_n)$, the cosine of the ith direction angle (θ_i) is defined as:

$$\cos \theta_i = \frac{x_i}{\left(\sum_{i=1}^{n} x_i\right)^{\frac{1}{2}}} \qquad (3.6.2)$$

Using this definition, the cosines of all the direction angles of x^{ℓ} are just $\frac{1}{\sqrt{n}}$.

Let the cosines of the direction angles of x^o be denoted $\cos \theta_i$ $i = 1, \ldots, n$. For the ith coordinate axis, if $\cos \theta_i > \frac{1}{\sqrt{n}}$ then the measure of x^o's ith direction angle is less than that of x^{ℓ}'s direction angle. Thus, to move x^o toward ℓ, the coordinate value x_i^o should be decreased by some amount.

On the other hand, if $\cos \theta_i < \frac{1}{\sqrt{n}}$, then x^o's direction angle is larger than x^{ℓ}'s with respect to the ith coordinate axis. Hence, x_i^o should be increased.

This concept is illustrated in Figure 3.6.

For any vector located on ℓ, both direction angles will be 45^o with cosines of $2^{-\frac{1}{2}}$. In this figure $\theta_1 > 45^o$ and $\theta_2 < 45^o$. Thus, $\cos \theta_1 < \frac{1}{2}$ so x_2^o should be decreased. The resultant vector of these two adjustments is a direction toward ℓ. If is the magnitude of the change in each doordinate value, then the initial search direction is $(\varepsilon, -\varepsilon)$.

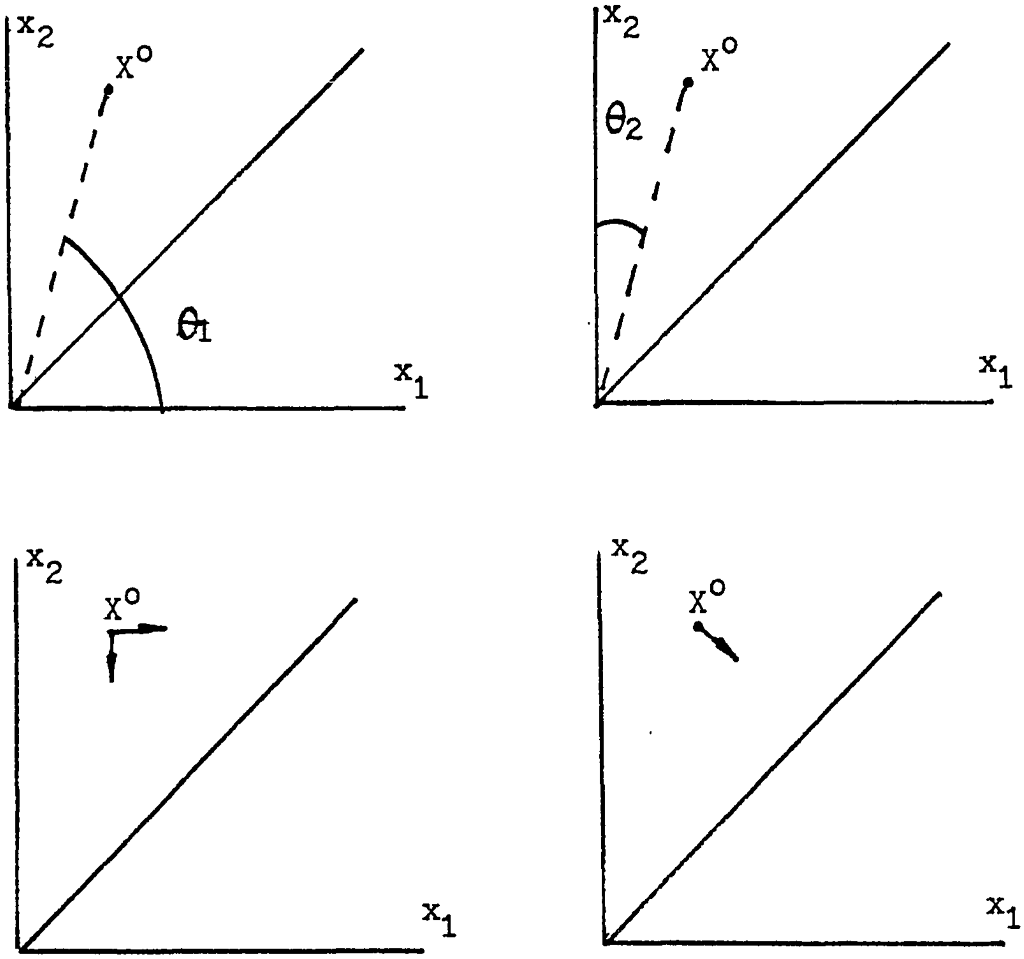


Figure 3.6. Determining a search direction.

3.7 The Algorithm

The complete algorithm for the multifacility location problem is stated below.

The input parameters are:

M - number of rectangular regions

N - number of new facilities

EPS - size of each component in initial direction vector

ALPHA - a scalar ≥ 1 to be used to increase a step size in a search direction

BETA - a scalar ($0<\beta<1$) to be used to decrease a step size in a search direction

EPSY - convergence criterion for terminating algorithm.

IMAX - maximum number of iterations allowed

1. Input the above parameters, the w_{ji}'s, and the intervals $[a_{i_1}, a_{i_2}] \times [b_{i_1}, b_{i_2}]$ for all regions i. (The first interval refers to the x-coordinates; the second to the y-coordinates.)

2. Solve a single facility location for each new facility j among the rectangular regions i where $w_{ji} > 0$. (This step entails proceeding through the algorithm in Section 3.5.) Denote these locations (x_j, y_j).

3. Input the v_{jk}'s.

The following steps will be addressed to the x-coordinate solution. They should then be repeated for the y-coordinate solution.

The Search Direction Vector--Direct

4. Compute $\frac{1}{\sqrt{n}}$.

5. Compute $\cos(j) = \frac{x_j}{(\Sigma x_i)^{\frac{1}{2}}}$ for each j.

6. If $\cos(j) < \frac{1}{\sqrt{n}}$, let Direct (j) = EPS

   If $\cos(j) > \frac{1}{\sqrt{n}}$, let Direct (j) = -EPS

   If $\cos(j) = \frac{1}{\sqrt{n}}$, let Direct (j) = 0

7. Evaluate objective function at point $(x_1, \ldots, x_n)$.

Direct Search Procedure

8. Let k = 1. Put the objective function value of Step 7 in OPT, the current best value and WOPT, previous best. Put the x_j's in the vector XOPT, the current best point and WXOPT, previous best point.

9. Let j = 1.

10. Compute x(j) = X(j) + Direct(j). Calculate the objective function value W(j) with this new value. If W(j) < OPT, go to 12.

11. Compute X(j) = X(j) - Direct(j). Calculate its objective function value W(j). If W(j) > OPT, go to 18.

12. Replace OPT with W(j). Replace XOPT(j) with X(j).

13. j = j+1. If j = n+1, go to 14. Otherwise, go to 10.

14. If k > IMAX, stop. Go to 20.

15. If |OPT - WOPT| < EPSY, stop. Go to 20.

16. k = k + 1.

17. Let X(j) = XOPT(j) + ALPHA·(XOPT(j) - WXOPT(j)) for all j. Go to Step 9.

18. Calculate and replace Direct(j) with BETA·Direct(j).

19. If |Direct(j)| < EPSY, stop and go to 20. Otherwise, go to Step 9.

Check for Interval Solutions

20. Let j = 1.

21. For i such that $w_{ij} > 0$, if XOPT(j) ε (a_{i_2}, a_{i+1_1}),

then XOPT(j) will have an interval solution on the above interval. If not, go to 22.

22. $j = j+1$. If $j < n$, go to 20. Otherwise, terminate the algorithm with the vector XOPT as the optimal locations and objective function value OPT.

These last three steps were necessary because in location problems with interval solutions, any point on the interval will produce the same objective function value. The search method would not perceive a point in the interval to be an improvement in the objective function value so it would distregard interval solutions.

The algorithm was verified by testing it with the examples of Wesolowsky and Love plus other examples contrived by the author. In all cases the objective function value and facility locations were all within a small neighborhood of optimal.

3.8 Computational Results

Computational results were developed for the multifacility location problem. The test problems were randomly generated. The interaction weights, w_{ji} and v_{jk}, were generated from a uniform distribution over the values [0,10]. The corner points of the rectangular regions, i.e., (a_{i_1},b_{i_1}) and (a_{i_2},b_{i_2}), were also randomly generated from a uniform distribution over the interval [0,100]. The problems were run on an IBM 370/158J computer.

The results are presented below in Figure 3.7. Three problems were generated and run for each particular combination of new facilities and existing facilities. The average of the execution times for all three problems is displayed.

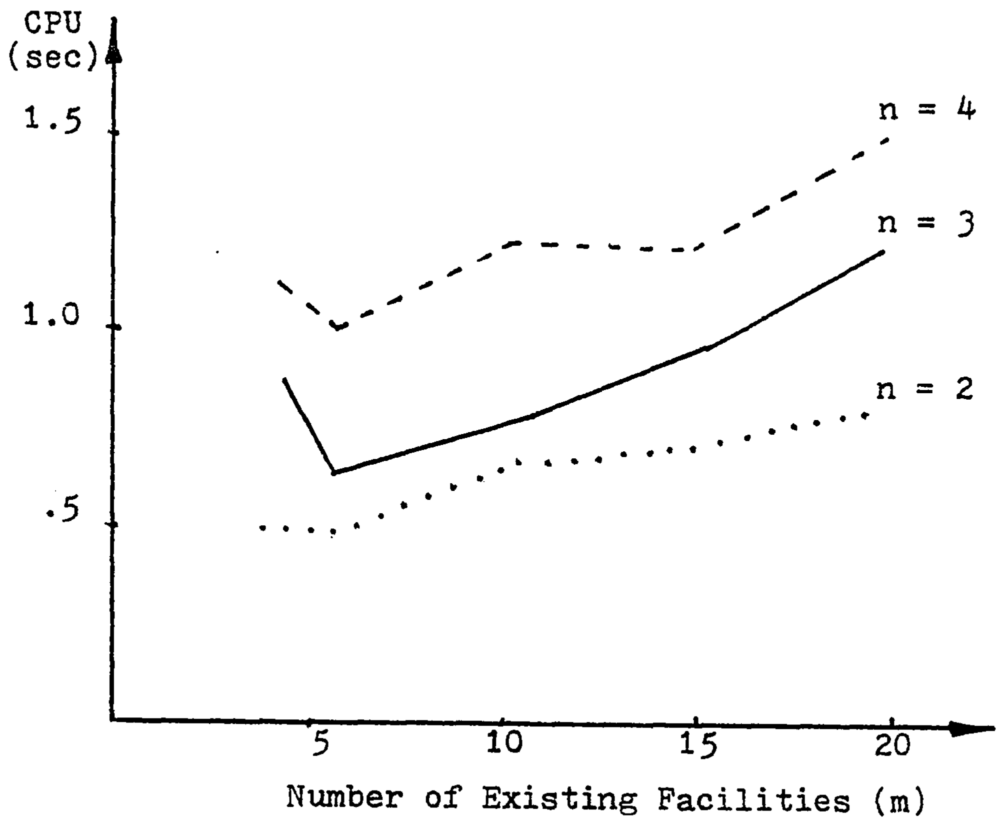


Figure 3.7. CPU time required to solve a multifacility location problem with n new facilities.

The results indicate, not surprisingly, that the execution time is a function of first, the number of new facilities and secondly, the number of existing facilities. Figure 3.7 indicates that at least in the case of m = 4, the CPU time rises as n approaches m. However, the CPU time was relatively insensitive to changes in m from 10 to 15. At m = 20, the times rise rapidly; it would be expected that the time would continue to rise at an exponential rate for values of m greater than 20.

Although there are no other results to serve as a comparison to these values, twenty existing facilities is usually considered a medium-scale location problem. The fact that this problem could be handled in about 1.5 seconds of CPU time is promising for the algorithm's performance on larger problems using only a reasonable amount of computational time.

3.9 Summary

In this chapter the location problem with respect to rectangular regions was considered. The centroid approach, a traditional approximation to the consideration of rectangular regions, was discussed in terms of metric. With Euclidean distance, the centroid approach could be used in conjunction with deterministic techniques to obtain a near optimal solution. The rectilinear distance metric was more sensitive to the use of centroids instead of rectangular regions.

The single facility location algorithm was a complete version of a gradient reduction procedure used by Wesolowsky and Love. The properties of the problem were developed as a basis for the completed algorithm.

The properties of the multifacility location problem were discussed to serve as a motivation for a solution technique. Because of discontinuities in the partial derivatives, a gradient-free direct search was used. Although it is a heuristic, the search yielded near-optimal solutions. Methods of developing a good initial point and search direction were also discussed as a part of the algorithm. Computational results were obtained for the algorithm.

The chief contributions of the algorithm are:

1. Location problems with a large number of "existing facility" regions and a larger number of new facilities than had ever been solved before were handled in a reasonably small amount of computational time.
2. Computational results are provided to serve as a basis for comparison for future works.
3. The direct search technique may be used as a model to handle other similarly "difficult" location problems.
4. The results of the direct search method give insight to the surface of the objective function in an area around the optimum. This may become a basis for some sensitivity analysis.

CHAPTER IV

THE LOCATION-ALLOCATION MODEL

4.1 Introduction

In Chapter III, the location problem was solved when the interaction between new and existing facilities was known a priori. In this chapter it will be assumed that only the locations of the existing facilities, their interaction costs with an arbitrary new facility, and the number of new facilities. The existing facilities must be allocated to the new facilities in a manner such that when the new facilities are located, the total cost of interaction between new and existing facilities is minimized. In this respect the location-allocation model will be built on the results of the location problems considered in the previous chapter.

4.2 Formulations

The general location-allocation model among rectangular regions is formulated as follows.

P4.2.1 minimize $\sum_{j=i}^{n} \sum_{i=1}^{m} \iint_{R_i} z_{ij} w_i |X_j - R_i|_{\ell_p} \, \theta(R_i) dR_i$

$$\text{subject to:} \quad \sum_{j=1}^{n} z_{ij} = 1 \qquad \text{for all } i$$
$$z_{ij} = 0,1 \qquad \text{for all } i \text{ and } j$$

where: n = number of new facilities

m = number of existing facilities

$X_j = (x_j, y_j)$ coordinate location of new facility j

R_i = existing rectangular region i

$\theta(R_i)$ = bivariate probability density function over R_i

w_i = interaction between region i and the new facility it will be allocated to

$$z_{ij} = \begin{cases} 1, & \text{if existing facility } i \text{ is allocated to new facility } j \\ 0, & \text{otherwise} \end{cases}$$

The particular problem to be emphasized in this chapter is the location-allocation problem among rectangular regions with bivariate uniform distributions. An expected value criterion will be used.

$$\text{P4.2.2} \quad \text{minimize} \quad \sum_{j=1}^{n} \sum_{i=1}^{m} \frac{z_{ij} w_i}{A_i} \int_{a_i} \int_{b_i} |X_j - R_i|_{\ell_p} \, da_i \, db_i$$

$$\text{subject to:} \quad \sum_{j=1}^{n} z_{ij} = 1 \qquad i = 1, \ldots, m$$
$$z_{ij} = 0,1 \qquad \text{for all } i \text{ and } j$$

where: (a_i, b_i) = general coordinate location in region R_i

A_i = area of region R_i

and n, m, w_i, X_j, R_i, and z_{ij} are as defined in P4.2.1.

Note that $\frac{I}{A_i}$ in P4.2.2 is just the bivariate uniform density function over R_i.

In Problems 4.2.1 and P4.2.2, the decision variables are the z_{ij}'s--reflecting the allocation aspects of the problem and the x_j's--reflecting the location aspects of the problem.

Note also that there is an implicit assumption that the new facilities have an infinite capacity to serve the existing facilities. Thus, each existing facility will be allocated to and subsequently interact with only the closest new facility. If there had been capacity restrictions on the new facilities, the problem would have been a transportation-location problem.

The assumptions of Section 3.5.1 will be somewhat relaxed. The first assumption that the w_i's may represent either deterministic values or expected values of random variables still holds. The second assumption will be relaxed. The regions must be rectangular, but they may be overlapping.

4.3 Related Works

There have been several works that have served as the state of the art for location-allocation (L-A) systems. Scott (1970) presented a review of the types of L-A models that had been formulated. Lea (1973) compiled an annotated bibliography of all works that dealt with an aspect of the L-A systems. In their comprehensive report on facilities location, Elshafei and Haley (1974) discussed the results of all types of research done with the model--whether it be solution techniques or applications. The reader is referred to these works for an overview of the L-A system.

The solution methods that have been developed for the deterministic version of the problem will be discussed in the rest of this section.

The use of dynamic programming was originally considered by Bellman (1965). He took the single facility location problem with Euclidean distance metric and transformed it to a dynamic programming problem by using quasi-linearization, a technique he had developed earlier which involved transforming the Euclidean metric to a linear form subject to both linear and nonlinear constraints. Love (1976) later revived the idea to solve the multifacility location problem when the existing facilities are all collinear. He used an absolute value distance metric, which is just the rectilinear metric applied to the one dimensional version of the problem.

Leamer (1968) studied the location-allocation model when the customers are uniformly distributed over a planar area. Since he was using Euclidean distances, he determined that the ideal situation would be to consider one allocation to be all customers located within a circle around a new facility. However, since circles do not divide the plane, the hexagon or a honeycomb shape would be the best geometrical figure to both divide the plane and minimize transportation costs. His work was done in an economic setting with each hexagon representing a market area.

In a series of articles, Cooper (1963, 1964, and 1967) developed heuristic algorithms for and properties of this model. The details of the heuristics will be discussed later. His results have served as a basis for subsequent research efforts by others interested in the location-allocation model.

Love and Morris (1975) developed an exact solution technique for the L-A problem with rectilinear distances. Their technique involved two phases: first, determination of all possible new locations for the new facilities and secondly, determining the optimal allocations and solving for the locations of the new facilities.

Their work was preceded by Kuenne and Soland (1972). The main contribution of their work was the use of a branch and bound technique to optimally solve the location-allocation problem. Their work will also be discussed later in greater detail.

Kuenne and Soland's work is one example of the larger role branch and bound techniques are playing in the solution of location problems. Versions of the L-A problem in discrete space, also known as the "plant location problem" or "fixed charge problem," had already been solved using branch and bound. Jarvinen et al. (1972) also applied branch and bound to solve the multifacility location problem on a graph. El-Shaieb (1973) also used branch and bound to solve a multifacility location problem with the restriction that a new facility must be located at the same site as an existing facility.

Ostresh (1975) worked on the Kuenne-Soland algorithm in terms of embellishing the bounding procedure. He applied the results of Wendell and Hurter (1973) to the two-facility problem. He determined that a simple test on the feasibility of an allocation scheme was to determine the convex hull of all the existing facilities allocated to one of the new facilities. The allocation scheme was feasible only if the two convex hulls were disjoint. However, this result could not easily be extended to location problems when the number of the new facilities, n, is greater than two.

In summary, all works reviewed here were addressed to the Euclidean distance metric except for that of Love and Morris (1975). All the research has been devoted to the L-A problem with the existing facilities being represented as deterministic points.

4.4 Characteristics of the Location-Allocation Problem

In this section several characteristics of the L-A problem will be discussed. These characteristics will explain the difficulties encountered in solving the L-A problem.

4.4.1 Non-Separability

The location-allocation problem is not separable. This means that the master problem cannot be separated into an independent allocation subproblem and an independent location subproblem, each of which is capable of being optimized. Instead, the two aspects of the problem are interdependent. If the optimal allocation scheme is known, the locations of the new facilities can be easily found by solving n single facility location problems among the existing facilities allocated to each of the n new facilities. Conversely, if the optimal locations of the new facilities are known, the allocation scheme can be determined by assigning each existing facility to the new facility which minimizes the weighted distance from it. This was the basis for Kuenne and Soland's heuristic algorithm CROSSCUT.

4.4.2 Non-Convexity

Cooper (1967) proved that the objective function of the location-allocation problem was neither convex nor concave because of the z variables. From the discussion above and the results of Chapter III it is obvious that if either the x's or the z variables assume values, the resulting location or allocation problem will be convex. This property indicates that a heuristic algorithm may find a local minimum that is not necessarily a global minimum. Thus, heuristic algorithms should be approached with some caution.

4.4.3 Insensitivities in Optimal Locations

The discussion in Chapter III indicated that many researchers have noticed that the multifacility location problem is extremely flat in the vicinity of the global optimum. Both Kuenne and Soland and Cooper observed that the same characteristic was true for the location-allocation problem. Hence, the objective function is more sensitive to inferior allocation schemes. An optimal allocation scheme will produce a near-optimal objective function value even if the new facilities are located at nonoptimal locations. Hence, a good solution technique will emphasize obtaining the optimal allocation.

4.4.4 Computational Burden

The total number of allocation schemes, or ways that m existing facilities can be assigned to n new facilities, is the Stirling number of the second kind--denoted $S(n,m)$.

$$S(n,m) = \sum_{k=0}^{n} \frac{(-1)^k (n-k)^m}{k!(n-k)!} \qquad (4.4.1)$$

Hence, for example, there are 63 ways to allocate seven existing facilities among two new facilities. This number increases exponentially with m. Therefore, any algorithm involving complete enumeration of all allocation schemes would quickly become computationally infeasible.

4.5 Heuristic Algorithms

Cooper promoted the use of heuristic algorithms during the 1960's because at the time there were no exact algorithms in existence. Most of the heuristics developed by others were modifications or adaptations of a heuristic originally proposed by Cooper.

Most of the heuristics involve the case where all the existing facilities have equal interaction weights. These heuristics are discussed below:

i) The destination subset algorithm involves considering all subsets of n existing facilities--that is, $\binom{m}{n}$ subsets in all--as a set of potential sites for the n new facilities. Although there was no guarantee, one of the subsets seemed likely to produce the optimal allocation. This heuristic coupled with an exact method of facility location, after the allocations were generated, produced the best objective function values of all the heuristics presented up to 1964. The computational times for this heuristic, however, were the worst requiring a prohibitive 3½ hours for a test problem where m = 60 and n = 4.

ii) The random destination algorithm involved the generation of n random numbers. Each random number was to be an integer between 1 and m. The n numbers indicated which existing facilities should become sites for the new facilities. The allocations were made and the resulting location problem solved. The procedure was repeated for a large number of trials until some arbitrary stopping criterion is met. This algorithm was deemed the best giving solutions only slightly higher than the previous algorithm at a fraction of the time. The test problem required only 8½ minutes.

iii) The successive approximations algorithm first solves a two facility location problem using the destination subset algorithm. If three new facilities are to be located, a third new facility is arbitrarily placed at the site of one of the existing facilities. The existing facilities are then reallocated and the location problem solved. This procedure continues as the third new facility is located at each of the m-1 other existing facilities, in turn. The best solution is considered to be the optimal allocation for locating three new facilities among the existing facilities. This procedure is then repeated for locating 4, 5, ... n new facilities. This algorithm gives the worst solutions averaging at least eight per cent of those given by the first algorithm. However, the computational times were the best. Only 48 seconds were required to solve the test problem. Kuenne and Soland modified this same algorithm to use weighted distances.

iv) The alternate location and allocation algorithm was initially proposed by Cooper (1963); it was initiated with an arbitrary allocation of existing facilities to new facilities. A single facility location problem was then solved for each of the subsets. Each existing facility was then reallocated to the closest new facility. Hence, there was a new allocation scheme. This procedure continued until no further improvement could be made. This algorithm produced solution values averaging a two per cent deviation from the first heuristic algorithm. Its computational times were second lowest with a time of just over two minutes for the test problem.

v) The elimination-alternate-correction heuristic was introduced in Cooper (1967) and was based on the previous heuristics. The elimination phase of the heuristic involved computing the sum of the distances between a given existing facility and all other existing facilities. The existing facility with the greatest sum is eliminated. The procedure is iterated until only n existing facilities remain. These facilities are taken as the locations of the new facilities. The alternate location and allocation algorithm is then applied until no further improvement is possible.

The correction phase is an attempt at further improvement. For each pair of existing facilities served by the same new facility, an interfacility distance is computed. The two existing facilities with the maximum interfacility distance between them become sites of new facilities. The new facility that serves them is then eliminated, leaving the system with n+1 new facilities. The two new facilities that are closest together are then identified. One of these new facilities is chosen to be eliminated from the system and the elimination-alternate procedure is used to attempt an improvement on the previous best solution. Then the other new facility is chosen to be eliminated and the process is repeated. This heuristic produced objective function values better than those produced by the destination subset algorithm. The computational time had improved to about 2.5 minutes required to solve a test problem where $m = 40$ and $n = 3$.

vi) The means-alternate-correction heuristic was applicable only to location-allocation problems with five or fewer new facilities. It involved locating five points. The first point was the coordinate values of the arithmetic means of the x and the y coordinates of all existing facilities. The other four points were the coordinates of the means of the existing facilities whose x-coordinates were less than the total mean; those whose x-coordinates were greater than the total mean; those whose y-coordinates were less than the total mean; and those whose y-coordinates were greater than the total mean. The elimination-alternate-correction algorithm is then used to eliminate the extra new facilities and improve the solution. This heuristic used in conjunction with the previous algorithm produced the best solution values.

vii) Cooper did not give a name to the heuristic he developed for solving the location-allocation problem with unequal interaction costs. Basically, it involved either just using the means-alternate-correction heuristic with weighted means or applying several of the heuristics in a progression. This meant one of the heuristics described above was chosen to generate an initial allocation. The alternate location and allocation algorithm was used several times to improve the solution. The correction phase of the elimination-alternate-correction algorithm is then used for further improvement. Cooper did not test this algorithm extensively.

Several conclusions can be drawn from the above discussion of heuristic approaches.

1. The heuristics were adapted to location-allocation systems where the existing facilities could be represented as deterministic points. The location of a new facility at the sites of existing facilities precluded the use of both probabilistic existing facilities and areal existing facilities. Modifications in the heuristics were necessary to accommodate these versions of the problem.

2. The heuristics were developed for the case when all existing facilities have equal interaction weights. Cooper was unable to develop a series of heuristics for the general problem with unequal interaction weights as he did for the case when all weights are equal. He discovered that the allocations produced by the best equal weight heuristics were usually never the optimal allocations for the L-A problem with unequal weights. Consequently, the heuristics that produced inferior solution values for the equal weights problem usually performed the best in obtaining allocations for the unequal weights problem. Hence, Cooper's results were not directly applicable to other problems.

3. There was no guarantee on the near-optimality of the heuristics. When Cooper did his work, there were no exact algorithms in existence. Thus, he had no way of knowing how close his results were to optimality. Heuristics could be judged only on the relative performance of other heuristics.

Eilon et al. (1971) studied the heuristics for a multifacility location-allocation problem with fifty existing facilities. They discovered that the per cent deviation from optimal of the worst local minimum generated by the heuristic ranged from 6.9 for n = 2 to 40.9 for n = 5.

The implications for the location-allocation problem are the following. The most important factor in obtaining an optimal solution is generating the optimal allocation scheme. The number of possible schemes is so large that complete enumeration is impossible. Heuristic algorithms will generate allocation schemes that correspond to local optima, but they may overlook the allocation scheme corresponding to the global optimum. Thus, with a heuristic algorithm there is no guarantee that even a near optimal solution will be produced.

Secondly, the fact that the best heuristics developed for a special case location-allocation system did not achieve the same success when applied to a general L-A system indicates that reliance on heuristic algorithms may dictate a need to generate new heuristic algorithms for every version of the location-allocation model.

In the next section an exact algorithm for the location-allocation model featuring implicit enumeration of all possible allocation schemes will be discussed. Since the branch and bound method was used by Kuenne and Soland (1972), the development of heuristic algorithms has been more or less, put to rest. Although other exact algorithms have been presented, there has been no comparison to date between heuristic and exact algorithms on the basis of computational results.

4.6 A Branch and Bound Approach

The branch and bound algorithm as developed by Kuenne and Soland (1972) offered an optimal solution to the location-allocation problem with at least the same computational efficiency as the heuristic algorithms. Two good review papers on general branch and bound methods are Lawler and Wood (1966) and Mitten (1970).

The algorithm was based on partitioning the set of all possible solutions to the location-allocation problem on the basis of the allocations of the existing facilities to the new facilities. For example, one set of solutions could be characterized by the assignment of the second and fourth existing facility to the second new facility. Hence, all possible solutions with allocation schemes having this one characteristic would be included in the set.

Any subset of solutions, denoted S, can be partitioned into at most n disjoint sets by considering the total number of ways a previously unallocated existing facility can enter the allocation scheme. (Here again n represents the total number of new facilities.) Suppose that in S the allocated existing facilities have been assigned to k new facilities where $k \leq n$. An unallocated existing facility is chosen.

If $k = n$, then S can be partitioned or separated into n subsets $S_1, S_2, \ldots, S_n$ where S_j is characterized by the assignment of the existing facility to new facility j. On the other hand, if $k < n$, then s may be partitioned into $k + 1$ subsets where S_j, $j=1, 2, \ldots, k$ is as described above. The subset S_{k+1} would be the assignment of the existing facility to a k+1th new facility. This k+1th new facility would have only one existing facility allocated to it.

The separation principle is illustrated below in Figure 4.1 for the case where $m = 4$ and $n = 3$. Each level i ($i = 1, \ldots, m$) represents the assignment of existing facility i to the new facilities. Each node is then labeled by the new facility that the existing facility is assigned to. At level i ($i = 2, \ldots, m$) all nodes emanating from the same node at level $i - 1$ represent the complete partition of their predecessor node. An allocation scheme is developed by choosing a node at the m^{th} level, tracing back to its predecessor node at the $m - 2$ level, etc. Even for this small problem, eighteen different allocation schemes were generated.

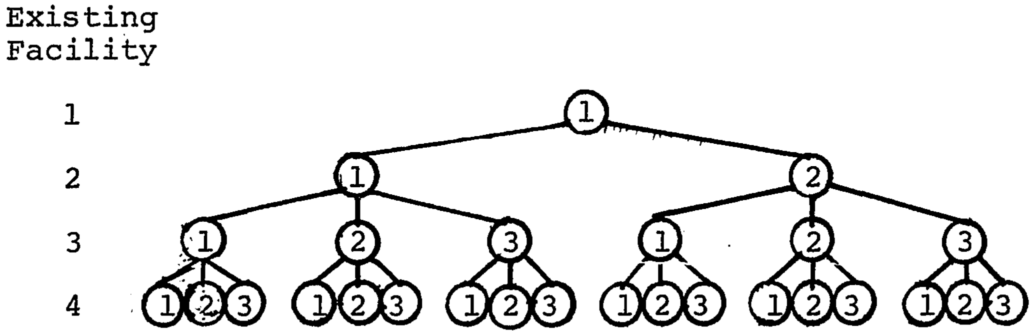


Figure 4.1. A complete enumeration of the ways 4 existing facilities can be allocated among 3 new facilities.

After a node or subset S has been partitioned, a lower bound is computed for each partition or succeeding node j. This bound is a lower bound on the objective function value that would be produced by any allocation scheme containing the allocations that have been made at this node j. The lower bound will be a sum of two values. The first value is the cost of optimally locating the new facilities among the existing facilities that have been allocated; this is just a multifacility location problem. The second value is just a lower bound on the cost of locating n new facilities among the unassigned existing facilities. If the sum of these two values is greater than the current upper bound on the optimal objective function value for the total L-A problem, then further consideration of the node j will be unprofitable. The partial allocation scheme it represents would never be part of the optimal allocation scheme, thus the node is fathomed or no longer considered as a candidate for partitioning.

If, on the other hand, the lower bound on node j is less than the current upper bound on the optimal objective function value, its partial allocation scheme is still a promising candidate as the optimal allocation scheme. The node is then partitioned and its successor nodes are placed under the same scrutiny for fathoming.

When the m^{th} level is reached a complete allocation scheme has been developed. Objective function values may be computed for each node at the m^{th} level. The node with the minimal objective function value that is still less than the current upper bound on the optimum represents the new best allocation scheme, and its objective function value becomes the new upper bound on optimal. Otherwise, these nodes are fathomed and nothing is changed. If all nodes have either been fathomed or partitioned, then the current upper bound is taken as the optimal objective function value.

There are two approaches to branch and bound: the depth approach and the breadth approach or "backtracking" and "jumptracking." The difference between the two approaches involves the criterion for selecting the node which is to be partitioned. The depth approach involves partitioning a node at level i. For its successor nodes at level i + 1, the lower bounds are computed and fathomed if possible. The unfathomed node with the minimal lower bound among these successors is chosen as the next node for partitioning. The other unfathomed nodes remain active. This procedure continues through the m^{th} level. The predecessor node at the m - 2 level is considered. If it has active nodes among its successors at the m - 1 level, then another attempt can be made to fathom these active nodes with the new upper bound. If active nodes still remain, the active successor with the minimal lower bound is chosen for partitioning. If no active nodes remain among the successors, the predecessor node at the m - 3 level is considered, and the process repeats.

The backtracking process terminates when the node at the first level has been considered as the predecessor node.

The breadth approach involves partitioning a node, computing the lower bounds, fathoming if possible, and then just labeling the successor nodes left as "active." The next node to be partitioned is chosen as the active node with the smallest lower bound among all active nodes in the tree. Thus, the selected node may be found at any level or may emanate from any node. Similarly, once a node at the m^{th} node is evaluated and a new upper bound is determined, all active nodes in the tree must be considered for fathoming. The next node for partitioning is then chosen as the active node with the smallest lower bound. This process continues until all nodes have been fathomed or separated.

The breadth approach has the advantage of potentially considering fewer nodes than the breadth approach because the opportunity to fathom is greater. However, the breadth approach requires greater storage on the computer since virtually all active nodes in the tree must be available at any time. The depth approach requires less storage since for any node, only its predecessor or its successor nodes need to be considered.

For the location-allocation model a depth approach is usually the better choice. Storage restrictions become a critical factor when the amount of storage required at each node is considered. The information to be stored would include information on the allocation of the existing facility chosen at this level, the locations of the new facilities, and the value of the lower bound. Thus, the L-A problem would be characterized as having many nodes since each partitioning of a node would produce, in general, n new nodes. The number of nodes and the storage requirement at each node indicate that storage is the crucial factor.

Secondly, a subsequent discussion on the lower bounding procedure to be used indicates that using the breadth approach could practically generate the whole tree before a feasible allocation scheme were generated. This would happen when the lower bounds on the nodes at level i increase significantly for the successor nodes at level i + 1. In a case like this it would be impossible to partition any node at level i + 1 until all nodes at level i had been partitioned.

4.7 A Branch and Bound Algorithm for the L-A Problem with Rectangular Regions and Rectilinear Distance

The preceding section discussed the use of branch and bound methods to get the optimal allocation scheme for a generalized L-A problem. Notice that the corresponding location problems are solved only to calculate lower and upper bounds. Thus, in the location-allocation system with rectangular regions, a deterministic approximation may perturb the lower or the upper bounds higher or lower so that it is possible that an optimal allocation scheme is fathomed.

The results of Chapter III indicate this problem is more pronounced with the rectilinear distance metric. Thus, the branch and bound algorithm will be developed for the L-A system with rectangular regions and a rectilinear distance metric.

4.7.1 The Branching Rule

The branching rule is the criterion used to choose the unallocated existing facility at each level whose assignment will be considered as the basis for making the partitions. Any rule may be used. For example, an unallocated existing facility could be chosen at random or the i^{th} existing facility could be chosen as the branching facility at the i^{th} level. However, an approach based on the properties of the problem may be more useful.

The rationale is that it may be best to first branch on the existing facilities that are farthest away from the current new facilities. These existing facilities will be the most troublesome so they are taken care of first. In this respect the new facilities will tentatively be located near the "remote" existing facilities during the first few iterations of the algorithm. As the clustered existing facilities are branched on, they will be allocated to the closest "remote" new facility and pull the location of the new facility toward them.

Kuenne and Soland (1972) tried two branching rules. The first was branching on the existing facility whose weighted distance to the closest new facility is a maximum. The second was branching on the existing facility whose weight times average distance is maximal. They found the latter case to work best.

For this problem where the sum of weighted expected distances is to be chosen, the expected distance from an existing facility to a new facility should be considered. Considering only the minimum distance or maximum distance between the existing facility and a new facility would disregard the size or variations of the locations contained in the existing facility.

The expected distance may be found by computing the following expression.

The expected distance between region i and new facility j is

$$\frac{w_i}{A_i} \int_{b_{i_1}}^{b_{i_2}} \int_{a_{i_1}}^{a_{i_2}} (|x_j-a_i| + |y_j-b_i|)\, da_i db_i \qquad (4.7.1)$$

where all parameters are defined as in P4.2.2.

This is equivalent to the following expression.

$$\frac{w_i}{a_{i_2}-a_{i_1}} \int_{a_{i_1}}^{a_{i_2}} |x_j-a_i| da_i + \frac{w_i}{b_{i_2}-b_{i_1}} \int_{b_{i_1}}^{b_{i_2}} |y_j-b_i| db_i \qquad (4.7.2)$$

Each expression in the sum may be computed independently.

Hence, there is an expected distance with respect to the x-coordinate and another with respect to the y-coordinate.

Theorem 4.7.1: If $x \notin (a_{i_1}, a_{i_2})$ and $y \notin (b_{i_1}, b_{i_2})$, then the expected distance from (x,y) to region i defined by $[a_{i_1}, a_{i_2}] \times [b_{i_1}, b_{i_2}]$ is equivalent to the rectilinear distance from (x,y) to the midpoint of these intervals $\left(\frac{a_{i_1}+a_{i_2}}{2}, \frac{b_{i_1}+b_{i_2}}{2}\right)$.

Proof: For the x coordinate, there are three cases to consider.

Case I. If $x \leq a_{i_1}$,

$$\text{then } \frac{1}{a_{i_2}-a_{i_1}} \int_{a_{i_1}}^{a_{i_2}} |x-a_i|\,da_i = \frac{(a_{i_2}-x)^2 - (a_{i_1}-x)^2}{2(a_{i_2}-a_{i_1})}$$

$$= \frac{(a_{i_2}-a_{i_1})(a_{i_2}+a_{i_1}-2x)}{2(a_{i_2}-a_{i_1})} = \frac{a_{i_2}+a_{i_1}}{2} - x. \qquad (4.7.3)$$

Case II. If $a_{i_1} < x < a_{i_2}$

$$\text{then } \frac{1}{a_{i_2}-a_{i_1}} \int_{a_{i_1}}^{a_{i_2}} |x-a_i|\,da_i = \frac{(a_{i_2}-x)^2 + (a_{i_1}-x)^2}{2(a_{i_2}-a_{i_1})}$$

$$= \frac{(a_{i_2}-a_{i_1})(a_{i_2}+a_{i_1}-2x) + 2x(x-2a_{i_1})}{2(a_{i_2}-a_{i_1})} = \frac{a_{i_2}+a_{i_1}}{2} - x$$

$$+ \frac{x^2-2a_{i_1}}{a_{i_2}-a_{i_1}} \;. \qquad (4.7.4)$$

Case III. If $a_{i_2} \geq x$

$$\text{then } \frac{1}{a_{i_2}-a_{i_1}} \int_{a_{i_1}}^{a_{i_2}} |x-a_i| da_i = \frac{(a_{i_1}-x)^2 - (a_{i_2}-x)^2}{2(a_{i_2}-a_{i_1})}$$

$$= \frac{(a_{i_2}-a_{i_1})(2x-a_{i_2}-a_{i_1})}{2(a_{i_2}-a_{i_1})} = x - \frac{(a_{i_2}+a_{i_1})}{2} . \qquad (4.7.5)$$

Thus, in Cases I and III only is the expected distance equivalent to $\left|\frac{a_{i_2}+a_{i_1}}{2} - x\right|$. The proof for the y coordinate is the same.

Theorem 4.7.1 serves two purposes. First, it presents simple computational expressions that can be used in both applying the branching rule and evaluating the objective function. Secondly, it indicates one source of deviation between the results of the L-A problem with rectangular regions and its deterministic equivalent.

4.7.2 Upper and Lower Bounds

4.7.2.1 Bounds on the Objective Function

When a branch and bound algorithm is initiated on a minimization problem, an upper bound on the objective function value should be developed so that it would be possible to fathom nodes before the computation of the first objective function value associated with an allocation scheme generated by the tree. In this case the objective function value associated with an arbitrary allocation scheme may serve as an upper bound. This upper bound may be improved by using a modification of Cooper's alternate location and allocation heuristic.

Consider the arbitrary allocation where existing facility i is allocated to new facility j where

$$j = \begin{cases} i \pmod{n} & \text{if } j \text{ is not divisible by } n \\ n & \text{otherwise} \end{cases} \tag{4.7.6}$$

By this definition existing facility n would be allocated to new facility n, but existing facility n + 1 would be allocated to new facility 1.

The location problem for this allocation is solved and the objective function value computed. This is an upper bound on the optimal solution. The upper bound could be tested for improvement by reallocating each existing facility to the new facility whose weighted expected distance from the former facility is a minimum. After the reallocations are made, the location problems are again solved and a new objective function value computed. If the new objective function is greater than the old objective function value, iterations cease. Otherwise, the reallocations start again. This heuristic may be iterated until no improvement is made or until a convergence criterion is met. The best objective function value from this heuristic becomes the upper bound on the optimal objective function value.

For a depth approach to branch and bound used to minimize an objective function value, a lower bound on the optimal is generally not needed. However, new developments in branch and bound such as the fictitious bound of Bazaraa and Elshafei (1977) require an initial lower bound. The fictitious lower bound is defined as

$$FUB = \alpha UB + (1-\alpha)\ LB \tag{4.7.6}$$

where: $0 \leq \alpha \leq 1$

UB = the upper bound

LB = the lower bound.

The fictitious lower bound may be used in place of the upper bound in order to fathom nodes quicker, thus hastening an in-depth tree search. In anticipation of future developments of this sort in branch and bound methods, a lower bound on the optimal will be developed.

An upper bound on n, the number of new facilities to be located is m, the number of existing facilities. If n = m, then a new facility could be located in each region to serve only that region. This poses the question of what is the minimum expected cost of serving a region.

Theorem 4.7.2: The minimum expected cost of serving region i from a point within i is $\frac{w_i}{4}(a_{i_2}-a_{i_1}+b_{i_2}-b_{i_1})$ where region i is defined as $[a_{i_1},a_{i_2}] \times [b_{i_1},b_{i_2}]$.

Proof: The expected cost of serving region i from (x,y) a point within i is

$$f(x,y) = w_i\left(\frac{(a_{i_2}-x)^2+(a_{i_1}-x)^2}{2(a_{i_2}-a_{i_1})} + \frac{(b_{i_2}-y)^2+(b_{i_1}-y)^2}{2(b_{i_2}-b_{i_1})}\right) \tag{4.7.7}$$

The partial derivatives of this expression are:

$$\frac{\partial f}{\partial x} = \frac{-2(a_{i_2}-x)-2(a_{i_1}-x)}{2(a_{i_2}-a_{i_1})} = \frac{2x-a_{i_2}-a_{i_1}}{a_{i_2}-a_{i_1}} \tag{4.7.8}$$

$$\frac{\partial f}{\partial y} = \frac{-2(b_{i_2}-y)-2(b_{i_1}-y)}{2(b_{i_2}-b_{i_1})} = \frac{2y-b_{i_2}-b_{i_1}}{b_{i_2}-b_{i_1}} \tag{4.7.9}$$

Setting expressions 4.7.8 and 4.7.9 equal to 0 and solving yields:

$$(x^*,y^*) = \left(\frac{a_{i_2}+a_{i_1}}{2}, \frac{b_{i_2}+b_{i_1}}{2}\right)$$

which is a minimum. Thus,

$$f\left(\frac{a_{i_2}+a_{i_1}}{2}, \frac{b_{i_2}+b_{i_1}}{2}\right) = \frac{w_i}{4}\,(a_{i_2}-a_{i_1}+b_{i_2}-b_{i_1}).$$

The lower bound may then be found by computing the expression:

$$\text{l.b.} = \sum_{i=1}^{m} T_i \tag{4.7.10}$$

where: $T_i = \frac{w_i}{4}\,(a_{i_2}-a_{i_1}+b_{i_2}-b_{i_1})$

4.7.2.2 A Lower Bound for Each Node

Computing a lower bound was discussed in Section 4.6 as a two-part process. The first part was solving the location problem for the allocated existing facilities and computing the corresponding existing facility. The second part involved underestimating the expected cost of locating the n new facilities among the unallocated existing facilities.

In order to develop the second expression consider two unallocated regions R_1 and R_2. Suppose that both are to be served by the same new facility $X = (x,y)$. The expected cost of serving these two regions is

$$f(X) = w_1E[|x-a_1|+|y-b_1|] + w_2E[|x-a_2|+|y-b_2|] \qquad (4.7.11)$$

where: (a_i, b_i) are random variables representing the points located in region i.

This expression can be considered the sum of the expected costs of serving the regions along the x-coordinate and the expected cost of serving the regions along the y-coordinate. These expressions are independent as indicated below.

$$f(x) = w_1E[|x-a_1|] + w_2E[|x-a_2|]$$
$$f(y) = w_1E[|y-b_1|] + w_2E[|y-b_2|] \qquad (4.7.12)$$

Thus, each one-dimensional case may be considered.

Notice that when the x-coordinate is considered

$$f(x) \geq \min\{w_1, w_2\}(E[|x-a_1|] + E[|x-a_2|]) \qquad (4.7.13)$$

Let a_1 and a_2 assume any values where $a_1 < a_2$ and consider the relative position of x. There are two ways to show $|x-a_1| + |x-a_2| \geq |a_1-a_2|$.

Case I. $a_1 \leq x \leq a_2$

a_1 x a_2 $\qquad |x-a_1| + |x-a_2| = |a_1-a_2|$

Case II. $x < a_1 < a_2$

x a_1 a_2 $\qquad |x-a_1| + |x-a_2| > |a_1-a_2|$

Case III. $a_1 < a_2 < x$

a_1 a_2 x $\qquad |x-a_1| + |x-a_2| > |a_1-a_2|$

Also, by the triangle inequality

$$|a_1-x_1| + |x-a_2| \geq |a_1-a_2| \qquad (4.7.14)$$

Since a_1 and a_2 are random variables, then

$$E[|x-a_1|] + E[|x-a_2|] \geq E[|a_1-a_2|] \qquad (4.7.15)$$

Substituting 4.7.15 into 4.7.13, a lower bound is produced.

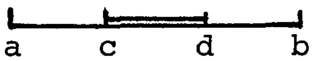
$$f(x) \geq \min\{w_1, w_2\} E[|a_1-a_2|] \qquad (4.7.16)$$

Thus, the expression given in 4.7.16 is an appropriate lower bound where $E[|a_1-a_2|]$ represents the expected distance between regions 1 and 2 along the x-coordinate

$$E[|a_1-a_2|] = \int_{a_{1_1}}^{a_{1_2}} \int_{a_{2_1}}^{a_{2_2}} |u-v| \, du dv. \qquad (4.7.17)$$

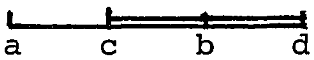
This expression may be evaluated for three cases. For ease in reading, let a represent a_{1_1}; b represent a_{1_2}; c represent a_{2_1}; and d represent a_{2_2} (The second interval a_2 is underlined).

Case I. $a < c < d < b$



$$E[|a_1-a_2|] = \frac{(a^2+b^2)(d-c) - (a+b)(d^2-c^2) + \frac{2}{3}(d^3-c^3)}{2(b-a)(d-c)} \qquad (4.7.18)$$

Case II. $a < c < b < d$



$$E[|a_1-a_2|]$$

$$= \frac{(b-c)[(a^2+c^2)-(a+b)(b+c)]+\frac{2}{3}(b^3-c^3)+(d-b)(b-a)(d-c)}{2(b-a)(d-c)} \qquad (4.7.19)$$

Case III. $a < b < c < d$

$$E[|a_1-a_2|] = \frac{(d^2-c^2)(b-a) - (d-c)(b^2-a^2)}{2(b-a)(d-c)} \qquad (4.7.20)$$

Corollary 4.7.1: If the two regions R_1 and R_2 have nonoverlapping x intervals, then the expected distance between the two is just $|\frac{d+c-b-a}{2}|$.

Proof: This can be proved in two ways. One is to simplify 4.7.20--the desired result will be obtained.

The other is to consider 4.7.15 and use the independence of the regions.

$$E[|a_1-a_2|] = |E[a_2-a_1]| = |E[a_2]-E[a_1]| = |\frac{d+c}{2} - \frac{b+a}{2}|$$

Thus, for any two rectangular regions R_i and R_j, the expression

$$\min\{w_i,w_j\}(E[|a_i-a_j|+|b_i-b_j|). \qquad (4.7.21)$$

can be computed as an underestimate of the expected cost of serving these two regions with the same new facility.

Directly applying a result of Cooper (1964), it can be shown that a lower bound on locating one new facility among s regions is

$$\frac{1}{s-1} \sum_i \sum_j \min\{w_i,w_j\}(E[|a_i-a_j|+|b_i-b_j|) \qquad (4.7.22)$$

Thus, the expression 4.7.21 is the building block for forming lower bounds. If there are p unallocated existing facilities, then there are $\frac{1}{2}p(p-1)$ different realizations of expression 4.7.21. Assume that all the expressions are placed in ascending order.

Let q_i be the i^{th} term in this progression where $i = 1, \ldots, \frac{1}{2}p(p-1)$. Next consider the expression 4.7.10 labeled L_i. Compute L_j for $j = 1, \ldots, p$ and arrange these expressions in ascending order.

Let r_i be the i^{th} term in this progression.

To underestimate the expected cost of allocating n new facilities among p existing facilities, the various combinations of allocations should be studied. For example, if $p \leq n$, then a new facility could be assigned to each of the p existing facilities. An underestimate of this cost would be the sum of all p of the r_i terms. This would follow since r_i represents a minimum expected cost for serving a region from a point in the region.

Another example is the case where $p = n + 4$. In this case, there are five possible combinations: four new facilities are allocated two existing facilities, all others are allocated one; one new facility is allocated three existing facilities, two are allocated two, and the others are allocated one; one new facility is allocated four existing facilities, one is allocated two, and all others are allocated one; two new facilities are allocated three existing facilities apiece, and all others are allocated one; and finally one new facility is allocated five existing facilities, and all others are allocated one.

Table 4.1 displays all lower bounds for these combinations using (4.7.22) and the definitions of the q_i's and the r_i's for different values of p-n.

It is obvious that as p-n becomes larger than five, the number of combinations to be considered also becomes large. Thus, a general lower bound will be used for values of p-n greater than five.

TABLE 4.1

LOWER BOUNDS FOR LOCATING N NEW FACILITIES AMONG P RECTANGULAR REGIONS

Value of P - N	Lower Bound
0 or less	$\sum_{i=1}^{m} r_i$
1	$q_1 + \sum_{i=1}^{n-1} r_i$
2	$\min\{q_1+q_2+\sum_{i=1}^{n-2} r_i, \frac{1}{2}(q_1+q_2+q_3)+\sum_{i=1}^{n-1} r_i\}$
3	$\min\{q_1+q_2+q_3+\sum_{i=1}^{n-3} r_i, q_i+\frac{1}{2}(q_2+q_3+q_4)+\sum_{i=1}^{n-2} r_i,$ $\frac{1}{3}(q_1+...+q_6)+\sum_{i=1}^{n-1} r_i\}$
4	$\min\{q_1+...q_4+\sum_{i=1}^{n-4} r_i, q_1+q_2+\frac{1}{2}(q_3+q_4+q_5)+\sum_{i=1}^{n-3} r_i,$ $q_1+\frac{1}{3}(q_2+...+q_7)+\sum_{i=1}^{n-2} r_i, \frac{1}{2}(q_1+...+q_6)+\sum_{i=1}^{n-2} r_i,$ $\frac{1}{4}(q_1+...+q_{10})+\sum_{i=1}^{n-1} r_i\}$
5	$\min\{q_1+...+q_5+\sum_{i=1}^{n-5} r_i, q_1+q_2+q_3+\frac{1}{2}(q_4+q_5+q_6)+\sum_{i=1}^{n-4} r_i,$ $q_1+q_2+\frac{1}{3}(q_3+...+q_8)+\sum_{i=1}^{n-3} r_i, q_1+\frac{1}{2}(q_2+...+q_7)+\sum_{i+1}^{n-3} r_i,$ $q_1+\frac{1}{4}(q_2+...+q_{11})+\sum_{i=1}^{n-2} r_i, \frac{1}{2}(q_1+q_2+q_3)+\frac{1}{3}(q_4+...+q_9)$ $+\sum_{i=1}^{n-2} r_i, \frac{1}{5}(q_1+...+q_{15})+\sum_{i=1}^{n-1} r_i\}$

Theorem 4.7.3: A general lower bound on locating n new facilities among p rectangular regions is $\frac{1}{2} \sum_{i=1}^{p-n} q_i$ where q_i is as defined above.

Proof: To prove the Theorem, four cases are considered.

Case I. No new facility has more than two existing facilities allocated to it. Here $2n \leq p$. So $(p-n)$ new facilities are allocated two existing facilities, and $2n-p$ new facilities are allocated one.

The lower bound is

$$\sum_{i=1}^{p-n} q_i + \sum_{i=1}^{2n-p} r_i > \frac{1}{2} \sum_{i=1}^{p-n} q_i .$$

Case II. No new facility has more than three existing facilities allocated to it.

Let k new facilities have three regions allocated to each of them. Then p-n-2k new facilities will have no more than two regions allocated to them.

A lower bound is

$$\sum_{i=1}^{p-n-2k} q_i + \frac{1}{2} \sum_{i=p-n-2k+1}^{p-n+k} q_i > \frac{1}{2} \sum_{i=1}^{p-n+k} q_i$$

Since $k \geq 1$

$$\frac{1}{2} \sum_{i=1}^{p-n+k} q_i - \frac{1}{2} \sum_{i=1}^{p-n+1} q_i > \frac{1}{2} \sum_{i=1}^{p-n} q_i$$

Case III. At least one new facility has more than three regions allocated to it.

If a new facility has four or more regions allocated to it, a lower bound is

$$\frac{1}{3}[q_1+\ldots+q_6] = \frac{1}{3}[q_1+q_2+q_3]+\frac{1}{3}[q_4+q_5+q_6]$$

$$\geq \frac{1}{2}[q_1+q_2]+\frac{1}{2}[q_4+q_5] \geq \frac{1}{2}[q_1+q_2+q_3+q_4]$$

A new facility with three regions allocated to it has a lower bound of

$$\frac{1}{2}[q_1+q_2+q_3].$$

A new facility with two regions allocated to it has a lower bound of

$$q_1.$$

Thus, if every two-region new facility is paired with a three-region new facility, their combined lower bound is

$$q_1 + \frac{1}{2}[q_2+q_3+q_4] > \frac{1}{2}[q_1+q_2+q_3+q_4].$$

This is the same as in the four region new facility. Hence, the new facilities with two and three regions, respectively, may be paired to create a situation equivalent to one new facility with four regions. This may be done $[\frac{p-n}{3}]$ times where $[\cdot]$ is the greatest integer function. A lower bound is

$$\frac{1}{2}\sum_{i=1}^{4[\frac{p-n}{3}]} q_i > \frac{1}{2}\sum_{i=1}^{p-n} q_i \quad \text{since } p-n \text{ is at least } 3.$$

Case IV. All new facilities have more than three regions allocated to them.

Obviously $p > 4n$.

A lower bound for one new facility with at least 4 regions as computed in Case III is

$$\tfrac{1}{2} \sum_{i=1}^{4} q_i$$

For all n new facilities, the lower bound is

$$\tfrac{1}{2} \sum_{i=1}^{4n} q_i > \tfrac{1}{2} \sum_{i=1}^{p-n} q_i .$$

Thus, $\tfrac{1}{2} \sum_{i=1}^{p-n} q_i$ is a general lower bound with the property that no considerations of combinations has to be made. This lower bound is well-suited for the cases when p-n is large. These cases will be levels 1,2, ..., m-5 of the tree. At these levels, the possibility of fathoming nodes is not as great as the other levels. This is because only a few existing facilities have been allocated, and the partial objective function value used in computing the lower bound will be far from the optimum. A tight lower bound would then involve considering all possible combinations of the unallocated facilities. To hasten the tree search, the general lower bound is used to quickly compute the lower bound and move to the next level.

On the other hand, in the last n+5 levels of the tree enough facilities have been allocated to identify unprofitable allocation schemes. Here the tighter lower bounds given in Table 4.1 should be used to fathom as many nodes as possible.

4.7.3 The Location-Allocation Branch-and-Bound Algorithm (LABB)

In this section the complete branch and bound algorithm for the location-allocation problem is given.

The input parameters are

N = number of new facilities

M = number of existing regions

x1(I) and x2(I) = left and right endpoints, respectively, of region I (R(I)) along x-axis.

y1(I) and y2(I) = lower and upper endpoints, respectively, along y-axis

w(I) = interaction cost for region I.

The parameters for computing bounds on the optimum value of the objective function are:

$\bar{z}$ = upper bound on optimum

$\underline{z}$ = lower bound on optimum

FX = current least upper bound on optimum

F = objective function value to be compared with Fx

ε = stopping criterion for alternate heuristic ($\varepsilon>0$)

The parameters for computing the branching facility are:

L = current level

J_L = index of branching facility chosen at level L

IJ_L = set of indices of unallocated facilities at level L

AED(I) = vector of average expected distances from region I to all other regions

AX(I) = vector of average distance of region I to the new facilities that have been currently located

The parameters for creating and fathoming new nodes are:

KL = number of new nodes to be created at current level

NODE = counter for nodes created

ND = node number of the last node created at previous level

IP(L) = the new facility the branching facility at level L was allocated to according to the node that was partitioned at level L

NL = number of new facilities at previous level

XX(J) = current location of new facility J

XLB(I) = lower bound at node I

Q(I) = the i^{th} smallest value of $\min\{w(j), w(k)\}\, E[|R(j)-R(k)|]$ for all $j<k$

R(I) = the i^{th} smallest value of $.25w(j)[x2(j) - x1(j) + y2(j) - y1(j)]$

Step 0. Initialize the input parameters. (Compute upper and lower bounds on optimum.)

Step 1. Let FX = ∞.

Step 2. Arbitrarily allocate region I to new facility $I\text{-}N\left(\frac{I-1}{N}\right)$.

Step 3. Solve the single facility location problem for each new facility XX(j), j = 1, ..., n among the regions allocated to new facility j.

Step 4. Evaluate F, the objective function value of the L-A problem, for the results of Step 3.

Step 5. If $FX-F > \varepsilon$, then replace FX with F. Otherwise, go to 7.

Step 6. For I = 1, ..., M, compute

$$k = \min_j\{w(I)\cdot E[|XX(j)-R(I)|].$$

Reallocate region I to new facility k. Go to 3.

Step 7. Let $\bar{z} = FX$.

Step 8. Compute $\underline{z} = .25 \sum_{I=1}^{M} w(I)\cdot[x2(I) - x1(I) + y2(I) - y1(I)]$

Step 9. If $\underline{z} = \bar{z}$, stop. Go to 31.

(Initialize for level 1)

Step 10. L = 1.

Step 11. For I = 1, ..., M, compute $AED(I) = \sum_{k=1}^{m} E[|R(I)-R(k)|]$

Step 12. Let $j_1 = \max_I AED(I)$. $IJ_1 = \{1,2,\ldots,M\} - j_1$

Step 13. Let NODE = 1. Assign region j_1 to new facility 1. Solve the location problem for xx(1). Let NL = 1. Let IP(1) = 1.

(Advance to next level)

Step 14. Let L = L + 1.

(Compute Branching Facility)

Step 15. Compute $AX(1) = \frac{1}{NL} \sum_{II=1}^{NL} E[|R(I) - XX(II)]$ for $I\varepsilon IJ_{L-1}$.

Step 16. Let $j_L = \max_I AX(I)$ and let $IJ_L = IJ_{L-1} - j_L$.

(Create New Nodes)

Step 17. Let KL = min(L,N). Let ND = NODE.

Step 18. Create KL new nodes ND+1, ..., ND+KL by allocation region j_L to new facility 1, ..., KL, respectively.

Let NODE = ND + KL.

(Compute Lower Bounds on Nodes)

Step 19. For node I = ND+1, ..., ND+KL, solve the location problem for the partial allocation scheme: region j_L allocated to new facility I-ND; j_k allocated to IP(k), k = L-1,L-2, ..., 1. Denote the objective function value LB(I).

Step 20. Compute the vectors Q(I) and R(I) using regions J, $J \varepsilon IJ_L$.

Step 21. If M-L-N > 5, let $Qx = \frac{1}{2} \sum_{I=1}^{M-L-N} Q(I)$.

If M-L-N ≤ 5, compute the lower bound for the value M-L-N as given in Table 3. Denote this value Qx.

Step 22. Let LB(I) = LB(I) + Qx I = ND+1, ..., ND+k.

If LB(I) ≥ $\bar{z}$, fathom node I.

Step 23. Among the unfathomed nodes in 22, choose I* as the value of I such that $LB(I^*) = \min_I LB(I)$.

If all nodes are fathomed, go to 27.

Step 24. Let IP(L) = I*-ND.

Step 25. If L<M, Set NL and XX(j) j = 1, ..., NL equal to the values found for I* in Step 19 and go to 14.

Step 26. If L=M, compare LB(I*) to $\bar{z}$.

If LB(I*) < $\bar{z}$ then $\bar{z}$ = LB(I*).

Fathom the newly created nodes at level M.

(Backtracking Procedure)

Step 27. Let L = L-1. If L=1, stop. Go to 31.

Step 28. Consider all nodes I at level L that are unfathomed and have not been partitioned such that their allocation scheme includes j_{L-1} allocated to IP(L-1). If there is a node I such that LB(I)< $\bar{z}$, go to 29. Otherwise go to 27.

Step 29. Choose I* such that LB(I*) = $\min_I$ LB(I) where I are the active nodes identified in 27. Let LL denote the new facility j_L was allocated to at I*. IP(L) = LL. Let NL and XX(j) become the appropriate values found in Step 19 for I*.

30. Go to Step 14.

31. The optimal allocation scheme is the one associated with $\bar{z}$, the optimal objective function value.

4.7.4 Verification of the Algorithm

The algorithm of Section 4.7.3, LABB, was coded in Fortran IV. The code was verified using Problem A.3 in Appendix A. A graph of this problem is presented in Figure 4.2.

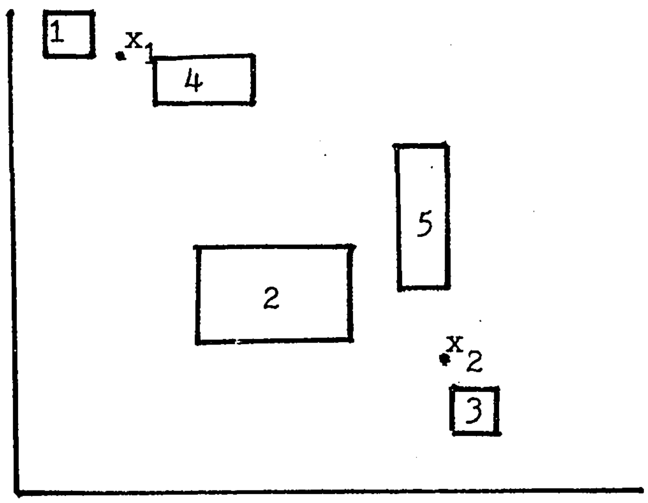


Figure 4.2. A graph of Problem A.3.

Both manual computation and the code produced the optimal allocation scheme to be: x_1 serves regions 1 and 4 and x_2 serves regions 2, 3, and 5. The two new facilities x_1 and x_2 were located at (2.5,9) and (9.5,1.5), respectively. The optimal objective function value was 18.5.

The same problem with a centroid approximation produced a different allocation scheme. It was x1 serves regions 1, 2 and 4 and x2 serves regions 3 and 5.

The new facilities were located at the points (4,8.5) and (9.5,1.5), the centroids of regions 4 and 3, respectively. These locations used in the objective function involving the rectangular regions produced a value of 37, a 100 per cent increase over the optimal locations.

Thus, in a simple test example the algorithm was verified. The impact of the sensitivity of the rectilinear distance metric to the centroid approach on the location-allocation problem is serious; it has produced a non-optimal allocation scheme and inferior locations for the new facilities. As in the multifacility location problem, the centroid approach does not even offer a good approximation to the solution of the location-allocation model.

4.7.5 Computational Results

The computational results given in this section represent experience with the branch and bound algorithm (LABB) for rectangular regions using a rectilinear distance metric. The problems were randomly generated from uniform distributions.

All w(I)'s were generated from a uniform [0,10] distribution. The x1(I)'s, x2(I)'s, y1(I)'s, and y2(I)'s were each generated from a uniform [0,100] distribution. All problems were run on an IBM 370/158J computer. The results are summarized in Table 4.2.

Not surprisingly, the required computational times reflects the number of nodes created which, in turn, is a function of the size of the problem and the number of active nodes.

As depicted in Figure 4.3 for $m \geq 11$, the case $n = 2$ is the critical case. Table 4.2 also affirms that for the larger values of m, it is faster to locate three or four new facilities than it is to locate two as was expected.

These results are compared to the results of Ostresh in Ruston et al. (1973). Using a branch and bound procedure for the deterministic Euclidean distance problem, the limits on the problems he could solve with an IBM 360/65 were:

$n = 2$ $m = 23$ CPU time = 23.26 sec.

$n = 3$ $m = 17$ CPU time = 78.01 sec.

$n = 4$ $m = 11$ CPU time = 10.28 sec.

Kuenne and Soland (1972) using a similar branch and bound algorithm worked problems of size $n = 4$ $m = 15$ in average times of 82.7 seconds and 54.2 seconds for problems with random and unit weights, respectively, on an IBM 360/91.

Cooper (1967) tested his heuristics by working problems of the size $n = 3$ $m = 40$ in times ranging from

TABLE 4.2

COMPUTATIONAL RESULTS FOR LOCATION-ALLOCATION PROBLEMS WHERE N = 2, 3, AND 4

M	No. of Problems	CPU Time (sec.)	No. of Nodes Created	Max. No. of Active Nodes	Optimal Node
N=2					
5	2	.585	11	1	10
6	3	1.01	27.3	2.7	19.7
7	4	1.00	25.75	3.5	20
9	3	1.99	94.33	4.67	79
11	3	5.37	240.3	8	119
15	2	8.55	330	11	219
20	2	11.33	332	18	39
25	2	19.08	349	23	69
30	1	33.02	517	28	59
35	1	51.15	541	33	69
N=3					
6	3	1.2	50	3.3	32
7	3	1.54	86	5.33	68
9	3	3.76	232	10.33	59
11	2	4.51	272	14	211.5
15	2	7.28	412.5	23	188
20	3	11.2	411	35.7	62.3
25	2	15.44	-	45	72
30	2	26.37	564	55	87
35	1	37.23	543	65	351
N=4					
7	3	2.73	129.33	7.3	87.33
9	3	3.01	223	11.3	118
11	2	4.8	316	23	38
15	2	6.73	416	36	54
20	4	11.77	586	48	74
25	1	13.26	458	66	94

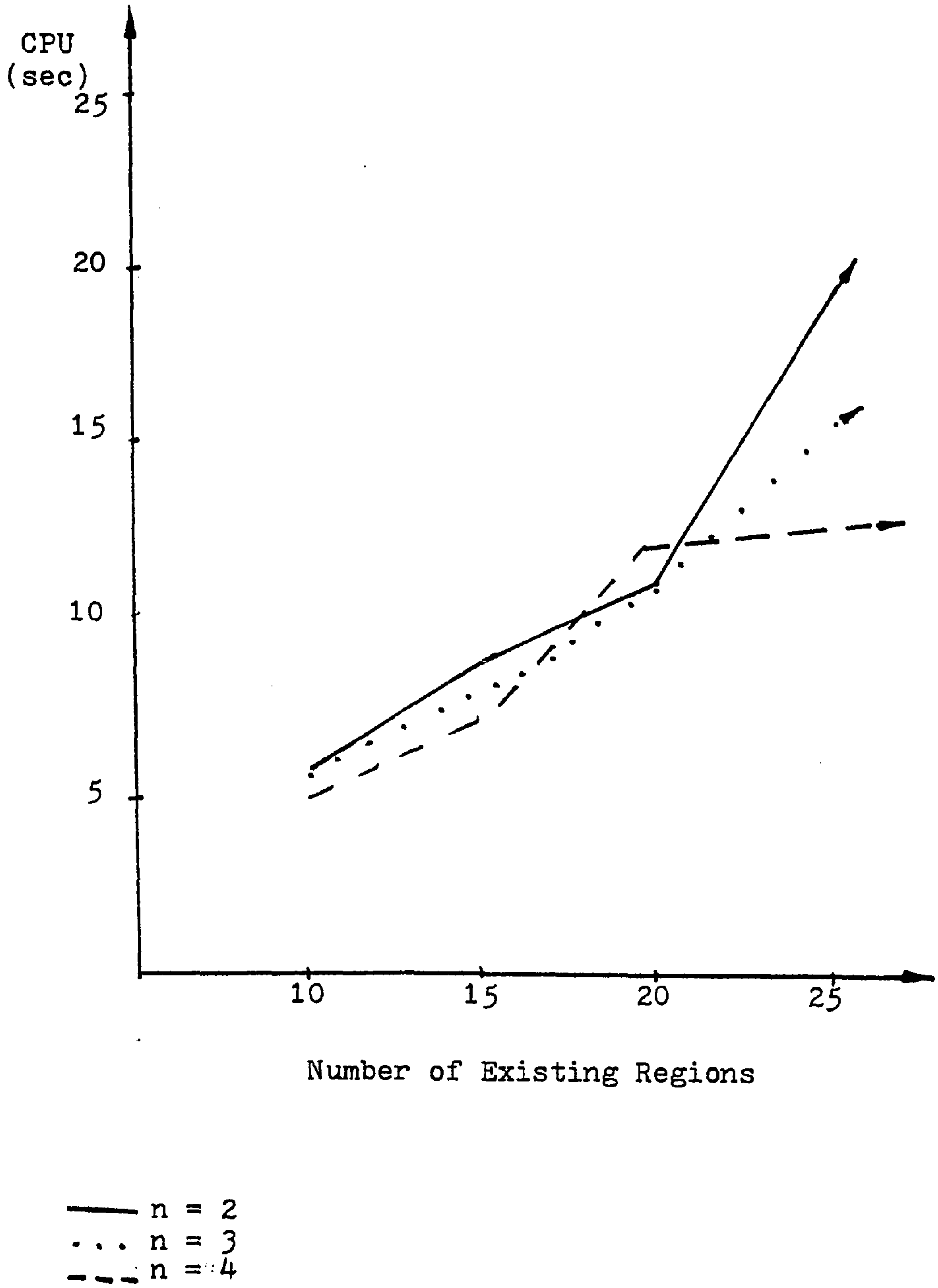


Figure 4.3. A plot of computational times for selected problems.

two minutes to 4.5 minutes on an IBM 7072.

The significance of the results is that location-allocation problems larger than those that have been solved before can be optimally solved in faster computational times. For the problems worked no computational time was over a minute. Further computational results of the algorithm will be given in the next section.

Another computational aspect is the performance of the lower bound. In each of the cases of m for n = 2, the optimal allocation was examined to determine what percentage of optimum was achieved by the lower bound at each level. In cases where m was large, the general lower bound was used at the first m - 6 levels. The lower bound improved rapidly from level to level; a typical improvement was ten per cent of optimum. Usually at the $m - 6^{th}$ level, the lower bound was within 85-90 percent of optimum. Thus, the switch to the combinatorial lower bounds for the last five levels represented less improvement from level to level, but convergence occurred rapidly.

On the smaller problems where the combinatorial lower bounds were used through most of the algorithm, the lower bounds increased erratically from level to level, and the convergence of the lower bound during the last several levels was not as marked.

To illustrate these results the average percentage of optimum was computed at levels 2-7 for problems where m = 7, and the same was done for all levels where m = 15. The results are plotted in Figure 4.4. In the case m = 15, the level where the combinatorial lower bound was first applied is circled.

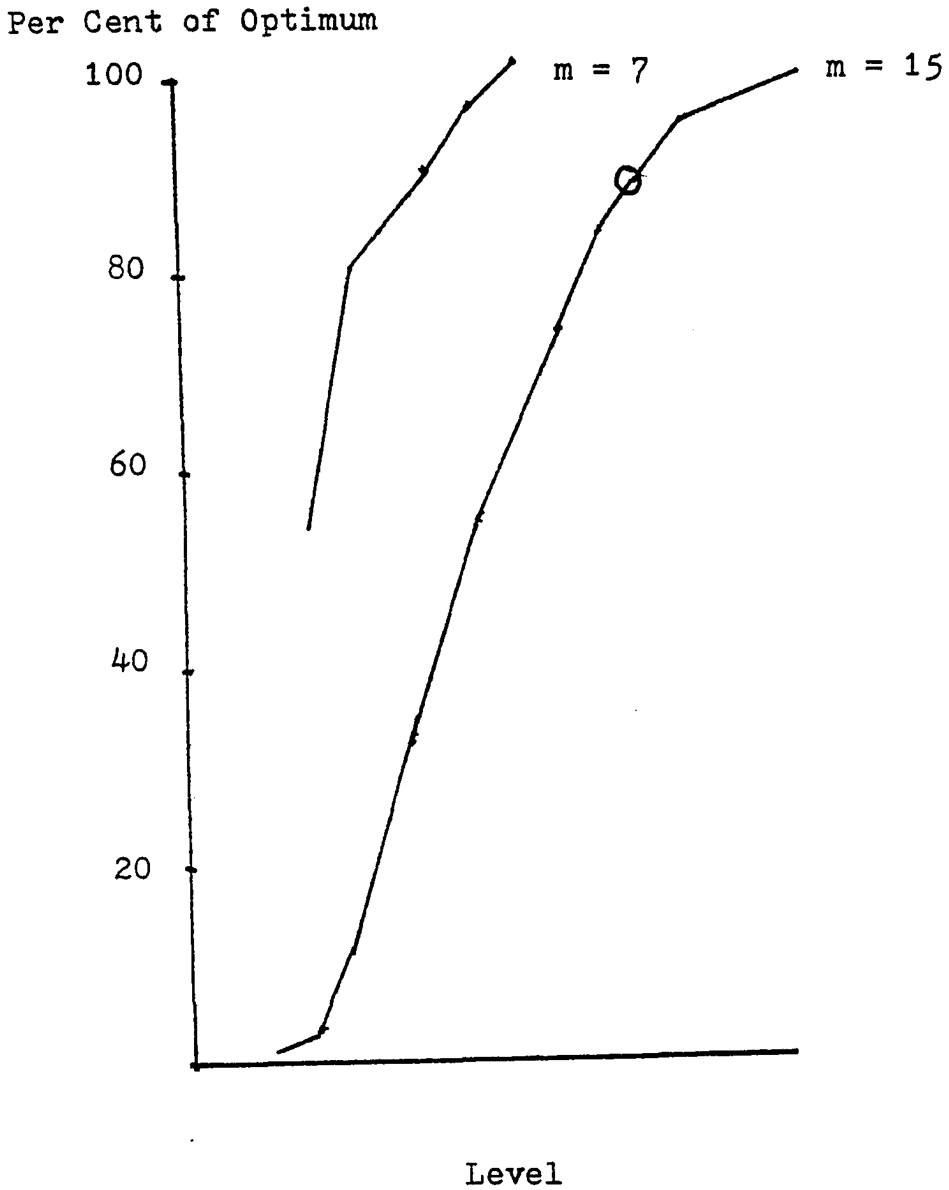


Figure 4.4. A plot of the lower bound at each level as a per cent of optimum.

4.8 Impact of the LABB Algorithm for Other Location Problems

The LABB algorithm is significant not only in the respect that it optimally solves the L-A problem for rectangular regions with rectilinear distance but also in its impact on similar deterministic problems and on the L-A problem with Euclidean distance. In this section the various impacts will be discussed.

4.8.1 The Impact on the Deterministic Version of the Problem

In considering the deterministic version of the L-A problem as formulated in Chapter I, it is immediately obvious that a branch and bound algorithm similar to the LABB could be developed. However, the deterministic case is just the limiting case of the probabilistic case as the areas of the rectangular regions become arbitrarily small. Thus, a new algorithm need not be developed; the same LABB algorithm, with minor modifications to the code, will handle the case of the rectangular region with area $\varepsilon > 0$.

Computational aspects of this adaptation of the algorithm were developed for three sets of randomly generated problems. Three values of ε were assumed--$(.025)^2$, $(.01)^2$, and $(.0001)^2$ for each problem of size (m,n). For all problems n was assumed to be two; this was because the results for the large rectangular regions indicated that the value n = 2 was the critical value of n.

Not surprisingly, the size of the tree search, or the number of nodes created was virtually the same for problems with small regions as opposed to problems with large regions. The computational times were different as displayed in Figure 4.5.

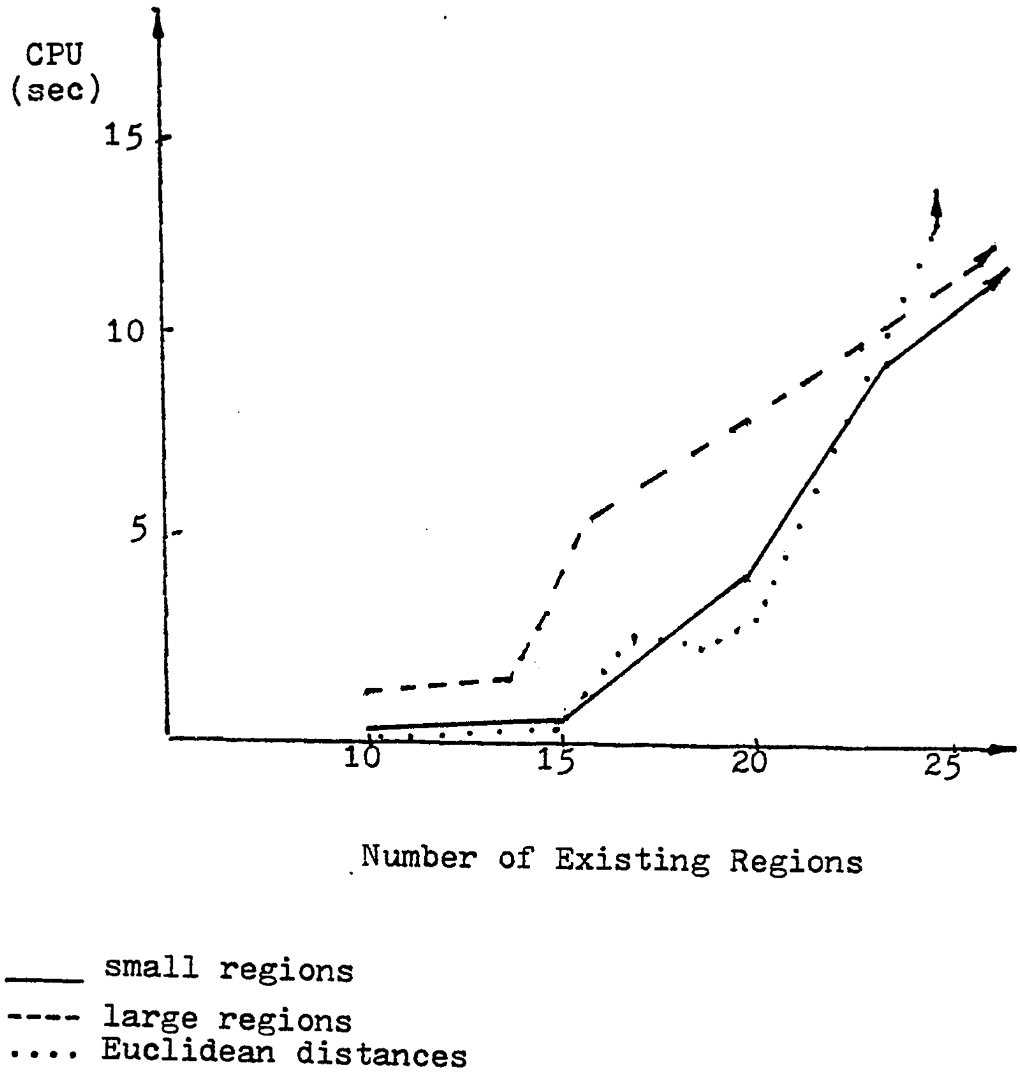


Figure 4.5. A plot of average computational times for a two facility L-A problem under three different assumptions.

The results for all three sets of problems were averaged, and some are presented below in Table 4.3 in comparison with the results of Love and Morris (1975) which were discussed in Section 4.3. Their results also reflect an average of randomly generated deterministic problems; they used a Univac 1110 computer.

TABLE 4.3

A COMPARISON OF COMPUTATIONAL TIMES (IN SECONDS) FOR TWO EXACT ALGORITHMS IN SOLVING A TWO FACILITY L-A PROBLEM

M =	14	15	16	18	20	25	30	35
LABB		4.29		9.6	10.9	17.64	28.09	44.5
Love and Morris	3.96		17.4	60	77	577	1995	5483

These results indicate a rather dramatic difference in results for values of m ≥ 18. The time requirement for the Love and Morris algorithm is an infeasible 1½ hours compared to a time of 44 seconds for the branch and bound method presented here. Thus, LABB algorithm performs better computationally than other existing exact algorithms.

Figure 4.5 displays the computational times for the small region problems versus the corresponding times for the large region problems. The computational results for the Kuenne-Soland algorithm for a deterministic problem with

Euclidean distances as programmed in the monograph of Rushton et al. (1973) are also presented.

An obvious result is that the algorithm is faster for small regions or deterministic problems than it is for large regions. This is also borne out for the cases of m = 25, 30, and 35 where the algorithm is about six seconds faster for small regions.

Thus, the LABB algorithm is computationally more efficient compared to the other existing exact algorithm in the deterministic case with rectilinear distance.

4.8.2 The Impact on Euclidean Distance L-A Problems

In Chapter II the relative insensitivities of the probabilistic location problem with Euclidean distance to the centroid approach were discussed. This result may be capitalized on in reference to the location-allocation problem for rectangular regions using a Euclidean metric. The same basic algorithm as in Section 4.7 may be used; however, the location problems may be solved with the centroid approach. Naturally, the objective function of the location problems would be evaluated with respect to the rectangular regions.

Kuenne and Soland who developed the branch and bound method for Euclidean distance deterministic problems noticed this insensitivity and used tentative new facility locations at several levels in order to avoid having to solve location problems at each node. The optimal allocation scheme was unaffected by the tentative new facility locations. Thus, it is reasonable to expect that the probabilistic problem would behave similarly to near-optimal facility locations.

It would be expected that the computational results for this modified branch and bound algorithm would be similar to the results for the deterministic problem based on the results of the probabilistic and deterministic rectilinear distance problems.

It has already been observed that the size of the largest deterministic Euclidean distance L-A problems that have been solved is far less than the sizes of the largest probabilistic rectilinear problems. The large computational times for the former problem are the prohibitive factor. The large computational times are the result of larger tree searches necessary to solve the Euclidean distance problem than that necessary to solve the rectilinear problem.

This phenomenon is best illustrated in the comparison of the results of branch and bound algorithm for rectilinear distance with the results of Kuenne and Soland (1972) in solving problems 1-8 presented in Cooper (1963). These problems are standards for testing location-allocation problems. Cooper tested his heuristics on these problems although he never reported any computational times. Others like Love and Morris (1975) also solved them. Some have reported their results in numbers of iterations, but Love and Morris did not do this nor did they report any times;

hence, a direct comparison is impossible.

Table 4.4 presents a comparison of the average results of the LABB algorithm for rectilinear distance with the results of Kuenne and Soland for Euclidean distances for the Cooper problems with unit weights.

TABLE 4.4

A COMPARISON BETWEEN THE RESULTS OF BRANCH AND BOUND ALGORITHMS FOR EUCLIDEAN AND RECTILINEAR DISTANCE METRICS

Metric	CPU Time (sec.)	No. of Nodes Created	Max. No. of Active Nodes
Euclidean	.5	28.3	6.1
Rectilinear	.51	19	4.1

Table 4.4 indicates that although the computational times are virtually the same, the tree search for the rectilinear metric is more efficient. As the Cooper problems are relatively small--m = 7, n = 2, the larger tree search is probably the critical factor for rendering the Euclidean problem relatively unsolvable in relation to the rectilinear problem as m and n become larger.

In order to help reduce the tree search, a large Euclidean distance L-A problem could first be solved using the rectilinear metric. Since rectilinear distance is an upper bound on Euclidean distance, the optimal objective function value from the rectilinear distance problem could be used as the initial upper bound on optimal for the Euclidean distance problem.

4.8.3 The Impact for Other Versions of the L-A Problem

The rectilinear branch and bound algorithm for the L-A problem has already shown itself to be versatile in several different situations. In this section other adaptations of the problem will be discussed.

4.8.3.1 Constrained Problems

A L-A problem may have constraints on the allocation scheme, on the locations of the new facilities, or on both. Constraints on the allocation scheme may include restrictions on the existing facilities that may be allocated to a new facility such as restrictions on the number of facilities allocated to a new facility or restrictions. In these cases, the constraints can be used as an additional test at each node as a basis for fathoming the node.

Constraints on the locations of the new facilities may be handled in the same manner that Love (1972) handled them. Any efficient nonlinear programming algorithm may be used to incorporate the constraints into the objective function. Thus, the objective function value of any allocation scheme violating the constraints will include a large penalty to warrant it fathomable.

4.8.3.2 Interfacility Weights among New Facilities

Although the general L-A problem assumes that all interfacility weights between new facilities, the v_{jk}'s, are zero, a problem with positive v_{jk} could be incorporated into the LABB algorithm. Since locations of new facilities may vary widely from level to level, consideration of the v_{jk} terms should be deferred until the m^{th} level, where the generated allocation scheme for the m facilities may be used to solve the multifacility location problem as described in Chapter III. The objective function value of the multifacility location problem is then tested against $\overline{z}$, the current upper bound. Using the interfacility weight at a level other than the m^{th} may result in the fathoming of a potentially optimal allocation on the basis of tentative locations. Thus, at these levels the case $v_{jk} = 0$, for all j and k, should be used as a lower bound on the case $v_{jk} > 0$.

4.8.3.3 Other Distributions

The branch and bound methods may be applied to location-allocation problems with distributions on existing facilities other than uniform. Since the branch and bound methods generate optimal allocation schemes no matter what type of objective function is used, the only difference would be the way the location problems would be solved at each node. Solution techniques were developed by Aly (1975) for existing facilities with a bivariate normal distribution in both the cases of a Euclidean and a rectilinear metric. Katz and Cooper (1976) and Wesolowsky (1977) developed solution techniques for the bivariate symmetric exponential distribution for the cases of Euclidean distance and rectilinear distance, respectively. It would be expected that the computational times to solve these related problems would be similar to the times for the uniform distribution with adjustments made on the basis of the speed of the individual solution technique.

4.9 Summary

In this chapter the location-allocation problem for existing facilities uniformly distributed over rectangular regions was considered. Previous works dealing with L-A systems were discussed, and the properties of the problem were developed. These results indicated that developing the optimal allocation scheme was the most important step in optimally solving the problem.

Because the number of possible allocations is so large, heuristic algorithms have been a common solution technique. The heuristics indicated that there was often a tradeoff between computational efficiency and near-otpimality. Heuristic algorithms locate local minima which often deviate significantly from global algorithms.

In order to implicitly enumerate all possible algorithms, a branch and bound algorithm was developed. It was initially developed for the rectilinear distance metric because the location problem under this metric is very sensitive to the assumption of rectangular regions. An example problem was solved to illustrate this sensitivity.

Computational results indicated that the algorithm was faster than heuristics and could solve problems larger than those that had been solved by other branch and bound algorithms. The algorithm proved its versatility by solving the deterministic problem with minimal modification. Computational results were reported and compared to other algorithms. The application of this algorithm to the Euclidean distance metric was discussed as both a solution technique and an upper bound. Applications to other location problems were also discussed.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Several conclusions can be drawn from this research effort regarding the consideration of new solution techniques for location problems. The following conclusions can be made:

1) Probabilistic formulations of location problems with the Euclidean metric were less sensitive to solution by the centroid approach than corresponding formulations with rectilinear metric.

In probabilistic formulations of the multifacility location problem, the centroid approach to solution produced objective function values that were near-optimal. The centroid approach serves as a heuristic algorithm with less computational burden than special purpose algorithms developed for a certain formulation. The rectilinear metric was too sensitive to rely on the deterministic approach, and it was necessary to develop a special purpose algorithm.

2) The multifacility location problem's objective function is convex but not differentiable. Problems of this type are amenable to gradient-free nonlinear search methods.

In the case of the multifacility location problem with rectilinear metric, the direct search method, a heuristic algorithm was computationally superior to a special purpose exact algorithm.

3) The optimal solution of the generalized location-allocation problem can be found only by an enumeration of all possible allocation schemes.

Heuristic algorithms could produce only a local minimum as a solution to the L-A problem, but could make no guarantees as far as percentage within optimality. Implicit enumeration procedures such as LABB produce optimal solutions with reasonable computational effort.

4) The branch and bound approach is versatile in that it may be used for any probability distribution, either metric, and deterministic formulations.

LABB, which originally was developed to solve the L-A problem among rectangular regions with a rectilinear metric, proved itself applicable to the deterministic formulations of the problem. Other versions of the L-A problem may also be solved with LABB.

Hopefully these conclusions may serve as a model for future developments in the area of location systems. As the various bibliographies indicate, the literature pertaining to location theory has grown rapidly; only a systematic approach to the location problem will provide a framework to evaluate future works.

Recommendations for further research in facility location among rectangular regions would be:

1) Development of a solution technique for the minimax location problem among rectangular distances.
2) Development of a solution technique for the location-allocation problem under the minimax criterion.
3) Development of a systematic method of sensitivity analysis for location problems.
4) Computational experience for the various adaptations of the branch and bound model considered in Chapter IV.
5) Considering the probabilistic formulations under dynamic location assumptions with both finite and infinite time horizons.

REFERENCES

Abernathy, W. J. and Hershey, J. C., "A Spatial-Allocation Model for Regional Health Services," Operations Research, Vol. 20 (1972), pp. 629-642.

Aly, A. A., "Probabilistic Formulations of Some Facility Location Problems," unpublished Ph.D. dissertation, Virginia Polytechnic Institute, 1975.

_______ and White, J. A., "Probabilistic Formulations of the Multifacility Weber Problem," Naval Research Logistics Quarterly (forthcoming).

_______ and _______, "Probabilistic Formulations of the Emergency Service Location Problem," Operations Research Quarterly (forthcoming).

Bazaraa, M. S. and Elshafei, A. N., "On the Use of Fictitious Bounds in Tree Search Algorithms," Management Science, Vol. 23 (1977), pp. 904-908.

Bellman, R. E., "An Application of Dynamic Programming to Location-Allocation Problems," Siam Review, Vol. 7 (1965), pp. 126-128.

Bennett, C. D. and Mirakhor, A., "Optimal Facility Location with Respect to Several Regions," Journal of Regional Science, Vol. 14 (1974), pp. 131-136.

Box, M. J., Davies, D., and Swann, W. H., Non-Linear Optimization Techniques, Monograph No. 5, Great Britain: Oliver and Boyd, Ltd., 1969.

Cabot, A. V., Francis, R. L., and Gary, M. A., "A Network Flow Solution to a Rectilinear Distance Facility," A.I.I.E. Transactions, Vol. 2 (1970), pp. 132-141.

Christofides, N., and Eilon, S., "Expected Distance in Distribution Problems," Operations Research Quarterly, Vol. 20 (1969), pp. 437-443.

Cooper, L., "Location-Allocation Problems," Operations Research, Vol. 11 (1963), pp. 331, 334.

_______, "Heuristic Methods for Location-Allocation Problems," Siam Review, Vol. 6 (1964), pp. 37-53.

_______, "Generalized Locational Equilibrium Models," Journal of Regional Science, Vol. 7 (1967), pp. 1-17.

_______, "An Extension of the Generalized Weber Problem," Journal of Regional Science, Vol. 8 (1968), pp. 181-197.

_______, "A Random Location Equilibrium Problem," Journal of Regional Science, Vol. 14 (1974), pp. 47-54.

Eilon, S., Watson-Gandy, C. D. T., and Christofides, N., Distribution Management, New York: Hafner Publishing Co., 1971.

Elshafei, A. N. and Haley, K. B., "Facilities Location: Some Formulations, Methods of Solution Applications and Computational Experience," O.R. Report No. 90, North Carolina State University, 1974.

El-Shaieb, A. M., "A New Algorithm for Locating Sources among Destinations," Management Science, Vol. 20 (1973), pp. 221-231.

Eyster, J. W., White, J. A., and Wierwille, W. W., "On Solving Multifacility Location Problems Using a Hyperboloid Approximation Procedure," A.I.I.E. Transactions, Vol. 5 (1973), pp. 1-6.

Francis, R. L., "A Note on the Optimum Location of New Machines in Existing Plant Layouts," Journal of Industrial Engineering, Vol. 14 (1963), pp. 57-59.

_______ and Cabot, A. V., "Properties of a Multifacility Location Problem Involving Euclidean Distances," Naval Research Logistics Quarterly, Vol. 19 (1972), pp. 335-353.

_______ and Goldstein, J. M., "Location Theory: A Selective Bibliography," Operations Research, Vol. 22 (1974), pp. 400-410.

_______ and White, J. A., Facility Layout and Location, Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1974.

Geoffrion, A. M., ed., Perspectives on Optimization, New York: Addison-Wesley Publishing Co., 1972.

Himmelblau, D. M., Applied Nonlinear Programming, New York: McGraw-Hill Book Co., 1972.

Hitchings, G. G., "Analogue Techniques for the Optimal Location of a Main Facility in Relation to Ancillary Facilities," International Journal of Production Research, Vol. 7 (1967), pp. 189-197.

Hooke, R. and Jeeves, T. A., "Direct Search Solution of Numerical and Statistical Problems," Journal of the Association of Computer Machinery, Vol. 8 (1961), pp. 212-221.

Jarvinen, P., Rajala, J., and Sinervo, H., "A Branch-and-Bound Algorithm for Seeking the p-Median," Operations Research, Vol. 20 (1972), pp. 173-178.

Katz, I. N., and Cooper, L., "An Always-Convergent Numerical Scheme for a Random Locational Equilibrium Problem," SIAM Journal of Numerical Analysis, Vol. 17 (1974), pp. 683-692.

_______ and _______, "Normally and Exponentially Distributed Locational Equilibrium Problems," Journal of Research of the National Bureau of Standards, Section B, Vol. 80 (1976), pp. 53-73.

Kay, D., Litwhiler, D., and Aly, A., "Location Dominance on Spherical Surfaces," Operations Research (forthcoming).

Kuenne, R. E. and Soland, R. M., "Exact and Approximate Solutions to the Multisource Weber Problem," Mathematical Programming, Vol. 3 (1972), pp. 193-209.

Kuhn, H. W. and Kuenne, E., "An Efficient Algorithm for the Numerical Solution of the Generalized Weber Problem in Spatial Economics," Journal of Regional Science, Vol. 4 (1962), pp. 21-33.

Larson, R. C. and Stevenson, K. A., "On Insensitivities in Urban Redistricting and Facilities Location," Operations Research, Vol. 20 (1972), pp. 595-612.

Lawler, E. and Wood, D., "Branch and Bound Methods: A Survey," Operations Research, Vol. 14 (1966), pp. 699-719.

Lea, A. C., Location-Allocation Systems: An Annotated Bibliography, Discussion Paper Series No. 13, University of Toronto, Dept. of Geography, 1973.

Leamer, E. E., "Locational Equilibria," Journal of Regional Science, Vol. 8 (1968), pp. 229-242.

Love, R. F., "A Computational Procedure for Optimally Locating a Facility with Respect to Several Rectangular Regions," Journal of Regional Science, Vol. 12 (1972), pp. 233-242.

_______, "One Dimensional Facility Location-Allocation Using Dynamic Programming," Management Science, Vol. 22 (1976), pp. 614-617.

_______ and Morris, J. G., "A Computation Procedure for the Exact Solution of Location-Allocation Problems with Rectangular Distances," Naval Research Logistics Quarterly, Vol. 22 (1975), pp. 441-453.

_______ and _______, "Solving Constrained Multi-Facility Location Problems Involving L_p Distances Using Convex Programming," Operations Research, Vol. 23 (1975), pp. 581-587.

Massam, B. H., "A Note on Expected Distances," Geographical Analysis, Vol. 6 (1974), pp. 377-381.

Mitten, L. G., "Branch and Bound Methods," Operations Research, Vol. 18 (1970), pp. 24-34.

Morris, J. G., "A Linear Programming Solution to the Generalized Rectangular Distance Weber Problem," Naval Research Logistics Quarterly, Vol. 22 (1975), pp. 155-164.

Ostresh, L. M., "An Efficient Algorithm for Solving the Two Center Location-Allocation Problem," Journal of Regional Science, Vol. 15 (1975), pp. 209-216.

_______, "The Multifacility Location Problem: Applications and Descent Theorems," Journal of Regional Science, Vol. 17 (1977), pp. 409-419.

Pritsker, A. A. B., "A Note to Correct the Procedure of Pritsker and Ghare for Locating New Facilities with Respect to Existing Facilities," AIIE Transactions, Vol. 5 (1973), pp. 84-86.

_______ and Ghare, P. M., "Locating New Facilities with Respect to Existing Facilities," AIIE Transactions, Vol. 2 (1970), pp. 290-297.

Rao, M. R., "On the Direct Search Approach to the Rectilinear Facilities Location Problem," AIIE Transactions, Vol. 5 (1973), pp. 256-264.

ReVelle, C., Marks, D., and Lieb man, J. C., "Private and Public Sector Location Models," Management Science, Vol. 16 (1970), pp. 692-707.

_______ and Swain, R. W., "Central Facilities Location," Geographical Analysis, Vol. 2 (1970), pp. 30-42.

Rushton, G., Goodchild, M. F. and Ostresh, L. M., eds., Computer Programs for Location-Allocation Problems, Monograph No. 6, University of Iowa, Dept. of Geography, 1973.

Scott, A. J., "Location-Allocation Systems: A Review," Geographical Analysis, Vol. 2 (1970), pp. 95-119.

Seppala, Y., "On a Stochastic Multi-Facility Location Problem," AIIE Transactions, Vol. 7 (1975), pp. 56-62.

Weber, A., Theory of the Location of Industries, C. J. Friedrich (translator), Chicago: University of Chicago Press, 1926.

Wendell, R. E. and Hurter, A. P., "Location Theory, Dominance, and Convexity," Operations Research, Vol. 21 (1973), pp. 314-320.

Wesolowsky, G. O., "Location in Continuous Space," Geographical Analysis, Vol. 5 (1973), pp. 95-119.

_______, "Probabilistic Weights in the One-Dimensional Facility Location Problem," Management Science, Vol. 24 (1977), pp. 224-229.

_______, "The Weber Problem with Rectangular Distances and Randomly Distributed Destinations," Journal of Regional Science, Vol. 17 (1977), pp. 53-59.

_______ and Love, R. F., "Location of Facilities with Rectangular Distances among Point and Area Destinations," Naval Research Logistics Quarterly, Vol. 18 (1971), pp. 83-90.

_______ and _______, "The Optimal Location of New Facilities Using Rectangular Distances," Operations Research, Vol. 19 (1971), pp. 124-130.

_______ and _______, "A Nonlinear Approximation Method for Solving a Generalized Rectangular Distance Weber Problem," Management Science, Vol. 18 (1972), pp. 656-663.

White, J. A. and Case, K. E., "On Covering Problems and the Central Facilities Location Problem," Geographical Analysis, Vol. 6 (1974), pp. 281-293.

APPENDIX A

DATA FOR EXAMPLE PROBLEMS

Problem A.1

Single Facility Location--Five Existing Facilities

Facility	Rectangular Region	Centroid	Weight
1	[5.0,7.5] x [7.5,10.0]	(6.25,8.75)	1
2	[10.0,14.0] x [5.0,7.5]	(12,6.25)	2
3	[16.0,18.5] x [3.5,7.5]	(17.25,5.5)	3
4	[12.5,15.0] x [0.5,3.5]	(13.75,2)	4
5	[7.5,11.0] x [1.0,3.5]	(9.25,2.25)	5

The constrained problem includes the constraints:

$$15 - x - y \geq 0$$

$$30 - 3x - y \geq 0$$

$$x, y \geq 0$$

where (x,y) is the new facility.

Problem A.2

Two Facility Location--Three Random Existing Facilities

i	1	2	3
μ_{x_i}	3	8	15
σ_{x_i}	2	5	4
μ_{y_i}	4	7	2
σ_{y_i}	2	5	4

$$E[w_{ji}] = \begin{array}{c|c} & i = \; 1 \quad 2 \quad 3 \\ j = \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 2 & 6 & 0 \\ 4 & 5 & 1 \end{bmatrix} \end{array}$$

$$E[v_{12}] = 3$$

Problem A.3

Location-Allocation Problem--Two New Facilities to be Located among Five Rectangular Regions

Facility	Rectangular Region	Centroid	Weight
1	[1,2] x [9,10]	(1.5,9.5)	2
2	[4,7] x [3,5]	(6.5,4)	1
3	[9,10] x [1,2]	(9.5,1.5)	2
4	[3,5] x [8,9]	(4,8.5)	2
5	8,9 x 4,7	(8.5,6.5)	1