INFORMATION TO USERS

This reproduction was made from a copy of a manuscript sent to us for publication and microfilming. While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. Pages in any manuscript may have indistinct print. In all cases the best available copy has been filmed.

The following explanation of techniques is provided to help clarify notations which may appear on this reproduction.

1. Manuscripts may not always be complete. When it is not possible to obtain missing pages, a note appears to indicate this.

2. When copyrighted materials are removed from the manuscript, a note appears to indicate this.

3. Oversize materials (maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or in black and white paper format.*

4. Most photographs reproduce acceptably on positive microfilm or microfiche but lack clarity on xerographic copies made from the microfilm. For an additional charge, all photographs are available in black and white standard 35mm slide format.*

*For more information about black and white slides or enlarged paper reproductions, please contact the Dissertations Customer Services Department.

UMI Dissertation Information Service
University Microfilms International
A Bell & Howell Information Company
300 N. Zeeb Road, Ann Arbor, Michigan 48106
Srivastava, Rajesh

ALGORITHMS FOR SOLVING THE LOCATION - ROUTING PROBLEM

The Ohio State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
ALGORITHMS FOR SOLVING THE
LOCATION - ROUTING PROBLEM

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

Rajesh Srivastava, B. Tech., M.B.A.

*****
The Ohio State University
1986

Dissertation Committee:
Dr. L. J. Krajewski
Dr. G. W. Milligan
Dr. W. C. Benton

Approved by

Dr. L. J. Krajewski
Adviser
Graduate Program in Business Administration
To my parents
ACKNOWLEDGEMENTS

I take this opportunity to express my sincere appreciation to Dr. L.J. Krajewiski for his guidance and interest throughout the research. I also express my appreciation to the other members of my advisory committee, Drs. G.W. Milligan and W.C. Benton for their helpful suggestions and comments. I also thank Dr. Milligan for his guidance during my doctoral program. To Dr. B.N. Srikar, who served as a member of my advisory committee during the initial stages and provided guidance, gratitude is expressed. The greatest debt I acknowledge to my wife Veena, who has been my source of strength and support during this period. Her undimining faith in me and her cheerful typing of the dissertation are in no small measure responsible for the successful completion of this dissertation.
VITA

December 18, 1954 ................ Born - Budaun, India

1977 ............................ B. Tech., Indian Institute of Technology, Kanpur, India

1977-1978 ........................ Production Planning and Control Engineer, Larsen & Toubro, Ltd., Bombay, India

1978-1980 ........................ Graduate Research Assistant, University of Wisconsin, Whitewater, Wisconsin

1980 ............................ M.B.A., University of Wisconsin, Whitewater, Wisconsin

1980-1984 ........................ Graduate Teaching Associate, Academic Faculty of Management Sciences, The Ohio State University, Columbus, Ohio

1984-1986 ........................ Assistant Professor, Department of Management, Marquette University, Milwaukee, Wisconsin

PUBLICATIONS


FIELDS OF STUDY

Major Field: Quantitative Methods

Minor Field: Production and Operations Management
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................... iii
VITA ........................................................................ iv
LIST OF TABLES ......................................................... x
LIST OF FIGURES ........................................................ xiii

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 The Location Problem</td>
<td>11</td>
</tr>
<tr>
<td>1.2 The Routing Problem</td>
<td>14</td>
</tr>
<tr>
<td>1.3 The Location-Routing Problem</td>
<td>15</td>
</tr>
<tr>
<td>1.4 Research Objectives and Overview</td>
<td>16</td>
</tr>
<tr>
<td>1.5 Research Design</td>
<td>19</td>
</tr>
<tr>
<td>1.6 Contributions, Assumptions and Limitations of This Research</td>
<td>24</td>
</tr>
<tr>
<td>1.7 Outline of Research Presentation</td>
<td>25</td>
</tr>
</tbody>
</table>

2. LITERATURE REVIEW

| 2.1 The Location Problem | 27 |
| 2.1.1 Location on a Plane | 28 |
| 2.1.1.1 Deterministic Single Facility Location on a Plane | 29 |
| 2.1.1.2 Deterministic Multifacility Location on a Plane | 31 |
| 2.1.1.3 Probabilistic Facility Location on a Plane | 32 |
| 2.1.2 Location on a Network | 33 |
| 2.1.2.1 Deterministic Single Facility Location on a Network | 34 |
| 2.1.2.2 Deterministic Multifacility Location on a Network | 34 |
| 2.1.2.3 Probabilistic Facility Location on a Network | 38 |

<p>| 2.2 The Routing Problem | 39 |
| 2.2.1 The Travelling Salesman Problem | 40 |
| 2.2.1.2 Solution Methods to the Travelling Salesman Problem | 40 |</p>
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Distribution Costs by Industry</td>
<td>6</td>
</tr>
<tr>
<td>3.1 Examples of Problem Size for the Optimal Model</td>
<td>71</td>
</tr>
<tr>
<td>3.2 Customer Data for Small Problem</td>
<td>120</td>
</tr>
<tr>
<td>3.3 General Parameters for the Small Problem</td>
<td>120</td>
</tr>
<tr>
<td>3.4 Comparison of the Optimal and Heuristic Solutions Using Total System Cost</td>
<td>121</td>
</tr>
<tr>
<td>4.1 Cost Ratios for Various Industries</td>
<td>138</td>
</tr>
<tr>
<td>4.2 Comparison of Heuristics to Optimal Model</td>
<td>144</td>
</tr>
<tr>
<td>4.3 Comparison of Heuristics to the Sequential Model; The Pilot Study</td>
<td>148</td>
</tr>
<tr>
<td>4.4 Main Effects for the $2^3$ Factorial Design Used in the Pilot Study</td>
<td>149</td>
</tr>
<tr>
<td>5.1 Results of the Savings-Drop Heuristic</td>
<td>155</td>
</tr>
<tr>
<td>5.2 Analysis of Variance Table for Savings-Drop Heuristic</td>
<td>156</td>
</tr>
<tr>
<td>5.3 Planned Comparisons Between the Group Means of Cost Ratio for Savings-Drop Heuristic</td>
<td>156</td>
</tr>
<tr>
<td>5.4 Results of the Savings-Add Heuristic</td>
<td>160</td>
</tr>
<tr>
<td>5.5 Analysis of Variance Table for Savings-Add Heuristic</td>
<td>161</td>
</tr>
<tr>
<td>5.6 Planned Comparison Between the Group Means of Cost Ratio for Savings-Add Heuristic</td>
<td>161</td>
</tr>
<tr>
<td>5.7 Results of the Cluster-Route Heuristic</td>
<td>164</td>
</tr>
<tr>
<td>5.8 Analysis of Variance Table for Cluster-Route Heuristic</td>
<td>165</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>6.10</td>
<td>Results of the Cluster-Route Heuristic</td>
</tr>
<tr>
<td>6.11</td>
<td>Two-way Results of the Cluster-Route Heuristic</td>
</tr>
<tr>
<td>6.12</td>
<td>Analysis of Variance Table for Cluster-Route Heuristic</td>
</tr>
<tr>
<td>6.13</td>
<td>Planned Comparisons Between the Group Means for Cluster-Route Heuristic</td>
</tr>
<tr>
<td>6.14</td>
<td>Results of the Best Heuristic Selection Rule</td>
</tr>
<tr>
<td>6.15</td>
<td>Two-way Results of the Best Heuristic Selection Rule</td>
</tr>
<tr>
<td>6.16</td>
<td>Analysis of Variance Table for the Best Heuristic Selection Rule</td>
</tr>
<tr>
<td>6.17</td>
<td>Planned Comparisons Between the Group Means for Best Heuristic Selection Rule</td>
</tr>
<tr>
<td>6.18</td>
<td>Summary of Statistical Results</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figures</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Breakdown of the Physical Distribution Dollar</td>
<td>3</td>
</tr>
<tr>
<td>1.2 The Location-Allocation Approach</td>
<td>9</td>
</tr>
<tr>
<td>1.3 The Location-Routing Problem</td>
<td>10</td>
</tr>
<tr>
<td>3.1 The Savings Approach for a Single Depot</td>
<td>77</td>
</tr>
<tr>
<td>3.2 The Multi-Depot Problem (Illustration of the problem with the use of the Standard Savings Approach)</td>
<td>78</td>
</tr>
<tr>
<td>3.3 Macro Flowchart for the Savings-Drop Model</td>
<td>93</td>
</tr>
<tr>
<td>3.4 Macro Flowchart for the Savings-Add Model</td>
<td>107</td>
</tr>
<tr>
<td>3.5 Macro Flowchart for the Cluster-Route Model</td>
<td>116</td>
</tr>
<tr>
<td>3.6 Small Test Problem</td>
<td>122</td>
</tr>
<tr>
<td>3.7 Solution from the Optimal Model for the Small Problem</td>
<td>123</td>
</tr>
<tr>
<td>3.8 Solution from the Savings-Drop Heuristic for the Small Problem</td>
<td>124</td>
</tr>
<tr>
<td>3.9 Solution from the Savings-Add Heuristic for the Small Problem</td>
<td>125</td>
</tr>
<tr>
<td>3.10 Solution from the Cluster-Route Model for the Small Problem</td>
<td>126</td>
</tr>
<tr>
<td>3.11 Solution from the Sequential Model for the Small Problem</td>
<td>127</td>
</tr>
<tr>
<td>6.1 Illustration of Interaction Between Cost Structure and Spatial Distribution for the Cluster-Route Heuristic</td>
<td>213</td>
</tr>
<tr>
<td>6.2 Illustration of Interaction Between Cost Structure and Spatial Distribution for the Best Heuristic Rule</td>
<td>221</td>
</tr>
</tbody>
</table>
INTRODUCTION

Key elements in the design of any physical distribution or logistics system are the location of depots and the distribution of goods from the depots (or warehouses) to the customers. Goods include any product as well as service which the firm sells to its customers. The location of depots, and the distribution of goods from the depots to the customers are also of importance in other areas, such as in materials management, production, and service sector applications. The location problem is to determine the number, size, and location of depots from a larger number of potential sites. Distribution of goods from the depots to the customers includes the problems of determining vehicle fleet size, allocating of customers to depots and routing of customers from the depots.

The importance of the correct problem identification in physical distribution system design can be realized by examining the key strategic planning areas in logistics management. Among the key areas identified by Ballou (6) and others in the field of physical distribution (42, 43, 93) are inventory deployment, facility location, and transportation selection/routing. Transportation selection/routing includes the issues of customer allocation to depots and the assignment and sequencing of vehicles to satisfy customer demands. Ballou's arguments are strengthened by examining the cost figures involving
physical distribution. It is estimated that in the United States the total distribution cost is 21.8% of sales (20). Of this percentage the outbound transportation costs account for 4.3%. Warehousing costs are only 1.6% of the total.

In terms of dollar volume during 1978, U. S. industry spent more than $200 billion on the transportation of goods, and more than $125 billion on warehousing. Overall, distribution accounted for 20% of the G.N.P.. Distribution has an impact on other factors too, such as productivity, interest rates, energy costs, etc. In a study by A. T. Kearney, Inc. in 1978, it was reported that an average company in the U.S. could improve its productivity by 20% or more by improving the distribution system. If the distribution process could be improved by 10% overall, the total savings to the U. S. industry would be $40 billion (94). The implication therefore, is that significant savings and efficiencies can be realized in the physical distribution system through realizing the interactions between the components and designing the system accordingly.

Johnson and Wood (80) identified the major elements of physical distribution and logistics costs. The single most important element is the transportation cost, which accounts for 45% of the total physical distribution costs averaged over all industries. Warehousing accounts for 19% of the total. The breakdown of the physical distribution dollar averaged over all industries is given in Figure 1.1. These figures are consistent with the earlier survey by Neuschel (116). Further, while these averages are based on actual company data and estimates, there can
Fig. 1.1 Breakdown of the Physical Distribution Dollar.

Adapted from:

be up to +20% variations in the costs for individual companies. It is therefore evident that transportation is the single highest distribution cost factor for most industries. Transportation costs include inbound (trunking) costs from the plants to the depots, as well as outbound distribution costs from the depot to the customers. They also include vehicle acquisition costs. Typically, the outbound cost is much more than the inbound cost (127). Warehousing (depot) costs include fixed and operating costs. Fixed costs generally include capital investment in land, building, machinery, and some maintenance and administrative costs. Operating or variable costs include all other costs such as labor, ordering, and material handling.

The relative breakdown of costs holds true in other transportation modes also, such as the air freight industry. The analysis of the distribution systems can therefore be extended to these transportation modes. In the air freight industry, warehousing accounts for 17.0% of the total physical distribution cost, and transportation accounts for 29.4% of the total physical distribution cost (80). Davis and Dillard (42) view the location of distribution centers (depots) as a significant logistic activity. In their analysis, while many factors influence the location decision, transportation plays an important role, since transportation cost is the single largest element of the total distribution cost. Therefore, in the location of depots, the actual transportation costs and method of transportation should be considered. Selecting the method of transportation includes routing decisions.
The relative importance of these key elements in the design of the distribution system depends on the specific system, or industry. For example, the distribution costs are typically much higher than the depot costs in the pharmaceuticals industry. On the other hand, in the electronics and electrical goods industry, the depot costs are much higher than the distribution costs. To make a valid comparison of the depot and outbound distribution costs, these costs should be amortized and considered over the same period of time. The industry figures on costs are provided as annualized costs of warehousing and of distribution. These costs in millions of dollars are summarized in Table 1.1, which has been derived from data collected in a 1976 survey by Lalonde and Zinszer (92). Similarly, the vehicle acquisition costs are also calculated on an annual basis. Such a comparison may allow for determination of the correct system design models to be used in different industries.

The relative importance of these elements also depends on whether the firm undertakes the particular distribution system element design itself or some outside source performs the function. In the location of depots, a firm can either locate its own (private) depots, or rent space in public depots (warehouses). Similarly, in considering the distribution of goods to the customers, private versus common carrier (for hire) options can be considered.

Other examples of logistics systems where the distribution costs are much higher are the air freight industry, newspaper delivery systems where depots may simply be drop off points; and schoolbus routing systems, where depots may be buildings on land owned by the community.
Table 1.1  Distribution Costs by Industry.  
(All costs are in millions of dollars)

<table>
<thead>
<tr>
<th>Industry</th>
<th>Total Distribution Cost</th>
<th>Warehousing Cost</th>
<th>Outbound Dist. Cost</th>
<th>Approximate LTL Distribution Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Manufacturing</td>
<td>22.10</td>
<td>3.7375</td>
<td>6.3375</td>
<td>3.8025</td>
</tr>
<tr>
<td>Chemical and Plastics</td>
<td>27.50</td>
<td>3.7057</td>
<td>6.4360</td>
<td>4.1190</td>
</tr>
<tr>
<td>Food</td>
<td>24.60</td>
<td>5.1400</td>
<td>10.2806</td>
<td>5.9627</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>9.00</td>
<td>0.4091</td>
<td>2.4545</td>
<td>1.5709</td>
</tr>
<tr>
<td>Electronics</td>
<td>10.90</td>
<td>1.4740</td>
<td>1.1473</td>
<td>0.1836</td>
</tr>
<tr>
<td>Paper</td>
<td>36.00</td>
<td>7.6460</td>
<td>18.4778</td>
<td>8.8694</td>
</tr>
<tr>
<td>Machine and Machine Tools</td>
<td>38.00</td>
<td>7.6000</td>
<td>13.3100</td>
<td>8.9110</td>
</tr>
<tr>
<td>All others</td>
<td>14.70</td>
<td>1.5638</td>
<td>4.3787</td>
<td>3.5030</td>
</tr>
<tr>
<td>All Merchandizing</td>
<td>16.90</td>
<td>2.0330</td>
<td>2.6680</td>
<td>1.5477</td>
</tr>
<tr>
<td>Consumer</td>
<td>18.80</td>
<td>2.5640</td>
<td>3.4181</td>
<td>1.6407</td>
</tr>
<tr>
<td>Industrial</td>
<td>5.80</td>
<td>0.3807</td>
<td>0.8062</td>
<td>0.5160</td>
</tr>
</tbody>
</table>

However, the location problem is important in these systems, simply because the location of depots directly affects the distribution costs. There are other systems, where the cost of locating and operating the depots is comparable to the distribution costs, when analyzed on an annual basis. Examples of such systems can be found in the food and in the soft drinks industry, the machine and machine tools industry, and in the paper and paper products industry as indicated by Table 1.1.

In such systems the correct design problem is to consider the design elements together, as a location-distribution problem, since total costs in the system are significantly affected by simultaneous considerations of the two elements. The problem is referred to, from hereon, as the location-routing problem, since routing of vehicles to customers is the major part of the distribution cost.

Previous research on the design of such distribution or logistics systems has focused on location and routing. However, each of these two elements has been examined independent of the other. Research on the location of the depots within the design framework of distribution systems has largely ignored the problem of acquisition and routing of vehicles as being a concurrent problem which affects the solution of the location problem. The usual approach taken in physical system design, which is termed as the sequential approach in this study, is to first solve the location-allocation problem in the system. The depots are located and the customers in the system are allocated to these depots. Next, the routing problem is solved with the customers already allocated to the depots. On the other hand, research on the acquisition and routing of vehicles within the design framework of the distribution
system has largely ignored the issue of location of depots. It is tenable that if the location of depots is considered simultaneously with the routing of vehicles, the distribution system design would be different and more efficient (42, 127).

The difference between previous location-allocation models and a simultaneous location-routing approach is shown in Figures 1.2 and 1.3. In Figure 1.2 the usual approach of location models is shown. From the available set of depot sites, using a location model, some depots are opened and customers are allocated to them. In this location-allocation approach each customer is assumed to be served in a separate trip from the depot to which the customer is assigned. The distance travelled (and therefore the cost) used in the models is based on the single customer round trips from the depot. Figure 1.3 illustrates the simultaneous location-routing approach. In this approach, the distance travelled (and therefore the cost) is based on round trips involving multiple customers. The customers are assigned to the depots based on their inclusion in routes formed simultaneously with the selection of the depots.

The major limitation of previous research in distribution systems has been to consider either the location or the routing element as the dominant problem in the distribution system design at all times. In reality, quite a few distribution systems exhibit an overlap of the two problems in the system design. This is, for example, especially true in distribution systems such as newspaper delivery where location of depots is short term and may shift frequently (79).
Fig. 1.2 The Location-Allocation Approach.
Fig. 1.3 The Location-Routing Problem
The purpose of this research is to gain a better understanding of the system design problem in logistics and distribution systems. In particular, two key elements of such systems, location of depots, and the allocation and routing of vehicles are examined. The problem is to simultaneously determine the location of depots as well as the allocation and routing of vehicles in designing the system. Solving for one element of the system within the solution framework of the other element may suboptimize. While each of the solutions may be optimal or "good", it is possible when considered together they may lead to poor solutions for distribution systems. A location-routing approach is therefore examined.

In the next three sections, the location, routing, and the combined location-routing problems are discussed. The general problem will be defined in each case and some examples provided.

1.1 The Location Problem

Location problems arise in the context of many distribution and logistics systems' design problems, whether it be the location of depots, plants, vehicles, people, or services. In general, the location problem may be defined as: Given a set of potential locations, select as facilities those which will satisfy the given constraints, while meeting the required objectives. Because of the nature of involvement, location problems can occur both in manufacturing industries, for example, the location of warehouses in distribution systems, as well as in service and public sectors. Examples of these would be emergency vehicles'
location, location of personnel (such as farm advisors, or medical advisors), and the location of blood banks.

Location models can be divided into two main structural categories. The first category is location on a continuous plane, or the infinite set approach. Classical location solution methods have taken this approach. Starting with the Weber location problem (53), the problem has been developed to include multiple facility location, and to consider stochastic factors such as customer demand. The usual problem is to find the location of facilities which minimize the sum of the weighted Euclidean distances of customers from these locations. The objective may not be specifically distance but may be some surrogate thereof.

The second category is location on a network. This location problem is characterized by a finite solution space consisting of points on the network. This is especially true of transportation based location problems, where the underlying network may be highways or streets (73). The distance measurement is according to the shortest paths between points on the network. In the past few years, more attention has been given to the problems of locating facilities on a network. Given a number of demand areas and alternative sites for the facilities the basic location problem is to determine where the facilities should be placed on the network, and which area is to be served by a given facility.

Location problems in the manufacturing sector and other private sector physical distribution systems share common characteristics. In the literature on location models, the objective of location has been
considered as mainly economic in nature, whether minimization of costs or maximization of profits. There may be various other non-quantifiable factors that affect the location decision. Some such factors might be the labor market, the union climate, community support, and even personal preference of the decision maker. However, these are difficult to quantify and formalize in a mathematical model.

The location problems in the service and public sectors have been treated somewhat differently in the literature. They differ from the problems in the manufacturing sector in that the goals and objectives may be different and not easily quantifiable. Examples of differing objectives may be the maximum coverage of customers, or the minimization of maximum response time in the case of emergency services. However, cost may still be defined as an objective in many instances, especially as a surrogate objective to represent the more nonquantifiable objectives. Services are further partitioned into those that are ordinary services, such as post offices and mail boxes, city bus stops, parks, public housing, some military applications (e.g. food centers, medical supplies); and those that are emergency services such as ambulance, fire and police. This is due to differing objectives and nature of constraints such as time and distance. Examples of current ordinary service sector location problems would be companies locating bank accounts to maximize float, locating switching centers in a communications network to minimize transmission costs, locating mailboxes to minimize the maximum distance travelled by a user, locating a shopping center to minimize total user cost of transportation and maximize potential customer coverage. Examples of emergency service
sector location problems are location of fire stations to minimize the maximum travel time, location of ambulances to minimize response time, location of police stations to maximize the population covered.

1.2 The Routing Problem

The general routing problem may be defined as finding the sequence of pickup (delivery) points which may be visited, starting and ending at some depot. This differs from scheduling which is the sequencing of the pickup (delivery) points, together with an associated set of arrival and departure times.

The routing problem arises in the larger context of the logistics or distribution problem design in both the manufacturing and the service sector, as part of the transportation subproblem in the distribution systems. Examples of routing problems would be truck dispatching in manufacturing industries when the customer demand is (LTL) (4), food and beverage industries, fuel oil delivery, schoolbus routing, and newspaper delivery.

The routing and scheduling of "vehicles" and their "crews" is an area of importance to both management scientists and transportation planners. From a practical viewpoint, effective routing and scheduling of "vehicles" and "crews" can result in tremendous savings by focusing on areas such as increased productivity, long-range planning, contract negotiation, and better vehicle utilization (67).

In the literature, the objective of the routing decision is mainly economic in nature, such as cost minimization or profit maximization. Other considerations in the routing decision include the fleet size of
vehicles, vehicle capacity, type and number of products being distributed, time windows on operation time, the planning horizon, and the nature of demand at the customers (whether fixed and known or stochastic).

The objectives and considerations of the routing problems generally hold for the service and public sectors also. However, besides cost there may be other objectives which may be used, such as level of service, or maximum coverage. Examples of routing and scheduling problems in the service sector are in schoolbus routing, municipal waste collection, crew scheduling in certain industries (such as airlines), police patrol routes, and routing of postmen.

1.3 The Location-Routing Problem

In many distribution and logistics systems a combined location-routing problem exists. "Commodities" are sent from the central facility to the depots. They are then distributed from the depots to the customers. Several customers may be served on one route from a depot. This is true when an individual customer's demand is less than a truck load (LTL), hence more than one customer's demand can be satisfied by a truckload. In this system design problem, the two elements of interest are the location of the depots and the routes to customers from the depot. However, the two elements are not considered in isolation. The decisions to be made are (79):

(a) number and location of depots,

(b) design of routes emanating from the depots to serve the customers.
The location-routing problem may be stated as:

Given a feasible set of potential depot sites and customer sites, find the location of the depots and the routes to customers from the depots such that the overall "cost" is minimized. The overall "cost" is the sum of location cost and distribution cost. The cost components of location cost and distribution cost have been defined earlier.

From a managerial viewpoint, the location-routing problem is significant because such a problem analysis allows for the distribution or logistics system to be considered from a more realistic viewpoint. For the systems which have been cited as examples earlier, location and routing problems considered in isolation may only suboptimize the system since the objective of each problem does not account for the objective of the other explicitly (79, 118, 127). In location problems in the literature, the distribution costs used ignore the fact that the actual costs are multi-stop routing costs, where many customers are on a single route. On the other hand, routing problems addressed in the literature fail to address the problem of how the depots are located. Thus, better alternative sites may have been ignored. Clearly, if it is known that vehicles will be used in routes for distribution purposes then the routing problem should be considered as part of the framework for locating depots, in the system design.

1.4 Research Objectives and Overview

This research provides a broader perspective on the design of distribution and logistics systems. Instead of treating the location and routing elements as separate design problems, a simultaneous solution to
these two problems is proposed using a location-routing model. Historically, the two problems have been treated in isolation. There are a few research studies which have addressed the problem (26,79,118,127). However, in these studies with the exception of Perl (127), no general solution method has been proposed, and the approaches are not simultaneous but iterative (79), iterating between location and routing solutions in a sequential manner. This research offers alternative solution methods for the location-routing problem. The objective is the minimization of the sum of location (depot) cost, the cost of acquiring the vehicles, and the vehicles' routing (distribution) cost. Cost is used as a surrogate for any other objective which may be proposed.

The most general form of the location-routing problem is a mixed integer program with all the integer variables being zero-one. By introducing the assumption that the set of feasible sites is finite and smaller than the entire set of demand points, optimal solutions can be derived for small test problems. The decision variables used are (1) the link variables, which indicate what demand points and depots are joined together in a route, and (2) the location variables to indicate which of the feasible depot sites are used for locating the depots.

For large-scale problems in both the routing and location areas, optimal solutions are computationally infeasible in the general case at the present time. For special types of problems, optimization approaches are available, such as the one proposed by Geoffrion and Graves for a multi-commodity distribution system design (59), but these cannot be generalized to all cases. In fact, these problems have been characterized as NP-hard in the theory of computational complexity (37,
This implies that for larger problems, optimal solutions are not feasible. Given that the component problems are not solvable optimally, the location-routing problem is not solvable optimally either. Thus, heuristic approaches are developed which share the same general framework, but may differ in terms of other environmental factors such as (1) the number of locations sites, (2) the route size (number of customers on a single route), and (3) any special ordering of the customers set, such as spatial clustering of customers. Such an approach within the framework of the location-routing problem allows for the solution procedures to take advantage of the special characteristics of the distribution system. Most previous approaches to the problem have failed to provide any general framework for the problem or a solution procedure which can be generalized to any distribution system environment.

The heuristic procedures developed in this study are used to obtain solutions to the location-routing problem. The procedures are compared with the conventional sequential decision-making procedure, where first the depots are located and then the vehicles are routed. To determine if the proposed heuristics obtain "good" solutions to the problem, a benchmark solution is needed. For small test problems, optimal solutions obtained from the mixed integer location-routing model are used as a benchmark against which heuristic solutions are compared. For large problems, evaluation of the heuristic solutions is an unresolved issue since there are no existing benchmark solutions to compare against, nor can any optimal solutions be generated. To evaluate the heuristics, the solution to the sequential decision making procedure is used as a
benchmark to compare the heuristics' solutions with. A lower bound on the solution can be obtained by relaxing the integer constraints in the mixed integer program model and solving the resultant linear program. The feasibility of such a bound is examined. The performance of the heuristics is compared against the benchmarks which are determined to be feasible and good.

In summary, the objectives to be achieved in this research are:

(1) The formulation of a location-routing model that includes the characteristics and objectives present in the design of a distribution system.

(2) The derivation of optimal solutions for small problems using the zero-one mixed integer location-routing model, which will also serve as a benchmark for testing the heuristic approaches.

(3) The development of heuristic approaches which provide effective solutions to the location-routing problem in distribution systems, when compared to the optimal approach, and are also computationally feasible for larger real world problems.

(4) The development of a sequential location-routing model based on existing solution methods in literature. This is considered the conventional approach in distribution system design, that of treating the location and routing problems in isolation, solving first for the location problem and then for the routing problem (127). The heuristic approaches are then evaluated against the sequential model for larger problems.

1.5 Research Design

In order to ascertain the impact of location costs versus distribution costs for which the location-routing problem is more relevant, the ratio of location costs for depots to the distribution costs will be included as a factor in the study. Other factors used in the study are the maximum number of depots possible in the distribution
system, the route size, (i.e., the number of customers on a route), and the spatial distribution of customers in the system, whether they occur in clusters or uniformly. These environmental factors are considered to have a greater impact on heuristic performance. There are other factors in distribution system design which have either been fixed to a single value in the study, such as number of customers in the system, the vehicle capacity, and vehicle costs. The factors used are not within the control of the firm, whereas the factors fixed in the study can be controlled. It is more important to measure the impact of those factors which the firm cannot control. Some other factors such as frequency of service to customers, percent of customers to be served, and stochasticity of demand are ignored. Their inclusion would increase the complexity of the study and may obscure the basic results.

Finally, in the absence of real data, factors such as location of customers, location of depot sites, the customer demands, and location costs have been randomized. This is not expected to influence the heuristic performance, and provides a method of performing an experiment to analyze the heuristics' performance.

There is little, if any, literature available on the location-routing problem. To the best of the author's knowledge, there have been no studies done to ascertain the effect of external environmental factors on the performance of solution methods for the problem. Hence, there is no indication as to what factors affect the solution quality.

In this study, it has been hypothesized that the above environmental factors used may affect the goodness of the heuristic
solutions. The factors along with the reasons for including them in the study are:

(1) Warehousing to distribution cost ratio (Cost Structure).
(2) Starting number of depot sites.
(3) Route size parameter.
(4) Spatial distribution of customers.

It is hypothesized that the heuristics will perform better than the sequential model when the ratio is close to 1.0, and be closer to the sequential model when the ratio is high. Further, the heuristics are expected to perform better than the sequential procedure for a larger number of starting depot sites. With the route size parameter it is hypothesized that larger routes may lead to better heuristic performance, since more savings in distance can be generated by travelling in a route. It is also expected that the heuristics will perform better than the sequential model for a uniformly distributed set of customers unless a heuristic explicitly considers the clustered distribution of customers.

To test the effectiveness of the heuristic procedures against the optimal solution, a formal experimental design will be set up. The effectiveness of the procedure is defined in terms of how good the solution is compared to the benchmark used. Further, the above hypotheses as to the suitability of the heuristic procedures under different factor levels are tested.

Problem size is used to examine the computational feasibility for larger problems. Specifically, the problem size, in terms of the number
of variables and constraints is varied by adjusting the number of customers and the number of depot sites.

The following questions are addressed to determine the effectiveness of the heuristic solution procedures.

(1) How effective are the heuristic procedures compared to the solutions obtained from the optimal model?

(2) How effective are the heuristic simultaneous location-routing procedures when compared to the sequential decision-making solution procedure as benchmark in the case of large problems.

(3) Under which environmental conditions (factor settings) are the simultaneous location-routing heuristic procedures different (and better) than the sequential process?

(4) How do different heuristic procedures compare under different environmental conditions?

The criterion used for the evaluation of the heuristics, which also appears as the objective function of the mixed integer program location model is total system cost. In the context of the location-routing problem, the system cost is the sum of the warehousing (depot) and distribution (routing) costs. These costs have been defined earlier.

As far as warehousing costs are concerned, there is no clear delineation between fixed and variable costs. The allocation of cost components to either varies with firms. The required facilities data includes annual space cost, inventory, administration, handling and storage equipment, order processing, personnel, other operating costs. Of this, the space costs and some administrative costs are generally fixed (127). Distribution cost data include the costs of vehicles,
operators and dispatching. The other costs include freight and local cartage (delivery or routing) costs (43).

The above data is needed in order to conduct a cost analysis in a distribution system. Most of this data can be gathered from existing sources. Warehousing data is relatively easy to obtain. Ton-mileage cost is used as a cost measure in the case of trunking (inbound to depots) where the shipments are usually in truck loads (TL). Cost is more difficult to measure in case of local delivery (outbound) shipments which in the location-routing problem consist of less than truckload (LTL) shipments. The absence of any distribution cost data base for allocation problems involving a truckload shipment for multiple customers poses problems in the development of a realistic measure of delivery cost (93). Route costs cannot be obtained accurately since they are dependent on the location of the depots and customer allocations to the depots. On the other hand, the routing of customers influences depot location choice, hence costs. Therefore, a ton-mile measure of cost is adopted as the best available measure of cost. In practice approximate delivery costs are developed based on a ton-mile measure, the average tour length, and the estimated number of tours annually (43).

In this study, a cost measure was developed along the ton-mile measure concept. An approximate minimum distance travelled can be obtained by routing each customer from the closest depot, assuming all depots are open. The location cost of the depots is then uniformly distributed, with the average cost based on the industry average. From the knowledge of the average industry distribution cost, and the total
the average industry distribution cost and the total minimum distance travelled, a cost per unit distance measure is developed.

The actual experimentation in this study is conducted in two phases. First a pilot study is run to examine the other factors besides Cost Structure which have been suggested as being relevant in the design of solution methods to the location routing problem. Given the results of the pilot study, a full research plan is developed in the second phase.

1.6 Contributions, Assumptions and Limitations of This Research

This research examines a significant problem in the design of distribution and logistics systems. The results of this research are expected to provide direction towards a more unified approach in the design of distribution systems by considering a simultaneous solution to two key elements of such systems - location and routing. Specifically, this study differs from past research in the following:

(1) The location problem and the routing problem are considered simultaneously in system design. Previous research has tended to consider each of these two elements of the distribution system in isolation from the other.

(2) A general framework is presented for considering the location-routing problem. A set of heuristics is developed based on the framework, to solve the location-routing problem.

(3) A study of factors which may affect the performance of solution procedures is undertaken. The analysis of these factors helps to identify which of the heuristics is more applicable to a given distribution system. This represents a new approach in system design.
(4) An evaluation procedure is provided to examine the effectiveness of the heuristic procedures for large problems, for which no benchmark solutions exist.

There are some limitations to this research. The integer model utilizes a single objective (cost) in the model. Details such as multiple products for distribution, both pickups and deliveries on a route, and separate capacities of different (i.e., vehicle containerization) products are not included in the model. It is also assumed that each customer demand can be satisfied by a single delivery, and the demand is periodic in nature. Average demand at each customer is assumed to be constant, i.e., the stochastic case is not considered in this research. Also, vehicles are routed back to their original depots. The integer model has the standard linear programming assumptions regarding linear cost relationships and resource utilization. Location and routing costs may not always be linear.

Finally, assumptions have been made regarding the nature of location and distribution costs. As is evident from literature, these are still areas being researched, and no accurate costing procedures or database is available.

When evaluating the results of this study, these assumptions must be considered. Relaxing these assumptions, and the effect of removing the limitations would be the basis for further research in this area.

1.7 Outline of Research Presentation

In this chapter, the location-routing problem in distribution and logistics systems has been introduced. The need for a simultaneous approach to solving the location and routing problems has been
emphasized. The research objectives of this study have been stated, together with a framework to achieve them. The contributions and assumptions of this study are also stated.

In Chapter 2, first the existing literature in location theory is reviewed. Second, the literature in routing theory is reviewed. Finally, the literature on location-routing problems is examined.

In Chapter 3, the zero-one mixed integer location-routing model and the heuristic procedures used to solve the location-routing problem are described in detail. The factors used in this study are discussed in Chapter 4, along with a description of the experimental designs. The pilot study is also discussed in Chapter 4. The results of the experiment to compare the optimal model to the heuristics and the results of the pilot study are presented in Chapter 5. Chapter 6 presents the results of the experiment to compare the heuristics with the sequential model. Finally, in Chapter 7, the results of this research are summarized and extensions to this research are suggested.
CHAPTER 2

The literature which addresses the location-routing design problem in logistics and distribution systems is limited (18, 26, 79, 102, 106, 119, 134, 146, 127). Most of the previous research has concentrated on studying the two simpler subproblems, location and routing. In this chapter, the subproblems of location and routing, and the problem of simultaneous location-routing are examined in order to gain a better understanding of the problem and provide a state-of-the-art knowledge base. This knowledge base will be utilized in the development of solution techniques for the complex location-routing problem.

2.1 The Location Problem

Given a set of feasible sites, the location problem is to determine the number of depots to be established and their locations, given the relevant costs of locating depots, operating the depots, and the distribution costs from the depots to the customers. Weber (53) provided one of the earliest mathematical formulations for the location of the manufacturing industry. Weber's model is for the location of a facility at any point on a plane. Since Weber's initial work on the location problem, this topic has grown into a well researched area. As the location literature is very large, no attempt is made to cover all the literature. Instead, those references that are relevant to the topic
under examination and provide a foundation for research in the location-routing problem will be discussed. The location literature will be categorized by the following areas:

(1) location on a plane versus location on a network (129),
(2) single facility location versus multiple facility location,
(3) probabilistic facility location.

2.1.1 Location on a Plane

Location on a plane is characterized by an infinite solution space, and the facilities may be located anywhere on the plane. The distance measurement is according to some given metric, such as Euclidean or Metropolitan. Weber's work falls into this category. The usual objective is the minimization of the cost of locating the depot and the cost of meeting the customer demands from the depot with the goods the customer requires. For the problem of locating more than one depot, the cost is given by summing over all the depots. The total cost function can be modified to include other costs, such as the cost of operating the depots. The usual assumption is that the transport costs from the depots to the customers are a monotonic function of distance, hence cost can be minimized by minimizing the sum of weighted straight line distances.

Eilon et al. (45) have identified the location problem as a combination of interrelated problems. Four of these are identified as:

(1) determine the number of depots,
(2) determine the location of sites for these depots,
(3) determine the allocation of customers to these depots,
(4) determine the size of the depot for each site, taking particular account of the economies of scale inherent in the operating costs of these depots.

Thus, the general approach has been to treat the location problem as a multistage decision process, in which the subproblems are optimized in sequence.

2.1.1.1 Deterministic Single Facility Location on a Plane

Often, the design problem on hand may be to determine the location of a depot in a single depot system, or in a subsystem of a larger distribution system. Even in well-established distribution systems, the need to reexamine the location of a depot may arise due to various reasons, such as change in demand pattern or customer requirements, expiration of lease or change in rental (lease) charges, or changes in transportation costs to certain customers or areas. In this section, the single depot location problem is examined and several methods for solving the problem are discussed.

A mathematical formulation of the model for the single depot problem is provided in Eilon, Watson-Gandy, and Christofides (45). The location of the depot is described by its Cartesian coordinates, and the depot is used to supply a set of customers, whose locations and demands are known. Further, the cost of transportation of goods from the depot to the customers is also assumed known (see Appendix A for the mathematical formulation). The objective is to find the depot location which minimizes total cost.
Among the solution methods, one of the most common has been the center of gravity method. The method is to simply use the center of gravity of the customers as the depot location. If differential transportation costs exists among the customers, then a weighted center of gravity method can be used. Because of the simplicity of the method, it still has some application. However, for the cost function shown in the mathematical formulation in Appendix A, the solution by this method may not be optimal, as Vergin and Rogers (145) point out. Further, Vergin and Rogers examine the square weighted mean as another choice for an initial location. The method of solution described involves moving the depot along the x and y directions by an amount which depends on the current gradients of the cost function with respect to x and y.

Haley (72) provides a proof of the convexity of the model given in Appendix A. He shows that there is a unique minimum for the one depot case when the customer points are not collinear.

Cooper (35, 36) and Kuhn and Kuenne (91) provide formulations and iterative solution procedures for the generalized Weber problem. The problem is to find a single point which minimizes the sum of the weighted Euclidean distances to that point. Cooper (36) also provides some heuristic procedures where the main emphasis is to find a suitable starting location before using the iterative location procedures. More recently, Oshtresh (123) has provided an iterative procedure to locate a facility.
2.1.1.2 Deterministic Multifacility Location on a Plane

In most distribution systems, the design problem usually involves location of more than one depot. While there has been more research done on location on networks, multi-depot location using the infinite set approach still remains an important and practical tool. This is due to reasons such as the size of problem, the large amount of data which needs to be collected and processed for the feasible set approach, and the fact that the feasible set of sites may have excluded sites which could provide better solutions. The mathematical model for the single depot problem provided in Appendix A can be extended to the multiple depot problem by summing the cost function over all depots.

Eilon et al. (45) present six models for the multi-depot location problem. These models differ in their treatment of the cost function, and the level of detail accorded to the cost function. The solution procedures developed for these models select a set of initial locations for the depots, and allocate the customers to the depots. Through an iterative process, the locations of depots are improved and customers reallocated as the location of the depots change. The iterative process is continued until no further improvement is noted.

Cooper (35, 36) has provided solutions to the location-allocation problem. The main emphasis is on obtaining suitable starting locations before using an iterative solution procedure which is termed the alternate location-allocation method. The method optimally locates the depots before testing the allocation of customers to the depots. Cooper's heuristics are complicated and timeconsuming (45). However, Cooper's study (35) shows that the objective function is shallow in the
region of the optimum. This is of importance since this implies that in practical situations the depot location could be moved from the optimal site to a more attractive site without significantly increasing costs.

Vergin and Rogers (145) also describe solution procedures for the multi-depot problem, similar to their solution procedure for the single depot problem. They also provide an algorithm for the location problem where the distances are rectangular.

Bellman (8) details the transformation of the cost function as given in Appendix A (modified for multi-depot problems) into a recursive maximizing form. This problem is then solved by dynamic programming.

Lawrence and Pengilly (95) developed a heuristic procedure to solve the problem. Much of their effort is aimed at developing a comprehensive and realistic cost function.

2.1.1.3 Probabilistic Facility Location on a Plane

Much of the research into the stochastic nature of the facility location problem has been recent. As Goodchild (69) points out, uncertainty can be present in location-allocation models in a number of ways, such as lack of information, measurement errors, uncertain forecasts of future patterns, and variability in the demand pattern. These can in turn affect the constraints, even the objective function in the model. Again, location-allocation models based on aggregate estimates of demand are subject to error because of loss of information in the aggregation process.

Wesolowsky (150) has provided a formulation of the probabilistic one dimensional facility location problem. Seppala (136) has used a
chance-constrained programming approach to solve the stochastic multifacility location problem. He also defines four different optimization criteria which can be used.

Aly and White (1) have also provided probabilistic formulations of the multifacility Weber problem. They consider both unconstrained and chance-constrained formulations. Other considerations of the problem are provided by Watson-Gandy and Eilon (147), and by Eilon et al. (45).

A more comprehensive bibliography of facility location on a plane can be found in Francis and Goldstein (45) for the period up to 1973.

2.1.2 Location on a Network

More attention has been devoted to the problems of locating facilities on a network. Revelle et al. (129) analyze and differentiate between private and public sector models. They have also distinguished between location on a plane and location on networks. Location on networks considers two factors which the problem of location on plane has not considered for most part. First, transportation costs may not in general be proportional to distance, second, the operating costs of depots may vary from location to location.

In the private sector, the objective in most cases is to minimize the sum of the transportation costs and the amortized facility costs. A general mathematical formulation of the facility location problem in the private sector is provided in Appendix B. The usual approach to the location problem has been solution by mixed integer programming. However, there are problems with this approach, such as presence of non-linearities in the cost functions in many cases. Also, the computer
storage and time required to solve mixed integer programs grows very quickly with problem size, and therefore, such an approach cannot be used for large problems. The approaches taken to solve the problem are broadly classified as heuristics, exact methods, and simulation. First, the single facility location problem is analyzed.

2.1.2.1 Deterministic Single Facility Location on a Network

Hakimi (70) was the first to study the single facility location problem. He examined the problem with regard to locating a single switching center in a communications network. He proved that there must exist an optimal location which is a node on the network and provided an algorithm to solve it.

Mirchandani (73) has shown that regardless of whether the network is oriented or not, an optimal solution exists on one node in the network.

2.1.2.2 Deterministic Multifacility Location on a Network

Most of the focus on solution procedures for the location problem on networks has been for the multifacility case. The model is presented in Appendix B, Problem 1. Here the solution procedures which fall under the categories of heuristic, exact methods, and simulation are discussed. One of the earliest heuristic approaches was taken by Manne (104). He used the Steepest Ascent One Point Move algorithm to obtain facility site selection. Manne's algorithm worked well on small problems, no results were reported for large problems. Other heuristic procedures based on the mixed integer programming model include the
Keuhn and Hamburger (90) heuristic. They employed an add routine to select the depots to be opened, starting with all depots closed. Depots were added one at a time. After opening the profitable depots, they employed a bump and shift routine to improve the solution. They assumed a linear cost function.

Subsequently, Feldman, Lehrer, and Ray (48) assumed a more general form of the facility cost, a continuous concave approach. Further, as opposed to the Keuhn and Hamburger "add" approach they employed a "drop" routine in their heuristic. Starting with all sites as open depots, they closed (or dropped) one depot at a time. A more recent study of the above approaches has been done by Cornuejols, Fisher, and Nemhauser (37). They have also presented extensive computational results on the Keuhn and Hamburger heuristic along with a discussion of the potential worst case performances of the heuristic.

Other heuristics along the same lines, based on the model presented earlier are those developed by Baumol and Wolfe (7) based on the transportation method; Balinski and Mills (2), who consider a concave cost function and Khumawala (84, 86). Khumawala’s heuristics are based on the exact branch and bound algorithm of Effroymson and Ray (44). He presents improvements which improve the computational times and also allow larger problems to be solved, than does the exact procedure, while allowing the solution to be optimal or near optimal in most cases. Khumawala and Whybark (87) also present a heuristic procedure to solve the dynamic warehouse location case.

Heuristics of a different nature are presented by Maranzana (105), and by Teitz and Bart (140). Maranzana’s procedure is based on a node-
partitioning scheme. For the m-median problems, 'm' medians are arbitrarily selected. The node set is partitioned into 'm' sets by assigning each node on the network to its nearest median. Next, an absolute median is determined for each set and the original 'm' medians are replaced by the new 'm' absolute medians and the procedure repeated until the median set does not change. Little computational work is reported for this heuristic. Teitz and Bart's heuristic is a node substitution (bump and shift) heuristic. Initially, nodes are selected as candidate medians in the m-median problem. Nonmedian nodes are examined one by one to see if any can improve the solution upon entering the candidate set. If a nonmedian node improves the objective function, it is entered, and one of the current 'm' medians, whose replacement yields the greatest improvement is removed. This procedure is continued until the solution can be improved no further.

Among exact procedures, Effroymson and Ray (44) were the first to use a branch and bound procedure to solve the mixed integer programming model given in Appendix B. Spielberg (138) also has a branch and bound algorithm for the solution of the location problem.

Among simulation procedures reported in literature is the one used by Shycon and Maffei (137) carried out for the H. J. Heinz company in America. In the public sector the location problem is usually different in the objective to be achieved. Revelle and Swain (130) consider the problem where the objective is to minimize the average time (or distance) people must travel to the facilities. They modify the earlier mixed integer program (in Appendix B, Problem 1) to solve the problem. Their modification is presented in Appendix B as Problem 2.
The location problem has been studied under time and/or distance constraints by Toregas and Revelle (143) and Khumawala (85). They limit the potential facility sites to be in the node set.

The other major class of location problems in the public sector are minimax location problems. Implicitly, two problems are considered: (a) to locate a given number of facilities to minimize the maximum distance between any facility and a demand point nearest it; (b) to determine the minimum number of facilities as well as their location, subject to an upper limit on the maximum distance between a demand point and the facility nearest it. These models are often appropriate in planning the location of emergency services facilities. They may also be extended to non emergency services, for example, location of mailboxes, libraries, etc. Handler and Mirchandani (73) and Minieka (110) develop procedures for solution of the problem.

Toregas et al (144) describe a solution method in the context of locating emergency service facilities. Patel (126) uses a similar approach in locating rural social service centers in India. Church and Revelle (31) formulated the Maximal Covering Location Problem which seeks to identify the locations of a prespecified number of facilities such that the maximum population is served within the maximal service distance.

The above solution procedures for minimax location problems are all based on solving a series of set covering problems. The major difficulty in their solution is the size of the generated set covering problems (27, 29, 55).
2.1.2.3 Probabilistic Facility Location on a Network

Mirchandani (111) has considered the case of locational decisions on stochastic networks. In the context of locating public facilities, implicit in deterministic modeling is the assumption that people always travel to the same facility by the same route, which may not be always true. Hence, there may be a need to model the stochastic nature. Further discussion on the probabilistic location problem on networks is provided in Handler and Mirchandani (73).

In concluding the survey of location models two points can be made. First, in many distribution problems, economic policy often dictates that several demands be processed in one tour of a service unit. In general, all the location approaches considered have assumed a one-demand-per-tour basis for service. Thus, there is a need to consider the several demand case, which could also be termed the Travelling Salesman Location Problem. This problem has been pointed out by Eilon et al. (45). Further, as Handler and Mirchandani note, this problem has received hardly any attention in the literature. Second, a reference must be made to the computational complexity of the location problem. Handler, in Christofides et al. (29), discusses the issues of complexity and efficiency. He notes that the problem is NP-complete. Cornuejols et al (37) also make the same observation. The implication is that it is highly unlikely that an exact algorithm will be obtained to solve all location problems of any size. This result strengthens the base for development of heuristics.
2.2 The Routing Problem

The routing and scheduling of vehicles and crews is of importance to both operations researchers and transportation planners. A well structured and costly activity which is present in both the public and private sector industries, it would appear to be a prime candidate for model-based planning and optimization (103). However, the combinatorial complexity of the routing problem has precluded the widespread use of optimization methods for this class of problems.

The problem considered is that in which a set of geographically dispersed customers with given requirements must be served by a fleet of vehicles stationed at given depots. It is assumed that all vehicles start and end at the depots. The vehicle routing problem is simply a generic name for a broad class of practical decision making problems, involving the visitation of "customers" by "vehicles". By vehicle route is implied a sequence of demand points which the vehicle must travel in order, starting and ending at a depot.

Applications of routing models are found in various areas, as diverse as fuel oil delivery (17), in clustering a data array, or in job-shop scheduling with no intermediate storage (96). The application to the tardiness problem in one-machine scheduling is discussed by Picard and Queyranne (128).

The vehicle routing problem has evolved from the Travelling Salesman Problem. In order to review the vehicle routing model, it is necessary to first examine the travelling salesman problem and its solution procedures.
2.2.1 The Travelling Salesman Problem

Most vehicle routing models are extensions of the Travelling Salesman Problem (TSP). The TSP is the most basic version of the Vehicle Routing Problem (VRP), and also one of the most intensely studies areas of combinatorial optimization. The problem is very easy to state, and very difficult to solve. Given a set of 'n' nodes, the problem is to form a tour of the 'n' nodes beginning and ending at the origin, say at node 1. A formulation of the TSP is provided in Appendix C as Problem 1. The TSP as stated is equivalent to a single depot, single vehicle VRP with no constraints on the vehicle in terms of capacity, time or distance.

Excellent discussions of the TSP are provided in Eilon et al. (45), and Christofides (23, 24). Bellmore and Nemhauser (11) present a thorough overview of the TSP and summarize the main research through the late 1960s on the TSP. In their extensive bibliography on deterministic network optimization, Golden and Magnanti (66) provide many references on the TSP.

2.2.1.2 Solution Methods to the Travelling Salesman Problem

There are many ways of obtaining solutions to the TSP. Some of the methods proposed are exact inasmuch as they provide the optimal solution if pursued to the end, while others are approximate, giving near optimal solutions upon termination. Four different ways of generating solutions are discussed:

(a) sequential tour building,
(b) tour-to-tour improvements,
(c) enumeration methods,
(d) subtour contraction.

Exact solution procedures based on sequential tour building have been developed by Held and Karp (74, 75), and by Little, Murty, Sweeney, and Karel (101). Held and Karp used a dynamic programming approach, while Little et al. have used a branch and bound approach.

Some approximation algorithms which are sequential tour building methods are as follows:

(a) Nearest Neighbor Algorithm: Starting with any node, find the node closest to node last added and add to the path the edge between the two nodes. The tour completes when the first and the last nodes join. Rosenkrantz (131) provides a bound on the worst case behavior of this algorithm.

(b) Clarke and Wright (32) Savings method: Saving is the distance travelled saved by visiting two customers joined in a link rather than visiting each of the two customers separately. The savings for all the links which can be formed are ordered, and starting from the largest savings the links are joined together to form a increasingly larger subtour until a tour is formed. Golden (64) has computed a bound on the worst case performance.

(c) Insertion methods: The tour is constructed by inserting a node into a subtour, thus increasing the subtour until it is finally a tour. The insertion method could be the nearest insertion, or the cheapest insertion (131). Other insertion methods are arbitrary insertion, and the farthest node insertion, where the node farthest from the subtour is inserted into the subtour first.
(d) Christofides (24) has developed a heuristic procedure which is graph-theoretic. The procedure uses the minimal spanning tree of the underlying graph. The worst case behavior of this heuristic is among the best known.

Next, some tour improvement based heuristics are examined. The best known of these tour improvement procedures are the branch exchange heuristics. Lin (99) developed the 2-opt and the 3-opt procedures, while Lin and Kernighan (100) generalized this procedure to the k-opt, for 'k' greater than or equal to three. The procedure involves finding some initial tour, which may be randomly chosen, and then perturbing it by exchanging 'k' branches in the tour by 'k' other branches to form a tour. If a better tour results, it is retained. Rosenkrantz et al (131) provide results on the worst case bounds. Tour improvement methods can also be included as part of algorithms which are composites of procedures, for example, in combination with a sequential tour building method (65).

The procedures developed by Dantzig et al. (40) and by Gilmore and Gomory (62) are based on the subtour contraction method. Gilmore and Gomory use a cutting plane method, which is also used by Crowder and Padberg (38), and by Miliotis (108).

However, as has been pointed out, the best known methods of solving the TSP take an amount of time exponential in the number of nodes. Furthermore, the problem has been characterized as NP-hard (131). Since the TSP has been shown to be NP-complete for Euclidean distances by Karp (83) and by Papadimitriou (124), it is unlikely that a polynomial bounded exact algorithm exists.
Because of the inability to use exact algorithms for large problems, heuristic procedures become attractive. Except for some instances (125), for most part they have been observed to perform well. These approximation algorithms are summarized in Rosenkrantz et al (131), and in Golden, Bodin, Doyle, and Stewart (65).

The formulation of subtour breaking constraints has great influence on the effectiveness of the algorithms. The most extensive analysis on the elimination of subtours, appearing in the literature is by Bellmore and Malone (10). They also developed some new techniques. Carpaneto and Toth (21) have also developed a subtour breaking approach.

2.2.1.3 The Multiple Travelling Salesman Problem

The logical extension of the TSP is the Multiple Travelling Salesman Problem (M-TSP) where more than one vehicle (tour) are considered from a single depot. This, therefore, comes closer to accommodating real world problems than the TSP.

2.2.1.4 Solution Methods to the Multiple Travelling Salesman Problem

For the M salesman, n demand points problem, Svestka and Huckfeldt (139) formulate the M-TSP as given in Appendix C, Problem 2. The distances used in the formulation are defined by a transformation on the original distances in the problem. They describe the computational experience with an exact branch and bound algorithm using both upper and lower bounds.

Other formulations and solution procedures have been given by Golden, Magnanti, and Nguyen (67), and by Gavish and Srikanth (58). Of
these, Gavish and Srikanth provide a good discussion of the problem and present an exact solution method. Their formulation is presented in Appendix C, Problem 3.

Another approach to the solution of the M-TSP is to transform it into the TSP. Bellmore and Hong (9) represent the M-TSP on a graph and show the transformation as an expanded graph for an equivalent TSP tour. For the M-TSP case, the expanded TSP tour will have M-1 more nodes than the original problem. The transformation is also provided in the case of the M-TSP with fixed charges by Hong and Padberg (77).

Russell (132) provides an effective heuristic approach to solving the M-TSP, with some side conditions. These side conditions, when imposed, make the M-TSP into a vehicle routing problem. The heuristic used by Russell is an extension of the tour exchange heuristic of Lin and Kernighan (100). The M-TSP solution procedures in general are no more difficult than the TSP procedures since the M-TSP can be transformed to the TSP.

2.2.2 The Vehicle Routing Problem

The TSP and the M-TSP discussion has been in the context of a single depot, with no constraints on the vehicles, in terms of capacity, distance, or time. In a more realistic situation, there are likely to be multiple depots, and vehicles would have constraints imposed on them, capacity being the most common one.

The single depot Vehicle Routing Problem (VRP) was first discussed by Dantzig and Ramser (41), who developed a heuristic approach using linear programming and aggregation of nodes ideas. The problem is stated
as: Find a set of delivery routes from a central depot to various demand points, each with known requirements, so as to minimize the total distance covered by the entire fleet of vehicles. It is assumed that all vehicles start and finish at the depot, thereby completing tours.

The allocation and routing of vehicles for collection or distribution of goods and services on a regular basis is an essential part of any distribution system. Some common examples appearing in literature are in newspaper delivery as applied by Golden et al (67), and by Jacobsen and Madsen (79), schoolbus routing applications by Bodin and Berman (14), Orloff (120), and by Saha (133) for commercial bus scheduling. Other applications are in municipal waste collection (12), and fuel oil delivery (17, 56). Other examples could be mail collection, laundry delivery services, preventive maintenance inspection tours, home library service, and machine scheduling and sequencing problems.

2.2.2.1 The Single Depot Vehicle Routing Problem

The relationship between the VRP and the TSP has been noted earlier. The single depot VRP with no vehicle constraints is the TSP. There are many formulations of the VRP appearing in literature. A general mixed integer program formulation of the VRP based on Christofides' formulation (29) is given in Appendix D, Problem 1. This formulation also includes scheduling constraints.

The main shortcoming of this model, and of other models formulated is that it is virtually impossible to introduce any of the additional practical constraints such as different types of products, subsets of vehicles that can be used to make deliveries to particular customers,
and priorities of customers. Similarly, vehicles may have additional constraints, such as compartmentalization. There are various other underlying assumptions in different VRPs which need to be considered. Bodin and Golden (15), and Schrage (134) summarize these assumptions. Bodin and Golden also provide a taxonomy for vehicle routing and scheduling problems.

2.2.2.2 Solution Methods to the Single Depot Vehicle Routing Problem

Almost all the heuristic methods suggested for solving the VRP are constructive in the sense that at any given stage, one complete route exists, which is extended in the following stages until it becomes the final route. The solution may then be improved by local optimization procedures, while maintaining feasibility. These constructive methods are classified according to:

(a) the criteria used for expanding the routes, (b) whether the routes are constructed simultaneously or sequentially. Some route expansion criteria are: (1) savings, (2) extra-mileage, (3) radial position, (4) composite criteria. Among sequential and parallel procedures, two categories exist, according to whether the number of routes being formed in parallel is fixed apriori, or whether smaller routes are coalesced into longer ones. First, the heuristic procedures for solution are discussed, followed by a discussion of optimization methods.

Without doubt, the most widely used heuristic is that developed by Clarke and Wright (32). This procedure has been discussed earlier in the context of the TSP. It is also the basis of many programs including the
IBM-VSPX package. This heuristic can be classified as a savings/insertion type, and can be used in parallel or sequential manner. The method is quite effective, as observed in Christofides and Eilon (25). Several extensions and modifications of the Clarke and Wright heuristic have been suggested. Holmes and Parker (76) provide an extension by considering solution improvement through a process of perturbation. Yellow (152) and Gaskell (57) suggest partial enumeration scheme and modified ranking criteria. Webb (149) noted that a minor change in ranking could lead to quite unrelated route structures. Sequential generation of vehicle routes leads to more efficient vehicle utilization than parallel generation, however, it also leads to longer distances. Recognizing this, Mole and Jameson (113) developed a more general savings criterion. Golden et al. (67) have used a version of the Clarke and Wright algorithm for a newspaper delivery system.

Another approach used is a cluster first-route second procedure. Demand nodes are clustered first, and the economical routes are designed over each cluster as the second step. This approach is based on the premise that petal-shaped routes are preferable in general. Golden et al. (67) superimposed a rectangular grid to cluster the customers. Wren and Holliday (151) generate a customer list in order of the angular coordinate from the depot, and align the coordinate axis along the most sparse direction. Customers are added such that the least additional route-metric is incurred, while feasibilities are preserved. Gillett and Miller (61), and Gillett and Johnson (60) also sort customers by angular coordinates. Their algorithm is known as the "sweep" algorithm.
Another approach, though less popular and not as efficient is a route first-cluster second approach. First a large, usually infeasible, route is constructed which includes all demand nodes. Next, the large route is partitioned into a number of smaller routes which are feasible. Bodin and Berman (14) applied this method for school bus routing to and from a single school, and Bodin and Kursh (16) for routing mechanical street sweepers.

The improvement or exchange procedure developed by Lin (99), and Lin and Kernighan (100) has also achieved some popularity. The method has been described in the context of TSP. It has been found to be particularly effective to use this procedure in conjunction with some other procedure which is first used to obtain a good starting solution (25, 132).

Interactive approaches have been suggested for situations where there is a high degree of human interaction incorporated into the problem-solving process. The papers by Krolak, Felts, and Nelson (89), and by Cullen, Jarvis and Ratliff (39) represent work in this area.

Other discussions of heuristic approaches are found in the case study by Knight and Hofer (88), and Christofides (24), who discusses a tree search heuristic and a two phase algorithm. Comparisons, surveys and evaluations of existing heuristics are found in O'Neil and Whybark (117), Mole (112), and Frederickson et al. (52).

A final class of solution procedures for the VRP are those which can be termed as exact. Some of the more effective approaches in literature are those of Held and Karp (74), Crowder and Padberg (38), and Christofides (24). Christofides discusses a set partitioning
formulation, which is also developed by Marsten and Shepardson (107) for crew scheduling problems.

Among the mathematical programming based approaches are the zero-one integer program of Balinski and Quandt (3) for a common carrier routing problem, Foster and Ryan's (50) integer program, which is solved using a revised simplex method, and Saha's (133) algorithm, which uses the underlying network.

Fisher and Jaikumar (49), and Magnanti (103), provide discussions on mathematical programming approaches and solution procedures. Lagrangean relaxation procedures are examined, and another approach suggested is Bender's decomposition.

Finally, there are some network approaches to solving the VRP. Orloff (119, 120, 121) has developed a general framework for the problem and has used the approach to develop a fast heuristic procedure for route constrained fleet scheduling.

2.2.2.3 The Multi-Depot Vehicle Routing Problem

A logical extension of the single depot VRP is the multi-depot VRP. Most realistic distribution systems involve more than one depot.

The formulation for the multi-depot VRP can be easily obtained from the formulation of the single depot VRP. The modifications in the formulation are presented in Appendix D, Problem 2. Whereas the VRP has been studied extensively, the multi-depot VRP has attracted less attention. Since optimal approaches for the single depot VRP are limited to small problems, realistic multi-depot problems are beyond the scope
of these optimal approaches. Hence, the solution procedures developed have been heuristic.

2.2.2.4 Solution Methods to the Multi-Depot Vehicle Routing Problem

As Bodin and Golden (15) point out, there are two obvious approaches to the problem. In the first, a cluster first-route second approach is taken. The nodes are arbitrarily divided over the depots, and then, several single depot VRPs are solved. If the vehicle constraints are violated, then the initial partitioning is revised, and new single depot VRPs solved. An example of this is the Gillett and Johnson (60) sweep algorithm. First the customers are assigned to the depots and then the sweep algorithm is used for the individual depot VRPs.

The second approach is a route first-cluster second approach. A VRP is solved over the entire problem and the minimum number of vehicles needed is obtained, disregarding the depot which would house each vehicle. Next, vehicle routes are assigned to depots, the objective being to minimize total deadhead time. Constraints are placed on the minimum and the maximum number of vehicles which can be housed at each depot.

Another class of heuristics used first generates a solution arbitrarily. The solution is then improved by exchanging nodes one at a time between routes, until no further improvement is noted. Examples of such heuristics can be found in Wren and Holliday (151), and Cassidy and Bennett's (22) TRAMP computer program developed for school meal delivery routes.
Another approach is suggested by Tillman and Cain (142). They extended Tillman's (141) modification of Clarke and Wright savings algorithm to account for multiple depots. The first assignment of nodes is to the nearest depot, subsequently, through their modified savings algorithm, they reassign nodes to build tours. They also obtain an upper bound to the solution by a penalty tour building approach similar to Little et al. (101). Golden et al. (67) present two algorithms for the multi-depot VRP. These algorithms are based on the savings method, and are capable of handling large problems. The algorithms manipulate data for fast computation and less storage requirements. A grid structure is used to reduce problem size.

This section is concluded by noting that there has not been much research in this area. As Bodin and Golden (15) note, the area of multi-depot routing is a promising area of research.

2.2.3 Probabilistic Vehicle Routing Problem

The stochastic VRP is one of current importance. Customer demand and travel times (or distance) are two characteristics of the problem which are often treated as stochastic in real world problems.

Tillman (141) presents a heuristic procedure for the multiple depot delivery problem with probabilistic demands. The approach used is a modification of Clarke and Wright's (32) savings approach, to account for multiple depots. Golden and Yee (68) provide a framework for probabilistic vehicle routing. Golden (63) has presented a statistical analysis of the problem for purposes of heuristic evaluation. Cook and Russell (34) provide a simulation and statistical analysis of the
stochastic vehicle routing problem with time constraints. Buxey (19) has outlined the use of Monte Carlo simulation as a way of solving the VRP. The method is a combination of Monte Carlo simulation and the Clarke and Wright (32) approach.

Finally, it is observed out that there are numerous variations of the VRP. Schrage (134) suggests some considerations in the design of realistic systems. These include considerations of frequency of requirements, time-dependent travel times, multidimensional capacity of vehicles, split deliveries and lumpy cargo, and the stochastic nature of the problem. He also suggests consideration of location of facilities with routing as a concurrent subproblem in the design of the overall system, for the system to be more realistic. Applications emphasizing various of these characteristics lead to a continuum of overlapping models such as location-routing models, as noted by Golden (67).

2.3 The Simultaneous Location-Routing Problem

As has been pointed out earlier, location models have generally never considered tours for variable costs in the facility location process (45). On the other hand, there has been no discussion in literature of the location aspect of depots when the routing problem has been considered. Generally, the routing problem has been solved after the location problem has been solved. Yet, there could be distribution systems where the simultaneous locating of depots and routing of vehicles is the real design problem.

The location-routing problem is stated as: Given a set of feasible depot sites, and a set of customers to be served, find the number and
location of depots, as well as the vehicle routes from the depots to serve all customers, so as to minimize the total system cost. The complex nature of both the location and the routing problems has been examined in the earlier sections of the chapter. Given the complex nature of the subproblems, the combined problem will be even more complex. This could be the reason in part, why not many solution procedures have been developed for the location-routing problem.

2.3.1 Solution Methods to the Simultaneous Location-Routing Problem

There is very little literature which deals with this practical problem which often occurs in distribution systems. The problem as stated earlier is to locate depots from which customers are served by tours rather than individually.

Madsen (102) provides a survey of the existing studies in literature which addresses the problem. As stated in his survey, on a broad level, among the decisions to be made are those concerning the number and location of depots, and the design of tours starting from these depots to serve the customers. In main, the article focuses on the study by Jacobsen and Madsen (79), though some of the other studies are also surveyed.

Christofides and Eilon (26) were among the first to consider the problem of locating a depot from which customers are served by tours rather than individually. They provide an approximation formula for the total length of the tours as a function of the average number of customers per vehicle, the size of the area served by the depot, and the sum of depot-customer distances. This formula is valid under the
assumption that customers are uniformly distributed.

Burness and White (18) consider the location-routing problem as a Travelling Salesman Location Problem. They consider the solution space to be continuous. Additionally, the set of customers to be visited in a given trip is considered to be a random variable. The solution procedure outlined is iterative, a new facility being inserted on a travelling salesman tour through ‘m’ customers. The approach uses the TSP as a subproblem within the location problem.

Watson-Gandy and Dohrn (146) describe an application of the problem to a food and soft drinks company in Southern England. The objective function is transformed using a Christofides and Eilon (26) like approximation as described above. Standard location-allocation techniques are then used to solve the location problem, following which, the routing problem is solved by a routing method such as the savings method. The method is sequential in nature.

Jacobsen and Madsen (79) do a comparative study of some heuristics developed for a two-level location-routing problem. They discuss the problem in the context of the design of a newspaper delivery system. They compare three different procedures. The first is a tour construction method which they term as Tree-Tour Heuristic (TTH). The problem is viewed as a spanning tree satisfying constraints at minimum costs (if return arcs are ignored). They conclude that the TTH would probably be useful in a situation where the constraints are tight, and the problem becomes one of finding a feasible solution. The second heuristic developed is called the ALA-SAV approach. It is basically a two-stage heuristic composed of the Alternate Location Allocation (ALA)
method of Cooper (36) and the savings method (SAV) of Clarke and Wright (32). First, a prespecified number of depots are located by ALA and customers allocated to depots. Then, secondary tours are formed by applying SAV sequentially to each depot and its allocated customers. The third method proposed is called the SAV-DROP procedure. First, tours are formed by a SAV procedure, i.e., the savings heuristic of Clarke and Wright (32). The facilities are then located on the tours formed in the first step using the DROP routine of Feldman, Lehrer, and Ray (48). Then, the tours are formed again using the new depot locations. This heuristic has the advantage of reducing the volume of data sequentially.

The computational results are described as fair.

Or and Pierskalla (118) have formulated a transportation location-allocation model for the location of blood banks from which vehicles are sent on tours to deliver blood to the hospitals. The model is called the Blood Transportation-Allocation Problem (BTAP). A feasible set approach is used for location, and the number of facilities required is assumed known. Vehicle capacity, and maximal travel distances are also considered.

The model developed is a complex mixed integer program which includes the vehicle routing formulation as a part of the program. Since the problem is a complex optimization one, it is solved by a two-stage heuristic procedure. The procedure alternates between solving a series of VRPs and allocating hospitals to blood banks based on these. By exchanging hospitals the allocations are changed and different VRPs run again. When all reasonable exchanges i.e., those that improve the solution are complete, the procedure stops. Two heuristic algorithms are
presented which work on the exchange idea. One deficiency noted in the solution procedures is that the location aspect of the problem has not been addressed, since the location of blood banks is assumed known in the solution.

More recently Perl (127) has presented a unified warehouse location-routing analysis. His work represents the first real step in obtaining solution methods to the simultaneous location-routing model. The interdependence between the location of distribution centers and vehicle routing is addressed in the study.

The three basic components of the distribution center location problem are identified as location, allocation, and routing. A mixed integer programming model of the warehouse location-routing problem is provided. A heuristic approach to solving the problem is proposed which is based on decomposing the problem into its parts. The three phase solution consists of solving a complete multi-depot vehicle dispatch problem in phase one, solving the warehouse location-allocation problem in phase two, and the multi-depot routing-allocation problem in phase three. Improvements in the initial solution are obtained by iterating through phases two and three. These iterations are continued until the improvement in solution becomes less than a prespecified value. An application of the proposed model to a national distribution company's operations at a regional level is presented.

Finally, Webb (148) points out that the use of simple cost functions as has traditionally been done in most of location literature, may give misleading results since most wholesale distribution is of the multiple-delivery routes type, and the cost of a delivery is thus influenced by
the occurrence of other deliveries.

2.4 Summary

In summary, in this chapter, the literature from the areas of location, vehicle routing, and simultaneous location-routing has been discussed. After examining these areas, the following statements can be made with regard to the state-of-the-art in the design of distribution systems.

(1) The location problem has been well researched, both with respect to location on a plane as well as location on a network. An issue, however, in the problem of location is the complexity of the problem. The implication is that for large problems, optimal solution procedures cannot be applied. Therefore, in most realistic distribution systems, heuristic solution procedures would have to be used to solve the location problem. Location problems have considered the distribution detail in solution methods, but the wrong distribution costs.

(2) The vehicle routing problem has also been well researched over the last two decades. However, as with location problems, the major issue with the VRP is the computational complexity of the problem. Almost all vehicle routing problems are NP-hard (83). Hence, for large realistic problems, optimal solution procedures will not work. Good heuristic procedures exist, which in many cases obtain near optimal solution results. However, the shortcoming of these VRP solution methods has been the exclusion of considerations present in a realistic system, such as differing frequency of requirements for customers, time-dependent travel time, split deliveries and simplistic cost functions. Again, the VRP has been considered as a separate problem in distribution system design, distinct from the location problem.

(3) Past research has failed to address the simultaneous location-routing problem adequately. The few existing studies are, for most, not really simultaneous in character since for most part their solution procedures imply sequential solution of the location and routing problems. Also, most of the literature deals with specific
applications and suggests no solution procedure for the general situation. However, it is also apparent from these studies that there are situations where there is a need to consider the simultaneous location-routing model.

(4) Given the complexity of the two subproblems of location and routing, it is obvious that the location-routing problem is no less complex. Thus, optimal solution procedures may not be possible for realistic problems. The thrust should be towards the development of good heuristic procedures. Further, since there exists no comparative literature base, the evaluation and analysis of any heuristics developed remains an issue.

Despite the complexity of the location-routing problem, and limited research in the area, several of the location and routing procedures discussed earlier provide approaches which can be used to create alternative heuristic solution procedures to the simultaneous location-routing problem. Of particular interest are the heuristics based on the savings approach and the cluster approach for the vehicle routing problem. The add and drop approaches towards depots in location literature, along with infinite set solution approaches are also of interest.

Since the focus is on developing heuristic procedures, there are some criteria which need to be considered for evaluation purposes. Ball and Magazine (5) develop a list of criteria, some of which are:

(a) how close does the solution come to the optimal?
(b) the computer running time and storage space for the heuristic,
(c) flexibility of the heuristics is an important consideration, since changes in the model required by different real world problems should be easily handled,
(d) simplicity and analyzability of the heuristics.
Besides these, the other issue in developing location-routing problems is the lack of a benchmark to compare against. For small problems, the optimal solution can be obtained for comparison with the heuristic solution. For larger problems, the sequential decision making process described in Chapter 1 is suggested. The sequential method may be considered as the normal practice in distribution system design. These are the benchmarks suggested against which to compare the heuristics developed in this research.

In Chapter 3, the mixed integer zero-one location-routing model is developed to evaluate the performance of the heuristic procedures for small problems. The sequential model is developed to provide the benchmark for heuristic evaluation in the case of larger problems. The heuristic procedures are also developed in detail.
CHAPTER 3
MODEL DEVELOPMENT AND SOLUTION PROCEDURES

The solution of the location-routing problem in distribution systems is quite complex, being a combination of two well known difficult problems, locating and routing. The literature review in the previous chapter indicates that in general, optimal or near optimal solutions to either the location, or the routing problem, are difficult to derive. In both cases, the review has shown that because of the difficulty in obtaining optimal solutions, heuristic procedures have been adopted. It has been shown by many researchers (37, 29, 98), that for large problems, it may not be possible to derive optimal solutions due to the computational complexity involved. Computational complexity occurs both in computer storage space and in computing time. Thus, the development of an optimal solution procedure to the simultaneous location-routing (SLR) model is highly unlikely for the case of realistic problems. Also, from the practitioner's point of view, the need is for techniques that can be understood and implemented with relative ease. The focus of this research is therefore, to develop effective solution procedures for the problem which can be used in practice in the design of distribution systems.

In this chapter, the SLR problem is first modeled as a zero-one mixed integer program (MIP). The objective in developing the model is to
provide a benchmark for comparing the proposed heuristics in the case of small problems. Moreover, the MIP model provides a specific statement of the location-routing problem. The objective function in this formulation minimizes the sum of the routing cost, the vehicle acquisition cost, and the cost of locating and operating the depots. Specifically, the model includes the location and routing costs, the number of depots, and the number of routes (which can be controlled by varying vehicle constraints, such as capacity or maximum distance travelled). The decision variables included in the model are arc variables between each pair of nodes, which indicate the path taken in each of the routes, and a set of variables to indicate which of the depot sites are open depots, and which are closed. The arc variables between each pair of nodes are zero-one variables. The value of the arc variable is 1 if there exists a connection between the two nodes for the vehicle route being created. Otherwise the value of the arc variable is 0. The nodes referred to here are all customer points and depot sites. A vehicle route is therefore defined by all the arc variables which have a value of 1 for the vehicle. A set of zero-one variables is used to represent the open and closed depots. A depot is considered open if the value of this variable is 1 and closed if the value of the variable is 0.

Since the computational requirements of the zero-one integer program developed are severe, the solutions are limited to very small problems. An analysis of the MIP model in the next section shows the computational complexity of the problem and the infeasibility of obtaining solutions to larger problems. In view of this, the heuristic
approach to solving the problem is undertaken. The heuristic procedures are techniques which can easily be applied to distribution systems. Another advantage with the heuristic approach is the greater flexibility afforded in system design when compared with the optimal approach. A wide range of sensitivity analysis can be performed using heuristic procedures, which is not possible with the optimal approach.

For larger problems, the mixed integer program cannot be used for comparison since no solutions can be obtained because of the large problem size. A comparison of the heuristic solutions against a sequential location, routing solution procedure is made. The sequential solution procedure has been defined in Chapter 1 as the conventional real world method of system design, whereby the location decision is made first, and the routing decision follows after the location of the depots is done. Since the objective of the study is to develop procedures to solve the distribution system design problem (the SLR problem), the effectiveness of such procedures can best be examined by comparing against existing practice under similar environmental conditions and measuring the degree of improvement.

In the following section, the zero-one integer formulation of the location-routing problem in distribution systems is presented, followed by a discussion of its limitations. The sequential solution procedure is discussed next, and finally the heuristic procedures are described.

3.1 The Location-Routing Model

A general mixed integer linear program formulation of the location-routing problem is discussed first. This model allows for M depots to be
located from a given finite set of feasible sites $R$, such that $N$ customers are served by one or more tours formed from each of the $M$ depots at minimum total system cost. The problem is defined as simultaneously establishing a given number of depots and determining the delivery routes from the depots, such that the total system cost is minimized. The vehicles may have capacity, travel distance, and maximum route time constraints. All vehicles finish at the depot at which they originate. In the next section a mathematical statement of the SLR is provided. The assumptions made in the development of the model are stated. A discussion of the objective function and the constraints used in the model is provided to explain the model.

3.1.1 A Mathematical Statement of the Location-Routing Problem

The parameters in the model are defined as:

$G = \{g_r | r = 1,\ldots,R\}$ is the set of $R$ feasible sites for locating the $M$ depots

$M =$ number of depots to be established, where $M \leq R$

$H = \{h_i | i = R+1,\ldots,R+N\}$ is the set of $N$ customers to be served from the depots

$S = (G) \cup (H)$ is the set of all feasible sites and customers

$V = \{v_k | k = 1,\ldots,K\}$ is the set of $K$ vehicles available for routing from the depots. Each of the vehicles may be routed from any one of the depots
\( C_{ij} \) = average annual cost of travelling from node \( i \) to node \( j \), 
\( i \in S, j \in S \). \( C_{ij} \) is assumed to be independent of type or capacity of vehicle used.

\( C_k \) = annual cost of acquiring vehicle \( k \) \( (k = 1, \ldots, K) \)

\( F_r \) = annual fixed and operating depot cost at site \( r \) \( (r = 1, \ldots, R) \)

\( q_j \) = average number of units demanded by customer \( j \), \( j \in H \)

\( Q_k \) = capacity of vehicle \( k \), \( (k = 1, \ldots, K) \)

\( d_{ij} \) = distance from node \( i \) to node \( j \), \( i \in S, j \in S \)

\( D_k \) = maximum distance that can be travelled in a route by vehicle \( k \) \( (k = 1, \ldots, K) \)

\( t_{ij} \) = time taken to travel from node \( i \) to node \( j \), \( i \in S, j \in S \)

\( u_i \) = time taken to deliver (or collect) at customer node \( i \), \( i \in H \), \( (u_i = 0, i \in G) \)

\( T_k \) = difference between the normal starting time and the normal quitting time for vehicle \( k \), \( k = 1, \ldots, K \)

The decision variables are defined as

\[
X_{ijk} = \begin{cases} 
1 & \text{if vehicle } k \text{ goes from node } i \text{ to node } j, \quad i \in S, j \in S, \\
0 & \text{otherwise,} \\
\end{cases} \quad k \in V, i \neq j
\]

\[
Z_r = \begin{cases} 
1 & \text{if a depot is established at site } r, \quad r \in G \\
0 & \text{otherwise} \\
\end{cases}
\]
\( R_i \) are continuous variables, used in the subtour breaking constraints, \( i \in G \cup H \).

For the above definitions the following basic assumptions are made. The costs are assumed to be on an annual basis. This is consistent with the literature (43), as has been discussed in Chapter 1. Thus, \( C_k \) is the equivalent annual cost of acquiring vehicle \( k \), not the full purchase cost. Similarly, the cost of establishing a depot is also considered on an equivalent annual basis. This depot cost includes both the fixed amortized costs and the operating costs. Fixed costs and operating costs can be further broken down into more detailed subcategories. Fixed costs include costs associated with land, building, maintenance, and some fixed administrative costs. Variable costs include labor, material handling, and order processing. Davis (43) provides a framework for the amortization of costs and their breakdown into specific categories. The fixed and operating costs have been combined in the objective function assuming that the depots are uncapacitated in the problem. Further, it is assumed that the distribution system under consideration is one where all the customers are demand points only and no pickup points exist. This situation is realistic for many distribution systems, such as newspaper delivery systems, and fuel oil delivery systems. Finally, the model assumes that the set of feasible depot sites is disjoint from the customer set to reduce the model complexity, though this may not be necessary in the general case.
A mixed integer programming formulation based on these parameters and decision variables is as follows:

Minimize \[ \sum_{i \in S} \sum_{j \in S} \sum_{k \in V} c_{ijk} x_{ijk} + \sum_{k \in V} c_k \left( \sum_{r \in G} \sum_{j \in H} x_{rjk} \right) + \sum_{r \in G} f_r z_r \]

Subject to

1. \[ \sum_{k \in V} \sum_{i \in S} x_{ijk} = 1 \quad \forall j \in H \]  
2. \[ \sum_{j \in H} \sum_{i \in S} q_{j} x_{ijk} \leq q_k \quad \forall k \in V \]  
3. \[ \sum_{j \in S} \sum_{i \in S} d_{ij} x_{ijk} \leq d_k \quad \forall k \in V \]  
4. \[ \sum_{j \in S} \sum_{i \in S} t_{ij} x_{ijk} + \sum_{j \in S} \sum_{i \in S} u_{ij} x_{ijk} \leq T_k \quad \forall k \in V \]  
5. \[ \sum_{i \in S} x_{ipk} - \sum_{j \in S} x_{pjk} = 0 \quad \forall k \in V, \quad p \in S \]  
6. \[ \sum_{r \in G} \sum_{j \in H} x_{rjk} \leq 1 \quad \forall k \in V \]  
7. \[ \sum_{r \in G} z_r = M \]
\[ \sum_{k \in V} x_{rmk} + z_r + z_m \leq 2 \quad m = 1, \ldots, R, \quad r \in G \] (8)

\[ \sum_{k \in V} \sum_{j \in H} x_{rjk} - z_r \geq 0 \quad V \quad r \in G \] (9)

\[ \sum_{j \in H} x_{rjk} - z_r \leq 0 \quad V \quad k \in V, \quad r \in G \] (10)

\[ R_i - R_j + (R+N) \sum_{k \in V} x_{ijk} \leq R+N-1 \quad V \quad i, j \in H, \quad i \neq j \] (11)

\[ x_{ijk} = 1 \text{ or } 0 \quad V \quad i, j \in S, \quad k \in V \] (12)

\[ z_r = 1 \text{ or } 0 \quad V \quad r \in G \] (13)

The objective function minimizes the total cost of routing, acquisition of vehicles and locating and operating depots. The first term in the objective function is the routing cost, the second term the acquisition cost of vehicles, and the last term represents the cost of locating and operating depots.

The first set of constraints ensures that each customer is visited by a vehicle, i.e., included in a route. It also ensures that each customer is visited by only one of the vehicles available.

The second, third and fourth set of constraints are vehicle type constraints imposed on the system. The second set of constraints are the vehicle capacity constraints. Since the model considers deliveries only,
the constraints ensure that the total demand on a route does not exceed the capacity of the vehicle assigned to the route. The constraints further consider that an individual customer's demand has to be less than a truckload in order for the customer to be included in a route along with other customers. This set of constraints can also be used to control the average number of customers in a route. The third set of constraints are included to ensure that a vehicle does not exceed the maximum distance specified for a route. Such constraints may arise occasionally, but do not always exist, unlike the vehicle capacity constraints. The fourth set of constraints are the total elapsed route time constraints. They limit the maximum time that a vehicle can spend on any single route, (such as in a newspaper delivery system where the time interval from press to street cannot be excessively long). Union contracts could be another reason for the existence of such constraints. The total route time constraints includes the travel time, as well as the time spent at the customers, loading or unloading.

Constraint set five ensures route continuity. This ensures that if a vehicle enters a node, whether customer or depot, it will also exit from that node. Constraint set six allows each vehicle available to be used from a maximum of one depot, and only once. This is not a serious constraint since it would in most cases be normal practice for vehicles to return to their depot. The constraint ensures that the same vehicle is not routed out of more than one depot in the same period of time. Of course, a vehicle may be routed from a depot more than once, by simply changing its identity in the model each time.
Constraint set seven fixes the number of depots to be established at M. It is not necessary to fix the number to M in order to obtain a solution to the model. If the number of depots open is set less than or equal to M, we will still obtain an optimal solution with the best number of depots. However, the number of computations increase drastically, making the problem even more difficult to solve. Thus, in the current model the number of open depots is set to M. To compare for different system designs in terms of number of depots, the parameter M can be varied from one run to another.

Constraint set eight is used to ensure that there are no links between any two depots. This is necessary to avoid routes including more than one depot. This allows the model to be consistent with earlier constraints. Constraint set 9 is used to require that each depot established is used also. What this set of constraints does is to force at least one customer to be included in a route from that particular depot. The constraint ensures that at least one vehicle will originate from an open depot. The tenth set of constraints complements the earlier one, and it is used to ensure that a vehicle does not start from a site which is a closed depot. The argument for this set of constraints is similar to the last set of constraints.

The eleventh set of constraints are the subtour-elimination constraints derived for the travelling salesman problem by Miller, Tucker, and Zemlin (109). These constraints ensure that no tour is formed which consists of customers exclusively. Each tour must include a depot from which the vehicle originates. Thus, they also force each tour to pass through a depot. These subtour elimination constraints work as
presented only if the triangle inequality holds for the distance matrix. Otherwise more general subtour breaking constraints have to be used. The advantage of using these constraints is that they are much less in number than the other forms of subtour elimination constraints.

The final two set of constraints are the integrality conditions to ensure that the decision variables are integer and zero-one.

3.1.2 Limitations. Complexity of the Model

An examination of the formulation presented in the previous section shows that even for a small problem of 3 location sites, 8 customer nodes, and 3 vehicles, the number of integer (zero-one) variables is very large. As can be seen from Table 3.1, there are 333 integer variables. Problem C in Table 3.1 illustrates the size problem. Hence, standard integer program optimal solution procedures cannot be used, except for very small distribution systems. An optimal mixed integer solution approach is thus impractical for most realistic distribution systems where there would be more than two or three of depot sites, and more than 8 to 10 customers. Again, all the restrictions associated with a linear program apply. Thus the solution is restrictive in nature, and is not flexible in accommodating changes which occur in realistic problems. Finally, it may also be pointed out that this problem belongs to the class of NP-hard problems, since both the subproblems have been shown to be NP-hard (37,83,98). The storage and time requirements grow in more than polynomial manner with respect to the problem size, as shown in Table 3.1. To obtain solutions to realistic problems, heuristic approaches have to be considered. In the following sections heuristic
Table 3.1 Examples of Problem Size for the Optimal Model

<table>
<thead>
<tr>
<th>Number of Customers (N)</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Depot Sites (R)</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Number of Vehicles Available (K)</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Total Number of Nodes (N+R=S)</td>
<td>8</td>
<td>11</td>
<td>19</td>
</tr>
<tr>
<td>Binary Variables</td>
<td>S(S-1)K+R</td>
<td>114</td>
<td>333</td>
</tr>
<tr>
<td>Continuous Variables</td>
<td>N</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

| Set 1 | N  | 6  | 8  | 15 |
| Set 2 | K  | 2  | 3  | 4  |
| Set 3 | K  | 2  | 3  | 4  |
| Set 4 | K  | 2  | 3  | 4  |
| Set 5 | SK | 16 | 33 | 76 |
| Set 6 | K  | 2  | 3  | 4  |
| Set 7 | 1  | 1  | 1  | 1  |
| Set 8 | R(R-1) | 2  | 6  | 12 |
| Set 9 | R  | 2  | 3  | 4  |
| Set 10 | RK | 4  | 9  | 16 |
| Set 11 | N(N-1) | 30 | 56 | 210 |
approaches to the location-routing problem are developed. The first algorithm developed is a sequential location, routing decision process. This is developed to serve as a benchmark for comparing the simultaneous heuristics developed in the later sections of the chapter. The sequential model is also assumed to be representative of the conventional real world distribution system design on the basis of the fact that all routing algorithms developed in literature and used in practice assume that the depot locations as well as the customer allocation to the depots are known. This implies that the location decision has already been made, prior to the routing decision.

3.2 The Sequential Approach to the Location-Routing Problem

A sequential location-routing algorithm is developed to serve as a benchmark for comparison of the simultaneous location-routing heuristics. The only vehicle constraint imposed is the capacity of the vehicles. The sequential approach has limitations. In the phase where the location decision is made, the distribution costs are assumed to be incurred on the basis of single customer routes from the depots. At best, assuming that it is known that customers will be served in a route, the demand for customers nearby can be aggregated, and their midpoint taken as an approximation to the distance travelled from the depots. Such an analysis leads to aggregation errors and loss of information. To avoid such problems, the sequential algorithm works in the following two phase manner.

In the first phase the algorithm utilizes an existing location algorithm for location of the depots and allocation of the customers to
the depots. Khumawala's (84) branch and bound algorithm is used. The procedure does not guarantee an optimal solution, though it does provide optimal to near optimal solutions. The distribution costs are at this stage calculated based on single customer trips from the depots to the allocated customers. These costs figure into the location decision.

In the second phase, the routing of vehicles on the allocated customers is performed using a modified Clarke and Wright (13) routing algorithm. Vehicle capacity constraints are also included in the routing detail. Clarke and Wright's original procedure was developed for a single depot route case, and is based on a savings scheme, where the savings is achieved by linking together two customers in a single route instead of two separate routes. This procedure has been modified to the case of single depot multiple routes. The problem may still be considered single depot, even though there are multiple depots, because the customers are already allocated to the depots, hence the routings are done for individual depots and their allocated customers. While this algorithm may not be the best performing vehicle routing algorithm, it is certainly one of the most popular methods and forms the basis of more commonly used commercial vehicle routing computer software such as the IBM VSPX.

The distribution costs computed in phase one are dropped, and only the depot costs are retained. The total routing cost combined with the total vehicle acquisition cost becomes the new distribution cost. This replaces the single customer trips based distribution cost calculated in phase one. The total system cost is the sum of the depot costs obtained in phase one, and the new distribution costs obtained in phase two.
3.3 Heuristic Approaches for the Location-Routing Problem

A set of heuristic procedures is developed to solve the SLR problem. It can be noted that distribution systems may differ not only in their objectives and in their cost structure, but may also differ in performance on other factors which have been included in this study. These factors have been identified in Chapter 1. It has been hypothesized in this study that such factors may influence the distribution system design by influencing the efficiency and quality of the solution procedures developed. There is no literature available on these environmental aspects of the SLR problem. The factors have been included in the study to examine their effect. Perhaps alternative procedures will interact differently with the various environmental factors. Three different heuristic procedures are developed in this chapter which differ in their approach. In these heuristic procedures, the concepts of routing model solution procedures have been integrated with location model solution procedures using the existing solution procedures for locating and routing as a base for further development.

3.3.1 The Savings-Drop Model

The savings-drop approach utilizes a modified version of the "travel time savings" heuristic developed by Clarke and Wright (32). The extension to the multi-depot situation is based on the multi-depot modification developed by Tillman (141). As has been noted in Chapter 2, this approach is one of the most popular and generally yields good solutions, besides being computationally easy. The drop idea used in the heuristic parallels the Feldman, Lehrer and Ray (48) approach, to the
extent that the idea of dropping depots from consideration is used. The development of this model is based on the fact that the two underlying approaches are among the most popular in vehicle routing and location, respectively. It is expected that such an approach will work well, especially with a smaller number of depot sites, and with a uniform spatial distribution of customers. The solution procedure simultaneously considers dropping depots and assigning the customers to routes developed from open depots and iterates until the final solution is obtained. A savings based drop routine identifies the depot to be closed, initially starting with all depots open. This incorporates the routing detail into the location decision. The routing of customer nodes is done utilizing a modified Clarke and Wright heuristic developed for the multiple vehicle routing problem by Benton and Srikar (13).

Tillman's approach is used in the multi-depot case for modifying the costs due to the following reason. The savings scheme utilized in vehicle routing was developed by Clarke and Wright (32) for the single depot routing problem. The Clarke and Wright travel cost savings is given by the following equation (see Figure 3.1).

\[ s_{ij}^1 = C_i^1 + C_j^1 - C_{ij} \]  

(1)

where

- \( s_{ij}^1 \) is the dollar savings obtained by serving customers \( i \) and \( j \) in a single tour from depot 1 rather than separately.
- \( C_i^1 \) is the cost of travelling to customer \( i \) from depot 1
- \( C_j^1 \) is the cost of travelling to customer \( j \) from the depot 1
$C_{ij}$ is the cost of travelling from customer $i$ to customer $j$

From the equation, it is easy to see that the further away a pair of neighboring customers are from the depot, the higher is the travel cost saving. The routes are built based on the rankings of the savings for all pairwise links of customers. In the multi-depot case this clearly leads to errors. From Figure 3.2 it can be seen that for the pair of customers $i$, $j$, maximum savings would occur when they are associated with depot 2 which is farthest away from them. However, it is also obvious that the actual distance travelled would be more in such a case, when compared to the distance travelled from the depot, which is nearest to the pair of customers. The result is higher actual travel distance (cost).

The reason for the error is that according to the basic Clarke and Wright savings equation, the alternative for routing customer $i$ in a link along with customer $j$, is to serve customer $i$ directly. While this is true in the single depot case, in the multi-depot case the alternative to the link $i$-$j$ from depot 2 is not directly serving $i$ from 2. Rather, the alternative should be to serve directly from the closest depot, depot 1 in this case, in Figure 3.2. The basic savings equation therefore needs to be modified for the multi-depot situation. This is achieved using Tillman’s (141) modified distance (cost) formula

$$
\bar{C}_{i}^{r} = \min_{s} C_{i}^{s} - (C_{i}^{r} - \min_{s} C_{i}^{s}) \quad s = 1, 2, \ldots, M \quad (2)
$$

where $\bar{C}_{i}^{r}$ is the modified cost of serving customer $i$ from depot $r$, given $M$ open depots. In Figure 3.2, then, $\bar{C}_{j}^{2} = 2 C_{j}^{1} - C_{j}^{2}$, assuming that $j$ is
(a) Radial Distance

Cost of travelling \( i \) and \( j \) = 

\[ 2C_i^1 + 2C_j^1 \]

Savings = \((2C_i^1 + 2C_j^1) - (C_i^1 + C_j^1 + C_{ij})\)

\[ S_{ij}^1 = C_i^1 + C_j^1 - C_{ij} \]

(b) Savings Link

Cost of travelling \( i \) and \( j \) = 

\[ C_i^1 + C_j^1 + C_{ij} \]

Figure 3.1 The Savings Approach for A Single Depot.
Fig. 3.2 The Multi-Depot Problem. (Illustration of the problem with the use of the standard savings approach).
closer to depot 1 than to depot 2. This modified value represents the savings in cost by travelling \( j \) in a route from depot 2, as opposed to travelling customer \( j \) on a straight and back basis from depot 1, which is the nearest to depot for customer \( j \). If \( r \) is the closest depot to \( i \), the modified cost is the true cost of \( r \) from \( i \). By this modification scheme, the alternative to serving customer \( i \) and \( j \) in a link from any depot is to serve \( i \) and \( j \) from their nearest depots, in a straight and back manner.

Using this modification scheme, the modified savings equation is

\[
S_{ij}^r = C_{ij}^r + C_{ij}^r - C_{ij}
\]

(3)

when customers \( i, j \) are considered in a link from depot \( r \).

The heuristic algorithm steps are outlined below. To start with, all sites are taken as open depots.

Define the following notation for the heuristic algorithm.

- \( H = \{h_i \mid i = R+1, \ldots, R+N\} \) is the set of all customers
- \( G = \{g_r \mid r = 1, \ldots, R\} \) is the set of all depots sites
- \( V = \{k \mid k = 1, \ldots, K\} \) is the set of all vehicles
- \( C_{ij} \) = cost of travelling from customer \( i \) to customer \( j \)
- \( C_{i}^r \) = cost of travelling from customer \( i \) from depot \( r \)
- \( I \setminus i = \{I\} \setminus \{i\} \) as set \( I \) minus element \( i \), for any set \( I \)
- \( I \cup i = \{I\} \cup \{i\} \) as set \( I \) plus element \( i \) for any set \( I \)
- \( S_{ij}^r \) = modified savings in cost obtained by serving customer \( i \)
  and \( j \) in a tour from depot \( r \) rather that the other depots
$I_1 =$ set of customers already assigned to routes

$I_2 =$ set of unassigned customers

$(I_1) U (I_2) = H$

$I^r =$ set of customers to be routed from depot $r$

$R_1 =$ set of open depots

$R_2 =$ set of closed depots

$(R_1) U (R_2) = G$

$Q_k =$ capacity of vehicle $k$

$RC^r =$ total routing cost at depot $r$

$TDC^r =$ total depot cost at depot $r$

$VC^r =$ total vehicle cost at depot $r$

$TSC(R_1) =$ total system cost for $R_1$ open depots

$F^r =$ annual fixed and operating cost of depot $r$, $r \in G$

$FREQ^{ir} =$ frequency of positive savings of customer $i$ with depot $r$

$q_i =$ periodic demand at customer $i$ per tour, $i \in H$

$LOAD_k =$ sum of all customer demands met by vehicle $k$

$V_r =$ set of all vehicles assigned to depot $r$. It is assumed that

the vehicles used are similar.

$V_a =$ set of available vehicles (unassigned)

$D_p =$ set of all end points of open routes, i.e. not connected to

the depot. $D_p = \{ e_{pk} \; \; p \in I_1, k \in V \}$
op_{ij}^r = \text{opportunity penalty for not connecting link i-j to best depot r but to next best depot}

P_{im}^r = \text{maximum opportunity penalty for customer i across all other customers for depot r}

M(x) = \text{Maximum value of x, } V x

Y(x) = \text{minimum value of x, } V x

dp^r = \text{overall penalty for keeping depot r closed}

Step 0: Initialization.

Set I_2 = H, I_1 = 0, R_1 = G, R_2 = 0, FREQ_{ir} = 0 V i,r

k = 1, V_r = 0 V r

Initially all customers are unassigned and all depots are open.

Step 1: Compute the modified costs of all customers i \in I, from the open depots r \in R_1 according to the modified distance (141) formula.

\overline{C}_i^r = Y(C_i^s) - (C_i^r - Y(C_i^s)) \quad s = 1, 2, \ldots \ldots M

where \overline{C}_i^r is the modified cost of serving customer i from depot r.

Step 2: Compute the modified savings array for each depot.

The modified travel cost saving is

\overline{S}_{ij}^r = \overline{C}_i^r + \overline{C}_j^r - C_{ij}, \quad V \ r \in R_1, \ i,j \in H

The modified travel cost savings are used in the next steps to create the routes from the open depots.
Step 3: Assign the customers to the depots, then route them.

For every \( i \in I_2, r \in R_1 \), compute \( FREQ_{ir} \)

For all \( i \), route customer \( i \) from depot \( r' \) if

\[
FREQ_{ir'} = M(FREQ_{ir}) \quad \forall r \in R_1
\]

\[
I_2 = I_2 \setminus i, \quad I^{r'} = I^{r'} \cup i
\]

Continue assigning customers until \( I_2 = 0 \)

In this step, the customers are assigned to depots. For all unassigned customers and open depots, a frequency of positive lavings of customer with depot are computed. Starting with the first unassigned customer, the customer is allocated to the depot with which its frequency is highest. The set of unassigned customers is updated to remove the assigned customer, and the set of assigned customers for the relevant depot is also updated. The process is continued until the set of unassigned customers is empty.

Step 4: Determine the number of vehicles needed, and assign vehicles to depots.

For every \( r \in R_1 \),

\[
\sum_{i \in I^r} q_i
\]

Start with first available vehicle.

\( k = 1, \quad V_r = V_r \cup \{k\}; \quad V_a = V_a \setminus k \)

Is \( Q_k > \sum_{i \in I^r} q_i \)

If yes, go to step 6. Else continue.
In this step, the total number of vehicles needed at the depot is determined, for each open depot. The total customer demand for the depot is computed by summing together the demands of all customers served by the depot. From the available vehicles, select the first vehicle and examine to see if the total demand at the depot can be satisfied by that vehicle, in terms of capacity. If not, then more vehicles are assigned to the depot in the next step. If capacity is sufficient, the routes will be formed next in step 6.

Step 5: Assign another vehicle to the depot.

\[ k = k + 1 \]. If \( k > K \), stop. The problem is vehicle infeasible.

Else, continue

\[ V_r = V_r \cup \{k\} \]

\[ V_a = V_a \setminus \{k\} \]

If \( \sum_{k \in V_r} Q_k > \sum_{i \in I_r} q_i \)

Go to Step 6,
else repeat Step 5.

The process of assigning vehicles to the depot continues until the sum of vehicle capacities exceeds the total customer demand at the depot.

Step 6: Route the customers. The routes are built simultaneously. The number of routes to be built is equal to the number of vehicles assigned to depot \( r \). All the link savings are ranked highest to lowest.
For each \( r \in \mathbb{R}_1 \) select \( i, j \) such that \( i, j \in I^r \) and \( \tilde{S}^r_{ij} = M(S^r_{ij}) \). 

Start route 1 by assigning link \( i, j \) to the first route (vehicle) from depot \( r \). Remove the selected link from further consideration.

Set \( I_1 = I_1 \cup \{i,j\}, I^r = I^r \setminus \{i,j\} \); set \( k = 1 \).

The vehicle is identified as:

Select first vehicle from \( V_r \) \( V_r = V_r \setminus k \)

\( \text{LOAD}_k = q_i + q_j \)

Is \( \text{LOAD}_k \leq Q^r \)? If yes, continue, else, select another vehicle which satisfies constraint.

Step 7: Start the next route.

Select \( i, j \) such that \( i, j \in I^r \), and

\( \tilde{S}^r_{ij} = M(S^r_{ij}) \) \( V i, j, \)

Set \( I_1 = I_1 \cup \{i,j\}; I^r = I^r \setminus \{i,j\} \)

\( k = k + 1 \)

Select vehicle \( k \) from available set \( V_r \)

\( V_r = V_r \setminus k \)

\( \text{LOAD}_k = q_i + q_j \)

Repeat Step 7 until \( V_r = 0 \). Then go to Step 8.
The routing part of the algorithm is accomplished one depot at a time. In Steps 6 and 7, for the current depot being considered, the initial links are established simultaneously for all routes which are to be formed from the depot. The number of routes equals the number of vehicles assigned to the depot. The routes are built in a simultaneous manner, since this leads to more balanced routes in terms of the number of customers, as compared to a sequential formation of routes. The advantage of this scheme has been shown (13). At this stage all routes are partial, i.e., they are not connected to the depot yet. The routes are considered open for further building at both ends. The set of assigned customers is updated to include the customers assigned to routes in these steps. The set of customers still to be routed from the depot is updated to remove the assigned customers. The vehicles are sequentially assigned to the routes as the routes are created, and the set of vehicles to be routed from the depot is updated to delete those which have been assigned to the partial routes. The loads for the customers already assigned to the partial routes are assigned to the respective vehicles.

Step 8: Check to see if each of the routes can be expanded by the addition of more customers subject to vehicle capacity.

If $I^r = 0$ go to Step 10, else

Add customer link $i-j$ to a route, such that $i$ or $j$ or both $\in I^r$

and $\delta_{ij}^r = M(\delta_{ij}^r) \forall i,j \in I^r$ over the list of available links.
If either customer i (or j) is already assigned to a route k, Check: $LOAD_k^+ + q_i$ (or $q_j$) \leq Q_k$.

If yes, add i (or j) to route of vehicle k

$LOAD_k = LOAD_k^+ + q_i$ (or $q_j$)

$I^r = I^r \setminus i$ (or j), $I_1 = I_1 \cup \{i, j\}$

Else delete link i-j from consideration.

If both customers i and j in selected link i-j are free, i.e., have not been assigned to any route, assign them to a route in the following manner.

From the list of partial routes, select that route which has an end point closest to either i or j. This partial route is the least added cost way of developing the route, and provides an improvement over the strictly simultaneous development of routes.

Add link i-j to route of vehicle k if

(i) $C_{me_{pk}}^m = Y(C_{me_{pk}}^m) m = i, j$

where $e_{pk}$ represents the end points on the partial route in existence, $V k \in V_r$

and

(ii) $LOAD_k + q_i + q_j \leq Q_k$

$I^r = I^r \setminus i,j \quad I_1 = I_1 \cup \{i, j\}$

$LOAD_k = LOAD_k + q_i + q_j$

and continue with Step 8,
Else go to Step 9.

In Step 8, the initial links are expanded to form routes by assigning each of the free customers to a route. All routes are formed simultaneously. The link with the maximum savings is selected from the available links. If in the link selected, one of the two customers is already assigned to a route, the link is assigned to the same route. However, if both the customers on the selected link are unassigned, the procedure explained above will be used. The load check is done to ensure that adding the customers to the route does not cause the total load to exceed the vehicle's capacity. If it does, then the customer is not added to the route, and Step 9 is considered. Otherwise, add the link, and update the free and assigned customer sets, and repeat Step 8.

Step 9: Check for an alternate route if the vehicle's capacity on the selected route is exceeded in Step 8 by adding free pair of customers i,j.

If for route k, \( \text{LOAD}_k + q_i + q_j > Q_k \), do not add customers i,j to the route.

Eliminate the end points of the partial route for k from consideration \( D_p = D_p \setminus \{ e_{pk} \} \) where \( e_{pk} \) = end points of the partial route for vehicle (route) k.

Find \( C_{me_{pk'}} = Y(C_{me_{pk}}) \) \( m = i, j, e_{pk} \in D_p \forall k \in V \)

Assign i,j to route \( k' \) with end point \( e_{pk'} \), selected above.
If $\text{LOAD}_{k'} + q_i + q_j < Q_{k'}$, add customer $i,j$ to route $k'$

$$\text{LOAD}_{k'} = \text{LOAD}_{k'} + q_i + q_j$$

$$I^r = I^r \setminus i,j \quad I_1 = I_1 \cup \{i,j\}$$

If $I_1 = H$ go to Step 10, else

Go to Step 8.

If in Step 8, the customers cannot be added because of least cost rule or vehicle capacity constraint, go to step 9 and examine the next best route according to the cost rule and vehicle capacity. Once the link is assigned to a route, the vehicle load for the route, the assigned and unassigned customer sets are updated, and Step 8 is repeated. When all customers are assigned, go to Step 10.

**Step 10:** $TDC^r = RC^r + VC^r + F^r$

The total depot cost at the current depot $r$ is computed as the sum of total routing cost, the total vehicle cost, and the depot cost.

**Step 11:** If $\bigcup_{r=1}^{R_1} I^r = 0$, continue, else go to Step 6 for next $r \in R_1$

Here, the union of all free customer sets is examined to see if any more depots have customers not assigned to routes. If there are, go back to Step 6 to assign the customers to routes. If all customers in the system have been placed on routes, go to the next step.
Step 12: TSC (Rₜ) = \sum_{r=1}^{Rₜ} TDₜᵣ

Compute the total system cost as the sum of the total depot costs for all open depots in the system.

At this stage, the routing of customers has been achieved for the current set of open depots. In the next stage of the algorithm, the set of open depots is evaluated to locate the depot to close. The closing of the depot is done in Steps 13 through 16.

Step 13: Calculate the penalty of closing one of the open depots.

For all i,j \in I, i \neq j

Find \( r^* \) such that \( S_{i,j}^{r^*} = M(S_{i,j}^r) \forall r \in k \)

Find \( r' \) such that \( S_{i,j}^{r'} = M(S_{i,j}^r) \forall r' \in R_{\perp} \setminus r^* \)

Compute \( o_{i,j}^{r^*} = S_{i,j}^{r^*} - S_{i,j}^{r'} \)

Set \( o_{i,j}^{r*} = 0 \)

and \( o_{i,j}^{r} = 0 \forall r \in R_{\perp} \setminus r^* \)

In Step 13, the penalty cost of closing the depot is calculated for all the open depots. The first step in finding the penalty cost is the computation of the opportunity penalty for each link i-j, over all open depots. The opportunity penalty for any link i-j is defined as the difference between the highest modified savings for the link over all open depots, and the
next highest savings ($r^*$ is the depot with the highest savings). Such a penalty would occur if the depot with the highest savings were to be closed, and the next highest savings depot was used. For the link $i$-$j$, the opportunity penalty is considered only for the highest savings depot $r^*$; it is set to zero for the reverse link $j$-$i$ for depot $r^*$ and for all the other open depots.

Step 14: For all $r \in R_1$

(a) $P_{im}^r = M(op_{ij}^r)$ if $i \neq j$; $i, j \in H$, is the maximum opportunity penalty for customer $i$ linked to any other customer for depot $r$, where $m = j$ for which $op_{ij}^r$ is maximum, $i \neq j$.

(b) Check to see if customer $i$ or customer $m$ appears in all $p_{im}^r > 0$ not more than twice.

If the above check holds true (i.e. less than equal to two), continue with Step 14(a).

If not, then for the customer $i^*$ who appears more than twice; find

$p_{yz}^r = Y(P_{im}^r)$ if $i^* = m$, $i \neq m$ and $y, z \in H$

If $y = i^*$ set $Y(P_{im}^r) = 0$ for

set $op_{yi}^r = 0$ for

If $z = i^*$ set $p_{yz}^r = 0$
set \( o^{r}_{yz} = 0 \)

For given \( i \), go to Step 14(a)

If all \( p^{r}_{im} \) have been computed go to Step 15.

The maximum opportunity penalty for a customer, across all its links is found for depot \( r \), where \( r \) is a member of the set of open depots. The maximum opportunity penalty is taken as the highest opportunity penalty a customer has across its links for the current depot \( r \). Concurrently, a check is made to ensure that in computing the maximum opportunity penalty for a given customer and depot, no more than two links of that customer with all other customers appear, since in a route a customer can have no more than two links.

Step 15: For all \( r \) in \( R \), compute overall depot penalty.

\[
d p^{r} = \sum_{i \in H} p^{r}_{im} - F^{r}
\]

For each of the open depots the overall depot penalty is computed as the sum of the maximum penalty costs over all the customers minus the fixed and operating costs of the depot. If a depot were to be closed, then the sum of the maximum penalty costs is the maximum amount in savings that would be lost. On the other hand, the depot costs represent the savings achieved from closing the depot. Therefore, the difference between the two is the overall depot penalty. In closing a depot therefore,
only negative values of $dp^r$ are considered, unless a fixed number of depots are desired open.

Step 16: Find $r^+$ such that $dp^r = Y(dp^r) \forall r \in R_1$

Close depot $r^+$

$R_1 = R_1 \setminus r^+, R_2 = R_2 \cup r^+$

If $R_1 = 0$ Stop

otherwise, set $I_2 = H, I^r = 0, \forall r \in R_1, I_1 = 0$

$FREQ_{ir} = 0, \forall i, r, V_r = 0, \forall r \in R_1, k = 1.$

Go to Step 1

The depot with the lowest overall penalty is closed. The set of open depots and closed depots are updated. The sets of assigned, unassigned customers, FREQ, and assigned vehicles are also reinitialized, and the algorithm shifts back to the routing stage. The algorithm alternates between the routing and depot closing stages until either all the depots have been closed, or a desired number of depots remain open.

The iterative procedure can be used to determine the minimum cost system by adjusting the desired number of open depots.

A macro flow diagram of the algorithm is shown in Figure 3.3.

3.3.2 The Savings-Add Model

The Savings-Add algorithm utilizes a scheme similar to the Savings-Drop approach. However, while the Savings-Drop model assumes that
START

INITIALIZE ALL DEPOTS OPEN

CALCULATE MODIFIED SAVINGS FOR ALL OPEN DEPOTS AND ALL CUSTOMERS

ASSIGN CUSTOMERS TO OPEN DEPOTS BASED ON MAXIMUM FREQUENCY OF POSITIVE SAVINGS

CREATE VEHICLE ROUTES FOR EACH OPEN DEPOT

COMPUTE TOTAL SYSTEM COST FOR THE CURRENT NUMBER OF OPEN DEPOTS

COMPUTE OPPORTUNITY PENALTY OF NOT USING BEST DEPOT FOR EACH CUSTOMER LINK I - J

COMPUTE MAXIMUM OPPORTUNITY PENALTY FOR EACH CUSTOMER WITH EVERY OPEN DEPOT

Figure 3.3 Macro Flowchart For The Savings-Drop Model
Figure 3.3 (Continued)
initially all feasible sites are open depots, the present approach assumes that all feasible sites are closed depots initially, and will proceed to open them one by one. In a system design where only a small number of depots will be established from a large set of feasible sites, this approach will be computationally easier and faster than the earlier savings-drop approach where all the depots would be open to start with. It could possibly lead to a better performance under such conditions.

The basis of this approach is the "travel time savings" idea of Clarke and Wright (32), which has been developed for vehicle routing problems and extended to the multi-depot case by Tillman (141). A similar savings approach has been used in the Savings-Drop model. The concept of modified costs and modified savings has already been treated in the previous section. Therefore, only the algorithm steps are discussed here.

The heuristic algorithm steps are outlined below. In the starting solution all the feasible sites are considered to be closed depots. Define the notation for the algorithm to be the same as for the Savings-Drop model. The steps for the Savings-Add model are:

**Step 0: Initialization.**

Set $I_2 = H, I_1 = 0, R_1 = 0, R_2 = G, \text{FREQ}_i = 0, V_{i,r}, k = 1$

$V_r = 0, V_{r}$,

In the initialization step all customers are declared unassigned and $R_2$, the set of closed depots is equal to $G$, the total number of sites.

The set of vehicles assigned to depots is also empty.
The first stage of the model consists of step 1 through step 5, in which the depot to be opened is selected.

Step 1: Compute the modified costs of all the customers $i \in H$, from all the closed depots $r \in R_2$ according to Tillman's modified cost formula

$$\bar{C}^r_i = Y(C^s_i) - (C^r_i - Y(C^s_i)) \quad V s \in R_2, r \in R_2$$

Step 2: Compute the modified savings array for each of the closed depots.

The modified travel cost savings are

$$\bar{S}^r_{ij} = \bar{C}^r_i + \bar{C}^r_j - C_{ij} \quad V r \in R_2, i, j \in H, i \neq j$$

Step 3: Calculate the penalty cost for not opening a depot; for all the closed depots. The first step is to calculate the opportunity penalty for each of the pairwise links of customers.

For all $i, j \in H, i \neq j$.

Find $r^*$ such that $\bar{S}^r_{ij} = M(\bar{S}^r_{ij}), \ V r \in R_2$

Find $r^'$ such that $S^r_{ij} = M(S^s_{ij}) \ V r \in R_2 \setminus r^*$

then, $op^r_{ij} = \bar{S}^r_{ij} - S^r_{ij}$

and $op^r_{ji} = 0,$

Also set $op^r_{ij} = 0, \ V r \in R_2 \setminus r^*$

In this step the opportunity penalty is computed for each pairwise customer link $i$-$j$. The opportunity penalty is defined
as the difference between the maximum savings for the i-j link over all closed depots being considered, and the next highest savings. The next highest savings is found by searching over the set of closed depots, excluding r, where r* is the depot with highest savings. By definition, the opportunity penalty is zero for link i, j for the rest of the depots, and for reverse link j, i for the depot r*.

Step 4a For all $r \in R_2$,

compute, $P_{im}^r = M (\delta P_{ij}^r)_{i \neq j; i, j \in H}$

as the maximum opportunity penalty for customer i over its links with all other customers m is defined as the customer j for which the opportunity penalty $\delta P_{ij}^r$ is maximum, $i \neq j$. This defines the maximum opportunity penalty which can be associated with any customer i and depot r. In this step, a check is also made to ensure that neither customer i nor m appears more than twice in the maximum penalty value links. This is done to ensure that no customer has more than two links associated with it, since in any route a customer has only two links.

If no customer appears more than twice for all such $P_{im}^r, i, m \in H$; go to step 5.

Step 4b Otherwise, for all customers i* who appear more than twice, find

$P_{yz} = Y(P_{im})_{V i \text{ or } m = i*; i \neq m \text{ and } V y, z \in H}$. 
if \( y = i^* \) set \( M(P_{im}^r) = 0 \), \( \forall m \), and set \( op_{yj}^r = 0 \), \( \forall j \)

if \( z = i^* \) set \( P_{yz}^r = 0 \), and set \( op_{yz}^r = 0 \)

Go back to the beginning of step 4a.

In the above step, for the customer \( i \) appearing more than twice, a search is carried out to determine the lowest maximum opportunity penalty \( P_{im}^r \) (since for each occurrence of customer \( i^* \), \( P_{im}^r \) is computed). The lowest penalty and the associated customer link are eliminated from further consideration. For the customer \( i \) (in \( P_{im}^r \)) whose link has been eliminated, a new link and thus, a new maximum opportunity penalty is computed. This step ensures that all customers have a maximum of two links with the customers, since in a route a customer can have only two links.

Step 5: For all \( r \in R_2 \), compute overall depot penalty for keeping the depot closed.

\[
dp^r = \sum_{i \in H} P_{im}^r - P^r
\]

The depot penalty is defined as the sum of the maximum opportunity penalties over all customers minus the fixed and operating costs of the depot. By opening a depot \( r \), \( dp^r \) amount of penalty is avoided.

Step 6: Find \( r^+ \) such that
The depot with the highest penalty is opened since it is desirable to avoid the penalty, open depot $r^+$. The depots with positive penalties are opened, unless a fixed number of depots are required in the system.

$R_2 = R_2 \setminus r^+ \quad R_1 = R_1 \cup r^+$

The sets of closed depots and open depots, $R_2$ and $R_1$ respectively, are updated. At this stage, the set of open depots has been updated (incremented by $r^+$) and the routing stage of the algorithm is implemented next for the new system of depots.

To allocate the customers to the depots, the following scheme is adopted.

Step 7: Compute $S^F_{1j} \quad V \ r \in R_1$

For the set of open depots $R_1$, the modified savings array is computed for all $r \in R_1$. These arrays are different from those computed in step 2, since now the open depots are considered.

Step 8: Rank the modified savings, highest to lowest, for each depot $r \in R_1$. Compute the frequency of positive savings for each customer and depot, and assign the customers to the depots as follows:

For every $i \in I_2$, $r \in R_1$, compute $FREQ_{ir}$.

For all $i$, allocate customer $i$ to depot $r'$ if
FREQ_{ir} = M (FREQ_{ir}) \forall r \in R_1.

Ties are broken arbitrarily.

Update the customer sets.

I_2 = I_2 \setminus i, I_r' = I_r' \cup i

Customer i is deleted from the set of unassigned customers and added to the set of customers assigned to depot r'.

Continue the step until I_2 = 0.

Step 9: Determine the number of vehicles needed, and assign vehicles to the depots.

For every r \in R_1,

\sum_{i \in I_r} q_i

Start with first available vehicle

k - 1, V_r = V_r \cup \{k\}, V_a = V_a \setminus k

Is Q_k > \sum_{i \in I_r} q_i

If yes, go to Step 11, else continue.

In this step, the total number of vehicles needed at the depot is determined for each open depot. The total customer demand for the depot is computed by summing together the demands of all customers served by the depot.

From the available vehicles, select the first vehicle and examine to see if the total demand at the depot can be satisfied by that vehicle, in terms of capacity. If not, then more
vehicles are assigned to the depot in the next step. If capacity is sufficient the routes should be formed next.

Step 10 Assign another vehicle to the depot.

\[ k = k + 1 \]

If \( k > K \), stop. The problem is vehicle infeasible.

Else continue.

\[ V_r = V_r \cup \{k\}, \quad V_a = V_a \setminus \{k\} \]

If

\[ \sum_{k \in V_r} Q_k > \sum_{i \in I_r} q_i \]

Go to Step 11,

else, continue with Step 10.

The process of assigning vehicles to the depot continues until the sum of the vehicle capacities assigned to the depot exceeds the total customer demand at the depot.

Step 11 Route the customers. The routes are built simultaneously with the number of routes to be built equal to the number of vehicles assigned to depot \( r \). All the link savings are ranked, highest to lowest. In this step the initial link or partial route is developed.

For each \( r \in R_1 \) select \( i,j \) such that \( i,j \in I^r \)

and \( \bar{S}_{ij}^r = M(\bar{S}_{ij}) \)

Start route 1 by assigning link \( i,j \) to the first route (vehicle) from depot \( r \). Remove the selected link from further consideration.
Set $I_1 = I_1 \cup \{i,j\}$. $I^r = I^r \setminus \{i,j\}$, $k = 1$

Select the first vehicle from $V_r$, the set of vehicles assigned to depot $r$. 

$V_r = V_r \setminus k$

$LOAD_k = q_i + q_j$

Is $LOAD_k \leq Q_k$. If yes, continue, else select another vehicle which satisfies constraint.

**Step 12** Start the next route.

Select $i,j$ such that $i,j \in I^r$, and 

$S^r_{ij} = M(S^r_{ij})$ from the links available in the ranked list.

Set $I_1 = I_1 \cup \{i,j\}$; $I^r = I^r \setminus \{i,j\}$

Select $k = k + 1$ vehicle from available set $V_r$

$V_r = V_r \setminus \{k\}$

$LOAD_k = q_i + q_j$

Repeat Step 12 until $V_r = 0$. Then go to Step 13.

The routing part of the algorithm is accomplished one depot at a time. In Steps 11 and 12, for the current depot being considered, the initial links are established simultaneously for all routes.

The set of all assigned customers is updated to include the customers assigned to the open routes in these steps. The set of customers still to be routed from the depot is updated to remove
the assigned customers. The vehicles are assigned to the routes and the set of vehicles to be routed from the depot is updated to delete those vehicles which have been assigned to the open routes. The loads for the initial pair of customers already assigned to the routes are added to the respective vehicle's load. At this step all initial partial routes have been established for the depot.

Step 13 Check to see if each of the partial routes can be expanded by the addition of more customers subject to vehicle capacity.

If \( R = 0 \), go to Step 15, else

Add customer link \( i-j \) to a route, such that \( i, j \) or both \( i \) and \( j \in R \)

and \( S_{ij}^R = M(S_{ij}) \) over the list of available links, \( R \).

If either customer \( i \) (or \( j \)) is already assigned to a route \( k \),

Check: \( LOAD_k + q_i \) (or \( q_j \)) \( \leq Q_k \)

If yes, add \( i \) (or \( j \)) to route of vehicle \( k \)

\( LOAD_k = LOAD_k + q_i \) (or \( q_j \))

\( I_k = I_k \setminus i \) (or \( j \)). \( I_1 = I_1 \cup \{i \text{ or } j\} \)

Else delete link \( i-j \) from consideration.

If both customers \( i \) and \( j \) in selected link \( i-j \) are free, i.e., have not been assigned to any route, assign them to a route in the following manner.

From the list of partial routes, select the route which has an end point closest to either end of the link \( i-j \). This partial
route is the least added cost way to developing the route, and provides an improvement over the strictly simultaneous development of routes.

Add link i-j to route of vehicle k if

(i) \( C_{m_i} = Y(C_{m_e}) \) \( m = i,j; e \in D_p \), where \( D_p \) has been defined as the set of end points on partial routes in existence, and

(ii) \( \text{LOAD}_{k} + q_i + q_j \leq Q_k \)

Then, \( I^r = I^r \setminus i,j; \quad I_1 = I_1 \cup \{i,j\} \)

\( \text{LOAD}_{k} = \text{LOAD}_{k} + q_i + q_j \)

continue with Step 13

else, go to Step 14.

In Step 13, the initial links are expanded to form larger partial routes by assigning each of the free customers to a route. All routes are formed simultaneously. The link with the maximum savings is selected from the available links, according to the savings method. If the link is selected, one of the two customers is already assigned to a route, the link is assigned to the same route. However, if both the customers on the selected link are unassigned, the procedure listed and explained above will be used. The load check is done to ensure that adding the customers to the route does not cause the total load to exceed the vehicle's capacity. If it does, then the customer is not added to the route, and Step 14 is considered. Otherwise,
add the link, update the free and assigned customer sets, and repeat Step 13.

Step 14 Check for an alternate route if the vehicle's capacity on the selected route is exceeded in Step 13 by adding a free pair of customers \( i, j \).

If for route \( k \), \( \text{LOAD}_k + q_i + q_j > Q_k \), do not add customers \( i, j \) to the route.

Eliminate the end points of the partial route for \( k \) from consideration since the customers cannot be added. \( D_p = D_p \setminus (e_{pk}) \) where \( e_{pk} \) = end points of the partial route for vehicle (route) \( k \).

Find \( C_{me_{pk}} = Y(C_{me_{pk}}) \) \( m = i, j; e_{pk} \in D_p, V k \in V_r \).

Assign \( i, j \) to route \( k' \) with end points \( e_{pk} \) selected above.

if \( \text{LOAD}_{k'} + q_i + q_j < Q_{k'} \), add customer \( i, j \) to route \( k' \)

\( \text{LOAD}_{k'} = \text{LOAD}_{k'} + q_i + q_j \)

\( \Gamma = \Gamma \setminus i, j \quad I_1 = I_1 \cup \{i, j\} \)

Go to Step 13.

If in Step 13, the customers cannot be added because of either the least cost rule or vehicle capacity constraint, in this step examine the next best route according to the cost rule and vehicle capacity. Once the link is assigned to a route, the vehicle load for the route, the assigned and unassigned customer sets are updated, and Step 13 is repeated.
Step 15 \( TDC^r = RC^r + VC^r + Fr \)

The total depot cost for depot \( r \) is computed as the sum of routing, vehicle acquisition, and depot costs.

Step 16 If \( U_1^r = 0 \) continue to Step 17, else go to Step 4 for the next depot \( r \) in \( R_1 \).

At this step, if all customers have been placed in routes continue to next step, otherwise route the customers still to be assigned to routes from the other open depots.

Step 17 \( TSC(R_1) = \sum_{r=1}^{R_1} TDC^r \)

The total system cost for \( R_1 \) open depots \( TSC(R_1) \) is the sum of the total depot costs for all open depots.

Step 18 If \( R_2 = 0 \) stop, else reinitialize the customer sets and go to Step 1.

All depots have been opened if \( R_2 \) is zero.

At this stage, the search has been completed over all depots. Alternatively, one can stop at a desired number of depots. A macro flow diagram of the algorithm is presented in Figure 3.4.

3.3.3 The Cluster-Routing Model

This heuristic algorithm has been developed on the assumption that if the customers are distributed in groups (clusters), a solution procedure which locates and routes accordingly should be more efficient.
START

INITIALIZE ALL DEPOTS CLOSED

CALCULATE MODIFIED SAVINGS FOR ALL CLOSED DEPOTS IF CUSTOMERS WERE TO BE ROUTED FROM THEM

COMPUTE OPPORTUNITY PENALTY OF NOT USING THE BEST DEPOT FOR EACH CUSTOMER LINK I - J

COMPUTE MAXIMUM OPPORTUNITY PENALTY FOR EACH CUSTOMER WITH EVERY CLOSED DEPOT

COMPUTE THE SUM OF THE MAXIMUM OPPORTUNITY PENALTY OF ALL CUSTOMERS FOR EACH CLOSED DEPOT

OPEN THE DEPOT WITH THE LARGEST PENALTY SUM

COMPUTE THE MODIFIED SAVINGS FOR ALL OPEN DEPOTS AND ALL CUSTOMERS

Figure 3.4 Macro Flowchart For The Savings-Add Model
ASSIGN CUSTOMERS TO OPEN DEPOTS ON THE BASIS OF MAXIMUM FREQUENCY OF POSITIVE SAVINGS

CREATE VEHICLE ROUTES FOR EACH OPEN DEPOT

COMPUTE TOTAL SYSTEM COST FOR THE GIVEN NUMBER OF OPEN DEPOTS

SELECT THE SYSTEM CONFIGURATION WITH LOWEST SYSTEM COST (OR WITH REQUIRED NUMBER OF DEPOTS)

STOP

Figure 3.4 (Continued)
The procedure identifies the cluster of customers by first generating the underlying minimal spanning tree of the network of customers, separating it into the desired number of clusters, and then refining the clusters further. The number of depots to be located is equal to the number of clusters identified, and is selected as the site nearest the cluster centroid. Routing is performed using a polar coordinates based technique. The routing is achieved by first selecting the customers to be included in a single route based on their polar coordinates, and then applying an optimizing travelling salesman solution method to obtain the best route.

The heuristic algorithm steps are outlined below. The heuristic does not sequentially open or close depots, however, the number of open depots in the system can be specified, thus the heuristic can be evaluated for any given system.

Define the notation for the algorithm to be the same as used in the earlier algorithms.

In addition, the following definitions are made.

\[ \text{NCL} = \text{Number of Clusters (depots)} \]
\[ \text{QQ}(I), \text{PP}(I) = \text{identifying numbers for two clusters (or points) from the customer set, which are closest to each other.} \]
\[ \text{T}(I) = \text{minimal distance between the two clusters.} \]
\[ \text{P}(i) = \text{cluster number of customer } i, \ i \in H. \]
\[ \text{LN} = \text{distance norm; Euclidean (LN = 2) is used.} \]
\[ \text{S}(j,L) = \text{centroid coordinate of clusters, } j = 1, \ldots, \text{NCL} \]
\[ \text{ANG}_{r}^{i} = \text{angle of customer } i \text{ with depot } r, \text{ polar coordinate.} \]
$\text{RAD}_i^r$ = radius of customer $i$ with depot $r$, polar coordinate.

$X(i,L)$ = coordinates of customers and depots $i = 1, \ldots, R+N$, $L=1, LN$

$E(j)$ = sum of distances between the cluster centroid and its cluster elements for cluster $j$, $j = 1, \ldots, NCL$

Step 0: Initialization

Set $I_2 = H$, $NCL = 1$, $k = 1$

Set all customers to be unassigned, the number of clusters (open depots) equal to one, and the first vehicle to be available for routing.

Step 1: From the location coordinates $X(i,L)$ of all customers and depot sites, calculate the Euclidean distance between all customers, and depot sites. The distance between any two points $i$ and $j$ is given by

$$d_{ij} = \left(\|X(i,1) - X(j,1)\|^2 + \|X(i,2) - X(j,2)\|^2\right)^{1/2}$$

Step 2: Obtain the Minimal Spanning Tree (MST) for the set of customers.

The technique used for clustering the customers is a density search technique, which seeks regions of high density in the data. The measure used in the clustering technique is Euclidean distance. A clustering algorithm is used to generate the minimal spanning tree (MST) for all customers, using single linkage clustering. Identifying numbers $QQ(I)$, $PP(I)$ are obtained for 2
clusters with minimal distance $T(I)$. These are used to construct the tree.

Step 3: Number of depots to be opened = $NCL$

Step 4: Divide the MST into $NCL$ partitions based on the identifying numbers $QQ(I)$ and $PP(I)$. The $QQ(I)$, $PP(I)$ numbers indicate which of the customers are closer to each other, the distances being recorded in $T(I)$ in a monotonic manner. Approximately, an equal number of customers are assigned to each of the clusters. Thus, initial $P(i)$ are obtained.

Step 5: The final clusters of customers are obtained. The partitioning of customers, $P(i)$ is such that

$$E(j) = \sum_{P(i)=j} \left[ \sum_{L=1}^{2} (S(j,L) - X(i,L))^2 \right]^{1/2}$$

is minimized, $j = 1, \ldots, NCL$

Through exchange of cluster members from initial partitions, $P(i)$ changes to a different final configuration. The coordinates of the cluster centroid, $S(j,1)$ and $S(j,2)$ are also obtained. The criterion minimizes the sum of the distances from the cluster center.

Step 6: For $j=1, NCL$, find depot site location $r$ with coordinates $X(i,L)$ such that

$$\left[ \sum_{L=1}^{2} (S(j,L) - X(i,L))^2 \right]^{1/2}$$

is minimized

$$R_1 = R_1 \cup r; \quad R_2 = R_2 \setminus r$$
The depot site nearest to the cluster center is located for each of the clusters as the depot to open.

Step 7: Update the customer and depot sets.

\[ I_2 = H, I_1 = 0, r = 1, \text{NCL}; \text{ if } P(i) = r, \text{ then } i \subseteq I^r. \]

The customer sets, set of open depots are updated. If the customer belongs to cluster \( r \), then the customer is added to the set of customers to be routed from depot \( r \).

Step 8: At this stage, the locations of open depots, and the customers to be routed from these depots are known. Start the routing as follows:

Find the number of vehicles needed at each depot \( r, r \notin R_l \).

Select first available vehicle from set of vehicles \( V_a (k - 1) \)

For depot \( r \) compute \( \sum_{i \in I^r} q_i \)

Is \( Q_k > \sum q_i, V \in I^r \)

If yes, then the vehicle meets all the customer demands.

\( V_r = V_r \cup (k) \quad V_a = V_a \setminus k \)

Go to Step 10, else continue with step 9.

The total number of vehicles needed at the depot are to be determined for each open depot. The total customer demand for the depot is computed by summing together the demands of all customers assigned to the depot.

From the available vehicles, select the first vehicle and examine to see if the total demand at the depot can be met by
that vehicle, in terms of capacity. If not, then more vehicles are assigned to the depot in the next step. If capacity is sufficient, the routes are formed next.

Step 9: Assign another vehicle to the depot.

\[ k = k + 1 \]

If \( k > K \), stop. The problem is vehicle infeasible.
Else continue.

\[ V_r' = V_r \cup \{ k \}, \quad V_a = V_a \setminus \{ k \} \]

Is \[ \sum_{k \in V_r'} Q_k > \sum_{i \in I^r} q_i \]

If yes, go to Step 10.
Else continue with Step 9.

The process of assigning vehicles to the depot continues until the sum of the vehicle capacities assigned to the depot exceeds the total customer demand at the depot.

Step 10: Route the customers for each depot.

For each \( r \in R_1 \), \( V_i \in I^r \), compute the polar coordinates \( \text{ANG}_{i}^r \) and \( \text{RAD}_{i}^r \) using the depot location as the center.

Rank order \( \text{ANG}_{i}^r \) lowest to highest.

Step 11: Select first vehicle \( k \) from the assigned set \( V_r' \). \( V_r' = V_r \setminus \{ k \} \)

Place customers on the first route as follows.

Start with the customer with lowest \( \text{ANG}_{i}^r \) for all \( i \in I^r \)

If \( \text{LOAD}_k + q_i > Q_k \) do not add customer to the vehicle's route,
go to Step 12,
else, place the customer on a list of customers whose demands will be met in a route by vehicle k.

\[ \text{LOAD}_k = \text{LOAD}_k + q_i \]

\[ I^R = I^R \setminus i \quad I_1 = I_1 \cup i, \quad I_2 = I_2 \setminus i \]

Continue with Step 11

Step 12: Route the customers assigned to vehicle k by Little's penalty tour building TSP approach.

Step 13: \( k = k + 1 \)

If \( I^R = 0 \), go to Step 14, else go to Step 11

The next vehicle is drawn from the list of vehicles assigned to depot r and again customers are assigned to its routing list as before and the route created.

Step 14: \( TDC^R = R^C^R + V^G^R + F^R \). The total depot costs for depot r are computed.

Step 15: If \( U \bigcap \bigcup_{r \in R_1} I^R \neq 0 \) go to Step 8.

The union of all sets for customers assigned to depots \( I^R \) is examined to determine if any customers remain assigned to depots but have not been placed on routes. This is done after routing from each depot is completed. If customers still remain to be routed from any depot, go back to Step 8 and create the routes from that depot.
Step 16: \[ TSC(R) = \sum_{r \in R_1} TDC^r. \]

The total system cost is obtained by summing open depots' cost.

Step 17: \[ NCL = NCL + 1 \]

If \( NCL \leq G \), go to Step 1,

else stop.

If the number of clusters is less than or equal to the total number of depot sites, continue with the algorithm for the new system configuration.

Thus, all system configurations (number of depots) are examined in the algorithm's run. The minimum cost system is selected. Alternatively, the algorithm can be run for a desired number of open depots.

3.4 Validation of the Models.

The optimal and heuristic procedures developed are validated by solving small test problems and manually solving the same problems. This is done to ensure that the logic of the algorithms is programmed correctly, and that there are no programming errors. A further validation check is made by using the test problem used in Perl's (127) study. The data from his study is used for input to the optimal and the heuristic models. This test problem is then run on the models. The optimal model lower bound solution, which is obtained by relaxing the integer constraints, is the same as in Perl's (127) study. Further, all three heuristic algorithms developed in this study provide the same solution to the test problem as has been obtained by Perl using his approach.
START

START WITH ALL OPEN DEPOTS (M)

FIND CUSTOMER, DEPOT SITE COORDINATES. CALCULATE THE EUCLIDEAN DISTANCES

GENERATE THE MINIMAL SPANNING TREE FOR CUSTOMER NETWORK. OBTAIN THE INITIAL CLUSTERS

USE INITIAL CLUSTERS FOR LOCATION-CLUSTERING ALGORITHM. OBTAIN FINAL CLUSTERS. # OF CLUSTERS = M

SET K = 1

SELECT CLUSTER K. LOCATE DEPOT AT THE SITE WHICH IS NEAREST THE CLUSTER CENTER

COMPUTE POLAR COORDINATES FOR CUSTOMERS IN THE CLUSTER USING DEPOT AS ORIGIN. RANK BY ANGLE

GROUP THE CUSTOMERS IN RADIAL RAYS FROM DEPOT SUCH THAT THE VEHICLE CAPACITY IS NOT EXCEEDED

Figure 3.5 Macro Flowchart For The Cluster-Route Model
ROUTE THE GROUPED CUSTOMERS USING LITTLE'S PENALTY TOUR BUILDING APPROACH

COMPUTE TOTAL SYSTEM COST FOR DEPOT, ADD TO SYSTEM COST.

SELECT NEXT CLUSTER

\[ K = K + 1 \]

ARE ALL CLUSTERS EVALUATED (IS \( K = M \))

YES

RECORD SYSTEM COST \( M = M - 1 \). REDUCE # OF DEPOTS BY 1

SELECT THE SYSTEM CONFIGURATION WITH LOWEST SYSTEM COST (OR WITH REQUIRED NUMBER OF DEPOTS)

STOP

Figure 3.5 (Continued)
3.5 Sample Results of a Test Problem

The heuristic results are shown for a small test problem. The test problem has eight customers and two depot sites. The customer coordinates are given in Table 3.2. The depot coordinates are shown in Table 3.3 and other problem parameters in Table 3.4. The system configuration for the optimal model and the three heuristic algorithms are shown in Figures 3.7 through 3.10. Figure 3.11 represents the sequential model solution.

For small problems of this size the heuristics are quite fast and run to completion in 2 to 3 seconds on the PRIME 550 mini computer. For this problem all three heuristics as well as the sequential model obtained both the one depot system and the two depot system solution in a total of two seconds each. For similar problems the optimal model takes a considerably longer time ranging from 200 to 1200 seconds for most problems. The amount of time taken to obtain the optimal solution grows at a much faster rate than increase in problem size. All optimal solutions have been obtained on the PRIME 550 minicomputer using SCICONIC, a fast mathematical programming package. In terms of time, therefore, the heuristics take less than one percent of the time taken to reach the optimal solution on the average.

Table 3.4 provides a comparison of the optimal and heuristic solutions for a test problem. The criterion used is the percent the total system cost is greater than the optimal. All heuristics perform well in this particular case; the results are within 1.13 percent of the optimal. This indicates the efficiency of the heuristics for small test problems, and a comparison of solution times.
A formal experiment is conducted to evaluate the performance of the heuristics in comparison to the optimal model. The experimental design and details of the experiment are developed in the next chapter. The results of this experiment are then presented in Chapter 5. In all test problems, the heuristics performed reasonably well, being close to optimal solution. Heuristic solution details to the small test problem are provided in Appendix E.

It has been shown in this chapter that due to the computational complexity of the problem, optimal solutions to large problems are infeasible using a mixed integer program. The performance of the heuristics is therefore evaluated by comparing the heuristic solution against the solution obtained from applying the sequential model to larger problems. An experiment is designed to compare the heuristic algorithms to the sequential model. The details of the experiment and the factors used in the study are developed in the next chapter.
### Table 3.2

Customer Data for Small Problem

<table>
<thead>
<tr>
<th>CUSTOMER</th>
<th>X-COORDINATE</th>
<th>Y-COORDINATE</th>
<th>DEMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>92</td>
<td>162</td>
<td>112</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>86</td>
<td>112</td>
</tr>
<tr>
<td>3</td>
<td>83</td>
<td>43</td>
<td>62</td>
</tr>
<tr>
<td>4</td>
<td>69</td>
<td>212</td>
<td>145</td>
</tr>
<tr>
<td>5</td>
<td>42</td>
<td>50</td>
<td>135</td>
</tr>
<tr>
<td>6</td>
<td>94</td>
<td>118</td>
<td>128</td>
</tr>
<tr>
<td>7</td>
<td>74</td>
<td>64</td>
<td>71</td>
</tr>
<tr>
<td>8</td>
<td>69</td>
<td>194</td>
<td>54</td>
</tr>
</tbody>
</table>

### Table 3.3

General Parameters for the Small Problem

Depot Data for the Small Problem

<table>
<thead>
<tr>
<th>NUMBER</th>
<th>X COORDINATE</th>
<th>Y COORDINATE</th>
<th>COST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>117</td>
<td>174</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>79</td>
<td>29</td>
<td>33</td>
</tr>
</tbody>
</table>

Vehicle Data for the Small Problem

| Vehicle Capacity | 714 |
| Vehicle Cost    | 26  |
TABLE 3.4

Comparison of the Optimal and Heuristic Solutions using Total System Cost

<table>
<thead>
<tr>
<th>MODEL</th>
<th>SYSTEM COST</th>
<th>% SOLUTION IS HIGHER THAN OPTIMAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal</td>
<td>265</td>
<td>0.0</td>
</tr>
<tr>
<td>Heuristic 1</td>
<td>268</td>
<td>1.13</td>
</tr>
<tr>
<td>(Savings Drop)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heuristic 2</td>
<td>268</td>
<td>1.13</td>
</tr>
<tr>
<td>(Savings Add)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heuristic 3</td>
<td>266</td>
<td>0.38</td>
</tr>
<tr>
<td>(Cluster Route)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential</td>
<td>268</td>
<td>1.13</td>
</tr>
</tbody>
</table>
Figure 3.6 Small Test Problem
Figure 3.7 Solution from the Optimal Model
Figure 3.8 Solution from the Savings Drop Heuristic for the Small Problem
Figure 3.9 Solution from the Savings Add Model for the Small Problem
Figure 3.10 Solution from the Cluster Routing Model for the Small Problem
Figure 3.11 Solution from the Sequential Model
CHAPTER 4

RESEARCH PLAN

In this chapter a plan is developed to examine the heuristic algorithms. The interaction of the other factors used in the study which have been outlined in Chapter 1 also needs to be examined with respect to heuristic performance. Very few test problems exist in the literature, and the inability to develop an optimal solution procedure for all but extremely small problems leads to difficulties in the evaluation of the heuristic procedures. Even for the small test problems, the mixed integer program model assumptions are necessarily restrictive in character. To evaluate the heuristic procedures' performance in relation to the optimal model, similar assumptions are imposed on the heuristics.

The optimal model solution can be considered to be a lower bound on the solution value of the heuristic procedure. Thus, for small problems the heuristics can be compared against the optimal solution. For larger problems, a different benchmark needs to be developed for evaluation of the heuristic performance. It is expected that a simultaneous location-routing method will outperform a conventional design approach, which is based on sequential decisions on location of depots and routing of vehicles from them.
The conventional design approach has been based on the actual process of first establishing the depot sites using one of the existing location modeling procedures, and then using a routing procedure to route the vehicles from the established depots. This approach therefore considers distribution costs in the location stage to be based on radial distances from the depots to the customers, and in the second stage replaces them with the distribution costs based on routes. The heuristic procedures developed consider the distribution costs to be based on routes formed in the location stage itself. They should, in general, outperform the conventional design model where the costs considered are only estimates in the location stage. This provides a method by which the simultaneous location-routing procedures can be judged. If these heuristic procedures are better than the sequential scheme, the implication is that better distribution system designs can be achieved, along with significant savings over the conventional method. The sequential scheme can therefore be used as a benchmark to compare the simultaneous procedures.

4.1 Evaluation of the Heuristics

For problems of small size, the optimal solution is used as a benchmark for comparison. The heuristics are evaluated in terms of total system cost, which is taken to be the sum of depot costs, vehicle costs, and distribution (routing) costs. Rather than consider the absolute value of system cost for comparison, the percent deviation of the heuristic solution from the optimal solution is used. The use of percent
deviation as a measure removes the effect of the arbitrary nature of costs that may be associated with the absolute cost measure.

For problems of larger size, the only reasonable benchmark which can be developed is the sequential model solution. A lower bound based on the optimal model was also considered as a benchmark. A lower bound may be obtained in the case of larger problems, by relaxing the integer constraints in the mixed integer programming optimal model to obtain a Linear Programming (LP) solution. However, using the SCICONIC (135) software package, even the LP solution is limited to relatively small sized problems (around 35-40 customers) because of the large number of variables involved.

Besides size, the second issue to examine in using the LP solution as a lower bound is the quality of the lower bound. A preliminary examination of the lower bound (LP solution) obtained for the small test problems shows that the quality of the lower bound is not very good. The MIP optimal solution cost was on the average around 33 percent higher than the LP solution cost. The variance of the difference between the MIP optimal solutions and the LP solutions was also high for the test problems. This is because most of the variables used are zero-one, and in the LP solution the variables have fractional values. This has also been indicated in other studies (127). Considering these two aspects of the LP solution obtained by relaxing the integer constraints in the MIP model, the LP solution did not appear to be a good lower bound on the solution obtained from the heuristic procedure. The LP solution was therefore not included in the study as a benchmark.
In order to run the optimal and heuristic procedures, some assumptions were made regarding the distribution system and the operating conditions. These were made to make the physical distribution system design amenable to model solving without increasing the models' complexity, and for convenience in obtaining the solutions. These assumptions are stated.

The assumptions made in the formulation and solution of the mixed integer programming model have been stated in Chapter 3. Some further assumptions are made in the solution procedures.

The following are assumptions about the customer demand. (1) A single customer's demand is considered to be less than or equal to the vehicle's capacity. If any customer's demand is more than a truckload (TL), a full truckload of the customer's demand will be assumed to be satisfied in a separate tour consisting of a single truck routed to the customer. The remaining demand at the customer, which will then be less than a truckload (LTL) will be satisfied in a tour along with other customers. (2) Split deliveries are not allowed. This means that any given customer's demand which is LTL is satisfied by a single truck (route) only.

In order to keep the number of variables low in the model and make the problem computationally feasible, the depot site set is considered to be disjoint from the customer set. This implies that the model does not consider demand at depots since such demand would be automatically satisfied at a very small cost. This is not a restrictive assumption since customer demand at depot sites can be modelled by introducing a
very small distance between the customer and the depot site to make them distinct in the model. The depots are considered to be uncapacitated.

Three vehicle related constraints have been provided in the general model in Chapter 3. These relate to vehicle capacity, the maximum distance a vehicle may travel in a route, and the maximum amount of time a vehicle can stay on a route. Not all of these constraints need exist in every distribution system. The most common constraint of these is the vehicle capacity constraint. Consequently, to decrease the computational effort involved in solving the optimal location-routing model, only the vehicle capacity constraints will be used in this study. The significance of the problem is not reduced since most distribution systems do not require all of these constraints to exist.

The heuristic procedures can be made more general and flexible than the optimal model. Different system configurations (number of depots) can be evaluated in a single run. Further, in the case of the heuristics, it is not necessary to have a requirement that all open depots have customers allocated to them. If in the system configuration there is an open depot with no customers assigned to it, the depot can be closed and its cost eliminated resulting in greater savings for the system under consideration. In the case of comparison of the heuristics with the optimal model, the requirement of customers being routed from open depots is imposed to ensure that the comparison is valid.

4.2 Research Questions and Criterion Development

In order to examine the heuristics' performance, several other factors are also examined. An experimental design is set up to examine
under what set of conditions the heuristics differ from the benchmark. The design is also used to investigate the factor settings over which the use of simultaneous location-routing procedures would be most valid, when compared to the sequential process. To enable setting up the experiment the following research questions are developed along with the criterion used.

The purpose of developing the research questions is to determine the effectiveness of the heuristic procedures. Their effectiveness is defined in terms of how much better they are as compared to the sequential benchmark, or how much worse they perform as compared to the optimal solution used as benchmark. The questions are also aimed at examining the effect of the various environmental factors used in the study on the heuristics' performance. Specifically, the following research questions are addressed in this study.

(1) How effective are the heuristic procedures compared to the solutions obtained from the optimal model used as benchmark. The objective in this case is to first establish that the heuristics do perform well and have merit before they are used for larger problems.

(2) How effective are the heuristic procedures when compared to the sequential decision making solution procedure used as benchmark. In this analysis, it will be seen how well the heuristics work on larger, more realistic problems. Further, it is to be investigated if the heuristic solution procedures provide an improvement significant in terms of cost savings over the sequential procedure.

(3) How do the various environmental factors used affect the heuristic solutions? Specifically, under what conditions (factor settings) are the heuristic procedures significantly better than the sequential procedure? Again, under what conditions does a given heuristic procedure perform best?

(4) Is any one heuristic procedure clearly better than the other procedures, or should a procedure be selected according to the prevailing environmental conditions?
In order to answer these questions, there has to be some criterion on which these comparisons and evaluations can be made. In this research, cost has been used as a (dependent variable) basic criterion. The specific criterion used in the study is the percent difference in total costs between the heuristic solution and the benchmark. While it may be true that not all distribution systems are concerned with minimizing cost, in most distribution systems cost does emerge as the single most quantifiable and clear objective. It may also be used as a surrogate for other criteria in system design.

When the literature on location and routing is reviewed, in almost all of the studies including practical design problems, cost has been used as an explicit criterion. In these studies annual cost figures have been used for comparison purposes and fixed long term costs have been amortized to yearly figures. In accordance with the established practice in the earlier studies (43), this study also uses amortized costs. The cost of locating and operating a depot can be amortized to an annual basis, given an estimate of the expected life of the facility. Similarly, the cost of acquiring the vehicles is amortized over the useful life of the vehicles. The usual approach taken in obtaining distribution costs is to develop the routing cost as a function of the distance travelled. Given the average demand of the customers, the average number of routes travelled by a vehicle in a year, and length of the average route, the total distance travelled in a year can be obtained. A conversion factor can be used to convert distance (in miles) to routing costs (in dollars). This has been the general practice in solving models found in the literature (43, 127). A similar approach is
adopted in this study. A more complete justification has been made earlier in Chapter 1 for the use of cost as a criterion, with reference to industry practice.

For all industries, the data on the distribution costs and the average number of warehouses used by a typical firm in that industry has been compiled by Lalonde and Zinszer (92). Their study deals with aggregate figures and does not consider specific existing literature models. Based on these figures a measure for cost per unit distance is developed for this study, similar to the cost per ton-mile concept. This measure is developed as follows. For any given industry, with the industry average number of depots open in the system, the minimum straight out and back distance to all customers is computed, which represents the total minimum distance travelled (cost) according to the existing warehouse location models in the literature. The cost per unit distance is obtained by dividing the distribution cost for the industry by the total minimum distance travelled. The average depot cost is obtained from the knowledge of total warehousing cost and average number of warehouses (depots). These calculations are made to obtain realistic data in this study. For any given firm, the actual costs would be known, or can be estimated from company data.

4.3 Experimental Factors

In this study the simultaneous location-routing problem addresses the cost structure and the environmental conditions of distribution systems. The environmental conditions for a distribution system which are used in this study are the spatial distribution of customers, the
distribution system size characteristics as defined by the number of depots in the system, and the size of the routes as defined by the average number of customers per route. The cost structure of a distribution system here refers to the breakdown of the total system cost between location costs and distribution (routing) costs. The cost structure is defined in this study as the ratio of warehousing costs to distribution (routing) costs for any given industry.

In order to answer the research questions two different experiments are conducted. The experimental designs used are factorial designs. The first experimental design is used to evaluate the performance of the heuristics with the optimal solution used as benchmark. Since the problem size is small in this case, not all of the environmental factors are used.

The second experimental design is used to evaluate the performance of the heuristic procedures for larger problems with the sequential model as benchmark. In addition, the effect of the environmental factors on the heuristics' solutions is examined in this experiment. The experiment is conducted in two stages. First, a pilot study is conducted to evaluate the factors to be included in the experiment. It is also run to establish the sample size needed to obtain a satisfactory level of power in the experiment. Next, the actual experiment is run.

The experimental factors are examined in detail in the following sections to assess their expected impact on the heuristics at various levels. Experimental designs are then constructed for the various cases and specific values assigned to the factor levels in the experiment. Other factors in the distribution system design have been fixed or
randomized in this study. These are also examined.

4.3.1 Cost Structure

The cost structure is of interest because the two cost components, location costs and distribution cost are expected to influence the effectiveness of the location-routing heuristics. The ratio of the warehousing (location) to distribution cost varies among different industries. The ratios of annual total warehousing cost to annual LTL (Less Than Truckload) distribution cost for the various industries are shown in Table 4.1. These industry figures are derived from data on distribution system components obtained from a 1976 survey of physical distribution in various industries (92). The relevant cost data has been presented earlier in Table 1.1, including warehousing and LTL distribution costs. From the available data, a ratio of warehousing cost to LTL distribution cost is developed. The derived ratios for various industries are shown in Table 4.1. The distribution cost used in the derivation of the ratios is the approximate percentage of the total distribution cost used in LTL shipments, since the routing problem arises only in the context of LTL shipments. For the cost structure the factor, specific cost ratios identified in Table 4.1 are used as levels of the factor in the study. These levels represent the cost ratios which would be typical of, for example, the pharmaceuticals industry (0.26), the paper industry (0.85) and the consumer merchandising industry (1.56). These ratios represent the low, middle, and high end respectively of the range for most industries. It is hypothesized that the effectiveness of the heuristic procedures may vary when compared to
Table 4.1. Cost Ratios for Various Industries.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Cost Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 All Manufacturing</td>
<td>0.9830</td>
</tr>
<tr>
<td>Chemical and Plastics</td>
<td>0.8997</td>
</tr>
<tr>
<td>Food</td>
<td>0.8620</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>0.2604 +</td>
</tr>
<tr>
<td>Electronics</td>
<td>8.0280</td>
</tr>
<tr>
<td>Paper</td>
<td>0.8620</td>
</tr>
<tr>
<td>Machine and Machine Tools</td>
<td>0.8529 +</td>
</tr>
<tr>
<td>Other</td>
<td>0.4464</td>
</tr>
</tbody>
</table>

2. All Merchandising

<table>
<thead>
<tr>
<th>Industry</th>
<th>Cost Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer</td>
<td>1.5627 +</td>
</tr>
<tr>
<td>Industrial</td>
<td>0.7378</td>
</tr>
</tbody>
</table>

* Derived Warehousing to Distribution Cost Ratio; using LTL Approximation.
+ Ratios used in the study as levels of the cost structure factor.
the sequential method, from when the cost ratio is low (lower warehousing cost compared to distribution cost; the routing problem dominates) to when the cost ratio is high (higher warehousing cost compared to distribution cost; the location problem dominates).

4.3.2 Spatial Distribution

It is hypothesized that in a physical distribution system the spatial distribution of customers could affect the solution procedures in terms of their effectiveness. In this study, the spatial distribution of customers is varied according to whether the customers are uniformly distributed in the system, or are clustered. It is expected that if the customers appear in clusters or groups, rather than being uniformly dispersed over the entire system, the performance of heuristics which account for this factor may be better than the performance of other heuristic algorithms compared to the sequential model.

To examine this factor, customer positions are generated to be within one of a desired number of clusters, keeping approximately an equal number of customers in each cluster. The cluster sizes are kept approximately equal so as to block the effect of unequal cluster sizes from entering the experiment. Two levels of spatial distribution are used. One level of spatial distribution is a uniform distribution of customers throughout the system. The other level of spatial distribution considered is a system with distinct clusters of customers. The specific levels used in the different experimental designs vary and are defined in the experimental designs in Tables 4.2, 4.3, and 6.1.
4.3.3 Number of Depot Sites

The number of available depot sites in the system could affect the performance of the location-routing models. In terms of computational efficiency, i.e. the cost of the system and the time taken to evaluate the entire system, the heuristic procedures may behave differently. The number of depot sites in the system could also affect the total physical distribution cost in the system, since with a larger number of depot sites more alternative system designs can be evaluated, therefore the heuristics may perform better.

In a system where location costs are low compared to routing costs, it is expected that there would be more open depots. The availability of a larger number of sites will allow more flexibility in choice and lower overall system cost. For systems with high location cost compared to routing cost, a lower number of depots would be expected. Two different factor level settings are established for this factor in the experimental designs. The specific levels are discussed in the context of the designs in Tables 4.2, 4.3, and 6.1.

4.3.4 Route Size

Route size has been defined as the average number of customers served in a route. This factor is also expected to affect the performance of the heuristics. By having smaller route sizes, more routes are created for the same total number of customers in the system. Some routing heuristics may be more efficient than others for different route sizes. This difference in performance may be reflected in the case of the location-routing models too. They may behave differently in terms
of computational efficiency. In this study, the route size is set by varying the vehicle capacity. Given the average customer demand and vehicle capacity, an average route size can be established. Two different route size settings are used, the specific values of which are indicated in the experimental designs in Tables 4.2 and 4.3.

4.3.5 Other Factors

There are other factors in the design of physical distribution systems besides the environmental factors used. Some of these have been fixed to specific values in the study. The fixed factors are the number of customers; set at eight for the first experiment (comparison to the optimal model), and at forty for the second experiment (comparison to the sequential model). The maximum number of vehicles used is set to two in the first, and ten in the second experiment. Vehicle capacity, and vehicle costs have also been fixed. The number of clusters in the clustered spatial distribution of customers has been set to five, the average number of warehouses across all industries.

In addition, there are some factors which have been ignored in this study. These include percentage of customers to be served (all customers are assumed to be served), frequency of service of each customer, and the stochastic nature of demands and costs. Further, the capacity of the depots have been ignored.

Finally, some factors have been randomized. These include the location of customers, the location of depot sites, customer loads, and location costs. The factors which have been fixed are usually within the control of the firm and can be fixed by the firm. Those ignored, if
included would add to the complexity of the models and would be an extension of the present study. Those randomized have been done so to provide replications in the experiment.

4.4 Experimental Design

To evaluate the heuristics' performance, the different experimental designs stated earlier are considered. One experiment is designed to evaluate the heuristics in comparison to the optimal model. The criterion used is the percent difference of the heuristic solution over the optimal solution. This experiment is performed to show the effectiveness of the heuristic procedures' solutions for small problems, in comparison to the optimal model solution. Once the effectiveness of the heuristic procedures is examined for small problems, their performance is evaluated for larger problems by means of another experiment. First, the experimental design for the comparison of the heuristic model to the optimal model is discussed.

4.4.1 Comparison of the Heuristic Models to the Optimal Model

In the comparison of the heuristic models to the optimal the main difficulty faced is problem size. It has been shown in Chapter 3 that only problems of a small size can be solved optimally. The problems used in this experiment are therefore of a small size, set at eight customers, two depot sites, and two vehicles. This leads to a total of one hundred and eighty two zero-one integer variables. Problems of a larger size than this are not considered because even slight increases in the problem parameters will lead to much larger problem size and
difficulty in obtaining an optimal solution in reasonable computing time (less than sixty minutes on the PRIME 550). Given the selected size of the problems for the experiment, it is not feasible to select more than one level for the factors used in the study, except for the cost structure. The number of depot sites has been set at two; with eight customers the spatial distribution factor cannot be clearly defined, and the route size parameter also cannot be effectively varied. The three levels of the cost ratio factor discussed earlier will be used in this experiment.

Each of the heuristic algorithms is run using the same experimental design. In this experiment, the design is a factorial design with cost ratio as the only factor.

A total of thirty problems are used in this experiment, ten for each of the three different cost structure levels. Each of the problems is run on the three location-routing algorithms, and the optimal model, resulting in a total of 120 observations. The problems are generated using a random number generator to generate the Euclidean distances (costs) of a uniformly distributed set of customers and depots. Customer demands are also randomly generated. The cost ratios and problem size parameters are the inputs required to generate the problems. By using a different random number seed every time, ten different problems are generated at each of the three cost ratios.

In the absence of any known studies or data, the level of power in the experiment is planned for by assuming one standard error worth of difference in the group means is satisfactory. For this difference, using the noncentral F distribution, at a level of significance of 0.05
TABLE 4.2

Comparison of Heuristics to Optimal Model

Factor Levels for the Factorial Design (1 Factor)

<table>
<thead>
<tr>
<th>Cost Structure</th>
<th>(Location to Routing Cost)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>Value</td>
</tr>
<tr>
<td>low</td>
<td>0.26</td>
</tr>
<tr>
<td>medium</td>
<td>0.85</td>
</tr>
<tr>
<td>high</td>
<td>1.56</td>
</tr>
</tbody>
</table>

Other Problem Parameters:

**Spatial Distribution**
Spatial distribution of customers and depot sites is uniform.

**Number of Depot Sites**
Number of depot sites is kept at 2.

**Route Size**
Not set at any level; maximum of 8 customers per route.

*The problem size is too small to achieve more than one level setting of these factors.*
the power of the experiment will be approximately 0.87. The experimental
design is shown in Table 4.2. The results of this experiment are
presented in the next chapter.

The comparison of the heuristic models to the optimal model is
followed by an analysis of the performance of the heuristics for larger
problems. First a pilot study is done. The experimental design for the
pilot study is presented in the next section. The heuristics are
compared to the sequential model in the study.

4.4.2 The Pilot Study

For large problems, the benchmark used is the solution from the
sequential model. An experimental design is set up to examine the
effectiveness of the heuristic procedures in comparison to the benchmark
under the different environmental conditions which may be obtained by
the various combinations of the experimental factors.

The first step taken in setting up the experiment to compare the
heuristic location-routing procedures to the sequential model is to
perform a pilot study. The pilot study is performed using the three
locating-routing heuristics and the sequential model at selected
settings of the factors to be used in the formal experiment. The pilot
study is done for two purposes. The first purpose is to determine if all
the factors hypothesized as affecting the heuristic's solution are
really significant. In general, this approach is important if there are
too many factors involved in the preliminary investigation and only an
important few are desired in the full study. In this study, the number
of factors is not too many. However, if the unimportant factors can be
dropped, it makes the final experiment smaller and makes the analysis of the factors simpler by eliminating higher order interactions.

The second purpose in conducting a pilot study is to examine the feasibility of running a large experiment such as the one entailed in a complete factorial design. If the run times and computations involved are excessive, a smaller experimental design such as a fractional factorial design may have to be used. The pilot study is also useful in determining the sample size needed to achieve a desired level of power in the experiment. The results of this pilot study are used to set up the final experimental design. The factors used in the pilot study and the settings used for the factors are given in Table 4.3. This is followed by a general experimental design for the full experiment.

Cost structure has been considered a desirable factor to retain in the final experiment. Using the different cost ratios will allow for the examination of the effectiveness of the heuristics for firms in various types of industries. Therefore, cost structure is kept fixed at one value in the pilot study. The ratio of total location cost to the total distribution cost computed as described earlier, is set at 1.00. This is close to the average cost ratio of all firms in manufacturing industries (See Table 4.1). The remainder of the study consisted of the following factors along with their settings.

For the spatial distribution factor two levels are set. The first level of spatial distribution used is a uniformly distributed customer and depot sites data set. The second level of the spatial distribution factor used is a clustered distribution, with five equal sized clusters of customers. Within individual clusters, the distribution of the
customers is kept uniform. A total of five clusters is selected since on an industry wide basis, the average number of warehouses established by a firm is five for the industries being considered. The depot sites are also distributed according to the clusters, with an equal number per cluster.

The depot sites factor is also considered at two levels. A high level of ten depot sites is selected, which is twice the average number of depots for the industries used in the study. A low level of five depot sites is used, which is the same as the average number of depots for the industries used in this study.

The route size parameter factor is also kept at two levels, a high level of an average of ten customers per route, and a low level of five customers on the average per route.

The above factor levels selected leads to a full factorial design with the three factors as between-subjects factors, i.e. a $2^3$ design. The experimental design is shown in Table 4.3.

There are a total of four algorithms included in the study; the three heuristic location-routing algorithms and the sequential model to be used as a benchmark. Every algorithm is run on the same set of problems to make the evaluation of performance valid, and in comparing the heuristics performance.

For the pilot study a total sample size of forty problems are generated. In the problem generation, the coordinates of the customers and of the depot sites are generated from a uniform distribution of random numbers. From the coordinates, Euclidean distances are computed
**TABLE 4.3**

Comparison of Heuristics to the Sequential Model: The Pilot Study

Factor Levels for the Complete Factorial Design

<table>
<thead>
<tr>
<th>Spatial Distribution</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
</tr>
<tr>
<td>uniform</td>
<td>1 cluster of customers</td>
</tr>
<tr>
<td>clustered</td>
<td>5 clusters of customers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Depot Sites</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>5 depot sites</td>
</tr>
<tr>
<td>high</td>
<td>10 depot sites</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Route Size</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>5 customers/route</td>
</tr>
<tr>
<td>high</td>
<td>10 customers/route</td>
</tr>
</tbody>
</table>

Other Parameters

*Cost Structure* *(Location to Routing Cost)*

Only one value set at 1.00 is used in this study

* Cost Structure is to be retained in the final study
TABLE 4.4

Main Effects for the $2^3$ Factorial Design used in the Pilot Study*

<table>
<thead>
<tr>
<th>Effect</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Spatial Distribution</td>
</tr>
<tr>
<td>B</td>
<td>Number of Depot Sites</td>
</tr>
<tr>
<td>C</td>
<td>Route Size</td>
</tr>
</tbody>
</table>

*Three different experiments are set up for the three heuristics.
between all customers and all depot sites. In all the problems used in the pilot study, the number of customers in the system is kept at forty.

4.4.3 Comparison of the Heuristic Models to the Sequential Model

Once the pilot study is completed, its results will be used to formalize the final experiment for this research study. The factors which are considered significant in the pilot study will be included in the final study. These results are presented in the next chapter.

The cost ratio factor is set at three levels in the final experiment. These are the ratios of 0.26, 0.85, and 1.56, which have been used earlier in the experimental comparison of the heuristic models to the optimal model.

The spatial distribution factor is kept at the same levels as in the pilot study. The reasons for the level settings used have been outlined earlier. For the industries which are typical of the above cost ratios, the average number of depots is around five. The number of depot sites factor is also be maintained at two levels. A high level of twelve depot sites is used, which is two to six times the industry average for all the industries selected in this experiment (based on their cost structures). A low level of eight sites is selected. The route size parameter, is dropped from the final experiment for reasons discussed in the next chapter.

The results of the experiments are presented in the following chapters. First, in chapter 5 the results of the comparison of the heuristics to the optimal model are presented. After the effectiveness of the heuristic procedures is established for the case of small
problems, the procedures are compared to the sequential model, which is considered to be the usual decision making method in practice. In making the comparison, the results of the pilot study are presented in chapter 5 to establish the final experimental design which is used to make the comparison of the heuristics to the sequential model. The results of this experiment are presented in chapter 6, and finally, conclusions are drawn based on the experiments which are conducted.
CHAPTER 5
PRELIMINARY ANALYSIS OF HEURISTICS' PERFORMANCE

In this chapter, the effectiveness of the heuristics is examined within the framework set in the previous chapter. The chapter is organized in two parts as follows. First, the heuristic models are compared with the optimal model to illustrate the performance of the heuristic models for small problem size, and to establish their merit. Next, the results of the pilot study are presented to examine the proposed factors and to allow for the final experimental design to be established, including the sample size.

The first experiment conducted in this study addresses the issue of the effectiveness of the heuristic solutions when compared to the optimal solution as benchmark. The experimental design has been described in Chapter 4 with cost structure as the only factor. This has three levels and the sample size is kept at ten. Optimal solutions to the zero-one mixed integer programming location-routing model are obtained for the set of thirty different problems using SCICONIC, a commercial mathematical programming package, and are used as benchmark. The three heuristic procedures are then applied to the same set of test problems and their solutions are compared with the benchmark solutions. In addition, a fourth heuristic which selects the best of the three heuristic procedures as the solution for each problem is also examined.
The idea is to select the heuristic which best performs in any given case and compare it with the optimal solution. In the case of the small test problems, the heuristic models have been shown to take a very small fraction (less than 5 percent) of the time taken by the optimal model to solve the problem. The total time taken to run all three heuristics is much less than the time taken to run the optimal model, and selecting the best heuristic solution provides a measure of the combined performance of the heuristics.

In order to normalize the results of the heuristics' and optimal solutions, the performance of the heuristic models is measured by the percent deviation of the heuristic model solution from the optimal solution. Percent deviation is defined as the ratio of heuristic solution minus optimal solution to the optimal solution. The model solution itself is measured as the total cost of the system, where the total cost is comprised of the depot, the vehicle and the routing costs.

In the following sections the results of each heuristic procedure are discussed. The basic statistics and the analysis of variance from the factorial designs used in the experiments are summarized and conclusions drawn. In the experiments, the assumptions of the analysis of variance, i.e., homogeneity of variance and normality are not of serious concern since the sample sizes are equal and the F-test is fairly robust under this condition.

5.1 Comparison of Savings-Drop Heuristic with the Optimal Model

A one-way analysis of variance is performed to examine the difference in the performance of the Savings-Drop heuristic for the
various ratio levels. The optimal solution is used as the benchmark. The criterion used in the comparison is the percent deviation of the heuristic from the benchmark. The solution results for the heuristic are displayed in Table 5.1. Overall, the heuristic procedure performs an average of 4.81% above the benchmark, the range of the percent deviation is 0.0 to 20.88. In the case of low cost ratio the average percent deviation of the heuristic solution from the benchmark is 3.77. The range of the percent deviation is 0.0 to 20.88. In two of the ten problems, the heuristic derives the benchmark solution. For the medium level of cost ratio the average percent deviation is 7.36, the range being 0.0 to 14.92. Again, the heuristic derives the benchmark solution in two of the ten problems. In the case of high level of cost ratio, the average percent deviation of the heuristic solution from the benchmark is 3.31, the range is 1.00 to 8.05 percent. In over two-thirds of all problems in the experiment, the percent deviation is less than 10.0.

In small problems, in almost all cases the number of depots opened by the heuristic procedure is the same as the optimal model. In almost all solutions, only one of the two depot sites in the problems is opened. Therefore, the difference in the performance can be attributed to the use of a heuristic procedure (the modified Clarke - Wright procedure) for routing the customers. It has been previously shown that the best available vehicle routing heuristic has a potential performance of 50% worst than the optimal. Therefore, such a difference in performance between the heuristic and the optimal model is expected. In small problems, a good heuristic should perform close to the optimal. Such performance is established for this heuristic since it is within 5%
Results of the Savings-Drop Heuristic

Total System Cost (in % deviation from Benchmark)*

<table>
<thead>
<tr>
<th>Cost Structure level</th>
<th>Average (%Deviation)</th>
<th>Std. Dev. (%Deviation)</th>
<th>Range (% Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>3.77</td>
<td>6.18</td>
<td>0.00 20.88</td>
</tr>
<tr>
<td>medium</td>
<td>7.36</td>
<td>5.58</td>
<td>0.00 14.92</td>
</tr>
<tr>
<td>high</td>
<td>3.31</td>
<td>2.68</td>
<td>1.00 8.05</td>
</tr>
<tr>
<td>Overall</td>
<td>4.81</td>
<td>5.21</td>
<td>0.00 20.88</td>
</tr>
</tbody>
</table>

Cell sample size is 10 at each level.

* % Deviation = \((\text{Heuristic Solution Cost} - \text{Benchmark Solution Cost}) \times 100 / \text{Benchmark Solution Cost}\)
### TABLE 5.2

Analysis of Variance Table for Savings-Drop Heuristic

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>97.96</td>
<td>48.98</td>
<td>4.20</td>
<td>0.166</td>
</tr>
<tr>
<td>Error</td>
<td>27</td>
<td>688.44</td>
<td>25.50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5.3

Planned Comparison Between the Group Means of Cost Structure

<table>
<thead>
<tr>
<th>Difference between means of Cost Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
</tr>
<tr>
<td>medium</td>
</tr>
<tr>
<td>high</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>low</th>
<th>medium</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>3.59</td>
<td>0.46</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>4.05</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value

\[ d = 5.74 \quad \text{at alpha} = .05 \quad \text{(experimentwise value of alpha)} \]

* Dunn's procedure is used.
of the optimal average.

The one-way analysis of variance is used to investigate if the variable cost structure can explain the variation in the difference between the heuristic and benchmark solutions. The analysis of variance table for the heuristic is given in Table 5.2. From the analysis of variance table it appears that the effect of cost structure is insignificant (the p value = 0.166). The power of the F-test is computed to be 0.42 at an alpha level of .05 using the tables of non-central F distribution. This is the power of the test that there is no difference between the group means of cost structure for the three levels of the factor, when the actual means are as found, i.e. 3.77, 7.36, and 3.31 respectively. These values represent the percent deviation of the heuristic solution from the optimal. Further, the coefficient of correlation between the criterion and cost structure is 0.35. Thus, there does not appear to be a strong relationship between the criterion (percent difference in cost between the heuristic and optimal solutions) and the factor cost structure.

Comparison among the means of the three groups of cost structure is carried out since it has been planned beforehand to investigate the difference between the three levels of cost structure. To carry out the three comparisons in this case, an a priori planned comparison method such as Dunn's multiple comparison procedure (114) may be used. The difference between group means is compared with a critical value termed 'd', in Dunn's procedure. A pair of group means are significantly different if the difference exceeds the critical value. The group means for the three levels have been shown in Table 5.1. Table 5.3 shows the
difference between the group means and \( d \), the critical difference value at an alpha level of 0.05. There does not appear to be any significant difference between any of the three cost structure levels.

Based on the preceding analysis, the following conclusions can be drawn regarding the performance of the Savings-Drop heuristic. For small problems, the heuristic works quite well, being within 5% of the optimal solution on the average. The number of depots opened is similar to the optimal solution. Further, in the case of small problems, the cost structure (of warehousing to distribution costs) does not affect the performance of the heuristic. There are no significant differences in its performance at the three levels. The implication is that the Savings-Drop heuristic seems to be a good heuristic to use in the practical situation, where it is not feasible to use the optimal procedure. Further, the use of the procedure seems equally applicable in the different industry environments as measured by cost ratio. The p-value for the test (0.166) does not indicate insignificance conclusively, and it may be that for larger problem size, differences may appear. Therefore, cost structure is retained as a factor, and the heuristic further evaluated for large sized problems.

5.2 Comparison of Savings-Add Heuristic with the Optimal Model

Table 5.4 shows the solution results for the heuristic. The average percent deviation from the benchmark solution for the Savings-Add approach is 4.84 percent, the range of the percent deviation is 0.0 percent to 20.88 percent. When analyzed further by the three groups of
cost structure, the results appear to be similar to those of Savings-Drop heuristic analyzed in the earlier section. Again, in two thirds of the runs the heuristic solution is within 10 percent of the benchmark solution.

In almost all cases the number of depots opened by the heuristic procedure is the same as the optimal model. The difference in performance compared to the optimal model can be attributed to the use of a heuristic procedure in the routing stage. Overall, on the average, this heuristic solution is only 4.84 percent higher than the optimal solution. This shows a good performance by the heuristic for small problems.

The analysis of variance table for the heuristic is given in Table 5.5. From the analysis of variance table it appears that the effect of cost ratio is insignificant (p = 0.116). Using the non-central F distribution tables, the power of this test has been computed to be 0.45 at an alpha level of .05. This is the power of the F-test which tests the hypothesis that there is no difference between the group means of cost structure, when the actual means are as found, in this instance, 3.49, 7.73, and 3.31 respectively for the three levels of cost structure. Further, the coefficient of correlation between the criterion and cost structure is 0.38. Thus, there does not appear to be a strong relationship between the criterion (percent deviation) and the cost structure factor.

Planned comparisons among the means of the three groups is carried out using Dunn's procedure. Table 5.6 shows the difference between the group means and the table value at an alpha level of 0.05. It appears
Results of the Savings-Add Heuristic

Total System Cost (in % deviation from Benchmark)*

<table>
<thead>
<tr>
<th>Cost Structure level</th>
<th>Average (%Deviation)</th>
<th>Std. Dev. (%Deviation)</th>
<th>Range (% Deviation)</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>3.49</td>
<td>6.18</td>
<td>0.00</td>
<td>0.00</td>
<td>20.88</td>
</tr>
<tr>
<td>medium</td>
<td>7.73</td>
<td>5.90</td>
<td>0.00</td>
<td>0.00</td>
<td>14.92</td>
</tr>
<tr>
<td>high</td>
<td>3.31</td>
<td>2.68</td>
<td>1.00</td>
<td>1.00</td>
<td>8.05</td>
</tr>
<tr>
<td>Overall</td>
<td>4.84</td>
<td>5.17</td>
<td>0.00</td>
<td>0.00</td>
<td>20.88</td>
</tr>
</tbody>
</table>

Cell sample size is 10 at each level

*%Deviation = (Heuristic Solution Cost - Benchmark Solution Cost) x 100

Benchmark Solution Cost
### TABLE 5.5

Analysis of Variance Table for Savings-Add Heuristic

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>124.92</td>
<td>62.46</td>
<td>2.34</td>
<td>0.116</td>
</tr>
<tr>
<td>Error</td>
<td>27</td>
<td>722.22</td>
<td>26.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5.6

Planned Comparison* Between the Group Means of Cost Structure

Difference between means of Cost Structure

<table>
<thead>
<tr>
<th></th>
<th>low</th>
<th>medium</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>4.24</td>
<td>0.18</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>4.42</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value

\[ d = 5.87 \text{ at alpha} = 0.05 \] (experimentwise value of alpha)

* Dunn's procedure is used.
that there is no significant difference between any of the three cost structure levels.

Based on the preceding analysis, the following conclusions can be drawn regarding the performance of the Savings-Add heuristic. For small problems, the heuristic works well, being within 5% of the optimal solution on the average. Cost structure does not significantly affect the performance of the heuristic. Therefore, the heuristic appears to be a good heuristic to use in practical situations where the optimal model cannot be applied, and seems equally applicable in different industry environments as measured by cost structure. There does not seem to be much difference from the Savings-Drop heuristic. However, the performance for large problems is still an issue, and the p-value of the test here (0.116) does not indicate insignificance conclusively. Therefore, the heuristic is retained for further evaluation on large problems.

5.3 Comparison of Cluster-Routing Heuristic with the Optimal Model

The results for the heuristic are shown in Table 5.7. The average percent deviation from the benchmark solution for the heuristic is 5.28 percent, the range of the percent deviation from the benchmark is 0.0 percent to 17.2 percent. Further analysis by the three groups of cost structure shows that the performance of this heuristic does not much differ from the other two heuristics. The average deviations for the low and medium levels of cost structure are similar to the other two heuristics, and for the high level the average percent deviation is slightly higher at 4.84 percent compared to the 3.31 percent deviation.
of the other two heuristics. In four-fifths of all runs, the heuristic solution is within 10 percent of the benchmark solution. Again, for this heuristic, in almost all cases, the number of depots opened by the heuristic's procedure is the same as the optimal model. The difference in performance is therefore due to the routing stage procedure. Overall, on the average, the heuristic is only 5.28% higher than the optimal solution, which shows a good performance by the heuristic for small problems.

The analysis of variance table is given in Table 5.8. From the analysis of variance table it appears that the effect of cost ratio is insignificant \( p = 0.162 \). Using the tables of non-central \( F \) distribution, the power of the test is computed to be 0.42 at an alpha level of 0.05. This is the power of the test that there is no difference in the three group means of the cost structure factor, when the actual means are 3.31, 7.68, and 4.84 respectively. Further, the coefficient of correlation between the criterion (percent deviation) and cost structure is 0.36. Thus, there does not appear to be a strong relationship between the criterion and the cost structure factor.

Comparisons among the three group means were carried out using Dunn's procedure. Table 5.9 shows the difference between the group means and also the critical 'd' value at \( \alpha = 0.05 \). There appears to be no significant difference between any of the cost ratio levels. This analysis indicates that cost structure does not affect the performance of the Cluster - Route heuristic significantly in the case of small problems, a finding similar to the other heuristics.
TABLE 5.7

Results of the Cluster Route Heuristic

Total System Cost (in % deviation from Benchmark)*

<table>
<thead>
<tr>
<th>Cost Structure level</th>
<th>Average (%Deviation)</th>
<th>Std. Dev. (%Deviation)</th>
<th>Range (% Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>3.31</td>
<td>4.30</td>
<td>0.00</td>
</tr>
<tr>
<td>medium</td>
<td>7.68</td>
<td>6.79</td>
<td>0.00</td>
</tr>
<tr>
<td>high</td>
<td>4.84</td>
<td>5.18</td>
<td>0.27</td>
</tr>
<tr>
<td>Overall</td>
<td>5.28</td>
<td>5.18</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Cell sample size is 10 at each level.

*Deviation = (Heuristic Solution Cost - Benchmark Solution Cost) x 100

Benchmark Solution Cost
### TABLE 5.8

Analysis of Variance Table for Cluster-Route Heuristic

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>98.25</td>
<td>49.13</td>
<td>1.95</td>
<td>0.162</td>
</tr>
<tr>
<td>Error</td>
<td>27</td>
<td>679.48</td>
<td>25.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5.9

Comparison Between the Group Means of Cost Structure

<table>
<thead>
<tr>
<th>Difference between means of Cost Structure</th>
<th>low</th>
<th>medium</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>4.37</td>
<td>1.53</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>2.84</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value

\[ d = 5.70 \] at alpha = 0.05 (experimentwise value of alpha)

* Dunn's procedure is used.
From the above analysis, the following conclusions can be drawn regarding the performance of the Cluster-Route heuristic. For small problems, the heuristic works as well as the other two considered; being 5.28% over the optimal solution on the average. Again, cost structure does not significantly affect the performance of the heuristic. The implication then is that for small problems, the heuristic performs similarly in all cost environments analyzed, and therefore can be used by any firm in these environments. However, the performance of the heuristic for larger problems still remains an issue, since for the problems analyzed here, the size is small and not much difference in performance may be expected for these simple problems. Therefore, this heuristic will be further run on larger problems.

5.4 Comparison of the Best Heuristic Rule with the Optimal Model

The best heuristic rule has been established as selecting the best solution among the three heuristics for each problem. The results from the best heuristic rule are shown in Table 5.10. The average percent deviation from the benchmark solution for the rule is 3.33 percent, the range of the percent deviation from the benchmark is 0.0 percent to 12.31 percent. An analysis by the three levels of cost structure shows that the performance of the best heuristic solution rule is better than any individual heuristic. In over four-fifths of all runs made, the percent deviation of this rule from the benchmark is less than 10 percent. The ranges of percent deviation in the various levels of cost ratio are also much narrower. For this heuristic rule which is a composite of the three heuristics developed, in almost all cases the
### TABLE 5.10.

Results of the Best Heuristic Solution Rule

Total System Cost (in % deviation from Benchmark)*

<table>
<thead>
<tr>
<th>Cost Structure level</th>
<th>Average (% Deviation)</th>
<th>Std. Dev. (% Deviation)</th>
<th>Range (% Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>0.77</td>
<td>0.63</td>
<td>0.00</td>
</tr>
<tr>
<td>medium</td>
<td>6.34</td>
<td>5.05</td>
<td>0.00</td>
</tr>
<tr>
<td>high</td>
<td>2.90</td>
<td>2.45</td>
<td>0.27</td>
</tr>
<tr>
<td>Overall</td>
<td>3.33</td>
<td>3.92</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Cell sample size is 10 at each level.

*% Deviation = \( \frac{(\text{Heuristic Solution Cost} - \text{Benchmark Solution Cost})}{\text{Benchmark Solution Cost}} \) x 100
### TABLE 5.11

Analysis of Variance Table for Best Heuristic Rule

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>158.06</td>
<td>79.03</td>
<td>7.43</td>
<td>0.003</td>
</tr>
<tr>
<td>Error</td>
<td>27</td>
<td>287.35</td>
<td>10.64</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5.12

Comparison Between the Group Means of Cost Structure

<table>
<thead>
<tr>
<th>Difference between means of Cost Structure</th>
<th>low</th>
<th>medium</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>5.57**</td>
<td>2.56</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>3.44</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value

\[ d = 3.71^{**} \quad \text{at alpha} = .05 \quad \text{(experimentwise value of alpha)} \]

* Dunn's procedure is used
depots opened under this rule is same as in the optimal model. The difference from the optimal model is due to the use of the modified Clarke-Wright approximate routing algorithm in the heuristics. Overall, however, the rule provides good results, being within 3.4 percent of the optimal solution on the average. The implication is that in the case of small problems, using the best of the three heuristics will provide good results.

The analysis of variance table is given in Table 5.11. From the analysis of variance table it appears that the effect of cost structure is significant when all heuristics are considered at the same time for their best solution in each case (p = 0.003). Using the non-central F distribution tables, the power of the test is computed to be 0.85 at an alpha level of .05. This represents the power of the test actually detecting the observed difference in the means, i.e. a difference of 5.57 between groups 1 and 2 under the test that there is no difference. Further, the coefficient of correlation between percent deviation and cost structure is 0.60, which indicates a moderate degree of relationship between the two.

Comparisons among the three group means were carried out using Dunn's procedure. Table 5.12 shows the difference between the group means and also the critical 'd' value at alpha = 0.05. There appears to be a significant difference between the low cost structure level and the medium cost structure level at a .05 significance level. This analysis indicates that cost structure may affect the overall heuristic performance as measured by the percent difference over the optimal, when
the best heuristic rule is employed in selecting the heuristic to use for a given problem.

From the preceding analysis, the following conclusions can be drawn regarding the performance of the best heuristic rule. For small problems, the heuristic works quite well, being within 3.4% of the optimal solution on the average. For this heuristic, cost structure factor does affect the performance of the heuristic. The analysis shows that the rule works particularly well for the low cost structure (0.26). At medium cost structure, the results are somewhat similar to the use of any individual heuristic. The implication here is that depending on the cost structure involved, the design of the distribution system should be achieved by examining different heuristics (as in low cost structure) and selecting the best, or by using any of the heuristics, (as in the medium structure). Since the best heuristic selection rule is only an average of 6.33% higher than the optimal solution in the worst case among the three cost ratio levels, it is likely that such a rule will perform well under all environmental conditions. However, problem size for this experiment is small, and the conclusions cannot be extended to large problems. Therefore, the rule will be further evaluated for large problems.

5.5 Comparison of the Sequential Model with the Optimal Model

Since the sequential model is used as a benchmark in the analysis of heuristic performance for large problems, the performance of the sequential model is analyzed in comparison to the optimal model. This helps to establish the validity of using the sequential model as
benchmark, especially if the sequential model also performs close to the optimal. The results for the sequential model are shown in Table 5.13. The average percent deviation from the benchmark solution for the rule is 6.50 percent, the range of the percent deviation from the benchmark is 0.0 to 28.97 percent. An analysis by the cost structure levels shows that the sequential model's performance is particularly poor for the low cost structure level, being 10.4% higher than optimal. The performance at the medium and high cost ratio levels is similar to the other heuristics. The variance in the percent deviation is also higher than the other heuristics, at all cost ratio levels. The sequential model performs slightly different from the heuristics; the number of depots opened under this rule differs from the optimal model in some problems. This may explain its poor performance at the low cost structure level. The other difference from the optimal model is in the use of a routing heuristic, as opposed to optimal routing of customers. Overall, however, the sequential model does not perform much worse on small problems, when compared to the other heuristics employed.

The analysis of variance table is shown in Table 5.14. From the analysis of variance table, it appears that the effect of cost ratio on percent difference of sequential model from the optimal model is significant \( p = 0.047 \). Using the non-central F distribution tables, the power of the test that there is no difference in the three group means of cost ratio is computed as 0.72 at an alpha value of 0.05. This would be the power of the test if in fact, the difference between the group means are as shown in Table 5.15. Further, the coefficient of correlation between the percent difference and Cost Structure is 0.45,
TABLE 5.13.

Results of the Sequential Model

Total System Cost (in % deviation from Benchmark)*

<table>
<thead>
<tr>
<th>Cost Structure level</th>
<th>Average (%Deviation)</th>
<th>Std. Dev. (%Deviation)</th>
<th>Range (% Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>10.42</td>
<td>9.72</td>
<td>1.13</td>
</tr>
<tr>
<td>medium</td>
<td>6.52</td>
<td>6.02</td>
<td>0.00</td>
</tr>
<tr>
<td>high</td>
<td>2.55</td>
<td>2.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Overall</td>
<td>6.50</td>
<td>7.26</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Cell sample size is 10 at each level.

* % Deviation = \( \frac{\text{Sequential Solution Cost} - \text{Benchmark Solution Cost}}{\text{Benchmark Solution Cost}} \times 100 \)
### TABLE 5.14

Analysis of Variance Table for Sequential Model

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>309.69</td>
<td>154.85</td>
<td>3.42</td>
<td>0.047</td>
</tr>
<tr>
<td>Error</td>
<td>27</td>
<td>1220.82</td>
<td>45.22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 5.15

Comparison Between the Group Means of Cost Structure

<table>
<thead>
<tr>
<th>Difference between means of Cost Structure</th>
<th>low</th>
<th>medium</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>3.90</td>
<td>7.87**</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>3.97</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value

\[ d = 7.64^{**} \quad \text{at alpha} = .05 \]

*Dunn's procedure is used*
which suggests only a slightly moderate relationship between percent difference and cost ratio.

Table 5.15 shows the difference between the group means and also the critical 'd' value at alpha = 0.05. There appears to be a significant difference between the low and high cost ratio levels. This indicates that cost ratio may affect the performance of the sequential model, with the model performing better at higher cost levels (i.e. higher warehousing to distribution costs).

From the above analysis the following conclusions can be drawn regarding the sequential model. For small problems, its overall performance is close to the heuristics, being only slightly worse. This suggests that the sequential model may be used as a benchmark for larger problems, since it has been shown to work well for smaller problems. The sequential model shows a clear pattern to the effect cost structure has on its performance. As the cost structure increases, the performance of the model improves in relation to the optimal. The implication is that at high cost ratio values, the sequential model may perform just as well as the heuristics developed. Therefore, the advantage gained from using the heuristics at such cost structures may be little. Whether such a trend holds true for larger problems should also be evaluated before definite conclusions can be reached.

5.6 Comparison of the Heuristics

A further comparison is made of the three heuristics and the sequential model to examine their behavior in the case of small problems. Such a comparison may reveal differences in performance of the
heuristics, which may serve as a basis for using the heuristics for larger problems. It will also allow for the comparison of the heuristics to the sequential model, and further validate the choice of the sequential model as a benchmark for large problems.

The Savings-Drop model and the Savings-Add model have similar solution logic but differ in their starting conditions. The former starts with all depot sites open and proceeds to close them one by one, whereas the latter starts with closed depots and opens them one by one. Overall, for small problems, the two heuristics' performance is quite similar, being the same on most test problems. However, both the heuristics are retained for further evaluation on large problems, because a research issue raised is that there may be a difference in performance based on the number of depot sites factor.

The Savings-Drop model is also compared to the Cluster-Routing heuristic here. The basic approach in the two heuristics is different. The comparison of the two heuristics shows that there is not much difference in performance. The Savings-Drop heuristic is slightly better at high cost ratio levels, while the other heuristic is better at low cost ratio levels. However, the difference between the heuristics at these levels is not significant. It may well be that the problem size is too small for such differences to be significant. Again, the spatial distribution factor has not been examined in this experiment. The Cluster-Routing model has been developed to account for clustered customer distribution. Therefore, this heuristic too is evaluated on large problems.
The Savings-Drop model, which showed the best performance of the three heuristics, though not by any significant amount is also compared to the sequential model here. Again, no significant difference in performance is found between the Savings-Drop and the sequential model, when a comparison of the means is done using a t-test with a significance level of 0.05. Neither is the difference significant at any of the three levels of cost ratio at alpha = 0.05. The most difference is found for the low cost ratio (p = 0.08).

Based on the analysis in the preceding sections as well as earlier conclusions drawn, it can be stated that the heuristics perform well for small problems. They perform better under certain cost ratios than others. The difference between the heuristics is not significant. However, it is expected that there may be some difference for large problems. Finally, the performance of the sequential model indicates that its performance is also quite good compared to the optimal model. Further, there are no significant differences observed between it and the heuristics. Therefore, it appears to be a reasonable benchmark to use for large problems.

5.7 The Pilot Study

The analysis in the previous sections indicates that the heuristics perform quite well when compared to the optimal solution for the case of small problems. The time required to obtain the solution is low (of the order of 2-4 seconds). Since the heuristics perform well for small problems, the next step is to investigate their performance for larger problems. The lack of optimal solutions for larger problems leads to the
The criterion used to evaluate the performance of the heuristics in the pilot study is the percent deviation from the sequential model's cost. Since the heuristic solution is expected to be lower than the benchmark, the criterion shows the percentage amount by which the heuristics are better than the benchmark. The design for the pilot study has been set up in the previous chapter and is a full factorial $2^3$ design with a sample of five. In investigating the results of the pilot study, simple statistics are first used to uncover the first order effects in the experiment. The results of the pilot study are summarized in Tables 5.16 through 5.18.

Table 5.16 summarizes the results for the Savings-Drop heuristic. The overall heuristic solution is 8.76 percent below the benchmark, with a range from -1.09 percent (the heuristic solution is 1.09 percent above the benchmark) to 22.74 percent. The heuristic performed equal to or worse than the benchmark in only two of the forty problems, with the worst case being 1.09 percent higher than the benchmark. The spatial distribution factor shows that there is a difference of more than 2.4 percent in the performance of the heuristic between the two levels of the factor. The heuristic also shows a difference in performance between the levels of the depot sites factor by over 2.1 percent. The performance of the heuristic is essentially the same across the different levels of the route size factor.
Table 5.16

Results of the Savings-Drop Heuristic

Total System Costs (in % Deviation* from Benchmark)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Average (of % Deviation)</th>
<th>Range (of % Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>low</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uniform</td>
<td>9.99</td>
<td>-1.09</td>
</tr>
<tr>
<td>Clustered</td>
<td>7.53</td>
<td>-0.62</td>
</tr>
<tr>
<td>Number of Depot Sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9.83</td>
<td>-0.62</td>
</tr>
<tr>
<td>10</td>
<td>7.69</td>
<td>-1.09</td>
</tr>
<tr>
<td>Route Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8.92</td>
<td>0.13</td>
</tr>
<tr>
<td>10</td>
<td>8.81</td>
<td>-1.09</td>
</tr>
<tr>
<td>Overall % Deviation</td>
<td>8.76</td>
<td>-1.09</td>
</tr>
</tbody>
</table>

* % Deviation = [(Sequential Solution - Heuristic Solution) / Sequential Solution] x 100
Table 5.17

Results of the Savings-Add Heuristic

Total System Costs (in % Deviation from Benchmark)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Average (of % Deviation)</th>
<th>Range (of % Deviation)</th>
<th>low</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uniform</td>
<td>6.74</td>
<td>-3.21</td>
<td>17.15</td>
<td></td>
</tr>
<tr>
<td>Clustered</td>
<td>4.89</td>
<td>-2.13</td>
<td>14.36</td>
<td></td>
</tr>
<tr>
<td>Number of Depot Sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7.35</td>
<td>-2.13</td>
<td>17.15</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>4.28</td>
<td>-3.21</td>
<td>16.88</td>
<td></td>
</tr>
<tr>
<td>Route Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.02</td>
<td>-0.75</td>
<td>17.15</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>5.60</td>
<td>-3.21</td>
<td>15.84</td>
<td></td>
</tr>
<tr>
<td>Overall % Deviation</td>
<td>5.60</td>
<td>-3.21</td>
<td>17.15</td>
<td></td>
</tr>
</tbody>
</table>
### Table 5.18

Results of the Cluster-Route Heuristic

Total System Costs (in % Deviation from Benchmark)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Average (of % Deviation)</th>
<th>Range (of % Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>low</td>
</tr>
<tr>
<td><strong>Spatial Distribution</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uniform</td>
<td>5.34</td>
<td>-2.63</td>
</tr>
<tr>
<td>Clustered</td>
<td>5.93</td>
<td>-1.75</td>
</tr>
<tr>
<td><strong>Number of Depot Sites</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.72</td>
<td>-2.63</td>
</tr>
<tr>
<td>10</td>
<td>4.55</td>
<td>-2.10</td>
</tr>
<tr>
<td><strong>Route Size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.73</td>
<td>-2.10</td>
</tr>
<tr>
<td>10</td>
<td>6.53</td>
<td>-2.63</td>
</tr>
<tr>
<td><strong>Overall % Deviation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.64</td>
<td>-2.63</td>
</tr>
</tbody>
</table>
The results of the Savings-Add heuristic are given in Table 5.17. Overall, the heuristic performs 5.8 percent lower than the benchmark, with a range of -3.21 percent to 17.15 percent. The heuristic performs at the same level or worse than the benchmark on four of the forty problems, the worst case being 3.2 percent higher than the benchmark. The spatial distribution factor shows a difference in performance of the heuristic by approximately 2 percent between the two levels of the factor. The heuristic also shows a difference in average percent deviation between the levels of the depot sites factor by over 3 percent. The performance of the heuristic does not differ by much between the two levels of the route size factor.

In evaluating the results of the Cluster-Route heuristic given in Table 5.18, the following observations are made. Overall, on the average the heuristic performs at about 5.6 percent below the benchmark, with a range of -2.63 percent to 16.98 percent. In only three of the forty problems does the heuristic perform equal to or worse than the benchmark, the worst case being 2.63 percent higher than the benchmark. The heuristic does not exhibit much difference between the levels of the spatial distribution factor, being only slightly better on the clustered level of spatial distribution. The depot sites factor shows the same performance by the heuristic as the earlier two heuristics. The Route Size factor also shows some difference in performance of the heuristics between the two levels, however, this difference is not large.
For each of the three heuristics a further analysis was conducted to examine the significance of the factors and the strength of relationship of the factors to the percent difference (the criterion). None of the factors or interaction effects turned out to be significant in any of the three analyses. The results are summarized in Table 5.19. The analysis of variance is not detailed since none of the results were significant. The coefficient of correlation showed that the route size factor had the lowest strength of relationship to the criterion (percent deviation from benchmark) for the heuristics. This indicates that the route size factor may not affect the quality of the solution since the value of route size factor is set by controlling the vehicle capacity. Tight vehicle capacity constraints can often lead to infeasible solutions in the vehicle routing process. These considerations led to the elimination of the route size factor from the final study. The other factors are retained.

The other objectives of the pilot study were to determine the feasibility of running a full factorial experiment in the final study, and the sample size needed to achieve sufficient power. On the first issue, the time needed to solve problems of the size used in the pilot study are small enough to allow a full factorial experiment to be conducted. With regard to the issue of power, the following observations are made. The existing sample size of 5 in the pilot study yields very low power for the tests of no difference between the levels for any of the factors. There is a high probability of a false null hypothesis of
factor not affecting the percent difference (criterion) being accepted. Based on the results obtained in the pilot study, a calculation of the sample size required in the final study is made. An alternate set of means are suggested for the factors. It is desired that there be a five percent difference in the mean value of the criterion between each pair of levels of a factor. In order that such a difference be detected in the final experiment, the sample size required would be 15 observations (per cell). This will allow for a power of more than 0.80 to be achieved for the cost ratio factor, at an alpha level of 0.05. For the other factors also, the power increases, though for a sample size of 15, the power still remains low (around 0.25). To obtain a high power for the factors Number of Depot Sites and Spatial Distribution, the sample size would have to be extremely large (around 100). Therefore, a sample size of 15 is selected for the final study.

The pilot study was useful in this research in deciding upon the number of factors to retain. The study also indicated that for the problem size used in this study, it would be feasible to run a full factorial design experiment upon the remaining factors. Further, the pilot study was used to examine the issue of power of the test and obtain the sample size necessary for the experiment to achieve a reasonable level of power.

On the basis of the pilot study runs, a full factorial study is conducted for the main experiment. Three factors are used in the experiment; cost ratio, spatial distribution, and number of depot sites.
### TABLE 5.19

Summary Results of the Analysis of Variance

The Pilot Study

<table>
<thead>
<tr>
<th>Factors</th>
<th>Heuristics</th>
<th>Savings-Drop</th>
<th>Savings-Add</th>
<th>Cluster-Route</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Distribution</td>
<td>Insignificant (p=0.28)</td>
<td>Insignificant (p=0.33)</td>
<td>Insignificant (p=0.43)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>r = 0.19</td>
<td>r = 0.16</td>
<td>r = 0.16</td>
<td></td>
</tr>
<tr>
<td>Number of Depot Sites</td>
<td>Insignificant (p=0.34)</td>
<td>Insignificant (p=0.115)</td>
<td>Insignificant (p=0.217)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>r = 0.16</td>
<td>r = 0.27</td>
<td>r = 0.21</td>
<td></td>
</tr>
<tr>
<td>Route Size</td>
<td>Insignificant (p=0.94)</td>
<td>Insignificant (p=0.83)</td>
<td>Insignificant (p=0.49)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>r = 0.01</td>
<td>r = 0.04</td>
<td>r = 0.14</td>
<td></td>
</tr>
</tbody>
</table>

Sample Size = 5

$r = \text{correlation coefficient}$
This leads to a 3x2x2 design for each of the three heuristics. The results of the heuristic are presented in the following chapter.

5.8 Summary of Results

This chapter has been devoted to examining the effectiveness of the heuristic procedures on small problems, when compared to the optimal solution. A pilot study also was conducted to establish the experimental design which defines the factors to be used in the study of the heuristics using larger problems. The results from the analysis can be summarized as:

(1) The Savings-Drop heuristic performed 4.81 percent above the optimum. The effect of the cost structure factor on the performance of the heuristic does not seem to be significant, however, a comparison of the group means shows that there might be some difference in the heuristic’s performance across different levels of cost structure.

(2) The Savings-Add heuristic performs 4.84 percent above the optimum on the average. The performance of this heuristic for the small problems is similar to the Savings-Drop heuristic’s. Again, there is no significant difference in performance of the heuristic across levels of cost structure, but comparison of means indicates that there might be some effect on the performance of the heuristic.

(3) The Cluster-Route heuristic performed 5.28 percent above the optimum. The performance of the heuristic is similar to the other heuristics. The cost ratio structure does not appear to significantly influence the performance of this heuristic. However, there appear to be differences between all the groups of cost structure.

(4) The best heuristic selection rule is only 3.33 percent above the optimal solution. This selection rule performs better than any single heuristic at all levels of the cost structure factor. The cost structure factor seems to affect the performance of this heuristic selection rule significantly. There appears to be a difference in performance at all levels of the cost structure factor. The power of the test is also
high. This analysis indicates that the heuristics perform well in comparison to the optimum. The figure of 3.33 percent is good when the complexity of the problem under study is considered.

(5) The use of the sequential model as a benchmark in the case of large problems seems reasonable, since the model performs well for small problems, being only 6.5% over the optimal. There is no significant difference between its performance and that of any heuristic in the case of small problems. Cost structure does seem to affect the performance of the sequential model.

(6) The cost structure factor did appear to have some effect on the performance of the heuristics. It affects the performance of some heuristics more than others. On this basis, it is included as a factor in the final experimental study.

(7) The pilot study results indicate the following: For problems of the size used it is feasible to run a full factorial experiment. The number of depot sites and Spatial Distribution factors appear to influence the performance of the heuristics more than the Route Size, though not significantly. However, in practical terms a two percent difference in performance may be significant, since a one percent difference may be equivalent to several thousand dollars of distribution cost. The Route Size parameter factor does not seem to influence the performance of the heuristics. This factor is therefore dropped from the final experiment.
CHAPTER 6

EXPERIMENTAL RESULTS

In this chapter the performance of the heuristics is examined by means of the experiment set up using the results of the pilot study in the previous chapter. The focus of this chapter is on analyzing the results of the experiment. The heuristic models are compared with the sequential model to illustrate the performance of the heuristics under various experimental conditions.

The size of the problems and the factor level settings have been established in the previous chapter. The design of the experiment is a full three-way factorial as shown in Table 6.1. The cost structure factor has the same settings as used in the experimental design for comparing the heuristics to the optimal model. This is done to examine whether the cost structure affects the criterion (percent difference in system cost), i.e. if the small problem size in the previous experiment was a limiting factor. The spatial distribution factor has been kept at the same levels as in the pilot study. The clustered distribution has five clusters, which represents the average number of warehouses per firm across all industries. Other values might prove to be more favorable to industries of one cost structure as opposed to another, i.e. closer to their average number of warehouses. The Number of Depot Sites factor has been set at values of 8 to 12. These values have been
selected since they represent 2 to 3 times the average number of industries selected by using the three levels of Cost Structure.

The criterion used to evaluate the heuristics is the same as used in the earlier chapter, the percent deviation from the benchmark, the sequential model solution. An issue of interest in the study is the amount of improvement that the heuristics exhibit over the benchmark. Using cost values will not provide a clear indication of the degree of improvement that could be obtained in an actual physical distribution system. On the other hand, percent deviation provides a clear indication of the amount of improvement and the accompanying savings in cost.

The three heuristics and the sequential model are applied to the same set of problems. For the $3 \times 2 \times 2$ design there are fifteen replications for a total of 180 observations for each of the heuristics. First, the performance of each of the heuristics is examined. This is followed by an evaluation of the combined performance of the three heuristics. This is done to examine the rule of selecting the best heuristic for any given problem which has been evaluated in the previous chapter for small problems. The best heuristic selection rule is compared with the benchmark to examine if it would be profitable to use such a rule. A comparison of the heuristics also is made to examine if they differ in performance. In the following sections the results of each of heuristic procedure are discussed. The basic statistics and the analysis of variance tables from the experiments are presented. The use of equal sample size ensures that the test is robust to violations of assumptions of analysis of variance.
TABLE 6.1

Comparison of Heuristics to the Sequential Model

Factor Levels for the Complete Factorial Design

**Factors**

**Cost Structure**

<table>
<thead>
<tr>
<th>Level</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.26</td>
</tr>
<tr>
<td>Medium</td>
<td>0.85</td>
</tr>
<tr>
<td>High</td>
<td>1.56</td>
</tr>
</tbody>
</table>

**Spatial Distribution**

<table>
<thead>
<tr>
<th>Level</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>1 cluster of customers</td>
</tr>
<tr>
<td>Clustered</td>
<td>5 clusters of customers</td>
</tr>
</tbody>
</table>

**Number of Depot Sites**

<table>
<thead>
<tr>
<th>Level</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>8</td>
</tr>
<tr>
<td>High</td>
<td>12</td>
</tr>
</tbody>
</table>
6.1. **Comparison of the Savings-Drop Heuristic with the Sequential Model**

A three-way analysis of variance is done to examine the difference in the performance of the Savings-Drop heuristic for the various experimental conditions. The total cost of the sequential solution is the benchmark, with the criterion being the percent deviation of the heuristic solution from the benchmark. The solution results for the three-way analysis of variance are shown in Tables 6.2 and 6.3.

Overall, the heuristic procedure averages 6.19 percent better than the benchmark. The range of the percent deviation is from -4.35 to 36.02 percent below the benchmark. This implies that in the worst case the heuristic solution is 4.35 percent higher than the benchmark. In the best case, the solution is lower than the benchmark by 36.02 percent. In only seventeen of the 180 problems run does the heuristic perform equal to or slightly worse than the benchmark. In ten of those problems, the superiority of the sequential model is less than 1%.

An analysis of the heuristic performance by individual factors gives the following results. For the cost ratio factor, the heuristic performance is best at the medium cost ratio level (0.85), with an average solution value of 6.89 percent below the benchmark, and is worst at the low level of cost ratio (0.26) with an average deviation of 5.32 percent below the benchmark. The number of depot sites factor shows a better performance by the heuristic at the low level (8 sites) than at the high level (12 sites). The heuristic performs better under a uniform distribution of customers, being on the average 7.4 percent better than the benchmark, while it is only an average of 5.01 percent better than the benchmark for the clustered distribution of the customers.
### TABLE 6.2

Results of the Savings-Drop Heuristic

**Individual Factors (% Deviation from Benchmark)**

**Simple Statistics**

<table>
<thead>
<tr>
<th>Cost Structure</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Range</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (.26)</td>
<td>5.32</td>
<td>6.64</td>
<td>-4.35</td>
<td>35.51</td>
<td></td>
</tr>
<tr>
<td>medium (.85)</td>
<td>6.89</td>
<td>6.80</td>
<td>-2.79</td>
<td>36.02</td>
<td></td>
</tr>
<tr>
<td>high (1.56)</td>
<td>6.37</td>
<td>5.38</td>
<td>-2.80</td>
<td>19.88</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of Depot Sites</th>
<th>Level</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Range</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (.8)</td>
<td>6.73</td>
<td>7.37</td>
<td>-4.35</td>
<td>36.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>high (.12)</td>
<td>5.66</td>
<td>5.01</td>
<td>-2.79</td>
<td>21.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spatial Distribution</th>
<th>Level</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Range</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>uniform</td>
<td>7.38</td>
<td>5.41</td>
<td>-4.35</td>
<td>21.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>clustered</td>
<td>5.01</td>
<td>6.92</td>
<td>-2.80</td>
<td>36.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Overall             | 6.19  | 6.31                | -4.35   | 36.02 |
TABLE 6.3

Results of the Savings Drop Heuristic
2-Way Results for Factors (Mean % Deviation from Benchmark)

<table>
<thead>
<tr>
<th>No. of Depot Sites</th>
<th>Low(8)</th>
<th>High(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (.26)</td>
<td>5.73</td>
<td>4.92</td>
</tr>
<tr>
<td>medium (.85)</td>
<td>7.85</td>
<td>5.93</td>
</tr>
<tr>
<td>high (1.56)</td>
<td>6.60</td>
<td>6.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spatial Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
</tr>
<tr>
<td>low (.26)</td>
</tr>
<tr>
<td>medium (.85)</td>
</tr>
<tr>
<td>high (1.56)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of Depot Sites</th>
<th>Uniform</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>low (8)</td>
<td>7.39</td>
<td>6.07</td>
</tr>
<tr>
<td>high (12)</td>
<td>7.37</td>
<td>3.95</td>
</tr>
</tbody>
</table>
TABLE 6.4

Analysis of Variance Table for Savings-Drop Heuristic

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>76.50</td>
<td>38.20</td>
<td>1.00</td>
<td>0.37</td>
</tr>
<tr>
<td>No. of Depots</td>
<td>1</td>
<td>51.20</td>
<td>51.20</td>
<td>1.34</td>
<td>0.25</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td>1</td>
<td>253.20</td>
<td>253.20</td>
<td>6.60</td>
<td>0.01</td>
</tr>
<tr>
<td>Structure x No. of Depots</td>
<td>2</td>
<td>17.10</td>
<td>8.50</td>
<td>0.22</td>
<td>0.80</td>
</tr>
<tr>
<td>Structure x Spatial Distribution</td>
<td>2</td>
<td>39.90</td>
<td>19.90</td>
<td>0.52</td>
<td>0.60</td>
</tr>
<tr>
<td>No. of Depots x Spatial Distribution</td>
<td>1</td>
<td>49.80</td>
<td>49.80</td>
<td>1.30</td>
<td>0.26</td>
</tr>
<tr>
<td>Structure x No. of Depots x Spatial Distribution</td>
<td>2</td>
<td>186.60</td>
<td>93.30</td>
<td>2.43</td>
<td>0.09</td>
</tr>
<tr>
<td>Error</td>
<td>168</td>
<td>6443.90</td>
<td>38.40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**TABLE 6.5**

Planned Comparisons between the Group Means* for the Savings-Drop Heuristic

1. Pairwise Difference between Means of Cost Structure

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>1.57</td>
<td>1.05</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>0.52</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value $d = 2.74$ at alpha = .05

2. Difference between Means of No. of Depots Sites

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>1.07</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value $d = 2.56$ at alpha = 0.05

3. Difference Between Means of Spatial Distribution

<table>
<thead>
<tr>
<th></th>
<th>Uniform</th>
<th>Clumped</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>-</td>
<td>2.37</td>
</tr>
<tr>
<td>Clustered</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value $d = 2.56$ at alpha = 0.05

*Dunn's procedure is used.*
The behavior of the heuristic can be partly explained by the performance of the sequential model. For low cost structure, the distribution costs are much higher, therefore the sequential model tends to locate with a view to minimizing distribution costs, i.e. it behaves more like the heuristics. Again, with a larger number of depot sites, the sequential model has a better chance of being able to locate nearer more customers, minimizing distribution costs. Thus, relatively, the performance of the heuristic is not as good. A similar argument can be extended regarding the performance of the sequential model in a clustered spatial distribution of customers, as opposed to a uniformly distributed customer set. With clustered customers, the routes are shorter, therefore the amount of error is distribution costs is not so high for the sequential model as in uniformly distributed customers, where the routes from the depot would be longer.

The two-way results in Table 6.3 show the combined effect of the factors on the performance of the heuristic. Based on the cell means the heuristic appears to perform best for combinations of medium cost structure, low number of depot sites, and uniform distribution of customers, being an average of 7.75 percent below the benchmark. The performance of the heuristic does not appear to be as good for factor level combinations of low cost structure, high number of depot sites, and clustered distribution of customers, with an average of only 4.7 percent below the sequential solution. However, even this performance may be considered good, since even a one percent improvement of the heuristic over the benchmark could mean a savings of thousands of dollars if the benchmark is the existing system design.
The analysis of variance is used to formally examine the effect of the factors on the performance of the heuristic. The analysis of variance table is shown in Table 6.4. The three-way interaction term of the factors shows that a mild interaction exists ($p = 0.09$). However, the result is not significant at an alpha level of 0.05. None of the two-way interactions between the factors is significant at an alpha level of 0.05. The overall main effects are significant at the 0.05 significance level ($p = 0.046$). An analysis of the three factors shows that the only significant factor is the Spatial Distribution factor ($p = 0.01$). The other factors are not even marginally significant. The power of the test is computed to be 0.71 at an alpha level of 0.05. Further, at the planned level of a difference of 5% between the group means, the power of this test is 0.61. This compares well with the estimated power of 0.25 for the test if the group means differ by 5%; which was developed in the pilot study in the previous chapter. The conclusion is that only the spatial distribution factor affects the heuristic to any significant degree. The mild three-way interaction between the factors is not significant.

Cost Structure has been included in the study as a factor of interest. A planned comparison among means of the three groups of cost structure is carried out to investigate the effect of the different cost structures. Dunn's procedure for multiple comparisons is used. The difference in group means is shown in Table 6.5. There does not appear to be a difference in the heuristic performance at the different cost ratio levels. The difference between the means for the other factors is also not significant, as shown in Table 6.5.
The overall conclusions which can be drawn from this analysis are as follows. The heuristic has been shown to perform well overall in comparison to the sequential model. An overall improvement of 6.2 percent can be considered as good for an existing physical distribution system, considering the magnitude of the distribution costs. The heuristic does not seem to be affected much by the environmental factors, except for spatial distribution of customers. In this case, the performance of the model is better for a uniformly distributed customer set than for a clustered one due to the reasons given earlier. Most of the problems in which the performance of the heuristic was not better than the benchmark were from clustered customer distribution. Finally, it is noted that overall, the heuristic tends to open fewer depots in the system than does the sequential model (the benchmark). Even in the clustered distribution of customers, with five clusters, the model opens less than five depots in most problems, whereas the benchmark opens close to five depots. This also improves the performance of the heuristic. The sequential model will open more depots than the heuristic since during the location phase it considers higher distribution costs than actual.

Further, the results of this experiment may be generalized to different values of the factors which have been fixed or randomized in the study. These include fixed factors such as number of customers, vehicle capacity, and randomized factors such as location of customers, and customer loads. These factors have been discussed in Chapter 4. The generalization may not apply to other factors ignored in study such as stochastic demand and percentage of customers served.
6.2 Comparison of the Savings-Add Heuristic with the Sequential Model

Tables 6.6 and 6.7 show the solution results for the Savings-Add heuristic. The criterion employed is the same as before. The average percent deviation from the benchmark solution is 2.67 percent, and the range of the percent deviation is -7.53 percent to 36.02 percent below the benchmark. This implies that in the worst case the heuristic solution is 7.53 percent above the benchmark, while in the best case the heuristic solution is 36.02 percent lower than the benchmark. In seventy five percent of the problems the heuristic outperformed the sequential model. For those problems where the sequential model was better, the average deviation was about two percent of total costs.

An analysis of the heuristic performance by individual factors in terms of means is given in Table 6.6. For the Cost Structure factor, the performance of the heuristic is best at the high level (1.56) at 4.76%, and worst for the low ratio (0.26) at 0.40 percent. In the case of Number of Depot Sites factor the performance of the heuristic is similar to the Savings-Drop heuristic. The heuristic performs better for a low number of depot sites. For the spatial distribution factor, this heuristic exhibits little difference in performance at the two levels.

The performance of the Savings-Add heuristic is better for the high cost structure since it proceeds to open depots starting with all closed. The higher location cost allows the heuristic to select the initial depots more accurately. The performance on the Number of Depot Sites factor is similar to the Savings-Drop for the same reason as explained in the last section. Its performance on the spatial distribution of customers, and in general, the poor performance compared
## TABLE 6.6

Results of the Savings-Add Heuristic

### Individual Factors (% Deviation from Benchmark)

#### Simple Statistics

<table>
<thead>
<tr>
<th>Cost Structure</th>
<th>Level</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>low (.26)</td>
<td>0.40</td>
<td>4.62</td>
<td>-7.53</td>
<td>18.13</td>
</tr>
<tr>
<td></td>
<td>medium (.85)</td>
<td>2.83</td>
<td>7.02</td>
<td>-5.15</td>
<td>36.02</td>
</tr>
<tr>
<td></td>
<td>high (1.56)</td>
<td>4.76</td>
<td>5.26</td>
<td>-4.90</td>
<td>21.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of Depot Sites</th>
<th>Level</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>low (.8)</td>
<td>3.51</td>
<td>6.56</td>
<td>-4.70</td>
<td>36.02</td>
</tr>
<tr>
<td></td>
<td>high (.12)</td>
<td>1.81</td>
<td>5.21</td>
<td>-7.53</td>
<td>22.53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spatial Distribution</th>
<th>Level</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>uniform</td>
<td>2.54</td>
<td>5.67</td>
<td>-7.53</td>
<td>24.53</td>
</tr>
<tr>
<td></td>
<td>clustered</td>
<td>2.78</td>
<td>6.27</td>
<td>-4.90</td>
<td>36.02</td>
</tr>
</tbody>
</table>

| Overall              |             | 2.67                | 5.97    | -7.53 | 36.02  |
TABLE 6.7

Results of the Savings Add Heuristic
2-Way Results for Factors (Mean \% Deviation from Benchmark)

<table>
<thead>
<tr>
<th>Cost Structure</th>
<th>No. of Depot Sites</th>
<th>Low(8)</th>
<th>High(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>low (.26)</td>
<td></td>
<td>1.46</td>
<td>-0.66</td>
</tr>
<tr>
<td>medium (.85)</td>
<td></td>
<td>3.53</td>
<td>2.14</td>
</tr>
<tr>
<td>high (1.56)</td>
<td></td>
<td>5.56</td>
<td>3.97</td>
</tr>
</tbody>
</table>

Spatial Distribution

<table>
<thead>
<tr>
<th>Cost Structure</th>
<th>Uniform</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>low (.26)</td>
<td>0.58</td>
<td>0.22</td>
</tr>
<tr>
<td>medium (.85)</td>
<td>3.24</td>
<td>2.42</td>
</tr>
<tr>
<td>high (1.56)</td>
<td>3.81</td>
<td>5.71</td>
</tr>
</tbody>
</table>

Spatial Distribution

<table>
<thead>
<tr>
<th>No. of Depot Sites</th>
<th>Uniform</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>low (8)</td>
<td>3.08</td>
<td>3.95</td>
</tr>
<tr>
<td>high (12)</td>
<td>2.01</td>
<td>1.62</td>
</tr>
</tbody>
</table>
### TABLE 6.8

Analysis of Variance Table for Savings-Add Heuristic

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>572.30</td>
<td>286.10</td>
<td>8.65</td>
<td>0.00</td>
</tr>
<tr>
<td>No. of Depots</td>
<td>1</td>
<td>130.00</td>
<td>130.00</td>
<td>3.93</td>
<td>0.05</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td>1</td>
<td>2.60</td>
<td>2.60</td>
<td>0.08</td>
<td>0.78</td>
</tr>
<tr>
<td>Structure x</td>
<td>1</td>
<td>4.20</td>
<td>2.10</td>
<td>0.06</td>
<td>0.94</td>
</tr>
<tr>
<td>No. of Depots x</td>
<td>2</td>
<td>64.20</td>
<td>32.10</td>
<td>0.97</td>
<td>0.38</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td>1</td>
<td>18.20</td>
<td>18.20</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>Structure x No. of Depots x Spatial Distribution</td>
<td>2</td>
<td>27.10</td>
<td>13.50</td>
<td>0.41</td>
<td>0.66</td>
</tr>
<tr>
<td>Error</td>
<td>168</td>
<td>5554.80</td>
<td>33.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 6.9

Planned Comparisons between the Group Means* for the Savings-Add Heuristic

1. **Pairwise Difference between Means of Cost Structure**

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>2.43</td>
<td>4.36**</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>1.93</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.54** at alpha = 0.05

2. **Difference between Means of No. of Depots Sites**

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>1.70</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.37 at alpha = 0.05

3. **Difference Between Means of Spatial Distribution**

<table>
<thead>
<tr>
<th></th>
<th>Uniform</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>-</td>
<td>0.16</td>
</tr>
<tr>
<td>Clustered</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.37 at alpha = 0.05

*Dunn's procedure is used.*
to the Savings-Drop model can be explained by observing that the add routine does not work as well as the drop routine in location, and using the add procedure, the heuristic ends up opening more depots than the Savings-Drop. The behavior of the heuristic is therefore, closer to the sequential model.

The two-way results in Table 6.7 show the combined effect of the factors on the performance of the heuristic. Based on the cell means, the heuristic appears to perform best under the conditions of high cost ratio level (1.56) a low number of depot sites, and a clustered distribution of customers. However, the performance of the heuristic for the spatial distribution factor in combination with the other factors is not well defined; since the performance of the heuristic is also the worst when a low cost ratio, high level of depot sites and clustered set of customers is concerned; being only an average of 0.5 percent better than the benchmark. This indicates that the spatial distribution does not affect the heuristic. Therefore, only the effect of the first two factors, i.e., Cost Structures and Number of Depot Sites should be considered. The worst performance of the heuristic appears to be under the conditions of low cost ratio (0.26) and for a high number of depot sites. Based on these observations it appears that the heuristic performs well only under certain conditions. Under others, it is no better than the sequential model.

The analysis of variance table for the heuristic is shown in Table 6.8. The three-way interaction is insignificant at an alpha value of 0.05. None of the two-way interaction effects are significant at an alpha level of 0.05. The overall main effects is highly significant at
the 0.05 significance level \( \text{p}=0.00 \). An analysis of the individual factors shows that the Cost Structure \( \text{p}=0.00 \) and the Number of Depot Sites \( \text{p}=0.05 \) factors are significant at the 0.05 level.

The power of the test for the Cost Structure factor is computed to be 0.75 at an alpha level of 0.05. This is the value for the alternative hypothesis that the actual group means are as observed. In the previous chapter, an alternative set of means was set to be such that a difference of 5% between the group means could be correctly identified. The power of the test here for that set of means is computed to be 0.97. This compares well with the planned power of approximately 0.82. For the Number of Depots factor, the power of the test to correctly discover the existing difference in means is 0.51 at an alpha level of 0.05. For the alternative set of means developed in the previous chapter, the power of the test has been estimated to be 0.25 for the sample size of 15. For this set of means (which differ by 5%) the power of the test here is 0.67, which compares well with the desired power level. The conclusion is that both the Cost Structure and the Number of Depot Sites factors affect the performance of the heuristic significantly.

A planned comparison of the mean percent deviations of the three groups of Cost Structure is shown in Table 6.9. Dunn's procedure for multiple comparisons is used to investigate the difference in the cost ratios. There appears to be a difference between the low cost ratio level (0.26) and the high cost ratio level (1.56) at the 0.05 significance level. The other groups are not significantly different from each other.
The overall conclusions which can be drawn from this analysis are as follows. Overall, in comparison to the benchmark the heuristic obtains a better solution. However, the improvement noted with this heuristic is not as high as for the Savings-Drop model, being only 2.7%. The performance of the Savings-Add heuristic is also affected by the environmental factors, in this case, Cost Ratio and Number of Depot Sites. This suggests that the use of this heuristic is more limited, and that it performs better than the benchmark only under a few combinations of factor settings. Again, in case of this heuristic, it tends to open fewer depots than does the benchmark, however, less frequently than the Savings-Drop heuristic. This tends to suggest that the add procedure does not work particularly well in comparison to the sequential. Further, the results of this study may be extended to the factor held fixed or randomized in the study. Such factors are not expected to influence or change the results obtained. The effectiveness of the procedure for the factors ignored in the study cannot be stated nor the results generalized.

6.3 Comparison of the Cluster-Route Heuristic with the Sequential Model

Tables 6.10 and 6.11 show the solution results for the Cluster-Route heuristic. The average percent deviation from the benchmark (Sequential Model) is 6.45 percent. The range of the percent deviation is -9.99 percent to 30.99 percent below the benchmark. This implies that in the worst case the heuristic solution is 9.99 percent over the benchmark, and in the best case the heuristic is 30.99 percent below the benchmark. In seventy five percent of all problems analyzed, the
### TABLE 6.10

Results of the Cluster-Route Heuristic

Individual Factors (% Deviation from Benchmark)

<table>
<thead>
<tr>
<th>Cost Structure</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (.26)</td>
<td>1.32</td>
<td>6.96</td>
<td>-9.99</td>
<td>30.99</td>
</tr>
<tr>
<td>medium (.85)</td>
<td>9.26</td>
<td>8.09</td>
<td>-2.01</td>
<td>25.35</td>
</tr>
<tr>
<td>high (1.56)</td>
<td>8.77</td>
<td>8.48</td>
<td>-8.75</td>
<td>26.86</td>
</tr>
</tbody>
</table>

| No. of Depot Sites | | | |
| Level              | | | |
| low (.8)           | 6.83 | 8.35 | -7.48 | 30.99 |
| high (.12)         | 6.07 | 8.94 | -9.99 | 25.75 |

| Spatial Distribution | | | |
| Level               | | | |
| uniform             | 4.33 | 5.09 | -8.75 | 25.35 |

### TABLE 6.11

Results of the Cluster Route Heuristic

2-Way Results for Factors (Mean % Deviation from Benchmark)

<table>
<thead>
<tr>
<th></th>
<th>No. of Depot Sites</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low(8)</td>
<td>High(12)</td>
<td></td>
</tr>
<tr>
<td><strong>Cost Structure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (.26)</td>
<td>1.94</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>medium (.85)</td>
<td>9.92</td>
<td>8.60</td>
<td></td>
</tr>
<tr>
<td>high (1.56)</td>
<td>8.62</td>
<td>8.92</td>
<td></td>
</tr>
<tr>
<td><strong>Spatial Distribution</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uniform</td>
<td>4.69</td>
<td>-2.05</td>
<td></td>
</tr>
<tr>
<td>Clustered</td>
<td>4.66</td>
<td>13.87</td>
<td></td>
</tr>
<tr>
<td>high (1.56)</td>
<td>3.64</td>
<td>13.90</td>
<td></td>
</tr>
<tr>
<td><strong>Spatial Distribution</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of Depot Sites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (8)</td>
<td>4.13</td>
<td>9.52</td>
<td></td>
</tr>
<tr>
<td>high (12)</td>
<td>4.53</td>
<td>7.62</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 6.12

Analysis of Variance Table for Cluster-Route Heuristic

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>2377.50</td>
<td>1188.70</td>
<td>27.82</td>
<td>0.001</td>
</tr>
<tr>
<td>No. of Depots</td>
<td>1</td>
<td>25.40</td>
<td>25.40</td>
<td>0.60</td>
<td>0.44</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td>1</td>
<td>809.90</td>
<td>809.90</td>
<td>18.95</td>
<td>0.001</td>
</tr>
<tr>
<td>Structure x No. of Depots</td>
<td>2</td>
<td>24.90</td>
<td>12.50</td>
<td>0.29</td>
<td>0.75</td>
</tr>
<tr>
<td>Structure x Spatial Distribution</td>
<td>2</td>
<td>2723.20</td>
<td>1361.60</td>
<td>31.86</td>
<td>0.00</td>
</tr>
<tr>
<td>No. of Depots x Spatial Distribution</td>
<td>1</td>
<td>59.80</td>
<td>59.80</td>
<td>1.40</td>
<td>0.24</td>
</tr>
<tr>
<td>Structure x No. of Depots x Spatial Distribution</td>
<td>2</td>
<td>145.20</td>
<td>72.60</td>
<td>1.70</td>
<td>0.19</td>
</tr>
<tr>
<td>Error</td>
<td>168</td>
<td>7179.30</td>
<td>42.70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 6.13

Planned Comparisons between the Group Means* for the Cluster-Route Heuristic

1. Pairwise Difference between Means of Cost Structure

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>7.94**</td>
<td>7.45**</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>0.49</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.89** at alpha = .05

2. Difference between Means of No. of Depots Sites

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>0.76</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.69 at alpha = 0.05

3. Difference Between Means of Spatial Distribution

<table>
<thead>
<tr>
<th></th>
<th>Uniform</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>-</td>
<td>4.24+</td>
</tr>
<tr>
<td>Clustered</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.69+ at alpha = 0.05

*Dunn's procedure is used.
heuristic performs better than the benchmark. In those cases where it does not, the deviation from the benchmark is approximately 2.2 percent.

An analysis of the heuristic performance by individual factors in terms of means is given in Table 6.10. For the Cost Structure factor, the performance of the heuristic is good at high and medium levels, and quite poor at the low level. There appears to be no difference in performance of the heuristic across different levels of Depot Sites. For the Spatial Distribution factor, the heuristic shows a better performance for the clustered customer data sets as compared to the uniformly distributed data sets.

The performance of the heuristic is poor at the low cost structure level compared to the other levels essentially due to the same reasons as the other two heuristics, namely, the better performance of the sequential model for the low cost structure. Since the heuristic searches for clusters of customers and locates near them, the number of depot sites should have little effect. It will perform better for a clustered spatial distribution of customers since it is designed to identify clusters and can thus form routes more effectively.

The two-way results in Table 6.11 show the combined effect of the factors on the heuristic. Based on the cell means the heuristic appears to perform best under the conditions of medium and high cost structure levels, and clustered spatial distribution of customers. The worst performance of the heuristic appears to occur for the conditions resulting from a low cost structure level. Number of Depot Sites factor appears to have no affect, while the performance of the heuristic for
Spatial Distributions is not clearly defined, having both its best and its worst performance for clustered data sets.

The analysis of variance table for the heuristic is given in Table 6.12. The three-way interaction is insignificant at alpha = 0.05 level. An examination of the two-way interactions show that the Cost Structure by Spatial Distribution interaction effect is significant (p = 0.001) at alpha = 0.05. A detailed analysis of this interaction shows that there is little difference in performance of the heuristic between the high and medium Cost Ratios at either level of Spatial Distribution. However, the performance is poor at the low cost ratio level, especially if the spatial distribution is clustered. When the spatial distribution is uniform there seems to be little difference in the performance of the heuristic at any Cost Structure level.

The interaction effect shown in Figure 6.1 is explained as follows. When the cost structure is medium or high, the weight is not excessively in favor of distribution costs in the sequential model. In the case of low cost structure, the distribution costs have more weight, and the sequential model opens less depots, similar to the heuristics. The heuristic itself should not behave much different at any cost structure level. When the spatial distribution is uniform, the heuristic cannot identify any well defined clusters. Since routing is dependent on clusters, the cost structure will not affect the heuristics. However, in the clustered spatial distribution of customers, when the cost structure is low, the sequential model is particularly efficient. Since the sequential model does not consider customers for routes based on exact clusters, it may be able to include customers just outside the border of
the cluster into routes to achieve greater savings. However, when cost structure is medium or high, location is a more important consideration with the sequential model, not the clusters of customers, which do provide some cost savings in distribution costs. On the other hand, the heuristic considers both.

The main effects which are significant are Cost Structure (p=0.001) and Spatial Distribution (p=0.001). Of more interest is their significant interaction effect which has been analyzed.

The power of the test for the interaction effect is computed to be close to 0.99 at an alpha level of 0.05. This would be the power of the test that there is no interaction when the cell means are as actually observed. For the 5% difference in means specified in the previous chapter, the power of this test is 0.75 which compares well with the planned power of 0.72. The power for the main effects Cost Ratio and Spatial Distribution are also good, being 0.95 and 0.56 respectively. These compared well with the planned levels of 0.82 and 0.25 respectively. The conclusion is that the Cost Structure and Spatial Distribution factors affect the performance of the heuristic and the heuristic is affected by their interaction. This can be useful to management since it provides a guideline as to when is it better to apply the heuristic as compared to the sequential.

A planned comparison of the mean percent deviations for the different factors is shown in Table 6.13. There appears to be a significant difference between the low Cost Structure level and the other Cost Structure levels. There is no difference between the means of the Depot Sites factor. The Spatial Distribution factor also shows a
Figure 6.1 Illustration of Interaction Between Cost Structure and Spatial Distribution for the Cluster-Route Heuristic
significant difference between the uniform and clustered levels.

The conclusions which can be drawn from this analysis are as follows. Overall, in comparison to the benchmark, the heuristic solution is much better, being 6.45 percent better on the average. Further, the performance of the heuristic is affected by Cost Structure and the Spatial Distribution factors. The results indicate that the heuristic is particularly effective when the cost structure levels are high and the spatial distribution of customers is clustered. Again, the results of the experiment may be extended to most factors randomized or fixed in the study without loss of generality. However, the results are limited to exclude those factors which have been ignored in the experiment. It is recognized that a specific heuristic such as the Cluster-Route may be influenced by a factor such as location of depot sites, which has been randomized in the study.

6.4 Comparison of the Best Heuristic Rule with the Sequential Model

The results from the best heuristic rule are shown in Tables 6.14 and 6.15. For this rule, the average percent deviation from the benchmark for the rule is 9.97 percent. In 175 out of 180 problems, this rule outperformed the sequential model. In the cases it did not, the average deviation is less than 1% of total costs. An analysis by the individual factors in terms of means shows that this rule performs better than any of the individual heuristics. Selection of the best heuristic solution rule exhibits essentially the same pattern of performance as the individual heuristics for the Cost Structure and Number of Depot Sites factors. The best performance of the rule is for
the medium level of cost structure. There is hardly any difference between the levels of the Number of Depot Sites factor. The rule shows a better performance for the clustered spatial distribution of customers than the uniform. The lowest average percent deviation occurs for the uniform distribution of customers, and in this case, it is 5.16 percent below benchmark. This indicates an overall superior performance by this rule.

The two-way results in Table 6.15 show the combined effect of the factors on the heuristic rule. This rule appears to yield the best results under the conditions of medium level of cost ratio and clustered groups of customers. The Depot Sites factor does not seem to affect the heuristic's performance.

In using the best heuristic rule, the Savings-Drop heuristic is the best or provides an equally good solution as any other heuristic in over fifty percent of all problems. The Cluster-Route heuristic is best or ties with some other heuristic in forty percent of all problems, while the Savings-Add heuristic is best or tied for best in less than twenty percent of all cases.

The analysis of variance table for the rule is given in Table 6.16. The three-way interaction is not significant at the 0.05 level of significance (p = 0.30). An examination of the two-way interactions shows that only the Cost Structure by Spatial Distance interaction effect is significant at an alpha level of 0.05 (p=0.001). A detailed analysis of this interaction shows that there is little difference in the performance of the heuristic when the cost ratio level is medium or high, regardless of the spatial distribution. However, when the Cost
### TABLE 6.14

Results of the Best Heuristic Selection Rule

<table>
<thead>
<tr>
<th>Individual Factors (% Deviation from Benchmark)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple Statistics</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cost Structure</th>
<th>Average % Deviation</th>
<th>Std Dev</th>
<th>Low</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (.26)</td>
<td>6.21</td>
<td>6.41</td>
<td>-0.83</td>
<td>35.51</td>
</tr>
<tr>
<td>medium (.85)</td>
<td>11.94</td>
<td>7.61</td>
<td>0.56</td>
<td>36.02</td>
</tr>
<tr>
<td>high (1.56)</td>
<td>11.77</td>
<td>6.85</td>
<td>0.00</td>
<td>19.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of Depot Sites</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (.8)</td>
<td>10.54</td>
<td>7.84</td>
<td>-0.83</td>
<td>36.02</td>
</tr>
<tr>
<td>high (.12)</td>
<td>9.41</td>
<td>7.00</td>
<td>-0.38</td>
<td>22.53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spatial Distribution</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>uniform</td>
<td>8.51</td>
<td>5.16</td>
<td>0.00</td>
<td>25.35</td>
</tr>
<tr>
<td>clustered</td>
<td>11.44</td>
<td>8.95</td>
<td>-0.83</td>
<td>36.02</td>
</tr>
</tbody>
</table>

| Overall                         | 9.97               | 7.43    | -0.83 | 36.02 |
### TABLE 6.15

Results of the Best Heuristic Selection Rule

2-Way Results for Factors (Mean % Deviation from Benchmark)

<table>
<thead>
<tr>
<th>No. of Depot Sites</th>
<th>Cost Structure</th>
<th>Uniform</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low(8)</td>
<td>low (.26)</td>
<td>6.71</td>
<td>5.71</td>
</tr>
<tr>
<td></td>
<td>medium (.85)</td>
<td>12.54</td>
<td>11.35</td>
</tr>
<tr>
<td></td>
<td>high (1.56)</td>
<td>12.37</td>
<td>11.17</td>
</tr>
<tr>
<td>High(12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spatial Distribution</td>
<td>Uniform</td>
<td>Clustered</td>
</tr>
<tr>
<td></td>
<td>low (.26)</td>
<td>8.56</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td>medium (.85)</td>
<td>8.72</td>
<td>15.16</td>
</tr>
<tr>
<td></td>
<td>high (1.56)</td>
<td>8.23</td>
<td>15.31</td>
</tr>
<tr>
<td>No. of Depot Sites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low (8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high (12)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 6.16

Analysis of Variance Table for the Best Heuristic Selection Rule

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>2</td>
<td>1276.80</td>
<td>638.40</td>
<td>15.97</td>
<td>0.001</td>
</tr>
<tr>
<td>No. of Depots</td>
<td>1</td>
<td>57.30</td>
<td>57.30</td>
<td>1.43</td>
<td>0.23</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td>1</td>
<td>388.20</td>
<td>388.20</td>
<td>9.71</td>
<td>0.002</td>
</tr>
<tr>
<td>Structure x No. of Depots</td>
<td>2</td>
<td>0.40</td>
<td>0.20</td>
<td>0.01</td>
<td>0.99</td>
</tr>
<tr>
<td>Structure x Spatial Distributions</td>
<td>2</td>
<td>1316.50</td>
<td>658.30</td>
<td>16.47</td>
<td>0.001</td>
</tr>
<tr>
<td>No. of Depots x Spatial Distribution</td>
<td>1</td>
<td>40.12</td>
<td>40.12</td>
<td>1.00</td>
<td>0.32</td>
</tr>
<tr>
<td>Structure x No. of Depots x Spatial Distribution</td>
<td>2</td>
<td>96.20</td>
<td>48.10</td>
<td>1.20</td>
<td>0.30</td>
</tr>
<tr>
<td>Error</td>
<td>168</td>
<td>6714.30</td>
<td>40.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 6.17

Planned Comparisons between the Group Means* for the Best Heuristic Selection Rule

1. Pairwise Difference between Means of Cost Structure

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>5.70**</td>
<td>5.56**</td>
</tr>
<tr>
<td>medium</td>
<td>-</td>
<td>-</td>
<td>1.79</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.79** at alpha = .05

2. Difference between Means of No. of Depots Sites

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>-</td>
<td>1.13</td>
</tr>
<tr>
<td>high</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.61 at alpha = 0.05

3. Difference Between Means of Spatial Distribution

<table>
<thead>
<tr>
<th></th>
<th>Uniform</th>
<th>Clustered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>-</td>
<td>2.93+</td>
</tr>
<tr>
<td>Clustered</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Critical value d = 2.61+ at alpha = 0.05

* Dunn's procedure is used.
Structure level is low there is a difference in the performance of the heuristic across the levels of Spatial Distribution. The heuristic performs worse for the clustered spatial distribution level for low cost structure being only 3.8 percent below the benchmark. In the case of uniform distribution across all three cost structure levels. These results are similar to those of the Cluster-Route heuristic and have a similar explanation in terms of the performance of the heuristic. The interaction is graphed in Figure 6.2.

The main effects which are significant at the 0.05 level of significance are the Cost Structure and the Spatial Distribution factors, \( p = 0.00 \) and \( 0.002 \) respectively. The power of the test for the interaction effect is 0.94 at an alpha level of 0.05. This is the power of the test that there is no interaction effect of Cost Structure and Spatial Distribution, when the actual cell means are as observed. For the 5\% difference which was specified for the alternate set of means of interest, the power of this test is around 0.80 which compares well with the planned power of 0.72 for a sample size of 15. The power for the Cost Structure and Spatial Distribution factors (main effects) is also good, being 0.94 and 0.93 respectively, compared to the planned levels of 0.82 and 0.25 respectively. The conclusion is that Cost Structure and Spatial Distribution factors jointly affect the performance of the heuristic.

The planned comparison between the group means of the Cost Structure factor is shown in Table 6.17. There appears to be a significant difference between the mean percent deviations for the low level and the other levels of Cost Ratio at the 0.05 significance level.
Figure 6.2 Illustration of Interaction Between Cost Structure and Spatial Distribution for the Best Heuristic Rule
The medium and high levels do not appear to be significantly different from each other. There is no difference between the means of the Depot Sites factor. The Spatial Distribution factor also shows a significant difference in the performance of the heuristic between the uniform and clustered levels.

The overall conclusions which can be drawn for this heuristic are as follows. Overall, in comparison with the benchmark, the heuristic is superior at all levels of factors. Thus, such a rule can be used at all times. In practical terms, the improvements obtained are significant, since each percent improvement can translate to thousands of dollars savings. Further, Cost Structure and Spatial Distribution factors do seem to affect the performance of the heuristic. This heuristic is effective under all environmental conditions, though less effective for the low cost structure and clustered customer distribution condition.

6.5 Comparison of the Savings-Drop and Savings-Add Heuristics

The two heuristics differ in their starting conditions. The Savings-Drop heuristic develops an initial solution of all open depots and proceeds to close them one by one. On the other hand, the Savings-Add heuristic starts with all closed depots and proceeds to open them one at a time. Overall, the Savings-Drop heuristic performed better than the other heuristic. However, neither heuristic outperformed the other in all cases. To examine the if the Savings-Drop heuristic did indeed perform better than the other heuristic, a t-test was used. A 't' statistic of 5.44 was obtained which is significant (p=0.0001). This confirms that the Savings-Drop heuristic did perform better than the
Savings-Add model. Based on earlier analysis and conclusions derived it appears that the Savings-Add model does not perform particularly well in most environmental conditions. In most cases it was outperformed by the other two heuristics. Therefore, the Savings-Add heuristic should not be considered in distribution system design. Other heuristics perform at least as well in the far majority of cases.

6.6 Comparison of the Savings-Drop and Cluster-Route Heuristics

The two heuristics differ in their basic approaches. On the average, the Savings-Drop heuristic performs equal to the Cluster-Route model. A 't' statistic of 0.33 was obtained for the hypothesis test that the Savings-Drop heuristic performs better than the Cluster-Route heuristic. This indicated that there was no statistically significant difference in the performance of the two heuristics. However, in terms of their performance under specific environmental conditions (combination of factor levels) there are some notable differences. The Savings-Drop heuristic consistently outperforms the Cluster-Route heuristic for uniform spatial distribution of customers. The Cluster-Route heuristic is less affected by the Number of Depot Sites factor. Again, the Cluster-Route heuristic does perform somewhat better for clustered spatial distribution of customers.

6.7 Comparison of the Savings-Drop Heuristic and the Best Heuristic Rule

A comparison of the two heuristics is made to examine if the use of the Best Heuristic Rule can significantly improve on the
performance of any single heuristic. A 't' statistic of 5.20 was obtained for the hypothesis test that the Best Rule performs better than the Savings-Drop model. This result is significant (p = 0.001) which implies that using the Best Heuristic Rule is preferable to using any single heuristic. Such a result is to be expected from this composite rule.

6.8 Summary of Results

This chapter was devoted to examining the performance of the heuristics under different environmental conditions, as well as for larger sized problems. A summary of the statistical results is given in Table 6.18. Major conclusions arrived at in this chapter are stated next.

(1) The Savings-Drop heuristic performs an average of 6.2 percent better than the Sequential model (which is used as the benchmark). The practical implication of a 6 percent improvement is a saving of thousands of dollars in an existing distribution system where the benchmark employed here has been the conventional design procedure used. In almost all problems, it is better than the benchmark. The performance of the heuristic is different across the levels of the various factors. The difference in performance is significant across the levels of the Spatial Distribution factor only. No other factor or interaction effect of factors was significant.

(2) The Savings-Add heuristic performs an average of 2.67 percent better than the Sequential model. The performance of this heuristic is not as strong as that of the Savings-Drop heuristic. In some problems, the heuristic does not do better than the Sequential model. The performance of this heuristic is affected significantly only by the Cost Structure and the Number of Depot Sites factors. Overall, the performance of the heuristic is not as good as the other two heuristics. Therefore, the Savings-Add heuristic may be dropped.

(3) The Cluster-Route heuristic is better than the Sequential model by 6.45 percent on the average. However, the overall performance of this heuristic is not as consistent as the Savings-Drop model, with the heuristic performing poorly...
TABLE 6.18
Summary of Statistical Results

<table>
<thead>
<tr>
<th>Factors</th>
<th>Savings-Drop</th>
<th>Savings-Add</th>
<th>Cluster-Route</th>
<th>Best Heuristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost Structure</td>
<td>Insignificant (p=0.37)</td>
<td>Significant (p=0.001)</td>
<td>Significant (p=0.0001)</td>
<td>Significant (p=0.001)</td>
</tr>
<tr>
<td>Number of Depot Sites</td>
<td>Insignificant (p=0.25)</td>
<td>Significant (p=0.05)</td>
<td>Insignificant (p=0.44)</td>
<td>Insignificant (p=0.23)</td>
</tr>
<tr>
<td>Spatial Distribution</td>
<td>Significant (p=0.01)</td>
<td>Insignificant (p=0.78)</td>
<td>Significant (p=0.001)</td>
<td>Significant (p=0.002)</td>
</tr>
<tr>
<td>Cost Structure x No. Of Depot Sites</td>
<td>Insignificant (p=0.80)</td>
<td>Insignificant (p=0.94)</td>
<td>Insignificant (p=0.75)</td>
<td>Insignificant (p=0.99)</td>
</tr>
<tr>
<td>Cost Structure x Spatial Distribution</td>
<td>Insignificant (p=0.60)</td>
<td>Insignificant (p=0.38)</td>
<td>Significant (p=0.001)</td>
<td>Significant (p=0.001)</td>
</tr>
<tr>
<td>No. of Depot Sites x Spatial Distribution</td>
<td>Insignificant (p=0.26)</td>
<td>Insignificant (p=0.46)</td>
<td>Insignificant (p=0.24)</td>
<td>Insignificant (p=0.32)</td>
</tr>
<tr>
<td>Cost Structure x No. of Depot Sites x Spatial Distribution</td>
<td>Insignificant (p=0.09)</td>
<td>Insignificant (p=0.66)</td>
<td>Insignificant (p=0.19)</td>
<td>Insignificant (p=0.30)</td>
</tr>
</tbody>
</table>
against the Sequential model on some occasions. There is more variation in its performance. The heuristic does show better performance for clustered customer data sets compared to the uniformly distributed data sets when the cost structures are higher, whereas the other two heuristics perform better for uniformly distributed data sets. The performance of this heuristic may be affected by the Cost Structure factor. Overall, both this heuristic, and the Savings-Drop heuristic perform well.

(4) Using the best heuristic selection rule provides stronger results than using any one heuristic. This rule performs 9.97 percent better than the Sequential model. In most of the problems, the best heuristic rule selected either the Savings-Drop or the Cluster-Route heuristic as the best heuristic to use. The interaction effect of Cost Structure and Spatial Distribution factors is significant, as is the Cost Structure factor, and the Spatial Distribution factor.

(5) The factors used in the study do seem to have some effect on the performance of the heuristics. The heuristics perform better under some combinations of the levels of these factors than others. In particular, the Cost Structure and the Spatial Distribution factors do affect the performance of the heuristics. The effect of the Number of Depot Sites is only significant in the case of the Savings-Add model, which itself does not perform well. Overall, the heuristics improve on the Sequential model by a good percentage. Further, in general, the results of the heuristics may be extended to other values of the factors which have been fixed or randomized in the study. However, the effect of factors specifically excluded from the study, such as stochastic demand, is uncertain, and the results cannot be extended to those situations.
CHAPTER 7
CONCLUSIONS

The problem of simultaneous location of depots and design of the vehicle routes from the depots has been studied in this research. The problems of depot location and vehicle scheduling from the depots are important subject areas in physical distribution. Presently, the techniques used in the design of the distribution system are based on a framework that separates the location of depots and the scheduling of vehicles into distinct decision making activities. In this research, a simultaneous depot location and vehicle routing model was formulated. The mathematical statement of the problem was presented in Chapter 3 as a zero-one mixed integer programming problem. It was shown that due to the computational complexity of the problem, derivation of solutions to this problem using the mixed integer programming model is limited to very small and simple problems. Given this limitation, a heuristic approach to solving the problem was adopted. Different heuristics were developed to solve the problem under different environmental conditions. The effectiveness of these heuristic procedures was examined within the context of the experimental design described in Chapter 4. The design included the various factors which were hypothesized as affecting the solution. The evaluation of the heuristics in comparison to the optimal mixed integer zero-one programming model was performed in Chapter 5. A
pilot study also was performed to examine which of the factors hypothesized as affecting the heuristic's solution should be included in a study of larger problems. The heuristics were evaluated in Chapter 6 in comparison with the sequential location-routing model in the context of the experimental design. The heuristics were shown to be effective in comparison with a sequential model. It is suggested that the sequential model is representative of the conventional real world decision making procedure in distribution system design. In the sequential scheme, the location-allocation problem is solved first, then the vehicles are routed on the customers already allocated to the depots. In Chapter 6 the impact of the different factors included in the research was also studied.

The results from this study provide important implications for the physical distribution system manager. First, the study clearly shows the benefits of using a simultaneous location-routing approach as opposed to the normal sequential location, routing approach. This is important because when the annual costs of physical distribution run into hundreds of thousands or millions or dollars, even a one percent improvement over the existing system can lead to a savings of thousands of dollars. Second, in this study, simple heuristic models have been developed, based on existing location and routing schemes, which are already in use in physical distribution systems. Therefore, the use of the heuristics should not be much more difficult. Third, this study has also analyzed the effect of environmental factors on the performance of the location-routing models. Such an analysis has not been done earlier, and the results from this study indicate that these factors do influence the
quality of the solution to the physical distribution system design problem. Therefore, by accounting for the environmental conditions, the proper system design model can be selected for greater effectiveness in system design and higher cost savings.

There are other factors in physical distribution system design which have been held fixed, randomized or ignored in this study. The effect of these on the performance of the heuristics has to be considered. The factors which have been held fixed or randomized, such as number of customers, vehicle capacity, location of customers, and customer loads are those that are not expected to affect the heuristics' effectiveness in comparison to the sequential (conventional) procedure. These factors will affect the heuristic and sequential model in a similar manner. Therefore, the results of this study can be generalized to the factors fixed or randomized in this study.

However, the effect of factors ignored in the study is uncertain. Factors such as stochastic nature of demand and percentage of customers served are more likely to impact heuristics which consider routes in the location decision. The results of the study may not be generalized to such situations. In this chapter, some further research extensions are suggested and finally, the contribution of the study are discussed.

7.1 Research Extensions

The location-routing model determines the depots to open, and the vehicles to route from them as well as the number of vehicles to use to minimize total system cost. Several environmental factors have been included in this study, such as cost ratio of location to routing,
number of depot sites relative to total population size, the spatial distribution of customers in the system, and the route size. Some assumptions have also been made in solving the problem, such as not allowing split deliveries, and considering a delivery only type of system. Uniformity in vehicle type and a single product are assumed. These environmental factors and assumptions need to be fully explored and further analysis of these may be considered. Other approaches to the problem may also be a topic for future research. Some research extensions are suggested here.

(1) The effect of the environmental factors should be studied further. Some factors have been included in this study. However, other factors also may be important. For example, the stochastic nature of the problem has been excluded and a deterministic approach used in this study. One extension to the study would be to include the stochastic aspects of the problem, such as customer demand. This may lead to a different approach to solve the problem than the one used in this study.

(2) More of the real world details may be incorporated in the study of the problem. Issues such as split deliveries, pickup and deliveries on the same route, and the issue of multiple products may be addressed in a real world application.

(3) The testing of the heuristic procedures in some real physical distribution problem may be considered a natural extension of this study. Such a testing would indicate how the heuristics may be improved. The application to an actual physical distribution problem would also allow for the examination of
the feasibility of using such models into the decision making scheme of the distribution system.

(4) Comparison of the models developed in this study with the models developed in earlier approaches to studying the problem, such as the ones listed in the literature survey.

(5) There exists a lack of any good lower bound to the location-routing problem. The LP relaxation to the mixed integer program is not likely to provide a good "tight" lower bound. Approaches to obtaining a good lower bound are necessary before the solution procedures can be evaluated fully. Other approaches such as Lagrangian relaxation, or a statistical estimation method for the lower bound should be examined.

(6) The problem addressed treats the depots as the primary supply points in the system. Future research may focus on the treatment of depots as transshipment points in the system, including the plants in the system design.

(7) Some companies use more than one type of depot or mode of delivery. Companies may use private or public warehouses. They may also own (or lease) the fleet of vehicles, or use common carriers. Companies may also use more than one form of transportation. Research could also focus on incorporating these different details in the solution procedures.

7.2 Contributions

The location-routing problem was shown to be a problem which is quite easy to state and very difficult to solve. This research addressed
the problem in a clearly defined framework and provides a method to approach the problem. In particular, the contributions of this study include:

(1) The formulation of a location-routing problem that captured the real simultaneous nature of the problem in the context of distribution system design.

(2) The derivation of optimal solutions to the zero-one mixed integer location-routing model, which served as a foundation for subsequent research in this area, and provided the direction for development of the problem. It also served as a benchmark to test the performance of the heuristic procedures.

(3) The development of heuristic procedures which:

(i) performed effectively against the optimal zero-one mixed integer program location-routing model and can be used to solve larger problems,

(ii) performed effectively against the conventional real world sequential location, routing decision making scheme.

(4) The identification of some of the environmental conditions which are likely to affect the solution quality in distribution system design.

7.3 Concluding Remarks

The primary objective of this research has been the development of a location-routing framework which can be used to improve distribution system design. The complexity of this problem has resulted in a lack of solution techniques to the problem. Almost all of the previous research
in the area has concentrated on solving either the location problem or the vehicle scheduling problem. The need to provide solution procedures to the actual problem of simultaneous location-routing exists. It is hoped that this research covers some of that ground and provides some impetus to move further in that direction.
LIST OF REFERENCES


APPENDIX A

Location of a Single Depot on a Plane.

Mathematical Formulation.

Define: \( x_i, y_i \) to be the coordinates of the depot \( i \), \((i = 1, \ldots, m)\)

\[ a_j w_j = \text{cost of serving customer } j \text{ with the units of goods } 'w_j' \]

that are required, \((j = 1, 2, \ldots, n)\)

\[ d_{ij} = \sqrt{(x_i-x_j)^2 + (y_i-y_j)^2} \]

\[ y_{ij} = \begin{cases} 1 & \text{if customer is served by depot } i \\ 0 & \text{otherwise} \end{cases} \]

Then, for \( M \) depot sites and \( n \) customers, the problem is:

Minimize \( H_i = \sum_{j=1}^{n} a_j w_j d_{ij} y_{ij} \) \((i = 1, \ldots, M)\)

subject to

\[ y_{ij} = 0,1 \]  \hspace{1cm} (1)
Location on a Network

Problem 1.
Define: \( m = \) number of proposed depots
\( n = \) number of demand points
\( x_{ij} = \) quantity shipped from depot \( i \) to demand node \( j \), \( i=1,\ldots,m, \ j=1,\ldots,n. \)
\( y_i = \) total amount shipped from depot \( i \), \( i=1,\ldots,m. \)
\( d_{ij}(x_{ij}) = \) cost of shipping \( x_{ij} \) from point \( i \) to point \( j \), \( i=1,\ldots,m, \ j=1,\ldots,n. \)
\( F_i(y_i) = \) fixed and operating costs of depot \( i \), \( i=1,\ldots,m \)
\( D_j = \) demand at node \( j \), \( j=1,\ldots,n \)

The model is formulated as:

\[
\text{Minimize } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} d_{ij}x_{ij} + \sum_{i=1}^{m} F_i y_i
\]

subject to

\[
\sum_{j=1}^{n} x_{ij} = y_i \quad i=1,\ldots,m
\]

\[
\sum_{i=1}^{m} x_{ij} = D_j \quad j=1,\ldots,n \quad (2)
\]

\[
x_{ij} > 0 \quad i=1,\ldots,m, \ j=1,\ldots,n \quad (3)
\]
Problem 2. The $m$-median Problem.

Define: $m =$ number of facilities to be established

$n =$ number of nodes in the network

$a_i =$ demand at node $i$, $i=1,\ldots,n$

$d_{ij} =$ shortest time (distance) from node $i$ to node $j$, $i=1,\ldots,n$, $j=1,\ldots,n$

$x_{ij} =$

$$
x_{ij} = \begin{cases} 
1 & \text{if node } i \text{ is assigned to center } j; \ i=1,\ldots,n, j=1,\ldots,n \\
0 & \text{otherwise}
\end{cases}
$$

The problem is formulated as

$$
\text{Minimize } Z \sum_{i=1}^{n} \sum_{j=1}^{n} a_i d_{ij} x_{ij}
$$

subject to

$$
\sum_{j=1}^{n} x_{ij} = 1 \quad i=1,\ldots,n \quad (1)
$$

$$
x_{jj} = m \quad (2)
$$

$$
x_{jj} > x_{ij} \quad (i, j=1,\ldots,n, i \neq j) \quad (3)
$$

$$
x_{ij} = 1 \text{ or } 0 \quad (4)
$$
Problem 1

The Travelling Salesman Problem (TSP)

The problem is to form a tour of the \( n \) nodes beginning and ending at the origin, say node 1.

Define: \( S = \{i| i=1,...,n\} \) be the set of all nodes in the network

\[
d_{ij} = \text{distance (cost) between nodes } i \text{ and } j, \forall i,j \in S
\]

\[
d_{ii} = \infty
\]

\[
x_{ij} = \begin{cases} 
1 & \text{if arc } i \rightarrow j \text{ is in the optimal tour } \forall i,j \in S \\
0 & \text{otherwise}
\end{cases}
\]

\( X = (x_{ij}| i,j=1,...,n) \) is the set for all \( x_{ij} \)

The problem is formulated as

Minimize \( Z = \sum_{i \in S} \sum_{j \in S} d_{ij}x_{ij} \)

\[
\sum_{i \in S} x_{ij} = 1 \quad \forall j \in S
\]

\[
\sum_{j \in S} x_{ij} = 1 \quad \forall i \in S
\]

\( V = \{(x_{ij}): \sum_{i \in Q} \sum_{j \in Q} x_{ij} \geq 1 \text{ for every nonempty proper subset of } Q \text{ of } S\} \)

\( X \in V \)

249
Constraint sets 1 and 2 force the tour to include every node (the tour enters and leaves every node).
Constraint set 3 is a set of constraints introduced to break any subtours.

Problem 2.
The Multi - Travelling Salesman Problem.
The basic M-TSP has been stated by Svetska and Huckfeldt (125) in the following manner:
Define : \( n \) = number of demand points
\( m \) = number of salesman
\( r = m + n - 1 \)
\( C_{ij} \) = distance from demand point \( i \) to demand point \( j \)
\( d_{ij} \) = new distance between \( i \) and \( j \) obtained by transformation on \( C_{ij} \)
\( x_{ij} \) = \( \begin{cases} 
1 & \text{if salesman travels from } i \text{ to } j \\
0 & \text{otherwise} 
\end{cases} \)
\( y_i \) = arbitrary real numbers used in subtour breaking constraints

The formulation of the problem is:
Minimize \( Z = \sum_{i=1}^{r} \sum_{j=1}^{r} d_{ij} x_{ij} \)
subject to
\[
\sum_{i=1}^{r} x_{ij} = 1 \quad j = 1, 2, \ldots, r \\
\sum_{j=1}^{r} x_{ij} = 1 \quad i = 1, 2, \ldots, r \\
y_i - y_j + \frac{1}{m} |x_{ij}| < \frac{1}{m} + m - 2 \quad \forall i \neq j, \text{ and } i, j \in I, \\
\text{where } I_0 = \{1, 2, \ldots, m\} \\
x_{ij} = 0 \text{ or } 1 \quad \forall i, j
\]

Constraint sets 1 and 2 together require the paths to include all the demand points.

Constraint set 3 blocks the formation of any infeasible tour.

d_{ij} are defined from the original c_{ij} by augmenting the \( ||c_{ij}|| \) matrix with m - 1 rows and m - 1 columns, where each new row and each new column is a duplicate of the first row and first column of the original matrix respectively.

**Problem 3.**

The Multi - Travelling Salesman Problem.

An alternative formulation to the problem is provided by Gavish and Srikanth (53).

Define : \( M = \) number of salesman

\( V = \{v_i\mid i = 1, \ldots, n\} \) as the sets of all nodes.

\( S = V - \{1\} = \{2, 3, \ldots n\} \)

\( c_{ij} = \) distance from node \( i \) to node \( j \)

\( A = \{(i, j) : 1 \leq i < j \leq n\} \) is the set of all arcs \( i, j \).
a subtour: a set of k arcs \( i_p \neq i_q, 1 \leq p, q \leq k \) is said to form a subtour of size k.

immediate subtour: \( ((1, i), (i, 1)) \)

\[
x_{ij} =
\begin{cases}
1 & \text{if } i \text{ and } j \text{ are connected } (i, j - S) \text{ or } i \text{ or } j \text{ is connected to node 1} \\
0 & \text{otherwise}
\end{cases}
\]

The formulation is

Minimize

\[
Z = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c_{ij} x_{ij} + \sum_{i \in S} c_{i1} x_{i1}
\]

subject to

\[
\sum_{j \in S} x_{ij} = M \quad \text{(1)}
\]

\[
\sum_{i \in S} x_{i1} = M \quad \text{(2)}
\]

\[
x_{ij} + \sum_{i<j} x_{ij} + \sum_{i>j} x_{ji} = 2 \quad \forall j \in S \quad \text{(3)}
\]

\[
\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_{ij} = n-1 \quad \text{(4)}
\]

\[
\sum_{i,j \in S} x_{ij} < |S_k| - 1 \quad \forall S_k \subseteq S \quad \text{(5)}
\]

\[
x_{ij} = 0 \text{ or } 1 \quad 1 \leq i < j \leq n \quad \text{(6)}
\]

\[
x_{i1} = 0 \text{ or } 1 \quad \forall i \in S \quad \text{(7)}
\]

The first constraint requires M salesman to depart from city 1 and the second requires M salesman to return to city 1. The third constraint ensures that a demand point is visited by one and only one salesman. The fourth constraint implies that any feasible solution to the problem
contains \( n + M - 1 \) arcs, of which \( M \) arcs represent salesman returning to city 1. The fifth set of constraints exclude the formation of any subtours which do not include city 1. The last two sets of constraints are the integrality constraints.
APPENDIX D

Problem 1

The Vehicle Routing Problem (VRP)

Define:

\( X = \{ x_i \mid i=1, \ldots, N \} \) is the set of customers, \( x_0 \) is depot

\( V = \{ v_k \mid k=1, \ldots, M \} \) is the set of vehicles at the depot

\( q_i \) = quantity of product demanded by customer \( x_i, i=1, \ldots, N \)

\( u_i \) = time required by a vehicle to visit customer \( x_i \) and unload

\( q_i \),

\( i=1, \ldots, N \)

\( Q_k \) = capacity of vehicle \( k, k=1, \ldots, M \)

\( T_{k}^{s} \) = start of working time for vehicle \( k, k=1, \ldots, M \)

\( T_{k}^{e} \) = end of working time for vehicle \( k, k=1, \ldots, M \)

\( C_k \) = fixed cost of vehicle \( k \)

\( C_{ij} \) = least cost path from \( i \) to \( j, i,j=0,1, \ldots, n, i \neq j \)

\( t_{ij} \) = travel time from \( i \) to \( j; i,j=0,1, \ldots, n, i \neq j \)

\( y_i \) = arbitrary real numbers

254
255

\[
X_{ijk} = \begin{cases} 
1 & \text{if vehicle } k \text{ visits customer } j \text{ immediately after visiting customer } i \\
0 & \text{otherwise}
\end{cases}
\]

The problem is formulated as:

Minimize \( Z = \sum_{i=0}^{M} \sum_{j=0}^{N} C_{ij} \sum_{k=1}^{M} X_{ijk} + \sum_{k=1}^{M} (C_{k} \sum_{j=1}^{N} X_{0kj}) \)

subject to

\[
\sum_{i=0}^{N} \sum_{k=1}^{M} X_{ijk} \leq 1, \quad j=1, \ldots, N \tag{1}
\]

\[
\sum_{i=0}^{N} X_{ipk} - \sum_{j=0}^{N} X_{pjk} = 0 \quad k=1, \ldots, M, \quad p=0,1, \ldots, N \tag{2}
\]

\[
\sum_{i=1}^{N} (q_{i} \sum_{j=0}^{N} X_{ijk}) \leq C_{k} \quad k=1, \ldots, M \tag{3}
\]

\[
\sum_{i=0}^{N} \sum_{j=0}^{N} t_{ij} X_{ijk} + \sum_{i=1}^{N} (u_{i} \sum_{j=0}^{N} X_{ijk}) \leq T_{k}^{f} - T_{k}^{s} \quad k=1, \ldots, M \tag{4}
\]

\[
\sum_{j=1}^{N} X_{0jk} \leq 1 \quad k=1, \ldots, M \tag{5}
\]

\[
y_{i} - y_{j} + N \sum_{k=1}^{M} X_{ijk} \leq N-1 \quad i,j = 1,2, \ldots, N \quad i \neq j \tag{6}
\]

\[
X_{ijk} = 0 \text{ or } 1 \quad \forall i, j, k \tag{7}
\]

Constraint set 1 implies that a customer can be visited at most once. Constraint set 2 requires that a vehicle visiting a customer must depart from it.
Constraint set 3 and 4 are vehicle related constraints, which require that the total load carried by a vehicle does not exceed the vehicle capacity, and the total working time does not exceed the allowable working time. Constraint set 5 implies that a vehicle can be used only once.

Constraint set 6 is a subtour breaking set of constraints, which force each vehicle route to pass through the depot. The last set of constraints are integrality constraints.

Problem 2

The Multi-Depot VRP

The multi-depot VRP formulation can be obtained by modifying the single depot VRP.

Denote $x_0, x_1, \ldots, x_{D-1}$ as the total of $D$ depots. Then, the new formulation is obtained by the following changes

(a) Change constraint (1) to

$$\sum_{i=0}^{N} \sum_{k=1}^{M} x_{ijk} \leq 1 \quad j=1, \ldots, D, D+1, \ldots, N$$

(b) Change constraint 5 to

$$\sum_{i=0}^{D-1} \sum_{j=D}^{N} x_{ijk} \leq 1, \quad k=1, \ldots, M$$

(c) Redefine the subtour elimination constraints (constraint 6) as

$$y_i - y_j + N \sum_{k=1}^{M} x_{ijk} \leq N-1 \text{ for } D \leq i \neq j \leq N$$
APPENDIX E

Solution to the Test Problem (Savings Drop Heuristic)

Cost Matrix ($C_{ij}$)

<table>
<thead>
<tr>
<th></th>
<th>d1</th>
<th>d2</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>c4</th>
<th>c5</th>
<th>c6</th>
<th>c7</th>
<th>c8</th>
<th>Depot #</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>999</td>
<td>54</td>
<td>10</td>
<td>47</td>
<td>49</td>
<td>22</td>
<td>52</td>
<td>22</td>
<td>43</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>D2</td>
<td>54</td>
<td>999</td>
<td>48</td>
<td>30</td>
<td>5</td>
<td>66</td>
<td>15</td>
<td>32</td>
<td>13</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>C1</td>
<td>10</td>
<td>48</td>
<td>999</td>
<td>38</td>
<td>43</td>
<td>20</td>
<td>44</td>
<td>16</td>
<td>36</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>C2</td>
<td>999</td>
<td>27</td>
<td>49</td>
<td>15</td>
<td>29</td>
<td>21</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C3</td>
<td>999</td>
<td>61</td>
<td>15</td>
<td>27</td>
<td>8</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C4</td>
<td>999</td>
<td>59</td>
<td>35</td>
<td>53</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C5</td>
<td>999</td>
<td>31</td>
<td>13</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C6</td>
<td>999</td>
<td>21</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C7</td>
<td>999</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>C8</td>
<td>999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>
### Hand Solution for the Problem

(1) Modified Costs (both depots open)

<table>
<thead>
<tr>
<th>Customer</th>
<th>$\bar{c}_i^1$</th>
<th>$\bar{c}_i^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>-28</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>-39</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>22</td>
<td>-22</td>
</tr>
<tr>
<td>5</td>
<td>-22</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>-17</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>19</td>
<td>-22</td>
</tr>
</tbody>
</table>
## Modified Savings

<table>
<thead>
<tr>
<th></th>
<th>D&lt;sub&gt;1&lt;/sub&gt;</th>
<th></th>
<th>D&lt;sub&gt;2&lt;/sub&gt;</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-53</td>
<td>-14</td>
<td>-24</td>
<td>+6</td>
</tr>
<tr>
<td>3</td>
<td>-78</td>
<td>-76</td>
<td>-44</td>
<td>-64</td>
</tr>
<tr>
<td>4</td>
<td>-59</td>
<td>+9</td>
<td>-48</td>
<td>+35</td>
</tr>
<tr>
<td>5</td>
<td>-31</td>
<td>-52</td>
<td>-56</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-16</td>
<td>+12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>-45</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Savings in Link Form</strong></td>
<td></td>
<td><strong>Opportunity Penalty</strong></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------------------</td>
<td>----------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$D_1$</td>
<td>$D_2$</td>
<td>$D_1$</td>
<td>$D_2$</td>
</tr>
<tr>
<td>1-2</td>
<td>-15</td>
<td>-36</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>1-3</td>
<td>-72</td>
<td>-66</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>1-4</td>
<td>+12</td>
<td>-70</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>1-5</td>
<td>-56</td>
<td>-57</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1-6</td>
<td>+16</td>
<td>-32</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>1-7</td>
<td>-43</td>
<td>-51</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>1-8</td>
<td>+15</td>
<td>-64</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>-53</td>
<td>+5</td>
<td>-</td>
<td>58</td>
</tr>
<tr>
<td>2-4</td>
<td>-14</td>
<td>-41</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>-24</td>
<td>+30</td>
<td>-</td>
<td>54</td>
</tr>
<tr>
<td>2-6</td>
<td>+6</td>
<td>+13</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>2-7</td>
<td>-25</td>
<td>+22</td>
<td>-</td>
<td>47</td>
</tr>
<tr>
<td>2-8</td>
<td>-11</td>
<td>-35</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>3-4</td>
<td>-78</td>
<td>-78</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3-5</td>
<td>-76</td>
<td>+5</td>
<td>-</td>
<td>81</td>
</tr>
<tr>
<td>3-6</td>
<td>-44</td>
<td>-10</td>
<td>-</td>
<td>34</td>
</tr>
<tr>
<td>3-7</td>
<td>-64</td>
<td>+10</td>
<td>-</td>
<td>74</td>
</tr>
<tr>
<td>3-8</td>
<td>-75</td>
<td>-72</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>4-5</td>
<td>-59</td>
<td>-66</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>4-6</td>
<td>+9</td>
<td>-45</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>4-7</td>
<td>-48</td>
<td>-62</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>4-8</td>
<td>+35</td>
<td>-50</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>
Frequency of Positive Savings (in all units)

<table>
<thead>
<tr>
<th></th>
<th>5-6</th>
<th>5-7</th>
<th>5-8</th>
<th>6-7</th>
<th>6-8</th>
<th>7-8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-31</td>
<td>-52</td>
<td>-56</td>
<td>-16</td>
<td>+12</td>
<td>-45</td>
</tr>
<tr>
<td></td>
<td>-4</td>
<td>+15</td>
<td>-60</td>
<td>+4</td>
<td>-39</td>
<td>-56</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>67</td>
<td>4</td>
<td>20</td>
<td>51</td>
<td>11</td>
</tr>
</tbody>
</table>

By rule of highest frequencies,

Route Customers 1, 4, 6, 8 from D₁; I₁ = {1, 4, 6, 8}

Route Customers 2, 3, 5, 7 from D₂; I₂ = {2, 3, 5, 7}
ROUTES:

(1) (a) No. of vehicles needed at $D_1$

$$\sum_{i \in I_1} q_i = q_1 + q_4 + q_6 + q_8$$

$$= 112 + 145 + 128 + 54$$

$$= 439$$

Select first vehicle from available set.

$k = 1 \quad Q_1 = 714$

$Q_1 > \sum_{i \in I_1} q_i$

Therefore, capacity is sufficient, $V_1 = \{1\}$, set of vehicles assigned to depot 1.

(b) Number of vehicles needed at $D_2$

$$\sum_{i \in I_2} q_i = q_2 + q_3 + q_5 + q_7$$

$$= 112 + 62 + 135 + 71$$

$$= 380$$

Select first vehicle from available set.

$k = 2 \quad Q_2 = 714$

$Q_2 > \sum_{i \in I_2} q_i$

Therefore capacity of vehicle 2 is sufficient to meet customer demand from depot 2.

$V_2 = \{2\}$: Set of vehicles assigned to depot 2.
Creating the Routes

For $D_1$

(i) Select link 4-8 with highest savings.

$I_1 = \{4, 8\} \quad I^1 = \{1, 6\}$

$LOAD_1 = q_4 + q_8$

$= 145 + 54$

$= 199$

(ii) Select link 1-6 with the next highest savings.

$I_1 = \{0\} \quad I^1 = \{1, 4, 6, 8\}$

$LOAD_1 = 199 + q_1 + q_6$

$= 439$

Since all customers have been assigned to the route, connect the open links to each other and to the depot so as to form a route.

Attach open links 1-4, thus the route is

$D_1 - 6 - 1 - 4 - 8 - D_1$

Costs

Travel cost $= C_i = 83$

Vehicle cost $= 26$

Depot cost $= 36$

Total Depot cost for depot $d_1 = 145$

Next, create the routes for open depot $D_2$

(1) Select link 2-5 with highest savings.

$I_1 = \{1, 2, 4, 5, 6, 8\}$
\( I_2 = (3, 7) \)

\[
\text{LOAD}_2 = q_2 + q_5 \\
= 112 + 135 \\
= 247
\]

(ii) Select the customer link with highest saving from the available links list.

Link 2-7 is selected.

This link can connect with link 2-5 only as 7 - 2 - 5.

\[
\text{LOAD}_2 = 247 + q_7 \\
= 247 + 71 \\
= 318
\]

\( I_1 = \{1, 2, 4, 5, 6, 7, 8\} \quad I_2 = \{3\} \)

(iii) Select the customer link with the highest savings from the available list.

In this case select link 3-7.

This link can only connect with existing partial route as 3 - 7 - 2 - 5.

\[
\text{LOAD}_2 = 318 + q_3 \\
= 318 + 62 \\
= 380
\]

\( I_1 = \{1, 2, 3, 4, 5, 6, 7, 8\} \quad I_2 = \emptyset \)

The route formed is

\( D_2 - 3 - 7 - 2 - 5 - D_2 \)
Costs

Travel cost = $C_{ij} = 64$

Vehicle cost = 25

Depot costs = 33

Total Depot Cost for depot 2 = 123

All customers have been assigned to routes from depots

\[ U = k^r = 0 \]

\[ r \in R_1 \]

Total System Cost for 2 depots = Total Depot 1 Cost + Total Depot 2 Cost

= 142 + 123

= 265

Penalty for closing one of the two open depots.

The opportunity penalties have been calculated.

Next, the maximum opportunity penalty is determined for each of the customers.
### Maximum Opportunity Penalty

<table>
<thead>
<tr>
<th>Customer</th>
<th>Depot D1</th>
<th>Depot D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>82</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>81</td>
</tr>
<tr>
<td>4</td>
<td>54</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>67</td>
</tr>
<tr>
<td>6</td>
<td>51</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>74</td>
</tr>
<tr>
<td>8</td>
<td>79</td>
<td>-</td>
</tr>
</tbody>
</table>

\[ \text{Overall Depot Penalty} \]

\[ dp_1^1 = 266 - F_1 = 266 - 36 = 230 \]

\[ dp_2^2 = 229 - F_2 = 229 - 33 = 196 \]

Since the depot with minimum overall penalty is depot 2, close depot 2.

\[ R_1 = \{1\} \quad R_2 = \{2\} \]

For depot 1, the only open depot, go back to Step 1.

Now, modified distances = actual distances (only 1 depot)
Modified Savings = Actual savings for the 1 depot case.

Savings for depot $D_1$

\[
\begin{array}{cccccccc}
\backslash & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
1 & 19 & 16 & 12 & 18 & 16 & 17 & 15 \\
2 & 69 & 20 & 84 & 40 & 69 & 23 \\
3 & 10 & 86 & 44 & 84 & 13 \\
4 & & 15 & 7 & 12 & 35 \\
5 & & & 43 & 82 & 18 \\
6 & & & & 44 & 12 \\
7 & & & & & 15 \\
8 & & & & & & \\
\end{array}
\]

Since there is only one depot open, all customers are to be routed from the depot.

First determine the number of vehicles needed at $D_1$

\[
\sum_{i \in I} q_i = q_1 + \ldots + q_8 = 819
\]

Select first vehicle from set of available vehicles

$k = 1 \quad Q_1 = 714$

Since $Q_1 < \sum q_i$

capacity is insufficient

Add another vehicle to the set of vehicles assigned to depot $D_1$

Select $k = 2 \quad Q_2 = 714$
\[ q_1 + q_2 = 1428 > \sum q_i \]

\[ V_1 = \{1, 2\} \]

Two vehicles are sufficient to route all customers from \( D_1 \), i.e. two routes need to be formed.

Next, the routes are created.

For route 1,
- Select link 3-5 with the highest saving.
- \( \text{LOAD}_1 = 62 + 135 \)
  \[ = 197 \]

For route 2,
- Select link 2-7 with the next highest saving, from available links.
- \( \text{LOAD}_2 = 112 + 71 \)
  \[ = 183 \]

Next, start building both the routes.

From the ranked list of savings select the next highest savings from feasible links. (Links such as 3-7, or 2-5 etc. are no longer feasible).

Select link 3-6.

This can only be added to route 1 since customer 3 is on route 1.

Select link 4-8 with the next highest savings. Assign to route 1.

Free links are assigned to available routes in order, one after the other. A check is made to ensure that the link is being added to a route where its distance from the existing partial route is smallest.

Next, select link 1-2.
This link can only be assigned to route 2 since customer 2 is already on route 2.
The partial route will be 1 - 2 - 7.
At this stage, all customers have been assigned to routes. The links are joined together to form routes.
The routes are:

Route 1: \( D_1 - 6 - 3 - 5 - 4 - 8 - D_1 \)

Route 2: \( D_1 - 1 - 2 - 7 - D_1 \)

\( \text{LOAD}_1 = 524 \)

\( \text{LOAD}_2 = 294 \)

Total travel cost = 260

Vehicles cost = 52

Depot costs = 36

Total depot cost for depot 1 = 348

Total depot cost for 1 depot = 348

In this problem, the system with two depot sites is selected with minimum total system cost.