12
Facility Location Models

12.1 Introduction

One of the most important aspects of logistics is deciding where to locate new facilities such as retailers, warehouses or factories. These strategic decisions are a crucial determinant of whether materials will flow efficiently through the distribution system.

In this chapter we consider several important warehouse location problems: the p-Median Problem, the Single-Source Capacitated Facility Location Problem and a distribution system design problem. In each case, the problem is to locate a set of warehouses in a distribution network. We assume that the cost of locating a warehouse at a particular site includes a *fixed* cost (e.g., building costs, rental costs, etc.) and a *variable* cost for transportation. This variable cost includes the cost of transporting the product to the retailers as well as possibly the cost of moving the product from the plants to the warehouse. In general, the objective is to locate a set of facilities so that total cost is minimized subject to a variety of constraints which might include:

- each warehouse has a capacity which limits the area it can supply.
- each retailer receives shipments from *one and only one* warehouse.
- each retailer must be within a fixed distance of the warehouse that supplies it, so that a reasonable delivery lead time is ensured.

Location analysis has played a central role in the development of the operations research field. In this area lie some of the discipline's most elegant results and theories. We note here the paper of Cornuéjols et al. (1977) and the two excellent books devoted to the subject by Mirchandani and Francis (1990) and Daskin (1995). Location problems encompass a wide range of problems such as the location of emergency services (fire houses or ambulances), the location of hazardous materials, problems in telecommunications network design, etc. just to name a few.

In the next section, we present an exact algorithm for one of the simplest location problems, the p-Median Problem. We then generalize this model and algorithm to incorporate additional factors important to the design of the distribution network, such as warehouse capacities and fixed costs. In Section 12.4, we present a more general model where all levels of the distribution system (plants and retailers) are taken into account when deciding warehouse locations. We also present an efficient algorithm for its solution. All of the algorithms developed in this chapter are based on the Lagrangian relaxation technique described in Chapter 5.3 which has been applied successfully to a wide range of location problems. Finally, in Section 12.5, we describe the structure of the optimal solution to problems in the design of large-scale logistics systems.

12.2 An Algorithm for the p-Median Problem

Consider a set of retailers geographically dispersed in a region. The problem is to choose where in the region to locate a set of p identical warehouses. We assume there are $m \geq p$ sites that have been preselected as possible locations for these warehouses. Once the p warehouses have been located, each of n retailers will get its shipments from the warehouse closest to it. We assume:

- there is no fixed cost for locating at a particular site, and
- there is no capacity constraint on the demand supplied by a warehouse.

Note that the first assumption also encompasses the case where the fixed cost is not site-dependent and therefore the fixed set-up cost for locating p warehouses is independent of where they are located.

Let the set of retailers be N where $N = \{1, 2, \ldots, n\}$, and let the set of potential sites for warehouses be M where $M = \{1, 2, \ldots, m\}$. Let w_i be the demand or flow between retailer i and its warehouse for each $i \in N$. We assume that the cost of transporting the w_i units of product from warehouse j to retailer i is c_{ij}, for each $i \in N$ and $j \in M$.

The problem is to choose p of the m sites where a warehouse will be located in such a way that the total transportation cost is minimized. This is the p-Median Problem.

The continuous version of this problem, where any point is a potential warehouse location, was first treated as early as 1909 by Weber. The discrete version was analyzed by Kuehn and Hamburger (1963) as well as Hakimi (1964), Manne (1964), Balinski (1965) and many others.

We present here a highly effective approach to the problem. Define the following decision variables:

$$Y_j = \begin{cases} 1, & \text{if a warehouse is located at site } j, \\ 0, & \text{otherwise,} \end{cases}$$

for $j \in M$, and

$$X_{ij} = \begin{cases} 1, & \text{if retailer } i \text{ is served by a warehouse at site } j, \\ 0, & \text{otherwise,} \end{cases}$$

for $i \in N$ and $j \in M$. The p-Median Problem is then:

$$\text{Problem } P: \text{Min } \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} X_{ij}$$

$$s.t. \quad \sum_{j=1}^{m} X_{ij} = 1, \quad \forall i \in N \tag{12.1}$$

$$\sum_{j=1}^{m} Y_j = p \tag{12.2}$$

$$X_{ij} \leq Y_j, \quad \forall i \in N,\ j \in M \tag{12.3}$$

$$X_{ij}, Y_j \in \{0, 1\}, \quad \forall i \in N,\ j \in M. \tag{12.4}$$

Constraints (12.1) guarantee that each retailer is assigned to a warehouse. Constraint (12.2) ensures that p sites are chosen. Constraints (12.3) guarantee that a retailer selects a site only from among those that are chosen. Constraints (12.4) force the variables to be integer.

This formulation can easily handle several side constraints. If a handling fee is charged for each unit of product going through a warehouse, these costs can be added to the transportation cost along all arcs leaving the warehouse. Also, if a particular limit is placed on the length of any arc between retailer i and warehouse j, this can be incorporated by simply setting the per unit shipping cost (c_{ij}) to $+\infty$. In addition, the model can be easily extended to cases where a set of facilities are already in place and the choice is whether to open new facilities or *expand* the existing facilities.

Let Z^* be an optimal solution to Problem P. One simple and effective technique to solve this problem is the method of Lagrangian relaxation described in Chapter 5.3.

As described in Chapter 5.3, Lagrangian relaxation involves relaxing a set of constraints and introducing them into the objective function with a multiplier vector. This provides a lower bound on the optimal solution to the overall problem. Then, using a subgradient search method, we iteratively update our multiplier vector in an attempt to increase the lower bound. At each step of the subgradient procedure (i.e., for each set of multipliers) we also attempt to construct a feasible solution to the location problem. This step usually consists of a simple and efficient subroutine. After a prespecified number of iterations, or when the solution found is within a fixed error tolerance of the lower bound, the algorithm is terminated.

To solve the p-Median Problem, we choose to relax constraints (12.1). We incorporate these constraints in the objective function with the multiplier vector $\lambda \in I\!R^n$. The resulting problem, call it P_λ, with optimal objective function value Z_λ, is:

$$\text{Min} \quad \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}X_{ij} + \sum_{i=1}^{n}\lambda_i\Big(\sum_{j=1}^{m} X_{ij} - 1\Big)$$

$$\text{subject to} \quad (12.2) - (12.4).$$

Disregarding constraint (12.2) for now, the problem decomposes by site, that is, each site can be considered separately. Let subproblem P_λ^j, with optimal objective function value Z_λ^j, be the following.

$$\begin{aligned} \text{Min} \quad & \sum_{i=1}^{n}(c_{ij} + \lambda_i)X_{ij} \\ s.t. \quad & X_{ij} \le Y_j, \quad \forall i \in N \\ & X_{ij} \in \{0,1\}, \quad \forall i \in N \\ & Y_j \in \{0,1\}. \end{aligned}$$

Solving Subproblem P_λ^j

Assume λ is fixed. In Problem P_λ^j, site j is either selected ($Y_j = 1$) or not ($Y_j = 0$). If site j is not selected, then $X_{ij} = 0$ for all $i \in N$ and therefore $Z_\lambda^j = 0$. If site j is selected, then we set $Y_j = 1$ and assign exactly those retailers i with $c_{ij} + \lambda_i < 0$ to site j. In this case:

$$Z_\lambda^j = \sum_{i=1}^{n} \min\{c_{ij} + \lambda_i, 0\}. \tag{12.5}$$

We see that P_λ^j is solved easily and its optimal objective function value is given by (12.5).

To solve P_λ, we must now reintroduce constraint (12.2). This constraint forces us to choose only p of the m sites. In P_λ, we can incorporate this constraint by choosing the p sites with smallest values Z_λ^j. To do this, let π be a permutation of the numbers $1, 2, \ldots, m$ such that

$$Z_\lambda^{\pi(1)} \le Z_\lambda^{\pi(2)} \le Z_\lambda^{\pi(3)} \le \cdots \le Z_\lambda^{\pi(m)}.$$

Then the optimal solution to P_λ has objective function value:

$$Z_\lambda \doteq \sum_{j=1}^{p} Z_\lambda^{\pi(j)} - \sum_{j=1}^{n} \lambda_j.$$

The value Z_λ is a lower bound on the optimal solution of Problem P for any vector $\lambda \in I\!R^n$. To find the best such lower bound, we consider the Lagrangian dual:

$$\max_\lambda\{Z_\lambda\}.$$

Using the subgradient procedure (described in Chapter 5.3), we can iteratively improve this bound.

Upper Bounds

It is crucial to construct good upper bounds on the optimal solution value as the subgradient procedure advances. Clearly, solutions to P_λ will not necessarily be feasible to Problem P. This is due to the fact that the constraints (12.1) (that each retailer choose *one and only one* warehouse) may not be satisfied. The solution to P_λ may have facilities choosing a number of sites. If, in the solution to P_λ, each retailer chooses only *one* site, then this must be the optimal solution to P and therefore we stop. Otherwise, there are retailers that are assigned to several or no sites. A simple heuristic can be implemented which fixes those retailers that are assigned to only one site, and assigns the remaining retailers to these and other sites by choosing the next site to open in the ordering defined by π. When p sites have been selected, a simple check that each retailer is assigned to its closest site (of those selected) can further improve the solution.

Computational Results

Below we give a table listing results of various computational experiments. The retailer locations were chosen uniformly over the unit square. For simplicity, we made each retailer location a potential site for a warehouse, thus $m = n$. The cost of assigning a retailer to a site was the Euclidean distance between the two locations. The values of w_i were chosen uniformly over the unit interval. We applied the algorithm mentioned above to many problems and recorded the relative error of the best solution found and the computation time required. The algorithm is terminated when the relative error is below 1% or when a prespecified number of iterations is reached. The numbers below "Error" are the relative errors averaged over ten randomly generated problem instances. The numbers below "CPU Time" is the CPU time averaged over the ten problem instances. All computational times are on an IBM Risc 6000 Model 950.

Table 1: Computational results for the p-Median algorithm

n	p	Error	CPU Time
10	3	0.3%	0.2s
20	4	1.7%	2.6s
50	5	1.4%	20.7s
100	7	1.3%	87.7s
200	10	2.4%	715.4s

12.3 An Algorithm for the Single-Source Capacitated Facility Location Problem

Consider the p-Median Problem where we make the following two changes in our assumptions.

- The number of warehouses to locate (p) is not fixed beforehand.
- If a warehouse is located at site j:
  - a fixed cost f_j is incurred, and
  - there is a capacity q_j on the amount of demand it can serve.

The problem is to decide where to locate the warehouses and then how the retailers should be assigned to the open warehouses in such a way that total cost is minimized. We see that the problem is considerably more complicated than the p-Median Problem. We now have capacity constraints on the warehouses and therefore a retailer will not always be assigned to its nearest warehouse. Allowing the optimization to choose the appropriate number of warehouses also adds to the level of difficulty.

This problem is called the single-source Capacitated Facility Location Problem (CFLP), or sometimes the Capacitated Concentrator Location Problem (CCLP). This problem was successfully used in Chapter 14 as a framework for solving the Capacitated Vehicle Routing Problem.

Using the same decision variables as in the p-Median Problem, we formulate the single-source CFLP as the following integer linear program.

$$\begin{aligned}
\text{Min} \quad & \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}X_{ij} + \sum_{j=1}^{m} f_j Y_j \\
\text{s.t.} \quad & \sum_{j=1}^{m} X_{ij} = 1 \qquad \forall i \in N && (12.6)\\
& \sum_{i=1}^{n} w_i X_{ij} \le q_j Y_j \qquad \forall j \in M && (12.7)\\
& X_{ij}, Y_j \in \{0,1\} \qquad \forall i \in N,\ j \in M. && (12.8)
\end{aligned}$$

Constraints (12.6) (along with the integrality conditions (12.8)) ensure that each retailer is assigned to exactly one warehouse. Constraints (12.7) ensure that the warehouse's capacity is not exceeded, and also that if a warehouse is not located at site j, no retailer can be assigned to that site.

Let Z^* be the optimal solution value of single-source CFLP. Note we have restricted the assignment variables (X) to be integer. A related problem, where this assumption is relaxed, is simply called the (multiple-source) Capacitated Facility Location Problem. In that version, a retailer's demand can be *split* between any number of warehouses. In the single-source CFLP, it is required that each retailer have only *one* warehouse supplying it. In many logistics applications, this is a realistic assumption since without this restriction optimal solutions might have a retailer receive many deliveries of the same product (each for, conceivably, a very small amount of the product). Clearly, from a managerial, marketing and accounting point of view, restricting deliveries to come from only one warehouse is a more appropriate delivery strategy.

Several algorithms have been proposed to solve the CFLP in the literature; all are based on the Lagrangian relaxation technique. This includes Neebe and Rao (1983), Barcelo and Casanovas (1984), Klincewicz and Luss (1986) and Pirkul (1987). The one we derive here is similar to the algorithm of Pirkul which seems to be the most effective.

We apply the Lagrangian relaxation technique by including constraints (12.6) in the objective function. For any $\lambda \in \mathbb{R}^n$, consider the following problem P_λ.

$$\begin{aligned}
\text{Min} \quad & \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}X_{ij} + \sum_{j=1}^{m} f_j Y_j + \sum_{i=1}^{n} \lambda_i\Big(\sum_{j=1}^{m} X_{ij} - 1\Big) \\
& \text{subject to} \quad (12.7) - (12.8).
\end{aligned}$$

Let Z_λ be its optimal solution and note that

$$Z_\lambda \le Z^*, \ \forall \lambda \in \mathbb{R}^n.$$

To solve P_λ, as in the p-Median Problem, we separate the problem by site. For a given $j \in M$, define the following problem P_λ^j, with optimal objective function value Z_λ^j:

$$\begin{aligned}
\text{Min} \quad & \sum_{i=1}^{n} (c_{ij} + \lambda_i) X_{ij} + f_j Y_j \\
\text{s.t.} \quad & \sum_{i=1}^{n} w_i X_{ij} \le q_j Y_j \\
& X_{ij} \in \{0,1\} \qquad \forall i \in N \\
& Y_j \in \{0,1\}.
\end{aligned}$$

Solving P_λ^j

Problem P_λ^j can be solved efficiently. In the optimal solution to P_λ^j, Y_j is either 0 or 1. If $Y_j = 0$, then $X_{ij} = 0$ for all $i \in N$. If $Y_j = 1$, then the problem is no more difficult than a single constraint 0-1 Knapsack Problem, for which efficient algorithms exist; see, for example, Nauss (1976). If the optimal knapsack solution is less than $-f_j$, then the corresponding optimal solution to P_λ^j is found by setting $Y_j = 1$ and X_{ij} according to the knapsack solution, indicating whether retailer i is assigned to site j. If the optimal knapsack solution is more than $-f_j$, then the optimal solution to P_λ^j is found by setting $Y_j = 0$ and $X_{ij} = 0$ for all $i \in N$.

The solution to P_λ is then given by

$$Z_\lambda \doteq \sum_{j=1}^{m} Z_\lambda^j - \sum_{i=1}^{n} \lambda_i.$$

For any vector $\lambda \in \mathbb{R}^n$, this is a lower bound on the optimal solution Z^*. In order to find the best such lower bound we use a subgradient procedure.

Note that if the problem has a constraint on the number of warehouses (facilities) that can be opened (chosen), this can be handled in essentially the same way as it was handled in the algorithm for the p-Median Problem.

Upper Bounds

For a given set of multipliers, if the values $\{X\}$ satisfy (12.6), then we have an optimal solution to Problem P, and we stop. Otherwise, we perform a simple subroutine to find a feasible solution to P. The procedure is based on the observation that the knapsack solutions found when solving P_λ give us some information concerning the benefit of setting up a warehouse at a site (relative to the current vector λ). If, for example, the knapsack solution corresponding to a given site is 0, that is, the optimal knapsack is empty, then this is most likely not a "good" site to select at this time. In contrast, if the knapsack solution has a very negative cost, then this is a "good" site. Given the values Z_λ^j for each $j \in M$, let π be a permutation of $1, 2, \ldots, m$ such that

$$Z_\lambda^{\pi(1)} \le Z_\lambda^{\pi(2)} \le \cdots \le Z_\lambda^{\pi(m)}.$$

The procedure we perform allocates retailers to sites in a myopic fashion. Let M be the minimum possible number of warehouses used in the optimal solution to CFLP. This can be found by solving the Bin-Packing Problem defined on the values w_i with bin capacities Q_j; see Section 3.2. Starting with the "best" site, in this case site $\pi(1)$, assign the retailers in its optimal knapsack to this site. Then, following the indexing of the knapsack solutions, take the next "best" site (say site $j \doteq \pi(2)$) and solve a new knapsack problem: one defined with costs $\bar{c}_{ij} \doteq c_{ij} + \lambda_i$ for each retailer i still unassigned. Assign all retailers in this knapsack solution to site j. If this optimal knapsack is empty, then a warehouse is not located at that site, and we go on to the next site. Continue in this manner until M warehouses are located.

The solution may still not be a feasible solution to P since some retailers may not be assigned to a site. In this case, unassigned retailers are assigned to sites that are already chosen where they fit with minimum additional cost. If needed, additional warehouses may be opened following the ordering of π. A local improvement heuristic can be implemented to improve on this solution, using simple interchanges between retailers.

Computational Results

We now report on various computational experiments using this algorithm. The retailer locations were chosen uniformly over the unit square. Again, for simplicity, we made each retailer location a potential site for a warehouse; thus, $m = n$. The fixed cost of a site was chosen uniformly between 0 and 1. The cost of assigning a retailer to a site was the Euclidean distance between the two locations. The values of w_i were chosen uniformly over the interval 0 to $\frac{1}{2}$ with warehouse capacity equal to 1. We applied the algorithm mentioned above to ten problems and recorded the average relative error of the best solution found and the average computation time required. The algorithm is terminated when the relative error is below 1% or when a prespecified number of iterations is reached. The numbers below "Error" are the relative errors averaged over the ten randomly generated problem instances. The numbers below "CPU Time" is the CPU time averaged over the ten problem instances. All computational times are on an IBM Risc 6000 Model 950.

Table 2: Computational results for the single-source CFLP algorithm

n	Error	CPU Time
10	1.2%	1.2s
20	1.0%	8.1s
50	1.1%	110.0s
100	1.1%	558.3s

12.4 A Distribution System Design Problem

So far the location models we have considered have been concerned with minimizing the costs of transporting products between warehouses and retailers. We now present a more realistic model that considers the cost of transporting the product from manufacturing facilities to the warehouses as well.

Consider the following warehouse location problem. A set of plants and retailers are geographically dispersed in a region. Each retailer experiences demands for a variety of products which are manufactured at the plants. A set of warehouses must be located in the distribution network from a list of potential sites.

The cost of locating a warehouse includes the transportation cost per unit from warehouses to retailers but also the transportation cost from plants to warehouses. In addition, as in the CFLP, there is a site-dependent fixed cost for locating each warehouse.

The data for the problem are the following.

- L = number of plants; we will also let $L = \{1, 2, \ldots, L\}$
- J = number of potential warehouse sites, also let $J = \{1, 2, \ldots, J\}$
- I = number of retailers, also let $I = \{1, 2, \ldots, I\}$
- K = number of products, also let $K = \{1, 2, \ldots, K\}$

- W = number of warehouses to locate
- $c_{\ell jk}$ = cost of shipping one unit of product k from plant ℓ to warehouse site j
- d_{jik} = cost of shipping one unit of product k from warehouse site j to retailer i
- f_j = fixed cost of locating a warehouse at site j
- $v_{\ell k}$ = supply of product k at plant ℓ
- w_{ik} = demand for product k at retailer i
- s_k = volume of one unit of product k
- q_j = capacity (in volume) of a warehouse at site j

We make the additional assumption that a retailer gets delivery for a product from one warehouse only. This does not preclude solutions where a retailer gets shipments from different warehouses, but these shipments must be for different products. On the other hand, we assume that the warehouse can receive shipments from any plant and for any amount of product.

The problem is to determine where to locate the warehouses, how to ship the product from the plants to the warehouses and also how to ship the product from the warehouses to the retailers. This problem is similar to one analyzed by Pirkul and Jayaraman (1996).

We again use a mathematical programming approach. Define the following decision variables:

$$Y_j = \begin{cases} 1, & \text{if a warehouse is located at site } j, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$U_{\ell jk} = \text{amount of product } k \text{ shipped from plant } \ell \text{ to warehouse } j,$$

for each $\ell \in L$, $j \in J$ and $k \in K$. Also define:

$$X_{jik} = \begin{cases} 1, & \text{if retailer } i \text{ receives product } k \text{ from warehouse } j, \\ 0, & \text{otherwise,} \end{cases}$$

for each $j \in J$, $i \in I$ and $k \in K$.

Then the Distribution System Design Problem can be formulated as the following integer program.

$$\text{Min} \quad \sum_{\ell=1}^{L}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{\ell jk}U_{\ell jk} + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} d_{jik}w_{ik}X_{jik} + \sum_{j=1}^{J} f_jY_j$$

$$\text{s.t.} \quad \sum_{j=1}^{J} X_{jik} = 1 \qquad \forall i \in I,\ k \in K \tag{12.9}$$

$$\sum_{i=1}^{I}\sum_{k=1}^{K} s_k w_{ik} X_{jik} \le q_j Y_j \qquad \forall j \in J \tag{12.10}$$

$$\sum_{i=1}^{I} w_{ik}X_{jik} = \sum_{\ell=1}^{L} U_{\ell jk} \qquad \forall j \in J,\ k \in K \tag{12.11}$$

$$\sum_{j=1}^{J} U_{\ell jk} \le v_{\ell k} \qquad \forall \ell \in L,\ k \in K \tag{12.12}$$

$$\sum_{j=1}^{J} Y_j = W \tag{12.13}$$

$$Y_j, X_{jik} \in \{0,1\} \qquad \forall i \in I, j \in J, k \in K \tag{12.14}$$

$$U_{\ell jk} \ge 0 \qquad \forall \ell \in L, j \in J, k \in K. \tag{12.15}$$

The objective function measures the transportation costs between plants and warehouses, between warehouses and retailers and also the fixed cost of locating the warehouses. Constraints (12.9) ensure that each retailer/product pair is assigned to one warehouse. Constraints (12.10) guarantee that the capacity of the warehouses is not exceeded. Constraints (12.11) ensure that there is a conservation of the flow of products at each warehouse; that is, the amount of each product arriving at a warehouse from the plants is equal to the amount being shipped from the warehouse to the retailers. Constraints (12.12) are the supply constraints. Constraints (12.13) ensure that we locate exactly W warehouses.

The model can handle several extensions such as a warehouse handling fee or a limit on the distance of any link used just as in the p-Median Problem. Another interesting extension is when there are a fixed number of possible warehouse types from which to choose. Each type has a specific cost along with a specific capacity. The model can be easily extended to handle this situation (see Exercise 12.1).

As in the previous problems, we will use Lagrangian relaxation. We relax constraints (12.9) (with multipliers λ_{ik}) and constraints (12.11) (with multipliers θ_{jk}). The resulting problem is:

$$\begin{aligned} \text{Min} \quad & \sum_{\ell=1}^{L}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{\ell jk}U_{\ell jk} + \sum_{j=1}^{J}\sum_{i=1}^{I}\sum_{k=1}^{K} d_{jik}w_{ik}X_{jik} + \sum_{j=1}^{J} f_jY_j \\ & + \sum_{j=1}^{J}\sum_{k=1}^{K} \theta_{jk}\Big[\sum_{i=1}^{I} w_{ik}X_{jik} - \sum_{\ell=1}^{L} U_{\ell jk}\Big] + \sum_{i=1}^{I}\sum_{k=1}^{K} \lambda_{ik}\Big[1 - \sum_{j=1}^{J} X_{jik}\Big], \end{aligned}$$

subject to $(12.10), (12.12) - (12.15)$.

Let $Z_{\lambda,\theta}$ be the optimal solution to this problem. This problem can be decomposed into two separate problems P_1 and P_2. They are the following.

$$\text{Problem } P_1 : Z_1 \doteq \text{Min} \quad \sum_{\ell=1}^{L}\sum_{j=1}^{J}\sum_{k=1}^{K}[c_{\ell jk} - \theta_{jk}]U_{\ell jk}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} U_{\ell jk} \le v_{\ell k}, \forall \ell \in L, \; k \in K \tag{12.16}$$

$$U_{\ell jk} \ge 0, \quad \forall \ell \in L, j \in J, k \in K.$$

$$\text{Problem } P_2 : Z_2 \doteq \text{Min} \quad \sum_{j=1}^{J}\sum_{i=1}^{I}\sum_{k=1}^{K}[d_{jik}w_{ik} - \lambda_{ik} + \theta_{jk}w_{ik}]X_{jik} + \sum_{j=1}^{J} f_j Y_j$$

$$\text{s.t.} \quad \sum_{i=1}^{I}\sum_{k=1}^{K} s_k w_{ik} X_{jik} \le q_j Y_j, \quad \forall j \in J \tag{12.17}$$

$$\sum_{j=1}^{J} Y_j = P, \tag{12.18}$$

$$Y_j, X_{jik} \in \{0,1\}, \quad \forall i \in I, j \in J, k \in K.$$

Solving P_1

Problem P_1 can be solved separately for each plant/product pair. In fact, the objective functions of each of these subproblems can be improved (without loss in computation time) by adding the constraints:

$$s_k \sum_{\ell=1}^{L} U_{\ell jk} \le q_j, \quad \forall j \in J, \; k \in K. \tag{12.19}$$

For each plant/product combination, say plant ℓ and product k, sort the J values $\bar{c}_j \doteq c_{\ell jk} - \theta_{jk}$. Starting with the smallest value of $\bar{c}_j$, say $\bar{c}_{j'}$, if $\bar{c}_{j'} \ge 0$, then the solution is to ship none of this product from this plant. If $c_{\ell j' k} < 0$, then ship as much of this product as possible along arc (ℓ, j') subject to satisfying constraints (12.16) and (12.19). Then if the supply $v_{\ell k}$ has not been completely shipped, do the same for the next cheapest arc, as long as it has negative reduced cost $(\bar{c})$. Continue in this manner until all of the product has been shipped or the reduced costs are no longer negative. Then proceed to the next plant/product combination repeating this procedure. Continue until all the plant/product combinations have been scanned in this fashion.

Solving P_2

Solving Problem P_2 is similar to solving the subproblem in the CFLP. For now we can ignore constraints (12.18). Then we separate the problem by warehouse. In the problem corresponding to warehouse j, either $Y_j = 0$ or $Y_j = 1$. If $Y_j = 0$, then $X_{jik} = 0$ for all $i \in N$ and $k \in K$. If $Y_j = 1$, then we get a Knapsack Problem with NK items, one for each retailer/product pair. Let Z_2^j be the objective function value when Y_j is set to 1 and the resulting knapsack problem is solved. After having solved each of these, let π be a permutation of the numbers $1, 2, \ldots, J$ such that

$$Z_2^{\pi(1)} \le Z_2^{\pi(2)} \le \cdots \le Z_2^{\pi(J)}.$$

The optimal solution to P_2 is to choose the W smallest values:

$$Z_2 \doteq \sum_{j=1}^{W} Z_2^{\pi(j)}.$$

For fixed vectors λ and θ, the lower bound is

$$Z_{\lambda,\theta} \doteq Z_1 + Z_2 + \sum_{i=1}^{I}\sum_{k=1}^{K} \lambda_{ik}.$$

To maximize this bound, that is,

$$\max_{\lambda,\theta}\{Z_{\lambda,\theta}\},$$

we again use the subgradient optimization procedure.

Upper Bounds

At each iteration of the subgradient procedure, we attempt to construct a feasible solution to the problem. Consider Problem P_2. Its solution may have a retailer/product combination assigned to several warehouses. We determine the set of retailer/product combinations that are assigned to one and only one retailer and fix these. Other retailer/product combinations are assigned to warehouses using the following mechanism. For each retailer/product combination we determine the cost of assigning it to a particular warehouse. After determining that this assignment is feasible (from a warehouse capacity point of view), the assignment cost is calculated as the cost of shipping all of the demand for this retailer/product combination through the warehouse plus the cost of shipping the demand from the plants to the warehouse (along one or more arcs from the warehouse to the plants). For each retailer/product combination we determine the penalty associated with assigning the shipment to its second best warehouse instead of its best warehouse. We then assign the retailer/product combination with the highest such penalty and update all arc flows and remaining capacities. We continue in this manner until all retailer/product combinations have been assigned to warehouses.

Computational results for this problem appear at the end of Chapter 17.

12.5 The Structure of the Asymptotic Optimal Solution

In this section we describe a region partitioning scheme to solve large instances of the CFLP.

Assume there are n retailers located at points $\{x_1, x_2, \ldots, x_n\}$. Each retailer also serves as a potential site for a warehouse of fixed capacity q. The fixed cost of locating a warehouse at a site is assumed to be proportional to the distance the site is from a manufacturing facility located at x_0 which is assumed (without loss of generality) to be the origin $(0,0)$. Retailer i has a demand w_i which is assumed to be less than or equal to q. Without loss of generality, we assume $q = 1$ and therefore $w_i \in [0,1]$ for each $i \in N$. Let α be the per unit cost of transportation between warehouses and the manufacturing facility, and let β be the per unit cost of transportation between warehouses and retailers.

We assume the retailer locations are independently and identically distributed in a compact region $A \subset I\!R^2$ according to some distribution μ. Assume the retailer demands are independently and identically distributed according to a probability measure ϕ on $[0,1]$. The bin-packing constant associated with the distribution ϕ (denoted by γ_ϕ or simply γ) is the asymptotic number of bins used per item in an optimal packing of the retailer demands into unit size bins, when items are drawn randomly from the distribution ϕ (see Section 4.2).

The following theorem shows that if the retailer locations and demand sizes are random (from a general class of distributions), then as the problem size increases, the optimal solution has a very particular structure. This structure can be exploited using a region partitioning scheme as demonstrated below.

Theorem 12.5.1 *Let x_k, $k = 1, 2, \ldots, n$ be a sequence of independent random variables having a distribution μ with compact support in $I\!R^2$. Let $\|x\|$ be the Euclidean distance between the manufacturing facility and the point $x \in I\!R^2$, and let*

$$E(d) = \int \|x\| d\mu(x).$$

Let the demands w_k, $k = 1, 2, \ldots, n$ be a sequence of independent random variables having a distribution ϕ with bin-packing constant equal to γ. Then, almost surely,

$$\lim_{n \to \infty} \frac{1}{n} Z_n^* = \alpha \gamma E(d).$$

This analysis demonstrates that simple approaches which consider only the geography and the packing of the demands can be very efficient on large problem instances. Asymptotically, this is in fact the optimal strategy. This analysis also demonstrates that, asymptotically, the cost of transportation between retailers and warehouses becomes a very small fraction (eventually zero) of the total cost.

12.6 Exercises

Exercise 12.1. In the Distribution System Design Problem, explain how the solution methodology changes when there are a fixed number of possible warehouse capacities. For example, at each site, if we decide to install a warehouse, we can install a *small*, *medium* or *large* one.

Exercise 12.2. Prove Theorem 12.5.1.

Exercise 12.3. Show how any instance of the Bin-Packing Problem (see Part I) can be formulated as an instance of the Single-Source CFLP.

Exercise 12.4. Consider Problem P_1 of Section 12.4.

(a) Show that this formulation can be strengthened by adding the constraints:

$$\sum_{\ell=1}^{L} \sum_{k=1}^{K} s_k U_{\ell jk} \leq q_j, \qquad \forall j \in J.$$

(b) Show that this new formulation can be transformed to a specialized kind of linear program called a transportation problem.

(c) Why might we not want to use this stronger formulation?

Exercise 12.5. (Mirchandani and Francis, 1990) Define the Uncapacitated Facility Location Problem (UFLP) in the following way. Let F_j be the fixed charge of opening a facility at site j, for $j = 1, 2, \ldots, m$.

$$\begin{aligned} \text{Problem UFLP} \quad : \text{Min} \quad & \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} X_{ij} + \sum_{j=1}^{m} F_j Y_j \\ s.t. \quad & \sum_{j=1}^{m} X_{ij} = 1, \quad \forall i \in N \\ & X_{ij} \leq Y_j, \quad \forall i \in N, \ j \in M \\ & X_{ij}, Y_j \in \{0,1\}, \quad \forall i \in N, \ j \in M. \end{aligned}$$

Show that UFLP is $\mathcal{NP}$-Hard by showing that any instance of the $\mathcal{NP}$-Hard Node Cover Problem can be formulated as an instance of UFLP. The Node Cover Problem is defined as follows: given a graph G and an integer k, does there exist a subset of k nodes of G that cover all the arcs of G? (Node v is said to cover arc e if v is an end-point of e.)

Exercise 12.6. (Mirchandani and Francis, 1990) It appears that the p-Median problem can be solved by solving the resulting problem UFLP (see Exercise 12.5) for different values of $F = F_j$, $\forall j$, until a value F^* is found where the UFLP opens exactly p facilities. Show that this method does not work by giving an instance of a 2-Median problem for which no value of F provides an optimal solution to UFCLP with two open facilities.

Part IV

VEHICLE ROUTING MODELS