Drexl, Andreas; Klose, Andreas

Working Paper
Facility location models for distribution system design

Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 546

Provided in Cooperation with:
Christian-Albrechts-University of Kiel, Institute of Business Administration

Suggested Citation: Drexl, Andreas; Klose, Andreas (2001) : Facility location models for distribution system design, Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 546, Universität Kiel, Institut für Betriebswirtschaftslehre, Kiel

This Version is available at:
http://hdl.handle.net/10419/149077

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes. You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Manuskripte
aus den
Instituten für Betriebswirtschaftslehre
der Universität Kiel

No. 546

Facility Location Models
for Distribution System Design

Andreas Drexl, Andreas Klose
Abstract

The design of the distribution system is a strategic issue for almost every company. The problem of locating facilities and allocating customers covers the core topics of distribution system design. Model formulations and solution algorithms which address the issue vary widely in terms of fundamental assumptions, mathematical complexity and computational performance. This paper reviews some of the contributions to the current state-of-the-art. In particular, continuous location models, network location models, mixed-integer programming models, and applications are summarized.

Keywords: Strategic planning; distribution system design; facility location; mixed-integer programming models

1 Introduction

Decisions about the distribution system are a strategic issue for almost every company. The problem of locating facilities and allocating customers covers the core components of distribution system design. Industrial firms must locate fabrication and assembly plants as well as warehouses. Stores have to be located by retail outlets. The ability to manufacture and market its products is dependent in part on the location of the facilities. Similarly, government agencies have to decide about the location of offices, schools, hospitals, fire stations, etc. In every case, the quality of the services depends on the location of the facilities in relation to other facilities.

The problem of locating facilities is not new to the operations research community; the challenge of where to best site facilities has inspired a rich, colorful and ever growing body of literature. To cope with the multitude of applications encountered in the business world and in the public sector, an ever expanding family of models has emerged. Location-allocation models cover formulations which range in complexity from simple linear, single-stage, single-product, uncapacitated, deterministic models to nonlinear probabilistic models. Algorithms include, among others, local search and mathematical programming-based approaches.

It is the purpose of this paper to review some of the work which has contributed to the current state-of-the-art. The focus is on the fundamental assumptions, mathematical models and specific references to solution approaches. For the sake of brevity, work which has been done using simulation is neglected; see, for instance, Conners et al. (1972).

The outline of the work is as follows: In Section 2 types of models are classified. Section 3 reviews continuous location models. Then, Section 4 is dedicated to network location models, while Section 5 provides mixed-integer programming models. Finally, Section 6 covers a variety of applications.

2 Types of Models

Facility location models can be broadly classified as follows:

1. The shape or topography of the set of potential plants yields models in the plane, network location models, and discrete location or mixed-integer programming models, respectively. For each of the subclasses distances are calculated using some metric.

2. Objectives may be either of the minsum or the minmax type. Minsum models are designed to minimize average distances while minmax models have to minimize maximum distances. Predominantly, minsum models embrace location problems of private companies while minmax models focus on location problems arising in the public sector.

3. Models without capacity constraints do not restrict demand allocation. If capacity constraints for the potential sites have to be obeyed demand has to be allocated carefully. In the latter case we have to examine whether single-sourcing or multiple-sourcing is essential.
4. Single-stage models focus on distribution systems covering only one stage explicitly. In multi-stage models the flow of goods comprising several hierarchical stages has to be examined.

5. Single-product models are characterized by the fact that demand, cost and capacity for several products can be aggregated to a single homogeneous product. If products are inhomogeneous their effect on the design of the distribution system has to be analyzed, viz. multi-product models have to be studied.

6. Frequently, location models base on the assumption that demand is inelastic, that is, demand is independent of spatial decisions. If demand is elastic the relationship between, e.g., distance and demand has to be taken into account explicitly. In the latter case cost minimization has to be replaced through, for example, revenue maximization.

7. Static models try to optimize system performance for one representative period. By contrast dynamic models reflect data (cost, demand, capacities, etc.) varying over time within a given planning horizon.

8. In practice model input is usually not known with certainty. Data are based on forecasts and, hence, are likely to be uncertain. As a consequence, we have either deterministic models if input is (assumed to be) known with certainty or probabilistic models if input is subject to uncertainty.

9. In classical models the quality of demand allocation is measured on Isolation for each pair of supply and demand points. Unfortunately, if demand is satisfied through delivery tours then, for instance, delivery cost cannot be calculated for each pair of supply and demand points separately. Combined location/routing models elaborate on this interrelationship.

Additional attributes such as single- vs. multiple objective models or desirable vs. undesirable facilities may be distinguished; see, for instance, Aikens (1985), Brandeau and Chiu (1989), Daskin (1995), and Revelle and Laporte (1996).

3 Continuous Location Models

Continuous location models (models in the plane) are characterized through two essential attributes: (a) The solution space is continuous, that is, it is feasible to locate facilities on every point in the plane. (b) Distance is measured with a suitable metric. Typically, the Manhattan or right-angle distance metric, the Euclidean or straight-line distance metric, or the $l_p$-distance metric is employed.

Continuous location models require to calculate coordinates $(x, y) \in \mathbb{R}^p \times \mathbb{R}^p$ for $p$ facilities. The objective is to minimize the sum of distances between the facilities and $m$ given demand points.

The subject of the Weber problem is to determine the coordinates $(x, y) \in \mathbb{R} \times \mathbb{R}$ of a single facility such that the sum of the (weighted) distances $w_k d_k(x, y)$ to given demand points $k \in K$ located in $(a_k, b_k)$ is minimized. The corresponding optimization problem

$$\nu(\text{SWP}) = \min_{(x, y)} \sum_{k \in K} w_k d_k(x, y) \quad \text{where} \quad d_k(x, y) = \sqrt{(x - a_k)^2 + (y - b_k)^2}$$

can be solved efficiently by means of an iterative procedure. This gradient-like search method was originally proposed by Weiszfeld (1937) and has been further improved by Miehle (1958). This simple problem has a century-long tradition for the case of $|K| = 3$ demand points and it
has been included in the famous book of Weber (1909) giving the problem its nowadays name. The history of the Weber problem is well documented in Wesolowsky (1993).

An extended version of the problem requires to locate \( p \), \( 1 < p < |K| \) facilities and to allocate demand to the chosen facilities. This problem, also denoted as multisource Weber problem (MWP), is NP-hard. It can be modelled as the nonlinear mixed-integer program

\[
\nu (\text{MWP}) = \min \sum_{k \in K} \sum_{j=1}^{p} (w_k d_k(x_j,y_j))z_{kj}
\]

s.t.: \[
\begin{align*}
\sum_{j=1}^{p} z_{kj} &= 1 \quad \forall k \in K \\
z_{kj} &\in \mathbb{B} \quad \forall k \in K, \ j = 1, \ldots, p \\
x, y &\in \mathbb{R}^p
\end{align*}
\]

where \( \mathbb{B} = \{0,1\} \) and \( z_{kj} \) equals 1 if demand point \( k \) is assigned to facility \( j \). Exact solution procedures reformulate the model as a set partitioning problem, the LP-relaxation of which can be solved by column generation; see Rösing (1992b) and du Merle et al. (1999). Fast heuristic algorithms have been proposed by Taillard (1996), Hansen et al. (1998) and Brimberg et al. (2000). The special case of \( p = 2 \) facilities has been analyzed by Ostresh (1973), Drezner (1984), Rösing (1992b) and Chen et al. (1998).

A couple of variants and extensions of continuous location problems have been investigated in literature. To mention a few: Problems with barriers are the subject of, e.g., Hamacher and Nickel (1994) and Käfer and Nickel (2001). The location of undesirable (obnoxious) facilities requires to maximize minimum distances; see, e.g., Melachrinoudis (1988), Erkut and Neuman (1989) and Brimberg and Mehrez (1994). Location models with both desirable and undesirable facilities have been analysed in, for instance, Chen et al. (1992). Minmax location models have been dealt with, among others, by Krarup and Pruzan (1979), Love et al. (1988, pp. 113 ff.) and Francis et al. (1992, pp. 217 ff.).

4 Network Location Models

In network location models distances are computed as shortest paths in a graph. Nodes represent demand points and potential facility sites correspond to a subset of the nodes and to points on arcs.

The network location model corresponding to the continuous multisource Weber model is called \( p \)-median problem. In the \( p \)-median problem \( p \) facilities have to be located on a graph such that the sum of distances between the nodes of the graph and the facility located nearest is minimized. Hakimi (1964, 1965) has shown that it is sufficient to restrict the set of potential sites to the set of nodes in the case of concave distance functions.

Let \( K \) denote the set of nodes, \( J \subseteq K \) the set of potential facilities, \( w_k d_{kj} \) the weighted distance between nodes \( k \) and \( j \), \( y_j \) a binary decision variable being equal to 1 if node \( j \) is chosen as a facility (0, otherwise), and \( x_{kj} \) a binary decision variable reflecting the assignment of demand node \( k \in K \) to the potential facility site \( j \). Then

\[
u (\text{PMP}) = \min \sum_{k \in K} \sum_{j \in J} (w_k d_{kj})z_{kj} \quad (1a)
\]

s.t.: \[
\begin{align*}
\sum_{j \in J} z_{kj} &= 1 \quad \forall k \in K \\
z_{kj} - y_j &\leq 0 \quad \forall k \in K, \ j \in J
\end{align*}
\]

where \( B = \{0,1\} \) and \( z_{kj} \) equals 1 if demand point \( k \) is assigned to facility \( j \). Exact solution procedures reformulate the model as a set partitioning problem, the LP-relaxation of which can be solved by column generation; see Rösing (1992b) and du Merle et al. (1999). Fast heuristic algorithms have been proposed by Taillard (1996), Hansen et al. (1998) and Brimberg et al. (2000). The special case of \( p = 2 \) facilities has been analyzed by Ostresh (1973), Drezner (1984), Rösing (1992b) and Chen et al. (1998).

A couple of variants and extensions of continuous location problems have been investigated in literature. To mention a few: Problems with barriers are the subject of, e.g., Hamacher and Nickel (1994) and Käfer and Nickel (2001). The location of undesirable (obnoxious) facilities requires to maximize minimum distances; see, e.g., Melachrinoudis (1988), Erkut and Neuman (1989) and Brimberg and Mehrez (1994). Location models with both desirable and undesirable facilities have been analysed in, for instance, Chen et al. (1992). Minmax location models have been dealt with, among others, by Krarup and Pruzan (1979), Love et al. (1988, pp. 113 ff.) and Francis et al. (1992, pp. 217 ff.).

4 Network Location Models

In network location models distances are computed as shortest paths in a graph. Nodes represent demand points and potential facility sites correspond to a subset of the nodes and to points on arcs.

The network location model corresponding to the continuous multisource Weber model is called \( p \)-median problem. In the \( p \)-median problem \( p \) facilities have to be located on a graph such that the sum of distances between the nodes of the graph and the facility located nearest is minimized. Hakimi (1964, 1965) has shown that it is sufficient to restrict the set of potential sites to the set of nodes in the case of concave distance functions.

Let \( K \) denote the set of nodes, \( J \subseteq K \) the set of potential facilities, \( w_k d_{kj} \) the weighted distance between nodes \( k \) and \( j \), \( y_j \) a binary decision variable being equal to 1 if node \( j \) is chosen as a facility (0, otherwise), and \( x_{kj} \) a binary decision variable reflecting the assignment of demand node \( k \in K \) to the potential facility site \( j \). Then

\[
u (\text{PMP}) = \min \sum_{k \in K} \sum_{j \in J} (w_k d_{kj})z_{kj} \quad (1a)
\]

s.t.: \[
\begin{align*}
\sum_{j \in J} z_{kj} &= 1 \quad \forall k \in K \\
z_{kj} - y_j &\leq 0 \quad \forall k \in K, \ j \in J
\end{align*}
\]
formally describes the p-median problem. Constraints (1b) guarantee that demand is satisfied, inequalities (1c) couple the location and the assignment decision, and constraint (1d) fixes the number of selected facilities to p. Solution methods for the p-median problem have been presented by, e.g., Christofides and Beasley (1982), Hanjoul and Peeters (1985), Beasley (1993) and Klose (1993).

Let us now consider the p-center problem the aim of which is to locate p facilities such that the maximum distance is minimized. Unfortunately, for the p-center problem we cannot restrict the set of potential facility sites to the set of nodes because the maximum of concave distance functions is no concave function any more. Fortunately, it suffices to consider a finite set of points on the arcs. These points can be determined as intersection points q for which the weighted distance \( w_{d_iq} \) between q and node i \( \in K \) equals the weighted distance \( w_{dkq} \) between q and another node k \( \in K \). Let J denote the set of intersection points. Then the discrete optimization model

\[
\nu(\text{PCP}) = \min r
\]
\[\text{s.t.: } r - \sum_{j \in J} w_{kdkj} z_{kj} \geq 0 \quad \forall k \in K
\]
\[
\sum_{j \in J} z_{kj} = 1 \quad \forall k \in K
\]
\[
z_{kj} - y_j \leq 0 \quad \forall k \in K, \ j \in J
\]
\[
\sum_{j \in J} y_j = p
\]
\[
z_{kj}, y_j \in \mathbb{B} \quad \forall j \in J
\]

formally describes the p-center problem which can be transformed into a sequence of covering problems; see, e.g., Handler (1979) and Domschke and Drexl (1996). We start with a given set \( S \subseteq J, |S| \leq p \), of centers with radius \( r = \max_{k \in K} \min_{j \in S} \{w_{kd_kj}\} \). Then the covering model

\[
\nu(\text{SCP}) = \min \sum_{j \in J} y_j
\]
\[\text{s.t.: } \sum_{j \in J} a_{kj} y_j \geq 1 \quad \forall k \in K
\]
\[
y_j \in \mathbb{B} \quad \forall j \in J
\]

with \( a_{kj} = 1 \) for \( w_{kd_kj} < r \) and \( a_{kj} = 0 \) for \( w_{kd_kj} \geq r \) computes a set of at most p centers with a radius smaller than r or shows that no such set exists.

Recently, Boland et al. (2001) considered the so-called “discrete ordered median problem” which contains, among others, p-median and p-center problems as special cases.

The models treated so far assume given demand and cost minimization as objective. On the contrary competitive facility location models aim at maximum sales or market shares. One of the first papers is due to Hotelling (1929). A survey and a classification can be found in Eiselt et al. (1993); see Dobson and Karmarkar (1987) and Bauer et al. (1993) also.

Given an undirected graph with arc and node weights two basic models can be described as follows. The nodes \( k \in K \) of the graph represent the customers with known demand \( b_k \) for a certain product. Two companies A and B producing that product compete for customers.
Company A (B) wants to locate r (p) facilities in order to satisfy customers. Originally none of the companies is present in the market. At first company A determines locations of r facilities, then company B does so for p facilities. Customers always choose the nearest facility; in case of ties demand is divided between A and B.

Let $A_r$ ($B_p$) denote the set of facilities of A (B). Furthermore, let $\nu(B_p|A_r)$ denote the market share which can be achieved by company B choosing $B_p$, given $A_r$, then both companies have to solve two different problems; see Hakimi (1983):

Given $A_r$, company B determines the set $B^*_p$ such that $\nu(B^*_p|A_r) = \max_{B_p} \{\nu(B_p|A_r)\}$. If $B_p$ can be chosen among all points of the graph we have to solve a (p|Ar)-medianoid problem, if this choice is restricted to the set of nodes it is called maximum capture problem. Models and methods for both cases can be found in ReVelle (1986).

Company A determines, given $B_p$, the set $A^*_r$ such that $\nu(A^*_r|B_p) = \max_{A_r} \{\nu(A_r|B_p)\}$. $A_r$ either can be chosen among all points of the graph or is restricted to the set of nodes. The problem at hand is called (p|Ar)-centroid problem. The reasoning is as follows: When A chooses his facilities no other facilities do already exist. A locates his facilities in such a way that the market share gained subsequently by B is minimized, i.e., A anticipates the reaction of his competitor. In fact A has to solve a minmax-problem, that is, he minimizes the maximum market share which can be gained subsequently by B. (p|Ar)-medianoid problems and (p|Ar)-centroid problems are NP-hard if r and p are not fixed in advance. Given p the (p|Ar)-medianoid problem can be solved in polynomial time if the choice is restricted to the node set; see Benati and Laporte (1994).

5 Mixed-Integer Programming Models

Starting with a given set of potential facility sites many location problems can be modelled as mixed-integer programming models. Apparently, network location models differ only gradually from mixed-integer programming models because the former ones can be stated as discrete optimization models. Yet network location models explicitly take the structure of the set of potential facilities and the distance metric into account while mixed-integer programming models just use input parameters without asking where they come from.

A rough classification of discrete facility location models can be given as follows: (a) single-vs. multi-stage models, (b) uncapacitated vs. capacitated models, (c) multiple- vs. single-sourcing, (d) single- vs. multi-product models, (e) static vs. dynamic models, and, last but not least, (f) models without and with routing options included.

5.1 Uncapacitated, Single-Stage Models

The most simple model of this category solely considers the tradeoff between fixed operating and variable delivery cost. Mathematically,\[
\nu(UFLP) = \min \sum_{k \in K} \sum_{j \in J} c_{kj} x_{kj} + \sum_{j \in J} f_j y_j \] (4a)
\[
s.t.: \sum_{j \in J} x_{kj} = 1 \quad \forall \, k \in K \] (4b)
\[
x_{kj} - y_j \leq 0 \quad \forall \, k \in K, \, j \in J \] (4c)
\[
0 \leq x_{kj} \leq 1 \quad \forall \, k \in K, \, j \in J \] (4d)
\[
y_j \in \mathbb{B} \quad \forall \, j \in J \] (4e)
describes the simple plant location problem (SPLP) or uncapacitated facility location problem (UFLP).

The UFLP can be formulated more compact by aggregating constraints (4c) to $\sum_{k \in K} x_{kj} \leq |K| y_j$. The LP-relaxation of this “weak” model can be solved analytically; see Efroymson and Ray (1966) and Khumawala (1972). Unfortunately, the lower bounds are very weak. Cornuejols and Thizy (1982) have shown that the restrictions (4b) and (4c) cover all clique cuts of the UFLP which accounts for the fact that the model (4) yields a tight LP-relaxation. Branch-and-bound algorithms for the UFLP based on dual ascent methods have been proposed by Erlenkotter (1978) and Körkel (1989). Guignard (1988) considers the addition of Benders’ inequalities within a Lagrangean ascent method for the UFLP.

Obviously, the p-median problem (1) and the UFLP are close to each other. While the number of facilities is fixed in the p-median problem, the number of open depots is part of the UFLP solution. Both models can be combined if cardinality constraints

$$pl \leq \sum_{j \in J} y_j \leq pu \tag{5}$$

are added to (4). Usually the outcome is called account location problem or generalized p-median problem. The aggregate capacity constraint

$$\sum_{j \in J} s_j y_j \geq d(K) \tag{6}$$

where $s_j > 0$ denotes the maximum capacity of depot $j$ and $d(K) = \sum_{k \in K} d_k$ total demand, ensures that facilities open in a feasible solution have enough capacity in order to satisfy demand. Adding constraint (6) to the UFLP

$$\nu(\text{APLP}) = \min \left\{ \sum_{k \in K} \sum_{j \in J} c_{kj} x_{kj} + \sum_{j \in J} f_j y_j : (4b)-(4e), (6) \right\} \tag{7}$$

yields the aggregate capacity plant location problem (APLP). Exact algorithms for solving the APLP have been developed by Ryu and Guignard (1992b), Thizy (1994) and Klose (1998). The APLP is not important as a stand-alone model but it has a dominant role as a relaxation when solving models presented in Section 5.2.

The UFLP is closely related to covering problems; see Balas and Padberg (1976). Formally the set covering problem (SCP) computes a minimal collection $\{M_j : j \in S\}$ of a family $\{M_j : j \in N\}$ of subsets of a set $M$ such that $\bigcup_{j \in S} M_j = M$ holds. Letting $a_{kj} = 1$ for $k \in M_j$ and $a_{kj} = 0$ for $k \notin M_j$ translates it into model (3). The SCP is closely related to the set partitioning problem (SPaP)

$$\nu(\text{SPaP}) = \min \sum_{j \in J} y_j \tag{8a}$$

s.t.: $\sum_{j \in J} a_{kj} y_j = 1 \quad \forall k \in K \tag{8b}$

$$y_j \in \mathbb{B} \quad \forall j \in J \tag{8c}$$

and to the set packing problem (SPP):

$$\nu(\text{SPP}) = \max \sum_{j \in J} y_j \tag{9a}$$

s.t.: $\sum_{j \in J} a_{kj} y_j \leq 1 \quad \forall k \in K \tag{9b}$
The covering model (3) itself is a location model: An optimal solution of (3) determines a minimal subset \( S = \{ j \in J : y_j = 1 \} \) of facilities such that every customer can be reached within a given maximal distance from one of the chosen depots. An important variant of (3) is denoted as the maximum covering location problem (MCLP):

\[
\nu(\text{MCLP}) = \max \sum_{k \in K} w_k z_k
\]

s.t.:

\[
\sum_{j \in J} a_{kj} y_j - z_k \geq 0 \quad \forall k \in K
\]

\[
\sum_{j \in J} y_j = p
\]

\[
z_k, y_j \in \mathbb{B} \quad \forall k \in K, j \in J
\]

The MCLP requires to calculate a subset \( S = \{ j \in J : y_j = 1 \} \) of facilities with cardinality \( p \) such that a maximum number of \( w_k \) weighted demand nodes \( k \in K \) can be covered through facilities \( j \in S \) within a given maximal distance; see, e.g., Schilling et al. (1993), Daskin (1995) and Galvaö (1996). Defining the parameters

\[
c_{kj} = \begin{cases} 
0 & \text{für } a_{kj} = 1 \\
\infty & \text{für } a_{kj} = 0
\end{cases} \quad \text{and} \quad f_j = 1 \forall j \in J
\]

states the SCP (3) as an UFLP. Additionally, because of

\[
\nu(\text{MCLP}) = \max \left\{ \sum_{k \in K} \sum_{j \in J} a_{kj} w_k z_{kj} : (1b)-(1e) \right\}
\]

\[
= \sum_{k \in K} w_k - \min \left\{ \sum_{k \in K} \sum_{j \in J} (1 - a_{kj}) w_k z_{kj} : (1b)-(1e) \right\}
\]

the MCLP (10) is equivalent to the p-median problem with the special “distance” measure \( d_{kj} = (1 - a_{kj}) w_k \). On the contrary substituting variables \( y_j \) in the UFLP through their complement \( y_j^c = 1 - y_j \) and adding slack variables \( r_{kj} \) in (4c) leads to the special SPaP

\[
\min \sum_{k \in K} \sum_{j \in J} c_{kj} z_{kj} - \sum_{j \in J} f_j y_j^c + \sum_{j \in J} f_j
\]

s.t.:

\[
\sum_{j \in J} z_{kj} = 1 \quad \forall k \in K
\]

\[
z_{kj} + y_j^c + r_{kj} = 1 \quad \forall k \in K, j \in J
\]

\[
z_{kj}, r_{kj}, y_j^c \in \mathbb{B} \quad \forall k \in K, j \in J
\]

which in turn can be transformed into the SPP

\[
\max \sum_{k \in K} \sum_{j \in J} (L_k - c_{kj}) z_{kj} + \sum_{j \in J} f_j y_j^c - \sum_{k \in K} L_k - \sum_{j \in J} f_j
\]

s.t.:

\[
\sum_{j \in J} z_{kj} \leq 1 \quad \forall k \in K
\]

\[
z_{kj} + y_j^c \leq 1 \quad \forall k \in K, j \in J
\]

\[
z_{kj}, y_j^c \in \mathbb{B} \quad \forall k \in K, j \in J
\]
by replacing min through max and penalizing the slack \( \sum_j z_{kj} - 1 \) with a sufficiently large \( L_k \). Guignard (1980), Cho et al. (1983) and Cornuejols and Thizy (1982) capitalize on the transformation from UFLP to SPAp and SPP in order to study the polyedral structure of the UFLP. These relationships date back to Krarup and Puzan (1983).

The UFLP can be transformed into the SCP as follows: Replace in (8) the restrictions \( \sum_j a_{kj}y_j = 1 \) through inequalities \( \sum_j a_{kj}y_j \geq 1 \) and gather the slack variables \( \sum_j a_{kj}y_j - 1 \) with sufficiently large penalties \( L_k \) in the objective function.

5.2 Capacitated, Single-Stage Models

If depots have scarce capacity, constraints

\[
\sum_{k \in K} d_kz_{kj} \leq s_jy_j \quad \forall j \in J
\]  

limiting transshipments \( \sum_k d_kz_{kj} \) for the depots selected \( (y_j = 1) \) to their capacity \( s_j \) have to be added. Hence, in the case of scarce capacity the UFLP mutates to the capacitated facility location problem (CFLP).

The extended formulation

\[
\nu(\text{CFLP}) = \min \sum_{k \in K} \sum_{j \in J} c_{kj}z_{kj} + \sum_{j \in J} f_jy_j
\]

s.t.

- \( \sum_{j \in J} z_{kj} = 1 \quad \forall k \in K \) (D)
- \( \sum_{k \in K} d_kz_{kj} - s_jy_j \leq 0 \quad \forall j \in J \) (C)
- \( z_{kj} - y_j \leq 0 \quad \forall k \in K, \forall j \in J \) (B)
- \( \sum_{j \in J} s_jy_j \geq d(K) \) (T)
- \( \sum_{j \in J_q} z_{kj} \leq 1 \quad \forall k \in K, \forall q \in Q \) (U)
- \( 0 \leq z_{kj} \leq 1, 0 \leq y_j \leq 1 \quad \forall k \in K, \forall j \in J \) (N)
- \( y_j \in \{0, 1\} \quad \forall j \in J \) (I)

of the CFLP is a nice starting point in order to study various relaxations. A common way to obtain lower bounds for the CFLP is to relax constraints (C) and/or (D) in a Lagrangean manner and to add some additional inequalities which are implied by the relaxed constraints and some of the other constraints. The valid inequalities which are usually considered for these purposes are the variable upper bound or trivial clique constraints (B) and the aggregate capacity constraint (T). Besides the two additional constraints (B) and (T), one may devise a number of valid inequalities which can be useful to sharpen a relaxation, provided that the resulting subproblem is manageable. One group of redundant constraints is easily constructed as follows. Let \( \{J_q : q \in Q\}, J_q \cap J_h = \emptyset, q \neq h \), denote a given partitioning of the set \( J \) of potential plant locations. Then the "clique constraints" (U) are implied by (D); however, they can be useful if constraints (D) are relaxed.

Without taking constraints (U) into account, Cornuejols et al. (1991) examine all possible ways of applying Lagrangean relaxation/decomposition to the CFLP. Following their notation, let

- \( Z_K^S \) denote the resulting lower bound if constraint set \( S \) is ignored and constraints \( R \) are relaxed in a Lagrangean fashion, and let

8
ZR\textsubscript{1}/R\textsubscript{2} denote the bound which results if Lagrangean decomposition is applied in such a way that constraints R\textsubscript{i} and \textit{i}\textsuperscript{2} are split into two subproblems.

Regarding Lagrangean relaxation, Cornuejols et al. (1991, Theorem 1) show that

\[ Z^{BiU} \leq Z^{IU} \leq Z^{tU}_{C} \leq Z^{U}_C \leq Z_I, \ Z^{IU} \leq Z_D^{U} \leq Z^{U}_C, \text{ and } Z^{BiU} \leq Z^{BU}_C \leq Z^{U}_D. \]

Furthermore, they provide instances showing that all the inequalities above can be strict. The subproblem corresponding to \( Z^{U}_D \) can be converted to a knapsack problem and is solvable in pseudo-polynomial time. Therefore, bounds inferior to \( Z^{U}_D \) seem not to be interesting. Furthermore, as computational experiments show, \( Z^{tU}_C = Z^{I}_C \) is usually not stronger than \( Z^{U}_D \). This leaves \( Z^{U}_D \) and \( Z^{U}_C = Z_C \) as candidate bounds. Since constraints (U) are implied by (D), constraints (U) can only be helpful if constraints (D) are relaxed. If the aggregate capacity constraint (T) is relaxed as well, the resulting Lagrangean subproblem decomposes into \(|Q| \) smaller CFLPs. Obviously,

\[
Z^{T}_D = \begin{cases} 
Z^{TU}_D = Z^{TU}_{DU} = Z^{IU} = Z^I, & \text{if } |Q| = |J| \\
Z & \text{if } |Q| = 1.
\end{cases}
\]

For \( 1 < |Q| < |J| \), however, the bound \( Z^{T}_D \) can be anywhere between the (strong) LP-bound \( Z^{IU} = Z^I \) and the optimum value \( Z \) of the CFLP, i.e. \( Z^I \leq Z^{TU}_D \leq Z \). Although the subproblem corresponding to \( Z^{TU}_D \) has the same structure as the CFLP, the bound \( Z^{T}_D \) may be advantageous, if the set of potential plant locations is large and if the capacity constraints are not very tight.

With respect to Lagrangean decomposition, Cornuejols et al. (1991, Theorem 2) proof that

\[
Z^{U}_{C/ID} = Z^{U}_{C/DB} = Z^{U}_{C/DBT} = Z^{U}_{C/DBT} = Z^{U}_C, \text{ max}\{Z^{TU}_C, Z^{I}_D\} \leq Z^{U}_{D/TC} \leq Z^{U}_C, \text{ and } Z^{U}_{D/BC} = Z^{U}_{D/TBC} = Z^{TU}_{D/BC} = Z^{U}_D.
\]

Since Lagrangean decomposition requires to solve two subproblems in each iteration and to optimize a large number of multipliers, Lagrangean decomposition should give a bound which is at least as strong as \( Z^{U}_D \). The only remaining interesting bound is, therefore, \( Z^{U}_{D/TC} \). As shown by Chen and Guignard (1998), the bound \( Z^{U}_{D/TC} \) is also obtainable by means of a technique called Lagrangean substitution, which substitutes the copy constraints \( x = x' \) by \( \sum_k d_k x_{kj} = \sum_k d_k x'_{kj} \). Compared to the Lagrangean decomposition, this reduces the number of dual variables from \(|K| \cdot |J| + |J| \) to \( 2|J| \).

In summary, interesting Lagrangean bounds for the CFLP are \( Z^{U}_D, Z^{U}_C, Z^{U}_{D/TC} \) and \( Z^{T}_D \). Compared to \( Z^{U}_C \), the computation of the bound \( Z^{U}_{D/TC} \) requires to optimize an increased number of dual variables. Furthermore, one of the subproblems corresponding to \( Z^{U}_{D/TC} \) is an UFLP while the subproblem corresponding to \( Z^{U}_C \) is an APLP. Since the bound \( Z^{U}_{D/TC} \) is no stronger than \( Z^{U}_C \) and since an APLP is often not much harder to solve than an UFLP, the bound \( Z^{U}_{D/TC} \) can be discarded. The computation of these bounds by means of column generation is described in detail in Klose and Drexl (2001).

In the CFLP demand \( d_k \) can be supplied from more than one depot. Given a certain set of depots the CFLP reduces to a simple transportation problem. Apparently, this implies transportation cost being proportional to shipment volumes. In many practical settings this assumption does not hold and, moreover, it is required that each customer is satisfied from exactly one depot. In this case additional constraints

\[
z_{kj} \in \mathbb{B} \quad \forall \ k \in K, \ j \in J
\]

yield a pure integer program, well-known as capacitated facility location problem with single sourcing (CFLPSS). Unfortunately, single sourcing constraints make the problem much harder.
to solve. For a given set $O$ of open depots an optimal solution of the NP-hard generalized assignment problem (GAP)

$$\nu(GAP) = \min \sum_{k \in K} \sum_{j \in O} c_{kj} z_{kj}$$

s.t.: $\sum_{j \in O} z_{kj} = 1 \quad \forall k \in K$ (13b)

$$\sum_{k \in K} d_k z_{kj} \leq s_j \quad \forall j \in O$$

$z_{kj} \in \mathbb{B} \quad \forall k \in K, j \in O$ (13d)

provides a minimum cost assignment of customers to depots. Note that the GAP usually is formulated in such a way that capacity requirements depend on the assignments also; see Martello and Toth (1990, pp. 189ff.).

Without surprise it is very difficult to calculate an exact solution for instances of realistic size. From an algorithmic point of view, both for the CFLP and the CFLPSS, Lagrangian relaxation (dual decomposition) plays a dominant role; see Geoffrion and McBride (1978), Nauss (1978), Christofides and Beasley (1983), Guignard and Kim (1983), Barcelo and Casanovas (1984), Klinewicz and Luss (1986), Beasley (1988), Shetty (1990), Barcelo et al. (1990), Cornuejols et al. (1991), Ryu and Guignard (1992a), Beasley (1993), Sridharan (1993), Sridharan (1995), Holmberg et al. (1999) and Díaz and Fernandez (2001); additionally, primal and primal-dual decomposition algorithms have been developed; see Van Roy (1986), Wentges (1994) and Wentges (1996).

5.3 Multi-Stage Models

Consider a distribution system consisting of facilities on several hierarchically layered levels. Facility locations on a higher level can be determined independently of the chosen locations on a lower level if the following conditions are met: Higher level nodes have a sufficiently high capacity and handling costs as well as transshipment costs associated with these nodes are proportional to the amount of items reloaded and shipped, respectively. Transshipment cost from the source to the depot then can be charged proportional to the cost of allocated demand. Otherwise transshipments covering several stages of the distribution system have to be considered explicitly. Clearly, multi-stage facility location problems are present if depots have to be located simultaneously on several layers of the distribution system.

The CFLP and the CFLPSS have to be generalized to a two-stage capacitated facility location model if the flow of products from a capacity-constrained predecessor stage (e.g., production facility, central distribution facility) to the potential depots is an additional decision variable. Let $x_{ij}$ denote the amount which has to be shipped from predecessor node $i \in I$ having capacity $p_i$ to a depot located in node $j$. Furthermore, let $t_{ij}$ denote the transshipment cost per unit (containing the handling cost at node $i$ also) then the following two-stage capacitated facility location problem (TSCFLP) arises:

$$\nu(TSCFLP) = \min \sum_{i \in I} \sum_{j \in J} t_{ij} x_{ij} + \sum_{k \in K} \sum_{j \in J} c_{kj} z_{kj} + \sum_{j \in J} f_j u_j$$

s.t.: (4b)-(4e), (6), (11)

$$\sum_{j \in J} x_{ij} \leq p_i \quad \forall i \in I$$ (14b)

$$\sum_{i \in I} x_{ij} = \sum_{k \in K} d_k z_{kj} \quad \forall j \in J$$ (14c)
\[ x_{ij} - p_i y_j \leq 0 \quad \forall i \in I, j \in J \]  
\[ x_{ij} \geq 0 \quad \forall i \in I, j \in J \]  
\[ (14d) \]
\[ (14e) \]

If single-sourcing of demand nodes is required constraints (12) have to be added. Constraints (14b) take care of limited capacities at higher level nodes while restrictions (14c) are flow conservation constraints. (14d) are redundant but useful in order to tighten some relaxations. If the capacity \( p_i \) of each node \( i \in I \) is sufficiently large in order to cover the total demand \( d(K) \) per period then the TSCFLP reduces to the CFLP or the CFLPSS, respectively.

If we introduce variables \( w_{ijk} \) which denote the fraction of demand \( d_k \) being routed via path \( i \to j \to k \) then an alternative TSCFLP model can be stated as follows for the case of single sourcing:

\[
\nu'(TSCFLP) = \min \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} q_{ijk} w_{ijk} \quad \forall j \in J, k \in K \\
s.t.: (4b)-(4e), (6), (11), (12) \\
\sum_{i \in I} w_{ijk} = z_{kj} \quad \forall j \in J, k \in K \\
\sum_{j \in J} d_k w_{ijk} \leq p_i \quad \forall i \in I \\
\sum_{k \in K} d_k w_{ijk} \leq p_i z_{kj} \quad \forall i \in I, \ j \in J \\
w_{ijk} \geq 0 \quad \forall i \in I, \ j \in J, \ k \in K 
\]

In this model \( q_{ijk} = t_{ij} d_k + c_{kj} \) defines the procurement cost of node \( k \in K \) via path \( i \to j \to k \). Note that this formulation allows to model situations where the cost depend on both the source node \( i \) and the sink node \( k \). Such cases occur in practice if for instance freight rates from source \( i \) to depot \( j \) are less than the sum of freight rates from \( i \) to \( j \) plus \( j \) to \( k \). Apparently, the first model is advantageous if the cost \( q_{ijk} \) can be split into two parts \( t_{ij} d_k \) and \( c_{kj} \), because it has far fewer decision variables while the values of the LP-relaxations of both models are identical.

If demand splitting is allowed then the variables \( z_{kj} \) can be eliminated in the second model. Accordingly, the demand constraint can be rewritten as \( \sum_{i \in I} \sum_{j \in J} w_{ijk} = 1 \) and the depot capacity restrictions as \( \sum_{i \in I} \sum_{k \in K} d_k w_{ijk} \leq s_j y_j \).

In general, models where facilities on several stages of a distribution system have to be located are called multi-level hierarchical facility location problems. In contrast to the TSCFLP the tightness of relaxations depends on whether variables for single links or variables covering whole paths of the network are used. If \( g_i \) denotes the fixed cost of facility \( i \in I \) on the highest level and \( \gamma_i \) corresponding decision variables then the TSCFLP generalizes to the following two-level capacitated facility location problem (TLCFLP).

\[
\nu'(TLCFLP) = \min \sum_{i \in I} \sum_{j \in J} t_{ij} x_{ij} + \sum_{k \in K} \sum_{j \in J} c_{kj} w_{kj} + \sum_{i \in I} g_i \gamma_i + \sum_{j \in J} f_j y_j \\
s.t.: (4b)-(4e), (6), (11), (14c), (14e) \\
\sum_{j \in J} x_{ij} \leq p_i \gamma_i \quad \forall i \in I \\
x_{ij} - \min\{p_i, s_j\} \gamma_i \leq 0 \quad \forall i \in I, \ j \in J \\
\sum_{i \in I} p_i \gamma_i \geq d(K) \\
\gamma_i \in \mathbb{B} \quad \forall i \in I 
\]

11
Analogously, the UFLP generalizes to the two-level uncapacitated facility location problem (TUFLP). An equivalent formulation of the TLCFLP based on path variables \( w_{ijk} \) can be given as follows:

\[
\nu(\text{TLCFLP}) = \min \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} q_{ijk} w_{ijk} + \sum_{i \in I} g_i \gamma_i + \sum_{j \in J} f_j y_j
\]

s.t.: \[ \sum_{i \in I} \sum_{j \in J} w_{ijk} = 1 \quad \forall k \in K \] (15a)

\[ \sum_{j \in J} \sum_{k \in K} d_k w_{ijk} \leq p_i \gamma_i \quad \forall i \in I \] (15b)

\[ \sum_{i \in I} \sum_{j \in J} d_k w_{ijk} \leq s_j y_j \quad \forall j \in J \] (15c)

\[ \sum_{j \in J} w_{ijk} \leq \gamma_i \quad \forall i \in I, k \in K \] (15d)

\[ \sum_{i \in I} w_{ijk} \leq y_j \quad \forall j \in J, k \in K \] (15e)

\[ \sum_{j \in J} s_j y_j \geq d(K) \] (15f)

\[ \sum_{i \in I} p_i \gamma_i \geq d(K) \] (15g)

\[ w_{ijk} \geq 0 \quad \forall i \in I, j \in J, k \in K \] (15h)

\[ \gamma_i, y_j \in \mathbb{B} \quad \forall i \in I, j \in J \] (15i)

Constraints (15b) guarantee that demand is satisfied completely. Constraints (15c) and (15d) take care of scarce capacities of facilities on both levels. Aggregate capacity constraints (15g) and (15h) are redundant but probably useful in order to tighten relaxations. The left hand side of (15f) corresponds to the variable \( z_{kj} \) in the former TLCFLP model and, hence, (15f) is equivalent to (4c). However, the left hand side of (15e) covers the fraction of demand \( d_k \) being shipped to \( k \in K \) indirectly from \( i \in I \). Note that this term cannot be incorporated in the former model because the flows on the two stages are modelled independently.


### 5.4 Multi-Product Models

The models discussed so far are based on aggregated demand, production, handling as well as distribution cost. Furthermore, capacity of production, depot and transshipment nodes must be given uniquely for all the products. Such an aggregation is no more valid if different products
make different claims on the capacities of some nodes of the network. In this case we must proceed to a multi-product model, where, for instance, the capacities of nodes, the demand as well as the flows are separated with respect to some homogeneous product groups. Such multi-product variants of the TSCFLP and TLCFLP have been presented in Geoffrion and Graves (1974), Hindi and Basta (1994), Hindi et al. (1998) and Pirkul and Jayaraman (1996, 1998).

Other types of multi-product models arise, e.g., if (a) different types of facilities have to be distinguished at some locations and/or if (b) fixed cost of locations depend on the product provided by a location. The first type is called “multi-type model” by Karkazis and Boffey (1981), Boffey and Karkazis (1984), Mirchandani et al. (1985), Lee (1996) and Mazzola and Neebe (1999) and the second is called “multi-activity model” by Klincewicz et al. (1986), Barros and Labbé (1992), Gao and Robinson, Jr. (1992, 1994) and Barros (1998). Both can be modelled uniquely if different types of facilities correspond to different products. Fixed cost depend on products if a specific infrastructure or equipment is required in order to provide a product or service at a specific location.

Let $I$ denote the set of product families $i \in I$ and (in addition to the fixed cost $f_j$) $g_{ij}$ the fixed product cost. Then an uncapacitated multi-activity model, also called multi-commodity or multi-activity uncapacitated facility location problem (MUFLP) can be given as follows:

$$
\nu(\text{MUFLP}) = \min \left\{ \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} q_{ijk} w_{ijk} + \sum_{i \in I} \sum_{j \in J} g_{ij} z_{ij} + \sum_{j \in J} f_j y_j \right\} \\
\text{s.t.:} \sum_{j \in J} w_{ijk} = 1 \quad \forall i \in I, k \in K
$$

$$(16a)$$

$$(16b)$$

$$(16c)$$

$$(16d)$$

$$(16e)$$

$$(16f)$$

Here, $z_{ij}$ is a binary variable which equals 1 if product/service type $i$ is provided at depot $j$. The variable $w_{ijk}$ denotes the fraction of demand $d_{ik}$ of demand node $k$ for product $i$ which is covered by depot $j$. Likewise $g_{ijk}$ denotes the cost of providing $d_{ik}$ units of product $i$ from depot $j$ to demand node $k \in K$. Constraints (16b) require that the demand of each customer is covered. The coupling constraints (16c) and (16d) forbid to assign products to closed depots and to deliver product $i$ to node $k$ from depot $j$ if product $i$ is unavailable at the depot. In the multi-type case each facility can provide one product or service and, hence, the constraints

$$
\sum_{i \in I} z_{ij} \leq 1 \quad \forall j \in J
$$

have to be added.

The model MUFLP adds product depot allocation decisions to the UFLP. Gao and Robinson, Jr. (1992, 1994) show that the MUFLP is a special two-stage hierarchical facility location model. To this end products $i \in I$ have to be viewed as locations on the higher level and customer-product pairs $(k, i)$ as single customers $k' \in K' = K \times I$ which have to be satisfied from combined locations $(i, j)$. Allocation cost $g_{ijk'}$ are prohibitive large if $k'$ does not correspond to product type $i$. The resulting formulation

$$
\nu(\text{MUFLP}) = \min \left\{ \sum_{i \in I} \sum_{j \in J} \sum_{k' \in K'} q_{ijk'} w_{ijk'} + \sum_{i \in I} \sum_{j \in J} g_{ij} z_{ij} + \sum_{j \in J} f_j y_j \right\} \\
\text{s.t.:} \sum_{i \in I} \sum_{j \in J} w_{ijk'} = 1 \quad \forall k' \in K'
$$

13
is a two-stage hierarchical facility location model in which the fixed cost $g_{ij}$ of lower level location $j \in J$ are determined completely through the assignment to a location $i \in I$ at the higher level.

5.5 Dynamic Models

In general, decisions about facility locations are made on a long-term basis. Depots, distribution centers and transshipment points once established shall be used for a couple of periods. However, factors influencing such decisions vary over time. In particular, demand (volume, regional distribution) and cost structures may change, but relocation and/or redimensioning of facilities can be quite costly. In order to cope with such issues dynamic location and allocation models have been developed. Dynamic location models are provided, for instance, by Schilling (1980), Erlenkotter (1981), Van Roy and Erlenkotter (1982), Frantzeskakis and Watson-Gandy (1989) and Shulman (1991).

In a dynamic version of the UFLP for every depot a close or open option is available in every period $t = 1, \ldots, T$ where $T$ denotes a given planning horizon. Fixed cost $g_{ij}^0$ and $g_{ij}^o$ for closing and opening depots are added to the fixed depot operating cost $f_{ij}$ for relocation purposes. Closing cost $g_{ij}^0$ have to be paid if depot $j \in J$ which is open in period $t - 1$, that is, $y_{t-1,j} = 1$, is closed in period $t$, i.e. $y_{t,j} = 0$; on the contrary opening cost $g_{ij}^o$ result if a depot which is closed in period $t - 1$, that is, $y_{t-1,j} = 0$, is openend in period $t$, i.e. $y_{t,j} = 1$. The following quadratic integer program is a dynamic version DUFLP of UFLP.

\[
\nu(\text{DUFLP}) = \min \sum_{t=1}^{T} \sum_{k \in K} \sum_{j \in J} c_{tkj} z_{tkj} + \sum_{t=1}^{T} \sum_{j \in J} f_{ij} y_{t,j} + \sum_{t=1}^{T} \left( \sum_{j \in J} g_{ij}^0 y_{t-1,j} (1 - y_{t,j}) + \sum_{j \in J} g_{ij}^o (1 - y_{t-1,j}) y_{t,j} \right) \\
\text{s.t.:} \sum_{j \in J} z_{tkj} = 1 \quad \forall k \in K, \ t = 1, \ldots, T \\
z_{tkj} - y_{t,j} \leq 0 \quad \forall k \in K, \forall j \in J, \ t = 1, \ldots, T \\
z_{tkj}, y_{t,j} \in \mathbb{B} \quad \forall k \in K, \forall j \in J, \ t = 1, \ldots, T
\]

DUFLP can be linearized by introducing the binary variables $u_{t-1,t,j} \equiv y_{t-1,j} y_{t,j}$ and the additional constraints

\[
u_{t-1,t,j} \leq y_{t-1,j}, \quad u_{t-1,t,j} \leq y_{t,j}, \quad u_{t-1,t,j} \geq y_{t-1,j} + y_{t,j} - 1.
\]

Supplementary constraints

\[
y_{t+\tau,j} \geq y_{t,j} \text{ for } \tau = 1, \ldots, \tau_1 \text{ or } y_{t+\tau_0,j} \leq y_{t,j} \text{ for } \tau = 1, \ldots, \tau_0
\]

achieve that the status of a depot opened (closed) in period $t$ remains open (close) for at least $\tau_1$ ($\tau_0$) periods. The dynamic UFLP variant of Van Roy and Erlenkotter (1982) further boils down opening/closing options. Closing a depot $j \in J_1$ being originally open is feasible in one period $t$ only. Similarly, opening a depot $j \in J_0$ being originally closed is feasible in one period $t$ only. Let the binary variable $y_{t,j}$ equal 1 (0) if a depot $j \in J_1$ ($j \in J_0$) is closed (opened). Furthermore,
let $F_{tj}$ denote the discounted cash flow of the period fixed cost for the periods $1, \ldots, t$ for $j \in J_1$ and for the periods $t, \ldots, T$ for $j \in J_0$. Then we get the following "simplified" dynamic version of the UFLP:

$$\begin{align*}
\text{min} & \sum_{t=1}^{T} \left( \sum_{k \in K} \sum_{j \in J} c_{tkj} z_{tkj} + \sum_{j \in J} F_{tj} y_{tj} \right) \\
\text{s.t.} & \sum_{j \in J} z_{tkj} = 1 \quad \forall k \in K, t = 1, \ldots, T \\
 z_{tkj} & \leq \sum_{\tau=1}^{t} y_{\tau j} \quad \forall k \in K, j \in J_0, t = 1, \ldots, T \\
 z_{tkj} & \leq \sum_{\tau=t}^{T} y_{\tau j} \quad \forall k \in K, j \in J_1, t = 1, \ldots, T \\
 z_{tkj}, y_{tj} & \in \mathbb{B} \quad \forall k \in K, j \in N, t = 1, \ldots, T
\end{align*}$$

Constraints (17c) achieve that demand nodes $k \in K$ can be assigned to locations $j \in J_0$ in period $t$ if the depot has been opened in period $\tau \leq t$. Likewise constraints (17d) prevent that demand nodes $k \in K$ are assigned to locations $j \in J_1$ in periods $t \geq \tau$ if the depot will be closed in period $\tau$. Note that demand allocation can be changed in every period while opening/closing of a depot is possible only once.

Dynamic models seem to be adequate in light of factors changing over time. However, their practical relevance seems to be limited. First, a "right" planning horizon does not exist. Second, the amount of data required is enormous. Third, "disaggregated" models are more sensitive to parameter/data adjustments than aggregated ones. Fourth, the complexity of dynamic models increases – compared with static models – dramatically and, hence, the chances to solve such models decrease.

### 5.6 Probabilistic Models

In practice some of the input data of location models are subject to uncertainty. Berman and Larson (1985), for instance, analyze queuing location models. Given certain distribution functions for the customer arrival process, waiting and service times are approximated. The waiting times are a function of the demand allocation and, hence, of facility location.

A stochastic variant of the $p$-median problem is discussed in Mirchandani et al. (1985). In particular, the input data demand and arc weights are supposed to be random variables. Under certain assumptions a finite number of states $i \in I$ of the graph with known probabilities can be enumerated. The objective of the model (18) is to minimize the expected sum of the weighted distances.

$$\begin{align*}
\text{min} & \sum_{i \in I} \sum_{k \in K} \sum_{j \in J} \pi_i c_{ikj} z_{ikj} \\
\text{s.t.} & \sum_{j \in J} z_{ikj} = 1 \quad \forall i \in I, j \in J \\
z_{ikj} - y_{j} & \leq 0 \quad \forall i \in I, k \in K, j \in J \\
\sum_{j \in J} y_{j} & = p \\
z_{ikj}, y_{j} & \in \mathbb{B} \quad \forall i \in I, k \in K, j \in J
\end{align*}$$

The symbol $c_{ikj}$ denotes the demand weighted distance between nodes $k \in K$ and $j \in J$ in state $i \in I$. The decision variables $z_{ikj}$ take care of the demand allocation in state $i \in I$, the variables
model the location decisions. The stochastic p-median model (18) can easily be reduced to (1) by replacing the variables \( z_{ikj} \) through variables \( z^l = k + |I|(i - 1) \), denoting the assignment of demand node \( k \in K \) in state \( i \in I \) with corresponding allocation cost \( c_{ij} = \pi_i c_{ikj} \). Similarly, stochastic variants of the UFLP and CFLP can be considered, but in the case of the CFLP capacity constraints prevent the reduction to a deterministic CFLP with an increased number of demand nodes. Further stochastic location models are, e.g., discussed in Laporte et al. (1994) and Listes and Dekker (2001). Laporte et al. (1994) develop a branch-and-cut algorithm for a location problem with stochastic demand; Listes and Dekker (2001) use stochastic models with recourse for the purposes of locating facilities in product recovery networks.

Unfortunately, stochastic models require a large amount of data in order to adapt empirically observed distributions to theoretical ones. Usually for strategic facility location problems such information is not available. Probably, calculating solutions, supported by sensitivity analysis, for some scenarios is useful. To gain insight into the effects of parameter changes is important. Furthermore, scenario analysis can be employed. This approach tries to find solutions which perform best over a set of scenarios with respect to some kind of regret measure; see, e.g., Owen and Daskin (1998) and Barros et al. (1998).

5.7 Hub Location Models

Recently, hub location models have received considerable attention. Usually, they are studied on hub-and-spoke networks with the following properties: For an undirected graph with node set \( K \) a flow exists between every pair \( i, j \in K \) of nodes. A subset of “central” nodes act as transshipment nodes (hubs); the other (terminal or non-hub) nodes are connected with an arc (spoke) starlike with one of the hubs. Flows from one node to another node travel directly if both nodes are hub nodes or if one node is a hub node and both are connected through a spoke. Otherwise, flow travels via at least another hub node.

Similar to p-median and facility location models the number of hubs can be fixed or subject to decision, capacities can be scarce or non-scarce. Additionally, the hub nodes can constitute a complete graph, a tree or a graph without special characteristics. The non-hub (terminal) nodes are linked via arcs with hub nodes (and probably with other non-hub nodes also). If each terminal node has to be connected with exactly one hub node the single allocation case is given. Otherwise, if terminal nodes have access to more than one hub the multiple allocation hub location problem arises.

In what follows we consider a multiple allocation hub location problem in more detail. Assume an undirected, complete graph with \( n = |K| \) nodes, arc set \( E \) and arc weights \( c \). For each pair of nodes \( i \) and \( j \) the volume of traffic (flow) equals \( v_{ij} \). Each node \( k \in H \) of a subset \( H \subseteq K \) of nodes can be chosen as hub node. The fixed cost of locating a hub in node \( k \in H \) equal \( f_k \).

The hub network is a complete subgraph. Flows between terminal nodes travel via at most two hubs, flows between terminal nodes are infeasible. \( c_{kj} \) denotes the arc weights (cost per unit). If one unit travels from terminal node \( i \) via hub nodes \( k \) and \( m \) to terminal node \( j \) then the cost are \( c_{ikmj} = c_{ik} + \alpha \cdot c_{km} + c_{mj} \); \( \alpha \) is scaling factor, \( 0 < \alpha \leq 1 \). Let \( x_{ikmj} \) denote the fraction of flow \( v_{ij} \) that travels via hub nodes \( k \) and \( m \). Furthermore, \( y_k \) is a binary decision variable for the selection of hubs. Then

\[
\nu(UHLP) = \min \sum_{i \in K} \sum_{k \in H} \sum_{m \in H} \sum_{j \in K} v_{ij} c_{ikmj} x_{ikmj} + \sum_{k \in K} f_k y_k 
\]

(19a)

\[
\text{s.t.: } \sum_{k \in H} \sum_{m \in H} x_{ikmj} = 1 \quad \forall i, j \in K 
\]

(19b)

\[
x_{ikmj} \leq y_k \quad \forall i, j \in K, k, m \in H 
\]

(19c)

\[
x_{ikmj} \leq y_m \quad \forall i, j \in K, k, m \in H 
\]

(19d)
formally describes the uncapacitated hub location problem (UHLP). Apparently, if the set of hubs is known a shortest path problem remains to be solved. Otherwise, the problem is NP-hard. Algorithms for solving the uncapacitated hub location problem have been developed, among others, by Klincewicz (1996), Ernst and Krishnamoorthy (1998), Hamacher et al. (2000) and Meyer and Wagner (2000). The capacitated case is studied by, e.g., Aykin (1994) and Ebery et al. (2000). A survey is given by Campbell (1994b). The problem of locating such interacting hub facilities arises in many applications some of which include airlines, see Campbell (1992), Aykin (1995b), the Civil Aeronautics Board, see O’Kelly (1986, 1987), emergency services, see Campbell (1994a) and postal delivery services, see Ernst and Krishnamoorthy (1996).

5.8 Routing Location Models

The application of the location models discussed so far requires that the cost \( C_{kj} \) for allocating the demand \( d_k \) of a customer \( k \in K \) to a depot can be allocated independently of the allocation of other demand points. A very complex form of service cost depending on each other arises if customers are satisfied within routes covering several customers simultaneously. In this case location and routing decisions are strongly interrelated. Unfortunately, the formulation and solution of routing location models is extremely complicated because of several reasons. First, optimization problems become very complicated. Second, the planning horizons inherent in both subproblems are different. Third, facility location requires to aggregate customers while routing does not. Moreover, besides the variety of facility location models there do exist many different routing models as well (for a survey see Fisher (1995) and Crainic and Laporte (1998)). Hence, a huge number of combined models is possible; to mention a few: Determine an optimal location for a traveling salesman; see Laporte et al. (1983), Simchi-Levi and Berman (1988), Branco and Coelho (1990). Combine an UFLP with a matching approach; see Gourdin et al. (2000). Integrate multi-stage facility location, multi-depot vehicle routing and scheduling and fleet mix models; see Jacobsen and Madsen (1980), Perl and Daskin (1985), Bookbinder and Reece (1988), Laporte et al. (1988), Nagy and Salhi (1996), Salhi and Fraser (1996), Bruns and Klose (1996), Bruns (1998). An in-depth discussion of combined routing location models can be found in Klose (2001). Aykin (1995a) studies hub location and routing problems.

6 Applications

Applications of facility location models are not restricted to the primary application area of this article, that is, the design of distribution systems. By contrast many other problems where location and allocation decisions are interdependent are covered also. For the sake of brevity some of them shall be sketched out as follows:

- Cluster analysis: The topic of cluster analysis is to group items in such a way that items belonging to one group are homogeneous and items belonging to different groups are heterogeneous. Location then means to select representative items from the overall set of items while allocation corresponds to the assignment of the remaining items to the chosen clusters. Mulvey and Crowder (1979) model the clustering task as a \( p \)-median problem. To the contrary, Rosing (1992a) uses a clustering algorithm in order to solve the MWP heuristically. Moreover, clustering is important in the problem setting of vehicle routing and scheduling, see Fisher and Jaikumar (1981) and Bramel and Simchi-Levi (1995), and in the area of combined routing location, see Klose and Wittmann (1995) and Klose (1996).

- Location of bank accounts: A company which has to pay suppliers has to decide which bank accounts to use for this purpose. Depending on the location of the used accounts
float can be optimized. Cornuejols et al. (1977) model this problem, the so-called account location problem, as an UFLP with the additional constraint (5). Nauss and Markland (1981) study the revers problem of locating bank accounts in order to receive customer payments, the so-called lock box location problem.

- Vendor selection: Each company must choose vendors for the supply of products. Vendor selection is based on multiple criteria such as price, quality, know-how, etc. Location in this setting means selecting some vendors from a given set of vendors. Allocation relates to the decision which product to buy from which vendor. Current and Weber (1994) discuss, among other topics, that this problem can be tackled using well-known location models such as the UFLP and the CFLP.

- Location and sizing of offshore platforms for oil exploration: Hansen et al. (1992, 1994) use a capacitated multi-type location model in order to locate offshore platforms for oil exploration. Different platform types relate to potential platform capacities.

- Database location in computer networks: Within a computer network databases can be installed on certain nodes. Installation and maintenance of databases gives raise to fixed cost while transmission times or cost may decrease, hence, once more, a certain location-allocation problem arises. Fisher and Hochbaum (1980) model this problem as an extended UFLP.

- Concentrator location: The design of efficient telecommunication and computer networks poses several complex, interdependent problems. Related surveys can be found in Boffey (1989), Gavish (1991) and Chardaire (1999). Starlike networks comprise a simple topology, connecting terminals with a central machine. Such a topology is inefficient in the case of many terminals and large distances. Probably, the installation of concentrators having powerful links to the central machine or another (backbone) network then is necessary. To determine the layout of a concentrator-based network results in a typical location-allocation problem, also called concentrator location problem by Mirzaian (1985) and Pirkul (1987). Chardaire (1999), Chardaire et al. (1999) and Chardaire et al. (1999) study the case where concentrators can be located on two different layers of the network.

- Index selection for database design: Databases comprise a set of tables, each of which consists of several arrays. Relating indices to arrays allows to store entries in a sorted manner yielding fast queries. Caprara and Salazar (1995, 1999) and Caprara et al. (1995) study the index selection problem as an important optimization problem in the physical design of relational databases. Moreover, it is shown that this problem can be formulated as an UFLP (4). Furthermore, efficient branch-and-bound and branch-and-cut algorithms are presented.

References


Bruns, A. D. and Klose, A. (1996). A locate first – route second heuristic for a combined...


