Logistics

LECTURE NOTES*

Maria Grazia Scutellà

DIPARTIMENTO DI INFORMATICA
UNIVERSITÀ DI PISA

September 2015

*These notes are related to the course of Logistics held by the author at the University of Pisa. They are based on some books from the literature, as acknowledged at the end of each chapter.

Chapter 4

Location models

Within the Operations Research and Management Science community there is a strong interest in location analysis and modeling due to:

1. location decisions are frequently made at all levels of human organizations;
2. location decisions are strategic: in fact, they involve large sums of capital resources, with long term economic effects, in the private and in the private sectors;
3. location decisions often determine economic externalities such as pollution, congestion. . . ;
4. location models are often very difficult to solve (at least to optimality), so great interest is aimed toward clever formulations and efficient implementations;
5. usually location models are application specific, i.e. their structural form depends on the specific application context (no general location model exists); see Table 3.1 in Z. Drezner and H. Hamacher (2004) for a list of possible applications; such a variety of applications has stimulated the interest in location modeling.

4.1 Basic facility location models

There exist eight basic location models, sharing some common characteristics:

1. the *underlying logistic network*: it is a directed graph whose nodes represent the locations of the clients (also denoted demands) to be served by the facilities and also the locations of the existing facilities (if any): the general location problem is to locate new facilities in the logistic network by optimizing a certain objective;
2. the *concept of distance* (or some measures related to distance such as travel time) is at the basis of the eight models.

   Depending on how "distance" is taken into account, in the literature the eight models are classified into:

- models based on maximum distance (four models);
- models based on total or average distance (four models).

4.2 Maximum distance models

In some location problems, a maximum distance exists "a priori". For example, elementary school students within a mile of their school must walk to school. As another example, some businesses guarantee a service within a pre-determined time, e.g. 20 minutes.

In the facility location terminology, a priori maximum distances are called *covering or coverage distances*: the demand of a client is considered fully satisfied if its nearest facility is within the coverage distance, and it is not satisfied (*uncovered*) if the closest facility is beyond the coverage distance. Please observe that being closer to a facility than the coverage distance does not improve the satisfaction level of the clients.

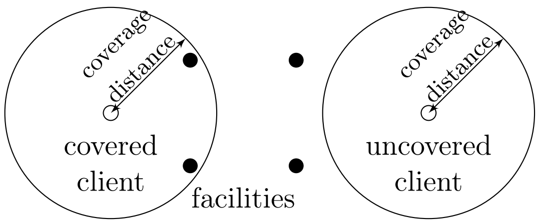

Figure 4.1: Covered vs. uncovered clients

4.2.1 Set covering location model

The Set Covering location problem is to locate the minimum number of facilities required to "cover" all clients, by considering the following input data:

- I = set of clients (or *demand nodes*), indexed by i;
- J = set of candidate *facility locations*, indexed by j;
- d_{ij} = distance between $i \in I$ and $j \in J$, $\forall i \in I, j \in J$;
- D_C = coverage distance;
- $N_i = \{j \in J : d_{ij} \leq D_C\}$, i.e. the set of all candidate facility locations that can cover i, $\forall i \in I$.

Example Consider Figure 4.2, where $I = \{1, 2, 3, 4\}$, $J = \{5, 6\}$, $D_C = 3$

Observe that to locate only in 5 (node 4 is not covered) or to locate only in 6 (1 and 2 are not covered) are not feasible solutions. Therefore we have to locate both in 5 and in 6.

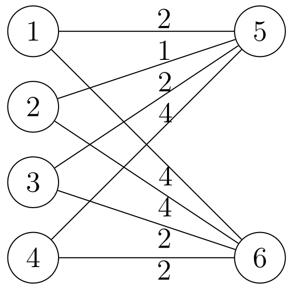

Figure 4.2: A Set Covering location example

ILP model

Define the decision variables as follows:

$$x_j = \begin{cases} 1 & \text{if we locate at site } j \\ 0 & \text{otherwise} \end{cases}, \ \forall j \in J.$$

Using these location variables the Set covering location problem can be stated as:

$$\begin{aligned} & \min \sum_{j \in J} x_j \\ & \sum_{j \in N_i} x_j \geq 1, \ \forall i \in I \quad \textit{covering constraints} \\ & x_j \in \{0,1\}, \ \forall j \in J \end{aligned} \tag{SCLP}$$

The objective function can be generalized by means of location costs:

$$\min \sum_{j \in J} c_j \, x_j.$$

The problem is *NP*-hard in both cases. However, the LP relaxation often returns an integer optimal solution.

Some possible row and column reduction rules exist, that may help in reducing the size of the model, so facilitating the model solution:

1. Let $M_j = \{i \in I : d_{ij} \leq D_C\}$ and $M_k = \{i \in I : d_{ik} \leq D_C\}$. If $M_k \subseteq M_j$, then variable x_k can be eliminated (a facility at j would cover all clients that a facility at k would cover). It can be said that "location j *dominates* k".
2. If $N_i \subset N_h$, then the covering constraint $\sum_{j \in N_h} x_j \geq 1$ can be eliminated, since it is redundant (by covering node i then the covering constraint for node h is automatically satisfied).

The (SCLP) model assumes that facilities can be located only at nodes of the logistics network (*Discrete facility location*). However, a lower cost could be achieved in some cases by locating also along the arcs. For example, refer to Figure 4.3. By allowing facility location along the arcs (Continuous facility location), we could locate only one facility (at •), whereas with (SCLP) we have to locate one facility at node A and one either at B or C. A possible (discrete) approach is to augment the network with a finite number of additional nodes along the arcs (i.e. enlarging set J).

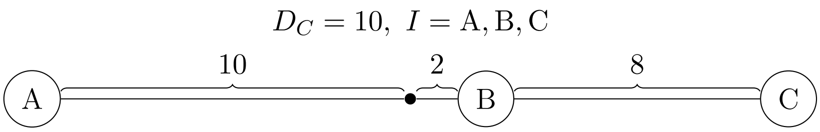

Figure 4.3: Example of continuous facility location

4.2.2 Maximal covering location model

In many facility planning situations (e.g. school districts) there is an upper limit on the number of facilities to be opened (budget constraint); therefore, not all demand nodes can be covered.

The problem is to locate a given number of facilities, say p, in such a way as to maximize the covered demand. We augment the definitions used in (SCLP) as follows:

- input data: h_i is the demand at node i, $\forall i \in I$, and p is the number of facilities to locate;
- decision variables:

$$z_i = \begin{cases} 1 & \text{if node } i \text{ is covered} \\ 0 & \text{otherwise} \end{cases}, \ \forall i \in I.$$

ILP model

$$\begin{aligned} & \max \sum_{i \in I} h_i \, z_i \\ & \sum_{j \in J} x_j = p \\ & \sum_{j \in N_i} x_j \geq z_i, \ \forall i \in I \\ & x_j \in \{0,1\}, \ \forall j \in J \\ & z_i \in \{0,1\}, \ \forall i \in I \end{aligned} \tag{MCLP}$$

The Maximal covering location problem is *NP*-hard. Please observe that, without loss of generality, we can replace the constraints $z_i \in \{0, 1\}$ with $z_i \leq 1, \ \forall i \in I$ (so sparing us some of the boolean variables).

A possible extension to the Maximal covering location problem is to add constraints about closeness between demand nodes and opened facilities. So denoting by D_m the maximum acceptable distance of any demand node from an opened facility, and defining $M_i = \{j \in J : d_{ij} \leq D_m\}$, we can state the Maximal covering location problem with mandatory closeness constraints as:

$$\begin{gathered} \text{(MCLP)} \\ + \\ \sum_{j \in M_i} x_j \geq 1, \ \forall i \in I \quad \textit{closeness contraints} \end{gathered} \tag{4.1}$$

4.2.3 p-center model

(SCLP) and (MCLP) assume that the covering distance, D_C, is a fixed standard. However, in many situations D_C is a target rather than a fixed standard. For example, in fire station planning we may want to minimize the maximum distance that a citizen has from such kind of facilities, for equity reasons.

The p-center problem (Hakimi, 1964–1965) is related to this issue. Given p facilities to be located, we want to minimize the maximum distance that a demand node is from its closest facility. Some variations of the basic model exist:

vertex p-center: facilities can be located only at the nodes;

absolute p-center: facilities can be located anywhere along the arcs;

unweighted: all demand nodes are treated equally;

weighted: distances d_{ij} are multiplied by a weight associated with demand node i.

Hereafter, we shall consider the weighted vertex p-center model. Define the following decision variables (in addition to x_{ij}):

$$y_{ij} = \begin{cases} 1 & \text{if demand node } i \text{ is assigned to a facility at } j \\ 0 & \text{otherwise} \end{cases}, \ \forall i \in I, j \in J.$$

Define also w as the maximum distance between a demand node and the facility to which it is assigned. Therefore we can state the following model:

$$
\begin{aligned}
&\min w \\
&\sum_{j\in J} x_j = p && && (1) \\
&\sum_{j\in J} y_{ij} = 1 && \forall i \in I && (2) \\
&y_{ij} \le x_j && \forall i \in I, j \in J && (3) \\
&\sum_{j\in J} h_i\, d_{ij}\, y_{ij} \le w && \forall i \in I && (4) \\
&x_j \in \{0,1\} && \forall j \in J \\
&y_{ij} \in \{0,1\} && \forall i \in I, j \in J
\end{aligned}
\qquad \text{(p-center)}
$$

The stated linear constraints have the following meaning:

① p facilities have to be located;

② each demand node must be assigned to exactly one facility (*assignment constraints*, stating the location–allocation rule);

③ nodes can be assigned only to open facilities (*linking between location and allocation*);

④ w is defined as an upper bound on the distance between each node and the facility to which it is assigned, and so on the "maximum" one; by minimizing w, we therefore minimize such a maximum distance.

The time complexity is $\mathcal{O}(N^p)$, with N number of candidate facility locations: in order to solve the problem, in the worst case, all sets of candidate locations must be evaluated. Therefore the problem is NP-hard for variable values of p.

Notice that (SCLP), (MCLP) and (p-center) are defined on a bipartite graph, which models one layer of the logistics network. See Figure 4.4 as an example.

4.2.4 Some examples

An instance of (SCLP)

Refer to the complete bipartite graph depicted in Figure 4.5. Let:

$I = \{1, 2, 3, 4\}, J = \{5, 6, 7\}, D_C = 20$ minutes, $d_{ij} =$

	5	**6**	**7**
1	20	25	60
2	30	15	30
3	30	15	30
4	30	15	10

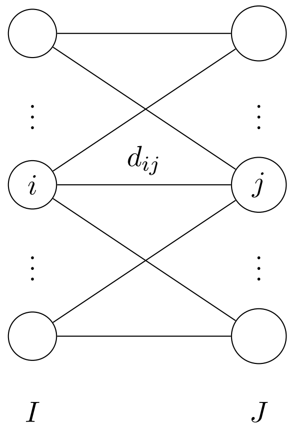

Figure 4.4: Bipartite graph in (SCLP), (MCLP) and (p-center) problems

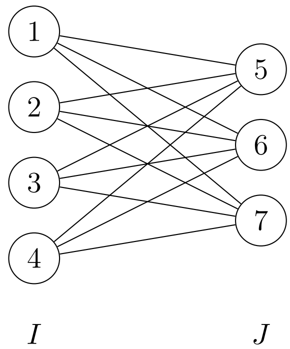

Figure 4.5: Bipartite graph for the (SCLP) example

where d_{ij} are expressed in minutes. Remember that $N_i = \{j \in J : d_{ij} \leq D_C\}$. So in this case it is $N_1 = \{5\}, N_2 = N_3 = \{6\}, N_4 = \{6, 7\}$.

The ILP model is therefore the one in Model 4.1.

The optimal solution can be easily found by inspection and is depicted in Figure 4.6: $x_5^* = 1, x_6^* = 1, x_7^* = 0$; the objective function value is 2. Please observe that facility 7 would be closer than 6 for node 4 (10 minutes instead of 15 minutes); however, in (SCLP) the target is just to guarantee a distance $\leq D_C = 20$ minutes.

Suppose now that, for budget reason, we can open only one facility ($p = 1$), and so we want to maximize the covered demand. Assume that $h_i = 1,\ i = 1, 2, 3, 4$. In such a case we get a Maximal covering location problem, and the model to be solved is the one in Model 4.2.

The new optimal solution, found by inspection, is depicted in Figure 4.7 and it is:

$$x_6^* = 1, x_5^* = 0, x_7^* = 0 \qquad z_1^* = 0, z_2^* = 1, z_3^* = 1, z_4^* = 1$$

Now, assume to have an additional information, i.e. the service is *not* acceptable beyond $D_m = 25$ minutes. Recall that $M_i = \{j \in J : d_{ij} \leq D_m\}$ and so in this example

$$
\begin{aligned}
\min x_5 + x_6 + x_7 & \\
x_5 \qquad\qquad & \geq 1 \quad \text{covering node 1} \\
x_6 \qquad & \geq 1 \quad \text{covering node 2} \\
x_6 \qquad & \geq 1 \quad \text{covering node 3 (redundant!)} \\
x_6 + x_7 & \geq 1 \quad \text{covering node 4} \\
x_5 \ , \ x_6 \ , \ x_7 & \in \{0, 1\}
\end{aligned}
$$

Model 4.1

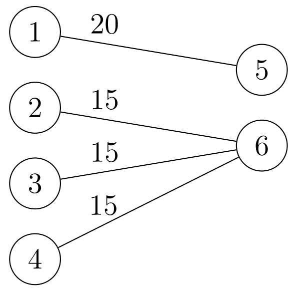

Figure 4.6: Optimal solution for the (SCLP) and p-center examples

$$
\begin{aligned}
\max z_1 + z_2 + z_3 + z_4 & \\
x_5 \qquad\qquad & \geq z_1 \\
x_6 \qquad & \geq z_2 \\
x_6 \qquad & \geq z_3 \\
x_6 + x_7 & \geq z_4 \\
x_5 + x_6 + x_7 & = 1 \\
x_5 \ , \ x_6 \ , \ x_7 & \in \{0, 1\} \\
z_1 \ , \ z_2 \ , \ z_3 \ , \ z_4 & \in \{0, 1\}
\end{aligned}
$$

Model 4.2

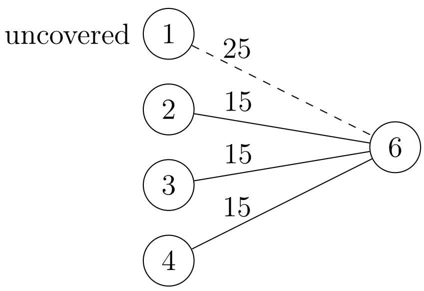

Figure 4.7: Optimal solution for the (MCLP) example

$M_1 = \{5,6\}, M_2 = M_3 = \{6\}, M_4 = \{6,7\}$. The problem is now an (MCLP) with mandatory closeness constraints, and is described in Model 4.3.

$$
\begin{aligned}
\max z_1 + z_2 + z_3 + z_4 & \\
x_5 &\geq z_1 \\
x_6 &\geq z_2 \\
x_6 &\geq z_3 \\
x_6 + x_7 &\geq z_4 \\
x_5 + x_6 + x_7 &= 1 \\
x_5 + x_6 &\geq 1 \\
x_6 &\geq 1 \\
x_6 + x_7 &\geq 1
\end{aligned}
$$

Model 4.3

The optimal solution is the same as before (one opened facility at site 6).

Consider now the same instance as before, with $h_i = 1, i = 1,2,3,4$, but change $p = 2$. We want to assign each demand node to one of the two opened facilities (to be decided), so as to minimize the maximum distance between demand node and facility. The problem has now become a p-center one and is described in Model 4.4.

$$
\begin{array}{llll}
\min w & & & \\
x_5 + & x_6 + & x_7 = 2 & \\
y_{15} + & y_{16} + & y_{17} = 1 & \\
y_{25} + & y_{26} + & y_{27} = 1 & \text{allocation} \\
y_{35} + & y_{36} + & y_{37} = 1 & \text{constraints} \\
y_{45} + & y_{46} + & y_{47} = 1 & \\
y_{15} & & \leq x_5 & \\
 & y_{25} & \leq x_5 & \text{linking constraints} \\
 & & \ddots \vdots & \\
\multicolumn{3}{r}{20y_{15} + 25y_{16} + 60y_{17} \leq w} & \text{maximum} \\
\multicolumn{3}{r}{30y_{25} + 15y_{26} + 30y_{27} \leq w} & \text{distance} \\
\multicolumn{3}{r}{30y_{35} + 15y_{36} + 30y_{37} \leq w} & \text{constraints} \\
\multicolumn{3}{r}{30y_{45} + 15y_{46} + 10y_{47} \leq w} & \\
x_5\ , & x_6\ , & x_7 \in \{0,1\} & \\
 & & y_{ij} \in \{0,1\}, & \forall i \in I, j \in J
\end{array}
$$

Model 4.4

The optimal solution, found by inspection, is the same found by solving Model 4.1 with

$D_C = 20$ minutes, and is depicted in Figure 4.6:

$$x_5^* = 1, x_6^* = 1, x_7^* = 0$$
$$y_{15}^* = 1, y_{26}^* = 1, y_{36}^* = 1, y_{46}^* = 1, y_{ij}^* = 0 \text{ otherwise}$$
$$w^* = 20 \text{ minutes}$$

The optimal solution for $p = 1$ would be to open a facility at site 6 ($w^* = 25$).

4.2.5 p-dispersion model

The previous models address the distance between demand nodes and new facilities; the assumption is that being close to a facility is desirable. The new problem differs from them in that it is concerned only with distances between facilities, since the objective is to maximize the minimum distance between any pair of facilities.

Possible applications of this kind of problem are siting of military installations (the separation makes them more difficult to attack) and outlet locations (separation may help storage management).

To express the objective function, we introduce a new decision variable D, that models the minimum separation distance between pairs of facilities.

Using D, the problem can be modelled as follows:

$$\begin{aligned} &\max D \\ &\sum_{j \in J} x_j = p \\ &D + (M - d_{ij})x_i + (M - d_{ij})x_j \leq 2M - d_{ij}, \;\; \forall i, j \in J : i < j \\ &x_j \in \{0, 1\}, \;\; \forall j \in J \end{aligned} \tag{p-dispersion}$$

where d_{ij} is the distance between i and j, and M denotes a very large value (often referred to as big M). For example we can set $M = \max_{i,j \in J} \{d_{ij}\}$. Let us explain the meaning of the complicating constraints:

1. if $x_i = x_j = 0$ (i.e. no facility in i and j), then the constraint related to i and j becomes $D \leq 2M - d_{ij}$, and therefore $D + d_{ij} \leq 2M$, which is always satisfied since both $D, d_{ij} \leq M$;
2. if $x_i = 1$ and $x_j = 0$ (i.e. only one facility at i), then $D + M \cancel{- d_{ij}} \leq 2M \cancel{- d_{ij}}$, i.e. $D \leq M$, which is always satisfied;
3. if $x_i = 0$ and $x_j = 1$, then the same explanation of case 2 applies;

4. if $x_i = x_j = 1$ (i.e. this is a pair of opened facilities), then the complicating constraint related to i and j becomes

$$D + \cancel{M} - \cancel{d_{ij}} + \cancel{M} - d_{ij} \leq \cancel{2M} - \cancel{d_{ij}} \iff D \leq d_{ij}.$$

In this last case, the constraint imposes that D be a lower bound on the distance between i and j; since there is a constraint of this type for any pair of opened facilities, D must be a lower bound on the smallest inter-facility distance: maximizing D has the effect of forcing this smallest distance to be as large as possible.

Please observe that the condition $i < j$ is added to take into account symmetries so as to avoid redundant constraints (e.g. the constraint for the pair $(3, 5)$ is equal to the one for $(5, 3)$; so, only one of the two constraints can be introduced within the model).

Example of p-dispersion $p = 4$

Consider the logistics network depicted in Figure 4.8.

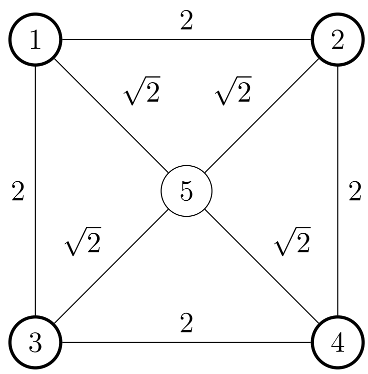

Figure 4.8: Example of p-dispersion

Note that the graph is not complete: for every $i, j \in J$, d_{ij} is the cost of the shortest path linking i and j; e.g. $d_{14} = 2\sqrt{2}$. The optimal solution is marked with thick nodes in the picture, and the optimum objective function value is $D^* = \min\{2, 2\sqrt{2}\} = 2$. The ILP model is presented in Model 4.5

Let's explicit some constraints for the optimal solution, $x_1^* = x_2^* = x_3^* = x_4^* = 1,\ x_5^* = 0$:

- pair $(1, 2) : D + (M - 2) + (M - 2) \leq 2M - 2 \iff D \leq 2 = d_{12}$;
- pair $(1, 3)$: analogous to the previous one: $D \leq 2 = d_{13}$;
- pair $(1, 4) : D + (M - 2\sqrt{2}) + (M - 2\sqrt{2}) \leq 2M - 2\sqrt{2} \iff D \leq 2\sqrt{2} = d_{14}$;
- pair $(1, 5)$: $D + (M - \cancel{\sqrt{2}}) + \cancel{(M - \sqrt{2})x_5^*} \leq 2M - \sqrt{2} \iff D \leq M$, which is always true if M is big.

$$
\begin{aligned}
&\max D \\
&x_1 + x_2 + x_3 + x_4 + x_5 = 4 \\
&D + (M-2)x_1 + (M-2)x_2 \leq 2M - 2 \quad \text{pair } (1,2) \\
&D + (M-2)x_1 + (M-2)x_3 \leq 2M - 2 \quad \text{pair } (1,3) \\
&D + (M-\underbrace{2\sqrt{2}}_{d_{14}})x_1 + (M-2\sqrt{2})x_4 \leq 2M - \underbrace{2\sqrt{2}}_{d_{14}} \quad \text{pair } (1,4) \\
&D + (M-\sqrt{2})x_1 + (M-\sqrt{2})x_5 \leq 2M - \sqrt{2} \quad \text{pair } (1,5) \\
&\qquad \text{(analogous constraints for the remaining pairs)} \\
&\qquad \vdots \\
&x_1, x_2, x_3, x_4, x_5 \in \{0,1\}
\end{aligned}
$$

Model 4.5

Therefore, $D \leq \min\{d_{12}, d_{13}, d_{14}, M\} = \min\{2, 2, 2\sqrt{2}, \overbrace{M}^{\text{redundant}}\} = 2$; so within the model D is a lower bound on the minimum inter-facility distances (clearly we have to consider all pairs).

4.3 Total or average distance models

Many facility location problems aim at minimizing the total travel distance between facilities and demand nodes: they privilege the “efficiency” objective as opposed to the “equity” objective, e.g. of minimizing the maximum distance.

4.3.1 p-median problem

This problem (Hakimi 1964–1965) is to determine the locations of p facilities (p is given) in order to minimize the total demand–weighted distance between the demand nodes and the facilities to which they are assigned.

The (location and allocation) variables as well as the constraints are common to the formulation of the (p-center) problem: only the objective function is different. The formulation also minimizes the average travel distance (it can be shown by dividing the objective function by $|I|$).

It can be formulated as follows:

$$
\begin{aligned}
&\min \sum_{i\in I}\sum_{j\in J} h_{ij}\, d_{ij}\, y_{ij} \\
&\sum_{j\in J} x_j = p \\
&\sum_{j\in J} y_{ij} = 1, \quad \forall i \in I \\
&y_{ij} \le x_j, \qquad \forall i \in I, j \in J \\
&x_j \in \{0,1\} \quad \forall j \in J \\
&y_{ij} \in \{0,1\} \quad \forall i \in I, j \in J
\end{aligned}
\tag{p-median}
$$

The problem is *NP*-hard for variable values of p.

An instance of p-median

Consider the graph depicted in Figure 4.9, and assume $p = 2$. Let $h_i = 1,\ i = 1,2,3,4,$ and

$d_{ij} =$

	5	**6**	**7**
1	20	25	60
2	30	15	10
3	30	15	10
4	30	15	10

The corresponding ILP model is presented in Model 4.6.

An optimal solution for this toy example can be found by inspection and is depicted in Figure 4.10:

$$
\begin{gathered}
x_5^* = 1, x_7^* = 1, x_6^* = 0 \\
y_{15}^* = 1, y_{27}^* = 1, y_{37}^* = 1, y_{47}^* = 1, y_{ij}^* = 0 \text{ otherwise} \\
\text{optimal value} = \text{minimum total distance} = 50
\end{gathered}
$$

Observe that this solution is also optimal for the p-center problem defined on the same instance: the minimum maximum distance is 20! However, (p-center) and (p-median) are *not* equivalent. Let us present a counterexample where an optimal solution to p-median is not necessarily optimal to p-center.

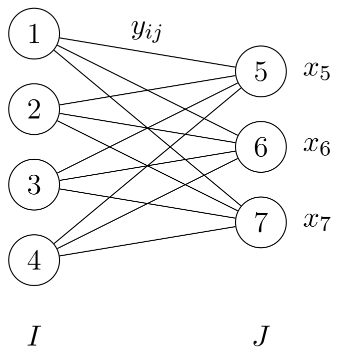

Figure 4.9: Bipartite graph for the p-median example

$$
\begin{aligned}
\min\ & 20y_{15} + 25y_{16} + 60y_{17} + 30y_{25} + \\
& + 15y_{26} + 10y_{27} + 30y_{35} + 15y_{36} + \\
& + 10y_{37} + 30y_{45} + 15y_{46} + 10y_{47} \\
& x_5 + \ x_6 + \ x_7 = 2 \\
& y_{15} + y_{16} + y_{17} = 1 \\
& y_{25} + y_{26} + y_{27} = 1 \qquad \text{allocation} \\
& y_{35} + y_{36} + y_{47} = 1 \qquad \text{constraints} \\
& y_{45} + y_{46} + y_{47} = 1 \\
& y_{15} \leq x_5 \\
& y_{16} \leq x_6 \qquad \text{linking between location} \\
& y_{17} \leq x_7 \qquad \text{and allocation} \\
& \text{(same for the other links)} \\
& \vdots \\
& x_5\ ,\ x_6\ ,\ x_7 \in \{0,1\} \\
& y_{ij} \in \{0,1\} \quad \forall i \in I, j \in J
\end{aligned}
$$

Model 4.6

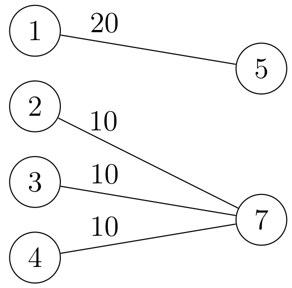

Figure 4.10: Optimal solution for the p-median example

Counterexample

Consider the graph in Figure 4.11, and assume $p = 1$. Let $h_i = 1,\ i = 1, 2, 3, 4$.

For this instance, the optimal solutions are:

- for p-median: locate at 5, with total distance 80 (note that the maximum distance is 50);
- for p-center: locate at 6, with maximum distance 30 (note that the total distance is 120).

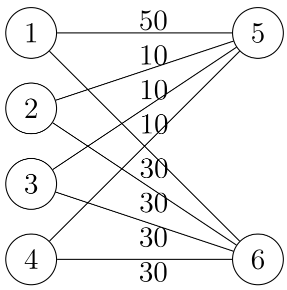

Figure 4.11

4.3.2 Fixed charge location problem

Location problems such as p-median make three assumptions:

1. each candidate site has the same fixed cost for locating a facility;
2. the facilities are *uncapacitated*, i.e. there is no upper limit on the demand they can serve;
3. the number of facilities to be opened, p, is known a priori.

These assumptions are not appropriate in some cases, so we may need to relax them.

The Fixed charge location problem is to determine the optimal number and locations of facilities, which are now capacitated, and determine the assignment of each demand node to an open facility, in order to minimize the total cost, i.e. fixed costs for opening facilities and transportation costs. Since the facilities have capacities, demand may not be assigned to its closest open facility (as in previous cases).

To state the problem let us introduce additional input data:

- f_j is the fixed cost of locating a facility at $j,\ \forall j \in J$;
- C_j is the capacity of a facility at $j,\ \forall j \in J$;
- α is the cost per unit demand per unit distance.

The ILP model is the following:

$$
\begin{aligned}
&\min \sum_{j\in J} f_j\, x_j + \alpha \sum_{i\in I}\sum_{j\in J} h_{ij}\, d_{ij}\, y_{ij} && & \\
&\sum_{j\in J} y_{ij} = 1 && \forall i \in I && \text{assignment constraints}\\
&y_{ij} \le x_j && \forall i \in I, j \in J && \text{linking between location–allocation constraints}\\
&\sum_{i\in I} h_i\, y_{ij} \le C_j\, x_j && \forall j \in J && \text{“new” constraints}\\
&x_j \in \{0,1\} && \forall j \in J &&\\
&y_{ij} \in \{0,1\} && \forall i \in I, j \in J &&
\end{aligned}
\tag{FCLP}
$$

The “new” constraints are highlighted within the model. The constraint related to j can be interpreted in the following way:

- if $x_j = 1$ (i.e. locate a facility at j) then $\sum_{i\in J} h_i\, y_{ij} \le C_j$, that is we put an upper bound on the total demand served by j;
- if $x_j = 0$ (i.e. no facility at j) then $\sum_{i\in I} h_i\, y_{ij} \le 0$, that is $y_{ij} = 0,\ \forall i \in I$, so no client is assigned to j. This was already fixed by the linking constraints between location and allocation, therefore $y_{ij} \le x_j,\ \forall i \in I, j \in J$ are redundant, and can be eliminated (although they enhance the LP relaxation).

Variants of the model

We present here two variants of the (FCLP) model:

1. by replacing (relaxing) $y_{ij} \in \{0,1\}$ with $0 \le y_{ij} \le 1,\ \forall i \in I, j \in J$, we allow node i to be served (partially) by multiple facilities (no more “single-source allocation”, as imposed by (FCLP));
2. if we remove the capacity constraints then we get the *Uncapacitated fixed charge location problem (UFCLP)*: in this case each client will be served by its closest (open) facility. Allocation then becomes an easy task.

An instance of (FCLP)

Consider now the same instance of (p-median), but with different values of h_i (see Figure 4.12). Recall that distances are:

$$d_{ij} = \begin{array}{c|ccc} & \mathbf{5} & \mathbf{6} & \mathbf{7} \\ \hline \mathbf{1} & 20 & 25 & 60 \\ \mathbf{2} & 30 & 15 & 10 \\ \mathbf{3} & 30 & 15 & 10 \\ \mathbf{4} & 30 & 15 & 10 \end{array}$$

Differently than (p-median), we have:

- facility capacities C_j (specified in Figure 4.12);
- fixed costs: $f_5 = 50, f_6 = f_7 = 100$;
- unit cost per demand per distance: $\alpha = 1$.

Observe that since the total demand is 15, we need to open at least two facilities (due to their capacities). Figure 4.13 presents an optimal solution for the example, found by inspection:

$$x_5^* = 1, x_6^* = 0, x_7^* = 1$$
$$y_{15}^* = 1, y_{25}^* = 1, y_{35}^* = 1, y_{47}^* = 1, y_{ij} = 0 \text{ otherwise.}$$

The minimum cost is $\overbrace{100 + 50}^{\text{fixed cost}} + 1 \cdot \overbrace{(40 + 150 + 90 + 50)}^{\text{transportation cost}} = 480$. Observe that nodes 2 and 3 are not served by their closest facility (as it would happen in p-median), due to capacity restrictions. The ILP model is presented in Model 4.7.

$$\begin{aligned} \min\ & 50x_5 + 100x_6 + 100x_7 + 2 \cdot 20y_{15} + 2 \cdot 25y_{16} + \dots \\ & y_{15} + y_{16} + y_{17} = 1 \\ & y_{25} + y_{26} + y_{27} = 1 \\ & y_{35} + y_{36} + y_{37} = 1 \\ & y_{45} + y_{46} + y_{47} = 1 \\ & 2y_{15} + 5y_{25} + 3y_{35} + 5y_{45} \leq 10x_5 \\ & 2y_{16} + 5y_{26} + 3y_{36} + 5y_{46} \leq 5x_6 \\ & 2y_{17} + 5y_{27} + 3y_{37} + 5y_{47} \leq 5x_7 \\ & x_5\ ,\ x_6\ ,\ x_7 \in \{0,1\} \\ & y_{ij} \in \{0,1\}\ \forall i \in I, j \in J \end{aligned}$$

Model 4.7

4.3.3 Hub location problems

Logistics systems such as the ones related to airline networks use hub and spoke systems, in order to exploit larger capacity or faster vehicles during the delivery from an origin

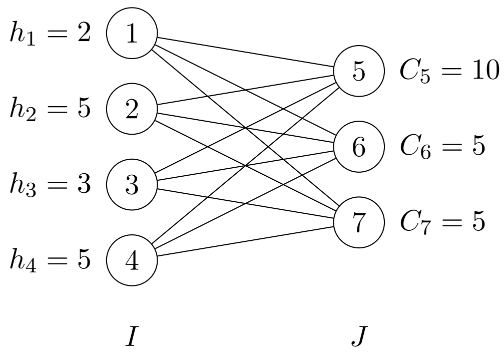

Figure 4.12: Bipartite graph for the FCLP example

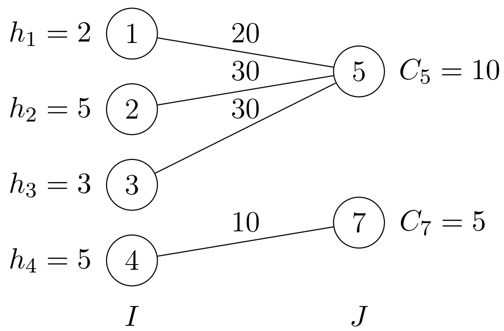

Figure 4.13: Optimal solution for the FCLP example

to a destination, so reducing the overall transportation cost. Figure 4.14 depicts an example of a delivery from an origin i to a destination j using two intermediate hubs. We call α the *discount factor* for transportation between hubs.

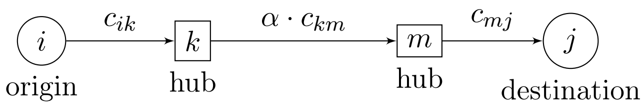

Figure 4.14: Example of a delivery using intermediate hubs

The basic p-hub location problem is to locate p hubs and assign each demand node to exactly one hub, so as to minimize the total transportation cost between the origin–destination pairs. As in the previous models, J denotes the set of the candidate facility locations, where hubs can be activated. On the other hand, the set of demand nodes I induces a set of origin–destination pairs (i, j), and we know the amount of flow h_{ij} that the origin i must send to j along the network (whereas the previous models assume to know the request of each demand nodes versus a facility, and not versus other demand nodes).

The input data consist of:

- h_{ij} are the units of product (flow) to be sent from i to j, $\forall i, j \in I$ (note that $h_{ij} = 0$ if no sending is required);

- c_{ij} is the unitary transportation cost from i to j, $\forall i, j \in J$;

- α is the discount factor for inter-hub sending.

Consider the following decision variables:

- $x_j = \begin{cases} 1 & \text{if we locate a hub at } j \\ 0 & \text{otherwise} \end{cases}, \ \forall j \in J;$

- $y_{ij} = \begin{cases} 1 & \text{if node } i \text{ is assigned to hub } j \\ 0 & \text{otherwise} \end{cases}, \ \forall i \in I, j \in J.$

The mathematical model is the following:

$$\min \sum_{i \in I} \sum_{j \in I} h_{ij} \left(\sum_{k \in J} c_{ik}\, y_{ik} + \sum_{m \in J} c_{mj}\, y_{jm} + \alpha \sum_{k \in J} \sum_{m \in J} c_{km} \overbrace{y_{ik}\, y_{jm}}^{\text{nonlinear}} \right)$$

$$\begin{array}{lll} \sum_{j \in J} x_j = p & & \text{exactly } p \text{ hubs} \\ \sum_{j \in J} y_{ij} = 1 & \forall i \in I & \text{allocation} \\ y_{ij} \leq x_j & \forall i \in I, j \in J & \text{linking between location–allocation} \\ x_j \in \{0,1\} & \forall j \in J & \\ y_{ij} \in \{0,1\} & \forall i \in I, j \in J & \end{array} \qquad \text{(p-hub)}$$

Two assumptions were made in formulating the problem: first, the hub portion of the logistics network is a complete graph, so the flow between any pair (i, j) will pass through at most two hubs. Second, each demand node is assigned to exactly one hub (we could relax $\sum_{j \in J} y_{ij} = 1,\ y_{ij} \in \{0,1\},\ \forall i \in I, j \in J$, to allow that nodes be served by more than one hub).

Some characteristics of the (p-hub) model are worth being pointed out:

- the model take into consideration node-to-node flows rather than simply demands at nodes (versus facilities);
- the model is not ILP since the objective function is nonlinear (it is quadratic);
- in the optimal solution, nodes are not necessarily assigned to the nearest hub, since the cost is measured in terms of origin–destination flows.

Example of p-hub location

Consider the network partially depicted in Figure 4.15 (the graph is in fact complete), where $I = \{1, 2, 3, 4\},\ J = \{5, 6, 7\},\ c_{ij} = 100, \forall i, j \in J, i \neq j,\ \alpha = \frac{1}{10}$.

Let $h_{13} = 5, h_{24} = 15, h_{ij} = 0$ otherwise. Furthermore, let $c_{ij} =$

	5	**6**	**7**
1	10	1000	10
2	10	1000	10
3	1000	10	10
4	1000	10	10

For $p = 2$, a feasible solution is the one depicted in Figure 4.16:

$$x_5^* = 1, x_6^* = 1, x_7^* = 0$$
$$y_{15}^* = 1, y_{25}^* = 1, y_{36}^* = 1, y_{46}^* = 1, y_{ij}^* = 0 \text{ otherwise}$$

The corresponding transportation cost is

$$\underbrace{5\left(10+10+100\frac{1}{10}\right)}_{\text{cost for pair (1,3)}}+\underbrace{15\left(10+10+100\frac{1}{10}\right)}_{\text{cost for pair (2,4)}}=5\cdot 30+15\cdot 30=600.$$

The ILP model is presented in Model 4.8.

$$\begin{aligned}
\min\ & 5\big[10y_{15}+1000y_{16}+10y_{17}+1000y_{35}+10y_{36}+10y_{37}+\\
&+\frac{1}{10}(0y_{15}y_{35}+100y_{15}y_{36}+\ldots)\big]+\\
&+15\big[10y_{25}+1000y_{26}+10y_{27}+1000y_{45}+10y_{46}+10y_{47}+\\
&+\frac{1}{10}(0y_{25}y_{45}+100y_{25}y_{46}+\ldots)\big]\\
& x_5+\ x_6+\ x_7=2\\
& y_{15}+y_{16}+y_{17}=1\\
& y_{25}+y_{26}+y_{27}=1 \qquad \text{allocation}\\
& y_{35}+y_{36}+y_{37}=1\\
& y_{45}+y_{46}+y_{47}=1\\
& y_{15}\le x_5\\
& y_{16}\le x_6 \qquad \text{linking between location–allocation}\\
& y_{17}\le x_7\\
& \vdots\\
& x_5\ ,\ x_6\ ,\ x_7\in\{0,1\}\\
& y_{ij}\in\{0,1\} \quad \forall i\in I, j\in J
\end{aligned}$$

Model 4.8

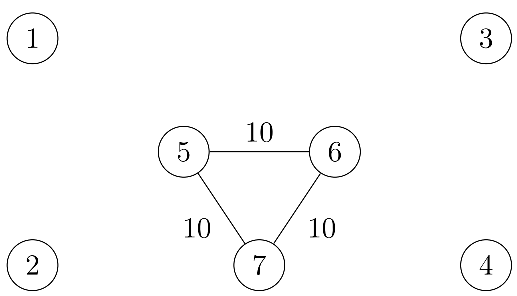

Figure 4.15: Example of p-hub

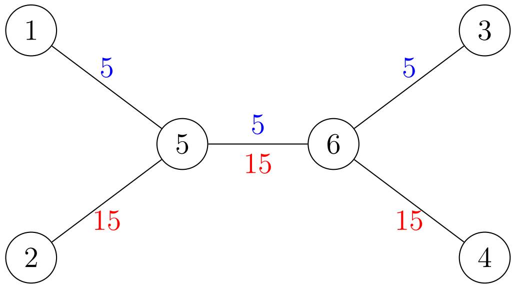

Figure 4.16: Example of p-hub. Values on edges indicate flows.

4.3.4 Maxisum location problem

The previous location problems involving both demand nodes and facilities assume that it is desirable to locate facilities as close as possible to demand nodes. This is not true in case of undesirable facilities (e.g. prisons, power plants, waste repositories...): in that case we would like to locate facilities far from demand nodes.

The problem is to determine the locations of p facilities (p is given) in order to maximize the total weighted distance between nodes and facilities.

In order to better define the problem, let us consider the p-median location problem introduced in Section 4.3.1 (see also Figure 4.9), where $p = 2$, and $h_i = 1,\ i = 1, 2, 3, 4,$ and

$d_{ij} =$

	5	**6**	**7**
1	20	25	60
2	30	15	10
3	30	15	10
4	30	15	10

What happens if we open the two facilities in 5 and 6 (as in Figure 4.17)? In this scenario, the most critical facility for node 1 is 5, since it is closer than 6. Instead, the most critical facility to the demand nodes 2, 3 and 4 is the one in 6. Therefore the total distance between demand nodes and their critical open facilities is 65: this is the objective function value associated with the considered feasible solution. If we opened the two facilities in 5 and 7, then the objective function value would be $20+10+10+10 = 50$; if we opened them in 6 and 7, it would be $25 + 10 + 10 + 10 = 55$. Since we want to "maximize" the total distance, opening in 5 and 6 gives the optimal solution.

In order to formulate the problem, let us introduce the following decision variables:

- $x_j = \begin{cases} 1 & \text{if we open at } j \\ 0 & \text{otherwise} \end{cases}, \ \forall j \in J;$
- z_i: the distance between i and the corresponding critical (i.e. closer) facility, $\forall i \in I$.

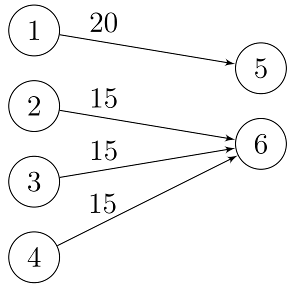

Figure 4.17: Optimal solution for the Maxisum example

The ILP model for Maxisum is therefore the following:

$$\max \sum_{i\in I} z_i$$
$$\sum_{j\in J} x_j = p \qquad \text{(Maxisum)}$$
$$(d_{ij} - M)x_j + M \geq z_i \quad \forall i \in I, j \in J$$
$$x_j \in \{0,1\} \quad \forall j \in J$$

where M denotes the usual very large value.

In order to explain the "$\geq$" constraints, consider an $i \in I$. For each $j \in J$:

- if $x_j = 1$, then $d_{ij} - \cancel{M} + \cancel{M} \geq z_i$, i.e. z_i is a lower bound to the distance between i and the opened facility in j;
- if $x_j = 0$, then $z_i \leq M$ (always satisfied), i.e. in this case the constraint "does not constrain" since in j there is no facility.

Therefore: $z_i \leq d_{ij}$, $\forall j \in J$ such that there is a facility, and so z_i is a lower bound to the minimum distance (which is the critical one) between i and the opened facilities, $\forall i \in I$. By maximizing $\sum_{i\in I} z_i$, the solver forces each z_i to be as large as possible, so maximizing the total maxisum distance, i.e. the total distance between each demand node and its closest facility (in the optimal solution, z_i^* gives such a minimum distance, $\forall i \in I$).

Let us resume the previous example. The model describing the problem is presented in Model 4.9. The optimal solution is:

$$x_5^* = 1, x_6^* = 1, x_7^* = 0,$$
$$z_1^* = 20, z_2^* = z_3^* = z_4^* = 15,$$
with optimal objective function value 65.

References Z. Drezner and H. Hamacher (2004): Chapter 3

$$
\begin{aligned}
&\max\ z_1 + z_2 + z_3 + z_4 \\
&x_5 + x_6 + x_7 = 2 \\
&\left.\begin{aligned}
(20 - M)x_5 + M \geq z_1 \\
(25 - M)x_6 + M \geq z_1 \\
(60 - M)x_7 + M \geq z_1
\end{aligned}\right\} \quad \text{demand node 1} \\
&\left.\begin{aligned}
(30 - M)x_5 + M \geq z_2 \\
(15 - M)x_6 + M \geq z_2 \\
(10 - M)x_7 + M \geq z_2
\end{aligned}\right\} \quad \text{demand node 2} \\
&\left.\begin{aligned}
(30 - M)x_5 + M \geq z_3 \\
(15 - M)x_6 + M \geq z_3 \\
(10 - M)x_7 + M \geq z_3
\end{aligned}\right\} \quad \text{demand node 3} \\
&\left.\begin{aligned}
(30 - M)x_5 + M \geq z_4 \\
(15 - M)x_6 + M \geq z_4 \\
(10 - M)x_7 + M \geq z_4
\end{aligned}\right\} \quad \text{demand node 4} \\
&x_5, x_6, x_7 \in \{0, 1\}
\end{aligned}
$$

Model 4.9

4.4 Location problems in the public sector

Location problems in the public sector include for instance the ones related to emergency services (ambulances, fire stations, police units), school system and postal facilities. The optimization criteria are often social cost minimization, universality of service, efficiency and equity (as opposed to profit maximization).

Since these objectives are difficult to measure, often they are "surrogated" by minimizing the costs needed for full coverage of the service (i.e. leading to SCLP), or searching for maximal coverage given a budget (i.e. leading to MCLP and its variants). The basic models in this sector are in fact SCLP and MCLP (also p-center and p-median).

Let us consider now a more complex example arising in Mobile Emergency Services: FLEET (Facility Location and Equipment Emplacement Technique). It is a generalization of MCLP as it envisions the simultaneous location of more types of facilities. The problem is to decide where to locate two different types of fire-fighting servers (i.e. pump or engine brigades, and ladder or truck brigades), as well as the depots housing them so as to cover the maximum number of people by both an engine company sited within a given distance E and a truck company sited within another distance T.

The following input data are given:

- $N_i^E = \{j \in J : d_{ij} \leq E\},\ \forall i \in I$;
- $N_i^T = \{j \in J : d_{ij} \leq T\},\ \forall i \in I$;
- p^S: the number of fire stations (depots) to open;

- p^{E+T}: the number of (engine or truck) companies to open;
- a_i: the number of inhabitants in i, $\forall i \in I$.

Let us introduce the following decision variables:

- $x_j^E = \begin{cases} 1 & \text{if an engine company is positioned in a fire station at } j \\ 0 & \text{otherwise} \end{cases}, \ \forall j \in J;$
- $x_j^T = \begin{cases} 1 & \text{if a truck company is positioned in a fire station at } j \\ 0 & \text{otherwise} \end{cases}, \ \forall j \in J;$
- $x_j^S = \begin{cases} 1 & \text{if a fire station (depot) is established at site } j \\ 0 & \text{otherwise} \end{cases}, \ \forall j \in J;$
- $z_i = \begin{cases} 1 & \text{if } i \text{ is covered} \\ 0 & \text{otherwise} \end{cases}, \ \forall i \in I$ (as in MCLP).

The ILP model is therefore the following:

$$
\begin{aligned}
&\max \sum_{i \in I} a_i \, z_i \\
&z_i \leq \sum_{j \in N_i^E} x_j^E && \forall i \in I \quad ① \\
&z_i \leq \sum_{j \in N_i^T} x_j^T && \forall i \in I \quad ① \\
&x_j^E \leq x_j^S && \forall j \in J \quad ② \\
&x_j^T \leq x_j^S && \forall j \in J \quad ② \\
&\sum_{j \in J} x_j^S = p^S \\
&\sum_{j \in J} x_j^E + \sum_{j \in J} x_j^T = p^{E+T} \\
&x_j^S, x_j^E, x_j^T \in \{0,1\} \ \forall j \in J \\
&z_i \in \{0,1\} && \forall i \in I
\end{aligned}
$$

The numbered constraints have the following meaning:

① z_i can be 1 only if at least one engine is within E and one truck is within T;

② housing of companies is allowed only at nodes with a depot.

References Z. Drezner and H. Hamacher (2004): Section 4.2.3

Textbooks

Bigi, G., A. Frangioni, G. Gallo, and M.G. Scutellà (2014). *Appunti di Ricerca Operativa*. Italian. URL: `http://didawiki.di.unipi.it/doku.php/ingegneria/ricercaoperativa1/start#testi_di_riferimento`.

Drezner, Z. and H.W. Hamacher (2004). *Facility Location: Applications and Theory*. Springer series in operations research. Springer. ISBN: 9783540213451.

Ghiani, G., G. Laporte, and R. Musmanno (2004). *Introduction to Logistics Systems Planning and Control*. Wiley Interscience Series in Systems and Optimization. Wiley. ISBN: 9780470091654.

Ragsdale, C.T. (2004). *Spreadsheet Modeling and Decision Analysis: A Practical Introduction to Management Science*. Cengage South-Western. ISBN: 9780324321777.

Toth, P. and D. Vigo (2002). *The Vehicle Routing Problem*. Monographs on Discrete Mathematics and Applications. Society for Industrial and Applied Mathematics. ISBN: 9780898715798.