

Chapter 1

Pioneering Developments in Location Analysis

H. A. Eiselt and Vladimir Marianov

1.1 Location Problems: Its Problem Statement, Its Components, and Applications

A mathematician would probably define a location problem as solving the following question: “given some metric space and a set of known points, determine a number of additional points so as to optimize a function of the distance between new and existing points.” A geographer’s explanation might be that “given some region in which some market places or communities are known, the task is to determine the sites of a number of centers that serve the market places or communities.” Students of business administration will want to determine “the location of plants and market catchment areas in the presence of potential customers,” while computer scientists (or, more specifically, analysts in computational geometry) may want to determine “the minimum number of equal geometrical shapes that are required to cover a certain area, and the positions of their centroids.”

All these views have in common the basic components of a location problem: a space, in which a distance measure is defined, a set of given points, and candidate locations for a fixed or variable number of new points. We refer to the known points as “customers” or “demands,” and the new points to be located as “facilities.”

As far as the space is concerned in which customers are and facilities are to be located, we distinguish between a subset of the d -dimensional real space (most prominently, the two-dimensional plane) and networks. Each of these two categories has two subcategories: one, in which the location of the facilities is continuous, and the other, in which it is discrete. These two subcategories will determine the toolkit needed by the researcher to solve problems. In discrete problems, the decision is

H. A. Eiselt (✉)

Faculty of Business Administration, University of New Brunswick,
Fredericton, NB E3B 5A3, Canada
e-mail: haeiselt@unb.ca

V. Marianov

Department of Electrical Engineering, Pontificia Universidad Católica de Chile,
Santiago, Chile
e-mail: marianov@ing.puc.cl

whether or not to locate a facility at that spot, thus modeling the decision with a binary variable is obvious. The result is a (mixed-)integer linear programming problem. On the other hand, continuous problems will have continuous variables associated with them, indicating the coordinates of the facilities that are to be located. Since functions of the distance are typically nonlinear, a nonlinear optimization problem will follow in this case. With the tools for the solution of the different types of problems being so different, it is little surprise that most researchers have decided to either work with continuous or with discrete models. Both types of models are represented in this book.

The space commonly corresponds to a geographical region. An obvious choice for representing such a region is the two-dimensional plane. In some cases, a one-dimensional space is used to simplify the analysis of complicated problems, as in the case that includes competition between firms. Location problems can also be defined over non-geographical spaces. For example, an issue space in which potential voters for a presidential election are located according to their positions in relation to the issues. The problem would be to optimally locate a candidate so to maximize his vote count in the presence of competing candidates, assuming that voters will vote for the candidate whose position is closest to them. Or consider a skill space in which a number of tasks (demands) have known locations, each location representing the combination of skills required to successfully accomplish the task. The location problem would be to locate company employees in such a way that all tasks are performed by sufficiently skilled employees, and no employee is overloaded with work.

In this book we deal with location problems, which are defined as models, in which the facilities and demands are very small as compared to the space they are located in. Examples of facility location problems are finding the location of an assembly plant within a country; or selecting the site of a school in one of several candidate points in a city. In these cases, the facilities can be considered dimensionless points. A generalization would be the location of lines in a plane, where the lines represent a new road. In some cases, location problems can be cast as geometrical problems: an example is finding the smallest circle that contains a known set of points on a plane, which is equivalent to locating a facility (the center of the circle) in such a way that the largest distance (the radius of the circle) between any customer (the given points) and the facility (the center of the circle) is minimized. In contrast to location problems, layout problems feature facilities, whose size is significant in comparison to the space the facility is to be located in. Examples of layout problems include the siting of a drilling machine in a workshop, the location of tables in a fast-food restaurant, or the location of supply rooms in a hospital. Other things being equal, layout problems are more difficult than location problems with similar features. One reason is that layout problems must consider the shape of the facility to be located, which is unnecessary in location models, where facilities are just points. Layout models are not discussed in this book.

Location problems can be formulated to answer to several different questions. Not only can the location of the facilities be unknown, but also the number of facilities and their capacities. Furthermore, when there is more than one facility, the solution of the location problem also requires finding the assignment of demands to

facilities, called the “allocation” problem. Examples of this problem are the assignment of individual customers to a particular warehouse from which grocery is to be delivered by a phone-order company, or the dispatching of specific ambulances to emergency sites. In these cases, the owner or operator of the facilities does the allocation, but there are other cases in which users decide what facility to patronize, such as the decision which theater to patronize or at what fast-food store to have lunch.

For the location problem to make sense, customers and facilities must be related by distance. Again, this distance can be measured either over a geographical region or over any kind of space. Typical goals are to minimize the distance between facilities and their assigned customers (minisum problems) or maximize the amount of demand (number of customers) that is within a previously specified distance from their assigned facilities. When all the demand is to be served, a goal could be to minimize the number of facilities needed for all the customers to be within a distance from their facility. If competing facilities are involved and distance is an issue for customers, the planner’s objective will be to locate his facilities closer than competitors’ facilities to as many customers as possible.

Proximity to facilities is, however, not always desirable. When locating a new landfill, most people will object if the facility is to be located close to their homes—the usual NIMBY (not in my back yard) argument applies, even though they usually understand that locating the landfill too far away will cost them, as they are, directly or indirectly, charged for the collection and transportation costs of the solid waste. In this case, the problem could be formulated with conflicting objectives: the landfill should be not too far from the area it serves, but most of the population should be as far as possible from it. These objectives have been described as “push” and “pull”.

Besides distances between facilities and demands, there are other factors that can be relevant when seeking good locations. Availability of services and skilled technicians at the candidate locations, land cost, existence of competitors and regional taxes are some examples. While such factors are maybe as important as or possibly even more important than proximity, most standard location problems do not consider them. This stresses the importance of considering the output of the standard location models as an input for the decision maker, who can then include any of the features that were ignored by the mathematical model.

Most of the applications of location models involve, decisions on the strategic level. As such, the decisions tend to be long-term, which implies that many of the data used in the decision-making process, will be quite uncertain. While there usually are few changes concerning distances, future demand tends to be highly uncertain. This problem is exacerbated by the high cost of locating and relocating a facility, meaning that once a location is chosen, it can only be changed at great expense. This argument implies that probabilistic and/or robust models are most likely to result in facility locations that are acceptable to decision makers.

Applications of location models are found in many different fields. Some of these applications are fairly straightforward, such as the locations of trucking terminals, blood banks, ambulances, motels and solid waste transfer points. Others are nontraditional and not at all obvious. Good examples are the location of measuring

points for glaucoma detection, the location of new employees in skill space, the location of advertisements in media and many others.

1.2 A Short History of the Early Developments in Location Analysis

In the early seventeenth century, a question was posed about how to solve the following puzzle: “Given three points in a plane, find a fourth point such that the sum of its distances to the three given points is as small as possible.” This question is credited to Fermat (1601–1665), who used to tease his contemporary mathematicians with tricky problems. The earliest (geometrical) solution is probably due to Torricelli (1598–1647), Fermat’s pupil and the discoverer of the barometer, although due to the many discoveries and rediscoveries of the problem and its solution, there are different opinions about the true origin of the problem and who solved it first. Because of the many scientists involved in the process, the problem has been called the Fermat problem, the Fermat-Torricelli problem, the Steiner problem, the Steiner-Weber problem, the Weber problem, and many variation thereof. The inclusion of Weber’s name follows his generalization of the problem by assigning different weights to the known points, so transforming the mathematical puzzle into an industrial problem, in which a plant is to be located (the unknown point) so to minimize transportation costs from suppliers and to customers (the known points) requiring different amounts of products (the weights). Weber’s name stuck, even though the formulation of the model in Weber’s book is in an appendix written by Pick. Even if the first formal occurrence of a location problem were due to Fermat or one of his contemporaries, location analysis must be much older than that. For centuries before, people may have wondered in which cave to live, where to build houses, villages, churches, and other “facilities.” And they solved these problems using some sort of heuristic method.

Location problems frequently require solving an associated allocation or assignment problem: if locations for more than one facility are known, which facility will serve what customers? A step towards the solution of the allocation problem was taken very early in the seventeenth century. Descartes imagined the universe as a set of vortices around each star—the “heavens”—and illustrated his theory with a drawing that made an informal use of what later would become known as Voronoi polygons and Voronoi diagrams. This concept was subsequently used by Dirichlet in 1850, extended to higher dimensions by Voronoi and rediscovered by Thiessen, both in the early twentieth century. While Thiessen’s application involves an improved estimate of precipitation averages, Voronoi diagrams are generally useful tools when consumers have to be assigned to their closest facilities.

Between the 1600s and the 1800s, there was no registered activity related to location problems other than a puzzle in the Ladies Diary or Woman’s Almanack in 1755 as reported in a book edited by Drezner and Hamacher in 2002. In 1826, the geographer von Thünen developed a theory concerning the allocation of crops on the land that surrounds a town. His point was that the agricultural activity should

be organized around the town according to transportation costs and value of the products. This theory results in crops cultivated in concentric circles about the town. Clearly, a location issue, although it could also be considered a land layout problem, as the areas or rings to be located around the central place are sizeable in comparison to the space considered.

In 1857, Sylvester asked another location-related question: “It is required to find the least circle which shall contain a given system of points in a plane.” This question was later answered by Sylvester himself in 1860 and by Chrystal in 1885. Nowadays, we would call this a one-center problem in the plane.

As in many other fields, the pace of discoveries increased dramatically in the twentieth century. In the late 1920s, the mathematical statistician-turned economist Hotelling wrote a seminal paper on competitive location models that spawned a rich diversity of models that are still under active discussion today. Also during that time, Reilly introduced gravity models into the fray as a way customers gravitate to facilities. The 1930s saw contributions by Christaller, who founded central place theory, and Weiszfeld, who developed his famed algorithm that solved Weber problems with an arbitrary number of customers.

Later important contributions included those by Lösch and the regional scientists Isard and Alonso. The birth of modern quantitative location theory occurred in the mid-1960s, when Hakimi wrote his path-breaking analysis of a location model on networks, which, in today’s parlance, is a p -median problem. Following Hakimi’s papers, ReVelle, Church, Drezner, Berman, and many others have made important contributions to location science. Some of those have also contributed to this book.

1.3 Some Standard Location Problems

Facility location problems can take a variety of forms, depending on the particular context, objectives and constraints. It is common to classify location problems in minimax, covering and minimax problems. Most of the numerous remaining facility location problems can be seen as combinations or modified versions of these key problems. We present here prototypes of these basic classes, as well as an assortment of other models that we consider of importance, either from a theoretical viewpoint, or because of their practical interest. This list of models is also a guide for understanding the following chapters and a way to introduce some standard notation, since in most of the chapters of this book, the original notation of the reviewed papers is preserved.

We first state a base formulation, which is used throughout this section to present some of the different problems and models for location on networks. This formulation applies only when demands lie at the nodes and facilities are to be located only at nodes of the network. The problem can be formulated as follows.

$$\text{Min } \sigma \left(\sum_{i,j} c_{ij}y_{ij} + \sum_j f_jx_j + gz \right) \quad (1.1)$$

$$\text{s.t. } \sum_{j \in N_i} y_{ij} = 1 \quad i = 1, 2, \dots, n \quad (1.2)$$

$$y_{ij} \leq x_j \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m. \quad (1.3)$$

$$y_{ij}, x_j \in \{0, 1\}, \quad z \in \mathbb{R}^+ \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m, \quad (1.4)$$

where the variables are:

- x_j : a location variable that equals 1, if a facility is located at node j , and 0 otherwise,
- y_{ij} : an allocation variable that equals 1, if customer i is assigned to a facility at j , and 0 otherwise, and
- z : a continuous variable that takes the value of a maximum or a minimum distance, depending on the problem being solved.

The subscripts are

- i, j : a subscript that indicates customers and potential facility sites, respectively,
- N_i : is the set of nodes. Its definition depends on the problem, and n and m denote the total number of customers and potential facility locations, respectively.

Finally, the parameters are

- σ : a parameter that takes the value 1 or (-1) , depending on whether the objective is minimized or maximized, and
- c_{ij}, f_j, g : parameters that depend on the problem being solved.

The objective (1.1) of the base problem optimizes linear combinations of the location and allocation variables and the distance variable. Depending on the problem, these linear combinations represent maximum, minimum or average distances, investment or transportation and manufacturing costs. The set of constraints (1.2) state that each customer's demand must be satisfied, and that all of it is satisfied from one facility. In other words, that each customer or demand node is assigned to or served by exactly one facility. The set of constraints (1.3) allows demand node i to be assigned or allocated to a facility at j only if there is an open facility in that location i.e., $x_j = 1$. Constraints (1.4) define the nature of all variables as continuous or binary. Note that these constraints define a feasible set of solutions of the problem of locating a (yet) undetermined number of facilities, choosing their sites from a number of known candidate locations, and finding the right assignment of demand nodes or customers to these facilities.

1.3.1 Minisum Problems

Minisum problems owe their name to the fact that a sum of facility-customer distances is minimized. These problems include single-facility, multiple-facility and

weighted versions of the Weber problem; the *simple plant location problem SPLP* and the p -median problem *PMP*.

1.3.1.1 Minisum Problems on the Plane: The Weber Problem

The Weber problem consists in finding the coordinates (x, y) of a single facility X that minimizes the sum of its distances to n known customer or demand points with coordinates (a_i, b_i) and weights w_i , $i = 1, 2, \dots, n$, all of them lying on the same plane. This is an unconstrained nonlinear optimization problem, and its general formulation is

$$\text{Min } z(X) = \sum_{i=1}^n w_i d(X, P_i),$$

where $d(X, P_i)$ is the distance between the facility and demand i . If Euclidean distances are used, this distance is $d(X, P_i) = \sqrt{(x - a_i)^2 + (y - b_i)^2}$.

An application of this problem is locating a single warehouse (the facility) that supplies different amounts (weights) of products to a number of dealers (the demands), in such a way that the total transportation cost is minimized, assuming that this cost depends on the distance and amount of transported product.

The solution of this problem is easily found using differential calculus. However, the coordinates of the optimal facility location turn out to be a function of the distance $d(X, P_i)$, which is unknown. The practical solution of this problem, found by Weiszfeld, involves the use of an iterative algorithm that takes as an initial solution the point that minimizes the sum of the squares of the distances.

A natural generalization of the Weber problem is its multiple-facility version. As more than one facility is to be located, an allocation problem must be solved together with the location problem, which makes the problem far more difficult. In the simplest case, the assumption is made that each demand is assigned to its closest facility. This problem was described and a heuristic proposed by Cooper in 1963.

1.3.1.2 Minisum Problems on Networks: Plant Location and Median Problems

The Simple Plant Location Problem

The simple plant location problem, sometimes also referred to as the *uncapacitated facility location problem* seeks minimizing production, transportation and site-related investment costs. The model assumes that the costs of opening the facilities depend on their location, and that the investment budget is not a constraint. Then, the number of facilities to be opened is left for the model to decide. Its integer programming formulation is the base model (1.1)–(1.4), with $\sigma = +1$, $c_{ij} = (e_j + \tilde{c}_{ij}d_{ij})w_i$,

where e_j denotes the production cost per unit, \tilde{c}_{ij} are the transportation costs from node i to j per unit, d_{ij} symbolizes the shortest distance between i and j , the parameters w_i denote the quantity of the product that must be shipped to customer i , the parameters f_j are the fixed costs of opening a facility at node j , and g equals zero. The set N_i in constraints (1.2) contains all possible candidates to location of facilities. The objective to be minimized, representing the total sum of variable and fixed costs, is now:

$$\text{Min } \sum_{i,j} c_{ij}y_{ij} + \sum_j f_jx_j.$$

Since the costs of locating facilities are included in the objective, the solution of the model prescribes the number of facilities to be located, depending on the tradeoff between transportation and fixed costs.

There is also a “weak” formulation of this problem that uses an aggregated version of constraints (1.3) for each facility location j , $\sum_j y_{ij} \leq nx_j$.

There is an important body of literature focusing on this problem, because of its versatility and practical interest. Due to its difficulty, many solution methods have been proposed. In Chap. 2 of this book, Verter describes this problem, and reviews in detail the classical dual ascent method by Erlenkotter, as well as the heuristic proposed by Kuehn and Hamburger.

The 1-Median Problem

The network equivalent of the single-facility Weber problem on the plane is the 1-median problem, whose formulation is exactly the same as that of the single-facility Weber problem. However, in the 1-median problem the demands occurs at the nodes of a network. Hakimi proved that there is always an optimal location of the facility at a node of the network. As a consequence, finding the solution reduces to searching the nodes of the network, and the problem can be stated as follows: find a node v^* such that for all nodes v_k , $k = 1, 2, \dots, n$,

$$\sum_{i=1}^n w_i d(v^*, v_i) \leq \sum_{i=1}^n w_i d(v_k, v_i).$$

This problem has been solved for locations on tree networks in the 1960s by Hua-Lo Keng, who was looking for the optimal location of a threshing floor for wheat fields, and rediscovered by Goldman in 1971.

The p -Median Problem

The p -median problem is a p -facility generalization of the 1-median problem, in which it is assumed that the decision maker knows how many facilities are to be lo-

cated, and that the cost of locating the facilities is the same no matter where they are located. As in the continuous case, the allocation problem must be solved together with the location problem, i.e., not only the location of the p facilities needs to be found, but also, for each demand, the facility to which this demand is assigned. Let X_p be a set of p points x_1, x_2, \dots, x_p . Define distance of a node v_i to its closest point in X_p as

$$d(v_i, X_p) = \min \{d(v_i, x_1), d(v_i, x_2), \dots, d(v_i, x_p)\}.$$

Then the set X_p^* is a “ p -median” of the network, if for every X_p on the network,

$$\sum_{i=1}^n w_i d(v_i, X_p^*) \leq \sum_{i=1}^n w_i d(v_i, X_p)$$

i.e. X_p^* is the set of p points on the graph such that, if facilities are open at these points, the total weighted distance between the demands and their closest facility would be minimized. For this problem, Hakimi proved that there is always a solution set containing only nodes of the network.

ReVelle and Swain formulated this problem as a linear integer programming problem. Their model uses the same equations (1.1)–(1.4) of the basic formulation above, and an additional constraint that requires exactly p facilities to be located:

$$\sum_j x_j = p.$$

In their formulation of the p -median, the objective minimizes the sum of the weighted facility-demand distances, i.e., in the objective (1.1) of the basic formulation, the parameters become $\sigma = +1$, $c_{ij} = w_i d_{ij}$, where w_i is the weight associated to each demand node (demand volume or number of customers) and d_{ij} is the shortest distance between nodes i and j , measured along routes on the network, $f_j = 0$, and $g = 0$. The objective (1.1) becomes

$$\text{Min} \sum_{i,j} w_i d_{ij} y_{ij}.$$

Also, the set N_i in constraint (1.2) contains all possible candidates to location of facilities. Note that this formulation assumes that each demand is assigned to its closest facility.

In Chap. 3 of this book, Marianov and Serra review the classic contributions of Hakimi, made in 1964 and 1965, including the definitions of the absolute median and the absolute p -medians of a network. Also, they review the paper in which the first integer programming formulation of the p -median problem was proposed by ReVelle and Swain in 1970.

1.3.2 Minimax Problems

The idea behind minimax problems is to minimize the longest distance between a customer and its assigned facility, expecting that a maximum equity or justice will be achieved. As before, these problems have been defined both on a plane and on a network.

1.3.2.1 Continuous Minimax Problems: Single and Multiple-Facility Center Problems on the Plane

The oldest known single-facility minimax problem was formulated by Sylvester in 1857 in a one-sentence reference. The problem consists in finding the smallest circle containing a set of given points. Finding the solution of this problem is equivalent to finding the location of the center of a circle that encloses all given points, in such a way that its radius is minimized. If a facility is located at the center of the circle, and the points represent demands, the maximum distance between the facility and a demand will be minimized. This problem is now known as the 1-center on the plane or the continuous 1-center. Several geometrical solution methods have been proposed to solve this problem, the first one due to the same Sylvester and rediscovered by Chrystal in 1885.

A natural extension of the 1-center problem is the p -center on the plane. Its formulation is as follows. Let X_p be a set of p points x_1, x_2, \dots, x_p on the plane. Define distance of a customer c_i to its closest point in X_p as

$$d(c_i, X_p) = \min \{d(c_i, x_1), d(c_i, x_2), \dots, d(c_i, x_p)\}.$$

Then the set X_p^* is a “ p -center”, if for every X_p on the plane,

$$\max d(c_i, X_p^*) = \max d(c_i, X_p).$$

In other words, X_p^* is the set of p points on the plane such that, if these points were facilities, the maximum distance between a demand and its closest facility would be minimized. Minimax problems on the plane, as well as their solution methods are analyzed in Chap. 4 of this book. In that chapter, Drezner describes the early research done in the nineteenth century, shows how geometrical methods were used to solve the continuous minimax problems, describes the algorithmic contributions by Elzinga and Hearn in the 1970s, and compares both approaches through new unpublished computational experience.

1.3.2.2 The 1-Center and p -Center Problems on Networks

The 1-center on a network has been first formulated and solved by Hakimi in 1964 and the solution method improved by Goldman in 1970 and in 1972. Hakimi for-

mulates the problem as follows: Let v_i be a node, w_i its potential demand, and x any point on the network. Define $d(v_i, x)$ as the minimum distance between v_i and x . Find a point x^* on the network such that for x ,

$$\max_i w_i d(x^*, v_i) \leq \max_i w_i d(x, v_i).$$

This point is called the “absolute center” of the network. Hakimi proved that his result for minisum problems does not apply to the 1-center problem. In other words, the solution can be on an arc of the network, which is even true if all demands are equal.

Later, the 1-center was extended to a multiple-facility problem: the p -center: let X_p be a set of p points x_1, x_2, \dots, x_p anywhere on the network. Define the distance between a node v_i and its closest point in X_p as

$$d(v_i, X_p) = \min \{d(v_i, x_1), d(v_i, x_2), \dots, d(v_i, x_p)\}.$$

Then the set X_p^* is a “ p -center”, if for every X_p on the network,

$$\max_i d(v_i, X_p^*) \leq \max_i d(v_i, X_p).$$

In other words, X_p^* is the set of p points on the network such that, if these points were facilities, the maximum distance between a demand and its closest facility would be minimized.

If facility locations are restricted to the nodes of the network, the p -center problem can be formulated as an integer programming problem using the base formulation (1.1)–(1.4), and setting $\sigma = +1$, $c_{ij} = 0$, $f_j = 0$, and $g = 1$. Variable z is the maximum distance between a demand and its assigned facility. Two extra constraints must be added:

$$d_{ij}y_{ij} \leq z \quad \text{for } i = 1, 2, \dots, n, j = 1, 2, \dots, m, \text{ and} \quad (1.5)$$

$$\sum_j x_j = p. \quad (1.6)$$

Constraint (1.5) requires z , the variable to be minimized, to take the value of the maximum distance between any customer and its assigned (closest) facility. Since constraint (1.2) forces assignment of each customer to exactly one facility, constraint (1.5) can take the aggregated form

$$\sum_j d_{ij}y_{ij} \leq z \quad \text{for } i = 1, 2, \dots, n.$$

Constraint (1.6) sets the number of facilities to be located to p . Pioneering developments on center problems on a network, specifically the works by Hakimi in 1964, Goldman in 1972, and Minieka in 1970 are the subject of Chap. 5 of this

book by Tansel. The chapter first defines the single facility p -center problem and its multiple-facility version and then describes their properties. Later, the three classic contributions are reviewed: Hakimi's definition of the absolute center, Goldman's algorithm for locating the center and Minieka's method for finding the p -centers on a network when locations are restricted to nodes.

1.3.3 Covering Problems

In minisum and minimax location models, the distance between each and every customer and his closest facility is explicitly considered. While minisum problems minimize the average distance between a customer and its closest facility, minimax problems minimize the longest customer-facility distance. In contrast, covering models do not explicitly include customer—facility distances in the model and these distances only matter if they exceed a preset value \bar{D} . A common example appears in emergency medical services: service is considered adequate if an ambulance reaches the site of the emergency in no more than, say 8 minutes. In the food business, a pizza parlor could offer free pizzas when the delivery time exceeds 30 minutes. In synthesis, the concept of coverage implies that a customer can and will be adequately served (“covered”) when a facility is located within a preset distance or travel time. Distance is no longer included in the objective function, but appears as a constraint.

1.3.3.1 The Location Set Covering Problem

The *location set covering problem* *LSCP* was first introduced by Hakimi, and formulated as an integer programming problem by Toregas et al. in 1971. Its goal is to find the minimum number of facilities and their locations, so that all customers are covered, meaning that all customers are no farther than a preset distance \bar{D} of their closest facility. Although the original formulation uses only location variables x_j , the problem can be reformulated using the base model (1.1)–(1.4). To reproduce the *LSCP* with this base model, we set $\sigma = +1$, $c_{ij} = 0 \forall i, j$, $f_j = 1 \forall j$, $g = 0$, and the set N_i in constraint (1.2) contains all candidates to location of facilities that are within distance \bar{D} , i.e., $N_i = \{j : d_{ij} \leq \bar{D}\}$. The objective is now

$$\text{Min } \sum_j x_j.$$

The formulation provided by Toregas et al. is simpler. In that formulation, only location variables are used and the allocation problem is not solved, i.e. the model does not prescribe which facility provides the service, which can be provided by any of the facilities within range. In consequence, it does not matter if a customer has more than one facility within covering distance.

Although at first sight there is no apparent relation between minimax and covering problems, Minieka found that the solution of the node constrained p -center can be found by solving a sequence of location set covering problems.

1.3.3.2 The Maximal Covering Location Problem (MCLP)

If the minimum number of facilities needed to cover all the demand cannot be achieved, because of a limited budget or any other constraints on the total number of facilities, the *maximal covering location problem* of Church and ReVelle solves the problem of maximizing covered demand when there are a limited number of facilities to be sited. In terms of the base model (1.1)–(1.4) $\sigma = -1$, $c_{ij} = h_i$ if $d_{ij} \leq \bar{D}$ and 0 otherwise, where h_i is the demand volume of customer i , the fixed costs $f_j = 0 \forall j$, the value of $g = 0$, and the set N_i in constraint (1.2) contains all candidates to location of facilities. The objective is now

$$\text{Max } \sum_{i,j} c_{ij} y_{ij}.$$

This objective maximizes the demand-weighted number of customers that have at least one facility located within distance \bar{D} .

Church and ReVelle's model was formulated using only location variables and covering variables. Again, the allocation or assignment problem is not solved by this model.

Seminal contributions on covering problems are reviewed in Chap. 6 of this book. In that chapter, Snyder thoroughly reviews the contributions by Hakimi in 1965, in which the location set covering problem is defined, the work by Toregas et al. in 1971 with the first integer programming formulation of the same problem, and the contribution by Church and ReVelle in 1974, which defines the maximum covering location problem and formulates it as an integer programming problem. The author of the chapter then adds some insight into these problems by running new computational experiments on variants of the original problems, including tradeoff between coverage and number of facilities; Lagrangean approach as a solution method; inclusion of budget constraints; and relationship to the p -median problem.

1.3.4 Other Relevant Location Problems

Minimax, minisum and covering models are considered the key formulations from which the remaining location problems are descendants. There are some nonstandard location problems and models that have attracted considerable attention from researchers. We briefly review some of these in the remainder of this section.

1.3.4.1 Competitive and Conditional Location Problems

None of the key location problems considers the existence of other facilities, either previously located or to be located in the same region by a competing firm. If this is the case, a whole new range of problems can be posed, in which the relationship between facilities plays an important role: competitive and conditional location problems.

In order to analyze the locational pattern of the competing firms, assumptions need to be made about consumers' behavior. A typical assumption is that customers prefer closest facilities over more distant ones, and facilities at which the goods or service can be obtained at a lower price over the more expensive facilities.

Competitive Location: Equilibrium Problems

Suppose a market over which the demand is distributed, either continuously or at discrete points. Two firms are entering the market, each one with one facility that can move at any time without incurring costs. There are two fundamental questions. The first asks that given that customers patronize the facility at which they can satisfy their demand for the lowest full price (i.e., mill price plus transportation costs), is there an equilibrium? In other words, is there a situation, in which both facilities have decided on a location and a price, so that neither facility has an incentive to unilaterally change its price and/or its location? The second question is then, what the location and price situation in such an equilibrium is, given that it exists.

These questions—whose answer is now known as Nash equilibrium—were first answered for two competing facilities by Hotelling in 1929, who studied the problem in its simplest form: a linear market (i.e., a simple line segment) with uniformly distributed customers. Each of the two firms locates a single branch, and both firms use mill pricing.

In Chap. 7, Eiselt discusses the details and applications of the problem of finding a locational equilibrium on a linear market; he provides a detailed review of Hotelling's seminal paper, assesses its impact and describes subsequent work on the subject, including the article by d'Aspremont et al. Their paper was published no less than fifty years after Hotelling's findings, and it invalidated some of Hotelling's results.

Conditional Location Problems

When facilities can move freely without costs, the search for Nash equilibrium is the natural approach. However, this is not always the case. Most facilities stay where they were first located, as relocation costs are significant. If a firm is planning to enter a market with *immobile* facilities, there are two different aspects to be considered. The first aspect concerns firms that intend to locate in the presence of

already existing competitors. this is a conditional location problems, i.e., the location of facilities, *given* that other facilities are already located in the region. This is the *follower's problem*, who will make his location choice based on whatever profit or revenue-maximizing objective he has.

On the other hand, the already existing firms are the leaders in the location game. The *leader's problem* is to locate under the consideration that after he has decided where to locate his branch(es) under the consideration that a follower will locate his own competing branches later. This means that the leader will have to take a cautious approach and guard against future entrants onto the market.

Solutions of this location game are usually referred to as von Stackelberg solutions, named after the German economist of that name. The main feature of the game is its sequential nature. In Chap. 8, Younies and Eiselt describe sequential location problems and explain the von Stackelberg concept of leader and follower firms. this is followed by a review of the classic application of these concepts in location analysis, made by Prescott and Visscher (1977), where they extend Hotelling results on a line to the sequential case.

In Chap. 9, Dasci discusses the conditional problems in a plane and a network, and offers a thoroughful revision of the seminal works by Drezner in 1982 and Hakimi in 1983. In these works, the follower's problem (resulting in what is now known as an $(r|X_p)$ -medianoid) and the leader's problem (dubbed an $(r|p)$ -centroid) are defined on the plane and a network, respectively, given that the leader locates p facilities and the follower locates r facilities.

1.3.4.2 Location of Undesirable or Semi-obnoxious Facilities

In many everyday situations, customers want to be as close as possible to the facilities that are to be located. It is the case of grocery stores and shopping centers, schools, primary care centers, and similar facilities. However, the situation changes when considering facilities such as prisons, landfills, or power plants. Facilities of this nature are considered undesirable by most people. In the early days, they were referred to as noxious and obnoxious. Some people nowadays refer to these facilities as semi-obnoxious, because although nobody wants them in the neighborhood, they cannot be too far away, because their operation becomes too expensive. However, while realistic solutions should incorporate desirable and undesirable features in a model, undesirable facility location models push facilities as far away from customers as possible. More realistic models will have to balance some of the usual criteria (minisum, minimax, and covering) with either distance maximization (maximin or maxisum), coverage minimization, or a minimization of the costs of compensating all the affected population.

An example of a maximin formulation of undesirable facility location problems is, again, derived from the base model (1.1)–(1.4), by setting the variable z to be the minimum distance between a demand and its assigned facility, $c_{ij} = 0, f_j = 0$, and $g = 1$, and maximizing the objective, which becomes simply z . The following two constraints must be added:

$$d_{ij}y_{ij} \geq z \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m. \quad (1.7)$$

$$\sum_j x_j = p. \quad (1.8)$$

Constraint (1.7) forces z to take the value of the smallest distance between a customer and his closest facility.

Early contributions on undesirable facility location are the subject of Chap. 10 by Melachrinoudis. The classical contribution of Church and Garfinkel 1978 (maximizing the sum of the distances from the closest facility) is reviewed, as well as the string of papers by Shamos published in 1975, Shamos and Hoey's work from 1975, Dasarathy and White's paper from 1980, and Drezner and Wesolowski's contribution from 1980, all of them dealing with the maximin problem, i.e., maximizing the shortest distance between a customer and its closest facility.

1.3.4.3 Probabilistic or Stochastic Location Problems

The pool of location problems would not be complete without the consideration of uncertainty. Future demand, the time of appearance of emergency calls, service times at the facilities, and travel times between demand nodes and facilities tend to be non-deterministic. Disruptions of facilities and connecting routes are also everyday random occurrences.

There is a large body of literature dealing with problems that include one or more of these sources of uncertainty. The design of emergency services is a popular application, as well as the location of immobile facilities that can become congested. For instance, service may be refused because the facility is busy, or lines are formed in which customers have to wait until they can be served. After the 9/11 events and the destruction of New Orleans by hurricane Kathrina, many researchers have turned their attention to the location of facilities so as to mitigate such catastrophic effects, or to the search for robust locations in the sense that disruptions do not interrupt their functioning.

Chapter 11 by Berman, Krass and Wang is devoted to the description of these problems. The original contribution by Frank from 1966 is reviewed in detail, and its impact analyzed. Generalizations of Frank's work are presented, and open problems are discussed.

1.3.4.4 Hub Location Problems

Hub location problems occur mainly in transportation and telecommunications networks. A hub is a transfer point at which either traffic from several origins is added up and forwarded to another hub, or disaggregated into several streams that are forwarded to their destination. Airlines use hub airports because they allow tak-

ing advantage of economies of scale, by using high capacity planes in high traffic routes, as opposed to using only small planes between every origin and destination city. Computer networks are frequently organized as a set of small hub-and-spoke networks, joined together by a backbone network.

Naturally, hub location problems have been formulated having in mind the same principles as the remaining location problems. Thus, there are p -hub location problems, median-hub location problems, uncapacitated hub location problems, p -hub-center location problems, and so on. The first formulation of a hub location problem by O’Kelly in 1982 was nonlinear. Later, different linearizations have been proposed. The base model (1.1)–(1.4) needs to be modified to represent an uncapacitated hub location problem (analogous to the uncapacitated facility location problem) as follows. Facilities are now hubs, and customers need to be assigned to hubs for both their incoming and outgoing traffic. We assume each customer is assigned to his closest hub. The traffic between two customers, h_{ij} , starts at a node i , first goes to the assigned hub k , continues from there to another hub m , and finally to the destination node j . The traffic into the opposite direction uses the same path. Both hubs could be the same, i.e., traffic could move through only one hub. We use constraints (1.2)–(1.4), but the objective is now:

$$\text{Min } \sum_i \sum_j h_{ij} \left(\sum_k c_{ik} y_{ik} + \alpha \sum_k \sum_m c_{km} y_{ik} y_{jm} + \sum_m c_{jm} y_{jm} \right) + \sum_j f_j x_j \quad (1.9)$$

where h_{ij} is the traffic between customers or demand nodes i and j , c_{ij} is the unit traffic cost on the arc (i, j) , i.e. on the leg of the trip that goes between nodes i and j , and α is a factor <1 , accounting for economies of scale. Note that the first term in parentheses in the objective function is the cost of sending traffic from the origin to the first hub. The second term is the cost of carrying traffic between hubs, where because of the traffic concentration, it is assumed that the cost is reduced by a factor $(1 - \alpha)$. The third term is the cost of sending the traffic from the second hub to the destination, and the last term is the cost of opening hubs. This objective is nonlinear, but different linear versions have been proposed in the literature.

Chapter 12 of this book, written by Kara and Taner, focuses on hub location problems. O’Kelly’s seminal paper is reviewed and its contribution assessed.

1.4 An Outline of Classic Solution Methods

Most of the problems related to facility location are known to **NP**-hard. It is natural then to look for procedures that can find adequate optimal or approximate solutions to these problems. Many solution methods have been proposed in the literature. It is important to describe some of these methods, either because they have a broader applicability or because they led the way to the discovery of a whole line of solution procedures. This book presents some of the more important techniques.

1.4.1 *Exact Techniques*

The best known exact solution technique for location problems written as integer programming problems is branch-and-bound. The technique was first proposed by Land and Doig in 1960. The first two contributions that apply this technique to the location field are those by Gavett and Plyter in 1966, who solve the quadratic assignment problem, and Efroymson and Ray in 1966, who solved the uncapacitated facility location problem or, as it is frequently called today, the simple plant location problem. These two contributions are reviewed in Chap. 13 by Lowe and Wendell. The authors describe the principles behind the branch-and-bound technique, *viz.*, partitioning the set of solution into smaller subsets, and provide a detailed description of the two seminal papers. They also show other techniques to solve the quadratic assignment problem. The Efroymson and Ray approach to solving the *SPLP* is complemented by other branch-and-bound techniques discovered later. One section of their contribution is devoted to branching strategies.

When the problem has a special structure, as is the case when the space is a tree-shaped network, efficient techniques can be developed to find solutions to the location problems more efficiently. An example of these techniques is the method for finding the 1-median on a tree. In Chap. 14, Kincaid first defines some graph theory concepts, presents a notation for location problems, and reviews three classical contributions: the Harary and Norman article of 1953, where central points on a tree are defined, the paper by Hua Lo-Keng and others in 1962—although known at an earlier date in Chinese—where the first method for location of a 1-median on a tree (and on networks with some cycles) is proposed and proven to be optimal, and the paper by Goldman in 1971, who rediscovered Hua Lo-Keng's findings.

1.4.2 *Heuristic Solution Methods*

Although the maximum size of the problems that can be solved using exact methods has increased in time almost like Moore's law for semiconductors (that states that the number of transistors in an integrated chip doubles every two years), exact methods are still not capable of finding a solution in many realistic cases, in which the number of variables is too large. Heuristics are methods that find a good solution, although optimality is not guaranteed or even sought. Many of these heuristics are built on principles that have been proposed a long time ago, but are still valid. Some of these, for location both in the plane and on networks, were sketched in a single paper by Cooper in 1963, and extended later by other authors. For facility location problems in the plane, a giant step forward was the heuristic for the Weber problem, which was discovered by Weiszfeld in the 1930s.

The heuristics by Cooper, and their formalization by Maranzana and by Teitz and Bart, are the subject of Chap. 15 of this book, by Brimberg and Hodgson. In that chapter, the authors describe in detail the several methods proposed by Cooper

for the continuous minimum problem, as well as the rediscoveries of these methods and their application to the discrete problem, made by Maranzana and by Teitz and Bart. Then, they assess the impact these early discoveries had, their application to other problems and the current status of the heuristic methods for the continuous and discrete p -median problem.

In Chap. 16, Plastria first offers a short history of the Weber problem, a brief listing of different methods that have been proposed to solve it, and a description of its optimality conditions. As a way to introducing the Weiszfeld algorithm, the problem is presented of finding the point at which the weighted squares of the distances to the given points is minimum. Then, the author describes in detail the method found by Weiszfeld, together with its properties and proofs of convergence. After that, he presents a modern view of Weiszfeld's algorithm and the rediscoveries, additions and improvements made by the researchers that followed Weiszfeld. The author then focuses on applications of the algorithm in other areas of knowledge.

Finally, there are heuristics that not only find a solution, but tell the user how far could be this solution from the real optimum. One such heuristic is Lagrangean Relaxation. As Galvão outlines in Chap. 17, this technique was first used in the field of location by Bilde and Krarup in 1967. Later, Diehr, and also Marsten, applied this technique also to location problems. The author then offers a brief review of the technique, surveys the seminal papers by Bilde and Krarup, Diehr and Marsten, and provides an account of the works that followed them.

1.5 Customer Choice and Location Patterns

Most location models that allow customers to choose among a number of facilities assume that customers patronize the closest facility or the facility they can satisfy their demand from most cheaply. However, these simplifying assumptions are not always true in practice. In fact, customers at a demand point can patronize different facilities at different times; or different customers at the same demand point can have different patterns of behavior, when choosing a facility. This phenomenon is the subject of a number of studies, and the problems that include a consideration to it, have been called models with customer choice.

Some studies are oriented to describe in the best possible form the behavior of customers. Some others, focus on the locational patterns and their interaction with markets and customer concentration.

1.5.1 Gravity Models or Spheres of Influence

Joseph and Kuby address gravity modeling in Chap. 18. In other words, they study models that include choice rules other than simple proximity or price. In particular, their focus is on that customers are attracted to facilities according to laws whose

expression resembles that of Newton's laws of physics. The authors describe the contribution by Reilly in 1931, who introduced the so-called *law of retail gravitation* and validated this model with an empirical study. They also review the contribution by Huff in 1963, who proposes a method to determine trade areas for retail, later known as the Huff model. In the chapter, other techniques are also sketched, and applications to location-allocation models are described.

1.5.2 *Voronoi Diagrams*

As mentioned above, Voronoi diagrams are a tool that allows modelers to allocate customers to facilities based on proximity. In Chap. 19, Burkey et al. first review the original work by Thiessen written in 1911. They then formally define Voronoi diagrams and explore some of their properties and extensions. This is followed by a case study of the Triad Region in North Carolina. Here, a number of weighted and unweighted Voronoi diagrams are constructed and compared using concentration indices and statistical tests.

1.5.3 *Central Places*

In the nineteenth and the first half of twentieth centuries, some authors attempted to explain why and how economic activities are located in relation to the markets. These contributions by von Thünen in 1826, Christaller in 1933 and Lösch in 1940, began long strings of work on land use, central places theory and regional planning. Truly, these must be considered as major works in the facility location field. In Chap. 20, Fischer synthesizes these works and assesses their contribution to the field.

Acknowledgements This contribution was in part supported by the National Research Council of Canada under grant number #9160, and by Instituto Milenio "Complex Engineering Systems" under grants ICM P-05-004-F and CONICYT FBO16. This support is gratefully acknowledged.