A FACILITY LOCATION MODEL FOR ELASTIC, SIZE-ATTRACTED, GRAVITY-TYPE DEMAND

João Bigotte[1], and António Pais Antunes[2]

ABSTRACT

Facility location models are mixed-integer optimization models aimed at determining the best location and size for any type of facilities. In this paper we present a facility location model where it is assumed that the demand for service decreases with the distance to the facility where it is provided (elastic demand) and that, when a user requires service, the probability of visiting some facility decreases with the distance to the facility and increases with the size of the facility (gravity-type assignment). It is also assumed that a minimum level of demand is needed in order to locate a facility (below that minimum level the facility would not be economically viable). The objective is to maximize the total served demand. The model is non-linear (in addition to being mixed-integer) and extremely difficult to solve. It addresses a type of problem arising with the location of facilities such as libraries, cinema complexes, etc. Because of its non-linearity the model must be solved by heuristic methods. We tested the classic *Add+Interchange* method and a simulated annealing algorithm. The former often failed to single out optimum or near-optimum solutions, while the latter often provided such solutions.

KEY WORDS

Facility location, optimization modelling, spatial interaction, gravity-type demand, elastic demand.

INTRODUCTION

When dealing with problems involving the location of facilities such as schools, fire stations, libraries or cinema complexes, planners may have to consider a large number of alternatives. If this is the case, the decision-making process will certainly be more efficient if they resort to facility location models (Daskin, 1995; Current *et al.*, 2002; ReVelle and Eiselt, 2005). These models are mixed-integer optimization models aimed at determining the best location and size for any type of facilities according to some objective or objectives (cost minimization, accessibility maximization, etc).

The classic models in the facility location literature (e.g., the *p*-median, the set covering, and the fixed charge models) are also aimed at determining the best user-to-facility assignment, assuming that the demand for service is inelastic (known in advance and

[1] Ph.D. Student, Civil Engineering Department, Pólo 2, Univ. of Coimbra, Coimbra, Portugal, Phone +351239797131, FAX +351239797142, jbigotte@dec.uc.pt

[2] Professor, Civil Engineering Department, Pólo 2, Univ. of Coimbra, Coimbra, Portugal, Phone +351239797139, FAX +351239797142, antunes@dec.uc.pt

independent of the location of facilities). Yet, in many cases, users are free to choose the facilities they want to patronize and demand for service is elastic to some extent. In these cases, the models must be based on assumptions regarding the behaviour of typical users. A frequent assumption is that users will patronize the closest facility (for instance, when locating elementary schools or post offices). However, when locating facilities such as libraries or cinema complexes do these assumptions (inelastic demand and closest assignment) still seem valid? Or, otherwise, does it seem reasonable to assume that demand for service decreases with the distance (or time, or cost) to the facility? Furthermore, which one is a more natural behaviour – do users always patronize the same facility? Or, otherwise, do they split their travel between the available facilities, with the probability of visiting some facility decreasing with the distance to the facility and increasing with the attractiveness of the facility, thus patronizing facilities according to some gravity-type pattern? The size of a facility is often seen as a measure of its attractiveness, because larger facilities normally offer a wider choice of services of the same type.

In this article, we present a facility location model for elastic, size-attracted, gravity-type demand, where a minimum level of demand is needed in order to locate a facility (below that level the facility would not be economically viable). Taken together, these features make the model non-linear (in addition to being mixed-integer) and extremely difficult to solve.

The plan for the article is as follows. The assumptions and the formulation of the model are described in Section 2. Section 3 presents the solution methods (both classic and modern heuristics) developed to solve the model, and a discussion of the methods from the standpoint of solution quality and computing time. Section 4 looks at the model results. Concluding remarks are given in Section 5.

MODEL PRESENTATION

We consider a set of N discrete population centres, each one of them being a candidate site for locating a facility. The potential number of facility users (or potential demand) in centre $j \in N$ is given by u_j. The summation of the potential demand over all centres is represented by U. The travel distance (or time, or cost) between centre j and centre k is designated by d_{jk}.

We address the problem of determining the best location ($y_k = 1$ if a facility is located at centre k, $y_k = 0$ otherwise) and size (z_k), for the facilities assuming that:

1. The demand from centre j for the services of a facility located at centre k (h_{jk}) is a decreasing function of the travel distance (Perl and Ho, 1990; Antunes *et al*., 2004):

$$h_{jk} = u_j \cdot \left(1 - \alpha \cdot d_{jk}^{\beta}\right)$$

   where $\alpha, \beta > 0$ are calibration parameters.

2. Users patronize facilities according to a gravity-type pattern. Following the spatial interaction literature, the utility for a user in centre j of a facility at centre k (w_{jk}) is given by the ratio of the attractiveness A_k of the facility to a function of the travel distance (O'Kelly, 1987; Berman and Krass, 2002):

$$w_{jk} = \frac{A_k}{e^{\gamma \cdot d_{jk}}}$$

where $\gamma>0$ is a calibration parameter.

The probability of users in centre j travelling to centre k to obtain service (p_{jk}) equals the relative utility of facility at centre k compared to other available facilities:

$$p_{jk} = \frac{w_{jk}}{\sum_{i \in N} w_{ji}}$$

This can also be interpreted as the proportion of users in centre j that obtain service from a facility located at centre k.

All in all, the probability/proportion of users visiting some facility decreases with the distance to the facility and increases with the attractiveness of the facility. As stated before, in our model the attractiveness of a facility is measured by its size.

3. The size of a facility is proportional to the number of users obtaining service from that facility. In order to be economically viable, a facility can only be located if the number of users it serves is equal to or above a minimum level $z_{\min}$.
4. The unit facility costs are assumed to be constant for facilities larger than $z_{\min}$.
5. The objective is to maximize the total served demand (that is, to maximize the number of users obtaining service from all facilities).

Given these assumptions, our model can be formulated as follows:

$$\max Z = \sum_{k \in N} z_k \tag{1}$$

subject to

$$z_k \leq U \cdot y_k \quad , \quad \forall\, k \in N \tag{2}$$

$$z_k \geq z_{\min} \cdot y_k \quad , \quad \forall\, k \in N \tag{3}$$

$$z_k = \sum_{j \in N} h_{jk} \cdot \frac{z_k \cdot e^{-\gamma \cdot d_{jk}}}{\sum_{i \in N} z_i \cdot e^{-\gamma \cdot d_{ji}}} \quad , \quad \forall\, k \in N \tag{4}$$

$$y_k \in \{0,1\} \, , \, z_k \geq 0 \, , \, \forall\, k \in N \tag{5}$$

The objective function (1) maximizes the total size of facilities. As stated before, we are assuming that the size of a facility is proportional to the number of users obtaining service from that facility. Therefore, the objective function maximizes the total served demand.

Constraint set (2) ensures that users will only be served at centre k if a facility is located at that centre, otherwise, when $y_k = 0$ then $z_k = 0$. Constraint set (3) requires that the size of a facility located at centre k exceeds the minimum level $z_{\min}$. Constraint set (4) specifies that the size of a facility located at centre k is proportional to the total number of users obtaining service from that facility. This is a non-linear constraint (due to the division of a z_k decision variable by the summation over k of z_k decision variables). Constraint set (5) reflects the binary nature of the location decisions and the non-negative nature of the size of a facility.

MODEL SOLVING

INTRODUCTION

In this section we present the approach adopted to solve the model. Because of its non-linearity, the model can not be solved to optimality by exact methods. Consequently, the only manner for one to be certain of finding an optimum solution is through complete enumeration. Obviously, this approach is unsuitable for large- and even moderate-size real-world applications because it would take an enormous amount of computing time to list and evaluate all possible solutions (starting from locating a single facility to locating the maximum number of facilities allowed by the minimum size constraint). Therefore, we tested some heuristic methods, both classic and modern (Gendreau and Potvin, 2005). Heuristic methods are approximate methods for solving optimization models. They usually involve a short computing time but there is no guarantee that the resulting solution is the optimum solution (they only guarantee to find a local optimum, not the global optimum). When deciding which heuristic method to use, it is necessary to ponder on both the computing time it takes and the quality of the solution it provides (that is, how close to the optimum the solution is). A large set of instances was created for testing the methods. We applied the classic *Add+Interchange* method and a simulated annealing algorithm (Dowsland, 1993; Antunes and Peeters, 2001). The former often failed to single out optimum or near-optimum solutions, while the latter often provided such solutions. Whichever the method, an important issue when solving the model is: once we have a solution vector for the location variables how can we compute the z_k variables? This can be done by reaching equilibrium in Constraint (4). We describe an iterative approach for doing it.

TEST INSTANCES

The solution methods were tested in a large set of instances. Those instances regard a region of 100×100 length units, and differ on the number of centres (recall that centres coincide with sites), the location of the centres, and the number of potential users at each centre. We considered fifty 20-centre instances and twenty 50-centre instances. The centre coordinates were created by generating random numbers uniformly distributed over [0, 100]. The number of potential users at each centre was generated in the same way, but this time considering the interval [10, 100]. The minimum level of demand was assumed to be 200.

Regarding the model parameters, we considered a linear decay function with $\alpha = 1/\sqrt{20000}$ (the inverse of the diagonal of the region under study) and $\beta = 1$, and a spatial interaction function with $\gamma = 0.05$ (a current value for Portugal).

HEURISTIC APPROACH

Our approach on solving the model presented above was to test at first the classic *Add+Interchange* method (even though we were not expecting to obtain very good solutions using this method we thought it would be useful for getting some insights). The main reason was that this method is fairly simple to program and known for providing rather good solutions in a short time. In addition, classic heuristic methods such as *Add+Interchange* have been applied successfully in many location models.

The *Add* heuristic allows building a solution from scratch. It starts by evaluating which site is the best, among all candidate sites, to locate a single facility (in our case, the best site is the one that entails the most served demand). Then, in successive iterations, it locates a facility at a time, by choosing among the closed sites, the one which allows the best feasible increase in the objective function value, until no further increase is possible (note that a site is called closed when there is no facility located there). The *Interchange* heuristic starts with the *Add* solution and, in successive iterations, considers moving a facility from any open site to any closed site, choosing the combination of site interchange which allows the best feasible increase in the objective function value, once again until no further increase is possible.

The simulated annealing (SA) algorithm is a modern improvement heuristic based on the physical process of cooling a material to low-energy states. This algorithm begins with a good (or, at least feasible) solution and, in successive iterations, accepts (or not) solutions generated in the neighbourhood of the current solution according to the Metropolis criterion (all solution improvements are accepted, solution deteriorations may be accepted with some probability depending on a parameter called temperature) until no further increase in the objective function value is possible.

The algorithm we developed uses the *Add* solution to begin the SA procedure. Then, in each iteration, it can either locate one more facility at any closed site (*Add* neighbourhood), or remove an existing facility from an open site (*Drop* neighbourhood), or move a facility from an open to a closed site (*Interchange* neighbourhood). There is an equal probability of visiting any one of these neighbourhoods. The choice of the neighbourhood to visit is made through the generation of random numbers. After the SA procedure, an exhaustive neighbourhood search is performed through *Drop*, *Add*, and *Interchange* procedures (Figure 1). The algorithm uses a range of parameters, designated as IT, TR, ME, MI and PW, which must be calibrated. The first parameter is necessary to determine the initial temperature – we assumed a certain probability (IT) of accepting a 10% deterioration in the starting solution (*Add* solution). The TR parameter gives the temperature reduction rate in each iteration. ME stands for the maximum number of solutions evaluated in each iteration without improving the objective function value. The stopping criterion is the maximum number of iterations without improving the objective function value (MI). We assumed a PW weight to penalize solutions that violated the minimum level of demand assumption (Constraints 3). The calibration of parameters was performed in two phases. In the first phase, we assumed a base value for each one of the parameters and then changed a single parameter at a time (either a decrease in the parameter value or an increase). In the second phase, the best set of parameters obtained from phase one was improved by decreasing each parameter at a time.

We performed a total of 50 runs for each set of parameters (ten 20-centre instances times five seeds to generate the random numbers). The solutions were analysed regarding the number of optimum solutions found, the average difference to the optimum solution, the maximum difference to the optimum solution, the average computing time, and the maximum computing time. The results are shown in Table 1. The best set of parameters obtained from the second phase was: IT = 0.3; TR = 0.9; ME = 20; MI = 5; and PW = 100.

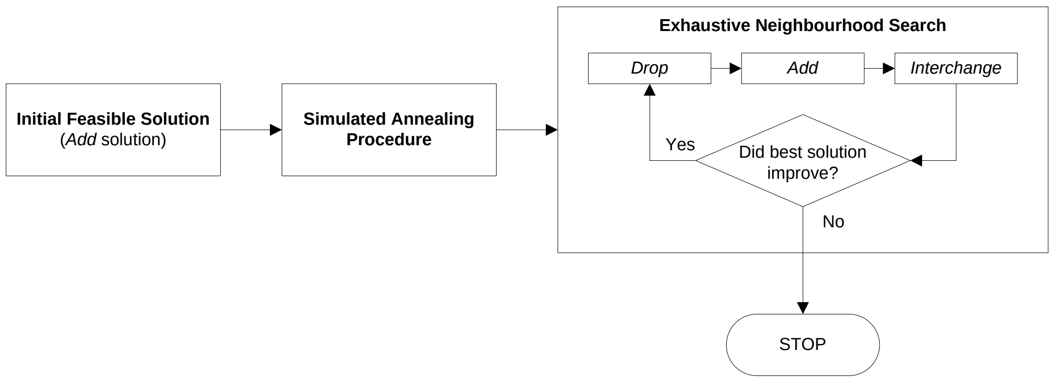


Figure 1: Flowchart of the Simulated Annealing Algorithm

Table 1: Calibration of Simulated Annealing Parameters

Phase	Parameter					Number of optimum solutions	Average difference to optimum	Maximum difference to optimum	Average computing time (sec.)	Maximum computing time (sec.)
	IT	TR	ME	MI	PW					
1	0.3	0.7	20	10	100	48	1.0%	1.7%	19	101
	0.1	0.7	20	10	100	44	1.3%	1.7%	7	22
	0.5	0.7	20	10	100	48	0.3%	0.3%	60	173
	0.3	0.5	20	10	100	44	1.2%	1.7%	14	99
	0.3	**0.9**	**20**	**10**	**100**	50	0.0%	0.0%	43	166
	0.3	0.7	10	10	100	33	1.0%	1.7%	3	6
	0.3	0.7	30	10	100	50	0.0%	0.0%	150	641
	0.3	0.7	20	5	100	48	1.0%	1.7%	18	100
	0.3	0.7	20	15	100	48	1.0%	1.7%	20	101
	0.3	0.7	20	10	50	48	1.0%	1.7%	20	104
	0.3	0.7	20	10	150	48	1.0%	1.7%	19	101
2	0.1	0.9	20	10	100	44	0.7%	1.3%	14	37
	0.3	0.9	10	10	100	35	0.8%	1.7%	5	14
	0.3	**0.9**	**20**	**5**	**100**	50	0.0%	0.0%	37	148
	0.3	0.9	20	10	50	50	0.0%	0.0%	44	165

In order to analyse the *Add+Interchange* method and the SA algorithm from the standpoint of solution quality and computing time, we performed a complete enumeration over each one of the 20-centre instances.

As expected, the *Add+Interchange* method did not prove efficient for solving the model – for the set of 20-centre instances, it provided only 16 optimum solutions and, in the remaining 34 instances, the maximum difference to the optimum solution was 10.2% (Table 2). Nevertheless, it showed extremely fast, with an average computing time of one second and a maximum of three seconds (while the complete enumeration took an average of nearly five minutes and a maximum of 25 minutes). The computations were made with a Pentium M processor, running at 1.73 GHz. In some cases, the *Add+Interchange* solution located the same number of facilities as in the optimum solution. When this occurred, the heuristic solution was optimum or close to optimum. However, in most of the cases that did not happen (it located either a larger or a smaller number of facilities than in the optimum solution), and so the *Add+Interchange* solution was quite different from the optimum solution. On the other hand, the SA algorithm provided optimum or very close to optimum solutions – it was capable of finding 42 optimum solutions for an average computing time of 27 seconds and a maximum of two minutes. In the remaining instances, the maximum difference to optimum was only 1.9%.

Table 2: Summary of Results for the 20-Centre Instances

Method	Number of optimum solutions	Average difference to optimum	Maximum difference to optimum	Average computing time (sec.)	Maximum computing time (sec.)
Complete Enumeration	50	-	-	277	1524
Add+Interchange	16	2.2%	10.2%	1	3
Simulated Annealing	42	0.7%	1.9%	27	129

The SA algorithm also performed globally better than the *Add+Interchange* method for the set of twenty 50-centre instances. It was not possible to perform a complete enumeration for these large instances. Consequently, it is impossible to know whether the heuristics solutions are optimum and, if not, how close to optimum they are. Therefore, the heuristic methods can only be compared with each other. The SA algorithm provided 14 better solutions than the *Add+Interchange* method (Table 3). The opposite happened only 3 times, and in the remaining 3 instances the methods provided equal solutions. In the cases where the SA algorithm performed better, the average difference of the *Add+Interchange* to the SA solutions was 0.4% and the maximum difference was 1.5%. The computing time increased considerably for the 50-centre instances. The *Add+Interchange* method took, on average, 63 seconds to solve the model, and a maximum computing time of 97 seconds, while the SA algorithm took an average computing time of about 40 minutes, and a maximum of one hour and fifteen minutes.

Table 3: Summary of Results for the 50-Centre Instances

Method	Number of best solutions	Average difference to other method	Maximum difference to other method	Average computing time (sec.)	Maximum computing time (sec.)
Add+Interchange	3	0.4%	1.5%	63	97
Simulated Annealing	14	0.1%	0.2%	2444	4595

USER-TO-FACILITY ASSIGNMENT

The user-to-facility assignment is the key innovation in our model. We assume that the demand for service decreases with the distance to the facility where it is provided (elastic demand) and that, when a user requires service, the probability of visiting some facility decreases with the distance to the facility and increases with the size of the facility (gravity-type assignment). Furthermore, we consider that the size of a facility is proportional to the total number of users obtaining service at that specific facility. In conclusion, the size of a facility depends on the number of users visiting the facility, and the number of users visiting a facility depends on the size of the facility. That is, z_k is a function of z_k: $z_k = f(z_k)$, as stated in Constraints (4). It can be proved that there is always one and only one solution for the z_k variables, with each z_k non-negative.

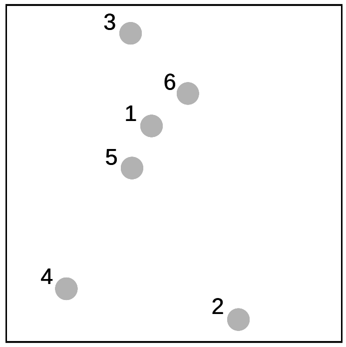
The equations in Constraints (4) must be solved iteratively (Figure 2). Assuming that the facilities are all of the same size (z_k), we first calculate for every j and k the proportion of users in centre j that obtain service from a facility located at centre k (p_{jk}). Then, in each iteration, we re-calculate the size of facilities (z_k*) using the previous p_{jk} and determine the new proportions (p_{jk}*). This process loops until equilibrium is reached. We assume to have reached equilibrium when the absolute value of the difference between p_{jk}* and p_{jk} is smaller than 0.01.

```
Select initial size z_k for facilities (same size)
Calculate initial p_jk
Repeat
        calculate z_k* (using p_jk)
        calculate p_jk*
        Δ = |p_jk*-p_jk|
        p_jk = p_jk*
Until Δ < 0.01
z_k=z_k*
```

Figure 2: Algorithm for the Computation of Equilibrium Assignment

MODEL RESULTS

In this section we illustrate the type of results that can be obtained through the application of the model. We consider a region with 6 population centres (Figure 3), model parameters $\alpha = 1/\sqrt{20000}$, $\beta = 1$, $\gamma = 0.05$, and a minimum size of 100.



Centre	Population
1	52
2	94
3	93
4	37
5	56
6	73
Total	405

Figure 3: Example Input Data

The optimum solution consists of locating two facilities, at centres 2 and 6 (Figure 4). In this solution, 86.7% of the potential demand is served (Table 4). It would be possible to locate one more facility, but the optimum solution for locating three facilities would only serve 79.6% of the potential demand.

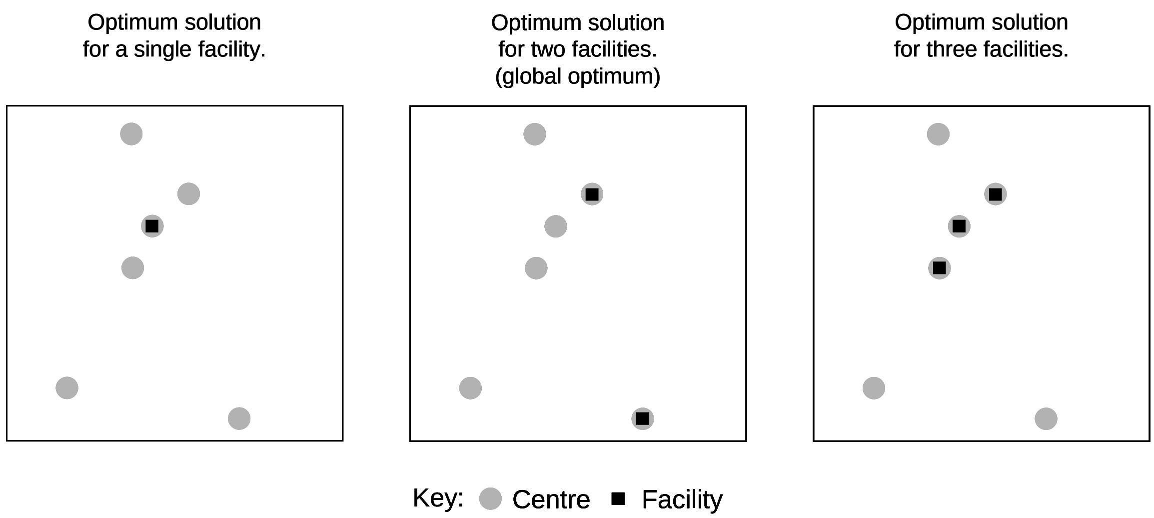


Figure 4 – Optimum solution and best locations for a single facility and for three facilities

Table 4: Solutions results

Number of facilities	Served demand	
	Total	Ratio
1	317	78.3%
2	351	86.7%
3	322	79.6%

This example shows how the model globally works: (1) the best site to locate a single facility is usually a very central one; (2) as more facilities are added, the open sites tend to be scattered over the region, what happens until the number of facilities comes close to the maximum number of possible facilities; (3) from that, the location of facilities tends to be more central again. While in the classic facility location models seeking the maximization of served demand the more facilities are located the more demand is served, in our model this is not necessarily true. For some instances, it is better to locate fewer facilities than the maximum possible number. The reason for this to occur is because the user-to-facility assignment is not straightforward. Demand depends both on the distance to the facility and on the size of the facility. Adding one more facility affects the size of the pre-existing ones. This means that the new facility may increase the demand from surrounding centres but, at the same time, it may have a global negative impact on the total demand served.

CONCLUSION

In this paper we presented a facility location model for elastic, size-attracted, gravity-type demand. The model assumes that the demand for service decreases with the distance to the facility where it is provided (elastic demand) and that, when a user requires service, the probability of visiting some facility decreases with the distance to the facility and increases with the size of the facility (gravity-type assignment). In the model, the attractiveness of a facility is measured by the size of the facility. This feature makes the model non-linear (in addition to being mixed-integer) and extremely difficult to solve. We also considered that a minimum level of demand is needed in order to locate a facility (below that minimum level the facility would not be economically viable). The objective is to maximize the total served demand. The model addresses a type of problem arising with the location of facilities such as libraries, cinema complexes, etc.

Because of its non-linearity the model must be solved by heuristic (approximate) methods. We tested the classic *Add+Interchange* method and a simulated annealing algorithm. The former often failed to single out optimum or near-optimum solutions, while the latter often provided such solutions.

We showed how the model globally works with a small example. An extremely important issue when solving the model is to find the optimum number of facilities to locate. In most facility location models seeking the maximization of served demand, the more facilities are located the more demand is served. In our model this is not necessarily true because the demand depends both on the distance to the facility and on the size of the facility. That is, the opening of one more facility may increase the demand from surrounding centres but, at the same time, it may have a global negative impact on the total served demand.

ACKNOWLEDGMENTS

The authors thankfully acknowledge support received from Fundação para a Ciência e a Tecnologia through grants POCTI/AUR/32599/2000 and SFRH/BD/18216/2004.

REFERENCES

Antunes, A., Gomes, C., and Ribeiro, A. (2004). "An Accessibility-maximization Approach to Public Facility Planning under Distance-sensitive Demand." *Studies in Regional & Urban Planning* 10 109-117.

Antunes, A. and Peeters, D. (2001). "On solving complex multi-period location models using simulated annealing." *European Journal of Operation Research* 130 (1) 190-201.

Berman, O. and Krass, D. (2002). "Locating Multiple Competitive Facilities: Spatial Interaction Models with Variable Expenditures." *Annals of Operations Research* 111 (1-4) 197-225.

Current, J., Daskin, M., and Schilling, D. (2002). "Discrete network location models." In Drezner, Z. and Hamacher, H. (Eds.), *Facility Location: Applications and Theory*, Springer-Verlag, Berlin (Germany), 81-118.

Daskin, M. (1995). *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley Interscience, New York (USA), pp. 498.

Dowsland, K. (1993). "Simulated Annealing." In Reeves, C. (Ed.), *Modern Heuristic Techniques for Combinatorial Problems*. John Wiley & Sons, New York (USA), 20-69.

Gendreau, M. and Potvin, J. (2005). "Metaheuristics in Combinatorial Optimization." *Annals of Operations Research* 140 (1) 189-213.

O'Kelly, M. (1987). "Spatial interaction based location-allocation models." In Ghosh, A. and Rushton, G. (Eds.), *Spatial Analysis and Location-Allocation Models*, Van Nostrand Reinhold, New York (USA), 302-326.

Perl, J. and Ho, R. (1990). "Public Facility Location under Elastic Demand." *Transportation Science* 24 (2) 117-136.

ReVelle, C. and Eiselt, H. (2005). "Location Analysis: A Synthesis and Survey." *European Journal of Operational Research* 165 (1) 1-19.