Theory and Methodology

On solving complex multi-period location models using simulated annealing

António Antunes [a,*], Dominique Peeters [b]

[a] *Departamento de Engenharia Civil, Universidade de Coimbra, 3030-290 Coimbra, Portugal*
[b] *CORE & Unité de Géographie, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium*

Abstract

This paper describes a study aimed at evaluating the capabilities of simulated annealing in dealing with complex, real-world multi-period location problems raised by school network planning in Portugal. The problems were formulated as mixed-integer linear optimization models. The models allow for facility closure or size reduction besides facility opening and size expansion, with sizes possibly limited to a set of pre-defined standards. They assume facility costs to be divided into a fix component and two variable components, respectively dependent on facility size and facility attendance. Results obtained through the study indicate that simulated annealing can be a useful tool for solving these kinds of models. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Location; Capacity expansion; Simulated annealing

1. Introduction

Location models are important tools for regional, urban, and sectorial planners who participate in decision-making processes regarding the location, size and catchment area of public facilities (like schools, sanitary landfills, libraries, swimming-pools and hospitals). It is not therefore surprising that a vast body of literature has accumulated and keeps growing on this subject

(Hansen et al., 1987; Daskin, 1995; Drezner, 1995).

The best-known (discrete) location models are probably the uncapacitated/capacitated facility location problem (UFLP/CFLP), also named simple/capacitated plant location problem (SPLP/CPLP), and the *p*-median problem. The former seeks a minimum-cost solution for the location of facilities within a given set of sites to satisfy the demands of a given set of centers; the latter seeks a maximum-accessibility solution with a given number of facilities. Together with these relatively simple (though NP-hard) problems, many others, sometimes much more complex, have been studied, where aspects like the dynamic or uncertain

* Corresponding author. Tel.: +351-239797139; fax: +351-239797146.
  E-mail address: antunes@dec.uc.pt (A. Antunes).

nature of the decision context, as well as design and organizational characteristics of particular facilities, have been taken into account.

Location models are mixed-integer linear programming models that, like any other models of this kind, can be solved with three types of algorithms: general (exact) algorithms, specialized (exact) algorithms, and (specialized) heuristic algorithms. General algorithms, like 'branch-and-bound', can be used to solve to exact optimality any (linear) facility location model, but they only are efficient in handling small problems (though, as shown by Ribeiro and Antunes (2000) much larger ones than some years ago). Good commercial software packages, like GAMS/CPLEX (Gams Development, 1996), XPRESS-MP (Dash Associates, 1997) or MINTO (Nemhauser et al., 1994), are now available for this purpose. Specialized algorithms can be used to solve large and complex problems of the particular type under consideration, exploiting their distinct mathematical structure. However, the best ones reported in the literature will normally require arduous programming work before they can be applied. Furthermore, it should be stressed here that it is often quite difficult to adjust a specialized algorithm developed for a given problem, to a slightly different problem. Heuristic algorithms also serve to solve large and complex problems, and the corresponding software is much easier to develop. However, they are not exact, in the sense that they only guarantee a local optimum, not necessarily close to the global optimum.

The whole field of location modeling started to grow in the early 60s, led by authors like Kuehn and Hamburger (1963), Feldman et al. (1966), and Teitz and Bart (1968). These authors proposed heuristics of a particular type, known as local search heuristics. Another kind of heuristics that may be used to solve location models became popular only recently, in the 80s, under the name of modern (or threshold) search heuristics. For a survey, see Reeves (1993) or Aarts and Leenstra (1997). Simulated annealing, the method dealt with here, belongs to this family of heuristics, together with tabu search, genetic algorithms and neural networks.

This paper describes a study aimed at evaluating the capabilities of simulated annealing in dealing with complex, real-world multi-period location problems raised by school network planning in Portugal. The second section contains a statement of the problem to be solved, and its formulation as a mathematical model. The third presents a detailed explanation about the approach adopted to solve it, before it was decided to apply simulated annealing. Then follows a section describing the criteria used to generate the partly random problems on which the solution methods were tested. The fifth section presents the general principles of simulated annealing and the main features of their implementation to our model. The sixth gives essential information about the results obtained in four real-world applications of the model. The final section summarizes the main conclusions of the study and indicates some directions for future research.

2. Location model

In 1986, the Portuguese Government decided to extend elementary education from 6 to 9 years. Such a decision could not be put into effect without a very significant expansion of infrastructure, and a program to know where, with what size and when should schools be built. Of course, in formulating the program, attention had to be paid to the fact that, in the future, given the 'fertility crisis' that is affecting Portugal, as much as the whole Western World, education demand tends to be drastically reduced.

Research on the kind of dynamic problems we were interested in was initiated by Roodman and Schwartz (1975, 1977), and has received several relevant contributions since then, particularly from authors like Van Roy and Erlenkotter (1982), Fong and Srinivasan (1981), Jacobsen (1990), and Shulman (1991). Further work specifically addressing school location problems has been done by Viegas (1987) and Greenleaf and Harrison (1987). However, to the best of our knowledge, the literature does not report work done with models accounting for facility closure and capacity reduction in addition to facility opening and

capacity expansion, particularly when capacity and assignment decisions and costs are dissociated. In circumstances such as those found in Portugal (capacity shortage with decreasing demand), these features are of crucial importance.

This led us to develop a multi-period model aimed at finding the minimum discounted cost solution for the evolution of a set of facilities through a given planning horizon in order to meet the demands for the corresponding services. The decisions to be made consist of opening new facilities, and expanding, reducing or closing existing facilities. The model is as follows:

$$\min C = \sum_{j \in \mathbf{J}} \sum_{k \in \mathbf{K}} \sum_{m \in \mathbf{M}} cvx_{jkm} x_{jkm}$$
$$+ \sum_{k \in \mathbf{K}} \sum_{m \in \mathbf{M}} cf_{km} \left(y_{km}^+ + y_{km}^- \right)$$
$$+ \sum_{k \in \mathbf{K}} \sum_{m \in \mathbf{M}} cvz_{km} \left(z_{km}^+ - z_{km}^- \right), \qquad (1)$$

subject to

$$\sum_{k \in \mathbf{K}} x_{jkm} = 1, \quad \forall j \in \mathbf{J}, \ m \in \mathbf{M}, \qquad (2)$$

$$\sum_{j \in \mathbf{J}} u_{jm} x_{jkm} \leqslant q_{k0} \left(y_{km}^+ + y_{km}^- \right) + q \left(z_{km}^+ - z_{km}^- \right),$$
$$\forall k \in \mathbf{K}, \ m \in \mathbf{M}, \qquad (3)$$

$$\left(q_{\min_k} - q_{k0} \right) y_{km}^+ \leqslant q \cdot z_{km}^+ \leqslant \left(q_{\max_k} - q_{k0} \right) y_{km}^+,$$
$$\forall k \in \mathbf{K}, \ m \in \mathbf{M}, \qquad (4a)$$

$$q \cdot z_{km}^- \leqslant \left(q_{k0} - q_{\min_k} \right) y_{km}^-, \quad \forall k \in \mathbf{K}^1, \ m \in \mathbf{M}, \qquad (4b)$$

$$y_{km}^+ + y_{km}^- \leqslant 1, \quad \forall k \in \mathbf{K}, \ m \in \mathbf{M}, \qquad (5)$$

$$y_{k,m-1}^+ \leqslant y_{km}^+, \quad \forall k \in \mathbf{K}^0, \ m \in \mathbf{M} \setminus \{1\}, \qquad (6)$$

$$y_{km}^+ \leqslant y_{k,m-1}^+, \quad \forall k \in \mathbf{K}^1, \ m \in \mathbf{M} \setminus \{1\}, \qquad (7a)$$

$$y_{km}^- \leqslant y_{k,m-1}^-, \quad \forall k \in \mathbf{K}^1, \ m \in \mathbf{M} \setminus \{1\}, \qquad (7b)$$

$$z_{k,m-1}^+ \leqslant z_{km}^+, \quad \forall k \in K, \ m \in \mathbf{M} \setminus \{1\}, \qquad (8)$$

$$z_{k,m-1}^- \leqslant z_{km}^-, \quad \forall k \in \mathbf{K}^1, \ m \in \mathbf{M} \setminus \{1\}, \qquad (9)$$

$$\sum_{k \in K^0} ef_{km} \left(y_{km}^+ - y_{k,m-1}^+ \right)$$
$$+ \sum_{k \in \mathbf{K}} evz_{km} \left(z_{km}^+ - z_{k,m-1}^+ \right) \leqslant b_m \quad \forall m \in \mathbf{M}, \qquad (10)$$

$$x_{jkm} \leqslant 0, \quad \forall j \in \mathbf{J}, \ k \in \mathbf{K}, \ m \in \mathbf{M},$$
$$y_{km}^+, y_{km}^- \in \{0, 1\}, \quad k \in \mathbf{K}, \ m \in \mathbf{M},$$
$$z_{km}^+, z_{km}^- \geqslant 0 \quad \text{and integer } \forall k \in \mathbf{K}, \ m \in \mathbf{M},$$

where $\mathbf{J}$ is the set of centers $(j = 1, \ldots, J)$, $\mathbf{K}$ the set of sites $(k = 1, \ldots, K)$, $\mathbf{K}^0$ the set of initially closed sites, $\mathbf{K}^1$ the set of initially opened sites, $\mathbf{M}$ the set of periods $(m = 1, \ldots, M)$, x_{jkm} the fraction of users from center j assigned to a facility located at site k in period m, $y_{km}^+ = 1$ if an expanding facility is located at site k in period m; $y_{km}^+ = 0$ otherwise, $y_{km}^- = 1$ if a reducing facility is located at site k in period m; $y_{km}^- = 0$ otherwise, z_{km}^+ the accumulated capacity expansion of the facility located at site k up to period m, z_{km}^- the accumulated capacity reduction of the facility located at site k up to period m, cvx_{jkm} the discounted attendance-variable cost of a facility located at site k in period m per user of center j (includes transport costs), cf_{km} the discounted fixed cost of a facility located at site k in period m (for initially opened facilities it includes the cost associated with a capacity of q_{k0}), cvz_{km} the discounted capacity-variable cost increase (or decrease) of expanding (or reducing) a facility located at site k in period m, u_{jm} the number of users located at center j in period m, q_{k0} the initial capacity of the facility located at site k, q the capacity of a module, $q_{\max_k}$ ($\geqslant q_{k0}$) the maximum capacity of a facility located at site k, $q_{\min_k}$ ($\leqslant q_{k0}$) the minimum capacity of a facility located at site k, ef_{km} the fixed setup expenditure for a facility to be located (or expanded) at site k in period m, evz_{km} the capacity-variable setup expenditure for a facility to be located at site k in period m, and b_m is the budget for period m.

This is a complicated mixed-integer linear programming model, to which we gave a name in the tradition of location analysis: 'dynamic modular capacitated facility location problem' (DMCFLP) model.

The geographical landscape assumed by the model is the one normally used in (discrete) location modeling. Demand for facilities is assumed to be concentrated in points named centers, which may represent regions, municipalities, towns or neighborhoods. Supply of facilities is assumed to be possible at points named sites, which represent either any one of the above geographical entities or

determined plots of land. The facilities are assumed to be composed of a fixed part and a variable number of modules of given fixed capacity. For schools, the module may be either one classroom or a group of classrooms.

Function (1) expresses the objective of minimizing the total discounted (socio-economic) costs of a set of facilities. Facility costs are divided into three parts: fixed costs, capacity-variable costs proportional to the number of modules, and attendance-variable costs proportional to the number of users, a significant part of which will normally consist of transport costs. With regard to existing facilities, initial capacity-variable costs are taken as fixed costs, which explains the minus sign applied to variables z_{km}^-.

Constraints (2) ensure that, in each period, the demand of any center will be met by some facility or facilities. Constraints (3) guarantee that the capacity of each facility in each period, resulting from adding the initial capacity to the accumulated capacity expansion occurred up to the period, will be large enough to meet the demand assigned to it. Constraints (4a) and (4b) ensure that maximum and minimum capacity limits, defined to avoid or exploit technical and economic scale advantages or disadvantages, will be taken into account at both expanding or reducing facilities. Constraints (5) guarantee that capacities may either be expanded or reduced, but not both. Constraints (6) ensure that facilities opened at initially closed sites, once opened, will remain open. Constraints (7a) and (7b) guarantee that facilities closed at initially opened sites, once closed, will remain closed. Constraints (8) ensure that the capacity of expanding facilities will never decrease. Constraints (9) guarantee that the capacity of reducing facilities will never increase. Finally, constraints (10) ensure that the budget constraints defined for each period will be satisfied. In these constraints, the variables z_{k0}^+ should be set equal to q_{k0}, and the variables y_{k0}^+ should be set equal to 0.

3. Solution approach

The model introduced in the previous section is quite a difficult one, something that we readily

understood after we tried, and failed, to solve some partly random 10-center × 10-site × 3-period problems using packages like SCICONIC, XPRESS-MP and CPLEX (early versions). Section 4 describes how the problems were generated.

Our first attempt to overcome these difficulties involved a 'myopic' approach, consisting of solving the first-period problem without taking into account future-period demands, and then solving the second-period problem given the optimum facility set identified for the first-period problem, and so forth. The results obtained using this approach, compared with those given by a 'panoramic' approach (within which short-term planning decisions would also reflect the long-term expected evolution for demand), were quite good on many occasions, though not in the presence of severe capacity shortage and rapidly decreasing demand. (Antunes, 1994, pp. 57–68.) This was precisely the situation found in many Portuguese regions when it was decided to extend elementary education.

The idea of using simulated annealing emerged when we realized that it would be relatively simple to apply this method to our model, and once we understood that our efforts to develop dual-based and lagrangean relaxation specialized methods were unlikely to be successful. Of course, we knew that tests on complex, large-constrained models had sometimes failed to meet expectations, but there were not many other promising options open to us.

Before applying simulated annealing to the model, we decided to try it on a simple location model, UFLP (Krarup and Pruzan, 1983; Cornuéjols et al., 1990), to see how it would compare with ADD and ADD + INTERCHANGE, two fast, well-known local search heuristics. ADD assumes all the sites to be initially closed and, in successive iterations, opens the facility which allows the best decrease in costs, until no further cost reduction is possible. INTERCHANGE works upon ADD and, in successive iterations, chooses the capacity transfer, from open to closed sites, which allows the best decrease in costs, once again until no further cost reduction is possible.

The results obtained with a good choice of annealing parameters (this concept is clarified in

Section 5) in solving a representative sample of partly random problems are summarized in Tables 1 and 2. The comparison with ADD was encouraging about the prospects of annealing's capabilities, because the solutions found were clearly better. But it was the comparison with ADD + INTERCHANGE, often regarded as being a good heuristic, that finally convinced us of its possibilities. In fact, for a set of small 20-center × 20-site problems, the annealing algorithm (called ANNEAL in abbreviated form) produced worse results and took longer on average than ADD + INTERCHANGE. But large 80-center × 80-site problems were solved better and faster, indicating the aptness of simulated annealing for dealing with problem size.

4. Random problems

As mentioned in the previous section, alternative methods were tested using partly random problems, i.e., problems including both deterministic and stochastic elements. There was no other way of building a representative sample, as multi-period location problems are relatively rare in the literature.

The deterministic part of problem data included the territory geometry (a square of 100 × 100 length units), the number of centers, sites and pe-

riods, and classroom capacity (25 students). It also included cost data. Setup costs were assumed to consist of a fixed part and a variable part, respectively, equal to 10 monetary units (μ) and 10 μ per classroom. Operating costs were also assumed to consist of a fixed part and a variable part, respectively equal to 10 μ per planning period of 5 years and 17 μ per classroom per period. Transport costs were taken to be 0.08 μ per period per student. A discount rate of 40% per period (equivalent to 7% per year) was taken to convert setup costs to a per period basis. These particular figures were the same as those to be used later in solving real-world problems of school network planning.

The stochastic part of problem data included the coordinates of centers and sites, the number of users in each center and its evolution through time, the location and capacity of existing facilities, and the maximum capacity of new facilities. The coordinates of centers and sites were assumed to be random integers uniformly distributed between 0 (zero) and 100. Thus, centers and sites were located inside or on the border of a 100 × 100 square, with integer coordinates. The users in each center were assumed to be a random integer uniformly distributed between 0 (zero) and 200. This number was taken to change through time according to the combined effect of a random global growth rate (varying from problem to problem)

Table 1
Comparison between ADD and ANNEAL results for UFLP problems

Problem size (centers × sizes)	ANNEAL better than ADD	ADD equal to ANNEAL	ADD better than ANNEAL	Computing time ANNEAL/ADD
20 × 20	24	89	12	6.31
40 × 40	58	50	17	3.27
80 × 80	103	4	17	1.87

Table 2
Comparison between ADD + INTERCHANGE and ANNEAL results for UFLP problems

Problem size (centers × sizes)	ANNEAL better than ADD+	ADD + better than ANNEAL	ADD+ equal to ANNEAL	Computing time ANNEAL/ADD+
20 × 20	5	106	14	1.87
40 × 40	14	84	27	1.06
80 × 80	50	38	37	0.64

and a random local growth rate (varying from center to center). Both rates were taken to follow a uniform distribution between −10% and 10% per period. Existing facilities were assumed to be located near the largest centers with probability proportional to center size. The capacity of these facilities was taken to be equal to a randomly selected number of classrooms, such that they would be able to accommodate a maximum of 4/3 and a minimum of 2/3 of the users living in the nearest center. The maximum capacity of new facilities was taken to be equal to a randomly selected number of modules, equal for all facilities, and such that global capacity would not exceed the total number of users multiplied by two.

The problems generated using these rules were quite 'realistic', at least with respect to the school network planning conditions encountered in Portugal.

5. Annealing implementation

After being successfully applied to the traveling salesman problem and other classic optimization problems, simulated annealing became widely known in the OR community.

The basic ideas behind the annealing algorithm are briefly described below. For a detailed presentation of the subject, see Dowsland (1993) or Aarts et al. (1997).

Suppose the cost of the current state or solution S of a given system is $C(S)$. A solution to a (discrete) public facility location problem is defined by the locations of open sites, the capacity of the corresponding facilities and the assignment of users to facilities. It can be shown that if the transition of the current state to a randomly selected neighboring solution S', with cost $C(S')$, is made according to some appropriate choice criteria, the system will tend towards the global least cost solution as the number of transition attempts increases.

The choice criterion most commonly used is the Metropolis criterion, built upon the Bolzmann-Gibbs distribution, according to which S would be selected with probability given by

$$p = \min\left\{1, \exp\left(-\frac{\Delta C}{\theta}\right)\right\}.$$

In this expression $\Delta C = C(S') - C(S)$ and θ is the temperature of the system, a parameter whose value may decrease during the annealing process. The function describing the evolution of temperature during the annealing process is known as the cooling schedule.

Notice that according to the Metropolis criterion solutions leading to cost decreases will always be selected ($p = 1$), while solutions leading to cost increases will be selected with larger probability at the beginning of the annealing process, especially if increases are small.

In general terms, the annealing algorithm consists of the following steps:
1. Choose S_1 {S_1 : initial solution}
2. Choose θ_1 {θ_1 : initial temperature}
3. Choose θ_f {θ_f : final temperature}
4. $j \rightarrow 0$
5. **Repeat**
   6. $j \leftarrow j + 1$
   7. Choose at random $S'_j \in N(S_j)$ {$N(S_j)$: neighborhood set}
   8. Choose at random $p \in [0, 1]$;
   9. **If** $p \leqslant \min\left\{1, \exp\left(-\frac{C(S'_j) - C(S_j)}{\theta_j}\right)\right\}$
      Then $S_{j+1} \leftarrow S'_j$
      Else $S_{j+1} \leftarrow S_j$
   10. Choose $\theta_{j+1} \leqslant \theta_j$
   Until $\theta_{j+1} < \theta_f$
11. **End**

Any practical implementation of simulated annealing requires two basic issues to be decided: the generation of candidate solutions; and the attributes of the cooling schedule.

The generation of candidate solutions is accomplished, in our implementation, through the procedure described in Fig. 1.

First, a site and a period are selected at random, with all the sites and periods having the same probability of being chosen. If the site is closed, it will become open, receiving a minimum capacity facility. When the site is open, it will go through a feasible transformation selected at random. This transformation may either consist of adding a new module to the existing facility, transferring a
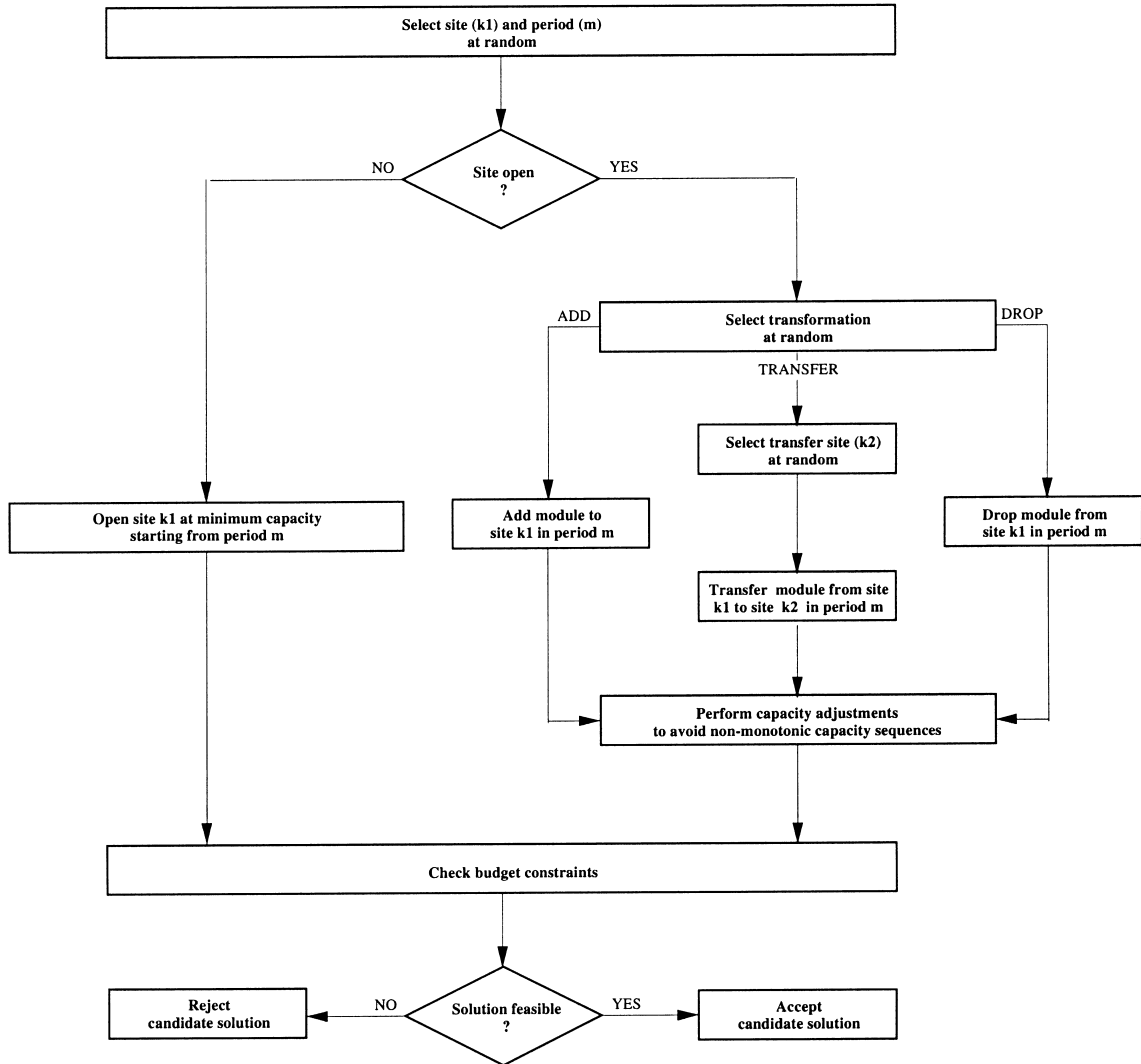
Fig. 1. Generation of candidate solutions for DMCFLP problems.

module to another facility (also selected at random), or dropping a module. For a facility at minimum capacity, the suppression of a module will naturally imply facility closure.

Second, if the site is open, the capacity sequence of facilities where modules have been added or dropped is analyzed and possibly adjusted, to ensure that the monotonic rules applying to the evolution of facilities are observed. The adjustments are made towards the front, starting from the initial period. For example, considering a 3-period problem, if the current capacity sequence for a facility is $\{z + q, z, z\}$ and the selected transformation consists of suppressing a module in the second period, capacity in this period would become equal to $z - q$ (where z represents a given capacity and q stands for the capacity of a module). Hence, we would have the sequence $\{z + q, z - q, z\}$. To avoid this non-monotonic sequence, capacity in the third period has to be set at $z - q$. Therefore, the candidate capacity sequence in the site would be $\{z + q, z - q, z - q\}$. Fig. 2
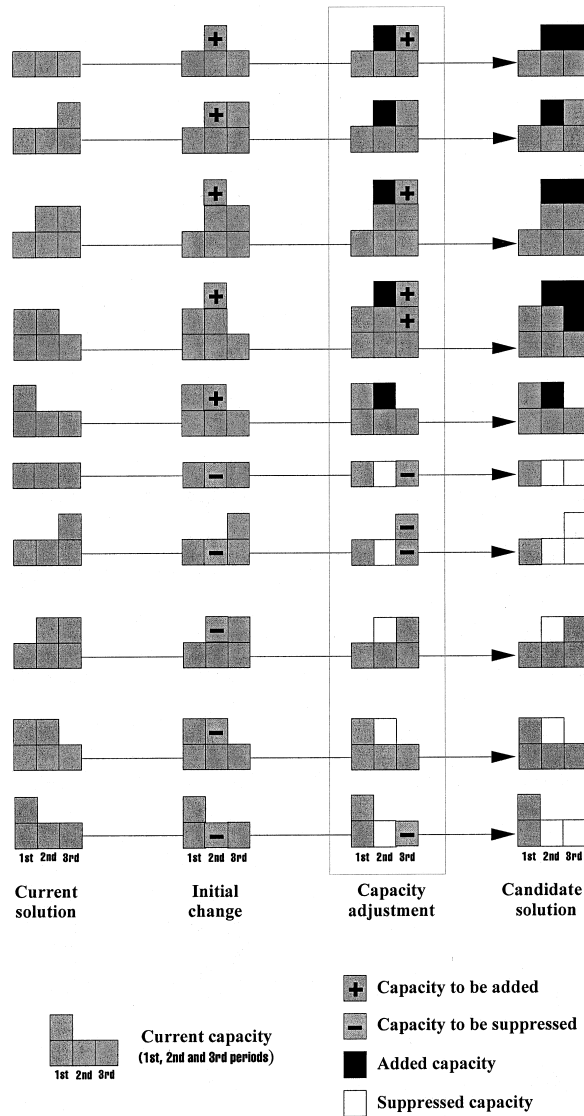
Fig. 2. Capacity adjustments in 3-period problems.

describes the capacity adjustments consecutive to several possible changes in the second period of a 3-period problem. Similarly, in a 4-period problem, if the current capacity sequence for a facility is $\{z + q, z, z, z - q\}$ and the selected transformation consists of adding a module in the third period, capacity in this period would become equal to $z + q$. Hence, we would have the sequence $\{z + q, z, z + q, z - q\}$. To avoid this non-

monotonic sequence, capacity in the third period has to be set back at z (i.e., the change would be rejected).

Finally, compliance with budget constraints is checked. If the candidate solution is accepted, the cost of the solution is calculated after assigning the users to the facilities by solving transportation problems (one for each period). The tree-indexing method employed for this purpose is explained in

Jacobsen (1978). Otherwise, a new candidate solution is generated.

The attributes of the cooling schedule were chosen following the principles adopted by Johnson et al. (1989) in their annealing algorithm for the graph partitioning problem. Those authors defined a schedule involving four parameters: the initial temperature, θ_1; the temperature length, λ; the cooling rate, γ; and the stopping number, σ. The initial temperature defines the rate at which candidate solutions with cost $x\%$ higher than the cost of the initial solution are retained. The temperature length is the minimum number of candidate solutions to be tried at each temperature. If the algorithm is unable to find at least one better single solution or a better average solution, the temperature is decreased. The cooling rate is the rate at which temperature is decreased. The stopping number is the maximum number of temperature reductions that may occur without finding any solution improvements. When this number is reached the system becomes "frozen", and the annealing process reaches the end. The links between the four parameters and the way they interact are shown in Fig. 3.

On the basis of a detailed empirical study carried out on the UFLP and the DMCFLP, described in detail in Antunes (1994, pp. 97–110, 114–124), we chose the following parameters:

- $\theta_1 = 0.13 \times C_1$, where C_1 is the cost of a randomly chosen feasible initial solution (this means that solutions with a cost 30% higher than the initial solution cost will be chosen with a probability of approximately 10%),
- $\lambda = 3KM$ (K is the number of sites and M is the number of periods),
- $\gamma = 0.3$,
- $\sigma = 6$.

In order to test our implementation of simulated annealing (ANNEAL), we compared the corresponding solutions with the solutions given by branch-and-bound (B–B) for a set of 50 random DMCFLP problems. The package XPRESS-MP was used to calculate the B–B solutions, because some preliminary tests indicated that this package would be more efficient than GAMS/CPLEX. The comparison was made on the basis of small 6-center × 6-site × 3-period problems, the
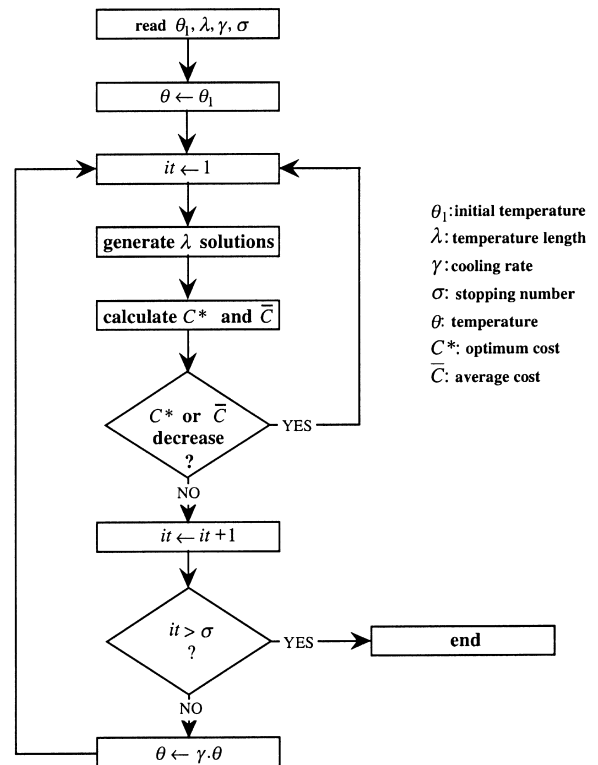

Fig. 3. The cooling schedule for the annealing process.

maximum size we were able to handle within reasonable computational effort.

The results are summarized in Table 3. Average annealing solutions taken for 5 random seeds (i. e., 5 different sets of random numbers) were inferior to the B–B solutions in 36 out of 50 problems, but only in 17 was ANNEAL unable to find the optimum B–B solution. However, it came close to the optimum solution in all these 17 problems, as the difference was always less than 1%. In the remaining 33 problems, ANNEAL was at least as efficient as B–B. The use of 'at least' is justified because, surprisingly, on 9 occasions, ANNEAL gave better solutions than B–B (the XPRESS-MP implementation of it), something that would be impossible if the corresponding optimum were really exact, as we expected them to be. According to the investigation we made, this unexpected result is associated with the default tolerances assumed by XPRESS-MP. Smaller tolerances would probably lead to better solutions, but we were not

Table 3
Comparison between ANNEAL and B–B results for DMCFLP problems

Solutions for five random seeds	Difference in cost between ANNEAL and B–B solutions (%)				
	<−1	[−1, 0[	0	]0, 1]	>1
Worst	1	5	6	21	17
Average	4	4	6	34	2
Best	4	5	24	17	0

able to complete the B–B search procedure within reasonable computational effort (and even if we were, we could not be sure they were true optimum). This made us decide to keep the default-tolerance solutions as a reference, and to see them as the best affordable B–B solutions.

These findings were quite encouraging especially because the computational effort required to run ANNEAL was about 10% of the effort required by B–B. Furthermore, the computing time, which averaged 530 seconds on a 40 MHz Macintosh Quadra 700, showed only limited fluctuation from problem to problem, their coefficient of variation being equal to just 21.5%.

6. Real-world applications

Following the results described in the previous section, ANNEAL was used to solve four facility location problems raised by school network planning in Portugal's Centro Region. These problems, and the context within which they arose, are described elsewhere (Antunes and Peeters, 2000). Therefore, we only include here the essential information about their data and results.

Three of the problems were defined for the secondary school (escolas secundárias (ES)) networks of the sub-regions of Baixo Vouga, Baixo Mondego and Pinhal Litoral. The other one was defined for the elementary school (escolas básicas (EB)) network of the municipality of Leiria, which is part of the Pinhal Litoral sub-region. The location of these four geographical areas is depicted in Fig. 4. All the problems were built considering three periods, representing the short-, the medium-and the long-term.

The ES problems were smaller in size, ranging from the 5-center × 12-site (×3-period) problem

Fig. 4. Location of real-world applications.

defined for the Pinhal Litoral, to the 8×20 and the 12×23 problems defined, respectively, for the Baixo Mondego and the Baixo Vouga problems. In these problems, each center represented a municipality, and each site represented a location within a municipality, either where facilities existed when the planning process began, or where facilities could be located during the process. Comparatively, the EB problem was quite large, involving 29 centers and 38 sites, with both centers and sites corresponding to communities (the smallest Portuguese administrative unit).

The Baixo Mondego problem was studied first, because it was used as a basis for adjusting the algorithm to certain specific needs of real-world applications (for instance, the needs associated with the presence of existing facilities built according to outdated maximum and minimum capacity requirements). This problem was solved 20 times (20 different random seeds), using a 200 MHz Macintosh Performa 6400. The 20 runs gave six different solutions. The best solution occurred six times. The difference in cost between the best and the worst solutions was 1.7%. It should however be said that the short-term intervention corresponding to the best solutions was obtained 14 times. This is an important fact to emphasize because, within a cyclical planning framework, the short-term decisions undoubtedly are the most relevant. Medium- and long-term decisions, which are not to be implemented immediately, can be changed later if necessary. As shown in Table 4, the number of candidate solutions investigated by ANNEAL was on average 5310, with a maximum of 8100 and a minimum of 2700. The average number of accepted solutions and solution improvements was 2740 (51.60%) and 72 (1.36%). The computing time was on average 170 seconds, with a maximum of 257 seconds (i.e., always less than five minutes).

The Pinhal Litoral, Baixo Vouga and Leiria problems were solved five times each. The main (relative) differences with regard to the Baixo Mondego problem occurred with computing time. The Pinhal Litoral problem took on average 43

seconds, with a maximum of 58 seconds (i.e., less than one minute). The computing time for the Baixo Vouga problem was on average 717 seconds, with a maximum of 1047 seconds (i.e., more than 15 minutes). The Leiria problem took on average 299 minutes to solve, with a maximum of 583 minutes (i.e., almost 10 hours). In spite of the large effort needed to compute the solutions for this problem, in one of the five runs a better solution was found through inspection, interchanging open and close sites.

7. Conclusion

The study described in this paper shows that simulated annealing may be a good resort when solving complex mid-size multi-period location problems like those raised by school network planning in Portugal. The corresponding algorithms are really easy to develop, and easy to adapt to new application conditions. The trade-off between cost and quality of solution in our case studies was quite interesting.

The study also revealed the limitations of simulated annealing when dealing with large-size problems. However, it must be said that our algorithm could be improved in respect to a few points. We used a relatively simple algorithm. It could be made more sophisticated, for instance through the introduction of penalty schemes or tabu lists commonly used today within modern search heuristics. Moreover, some elementary procedures could be improved. One such improvement would be achieved by solving transportation problems more quickly, using solutions found in previous iterations. As the algorithm will normally involve solving thousands of transportation problems, the impact of this improvement will be considerable. In the future, we will direct part of our research activities to this kind of issues.

Table 4
Characteristics of the annealing process for the Baixo Mondego problem

Characteristic	Maximum	Average	Minimum
Number of candidate solutions	8100	5310	2700
Number of accepted solutions	3885	2740	1716
Number of solution improvements	111	72	53
Computing time (s)	257	170	86

References

Aarts, E., Leenstra, J. (Eds.), 1997. Local Search in Combinatorial Optimization, Wiley, Chichester, pp. 91–120.

Aarts, E., Korst, J., Laarhoven, P., 1997, Simulated annealing. In: Aarts, E., Leenstra, J. (Eds.), Local Search in Combinatorial Optimization, Wiley, Chichester, pp. 91–120.

Antunes, A., 1994. De la planification optimale de l'équipement scolaire (on optimal planning of school networks). Ph.D. Thesis, Université Catholique de Louvain, Belgium.

Antunes, A., Peeters, D., 2000. A dynamic optimization model for school network planning. Socio-Economic Planning Sciences (forthcoming).

Cornuéjols, G., Nemhauser, G., Wolsey, L., 1990. The uncapacitated facility location problem. In: Mirchandani P., Francis, R. (Eds.), Discrete Location Theory, Wiley, New York, pp. 119–171.

Dash Associates, 1997. XPRESS-MP Release 10 User Guide and Reference Manual.

Daskin, M., 1995. Network and Discrete Location, Wiley, New York.

Dowsland, K., 1993. Simulated annealing. In: Reeves, C. (Ed.), Modern Heuristic Techniques for Combinatorial Problems. Wiley, New York, pp. 20–69.

Drezner, Z. (Ed.), 1995. Facility Location. Springer, New York.

Feldman, E., Lehrer, F., Ray, T., 1966. Warehouse location under continuous economies of scale. Management Science 12, 670–684.

Fong, C., Srinivasan, V., 1981. The multi-region dynamic capacity expansion problem – Part II. Operations Research 29, 801–816.

GAMS Development Corporation, 1996. GAMS/CPLEX 4.0 User Notes.

Greenleaf, N., Harrison, T., 1987. A mathematical programming approach to elementary school facility decisions. Socio-Economic Planning Sciences 21, 395–401.

Hansen, P., Labbé, M., Peeters, D., Thisse, J.-F., 1987. Facility location analysis. In: Lesourne, J., Sonnenschein, H. (Eds.), Systems of Cities and Facility Location, Harwood, London, pp. 1–70.

Krarup, J., Pruzan, P., 1983. The simple plant location model: Survey and synthesis. European Journal of Operational Research 4, 256–269.

Kuehn, A., Hamburger, M., 1963. A heuristic program for locating warehouses. Management Science 9, 643–666.

Jacobsen, S., 1978. On the use of tree-indexing methods in transportation algorithms. European Journal of Operational Research 2, 54–65.

Jacobsen, S., 1990. Multiperiod capacitated location models. In: Mirchandani, P., Francis, R. (Eds.), Discrete Location Theory, Wiley, New York, pp. 173–208.

Johnson, D., Aragon, C., McGeoch, L., Schevon C, 1989. Optimization by simulated annealing: An experimental evaluation – Part I. Graph partitioning. Operations Research 37, 865–892.

Nemhauser, G., Savelsberg, M., Sigismondi, G., 1994. MINTO: A mixed integer optimizer. Operations Research Letters 15, 47–58.

Reeves, C. (Ed.), 1993. Modern Heuristic Techniques for Combinatorial Problems. Wiley, New York.

Ribeiro, A., Antunes, A., 2000. On solving public facility location problems using general mixed-integer programming methods. Engineering Optimization 32 (4).

Roodman, G., Schwartz, L., 1975. Optimal and heuristic facility phase-out strategies. AIIE Transactions 7, 177–184.

Roodman, G., Schwartz, L., 1977. Extensions of the multiperiod facility phase-out model: New procedures and applications to a phase-in/phase-out problem. AIIE Transactions 9, 103–107.

Shulman, A., 1991. An algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. Operations Research 39, 423–436.

Teitz, M., Bart, P., 1968. Heuristic methods for estimating the generalized vertex median of a graph. Operations Research 16, 955–961.

Van Roy, T., Erlenkotter, D., 1982. A dual-based procedure for dynamic facility location. Management Science 28, 1091–1105.

Viegas, J., 1987. Short and mid-term planning of an elementary school network in a suburb of Lisbon. Sistemi Urbani 1, 57–77.