

UNCAPACITATED FACILITY LOCATION PROBLEMS: CONTRIBUTIONS

Roberto Diéguez Galvão

Programa de Engenharia de Produção / COPPE

Universidade Federal do Rio de Janeiro

Rio de Janeiro – RJ

galvao@pep.ufrj.br

Recebido em 05/2003; aceito em 09/2003

Received May 2003; accepted September 2003

Abstract

The objective of the present paper is to review my personal contributions in the field of uncapacitated facility location problems. These contributions took place throughout my academic career, from the time I was a Ph.D. student at Imperial College to the present day. They cover approximately 30 years, from 1973 to 2003; they address: algorithms developed for the p -median problem and for a general formulation of uncapacitated location problems; the study of dynamic location models; covering and hierarchical location problems; queuing-based probabilistic location models. The contributions encompass theoretical developments, computational algorithms and practical applications. All work took place in an academic environment, with the invaluable collaboration of colleagues (both in Brazil and abroad) and research students at COPPE. Each section in the paper is dedicated to a topic that involves a personal contribution. Every one of them is placed within the context of the existing literature.

Keywords: uncapacitated facility location; dynamic location models; covering location problems; hierarchical location models; probabilistic location models.

Resumo

O objetivo do presente artigo é fazer uma revisão de minhas contribuições na área de problemas de localização não-capacitados. Estas contribuições foram realizadas ao longo de minha carreira acadêmica, do tempo em que eu era estudante de doutorado no Imperial College aos dias atuais. Cobrem aproximadamente 30 anos, de 1973 a 2003; são referentes a: algoritmos para o problema das p -medianas e para uma formulação geral para problemas não-capacitados; estudo de modelos de localização dinâmicos; problemas de localização com cobertura e problemas hierárquicos; modelos de localização probabilísticos com base em teoria das filas. As contribuições incluem desenvolvimentos teóricos, algoritmos computacionais e aplicações práticas. O trabalho foi realizado em ambiente acadêmico, com a valiosa colaboração de colegas (no Brasil e no exterior) e de estudantes de pesquisa da COPPE. Cada seção do artigo é dedicada a um tópico envolvendo uma contribuição pessoal, colocada no contexto da literatura existente sobre o assunto.

Palavras-chave: problemas de localização não-capacitados; modelos de localização dinâmicos; problemas de localização com cobertura; modelos de localização hierárquicos; modelos de localização probabilísticos.

1. Introduction

Uncapacitated facility location problems take a great variety of forms, depending on the nature of the objective function (minisum, minimax, problems with covering constraints), on the time horizon under consideration (static, dynamic), on the existence of hierarchical relationships between the facilities and on the inclusion or not of stochastic elements in their formulation (probabilistic, deterministic). When we consider the possible combinations of the categories above, numerous different types of problem can be defined.

Location problems initially studied in the literature were related to industrial contexts, referring to the supply of a single commodity from a set of potential locations, where facilities may be placed, to clients of known locations and demands, at minimum cost. These problems consist of determining the locations of the facilities and the flows of the commodity from facilities to clients, such that the sum of fixed (establishment) and variable (operational and transportation) costs are minimized.

Uncapacitated problems assume that each facility can produce and ship unlimited quantities of the commodity under consideration. A great variety of mathematical models were proposed for this problem. The first models date back to the 60's, when the *Simple Plant Location Problem (SPLP)*, Kuehn & Hamburger, 1963; Balinski, 1965) and the *p-Median Problem (PM)*, Hakimi, 1964, 1965) were defined. It was later shown that these two *prototype models* are particular cases of a more general formulation for deterministic, static, uncapacitated problems having a minisum objective function (Cornuejols, Fisher & Nemhauser, 1977; Galvão & Raggi, 1989).

Problems with covering constraints were defined in the 70's. The objective in this case is to locate facilities such that demand areas (clients) are covered. A demand area is said to be covered by a facility (server) if it is within a critical, pre-defined distance (time) from this facility. The simplest of these models seeks to find the minimum number of facilities (and their locations) such that all demand areas are covered by at least one facility. Church & ReVelle (1974) defined a model in which the number of facilities to be located is fixed, but in this case coverage of all demand areas is not guaranteed.

The covering location models described above do not take into account the stochastic nature of many problems encountered in practical applications, when the unavailability of a server at the time of request (it may be busy servicing another call) may cause a demand area theoretically covered to be left uncovered in practice. This type of situation was initially dealt with through the development of *back-up covering models* (Hogan & ReVelle, 1986; ReVelle, 1989) and finally through the definition of *probabilistic location models* (Daskin, 1983; ReVelle & Hogan, 1989).

Hierarchical systems generally consist of k (≥ 2) distinct types of facility that are hierarchically related. For example, health care systems may consist of clinics and hospitals; higher education systems may consist of technical schools and universities; production-distribution systems may consist of factories and warehouses, with a given product shipped to a client directly from the factory or through one of the warehouses. Two hierarchical models are reviewed in this paper: a model developed for the location of maternal and perinatal health care facilities in the municipality of Rio de Janeiro (Galvão, Espejo & Boffey, 2002; Boffey, Yates & Galvão, 2003) and a hierarchical covering location model (Espejo, Galvão & Boffey, 2003).

All problems defined in the preceding paragraphs are of a static nature, that is, they solve a given problem for a specific point in time. They do not take into consideration the dynamic nature of real world systems, where data and parameters change with time. When it is necessary to take into account a *planning horizon* in the solution of a location problem, and conditions that change along this horizon, it is necessary to define *dynamic location models* (Van Roy & Erlenkotter, 1982; Galvão & Santibañez-Gonzalez, 1990, 1992).

The objective of the present paper is to review my personal contributions in the field of uncapacitated facility location problems. These contributions took place throughout my academic career, from the time I was a Ph.D. Student at Imperial College in London to the present day. They cover approximately 30 years, from 1973 to 2003 (when I became 60 years old), and address algorithms developed for the p -median problem and for a general formulation of uncapacitated location problems, the study of dynamic location models, covering and hierarchical location problems and queuing-based probabilistic location models.

The contributions encompass theoretical developments, computational algorithms and practical applications. All work took place in an academic environment, with the invaluable collaboration of colleagues (both in Brazil and abroad) and research students at COPPE, Federal University of Rio de Janeiro. Each section in the paper is dedicated to a topic that involves a personal contribution. Every one of them is placed within the context of the existing literature. Section 2 covers Minisum Uncapacitated Location Models, Section 3 Dynamic Models and Section 4 Covering Models. Hierarchical Models are the object of Section 5, which is followed by Probabilistic Models in Section 6. Conclusions (Section 7) close the paper.

2. Minisum Uncapacitated Static Location Models

We start by giving the mathematical formulation of a general model for static uncapacitated facility location problems with a minisum objective function (UFP). The general model is that of Galvão & Raggi (1989). It is defined by:

(UFP)

$$v(UFP) = \min \left\{ \sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \right\} \quad (1)$$

s. to

$$\sum_{i \in I} x_{ij} = 1, \quad j \in J \quad (2)$$

$$\sum_{i \in I} y_i \leq p \quad (3)$$

$$x_{ij} \leq y_i, \quad i \in I, \quad j \in J \quad (4)$$

$$x_{ij} \in \{0,1\}, \quad i \in I, \quad j \in J \quad (5)$$

$$y_i \in \{0,1\}, \quad i \in I, \quad (6)$$

where $I = \{1, \dots, n\}$ is the set of candidate locations at which facilities may be established, $J = \{1, \dots, m\}$ is the set of demand points, f_i is the fixed cost of establishing a facility in $i \in I$, c_{ij} is the total cost of supplying demand $j \in J$ from a facility located in $i \in I$ and p is the maximum number of facilities allowed in the solution. The decision variable x_{ij} is the fraction of the demand of $j \in J$ supplied from $i \in I$ and y_i is a binary location variable: $y_i = 1$ if a facility is located in $i \in I$, $y_i = 0$ otherwise.

Restrictions (2) ensure that all demand is satisfied. Constraint (3) limits the number of facilities in the solution to p . Restrictions (4) ensure that a customer j can be served from a facility $i \in I$ only if a facility is established at i . Constraints (5)-(6) define the decision variables as zero-one. Note that, in general, we would define the x_{ij} 's as $x_{ij} \geq 0$, $i \in I$, $j \in J$; however, due to the single assignment property of (UFP) (see Krarup & Pruzan, 1983), a customer is always entirely supplied from its nearest facility.

It is easy to see that (UFP) may reproduce well-known formulations of both (SPLP) and (PM). If in (3) we make $p = |J|$, restriction (3) becomes redundant and can be eliminated from the formulation, which will then correspond to that of (SPLP), with the number of facilities in the solution determined by the equilibrium between the c_{ij} 's and the f_i 's. On the other hand, if $I \equiv J$ correspond to the vertices of a network, $[c_{ij}] \equiv [d_{ij}]$ correspond to distances d_{ij} measured along the arcs of the network and $f_i = 0$ for all vertices $i \in I$, then exactly p facilities will be sited ($\sum_{i \in I} y_i = p$) and the formulation corresponds to that of a p -median problem. The reader should keep in mind the assumption $I \equiv J$ for p -median type problems.

Our contributions have been more related to (PM), the subject of our Ph.D. thesis (Galvão, 1977), although recently (SPLP) has also been addressed by us on the basis of reduction tests and ADD/DROP heuristics (Bornstein *et al.*, 2004). The reduction tests determine, *a priori*, if some facilities must be kept open or closed in the optimal solution. The term *reduction* is used due to the fact that it is possible, in this manner, to reduce the size of the original problem. Reduction tests unfortunately are generally unable to fix the status of all facilities; ADD/DROP heuristics have therefore to be additionally used. Bornstein *et al.* (2004) developed six of these heuristics to complement the reduction tests in their algorithm for (SPLP).

This algorithm was tested with both Euclidean and non-Euclidean data sets. The Euclidean sets include adapted Beasley (1990), Cornuejols, Sridharan & Thizy (1991) and Karg & Thompson (1964) test problems. Although the algorithm performed well for these problems, producing good quality solutions in reduced computing times, there is not a clear advantage of using it instead of optimal procedures, for example the algorithm of Galvão & Raggi (1989).

The non-Euclidean data sets were randomly generated. They correspond to combinatorial optimization problems that are much harder to solve than the Euclidean problems. In this case the algorithm shows a clear advantage over the use of optimal procedures (e.g., the algorithm of Galvão & Raggi), when computing times are taken into consideration. It produced, on average, solutions of very good quality in very reduced computing times. For problems for which the optimal solution was available, the maximum percentage deviation from the optimal was 2.09%. The computing times were more than one order of magnitude lower than the corresponding computing times obtained with the algorithm of Galvão & Raggi.

2.1 Solution methods for the p -median problem

We shall now consider the following mathematical programming formulation of (PM):

(PM)

$$v(PM) = \min \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \quad (7)$$

s. to

$$\sum_{i \in I} x_{ij} = 1, \quad j \in J \quad (8)$$

$$\sum_{i \in I} y_i = p \quad (9)$$

$$x_{ij} \leq y_i, \quad i \in I, \quad j \in J \quad (10)$$

$$x_{ij} \in \{0,1\}, \quad i \in I, \quad j \in J \quad (11)$$

$$y_i \in \{0,1\}, \quad i \in I. \quad (12)$$

The first attempts to solve (PM) to optimality were through branch-and-bound algorithms that used bounds derived from the structure of the cost matrix, see Järvinen, Rajala & Sinervo (1972); El-Shaieb (1973), or through linear programming-related approaches, see ReVelle & Swain (1970); Garfinkel, Neebe & Rao (1974); Swain (1974). The problems that these approaches solved did not exceed the size (55x55), either because the bounds were not tight enough or because the very large linear programs that resulted could not be solved efficiently due to the very degenerate nature of the corresponding formulations, see Galvão (1981).

Dual-based approaches

Dual-based approaches have been very effective in the solution of (SPLP), see Bilde & Krarup (1977); Erlenkotter (1978); Körkel (1989); Guignard (1988) and Tcha, Ho & Yo (1988), among others; perhaps for this reason we are not aware of attempts to solve (SPLP) by the direct use of Lagrangean relaxation. The same is not true for (PM). The use of dual-based procedures to solve (PM) was initially exploited in two unpublished papers by Diehr (1972) and Marsten (1972). It was on the basis of their work that this author developed a dual-bounded algorithm to solve (PM) during 1975-1976, although the corresponding paper was only published in 1980, see Galvão (1980). This work was later reviewed as a modification of Erlenkotter's algorithm for (SPLP). It may indeed be viewed as a specialization of Erlenkotter's algorithm, but it was in fact developed independently, before Erlenkotter's paper was published, at a time when ideas about dual-based procedures for these problems were just beginning to be considered.

The algorithm of Galvão (1980) for (PM) is a dual ascent algorithm that solves the dual of the linear programming relaxation of (7)-(12). This procedure produces sharp lower bounds and was embedded into a branch-and-bound algorithm. The computational results, while strong at the time, are modest by today's standards, having solved to optimality problems of up to the size (40 x 40) in reasonable computing times. Thus, it may be concluded that dual-based approaches for (PM) have been far less successful than corresponding approaches for (SPLP), mainly because the inclusion of a restriction on the maximum number of facilities in the solution (restriction (9)) complicates the solution of the dual (see Galvão, 1993).

Methods based on Lagrangean relaxations of (PM)

Lagrangean relaxation has been used quite extensively to solve (PM). Narula, Ogbu & Samuelsson (1977) were the first authors to use this approach for (PM). They relaxed constraints (8) and used subgradient optimization to solve the Lagrangean dual; they tested their algorithm solving the 10-, 20- and 30-vertex city problems given in El-Shaieb (1973). In all problems an optimal integer solution was obtained at the end of the subgradient optimization procedure.

Christofides & Beasley (1982) used a “weak” formulation of (PM), obtained when constraints (10) are replaced by constraints

$$\sum_{j \in J} x_{ij} \leq n_i y_i, \quad i \in I, \quad (10a)$$

where n_i is the cardinality of the set of customers that can be supplied from facility $i \in I$. If every customer can be supplied from any facility, then $n_i = n \forall i$. They considered two different Lagrangean relaxations of this formulation, the first dualizing constraints (8), the second dualizing constraints (10a). Both relaxations have the integrality property, but the authors improved the value of the bounds by the use of penalties in both relaxations and developed a tree search algorithm to close duality gaps. Beasley (1985) enhanced this algorithm and was able to solve problems in networks of up to 900 vertices (for selected values of p) on a Cray-1S ‘super-computer’, within a self-imposed time limit of 600 secs.

The original formulation (7)-(12) of (PM) was revisited by Hanjoul & Peeters (1985), who tested a new relaxation based on the dualization of constraint (9). This was inspired by Erlenkotter’s successful approach in solving (SPLP); dualization of (9) yields a (SPLP). The authors used DUALOC to solve this relaxation and embedded the Lagrangean dual in a tree search procedure to close eventual duality gaps.

Mirchandani, Oudjit & Wong (1985) used the same relaxation to solve (PM) and some of its extensions: a stochastic version, a multi-commodity version and a multi-objective version. The authors show that these multidimensional versions of (PM) simplify to the classical p -median problem via a suitable transformation of variables, but with a κ -fold increase in the number of nodes of the underlying network, where κ is defined as the number of dimensions of the network. They also used DUALOC to solve the Lagrangean problems, but did not define a branch-and-bound procedure to close duality gaps.

The idea of relaxing constraint (9) to solve (PM) was also investigated by Boffey & Karkazis (1984), but without formally defining a Lagrangean relaxation-based algorithm for the problem. They also discuss a multi-commodity version of (PM) and show that this problem may be solved as a p -median problem of larger size. They tested their algorithm using data related to the 20 and 30 largest cities in the United States; also was used data related to the placement of stations in the Ruhr district of Germany.

Our contribution towards using Lagrangean relaxation to solve (PM) comes as a specialization of the 3-phase optimal procedure developed by Galvão & Raggi (1989) for (UFP). This method is described in some detail in Section 2.2. Regarding its use to solve (PM), the algorithm was tested with both Euclidean and non-Euclidean data (the latter corresponding to randomly generated networks). When the computational results shown in that paper are analyzed, it becomes clear that (PM) is the “hardest” problem to solve by the method (after results obtained using the algorithm to solve (UFP) and (SPLP) are taken into account).

The method is quite effective in solving p -median problems of large size in reduced computing times. Results for non-Euclidean problems are shown in Galvão & Raggi (1989). Regarding Euclidean problems, the algorithm was tested against problems obtained from Beasley's OR-Library (see Beasley, 1990). For example, it solved the 900-vertex Beasley problems mentioned a few paragraphs above in less than 600 seconds on a CDC 300 Workstation of old generation, for several values of p (see Galvão, Ferreira Filho & Rivas, 1996).

Heuristics and Meta-heuristics

The first heuristic methods developed for (PM) were the algorithms of Maranzana (1964) and Teitz & Bart (1968), the latter describing a method based on single vertex substitution. Pizzolato (1994) and Pizzolato & Fraga da Silva (1997) developed a heuristic for large weighted graphs and used their algorithm to locate schools in the metropolitan area of Rio de Janeiro, Brazil.

The vertex substitution method of Teitz & Bart is in fact one of a family based on local optimization and the idea of β -optimality, which was first introduced by Lin (1965) for the traveling salesman problem and later extended by others for a variety of combinatorial problems (see Christofides & Eilon, 1969, 1972; Kerningan & Lin, 1970). Eilon & Galvão (1978) extended the method of Teitz & Bart by introducing the idea of β -substitution of vertices.

The β -substitution procedure of Eilon & Galvão for (PM) involves examining a given solution S consisting of a set of p points by exchanging β of its points (where $\beta \leq p$) with β points taken from the total set X of the n points in the network. The replacement set of β vertices chosen from X must obviously satisfy the condition that at least one of the β points does not belong to S . In (PM) a set \bar{S} of p vertices is called β -optimal if the substitution of any β vertices in \bar{S} does not improve the solution corresponding to \bar{S} . In this context, the answer produced by the single vertex substitution algorithm of Teitz & Bart relates to $\beta = 1$ and may be called 1-optimal.

Eilon & Galvão (1978) used a "greedy" vertex addition heuristic as a "pre-processor" for their β -optimal vertex substitution procedure. The idea behind this "combined approach" is that since β -optimal substitution algorithms must start from a set S of p vertices, some advantage might be gained by starting from a "good" initial set. Due to the high computing times involved, Eilon & Galvão had to limit their experiments to a maximum value of $\beta = 2$. They observed that 2-optimal substitution solutions are only slightly better than the corresponding 1-optimal solutions. From a cost-effectiveness point of view, the combination of the "greedy" vertex addition procedure with the 1-optimal substitution method appears to be the best of the methods studied.

A variety of meta-heuristic approaches for (PM) have been proposed in recent years. These include the two stage construction heuristic of Rosing & ReVelle (1997); the tabu search procedures of Mladenovic, Moreno & Moreno-Vega (1996), Voss (1996) and Rolland, Schilling & Current (1996); and the variable neighborhood search approaches of Hansen, Mladenovic & Perez-Brito (1998) and Hansen & Mladenovic (1998).

Chiyoishi & Galvão (2000) made a statistical analysis of simulated annealing applied to the p -median problem. Their algorithm combines elements of the vertex substitution method of

Teitz & Bart with the general methodology of simulated annealing. The cooling schedule adopted incorporates the notion of temperature adjustments rather than just temperature reductions. The authors give computational results for test problems ranging from 100 to 900 vertices, retrieved from Beasley's OR-Library for combinatorial problems. Each problem was run for a maximum of 100 different streams of random numbers. Optimal solutions were obtained for 26 of the 40 problems tested, although high optimum hitting rates were obtained for only 20 of them. The worst gap in relation to the optimal solution was 1.62%, after all runs for each of the test problems were computed.

It is not an easy task to compare the meta-heuristic procedures listed above for (PM), due to the diversity of objectives sought by authors and the different machines, programming languages and test problems used in the corresponding papers. Rosing & ReVelle (1997), for example, state that their objective was not to develop a *faster* algorithm for (PM), but to obtain improved solutions through a heuristic procedure. In their comparative paper Rosing *et al.* (1998) conclude that, for test problems that they regard as being of a particularly challenging character, the heuristic concentration of Rosing & ReVelle (1997) finds better solutions than the tabu search of Rolland, Schilling & Current (1996) in 95% of the cases; no general conclusion on the relative computing times of the two approaches is drawn.

The paper of Chiyoshi & Galvão (2000) was aimed at exploring the capability of simulated annealing in solving (PM); no particular concern was devoted to the efficiency of the algorithm in terms of computing times. It was felt, however, that obtaining an estimate of computing times could be useful and to this end a regression model was developed. It was found, for example, that for n ranging from 500 to 700 the computing time t regresses on $p(n-p)$ with a correlation coefficient of 0.972, the estimation equation being t (in sec) = $p(n-p)/715$.

2.2 A solution method for minisum uncapacitated facility location problems

We refer now to the formulation of (UFP) given by equations (1)-(6) above. Galvão & Raggi (1989) proposed a three-phase method to solve this problem. The method is composed of a primal-dual algorithm, a subgradient optimization to solve a Lagrangean dual and a branch-and-bound algorithm. It has a hierarchical structure, with a given stage being activated only if the optimal solution could not be identified in the preceding stage. While we concentrate on the Lagrangean relaxation-based algorithm (phase 2 of the method), we give a brief overview of the complete procedure.

The primal-dual algorithm consists of interactively obtaining upper and lower bounds for (UFP). The upper bounds are given by a primal procedure that consists of vertex addition and substitution; see Teitz & Bart (1968); Whitaker (1983). This procedure finds approximate solutions for the problem. The lower bounds are found through an ascent procedure used to solve the dual of the linear programming relaxation of (UFP). The interactive procedure starts with the generation of an initial primal solution through a vertex addition heuristic. This primal solution defines the values of the initial dual variables, which allows the dual ascent procedure to be activated. The solution of the dual in turn allows the identification of a reduced set of vertices that are strong candidates to enter a new solution of the primal (through vertex substitution).

A new iteration of the primal-dual algorithm takes place whenever better bounds are produced by either the primal or the dual. If at any stage the upper and lower bounds

coincide, an optimal solution has been identified. If non-coinciding upper and lower bounds are repeated, the Lagrangean relaxation-based algorithm (phase 2) is activated.

Consider now the (UFP) formulation, and note that restrictions (2) can be replaced by

$$\sum_{i \in I} x_{ij} \geq 1, \quad j \in J, \quad (2a)$$

in an equivalent formulation of the problem: for non-negative c_{ij} 's and f_i 's the inequality form holds with equality at the optimum. Now let $\lambda \geq 0$ be a vector of Lagrange multipliers associated with restrictions (2a). We can define:

(UFP $_{\lambda}$)

$$\begin{aligned} v(UFP_{\lambda}) &= \min \left\{ \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{i \in I} f_i y_i + \sum_{j \in J} \lambda_j \left(1 - \sum_{i \in I} x_{ij} \right) \right\} \\ &\equiv \min \left\{ \sum_{j \in J} \lambda_j - \sum_{i \in I} \left(\sum_{j \in J} (\lambda_j - c_{ij}) x_{ij} - f_i y_i \right) \right\}, \end{aligned} \quad (13)$$

subject to (3)-(6).

Problem (UFP $_{\lambda}$) can be solved by inspection for a fixed vector λ . It is not difficult to see that there is an optimal solution in which the x_{ij} 's satisfy

$$x_{ij} = \begin{cases} y_i & \text{if } \lambda_j - c_{ij} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Now if we define $s_i(\lambda) = \sum_{j \in J} \max(0, \lambda_j - c_{ij}) - f_i$, from (13) it follows that the optimal y_i 's are obtained by solving the reduced problem $v(UFP_{\lambda}) = \max \sum_{i \in I} s_i(\lambda) y_i$, subject to $1 \leq \sum_{i \in I} y_i \leq p$, $y_i \in \{0, 1\}$, $i \in I$. To show that the reduced problem can be solved by inspection, let $J(\lambda) = \{i \in I \mid s_i(\lambda) \geq 0\}$ and let $J_p(\lambda)$ be a set such that:

- (a) If $1 \leq |J(\lambda)| \leq p$, then $J_p(\lambda) = J(\lambda)$;
- (b) If $|J(\lambda)| = 0$, $J_p(\lambda)$ has only one element which corresponds to the index of a largest $s_i(\lambda)$;
- (c) If $|J(\lambda)| > p$, then $J_p(\lambda)$ is a set of indices that correspond to p largest $s_i(\lambda)$'s.

The solution of the reduced problem is given by $y_i = 1$ if $i \in J_p(\lambda)$, $y_i = 0$ otherwise.

Galvão & Raggi used the subgradient optimization method to solve the Lagrangean dual. They decided, however, to continuously update the step size parameter α_k by a function that approaches the shape of the positive half of the normal curve, obtaining better convergence results than through the traditional scheme (given by Held, Wolfe & Crowder, 1974) of halving the value of α_k until it reaches a very small value. In the event of a duality gap at the end of phase 2, a branch-and-bound algorithm that uses information provided by phase 2 finds an optimal solution for (UFP).

The algorithm of Galvão & Raggi (1989) was tested with two sets of data: data available in the literature and randomly generated networks. These data were used in the solution of problems (UFP), (SPLP) and (PM), for problems up to the size 200 customers x 200 potential facility sites. Only problems (PM) of larger size required the branch-and-bound algorithm to close duality gaps. For (SPLP) this method was compared with the algorithm of Erlenkotter (1978), with data used by him. Although the method of Erlenkotter is generally faster for the larger problems, the three-phase method was faster in selected large problems (of size 100 x 100), which leads to the conclusion that the results are data dependent to a great extent.

2.3 A practical application: distribution of benefits in the Brazilian Social Security System

The distribution of benefits in the Brazilian Social Security System is done by Government-run facilities located both in small and large urban areas. In the large urban areas a capacitated location problem must be solved; in the smaller centres the solution of an uncapacitated problem is sufficient. In the latter case the p -median model was used to locate these facilities in the state of Espírito Santo, in what constituted a pilot project for the application of the methodology, see Galvão & Nascimento (1990).

In Brazil states are divided into municipalities, which are the smallest regional demographic units for data aggregation purposes. Municipalities with a large population will evidently require more than one facility; this case was dealt with separately by using a capacitated model. For municipalities with small populations, however, it is not economically feasible to locate a benefit-distributing facility in each municipality. In the case of these municipalities it became an objective to determine, as function of a previously defined social criterion, and given an investment budget, in which municipalities to locate facilities to provide service for the neighbouring populations. It was decided that capacities should not be taken into account in this case.

In wide geographical areas with sparse population, a social criterion that was considered satisfactory for the location of the facilities was the minimization of the average distance travelled by the user. A solution to this problem can be obtained through the use of an uncapacitated location model (for example, the p -median model). The problem thus consists of locating the facilities so as to minimize distances travelled along the existing transportation network. These distances must be weighted by the population of each municipality, centred in its administrative unit.

The p -median model was used to locate benefit-distributing facilities in the state of Espírito Santo. The results obtained correspond to official population data by municipality in the year of 1985. The transportation network considered corresponds to federal and state roads, plus the part of the Vitória-Minas railway that serves the state.

The state of Espírito Santo is divided into 53 municipalities. With the objective of comparing the existing network at the time of the study with the network proposed by the methodology we run the model for $p = 11$, which was the number of municipalities in which benefit-distributing facilities were then located. When the existing network was compared with the network proposed by the methodology, a 20% reduction in average distance travelled by the user was obtained with the locations proposed by the model. This reduction appears to be of some significance, although facility relocation costs would also have to be considered before deciding whether to implement the results suggested by the model.

3. Dynamic Models

Scott (1971) pioneered in the study of uncapacitated dynamic location problems. His dynamic model is an extension of a static location-allocation model that does not include fixed costs; it only allows one facility to be opened per time period. Warszawski (1973) addressed two multi-dimensional location problems, a multi-commodity problem and a multi-stage problem. His multi-stage problem includes installation, maintenance and transportation costs and the corresponding formulation has nonlinear terms in the objective function (see Galvão, 1993).

Roodman & Schwarz (1975, 1977) worked with a model similar to that of Warszawski, in which the facilities subject to opening and closing are limited to predefined sets I_o and I_c . Van Roy & Erlenkotter (1982) defined a formulation closely related to that of Roodman & Schwarz. In their formulation the sum of establishment and maintenance costs is represented by the symbol f_{ik} , which is defined as the fixed cost of having the facility at $i \in I$ open in period k . They use the following mathematical formulation:

(DSPLP)

$$v(DSPLP) = \min \left\{ \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} c_{ijk} x_{ijk} + \sum_{i \in I} \sum_{k \in K} f_{ik} y_{ik} \right\} \quad (14)$$

s. to

$$\sum_{i \in I} x_{ijk} = 1, \quad j \in J, \quad k \in K \quad (15)$$

$$x_{ijk} \leq y_{ik}, \quad i \in I, \quad j \in J, \quad k \in K \quad (16)$$

$$y_{ik} \leq y_{i(k+1)}, \quad i \in I_o, \quad k \in K, \quad 1 \leq k \leq r-1 \quad (17)$$

$$y_{ik} \geq y_{i(k+1)}, \quad i \in I_c, \quad k \in K, \quad 1 \leq k \leq r-1 \quad (18)$$

$$x_{ijk}, y_{ik} \in \{0,1\}, \quad i \in I, \quad j \in J, \quad k \in K, \quad (19)$$

where $K = \{k \mid k = 1, \dots, r\}$ is the set of time periods, c_{ijk} is the total transportation cost of supplying customer j from facility i in period k and

$$x_{ijk} = \begin{cases} 1 & \text{if client } j \text{ is supplied from facility } i \text{ in period } k, \\ 0 & \text{otherwise;} \end{cases}$$

$$y_{ik} = \begin{cases} 1 & \text{if a facility } i \text{ is operating in period } k, \\ 0, & \text{otherwise.} \end{cases}$$

Scott (1971) used dynamic programming to obtain a solution to his formulation. Warszawski (1973) proposed an optimal recursive procedure based on dynamic programming and a heuristic based on facility interchange which he called “the highest marginal savings procedure”. Roodman & Schwarz (1975, 1977) used two “phase-in/phase-out” algorithms, one exact, the other heuristic. Finally, Van Roy & Erlenkotter (1982) extended, for the dynamic problem, the dual-based approach developed by Bilde & Krarup (1977) and Erlenkotter (1978) for (SPLP).

p -Median dynamic location models

Wesolowsky & Truscott (1975) introduced dynamic location models with restrictions on the number of facilities in the solution. Their model includes both opening and closing fixed costs in the objective function; maintenance costs however are not considered. A facility may be closed in any period $k \in K$ of the planning horizon, but at any given period k exactly p facilities must be in operation. In order to reflect acceptable levels of organizational disruption, the authors also include a restriction on the number of facility location changes in any period k .

Galvão & Santibañez-Gonzalez (1990) formulated a model in which the required number of operating facilities changes along the planning horizon. The cost of closing facilities is considered negligible and therefore not included in the objective function; maintenance costs are also not considered. The problem is defined within a network structure; therefore $I \equiv J$. It is also assumed that $c_{iik} = 0 \forall i, k$. The original formulation of the problem is given by:

(DLP- p_k)

$$v(DLP-p_k) = \min \left\{ \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} c_{ijk} x_{ijk} + \sum_{i \in I} f_{i1} x_{i1} + \sum_{i \in I} \sum_{\substack{k \in K \\ k \neq 1}} f_{ik} x_{iik} (1 - x_{ii(k-1)}) \right\} \quad (20)$$

s.to

$$\sum_{i \in I} x_{ijk} = 1, \quad j \in J, \quad k \in K \quad (21)$$

$$\sum_{i \in I} x_{iik} = p_k, \quad k \in K \quad (22)$$

$$x_{ijk} \leq x_{iik}, \quad i \in I, \quad j \in J, \quad k \in K, \quad i \neq j \quad (23)$$

$$x_{ijk} \in \{0,1\}, \quad i \in I, \quad j \in J, \quad k \in K, \quad (24)$$

where p_k is the number of facilities operating in period k . It is possible to linearize the objective function by defining new variables y_{ik} as $y_{ik} = 1$ if a facility at $i \in I$ is opened in period k , $y_{ik} = 0$ otherwise, and adding the constraints

$$y_{ik} - x_{iik} + x_{ii(k-1)} \geq 0, \quad i \in I, \quad k \in K, \quad (25)$$

with $x_{ii0} = 0, i \in I$. The problem can be rewritten as

(DLP- p_k)

$$v(DLP-p_k) = \min \left\{ \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} c_{ijk} x_{ijk} + \sum_{i \in I} \sum_{k \in K} f_{ik} y_{ik} \right\} \quad (26)$$

s.to (21)-(25).

Note that variables x_{iik} and y_{ik} have different meanings. $x_{iik} = 1$ implies that a facility at $i \in I$ is operating in period k (whether or not it has been opened in this period), while $y_{ik} = 1$ means that a facility was opened at $i \in I$ in period k . Galvão & Santibañez-Gonzalez (1990) developed a two-phase heuristic solution method based on dynamic programming to solve (DLP- p_k). In the first phase the number of states in each period k is reduced through the use

of upper and lower bounds (obtained by solving static p -median problems) and the solution of a restricted p -median problem. In phase 2 a forward dynamic programming algorithm is used to find approximate solutions for (DLP- p_k). The authors were able to solve problems of up to the size 70 customers x 70 potential facility sites x 8 time periods (70 x 70 x 8) in less than 3.5 min in a Burroughs A9 computer.

In a later paper Galvão & Santibañez-Gonzalez (1992) used a Lagrangean heuristic to solve (DLP- p_k). Two Lagrangean Relaxations were obtained for the problem using the formulation given by (26) subject to (21)-(25). The first relaxation results from the dualization of (21), using the vector of Lagrangean multipliers $\lambda 1 = \{\lambda_{ik}\} \geq 0$, $\lambda 1$ being non-negative since restrictions (21) could be replaced by $\sum_{i \in I} x_{ijk} \geq 1, j \in J, k \in K$, in an equivalent formulation of the problem, and also from the dualization of (25), using vector $\mu = \{\mu_{ik}\} \geq 0$. The second Lagrangean relaxation is obtained through the dualization of (23) (vector $\lambda 2 = \{\lambda_{ijk}\} \geq 0$) and (25). For further details of these two relaxations the reader is referred to Galvão & Santibañez-Gonzalez (1992).

Galvão & Santibañez-Gonzalez (1992) tested their algorithm using both relaxations, but the use of the second relaxation did not produce good results, in terms both of quality of bounds and computing times. With the first relaxation they were able to solve problems of size up to 50 x 50 x 7 in less than 18 min on a Burroughs A9 computer. Comparison of these results with those of Galvão & Santibañez-Gonzalez (1990) shows that the lower bounds obtained through the Lagrangean heuristic are of considerably better quality, while the upper bounds only represent a slight improvement over those of the earlier paper. The computing times, however, are up to 20 times longer for the Lagrangean heuristic.

4. Covering Models

The objective of location covering models is to provide *coverage* to demand areas. A demand area is said to be covered by a facility if it is within a required distance or time (*critical* or *service distance* [SD]) from the facility. There is a vast literature on models of this type, which generally address the location of urban public facilities, especially emergency facilities. It is not our intention to review these models extensively; for good reviews of the subject the reader is referred to ReVelle (1987, 1989).

The first covering models studied were deterministic. The simplest of these models is the location set covering problem (LSCP), which seeks to determine and position the minimum number of facilities that are necessary to cover all demand areas within SD distance or time units. A related problem is the p -center problem (PCP), that seeks the location of p facilities such that the maximum distance (time) from any demand area to its nearest facility is minimized.

(LSCP) requires that all demand areas be covered, and this may demand excessive resources not always available to the public authorities. Recognizing this fact, Church & ReVelle (1974) developed the maximal covering location problem (MCLP), which does not require that all demand areas be covered. White & Case (1974) worked on a similar problem that seeks to locate p facilities that can cover the maximum number of demand areas (rather than the maximum population). In the case of MCLP the objective is to locate p facilities in such a way that the maximum possible population is covered within distance (time) SD . Its mathematical formulation is given by:

(MCLP)

$$v(MCLP) = \max \sum_{j \in J} pop_j \xi_j \quad (27)$$

s.to

$$\sum_{i \in I} a_{ij} y_i - \xi_j \geq 0, \quad j \in J \quad (28)$$

$$\sum_{i \in I} y_i = p \quad (29)$$

$$\xi_j \in \{0,1\}, \quad j \in J \quad (30)$$

$$y_i \in \{0,1\}, \quad i \in I, \quad (31)$$

where pop_j is the population of demand area $j \in J$; $a_{ij} = 1$ if demand area j can be covered by a facility located at $i \in I$ within critical distance SD ($a_{ij} = 0$ otherwise). $\xi_j = 1$ if demand area j is covered ($\xi_j = 0$ otherwise) and y_i is the usual location variable.

In the formulation above the objective function seeks to maximize the total population covered. Constraints (28) state that a demand area $j \in J$ is covered if there is at least one facility within distance SD from it. Restriction (29) limits the number of facilities in the solution to p . Finally, constraints (30)-(31) define the binary nature of the decision variables.

Since its proposal (MCLP) has been generalized in different ways (see Boffey & Narula, 1997). Applications are found both in the public and private sectors. Chung (1986) reviews several applications of (MCLP). In relation to emergency services, Eaton *et al.* (1986) used it to determine ambulance deployment in Santo Domingo (Dominican Republic), Current & O'Kelly (1992) to locate warning sirens in cases of emergency. In the private sector (MCLP) has been used to locate bank branches, see Pastor (1994). Other applications of (MCLP) can be found in Dwyer & Evans (1981) [selection of mailing lists]; Daskin, Jones & Lowe (1990) [flexible manufacturing]; and Houglund & Stephens (1976) [air pollution control], among others.

Early solution methods proposed for (MCLP) include the Linear Programming (LP) relaxation of the 0-1 integer formulation of the problem and a "greedy"-interchange heuristic (see Church & ReVelle, 1974). Galvão & ReVelle (1996) developed a Lagrangean heuristic for the problem; they report computational experience using data from the literature and randomly generated networks. Exact methods include the algorithm of Dwyer & Evans (1981), developed for the particular case where all demand areas have equal weight, and the dual-based algorithm of Downs & Camm (1996). The latter authors present an extensive computational evaluation of their method, in terms of both variety of applications and problem size.

A second generation of location covering models focussed on additional coverage. These models emphasize the importance of additional coverers for the demand areas, given the possibility that in congested systems the first server, possibly the only server in a particular coverage area, might not be available when requested. Several such models were developed, as for example in Daskin & Stern (1981), Eaton *et al.* (1981), Hogan & ReVelle (1986) and Batta & Mannur (1990). Probabilistic covering models are a natural extension of the second generation models; these are discussed in Section 6.

Lagrangean and surrogate relaxations of (MCLP)

Galvão, Espejo & Boffey (2000) compare heuristics based on Lagrangean and surrogate relaxations of (MCLP). Let $\lambda_j \geq 0$ be multipliers associated with constraints (28). The Lagrangean relaxation (MCLP $_{\lambda}$) is defined in the following way:

(MCLP $_{\lambda}$)

$$\begin{aligned} v(\text{MCLP}_{\lambda}) &= \max \left\{ \sum_{j \in J} \text{pop}_j \xi_j + \sum_{j \in J} \lambda_j \left[\sum_{i \in I} a_{ij} y_i - \xi_j \right] \right\} \\ &\equiv \max \left\{ \sum_{j \in J} (\text{pop}_j - \lambda_j) \xi_j + \sum_{i \in I} \sum_{j \in J} a_{ij} \lambda_j y_i \right\}. \end{aligned}$$

Or, making $\alpha_i = \alpha_i(\lambda) = \sum_{j \in J} a_{ij} \lambda_j$,

$$v(\text{MCLP}_{\lambda}) = \max \left\{ \sum_{j \in J} (\text{pop}_j - \lambda_j) \xi_j + \sum_{i \in I} \alpha_i y_i \right\}, \quad (32)$$

s.to (29)-(31).

The solution of this Lagrangean problem is straightforward and its optimal value is given by

$$v(\text{MCLP}_{\lambda}) = \sum_{j \in J} \max(0, \text{pop}_j - \lambda_j) + \sum^p \alpha_i,$$

where $\sum^p \alpha_i$ is the sum of p largest α_i , ties being broken arbitrarily, with $\xi_j = 1$ if $\lambda_j \leq \text{pop}_j$ ($\xi_j = 0$ otherwise) and $y_i = 1$ for p largest α_i ($y_i = 0$ otherwise). It is clear that this relaxation has the integrality property. A primal solution v_{primal} can be readily obtained from the optimal y_{λ}^* 's of (MCLP $_{\lambda}$): for each of the p $y_{\lambda i}^* = 1$ make $\xi_j^{\text{primal}} = 1$ if $a_{ij} = 1$, all other $\xi_j^{\text{primal}} = 0$. Then $v_{\text{primal}} = \sum_{j \in J} \text{pop}_j \xi_j^{\text{primal}}$ is a lower bound for (MCLP).

Galvão, Espejo & Boffey (2000) prove a theorem that states that there is an optimal solution of the Lagrangean dual

(DMCLP $_{\lambda}$)

$$v(\text{DMCLP}_{\lambda}) = \min_{\lambda \geq 0} v(\text{MCLP}_{\lambda})$$

for which $\lambda_j \leq \text{pop}_j$ for all $j \in J$; a corresponding Corollary asserts that $v(\text{DMCLP}_{\lambda}) = \sum_{j \in J} (\text{pop}_j - \lambda_j) + \sum^p \alpha_i$ for some λ with $0 \leq \lambda_j \leq \text{pop}_j$ for all $j \in J$. This allows the definition of a set of initial Lagrangean multipliers that considerably improves the solution of the Lagrangean dual.

Consider now the following surrogate relaxation of (MCLP). For any vector $\mu \geq 0$, $\mu \neq 0$ the authors combine cover constraints (28) into a single knapsack constraint:

(MCLP $^\mu$)

$$v(\text{MCLP}^\mu) = \max \sum_{j \in J} \text{pop}_j \xi_j$$

s.to

$$\sum_{j \in J} \mu_j \left\{ \sum_{i \in I} a_{ij} y_i - \xi_j \right\} \geq 0 \equiv \sum_{i \in I} \gamma_i y_i - \sum_{j \in J} \mu_j \xi_j \geq 0$$

and (29)-(31), where $\gamma_i = \sum_{j \in J} a_{ij} \mu_j$. Clearly those y_i corresponding to p largest γ_i should be set to 1. Call $y_{\mu i}^*$ the y_i 's set in this way and make $c = \sum_{i \in I} \gamma_i y_{\mu i}^*$; the resulting problem in ξ_j 's is the following 0-1 knapsack problem:

$$v(\text{MCLP}^\mu) = \max \sum_{j \in J} \text{pop}_j \xi_j$$

s.to

$$\sum_{j \in J} \mu_j \xi_j \leq c$$

and (30).

Galvão, Espejo & Boffey (2000) note that as the 0-1 knapsack problem is NP-hard (Garey & Johnson, 1979), a strategy often used (as for example in Lorena & Lopes, 1994; Lorena & Narciso, 1996) is to relax the integrality of the ξ_j 's. The solution of the resulting problem, called (RMCLP $^\mu$) by the authors, is straightforward (see Dudzinski & Waluckiewicz, 1984, 1987) with at most a single fractional variable.

The authors finally note that if $\mu' = \Psi \mu$, Ψ a strictly positive number, (MCLP $^\mu$) and (MCLP $^{\mu'}$) are equivalent problems and it follows that there is an optimal solution μ of the surrogate dual (DMCLP $^\mu$) for which $\mu_j \leq \text{pop}_j$ for all j (i.e., the theorem can also be used in this case). Similar to (MCLP $_\lambda$), a primal solution v_{primal} can be readily obtained from the optimal $y_{\mu i}^*$'s of (MCLP $^\mu$).

Having as a basis the developments shown above, Galvão, Espejo & Boffey (2000) compared the Lagrangean and surrogate heuristics (in the surrogate heuristic they solved the LP relaxation of the original 0-1 knapsack problem) using 331 test problems available in the literature, corresponding to networks ranging from 55 to 900 vertices. The gaps obtained with both heuristics were very low and did not differ substantially among themselves for the several problem sets used, in accordance with theoretical results reviewed in the paper. When the initial set of multipliers was determined using a valid bound for (MCLP) [cf. theorem] the computing times did not differ significantly between the Lagrangean and surrogate heuristics. For a further discussion on the use of Lagrangean and surrogate relaxations in the solution of integer programming problems the interested reader is referred to Espejo & Galvão (2002).

5. Hierarchical Models

Narula (1984) proposed a classification scheme for hierarchical location-allocation problems. The Narula classification scheme takes into account the facility hierarchy (relationship among various types of facility), the arc flow discipline and the node flow discipline. He defines two types of facility hierarchy: successively inclusive and successively exclusive.

In a *successively inclusive* hierarchy a facility provides its own level of service *and all lower levels of service*. The 2-level hierarchical extension of the maximal covering location problem (MCLP) that Moore & ReVelle (1982) applied to the health services of Honduras is an example of a successively inclusive hierarchy. Other examples of such hierarchy are banking and postal systems (branch offices, main offices). In a *successively exclusive* hierarchy a facility of a given type offers services unique to it, as for example in electricity distribution and telephone systems. Tien, El-Tell & Simons (1983) introduced the idea of a *locally inclusive hierarchy*, in which a facility of level ζ offers services of levels 1, 2, ..., ζ to users located in its neighbourhood, and only service of level ζ to users outside its neighbourhood.

Hierarchical models may have either a minimum objective, in which the total weighted travel distance to the facilities is minimized, or a maximal covering objective, in which the population covered is maximized, subject to service distance constraints. Narula & Ogbu (1985) developed a successively inclusive minimum model for an uncapacitated 2-hierarchical problem where p_1 level 1 facilities and p_2 level 2 facilities are to be located among n ($\geq p_1 + p_2$) potential sites, where a fraction θ ($0 \leq \theta \leq 1$) of the demand from a level 1 facility is referred to a level 2 facility and at most one facility may be located at a given site. They formulated their model as a mixed-integer programming problem and used Lagrangean relaxation to solve it.

Covering-type hierarchical models are often linked to emergency medical services (EMS), which generally consist of basic life support (BLS) units and advanced life support (ALS) units. EMS systems are usually defined in a successively inclusive hierarchical context. They are referred to as two-tiered EMS systems, see Mandell (1996). In a non-EMS context, Rahman & Smith (1999) studied the location of *Health and Family Welfare Centres* and *Community Clinics* in Bangladesh. Although the authors describe this as a hierarchical problem, they solved it by successively applying MCLP models in a non-hierarchical context.

The location of perinatal health care facilities in the municipality of Rio de Janeiro

Galvão, Espejo & Boffey (2002) developed a 3-level hierarchical model for the location of maternal and perinatal health care facilities in the municipality of Rio de Janeiro. The model was part of a project jointly funded by the Brazilian National Research Council (CNPq) and the British Council, and was carried out by the Department of Mathematical Sciences of the University of Liverpool and the Post-Graduate Department of Production Engineering of the Federal University of Rio de Janeiro. This project aimed at the development of mathematical models for the location of perinatal facilities in the municipality of Rio de Janeiro, with the overall objective of seeking to contribute to the reduction of perinatal mortality in the municipality, through a better spatial distribution of health care facilities.

In maternal and perinatal health care both mothers and babies may be classified in different categories of risk, according to certain clinical criteria. Vasconcellos (1997) proposes low, medium and high risk categories for babies and low and high risk categories for mothers. There are presently four main levels of facilities associated with maternal and perinatal care in Rio de Janeiro. These are:

- (i) *Basic Units* (Level 1) – These are low technology units, responsible for providing prenatal medical care to low risk mothers and babies.
- (ii) *Maternity Homes* (Level 2) – Maternity homes provide the basic care of level 1 units, plus prenatal care to medium risk babies. They are also responsible for routine births and neonatal assistance to low risk mothers and babies and to medium risk babies.
- (iii) *Neonatal Clinics* (Level 3) – They have the technological capability of level 1 and 2 units, plus that needed for non-routine births and perinatal assistance to high risk babies.
- (iv) *General Hospitals* (Level 4) – General Hospitals incorporate the technological capability of level 1, 2 and 3 units and are responsible for providing health care to high risk mothers. Given that general hospitals are concerned with much else other than child birth, their location was not considered in the model developed by the authors.

Considering the description of the different types of facilities above the authors developed a *successively inclusive* model, i.e., level 1 service may be obtained at a basic unit, maternity home or neonatal clinic and level 2 service may be obtained at a maternity home or neonatal clinic; level 3 service may only be obtained at a neonatal clinic.

The basic assumptions of the mathematical model are as follows: (i) Each mother-to-be will go to a basic unit, maternity home or neonatal clinic to obtain level 1 service; (ii) each mother-to-be will go to a maternity home or neonatal clinic to give birth; (iii) a proportion of mothers at maternity homes will be *referred* to a neonatal clinic and will be transported there by ambulance; (iv) a proportion of mothers-to-be attending basic unit service will be advised to go to a neonatal clinic to give birth. They will travel directly from home to the clinic; (v) since the model is successively inclusive, location of basic units, maternity homes and neonatal clinics at the same site is prohibited; (vi) travel is always to a nearest facility of appropriate level.

In their mathematical formulation of the model the authors used the following notation:

Index Sets

I : Set of mothers-to-be, $I = \{i | i = 1, 2, \dots, m\}$;

J : Set of potential facility sites, $J = \{j | j = 1, \dots, n\}$;

T : Set of types of service offered, $T = \{t | t = 1, 2, 3\}$.

Data

W_i : Number of mothers-to-be located in $i \in I$;

d_{ij} : shortest distance between demand point $i \in I$ and facility site $j \in J$;

d_{jk} : shortest distance between facility site $j \in J$ and facility site $k \in J$.

Parameters

α : factor to account for differences in travel costs ($0 < \alpha \leq 1$), given that travel of patients referred from maternity homes to neonatal clinics is by ambulance;

ϕ : proportion of mothers attending basic unit service that will be advised to go directly to a neonatal clinic to give birth;

θ : proportion of mothers at maternity homes referred to a neonatal clinic;

p_1, p_2 and p_3 : maximum number of basic units, maternity homes and neonatal clinics considered in the model;

M : big number.

Flow and Location Variables

x_{ij}^t : flow of mothers from demand point $i \in I$ to facility $j \in J$, to receive service level $t \in T$;

x_{jk}^4 : flow of mothers referred from maternity home $j \in J$ to neonatal clinic $k \in J$;

$y_j^t = \begin{cases} 1, & \text{if a facility of level } t \in T \text{ is located at } j \in J; \\ 0, & \text{otherwise.} \end{cases}$

A flow diagram of mothers is shown in Figure 1, where x_{ij}^1 , x_{ij}^2 and x_{ij}^3 represent respectively the flow of mothers to service levels 1, 2 and 3. As the hierarchy is successively inclusive, each facility is represented as a set of pseudo-clinics, each one offering a single level of service.

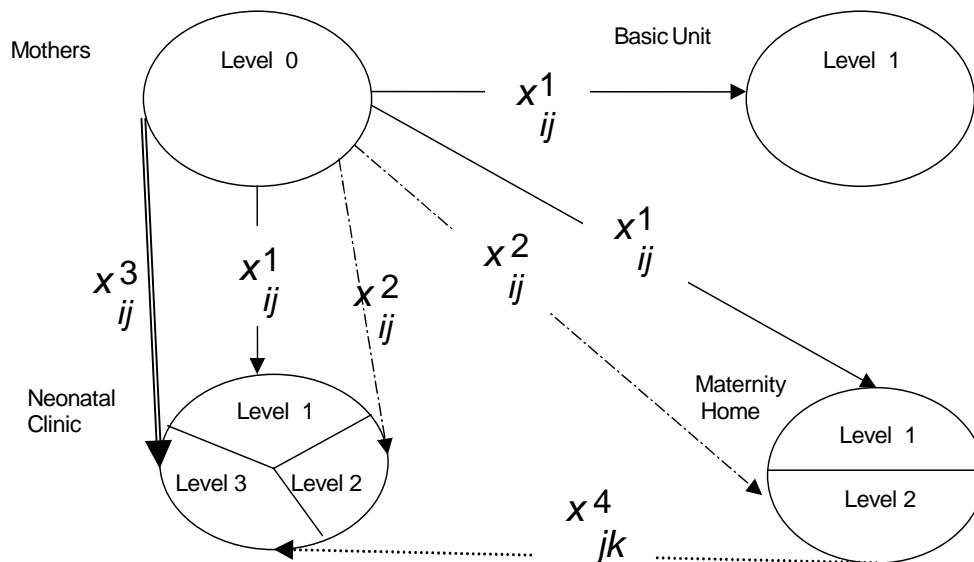


Figure 1 – Flow Diagram of Mothers

The mathematical formulation of the model (*HLPF*) is given by:

(HLPF)

$$v(HLPF) = \min \left\{ \sum_{i \in I} \sum_{j \in J} d_{ij} (x_{ij}^1 + x_{ij}^2 + x_{ij}^3) + \alpha \sum_{j \in J} \sum_{k \in J} d_{jk} x_{jk}^4 \right\} \quad (33)$$

s.to

$$\sum_{j \in J} x_{ij}^1 = W_i, \quad i \in I \quad (34)$$

$$\sum_{j \in J} x_{ij}^2 = (1 - \phi) W_i, \quad i \in I \quad (35)$$

$$\sum_{j \in J} x_{ij}^3 = \phi W_i, \quad i \in I \quad (36)$$

$$\sum_{k \in J} x_{jk}^4 = \theta \sum_{i \in I} x_{ij}^2, \quad j \in J \quad (37)$$

$$\sum_{i \in I} x_{ij}^1 \leq M(y_j^1 + y_j^2 + y_j^3), \quad j \in J \quad (38)$$

$$\sum_{i \in I} x_{ij}^2 \leq M(y_j^2 + y_j^3), \quad j \in J \quad (39)$$

$$\sum_{i \in I} x_{ij}^3 + \sum_{k \in J} x_{kj}^4 \leq M y_j^3, \quad j \in J \quad (40)$$

$$y_j^1 + y_j^2 + y_j^3 \leq 1, \quad j \in J \quad (41)$$

$$\sum_{j \in J} y_j^1 \leq p_1 \quad (42)$$

$$\sum_{j \in J} y_j^2 \leq p_2 \quad (43)$$

$$\sum_{j \in J} y_j^3 \leq p_3 \quad (44)$$

$$x_{ij}^1, x_{ij}^2, x_{ij}^3, x_{jk}^4 \geq 0, \quad i \in I, \quad j \in J, \quad k \in J \quad (45)$$

$$y_j^1, y_j^2, y_j^3 \in \{0, 1\}, \quad j \in J. \quad (46)$$

In the formulation above restrictions (34) ensure level 1 service for all mothers-to-be and restrictions (35)-(36) state that births take place at either a maternity home or a neonatal clinic (restrictions (36) reflect the proportion of mothers-to-be needing level 3 service to give birth). Restrictions (37) are the referral constraints. Restrictions (38) to (40) state, respectively, that level 1, 2 and 3 service can only be obtained at points where appropriate level facilities are located. Restrictions (41) avoid location of different types of facility at the same site, restrictions (42) to (44) are budget constraints and restrictions (45) and (46) define the nature of the variables. Notice that as the only “costs” in the objective function are distances (no fixed costs are present), restrictions (42) to (44) will be always satisfied as equalities in the optimal solution.

Two basic heuristics were developed to solve the 3-level hierarchical model to locate perinatal facilities in the municipality of Rio de Janeiro: a Lagrangean Heuristic (LH) and a

heuristic based on the solution of 3 successive p -median problems (the 3 p -Median Heuristic). LH was then modified to include an initial upper bound calculated by the 3 p -median heuristic; new strategies were also tested to update the step size, resulting in a Modified Lagrangean Heuristic (MLH). These three heuristics were tested on problems available in the literature.

The qualities of the solutions produced by the three procedures did not differ appreciably among themselves, although MLH consistently produced tighter lower bounds. The model was then tested in a case study that used real 1995 data of the municipality of Rio de Janeiro. The results for the different scenarios considered in the case study, when compared with the actual location of facilities in 1995, showed an improved spatial distribution of the units at all three levels of the hierarchy.

An alternative approach towards solving the 3-level hierarchical problem was taken by Boffey, Yates & Galvão (2003), who used a model that regarded each facility as a collection of (*pseudo*) *clinics*, each providing just one 'level of service'. The objective function of their formulation includes a non-linear term and the authors developed a genetic algorithm to solve the problem. The results obtained with the genetic algorithm were of very similar quality to those obtained by Galvão, Espejo & Boffey (2002) with their Lagrangean heuristics. This added to the authors' confidence that both approaches yield near optimal solutions.

An important issue that arose during the course of the research was the need to include some form of capacity constraints into the model, especially in the higher, resource intensive level of the hierarchy. This issue is discussed in Galvão *et al.* (2004). In the capacitated model discussed in that paper two different situations arise in practice: (i) existing capacity at level 3 is appropriate, in which case the problem becomes one of load balancing among the level 3 services; (ii) existing capacity is insufficient; in this case the problem becomes one of how to locate these services and allocate with equity, among the population, the demand that can be met. In the case of situation (i) the problem takes the form of load balancing among level 3 services. The capacitated model was extended by Galvão *et al.* (2004) to deal with load balancing, taking the form of a bi-criterion model that seeks to minimize both total distance traveled and load imbalance among level 3 services.

Although in theoretical terms the research may be classified as a success, the models developed were not implemented by the municipality health authorities. Among the reasons for this outcome we may mention political motivations and the lack of a stable civil service in developing countries. On the positive side we may point out a long term cooperation that was established between British and Brazilian OR scientists and the development of models and solution procedures that may prove useful in other circumstances.

A hierarchical covering location model

Espejo, Galvão & Boffey (2003) considered the 2-level hierarchical extension of the Maximal Covering Location Problem proposed by Moore & ReVelle (1982) for the health services of Honduras. In this model the lower level facilities (clinics) provide only a level one service, whereas the higher level facilities (hospitals) provide *both* a level two service *and* the level one service. The hierarchy is therefore successively inclusive. Moore & ReVelle formulated the problem as a 0-1 programming problem and used the linear programming relaxation of their formulation to solve it. Their computational experience is however restricted to a test network developed from provinces in Honduras.

It is important to note that in the Moore & ReVelle model coverage is defined in terms of access to services, not in terms of access to facilities. Thus a demand area is considered to be covered if it has access, within the defined service distances, to *both* level 1 *and* level 2 services in the two-level hierarchy. As the hierarchy of this model is successively inclusive, the lower level facilities offer level 1 service, whereas the higher level facilities offer *both* level 1 *and* level 2 services. The service distances defined by Moore & ReVelle are different for the two levels of service, and the service distance for level 1 service is permitted to be different at the two types of facilities.

The rationale behind the definition of the service distances is the following. Let R_1 be the service distance for level 1 service provided by the lower level facility. In principle this service distance could be considered to be equal for the same type of service provided by the higher level facility, but in practice people may be prepared to travel an extra distance to obtain the same service from a facility with more resources. So T_1 , the service distance for level 1 service provided at the higher level facility, is supposed to satisfy $T_1 > R_1$. On the other hand, let R_2 be the service distance for level 2 service. This type of service is offered only by the higher level facilities and in practice people will be prepared to travel greater distances to obtain the more sophisticated level 2 services, so we will consider in this model that $R_2 > T_1 > R_1$.

Espejo, Galvão & Boffey (2003) gave the following mathematical programming formulation for this hierarchical problem, which they called the Hierarchical Covering Location Problem (HCLP):

(HCLP)

$$v(HCLP) = \text{Max} \left\{ \sum_{j \in J} \text{pop}_j \xi_j \right\} \quad (47)$$

s.to

$$\sum_{i \in I} a_{ij} y_i + \sum_{i \in I} b_{ij} z_i - \xi_j \geq 0, \quad j \in J \quad (48)$$

$$\sum_{i \in I} c_{ij} z_i - \xi_j \geq 0, \quad j \in J \quad (49)$$

$$\sum_{i \in I} y_i = p \quad (50)$$

$$\sum_{i \in I} z_i = q \quad (51)$$

$$\xi_j \in \{0, 1\}, \quad j \in J \quad (52)$$

$$y_i, z_i \in \{0, 1\}, \quad i \in I, \quad (53)$$

where $J = \{1, 2, \dots, m\}$ is the set of demand areas, $I = \{1, 2, \dots, n\}$ is the set of potential facility sites, pop_j is the population of demand area j , $a_{ij} = 1$ if demand area j can be covered by level 1 service (within distance R_1) offered at a lower level facility located at $i \in I$ ($a_{ij} = 0$ otherwise), $b_{ij} = 1$ if demand area j can be covered by level 1 service (within distance T_1) offered at a higher level facility located at $i \in I$ ($b_{ij} = 0$ otherwise), $c_{ij} = 1$ if demand area j can

be covered by level 2 service (within distance R_2) offered at a higher level facility located at $i \in I$ ($c_{ij} = 0$ otherwise), p is the number of lower level facilities to be located, q is the number of higher level facilities to be located and ξ_j , y_i and z_i are the decision variables. $\xi_j = 1$ if demand area j is covered ($\xi_j = 0$ otherwise); $y_i = 1$ means that a lower level facility is located at site $i \in I$ ($y_i = 0$ otherwise); $z_i = 1$ means that a higher level facility is located at site $i \in I$ ($z_i = 0$ otherwise).

In the formulation above the objective function to be maximized represents the total population covered by *both* level 1 *and* level 2 services. Constraints (48) state that a demand area $j \in J$ is covered by level 1 service if there is at least *either* one lower level facility within distance R_1 *or* one higher level facility within distance T_1 . Constraints (49) state that a demand area $j \in J$ is covered by level 2 service if there is at least one higher level facility within distance R_2 . Constraint (50) limits the number of lower level facilities in the solution to p , whereas constraint (51) limits the number of higher level facilities in the solution to q . Finally, constraints (52)-(53) define the binary nature of the decision variables.

Espejo, Galvão & Boffey developed a combined Lagrangean-surrogate (L-S) relaxation of (HCLP), which reduces to a 0-1 knapsack problem. Tests were carried out using a subgradient-based heuristic incorporating the L-S relaxation, with the resulting knapsack problems being solved both with and without the integrality constraints relaxed. Results were obtained for test problems available in the literature ranging from 55-node to 700-node networks. These were compared, where possible, with exact results obtained using CPLEX. The computational experience reported suggests that, in practical terms, solving (HCLP) has a similar degree of difficulty to that previously reported for (MCLP) by Galvão, Espejo & Boffey (2000).

6. Probabilistic Models

Probabilistic location problems deal with the stochastic nature of real-world systems. In these systems some parameters, such as for example travel times, the location of clients, demand and the availability of servers are treated as random variables. The objective is to determine robust server/facility locations that optimize a given utility function, for a range of values of the parameters under consideration. According to Owen & Daskin (1998), the corresponding models capture the stochastic aspects of the problems through the explicit consideration of the probability distributions of the random variables; some authors incorporate these distributions into standard mathematical programming formulations, while others use them within a queuing framework.

A detailed review of probabilistic location problems can be found in Owen & Daskin (1998). Swersey (1994), Chiyoshi, Galvão & Morabito (2000) and Brotcorne, Laporte & Semet (2003) also examine some probabilistic models. When modeling emergency systems the use of simplifying assumptions may allow the definition of mathematical programming models, for example through the definition of chance constraints. Situations in which the simplifying assumptions are not applicable lead to models based on spatially distributed queues, in which Larson's Hypercube Model (1974, 1975) is of paramount importance.

The hypercube model, proposed by Larson (1974) and studied by several authors (see Burwell, Jarvis & McKnew, 1993; Swersey, 1994), is an important tool for planning service systems, especially urban systems in which servers travel to offer some type of service to

clients (server-to-customer service). The model treats geographical and temporal complexities of the region under study, based on the theory of spatially distributed queues. Basically, the idea is to expand the description of the state space of a queuing system with multiple servers, in order to represent each server individually and incorporate more complex dispatch policies.

Our interest lies in models that treat the availability of servers as a random variable within a queuing framework, of particular importance for the location of emergency services. Strictly speaking, probabilistic location models cannot be considered uncapacitated. When only queues of finite capacity are allowed (note that no queues allowed represents the particular case of zero queue capacity), the number of servers, plus the maximum allowed queue size, may be considered to represent the capacity of the system. Even when infinite queues are allowed, long waiting times may render these systems capacitated in practice.

The reason we decided to include a short section on probabilistic models is that our work on them constituted a natural continuation of our work on deterministic covering models. A limitation of the deterministic models is that they assume that servers are available when requested, which is not always true in practical situations. In non-congested systems, with little demand, the assumption is reasonable, but in congested systems, in which frequent calls for service may for example keep ambulances busy 20% to 30% of the time, the assumption is totally unjustifiable. Congestion in emergency services, which may cause the unavailability of servers within the critical distance when a call is placed, lead to the development of probabilistic covering models. Non-homogeneous servers were studied in Chiyoshi, Galvão & Morabito (2001).

Important models that treat the availability of servers as a random variable are the Maximum Expected Covering Location Problem (MEXCLP) of Daskin (1983) and the Maximum Availability Location Problem (MALP) of ReVelle & Hogan (1989). In both models simplifying assumptions lead to the definition of mathematical programming models: the assumption that servers operate independently is a common feature to both models. Daskin also assumes that each server has the same busy probability. ReVelle & Hogan define two variations for MALP: MALPI, where the authors assume, as Daskin, that each server has the same busy probability, and MALPII, where they allow busy fractions to be different in the various sections of a region under consideration.

Batta, Dolan & Krishnamurthy (1989) suggest that an approximate way to relax the server independence assumption of MEXCLP is to use the hypercube correction factor developed by Larson (1975). This correction factor, applied to the MEXCLP objective function, leads to an “adjusted” model, which the authors called AMEXCLP. These authors also relaxed MEXCLP’s simplifying assumptions by embedding the hypercube model into a single node vertex substitution heuristic procedure, defining an extended model (which we call EMEXCLP) that seeks to determine a set of server locations that maximizes expected coverage.

Chiyoshi, Galvão & Morabito (2003) observe that the three models compared by Batta, Dolan & Krishnamurthy are in fact not strictly comparable: by analyzing their objective functions it can be seen that, while both MEXCLP and AMEXCLP are restricted to coverage arising from unqueued calls, the embedding of the hypercube model into a vertex substitution heuristic takes into account unqueued as well as queued calls to predict expected coverage. Unless the system is operating at a very low overall workload or under very restrictive cover constraints, in which case no significant contribution to coverage is

expected from queued calls, the models should produce different expected coverages due to the very nature of their objective functions.

Galvão *et al.* (2003) used the same approach as Batta, Dolan & Krishnamurthy to relax the simplifying assumptions of the MALPI model, obtaining an extended model (EMALP) that seeks to maximize the population covered with a predefined reliability α . In both cases the extended models (EMEXCLP, EMALP) are able to deal with server cooperation and the definition of busy fractions for each individual server, which reflects more precisely the situation in real-world systems. The idea in both cases is to reproduce conditions that are closer to those expected in practical applications.

Finally, Galvão, Chiyoshi & Morabito (2003) give a unified view of the MEXCLP and MALPI models, identifying similarities and dissimilarities between these models and showing how they relate to each other. They also develop a mathematical formulation of EMALP. Even though this formulation is of little practical use because of its complexity, it has a structure that allows the development of heuristic solution methods.

Galvão, Chiyoshi & Morabito used Simulated Annealing (SA) in an attempt to enhance the local searches used by Batta, Dolan & Krishnamurthy (1989) and by Galvão *et al.* (2003) for the extended MEXCLP and MALPI models, respectively. The results produced by the SA algorithm are particularly important for the extended MALPI model: since in this model the coverage provided by different solutions is at great variance, the better solutions produced by the SA methodology are of considerable practical importance.

7. Conclusions

My research activities have been concentrated, over the past 30 years, in the area of location and distribution models. Discrete location models in general, and uncapacitated models in particular, have become my main area of research within this time period. My interests have evolved from simpler prototype models, such as (PM) and (SPLP), to more complex and sophisticated models: dynamic, hierarchical and probabilistic location models. This has been an evolutionary process, since the simpler models are often embedded into the more sophisticated ones.

The techniques used to solve the models developed encompass both exact and heuristic (meta-heuristic) methods. Dual-based methods and different types of relaxation (Lagrangean, surrogate and combined Lagrangean-surrogate) have been widely used. In some applications these were used in procedures to find near-optimal solutions, in some others they were complemented by primal procedures (such as branch-and-bound algorithms) that sought optimal solutions of the problems under consideration.

My contributions encompass theoretical developments, computational algorithms and practical applications. Practical applications, unfortunately, represent only a small fraction of the work conducted and were not always implemented by the client. The location of facilities for the distribution of benefits in the Brazilian Social Security System (Galvão & Nascimento, 1990) is a success story, but the hierarchical models developed for the location of maternal and perinatal health care facilities in the municipality of Rio de Janeiro (Galvão, Espejo & Boffey, 2002; Boffey, Yates & Galvão, 2003) were not implemented by the municipality health authorities. Among the reasons for this outcome we may mention political motivations and the lack of a stable civil service in developing countries.

All work took place in an academic environment, with the invaluable collaboration of colleagues (both in Brazil and abroad) and research students at COPPE, Federal University of Rio de Janeiro. It has been my pleasure to work all these years in such intellectually stimulating environment and I take this opportunity to thank all my colleagues and students for their help and encouragement.

Acknowledgements

This research was supported by the Brazilian National Research Council (CNPq grant 471919/01-2) and by the Research Foundation of the state of Rio de Janeiro (FAPERJ grant E26/151.934/2000).

References

- (1) Balinski, M.L. (1965). Integer programming: methods, uses, computation. *Management Science*, **12**, 253-313.
- (2) Batta, R.; Dolan, J.M. & Krishnamurthy, N.N. (1989). The maximal expected covering location problem: Revisited. *Transportation Science*, **23**, 277-287.
- (3) Batta, R. & Mannur, N.R. (1990). Covering-location models for emergency situations that require multiple response units. *Management Science*, **36**, 16-23.
- (4) Beasley, J.E. (1985). A note on solving large p -median problems. *European Journal of Operational Research*, **21**, 270-273.
- (5) Beasley, J.E. (1990). OR-Library: distribution of test problems by electronic mail. *Journal of the Operational Research Society*, **41**, 1069-1072.
- (6) Bilde, O. & Krarup, J. (1977). Sharp lower bounds and efficient algorithms for the simple plant location problem. *Annals of Discrete Mathematics*, **1**, 79-97.
- (7) Boffey, T.B. & Karkazis, J. (1984). P -medians and multi-medians. *Journal of the Operational Research Society*, **35**, 57-64.
- (8) Boffey, T.B. & Narula, C.S. (1997). Multiobjective covering and routing problems. **In:** *Essays in Decision Making: A Volume in Honor of Stanley Zionts* [edited by M. Karwan, J. Sprong & J. Wallenius], Springer, Berlin, 342-370.
- (9) Boffey, B.; Yates, D. & Galvão, R.D. (2003). An algorithm to locate perinatal facilities in the municipality of Rio de Janeiro. *Journal of the Operational Research Society*, **54**, 21-31.
- (10) Bornstein, C.T.; Pereira, A.B.; Galvão, R.D. & Campêlo Neto, M.B. (2004). An algorithm for the simple plant location problem based on reduction tests and add/drop heuristics. Submitted to the *Journal of Heuristics*.
- (11) Brotcorne, L.; Laporte, G. & Semet, F. (2003). Ambulance location and relocation models. *European Journal of Operational Research*, **147**, 451-463.
- (12) Burwell, T.H.; Jarvis, J.P. & McKnew, M.A. (1993). Modeling co-located servers and dispatch ties in the hypercube model. *Computers & Operations Research*, **20**, 113-119.

- (13) Chiyoshi, F.Y. & Galvão, R.D. (2000). A statistical analysis of simulated annealing applied to the p -median problem. *Annals of Operations Research*, **96**, 61-74.
- (14) Chiyoshi, F.Y.; Galvão, R.D. & Morabito, R. (2000). O uso do modelo hipercubo na solução de problemas de localização probabilísticos (The use of the hypercube model in the solution of probabilistic location problems). *Gestão & Produção*, **7**, 146-174.
- (15) Chiyoshi, F.Y.; Galvão, R.D. & Morabito, R. (2001). Modelo hipercubo: análise e resultados para o caso de servidores não-homogêneos (Hypercube model: analysis and results for the case of non-homogeneous servers). *Pesquisa Operacional*, **21**, 199-218.
- (16) Chiyoshi, F.Y.; Galvão, R.D. & Morabito, R. (2003). A note on solutions to the maximal expected covering location problem. *Computers & Operations Research*, **30**, 87-96.
- (17) Christofides, N. & Beasley, J.E. (1982). A tree search algorithm for the p -median problem. *European Journal of Operational Research*, **10**, 196-204.
- (18) Christofides, N. & Eilon, S. (1969). An algorithm for the vehicle dispatching problem. *Operational Research Quarterly*, **20**, 309-318.
- (19) Christofides, N. & Eilon, S. (1972). Algorithms for large scale traveling salesman problems. *Operational Research Quarterly*, **23**, 511-518.
- (20) Chung, C.H. (1986). Recent applications of the Maximal Covering Location Problem (MCLP) model. *Journal of the Operational Research Society*, **37**, 735-746.
- (21) Church, R.L. & ReVelle, C.S. (1974). The maximal covering location problem. *Papers of the Regional Science Association*, **32**, 101-118.
- (22) Cornuejols, G.; Fisher, M.L. & Nemhauser, G.L. (1977). Location of bank accounts to optimize float: an analytic study of exact and approximate algorithms. *Management Science*, **23**, 789-810.
- (23) Cornuejols, G.; Sridharan, R. & Thizy, J.M. (1991). A comparison of heuristics and relaxations for the capacitated plant location problem. *European Journal of Operational Research*, **50**, 280-297.
- (24) Current, J.R. & O'Kelly, M. (1992). Locating emergency warning sirens. *Decision Sciences*, **23**, 221-234.
- (25) Daskin, M. & Stern, E. (1981). A hierarchical objective set covering model for emergency medical service vehicle deployment. *Transportation Science*, **15**, 137-152.
- (26) Daskin, M. (1983). The maximal expected covering location model: formulation, properties and heuristic solution. *Transportation Science*, **17**, 48-70.
- (27) Daskin, M.S.; Jones, P.C. & Lowe, T.J. (1990). Rationalizing tool selection in a flexible manufacturing system for sheet metal products. *Operations Research*, **38**, 1104-1115.
- (28) Diehr, G. (1972). An algorithm for the p -median problem. Working Paper No. 191, Los Angeles: Western Management Science Institute, UCLA.
- (29) Downs, B.T. & Camm, J.D. (1996). An exact algorithm for the maximal covering location problem. *Naval Research Logistics Quarterly*, **43**, 435-461.

-
- (30) Dudzinski, K. & Waluckiewicz, S. (1984). A fast algorithm for the linear multiple choice knapsack problem. *Operational Research Letters*, **3**, 205-209.
- (31) Dudzinski, K. & Waluckiewicz, S. (1987). Exact methods for the knapsack problem and its generalizations. *European Journal of Operational Research*, **28**, 3-21.
- (32) Dwyer, F.R. & Evans, J.R. (1981). A branch and bound algorithm for the list selection problem in direct mail advertising. *Management Science*, **27**, 658-667.
- (33) Eaton, D.; Church, R.L.; Bennet, V. & Namon, B. (1981). On deployment of health resources in rural Colombia. *TIMS Studies in Management Science*, **17**, 331-359.
- (34) Eaton, D.; Hector, M.; Sanchez, V.; Latingua, R. & Morgan, J. (1986). Determining ambulance deployment in Santo Domingo, Dominican Republic. *Journal of the Operational Research Society*, **37**, 113-126.
- (35) Eilon, S. & Galvão, R.D. (1978). Single and double vertex substitution in heuristic procedures for the p -median problem. *Management Science*, **24**, 1763-1766.
- (36) El-Shaieb, A.M. (1973). A new algorithm for locating sources among destinations. *Management Science*, **20**, 221-231.
- (37) Erlenkotter, D. (1978). A dual-based procedure for uncapacitated facility location. *Operations Research*, **26**, 992-1009.
- (38) Espejo, L.G.A. & Galvão, R.D. (2002). O uso das relaxações Lagrangeana e *surrogate* em problemas de programação inteira (The use of Lagrangean and surrogate relaxations in integer programming problems). *Pesquisa Operacional*, **22**, 387-402.
- (39) Espejo, L.G.A.; Galvão, R.D. & Boffey, B. (2003). Dual-based heuristics for a hierarchical covering location problem. *Computers & Operations Research*, **30**, 165-180.
- (40) Galvão, R.D. (1977). The optimal location of facilities on a network. *Ph.D. Thesis*, Imperial College of Science and Technology, London, U.K.
- (41) Galvão, R.D. (1980). A dual-bounded algorithm for the p -median problem. *Operations Research*, **28**, 1112-1121.
- (42) Galvão, R.D. (1981). A note on Garfinkel, Neebe and Rao's LP decomposition for the p -median problem. *Transportation Science*, **15**, 175-182.
- (43) Galvão, R.D. & Raggi, L.A. (1989). A method for solving to optimality uncapacitated location problems. *Annals of Operations Research*, **18**, 225-244.
- (44) Galvão, R.D. & Nascimento, E.M. (1990). The location of benefit-distributing facilities in the Brazilian social security system. *Operational Research '90 (Proceedings of the IFORS 1990 Conference)*, 433-443.
- (45) Galvão, R.D. & Santibañez-Gonzalez, E. del R. (1990). The p_k -median dynamic location problem: formulation and a heuristic solution method. *Investigación Operativa*, **1**, 295-308.
- (46) Galvão, R.D. & Santibañez-Gonzalez, E. del R. (1992). A Lagrangean heuristic for the p_k -median dynamic location problem. *European Journal of Operational Research*, **58**, 250-262.

- (47) Galvão, R.D. (1993). The use of Lagrangean relaxation in the solution of uncapacitated facility location problems. *Location Science*, **1**, 57-79.
- (48) Galvão, R.D.; Ferreira Filho, V.J.M. & Rivas M.P.A. (1996). Some computational aspects of p -median type location problems. *Proceedings of VIII CLAIO/XXVIII Symposium of the Brazilian O.R. Society*, Rio de Janeiro, Brazil, 1266-1271.
- (49) Galvão, R.D. & ReVelle, C.S. (1996). A Lagrangean heuristic for the maximal covering location problem. *European Journal of Operational Research*, **88**, 114-123.
- (50) Galvão, R.D.; Espejo, L.G.A. & Boffey, B. (2000). A comparison of Lagrangean and surrogate relaxations for the maximal covering location problem. *European Journal of Operational Research*, **124**, 377-389.
- (51) Galvão, R.D.; Espejo, L.G.A. & Boffey, B. (2002). A hierarchical model for the location of perinatal facilities in the municipality of Rio de Janeiro. *European Journal of Operational Research*, **138**, 495-517.
- (52) Galvão, R.D.; Espejo, L.G.A.; Boffey, B. & Yates, D. (2004). Load balancing and capacity constraints in a hierarchical location model. Submitted to the *European Journal of Operational Research*.
- (53) Galvão, R.D.; Chiyoshi, F.Y. & Morabito, R. (2003). Towards unified formulations and extensions of two classical probabilistic location models. To appear in *Computers & Operations Research*.
- (54) Galvão, R.D.; Chiyoshi, F.Y.; Espejo, L.G.A. & Rivas M.P.A. (2003). Solução do problema de localização de máxima disponibilidade utilizando o modelo hipercubo (Solution of the maximum availability location problem using the hypercube model). *Pesquisa Operacional*, **23**, 61-78.
- (55) Garey, M.R. & Johnson, D.S. (1979). *Computers and Intractability: A Guide to the Theory of NP-completeness*. W.H. Freeman and Co., San Francisco.
- (56) Garfinkel, R.S.; Neebe, A.W. & Rao, M.R. (1974). An algorithm for the M -median plant location problem. *Transportation Science*, **8**, 217-236.
- (57) Guignard, M. (1988). A Lagrangean dual ascent algorithm for simple plant location problems. *European Journal of Operational Research*, **35**, 193-200.
- (58) Hakimi, S.L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, **12**, 450-459.
- (59) Hakimi, S.L. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research*, **13**, 462-475.
- (60) Hanjoul, P. & Peeters, D. (1985). A comparison of two dual-based procedures for solving the p -median problem. *European Journal of Operational Research*, **20**, 287-296.
- (61) Hansen, P. & Mladenovic, N. (1998). Variable neighborhood search for the p -median. *Location Science*, **5**, 207-226.
- (62) Hansen, P.; Mladenovic, N. & Perez-Brito, D. (1998). Variable neighborhood decomposition search. Les Cahiers du GERARD, G-98-53, Montreal.

- (63) Held, M.; Wolfe, P. & Crowder, H. (1974). Validation of subgradient optimization. *Mathematical Programming*, **6**, 62-88.
- (64) Hogan, K. & ReVelle, C.S. (1986). Concepts and applications of backup coverage. *Management Science*, **32**, 1434-1444.
- (65) Houghland, E.S. & Stephens, N.T. (1976). Air pollutant monitor siting by analytical techniques. *Journal of the Air Pollution Control Association*, **26**, 52-53.
- (66) Järvinen, P.; Rajala, J. & Sinervo, H. (1972). A branch-and-bound algorithm for seeking the p -median. *Operations Research*, **20**, 173-178.
- (67) Karg, R.L. & Thompson, G.L. (1964). A heuristic approach to solving traveling salesman problems. *Management Science*, **10**, 225-248.
- (68) Kerningan, B.W. & Lin, S. (1970). An efficient heuristic procedure for partitioning graphs. *Bell System Technical Journal*, **49**, 291-307.
- (69) Körkel, M. (1989). On the exact solution of large-scale simple plant location problems. *European Journal of Operational Research*, **39**, 157-173.
- (70) Krarup, J. & Pruzan, P.M. (1983). The simple plant location problem: survey and synthesis. *European Journal of Operational Research*, **12**, 36-81.
- (71) Kuehn, A.A. & Hamburger, M.J. (1963). A heuristic program for locating warehouses. *Management Science*, **9**, 643-666.
- (72) Larson, R.C. (1974). A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research*, **1**, 67-95.
- (73) Larson, R.C. (1975). Approximating the performance of urban emergency service systems. *Operations Research*, **23**, 845-868.
- (74) Lin, S. (1965). Computer solutions of the traveling salesman problem. *Bell System Technical Journal*, **44**, 2245-2269.
- (75) Lorena, L.A.N. & Lopes, F.B. (1994). A surrogate heuristic for set covering problems. *European Journal of Operational Research*, **79**, 138-150.
- (76) Lorena, L.A.N. & Narciso, M.G. (1996). Relaxation heuristics for a generalized assignment problem. *European Journal of Operational Research*, **91**, 600-610.
- (77) Mandell, M.B. (1996). Covering models for two-tiered emergency medical services systems. *Location Science*, **6**, 355-368.
- (78) Maranzana, F.E. (1964). On the location of supply points to minimize transport costs. *Operational Research Quarterly*, **15**, 261-270.
- (79) Marsten, R.E. (1972). An algorithm for finding almost all of the medians of a network. Discussion Paper No. 23, Northwestern University.
- (80) Mirchandani, P.B.; Oudjit, A. & Wong, R.T. (1985). 'Multidimensional' extensions and a nested dual approach for the m -median problem. *European Journal of Operational Research*, **21**, 121-137.
- (81) Mladenovic, N.; Moreno, J.P. & Moreno-Vega, J. (1996). A chain-interchange heuristic method. *Yugoslav Journal of Operations Research*, **6**, 41-54.

- (82) Moore, G.C. & ReVelle, C.S. (1982). The hierarchical service location problem. *Management Science*, **28**, 775-780.
- (83) Narula, S.C. (1984). Hierarchical location-allocation problems: a classification scheme. *European Journal of Operational Research*, **15**, 93-99.
- (84) Narula, S.C.; Ogbu, U.I. & Samuelsson, H.M. (1977). An algorithm for the p -median problem. *Operations Research*, **25**, 709-713.
- (85) Narula, S.C. & Ogbu, U. I. (1985). Lagrangean relaxation and decomposition in an uncapacitated 2-hierarchical location-allocation problem. *Computers & Operations Research*, **12**, 169-180.
- (86) Owen, S.H. & Daskin, M.S. (1998). Strategic facility location: a review. *European Journal of Operational Research*, **111**, 423-447.
- (87) Pastor, J.T. (1994). Bi-criterion programs and managerial locations: application to the banking sector. *Journal of the Operational Research Society*, **45**, 1351-1362.
- (88) Pizzolato, N.D. (1994). A heuristic for large size p -median location problems with application to school location. *Annals of Operations Research*, **50**, 473-485.
- (89) Pizzolato, N.D. & Fraga da Silva, H.B. (1997). The location of public schools: evaluation of practical experiences. *International Transaction in Operations Research*, **4**, 13-22.
- (90) Rahman, S. & Smith, D.K. (1999). Deployment of rural health facilities in a developing country. *Journal of the Operational Research Society*, **50**, 892-902.
- (91) ReVelle, C.S. (1987). Urban facility location. **In:** *Handbook of Urban and Regional Economics* [edited by E. Mills], **2**, North-Holland, Amsterdam.
- (92) ReVelle, C.S. (1989). Review, extensions and predictions in emergency service siting models. *European Journal of Operational Research*, **40**, 58-69.
- (93) ReVelle, C.S. & Swain, R.W. (1970). Central facilities location. *Geographical Analysis*, **2**, 30-42.
- (94) ReVelle, C.S. & Hogan, K. (1989). The maximum availability location problem. *Transportation Science*, **23**, 192-200.
- (95) Rolland, E.; Schilling, D.A. & Current, J.R. (1996). An efficient tabu search procedure for the p -median problem. *European Journal of Operational Research*, **96**, 329-342.
- (96) Roodman, G.M. & Schwarz, L.B. (1975). Optimal and heuristic facility phase-out strategies. *AIIE Transactions*, **7**, 177-184.
- (97) Roodman, G.M. & Schwarz, L.B. (1977). Extensions of the multi-period facility phase-out model: new procedures and applications to a phase-in/phase-out problem. *AIIE Transactions*, **9**, 103-107.
- (98) Rosing, K.E. & ReVelle, C.S. (1997). Heuristic concentration: two stage solution construction. *European Journal of Operational Research*, **97**, 75-86.
- (99) Rosing, K.E.; ReVelle, C.S.; Rolland, E.; Schilling, D.A. & Current, J.R. (1998). Heuristic concentration and tabu search: a head to head comparison. *European Journal of Operational Research*, **104**, 93-99.

- (100) Scott, A.J. (1971). Dynamic location-allocation systems: some basic planning strategies. *Environment and Planning*, **3**, 73-82.
- (101) Swain, R.W. (1974). A parametric decomposition approach for the solution of uncapacitated location problems. *Management Science*, **21**, 189-198.
- (102) Swersey, A.J. (1994). The deployment of police, fire and emergency medical units. **In:** *Handbooks in OR & MS* [edited by S.M. Pollock *et al.*], **6**, Elsevier Science B.V., 151-200.
- (103) Tcha, D.-W.; Ro, H.-B. & Yoo, C.-B. (1988). A dual-based add heuristic for uncapacitated facility location. *Journal of the Operational Research Society*, **39**, 873-878.
- (104) Teitz, M.B. & Bart, P. (1968). Heuristic methods for estimating the generalized vertex median of a weighted graph. *Operations Research*, **16**, 955-961.
- (105) Tien, J.M.; El-Tell, G.R. & Simons, G.R. (1983). Improved formulations to the hierarchical health facility location-allocation problem. *IEEE Transactions on Systems, Man and Cybernetics*, **SMC-13**, 1128-1132.
- (106) Van Roy, T.J. & Erlenkotter, D. (1982). A dual-based procedure for dynamic facility location. *Management Science*, **28**, 1091-1105.
- (107) Vasconcellos, M.M. (1997). Modelos de localização e Sistemas de Informações Geográficas na assistência materna e perinatal: uma aplicação no município do Rio de Janeiro (Location models and Geographical Information Systems in maternal and perinatal assistance: an application in the municipality of Rio de Janeiro). *Ph.D. Thesis*, COPPE/UFRJ, Rio de Janeiro, Brazil.
- (108) Voss, S. (1996). A reverse elimination approach for the p -median problem. *Studies in Locational Analysis*, **8**, 49-58.
- (109) Warszawski, A. (1973). Multi-dimensional location problems. *Operational Research Quarterly*, **24**, 165-179.
- (110) Wesolowsky, G.O. & Truscott, W.G. (1975). The multiperiod location-allocation problem with relocation of facilities. *Management Science*, **22**, 57-65.
- (111) Whitaker, R. (1983). A fast algorithm for the greedy interchange for large-scale clustering and median location problems. *INFOR*, **21**, 95-108.
- (112) White, J. & Case, K. (1974). On covering problems and the central facility location problem. *Geographical Analysis*, **6**, 281-293.