Belgian Journal of Operations Research, Statistics and Computer Science, Vol 26, n° 2

MAXIMIZING USER CONVENIENCE AND POSTAL SERVICE EFFICIENCY IN POST BOX LOCATION

M. LABBE

CEME, Université Libre de Bruxelles, CP 139, B-1050 Bruxelles, Belgique

G. LAPORTE

GERAD, Ecole des Hautes Etudes Commerciales de Montréal, 5255 avenue Decelles, Montréal H3T 1V6, Canada

ABSTRACT

This paper deals with the optimal location of post boxes in an urban or in a rural environment. The problem consists of selecting sites for post boxes which will maximize an appropriate linear combination of user convenience and postal service efficiency. Several variants of the problem are considered and appropriate models and algorithms are developed.

1. INTRODUCTION

This paper deals with the optimal location of post boxes in an urban or in a rural environment. More specifically, we first consider a territory partitioned into n zones; from statistical surveys on mail volumes, it is possible to determine for every zone k the minimum number ϑ_k of required post boxes as well as candidate sites for these boxes. In a general situation, some of these sites will be chosen a priori as a post box location whereas the remaining sites will only represent potential locations. The cost associated with a certain selection of sites may be decomposed into two main components; (i) the routing cost associated with post box collection operations by van; (ii) the inconvenience cost incurred by users having to travel to and from their home or place of employment to the nearest post box. The problem consists of selecting sites for post boxes which will minimize an appropriate linear combination of these two costs, subject to some hypotheses and constraints. Equivalently, the problem may be posed in terms of maximization of user convenience and of postal service efficiency.

Insofar as we seek a selection of sites that will minimize routing costs, this problem may be viewed as a location-routing problem (see Laporte et al. [12,14,15] for some examples). It is also a combination of optimal locations of services and of optimal allocations of users to these services. Hence, it belongs to the family of location-allocation problems [4,9]. It is in fact a location-allocation-routing problem. This problem is part of the large class of network optimization problems defined by Golden et al. [5] and more specifically, of the family of problems consisting of optimally locating a path in a network (since the post box locations will determine the collection route). Two recent examples of such problems are provided by Current et al. [1] and by Vandale [18]. Savas [17] studied the most frequent criteria used in such contexts while Labbé [10, chap.4] concentrated on minisum-max problems. Finally, Halpern's papers [6,7,8] and Lowe's article [16] deal with multiobjective problems belonging to this family.

We examine, in the following sections, some important cases of the post box location problem.

2. PROBLEMS WITH PREDEFINED ZONES

Consider a territory subdivided into disjoint zones corresponding to conveniently chosen administrative units such as postal zones, census tracts, etc. In a city, these zones will likely be contiguous (see figure 1) whereas in the country, there will often exist an empty area between the zones (see figure 2). But from our point of view, these two situations can be treated in exactly the same way.

The problem is defined over a certain planning horizon normally consisting of a fraction of a day and in which all post boxes are visited once by a collection van. It is assumed that the collection route consists of a single Hamiltonian circuit passing through the post office which can be identified with one of the prespecified post boxes. Let L be the set of users: here a user can be thought of as an individual or as a group of people, for example all inhabitants of the same city block or of the same stretch of road. Over the planning horizon, user $\ell \in L$ makes a certain number of trips to the closest post box located within his zone. Let L_k be the set of users located within zone k. If a post box is located at site i, it is then convenient to define $d_{i\ell}$ as the total distance travelled by user ℓ to post box i over the planning horizon.

Also define:

S_k : the set of potential post box locations in zone k;

T_k : the set of prespecified post box locations in S_k ($T_k \subsetneq S_k$ and $|T_k| \leq \vartheta_k$);

$N = S_1 \cup \ldots \cup S_n$;

$M = T_1 \cup \ldots \cup T_n$;

c_{ij} : the length of a shortest path from site i to site j ($i,j \in N$);

λ : a real number in $[0,1]$.

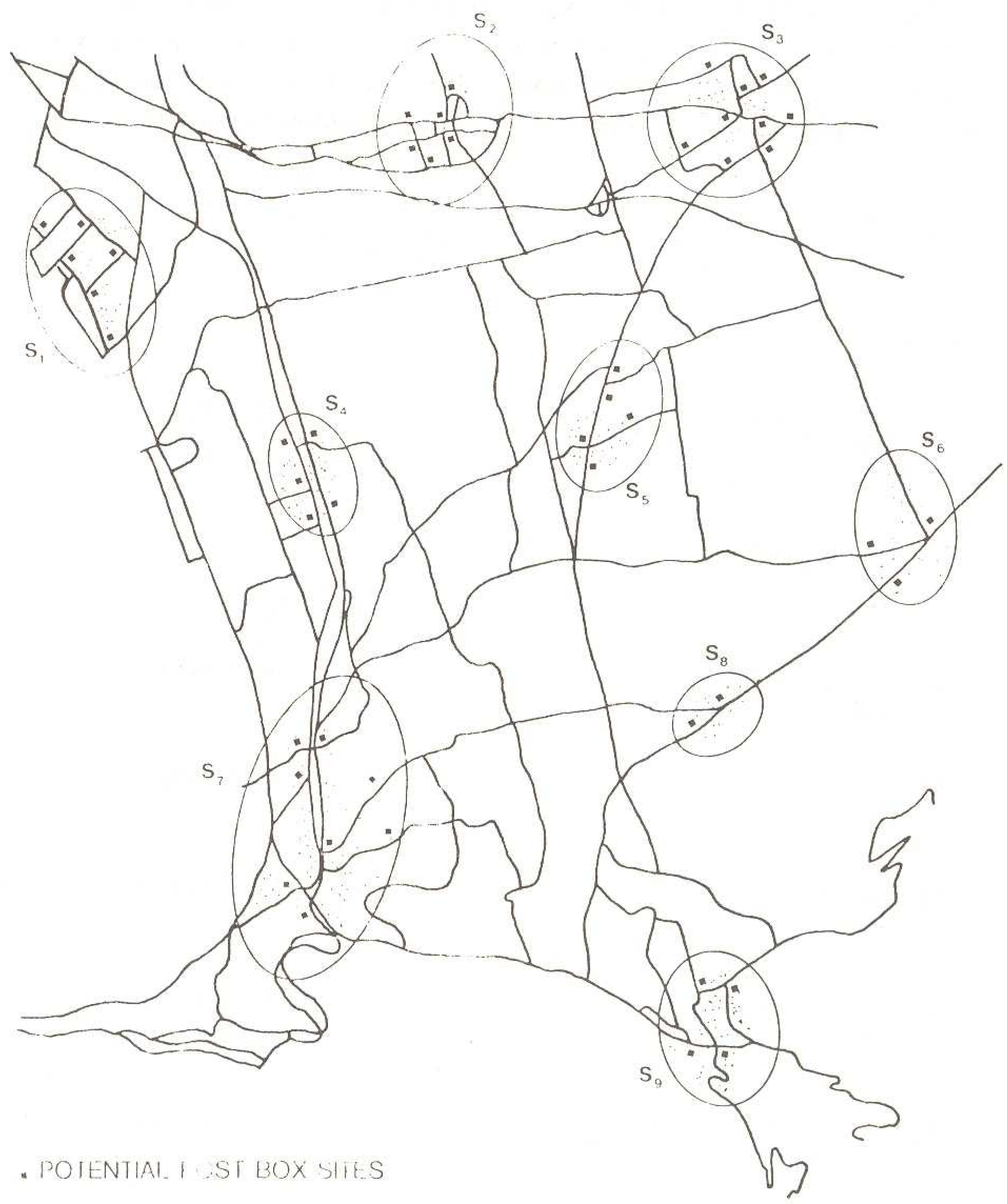
FIGURE 1

Subdivision of the St.Josse district (Brussels) into 10 zones

▪ POTENTIAL POST BOX SITES

FIGURE 2

Definition of 9 postal zones in a rural area
(near Pompei)



We formulate the problem as an integer linear program: let $x_{ij}(i,j \in N,\ i \neq j)$ be a binary variable indicating whether the van travels directly from i to j in the optimal solution ($x_{ij} = 1$) or not ($x_{ij} = 0$); let $x_{ii}(i \in N\text{-}M)$ be a binary variable indicating whether a post box is located at site i ($x_{ii} = 0$) or not ($x_{ii} = 1$) in the optimal solution. Finally, let $y_{i\ell} = 1$ if user ℓ is associated with post box i and $y_{i\ell} = 0$ otherwise; in this model, it is sufficient to define $y_{i\ell}$ only if i and ℓ belong to the same zone. The problem is then to

$$\text{(P1)} \quad \text{minimize} \quad \lambda \sum_{\substack{i,j \in N \\ i \neq j}} c_{ij}\, x_{ij} + (1-\lambda) \sum_{i \in N} \sum_{\ell \in L} d_{i\ell}\, y_{i\ell}$$

subject to

$$(1) \quad \sum_{\substack{i \in N \\ i \neq j}} x_{ij} = 1 \qquad (j \in M)$$

$$(2) \quad \sum_{\substack{j \in N \\ j \neq i}} x_{ij} = 1 \qquad (i \in M)$$

$$(3) \quad \sum_{i \in N} x_{ij} = 1 \qquad (j \in N\text{-}M)$$

$$(4) \quad \sum_{j \in N} x_{ij} = 1 \qquad (i \in N\text{-}M)$$

$$(5) \quad \sum_{i \in S_k} y_{i\ell} = 1 \qquad (\ell \in L_k;\ k=1, \ldots, n)$$

$$(6) \quad y_{i\ell} \leq 1 - x_{ii} \qquad (i \in S_k,\ \ell \in L_k;\ k=1, \ldots, n)$$

$$(7) \quad \sum_{i \in S_k - T_k} x_{ii} \leq |S_k| - \vartheta_k \qquad (k=1, \ldots, n)$$

$$(8) \quad \sum_{\substack{i,j \in S \\ i \neq j}} x_{ij} \leq |S| - 1 \qquad (S \subset N;\ S \text{ intersects with some but not all } S_k\text{'s})$$

(9) $x_{ij} = 0, 1$ $(i,j \in N)$

(10) $y_{i\ell} = 0, 1$ $(i \in S_k, \ell \in L_k;$
$k=1, \ldots, n)$

In the above formulation, constraints (1) to (4) specify the degree of each node $i \in N$: potential sites chosen a priori as post box locations will have to be entered and left exactly once (constraints (1) and (2)); constraints (3) and (4) state that other sites will be entered and left $1 - x_{ii}$ times, i.e. they will be visited only if a post box is located at site i ($x_{ii} = 0$). Constraints (5) ensure that every user is assigned to exactly one post box while constraints (6) mean that user ℓ can only be assigned to post box i if site i is used to locate a post box. Constraints (7) ensure that at least ϑ_k post boxes are located in zone k: since at least $\vartheta_k - |T_k|$ sites will be opened in $S_k - T_k$, then at most $(|S_k| - |T_k|) - (\vartheta_k - |T_k|) = |S_k| - \vartheta_k$ sites will be closed. Constraints (8) guarantee that the solution consists of a single Hamiltonian circuit: they prohibit the formation of subtours involving some but not all S_k's. Finally, constraints (9) and (10) specify that all variables are equal to 0 or 1.

The problem defined by (P1) is NP-complete. This is easily shown as follows. In the particular case where the second term of the objective function is dropped, (P1) reduces to a GTSP (generalized travelling salesman through n sets of nodes [11,13]) which consists of determining the shortest Hamiltonian circuit passing through every S_k at least ϑ_k times. But the GTSP is a generalization of the travelling salesman problem [2] which is itself NP-complete.

It is therefore unlikely that exact optimal solution to (P1) can be obtained for problems of realistic dimensions. We suggest the following procedure which exploits the fact that the objective function of (P1) represents a compromise between short travelling distances for the postal van and for users. Two extreme cases are in fact obtained by setting $\lambda = 0$ and $\lambda = 1$.

Step 1 : Select a first set of ϑ_k post box locations in every cluster k. This is done by solving in each case a v-median problem [3] in $S_k - T_k$ where $v = \vartheta_k - |T_k|$ and the distances used are defined by

$$(11) \quad d'_{i\ell} = \min_{i \in T_k \cup \{i\}} \{d_{i\ell}\} \qquad (i \in S_k - T_k,\ \ell \in L_k)$$

This ensures that the total distance travelled by users to their nearest post box is minimized. Let J_k be the set of selected sites in zone k (including those of T_k). For each user ℓ, compute e_ℓ, his distance to the closest site in J_k. For each site $i \in N$, define a weight $D_i = 0$. Set $J = \bigcup_{k=1}^{n} J_k$.

Step 2 : Solve a GTSP with the objective function

$$(12) \quad \lambda \sum_{\substack{i,j \in N \\ i \neq j}} c_{ij}\, x_{ij} + (1-\lambda) \sum_{i \in N-M} D_i (1 - x_{ii})$$

Efficient algorithms for the GTSP are described in [11,13]. Let I_k be the set of sites of zone k included in the GTSP solution. Let $I = \bigcup_{k=1}^{n} I_k$. If $I = J$ or if cycling occurs, stop. Otherwise, go to step 3.

Step 3 : Update the weights D_i as follows.

(i) Consider in turn all clusters k. Allocate all users $\ell \in L_k$ to the closest site $i \in I_k$. Ties are broken arbitrarily. Let $L_{k,i}$ be the set of users of L_k allocated to site i of I_k.

(ii) Define

$$(13) \quad D_i = \begin{cases} \sum_{\ell \in L_{k,i}} (d_{i\ell} - e_\ell) & (i \in I_k;\ k=1, \ldots, n) \\ 0 & (i \notin I_k;\ k=1, \ldots, n) \end{cases}$$

(iii) Set $J = I$.

Go to step 2.

Remarks

(i) If this algorithm is used, the objective function in step 2 can be simplified by distributing the cost $(1-\lambda)D_i$ associated with site i over its incoming and outgoing arcs. This is done by dropping the second term from the objective and by replacing the c_{ij}'s by

(14) $$c'_{ij} = c_{ij} + \tfrac{1}{2}(1-\lambda)\,(D_i + D_j)$$

Computational experience with the GTSP [13] indicates that D_i's having a large variance tend to produce easier problems since the optimal site selection then becomes more obvious and leads to smaller search trees if the problem is solved by branch and bound.

(ii) One interesting feature of the GTSP is that when $C = (c_{ij})$ satisfies the triangle inequality (i.e. $c_{ij} \le c_{ik} + c_{kj}$ for all $i,j,k \in N$), then only ϑ_k sites are used in cluster k in the optimal solution; the "$\le$" sign in constraints (7) can be replaced by an equality sign. And if $\vartheta_k = 1$, all variables x_{ij} for which i and $j \in S_k$ can then be dropped. Finally, observe that if C satisfies the triangle inequality, then so does $C' = (c'_{ij})$.

(iii) A relatively easy special case of (P1) occurs when C satisfies the triangle inequality, $T_k = \phi$ and $\vartheta_k = 1$ for all k. Then the objective function is defined by (12) with

(15) $$D_i = \sum_{\ell \in L_k} d_{i\ell} \qquad (i \in S_k)$$

Then (P1) can be solved directly by means of a GTSP algorithm. For this, it suffices to replace the c_{ij}'s by the c'_{ij}'s defined as in (14).

In the general case, setting a meaningful and easy to interpret value for λ presents a certain difficulty since the two parts of the objective function do not represent the same type of costs. One way to circumvent this difficulty is to first compute z_1^* and z_2^*, where

- z_1^* is the value of the optimal solution to the problem consisting of determining the best sites and the associated collection route; this problem consists of minimizing the objective function of (P1) with $\lambda=1$, under constraints (1) to (4) and (7) to (9);

- z_2^* is the value of the optimal solution of the problem consisting of assigning users to sites; this problem is obtained from (P1) by setting all x_{ij}'s and λ equal to zero and by retaining constraints (5), (6) and (10); the value of z_2^* can again be obtained by solving a $(\vartheta_k - |T_k|)$-median problem for $k=1, \ldots, n$.

The objective function of (P1) can then be replaced by

$$\lambda \left[\frac{\sum\limits_{i \neq j} c_{ij} \, x_{ij}}{z_1^*} \right] + (1-\lambda) \left[\frac{\sum\limits_{i \in N} \sum\limits_{\ell \in L} d_{i\ell} \, y_{i\ell}}{z_2^*} \right]$$

This transformation removes the unit of measurement effect: the expressions in brackets now measure the departure of each component of the objective function from its optimal value. This provides a useful interpretation of the final solution. The value of λ is easier to set since it now reflects more directly the weight attached to the satisfaction of the postal service objective.

3. PROBLEMS WITHOUT PREDEFINED ZONES

In contexts where predefined zones are not available, it may be preferable to treat the problem globally rather than arbitrarily defining zones which might unduly restrict the solution space and thus yield suboptimal solutions. In such cases, a variety of approaches are available.

The first strategy requires the specification of a number p of post boxes required on the whole territory. The problem then consists of (i) solving a p-median problem or a p-center problem (see Erlenkotter [3] or Krarup and Pruzan [9] for algorithms relative to these two problems) and of (ii) determining the optimal route through the p selected sites: this latter problem is a simple TSP. This approach is fundamentally user-oriented as the users' requirements are first met without any consideration for the routing costs which are only treated as a subsidiary objective.

The second proposed strategy consists of optimizing the routing costs under the constraint that the distance between a user and its nearest post box does not exceed a prespecified limit r. The algorithm corresponding to this strategy can be described as follows:

Step 1 : Determine, for every user ℓ, the set S_ℓ of post boxes whose distance from ℓ does not exceed r. S_ℓ may contain a set T_ℓ of already located post boxes. Let $\vartheta_\ell = \max \{1, |T_\ell|\}$. At the end of this step, we have obtained n distinct but not necessarily disjoint clusters. Since two different users may define the same cluster, it follows that $n \leq |L|$.

Step 2 : Solve (P1) with $\lambda = 1$. This problem is a GTSP. Since the clusters may intersect, the number of post boxes used in the optimal solution is not obvious, even if the distance matrix satisfies the triangle inequality. It would therefore seem appropriate to replace C by $F = (f_{ij})$, a matrix of travelling times, and to consider in the objective the time f necessary to unload a post box. The objective would then become

$$\text{(P2)} \quad \text{minimize} \sum_{\substack{i,j \in N \\ i \neq j}} f_{ij} x_{ij} + \sum_{i \in N-M} f(1 - x_{ii}) + f|M|$$

The constraints are those of (P1).

Step 3 :   Allocate users to their nearest post box.

Apart from constituting a fair compromise between minimizing user's inconvenience and routing costs, this approach has the advantage that its difficulty reduces to that of a GTSP. This problem is of course of considerable complexity but it can nevertheless be solved to optimality for up to about 100 sites [11].

4. EXTENSION

We have assumed, in the preceding sections, that only one van is used to empty the post boxes. In general, this is not so in practice since it is usually infeasible for a single van to cover the whole of the territory within a reasonable time. Usually, the operations are carried out by a fleet of vehicles based at a central depot (for example, a sorting office). These vans are then dispatched in order to minimize the sum of their fixed and running costs, while ensuring that their capacity is never exceeded and that the length of any route is at most equal to a given limit. In order to avoid a suboptimal solution, the allocation of vans to post boxes must be made simultaneously with the routing and the determination of the best post office sites.

These additional constraints considerably amplify the difficulty of the problem but, at the same time, bring the model closer to reality. While the computational difficulty of this extended model far exceeds that of the original one, preliminary results indicate that relatively large problems can still be solved to optimality. These results will be reported in a subsequent paper.

5. ACKNOWLEDGEMENTS

The authors are grateful to the Belgian Government (Action de Recherche Concertée, contract 84/8965) and to the Canadian Natural Sciences and Engineering Research Council (grant A4747) for their financial support. Thanks are also due to an anonymous referee for his valuable comments.

REFERENCES

[1] CURRENT, J., ReVELLE, C.S., COHON, J.L., "The Maximum Covering/ Shortest Path Problem: a Multiobjective Network Design and Routing Formulation", European Journal of Operational Research, vol.21, pp.189-199, 1985.

[2] DANTZIG, G.B., FULKERSON, D.R., JOHNSON, S.M., "Solution of a Large Scale Travelling Salesman Problem", Operations Research, vol.2, pp.393-410, 1954.

[3] ERLENKOTTER, D., "A Dual-Based Procedure for Uncapacitated Facility Location", Operations Research, vol.26, pp.992-1009, 1978.

[4] FRANCIS, R.L., McGINNIS, L.F., WHITE, J.A., "Locational Analysis", European Journal of Operational Research, vol.12, pp.220-252, 1983.

[5] GOLDEN, B.L., BALL, M., BODIN, L.D., "Current and Future Research Directions in Network Optimization", Computers & Operations Research, vol.8, pp.71-81, 1981.

[6] HALPERN, J., "The Location of a Center-Median Convex Combination on an Undirected Tree", Journal of Regional Science, vol.16, pp.237-245, 1976.

[7] HALPERN, J., "Finding Minimal Center-Median Convex Combinations (Cent-Dian) of a Graph", Management Science, vol.24, pp.535-544, 1978.

[8] HALPERN, J., "A Simple Edge Elimination Criterion in a Search for the Center of a Graph", Management Science, vol.25, pp.105-107, 1979.

[9] KRARUP, J., PRUZAN, P.M., "The Simple Plant Location Problem: Survey and Synthesis", European Journal of Operational Research, vol.12, pp.36-81, 1983.

[10] LABBE, M., "Essays in Network Location Theory", Ph.D. Thesis, Université Libre de Bruxelles, published in Les Cahiers du CERO, vol.27, pp.50-130, 1985.

[11] LAPORTE, G., MERCURE, H., NOBERT, Y., "Generalized Travelling Salesman Problem through n Sets of Nodes: the Asymmetrical Case", Cahiers du GERAD G-85-07, Ecole des Hautes Etudes Commerciales de Montréal, 1985.

[12] LAPORTE, G., NOBERT, Y., "An Exact Algorithm for Minimizing Routing and Operating Costs in Depot Location", European Journal of Operational Research, vol.6, pp.224-226, 1981.

[13] LAPORTE, G., NOBERT, Y., "Generalized Travelling Salesman through n Sets of Nodes: an Integer Programming Approach", INFOR, vol.21, pp.61-75, 1983.

[14] LAPORTE, G., NOBERT, Y., ARPIN, D., "An Exact Algorithm for Solving a Capacitated Location-Routing Problem", to appear in Location Theory: Methodology and Applications, J.C. Baltzer AG, Scientific Publishing Co., Basel, Switzerland, 1986.

[15] LAPORTE, G., NOBERT, Y., PELLETIER, P., "Hamiltonian Location Problems", European Journal of Operational Research, vol.12, pp.82-89, 1983.

[16] LOWE, T.J., "Efficient Solutions in Multiobjective Tree Network Location Problems", Transportation Science, vol.12, pp.298-316, 1979.

[17] SAVAS, E.S., "On Equity in Providing Public Services", Management Science, vol.24, pp.800-808, 1978.

[18] VANDALE, F., "Model for Creating Routes and Schedules for a Public Transportation Network", Paper presented at the Third National Congress on Quantitative Methods for Decision Making, Brussels, 1985.