Journal of the Operational Research Society (1996) 47, 1461–1467  0160-5682/96 $12.00

# Some Applications of the Generalized Travelling Salesman Problem

GILBERT LAPORTE[1,2], ARDAVAN ASEF-VAZIRI[3], and CHELLIAH SRISKANDARAJAH[3]
[1]GERAD and École des Hautes Études Commerciales, Montréal, Canada, [2]Centre de recherche sur les transports, Université de Montréal, Montréal, Canada, and [3]University of Toronto, Canada.

In the Generalized Travelling Salesman Problem (GTSP), the aim is to determine a least cost Hamiltonian circuit or cycle through several clusters of vertices. It is shown that a wide variety of combinatorial optimization problems can be modelled as GTSPs. These problems include location-routeing problems, material flow system design, post-box collection, stochastic vehicle routeing and arc routeing.

*Key words:* generalized travelling salesman problem

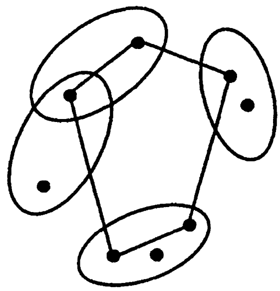
## INTRODUCTION

The purpose of this article is to show how a wide variety of combinatorial optimization problems can be modelled as a Generalized Travelling Salesman Problem (GTSP), a well-known extension of the Travelling Salesman Problem (TSP). The GTSP can be formally defined as follows. Let $G=(V,A)$ be a graph where $V = \{v_1, \ldots, v_n\}$is the vertex set, and $A = \{(v_i, v_j): i \neq j, v_i, v_j \in V\}$ is the arc set. A *cost* or *distance* matrix $C = (c_{ij})$ is defined on $A$. If $C$ is symmetric, arc directions are irrelevant and it is then customary to replace each arc $(v_i, v_j)$ by an edge $(v_i, v_j)$ where $i<j$. In the GTSP, $V$ is the union of $p$ clusters, i.e., $V = \{V_1, \ldots, V_p\}$. One version of the problem consists of determining the shortest Hamiltonian *circuit* passing through each cluster at least once. (In symmetric problems, a shortest Hamiltonian *cycle* is sought.) In another version of the GTSP, the condition 'at least once' is replaced by 'exactly once'. In both cases, this means that some vertices of a cluster may be left unvisited (see Figure 1).

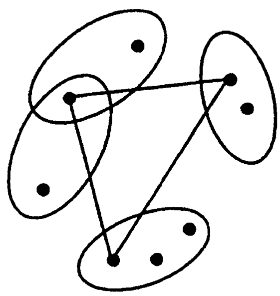
The GTSP was introduced by Henry-Labordere[1], Saksena[2] and Srivastava *et al.*[3] in relation with record balancing problems arising in computer design and with the routeing of clients through agencies providing various services. For example, a person may wish to obtain $p$ different services provided by government agencies. In this context, $V_k$ is the set of agencies dispensing the $k$th type of service. If each institution offers only one type of service, the clusters $V_k$ are disjoint; otherwise, they intersect. These authors provided a simple dynamic programming formulation for the problem, all having limited computational efficiency. Laporte and Nobert[4] proposed a constraint relaxation algorithm for the symmetric version of the problem. A Lagrangean relaxation algorithm valid for the asymmetric case is presented by Noon and Bean[5], while Fischetti, Salazar and Toth[6] developed a branch and cut for symmetric instances. An interesting construction is that of Noon and Bean[7]. These authors show that the GTSP with disjoint clusters defined on a directed graph can readily be transformed into an equivalent TSP involving the same number of vertices. This finding does not imply, however, that asymmetric GTSPs can be solved efficiently as symmetric TSPs since the transformation induces a fair amount of degeneracy.

As will be seen, the GTSP provides an attractive way of modelling a wide range of situations. These are more involved than the agencies scenario just described. They arise in areas such as location-routeing, loop material flow system design, post-box collection, stochastic vehicle routeing and arc routeing. Our purpose is to illustrate the versatility of the GTSP as a modelling tool, without reference to algorithmic aspects which have already received broad attention. Each of the following five sections of this paper will be devoted to one of the situations just mentioned. Some of these applications have been described in scientific journal articles or in working papers, while two are entirely new.

*Correspondence: G. Laporte, Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Canada H3C 3J7*



a) Each cluster is visited at least once



b) Each cluster is visited exactly once

FIG. 1. *Two versions of the GTSP.*

## THE COVERING TOUR PROBLEM

Location-routeing problems arise in a variety of practical contexts where it is required to simultaneously locate facilities and design vehicle routes through these facilities and a number of other sites (see Laporte[8]). For example, Watson-Gandy and Dohrn[9] solve the problem of locating depots and of constructing vehicle delivery routes through grocery stores. GTSP applications occur when only one vehicle is involved, as in the Covering Tour Problem (CTP) recently analyzed by Gendreau, Laporte and Semet[10]. The CTP is defined on a graph $H = (V \cup W, E)$, where $E = \{(v_i, v_j): i < j, v_i, v_j \in V \cup W\}$, and the distance matrix $C$ is defined on $E$. The aim is to construct a least length tour through a subset of $V$ in such a way that each vertex of $W$ is within a distance $c$ of a vertex of the tour. Letting $V_k = \{v_i \in V: c_{ik} \leqslant c\}$, for each $v_k \in W$, the CTP can then be stated as a GTSP on the subgraph of $G$ of $H$ induced by $\cup_{v_k \in W} V_k$. Examples of the CTP include post-box location problems[11] in which $W$ is a set of population centroids ($v_k \in W$ would represent, for example, all people living in city block $k$), $V_k$ is a set of potential post-box locations within walking distance $c$ from $v_k$, and the objective is to design a minimum length tour for a postal vehicle visiting the selected locations on a regular basis. The CTP is sometimes referred to as the Travelling Circus Problem as it is the problem faced by circus managers wishing to tour at minimum cost a set of villages, while ensuring that any village not on the tour is within $c$ distance limits from it.

## MATERIAL FLOW SYSTEM DESIGN

An important problem in a number of production settings is the design of loop material flow systems. Consider the block layout of a production plant partitioned into $p$ polygonal zones $Z_1, \ldots, Z_p$ as illustrated in Figure 2. These zones need not be convex, but they only contain 90° and 270° angles. The problem is to design a minimal length circular path (for example for an automated monorail or an automated guided vehicle) along the edges of the zones. There are two versions to this problem. In the first version, the path must contain at least one vertex from each zone, but it need not include an edge from any zone. In the second version, the path must include at least one edge from each zone.

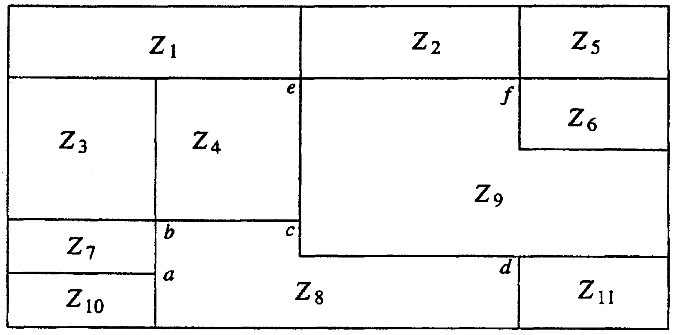



FIG. 2. *Production plant partitioned into eleven zones.*

The first case is easily formulated as a GTSP. First define the vertex set of $V_k$ of a zone $Z_k$ as the set of vertices of zone $Z_k$ with a degree at least equal to 3. Vertices of degree 2 play no structural role in this problem and can be disregarded. Indeed, if a vertex of degree 2 belongs to the loop, its two neighbours also belong to it. Then a minimal length circular loop is simply determined by solving a GTSP over the clusters $V_1, \ldots, V_p$.

The second case is slightly more complicated. First observe that all zone vertices have a degree equal to 2, 3, or 4. Again, vertices of degree 2 need not be considered. If the loop passes through a vertex of degree 3, then it must necessarily contain an edge in each of the three zones to which belongs. Thus, if all vertices have degree 3, solving the problem as in the first case ensures that the loop contains an edge from each zone. This is not the case for vertices of degree 4. For example, if the loop follows the path $(\ldots, a, b, c, d, \ldots)$ in Figure 2, it contains a vertex from $Z_3$ (vertex $b$), but not necessarily an edge from that zone. A transformation is thus required for such vertices. Consider in turn all edges $(v_i, v_j)$ where $v_i$ and $v_j$ are both of degree 4. Then introduce a vertex on $(v_i, v_j)$ half-way between $v_i$ and $v_j$, and include it in $V_r$ and $V_s$, where $Z_r$ and $Z_s$ are the two zones sharing edge $(v_i, v_j)$. Then remove all vertices of degree 4 from the clusters $V_k$. In Figure 2, a new vertex would have to be introduced between $e$ and $f$, but not between $b$ and $c$ since $c$ is of degree 3. Then again the loop problem can be solved as a GTSP, where it is required to go through each of the clusters $V_1, \ldots, V_p$ at least once.

## POST-BOX COLLECTION

The GTSP can be used to model a number of vehicle routeing problems in which the distance $c_{ij}$ from $v_i$ to $v_j$ depends on the vertex preceding $v_i$. Such a problem is encountered in the planning of post-box collection routes. It was first mentioned to the first author by Rousseau[12] in relation to a study undertaken for the Canada Post Corporation (CPC) in the late 1980's. A full description is provided in Laporte *et al.*[13] It will now be summarized. Postal services throughout the world must empty on a regular basis post-boxes located along rural roads or city streets. This operation entails large costs and substantial savings can generally be achieved through better planning of collection vehicle routes.

The problem posed by CPC arises in urban contexts and stems from a work of agreement clause which states that drivers may not cross a street on foot in order to empty a post-box. To show how this simple rule can affect the planning of collection routes, consider the situation depicted in Figure 3. Assume two-way traffic on all streets, right-hand side driving and U-turn prohibitions. Consider two post-boxes located at $v_i$ and $v_j$. The problem is to determine the value of the shortest distance $c_{ij}$ from $v_i$ to $v_j$. If the vehicle approaches $v_i$ from the South, then $c_{ij}$ is simply the length of the dashed line from $v_i$ to $v_j$. If, on the other hand, $v_i$ is visited coming from the West, then the shortest way to reach $v_j$ is to follow the dotted line round the block, resulting in a longer distance. As a result, $c_{ij}$ is not well defined. A way round this problem is to duplicate all post-boxes located at street corners and locate each copy of the post-box on a different edge, slightly off the corner. This is illustrated in Figure 4: $v_i$ is replaced with $v_i'$ and $v_i''$, and $v_j$ is replaced with $v_j'$ and $v_j''$. Then the four distances $c_{i'j'}$, $c_{i'j''}$, $c_{i''j'}$, $c_{i''j''}$ are clearly defined and the problem of determining a shortest Hamiltonian circuit through all post-boxes can be stated as a

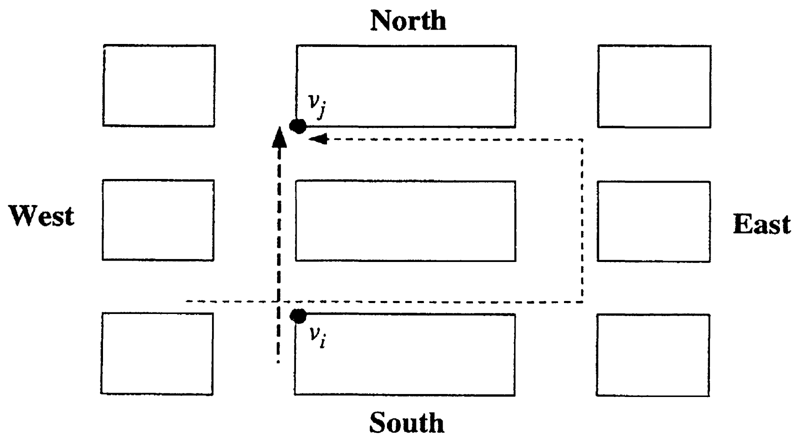



FIG. 3. *City blocks with two post- boxes $v_i$ and $v_j$ located at street corners.*

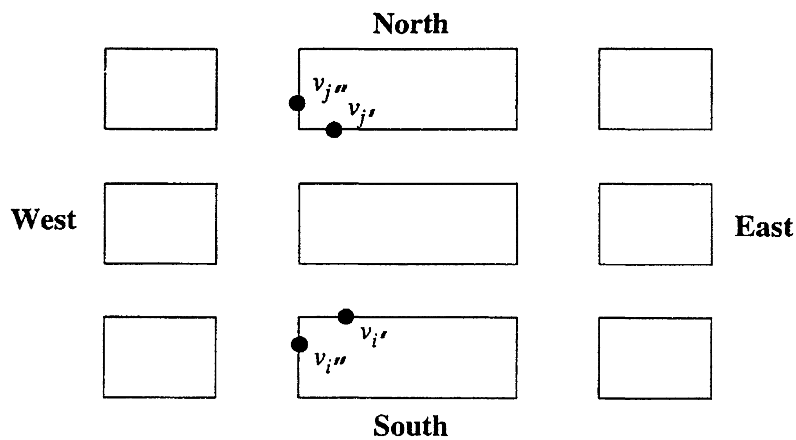



FIG. 4. *City blocks with duplicated post-boxes.*

GTSP in which $V_i = \{v_i\}$ if $v_i$ has not been duplicated (if it could only be approached from one direction), and $V_i = \{v_i', v_i''\}$ if $v_i$ has been duplicated.

## STOCHASTIC VEHICLE ROUTEING

The following application of the GTSP arises from a stochastic vehicle context. Let $v_1$ represent a depot, and $v_j (j = 2, \dots, n)$ be customers with supplies $\xi_j$ at which collections must be made by a single vehicle of capacity $Q$ based at the depot. If all supplies $\xi_j$ are known constants and $\sum_{j=2}^{n} \xi_j \leqslant Q$, then the problem of designing an optimal collection route through the depot and all customers is a TSP. There are contexts, however, when the $\xi_j$s are random variables. This gives rise to a stochastic routeing problem (see Gendreau, Laporte and Séguin[14] for a recent survey). Here, a *planned route* through $v_1, \dots, v_n$ is first established. A vehicle follows it until its capacity becomes attained or exceeded, at which point *route failure* is said to occur. A *recourse action* must then be taken, for example, return to the depot to unload and resume collections as planned. The problem is to determine a delivery sequence such that the cost of the planned route plus the expected recourse cost is minimized. Such a problem was studied by Dror, Laporte and Louveaux[15]. As argued by these authors, it would make little managerial sense to plan routes that systematically fail and it is meaningful to study the special case where failure can only occur at the last visited customer, namely

$$P\left(\sum_{v_i \in V/\{v_1, v_j\}} \xi_i \leqslant Q\right) = 1 \qquad (j = 1, 2, \ldots, n).$$

Letting $\alpha = P(\sum_{v_i \in V \setminus \{v_1\}} \xi_i > Q)$, the problem can then be transformed as a TSP in which all costs $c_{j1}$ are replaced by $c_{j1} + \alpha(c_{j1} + c_{1j})(j = 2, \ldots, n)$.

A more sophisticated policy would be to allow for a preventive break at some customer $v_l$ along the planned route. There, upon reaching $v_l$, the vehicle would return to the depot to unload before resuming collections, even if its capacity is not reached. This would make sense if, for example, $v_l$ is close to the depot, $\alpha$ is large, and the last customer on the planned route is far from the depot. Such a course of action may be viewed as a form of insurance policy against failure. It can be modelled as a TSP by introduction of an artificial depot $v_{n+1}$, with all original $c_{ij}$s remaining the same, and

$$\begin{aligned}
c_{1,n+1} &:= c_{n+1,1} := \infty \\
c_{i,n+1} &:= c_{i1} \qquad (i = 2, \ldots, n) \\
c_{n+1,j} &:= c_{1j} \qquad (j = 2, \ldots, n).
\end{aligned}$$

The decision whether to risk failure at the last customer or include a preventive break on the route can therefore be made by solving two separate TSPs. However, it is possible to formulate the problem as a GTSP with clusters $V_k = \{v_k\}(k = 1, \ldots, n)$ and $V_{n+1} = \{v_{n+1}, v_{n+2}\}$, where $v_{n+1}$ and $v_{n+2}$ are two vertices with the following meanings: if $v_{n+1}$ is entered from a vertex of $V \setminus \{v_1\}$, the solution includes a preventive break; if $v_{n+2}$ is entered from a vertex of $V \setminus \{v_1\}$, the solution includes no preventive break, but failure may occur at the last customer. In the GTSP, only one of $v_{n+1}$ or $v_{n+2}$ must be visited since risking failure and planning a preventive break are two mutually exclusive scenarios. The costs related to $v_{n+1}$ and $v_{n+2}$ are as follows:

$$\begin{aligned}
c_{i,n+1} &:= c_{i1} \qquad (v_i \in V \setminus \{v_1\}) \\
c_{n+1,j} &:= c_{ij} \qquad (v_j \in V \setminus \{v_1\}) \\
c_{1,n+1} &:= c_{n+1,1} := \infty \\
c_{i,n+2} &:= c_{ij} + \alpha(c_{i1} + c_{1i}) \qquad (v_i \in V \setminus \{v_1\}) \\
c_{1,n+2} &:= \infty \\
c_{n+2,1} &:= 0 \\
c_{n+1,n+2} &:= c_{n+2,n+1} := \infty.
\end{aligned}$$

## ARC ROUTEING

Finally, we show how the GTSP constitutes a convenient way of transforming arc routeing problems into vertex routeing problems. In arc routeing problems, the aim is to determine a least cost traversal of a given subsets of the arcs or edges of a graph $G = (V, A \cup E)$, where $A$ is an arc set and $E$ is an edge set. Let $c_{ij}$ be the cost of arc $(v_i, v_j)$ and let $d_{ij}$ be the cost of traversing edge $(v_i, v_j)$ from $v_i$ to $v_j$. Let $A'$ and $E'$ represent the subsets of required arcs and edges, respectively, i.e., those arcs or edges that must belong to the solution. Using this notation, we can define several classes of well-known problems.

### *Chinese Postman Problems (CPPs)*

Here, all arcs and edges of the graph must belong to the solution (namely $A' = A$ and $E' = E$).

- Directed CPP: $E = \varnothing$;
- Undirected CPP: $A = \varnothing$ and $d_{ij} = d_{ji}$ for all edges $(v_i, v_j)$;

- Mixed CPP: $A \neq \varnothing$, $E \neq \varnothing$ and $d_{ij} = d_{ji}$ for all edges $(v_i, v_j)$;
- Windy CPP: $A = \varnothing$ and $d_{ij} \neq d_{ji}$ for some edge $(v_i, v_j)$.

Both the undirected and directed CPP can be solved in polynomial time, but the mixed CPP and the windy CPP are both NP-hard (see Eiselt, Gendreau and Laporte[16] for a recent survey).

*Rural Postman Problems (RPPs)*

The undirected, directed, mixed and windy RPPs are defined in a similar fashion, except that now $A' \subset A$ or $E' \subset E$. The *Stacker Crane Problem* is a particular mixed RPP in which $A' = A$ and $d_{ij} = d_{ji}$ for all edges $(v_i, v_j)$. All these problems are NP-hard[16].

It is quite straightforward to transform completely directed CPPs and RPPs into equivalent vertex routeing problems by working on an auxiliary graph $H = (W, B)$, where $W$ contains a vertex $w_{ij}$ for each-arc $(v_i, v_j) \in A'$, $B = \{(w_{ij}, w_{kl}): w_{ij}, w_{kl} \in W, w_{ij} \neq w_{kl}\}$, and the cost of arc $(w_{ij}, w_{kl})$ is equal to the cost of a shortest path from $v_j$ to $v_k$ in $G$. Then solving a directed TSP on $G$ solves the arc routeing problem. This was the approach used by Beltrami and Bodin[17], for example. Note that this transformation is of no use if the original arc routeing problem is a directed CPP since it would replace a polynomially solvable problem with an NP-hard problem.

If $G$ is not completely directed, the original arc routeing problem can be transformed into a GTSP as follows. For this, first replace in $G$ each edge $(v_i, v_j)(i < j)$ by two arcs $(v_i, v_j)$ and $(v_j, v_i)$, with respective costs $d_{ij}$ and $d_{ji}$. Then construct $H$ as above and define clusters $V_{ij}$ as follows: if $(v_i, v_j)$ was an arc in $G$, then $V_{ij} = \{w_{ij}\}$; if $(v_i, v_j)$ was an edge in $G$, then $V_{ij} = \{w_{ij}, w_{ji}\}$. We have thus transformed the original arc routeing problem into a GTSP in which it is required to visit exactly one element per cluster. Whether this transformation can yield efficient algorithms for arc routeing problems remains to be investigated, but we suspect it should work well for graphs that are highly asymmetric, either because $|E'|$ is small, or because $d_{ij} \neq d_{ji}$ for several edges. Again, this transformation is only useful if the original problem is not an undirected CPP.

## CONCLUSIONS

We have shown, through five examples, that the GTSP constitutes a versatile and elegant tool to model several classes of combinatorial optimization problems. Routeing problems involving locational choices usually lend themselves to a GTSP formulation. These include classical location-routeing problems and the material flow design problem described in the paper. In other cases, such as the post-box collection problem, stochastic routeing or arc routeing the purpose of the GTSP formulation is to represent various ways of computing routeing costs. The emphasis of this paper is on modelling, not on algorithms. No claim of computational efficiency is made, but given the recent algorithmic advances made in the area of the GTSP[5,6] there is hope that the proposed models could constitute the basis for efficient algorithms.

*Acknowledgements*—This work was partially supported by the Canadian Natural Sciences and Engineering Research Council under grants OPG0039682 and OPG0104900. This support is gratefully acknowledged.

## REFERENCES

1. A. L. Henry-Labordere (1969) The record balancing problem: a dynamic programming solution of a generalized travelling salesman problem. *RIRO B* **2**, 43–49.
2. J. P. Saskena (1970) Mathematical model for scheduling clients through welfare agencies. *CORS J.* **8**, 185–200.
3. S. S. Srivastava, S. Kumar, R. C. Garg and P. Sen (1969) Generalized travelling saleman problem through $n$ sets of nodes. *CORS J.* 7, 97–101.
4. G. Laporte and Y. Nobert (1983) Generalized travelling salesman problem through $n$ sets of nodes: an integer programming approach. *INFOR* **21**, 61–75.
5. C. E. Noon and J. C. Bean (1991) A Lagrangian based approach for the asymmetric generalized travelling salesman problem. *Opns Res.* **39**, 623–632.
6. M. Fischetti, J. J. Salazar and P. Toth (1994) A branch-and-cut algorithm for the symmetric generalized travelling salesman problem. Working paper, DEIS, University of Bologna.

7. C. E. NOON and J. C. BEAN (1993) An efficient transformation of the generalized travelling salesman problem. *INFOR* **31**, 39–44.
8. G. LAPORTE (1988) Location-routing problems. In *Vehicle Routeing: Methods and Studies* (B. L. GOLDEN and A. A. ASSAD, Eds.) pp. 163–197. North-Holland, Amsterdam.
9. C. D. T. WATSON-GANDY and P. J. DOHRN (1993) Depot location with van salesman—a practical approach. *Omega* **1**, 321–329.
10. M. GENDREAU, G. LAPORTE and F. SEMET (1996) The covering tour problem. *Opns Res.* Forthcoming
11. M. LABBÉ and G. LAPORTE (1986) Maximizing user convenience and postal service efficiency in post-box location. *Belgian J. Opns Res., Stat. and Comp. Sci.* **26**, 21–35.
12. J.-M. ROUSSEAU (1985) Private communication.
13. G. LAPORTE, S. CHAPLEAU, P.- É. LANDRY and H. MERCURE (1989) An algorithm for the design of mailbox collection routes in urban areas. *Trans. Res. B* **23B**, 271–280.
14. M. GENDREAU, G. LAPORTE and R. SÉGUIN (1996) Stochastic vehicle routing. *Eur. J. Opl Res.* **88**, 3–12.
15. M. DROR, G. LAPORTE and F. V. LOUVEAUX (1993) Vehicle routing with stochastic demands and restricted failures. *Zeitschrift für Opns Res.* **37**, 273–283.
16. H. A. EISELT, M. GENDREAU and G. LAPORTE (1995) Arc routing problems, part I: The Chinese postman problem. *Opns Res.* **43**, 231–242.
17. E. L. BELTRAMI and L. D. BODIN (1974) Networks and vehicle routing for municipal waste collection. *Networks* **4**, 65–94.


*Received November 1995; accepted May 1996 after one revision*