Survey Paper

Time Window Constrained Routing and Scheduling Problems

MARIUS M. SOLOMON
Northeastern University, Boston, Massachusetts 02115

JACQUES DESROSIEUX
University of Montreal, HEC, Montreal, Quebec, Canada

We have witnessed recently the development of a fast growing body of research focused on vehicle routing and scheduling problem structures with time window constraints. It is the aim of this paper to survey the significant advances made for the following classes of routing problems with time windows: the single and multiple traveling salesman problem, the shortest path problem, the minimum spanning tree problem, the generic vehicle routing problem, the pickup and delivery problem including the dial-a-ride problem, the multiperiod vehicle routing problem and the shoreline problem. Having surveyed the state-of-the-art in this area, we then offer some perspectives on future research.

The effective and efficient management of the distribution of goods or services is becoming increasingly important in both the private and public sectors. A very important segment of many distribution and transportation systems costs is associated with the routing and scheduling of vehicles.

Due to the intrinsic complexity of distribution problems involving routing components, the use of mathematical-programming-based models and algorithms is needed when analyzing and solving such problems to permit the realization of cost reductions or profit improvement.

The vehicle routing problem (VRP) involves the design of a set of minimum cost routes, originating and terminating at a central depot, for a fleet of vehicles which services a set of customers with known demands. Each customer is serviced exactly once and, furthermore, all the customers must be assigned to vehicles such that the vehicle capacities are not exceeded. The VRP area has been intensively studied in the literature. We refer to BODIN et al. [6] for a comprehensive survey of the VRP and its variations which also describes the many practical occurrences of these problems. We also refer to the surveys of MAGNANTI [48] and more recently LAPORTE and NOBERT [44] on optimization methods for VRP and related problems.

In the VRP with time windows (VRPTW), the above issues have to be dealt with under the added complexity of allowable delivery times, or time windows, stemming from the fact that some customers impose service deadlines and earliest service time constraints. In these problems, the spacial aspect of routing is blended with the temporal aspect of scheduling, which must be performed to ensure the satisfaction of the time window constraints. The service of a customer, involving pickup and/or delivery of goods or services, can begin within the time window defined by the earliest time and the latest time when the customer will permit the start of service. Note that the times at which services begin are decision variables.

It is worth distinguishing now between hard and soft time windows. In the hard time window case, if a vehicle arrives at a customer too early, it will wait. In addition, due dates cannot be violated. In contrast, in the soft time window case, the time window constraints can be violated at a cost, in a manner similar to dualizing constraints in a Lagrangian relaxation context.
Time windows arise naturally in problems faced by business organizations which work on fixed time schedules. Specific examples of problems with hard time windows include bank deliveries, postal deliveries, industrial refuse collection and school-bus routing and scheduling. Among the soft time window problem instances dial-a-ride problems play a central role.

Concerning the vehicle fleet size and mix, most of the literature has assumed a homogeneous fleet. However, many of the methods proposed can be extended to consider a heterogeneous fleet size. Furthermore, most authors have assumed that the number of vehicles used is free, i.e., the fleet size is determined simultaneously with the best set of routes and schedules rather than being fixed a priori.

Additional complexities encountered in the VRPTW are length of route constraints arising from depot time windows. Note that the vehicle departure times from the depot are decision variables. Furthermore, precedence relationships among certain customers are encountered in some problems, such as those involving both pickups and deliveries.

Most of the effort has been directed at the operational problem of determining the best set of routes and schedules. In the presence of time windows, the total routing and scheduling costs include not only the total travel distance and time costs considered for routing problems, but also the cost of waiting time which is incurred when a vehicle arrives too early at a customer location or when the vehicle is loaded or unloaded. Several authors have also analyzed the strategic question of the optimal fleet size. Other problem variants, involving different objective functions, have been scantily addressed.

The VRPTW has emerged as an important area for progress in handling realistic complications and generalizations of the basic routing model (SCHRAGÉ [65], Bodin et al. [86]). In the last few years we have witnessed the development of a fast growing body of research focused on vehicle routing and scheduling problem structures with time windows. It is the aim of this paper to survey the significant advances made and offer some perspectives on future research. We concentrate on the development of algorithms and their computational capabilities. For the sake of brevity, given the large number of model variations, these models will not be examined in detail. However, to focus the problem and illustrate its difficulty we will present one general model. This model will then be used to provide a taxonomy for this class of problems.

The paper is organized as follows. In Section 1, we present a model for the pickup and delivery problem with time windows and show the many problem variants which can be derived. We then review the research relating to the study of special problem structures with time windows such as, the traveling salesman problem (Section 2), the multiple traveling salesman problem (Section 3), the shortest path problem (Section 4), and the minimum spanning tree problem (Section 5). We then examine the body of research developed for the generic vehicle routing problem with time window constraints (Section 6). Pickup and delivery problems including the dial-a-ride problem (Section 7), the multiperiod vehicle routing problem (Section 8), and the shoreline problem (Section 9) are then described. Finally, Section 10 offers some concluding remarks and suggests future research directions.

1. SOME MATHEMATICAL FORMULATIONS

In this section we present a formulation for the pickup and delivery problem with time windows (PDPTW) to focus the problem and illustrate the difficulty of problems with time windows. We then use this model to develop a taxonomy for routing and scheduling problems with time windows. For simplicity, we develop the formulation for the PDPTW involving only a single depot and a homogeneous fleet. However, formulations for more general variants involving multiple depots and different types of vehicles have been considered (see GAVISH and SRIKANTH [94] and Dumas [96]).

Let there be $n$ customers indexed by $i$. Associate to the pickup location of customer $i$ a node $i$ and to his delivery location a node $n + i$. Also associate to the depot nodes 0 and $2n + 1$. This creates a clear distinction between customers, their associated locations and the nodes of the network. Note, however, that different nodes may correspond to the same physical location. Therefore, $N = \{0, 1, 2, \ldots, n, n + 1, n + 2, \ldots, 2n, 2n + 1\}$ is the node set for our network, and $P = P^+ \cup P^-$, $P^+ = \{1, 2, \ldots, n\}$, $P^- = \{n + 1, n + 2, \ldots, 2n\}$ is the set of nodes other than the depot nodes. We will index both $N$ and $P$ by $u$ and $w$. Customer $i$ demands that $d_i$ units be shipped from node $i$ to node $n + i$. Next, let $[a_i, b_i]$ denote the pickup time window for customer $i$ and let $[a_{n+i}, b_{n+i}]$ denote his delivery time window. Let $V = \{1, 2, \ldots, |V|\}$ be the set of vehicles to be routed and scheduled; index this set by $v$. Let $D$ be the capacity of each vehicle. Let also $[a_0, b_0]$ denote the vehicles' departure time window from the depot and $[a_{2n+1}, b_{2n+1}]$ their time window for arrival back at the depot.

For each distinct $u, w$ in $N$, let $t_{uw}$ and $c_{uw}$ represent the travel time and the travel cost from $u$ to $w$, respectively. Retain only those arcs $(u, w)$ for which (1) $a_u + s_u + t_{uw} \leq b_w$, where $s_u$ is the service time at customer $u$, and either (2) $d_u + a_u \leq D$, $u, w \in P^+$, $u \neq w$, or (2') $d_{n-u} + d_{n-w} \leq D$, $u, w \in P^-$, $u \neq w$. 
Finally, let $K$ be the fixed cost associated with each vehicle, incurred if this vehicle is utilized.

Three types of variables are used in the mathematical formulation: binary flow variables $X_{uw}, u, w \in N, v \in V, u \neq w$, time variables $T_u, T_0, T_{2n+1}, u \in P, v \in V$, and load variables $Y_u, u \in P$. Let the binary decision variable $X_{uw}$ equal to one if vehicle $v$ travels from node $u$ to node $w$, and equal to zero otherwise, $v \in V, u, w \in N, u \neq w$. Let $T_u$ be the time at which service at customer $u$ begins, $u \in P$. Let also $T_0$ be the time at which vehicle $v$ leaves the depot and $T_{2n+1}$ be the time at which it returns to the depot, $v \in V$. Next, let $Y_u$ be the total load on the vehicle just after it leaves node $u$, $u \in P$. It is assumed that the vehicles depart empty from the depot, i.e., $Y_0 = 0$. These are decision variables with nonnegativity constraints.

Finally, let $f_u(T_u)$ denote the penalty associated with beginning service at node $u$ at time $T_u, u \in P$, and $g_v(T_{2n+1} - T_0)$ be the penalty associated with a route of the indicated duration when traversed by vehicle $v \in V$. These penalties must be expressed in the same units as the costs $c_{uw}$. We are now in a position to present the pickup and delivery problem with time windows formally. The mathematical formulation is:

$$
\begin{align*}
\text{Min} & \quad \sum_{u \in V} \sum_{w \in N} \sum_{v \in N} c_{uw} X_{uw} \\
& + \sum_{u \in P} f_u(T_u) + \sum_{v \in V} g_v(T_{2n+1} - T_0) \\
\text{subject to} & \quad \sum_{v \in V} \sum_{w \in N} X_{uw} = 1, \quad u \in P \\
& \quad \sum_{v \in N} X_{0w} - \sum_{v \in N} X_{uw} = 0, \quad u \in P, \quad v \in V \\
& \quad \sum_{v \in P^+} X_{uw} = 1, \quad u \in P, \quad v \in V \\
& \quad \sum_{v \in P^+} X_{w,2n+1} = 1, \quad v \in V \\
& \quad \sum_{v \in N} X_{uw} - \sum_{v \in N} X_{u,w,n+u} = 0, \quad u \in P^+, \quad v \in V \\
& \quad T_u + s_u + t_{uw} - T_w \leq 0, \quad u \in P^+, \quad w \in V \\
& \quad Y_u = 1 \\
& \quad X_{uw} = 1 \Rightarrow T_u + s_u + t_{uw} \leq T_w, \quad u, w \in P, \quad v \in V \\
& \quad X_{0w} = 1 \Rightarrow T_0 + t_{0w} \leq T_w, \quad w \in P^+, \quad v \in V \\
& \quad X_{u,2n+1} = 1 \Rightarrow T_u + s_u + t_{u,2n+1} \leq T_{2n+1}, \quad u \in P^-, \quad v \in V \\
& \quad a_u \leq T_u \leq b_u, \quad u \in P \\
& \quad a_0 \leq T_0 \leq b_0, \quad v \in V \\
& \quad a_{2n+1} \leq T_{2n+1} \leq b_{2n+1}, \quad v \in V
\end{align*}
$$

We seek to minimize the sum of the total travel cost, the total penalty associated with servicing customers too early or too late, and the total penalty associated with routes exceeding a given duration. Constraints (2)–(5) and (18) form a multicommodity minimum cost flow problem. Constraints (6) ensure that the same vehicle, $v$, visits both $u$ and $n + u$. Constraints (7) are precedence constraints which force node $u$ to be visited before node $n + u$. Next, constraints (8)–(10) describe the compatibility requirements between routes and schedules, while constraints (11)–(13) are the time window constraints. Finally, constraints (14)–(16) express the compatibility requirements between routes and vehicle loads, while constraints (17) are the capacity constraints.

Note that the formulation includes a route duration restriction $(b_{2n+1} - a_0)$. Note also that the formulation does not include subtour elimination constraints. Constraints (8) ensure subtour elimination (see also Solomon, and Desroiers et al.[151]). These constraints can be rewritten in an equivalent linear form using a large constant $M$:

$$
T_u + s_u + t_{uw} - T_w \leq (1 - X_{uw})M, \quad u, w \in P, \quad v \in V
$$

The subtour elimination constraints proposed by Miller et al.[90] for the traveling salesman problem are a special case of constraints (19) when all $t_{uw} = 1$, $s_u = 0$, and $M = |P|$. Let us now examine how models for different problem variants can be obtained from this formulation. The above formulation is the well known multivehicle, many-to-many dial-a-ride problem with time windows. If we let $|V| = 1$ and eliminate constraint (6) we obtain the single vehicle variant. To also consider the minimization of the fleet size one must add to the objective function the term $\sum_{v \in V} \sum_{u \in P^+} K X_{uw}$ and replace the equality sign with a less than or equal to sign in constraints (4) and (5). The formulation can also handle soft time windows. For this, one needs to eliminate constraints (11)–(13). Redefine now $N = PU[0, n + 1]$, and $P = P^+ = P^- = \{1, 2, \ldots, n\}$. To obtain the VRPTW we delete constraints (6), (7) and (15). Then constraints (8), as
well as (14) ensure subtour elimination. If the VRPTW involves only pickups, define \( d_u \geq 0 \). For delivery VRPTW, \( d_u \leq 0 \), and in constraint (17), set \( Y_u = D \) and eliminate \( Y_u \leq D \). The multiperiod VRP can be viewed as a VRPTW where the customers are replaced with requests for service and for each request there is a time window. By deleting constraints (6), (7) and the capacity constraints (14)–(17) we obtain a formulation for the multiple traveling salesman problem with time windows. Letting \( |V| = 1 \) we obtain the traveling salesman problem with time windows. The shoreline problem is obtained as follows. This structure is a restriction of the triangle-inequality metric where, if \( 1 \leq i \leq k \leq j \leq n \) are points located on the shoreline, then \( t_{ij} \geq t_{ik} + t_{kj} \). Furthermore, different shortest path problems with time windows can be obtained from the above formulation. For example, constraints (1), (3)–(5), (8)–(13), (18) and \( |V| = 1 \) constitute one shortest path problem with time windows. Others, such as shortest path problems with capacity, precedence, and time windows, can be obtained similarly. Finally, constraints (1), (2), (8)–(13), (18) and \( |V| = 1 \) form a spanning tree with time windows model. We now review the existing literature for routing problems with time windows.

2. THE TRAVELING SALESMAN PROBLEM (TSPTW)

The TSPTW is a VRPTW involving only one uncapacitated vehicle. Research on the TSPTW has focused on exact algorithms to minimize the total distance traveled. Christofides et al. discuss state space relaxations for dynamic programming approaches to this problem. The lower bounds obtained from the relaxed recursions could be maximized by the use of subgradient optimization and “state space ascent” and used in branch-and-bound algorithms for these problems. The authors provide a short summary of computational results. Problems with up to 50 customers are considered. For the largest problems, an average (over 5 problems) of 21 seconds, and a maximum of 42.2 seconds of CDC 6600 CPU time were used. It is claimed that “moderately tight” time windows are used, but no information is provided on the construction of the problem set. The reader may also refer to Frisch for a Pascal code of this algorithm.

Baker presents a branch-and-bound algorithm for a new, time oriented formulation of the TSPTW. The algorithm exploits the structure of the dual of a relaxation of the proposed model, which is a longest path problem on an acyclic network. This provides lower bounds on the original problem. Branching from a node in the enumeration tree creates two subproblems corresponding to the case when arrival at customer \( i \) precedes the arrival at customer \( j \) and vice versa. Computational experience is reported on a problem set derived from that of Eilon et al. and involving from 12 to 50 customers with tight time windows. The tightness of the time windows creates distinct allowable delivery times; in turn, these introduce precedence relationships between customer arrival times which leads to fewer vertices being examined in the branch-and-bound tree. The algorithm performed well on problems with distinct time windows with up to 50 nodes. The authors report solving 50-customer problems in less than 1 minute on a UNIVAC 1100/81A.

Finally, Savelsbergh proved that even finding a feasible solution to the TSPTW is an NP-complete problem. In this light, the author presents heuristic algorithms based on the \( k \)-interchange concept.

3. THE MULTIPLE TRAVELING SALESMAN PROBLEM (m-TSPTW)

The \( m \)-TSPTW is a VRPTW involving \( m \) uncapacitated vehicles. When the time windows consist of a single point, i.e., the arrival time at each node is prespecified, the problem of minimizing the number of vehicles is easily solved by appealing to Dilworth’s decomposition theorem, while the more general case involving also total distance minimization is easily tackled by using a network flow formulation (see Ford and Fulkerson).

In an early paper, Gerstbakh and Stern discuss a special case of the \( m \)-TSPTW where the objective is to minimize the fleet size; the authors present a heuristic based on the entropy principle of informational smoothing and an exact algorithm based on an integer programming formulation.

Aircraft, urban-bus and school-bus scheduling have proven to be very fruitful grounds for dealing with time windows. Minimizing the number of vehicles needed, Levin presents in the context of aircraft scheduling, an integer programming formulation with the time windows discretized. The method can be viewed as a heuristic for the problem with continuous time windows. Solutions are obtained by solving the LP relaxation which yields integer values quite often. In the same area, Martin presents a forward enumeration-type algorithm with exclusion rules to generate feasible schedules when the number of aircraft is fixed. Using the same integer programming formulation and approach as Levin, Sweeney and Ballard solved school-bus scheduling problems and have been successful in manually adjusting the few fractional values obtained without increasing the number of buses needed. Computational experience is reported using actual data and on 30 random problems. The number of buses was reduced by about 25%
when compared to the method in current use. The sensitivity of the number of buses to small changes in schools’ arrival time windows is also examined. The more general problem of how to assign school starting times in a district is examined in detail by Desrosiers et al.\cite{desrosiers1996}

In another early paper, Orlin\cite{orlin1986} presents a two-phase heuristic for school-bus scheduling. First, a schedule is created by building longer and longer segments through the solution of a series of matching problems. A 3-opt routine is then used to improve on a feasible schedule. Graham and Nuttle\cite{graham1987} report numerical tests with heuristics adapted from the previous procedures. The computational results suggest that each of the heuristics can provide good schedules. However, each has its shortcomings, as well.

With a more general objective function, which includes not only the fixed costs related to the number of vehicles, but also the travel time between trips, the school-bus scheduling problem is equivalent to the $m$-TSPTW. Bodin and Berman\cite{bodin1988} and Desrosiers et al.\cite{desrosiers1996} describe similar heuristic procedures for this problem. The approach in the first paper is to partition the time horizon and to link the routes in adjacent time periods by solving a transportation problem. In the second paper, the time horizon is not partitioned a priori. Soumis et al.\cite{soumis1992} propose an exact solution procedure. They relax the scheduling constraints, solve the underlying network problem and use the Carpaneto and Toth\cite{carpaneto1988} branching on flow variables to satisfy the time window constraints. Urban-bus scheduling problems of up to 128 and 158 trips are easily solved by this algorithm in less than 25 seconds on a CYBER 173. A superior branching strategy when the time windows are wider (e.g., school-bus scheduling) was found to be one based on time window partitioning (Desrosiers et al.\cite{desrosiers1985}).

A different primal algorithm which uses a column generation approach for a set partitioning formulation is presented in Desrosiers et al.\cite{desrosiers1985}. The columns are generated by solving a shortest path problem with time windows denoted SPPTW (see Section 3). The authors’ experience indicates that the LP relaxation of the set partitioning problem gives a very good lower bound even when the number of trips increases or the time windows are widened. Furthermore, tightening this bound by a cut requiring that the number of vehicles be rounded up to the next larger integer, led to the optimal fleet size in all the cases considered. The most difficult problem solved optimally (151 trips) took almost 20 minutes of CYBER 173 CPU time. Unpublished recent numerical results indicate that by using a better simplex code, the computational times decreased by at least 50%. Furthermore, the column generation scheme has been generalized to handle problems with several depots and with availability constraints on different pieces of work for drivers in Desrosiers et al.\cite{desrosiers1985}

Building on their previous efforts, Desrosiers et al.\cite{desrosiers1985} examine the use of Lagrangian relaxation methods for obtaining lower bounds on the minimum fleet size. The Lagrangian subproblem is again a shortest path problem with time windows while the most successful dual ascent method used was the augmented Lagrangian method in inequality form. This approach was successful in reducing the average computation time required to reach the best lower bound to 30% of that required by the primal column generation algorithm.

4. THE SHORTEST PATH PROBLEM (SPPTW)

This problem arises naturally as a subproblem in solving the $m$-TSPTW by employing a column generation scheme based on a set partitioning formulation or in Lagrangian relaxation procedures. The shortest path problem with time window constraints has been extensively studied by Desrosiers et al.\cite{desrosiers1985} and Desrochers and Soumis.\cite{desrochers1988, soumis1988}

Desrosiers et al.\cite{desrosiers1985} developed a generalization of the Ford-Bellman-Moore dynamic programming algorithm approach to the classical shortest path problem. They use a list of two-dimensional labels (time, cost) at each node of the network and discard dominated labels to reduce the storage requirements. Desrochers and Soumis\cite{desrochers1988} modified this algorithm by introducing a labels’ treatment order which generalizes Dijkstra’s\cite{dijkstra1959} algorithm, thus reducing the computation time. This algorithm solved problems having up to 25,000 nodes and 250,000 arcs in less than 1 minute on CYBER 173. Desrochers and Soumis\cite{desrochers1988} also present a primal-dual reoptimization algorithm which is particularly suitable when generating multiple disjoint routes in a column generation scheme for the $m$-TSPTW. Finally, Desrochers\cite{desrochers1988} generalizes the above primal-dual approach to the shortest path problem with multiple resource constraints. If one considers the multiple resource constraints as being vehicle capacity and time window constraints, the algorithm could be used to solve the VRPTW.

Another constrained shortest path problem is the one which occurs in the PDPTW. Dumas and Desrosiers\cite{dumas1989} present an efficient dynamic programming algorithm which was used to solve problems involving up to 55 requests (110 nodes).

As seen from the previous discussion, dynamic programming algorithms have been used with great success to obtain integer optimal solutions to the above constrained shortest path problems. These shortest path problems with tight constraints are easy to solve.
by dynamic programming since the sets of feasible states are relatively small. Furthermore, solving the subproblems with integer variables allows the exploration of the integrality gaps of the original problems and produces better bounds than the solution of a convex subproblem obtained by relaxing certain constraints as well as the integrality constraints. This gain on the quality of the bound is particularly noticeable for tightly constrained problems which generally present a large duality gap.

5. THE MINIMUM SPANNING TREE PROBLEM (MSTPTW)

A related line of research has dealt with the minimum spanning tree problem with time window constraints (MSTPTW). The importance of this research rests with the fact that it indicates that time window constraints alter the computational complexity of even “easy” problems involving routing components. While the minimum spanning tree problem can be solved optimally in $O(n^2)$ time (e.g., PRIM$^{[56]}$), in SOLOMON$^{[70]}$ it is shown that the problem with time windows is NP-hard by a polynomial transformation from the bounded diameter spanning tree problem.

Computational experiments conducted with two types of methodologies, one stemming from the minimum spanning tree problem solution (an extension of Prim's$^{[56]}$ algorithm), the other based on VRPTW approaches (an insertion heuristic) indicates that the insertion heuristic uniformly performed better than Prim's algorithm on both the solution quality and running time dimensions. Based on the above findings, a tentative conclusion is that the structure of this problem class is more closely related to that of the VRPTW than that of the minimum spanning tree.

6. THE VEHICLE ROUTING PROBLEM WITH TIME WINDOWS (VRPTW)

The early literature on the generic VRPTW deals mainly with case studies. PULLEN and WEBB$^{[56]}$ describe a system developed for duty scheduling of van drivers in a heavily time constrained environment. Of concern was the identification of when and where idle time occurs in schedules. The heuristic solution presented focuses on the allocation of jobs to vehicles to reduce idle- and empty-running time between jobs. Simulation is used to evaluate the procedure. KNIGHT and HOFER$^{[42]}$ present a case study involving a contract transport company. The problem is dominated by time windows ranging from 15 minutes to a whole day and averaging 1 to 2 hours. A heuristic manual system is developed to increase the utilization of vehicles as measured by the average number of calls per vehicle hour and the total number of vehicles used.

MADSEN$^{[47]}$ develops a simple algorithm based on Monte Carlo simulation to solve a routing problem with tight due dates faced by a large newspaper and magazine distribution company.

Given the intrinsic difficulty of this problem class, the later work on the generic VRPTW has typically focused on heuristic methods. The findings of SOLOMON and SAVELSBERGH$^{[84]}$ indicate that this problem class is fundamentally more difficult than the VRP. The VRPTW is NP-hard (by reduction from the VRP) and heuristics have so far offered the most promise for solving realistic size problems. SOLOMON$^{[70, 71]}$ has designed and analyzed a variety of route construction heuristics for the VRPTW. The results of the extensive computational study reported in Solomon$^{[70]}$ indicate that a sequential time-space insertion algorithm (a generalization of the Golp and Jameson$^{[6]}$ approach to the VRP) proved to be very successful in a number of important VRPTW environments. Very good initial solutions to a large number of 100-customer test problems were obtained in a matter of a few seconds of DEC10 CPU time. Several authors, including BAKER and SCHAEFFER$^{[9]}$ and SORENSEN$^{[44]}$ have confirmed the success of such approaches.

The excellent performance of such heuristics can be explained by realizing that while routing problems seem to be driven by the assignment of customers to vehicles—as indicated by the success of the FISHER and JAIKUMAR$^{[28]}$ generalized assignment heuristic—the sequencing aspect of the problem seems to drive routing problems dominated by time windows. It is this aspect of the problem that insertion approaches capture so well.

In related work, Baker and Schaffer$^{[23]}$ have conducted a computational study of route improvement procedures which were applied to heuristically generated initial solutions. The methods considered were extensions to the VRPTW of the 2-opt and 3-opt branch exchange procedures of LIN$^{[49]}$. The computational experience reported in this work suggests that while these methods were effective in improving the quality of the solutions generated by the initialization heuristics, the processing time required to obtain these results was large. A very large computational requirement was also reported by COOK and RUSSELL$^{[123]}$ in early work on the VRPTW for a $k$-optimal improvement heuristic, M-Tour (RUSSELL$^{[62]}$), which was effective in solving an actual problem with a few time constrained customers (the 163-customer problem with 15% time constrained customers was solved in 1.27 minutes of IBM 370/168 CPU time).

Several efficient implementations of branch exchange solution improvement procedures for VRPTW are developed in SOLOMON et al.$^{[73]}$ These implemen-
tations have resulted in an average decrease of 57% in required processing time. The efficiency of the computationally burdensome 3-opt procedure was shown to be drastically improved by the incorporation of pre-processing. Furthermore, the OROPT procedure (OR\textsuperscript{50}) was shown to produce improved solutions of comparable quality to those produced by the 3-opt, while requiring significantly less execution time. This procedure which was extended to the VRPTW to incorporate only orientation preserving branch exchanges, uses push-forward and backward shifts in customer arrival times for time window violation checks, the lexicographic ordering for processing suggested by Savelsbergh\textsuperscript{64} and it has a pre-processing option. The solutions produced from a 2-optimal starting solution by the OROPT procedure were, overall, within 1% of the solution produced by the full 3-opt from the same starting point, while OROPT was, on the average, 74% more efficient. All the codes were written in the C programming language and executed on an IBM-PC AT microcomputer. This is characteristic of the computing environment for much of the development and implementation currently underway for vehicle routing models.

Analytical results concerning the behavior of VRPTW approximation methods are derived through worst-case analysis of heuristics in Solomon\textsuperscript{71}. For a variety of heuristics, including improvement methods, it is shown that their worst-case behavior on n-customer problems, is at least order of n for minimizing the number of vehicles used, the total distance traveled, and the total schedule time, respectively. These results parallel similar findings by Frieze et al.\textsuperscript{32} for the asymmetric traveling salesman problem. To our knowledge, this is the first attempt to systematically develop and analyze heuristic approaches for the general VRPTW.

While heuristics have been found to be very effective and efficient in solving a wide range of practical size VRPTW, optimal approaches have lagged considerably behind. The only work that we are aware of on exact methods for this problem is that of Kolen et al.\textsuperscript{43} and Jornsten et al.\textsuperscript{38}. The former authors extend the shortest q-path relaxation algorithm of Christofides et al.\textsuperscript{9} to the problem with time windows. The largest problem solved to optimality involved 4 vehicles servicing 14 customers with tight time windows. With no branching required to obtain this solution, the algorithm took 0.58 minute of DEC 20/60 CPU time. Jornsten et al.\textsuperscript{38} propose a Lagrangian relaxation for the computation of a lower bound for the VRPTW. Using variable splitting, this relaxation necessitates the integer solution of two types of subproblems: the SPPTW and the generalized assignment problem. While it is conjectured that the method should produce good bounds, computational results have not yet been reported.

7. PICKUP AND DELIVERY PROBLEMS (PDPTW)

PDPTW is a generalization of the VRPTW which is concerned with the construction of optimal routes to satisfy transportation requests, each requiring both pickup and delivery under precedence, capacity and time window constraints (its formulation is presented in Section 1). Note that the VRPTW is the particular case of PDPTW where the destinations are all the common depot.

An optimal algorithm to minimize the total distance traveled has been designed in the context of goods transportation by Dumas\textsuperscript{84}. This algorithm uses a column generation scheme with a constrained shortest path subproblem (Dumas and Desrosiers\textsuperscript{29}). It has been tested on problems involving up to 55 requests (110 nodes) with tight vehicle capacity constraints. In addition, the algorithm can handle multiple depots and different types of vehicles.

Sexton and Choi\textsuperscript{86} present a heuristic Benders decomposition procedure to the single vehicle version with soft time windows. The authors consider an objective function minimizing a linear combination of total vehicle operating time and total customer penalty due to missing any of the time windows. The methodology involves a two-phase routing and scheduling procedure similar to the one developed in the dial-a-ride context (see below).

Most of the literature on pick-up and delivery problems has appeared for the dial-a-ride problem (DARP) for which Gavish and Srikanth\textsuperscript{84} give mathematical formulations. The DARP was first examined by Wilson et al.\textsuperscript{98-80} in connection with the development of real-time algorithms and many concepts such as sequential insertion of customers and form of the objective function are derived from that work. In the rest of this section we will focus first on the single vehicle case and then examine the multivehicle case.

It can be seen that the single vehicle DARPTW is a constrained version of the classical traveling salesman problem. Psarasftis\textsuperscript{56,57} presents two dynamic programming algorithms using respectively backward and forward recursions. The objective function minimizes the total customer inconvenience and “maximum position shifts” define time window constraints. These algorithms require \(O(n^3)\) time (where \(n\) is the number of customers), a fact which limits the tractable problem size to no more than 8–10 customers.

Sexton and Bodin\textsuperscript{86} give a mixed integer nonlinear programming formulation for the single vehicle dial-a-ride problem with given maximum delivery times or specified minimum pickup times to minimize
customer inconvenience. The authors' formulation cannot handle the "mixed" problem, where both types of constraints are present. Benders' decomposition is used to partition the problem into a routing component (a nonlinear integer program) and a scheduling component (a linear program). For the scheduling problem, an optimal algorithm which exploits the network flow structure of its dual is developed. The overall procedure is heuristic due to the difficulty of solving the routing problem exactly. Its computational tractability stems from the noniterative nature of the scheduling algorithm. Successful computational experience on moderately sized real data is reported. Problems with 7 to 20 customers are solved in an average of 18 seconds of UNIVAC 1100/81A CPU time.

Finally, Desrosiers et al. consider the objective function of minimizing the total distance traveled while respecting vehicle capacity and time windows. This objective function is less general than others proposed for minimizing user inconvenience. However, it is possible to take user inconvenience issues into account through the definition of the time windows. The paper presents an optimal forward dynamic programming algorithm which was capable of solving problems involving up to 40 customers (80 nodes) in less than 6 seconds on CYBER 173. Using a two-dimensional (time, cost) labeling, infeasible states are not defined and dominated states are discarded. Also, feasible states which due to the presence of time windows could not possibly be part of the optimal solution, are eliminated. This explains to a large extent why this algorithm performs so well.

For the multivehicle DARPTW version, we will present six studies. The Hung et al. procedure schedules one vehicle at a time and determines the best customer to be inserted into the route. When no more customers can be added to the route, this route is frozen and the process is repeated over the remaining set of unassigned customers. This is a sequential insertion procedure and generally, the routes generated first are the most productive, while the routes generated last tend to be inferior in quality.

The work of Roy et al. is based on a parallel insertion procedure. This algorithm takes a set of known requests and simultaneously constructs routes for all vehicles starting at the beginning of the day by using proximity measures. It can also insert new requests into a set of existing routes with the possibility of initializing new routes as needed.

Jaw et al. also use a parallel insertion algorithm which appears to be very effective and efficient in minimizing a weighted combination of customer dissatisfaction and system costs. The authors report computational experience with simulated and real data on the VAX 11/750. The former problems, involving 250 customers and 10–14 vehicles, took in the range of 20 seconds each. The latter problem, of much larger size than ever attempted before, involved about 2600 customers and 20 vehicles and was successfully solved in 12 minutes.

The Bodin and Sexton procedure is a traditional "cluster first, route second" approach. For a fixed fleet size, it partitions the set of requests into vehicle clusters and solves the resulting single vehicle dial-a-ride problems using the heuristic based on Benders' decomposition (Sexton and Bodin). Requests are then moved one at a time from one vehicle to another while attempting to reduce total user inconvenience.

Finally, Desrosiers et al. solve the problem by miniclustering first and optimal routing second. A heuristic algorithm groups together customers who can be efficiently served by the same vehicle route segment to form miniclusters. An optimal algorithm (see Section 2 on the m-TSPTW) then constructs routes corresponding to drivers' workdays by stringing together these vehicle route segments. In a methodology of "cluster first, route second" type, a cluster is formed from the set of customers assigned to a vehicle, and routing is carried out separately for each vehicle. The routing problems are usually fairly easy to solve. However, the most important decisions are taken at the clustering stage and it is very difficult to globally construct a good set of clusters. The authors' approach moves a part of the clustering problem into the routing problem. Problems with up to 200 customers and 85 mini-clusters are easily solved in less than 5 minutes on CYBER 173. Larger problems are solved by dividing the day into time slices and applying the algorithm several times. A real-life problem involving 880 customers and 282 miniclusters has been solved in 22 minutes on the same computer.

8. THE MULTI-PERIOD VEHICLE ROUTING PROBLEM (MPVRP)

Substantial progress has also been made for the MPVRP, a problem related to the VRPTW. Here, the time windows are full days and a service activity must occur on a specified number of days of the planning horizon. Hence, this problem can be viewed as a VRPTW where each customer has to be visited a given number of times within its multiple time windows. Most of the existing literature is based on applications that require the solution of large-scale problems.

Beltrami and Bodin address this problem in the context of routing hoist compactor trucks. The problem, which involves customers requiring service either 3 or 6 times per week, is solved heuristically. One approach involves developing the routes first and then
assigning them to days of the week. In the second approach, routing is performed after the customers have been randomly preassigned to days of the week.

Russell and Igo extend several routing heuristics to assign accounts to days of the week such that the weekly travel distance is minimized. The refuse collection problem considered involves the routing of 4 trucks through 490 locations requiring service from 1 to 6 times per week. The original problem is transformed into a pure routing problem by creating multiple copies of each account that requires service several times a week, resulting in a 776 customer problem. The best approach found was to assign the locations to days of the week using a simple heuristic, solve the resulting problem using a modified Clarke and Wright method, where the spacing conditions are enforced, and improve the solution through the use of a $k$-interchange heuristic (M-Tour, Russell). The use of M-Tour proved once again to be computationally prohibitive. A 192-customer, clustered version of the original problem required 6 minutes of IBM 370/168 CPU time.

Federgruen and Legeweg investigate the MPVRP as one of the issues arising in hierarchical distribution systems planning. They describe a heuristic for the initial customers-to-days assignment which was used in an application undertaken for a large producer and distributor of industrial gases. The basic idea of the heuristic is to perform the assignment sequentially, in decreasing order of frequency of service. Within each frequency class, measures of the benefit from having a given demand point being serviced on a certain day of the planning horizon are used to select the next point to be assigned.

Christofides and Beasley develop new heuristics for the MPVRP based on median problem and traveling salesman problem relaxations. The algorithms are characterized by the use of an interchange procedure which improves on initial choices of delivery-days combinations for customers.

Tan and Beasley present an extension of the Fisher and Jaikumar vehicle routing model and heuristic to the MPVRP. The computational experience reported seems to indicate the superiority of these approaches over that of Russell and Igo. Furthermore, the traveling salesman based algorithm of Christofides and Beasley seems to provide better quality solutions than that of Tan and Beasley at the expense of much higher CPU time.

A very successful application was undertaken by Fisher et al. (see also Bell et al.) for a major manufacturer of liquid oxygen and nitrogen. The authors present a mixed integer programming formulation and an algorithm based on Lagrangian relaxation, coupled with a multiplier adjustment method, for the problem of scheduling a fleet of vehicles that make bulk deliveries. A very interesting twist in this problem is that the frequency of delivery and the amount to be delivered to a customer are decision variables. The route-generation/route-selection approach described is made tractable by the fact that routes can have at most 4 customers. The authors exemplify their computational experience with a problem involving 15 vehicles, 54 customers and a planning horizon of 5 days. They report solving the problem within 0.5% of optimality in 2.52 minutes of AMDAHL-V8 CPU time. The multiplier adjustment method used in the above work to set dual variables has been shown to provide excellent bounds in a branch and bound algorithm for the set partitioning problem, as reported in Kedia. The success of dual ascent approaches to solve the linear set partitioning problem is confirmed by the work of Desrosiers et al.

## 9. The Shoreline Problem (SHP)

Recently, research has been initiated on an important new routing structure, the shoreline network. Shoreline problems arise in several transportation environments. For instance, in the routing and scheduling of cargo ships, the routing structure is “easy” because the ports to be visited are usually located along a real-world shoreline. However, because earliest pickup times of cargoes at ports generally complicate the routing structure, the resulting routing and scheduling problem is nontrivial.

Psaraftis et al. examine computational complexity issues and develop algorithms for a class of shoreline single-vehicle routing and scheduling problems with earliest pickup time constraints. For the straight-line case (a restriction of the shoreline case where $t_{ij} = t_{ik} + t_{kj}$), the authors show that the problem of minimizing maximum completion time can be solved exactly in quadratic time by dynamic programming.

For the general shoreline case, the traveling salesman problem without time constraints is polynomially solvable. On the other hand, as is well known, the Euclidean traveling salesman problem is NP-complete even in the absence of time windows (Papadimitriou). Nevertheless, the computational complexity of the shoreline problem with time windows remains open at this time. Psaraftis et al. develop and analyze heuristic algorithms for its solution. The authors derive data-dependent worst-case performance ratios for these heuristics and also discuss how these algorithms perform on practical data. Furthermore, they examine the computational complexity of other problem variants involving alternative objective functions and different types of time window constraints.
Additional work on this problem class can be found in Kim.\cite{61}

10. CONCLUDING REMARKS

In this paper, we have surveyed the research carried out on routing and scheduling problems with time windows, highlighting the significant breakthroughs in solution methodology and their analysis. While the up-to-date research provides an important advancement of the state-of-the-art in the routing and scheduling of vehicles, by no means is this problem class "well solved." There are many open research avenues and we describe some of them next.

To begin with, it is apparent from our survey that many methods have been proposed for the different variants of the problem. An immediate need is that of standardization of the computational experiments to be conducted. The comparison of different methods proposed on the same problem set is a must. For the generic VRPTW, an extensive test data bank proposed by Solomon\cite{69} has already been utilized by several researchers such as Baker and Schaffer\cite{58} and Sorensen.\cite{74} The trend of using the same test problem should be further expanded in this and other problem areas.

Concerning algorithmic development, the m-TSPTW area is a prime candidate. Efficient optimal algorithms are needed here for problems with many (more than 10) customers per route. Removal of some of the computational barriers now existing in this field for optimal approaches to the generic VRPTW should be a priority. One way to bring this about is to explore further the algorithmic possibilities created by the time window constraints in conjunction with recent advances in integer programming such as Lagrangian relaxation and column generation.

As this field matures, we expect to see an increasing degree of sophistication in the design of successful approximation methods; optimization-based heuristics are a prime candidate in this respect, especially column generation techniques which can easily handle multiple depots, different types of vehicles and cost functions. For routing problems, successful heuristics have utilized critical customers or "seeds" to initialize the procedures. The selection of seed customers permits the identification of geographic clusters in the customer set. For the VRPTW it is of interest to determine such clusters in a time-space framework.

There are important questions left open concerning the theoretical assessment of heuristics for the vehicle routing and scheduling problem with time windows. For example, is there a polynomial time heuristic with worst-case performance ratio bounded by a constant? Based on the results of Savelsbergh,\cite{64} Solomon,\cite{71} and Psaraftis et al.,\cite{58} we believe that, unless additional regularity conditions are imposed on the time windows, any heuristic will behave at least as badly as $O(n)$ on $n$-customer problems. Furthermore, since the probabilistic analysis of heuristics has the potential to explain the gap between the empirical and the worst-case performance of a heuristic, this type of analysis should be given serious consideration.

In many real-world settings, deviations from the standard routing and scheduling objectives have been observed. Examples include meeting due dates and the minimization of idle time. Problems with routing components have proved so challenging that no serious efforts have been made to study other criteria in models with sequence dependent processing times. Given their importance, additional criteria should be included in the future research agenda. For example, one could consider the problem of minimizing the number of tardy customer orders. As far as the minimization of waiting time is concerned, it is conjectured based on the results in Solomon\cite{69} that insertion heuristics will perform well in environments where this is the main objective.

Other research paths lead to the investigation of related problem structures. It would be of interest to examine further the "soft" problem, where some of the constraints can be violated at a cost. This problem can be viewed as a Lagrangian relaxation of the "hard" problem, where all the constraints must be satisfied.

Furthermore, the shoreline routing structure has opened an important research path from both a practical and a theoretical perspective. It would be of interest to examine the computational complexity and develop algorithms for the shoreline problem with both earliest pickup times and due dates.

As the algorithmic development work shifts more and more toward a PC-based environment, the appeal of interactive optimization increases. While this field has had a strong impact on the VRP (see for example Fisher et al.,\cite{50}), very little has yet been done for the VRPTW. Some preliminary research conducted by Sorensen\cite{74} indicates the advantages to be derived from interactive methods for the VRPTW. However much more research needs to be performed in this area.

The interface between the VRPTW and artificial intelligence is yet another area of strong interest. Given the complexity of routing problems with time window constraints, a VRPTW expert system is an immediate need. Different problem parameters such as routing structure, time windows' density and tightness, vehicle capacities, length of scheduling horizon, and number of customers per route would be important knowledge acquisition tools for the selec-
tion of the best method to use given the problem environment.

As can be seen from the above discussion, the multitude of possible research directions for vehicle routing and scheduling problems with time windows is perhaps only equaled by the number of challenges they pose. It is hoped that the present paper has done justice to the prior research and has stimulated a number of interesting research questions.

ACKNOWLEDGMENTS

The authors wish to thank Francois Soumis for his invaluable comments on an earlier draft of this paper. The authors also wish to thank two anonymous referees whose comments have substantially improved the paper. The research of the first author was supported, in part, by a Northeastern University, College of Business Administration, Research Grant. The research of the second author was supported, in part, by the Federal Government of Canada (NSERC Program) and by the Quebec Government (FCAR Program).

REFERENCES

52. I. Or, "Travelling Salesman-Type Combinatorial Problems and Their Relation to the Logistics of Blood Banking," Ph.D. dissertation, Department of Industrial Engineering and Management Sciences, Northwestern University, 1976.
TP 5598E, Transport Canada, Transport Development Center, Montreal, Canada, 1984.


(Received, February 1987; revision received August 1987; accepted September 1987)