Strategic Allocation of Time Windows in Vehicle Routing Problems under Uncertainty

Anirudh Subramanyam and Chrysanthos E. Gounaris*

Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213

Abstract

We consider the problem of allocating long-term delivery time windows to customers in the context of vehicle routing operations. Once the time windows have been assigned, the distributor must attempt to meet them on a daily basis as well as possible. Given that operational parameters, such as customer demands or travel times, vary from day to day, assigning time windows in a way that minimizes the expected routing costs is non-trivial. This problem can be modeled as a two-stage stochastic program where the time window assignments constitute first-stage decisions and the optimal vehicle routes constitute second-stage decisions that depend on the realizations of the uncertain parameters. We show that a sampled deterministic equivalent of this stochastic model can be reduced to a variant of a large class of problems known as Consistent Vehicle Routing Problems in the literature. We adapt an algorithmic framework we had previously developed for the latter class of problems and solve to guaranteed optimality instances of the former. Our approach is highly competitive when compared to the previous state-of-the-art as it achieves the fastest computation time on all literature benchmark instances. Moreover, and in contrast to existing approaches, we can readily incorporate uncertainty in customer presence and travel times.

Keywords

Vehicle Routing Problems, Time Window Assignment, Uncertainty, Service Consistency.

Introduction

The delivery (or pickup) of goods within a scheduled time window is widespread in several real world distribution networks. In many industries, these time windows are mutually agreed upon by the distributor and customer through long-term delivery contracts. For example, it is very common that deliveries are always made on the same day of the week and at about the same time of the day for an entire year. This is crucial for efficient inventory management as well as scheduling of personnel. Examples of applications where such operations are typical include among others vendor-managed maritime inventory routing (Zhang et al., 2015), online retail (Spliet and Gabor, 2015), attended home delivery (Agatz et al., 2011) and courier services (Jabali et al., 2015).

Once a time window has been agreed upon and communicated to the customer, the distributor must attempt to meet these time windows as well as possible

on an operational (e.g., daily) basis. The actual vehicle routes on any given day are obtained through the solution of a (capacitated) Vehicle Routing Problem with Time Windows (VRPTW) using parameter values (e.g., customer demands) that are realized on the day of delivery. This operational-level information is not known with certainty at the tactical planning stage when the time windows are to be decided. However, companies often have large amounts of historical data which can be used to obtain forecasts of different parameter scenarios (e.g., as perturbations from the nominal values). It is possible to take advantage of this information to generate better risk-averse solutions. Indeed, failure to take into account this information at the planning stage and utilizing nominal forecast values can lead to situations where the routing costs are unacceptably high.

The Time Window Assignment Vehicle Routing Problem (TWAVRP), was first introduced by Spliet and Gabor (2015). The authors considered the problem of assigning time windows of pre-specified width within some exogenous time windows (for e.g., arising from hours-of-work regulations) to a set of known customers under demand uncertainty. They assumed that a finite set of scenarios, each describing a realization of demand for every customer, is given with known probability of occurrence. They formulated the problem of designing time windows as well as vehicle routes satisfying these time windows for each of the postulated scenarios such that the expected routing costs are minimized, and designed a branch-price-and-cut algorithm that could solve instances with 25 customers and 3 scenarios to optimality. A similar approach was followed in Spliet and Desaulniers (2015), with the difference that the time windows were selected from a discrete set of a priori constructed windows.

Jabali et al. (2015) considered the assignment of time windows under deterministic customer demands and uncertain travel times in order to determine a single routing plan to be executed unchanged on each day of the year. They developed a Tabu Search heuristic method to design routes and a linear program to determine the optimal time windows under the assumption that at most one arc will be disrupted on each vehicle route. Their objective was to minimize the sum of traveling costs and expected overtime and tardiness penalty costs. Finally, in the context of e-commerce, Agatz et al. (2011) considered the problem of designing a discrete set of time windows (instead of just one) to offer to potential customers in different zip code areas under customer and demand uncertainty. They do not incorporate a detailed routing model but instead choose to estimate the expected routing costs using continuous approximation methods and aggregate-level routing models.

Seemingly different from the above problems and motivated in the context of operational level planning, the Consistent Vehicle Routing Problem (ConVRP) (Groër et al., 2009) aims to design minimum cost vehicle routes over a finite, multi-day horizon to serve a set of customers with known demands. The goal is to design routes that are *consistent* over time; this translates to satisfying any of the following requirements each time service is provided to a customer: (i) arrival-time consistency, wherein the customer should be visited at roughly the same time during the day, (ii) person-oriented consistency, in which the customer should be visited by the same driver, and whenever applicable, (iii) delivery consistency, for which a customer should receive roughly the same quantity of goods. We refer the reader to the survey by Kovacs et al. (2014) for an overview of this problem and its applications.

For the purposes of this study, we only need to focus on the ConVRP with the arrival-time consistency requirement (i). In this problem, every customer must be visited at roughly the same time on each day for which service is requested. The exact time of service is a decision variable but the difference between the earliest and the latest arrival times at each customer location must differ by no more than some pre-specified constant bound. Exact algorithms for the single-vehicle variant of the arrival-time ConVRP (a.k.a., Consistent Traveling Salesman Problem) were proposed by Subramanyam and Gounaris (2016a,b), and the authors were able to address to guaranteed optimality realistic instances containing up to 100 customers that require service over a 5-day horizon. Notably, the algorithm described in the second reference allows also for route duration limits and the option for a vehicle to idle en route.

In this paper, we show that the deterministic equivalent of any TWAVRP instance can be reduced to an instance of the arrival-time ConVRP. This has two consequences: (i) any algorithm developed for the latter can be used to obtain solutions for the former, and (ii) in the context of strategic time window allocation, a larger subset of parameters can be addressed under uncertainty, since the scope of the ConVRP is wider. Furthermore, we extend the exact algorithm of Subramanyam and Gounaris (2016b) developed for the Consistent Traveling Salesman Problem to the TWAVRP and solve benchmark instances to guaranteed optimality. A computational study shows that our proposed approach strongly outperforms the existing state-of-the-art method of Spliet and Gabor (2015) and can be readily extended to handle the case of uncertain customers.

Problem Definition

Let G = (V, A) denote a complete directed graph with nodes $V = \{0, 1, ..., n\}$ and arcs $A = \{(i, j) \in V \times V : i \neq j\}$. Node $0 \in V$ represents the unique depot, and each node $i \in V \setminus \{0\}$ corresponds to a customer with demand $q_i \in \mathbb{R}_+$. The depot has operating time windows $[e_0, \ell_0]$ and is equipped with an unlimited number of homogeneous vehicles of capacity Q^{1} . Each vehicle incurs a transportation cost $c_{ij} \in \mathbb{R}_+$ and a travel time $t_{ij} \in \mathbb{R}_+$ if it traverses the arc $(i, j) \in A$. Service times $s_i \in \mathbb{R}_+$ can be incorporated in the travel times

¹Although this exposition is in line with existing literature, our approach can readily address the limited fleet case.

via the operation $t_{ij} \leftarrow t_{ij} + s_i$ for all $(i, j) \in A$. For notational convenience, we refer to the set of customers $V \setminus \{0\}$ as V_C .

A set of routes $\mathbf{R} = (R_1, \ldots, R_m)$ represents a partition of the customer set V_C into $m \ge 1$ vehicle routes. Here, $R_k = (R_{k,1}, \ldots, R_{k,n_k})$ represents the k^{th} vehicle route, $R_{k,l}$ the l^{th} customer and n_k the number of customers visited by vehicle k. We assume that each customer must be visited exactly once by a single vehicle, i.e, split deliveries are not allowed. The cost of a route is evaluated as $c(\mathbf{R}) = \sum_{k=1}^m \sum_{l=0}^{n_k} c_{R_{k,l},R_{k,l+1}}$, where we define $R_{k,0} = R_{k,n_k+1} = 0$, for all $k \in K = \{1, \ldots, m\}$, i.e., each route starts and ends at the depot.

Let $w_i \in \mathbb{R}_+$ denote the pre-specified width of the time window to be assigned to customer $i \in V_C$. The decision-maker must then decide the starting time $x_i \in$ $[e_i, \ell_i]$ of the time window $[x_i, x_i + w_i]$ to be allocated to customer $i \in V_C$.² A route set $\mathbf{R} = (R_1, \ldots, R_m)$ is feasible under a fixed time window assignment [x, x + w] and under a specific demand realization q, if (i) all capacity constraints are satisfied, i.e., $\sum_{i \in R_k} q_i \leq Q$ for all $k \in K$, and (ii) all time window constraints are satisfied, i.e., there exists a vector of arrival times, $a \in \mathbb{R}^n_+$, that satisfies the following linear system:

$$a_{R_{k,1}} \ge e_0 + t_{0,R_{k,1}} \qquad \forall k \in K \tag{1}$$

$$a_{R_{k,l+1}} - a_{R_{k,l}} \ge t_{R_{k,l},R_{k,l+1}} \ \forall \ l \in \{1,\dots,n_k\}, \forall \ k \qquad (2)$$

 $a_{R_{k,n_k}} \le \ell_0 - t_{R_{k,n_k},0} \qquad \forall k \in K \tag{3}$

$$a \in [x, x+w] \tag{4}$$

Observe that, by this definition, if a vehicle arrives at customer location $i \in V_C$ at a time earlier than x_i , then it is allowed to wait until x_i . However, arriving later than $x_i + w_i$ is not permitted. We denote by $\mathcal{R}(x, q)$ the set of all feasible route sets for a given demand realization q and a fixed time window assignment [x, x + w]. Note that we suppress the dependence on parameters other than q because we assume for now that only q is uncertain.

If the probability distribution P of the uncertain demand vector q is known, then the goal of the TWAVRP is to design a vector of time windows [x, x+w] such that the expected cost of the associated routing instance is minimized:

$$\min_{x \in [e,\ell]} \mathbb{E}_P[\mathsf{VRPTW}(x,q)] \tag{5}$$

where
$$\operatorname{VRPTW}(x,q) = \min_{\mathbf{R} \in \mathcal{R}(x,q)} c(\mathbf{R})$$
 (6)

In practice, the exact probability distribution P is not available or is hard to obtain. Indeed, even if is exactly known, computing the objective function involves integrating the recourse function $\mathsf{VRPTW}(x, \cdot)$, which is almost impossible considering that the solution of the deterministic problem is itself challenging. Instead, we assume that we are given a finite set of demand scenarios q^s , $s \in S$, along with associated probabilities of occurrence p^s . In this case, we are interested in optimizing the deterministic equivalent of problem (5), obtained by replacing the expectation with a sample average:

$$\min_{x \in [e,\ell]} \sum_{s \in S} p^{s} \operatorname{VRPTW}(x, q^{s}) = \min_{\substack{x \in [e,\ell] \\ \mathbf{R}^{s} \in \mathcal{R}(x, q^{s})}} \sum_{s \in S} p^{s} c\left(\mathbf{R}^{s}\right) (7)$$

We denote by $(x, {\mathbf{R}^s}_{s \in S})$ a feasible solution to problem (7), i.e, a combination of time window decisions xand route sets \mathbf{R}^s feasible for the VRPTW with demand vectors q^s and time window vector [x, x + w].

We shall now construct a problem instance of the arrival-time ConVRP and show that the optimal value of this problem coincides with the optimal value of (7). Note that the ConVRP is a multi-period problem. The set of time periods in our ConVRP instance is S; that is, each scenario in the TWAVRP corresponds to a time period in the ConVRP. The set of customer nodes is V_C and the depot node is 0. Each customer $i \in V_C$ requires service in every period $s \in S$ of the planning horizon, within a time window $[e_i, \ell_i + w_i]$ that is common across periods, and requests demand quantity q_i^s in each period $s \in \mathcal{S}$. The set of arcs is identical to A, with the exception that the transportation costs are period-specific: each vehicle incurs a transportation cost $p^{s}c_{ij}$ when it traverses the arc $(i, j) \in A$ in time period $s \in S$. The travel times, service times and vehicle capacities remain the same as in the TWAVRP.

The maximum allowable arrival-time differential in our ConVRP instance is customer-specific and is equal to w_i for customer $i \in V_C$. Note that this represents the maximum allowable difference between the earliest and latest arrival times at location $i \in V_C$. More specifically, a set of routes $\{\mathbf{R}^s\}_{s\in\mathcal{S}}$ is feasible for the arrivaltime ConVRP if and only if (i) all capacity constraints are satisfied, i.e., $\sum_{i\in R_k^s} q_i^s \leq Q$ for all $k \in K, s \in \mathcal{S}$ and (ii) all arrival-time consistency constraints are satisfied; that is, there exist a vector of time window feasible

 $^{{}^{2}[}e_{i}, \ell_{i} + w_{i}]$ is the exogenous time window associated with $i \in V_{C}$. If no such time windows are applicable, we can set $e_{i} = 0$ and l_{i} to an appropriately large value.

arrival-times whose difference at any customer location is bounded by the maximum allowable value, i.e, there exist feasible solutions $a^s \in \mathbb{R}^n_+$ to the linear system (1)– (3), for each $s \in S$, that satisfy the following inequalities:

 $a_i^s \in [e_i, \ell_i + w_i] \qquad \forall i \in V_C, \forall s \in \mathcal{S}$ (8)

$$\max_{s \in \mathcal{S}} a_i^s - \min_{s \in \mathcal{S}} a_i^s \le w_i \quad \forall \, i \in V_C \tag{9}$$

We denote by \mathcal{RC} the set of all feasible ConVRP routes $\{\mathbf{R}^s\}_{s\in\mathcal{S}}$ as per the above definition. The cost of a feasible solution is evaluated as the sum of transportation costs across all time periods $\sum_{s\in\mathcal{S}} p^s c(\mathbf{R}^s)$. The arrival-time ConVRP can now be stated as follows:

$$\min_{\{\mathbf{R}^s\}_{s\in\mathcal{S}}\in\mathcal{RC}} \sum_{s\in S} p^s c\left(\mathbf{R}^s\right) \tag{10}$$

Algorithmic Framework

Observe that every feasible solution $(x, {\mathbf{R}^s}_{s \in S})$ to problem (7) corresponds to a feasible solution ${\mathbf{R}^s}_{s \in S}$ in problem (10). Moreover, it can be shown that for every feasible solution ${\mathbf{R}^s}_{s \in S}$ in problem (10), there exists some $x \in [e, \ell]$ such that $(x, {\mathbf{R}^s}_{s \in S})$ is feasible in problem (7). Therefore, since the objective functions of the two problems are identical, the two statements together imply that the optimal solutions of problems (7) and (10) coincide.

Thus, we can solve the deterministic equivalent of any TWAVRP instance by reducing it to an equivalent arrival-time ConVRP instance. We remark that the solution of the corresponding ConVRP instance does not give us the time window assignments x explicitly. Nevertheless, it can be easily computed post-optimization. Indeed, if $\{\mathbf{R}^s\}_{s\in S}$ is feasible in problem (10), then by definition, there exists a vector of arrival-times $\{a^s\}_{s\in S}$ that is feasible for the system of linear inequalities consisting of (1)–(3), imposed for each $s \in S$ and (8)–(9). Using any such vector of arrival-times, we can compute the explicit time window assignments by setting $x_i = \min\{\ell_i, \min_{s\in S} a_i^s\}$, for each $i \in V_C$. It is not difficult to show that the corresponding vector of arrivaltimes would satisfy $a^s \in [x, x + w]$ for each $s \in S$.

Note that in the most general case, the ConVRP allows for the possibility that not all customers require service in all time periods and that travel and service times may differ from one time period to the other. Translated in the context of a TWAVRP, this allows for the possibility to consider uncertainty in the existence of customer orders by constructing scenarios where different subsets of customers have been removed from the graph G. Furthermore, it allows for the possibility to consider uncertainty in travel times by constructing scenarios which represent perturbations from their nominal values. While these observations enable the simultaneous consideration of uncertainty in all of these parameters, it should be mentioned that doing so may come at the cost of an explosion in the number of scenarios that have to be considered.

In our computational experiments, we adapt the decomposition algorithm of Subramanyam and Gounaris (2016b), developed for the Consistent Traveling Salesman Problem, to solve our reduced TWAVRP instances to optimality. We shall now briefly summarize the main ingredients of the algorithm, translated into our context. The algorithm uses a branch-and-bound tree search to implicitly enforce the scenario-linking constraints (9) by solving within each node a set of *separable* VRPTW instances. The tree is initialized with the original problem instance where only the exogenous time windows (8) are enforced. The following linear program is then solved to check if the optimal node solution $\{\mathbf{R}^s\}_{s\in\mathcal{S}}$ is feasible in problem (10):

$$\min_{d \in \mathbb{R}^n} d \tag{11}$$

s.t.
$$d \ge \max_{s \in \mathcal{S}} a_i^s - \min_{s \in \mathcal{S}} a_i^s - w_i \quad \forall i \in V_C$$
 (12)

$$a^s \in \mathbb{R}^n$$
; Eq. (1) – (3), (8) $\forall s \in \mathcal{S}$ (13)

If the optimal solution of this linear program satisfies $d^* \leq 0$, then $\{\mathbf{R}^s\}_{s \in S}$ is a feasible solution for problem (10). Otherwise, if $d^* > 0$, there is at least one customer $i \in V_C$ which violates its corresponding constraint (9): $\max_{s \in S} a_i^{s*} - \min_{s \in S} a_i^{s*} > w_i$. In this case, the algorithm creates two new nodes by using the disjunction (14) with $\beta = (\max_{s \in S} a_i^{s*} + \min_{s \in S} a_i^{s*})/2$ as a branching rule.

Observe that for every feasible solution $\{\mathbf{R}^s\}_{s\in\mathcal{S}}$ in problem (10), there exists an arrival-time vector $\{a^s\}_{s\in\mathcal{S}}$ which satisfies the following disjunctive constraints:

$$[a_i^s \ge \beta - w_i/2 \,\forall s \in \mathcal{S}] \lor [a_i^s \le \beta + w_i/2 \,\forall s \in \mathcal{S}] \forall \beta \in \mathbb{R}, \,\forall i \in V_C$$
(14)

Since the expressions within each disjunct in (14) apply for all scenarios simultaneously, each node of the branchand-bound tree corresponds to solving a set of separable VRPTW instances, which is advantageous from a numerical point of view. In our implementation, we use the branch-and-cut algorithm proposed by Kallehauge et al. (2007) for the solution of these instances. Note that, although we did not pursue it in this work, the solution of these VRPTW instances can be conducted in parallel so as to expedite the overall algorithm.

Computational Results

We implemented our algorithm in C++. The runs were conducted on a single-core of an Intel Xeon 2.8 GHz processor and the C Callable Library of CPLEX 12.6 was used to implement the branch-and-cut based VRPTW algorithms of Kallehauge et al. (2007). In the implementation of these branch-and-cut algorithms, all CPLEX-generated cuts were disabled.

We tested our algorithmic framework on the set of 40 benchmark instances proposed by Spliet and Gabor (2015). This benchmark set contains instances with 10, 15, 20, and 25 customers (10 instances each). Each instance consists of three scenarios (representing low, medium, and high demand realizations) with equal probabilities of occurrence. The instances are inspired by a Dutch retail chain and we refer the reader to that paper for further details.

We now present a comparison of our method with the branch-price-and-cut method presented in Spliet and Gabor (2015). Their algorithm was implemented on an Intel Core i5-2450M CPU 2.5 GHz processor using CPLEX 12.4. In order to make a fair comparison, we use the same CPU time limit of 1 hour. In Tables 1 and 2, we summarize the performance of the two algorithms across the entire dataset of 40 instances.

Table 1. Summary of computational results of our decomposition algorithm.

	Proven Optimal		Residual Gap		
n	#	t (sec)	#	%	No UB
10	10	1.3	0	_	0
15	9	17.8	1	0.49	0
20	10	2.8	0	_	0
25	8	61.0	2	0.21	0
All	37	18.6	3	0.30	0

For each of the two approaches, we report the number of instances (out of 10) for which optimality was proved as well as the computational time required, averaged across the same instances. For those instances which could not be solved within the imposed time limit, we report the average residual gap defined as $\frac{\text{ub-lb}}{\text{ub}} \times 100\%$,

Table 2. Summary of results of the branch-price-and-cut algorithm of Spliet and Gabor (2015).

	Proven Optimal		Residu	Residual Gap	
n	#	t (sec)	#	%	No UB
10	10	25.4	0	_	0
15	8	83.6	1	0.25	1
20	8	83.6	1	0.25	1
25	5	792.4	0	_	5
All	31	188.0	2	0.25	7

where ub is the global upper bound and lb is the global lower bound of the branch-and-bound tree, whenever $ub < +\infty$ (i.e., an incumbent solution was found during the search process). The last column reports the number of instances for which a feasible solution could not be found in the imposed time limit.

Our algorithm is able to prove the optimality of 37/40 instances utilizing an average computation time of 18.6 seconds; of these instances, 6 were unsolved by the best previous method. Moreover, for the considered dataset, our approach always finds a feasible solution within the first ten seconds and the optimal solution within the first minute; the remaining time is spent trying to close the optimality gap. The new algorithm thus strongly outperforms the existing method, solving more problems and achieving the fastest computation time in almost all instances. These results showcase that the solution of TWAVRP instances via their reduction to equivalent ConVRP instances is also promising from a computational standpoint.

Conclusions

The strategic allocation of delivery time windows to customers is an important problem that arises in the context of vehicle routing operations. These long-term decisions can significantly impact daily operations: the distributor must strive to meet them as well as possible, since failure to do so may result in significant economic and reputational repercussions. In this regard, it is important to recognize and incorporate at the planning stage the inherent uncertainty in parameters like customer demands and travel times, whose actual values will be realized only on the day of operation.

In this paper, we studied the Time Window Assignment Vehicle Routing Problem, which can be modeled as a two-stage stochastic program where the time window assignments constitute first-stage decisions and the actual daily vehicle routes constitute second-stage decisions. We showed that a sampled deterministic equivalent of this stochastic model can be reduced to a wellstudied class of Consistent Vehicle Routing Problems with arrival-time consistency requirements. We adapted an exact algorithm for the latter class of problems and solved to guaranteed optimality TWAVRP benchmark literature instances that considered demand uncertainty. In addition to being numerically superior to the existing state-of-the-art method, our approach enjoys the benefit of being able to readily incorporate uncertainty in other parameters like customer presence and travel times. The precise incorporation of these uncertain parameters and the consideration of other time window assignment policies (see, e.g., Spliet and Desaulniers (2015); Agatz et al. (2011)) remains to be investigated in the future.

Acknowledgments

The authors gratefully acknowledge support from the U.S. National Science Foundation (award number CMMI–1434682). Anirudh Subramanyam also acknowledges support from the John and Claire Bertucci Graduate Fellowship Program.

References

- Agatz, N., Campbell, A., Fleischmann, M., and Savelsbergh, M. (2011). Time Slot Management in Attended Home Delivery. *Transportation Science*, 45(3):435–449.
- Groër, C., Golden, B., and Wasil, E. (2009). The Consistent Vehicle Routing Problem. Manufacturing & Service Operations Management, 11(4):630–643.
- Jabali, O., Leus, R., Van Woensel, T., and de Kok, T. (2015). Self-imposed time windows in vehicle routing problems. OR Spectrum, 37(2):331–352.
- Kallehauge, B., Boland, N., and Madsen, O. B. (2007). Path inequalities for the vehicle routing problem with time windows. *Networks*, 49(4):273–293.
- Kovacs, A. A., Golden, B. L., Hartl, R. F., and Parragh, S. N. (2014). Vehicle routing problems in which consistency considerations are important: A survey. *Networks*, 64(3):192–213.
- Spliet, R. and Desaulniers, G. (2015). The discrete time window assignment vehicle routing problem. *European Jour*nal of Operational Research, 244(2):379–391.

- Spliet, R. and Gabor, A. F. (2015). The Time Window Assignment Vehicle Routing Problem. *Transportation Sci*ence, 49(4):721–731.
- Subramanyam, A. and Gounaris, C. E. (2016a). A branchand-cut framework for the consistent traveling salesman problem. *European Journal of Operational Research*, 248(2):384–395.
- Subramanyam, A. and Gounaris, C. E. (2016b). A Decomposition Algorithm for the Consistent Traveling Salesman Problem with Vehicle Idling. *Transportation Science*, Accepted for publication.
- Zhang, C., Nemhauser, G., Sokol, J., Cheon, M.-S., and Papageorgiou, D. (2015). Robust Inventory Routing with Flexible Time Window Allocation. Available on Optimization Online.