

# A Stochastic Vehicle Routing Problem with Travel Time Uncertainty: Trade-Off Between Cost and Customer Service

Junlong Zhang · William H. K. Lam · Bi Yu Chen

Published online: 30 May 2013  
© Springer Science+Business Media New York 2013

**Abstract** On-time shipment delivery is critical for just-in-time production and quick response logistics. Due to uncertainties in travel and service times, on-time arrival probability of vehicles at customer locations can not be ensured. Therefore, on-time shipment delivery is a challenging job for carriers in congested road networks. In this paper, such on-time shipment delivery problems are formulated as a stochastic vehicle routing problem with soft time windows under travel and service time uncertainties. A new stochastic programming model is proposed to minimize carrier's total cost, while guaranteeing a minimum on-time arrival probability at each customer location. The aim of this model is to find a good trade-off between carrier's total cost and customer service level. To solve the proposed model, an iterated tabu search heuristic algorithm was developed, incorporating a route reduction mechanism. A discrete approximation method is proposed for generating arrival time distributions of vehicles in the presence of time windows. Several numerical examples were conducted to demonstrate the applicability of the proposed model and solution algorithm.

**Keywords** Vehicle routing · Time window · Customer service · Stochastic programming · Tabu search · Discrete approximation

## 1 Introduction

The Vehicle Routing Problem (VRP), introduced by Dantzig and Ramser (1959), involves the design of a set of minimum-cost routes for the vehicles of a logistics

---

J. Zhang (✉) · W. H. K. Lam · B. Y. Chen  
Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University,  
Hung Hom, Kowloon, Hong Kong  
e-mail: junlong.zhang@connect.polyu.hk

B. Y. Chen  
State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,  
Wuhan University, Wuhan 430079, China

company. The routes must start at the depot, serve a group of geographically scattered customers, and finally return to the depot. Each customer can only be visited once by one single vehicle. The total demand of each route cannot exceed the capacity of the vehicle and the duration of each route cannot exceed a given limit. VRP has broad applications in distribution and logistics management fields. In the last 50 years, strides have been made in the development of efficient and effective solution algorithms using both exact and heuristic approaches (Laporte 2009). More than one hundred software companies are now selling commercial vehicle routing software and thousands of logistics companies are using VRP software (Drexl 2012).

In the literature, a number of VRP variants have been intensively studied (Toth and Vigo 2002; Golden et al. 2008; Leung et al. 2011; Norouzi et al. 2012; Yu et al. 2011; Escuín et al. 2012; Li et al. 2012). Vehicle routing problems with time windows (VRPTW) form a large proportion of the matters studied. One example is the request for a vehicle to start service within a given time interval (i.e. time window). The time window constraints can be modeled as either hard or soft. In the hard time window case, customers refuse the service of late arrival vehicles (Solomon 1987; Savelsbergh 1992; Cordeau et al. 2002; Nagata et al. 2009; Yu and Yang 2011). In the soft time window case, customers accept the vehicle service regardless of arrival time, but nonetheless penalties for earliness or tardiness are incurred (Koskosididis et al. 1992; Balakrishnan 1993; Taillard et al. 1997; Chiang and Russell 2004; Liberatore et al. 2011). In both the hard and soft cases, early arrival vehicles must wait until the customer's requested service time arrives.

In most of VRP studies, the three elements including demand, customers and travel times are assumed to be deterministic. In reality, however, all these elements can be highly stochastic due to the complexity of the real logistics applications. For example, travel times in urban road networks are highly stochastic due to roadway capacity variations and traffic demand fluctuations (Lam et al. 2008; Chen et al. 2011, 2012; Li et al. 2012; Wei et al. 2012). Ignoring the stochastic nature of these elements may lead to sub-optimal even infeasible delivery solutions. In view of this, researchers have investigated the stochastic vehicle routing problem (SVRP) by considering stochastic demands and/or customers (Bertsimas and van Ryzin 1991; Gendreau et al. 1995; Laporte et al. 2002; Lei et al. 2011) and uncertain travel times (Laporte et al. 1992; Lambert et al. 1993; Kenyon and Morton 2003; Zhang et al. 2012).

Some researchers have also investigated SVRP with time windows under travel time uncertainties. Ando and Taniguchi (2006) studied SVRP with soft time window constraints. A model was proposed to minimize carrier's total cost, which is comprised of the fixed vehicle employment cost, operating cost and penalty cost. The operating cost was assumed to be proportional to the total mean travel time, while the penalty cost was formulated as the expected earliness and tardiness of vehicles at customer locations. Travel time distributions were estimated from probe vehicle data.

Russell and Urban (2008) developed a multiple-objective model for SVRP with soft time window constraints. Priorities among different objectives were assumed in order of the number of required vehicles, total distance traveled and time-window penalties incurred. The model minimizes a weighted average of these objectives. A tabu search heuristics was developed to solve the model. To reduce the number of required vehicles, the fixed vehicle employment cost was multiplied by a large weighting parameter in the objective function. The earliness and tardiness penalties

due to time window violation were deduced based on the assumption of Erlang travel time distributions.

Li et al. (2010) investigated SVRP in both the hard and soft time window cases. Uncertain service times were also considered. In the soft time window case, a two-stage stochastic programming with recourse model was built. The objective of the model is to design a set of routes in the first stage and to minimize the expected costs in the second stage when random travel and service times are realized. A tabu search heuristics was also developed to solve the model. Travel and service times were assumed to be normally distributed and stochastic simulation was used for probability check and computing expected values. In the hard time window case, a chance-constrained model was proposed to ensure that the probability of vehicles arriving at customers within the time windows is at least a pre-specified value. Expected earliness and tardiness of vehicles were not included in carrier's total cost in the hard time window case.

On the basis of the previous works, this paper aims to investigate SVRP with soft time window constraints under travel and service time uncertainties. In the rest of the paper, the SVRP with soft time window constraints is referred to as SVRPSTW. The previous works is extended in the following two aspects.

- (1) A new stochastic programming model is proposed in this paper not only to minimize carrier's total cost, but also to guarantee a minimum on-time arrival probability at each customer location. The previous SVRPSTW models mainly focused on reducing carrier's total cost. This optimization approach is essentially formulated from the perspective of the carrier in order to provide shipment delivery service at minimum total cost. By this approach, the probability of late shipment delivery for some customers may be quite high. In practice, inventories are limited due to high holding costs and therefore production or sales processes may be disrupted by frequently delayed shipments. It is common for customers to require a certain probability of on-time shipment delivery, though late service is at times permitted. In this paper, the on-time arrival probability at each customer location is explicitly formulated in the proposed model by the introduction of a customer service level constraint. The probability of on-time shipment delivery to each customer can be then ensured. The advantages of the proposed model are listed below.
  - The proposed model is a generalization of the conventional recourse models for SVRPSTW in the literature (Russell and Urban 2008; Li et al. 2010). When no customer service level constraint is imposed, solutions of the proposed model are the same as the previous SVRPSTW models.
  - The proposed model provides an easy way of exploring trade-offs between carrier's total cost and customer service level, simply by adjusting the customer service level constraints.
  - The proposed model can easily be adapted to multi-class customers with various service level preferences by imposing appropriate customer service level constraints.
- (2) An iterated tabu search heuristic algorithm is developed to solve the proposed model. A route reduction mechanism is designed and incorporated in the developed heuristics. When performing a neighborhood search, trial move costs that remain unchanged in one iteration are kept in the memory for use in the next

iteration. On the basis of these cost records, every route that can be decomposed and inserted into other routes is identified in an SVRPSTW solution. In this way, the number of routes can be reduced. Additionally, inspired by Miller-Hooks and Mahmassani (1998) and Chen et al. (unpublished), an approximation method called  $\alpha$ -discrete is proposed in this paper for generating arrival time distributions of vehicles in the presence of time windows. Using this approximation method, SVRPSTW solutions can be evaluated without suffering restrictions on the assumption of travel and service time distributions.

The remainder of this paper is organized as follows. The proposed model is formulated in the following section. The iterative tabu search heuristics designed for solving the proposed model is presented in Section 3. The approximation method for estimating arrival time distributions is described in Section 4. Computational results are shown in Section 5. Finally, concluding remarks are given in Section 6.

## 2 Model Formulation

Let  $G=(V_0, A)$  be a complete digraph, where  $V_0 = \{0, \dots, n\}$  is the vertex set and  $A = \{(i, j) : i, j \in V_0, i \neq j\}$  is the arc set. Vertex 0 represents the depot where  $m_0$  identical vehicles with capacity  $Q$  are available. The customer set is denoted as  $V = V_0 / \{0\} = \{1, \dots, n\}$ . Each customer  $i \in V$  has a nonnegative demand  $q_i$ , a service time  $S_i$  and a time window  $[e_i, l_i]$ . It is expected that service at customer  $i$  begins within  $[e_i, l_i]$ . If the vehicle arrives at customer  $i$ 's location before  $e_i$ , it has to wait until  $e_i$ ; if it arrives at customer  $i$ 's location after  $l_i$ , a penalty proportional to the lateness must be paid. A time window  $[e_0, l_0]$  is also associated with the depot, where  $e_0$  represents the earliest possible departure time from the depot and  $l_0$  represents the latest possible arrival time at the depot. A travel time  $T_{ij}$  is associated with each arc  $(i, j) \in A$ . Both  $T_{ij}$  and  $S_i$  are random variables with distributions assumed to be known and independent of everything else. Additional assumptions are:  $Q \geq q_i, i \in V$  (i.e. each vehicle can serve at least one customer) and  $m_0$  is big enough (i.e. there are sufficient vehicles at the depot). Additional notation is listed as follows:

- $M$  a sufficiently large number
- $f$  fixed cost of employing one vehicle
- $m$  number of required vehicles in a feasible solution,  $m \leq m_0$
- $m^*$  number of required vehicles in the optimal solution,  $m^* \leq m$
- $\mathbf{K}$  the set of required vehicles in a feasible solution,  $\mathbf{K} = \{1, 2, \dots, m\}$
- $x_{ijk}$  a binary variable associated with each arc  $(i, j) \in A$ . It is equal to 1 if and only if arc  $(i, j)$  is traversed by vehicle  $k$  and 0 otherwise,  $k \in \mathbf{K}$
- $\mathbf{R}_k$  route  $k$  defined as  $\mathbf{R}_k = \{r_0 = 0, r_1, \dots, r_j, r_{j+1}, \dots, r_{n_k}, r_{n_k+1} = 0\}$ , where  $n_k$  is the number of customers assigned to vehicle  $k$ ,  $r_j \in V_0, 0 \leq j \leq n_k + 1, k \in \mathbf{K}$
- $A_{r,j,k}$  arrival time of vehicle  $k$  at vertex  $r_j$ 's location,  $r_j \in \mathbf{R}_k, 1 \leq j \leq n_k + 1, k \in \mathbf{K}$
- $T_{r_j,k}^{ss}$  service start time of vehicle  $k$  at customer  $r_j$ ,  $r_j \in \mathbf{R}_k, 1 \leq j \leq n_k, k \in \mathbf{K}$
- $d_{0k}$  departure time of vehicle  $k$  from the depot,  $k \in \mathbf{K}$

- $D_{r_j,k}$  departure time of vehicle  $k$  from customer  $r_j$ 's location,  $r_j \in \mathbf{R}_k, 1 \leq j \leq n_k, k \in \mathbf{K}$
- $W_{r_j,k}$  earliness (waiting time) of vehicle  $k$  at customer  $r_j$ 's location,  $r_j \in \mathbf{R}_k, 1 \leq j \leq n_k, k \in \mathbf{K}$
- $P_{r_j,k}$  tardiness of vehicle  $k$  at vertex  $r_j$ 's location,  $r_j \in \mathbf{R}_k, 1 \leq j \leq n_k + 1, k \in \mathbf{K}$
- $B_k$  duration of route  $k, k \in \mathbf{K}$
- $h$  upper bound of the duration of each route
- $Z_k$  excess route duration for vehicle  $k, k \in \mathbf{K}$
- $\lambda_{1i}$  penalty coefficient for earliness at customer  $i$ 's location,  $i \in \mathbf{V}$
- $\lambda_{2i}$  penalty coefficient for tardiness at customer  $i$ 's location,  $i \in \mathbf{V}$
- $\lambda_{2,0}$  penalty coefficient for tardiness when a vehicle returns to the depot
- $\lambda_3$  penalty coefficient for excess route duration
- $\alpha_i$  required probability of on-time shipment delivery (i.e. required service level) by customer  $i, i \in \mathbf{V}$
- $\alpha_0$  required on-time arrival probability when a vehicle returns to the depot
- $\beta$  required probability that duration of each route is smaller than  $h$ .

The stochastic programming model for SVRPSTW proposed in this paper is given below.

Min

$$Mf \cdot m + \sum_{(i,j) \in \mathbf{A}} \sum_{k \in \mathbf{K}} E(T_{ij})x_{ijk} + \sum_{k \in \mathbf{K}} \left( \sum_{j=1}^{n_k} \lambda_{1r_j} E(W_{r_j,k}) + \sum_{j=1}^{n_k+1} \lambda_{2r_j} E(P_{r_j,k}) + \lambda_3 E(Z_k) \right) \quad (1)$$

Subject to

$$\sum_{j \in \mathbf{V}_0} \sum_{k \in \mathbf{K}} x_{ijk} = 1, \quad \forall i \in \mathbf{V} \quad (2)$$

$$\sum_{j \in \mathbf{V}} x_{0jk} = 1, \quad \forall k \in \mathbf{K} \quad (3)$$

$$\sum_{i \in \mathbf{V}} x_{i0k} = 1, \quad \forall k \in \mathbf{K} \quad (4)$$

$$\sum_{i \in \mathbf{V}_0} x_{ijk} - \sum_{i \in \mathbf{V}_0} x_{jik} = 0, \quad \forall j \in \mathbf{V}, k \in \mathbf{K} \quad (5)$$

$$\sum_{i \in \mathbf{V}} q_i \sum_{j \in \mathbf{V}_0} x_{ijk} \leq Q, \quad \forall k \in \mathbf{K} \quad (6)$$

$$P\{A_{r_j,k} \leq l_{r_j}\} \geq \alpha_{r_j}, \quad \forall r_j \in \mathbf{R}_k, 1 \leq j \leq n_k + 1, k \in \mathbf{K} \quad (7)$$

$$P\{B_k \leq h\} \geq \beta, \quad \forall k \in \mathbf{K} \quad (8)$$

$$x_{ijk} = \{0, 1\}, \quad \forall i, j \in \mathbf{V}_0, k \in \mathbf{K} \tag{9}$$

The objective function Eq. (1) consists of three parts: 1. fixed vehicle employment cost, 2. total mean travel time as the operating cost, and 3. weighted expected earliness, tardiness and excess route duration as penalty cost.  $W_{r,j,k}$ ,  $P_{r,j,k}$  and  $Z_k$  are given in Eqs. (10), (11) and (13) respectively. Illustrations of the penalty coefficients  $\lambda_{1i}$  and  $\lambda_{2i}$ ,  $i \in \mathbf{V}$  can be found in Fig. 1. The proposed model has a hierarchical optimization objective: the primary objective is to minimize the number of required vehicles to satisfy constraints (2) to (9); the secondary objective is to minimize the operating and penalty costs given the minimized number of vehicles  $m$ . This hierarchical optimization objective implies that one SVRPSTW solution with fewer routes but higher operating and penalty costs is better than another with more routes but lower operating and penalty costs.

$$W_{r,j,k} = \max\{e_{r_j} - A_{r,j,k}, 0\}, \quad r_j \in \mathbf{R}_k, 1 \leq j \leq n_k, k \in \mathbf{K} \tag{10}$$

$$P_{r,j,k} = \max\{A_{r,j,k} - l_{r_j}, 0\}, \quad r_j \in \mathbf{R}_k, 1 \leq j \leq n_k + 1, k \in \mathbf{K} \tag{11}$$

$$B_k = T_{r_0 r_1} + \sum_{j=1}^{n_k} (W_{r_j,k} + S_{r_j} + T_{r_j r_{j+1}}), \quad r_j \in \mathbf{R}_k, k \in \mathbf{K} \tag{12}$$

$$Z_k = \max\{B_k - h, 0\}, \quad k \in \mathbf{K} \tag{13}$$

Equation (2) indicates that each customer must be visited exactly once by one vehicle. Equations (3) and (4) ensure that each vehicle starts and ends its route at the depot. Equation (5) ensures that each vehicle departs from a customer location after it visits the customer. Equation (6) is the capacity constraint. Equation (7) is the customer service level constraint and ensures that the probability of on-time shipment delivery (i.e. service level) to each customer is at least a predefined value  $\alpha_i$  (i.e. required service level). Similarly, when a vehicle returns to the depot, the on-time arrival probability at the depot must be larger than a threshold  $\alpha_0$ . The required service level of each customer can be adjusted according to practical requests. Equation (8) ensures that each route is completed within  $h$  with at least probability  $\beta$ . If  $\alpha_i$  in Eq. (7) and  $\beta$  in Eq. (8) are set to zero, the proposed model then reduces to the conventional recourse models for SVRPSTW (Russell and Urban 2008; Li et al. 2010). Equation (9) defines the domain of the decision variables.

Figure 1 shows the relationship between the time window specified by customer  $i$  and possible arrival times of vehicle  $k$ . If the vehicle arrives earlier (later) than  $e_i$  ( $l_i$ ), an earliness (tardiness) penalty cost will then be incurred with unit earliness (tardiness)

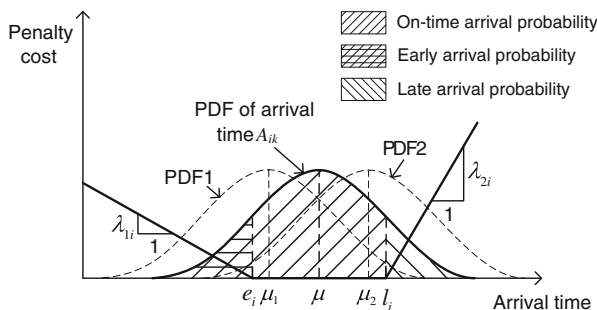


Fig. 1 Time window  $[e_i, l_i]$  and arrival time  $A_{ik}$  of vehicle  $k$  at customer  $i$ 's location

penalized by  $\lambda_{1i}$  ( $\lambda_{2i}$ ). The probability density functions (PDF) of three possible vehicle arrival times are also shown in Fig. 1. In deterministic cases, arrival times  $\mu_1$  and  $\mu_2$  may be accepted because both lie within the time window. While under travel time uncertainties, arrival times with PDF1 and PDF2 might not be accepted since PDF1 can possibly lead to a long waiting time and PDF2 may result in a large late arrival probability. In this sense, it is reasonable and necessary to incur an earliness penalty cost and to impose a constraint on the vehicle's minimum on-time arrival probability.

In the proposed model, objective function Eq. (1) represents carrier's total cost and constraint Eq. (7) guarantees customers' required service level. In order to highlight the primary optimization objective of the proposed model, fixed vehicle employment cost is multiplied by a sufficiently large number  $M$  in Eq. (1). By adjusting  $\alpha_i, i \in \mathbf{V}$  in Eq. (7), various trade-offs between carrier's total cost and customer service level can be obtained.

An alternative way to trade off between carrier's total cost and customer service level is to adjust the penalty coefficient  $\lambda_{2i}, i \in \mathbf{V}$  in Eq. (1), without imposing any customer service level constraint as Eq. (7). If  $M$  in Eq. (1) is a moderate number,  $\lambda_{2i}$  can be then increased to a level such that tardiness penalty cost is comparable to the fixed vehicle employment cost. In this way the trade-off between cost and service can also be found. However, the increase of  $\lambda_{2i}$  may be irrational as fixed cost of employing one vehicle is usually much larger than the unit penalty cost (e.g. 1000 versus 0.5 in Russell and Urban 2008). Irrational increase of  $\lambda_{2i}$  may lead to unrealistic cost coefficients and even irrational routing solutions (Koskosidis et al. 1992). Compared with this method, the proposed model in this paper provides a more straightforward way of achieving the goal of exploring trade-offs between carrier's total cost and customer service level. The probability of on-time arrival to each customer is explicitly formulated in the proposed model and guaranteed by customer service level constraints. This has a clear implication in practical applications.

### 3 Solution Algorithm

In this section, a heuristic algorithm based on an iterated tabu search (ITS) by Cordeau and Maischberger (2012) is developed for solving the proposed model. In ITS, an iterated local search (Lourenço et al. 2002) is used as the general framework and a tabu search is adopted as the local search improvement method. The major strengths of ITS are simplicity, flexibility and efficiency. The neighborhood structure of the tabu search heuristics in ITS is simple and a single type of solution perturbation is adopted. ITS is also flexible as it can solve many VRP variants without changing the methodology and parameter settings. Finally, ITS is reasonably fast and effective. Cordeau and Maischberger (2012) reported solving classical VRP and seven VRP variants by using ITS, and showed the competitiveness of ITS with other heuristics for each particular problem variant.

The heuristic algorithm developed in this section is a modified version of ITS, denoted as SVRP-ITS. Minor extensions of ITS are made in SVRP-ITS: (a) a direct route reduction mechanism is incorporated in SVRP-ITS, aiming to reduce the number of required vehicles in a SVRPSTW solution; (b) when performing a neighborhood search, trial move costs that remain unchanged in one iteration are kept in the memory for use in the next iteration. These cost records can be used to avoid possible repetitive computing in both neighborhood search and route reduction

processes. In the following sections, the main framework of SVRP-ITS is shown first, followed by brief descriptions of each component.

### 3.1 Main Framework of SVRP-ITS

The main framework of SVRP-ITS is summarized in Algorithm SVRP-ITS. Firstly a feasible initial solution  $s_0$  is constructed (Section 3.2). In each iteration of SVRP-ITS, tabu search (Section 3.3) is then used to improve the current solution  $s'$ , resulting in an improved solution  $\tilde{s}$ . The best feasible solution  $s^*$  is updated accordingly. Before the start of the next iteration,  $s'$  is renewed employing a perturbation mechanism (Section 3.5). If  $\tilde{s}$  satisfies the acceptance criterion (Section 3.4),  $\tilde{s}$  is then perturbed, otherwise  $s^*$  is perturbed. Let  $\eta$  be the maximum number of tabu search iterations allowed to be performed during the entire search process of SVRP-ITS.  $\zeta$  represents the total number of tabu search iterations performed so far. SVRP-ITS stops if  $\zeta$  is larger than  $\eta$ . Note that a tabu search iteration denotes an iteration within a tabu search process (see Step 2 in Procedure Tabu Search);  $c(s)$  is set to the objective function Eq. (1).

---

#### Algorithm SVRP-ITS

##### Step 1. Initialization.

$s_0 \leftarrow$  Initial Solution Construction()

$s^* \leftarrow s_0, s' \leftarrow s_0$

##### Step 2. Iterated local search.

while  $\zeta$  is smaller than  $\eta$  do

##### Step 2.1 Local search improvement.

$\tilde{s} \leftarrow$  Tabu Search( $s'$ )

##### Step 2.2 Update the best feasible solution $s^*$ found so far.

if  $\tilde{s}$  is feasible and  $c(\tilde{s}) < c(s^*)$  then

$s^* \leftarrow \tilde{s}$

end if

##### Step 2.3 Renew the current solution $s'$ employing a perturbation mechanism.

if Acceptance Criterion( $\tilde{s}$ ) is satisfied then

$s' \leftarrow$  Solution Perturbation( $\tilde{s}$ )

else

$s' \leftarrow$  Solution Perturbation( $s^*$ )

end if

end while

##### Step 3. Return the best feasible solution $s^*$ .

---

### 3.2 Initial Solution Construction

In ITS (Cordeau and Maischberger 2012), an initial solution with  $m^*+1$  routes was constructed for a benchmark problem, where  $m^*$  is the number of routes in the best known solution of the benchmark problem. For SVRPSTW, no benchmark problem exists and  $m^*$  is unavailable. Alternatively, Solomon's Insertion heuristics (Solomon 1987) is used to construct an initial solution for SVRP-ITS. Both the two criteria  $c_1(i,u,j)$  and  $c_2(i,u,j)$  in



Solomon (1987) are set to  $c(s)$  in SVRP-ITS. The resulting initial solution is feasible and the number of routes in the initial solution is an upper bound of  $m^*$ .

### 3.3 Tabu Search

The tabu search heuristics employed in SVRP-ITS is summarized in the Procedure Tabu Search. The input solution  $s'$  (may be infeasible) is inherited from Step 2.1 in Algorithm SVRP-ITS. At the beginning of each tabu search iteration, route reduction procedure (Section 3.3.6) is performed on the current solution  $s$ . If the procedure fails, the best non-tabu neighborhood solution  $\bar{s}$  in  $\mathbf{N}(s)$  (Section 3.3.1) is selected to renew  $s$ , considering tabu tenure and aspiration criterion (Section 3.3.4).  $\bar{s}$  may deteriorate compared with  $s$ . Let  $\zeta_1$  be the number of consecutive iterations in which  $\bar{s}$  has not been improved. The tabu search stops if the best feasible solution  $\tilde{s}$  has not been improved for  $\eta_1$  consecutive iterations.

---

**Procedure** Tabu search( $s'$ )

**Step 1.** Initialization.

```

 $\tilde{s} \leftarrow s', s \leftarrow s'$ 
if  $\tilde{s}$  is infeasible then
     $c(\tilde{s}) \leftarrow +\infty$ 
end

```

Initialize parameters and cost records.

**Step 2.** Tabu search iterations.

```

while  $\zeta_1$  is smaller than  $\eta_1$  do

```

**Step 2.1** Perform route reduction procedure.

```

 $s \leftarrow$  Route reduction( $s$ )
if route reduction succeeds then
     $\tilde{s} \leftarrow s$ , return  $\tilde{s}$ 
end if

```

**Step 2.2** Perform neighborhood search.

```

for each solution  $\bar{s}$  in neighborhood  $N(s)$  do
     $f(\bar{s}) \leftarrow$  Solution evaluation( $\bar{s}$ )
     $p(\bar{s}) \leftarrow$  Diversification( $\bar{s}$ )
     $g(\bar{s}) \leftarrow f(\bar{s}) + p(\bar{s})$ 
end for

```

Select one solution  $\bar{s}$  that minimizes  $g(\bar{s})$  and is not tabu or satisfies aspiration criterion.

**Step 2.3** Update the best feasible solution  $\tilde{s}$  found so far.

```

if  $\bar{s}$  is feasible and  $c(\bar{s}) < c(\tilde{s})$  then
     $\tilde{s} \leftarrow \bar{s}$ 
end if
 $s \leftarrow \bar{s}$ 

```

**Step 2.4** Update parameters and cost records.

```

end while

```

**Step 3.** Return  $\tilde{s}$ .

---

### 3.3.1 Neighborhood Structure

A neighborhood solution of  $s$  is constructed by removing a customer from its original route and then inserting it into another route. This process is performed on every customer in  $s$  and the neighborhood  $\mathbf{N}(s)$  is then obtained. When a customer is removed, its original route is reconstructed by connecting the predecessor and successor vertices of the customer. The customer is inserted into another route with minimum increase in the value of  $f(s)$  (Section 3.3.2).

The definition of attribute was used in Cordeau et al. (2001). If customer  $i$  is assigned to vehicle  $k$  in solution  $s$ , then  $s$  is said to have an attribute  $(i, k)$  and all the attributes of  $s$  form the attribute set  $\mathbf{B}(s)$ . The generation of a neighborhood solution of  $s$  is just the substitution of one attribute in  $\mathbf{B}(s)$  with a new attribute.

### 3.3.2 Solution Evaluation

To diversify the search, capacity, route duration and customer service level constraints are relaxed during the tabu search, with their violations penalized in the objective function. The augmented objective function is  $f(s) = c(s) + \gamma_1 q(s) + \gamma_2 d(s) + \gamma_3 w(s)$ , where  $q(s)$ ,  $d(s)$  and  $w(s)$  are violations of capacity, route duration and customer service level constraints respectively.  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$  are three parameters adjusted dynamically. These parameters are initially set to 1 and any one of them is divided by  $1 + \delta$  if the corresponding constraint is satisfied, otherwise it is multiplied by  $1 + \delta$ .  $\delta$  is set to 0.5 in SVRP-ITS. Different from Cordeau et al. (2001),  $\gamma_1$ ,  $\gamma_2$  and  $\gamma_3$  are restricted in the interval  $[10^{-4}, 10^4]$  in SVRP-ITS. A similar restriction was applied by Brandão (2004).

### 3.3.3 Diversification

Consider a move that substitutes attribute  $(i, k')$  with  $(i, k)$  in  $\mathbf{B}(s)$ . To direct the search into less explored search space, if this move generates a non-improving neighborhood solution  $\bar{s}$  with respect to  $s$ , then  $\bar{s}$  is penalized by a term proportional to the frequency ( $\rho_{ik}$ ) that attribute  $(i, k)$  is added into the current solution in the tabu search process. The penalty term is  $p(\bar{s}) = 0.015c(\bar{s})\sqrt{nm}\rho_{ik}$ . The new evaluation function for  $\bar{s}$  is  $g(\bar{s}) = f(\bar{s}) + p(\bar{s})$ ,  $p(\bar{s}) = 0$  if  $f(\bar{s}) < f(s)$ .

### 3.3.4 Tabu Tenure and Aspiration Criterion

To prevent cycling, moves that lead to previously encountered solutions are forbidden for a number of iterations. For example, if attribute  $(i, k')$  is replaced with  $(i, k)$  in a particular move and the resulting solution  $\bar{s}$  is selected to renew  $s$  for the next iteration,  $(i, k')$  is then assigned a tabu status and can not be added to the current solution for the next  $\theta$  iterations.  $\theta$  is set to  $\lceil 7.5 \log_{10} n \rceil$ , where  $\lceil x \rceil$  is the integer nearest to  $x$ . In following iterations, if a move adding attribute  $(i, k')$  to  $\mathbf{B}(s)$  leads to a better solution than the best solution so far identified having attribute  $(i, k')$ , the move is then allowed to be performed even if the tabu status of attribute  $(i, k')$  still exists.

In SVRP-ITS, the aspiration criterion is further refined to account for both feasible and infeasible solutions. During the tabu search process, both the best feasible and infeasible solutions identified having attribute  $(i, k')$  are recorded. If a move adding attribute  $(i, k')$  results in a better feasible or infeasible solution, the tabu status of attribute  $(i, k')$  can be then overridden.

### 3.3.5 Cost Records

In SVRP-ITS, cost records are used for storing the costs of removing each customer from its original route and inserting it into every possible position in another route. These costs can be used in both neighborhood search and route reduction processes. At the end of each tabu search iteration, only two routes in the current solution  $s$  will be changed. Therefore, only the costs related to the two changed routes need to be updated and the remainder can still be used in the next iteration. This mechanism aims to avoid possible repetitive computing and reduce the computational effort of SVRP-ITS.

Suppose that there are  $m$  routes in a particular solution and  $n$  customers in total. On average there will be  $n/m$  customers per route. If no mechanism such as cost records exists, then approximately  $n \cdot (n/m) \cdot (m-1)$  trial moves need to be evaluated in each tabu search iteration. If a mechanism such as cost records is adopted, about  $n \cdot (n/m) \cdot 2$  trial moves then need to be evaluated in each iteration. When  $m$  is larger than 3, a reduction of about  $n \cdot (n/m) \cdot (m-3)$  in computing occurs. For SVRPSTW with customer service level constraints in this paper, it is normal for  $m$  to be larger than 3. A similar mechanism can be found in Brandão (2004) and the author reported a reduction of at least 30 % of the computational effort using this mechanism.

### 3.3.6 Route Reduction

The original ITS does not have a direct mechanism to minimize the number of routes (Cordeau and Maischberger 2012). The performance of ITS in solving VRPTW was tested on Solomon's benchmark problems (Solomon 1987), for which best known solutions exist in the literature. The number of routes  $m^*$  in these best known solutions was utilized in ITS for solving the benchmark problems. With this useful information of  $m^*$ , the authors provided an indirect route reduction mechanism in ITS. Details can be found in Cordeau and Maischberger (2012).

For SVRPSTW, no benchmark problem exists in the literature and the information of  $m^*$  is unavailable. The number of routes in an initial solution of SVRP-ITS is an upper bound of the optimal value  $m^*$  (Section 3.2). To reduce the number of routes to  $m^*$  in a SVRPSTW solution, a direct route reduction mechanism for SVRP-ITS is developed in this section. With the use of cost records, the mechanism checks every route in each tabu search iteration to find if all customers on that route can be removed and feasibly inserted into other routes. The mechanism is briefly described below and summarized in the Procedure Route Reduction.

The procedure begins with checking the feasibility of the input solution  $s$ , as an infeasible solution has little possibility of becoming a feasible one with fewer routes. The procedure then examines whether there is such a route that all customers on that

route can be feasibly inserted into other routes. If such a route exists, then ways of inserting the customers are investigated and the one with the least cost is selected. Owing to cost records, this procedure does not need much computational effort. In most cases the procedure stops at an early stage because either, the input solution is found not feasible or the procedure fails to find a route that can be decomposed.

### 3.4 Solution Perturbation

In Step 2.3 of Algorithm SVRP-ITS,  $s'$  is perturbed for a tabu search restart. More specifically, firstly a seed customer is selected at random. The seed customer and its  $\pi$  closest customers are then removed and reinserted into  $s'$  with minimum cost in a random order, where  $\pi \sim U(0, \sqrt{n})$ . This cluster removal heuristics has the potential to prevent the removed customers from being reinserted into their original routes (Pisinger and Ropke 2007).

---

#### Procedure Route Reduction( $s$ )

**if**  $s$  is infeasible **then**

↑ **return**  $s$

**end if**

$\hat{s} \leftarrow s$ ,  $c(\hat{s}) \leftarrow +\infty$

**for**  $k=1:m$  **do**

↑ **if** all customers in route  $k$  can be feasibly inserted into other routes **then**

↑ Enumerate all insertion possibilities in terms of target route.

↑ **for** each target route possibility **do**

↑ Divide the customers in route  $k$  into groups so that customers in the same group will be inserted into the same target route.

↑ **for** each group **do**

↑ Initialize the base route set with the target route.

↑ **for** each customer in the group **do**

↑ Insert the customer into every possible position of each route in the base route set.

↑ **if** no feasible route exists after the trial insertions **then**

↑ The current target route possibility fails. Move to the next target route possibility.

↑ **else**

↑ Replace the routes in the base route set with the resulting feasible routes.

↑ **end if**

↑ **end for**

↑ Select the route with minimum cost in the base route set. Record the insert positions of the customers in the route.

↑ **end for**

↑ Construct solution  $\check{s}$  according to the insert position record.

↑ **if**  $c(\check{s}) < c(\hat{s})$  **then**

↑  $\hat{s} \leftarrow \check{s}$ ,  $c(\hat{s}) \leftarrow c(\check{s})$

↑ **end if**

↑ **end for**

↑ **end if**

**end for**

$s \leftarrow \hat{s}$ , **return**  $s$ .

---

### 3.5 Acceptance Criterion

When performing solution perturbation, the improved solution  $\tilde{s}$  has a  $1-(\zeta/\eta)^2$  probability of being the target to be perturbed. A second choice is the best feasible solution so far  $s^*$ . At the early stage of SVRP-ITS,  $\tilde{s}$  is very likely to be selected and perturbed, aiming at exploring a broad solution space. As SVRP-ITS approaches the end, it will revert to  $s^*$ .

## 4 Arrival Time Distribution Approximation

When evaluating a SVRPSTW solution  $s$ , (e.g. calculating  $c(s)$  and  $f(s)$  in Section 3.3.2), arrival time distributions of vehicles need to be computed first. Consider route  $\mathbf{R}_k$  for vehicle  $k$ , in the soft time window case, the arrival time  $A_{r_{j+1}k}$  of vehicle  $k$  at customer  $r_{j+1}$ 's location can be recursively computed by Eqs. (14) to (17).

$$A_{r_1k} = d_{0k} + T_{r_0r_1} \tag{14}$$

$$T_{r_jk}^{ss} = \max\{A_{r_jk}, e_{r_j}\}, \quad 1 \leq j \leq n_k \tag{15}$$

$$D_{r_jk} = T_{r_jk}^{ss} + S_{r_j}, \quad 1 \leq j \leq n_k \tag{16}$$

$$A_{r_{j+1}k} = D_{r_jk} + T_{r_jr_{j+1}}, \quad 1 \leq j \leq n_k \tag{17}$$

For VRPTW,  $A_{r_{j+1}k}$  can be predicted deterministically. For SVRPSTW, determining the distribution of  $A_{r_{j+1}k}$  is difficult in the presence of time windows. In Chang et al. (2009), travel and arrival times of vehicles were assumed to be normally distributed. The mean and variance of departure time  $D_{r_jk}$  were estimated based on the distribution of  $\max\{A_{r_jk}, e_{r_j}\} + S_{r_j}$ . In Thompson et al. (2011), assumptions of normal distribution were also made. The mean and variance of arrival time distributions were estimated based on truncated normal distributions. Travel time distributions, however, are far from normal in some real-life cases (van Lint et al. 2008; Fosgerau and Karlström 2010). Methods based on normal assumptions may lead to large estimation errors in those particular cases. Even if travel times do follow normal distribution, arrival time distributions are not normal in the presence of time windows (Thompson et al. 2011). In the remainder of this section, the approximation method,  $\alpha$ -discrete is presented as a new method for estimating vehicle arrival time distributions in the presence of time windows.

Take route  $\mathbf{R}_k$  for example. For travel time  $T_{r_jr_{j+1}}$ , a set of  $L$  discrete points  $0.5\epsilon, 1.5\epsilon, \dots, L\epsilon - 0.5\epsilon$  in the space of cumulative probability  $[0, 1]$  is considered, where  $\epsilon = 1/L$ . Correspondingly, a sequence of discrete travel times  $b_\tau^{T_{r_jr_{j+1}}}$ ,  $\tau = 1, \dots, L$  can be generated by Eq. (18), where  $F^{-1}(\cdot)$  is the inverse cumulative distribution function (CDF) of  $T_{r_jr_{j+1}}$ . The resulting approximated inverse CDF of  $T_{r_jr_{j+1}}$  is shown in Eq. (19).

Distribution of service time  $S_{r_j}$  can be estimated in the same way as  $T_{r_j r_{j+1}}$ .

$$b_{\tau}^{T_{r_j r_{j+1}}} = F^{-1}\left(\tau\varepsilon - \frac{\varepsilon}{2}\right), \quad \tau = 1, \dots, L \tag{18}$$

$$v^{T_{r_j r_{j+1}}}(\tau\varepsilon) = b_{\tau}^{T_{r_j r_{j+1}}}, \quad \tau = 1, \dots, L \tag{19}$$

Given departure time  $d_{0k}$  and distribution of  $T_{r_0 r_1}$ , distribution of arrival time  $A_{r_1 k}$  in Eq. (14) can be computed as shown in Eq. (20).

$$v^{A_{r_1 k}}(\tau\varepsilon) = b_{\tau}^{T_{r_0 r_1}} + d_{0k}, \quad \tau = 1, \dots, L \tag{20}$$

In Eq. (15), given  $e_{r_j}$  and distribution of  $A_{r_j k}$ , distribution of  $T_{r_j k}^{SS}$  can be estimated using Eqs. (21) to (23).

$$v^{T_{r_j k}^{SS}}(\tau\varepsilon) = e_{r_j}, \quad \tau = 1, \dots, L, \text{ if } e_{r_j} \geq b_L^{A_{r_j k}} \tag{21}$$

$$v^{T_{r_j k}^{SS}}(\tau\varepsilon) = \begin{cases} e_{r_j}, & \tau = 1, \dots, \tau_{ej} \\ b_{\tau}^{A_{r_j k}}, & \tau = \tau_{ej} + 1, \dots, L, \end{cases} \text{ if } b_{\tau_{ej}}^{A_{r_j k}} \leq e_{r_j} < b_{\tau_{ej}+1}^{A_{r_j k}}, \quad 1 \leq \tau_{ej} \leq L-1 \tag{22}$$

$$v^{T_{r_j k}^{SS}}(\tau\varepsilon) = b_{\tau}^{A_{r_j k}}, \quad \tau = 1, \dots, L, \text{ if } e_{r_j} < b_L^{A_{r_j k}} \tag{23}$$

In Eq. (16), based on distributions of  $T_{r_j k}^{SS}$  and  $S_{r_j}$ , a set of  $L^2$  discrete values  $\bar{b}_{\tau}$ ,  $\tau = 1, \dots, L^2$  can be generated using Eq. (24). Another set of discrete values  $b_{\tau}$ ,  $\tau = 1, \dots, L^2$  can be obtained by sorting the elements in  $\bar{b}_{\tau}$ ,  $\tau = 1, \dots, L^2$  in an ascending order. Distribution of  $D_{r_j k}$  is then computed as shown in Eq. (25). Given distributions of  $D_{r_j k}$  and  $T_{r_j r_{j+1}}$ , distribution of  $A_{r_{j+1} k}$  in Eq. (17) can be computed in the same way as  $D_{r_j k}$ .

$$\bar{b}_{\tau} = v^{T_{r_j k}^{SS}}(\tau_1\varepsilon) + v^{S_{r_j}}(\tau_2\varepsilon), \quad \tau_1 = 1, \dots, L, \quad \tau_2 = 1, \dots, L, \quad \tau = \tau_1\tau_2 \tag{24}$$

$$v^{D_{r_j k}}(\tau_1\varepsilon) \leftarrow \frac{1}{L} \sum_{\tau_2=1}^L b_{(\tau_1-1)L+\tau_2}, \quad \tau_1 = 1, \dots, L \tag{25}$$

Expected values  $E(W_{r_j k})$ ,  $E(P_{r_j k})$  and  $E(Z_k)$  in Eq. (1) can be estimated as shown in Eqs. (26), (27) and (28) respectively, given the distribution of  $A_{r_j k}$ .

$$E(W_{r_j k}) = \begin{cases} 0, & \text{if } e_{r_j} < b_1^{A_{r_j k}} \\ \sum_{\tau=1}^{\tau_{ej}} \varepsilon (e_{r_j} - b_{\tau}^{A_{r_j k}}), & \text{if } b_{\tau_{ej}}^{A_{r_j k}} \leq e_{r_j} < b_{\tau_{ej}+1}^{A_{r_j k}}, \quad 1 \leq \tau_{ej} \leq L-1 \\ \sum_{\tau=1}^L \varepsilon (e_{r_j} - b_{\tau}^{A_{r_j k}}), & \text{if } e_{r_j} \geq b_L^{A_{r_j k}} \end{cases} \tag{26}$$

$$E(P_{r_jk}) = \begin{cases} \sum_{\tau=1}^L \varepsilon (b_{\tau}^{A_{r_jk}} - l_{r_j}), & \text{if } l_{r_j} < b_1^{A_{r_jk}} \\ \sum_{\tau=\tau_{lj}+1}^L \varepsilon (b_{\tau}^{A_{r_jk}} - l_{r_j}), & \text{if } b_{\tau_{lj}}^{A_{r_jk}} \leq l_{r_j} < b_{\tau_{lj}+1}^{A_{r_jk}}, \quad 1 \leq \tau_{lj} \leq L-1 \\ 0, & \text{if } l_{r_j} \geq b_L^{A_{r_jk}} \end{cases} \quad (27)$$

$$E(Z_k) = \begin{cases} \sum_{\tau=1}^L \varepsilon (b_{\tau}^{A_{r_jk}} - d_{0k} - h), & \text{if } d_{0k} + h < b_1^{A_{r_jk}}, \quad j = n_k + 1 \\ \sum_{\tau=\tau_h+1}^L \varepsilon (b_{\tau}^{A_{r_jk}} - d_{0k} - h), & \text{if } b_{\tau_h}^{A_{r_jk}} \leq d_{0k} + h < b_{\tau_h+1}^{A_{r_jk}}, \quad 1 \leq \tau_h \leq L-1, \quad j = n_k + 1 \\ 0, & \text{if } d_{0k} + h \geq b_L^{A_{r_jk}}, \quad j = n_k + 1 \end{cases} \quad (28)$$

The probabilities in Eqs. (7) and (8) can also be estimated, as shown in Eqs. (29) and (30) respectively.

$$P\{A_{r_jk} \leq l_{r_j}\} = \begin{cases} 0, & \text{if } l_{r_j} < b_1^{A_{r_jk}} \\ \tau_{lj}\varepsilon, & \text{if } b_{\tau_{lj}}^{A_{r_jk}} \leq l_{r_j} < b_{\tau_{lj}+1}^{A_{r_jk}}, \quad 1 \leq \tau_{lj} \leq L-1 \\ 1, & \text{if } l_{r_j} \geq b_L^{A_{r_jk}} \end{cases} \quad (29)$$

$$P\{B_k \leq h\} = \begin{cases} 0, & \text{if } d_{0k} + h < b_1^{A_{r_jk}}, \quad j = n_k + 1 \\ \tau_h\varepsilon, & \text{if } b_{\tau_h}^{A_{r_jk}} \leq d_{0k} + h < b_{\tau_h+1}^{A_{r_jk}}, \quad 1 \leq \tau_h \leq L-1, \quad j = n_k + 1 \\ 1, & \text{if } d_{0k} + h \geq b_L^{A_{r_jk}}, \quad j = n_k + 1 \end{cases} \quad (30)$$

Expressions for estimating violations of route duration and service level constraints (see Section 3.3.2) in solution  $s$  are shown in Eqs. (31) and (32) respectively.

$$d(s) = \sum_{k \in \mathbf{K}} \max\{v^{A_{r_jk}}([\beta/\varepsilon] \cdot \varepsilon) - d_{0k} - h, 0\}, \quad r_j \in \mathbf{R}_k, \quad j = n_k + 1 \quad (31)$$

$$w(s) = \sum_{k \in \mathbf{K}} \sum_{j=1}^{n_k+1} \max\{v^{A_{r_jk}}([\alpha_{r_j}/\varepsilon] \cdot \varepsilon) - l_{r_j}, 0\}, \quad r_j \in \mathbf{R}_k \quad (32)$$

## 5 Computational Results

### 5.1 Accuracy of the $\alpha$ -Discrete Approximation Method

In this section, the accuracy of the proposed  $\alpha$ -discrete approximation method was tested by comparison with Chang’s method (Chang et al. 2009). Stochastic simulation (Li et al. 2010) provided the ground true value for the comparison. The test example

is a simple partial route shown in Fig. 2, where  $D$  denotes the depot and  $C_i$  represents the customer,  $i=1, \dots, 5$ . In Fig. 2, numbers within square brackets are customers' soft time windows, whereas numbers within round brackets represent the mean and variance of link travel times. Departure time from  $D$  was set to zero.

For the proposed  $\alpha$ -discrete approximation method,  $L$  was set to 100. For stochastic simulation, a total of  $10^6$  iterations were performed. Since arrival times were assumed to be normally distributed in Chang's method (Chang et al. 2009), expressions for computing  $E(W_{r_j,k})$ ,  $E(P_{r_j,k})$  and customer service levels under the assumption of normal arrival time distribution are given below,

$$E(W_{r_j,k}) = \sqrt{\text{Var}(A_{r_j,k})}[\phi(z_1) + z_1\Phi(z_1)], \quad r_j \in \mathbf{R}_k, \quad 1 \leq j \leq n_k \quad (33)$$

$$E(P_{r_j,k}) = \sqrt{\text{Var}(A_{r_j,k})}\{\phi(z_2) - z_2[1 - \Phi(z_2)]\}, \quad r_j \in \mathbf{R}_k, \quad 1 \leq j \leq n_k + 1 \quad (34)$$

$$P\{A_{r_j,k} \leq l_{r_j}\} = \Phi(z_2), \quad r_j \in \mathbf{R}_k, \quad 1 \leq j \leq n_k + 1 \quad (35)$$

where  $z_1 = (e_{r_j} - E(A_{r_j,k})) / \sqrt{\text{Var}(A_{r_j,k})}$ ;  $z_2 = (l_{r_j} - E(A_{r_j,k})) / \sqrt{\text{Var}(A_{r_j,k})}$ ;  $\phi(\cdot)$  is the probability density function of standard normal distribution;  $\Phi(\cdot)$  is the cumulative distribution function of standard normal distribution.

The test was conducted in two scenarios. Link travel times in Fig. 2 were assumed to follow normal distribution in Scenario 1 and lognormal distribution in Scenario 2. Given mean and variance of link travel times, parameters for the lognormal distribution of link travel times can be computed using Eqs. (36) and (37). In both scenarios, service times at customers were assumed to follow the same normal distribution  $N(10, 5^2)$ . The test results are shown in Tables 1 to 4.

$$\mu(T_{r_j,r_{j+1}}) = \log\left(E(T_{r_j,r_{j+1}})^2 / \sqrt{\text{Var}(T_{r_j,r_{j+1}}) + E(T_{r_j,r_{j+1}})^2}\right), \quad r_j \in \mathbf{R}_k, \quad 0 \leq j \leq n_k \quad (36)$$

$$\sigma(T_{r_j,r_{j+1}}) = \sqrt{\log\left(\text{Var}(T_{r_j,r_{j+1}}) / E(T_{r_j,r_{j+1}})^2 + 1\right)}, \quad r_j \in \mathbf{R}_k, \quad 0 \leq j \leq n_k \quad (37)$$

In Scenario 1, it can be seen from Tables 1 to 3 that Chang's method performed well at the first customer  $C_1$ , but not at the rest of the customers. This result is expected, since the arrival time at  $C_1$  is exactly normally distributed in Scenario 1, but arrival times at customers  $C_2$  to  $C_5$  are not, due to the impact of time windows. In Scenario 2, the estimation error is even larger than that in Scenario 1 by Chang's method. For example, the relative error between the estimated expected tardiness at  $C_3$  by Chang's method and

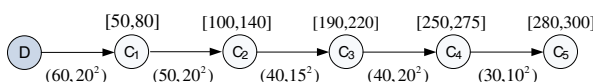


Fig. 2 A simple partial route



**Table 1** Estimated service level at each customer by different methods

Type of travel time distribution	Approximation method	Estimated service level (%) at each customer (relative error %)				
		$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
Normal (Scenario 1)	Stochastic simulation	84.07	74.31	93.42	89.10	63.03
	Chang's method	84.13( <b>0.07</b> )	73.60( <b>-0.96</b> )	94.79(1.47)	89.85(0.84)	56.92(-9.69)
	$\alpha$ -discrete	83.50(-0.68)	73.50(-1.10)	92.50( <b>-0.99</b> )	88.50( <b>-0.67</b> )	62.50( <b>-0.84</b> )
Lognormal (Scenario 2)	Stochastic simulation	85.26	77.04	92.24	87.69	65.73
	Chang's method	84.13(-1.32)	73.60(-4.47)	94.79(2.77)	89.85(2.47)	56.92(-13.40)
	$\alpha$ -discrete	84.50( <b>-0.89</b> )	76.50( <b>-0.70</b> )	91.50( <b>-0.81</b> )	87.50( <b>-0.21</b> )	65.50( <b>-0.35</b> )

A bolded number represents the smallest one in relevant rows of a particular column

that by stochastic simulation is -37.09 % in Scenario 1, while it is -63.52 % in Scenario 2 (Table 3). The reason is clear that travel and arrival times are not normally distributed in Scenario 2. This, however, is a basic assumption of Chang's method.

For the proposed  $\alpha$ -discrete approximation method, it was found that larger errors were produced when estimating expected tardiness. The largest relative error between the results of  $\alpha$ -discrete and stochastic simulation is -1.44 % in Tables 1 and 2, while it is -9.89 % in Table 3. However, the estimation error by  $\alpha$ -discrete is still much smaller compared with that of Chang's method, as shown in Tables 1 to 3. In addition, the accuracy of  $\alpha$ -discrete can be further improved by increasing the value of  $L$  (e.g. 200).

Table 4 shows that the difference between the estimation error by Chang's method and that by  $\alpha$ -discrete is small in terms of the total cost of the partial route (-0.52 %

**Table 2** Estimated expected earliness at each customer by different methods

Type of travel time distribution	Approximation method	Estimated expected earliness at each customer (relative error %)				
		$C_1$	$C_2$	$C_3$	$C_4$	$C_5$
Normal (Scenario 1)	Stochastic simulation	3.95	2.14	19.63	11.82	0.76
	Chang's method	3.96( <b>0.11</b> )	2.36(10.57)	18.92(-3.62)	11.90(0.67)	1.25(64.98)
	$\alpha$ -discrete	3.94(-0.42)	2.11( <b>-1.22</b> )	19.61( <b>-0.12</b> )	11.81( <b>-0.09</b> )	0.75( <b>-1.19</b> )
Lognormal (Scenario 2)	Stochastic simulation	3.17	1.61	21.43	12.00	0.56
	Chang's method	3.96(24.62)	2.36(46.30)	18.92(-11.70)	11.90(-0.83)	1.25(125.57)
	$\alpha$ -discrete	3.16( <b>-0.44</b> )	1.61( <b>-0.03</b> )	21.46( <b>0.14</b> )	11.98( <b>-0.15</b> )	0.55( <b>-1.44</b> )

A bolded number represents the smallest one in relevant rows of a particular column

**Table 3** Estimated expected tardiness at each customer by different methods

Type of travel time distribution	Approximation method	Estimated expected tardiness at each customer (relative error %)				
		C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>
Normal (Scenario 1)	Stochastic simulation	1.67	4.28	0.94	1.65	5.64
	Chang's method	1.67( <b>0</b> )	4.08(-4.68)	0.59(-37.09)	1.13(-31.92)	5.21(-7.52)
	$\alpha$ -discrete	1.65(-1.51)	4.22(-1.37)	0.88(-6.07)	1.58(-4.59)	5.55(-1.61)
Lognormal (Scenario 2)	Stochastic simulation	2.29	4.91	1.61	2.83	6.82
	Chang's method	1.67(-27.31)	4.08(-16.95)	0.59(-63.52)	1.13(-60.20)	5.21(-23.53)
	$\alpha$ -discrete	2.25(-2.04)	4.78(-2.56)	1.45(-9.89)	2.59(-8.48)	6.53(-4.30)

A bolded number represents the smallest one in relevant rows of a particular column

versus -0.14 % in Scenario 1 and -2.23 % versus -0.42 % in Scenario 2). This is because penalty cost takes up only a small part of the total cost in this test example and mean travel times are the same for both methods. Even though this difference can be ignored, however, the difference in estimating service levels at each customer by these two methods is sometimes large (Table 1) and thus can not be neglected.

In conclusion, the accuracy of the proposed  $\alpha$ -discrete approximation method is higher than that of Chang's method, especially when travel times are not normally distributed. Thus  $\alpha$ -discrete is selected as the approximation method for solution evaluation and used in the remainder of this paper.

**Table 4** Estimated cost of the partial route by different methods

Type of travel time distribution	Approximation method	Estimated cost of the partial route (relative error %)				
		No. of vehicles used	Mean travel time	Penalty cost		Sum of mean travel time and penalty cost
				Earliness penalty	Tardiness penalty	
Normal (Scenario 1)	Stochastic simulation	1	220	38.30	14.18	272.48
	Chang's method	1	220	38.39(0.23)	12.67(-10.65)	271.06(-0.52)
	$\alpha$ -discrete	1	220	38.22(-0.21)	13.88(-2.12)	272.10(-0.14)
Lognormal (Scenario 2)	Stochastic simulation	1	220	38.77	18.46	277.23
	Chang's method	1	220	38.39(-0.98)	12.67(-31.37)	271.06(-2.23)
	$\alpha$ -discrete	1	220	38.76( <b>0</b> )	17.60(-4.66)	276.36(-0.42)

A bolded number represents the smallest one in relevant rows of a particular column

## 5.2 Performance of the Proposed Model for SVRPSTW

In this section, several of Solomon's benchmark problems (Solomon 1987) were adapted as test instances to demonstrate the performance of the proposed model. In the remainder of this section, the test dataset and experiment settings are introduced first, followed by the computational results of the proposed model on this dataset.

### 5.2.1 Test Dataset and Experiment Settings

The well-known Solomon's benchmark problems (Solomon 1987) have been chosen as test datasets by many previous studies on VRPTW (Taillard et al. 1997; Cordeau et al. 2001; Chiang and Russell 2004; Nagata et al. 2009). Several factors were considered when these benchmark problems were generated, such as geographical locations of customers and tightness of time windows. In this study, seven of the benchmark problems were chosen and adapted for the demonstration of the proposed model. Their major characteristics are listed in Table 5.

Three types of problems according to the geographical locations of customers are shown in Table 5. For each problem type, two or three instances were chosen with different time window width sizes. For each problem instance, there are a total of 20 customers and the sum of customer demands is listed in Table 5. Vehicle capacity is 200 units in each problem instance. Travel times were assumed to follow lognormal distribution and service times were assumed to be normally distributed. The COV of the travel times was randomly generated from [0.2, 0.6]. Given the mean and variance of travel times, parameters for the lognormal distribution of travel times can be computed using Eqs. (36) and (37). This dataset is chosen to demonstrate the applicability of the proposed model on problem instances with different characteristics. Additional data (e.g. coordinates and time windows) of the dataset can be found on <http://web.cba.neu.edu/~msolomon/problems.htm>.

The proposed solution algorithm SVRP-ITS was implemented in Matlab 7.5.0 and run on a PC with a four-core Inter Core i7 3.40 GHz CPU (only one core was used) and 4 GB RAM. The maximum number of tabu search iterations  $\eta$  was set to 2000. In local search improvement phases, tabu search stops if the best feasible solution has not been improved for 200 consecutive iterations. For the proposed  $\alpha$ -discrete approximation method,  $L$  was set to 100.

Each problem instance was tested in four scenarios, with parameter settings of the proposed model in these scenarios shown in Table 6. In the last scenario, customers were divided into two groups: the first 6 customers being the first group and the rest 14 the second group. It was assumed that customers in the first group require higher service levels than those in the second group.

### 5.2.2 Computational Results of the Proposed Model on a Typical Problem Instance

In this section, one of the problem instances RC101.20 in Table 5 was selected as a typical example to demonstrate the performance of the proposed model. RC101.20 was considered typical because of the mixed (randomly distributed

**Table 5** Characteristics of the test dataset

Problem	Geographical locations of customers	Time window width	Sum of customer demands	Travel time		Service time			
				Distribution	Mean	s.d.	Distribution	Mean	s.d.
<b>R105.20</b>	Random	30	265	Lognormal	Dist. <sup>a</sup>	0.2~0.6	Normal	12	0.4
<b>R109.20</b>	Random	37~83	265	Lognormal	Dist.	0.2~0.6	Normal	12	0.4
<b>C101.20</b>	Clustered	37~89	360	Lognormal	Dist.	0.2~0.6	Normal	90	0.4
<b>C106.20</b>	Clustered	75~177	360	Lognormal	Dist.	0.2~0.6	Normal	90	0.4
<b>RC101.20</b>	Rand./Clus. <sup>b</sup>	30	430	Lognormal	Dist.	0.2~0.6	Normal	12	0.4
<b>RC106.20</b>	Rand./Clus.	60	430	Lognormal	Dist.	0.2~0.6	Normal	12	0.4
<b>RC107.20</b>	Rand./Clus.	41~155	430	Lognormal	Dist.	0.2~0.6	Normal	12	0.4

A bolded number represents the smallest one in relevant rows of a particular column

<sup>a</sup> With vehicle velocity set to 1, the value of mean travel time is equal to the value of the distance traveled;

<sup>b</sup> Geographical locations of customers in this problem type are randomly distributed and/or clustered

**Table 6** Parameter settings of the proposed model for each problem instance in different scenarios

Scenario	Parameter settings of the proposed model								
	$\lambda_{1,i}, i \in \mathbf{V}$	$\lambda_{2,i}, i \in \mathbf{V}$	$\alpha_i, i \in \mathbf{V}$	$\lambda_{1,0}$	$\lambda_{2,0}$	$\lambda_3$	$\alpha_0$	$\beta$	$f$
1	0.5	2	0	0	1	0	0	0	1000
2	0.5	2	0.5	0	1	0	0	0	1000
3	0.5	2	0.8	0	1	0	0	0	1000
4	0.5	4 <sup>a</sup> , 2 <sup>b</sup>	0.8 <sup>a</sup> , 0.5 <sup>b</sup>	0	1	0	0	0	1000

A bolded number represents the smallest one in relevant rows of a particular column

<sup>a</sup> Parameter setting for customers in the first group

<sup>b</sup> Parameter setting for customers in the second group

and/or clustered) geographical locations of customers in that problem and because the time window width (30 min) is common in just-in-time production (Chang et al. 2009). Computational results of the proposed model on RC101.20 are shown in Table 7.

Table 7 shows that the number of vehicles used  $m$  in Scenario 1 is the smallest among all the four scenarios. Because in achieving the primary objective of the proposed model (i.e. minimizing  $m$  while satisfying all the constraints, Section 2), constraints are the weakest in Scenario 1 with no customer service level constraints imposed (Table 6). In such a case, capacity constraint Eq. (6) determines the number of required vehicles in the solution. As the sum of customer demands is 430 units in RC101.20 (Table 5) and the vehicle capacity is 200, at least 3 vehicles are required in the solution, as shown in Table 7. With limited number of vehicles used in Scenario 1, customer service levels cannot be ensured. The tardiness penalty cost is the largest and the mean customer service level is the lowest in this scenario.

**Table 7** Computational results of the proposed model on RC101.20

Scenario	Computational results of the proposed model						
	No. of vehicles used	Mean travel time	Penalty cost		Mean service level <sup>a</sup>	Sum of mean travel time and penalty cost	Total cost <sup>b</sup>
			Earliness penalty	Tardiness penalty			
RC101.20_1	<b>3</b>	<b>338.33</b>	<b>29.72</b>	176.53	<b>69.30%</b>	544.58	<b>3544.58</b>
RC101.20_2	4	429.31	51.05	22.62	93.92 %	<b>502.99</b>	4502.99
RC101.20_3	5	455.75	85.21	<b>6.49</b>	98.15 %	547.45	5547.45
RC101.20_4	4	433.58	47.26	40.93	91.27 %	521.77	4521.77

A bolded number represents the smallest one in relevant rows of a particular column

<sup>a</sup> Mean value of the service levels at the 20 customers in the problem;

<sup>b</sup> Total cost equals to the sum of fixed vehicle employment cost, mean travel time and penalty cost

In Scenario 2, a minimum service level of 50 % was specified for each customer in RC101.20. Consequently, the number of required vehicles increases in this scenario (Table 7). With one more vehicle used, the mean customer service level is greatly improved compared with that in Scenario 1. This result implies that carriers may consider adding one more vehicle to their fleet to improve their service quality. In Scenario 3, required customer service levels were further increased to 80 %, resulting in a high mean customer service level (Table 7). Whereas the number of required vehicles in this scenario is the largest among all the four scenarios.

In Scenario 4, only thirty percent of the customers in RC101.20 (i.e. customers in the first group) were guaranteed a minimum service level of 80 %, while for others (i.e. customers in the second group) the value was 50 %. This differentiation among customers according to their diverse service level preferences leads to a reduction in the number of required vehicles in Scenario 4, compared with that in Scenario 3 (Table 7).

It can be also found from Table 7 that mean travel time increases when more vehicles are used. For example, mean travel time increases by 26.89 % in Scenario 2 compared with that in Scenario 1. Earliness penalty cost also increases in line with the increase in the number of vehicles used. The reason is that, in such a case, on average, fewer customers will be assigned to one vehicle. Additionally, the duration of each route will not necessarily become shorter, due to the existence of time windows. Therefore more time might be spent in waiting at customer locations. Finally, it is shown in Table 7 that the total cost is the highest in Scenario 3, since the number of vehicles used is the largest in that scenario.

### 5.2.3 Computational Results of the Proposed Model on Other Problem Instances

To demonstrate the applicability of the proposed model on different types of problems, the proposed model was solved on the other six problem instances in Table 5 (except RC101.20) in this section. The computational results are displayed in Table 8.

A large increase in the number of vehicles used in the solution (from 2 to 5) is found from Scenario 1 to Scenario 2 for problem R105.20 in Table 8. The trade-off is that the mean customer service level is greatly improved (from 30.65 % in Scenario 1 to 90.40 % in Scenario 2). For Scenarios 2 to 4 of problem R105.20, the number of vehicles used is the same, although the required customer service levels differ in these scenarios (Table 5). In these cases, the trade-off is between the mean travel time and the tardiness penalty cost, with the number of required vehicles kept the same. Similar trends can be found for problem R109.20, except that the increase in the number of vehicles used from Scenario 1 to Scenario 2 is smaller for R109.20 than that for R105.20. The reason is that the time windows specified by customers were tighter in Problem R105.20 than those in R109.20 (Table 5).

For problems C101.20 and C106.20, the total cost is the smallest in Scenario 2 among all four scenarios as shown in Table 8. This is because the cost paid for additional vehicles is compensated by the reduction in the tardiness penalty cost. In Scenario 4 of these two problem instances, a reduction in the number of required vehicles is seen again because of the differentiation among customers according to their diverse service level preferences.

**Table 8** Computational results of the proposed model on the other 6 problem instances

Scenario	Computational results of the proposed model						
	No. of vehicles used	Mean travel time	Penalty cost		Mean service level <sup>a</sup>	Sum of mean travel time and penalty cost	Total cost <sup>b</sup>
			Earliness penalty	Tardiness penalty			
R105.20_1	<b>2</b>	<b>350.78</b>	<b>12.13</b>	1893.61	<b>30.65%</b>	2256.52	<b>4256.52</b>
R105.20_2	5	455.91	84.78	32.62	90.40 %	<b>573.31</b>	5573.31
R105.20_3	5	524.18	76.29	<b>24.80</b>	94.47 %	625.27	5625.27
R105.20_4	5	487.19	79.78	39.44	91.43 %	606.41	5606.41
R109.20_1	<b>2</b>	<b>318.0</b>	<b>10.06</b>	1259.10	<b>48.35%</b>	1587.16	<b>3587.16</b>
R109.20_2	4	392.16	42.01	20.05	95.40 %	<b>454.21</b>	4454.21
R109.20_3	4	428.62	27.46	<b>13.28</b>	96.65 %	469.36	4469.36
R109.20_4	4	394.15	57.63	24.94	93.93 %	476.71	4476.71
C101.20_1	<b>2</b>	<b>160.82</b>	<b>29.45</b>	1612.38	<b>47.0%</b>	1802.64	3802.64
C101.20_2	3	250.18	217.52	82.81	90.85 %	<b>550.51</b>	<b>3550.51</b>
C101.20_3	4	283.04	252.36	<b>49.40</b>	93.80 %	584.80	4584.80
C101.20_4	3	275.17	205.02	108.69	90.55 %	588.88	3588.88
C106.20_1	<b>2</b>	<b>160.82</b>	<b>26.34</b>	1505.31	<b>50.95%</b>	1692.46	3692.46
C106.20_2	3	248.76	194.60	78.02	91.70 %	<b>521.37</b>	<b>3521.37</b>
C106.20_3	4	281.62	230.93	<b>45.17</b>	94.55 %	557.72	4557.72
C106.20_4	3	250.18	201.11	81.56	92.15 %	532.85	3532.85
RC106.20_1	<b>3</b>	<b>329.76</b>	<b>22.33</b>	65.0	<b>91.92%</b>	<b>417.09</b>	<b>3417.09</b>
RC106.20_2	<b>3</b>	<b>329.76</b>	<b>22.33</b>	65.0	<b>91.92%</b>	<b>417.09</b>	<b>3417.09</b>
RC106.20_3	4	387.13	32.47	<b>9.19</b>	98.25 %	428.79	4428.79
RC106.20_4	<b>3</b>	<b>329.76</b>	<b>22.33</b>	66.42	<b>91.92%</b>	418.51	3418.51
RC107.20_1	<b>3</b>	<b>289.67</b>	<b>25.12</b>	36.46	93.75 %	<b>351.25</b>	<b>3351.25</b>
RC107.20_2	<b>3</b>	<b>289.67</b>	<b>25.12</b>	36.46	93.75 %	<b>351.25</b>	<b>3351.25</b>
RC107.20_3	<b>3</b>	353.48	25.22	<b>23.90</b>	96.37 %	402.60	3402.60
RC107.20_4	<b>3</b>	307.59	25.22	45.88	<b>93.22%</b>	378.69	3378.69

A bolded number represents the smallest one in relevant rows of a particular column

<sup>a</sup> Mean value of the service levels at the 20 customers in the problem;

<sup>b</sup> Total cost equals to the sum of fixed vehicle employment cost, mean travel time and penalty cost

For problems RC106.20 and RC107.20, it can be seen from Table 8 that solutions in some of the scenarios are the same. For example, the solution of problem RC106.20 in Scenario 1 is the same as that in Scenario 2. The reason is that for problem RC106.20 at least 3 vehicles were required in Scenario 1 with no customer service level imposed (for the same reason as stated in Section 5.2.2 for problem RC101.20). With this number of vehicles used, customer service levels in Scenario 1

already reached the required level in Scenario 2. Therefore the problem solution for Scenario 2 is the same as that in Scenario 1.

## 6 Conclusions

A stochastic vehicle routing problem with soft time windows considering travel and service time uncertainties is presented in this paper. A new stochastic programming model has been proposed to minimize carrier's total cost while ensuring a minimum on-time arrival probability at each customer location. To solve the proposed model, an iterated tabu search heuristic algorithm was developed. A route reduction mechanism was incorporated in this heuristics to reduce the number of required vehicles. A discrete approximation method has been proposed to estimate the arrival time distributions of vehicles in the presence of time windows.

Computational results showed that the accuracy of the proposed  $\alpha$ -discrete approximation method is higher than that of Chang's method (Chang et al. 2009), especially when travel times are not normally distributed. Although travel times were assumed to be normal or lognormal in the test example, it is noted that  $\alpha$ -discrete can also be applied for other travel time distributions. In addition, several numerical examples were carried out to demonstrate the performance of the proposed SVRPSTW model. Computational results confirmed that the proposed model can be applied on different types of problems. Various trade-offs between carrier's total cost and customer service level can be explored using the proposed model. Results indicated that the proposed model can easily be adapted to multi-class customers with various service level preferences by imposing appropriate customer service level constraints.

The computational efficiencies of the proposed  $\alpha$ -discrete approximation method, Chang's method (Chang et al. 2009) and stochastic simulation (Li et al. 2010) will be investigated in future work. More effort is also required to improve the efficiency of the proposed solution algorithm for solving larger problem instances. In addition, it is noted that VRP has been studied in combination with other problems. For example, location-routing problems (Bozkaya et al. 2010; Toyoglu et al. 2012; Silva and Gao 2012) combine the problems of facility location and vehicle routing. It is interesting to investigate the effects of travel time uncertainties on these problems.

**Acknowledgments** The work described in this paper was jointly supported by a research grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 5196/10E) and a research grant from the National Science Foundation of China (41201466). The first author is sponsored by a Postgraduate Stipend of the Hong Kong Polytechnic University. Thanks are due to two anonymous reviewers for their valuable comments that improved both the content and presentation of the paper.

## References

- Ando N, Taniguchi E (2006) Travel time reliability in vehicle routing and scheduling with time windows. *Netw Spat Econ* 6:293–311
- Balakrishnan N (1993) Simple heuristics for the vehicle routing problem with soft time windows. *J Oper Res Soc* 44:279–287



- Bertsimas DJ, van Ryzin G (1991) A stochastic and dynamic vehicle routing problem in the Euclidean plane. *Oper Res* 39:601–615
- Bozkaya B, Yanik S, Balcişoy S (2010) A GIS-based optimization framework for competitive multi-facility location-routing problem. *Netw Spat Econ* 10:297–320
- Brandão J (2004) A tabu search algorithm for the open vehicle routing problem. *Eur J Oper Res* 157:552–564
- Chang T-S, Wan Y-W, Ooi WT (2009) A stochastic dynamic traveling salesman problem with hard time windows. *Eur J Oper Res* 198:748–759
- Chen BY, Lam WHK, Sumalee A, Shao H (2011) An efficient solution algorithm for solving multi-class reliability-based traffic assignment problem. *Math Comput Model* 54:1428–1439
- Chen BY, Lam WHK, Sumalee A, Li QQ, Shao H, Fang Z (2012) Finding reliable shortest paths in road networks under uncertainty. *Netw Spat Econ*. doi:10.1007/s11067-012-9175-1
- Chiang W-C, Russell RA (2004) A metaheuristic for the vehicle-routing problem with soft time windows. *J Oper Res Soc* 55:1298–1310
- Cordeau J-F, Maischberger M (2012) A parallel iterated tabu search heuristic for vehicle routing problems. *Comput Oper Res* 39:2033–2050
- Cordeau J-F, Laporte G, Mercier A (2001) A unified tabu search heuristic for vehicle routing problems with time windows. *J Oper Res Soc* 52:928–936
- Cordeau J-F, Desaulniers G, Desrosiers J, Solomon MM, Soumis F (2002) VRP with time windows. In: Toth P, Vigo D (eds) *The vehicle routing problem*. SIAM, Philadelphia, pp 157–193
- Dantzig GB, Ramser JH (1959) The truck dispatching problem. *Manage Sci* 6:80–91
- Drexl M (2012) Synchronization in vehicle routing—A survey of VRPs with multiple synchronization constraints. *Transp Sci* 46:297–316
- Escuín D, Millán C, Larrodé E (2012) Modelization of time-dependent urban freight problems by using a multiple number of distribution centers. *Netw Spat Econ* 12:321–336
- Fosgerau M, Karlström A (2010) The value of reliability. *Transp Res B Methodol* 44:38–49
- Gendreau M, Laporte G, Séguin R (1995) An exact algorithm for the vehicle routing problem with stochastic demands and customers. *Transp Sci* 29:143–155
- Golden B, Raghavan S, Wasil E (2008) *The vehicle routing problem: Latest advances and new challenges*. Springer, New York
- Kenyon AS, Morton DP (2003) Stochastic vehicle routing with random travel times. *Transp Sci* 37:69–82
- Koskosis YA, Powell WB, Solomon MM (1992) An optimization-based heuristic for vehicle routing and scheduling with soft time window constraints. *Transp Sci* 26:69–85
- Lam WHK, Shao H, Sumalee A (2008) Modeling impacts of adverse weather conditions on a road network with uncertainties in demand and supply. *Transp Res B Methodol* 42:890–910
- Lambert V, Laporte G, Louveaux F (1993) Designing collection routes through bank branches. *Comput Oper Res* 20:783–791
- Laporte G (2009) Fifty years of vehicle routing. *Transp Sci* 43:408–416
- Laporte G, Louveaux F, Mercure H (1992) The vehicle routing problem with stochastic travel times. *Transp Sci* 26:161–170
- Laporte G, Louveaux F, van Hamme L (2002) An integer L-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Oper Res* 50:415–423
- Lei H, Laporte G, Guo B (2011) The capacitated vehicle routing problem with stochastic demands and time windows. *Comput Oper Res* 38:1775–1783
- Leung SCH, Zhou X, Zhang D, Zheng J (2011) Extended guided tabu search and a new packing algorithm for the two-dimensional loading vehicle routing problem. *Comput Oper Res* 38:205–215
- Li X, Tian P, Leung SCH (2010) Vehicle routing problems with time windows and stochastic travel and service times: models and algorithm. *Int J Prod Econ* 125:137–145
- Li X, Leung SCH, Tian P (2012) A multistart adaptive memory-based tabu search algorithm for the heterogeneous fixed fleet open vehicle routing problem. *Expert Syst Appl* 39:365–374
- Li Z-C, Huang H-J, Lam WHK (2012) Modelling heterogeneous drivers' responses to route guidance and parking information systems in stochastic and time-dependent networks. *Transportmetrica* 8:105–129
- Liberatore F, Righini G, Salani M (2011) A column generation algorithm for the vehicle routing problem with soft time windows. *4OR-Q J. Oper Res* 9:49–82
- Lourenço HR, Martin OC, Stützle T (2002) Iterated local search. In: Glover F, Kochenberger G (ed) *Handbook of metaheuristics*. International series in operations research & management Science, vol. 57. Kluwer Academic Publishers, Norwell, MA, pp 321–353
- Miller-Hooks E, Mahmassani HS (1998) Optimal routing of hazardous materials in stochastic, time-varying transportation networks. *Transp Res Rec* 1645:143–151

- Nagata Y, Bräysy O, Dullaert W (2009) A penalty-based edge assembly memetic algorithm for the vehicle routing problem with time windows. *Comput Oper Res* 37:724–737
- Norouzi N, Tavakkoli-Moghaddam R, Ghazanfari M, Alinaghian M, Salamatbakhsh A (2012) A new multi-objective competitive open vehicle routing problem solved by particle swarm optimization. *Netw Spat Econ* 12:609–633
- Pisinger D, Ropke S (2007) A general heuristic for vehicle routing problems. *Comput Oper Res* 34:2403–2435
- Russell RA, Urban TL (2008) Vehicle routing with soft time windows and Erlang travel times. *J Oper Res Soc* 59:1220–1228
- Savelsbergh MWP (1992) The vehicle routing problem with time windows: minimizing route duration. *INFORMS J Comput* 4:146–154
- Silva F, Gao L (2012) A joint replenishment inventory-location model. *Netw Spat Econ*. doi:10.1007/s11067-012-9174-2
- Solomon MM (1987) Algorithms for the vehicle routing and scheduling problems with time window constraints. *Oper Res* 35:254–265
- Taillard E, Badeau P, Gendreau M, Guertin F, Potvin J-Y (1997) A tabu search heuristic for the vehicle routing problem with soft time windows. *Transp Sci* 31:170–186
- Thompson RG, Taniguchi E, Yamada T (2011) Estimating the benefits of considering travel time variability in urban distribution. *Transp Res Rec* 2238:86–96
- Toth P, Vigo D (2002) *The vehicle routing problem*. SIAM, Philadelphia
- Toyoglu H, Karasan OE, Kara BY (2012) A new formulation approach for location-routing problems. *Netw Spat Econ* 12:635–659
- van Lint JWC, van Zuylen HJ, Tu H (2008) Travel time unreliability on freeways: why measures based on variance tell only half the story. *Transp Res A Policy Pract* 42:258–277
- Wei C, Asakura Y, Iryo T (2012) A probability model and sampling algorithm for the inter-day stochastic traffic assignment problem. *J Adv Transp* 46: 222–235
- Yu B, Yang ZZ (2011) An ant colony optimization model: the period vehicle routing problem with time windows. *Transp Res E Logist Transp Rev* 47:166–181
- Yu B, Yang ZZ, Xie JX (2011) A parallel improved ant colony optimization for multi-depot vehicle routing problem. *J Oper Res Soc* 62:183–188
- Zhang T, Chaovallitwongse WA, Zhang Y (2012) Scatter search for the stochastic travel-time vehicle routing problem with simultaneous pick-ups and deliveries. *Comput Oper Res* 39:2277–2290