



HAL
open science

Optimization of Chartered Buses Routes Under Uncertainties Using Probabilistic Vehicle Routing Problem Modeling

Zied Bouyahia, Hedi Haddad, Nafaâ Jabeur, Ahmed Nait-Sidi-Moh

► **To cite this version:**

Zied Bouyahia, Hedi Haddad, Nafaâ Jabeur, Ahmed Nait-Sidi-Moh. Optimization of Chartered Buses Routes Under Uncertainties Using Probabilistic Vehicle Routing Problem Modeling. *Procedia Computer Science*, 2018, 130, pp.644-651. 10.1016/j.procs.2018.04.115 . hal-01993391

HAL Id: hal-01993391

<https://hal.science/hal-01993391>

Submitted on 24 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



The 9th International Conference on Ambient Systems, Networks and Technologies
(ANT 2018)

Optimization of Chartered Buses Routes Under Uncertainties Using Probabilistic Vehicle Routing Problem Modeling

Zied Bouyahia^{a,*}, Hedi Haddad^a, Nafâa Jabeur^b, Ahmed Nait Sidi Moh^c

^aDepartment of Computer Science, College of Arts and Applied Sciences, Dhofar University, Salalah, Oman.

^bGerman University of Technology, P.O. Box 1816, Athaibah, PC 130, Sultanate of Oman

^cUniversité de Picardie Jules Verne, INSSET Saint Quentin, France

Abstract

Most of the current Probabilistic Vehicle Routing (PVRP) Problem models simultaneously address only one stochastic aspect of the problem, and there is a need of more realistic PVRP models that can take into consideration more than one stochastic aspect in the same time. In this paper we propose a new stochastic PVRP algorithm that takes into consideration both uncertain transport demand and travel time. We propose a priori generalization strategy that can be either rigid or flexible in order to provide decision makers with rapid and adjustable solution schemes. A simulated annealing algorithm has been implemented to solve the PVRP with stochastic travel times in the context of chartered buses, and the results are quite satisfactory.

© 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the Conference Program Chairs.

Keywords: Probabilistic Vehicle Routing Problem; Chartered Buses; Stochastic demand and travel time; Transport cost optimization

1. Introduction

Vehicle Routing Problem (VRP) is a widely studied topic in the field of Operations Research. It has been extensively applied to the transportation and logistics domain to solve the generic problem of satisfying transportation requests of spatially dispersed customers using a fleet of vehicles. The challenge consists in the identification of the best configuration of customers' pick up orders, time and routes that minimizes the total transportation cost⁵. Many varieties of VRP models have been presented in the literature, but the basic distinction is between deterministic (DVRP) and probabilistic/stochastic (PVRP or SVRP) models. Stochastic VRP models are proposed to reflect the intrinsic real-world uncertainty. In fact, in a real context, many aspects of the VRP problem cannot be exactly known in advance. Consequently, many PVRP models have been proposed to address every stochastic element of the VRP, such that VRP models with stochastic demand, stochastic customers, stochastic pickup and delivery, stochastic travel times and stochastic service time (see^{7,8} for complete reviews). However, most of the proposed probabilistic models simultaneously address only one stochastic aspect, and except few works⁹, there is a need of more realistic PVRP

* Corresponding author. Tel.: +968-2323-7215 ; fax: +968-2323-7215.
E-mail address: zbouyahia@du.edu.om

models that can take into consideration more than one stochastic aspect in the same time ⁴.

In this paper, we propose a new stochastic algorithm that takes into consideration both uncertain transport demand and travel time. As an application domain, we address the problem of chartered buses, which has been adopted by companies with a large number of employees to alleviate the costs of transportation of working force and to reduce the negative impact of lateness caused by public transportation. It has been also widely used by school buses' companies to daily drive students to their schools. Chartered buses is a typical example of PVRP with both uncertain demand and travel times. Regarding demand, customers - especially students- can be absent without pre-notification, or they may exceptionally ride with their friends or parents. Traveling times are uncertain too, mainly because of the dynamic nature of traffic state, especially delays caused by unusual events (accidents, heavy fog, urgent roads' maintenance, etc.). Both uncertainties need to be taken into consideration to calculate the optimal routing order of the VRP.

The remainder of this paper is organized as follows. In Section 2, we present an overview of the general elements of basic PVRP models. In Section 3, we recall the solution schemes that we use to solve the PVRP, which are inspired by the work of Bertsimas². In Section 4, we present and formalize our new algorithm for PVRP with both stochastic demand and travel time. In Section 5, we report the results of an experimental study that we performed to evaluate our algorithm. In Section 6, we stress the main contributions of our work and we highlight the extensions that shall be developed further in future work.

2. The probabilistic vehicle routing problem (PVRP)

From a conceptual perspective, basic VRP models suppose that all vehicles start from an initial location called *depot*, collect the demands according to the selected order, and once their capacities are attained they return back to the depot. The problem is modeled as a complete graph, where nodes represent customers' locations and weighted arcs correspond to the routes between them. To formulate the VRP, let us consider a complete network $\mathcal{G} = \{C, \mathcal{A}\}$ where C is a set of n nodes $0, 1, \dots, n$. Let the node 0 be considered as a depot such that all routes traveled by the vehicles should start and end with 0. Nodes set C correspond to customers pick up locations with respective demands D_i . $\mathcal{A} = \{a_{ij}\}$ is the set of arcs a_{ij} , where i and j are nodes in C . Let us denote by $\delta(i, j)$ the cost associated to a_{ij} , such that $\delta(i, j) \leq \delta(i, k) + \delta(k, j)$. For the sake of simplification, let us assume that costs are symmetric i.e. $\delta(i, j) = \delta(j, i)$. The case of asymmetric distances can be easily derived from the symmetric one¹. On the other hand, let us consider a set of f vehicles denoted as $\mathcal{F} = \{V_1, \dots, V_f\}$ with capacities $Q_j, 1 \leq j \leq f$ such that $\forall 1 \leq j \leq f, Q_j \geq \max_i D_i$. In this paper, we consider the case of equal-capacities vehicles, i.e. $\forall 1 \leq i \leq f, Q_i = Q$. The vehicle routing problem (VRP) is stated as follows: Given the network \mathcal{G} and the fleet \mathcal{F} , the objective is to find a set of minimum-length routes to meet the demands of the customers such that each route starts and finishes at the depot node 0 whenever its capacity is reached.

Despite its high potential to solve a wide range of practical problems dealing with transportation, logistics, etc., in the deterministic definition of VRP, we assume that the customers' demands are known when the solution is designed. However, this assumption is not realistic since these demands can vary from time to time and some of them might be canceled. A natural modeling for this situation is to consider the number of customers as a random variable which can be done by assigning to each customer node i , a probability p_i of having non-zero demand. The solution cost under this assumption is then a random variable and the associated goal is to minimize its expectancy. This probabilistic generalization of the vehicle routing problem was first proposed by D. Bertsimas² motivated by applications in strategic planning and distribution systems. Instead of resolving the problem when the demand becomes known, the author proposes to design an a priori sequence among all customers of minimal expected total length. The model that we propose in this paper is inspired by Bertsimas' solution and is detailed in the following section.

3. Solving the PVRP

3.1. Solution strategies

The PVRP can be solved using two different strategies: reoptimization and a priori. The reoptimization scheme consists in finding a minimum cost solution for every possible instance of the probabilistic problem. Each reduced version of the VRP instance is considered as a deterministic VRP defined over the subset $I \subseteq C$ of customers requiring visits. The optimum cost for the subset I is denoted by $L(I, Opt)$. Using the reoptimization scheme, the associated cost for the PVRP is a random variable denoted by $L(Opt)$ and its expected value is given by:

$$\mathbb{E} [L(Opt)] = \sum_{I \subseteq C} P(I)L(I, Opt) \tag{1}$$

As a matter of fact, reoptimization scheme provides an optimal solution albeit not workable since the number of possible subsets (i.e. $I \subseteq C$) is 2^n and, consequently, the computational complexity rises exponentially by increasing the number of nodes in C . In the a priori scheme, an a priori solution is first designed for the whole set of nodes in C regardless their probabilities to require visits. Then, for every subset $I \subseteq C$, the a priori solution is rapidly updated using a modification algorithm \mathcal{U} . The cost under the a priori scheme is a random variable denoted by $L(I, \mu)$ and its expectancy is given by the following expression:

$$\mathbb{E} [L(\mathcal{U})] = \sum_{I \subseteq C} P(I)L(I, \mathcal{U}) \tag{2}$$

Bertsimas proposed and studied two implementations of the a priori strategy depending on the modification method \mathcal{U} . These strategies are baptized rigid and flexible with respect to two modification methods \mathcal{U}_A and \mathcal{U}_B . Under the rigid strategy, the vehicle visits all the assigned nodes in the a priori route and stops only at nodes when the customer has non-zero demand. When the capacity of the vehicle is attained, the vehicle returns to the depot and then resumes the route from the next node in the designed route. In the flexible strategy, the vehicle visits only the nodes that require a visit. When the maximum capacity of a vehicle is reached at a given node i , the vehicle returns to the depot to unload; then resumes its route from the next node requiring a visit. Figure 1 illustrates the two a priori strategies.

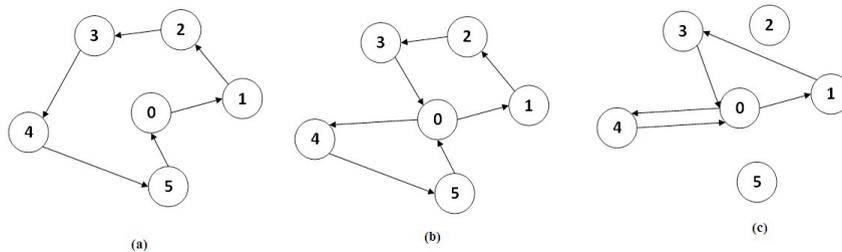


Fig. 1. Example of a priori strategy for a PVRP instance with 5 customers having binary demands and one vehicle of capacity 2. The a priori route (a), the modified routes when only customers 2 and 5 have non-zero demands under the rigid strategy (b) and the flexible strategy (c).

3.2. Objective function

Under the rigid a priori strategy relying on \mathcal{U}_A , the expected length of a feasible route is given by the sum of the expected cost of the a priori route $\sum_{i=0}^n \delta(i, i + 1)$ and the expected cost of the additional route the vehicle has to travel when its maximum capacity has been reached. To calculate the cost of the additional route, we assume without loss of generality that the maximum capacity of the vehicle is reached at the node i with a probability α_i . Whence, the vehicle shall return to the depot (which corresponds to the cost $\delta(i, 0)$) and then resume the route from the node $i + 1$ regardless its actual demand which corresponds to $\delta(0, i + 1)$. The length of this additional trip is $\delta(i, 0) + \delta(0, i + 1) - \delta(i, i + 1)$. The probability α_i is calculated using the following polynomial-time scheme:

- $\alpha_i = 0, \forall i = 0, \dots, Q - 1$
- $\alpha_{rQ+i} = p_{rQ+i} \sum_{k=1}^r \gamma(rQ + i - 1, kQ - 1), 0 \leq i \leq Q,$
 where $\gamma(m, c) = P(c \text{ among the } m \text{ customers have non-zero demand})$ and p_{rQ+i} is the probability that node $rQ+i$ has non-zero demand.

The computation of $\gamma(m, c)$ is achieved using the following recursive scheme:

- Boundary conditions: $\gamma(m, m) = \prod_{i=1}^m p_i$ and $\gamma(m, 0) = \prod_{i=1}^m (1 - p_i)$
- Induction: $\gamma(m, c) = p_m \gamma(m - 1, c - 1) + (1 - p_m) \gamma(m - 1, c)$

Finally, under the rigid strategy, the objective function is given by the following formula:

$$\mathbb{E}_{\mathcal{U}_A} [L_r] = \sum_{i=0}^n \delta(i, i + 1) + \sum_{i=1}^n \alpha_i s(i, i + 1), \tag{3}$$

where $s(i, j) = \delta(0, i) + \delta(j, 0) - \delta(i, j)$.

Under the flexible strategy, the objective function is calculated based on the minimum expected length of Hamiltonian tour in the sense of Probabilistic Traveling Salesman Problem⁶. Without loss of generality, let us assume that i is the first node in the a priori solution and that j is the last node before the vehicle has reached its maximum capacity. The expected costs to travel from and to the depot from the two aforementioned nodes i and j are respectively $\delta(0, i) p_i \prod_{k=1}^{i-1} (1 - p_k)$ and $\delta(j, 0) p_j \prod_{k=i+1}^m (1 - p_k)$. Contrarily to the reoptimization strategy, when the vehicle reaches its maximum capacity at a node i and returns to the depot, it skips all nodes with zero-demand following i and resumes its route from the first node with non zero-demand and therefore discards nodes $i + 1, \dots, j - 1$ from its route. Hence the objective function is defined as follows:

$$\begin{aligned} \mathbb{E}_{\mathcal{U}_B} [L_r] = & \sum_{i=1}^n \delta(0, i) p_i \prod_{k=1}^{i-1} (1 - p_k) + \sum_{j=1}^n \delta(j, 0) p_j \prod_{k=i+1}^m (1 - p_k) + \\ & \sum_{i=1}^n \sum_{j=i+1}^n \delta(i, j) p_i p_j \prod_{k=i+1}^{j-1} (1 - p_k) + \sum_{i=1}^n \sum_{j=i+1}^n \alpha_i s(i, j) p_j \prod_{k=i+1}^{j-1} (1 - p_k) \end{aligned} \tag{4}$$

Time complexity. The computation of the expected costs under both reoptimization and a priori strategies relies on the calculation of the probabilities λ_i . Since $\lambda_i = 0, \forall 0, \dots, Q - 1$, then $n - Q$ probabilities require computation, each consisting in the summation of a maximum of $\frac{n}{Q}$ terms. Hence, the computation cost of λ_i is $O\left((n - Q) \frac{n}{Q}\right) = O(n^2)$.

Remark. Let us note that in the particular case where all customers have the same probability to require visit, the probability $\gamma(m, c)$ is given by the following closed form formula:

$$\gamma(m, c) = \binom{m}{c} p^c (1 - p)^{(m-c)}$$

This case is trivial but serves for customer clustering purposes when the vehicle number and capacities are limited. During the pooling stage, customers having the same probability of requiring visit can be assigned to the same fleet. Another interest of this case is that it shows the similarity of this problem with other probabilistic combinatorial problems such as probabilistic traveling salesman problem (PTSP)⁶, the probabilistic scheduling problem (PSP)³, etc. This similarity allows for the utilization of combinatorial properties of well-studied problems and therefore the design of efficient solution schemes.

4. PVRP with stochastic travel times

In this section we present our proposed PVRP with both stochastic demand and travel time. We first present the objective function, then we elaborate the lower bounds before elaborating the steps of the proposed algorithm.

4.1. Objective function

From an application point of view, the costs of the arcs linking the nodes in \mathcal{G} are usually varying depending on the traffic state between the visited nodes. The use of constant distances as costs for the links is not realistic and leads to an inaccurate estimation of routes costs and a significant discrepancy between the model and the ground-truth data. Instead of considering the physical length of the routes between two nodes i and j , we propose to use the travel time. We assume that the travel time between nodes i and j is a random variable following an exponential distribution of parameter λ_{ij} . To illustrate the impact of travel time randomness on the solution design let us consider the following example. Consider a complete graph \mathcal{G} with one depot and three customer nodes 1, 2, 3. Let us also consider a fleet with 2 vehicles. For the sake of simplification, let us assume that travel times between these nodes are symmetrical and that travel times from depot to customer nodes are identical ($\delta(0, i) \sim e(\lambda_0)$). Let us consider a feasible route in which the nodes 1 and 3 are visited first (by the two vehicles). After nodes 1 and 3 have been visited, the next node to be visited is 2 with a probability p . The next vehicle to visit node 2 is the one that has finished his visit which corresponds to $\min_{j=1,3} \delta(0, j)$. Since $\delta(0, j) \sim e(\lambda_{0j})$ then $\min_{j=1,3} \delta(0, j) \sim e(\lambda_{0,1} + \lambda_{0,3})$. Then if we denote by τ_2 , the moment node 2 is visited by one of the vehicles, then the following cases are possible (see Figure 2):

- Nodes 1 and 3 do not require visit and node 2 can be visited by one of the two available vehicles.
- Only one node among $\{1, 3\}$ requires visit. Then possible cases are (i) node 2 does not require visit, or node 2 is visited after 1 (ii) or after 3 (iii).
- Both nodes 1 and 3 are visited and node 2 is visited afterwards. In this case, we distinguish two main scenarios depending on which node (1 or 3) is reached first. Then depending on the probability p node 2 shall or not require visit. Hence there are 6 possible routes.

Let ϕ_i be the random variable referring to the time point at which the next visit has to be planned, we have:

$$\mathbb{E}[\phi_1] = \frac{p^2}{\lambda_{01} + \lambda_{03}} + p^2(1 - p) + \left(\frac{1}{\lambda_{01} + \lambda_{12}}\right) + p(1 - p)^2 \left(\frac{1}{\sum_{j=1,3} \lambda_{0j} + \sum_{j=1,3} \lambda_{j2}}\right).$$

Since $\mathbb{E}[\phi_2] = \mathbb{E}[\phi_1] + \mathbb{E}[\phi_2 - \phi_1]$, we have:

$$\begin{aligned} \mathbb{E}[\phi_2 - \phi_1] &= \frac{p^2}{\lambda_{01} + \lambda_{03}} \mathbb{E}[\phi_2 - \phi_1 \mid \text{nodes 1 and 3 are visited and node 1 is visited first}] + \\ &\frac{p^2}{\lambda_{01} + \lambda_{03}} \mathbb{E}[\phi_2 - \phi_1 \mid \text{nodes 1 and 3 are visited and node 3 is visited first}] + \\ &p(1 - p) \mathbb{E}[\phi_2 - \phi_1 \mid \text{Only node 1 is visited}] + p(1 - p) \mathbb{E}[\phi_2 - \phi_1 \mid \text{Only node 3 is visited}] + \\ &(1 - p)^2 \mathbb{E}[\phi_2 - \phi_1 \mid \text{nodes 1 and 3 are both absent}] \end{aligned}$$

Hence:

$$\begin{aligned} \mathbb{E}[\phi_2] &= \frac{p^2}{\lambda_{01} + \lambda_{03}} + p^2(1 - p) + \left(\frac{1}{\lambda_{01} + \lambda_{12}}\right) + p(1 - p)^2 \left(\frac{1}{\sum_{j=1,3} \lambda_{0j} + \sum_{j=1,3} \lambda_{j2}}\right) + \\ &\frac{p^2}{\lambda_{01} + \lambda_{03}} + p^2(1 - p) \frac{\lambda_{03}}{\lambda_{01} + \lambda_{03}} p \left(\frac{1}{\lambda_{01} + \lambda_{12}} + (1 - p) \frac{1}{\lambda_{01}}\right) + \\ &p^2(1 - p) \left(\frac{1}{\lambda_{02} + \lambda_{03}} \left(\frac{\lambda_{03}}{\lambda_{03}} + \frac{\lambda_{02}}{\lambda_{02}}\right)\right) + p^2(1 - p) \left(\frac{1}{\lambda_{02} + \lambda_{03}} (\lambda_{03} \lambda_{02} + \frac{\lambda_{03}}{\lambda_{02}})\right) + \mathbb{E}[\phi_1] \end{aligned}$$

Since the distances $\delta(i, j)$, $1 \leq i, j \leq n$ follow exponential distributions with parameters λ_{ij} , then $\min_{i,j} \delta_{ij} \sim e(\lambda)$, $\lambda = \sum_{i,j} \lambda_{ij}$. Therefore, $p(\min_{i,j} \{\lambda_{ij} = \lambda_{kl}\}) = \frac{\lambda_{kl}}{\lambda}$. Then, under the rigid a priori strategy and random distances:

$$\mathbb{E}[\mathbb{E}_{\mathcal{U}_A} [L_r, \{\}]] = \sum_{I \subseteq C} p(I) \frac{|I|}{\lambda_I} + \sum_{I \subseteq C} \sum_{k \in I | k \text{ is served}} \sum_{i \in \text{Pred}(k)} \frac{\lambda_{ik}}{\lambda_I} \mathbb{E}_{\mathcal{U}_A} [I - \{k\}] \tag{5}$$

Remark. In the particular case, where all arcs' values follow the same exponential distribution $e(\lambda)$, the distances $\delta(i, j)$ follow a gamma distribution $\Gamma(1, \frac{1}{\lambda})$. Whence, $\delta(i, k) + \delta(k, j) \sim \Gamma(2, \frac{1}{\lambda})$. Based on an expectancy-wise comparison between $\delta(i, j)$ and $\delta(i, k) + \delta(k, j)$, the triangle inequality holds.

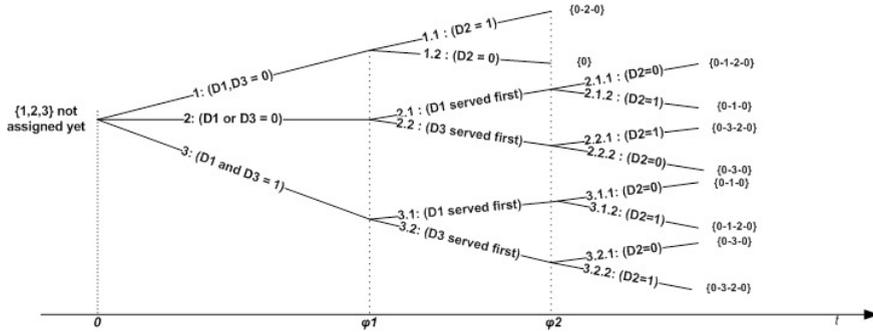


Fig. 2. Allocation scenarios for a PVRP instance with three customers having binary demands, vehicle capacity $Q = 2$ and distances following an exponential distribution

4.2. Lower bounds

Let us consider the particular case of unit-demands. This configuration corresponds to a door-to-door visit strategy. We assume that the objective is to design a set of routes to pick travelers at their residence locations. Under this assumption, we derive two lower bounds for both a priori and reoptimization schemes. Assume that each customer has a probability p_i of requiring a ride from node $i \in \mathcal{G}$. Let us first consider the reoptimization scheme and let us also consider a feasible solution for the problem in which the vehicle (bus) has to collect a subset $I_u \subset \mathcal{N}$ such that $|I_u| \leq Q$. After $|I_u|$ customer have been driven to the depot node, the vehicle resumes its tour with the first node in the next subset $|I_{u+1}|$, $I_u \cap I_{u+1} = \emptyset$. If L_u^* is the cost of the route to visit all nodes in $|I_u|$ in the optimal solution, then $L_u^* \geq \max_{i \in I_u} \delta(0, i) + \delta(i, 0)$. Since $\max_{i \in I_u} \delta(0, i) \geq \frac{\sum_{i \in I_u} \delta(0, i)}{|I_u|}$ and $\max_{i \in I_u} \delta(i, 0) \geq \frac{\sum_{i \in I_u} \delta(i, 0)}{|I_u|}$ and $|I_u| \leq Q$, then we can write $L_u^* \geq \frac{\sum_{i \in I_u} \delta(0, i) + \delta(i, 0)}{Q}$. Hence, $R^* \geq 2 \frac{\sum_{i \in \mathcal{N}} \delta(0, i) + \delta(i, 0)}{Q}$. Finally, $\mathbb{E}[L_{S^*}] \geq \frac{1}{Q} \sum_{\mathcal{N} \subset C} P(\mathcal{N}) \sum_{i \in C} \delta(i, 0) + \delta(0, i) = \frac{1}{Q} \sum_{i=1}^n \delta(i, 0) + \delta(0, i)p_i$.

Let us note that since the costs are calculated based on the travel time between nodes in C and these costs are considered random variables taking into account the fluctuating traffic status, the triangle inequality (i.e. $\delta(i, j) \leq \delta(i, k) + \delta(k, j)$) does not apply. If we denote by λ_{ij} , λ_{ik} and λ_{kj} the parameters of the exponential distributions of $\delta(i, j)$, $\delta(i, k)$ and $\delta(k, j)$ respectively, the density of $\delta(i, k) + \delta(k, j)$ is $f_{\delta(i,k)+\delta(k,j)}(x) = (\lambda_{ik}\lambda_{kj}) \left(\frac{e^{-\lambda_{ik}\lambda_{kj}x}}{(\lambda_{kj}-\lambda_{ik})} \right)$.

In order to computer tighter lower bounds when the triangle inequality does not hold for random arcs values, we set $\delta'(i, j) = \delta(i, j) - (d_i + d_j)$ such that $\delta'(i, j) \geq 0$ (one can choose $d_i = \min \delta(i, j)$). Let $S = 0, 1, \dots, n, 0$ be a minimum-cost solution for the PVRP with respect to $\delta(i, j)$, $0 \leq i, j \leq n$ under the a priori strategy and let us denote by $\mathbb{E}[S]$ and $\mathbb{E}[S']$ the expected costs of S and S' with respect to $\delta(i, j)$, $0 \leq i, j \leq n$ and $\delta'(i, j)$, $0 \leq i, j \leq n$. Then, we have:

$$\mathbb{E}[S'] = \mathbb{E}[S] - \sum_{i,j=1}^n (d_i + d_j) - d_0(1 + \sum_{i=1}^n \alpha_i)$$

Since $\mathbb{E}[S'] \geq 0$ then:

$$\mathbb{E}[S] \geq \sum_{i,j=1}^n (d_i + d_j) - d_0(1 + \sum_{i=1}^n \alpha_i) \tag{6}$$

4.3. Solution procedure

The VRP is NP-complete and its exact resolution is not workable for large instances. The proposed PVRP is as at least as complex as its deterministic counterpart. An heuristic based algorithm should be used to yield near-optimal

solutions. In this work, we outline a simulated annealing algorithm for the solution design of the stochastic vehicle routing problem with uncertain demands and varying inter-nodes costs. The proposed SA algorithm is achieved in five main steps:

Step 1 First, an initial route is randomly generated using as follows:

1. Generate a random number of nodes in each route τ_r , $1 \leq r \leq R$ such that $\sum_{i \in \tau_r} D_i \leq Q$.
2. For each route τ_r , $1 \leq r \leq R$, assign first nodes with highest probabilities of presence p_i .

Step 2 Using a constructive heuristic, a new set of routes is generated based on the initial route. Let $\mathbb{E}[\mathbb{E}_{\mathcal{U}}[L_r]]^q$ be the expected cost of the new solution and $\mathbb{E}[\mathbb{E}_{\mathcal{U}}[L_r]]^{q-1}$ be the old one. The perturbation in the so-called system energy is $\Delta \mathbb{E}_{\mathcal{U}} = \mathbb{E}[\mathbb{E}_{\mathcal{U}}[L_r]]^q - \mathbb{E}_{\mathcal{U}}[L_r]^{q-1}$. We proceed to step 3 if there is an improvement in the expected solution cost (i.e. $\Delta \mathbb{E}_{\mathcal{U}} \geq 0$). Otherwise, step 4 should follow.

The constructive heuristic used in this work relies on the probabilities p_i and relies on three functions: **Swap**, **Insert** and **Revert**.

At each iteration, one of the three functions is randomly selected to perturb the current solution. **Swap** is used to randomly swap nodes in the a priori route, **Insert** is used to insert a random sequence of nodes in the a priori solution and **Revert** inverts the order of visits of a random number of nodes.

Step 3 Generate a random number $R \sim U[0, 1]$ and compute the probability P to accept the solution built in Step 2 such that $P = e^{-\frac{\Delta \mathbb{E}_{\mathcal{U}}}{T}}$. If $R < P$ then step 4 follows, otherwise, the old set of routes should be kept and reapply step 2.

Step 4 The obtained solution $S^q = \{r_1^q, \dots, r_m^q\}$ is considered for the next iteration and $\mathbb{E}[\mathbb{E}_{\mathcal{U}_B}[L_r]]^{q-1} = \mathbb{E}[\mathbb{E}_{\mathcal{U}_B}[L_r]]^q$.

Step 5 Let us denote by $\mathbb{E}[L]^q$ (resp. $\mathbb{E}[L]^{q-1}$) the expected cost at the current (resp. previous) SA iteration, with $q > 1$. We compute the thermal equilibrium condition defined as $\chi = \frac{\mathbb{E}[L]^q - \mathbb{E}[L]^{q-1}}{\mathbb{E}[L]^{q-1}}$. If $\chi \leq \varepsilon$ (ε is a predefined parameter), the thermal equilibrium is attained and then the temperature T is decreased. Then step 2 follows and $q \leftarrow q + 1$.

The algorithm stops when the temperature threshold T_{STOP} is reached.

5. Experimental Study

5.1. Experimental setup

In this section we perform an experimental evaluation on randomly generated benchmarks, in order to evaluate the performances of the re-optimization strategy and the a priori strategy. The experimental study consists in considering three different values of fleet size $f = 2, 4, 6$ and generating, for each value, a complete graph \mathcal{G} with n customer nodes ($n = 20, 50, 70, 100, 200, 500$) where the arcs a_{ij} , $1 \leq i, j \leq n$ are valued by random travel times following exponential distributions with parameters $\lambda_{ij} \in [\lambda_{min}, \lambda_{max}]$ ($\lambda_{min} = 5, \lambda_{max} = 20$). Moreover, we assume that all vehicles have the same capacity $Q = 10$. For the computation of a mean value of the expected routing costs under a priori strategy, we generate, for each initial problem of size n , a subset of 1000 "absent" nodes to be removed from vehicles' routes under both re-optimization and a priori strategies. We also assign to each node i a probability p_i assumed to follow a uniform distribution in $[0, 1]$.

5.2. Preliminary results

The quality assessment of the a priori strategies is performed by measuring the deviation of a solution found by means of a priori and re-optimization schemes compared to lower bounds. If we denote by LB the lower bound calculated in eq. (6), we calculate for each solution cost $L(I, \mathcal{U}_A)$ and $L(I, \mathcal{U}_B)$, the following ratios:

$$R_{\mathcal{U}_A} = \frac{\sum_{i=1}^{1000} L(I, \mathcal{U}_A)}{1000 * LB} \times 100, \text{ and } R_{\mathcal{U}_B} = \frac{\sum_{i=1}^{1000} L(I, \mathcal{U}_B)}{1000 * LB} \times 100 \quad (7)$$

Table 1. A priori strategy based solutions' costs against lower bounds for different problem sizes $n = 20, 50, 70, 100, 200, 500$ and fleet size $f = 2, 4, 6$. (1.0 is 1%).

n	20		50		70		100		200		500	
	\mathcal{U}_A	\mathcal{U}_B										
f=2	13.6	9.22	11.58	8.66	10.25	8.55	10.11	8.23	8.11	6.25	4.12	3.99
f=4	15.56	13.62	12.58	9.13	12.15	9.08	11.88	9.11	9.18	7.11	5.15	4.50
f=6	16.42	12.85	14.15	11.69	13.15	11.85	12.89	11.02	10.87	8.55	7.04	5.98

Table 1 reports the results of the conducted experiments. We notice that the higher is the number of vehicles, the higher is the deviation of a priori strategy from optimal schedules. We also notice that the flexible a priori strategy outperforms the rigid one regardless the size of the problem. While the number of vehicles slightly affects the expected cost, the number of customers have a major impact on solution quality. Under the rigid strategy, the ratio $R_{\mathcal{U}_A}$ does not exceed 8% for instances sizes larger than 500 (The results corresponding to $n > 500$ have been discarded). The obtained results suggest that asymptotically, the a priori strategy converges towards an optimal solution. An ongoing asymptotic analysis is being conducted to confirm the experimental findings.

6. Conclusion

In this paper, we proposed a stochastic generalization of a classical vehicle routing problem in which both customer demands as well as the distances between customers are assumed to be random variables. This stochastic generalization allows for a more realistic modeling and provides long term robust solutions for several applications in the context of transport and logistics. We explicitly introduced probabilities in the problem definition in order to cope with uncertainties dealt with on a daily basis in the context of chartered and private buses route planning. Depending on the availability of required customers demands data, the a priori strategy can be either rigid or flexible in order to provide decision makers with rapid and adjustable solution schemes. The distances between nodes have been assumed to be random with known distributions to conform with road network state fluctuations. A simulated annealing algorithm has been implemented to solve the PVRP with stochastic travel times and the results are quite satisfactory. A deeper study is being conducted to improve the proposed lower bounds.

References

1. D. Bertsimas. The probabilistic vehicle routing problem. *Sloan School of Management, Massachusetts Institute of Technology*, 1988.
2. D. Bertsimas. A vehicle routing problem with stochastic demand. *Operations Research*, 40:574–585, 1992.
3. Z. Bouyahia, M. Bellalouna, P. Jaillet, and K. Ghedira. A priori parallel machines scheduling. *Computers and Industrial Engineering*, 58(3):488 – 500, 2010.
4. K. Braekers, K. Ramaekers, and I. V. Nieuwenhuysse. The vehicle routing problem: State of the art classification and review. *Computers and Industrial Engineering*, 99:300 – 313, 2016.
5. B. Eksioğlu, A. V. Vural, and A. Reisman. The vehicle routing problem: A taxonomic review. *Computers and Industrial Engineering*, 57(4):1472 – 1483, 2009.
6. P. Jaillet. A priori solution of a traveling salesman problem in which a random subset of the customers are visited. *Operations Research*, 36(6):929–936, 1988.
7. J. Oyola, H. Arntzen, and D. L. Woodruff. The stochastic vehicle routing problem, a literature review, part i: models. *EURO Journal on Transportation and Logistics*, Oct 2016.
8. J. Oyola, H. Arntzen, and D. L. Woodruff. The stochastic vehicle routing problem, a literature review, part ii: solution methods. *EURO Journal on Transportation and Logistics*, 6(4):349–388, Dec 2017.
9. K. Sörensen and M. Sevaux. A practical approach for robust and flexible vehicle routing using metaheuristics and monte carlo sampling. *Journal of Mathematical Modelling and Algorithms*, 8(4):387, Jun 2009.