Phase-type distributions and their applications to the vehicle routing problem with stochastic travel and service times

Andrés Gómez, Raha Akhavan-Tabatabaei, Andrés Medaglia, Jorge E. Mendoza

Phase-type distributions and their application to the vehicle routing problem with stochastic travel and service times

Andrés Gómez[1], Raha Akhavan-Tabatabaei[1], Andrés L. Medaglia[1], Jorge E. Mendoza[2]

[1] Centro para la Optimización y Probabilidad Aplicada (COPA), Departamento de Ingeniería Industrial, Universidad de los Andes, Bogotá, Colombia
agomeze8@gmail.com, {r.akhavan, amedagli}@uniandes.edu.co
[2] Université François-Rabelais de Tours, CNRS, LI EA 6300, OC ERL CNRS 6305, Tours, France
jorge.mendoza@univ-tours.fr

Keywords: *Stochastic vehicle routing, Phase Type distributions*

1 Introduction

The vehicle routing problem with stochastic travel and service times (VRPSTT) is defined on a complete graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{0, v_1, v_2, \ldots, v_n\}$ is the vertex set and $\mathcal{E} = \{e = (v_i, v_j) : v_i, v_j \in \mathcal{V}, v_i \neq v_j\}$ is the edge set. Vertices $v_i \in \mathcal{V} \setminus \{0\}$ represent the customers and vertex 0 represents the depot. An edge weight $\tilde{t}_e$, associated with edge $e = (v_i, v_j)$, represents the random travel time along edge e. Each customer $v_i \in V \setminus \{0\}$ has a random service time $\tilde{s}_{v_i}$ and a known demand d_{v_i} for a given product. Both travel and service times are assumed to follow known distributions. Customers are served by an unlimited fleet of homogeneous vehicles located at the depot, each with a maximum capacity Q and a maximum route duration T. The objective is to design a route set $\mathcal{R}$ of minimum total expected duration $E\left[\tilde{T}(\mathcal{R})\right] = \sum_{r\in\mathcal{R}} E\left[\tilde{T}_r\right]$, where $\tilde{T}(\mathcal{R})$ is the total (random) duration of the route set, $\tilde{T}_r$ is the (random) duration of route r, and $E[\cdot]$ denotes the expected value. Each route $r \in \mathcal{R}$ is an ordered set $r = \left(0, v_{(1)}, \ldots, v_{(i)}, \ldots, v_{(n_r)}, 0\right)$, where $v_{(i)} \in \mathcal{V} \setminus \{0\}$ is the i-th customer visited in the route, n_r is the number of customers serviced by the route, and $(v_{(i)}, v_{(i+1)}) \in \mathcal{E}$ (with $v_{(0)} = v_{(n_r+1)} = 0$). We will refer to route r, depending on the context, either as the sequence of vertices or edges in the route. In the baseline version of the problem, aside from the classical capacity constraint, each route $r \in \mathcal{R}$ satisfies a duration constraint stating that:

$$P\left(\tilde{T}_r \leq T\right) \geq \beta \quad \forall r \in \mathcal{R} \tag{1}$$

where the left-hand side is the probability that the route completes before T and $\beta \in [0, 1]$ is a minimum acceptance threshold. The latter represents the decisions maker's aversion towards violations to the duration constraint. Note that the total duration of a route $\tilde{T}_r$ is the convolution of several random variables representing the travel and service times in the route.

One of the main challenges when solving the VRPSTT is selecting the appropriate distributions to model random variables. First, the selected distributions should be able to accurately model travel and service times. Second, computing their convolutions should not add heavy computational burden for optimization algorithms. The most widely-used approach in the stochastic VRP literature is to model random variables using families of additive distributions such as the normal or gamma distributions [4, 3, 2, 1, 10, 11]. Using these distributions, computing the convolution of random variables can be done efficiently. On the other hand, recent studies suggest that often these distributions inaccurately model travel times [5, 9].

Computing the convolutions of travel and service times when the distributions are non-additive is not a trivial task. To overcome this difficulty, researchers have developed approaches based on Monte Carlo simulation [3, 6] and queuing theory [12, 13]. In this talk we present a new approach based on Phase-type (PH) distributions. Using this family of distributions one can closely approximate any positive, continuous distribution with arbitrary precision and compute their convolutions in an exact and efficient manner. We show how PH distributions can be used to build a route evaluator that can be embedded into any optimization (search) algorithm for the VRPSTT. To assess the benefits of our approach, we compare the performance of our PH route evaluator with route evaluators based on normal distributions and Monte-Carlo simulation.

2 Building the PH route evaluator

2.1 Why PH distributions?

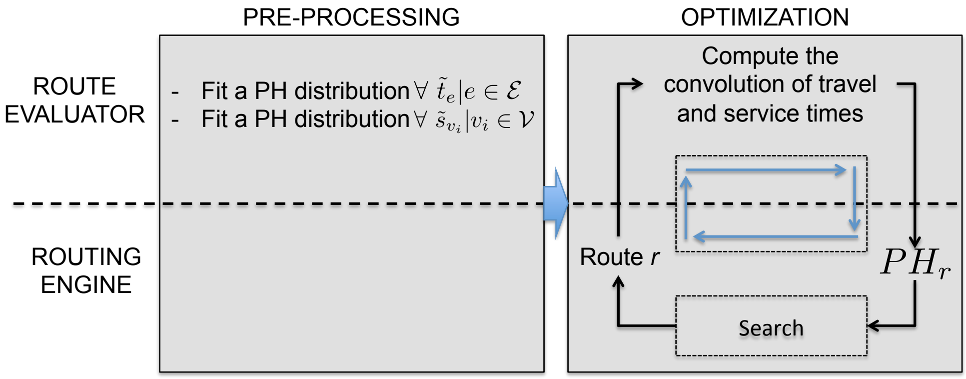
PH distributions are dense in the set of continuous density functions with support on $[0, \infty)$, meaning that there exists a PH distribution arbitrarily close to any positive distribution. Finding such a PH distribution is a process known as *fitting*, for which there exist efficient algorithms. In other words, since travel and service times are always positive, they can be accurately modeled using PH distributions. From the computational point of view, PH distributions have two properties that make them a good fit for optimization algorithms: i) the convolution of PH distributions is again a PH distribution; and, ii) the distribution function, the expected value, and higher-order moments of a PH distribution can be found in closed form.

2.2 The route evaluator

Search algorithms for the VRPSTT can be decoupled into two components: the *routing engine* and the *route evaluator*. The routing engine is responsible for exploring the solution space, unveiling new routes to make up a solution. On the other hand, the route evaluator is responsible for extracting the *performance metrics* of a route such as the probability of satisfying the duration constraint in (1). In our approach, the route evaluator builds the PH distribution of the total duration of the route (PH_r) and uses it to extract the performance metrics.

Figure 1 illustrates the interaction between our route evaluator and the routing engine. During the pre-processing phase, the route evaluator fits a PH distribution to each random variable (travel time or service time) in the problem instance. Then, during the optimization phase, the routing engine invokes the route evaluator every time the evaluation of a route r is needed. The route evaluator computes and returns PH_r as the convolution of the PH distributions fitted (during the pre-processing phase) to the travel and service times that appear in the route.

Figure 1: Route evaluator general structure



3 Computational experiments

Since no publicly available benchmark exists for the VRPSTT, we adapted four instances from the classical Christofides, Mingozzi, and Toth Capacitated VRP instances. For each instance we built six versions with different travel and service time distributions. These distributions were selected to model different real-world scenarios. For instance, we built a version of each instance in which travel times are lognormally distributed, trying to reproduce the travel times of congested links in uninterrupted traffic.

To compare our approach to alternatives from the literature we implemented three route evaluators: the proposed PH evaluator; an evaluator that assumes that all random variables are normally distributed; and, an evaluator that computes the performance metrics using Monte Carlo simulation. We embedded each of these evaluators in the multi-space sampling heuristic (MSH) proposed by Mendoza and Villegas for the VRP with stochastic demands [7]. The algorithm follows a two-phase solution strategy. In the first phase, it samples multiple solution representation spaces; while in the second it assembles the best possible solution using parts of the sampled elements. The approach operates as follows. At each iteration k, the algorithm selects a *sampling heuristic* from a set $\mathcal{H}$ of randomized traveling salesman problem (TSP) heuristics and uses it to build a giant tour p^k visiting all customers. Then, the algorithm makes a call to a *splitting procedure*, similar to the one introduced by Prins [8], to retrieve a tuple $\langle \Omega^k, s^k \rangle$, where Ω^k is the set of all feasible routes (in terms of not exceeding the vehicle capacity Q and satisfying the constraints in $\mathcal{C}$) that can be extracted from p^k without altering the order of the customers, and s^k is the best solution that can be built using routes from Ω^k. The routes in Ω^k join a set of sampled routes Ω (i.e., $\Omega \leftarrow \Omega \cup \Omega^k$), while s^k is used to update an upper bound $f(s^*)$ on the objective function of the final solution, where s^* is the best solution found so far. After a total of K iterations, the heuristic proceeds to the assembly phase, which consists in solving a set partitioning (column-oriented) formulation of the underlying routing problem over Ω, using $f(s^*)$ as an upper bound.

To analyze the performance delivered by MSH running with the three route evaluators, we evaluate each solution over 1M realizations of the stochastic travel and service times. Our results suggest that assuming normal distributions is a valid alternative when the underlying travel times do not exhibit large skewness. However, in practice, more often than not, travel times have a positive skewness and normal distributions tend to lead to poor routing decisions. According to our experiments, using Monte-Carlo simulation to evaluate routes leads to overly optimistic solutions that do not satisfy the duration constraints in practice. On the other hand, using PH distributions leads to solutions that are consistently more reliable and have similar total expected duration. In conclusion PH distributions are the most accurate choice when modeling travel and service times for all distributions used in our experiments.

References

[1] N. Chen, M. H. HÃă, A. Langevin, and M. Gendreau. Optimizing road network daily maintenance operations with stochastic service and travel times. *Transportation Research Part E: Logistics and Transportation Review*, 64(0):88 – 102, 2014.

[2] M. A. Figliozzi. The impacts of congestion on commercial vehicle tour characteristics and costs. *Transportation Research Part E: Logistics and Transportation Review*, 46(4):496 – 506, 2010.

[3] A. S. Kenyon and D. P. Morton. Stochastic vehicle routing with random travel times. *Transportation Science*, 37(1):69–82, 2003.

[4] G. Laporte, F. Louveaux, and H. Mercure. The vehicle routing problem with stochastic travel times. *Transportation Science*, 26(3):161–170, 1992.

[5] C. Lecluyse, T. Van Woensel, and H. Peremans. Vehicle routing with stochastic time-dependent travel times. *4OR: A Quarterly Journal of Operations Research*, 7:363–377, 2009.

[6] X. Li, P. Tian, and S. C. H. Leung. Vehicle routing problems with time windows and stochastic travel and service times: Models and algorithm. *International Journal of Production Economics*, 125:137–145, 2010.

[7] J. E. Mendoza and J. G.. Villegas. A multi-space sampling heuristic for the vehicle routing problem with stochastic demands. *Optimization Letters*, 7(7):1503–1516, 2013.

[8] C. Prins. A simple and effective evolutionary algorithm for the vehicle routing problem. *Computers & Operations Research*, 31(1):1985 – 2002, 2004.

[9] S. Susilawati, M. Taylor, and S. Somenahalli. Distributions of travel time variability on urban roads. *Journal of Advanced Transportation*, 47(8):720–736, 2013.

[10] D. Taş, N. Dellaert, T. van Woensel, and T. de Kok. Vehicle routing problem with stochastic travel times including soft time windows and service costs. *Computers & Operations Research*, 40(1):214 – 224, 2013.

[11] D. Taş, M. Gendreau, N. Dellaert, T. Van Woensel, and A.G. de Kok. Vehicle routing with soft time windows and stochastic travel times: A column generation and branch-and-price solution approach. *European Journal of Operational Research*, 236(3):789 – 799, 2014.

[12] T. Van Woensel, L. Kerbache, H. Peremans, and N. Vandaele. A vehicle routing problem with stochastic travel times. In *Fourth Aegean International Conference on Analysis of Manufacturing Systems, Samos, Greece*, 2003.

[13] T. Van Woensel, L. Kerbache, H. Peremans, and N. Vandaele. A queueing framework for routing problems with time-dependent travel times. *Journal of Mathematical Modelling and Algorithms*, 6:151–173, 2007.