CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

The Multi-Path Traveling Salesman Problem with Stochastic Travel Costs

Roberto Tadei
Guido Perboli
Francesca Perfetti

January 2013

CIRRELT-2013-01

Bureaux de Montréal :
Université de Montréal
C.P. 6128, succ. Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

The Multi-Path Traveling Salesman Problem with Stochastic Travel Costs

Roberto Tadei[1], Guido Perboli[1,2,*], Francesca Perfetti[1]

[1] Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

Abstract. Given a set of nodes, where each pair of nodes is connected by several paths and each path shows a stochastic travel cost with unknown distribution, the multipath Traveling Salesman Problem (TSP) with stochastic travel costs aims at finding an expected minimum Hamiltonian tour connecting all nodes. Under a mild assumption on the unknown probability distribution a deterministic approximation of the stochastic problem is given. The comparison of such approximation with a Montecarlo simulation shows both the accuracy and the efficiency of the deterministic approximation, with a mean percentage gap around 2% and a reduction of the computational times of two orders of magnitude.

Keywords: Traveling salesman problem (TSP), multiple paths, stochastic travel costs, deterministic approximation.

Acknowledgements. Partial funding was provided by the Ministero dell'Istruzione, Università e Ricerca (MIUR) (Italian Ministry of University and Research), under the 2009 PRIN Project “Methods and Algorithms for the Logistics Optimization” and by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Industrial Research Chair and Discovery Grants programs.

---

* Corresponding author: Guido.Perboli@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2013

1 Introduction

In the past decade, City Logistics pushed researchers towards the definition of a new paradigm of transportation and supply chain integration in urban areas. In recent years, this paradigm has been extended with the introduction of the concept of Smart City (Chourabi et al., 2012), where "smart" implies to incorporate a plethora of methods and disciplines in a holistic vision in order to mitigate the problems generated by population growth and its rapid urbanization. In this context, new transportation issues emerge, bringing researchers to define new transportation problems and, in particular, to incorporate information about uncertainty and multiple attributes (Perboli et al., 2012, 2011; Tadei et al., 2012). In this paper we present the multi-path Traveling Salesman Problem with stochastic travel costs (mpTSP$_s$), a new stochastic variant of the Traveling Salesman Problem. Given a set of nodes, where each pair of nodes is connected by several paths and each path shows a stochastic travel cost with unknown distribution, the mpTSP$_s$ aims at finding an expected minimum Hamiltonian tour connecting all nodes.

The mpTSP$_s$ arises in City Logistics applications when one has to design tours to provide services such as garbage collection, periodic delivery of goods in urban grocery distribution, and periodic checks of shared resources as in bike sharing services. In these situations the decision maker must provide tours that will be used for a time horizon which spans from one to several weeks. In this case the different paths connecting pairs of nodes in the city are affected by the uncertainty due to the different time dependent travel time distributions of the different paths. Moreover, in many cases even an approximated knowledge of the travel time distribution is made difficult by the large size of the data involved and the high variance of the travel times. In more detail, the introduction of this problem is also motivated by a real-life Smart City application, the PIE_VERDE project, a project funded by the ERDF - European Regional Development Fund for the development of new planning tools for freight delivery in urban areas by means of electrical vehicles. In this project one of the goals is to plan and manage a two-echelon delivery service, where the trucks are not allowed to directly enter inside the city, but the freight is consolidated in small peripheral depots and from them brought to customers by means of environmental-friendly vehicles (Perboli et al., 2011). In this application context, a crucial part is played by the planning of periodic tours between recurrent nodes. In this case one aims to plan a tour for each vehicle which is valid for a given time horizon. Unfortunately, at the time of planning, the decision maker has only a rough idea of the different paths interconnecting any pair of nodes of the transportation network. Moreover, due to congestion the travel time profile of these paths rapidly changes during the day.

Similar problems can be also found in other applications, like garbage collection or periodic replenishment of medium-sized grocery stores, where the tours of the trucks are designed in advance and cannot be changed for a fixed number of weeks.Recently, the European project Citylog (CITYLOG Consortium, 2010), a joint project between IVECO and TNT, presented the BentoBox, a modular system of containers for envelopes delivery. The BentoBox are usually placed in malls and shopping centers and the company needs to design fixed tours in order to store the envelopes in the BentoBox.

This paper introduces the formulation of the mpTSP$_s$. From this formulation a deterministic approximation is derived. In particular, under a mild hypothesis on the unknown probability distribution of the travel time for the different paths, the deterministic approximation becomes a TSP problem where the minimum expected total travel cost is equivalent to the maximum of the logarithm of the total accessibility of the Hamiltonian tours to the path set. The quality of the deterministic approximation is then evaluated by comparing it with the results of a Montecarlo simulation of the stochastic model. The comparison shows a good accuracy of the deterministic approximation, with a reduction of the computational times of two orders of magnitude.

The paper is organized as follows. In Section 2 a relevant literature is recalled. Section 3 presents the stochastic model of the mpTSP$_s$ and Section 4 derives its deterministic approximation. In Section 5 we compare the results of the deterministic approximation with the results of a Montecarlo simulation of the stochastic problem. Finally, in Section 6 conclusions are drawn.

2 Literature review

While different stochastic and/or dynamic variants of TSP (and more in general of vehicle routing problem) are present in the literature, the mpTSP$_s$ is absent (Gendreau et al., 1996; Golden et al., 2008). For this reason, we will consider some relevant literature on similar problems, highlighting the main differences with the problem faced in this paper.

In the literature several stochastic TSP problems can be found. In these problems a known distribution affecting some problem parameters is given and the theoretical results are strongly connected with the hypotheses on such distribution. The main sources of uncertainty are related to the arc costs (Leipala, 1978; Toriello et al., 2012)and the location and the subset of cities to be visited (Jaillet, 1988; Goemans and Bertsimas, 1991).

If we consider general routing problems, different types of uncertainty and dynamics can be considered. The most studied variants are related to the online arrival of the customers, with the requests being both goods (Hvattum et al., 2006, 2007; Ichoua et al., 2006; Mitrović-Minić and Laporte, 2004) and services

(Beaudry et al., 2010; Bertsimas and Van Ryzin, 1991; Gendreau et al., 1999; Larsen et al., 2004). Only in recent years the dynamics related to the travel time has been considered in the literature (Chen et al., 2006; Fleischmann et al., 2004; Güner et al., 2012; Kenyon and Morton, 2003; Tagmouti et al., 2011; Taniguchi and Shimamoto, 2004), while, to the best of our knowledge, service time has not been explicitly studied. The last variants of vehicle routing problems are related to the dynamically revealed demands of a known set of customers (Novoa and Storer, 2009; Secomandi and Margot, 2009) and the vehicle availability (Li et al., 2009a,b; Mu et al., 2011). For a recent review, the reader can refer to (Pillac et al., 2013).

All the papers presented in this survey deal with uncertainty and/or dynamic aspects of the routing problems where the magnitude of the uncertainty is limited and the values of the parameters are revealed in a time interval compatible with the operations optimization. Then, even if multiple paths can be present between two given nodes, the multi-path aspects can be ignored, being possible an a priori choice of the path connecting the two points. In our case, the mpTSP$_s$ is thought to be used for the planning of a service. Thus, the enlarged time horizon as well as strong dynamic changes in the travel times due to traffic congestion and other nuisances typical of the urban transportation force the presence, in the decision problem, of multiple paths connecting every pair of nodes, each one with its stochastic cost. This is, to our knowledge, an aspect of the transportation literature considered only in transshipment problems, where the routing aspect is heavily relaxed (Baldi et al., 2012; Tadei et al., 2012).

3 The mpTSP$_s$

Given any pair of nodes, we consider the set of paths between the two nodes. Each path is characterized by a travel cost which is composed by a deterministic travel cost plus a random term, which represents the travel cost oscillation due to the path congestion. In practice, such travel cost oscillations randomly depend on some time scenarios and are actually very difficult to be measured. This implies that the probability distribution of these random terms must be assumed as unknown.

Let it be

- N: set of nodes
- L: set of time scenarios
- K_{ij}: set of paths between nodes i and j
- c_{ij}^k: unit deterministic travel cost of path $k \in K_{ij}$

- $\tilde{\theta}_{ij}^{kl}$: random travel cost oscillation of path $k \in K_{ij}$ under time scenario $l \in L$
- $\tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{kl}) = c_{ij}^{k} + \tilde{\theta}_{ij}^{kl}$: unit random travel cost of path $k \in K_{ij}$ under time scenario l
- x_{ij}^{k}: boolean variable equal to 1 if path $k \in K_{ij}$ is selected, 0 otherwise.
- y_{ij}: boolean variable equal to 1 if node j is visited just after node i, 0 otherwise.

The mpTSP$_s$ is formulated as follows

$$\min_{\{y,x\}} \mathbb{E}_{\{\tilde{\theta}_{ij}^{kl}\}} \left[\sum_{i\in N} \sum_{j\in N} y_{ij} \sum_{k\in K_{ij}} \sum_{l\in L} \tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{kl}) x_{ij}^{k} \right] \tag{1}$$

subject to

$$\sum_{j\in N: j\neq i} y_{ij} = 1 \qquad i \in N \tag{2}$$

$$\sum_{i\in N: i\neq j} y_{ij} = 1 \qquad j \in N \tag{3}$$

$$\sum_{i\in U} \sum_{j\notin U} y_{ij} \geq 1 \qquad \forall U \subset N \tag{4}$$

$$x_{ij}^{k} \in \{0,1\} \qquad k \in K_{ij}, \quad i \in N, \quad j \in N \tag{5}$$

$$y_{ij} \in \{0,1\} \qquad i \in N, \quad j \in N \tag{6}$$

The objective function (1) expresses the minimization of the expected total travel cost; (2) and (3) are the standard assignment constraints; (4) are the subtour elimination constraints. Finally, (4)-(5) are the integrality constraints.

Let us assume that the paths are enough disjoint each other to consider $\tilde{\theta}_{ij}^{kl}$ as independent and identically distributed (i.i.d.) random variables. The common and unknown probability distribution of these random variables is given by the following cumulative right distribution function

$$F(x) = Pr\{\tilde{\theta}_{ij}^{kl} \geq x\} \tag{7}$$

Following (Tadei et al., 2012), we define $\tilde{\theta}_{ij}^{k}$ as the minimum of the random travel cost oscillations $\tilde{\theta}_{ij}^{kl}$ of path $k \in K_{ij}$ among the alternative time scenarios $l \in L$

$$\tilde{\theta}_{ij}^{k} = \min_{l\in L} \tilde{\theta}_{ij}^{kl} \qquad k \in K_{ij}, \quad i \in N, \quad j \in N \tag{8}$$

Because $F(x)$ is unknown, $\tilde{\theta}_{ij}^{k}$ are still of course random variables with a common unknown probability distribution given by

$$B(x) = \Pr\left\{\tilde{\theta}_{ij}^{k} \geq x\right\} \tag{9}$$

As, for any path $k \in K_{ij}$, $\tilde{\theta}_{ij}^{k} \geq x \Longleftrightarrow \tilde{\theta}_{ij}^{kl} \geq x,\ l \in L$ and $\tilde{\theta}_{ij}^{kl}$ are independent, using (7) one gets

$$B(x) = \prod_{l \in L} \Pr\left\{\tilde{\theta}_{ij}^{kl} \geq x\right\} = \prod_{l \in L} F(x) = [F(x)]^{|L|} \tag{10}$$

We assume that the routing is efficiency-based so that among the alternative time scenarios $l \in L$ the one which minimizes the random travel cost $\tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{kl})$ will be selected.

Then, the random travel cost of path $k \in K_{ij}$ becomes

$$\tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{k}) = \min_{l \in L} \tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{kl}) = c_{ij}^{k} + \min_{l \in L} \tilde{\theta}_{ij}^{kl} = c_{ij}^{k} + \tilde{\theta}_{ij}^{k} \qquad k \in K_{ij}, \quad i \in N, \quad j \in N \tag{11}$$

The minimum travel cost oscillation $\tilde{\theta}_{ij}^{k}$ can be either positive or negative, but, in practice, its absolute value does not overcome the travel cost c_{ij}^{k}, so that $\tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{k})$ is always non negative.

For each pair of node (i, j), let us consider the path k^{*} (for the sake of simplicity, we assume it is unique) which gives the minimum random travel cost.

The minimum random travel cost between i and j is then

$$\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^{*}}) = \min_{k \in K_{ij}} \tilde{c}_{ij}^{k}(\tilde{\theta}_{ij}^{k}) \qquad i \in N, \quad j \in N \tag{12}$$

and the optimal variables $\left\{x_{ij}^{k}\right\}$ of problem (1)-(6) become

$$x_{ij}^{k} = \begin{cases} 1, & \text{if} \quad k = k^{*} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

Using (12), (13), and the linearity of the expected value operator $\mathbb{E}$, the objective function (1) becomes

$$\min_{\{y\}} \mathbb{E}_{\{\tilde{\theta}_{ij}^{k^{*}}\}} \left[\sum_{i \in N} \sum_{j \in N} y_{ij} \tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^{*}})\right] = \min_{\{y\}} \sum_{i \in N} \sum_{j \in N} y_{ij} \mathbb{E}_{\{\tilde{\theta}_{ij}^{k^{*}}\}} \left[\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^{*}})\right] = \min_{\{y\}} \sum_{i \in N} \sum_{j \in N} y_{ij} \hat{c}_{ij} \tag{14}$$

where

$$\hat{c}_{ij} = \mathbb{E}_{\{\tilde{\theta}_{ij}^{k^*}\}}\left[\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*})\right] \qquad i \in N, \quad j \in N \tag{15}$$

The mpTSP$_s$ then becomes

$$\min_{\{y\}} \sum_{i \in N} \sum_{j \in N} y_{ij} \hat{c}_{ij} \tag{16}$$

subject to (2)-(6).

However, the calculation of $\hat{c}_{ij}$ in (16) requires to know the probability distribution of the minimum random travel cost between i and j, i.e. $\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*})$, which will be derived in the next section.

4 The deterministic approximation of the mpTSP$_s$

By (11) and (12), let

$$G_{ij}(x) = Pr\left\{\tilde{c}_{ij}(\tilde{\theta}_{ij}^{k^*}) \geq x\right\} = Pr\left\{\min_{k \in K_{ij}} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x\right\} \qquad i \in N, \quad j \in N \tag{17}$$

be the cumulative right distribution function of the minimum random travel cost between i and j.

As, for any pair of nodes (i, j), $\min_{k \in K_{ij}} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x \iff \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x, \quad k \in K_{ij}$, and the random variables $\tilde{\theta}_{ij}^k$ are independent (because $\tilde{\theta}_{ij}^{kl}$ are independent), due to (9) and (10), $G_{ij}\{x\}$ in (17) becomes a function of the total number $|L|$ of time scenarios as follows

$$\begin{aligned}
G_{ij}(x, |L|) \quad &= Pr\left\{\min_{k \in K_{ij}} \tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x\right\} = \prod_{k \in K_{ij}} Pr\left\{\tilde{c}_{ij}^k(\tilde{\theta}_{ij}^k) \geq x\right\} \\
&= \prod_{k \in K_{ij}} Pr\left\{\tilde{\theta}_{ij}^k \geq x - c_{ij}^k\right\} = \prod_{k \in K_{ij}} B\left(x - c_{ij}^k\right) \\
&= \prod_{k \in K_{ij}} \left[F\left(x - c_{ij}^k\right)\right]^{|L|} \qquad i \in N, \quad j \in N
\end{aligned} \tag{18}$$

First, let us consider the following aspect: the optimal solution of problem (1)-(6) does not change if any arbitrary constant is added or subtracted to the random variables $\tilde{\theta}_{ij}^{kl}$.

Let us choose this constant as the root a of the equation

$$1 - F(a) = 1/|L| \tag{19}$$

Let us assume that $|L|$ is large enough to use the asymptotic approximation $\lim_{|L|\to+\infty} G_{ij}(x,|L|)$ as a good approximation of $G_{ij}(x)$, i.e.

$$G_{ij}(x) = \lim_{|L|\to+\infty} G_{ij}(x,|L|)) \qquad i \in N, \quad j \in N \tag{20}$$

The calculation of the limit in (20) would require to know the probability distribution $F(.)$ in (7), which is unknown. From (Tadei et al., 2012), we know that under a mild assumption on the shape of the unknown probability distribution $F(.)$ (i.e. it is asymptotically exponential in its left tail), the limit in (20) tends towards the following Gumbel Gumbel (1958) probability distribution

$$G_{ij}(x) = \lim_{|L|\to+\infty} G_{ij}(x,|L|)) = \exp\left(-A_{ij}e^{\beta x}\right) \qquad i \in N, \quad j \in N \tag{21}$$

where $\beta > 0$ is a parameter to be calibrated and

$$A_{ij} = \sum_{k\in K_{ij}} e^{-\beta c_{ij}^k} \qquad i \in N, \quad j \in N \tag{22}$$

is the *accessibility*, in the sense of Hansen (Hansen, 1959), of the pair of nodes i, j to the set of paths between i and j.

Using the probability distribution $G_{ij}(x)$ given by (21), after some manipulations, $\hat{c}_{ij}$ in (15) becomes

$$\hat{c}_{ij} = -\int_{-\infty}^{+\infty} x dG_{ij}(x) = \int_{-\infty}^{+\infty} x \exp\left(-A_{ij}e^{\beta x}\right) A_{ij}e^{\beta x}\beta dx = -\frac{1}{\beta}(\ln A_{ij}+\gamma) \qquad i \in N, \quad j \in N \tag{23}$$

where $\gamma \simeq 0.5772$ is the Euler constant.

By (23) and up to the constant $-\frac{\gamma}{\beta}\sum_{i\in N}\sum_{j\in N} y_{ij} = -\frac{\gamma}{\beta}|N|$, (16) becomes

$$\begin{aligned}
&\min_{\{y\}} \sum_{i\in N}\sum_{j\in N} -\frac{1}{\beta} y_{ij} \ln A_{ij} = \\
&= \frac{1}{\beta} \max_{\{y\}} \sum_{i\in N}\sum_{j\in N} \ln A_{ij}^{y_{ij}} = \\
&= \frac{1}{\beta} \max_{\{y\}} \ln \prod_{i\in N}\prod_{j\in N} A_{ij}^{y_{ij}} = \\
&= \frac{1}{\beta} \max_{\{y\}} \ln \Phi
\end{aligned} \tag{24}$$

subject to (2)-(6), where $\Phi = \prod_{i\in N}\prod_{j\in N} A_{ij}^{y_{ij}}$ is the *total accessibility* of the set of arcs of an optimal Hamiltonian tour to the global set of paths.

From (24), it is interesting to observe that the expected minimum total travel cost is equivalent, but the constant $\frac{1}{\beta}$, to the maximum of the logarithm of the total accessibility.

5 Computational results

In this section, we present and analyze the results of the computational experiments. The goal is to evaluate the effectiveness of the deterministic approximation of the mpTSP$_s$ we derived.

We do that by comparing our deterministic approximation with a Montecarlo simulation performed on the stochastic problem. The Montecarlo simulation is implemented in C++, with the underlying TSP instances solved by means of the Concorde TSP solver (Applegate et al., 2007; Cook, 2012). Experiments were performed on an Intel I7 2 GHz workstation with 8 GB of RAM.

Section 5.1 introduces the instance sets. The details of the Montecarlo simulation are presented in Section 5.2. The calibration of the parameter involved in the deterministic approximation of the mpTSP$_s$ is described in Section 5.3, whilst the comparison between the Montecarlo simulation and the approximated results is given in Section 5.4.

5.1 Instance sets

No real-life instances are present in the literature for this stochastic version of the TSP problem. Then, we generated instances, partially based on those available in the TSPLIB (Reinelt, 1991) for the deterministic TSP problem. According to the literature, we generated the stochastic costs according to the guidelines presented in (Kenyon and Morton, 2003):

- Instances. In order to limit the computational time, which is mainly due to the Montecarlo simulation, we considered all instances with a number of nodes up to 200 in the TSP Library set. In particular, we split those instances into two sets: 11 instances with up to 100 nodes (set S100) and 15 instances with number of nodes between 101 and 200 (set S200).

- Nodes. The nodes and their position on the plane are the same of the original TSP instances.

- Multiple paths. The number of paths is set to 1, 3, and 5. Although the mpTSP$_s$ hypothesizes that several paths are present between any pair of nodes, we decided to also test the case where only one path is available. In fact, it is interesting to observe the behavior of the approximation in an extreme situation where the aspect characterizing the problem is just the stochasticity of the travel times on a single path.

- Path costs. The cost c_{ij}^k associated to each path k between nodes i and j is considered as a function of the Euclidean distance between i and j. In details, this cost has been drawn from $U(EC_{ij}, 3EC_{ij})$, where EC_{ij} is the Euclidean distance between i and j and U is the uniform distribution. The random travel cost oscillations θ_{ij}^k have been drawn as $\mathcal{D}(-c_{ij}^k/2, 2c_{ij}^k)$, where $\mathcal{D}$ is a probability distribution with its support limited to 50% and 100% of the corresponding deterministic cost, such that $c_{ij}^k + \theta_{ij}^k \geq EC_{ij}$. For $\mathcal{D}$ we have considered both the Uniform and the Gumbel distribution.

5.2 Montecarlo simulation

In order to evaluate the stochastic objective function of our problem, we used a Montecarlo simulation. Our Montecarlo simulation repeats I times the following overall process:

- Create S scenarios with the random costs θ_{ij}^k generated as described in 5.1.
- Solve each scenario as follows. Build a TSP with the node set equal to the node set of the stochastic problem. Set the cost c_{ij} between nodes i and j as $c_{ij} = \min_k(c_{ij}^k + \theta_{ij}^k)$. Indeed, when a cost scenario becomes known, its optimal solution is obtained by using, as path between the two nodes, the path with the minimum random travel cost. The scenarios are solved to optimality by means of the Concorde TSP solver.
- Given the scenario optima, compute the expected value of the total cost.
- Compute the distribution of the expected value of the total cost for the scenario-based simulations.

In order to obtain the most reliable results of the Montecarlo simulation, we performed a set of tuning testbeds by using a subset of instances (5 from S100 and 5 from S200). The values for the parameters I (number of repetitions) and S (number of scenarios) have been set such that the standard deviation of the distribution of the expected value was less than 1% of its mean. These values were $I = 10$ and $S = 100$.

5.3 Calibration of the β parameter

The deterministic approximation of the mpTSP$_s$ requires, see (24), an appropriate value of the parameter β. This parameter describes the propensity of the model to choose among the set of the paths characterized by different random travel costs.

β is obtained by calibration as follows. Let us consider the standard Gumbel distribution $G(x) = \exp(e^{-x})$. If an approximation error of 2‰ is accepted, then $G(x) = 1 \Leftrightarrow x = 6.08$ and $G(x) = 0 \Leftrightarrow x = -1.76$. Let us consider the distribution range $[m, M]$. The following equations hold

$$\beta(m - \zeta) = -1.76 \tag{25}$$

$$\beta(M - \zeta) = 6.08 \tag{26}$$

where ζ is the mode of the Gumbel distribution $G(x) = \exp\left(e^{-\beta(x-\zeta)}\right)$.

By subtracting (25) from (26) one gets for β the value

$$\beta = \frac{6.08 - (-1.76)}{M - m} = \frac{7.84}{M - m} \tag{27}$$

According to our random oscillations rule, m is set equal to $\min_{i,j} EC_{ij}$. In order to calculate M we need to know the order of magnitude of the travel cost oscillations in the final solution. This is needed to avoid considering those arcs with travel costs very far from the travel costs in the solution, which could lead us to overestimate M. In fact, the presence of arcs with a travel cost much greater of the mean travel cost is a quite common situation in the TSP and VRP problems.

M has been calculated as follows

- Solve a TSP instance with the same node set of the stochastic problem and the cost of each arc determined as $c_{ij} = \min_k c_{ij}^k$. Let us call C_D the optimum of this deterministic instance.
- Set $m = \min_{i,j} EC_{ij}$ and $M = \frac{|N|}{2KC_D}$, where $|N|$ is the number of nodes and K is the number of paths. The rationale of the formula for calculating M is that $C_D/|N|$ gives us the order of magnitude of the mean deterministic cost, which, given the rules we used to generate the instances, can have a maximum oscillation of 100%. The number of paths K is used for normalizing the accessibility effect when the path cardinality increases.

More sophisticated methods to calibrate β can be found in (Galambos et al., 1994).

5.4 Comparison of Montecarlo simulation and deterministic approximation results

Here we summarize the results for all instances with different combinations of the parameters. The performance, in terms of percentage gap, is defined as the relative percentage error of the approximated optimum when compared to the mean of the expected value distribution given by the Montecarlo simulation.

Table 1 reports the percentage gap for all combinations of the parameters, while varying the probability distribution (either Uniform or Gumbel). The first two columns display the instance set and the number of paths between any pair of nodes, while Columns 3-4 report the mean of the percentage gaps. The best mean values are obtained for the Gumbel distribution. For both distributions the best results are obtained with one path between the nodes, with a gap of less than 1% for the Gumbel distribution. This gap increases with the number of paths. The quality of the approximation seems to be inversely correlated with the number of nodes. However, the percentage gap is, in all cases, quite limited, with a worst case of 7.77% for the Uniform and 4.46% for the Gumbel distribution.

Computational times, expressed in seconds, are reported in Table 2. Notice that, being the computational time in both cases (Montecarlo and deterministic approximation) mainly given by the TSP instances computational time and being the number of the TSP instances independent from the number of multiple paths, also the computational times are independent from the number of multiple paths. Thus, the results are summarized by considering the instance set only. The Montecarlo simulation needs a computational time of about 2 orders of magnitude greater than the deterministic approximation. This makes the deterministic approximation more and more interesting when applied to large instances, where the Montecarlo simulation becomes impracticable.

Table 1: Percentage gap between the deterministic approximation and the Montecarlo simulation

Set	**Path**	**Uniform**	**Gumbel**
S100	1	1.32	0.62
	3	3.41	1.86
	5	4.01	2.22
Avg		**2.91**	**1.57**
S200	1	0.71	0.35
	3	7.46	3.13
	5	7.77	4.46
Avg		**5.31**	**2.64**
Global avg		**4.30**	**2.19**

In conclusion, the results seem very promising. The deterministic approximation performs quite well for all types of instances and distributions and guarantees a good accuracy. The best performance is obtained when the random travel costs have a Gumbel distribution, that is usually the case for real travel cost random

Table 2: Computational times in seconds of the Montecarlo simulation and the deterministic approximation

Set	**Montecarlo**	**Approx**
S100	523.40	5.52
S200	1507.71	14.54
Global avg	1091.27	10.72

oscillations.

6 Conclusions

In this paper we have addressed the multi-path Traveling Salesman Problem with stochastic travel costs, which consists in finding an expected minimum Hamiltonian tour connecting all nodes, when each pair of nodes is connected by several paths and each path shows a stochastic travel cost with unknown distribution.

From a theoretical perspective, the paper shows that, under a mild assumption, the probability distribution of the minimum random travel cost between any pair of nodes becomes a Gumbel distribution. Moreover, the expected minimum total travel cost is proportional to the maximum of the logarithm of the total accessibility of the Hamiltonian tours to the path set.

The deterministic approximation of the stochastic model provides very promising results on a large set of instances in negligible computational times.

In conclusion, the performance of the methodology proposed is particularly good when the probability distribution of the random travel costs of the stochastic model is a Gumbel distribution, even if good results are also provided with the Uniform distribution. This feature makes our deterministic approximation a good predictive tool for addressing stochastic travel costs in multi-path networks.

7 Acknowledgements

Partial funding was provided by the Ministero dell'Istruzione, Università e Ricerca (MIUR) (Italian Ministry of University and Research), under the 2009 PRIN Project "Methods and Algorithms for the Logistics Optimization" and by the Natural Sciences and Engineering Council of Canada (NSERC), through its Industrial Research Chair and Discovery Grants programs.

References

D.L. Applegate, R.E. Bixby, V. Chvtal, W.J. Cook, The Traveling Salesman Problem: A Computational Study, Princeton University Press, Princeton, NJ, USA, 2007.

M.M. Baldi, M. Ghirardi, G. Perboli, R. Tadei, The capacitated transshipment location problem under uncertainty: A computational study, Procedia - Social and Behavioral Sciences 39 (2012) 424–436.

A. Beaudry, G. Laporte, T. Melo, S. Nickel, Dynamic transportation of patients in hospitals, OR Spectrum 32 (2010) 77–107.

D. Bertsimas, G. Van Ryzin, A stochastic and dynamic vehicle-routing problem in the euclidean plane, Operations Research 39 (1991) 601–615.

H.K. Chen, C.F. Hsueh, M.S. Chang, The real-time time-dependent vehicle routing problem, Transportation Research Part E: Logistics and Transportation Review 42 (2006) 383–408.

H. Chourabi, T. Nam, S. Walker, J.R. Gil-Garcia, S. Mellouli, K. Nahon, T.A. Pardo, H.J. Scholl, Understanding smart cities: An integrative framework, in: Proceedings of the 2012 45th Hawaii International Conference on System Sciences, HICSS '12, IEEE Computer Society, Washington, DC, USA, 2012, pp. 2289–2297.

CITYLOG Consortium, CITYLOG European project, 2010. URL: `http://http://www.city-log.eu/`.

W.J. Cook, In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation, Princeton University Press, Princeton, NJ, USA, 2012.

B. Fleischmann, S. Gnutzmann, E. Sandvoss, Dynamic vehicle routing based on online traffic information, Transportation Science 38 (2004) 420–433.

J. Galambos, J. Lechner, E. Simiu, Extreme Value Theory and Applications, Kluwer, 1994.

M. Gendreau, F. Guertin, J.Y. Potvin, E. Taillard, Parallel tabu search for real-time vehicle routing and dispatching, Transportation Science 33 (1999) 381–390.

M. Gendreau, G. Laporte, R. Sguin, Stochastic vehicle routing, European Journal of Operational Research 88 (1996) 3–12.

M.X. Goemans, D. Bertsimas, Probabilistic analysis of the Held and Karp lower bound for the Euclidean traveling salesman problem, Mathematics of Operations Research 16 (1991) 72–89.

B.L. Golden, S. Raghavan, E.A. Wasil, The Vehicle Routing Problem: Latest Advances and New Challenges, Springer, 2008.

E.J. Gumbel, Statistics of Extremes, Columbia University Press, 1958.

A.R. Güner, A. Murat, R.B. Chinnam, Dynamic routing under recurrent and non-recurrent congestion using real-time ITS information, Computers & Operations Research 39 (2012) 358–373.

W. Hansen, How accessibility shapes land use, Journal of the American Institute of Planners 25 (1959) 73–76.

L.M. Hvattum, A. Lokketangen, G. Laporte, Solving a dynamic and stochastic vehicle routing problem with a sample scenario hedging heuristic, Transportation Science 40 (2006) 421–438.

L.M. Hvattum, A. Lokketangen, G. Laporte, A branch-and-regret heuristic for stochastic and dynamic vehicle routing problems, Networks 49 (2007) 330–340.

S. Ichoua, M. Gendreau, J.Y. Potvin, Exploiting knowledge about future demands for real-time vehicle dispatching, Transportation Science 40 (2006) 211–225.

P. Jaillet, A priori solution of a traveling salesman problem in which a random subset of the customers are visited, Operations Research 36 (1988) 929–936.

A.S. Kenyon, D.P. Morton, Stochastic vehicle routing with random travel times, Transportation Science 37 (2003) 69–82.

A. Larsen, O.B.G. Madsen, M.M. Solomon, The a priori dynamic traveling salesman problem with time windows, Transportation Science 38 (2004) 459–472.

T. Leipala, On the solutions of stochastic traveling salesman problems, European Journal of Operational Research 2 (1978) 291–297.

J.Q. Li, P.B. Mirchandani, D. Borenstein, A Lagrangian heuristic for the real-time vehicle rescheduling problem, Transportation Research Part E: Logistics and Transportation Review 45 (2009a) 419–433.

J.Q. Li, P.B. Mirchandani, D. Borenstein, Real-time vehicle rerouting problems with time windows, European Journal of Operational Research 194 (2009b) 711–727.

S. Mitrović-Minić, G. Laporte, Waiting strategies for the dynamic pickup and delivery problem with time windows, Transportation Research Part B: Methodological 38 (2004) 635–655.

Q. Mu, Z. Fu, J. Lysgaard, R. Eglese, Disruption management of the vehicle routing problem with vehicle breakdown, Journal of the Operational Research Society 62 (2011) 742–749.

C. Novoa, R. Storer, An approximate dynamic programming approach for the vehicle routing problem with stochastic demands, European Journal of Operational Research 196 (2009) 509–515.

G. Perboli, R. Tadei, M.M. Baldi, The stochastic generalized bin packing problem, Discrete Applied Mathematics 160 (2012) 1291–1297.

G. Perboli, R. Tadei, D. Vigo, The two-echelon capacitated vehicle routing problem: Models and math-based heuristics, Transportation Science 45 (2011) 364–380.

V. Pillac, M. Gendreau, C. Guret, A.L. Medaglia, A review of dynamic vehicle routing problems, European Journal of Operational Research 225 (2013) 1–11.

G. Reinelt, TSPLIB- a traveling salesman problem library, ORSA Journal of Computing 3 (1991) 376–384.

N. Secomandi, F. Margot, Reoptimization approaches for the vehicle routing problem with stochastic demands, Operations Research 57 (2009) 214–230.

R. Tadei, G. Perboli, N. Ricciardi, M.M. Baldi, The capacitated transshipment location problem with stochastic handling utilities at the facilities, International Transactions in Operational Research 19 (2012) 789–807.

M. Tagmouti, M. Gendreau, J.Y. Potvin, A dynamic capacitated arc routing problem with time-dependent service costs, Transportation Research Part C: Emerging Technologies 19 (2011) 20–28.

E. Taniguchi, H. Shimamoto, Intelligent transportation system based dynamic vehicle routing and scheduling with variable travel times, Transportation Research Part C: Emerging Technologies 12 (2004) 235–250.

A. Toriello, W.B. Haskell, M. Poremba, A Dynamic Traveling Salesman Problem with Stochastic Arc Costs, Technical Report, Optimization Online, 2012. URL: http://www.optimization-online.org/DB_FILE/2012/09/3593.pdf.