J Optim Theory Appl manuscript No.
(will be inserted by the editor)

Solution Approaches for the Stochastic Capacitated Traveling Salesmen Location Problem with Recourse

Luca Bertazzi · Francesca Maggioni

Received: date / Accepted: date

Abstract A facility has to be located in a given area to serve a given number of customers. The position of the customers is not known. The service to the customers is carried out by several traveling salesmen. Each of them has a capacity in terms of the maximum number of customers that can be served in any tour. The aim is to determine the *service zone* (in a shape of a circle) that minimizes the expected cost of the traveled routes. The centre of the circle is the location of the facility. Once the position of the customers is revealed, the customers located outside the service zone are served with a recourse action at a greater unit cost. For this problem, we compare the performance of two different solution approaches. The first is based on a heuristic proposed for the *Capacitated Traveling Salesman Problem* and the second on the optimal solution of a stochastic second–order cone formulation with an approximate objective function.

L. Bertazzi
Department of Economics and Management, University of Brescia, Contrada S. Chiara 50, 25122 Brescia ITALY

F. Maggioni (corresponding author)
Department of Management, Economics and Quantitative Methods, University of Bergamo, Via dei Caniana 2, 24127 Bergamo ITALY E-mail: francesca.maggioni@unibg.it

Keywords Facility Location · Traveling Salesmen Location Problem · Stochastic Second–Order Cone

Mathematics Subject Classification (2000) 90C15 · 65K05 · 90B05 · 90B06

1 Introduction

Facility location problems play an important role in the strategic design of logistic networks. We refer to [1–3], for recent surveys on facility location problems and to [4, 5] for surveys on facility location problems under uncertainty.

A relatively new class of problems, in which the classical facility location and the vehicle routing problems are integrated, is called *Location-Routing Problem* (LRP). We refer to [6] for a recent survey on vehicle routing problems. Although it is well known that facility location and routing are often interrelated, LRP received little attention in the past. However, in the last years, several papers about deterministic location–routing problems have been published. Instead, a few papers have been devoted to location–routing problems with stochastic demand.

The simplest location–routing problem with stochastic demand is the *Traveling Salesman Location Problem* (TSLP). A set of customers is served by a single facility. At each time, only a subset of customers has to be served. A TSP is built to serve this subset of customers. The aim is to determine where to locate the facility in order to minimize the expected cost of the TSP. This problem has several applications in many service systems, such as delivery services, customer pickup services, repair vehicles. It has been introduced by [7] and studied by [8–11]. It has been extended to the case, referred to as the *Capacitated Traveling Salesmen Location Problem*, with several capacitated salesmen, by [12]. The main difficulty of these problems is that there exist

$2^n - 1$ different subsets of n customers that can require to be served and therefore an exponential number of TSP has to be solved. A polynomial time heuristic algorithm, with known relative worst-case error, has been proposed in [12]. A different case is given by the *Probabilistic Traveling Salesman Location Problem* (PTSLP). In this case, first an a priori TSP visiting all customers is computed. Then, for each time, the TSP to serve the subset of customers that have to be visited is obtained by just skipping in the a priori tour the customers that have not to be served at that time. The aim is to simultaneously determine where to locate the facility, and the a priori TSP that minimize the expected cost. This problem has been introduced by [8] and has been studied by [13]. Recent papers are [14] and [15]. For a vehicle routing and districting problem with stochastic customers see [16].

We study a problem in which a single facility (typically a service station) has to be located in a given area. This facility is used to serve a given number of customers. The position of the customers is not known. The service to the customers is carried out by several traveling salesmen. Each of them has a capacity in terms of the maximum number of customers that can be served in any tour. The aim is to determine the *service zone* (in a shape of a circle) that minimizes the expected cost of the traveled routes. The centre of the circle is the location of the facility. Once the position of the customers is revealed, the customers located outside the service zone are served with a recourse action at a greater unit cost. We refer this problem to as the *Stochastic Capacitated Traveling Salesmen Location Problem with Recourse* (SCTSLP–R). The cost to select the service zone is chosen to be proportional to the cost of the corresponding *Capacitated Vehicle Routing Problem* (CVRP) estimated by using the formula inspired to [17] and to [18].

We propose two different solution approaches. The first is based on the heuristic proposed by Simchi–Levi (1991) for the *Capacitated Traveling Salesmen Problem* (see [12]). We prove that, in the worst case, this algorithm can give a solution infinitely worse with respect to the optimal one. The second approach is obtained by optimally solving a two–stage stochastic optimization model (see [19–21]) with an approximate objective function. In particular, we first model the problem as a two-stage stochastic semidefinite programming problem (see [22–25]), and then as a two-stage stochastic second-order cone programming problem (see [26, 27]), in which the first stage decisions are the facility location and the radius of the service zone. The input data of problems affected by uncertainty are usually described by stochastic processes that can be represented using discrete random variables. With the term realization, we denote all possible discrete values that the random variables can assume. In our problem, each realization is represented by a set of ellipses. Each ellipse represents the area covered by a traveling salesman to supply its customers, randomly generated by uniform and normal distributions in a neighborhood of its starting position inside a given region. The costs in the objective function are assumed to be proportional to the area of the service zone in the first stage, and proportional to the difference between the area covered on the basis of the recourse and the area of the service zone in the second stage. Computational results show that the solutions obtained by this second approach significantly dominate the one obtained by the first.

The paper is organized as follows. In Section 2, the problem we study is formally described and formulated. In Section 3, the approach based on the heuristic proposed by Simchi–Levi (1991) [12] for the *Capacitated Traveling Salesmen Problem* is described and analyzed from the worst–case point of view. In Section 4, the approach based on the optimal solution of a second–order cone formulation of the problem with an approximate objective function is proposed. Finally, in Section 5, the two approaches are evaluated and compared on the basis of randomly generated problem instances.

2 Problem Description

A facility has to serve R customers in a given square $\mathscr{A}$ with side length L. The position of the customers is not known. A set $\mathcal{N} = \{1, 2, \ldots, N\}$ of traveling salesmen is available to serve the customers. Each traveling salesman $i \in \mathcal{N}$ has a given initial position and a given capacity, in terms of number of customers that can be served in a tour, q. We define a *realization* of the problem, the fulfillment of customers positions. Each realization is modeled by a given number of ellipses. Each ellipse corresponds to the area covered by a traveling salesman. In particular, let $\mathcal{K} = \{1, 2, \ldots, K\}$ be the set of realizations. Each realization $k \in \mathcal{K}$ is composed of N random ellipses and has probability p_k. The ellipse E_k^i corresponds to the area covered by the traveling salesman i at realization k and it covers q customers. Therefore, the total number of customers R is KNq. Note that at each realization the same number of customers Nq is served. The difference among the realizations is in the customers' positions. The described approach corresponds to model distributed demands with different degree of density since we generate the same number of customers inside areas of different extension.

The problem is to determine the service zone that minimizes the total expected cost.

The *service zone* is defined as a circle C, with centre $\tilde{\mathbf{u}}$ and radius r:

$$C := \{\mathbf{u} \in \mathbb{R}^2 : \mathbf{u}^T\mathbf{u} - 2\tilde{\mathbf{u}}^T\mathbf{u} + \gamma \leq 0\}, \tag{1}$$

where the centre $\tilde{\mathbf{u}} \in \mathbb{R}^2$ (with $\mathbb{R}^2$ the space of 2-dimensional real vectors) and γ are decision variables of the problem. The corresponding radius r is $\sqrt{\tilde{\mathbf{u}}^T\tilde{\mathbf{u}} - \gamma}$ and is not greater than $r_{max} = \sqrt{2}L$. The cost to select the circle C is α times the cost of the corresponding *Capacitated Vehicle Routing Problem* (CVRP), estimated by using the formula inspired to [17] and [18]:

$$D_1(r) := 2\bar{d}\frac{R\frac{r}{r_{max}}}{q} + 0.57\sqrt{AR\frac{r}{r_{max}}} = r\left(\frac{R\frac{r}{r_{max}}}{q} + 0.57\sqrt{\pi R\frac{r}{r_{max}}}\right), \tag{2}$$

where $\bar{d} = \frac{r}{2}$ is the average distance between the customers and the facility, $R\frac{r}{r_{max}}$ is an estimate of the number of customers expected in the service zone C and $A = \pi r^2$ is the area of the service zone C. This formula reduces to $D_1(r) = r\left(\frac{R}{q} + 0.57\sqrt{\pi R}\right)$ whenever the service zone is able to cover all possible customers. An example is when $\tilde{\mathbf{u}}$ is equal to the centre of the square $\mathscr{A}$ and $r = \frac{\sqrt{2}L}{2}$.

In any realization k, if all customers are covered by C, no further action is needed. Otherwise, a recourse action is needed to enlarge the service zone C to the new circle

$$C_k := \{\mathbf{u} \in \mathbb{R}^2 : \mathbf{u}^T\mathbf{u} - 2\tilde{\mathbf{u}}^T\mathbf{u} + \tilde{\gamma}_k \leq 0\}, \tag{3}$$

with the same centre $\tilde{\mathbf{u}}$ of C and radius $r_k = \sqrt{\tilde{\mathbf{u}}^T\tilde{\mathbf{u}} - \tilde{\gamma}_k} > r$ enough to cover all customers under realization k. The cost of this new circle is β times, with $\beta > \alpha$, the cost of the corresponding CVRP, estimated by

$$D_2(r, r_k) := 2\left(r + \frac{r_k - r}{2}\right)\frac{R(1 - \frac{r}{r_{max}})}{q} + 0.57\sqrt{\pi(r_k^2 - r^2)R\left(1 - \frac{r}{r_{max}}\right)}, \tag{4}$$

where $r + \frac{r_k - r}{2}$ is the average distance between the customers and the facility, $R\left(1 - \frac{r}{r_{max}}\right)$, is an estimate of the number of customers expected outside the service zone C, and $\pi(r_k^2 - r^2)$ is the difference between the area of C_k and the area of the service zone C.

The aim of the problem is to determine $\tilde{\mathbf{u}}$, γ and $\tilde{\gamma}_k$ for all realizations $k \in \mathcal{K}$, such that the expected cost

$$\alpha D_1(r) + \beta \sum_{k \in \mathcal{K}} p_k D_2(r, r_k) \tag{5}$$

is minimized. Due to the non-linear dependence of CVRP costs by the decision variables $\tilde{\mathbf{u}}$, γ and $\tilde{\gamma}_k$, a second-oder cone formulation with linear objective function over the intersection of an affine set and product of second-order cones cannot be formulated for this problem with CVRP costs. For this reason we propose and compare two different solution approaches for the same problem, of increasing refinement.

An alternative policy could be to choice at the first stage only the centre of the circle and at the second stage the radius according to the particular realization of the random variable (customers' positions). In such a case eq. (5) reduces to

$$\beta \sum_{k \in \mathcal{K}} p_k r_k \left(\frac{R}{q} + 0.57\sqrt{\pi R} \right) \ , \tag{6}$$

which is less convenient than the proposed approach since it requires to pay a recourse price β for the whole second stage radius r_k instead of the difference $r_k - r$. However, a second-order cone formulation with linear objective function over the intersection of an affine set and product of second-order cones can be formulated for this problem without approximation of the costs in the objective function, and it will be investigated in a separate paper.

3 The Approach based on Simchi–Levi (1991)

The *Capacitated Traveling Salesmen Problem* is identical to the SCTSLP–R problem, but the position of the R customers is known, each customer has a probability to require a service at each time, the routing cost is paid and no recourse is possible. In this problem, it is optimal to locate the facility at the position of one of the customers.

The second approach we consider, referred to as *SL* approach, is based on the heuristic proposed by [12] for the *Capacitated Traveling Salesmen Problem*, which select the customer at which to locate the facility. Given the location of the facility, the *SL* selects, as service zone, the minimum enclosing circle having centre at the selected location. Let $\mathcal{R}$ be the set of the R customers to serve, h_a be the probability that customer $a \in \mathcal{R}$ requires to be served at a given time. Since h_a is not a data of the SCTSLP–R problem, we assume $h_a = \frac{1}{R}$. The *SL* can be described as follows.

SL

1. For each customer $a \in \mathcal{R}$:
   (a) Order the customers $b \in \mathcal{R}$ in the non–decreasing order of $d(a,b)$. Rename them as $\gamma_1, \gamma_2, \ldots, \gamma_R$. Note that $\gamma_1 = a$.
   (b) Compute

$$\Phi(a) := \frac{1}{q}\sum_{b\in\mathcal{R}} h_b d(a,b) + \left(1 - \frac{1}{q}\right)\left(\sum_{\delta=2}^{R} d(a,\gamma_\delta) h_{\gamma_\delta} \prod_{\tau=1}^{\delta-1}(1 - h_{\gamma_\tau})\right).$$

2. Select the customer $a^* = \arg\min_{a\in\mathcal{R}} \Phi(a)$ as facility location.
3. Compute the radius r of the minimum enclosing circle having centre in a^*.

We now prove worst–case results about the performance of the *SL*. Consider first instances with just one realization k. Let $r^{SL}(k)$ and $z^{SL}(k)$ be the radius and the cost of the solution obtained by applying the *SL*, and $z^*(k)$ be the optimal cost of the SCTSLP-R problem. Recall that $\bar{r}_k$ is the distance between the two customers having maximum distance in the realization k. The following theorem states that the *SL* has a cost at least 73.205% greater than the optimal cost of the SCTSLP-R problem, whenever $r^{SL}(k) = \bar{r}_k$. This happens, for instance, when the the facility is located at one the two customers corresponding to $\bar{r}_k$ in the solution obtained by the *SL*.

Theorem 3.1 *If* $r^{SL}(k) = \bar{r}_k$*, then* $\frac{z^{SL}(k)}{z^*(k)} \geq \sqrt{3} \approx 1.73205.$

Proof Given that $r^{SL}(k) = \bar{r}_k$, the cost of the solution obtained by applying the *SL* is $z^{SL}(k) = \alpha\bar{r}_k\left(\frac{R}{q} + 0.57\sqrt{\pi R}\right)$.

The cost $z^*(k)$ generated by the SCTSLP-R problem is not greater than the one corresponding to the smallest first stage enclosing circle, i.e. the circle at minimum radius covering all customers. Remember that Jung's Theorem states that this circle has a radius not greater than $\frac{\bar{r}_k}{\sqrt{3}}$.
Therefore, $z^*(k) \leq \alpha\frac{\bar{r}_k}{\sqrt{3}}\left(\frac{R}{q} + 0.57\sqrt{\pi R}\right)$ and

$$\frac{z^{SL}(k)}{z^*(k)} \geq \frac{\alpha \bar{r}_k \left(\frac{R}{q} + 0.57\sqrt{\pi R}\right)}{\alpha \frac{\bar{r}}{\sqrt{3}} \left(\frac{R}{q} + 0.57\sqrt{\pi R}\right)} = \sqrt{3}.$$

□

We now show the worst–case performance bound of any realization k such that $r^{SL}(k) = \sqrt{2}L$, i.e. the realization in which the *SL* has maximum cost.

Theorem 3.2 *If* $r^{SL}(k) = \sqrt{2}L$*, then* $\frac{z^{SL}(k)}{z^*(k)} \leq \frac{6\sqrt{3}}{5} \approx 2.07846$ *and the bound is tight.*

Proof Given that $r^{SL}(k) = \sqrt{2}L$, the cost of the solution obtained by applying the *SL* is $z^{SL}(k) = \alpha\sqrt{2}L\left(\frac{R}{q} + 0.57\sqrt{\pi R}\right)$.

Let us now compute a lower bound on the optimal cost $z^*(k)$. Consider first the two extreme cases with $r = 0$ and $r = \frac{\sqrt{2}L}{2}$. In both cases, the cost of the realization k is not smaller than $\alpha\frac{\sqrt{2}L}{2}(\frac{R}{q} + 0.57\sqrt{\pi R})$. Consider now the case $0 < r < r_k$. In the recourse action, the smallest enclosing circle C_k has radius $r_k = \frac{\sqrt{2}L}{2}$, as $r^{SL}(k) = \sqrt{2}L$ implies $\bar{r}_k = \sqrt{2}L$. Since $\beta > \alpha$ and $r_{max} = \frac{\sqrt{2}L}{2}$, the total cost of the realization k is

$$\alpha D_1(r) + \beta D_2(r, \frac{\sqrt{2}L}{2}) \geq \alpha(D_1(r) + D_2(r, \frac{\sqrt{2}L}{2})) =$$

$$= \alpha\left(\frac{\sqrt{2}L}{2}\frac{R}{q} + 0.57\sqrt{\pi R}\left(r\sqrt{\frac{r}{\frac{\sqrt{2}L}{2}}} + \sqrt{((\frac{\sqrt{2}L}{2})^2 - r^2)(1 - \frac{r}{\frac{\sqrt{2}L}{2}})}\right)\right).$$

This lower bound reaches its minimum value for $r = \frac{\sqrt{2}L}{6} \approx 0.23570L$.

Therefore, given $f(r) = r\sqrt{\frac{r}{\frac{\sqrt{2}L}{2}}} + \sqrt{((\frac{\sqrt{2}L}{2})^2 - r^2)(1 - \frac{r}{\frac{\sqrt{2}L}{2}})}$, since $f(\frac{\sqrt{2}L}{6}) = \frac{5\sqrt{6}L}{18} < \frac{\sqrt{2}L}{2}$, in any solution

$$z^*(k) \geq \alpha\left(\frac{\sqrt{2}L}{2}\frac{R}{q} + \frac{5\sqrt{6}L}{18}0.57\sqrt{\pi R}\right).$$

Since $\frac{\alpha\sqrt{2}L\frac{R}{q}}{\alpha\frac{\sqrt{2}L}{2}\frac{R}{q}} = 2$ and $\frac{\alpha\sqrt{2}L0.57\sqrt{\pi R}}{\alpha\frac{5\sqrt{6}L}{18}0.57\sqrt{\pi R}} = \frac{6\sqrt{3}}{5} \approx 2.07846$, then

$$z^{SL}(k) \leq \frac{6\sqrt{3}}{5}\alpha\left(\frac{\sqrt{2}L}{2}\frac{R}{q} + \frac{5\sqrt{6}L}{18}0.57\sqrt{\pi R}\right).$$

Therefore,

$$\frac{z^{SL}(k)}{z^*(k)} \leq \frac{\alpha\sqrt{2}L\left(\frac{R}{q}+0.57\sqrt{\pi R}\right)}{\frac{\alpha\sqrt{2}L}{2}\left(\frac{R}{q}+\frac{5\sqrt{6}L}{18}0.57\sqrt{\pi R}\right)} \leq \frac{6\sqrt{3}}{5}.$$

We now prove that the bound is tight. Consider the following realization: $L=1$, centre of the area $\mathscr{A}$ in $(0,0)$, $N=1$, $R \geq 2$, $q=R$, $\alpha=1$, $\beta=1+\frac{1}{R}$, customer 1 located at $(-0.5-0.5)$, customer 2 at $(0.5, 0.5)$, the remaining customers located at $(0.5-\frac{1}{R}\rho_\tau, 0.5-\frac{1}{R}\rho_\tau)$, where τ is the index of the customers and ρ_τ is a random number between 0 and 1.

In the *SL*, the facility is located at one of the customers. The location that minimizes the cost is the one of the customer closest to $(0,0)$. Since $\rho_\tau \leq 1$, the coordinates of this customer are $(0.5-\frac{1}{R}, 0.5-\frac{1}{R})$, at best. Therefore, $r^{SL}(k) \geq \sqrt{2(-1+\frac{1}{R})^2}$ and $z^{SL}(k) \geq \sqrt{2(-1+\frac{1}{R})^2}\left(1+0.57\sqrt{\pi R}\right)$.

The optimal cost $z^*(k)$ is not greater than the one of the solution in which the circle C has centre in $(0,0)$ and $r=\frac{\sqrt{2}}{6}$, while $r_k=\frac{\sqrt{2}}{2}$, that is,

$$\begin{aligned} z^*(k) &\leq \alpha D_1(\frac{\sqrt{2}}{6}) + \beta D_2(\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{2}) = \\ &= \frac{\sqrt{2}}{6}(\frac{1}{3}+\frac{1}{\sqrt{3}}0.57\sqrt{\pi R}) + (1+\frac{1}{R})(\frac{4\sqrt{2}}{9}+\frac{2\sqrt{6}}{9}0.57\sqrt{\pi R}) \ . \end{aligned}$$

Therefore, in this realization

$$\frac{z^{SL}(k)}{z^*(k)} \geq \frac{\sqrt{2(-1+\frac{1}{R})^2}\left(1+0.57\sqrt{\pi R}\right)}{\frac{\sqrt{2}}{6}\left(\frac{1}{3}+\frac{1}{\sqrt{3}}0.57\sqrt{\pi R}\right)+(1+\frac{1}{R})\left(\frac{4\sqrt{2}}{9}+\frac{2\sqrt{6}}{9}0.57\sqrt{\pi R}\right)} \rightarrow \frac{6\sqrt{3}}{5}$$

for $R \rightarrow \infty$. □

We now consider the general case, i.e., instances with several realizations. We show that the *SL* can have a performance infinitely worse than the one of the optimal solution of the SCTSLP-R problem. Let r^{SL} and z^{SL} be the radius and the total cost obtained by the *SL*, respectively.

Theorem 3.3 *There exists an instance such that* $\frac{z^{SL}}{z^*} \rightarrow \infty$.

Proof Consider the following instance: $L = 1$, centre of the area $\mathscr{A}$ in $(0,0)$, $R \geq 2$, $N = 1$, $q = 2$, $\alpha = 1$, $\beta = 1 + \frac{1}{R}$, customer 1 located at $(-0.5 - 0.5)$, customer 2 at $(0.5, 0.5)$, the remaining customers located at $(0.5 - \frac{1}{R}\rho_\tau, 0.5 - \frac{1}{R}\rho_\tau)$, where τ is the index of the customers and ρ_τ is a random number between 0 and 1. $\frac{R}{2}$ realizations are available. Each realization has the same probability $\frac{2}{R}$ and is composed of $q = 2$ customers served by the same traveling salesman. In particular, realization 1 is just composed of customers 1 and 2, while the remaining $\frac{R}{2} - 1$ realizations are composed of two of the remaining customers at each realization.

In the *SL*, the facility is located at one of the customers. The location that minimizes the cost is the customer closest to $(0,0)$. Since $\rho_\tau \leq 1$, the coordinates of this customer are $(0.5 - \frac{1}{R}, 0.5 - \frac{1}{R})$, at best.
Therefore, $r^{SL} \geq \sqrt{2(-1+\frac{1}{R})^2}$ and $z^{SL} \geq \sqrt{2(-1+\frac{1}{R})^2}\left(1 + 0.57\sqrt{\pi R}\right)$.

The optimal cost is not greater than the cost of the solution in which the circle C has centre in $(0,0)$ and $r = 0$, that is,

$$z^* \leq (1+\frac{1}{R})\left(\frac{2}{R}\left(\frac{\sqrt{2}}{2}(\frac{R}{2} + 0.57\sqrt{\pi R})\right) + (1-\frac{2}{R})\left(\frac{\frac{\sqrt{2}}{R}}{\sqrt{3}}\left(\frac{R}{2} + 0.57\sqrt{\pi R}\right)\right)\right).$$

Therefore, in this instance

$$\frac{z^{SL}}{z^*} \geq \frac{\sqrt{2(-1+\frac{1}{R})^2}\left(1 + 0.57\sqrt{\pi R}\right)}{(1+\frac{1}{R})\left(\frac{2}{R}\left(\frac{\sqrt{2}}{2}(\frac{R}{2} + 0.57\sqrt{\pi R})\right) + (1-\frac{2}{R})\left(\frac{\frac{\sqrt{2}}{R}}{\sqrt{3}}\left(\frac{R}{2} + 0.57\sqrt{\pi R}\right)\right)\right)} \to \infty$$

for $R \to \infty$. □

We can note that the feasible solutions of the SCTSLP–R problem used in the proofs of the previous theorems were not trivial. Therefore, the optimal solution can be not easy to be found.

4 The Approach based on the Stochastic Second–Order Cone Formulation

In this section, we propose a solution approach based on the idea to optimally solve a Stochastic Second–Order Cone Formulation of the problem with an approximate objective function. This approach is referred to as *SSOCP* approach.

4.1 Approximation of the Costs

The first step is to approximate $D_1(r)$ by a function proportional to the area of the circle C, that is $\theta\pi r^2$, where θ is a parameter to be estimated. The optimal value of θ, that is θ^*, is obtained by solving the following minimum least square problem:

$$\min_{\theta} \sum_{\omega=0}^{\lfloor \frac{r_{\max}}{\Delta} \rfloor} \Big(\theta\pi\,(\omega\Delta)^2 - D_1\,(\omega\Delta)\Big)^2 ,$$

where Δ is a given stepsize. Therefore, the first stage cost $\alpha D_1(r)$ is approximated by $\alpha\theta^*\pi r^2$.

We then approximate the second stage cost by a function proportional to the difference between the area of C_k and the area of C, that is $\phi\pi(r_k^2 - r^2)$, where ϕ is a parameter to be estimated. The optimal value of ϕ, that is ϕ^*, is obtained by solving the following minimum least square problem:

$$\min_{\phi} \sum_{\omega=0}^{\lfloor \frac{r_{\max}}{\Delta} \rfloor} \sum_{k=\omega+1}^{\lfloor \frac{r_{\max}}{\Delta} \rfloor} \Big(\phi\pi((k\Delta)^2 - (\omega\Delta)^2) - D_2(\omega\Delta, k\Delta)\Big)^2 .$$

The second stage cost $\beta\sum_{k\in\mathcal{K}} p_k D_2(r, r_k)$ is approximated by

$$\beta \sum_{k\in\mathcal{K}} p_k \phi^* \pi (r_k^2 - r^2) \ .$$

4.2 The Stochastic Semidefinite Model

We first model the problem SCTSLP–R as a stochastic semidefinite model. A similar approach has been adopted by [22] and [25] for another type of application: the Stochastic location-aided routing (SLAR) in wireless ad-hoc networks.

Due to the different type of application, several are the differences in the proposed formulation, with respect to the one presented in [22] and [25]:

1. The source node defined in SLAR does not exist in the context of SCTSLP-R problem since signals have not to be sent.
2. Multiple customers are considered in the context of SCTSLP-R problem instead of a single destination node as in SLAR.
3. An initial circle defined in SLAR, centred at the early location of a customer, with radius corresponding to the minimal velocity the client is supposed to move, does not make any sense in the context of SCTSLP-R problem and it is not considered.
4. Each realization is modelled by a given number of ellipses corresponding to the area covering by each traveling salesman, instead of a single ellipse describing the random movement of a single destination node as in SLAR.
5. Information on the number of customers served by each traveling salesman is included in the context of SCTSLP-R problem.
6. The first and second stage costs are proportional to the routing cost of the corresponding Capacitated Vehicle Routing Problem (CVRP) in the context of SCTSLP-R problem.

In the following, $\mathbb{R}^n$ denotes the space of all $n-$dimensional reals vectors, $\mathbb{R}^{n\times n}$ denotes the vector space of real $n \times n$ matrices, lower case boldface letters $\mathbf{x}, \mathbf{c}$ etc. for column vectors, and uppercase letters A, X etc. for matrices. Sub-scripted vectors, such as $\mathbf{x_i}$, represent the i^{th} block of $\mathbf{x}$. The j^{th} component of the vectors $\mathbf{x}$ and $\mathbf{x_i}$ are indicated by x_j and x_{ij}, respectively. We use

$\mathbf{0}$ and $\mathbf{1}$ for the zero vector and vectors of all ones, respectively, and 0 and I for the zero and identity matrices.

Let $\mathbb{R}_s^{n\times n}$ denotes the vector space of real $n \times n$ symmetric matrices, for $A, B \in \mathbb{R}_s^{n\times n}$ we write $A \succeq 0$ ($A \succ 0$) to mean that A is positive semidefinite (positive definite) and $A \succeq B$ ($A \succ B$) to mean that $A - B \succeq 0$ ($A - B \succ 0$). For $A, B \in \mathbb{R}^{n\times n}$ we denote by $A \bullet B$ the Frobenius inner product between A and B: $A \bullet B = trace(A^T B)$.

With the above notation, a stochastic semidefinite programming problem with recourse (SSDP) in standard primal form (see [22]), is given by

$$\begin{aligned} &\min_{X\in\mathbb{R}_s^{n_1\times n_1}} C \bullet X + E[Q(X,\omega)] \\ &\text{s.t.}\ \ A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m_1, \quad X \succeq 0, \end{aligned}$$

where $X \in \mathbb{R}_s^{n_1\times n_1}$ is the first-stage decision variable, $C \in \mathbb{R}_s^{n_1\times n_1}$ is a given matrix, $\mathbf{b} \in \mathbb{R}^{m_1}$ another given vector, $A \in \mathbb{R}_s^{n_1\times n_1}$, $\mathbf{c}$, $\mathbf{b}$ and A are deterministic data. $Q(X, \omega)$ is the minimum of the second stage problem

$$\begin{aligned} &\min_{Y(\omega)\in\mathbb{R}_s^{n_2\times n_2}} Q(\omega) \bullet Y \\ &\text{subject to}\ \ T_i(\omega) \bullet X + W_i(\omega) \bullet Y = h_i(\omega) \quad i = 1, 2, \ldots, m_2, \quad Y \succeq 0, \end{aligned}$$

and

$$E[Q(X,\omega)] = \textstyle\int_\Omega Q(X,\omega) P(\mathrm{d}\omega), \tag{7}$$

where $Y(\omega) \in \mathbb{R}_s^{n_2\times n_2}$ is the second-stage decision vector, $Q \in \mathbb{R}_s^{n_2\times n_2}$, $T_i(\omega) \in \mathbb{R}_s^{n_1\times n_1}$, $W_i(\omega) \in \mathbb{R}_s^{n_2\times n_2}$, $\mathbf{h} \in R^{m_2}$ and $\omega \in \Omega$ is a random outcome with known probability distribution P, whose realizations will affect the coefficient matrices of the problem.

Using the above notation, the two stage stochastic semidefinite programming problem for SCTSLP-R can be summarized with the following decision variables:

$$\mathbf{x} = [d_2, \tilde{\mathbf{u}}, \gamma]^T \ , \tag{8}$$

$$\mathbf{y} = [\mathbf{z}, \tilde{\boldsymbol{\gamma}} \ , \boldsymbol{\delta}]^T \ , \tag{9}$$

where $\mathbf{x}$ is the first stage decision variable with components

- d_2: is an upper bound on square of the radius of the circle C;
- $\tilde{\mathbf{u}} \in \mathbb{R}^2$: is the centre of circle C;
- γ: is the coefficient in the equation of circle C,

and $\mathbf{y}$ is the second stage decision vector whose components are

- $\mathbf{z} \in \mathbb{R}^K$: is the vector of the upper bounds, at realization k, on the quantity to enlarge the circle C to cover all the ellipses E_k^i by means of C_k (see eq. 3);
- $\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^K$: is the vector of the coefficients $\tilde{\gamma}_k$ of the second stage circles C_k, $k \in \mathcal{K}$;
- $\boldsymbol{\delta} \in \mathbb{R}^K$: is a vector of non-negative parameters (see [24, 28]).

The coefficients of the decision variables in the objective function are given by:

$$\mathbf{c} = [\alpha\theta^*\pi, \mathbf{0}, 0]^T \ , \tag{10}$$

$$\mathbf{q} = [\boldsymbol{\beta}\phi^*\pi, \mathbf{0}, \mathbf{0}]^T \ , \tag{11}$$

where $\alpha > 0$ is the cost per unit of the area of C and $\boldsymbol{\beta}$ is the vector of identical costs β per unit increase of the area of C to the area of C_k.

The costs of choosing the service area C is

$$\alpha\theta^*\pi \left(\tilde{\mathbf{u}}^T\tilde{\mathbf{u}} - \gamma\right) . \tag{12}$$

We note that minimizing the nonlinear cost function (12) is equivalent to minimizing the linear cost function $\alpha\theta^*\pi d_2$ under the constraint $d_2 \geq \tilde{\mathbf{u}}^T\tilde{\mathbf{u}} - \gamma$.

Finally, by noting that the latter constraint can be written as

$$\begin{pmatrix} I & \tilde{\mathbf{u}} \\ \tilde{\mathbf{u}}^T & d_2+\gamma \end{pmatrix} \succeq 0\,, \tag{13}$$

and the condition of inclusion into a second stage circle C_k of the ellipse E_k^i

$$E_k^i = \{\mathbf{u} \in \mathbb{R}^2 : \mathbf{u}^T H_k^i \mathbf{u} + 2{\mathbf{g}_k^i}^T \mathbf{u} + v_k^i \leq 0\}, \quad k \in \mathcal{K},\ i \in \mathcal{N}\ , \tag{14}$$

(with $H_k^i \in \mathbb{R}_s^{2\times 2}$, $H_k^i \succ 0$, $\mathbf{g}_k^i \in \mathbb{R}^2$ and $v_k^i \in \mathbb{R}$) as

$$\begin{pmatrix} I & -\tilde{\mathbf{u}} \\ -\tilde{\mathbf{u}}^T & \tilde{\gamma}_k \end{pmatrix} \preceq \delta_k \begin{pmatrix} H_k^i & \mathbf{g}_k^i \\ {\mathbf{g}_k^i}^T & v_k^i \end{pmatrix} \quad , \quad k \in \mathcal{K},\ i \in \mathcal{N}\ , \tag{15}$$

we deduce that the problem can be written as a stochastic semidefinite programming model with recourse (SSDP) as follows:

$$\min_{\mathbf{x}\in\mathbb{R}^4}\ \mathbf{c}^T\mathbf{x} + E[Q(\mathbf{x},\omega)] \tag{16}$$

$$\text{s.t.} \quad 0 \preceq \begin{pmatrix} I & \tilde{\mathbf{u}} \\ \tilde{\mathbf{u}}^T & d_2+\gamma \end{pmatrix}\ ,$$

where $Q(\mathbf{x},\omega)$ is the minimum of the problem

$$\begin{aligned} &\min_{\mathbf{y}\in\mathbb{R}^{3K}}\ \mathbf{q}^T\mathbf{y} \\ &\text{s.t.} \quad \begin{pmatrix} I & -\tilde{\mathbf{u}} \\ -\tilde{\mathbf{u}}^T & \tilde{\gamma}_k^i \end{pmatrix} \preceq \delta_k \begin{pmatrix} H_k & \mathbf{g}_k^i \\ {\mathbf{g}_k^i}^T & v_k^i \end{pmatrix} \quad , \quad k \in \mathcal{K}\ ,\ i \in \mathcal{N}\ , \\ &\qquad 0 \leq \delta_k\ , \quad k \in \mathcal{K}\ , \\ &\qquad 0 \leq \gamma - \tilde{\gamma}_k^i \leq z_k\ , \quad k \in \mathcal{K}\ \ i \in \mathcal{N}\ . \end{aligned} \tag{17}$$

4.3 The Stochastic Second-Order Cone Model

A special case of semidefinite programming (SDP) is given by second-order cone programming (SOCP). SOCP problems consist in convex optimization problems in which a linear function is minimized over the intersection of an affine set and the product of second-order (Lorentz) cones:

$$\mathscr{K}_n := \{\mathbf{x} = (x_0; \bar{\mathbf{x}}) \in \mathbb{R}^n : x_0 \geq \|\bar{\mathbf{x}}\|\} \ , \tag{18}$$

where $\|\cdot\|$ refers to the standard Euclidean norm and n the dimension of $\mathscr{K}_n$ (see [26]).

A second-order cone can be embedded in the cone of positive semidefinite matrices since a second-order cone constraint is equivalent to a linear matrix inequality according to the following relation:

$$Arw(\mathbf{x}) := \begin{pmatrix} x_0 & -\bar{\mathbf{x}}^T \\ -\bar{\mathbf{x}} & x_0 I \end{pmatrix} \succeq 0 \Leftrightarrow \ x_0 \geq \|\bar{\mathbf{x}}\| \ . \tag{19}$$

In fact, $Arw(\mathbf{x}) \succeq 0$ if and only if either $\mathbf{x} = \mathbf{0}$, or $x_0 > 0$ and it holds true the Shur Complement $x_0 - \bar{\mathbf{x}}^T (x_0 I)^{-1} \bar{\mathbf{x}} \geq 0$.

From a computational point of view, the effort per iteration required by interior-point method to solve stochastic second order cone problems is lower than the one required to solve stochastic semidefinite problems of similar size and structure (see [26]).

The aim of this section is to formulate the semidefinite stochastic problem presented in the previous section as a stochastic second-order cone *SSOCP* problem. See [27] for a similar reformulation for a stochastic location-aided routing (SLAR) model in wireless ad-hoc networks.

We rewrite each semidefinite constraint as a second order cone one. The constraint:

$$0 \preceq \begin{pmatrix} I & -\tilde{\mathbf{u}} \\ -\tilde{\mathbf{u}}^T & d_2 + \gamma \end{pmatrix} \Leftrightarrow d_2 + \gamma \geq \tilde{\mathbf{u}}^T \tilde{\mathbf{u}} \Leftrightarrow \begin{pmatrix} \sqrt{d_2 + \gamma} \\ \tilde{\mathbf{u}} \end{pmatrix} \in \mathscr{K}_3 \ ; \tag{20}$$

the second stage constraint

$$\begin{pmatrix} I & -\tilde{\mathbf{u}} \\ -\tilde{\mathbf{u}}^T & \tilde{\gamma}_k \end{pmatrix} \preceq \delta_k \begin{pmatrix} H_k^i & \mathbf{g}_k^i \\ {\mathbf{g}_k^i}^T & v_k^i \end{pmatrix} \ , \quad k \in \mathcal{K}, \ i \in \mathcal{N} \tag{21}$$

is equivalent to

$$M_k^i := \begin{pmatrix} \delta_k H_k^i - I & \delta_k \mathbf{g}_k^i + \tilde{\mathbf{u}} \\ \delta_k {\mathbf{g}_k^i}^T + \tilde{\mathbf{u}}^T & \delta_k v_k^i - \tilde{\gamma}_k \end{pmatrix} \succeq 0 \ , \qquad k \in \mathcal{K},\ i \in \mathcal{N} \ . \tag{22}$$

Following [26], let $H_k^i = Q_k^i \Lambda_k^i {Q_k^i}^T$ be the spectral decomposition of H_k^i, $\Lambda_k^i = Diag\left(\lambda_{k1}^i; \ldots; \lambda_{kn}^i\right)$ and $\mathbf{h}_k^i = {Q_k^i}^T \left(\delta_k \mathbf{g}_k^i + \tilde{\mathbf{u}}\right)$, for $k \in \mathcal{K},\ i \in \mathcal{N}$. Then

$$\bar{M}_k^i := \begin{pmatrix} {Q_k^i}^T & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} M_k^i \begin{pmatrix} Q_k^i & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \delta_k \Lambda_k^i - I & \mathbf{h}_k^i \\ {\mathbf{h}_k^i}^T & \delta_k v_k^i - \tilde{\gamma}_k \end{pmatrix} \succeq 0 \ , \tag{23}$$

for $k \in \mathcal{K},\ i \in \mathcal{N}$ and $M_k^i \succeq 0$ if and only if $\bar{M}_k^i \succeq 0$. It holds if and only if $\delta_k \geq \frac{1}{\lambda_{min}\left(\Lambda_k^i\right)}$, i.e., $\delta_k \lambda_{kj}^i - 1 \geq 0 \ \forall\, k, j, i$, $h_{kj}^i = 0$ if $\delta_k \lambda_{kj}^i - 1 = 0$ and the Shur complement of the columns and rows of $\bar{M}_k^i$ that are not zero

$$\delta_k v_k^i - \tilde{\gamma}_k - \sum_{\delta_k \lambda_{kj}^i > 1} \frac{{h_{kj}^i}^2}{\delta_k \lambda_{kj}^i - 1} \geq 0 \ . \tag{24}$$

If we define $\boldsymbol{s}_k^i := (s_{k1}^i; \ldots; s_{kn}^i)$, where $s_{kj}^i = \frac{{h_{kj}^i}^2}{\delta_k \lambda_{kj}^i - 1}$, for all j such that $\delta_k \lambda_{kj}^i > 1$ and $s_{kj}^i = 0$, otherwise, then (24) is equivalent to

$$\tilde{\gamma}_k \leq \delta_k v_k^i - \mathbf{1}^T \ \boldsymbol{s}_k^i \ . \tag{25}$$

Since we are minimizing the radius of the second stage circles C_k, $\sqrt{\tilde{\mathbf{u}}^T \tilde{\mathbf{u}} - \tilde{\gamma}_k}$, we can relax the definition of s_{kj}^i replacing it by ${h_{kj}^i}^2 \leq s_{kj}^i \left(\delta_k \lambda_{kj}^i - 1\right)$, $k \in \mathcal{K},\ i \in \mathcal{N},\ j = 1, 2$. Combining all of the above constraints (21) is equivalent to the following formulation involving only linear and restricted hyperbolic second-stage constraints:

$$\mathbf{h}_k^i = {Q_k^i}^T \left(\delta_k \mathbf{g}_k^i + \tilde{\mathbf{u}}\right) \ , \qquad k \in \mathcal{K},\ i \in \mathcal{N} \ , \tag{26}$$

$${h_{kj}^i}^2 \leq s_{kj}^i \left(\delta_k \lambda_{kj}^i - 1\right) \ , \qquad k \in \mathcal{K},\ i \in \mathcal{N} \quad j = 1, 2 \tag{27}$$

$$\tilde{\gamma}_k \leq \delta_k v_k^i - \mathbf{1}^T \ \boldsymbol{s}_k^i \ , \qquad k \in \mathcal{K},\ i \in \mathcal{N} \ , \tag{28}$$

$$\delta_k \geq \frac{1}{\lambda_{min}\left(\Lambda_k^i\right)} \ , \qquad k \in \mathcal{K},\ i \in \mathcal{N} \ . \tag{29}$$

Note that the restricted hyperbolic constraint (27) is equivalent to the following $2NK$ 3-dimensional second-order cone inequalities:

$$\left\| \begin{pmatrix} 2h^i_{kj} \\ s^i_{kj} - \delta_k \lambda^i_{kj} + 1 \end{pmatrix} \right\| \leq s^i_{kj} + \delta_k \lambda^i_{kj} - 1 \tag{30}$$

$$\Updownarrow$$

$$\begin{pmatrix} s^i_{kj} + \delta_k \lambda^i_{kj} - 1 \\ 2h^i_{kj} \\ s^i_{kj} - \delta_k \lambda^i_{kj} + 1 \end{pmatrix} \in \mathscr{K}_3, \ k \in \mathcal{K}, \ i \in \mathcal{N}, \ j = 1, 2, \tag{31}$$

and each of the linear constraints (28) and (29) are NK 1-dimensional second-order cone constraints. In conclusion, the model (16)-(17) can be formulated as a stochastic second-order cone $SSOCP$ problem in the following way:

$$\begin{aligned}
\min_{\mathbf{x} \in \mathbb{R}^4, \ \mathbf{y} \in \mathbb{R}^{3K}} \quad & \mathbf{c}^T \mathbf{x} + \sum_{k=1}^{K} p_k \mathbf{q}^T \mathbf{y} \\
\text{s.t.} \quad & d_2 + \gamma \geq \|\tilde{\mathbf{u}}\|^2 , \\
& \left(s^i_{kj} + \delta_k \lambda^i_{kj} - 1\right)^2 \geq \left\| \begin{pmatrix} 2h^i_{kj} \\ s^i_{kj} - \delta_k \lambda^i_{kj} + 1 \end{pmatrix} \right\|^2 , k \in \mathcal{K}, i \in \mathcal{N}, j = 1, 2, \\
& \mathbf{h}^i_k = {Q^i_k}^T \left(\delta_k \mathbf{g}^i_k + \tilde{\mathbf{u}}\right) , \quad k \in \mathcal{K}, \ i \in \mathcal{N} , \\
& \tilde{\gamma} \leq \delta_k v^i_k - \mathbf{1}^T \mathbf{s}^i_k , \quad k \in \mathcal{K}, \ i \in \mathcal{N} , \\
& \delta_k \geq \frac{1}{\lambda_{min}\left(\Lambda^i_k\right)} , \quad k \in \mathcal{K}, \ i \in \mathcal{N} , \\
& \delta_k \geq 0 , \quad k \in \mathcal{K} , \\
& \gamma - \tilde{\gamma}_k \geq 0 , \quad k \in \mathcal{K} , \\
& \gamma - \tilde{\gamma}_k \leq z_k , \quad k \in \mathcal{K} , \\
& d_2 \geq 0 .
\end{aligned} \tag{32}$$

5 Numerical Results

In this section, we first describe the methodology we use to generate the realizations. Then, we analyze the sensitivity of the exact *SSOCP* model solution with respect to input parameters. Finally, we compare the solutions of the *SSOCP* approach with the *SL* approach. The *SSOCP* model is implemented in GAMS 22.5, by using the second order cone programming solver from the software package Mosek (http://www.mosek.com/), while the *SL* approach is implemented in C++.

5.1 Uncertainty generation

In our numerical experiments, we generate ellipses of the form

$$E = \{\mathbf{u} \in \mathbb{R}^2 : \mathbf{u}^T H \mathbf{u} + 2\mathbf{g}^T \mathbf{u} + v \leq 0\}, \tag{33}$$

where $H \in \mathbb{R}_s^{2\times 2}$ is a given positive definite matrix, $\mathbf{g} \in \mathbb{R}^2$ is a given vector, and $v \in \mathbb{R}$ is a given real number such that

$$v < \mathbf{g}^T H^{-1} \mathbf{g} \ . \tag{34}$$

Equation (33) can also be written as

$$E = \{\mathbf{u} \in \mathbb{R}^2 : \left\| H^{1/2}(\mathbf{u} - \mathbf{u}^0) \right\| \leq 1\} \ , \tag{35}$$

with $H = Q\Lambda Q^T$ the spectral decomposition of H, where Q is the matrix whose columns are the eigenvectors of H and $\Lambda = Diag\,(\lambda_1; \ldots; \lambda_n)$ is the diagonal matrix of the corresponding eigenvalues with

$$\mathbf{u}^0 = -H^{-1}\mathbf{g}, \quad \rho = \sqrt{\mathbf{g}^T H^{-1}\mathbf{g} - v} \ .$$

Note that an ellipse is completely defined by its centre $\mathbf{u}^0 = (u_1^0, u_2^0)$, the angle φ between the first axis of the ellipse and the u_1-axis of the coordinate system, and the lengths σ_1, σ_2 of the two semi-axes. For given φ, $\mathbf{u}^0 = (u_1^0, u_2^0)$, σ_1, σ_2, we can represent the ellipse in the form (33) by setting

$$H := Q \begin{pmatrix} \sigma_1^{-2} & 0 \\ 0 & \sigma_2^{-2} \end{pmatrix} Q^T \ , \quad \mathbf{g} = -H\mathbf{u}^0 \ . \quad v = \mathbf{u}^{0T} H \mathbf{u}^0 - 1 \ . \qquad (36)$$

We randomly generate these quantities to obtain ellipses of the form (33) by using (36). The quantities

$$u_{1,k}^{0,i}, u_{2,k}^{0,i}, \varphi_k^i, \sigma_{1,k}^i, \sigma_{2,k}^i, \qquad k \in \mathcal{K} \ , i \in \mathcal{N} \qquad (37)$$

corresponding to the ellipse

$$E_k^i = \{\mathbf{u} \in \mathbb{R}^2 : \mathbf{u}^T H_k^i \mathbf{u} + 2{\mathbf{g}_k^i}^T \mathbf{u} + v_k^i \leq 0\} \ , \qquad k \in \mathcal{K}, \ i \in \mathcal{N} \ , \qquad (38)$$

(with $H_k^i \in \mathbb{R}_s^{2\times 2}$, $H_k^i \succ 0$, $\mathbf{g}_k^i \in \mathbb{R}^2$ and $v_k^i \in \mathbb{R}$), which represents the area covered by the traveling salesman $i \in \mathcal{N}$ at realization $k \in \mathcal{K}$, are obtained by considering:

- the uniform distribution in the interval $\left(\hat{u}_1^i - 2, \hat{u}_1^i + 2\right)$ for generating $u_{1,k}^{0,i}$;
- the normal distribution $\mathcal{N}\left(\hat{u}_2^i, 0.5\right)$ for generating $u_{2,k}^{0,i}$;
- the uniform distribution in the interval $[0, \pi]$ for generating φ_k^i;
- the normal distribution $\mathcal{N}(2, 1)$ for generating $s_{u_1,k}^i$ and $s_{u_2,k}^i$,

with $\hat{\mathbf{u}}_1 \in \mathbb{R}^N$ and $\hat{\mathbf{u}}_2 \in \mathbb{R}^N$ representing the coordinates of the initial position of the N traveling salesmen. The ellipses E_k^i are randomly generated in MATLAB 7.4.0 inside the square $[-10, 10] \times [-10, 10]$. Each ellipse E_k^i covers q customers randomly generated by a uniform distribution. Note that at each realization the same number of customers Nq should be served. The difference among the realizations is in the customers' positions. Note that since we generate the same number of customers inside areas of different extension, the described approach corresponds to model distributed demands with different degree of density. In our computational experiment, we consider the cases in which each traveling salesman respectively serves 1, 5, and 25 customers

($q = 1, 5, 25$). Furthermore, the K realizations are supposed to be equiprobable, with probability $1/K$.

5.2 Sensitivity Analysis

In this section, we carry out a sensitivity analysis of the *SSOCP* solution with respect to some of the input parameters (see [29]). In particular, we consider:

1. The number K of realizations in the realization tree;
2. The cost β of making corrections on the radius of the second stage circles.

To carry out the analysis, we assume that the number of traveling salesmen is $N = 5$, with initial positions as follows:

$$\hat{\mathbf{u}}_1 = (5, -5, 0, 0, 5)^T \qquad \hat{\mathbf{u}}_2 = (0, 0, 5, -5, 5)^T \ .$$

Furthermore, in order to make a consistent comparison in terms of number of realizations, the approximation of the costs according to the corresponding *Capacitated Traveling Salesmen Problem* described in Section 4.1 is here not considered. The exact solutions obtained by optimally solving the deterministic equivalent SCTSLP-R with increasing number of realizations or second stage cost β are then compared.

The results of the analysis are the following:

(1) *Number K of realizations:* Figure 1 shows the stabilization of the objective function in the *SSOCP* model as the number of realizations increases from 3 up to 20335. An *in-sample stability* is observed, (see [30]) with a stabilization of the objective function around the value of 267.

(2) *Cost β:* The sensitivity analysis of the first stage objective function and total costs versus the second stage cost $\beta\phi\pi \in [0, 4\pi]$ is reported in Figure 2. The results show that, for a low value of $\beta\phi\pi \in [0, 2.51]$ the first stage cost is zero; this means that it is more convenient to enlarge the radius

of each recourse zone C_k instead of moving the location of the facility. Conversely, for $\beta\phi\pi \geq 8.4$, the service zone C already contains all the routes of traveling salesmen E^i_k, $(k \in \mathcal{K},\, i \in \mathcal{N})$; the total cost reduces to the first stage total cost and it stabilizes on the value of 287.78.

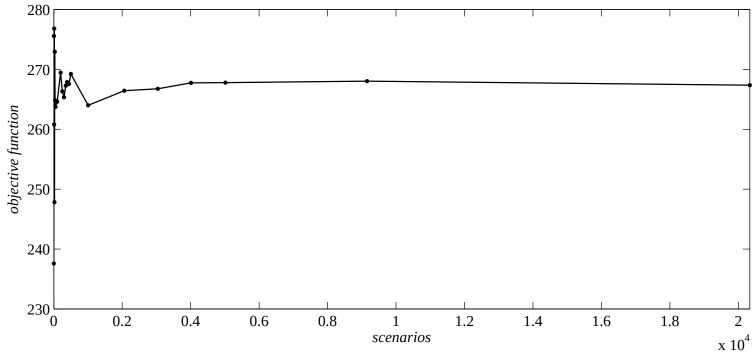


Fig. 1 Sensitivity analysis of the optimal function value versus the number realizations for the exact *SSOCP* model.

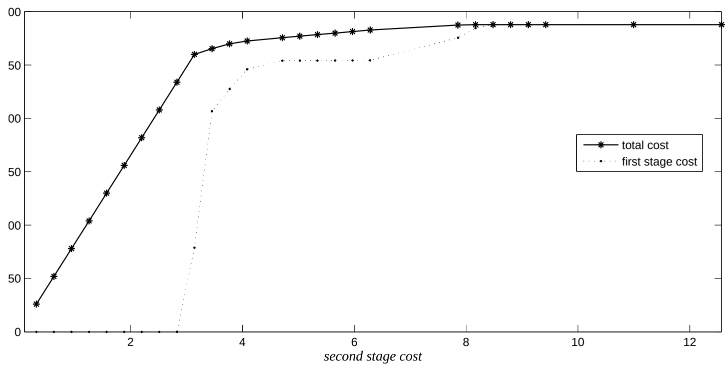


Fig. 2 Case of 5 realizations: sensitivity analysis of the objective function and first stage costs versus the second stage cost $\beta\phi\pi \in [0, 4\pi]$ for the for the exact *SSOCP* model.

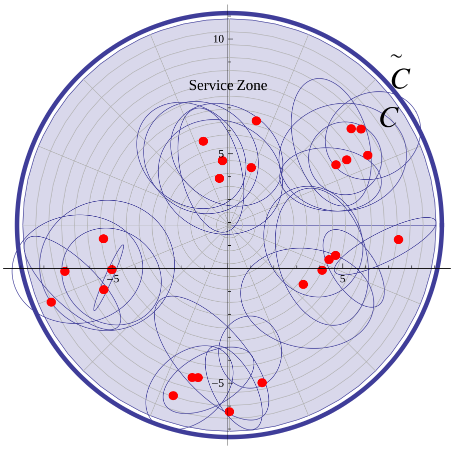
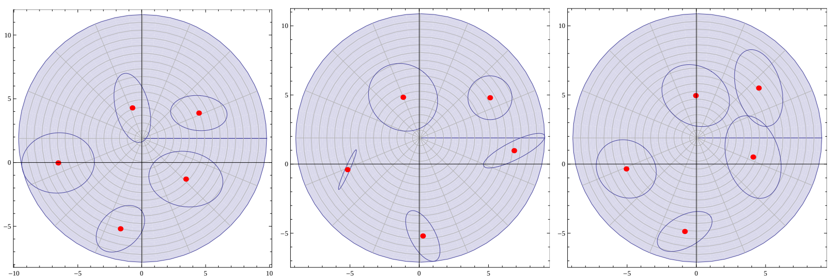
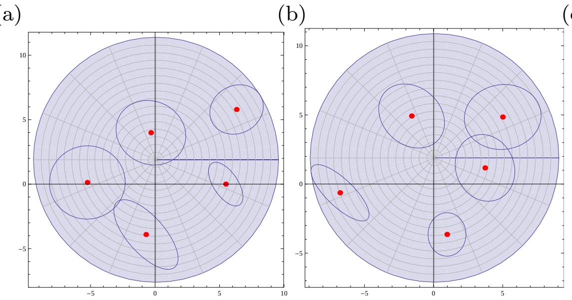


Fig. 3 Case of 25 random customers supplied by 5 traveling salesmen with capacity $q = 1$ in 5 realizations: service zone C and requested service zone $\tilde{C}$ solutions for the exact *SSOCP* model.



(a) (b) (c)



(d) (e)

Fig. 4 Second stage zones solutions C_k for the exact *SSOCP* model in the case of 25 random customers supplied by 5 traveling salesmen with capacity $q = 1$ in each of the five realizations (a) $k = 1$, (b) $k = 2$, (c) $k = 3$, (d) $k = 4$ and (e) $k = 5$. Note that, at realization k, the second stage disk C_k contains all the ellipses E_k^i $(i = 1, \ldots, 5)$.

5.3 Comparison with the *SL* approach

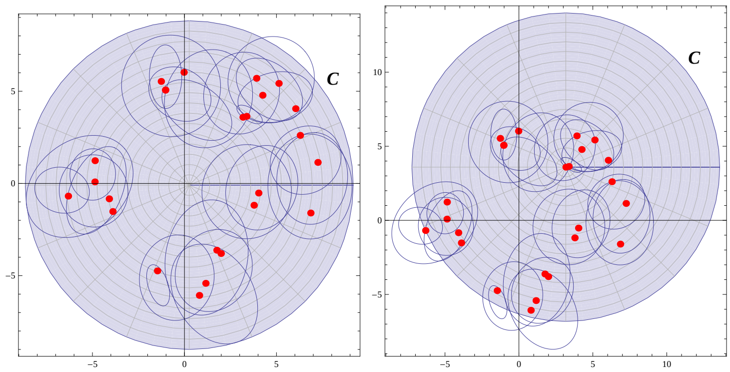
In this section, we show the improvement in the cost obtained by the *SSOCP* approach with respect to the *SL*. We consider the cases of 5 and 25 traveling salesmen ($N = 5$ and $N = 25$) with capacity $q = 1, 5, 25$ (see Tables 2 and 3). The number of realizations is set to $K = 5$, each of them contains N traveling salesmen serving q different customers, for a total number of customers $R = KNq$. In order to make a consistent comparison between the two solutions approaches, the costs θ and ϕ in the *SSOCP* model are now computed according to the cost of the corresponding *Capacitated Vehicle Routing Problem* (CVRP), by solving the minimum least square problem explained in Section 4.1. The first stage cost $\alpha = 1$ and second stage cost $\beta = 1.5$ paying more for the recourse action. Tables 2 and 3 show the average value over 10 simulations of the out-of-sample costs of the *SL* solutions, obtained as follows:

- For given capacity $q = 1, 5, 25$ of traveling salesmen, we uniformly generate 10 different samples of Nq customers's position in the interior or at the border of the ellipses E_k^i, $k = 1, \dots, 5$, $i = 1, \dots, 5$, which represent the realizations for the corresponding *SSOCP* model. Note that the ellipses E_k^i, $k = 1, \dots, 5$, $i = 1, \dots, 5$, are fixed in all the 10 simulations, only customers' position change (see Table 2).
- The solutions obtained by the *SL* are then evaluated in terms of *SSOCP* costs (out-of-sample analysis), by fixing *SL* solution variables $\tilde{u}$ and r obtained in each of the 10 simulations.
- The average costs over 10 simulations are then computed (see sixth and eighth column of Tables 2 and 3) and compared with the corresponding *SSOCP* formulation.

The results show that in all the cases, the solution obtained by the *SL* performs worse than the one obtained by the *SSOCP* approach, forcing the facility to be fixed in one of the customers location (see Table 1 and Figure 5). Note that, in order to reach all the customers, the service zone radius in the *SL* has to be larger (10.41 instead of 8.91). The computational times of the two approaches are not comparable since the *SSOCP* and *SL* models are implemented in different workspaces. However in both the cases, the solution is obtained in less than one CPU second.

	d_2	$\tilde{u}_1$	$\tilde{u}_2$	r	$\tilde{r}$	1^{st} stage c.	Obj. value
SSOCP	79.42	0.26	-0.09	8.91	9.06	161.60	171.32
SL	108.35	3.17	3.59	10.41	11.80	220.45	335.40

Table 1 Comparison between the *SSOCP* approach and out-of-sample evaluation of solution of the *SL* in the case of a total number of $R = 25$ customers served by $N = 5$ traveling salesmen with capacity $q = 1$ over $K = 5$ realizations.



(a) (b)

Fig. 5 Case of 25 random customers supplied by 5 traveling salesmen with capacity $q = 1$ in 5 realizations: service zone C solution of the (a) *SSOCP* model and (b) *SL*.

q	N	R	θ	ϕ	1^{st} s. *SL*	1^{st} s. *SSOCP*	Obj. *SL*	Obj. *SSOCP*
1	5	25	0.647	0.774	208.54	161.60	300.49	171.32
5	5	125	0.792	0.934	382.63	197.72	520.97	209.50
25	5	625	1.116	1.253	440.34	268.16	463.82	294.19

Table 2 Case of $N = 5$ traveling salesmen serving respectively $q = 1, 5, 25$ customers over $K = 5$ realizations: comparison between the *SSOCP* model costs (seventh and ninth columns) and average costs of the *SL* over ten simulations (sixth and eighth columns).

q	N	R	θ	ϕ	1^{st} s. *SL*	1^{st} s. *SSOCP*	Obj. *SL*	Obj. *SSOCP*
1	25	125	2.914	3.639	657.82	728.01	2049.45	772.88
5	25	625	3.238	3.958	1416.64	1368.43	1615.01	1415.74
25	25	3125	3.963	4.67	1982.71	1674.57	2036.154	1730.45

Table 3 Case of $N = 25$ traveling salesmen serving respectively $q = 1, 5, 25$ customers over $K = 5$ realizations: comparison between the *SSOCP* approach (seventh and ninth columns) and average costs of the *SL* over ten simulations (sixth and eighth columns).

5.4 Comparison on a TSP benchmark instance

We now show the results we obtained by applying the *SL* approach and the *SSOCP* approach to a TSP benchmark instance with 13,509 nodes (see TSPLIB, instance name: usa13509.tsp). We refer to [31] for more details.

The realizations are obtained by splitting the 13,509 customers in the United States into four subareas according to their longitude. Then, subsets of $q = 100$ nodes are extracted from the four subareas on the basis of a uniform distribution. Minimum covering ellipses, representing each the traveling salesman i at realization k visiting a subset of q random customers, have been generated by using MATLAB 7.4.0 environment according to [28]. Fig. 6 shows an example of one of these realizations with four traveling salesmen (the ellipses) supplying 400 customers. Bullet stars represent the customers, while bullet squares the centers of the 4 ellipses.

We consider $K = 34$ realizations, such that the total number of customers is $R = KNq = 13,509$. Table 4 shows the out-of-sample costs of the *SL* ap-

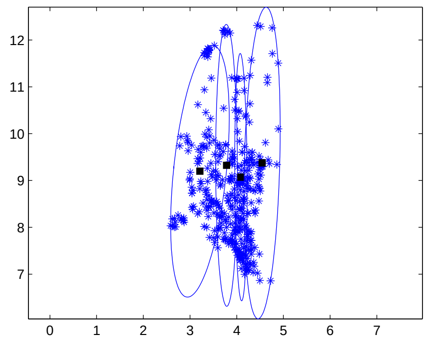


Fig. 6 Example of one realization in the usa13509.tsp instance: 400 random customers supplied by four traveling salesmen (the ellipses) with capacity $q = 100$. Bullet stars represent the centers of the 4 ellipses. Note that each customer is served by a unique traveling salesman

Approach	d_2	$\tilde{u}_1$	$\tilde{u}_2$	γ	τ	r	1^{st} s. cost	2^{st} s. cost
SL	20	3.91	8.73	71.41	0	4.47	654.66	786.37
SSOCP	14.62	4.14	9.74	97.32	0	3.82	478.46	530.22

Table 4 Comparison between the out-of-sample evaluation of the *SL* approach and the *SSOCP* approach in the usa13509.tsp instance

proach, computed as follows: different subsets of $q = 100$ possible customers are uniformly extracted from the four different longitude subareas in the United States and the corresponding minimum covering ellipses are generated. The process is repeated for the 34 realizations such that all the 13,509 customers are visited. The solution obtained by the *SL* approach is then evaluated in terms of *SSOCP* costs (out-of-sample analysis, obtained by fixing the variables obtained in the *SL* solution) and compared with the corresponding *SSOCP* total costs.

The result shows that the *SL* approach performs significantly worse than the *SSOCP* approach. In fact, the percent increase in the total cost of the *SL* approach with respect to the *SSOCP* approach is about 42.9%. This is mainly due to the fact that in the *SL* approach the facility is forced to be fixed in one of the customer locations in Indiana, instead than in a point (different than a customer location) in Nebraska, as obtained in the *SSOCP* approach (see Fig. 7).

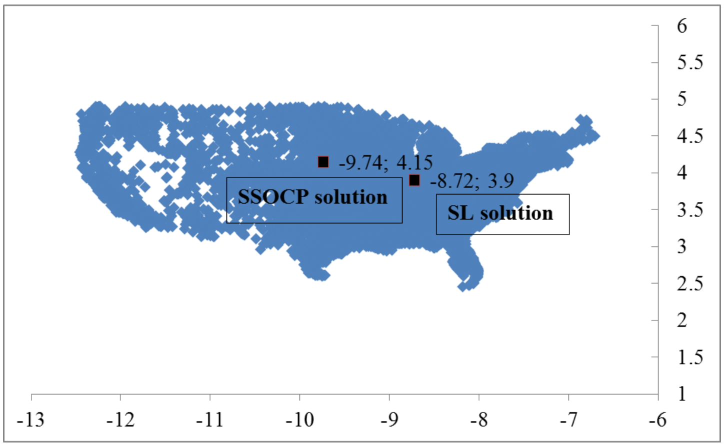


Fig. 7 Comparison of the *SSOCP* approach (-9.74,4.15) and the *SL* approach (-8.72,3.9) facility location solutions in the usa13509.tsp instance. The *SSOCP* solution refers to the case of $K = 150$ realizations, $N = 4$ traveling salesmen with capacity $q = 100$.

6 Conclusions

We studied the problem of a single facility serving a given number of customers in a given area. The position of the customers is not known. The service to the customers is carried out by several capacitated traveling salesmen. The aim is to determine the *service zone* (in a shape of a circle) that minimizes the expected cost of the traveled routes. The centre of the circle is the location of the facility. Once the position of the customers is revealed, the customers located outside the service zone are served with a recourse action at a greater unit cost. We showed that the solutions obtained by applying a simpler approach, namely *SL*, can be infinitely worse than the optimal solution. Then, we proposed a solution based on the optimal solution of a stochastic second–order cone formulation with an approximate objective function (*SSOCP* approach) and showed that this solution is significantly better than the previous solution in a large set of randomly generated problem instances.

Acknowledgements The authors wish to thank the Referees and Guest Editors for useful suggestions that allowed us to improve the paper.

References

1. Klose, A., Drexl, A.: Facility location models for distribution system design, European Journal of Operational Research, 162, 4–29 (2005)
2. ReVelle, C.S., Eiselt, H.A.: Location analysis: a synthesis and survey, European Journal of Operational Research, 165, 1–19 (2005)
3. Melo, M.T., Nickel, S., Saldanha-da-Gama, F.: Facility location and supply chain management - A review, European Journal of Operational Research, 196, 401–412 (2009)
4. Snyder, L.V.: Facility location under uncertainty: a review, IIE Transactions 38, 537–554 (2006)
5. Nagy, G., Salhi, S.: Location-routing: issues, models and methods, European Journal of Operational Research, 177, 649–672 (2007)
6. Laporte, G.: Fifty years of vehicle routing, Transportation Science, 43, 408–416 (2009)
7. Burness, R.C., White, J.A.: The traveling salesman location problem, Transportation Science, 10/4, 348–360 (1976)
8. Berman, O., Simchi-Levi, D.: Minisum location of a traveling salesman, Networks, 16, 239–254 (1986)
9. Simchi-Levi, D., Berman, O.: Heuristics and bounds for the travelling salesman location problem on the plane, Operations Research Letters, 6, 243–248 (1987)
10. Simchi-Levi, D., Berman, O.: A heuristic algorithm for the traveling saleman location problem on networks, Operation Research, 36/3 478–484

(1998)

11. McDiarmid, C.: Probability modelling and optimal location of a travelling salesman, Journal of Operational Research Society, 43/5, 533–538 (1992)
12. Simchi-Levi, D.: The capacitated traveling salesmen location problem, Transportation Science, 25/1, 9–18 (1991)
13. Bertsimas, D.J.: Traveling salesman facility location problems, Transportation Science, 23/3, 184–191 (1989)
14. Klibi, W., Lasalle, F. Martel, A. Ichoua, S.: The stochastic multi-period location-transportation problem, Transportation Science, 44, 221–237 (2010)
15. Santoso, T., Ahmed, S., Goetschalckx, M., Shapiro, A.: A stochastic programming approach for supply chain network design under uncertainty, European Journal of Operational Research, 167, 96–115 (2005)
16. Lei, H., Laporte, G., Guo, B.: Districting for routing with stochastic customers, EURO J Transp. Logist. 1, 67–85 (2012)
17. Beardwood, J., Halton, J.H., Hammersley, J.M.: The shortest path through many points, Mathematical Proceedings of the Cambridge Philosophical Society, 55/4. Cambridge University Press, 299–327 (1959)
18. Daganzo, C.F.: The distance traveled to visit n-points with a maximum of c-stops per vehicle - an analytic model and an application, Transportation Science, 18, 331–350 (1984)
19. Birge, J.R., Louveaux, F.: Introduction to stochastic programming, Springer-Verlag, New York (1997)
20. Kall, P., Wallace, S.W.: Stochastic programming, Wiley, Chichester (1994)
21. Shapiro, A.: Stochastic programming approach to optimization under uncertainty, Mathematical Programming, Ser. B, 112/1 183–220 (2008)
22. Ariyawansa, K.A., Zhu. Y.: Stochastic semidefinite programming a new paradigm for stochastic optimization, Online, Springer (2006)

23. Pataki G.: (editor). Computational semidefinite and second order cone programming: the state of the art, Springer-Verlag, Heidelberg (2003)
24. Vandenberghe L., Boyd, S.: Semidefinite programming, SIAM Review, 38, 49–95 (2004)
25. Zhu, Y., Zhang, J., Partel, K.: Location-aided routing with uncertainty in mobile ad hoc networks: A stochastic semidefinite programming approach, Mathematical and Computer Modelling, 52, 2192–2203 (2011)
26. Alizadeh, F., Goldfarb, D.: Second-order cone programming, Mathematical Programming, Ser. B, 95, 3–51 (2003)
27. Maggioni, F., Potra, F., Bertocchi, M., Allevi, E.: Stochastic second-order cone programming in mobile ad hoc networks, Journal of Optimization, Theory and Applications, 143, 309–328 (2009)
28. Sun P., Freund, R.M.: Computation of minimum-costcovering elliposoids, Operations Research, 52/5, 690–706 (2004)
29. Maggioni, F., Wallace, S.W. Bertocchi, M. Allevi, E.: Sensitivity analysis in stochastic second-order cone programming for mobile ad hoc networks, Procedia Social and Behavioral Sciences, 2, 7704–7705 (2010)
30. Kaut, M., Wallace, S.W.: Evaluation of scenario generation methods for stochastic programming, Pacific Journal of Optimization, 3/2, 257–271 (2007)
31. Bertazzi, L. Maggioni, F.: The Stochastic Capacitated Traveling Salesmen Location Problem: a computational comparison for a United States instance, Procedia Social and behavioral Sciences 108, 47–56 (2014)