

A dynamic scheme for stochastic vehicle routing

Alan L. Erera

*Georgia Institute of Technology
School of Industrial and Systems Engineering
Atlanta, GA 30332-0205*

Carlos F. Daganzo

*University of California, Berkeley
Department of Civil and Environmental Engineering
Berkeley, CA 94720-1720*

September 15, 2003

Abstract

This paper considers stochastic load-constrained vehicle routing systems in which a fleet of homogeneous vehicles is dispatched from a central depot to serve the demands of surrounding customers. Customer locations and load sizes are known only in distribution before operations. Variants of this problem are often treated with very simple uncoordinated operating schemes, which are mathematically tractable but not very flexible. This research proposes a more flexible scheme denoted *threshold global sharing* (TGS) that utilizes a real-time reoptimization control to significantly reduce costs. The TGS scheme jointly replans the operations of the entire vehicle fleet at a single decision epoch during the operating period. Exact models for determining an optimal configuration for such a scheme are intractable; thus, a continuous approximation approach is employed to generate near-optimal configurations for large-scale problems with many customers and vehicles. TGS is shown to improve substantially on current schemes.

1 Load-capacitated vehicle routing with uncertain demand

This paper considers single-period, single-depot vehicle routing systems in which a fleet of homogeneous vehicles is dispatched from a central depot to serve the demands of surrounding customers. All customers are served during the period, and vehicles return to the depot after their tasks are completed. Such a system is *load-capacitated* if the set of tasks each vehicle can complete is constrained primarily by the size of the maximum load that each vehicle can transport.

In *stochastic load-capacitated vehicle routing systems*, customer locations and/or shipment demands are not known with certainty during planning. Many real-world collection and distribution routing problems are best modeled as stochastic load-constrained systems (see Erera (2000) for application examples). Operators of such systems must design an operating *scheme* for the vehicle fleet that serves all customers with minimum long-run average cost. Transportation cost is typically modeled as a non-decreasing function of the number of vehicles dispatched and the total distance traveled.

This paper considers a collection system where customer locations and load sizes are known only in distribution before operations. Variants of this problem are often treated with very simple uncoordinated operating schemes, which are mathematically tractable but not very flexible. This research proposes a more flexible scheme denoted *threshold global sharing* (TGS) that utilizes a real-time reoptimization control to significantly reduce costs. The TGS scheme jointly replans the operations of the entire vehicle fleet at a single decision epoch during the operating period. Exact models for determining an optimal configuration for such a scheme are intractable; thus, a continuous approximation approach is employed to generate near-optimal configurations for large-scale problems with many customers and vehicles. TGS is shown to improve substantially on current schemes.

The remaining sections of this paper are organized as follows. Section 2 classifies operating schemes for stochastic vehicle routing systems, and reviews relevant research. Section 3 then defines the TGS operating scheme. Section 4 defines the scheme configuration problem, and develops an idealized problem setting for the subsequent analysis. Section 5 develops an approximation model for the standard detour-to-depot scheme, and Section 6 for the TGS scheme. Section 7 briefly discusses a simulation model used to validate the approximation model. Finally, section 8 compares results for the two schemes on randomized problems.

2 Classification of operating schemes

Determining an effective fleet operating scheme is typically very difficult for stochastic vehicle routing systems. The characteristics of these systems that create challenges include:

1. *Customer service*: Service constraints that guarantee that all customers are completely served during the operating period can often only be enforced probabilistically when demands are uncertain.

2. *Real-time control*: It may be very costly or infeasible to execute preplanned vehicle operations that meet service requirements. Thus some form of real-time vehicle control is typically necessary.
3. *Information dynamics*: Uncertain customer locations and shipment sizes are revealed dynamically during the operating period; the availability of information affects how and when decisions can be made.

Schemes without provision for real-time control are denoted *no-control* schemes. Consider this example: vehicle tours are planned serving all potential customer locations prior to the operating period, and each tour is operated as planned until completion or until the occurrence of a *failure* (*i.e.*, a vehicle unable to serve a planned customer due to insufficient remaining capacity). Upon failure, a vehicle returns to the depot; remaining customers are left unserved. Stewart and Golden (1983) and Laporte et al. (1989) propose chance-constrained probabilistic programming approaches for the optimal configuration of such schemes.

Of course, very few real-world systems could operate without some form of real-time control, since they would likely fail with undesirable frequency, or require exorbitant cost. One form of real-time control is to plan *recovery* operations after failures occur. Early research in Stewart and Golden (1983) and Dror and Trudeau (1986) explores the use of penalty costs to approximate the expected additional costs required by recovery operations.

More recent research attempts to model real-time control decisions and costs more explicitly. Since operating schemes that utilize real-time control vary widely in both complexity of implementation and analysis, it is useful to propose a simple categorization. The two most important differentiators specify *when* and *how* controls are applied:

Control epochs. Real-time control decisions are determined and applied at *control epochs* during the operating period. Such epochs may be determined by events, such as a vehicle becoming full. Alternately, controls may be applied at predetermined time epochs. Mixed designs are also possible. The symbols $t = E, T$, and M are used respectively to differentiate such designs.

Vehicle coordination. Controls typically do not replan all vehicle operations jointly at each epoch. A p -coordinated control indicates that replanning is performed for independent vehicle subsets of size p . For example, a scheme in which vehicles are controlled in pairs is 2-coordinated.

It is convenient to refer to a scheme by the symbol combination t/p . For example, an event-based control in which vehicles are controlled in independent sets of size 2 is an $E/2$ scheme. Most research in stochastic vehicle routing addresses the problem of optimally-configuring $E/1$ schemes in which each vehicle is controlled independently of all others. Extending work on probabilistic traveling salesman problems (see Jaillet (1988 and 1993)), Bertsimas (1992) and Bertsimas and Howell (1993) consider a single-vehicle stochastic vehicle routing system with uncertain customer locations and demands, and are the first to propose the $E/1$ detour-to-depot operating scheme. Under detour-to-depot, an *a priori* tour ordering the potential customer locations is determined before operations.

The real-time control then specifies that the vehicle serve these locations in order, detouring to the depot to unload/reload whenever capacity is reached. The papers develop asymptotic worst-case bounds for efficient heuristics, and present an asymptotically-optimal heuristic.

Building on this work, researchers have investigated this scheme for systems with multiple vehicles. Since the real-time control in detour-to-depot is specified in a set of predetermined rules that affect each vehicle independently, such schemes can be modeled using two-stage stochastic integer programming methods. Following Laporte and Louveaux (1993) and Laporte and Nobert (1980), Gendreau et al. (1995) develops a method for determining *a priori* tours that minimize expected travel costs for the detour-to-depot scheme. Computational results are limited to problems with two vehicles. Gendreau et al. (1996) develops a tabu search heuristic for the problem. Recent computational work in Verweij et al. (2003) on the sample average approximation solution approach applied to related problems may provide a method for developing solutions to problems with larger vehicle fleets.

$E/1$ operating schemes are appealing for their simplicity, since vehicles can operate independently in real-time with minimal communication requirements, but they may not be as efficient as $*/p$ schemes where $p > 1$. By coordinating groups of vehicles, we create beneficial *risk-pooling* effects that allow smaller vehicle fleets to be used. These benefits increase with p ; therefore, we would like to pool the capacity of the entire fleet of m vehicles. Such coordination should be possible given current advances in communication technology.

Many $*/m$ system schemes are conceivable; the key is to develop a scheme that provides the risk-pooling benefit without complex coordination requirements. This research proposes a simple E/m scheme called *threshold global sharing* (TGS).

3 Threshold global-sharing scheme

The primary idea of the TGS scheme is to partition the service region *a priori* into two subregions, *region 1* and *region 2*. Region 1 is defined as all potential customer locations located beyond a *threshold* distance r_T from the depot, while region 2 consists of the locations within the threshold.

Potential customer locations in region 1 are each uniquely assigned to a *phase 1* vehicle route, and each vehicle in the fleet operates such a route. All customers in region 2 are left unassigned during phase 1. A phase 1 route is completed when either all customers on the route have been served, or when the vehicle has reached capacity. When all phase 1 routes have been completed, a single real-time control decision is made to assign each unserved customer to a vehicle with remaining capacity to create *phase 2* routes. Note that vehicles only return to the depot when they reach capacity or they complete their phase 2 route. Any customers that remain unserved after phase 2 are served by *phase 3* overflow tours from the depot.

When configured near-optimally, the TGS scheme should result in very efficient m -vehicle risk pooling. Phase 2 routes should require little additional travel distance, since vehicles should only need minor lateral detours to serve customers en route back to the depot. Additionally, vehicles

that reach capacity are more likely to do so during phase 2. Thus, customers requiring service by a phase 3 tour will be located close to the depot, and can be served with minimal additional travel distance.

Like any $*/m$ scheme, TGS is difficult to analyze exactly. Therefore, approximations will be used to configure the scheme near-optimally.

4 Problem specification

Load-capacitated collection vehicle routing systems serve customers located in a service region \mathbf{A} with a fleet of m vehicles that remains fixed across operating periods. Customers require freight transport service to and/or from a depot location \mathbf{x}_0 . During a single operating period (*e.g.*, a day), each customer i requires pickup of an unsplittable shipment of q_i units. Each depot-based vehicle has load capacity to transport Q units at a time. The total transportation cost accrued by the system during an operating period, c_T , is given by:

$$c_T = c_v m + c_d d_T \tag{1}$$

where c_v and c_d are non-negative, and d_T is the total distance traveled by all vehicles.

The uncertainty aspects in our problem are identical to those in Gendreau et al. (1995). Thus, we assume the following:

Customer locations: The set of customers requiring service during each operating period is not known with certainty until immediately prior to vehicle dispatch.

Customer shipments: The shipment size q_i of each customer i is not known with certainty until the arrival of a vehicle.

Problem solutions for such systems specify a fleet size m and configure an operating scheme. This research considers *full-service* schemes in which the demand of each customer is served fully with very high probability. An optimal solution will be a scheme with minimum expected total cost per operating period:

$$C_T \equiv E[c_T] = c_v m + c_d D_T \tag{2}$$

where $D_T \equiv E[d_T]$. Risk neutral cost expectations are used, since it is assumed that the systems will be operated for many consecutive operating periods.

4.1 Idealized problem

The analysis in this research uses a problem idealization in order to simplify the discussion, but the framework can be applied to real-world systems. Suppose as depicted in Figure 1 that \mathbf{A} is a disk in the plane with radius R centered at the origin. Define $|\mathbf{A}|$ to be the area of \mathbf{A} , and let

$A = |\mathbf{A}|$. Suppose that the depot is located at the origin: $\mathbf{x}_0 = (0, 0)$. Transportation distances between points are given by a ring-radial distance function:

$$d(\mathbf{x}_1, \mathbf{x}_2) = \min\{r_1 + r_2, |r_1 - r_2| + \min\{r_1, r_2\} \min\{|\theta_1 - \theta_2|, 2\pi - |\theta_1 - \theta_2|\}\} \quad (3)$$

This distance function is a good approximation of systems where local travel between nearby points occurs on a street grid while line-haul travel to and from the depot is more direct.

Suppose customer locations are uniformly scattered in the region in every operating period, and that customer shipment sizes V_i are independent and identically-distributed random variables with mean μ and variance σ^2 . Locations are therefore modeled as the realization of a homogeneous spatial Poisson process with rate δ (*customers/area*). Demands can then be modeled using a general spatial stochastic process Λ with independent, nonnegative increments characterized by a rate λ (*units/area*) and a dispersion γ (*units*):

$$\lambda = \mu\delta \quad (4)$$

$$\gamma = \left(\mu + \frac{\sigma^2}{\mu} \right) \quad (5)$$

The total customer shipment size demand in subregion $\mathbf{A}_Z \subseteq \mathbf{A}$ can thus be approximated by a random variable $V(\mathbf{A}_Z)$ with mean and variance given by:

$$E[V(\mathbf{A}_Z)] = \lambda A_Z \quad (6)$$

$$\text{var}(V(\mathbf{A}_Z)) = \gamma \lambda A_Z \quad (7)$$

where $A_Z = |\mathbf{A}_Z|$. We will assume that $V(\mathbf{A}_Z)$ is a normal random variable, which is reasonable in many applications.

5 Approximations for detour-to-depot scheme

This section develops an approximation of the expected cost of the detour-to-depot operating scheme (rule ‘a’ in Bertsimas (1988)) for problems with many customers and vehicles. The approximation can be used to configure the scheme near-optimally.

We begin by modifying the analysis presented for vehicle routing problems where vehicles can make a fixed number of stops in Daganzo (1984a) and Daganzo (1984b). In the context of this paper, the model in these references corresponds to the case where customer lot sizes are deterministic. The references show that near-optimal solutions to the VRP are obtained by a homogeneous partition of the service region into similar, approximately rectangular *zones* \mathbf{A}_Z with length ℓ and width w , each served by a single vehicle. It is assumed that $\ell \geq w$ and that the long side of every zone points toward the depot. A vehicle first serves customers in order of increasing distance from the depot in a swath of width $\frac{w}{2}$, then turns back toward the depot and serves the remaining customers in order of decreasing distance; such a tour is called an out-and-back tour. Refer to Figure 2.

When load sizes are random, a similar result can be obtained for the detour-to-depot scheme. We now show that if we use a homogeneous partition with optimally-elongated zones of size A_Z then the expected total distance of the detour-to-depot scheme, D_T^{DD} , is given by a simple formula.

Proposition 1 (Expected Detour-to-Depot Distance)

$$D_T^{DD}(A_Z) \approx \frac{4\pi R^3}{3} \frac{f_E(A_Z)}{A_Z} + \sqrt{\frac{2}{3}} \delta^{1/2} \pi R^2 \quad (8)$$

where $f_E(A_Z)$ is the expected number of vehicle trips required to serve one zone.

Proof. Daganzo (1984a) shows that if the total demand in \mathbf{A}_Z is no greater than Q , then the expected distance per period D_t^{VRP} of a tour through \mathbf{A}_Z under an L_1 metric is:

$$D_t^{\text{VRP}} \approx \begin{array}{ll} 2r + \ell & + \frac{w}{6}(\delta w \ell) \\ \text{(radial)} & \text{(lateral)} \end{array} \quad (9)$$

where r is the distance from the center of \mathbf{A}_Z to the depot. However, since total customer demand may exceed Q , detours to the depot may be required.

Let B be the random number of detours for \mathbf{A}_Z in a particular period. Since detours to the depot do not add any lateral distance, a reasonable approximation is that each detour adds $2r$ distance units. Therefore the approximate total expected tour distance is:

$$D_t^{DD} \approx 2r(E[B] + 1) + \ell + \frac{w}{6}(\delta w \ell) \quad . \quad (10)$$

Since our problem is spatially-homogeneous, the expected number of vehicle trips, $E[B]+1$, required to serve the demand in the zone is a function of its area: $f_E(A_Z)$. Therefore, the expected distance per unit area per period, $D_a^{DD} = \frac{D_t^{DD}}{A_Z}$, is:

$$D_a^{DD} \approx 2r \frac{f_E(A_Z)}{A_Z} + \frac{1}{w} + \frac{\delta w}{6} \quad . \quad (11)$$

For fixed A_Z , there is a zone width w that minimizes (11). The optimum is $w^{DD*} = \sqrt{\frac{6}{\delta}}$, which is independent of A_Z and Q , and is therefore the same as in the deterministic case. Substituting into (11) yields the optimal expected distance:

$$D_a^{DD*}(A_Z) \approx 2r \frac{f_E(A_Z)}{A_Z} + \sqrt{\frac{2}{3}} \delta^{1/2} \quad . \quad (12)$$

Since region \mathbf{A} is partitioned into zones of roughly equal sizes and shapes, like \mathbf{A}_Z , then the sum of (12) across the service region is an approximation for the total expected distance per operating period:

$$D_T^{DD}(A_Z) \approx \left(2 \int_{\mathbf{A}} r(\mathbf{x}) d\mathbf{x} \right) \frac{f_E(A_Z)}{A_Z} + \sqrt{\frac{2}{3}} \delta^{1/2} A \quad (13)$$

where $r(\mathbf{x})$ is the distance from \mathbf{x} to the depot. For our homogeneous disk, the parenthetical expression is simply $\frac{4\pi R^3}{3}$. The result in the proposition follows by substitution. \square

Since we use one vehicle per zone, the total fleet size required under this scheme is:

$$m^{DD}(A_Z) \approx \frac{\pi R^2}{A_Z} \quad (14)$$

and the total expected cost per period is:

$$C_T^{DD}(A_Z) \approx c_v \frac{\pi R^2}{A_Z} + c_d \frac{4\pi R^3}{3} \frac{f_E(A_Z)}{A_Z} + c_d \sqrt{\frac{2}{3}} \delta^{1/2} \pi R^2 \quad . \quad (15)$$

For most reasonable parameter values, (15) will be a decreasing function in A_Z , and thus achieves its minimum at $A_Z^* = \pi R^2$; *i.e.*, for a single vehicle. This finding is consistent with Bertsimas (1988), which shows that a single-vehicle detour-to-depot operating strategy is asymptotically optimal as the number of customers grows large. Unfortunately, however, these results are only of mathematical interest since most systems cannot be operated with many detours for obvious practical reasons.

A more reasonable approach would constrain the probability that a single vehicle trip is sufficient to fully serve a zone \mathbf{A}_Z :

$$\Pr \{V(\mathbf{A}_Z) \leq Q\} \geq (1 - p_\alpha) \quad (16)$$

where p_α is a failure probability. By choosing a small p_α , we can limit the number of detours. Under the normality assumption, (16) becomes:

$$\lambda A_Z + \alpha \sqrt{\gamma \lambda A_Z} \leq Q \quad (17)$$

where $\alpha = \Phi^{-1}(1 - p_\alpha)$.¹

The constrained minimum of $C_T^{DD}(A_Z)$ is now achieved when (17) holds with equality. Therefore:

$$A_Z^*(\alpha) = \frac{Q}{\lambda} - \frac{\alpha^2 \gamma}{2\lambda} \left(\sqrt{1 + \frac{4Q}{\alpha^2 \gamma}} - 1 \right) \quad . \quad (18)$$

The first term $\frac{Q}{\lambda}$ is the optimal area for a problem with deterministic lot sizes ($\gamma \rightarrow 0$), and the second term is the reduction in service zone size due to uncertainty.

For small values of p_α (*i.e.*, $p_\alpha < 0.3$), more than one detour is unlikely. Thus, a reasonable approximation for f_E is:

$$f_E(A_Z) \gtrsim 1 + p_\alpha = 1 + \Phi(-\alpha) \quad . \quad (19)$$

Expressions (18), (19), (14), and (15) can be used to approximate the optimal fleet size and the resulting optimal total expected transportation cost for \mathbf{A} .

To provide a conservative estimate of the benefits of the TGS scheme, we now present a simple lower bound \underline{D}_T^{DD} for the expected total distance required by detour-to-depot for the case where (17) holds with equality.

¹In this paper, ϕ is the standard normal density and Φ is the standard normal cumulative distribution.

Proposition 2 (Expected Detour-to-Depot Distance Lower Bound)

$$\underline{D}_T^{DD*} \approx \frac{4\pi R^3}{3} \frac{(1+p_\alpha)}{\bar{A}_Z} + \sqrt{\frac{2}{3}} \delta^{1/2} \sqrt{1-p_\alpha} \pi R^2 \quad (20)$$

where

$$p_\alpha = \Phi \left(\frac{\lambda \bar{A}_Z - Q}{\sqrt{\lambda \gamma \bar{A}_Z}} \right). \quad (21)$$

Proof. Simply repeat the derivation of Proposition 1 assuming that a zone requires a single detour with probability p_α and that the detour trip only travels to the depot from the edge of the zone closest to the depot and back. In this case, the right-hand side of (10) becomes $2r(1+p_\alpha) + (1-p_\alpha)\ell + \frac{w}{6}\delta w\ell$, and the optimal zone width becomes $\sqrt{\frac{6}{\delta(1-p_\alpha)}}$. The result follows. \square

6 Approximation model for TGS scheme

Figure 3 depicts the TGS scheme graphically for the idealized problem. In this case, region 1 is the ring $\{\mathbf{x} \in \mathbf{A} : r_T \leq r(\mathbf{x}) \leq R\}$, while region 2 is the disk $\{\mathbf{x} \in \mathbf{A} : r(\mathbf{x}) < r_T\}$. Configuration of the TGS scheme requires specification of procedures for:

1. Determination of phase 1 routes for each vehicle
2. Real-time assignment of phase 2 customers to eligible vehicles and determination of phase 2 routes
3. Determination of phase 3 overflow tours

In this analysis, we introduce simple procedures for these three tasks. Based on these procedures, a total expected distance function of m and r_T is developed which is then minimized to determine near-optimal values of these configuration parameters. By insisting on deriving approximate formulae, we acknowledge that we use suboptimal procedures but this gains us the ability to design the overall system in a more systematic way.

To do this, we decompose the vehicle trip of a single vehicle, as depicted in Figure 4. The vehicle first travels radially from the depot to its region 1 zone, and second serves customers in this zone. Third, the vehicle travels to the threshold radius, and fourth, is repositioned before beginning its phase 2 route. Fifth, the vehicle may serve some region 1 customers that remain unserved due to other vehicle failures. Sixth and finally, the vehicle visits region 2 customers on its return trip to the depot. If the vehicle loads to capacity Q at any time, it immediately returns directly to the depot. The following subsections present formulae for these distance components.

6.1 Phase 1 tour distance

Region 1 is partitioned into m zones of equal size that remain fixed over all operating periods. When customer locations become known immediately prior to dispatch, a phase 1 route is generated for each vehicle within each zone. These zones are designed as in Daganzo (1984a). Thus, zones are nearly rectangular with width $\sqrt{\frac{6}{\delta}}$, and vehicles operate out-and-back tours in swaths of width $\sqrt{\frac{3}{\delta}}$. For our homogeneous problem, the area of these zones is $A_{R1} = \frac{\pi(R^2 - r_T^2)}{m}$.

We now calculate D_T^{R1} , the expected total travel distance for all vehicles to travel from the depot to their region 1 zones, serve all region 1 customers as if $Q = \infty$, and then return to the depot:

Proposition 3 (Expected Phase 1 Tour Distance - Assuming $Q = \infty$)

$$D_T^{R1} \approx \frac{4m(R^3 - r_T^3)}{3(R^2 - r_T^2)} + \sqrt{\frac{2}{3}}\pi\delta^{1/2}(R^2 - r_T^2) \quad (22)$$

Proof. Region 1 is partitioned into equal-area routing zones to minimize the maximum probability that an individual vehicle overflows. For a zone of area A_{R1} , the total expected tour length is:

$$D_t^{R1} \approx 2r + A_{R1}\sqrt{\frac{2}{3}}\delta^{1/2} \quad (23)$$

where r is the distance from the depot to the zone centroid. For equal-area zones of size A_{R1} , summing (23) for all vehicles yields:

$$D_T^{R1} \approx 2m\bar{r}_{R1} + \sqrt{\frac{2}{3}}\pi\delta^{1/2}(R^2 - r_T^2) \quad (24)$$

where \bar{r}_{R1} is the average distance from the depot to a zone centroid. For our spatially-homogeneous problem, \bar{r}_{R1} is approximately the average distance between a region 1 customer and the depot:

$$\bar{r}_{R1} \approx \frac{2(R^3 - r_T^3)}{3(R^2 - r_T^2)} \quad (25)$$

Substitution of this value into (24) provides the result. □

6.2 Real-time control: assignment and routing procedures

Vehicles with remaining capacity at the end of phase 1 are assigned to a set of phase 2 customers and new routes are created. This subsection describes these real-time control procedures, and the following subsections develop the corresponding distance formulae.

Let q_k^{R1} be the total units of demand collected from region 1 customers by vehicle k , and let M_R be the set of available vehicles for which $Q - q_k^{R1} > 0$. Further, let $\xi_k \forall k \in M_R$ be the polar coordinate angle of the final customer in the phase 1 route for vehicle k ; suppose for clarity that these angles are unique. Let M_F be the set of region 1 zones with unserved customers. Let φ_k be

the polar angle of the centroid of the remaining customers in $k \in M_F$, and let o_k^{R1} be their overflow total demand. Note that we assume in the analysis to follow that o_k^{R1} is known with certainty. This assumption is made for clarity only. The reassignment procedures in this research rely on cumulative curves of remaining vehicle capacity.

Definition 1 (Cumulative Capacity Functions)

1. The available capacity function $x_A(\theta)$ is given by:

$$x_A(\theta) = \sum_{k \in M_R | \xi_k \leq \theta} Q - q_k^{R1} \quad \forall \theta \in [0, 2\pi)$$

2. The overflow demand function x_F is:

$$x_F(\theta) = \sum_{k \in M_F | \varphi_k \leq \theta} o_k^{R1} \quad \forall \theta \in [0, 2\pi)$$

3. The net remaining capacity cumulative function y_R is:

$$y_R(\theta) = x_A(\theta) - x_F(\theta) \quad \forall \theta \in [0, 2\pi]$$

4. The mean remaining capacity cumulative function $\bar{y}_R(\theta)$ is:

$$\bar{y}_R(\theta) = \frac{y_R(2\pi)}{2\pi} \theta \quad \forall \theta \in [0, 2\pi)$$

Figure 5 depicts an example y_R , where a smooth curve has been superimposed as an approximation of the step function y_R .

Reassignment procedure for overflow region 1 customers

Region 1 customers with remaining demand are first assigned to vehicles using a greedy procedure based on curves y_R and \bar{y}_R . Down steps in y_R identify angles where vehicles are needed to serve overflows, and up steps identify angular positions of vehicles with available capacity. The goal of a procedure is to assign vehicles to overflows such that the resulting net remaining capacity curve is monotonic. A reasonable method to do so is assign close vehicles to overflows such that vehicles are repositioned in the direction of remaining capacity balance, *i.e.*, towards \bar{y}_R .

Although it will likely be rare, some vehicles with remaining capacity may become completely full after assignment to overflow customers. Let M_R^U represent this vehicle set. Consider now the following algorithm:

Algorithm 1 (Greedy Region 1 Overflow Reassignment)

1. If no overflow region 1 zones exist, stop.
2. Otherwise:
 - (a) Order zones $k \in M_F$ by nondecreasing φ_k : $\{k_1, k_2, \dots, k_{|M_F|}\}$.
 - (b) Let $M_R^U = \emptyset$.

3. Forward assignment

- (a) Let $i = 1$.
- (b) If $y_R(\varphi_{k_i}) \geq \bar{y}_R(\varphi_{k_i})$, then:
 - i. Let $\theta_{\max} = \varphi_{k_i}$.
 - ii. Let ℓ^* be the solution to $\min_{\ell \in M_R} \theta_{\max} - \xi_\ell$ such that $y_R(\xi_\ell) \geq y_R(\varphi_{k_i})$.
 - iii. Assign vehicle $\ell \in M_R \setminus M_R^U$ to overflow if $\xi_{\ell^*} \leq \xi_\ell \leq \theta_{\max}$.
 - iv. Add vehicle $\ell \in M_R \setminus M_R^U$ to the full vehicle set M_R^U if:
 - A. $\xi_{\ell^*} < \xi_\ell \leq \theta_{\max}$, or
 - B. $\ell = \ell^*$ and $y_R(\xi_{\ell^*}) = y_R(\varphi_{k_i})$.
- (c) Let $i = i + 1$. If $i > |M_F|$, go to Step 4. Else, go to Step 3b.

4. Backward assignment

- (a) Let $i = |M_F|$.
- (b) If $y_R(\varphi_{k_i}) < \bar{y}_R(\varphi_{k_i})$, then:
 - i. Let $\theta_{\min} = \varphi_{k_i}$.
 - ii. Let ℓ^* be the solution to $\min_{\ell \in M_R} \xi_V - \theta_{\min}$ such that $y_R(\xi_\ell) \geq y_R(\varphi_{k_i}) + o_{k_i}^{R1}$.
 - iii. Assign vehicle $\ell \in M_R \setminus M_R^U$ to overflow if $\theta_{\min} \leq \xi_\ell \leq \xi_{\ell^*}$.
 - iv. Add vehicle $\ell \in M_R \setminus M_R^U$ to the full vehicle set M_R^U if:
 - A. $\theta_{\min} \leq \xi_\ell < \xi_{\ell^*}$, or
 - B. $\ell = \ell^*$ and $y_R(\xi_{\ell^*}) = y_R(\varphi_{k_i}) + o_{k_i}^{R1}$.
- (c) Let $i = i - 1$. If $i = 0$, Stop. Else, go to Step 4b.

The result of this procedure on an example is depicted graphically in Figure 6, where vertical segments are added to the step function $y_R(\theta)$. Vehicle A is assigned to overflow k_1 in the forward assignment pass. In the backward pass vehicle C is assigned to k_3 , and then B and C are both assigned to k_2 .

Assignment procedure for region 2 customers

Let q_k^1 be the total units of demand for the overflow region 1 customers assigned to vehicle $k \in M_R$ by Algorithm 1. Then the remaining capacity of vehicle k to serve region 2 customers is:

$$Q_k^R = Q - q_k^{R1} - q_k^1 \quad ,$$

and the total remaining capacity of all vehicles Q_T^R is:

$$Q_T^R = \sum_{k \in M_R} Q_k^R \quad .$$

To ensure a high probability that vehicles do not overflow during phase 2, we assign a number of customers to each vehicle in proportion to its relative remaining capacity, that is, if N_R^2 is the set of region 2 customers, we assign

$$n_k^2 = \frac{Q_k^R}{Q_T^R} |N_R^2|$$

to vehicle k . This assignment algorithm is then:

Algorithm 2 (Region 2 Sweep Assignment)

1. Order vehicle set M_R by nondecreasing ξ_k , and let the vehicle order be $\{k_1, k_2, \dots, k_{|M_R|}\}$.
2. Let $i = 1$, and $\theta^S = \theta^E = 0$.
3. while $i < |M_R|$ do:
 - (a) Let $n = \lfloor n_{k_i}^2 + \frac{1}{2} \rfloor$
 - (b) Increase θ^E until n customers in N_R^2 have polar angle in $[\theta^S, \theta^E)$; assign these customers to vehicle k_i .
 - (c) Let $\theta_S = \theta_E$, $i = i + 1$.
4. Assign customers in N_R^2 with polar angle in $[\theta_S, 2\pi)$ to vehicle $k_{|M_R|}$.

Routing procedure for phase 2

Given the assignments made by Algorithms 1 and 2, a phase 2 tour for each vehicle is created as follows. First, if the vehicle is assigned any unserved region 1 customers, they are added to the beginning of its phase 2 tour in order of decreasing radial distance to the depot; when the probability of phase 1 zone failures is small, these customers are served in the original planned order. Second, assigned region 2 customers are added to the tour in order of decreasing radius. The vehicle serves customers in the specified order beginning from its position along the circle r_T , and then returns to the depot.

6.3 Repositioning distance

Most vehicles will require some circumferential repositioning distance before reaching their first phase 2 customer. Suppose conservatively that this travel is conducted along the circumference at the threshold radius, r_T .

We now approximate the expected total repositioning distance D_T^{RE} using the curves y_R and \bar{y}_R :

Proposition 4 (Expected Repositioning Distance)

$$D_T^{RE} \approx (1 - p_F^{R1}) r_T \frac{\sigma_R}{\mu_R} \sqrt{\frac{m\pi^3}{8}} \left(1 + \frac{\sigma_R^2}{m\mu_R^2} \right) \quad (26)$$

where $\mu_R = Q - \lambda A_{R1}$, $\sigma_R^2 = \gamma \lambda A_{R1}$, and $p_F^{R1} = \Phi\left(-\frac{\mu_R}{\sigma_R}\right)$, the probability that a region 1 zone exceeds vehicle capacity.

Proof. The expected number of vehicles used in phase 2 is $(1 - p_F^{R1})m$. Since each vehicle travels along the threshold circumference, the total expected repositioning distance is:

$$D_T^{\text{RE}} = (1 - p_F^{R1})mr_T E[\overline{H}_\theta] \quad (27)$$

where $E[\overline{H}_\theta]$ is the expected angle traversed by each vehicle. To approximate this quantity, consider a smoothed realization of y_R and \overline{y}_R , as shown in Figure 7. The labeled points on the figure represent sample vehicles. Algorithms 1 and 2 move vehicles in the direction of balanced capacity during assignment. Thus, a remaining cumulative capacity curve constructed *after* vehicles have been repositioned should resemble \overline{y}_R closely, and $E[\overline{H}_\theta]$ can be approximated by the expected mean angular separation $E[\widehat{H}_\theta]$ between y_R and \overline{y}_R .

For any realization, the mean angular separation between these curves is:

$$\widehat{H}_\theta = \frac{1}{y_R(2\pi)} \int_0^{2\pi} |y_R(x) - \overline{y}_R(x)| dx \quad (28)$$

which is simply the area between the curves divided by the total remaining capacity in items. Let V_k^{R1} be the random total demand of all region 1 customers assigned to vehicle k ; V_k^{R1} has the distribution of $V(A_{R1})$. Suppose now that y_R is the realization of a stochastic process, $Y_R(\theta)$, which sums $Q - V_k^{R1}$ for all k whose region 1 zone centroid angle lies in $[0, \theta]$. Clearly, $Y_R(\theta)$ is normal for all θ , with the following mean and variance:

$$\begin{aligned} E[Y_R(\theta)] &= \frac{\theta m}{2\pi} \mu_R \\ \text{var}(Y_R(\theta)) &= \frac{\theta m}{2\pi} \sigma_R^2 \end{aligned}$$

We now calculate $E[\widehat{H}_\theta]$ by conditioning on $Y_R(2\pi)$:

$$E[\widehat{H}_\theta] = E[E[\widehat{H}_\theta|Y_R(2\pi)]] = E\left[\frac{1}{Y_R(2\pi)} E[Z_Y|Y_R(2\pi)]\right] \quad (29)$$

where $Z_Y = \int_0^{2\pi} |Y_R(x) - \overline{Y}_R(x)| dx$ and $\overline{Y}_R(x) = \frac{Y_R(2\pi)}{2\pi}x$. The conditional expected area $E[Z_Y|Y_R(2\pi)]$ can be determined using the conditional process, $Y_R(\theta)|Y_R(2\pi)$. First, recognize that $Y_R(\theta)|Y_R(2\pi)$ is normally-distributed for all θ (see Appendix for more detail) with mean and variance:

$$E[Y_R(\theta)|Y_R(2\pi)] = \frac{Y_R(2\pi)\theta}{2\pi} \quad (30)$$

$$\text{var}(Y_R(\theta)|Y_R(2\pi)) = m\sigma_R^2 \frac{\theta}{2\pi} \left(1 - \frac{\theta}{2\pi}\right) \quad (31)$$

Since $E[Y_R(\theta)|Y_R(2\pi)] = \overline{Y}_R(\theta)$, $E[Z_Y|Y_R(2\pi)]$ simplifies to:

$$E[Z_Y|Y_R(2\pi)] = \int_0^{2\pi} E\left[|Y_R(x)|Y_R(2\pi) - E[Y_R(\theta)|Y_R(2\pi)]\right] dx \quad (32)$$

where the integrand is the mean absolute deviation of a normal random variable for all x . Letting $\sigma_{Y_R|Y_R(2\pi)} = \sqrt{\text{var}(Y_R(\theta)|Y_R(2\pi))}$, this integral reduces to:

$$\begin{aligned} E[Z_Y|Y_R(2\pi)] &= \int_0^{2\pi} \sqrt{\frac{2}{\pi}} \sigma_{Y_R|Y_R(2\pi)} dx \\ &= \sigma_R \sqrt{\frac{m\pi^3}{8}} . \end{aligned} \quad (33)$$

We can complete an approximation for $E[\widehat{H}_\theta]$ using (33) and an estimate for $E\left[\frac{1}{Y_R(2\pi)}\right]$. To do so, let $S = Y_R(2\pi)$ and expand $\frac{1}{S}$ in a Taylor series about $E[S]$. Using the first three terms, we have:

$$E\left[\frac{1}{S}\right] \approx \frac{1}{E[S]} + \frac{\text{var}(S)}{E[S]^3} ,$$

therefore:

$$E[\overline{H}_\theta] \approx E[\widehat{H}_\theta] \approx \frac{\sigma_R}{m\mu_R} \sqrt{\frac{m\pi^3}{8}} \left(1 + \frac{\sigma_R^2}{m\mu_R^2}\right) . \quad (34)$$

Substituting (34) into (27) yields the proposed result. \square

6.4 Extra radial distance in region 1

After repositioning, some vehicles serve overflow region 1 customers. The lateral distance required to serve these customers is already included in D_T^{R1} . Therefore, we now give an approximation for the expected radial distance D_T^{OF1} :

Proposition 5 (Expected Overflow Radial Distance in Region 1)

$$D_T^{OF1} \approx 2p_F^{R1} m \left(1 + e^{-0.2097 - 1.0067\alpha_Q - 0.1325\alpha_Q^3 + 0.0035\alpha_Q^5}\right) \left(\frac{2(R^3 - r_T^3)}{3(R^2 - r_T^2)} - r_T\right) \quad (35)$$

where $\alpha_Q = \frac{\mu_R}{\sigma_R}$.

Proof. Region 1 zones are equally likely to require overflow service, therefore the approximate expected extra radial distance per vehicle is twice the difference between the average region 1 zone centroid radius and the threshold radius.

To approximate the expected number of vehicles required to serve region 1 overflow zones, $E[m^{OF1}]$, we note that:

$$E[m^{OF1}] = p_F^{R1} m E[m_Z^{OF1}] \quad (36)$$

where $E[m_Z^{OF1}]$ is the expected number of vehicles required to serve a single overflow zone. The excess demand of a region 1 zone is $\max\{V(A_{R1}) - Q, 0\}$ and the remaining capacity of a vehicle is $\max\{Q - V(A_{R1}), 0\}$. Using the conditional distributions of these quantities for $V(A_{R1}) > Q$ and $V(A_{R1}) < Q$, Erera (2000) uses a Monte Carlo approach to estimate $E[m_Z^{OF1}]$ yielding:

$$E[m_Z^{OF1}] \approx 1 + e^{-0.2097 - 1.0067\alpha_Q - 0.1325\alpha_Q^3 + 0.0035\alpha_Q^5} . \quad (37)$$

□

6.5 Lateral distance in region 2

Algorithm 2 creates pie-shaped zones partitioning the disk inside r_T . The radial distance required to serve these zones is again already included in D_T^{R1} . We now show that the expected lateral distance D_T^{R2} required is given by:

Proposition 6 (Expected Lateral Distance in Region 2)

$$D_T^{R2} = mE[D_Z^{R2}] \approx m \frac{\delta r_T^3}{9} \left(\frac{2\pi}{m} \right)^2 \left(\frac{\mu_R^2 + \sigma_R^2}{\mu_R^2} \Phi \left(\frac{\mu_R}{\sigma_R} \right) + \frac{\sigma_R}{\mu_R} \phi \left(\frac{\mu_R}{\sigma_R} \right) \right) \left(1 + \frac{3\sigma_R^2}{m\mu_R^2} \right) \quad (38)$$

where $E[D_Z^{R2}]$ is the expected region 2 lateral distance per vehicle.

Proof. In our spatially-homogeneous problem, region 2 customers will be uniformly-scattered in pie-shaped zones partitioning the disk inside r_T . Under (3), the total expected circumferential (lateral) travel distance in a pie of angular width ω_θ can be estimated using a continuous approximation:

$$\begin{aligned} E[D_Z^{R2} | \omega_\theta] &\approx \int_0^{r_T} (\delta \omega_\theta x) \left(\frac{1}{3} \omega_\theta x \right) dx \\ &= \frac{\delta r_T^3 \omega_\theta^2}{9} . \end{aligned} \quad (39)$$

Removing the conditioning, the expected region 2 lateral distance per vehicle is:

$$E[D_Z^{R2}] = \frac{\delta r_T^3}{9} E[\omega_\theta^2] . \quad (40)$$

The pie width ω_{θ_k} served by vehicle k is a random variable whose moments can be approximated using Y_R and \bar{Y}_R . Since each vehicle is assigned a pie with a width proportional to its remaining capacity, ω_{θ_k} is:

$$\omega_{\theta_k} \approx \max \left(\frac{2\pi}{Y_R(2\pi)} (Q - V_k^{R1}), 0 \right) . \quad (41)$$

Since V_k^{R1} are independent and identically-distributed, the expectation of these pie angles is non-varying:

$$\begin{aligned} E[\omega_\theta^2] &\approx E \left[\left(\frac{2\pi}{Y_R(2\pi)} \max(Q - V_k^{R1}, 0) \right)^2 \right] \\ &= 4\pi^2 E[\max(Q - V_k^{R1}, 0)^2] E \left[\frac{1}{Y_R(2\pi)^2} \right] . \end{aligned} \quad (42)$$

The first expectation in (42) follows from the normal distribution of $Q - V_k^{R1}$,

$$E[\max(Q - V_k^{R1}, 0)^2] = (\mu_R^2 + \sigma_R^2) \Phi \left(\frac{\mu_R}{\sigma_R} \right) + \mu_R \sigma_R \phi \left(\frac{\mu_R}{\sigma_R} \right) . \quad (43)$$

To approximate $E\left[\frac{1}{Y_R(2\pi)^2}\right]$, again let $S = Y_R(2\pi)$. Using the first three terms of the Taylor series approximation for $\frac{1}{S^2}$ yields:

$$E\left[\frac{1}{S^2}\right] \approx \frac{1}{E[S]^2} + \frac{3\text{var}(S)}{E[S]^4} \quad ,$$

thus:

$$E\left[\frac{1}{Y_R(2\pi)^2}\right] \approx \frac{1}{m^2\mu_R^2} \left(1 + \frac{3\sigma_R^2}{m\mu_R^2}\right) \quad . \quad (44)$$

Substituting (44) and (43) into (42), then (42) into (40) yields the result. \square

6.6 Phase 3 distance

Any customers that still remain unserved after phase 2 are grouped into tours based from the depot. Since the expected distance D_T^{OF2} required by such tours will be small, a rough approximation is developed.

Proposition 7 (Phase 3 Overflow Distance)

$$D_T^{OF2} \approx E[m^{OF2}] \left(2r_U + \frac{4\pi^2\delta r_U^3}{9E[m^{OF2}]^2}\right) \quad (45)$$

where r_U is the radius of the disk containing overflow customers approximated by Expression (47), and $E[m^{OF2}]$ is the expected number of phase 3 tours required approximated by Expression (48).

Proof. Assume that each vehicle tour for an overflow serves an equal-sized pie-shaped zone inside of the disk with radius r_U . Suppose each vehicle travels from the depot radially to r_U , and then serves customers in order of decreasing distance to the depot. Then, (45) follows directly from (40)

To determine r_U and $E[m^{OF2}]$, suppose that each region 2 zone is of equal size and that each of the m vehicles serves such a zone. Then the expected overflow demand $E[V_O^{R2}]$ for a region 2 zone is given by the expected overage of a normal random variable:

$$E[V_O^{R2}] = \sqrt{\frac{\gamma\lambda\pi r_T^2}{m}} \Psi\left(\frac{mQ - \lambda\pi r_T^2}{r_T\sqrt{m\gamma\lambda\pi}}\right) \quad (46)$$

where $\Psi(x) = \phi(x) - x\Phi(-x)$.

Radius r_U can now be determined roughly by assuming that overflows are distributed uniformly across angles:

$$r_U \approx \sqrt{\frac{mE[V_O^{R2}]}{\pi\lambda}} \quad , \quad (47)$$

and an approximation of the expected number of vehicle tours required is simply:

$$E[m^{OF2}] \approx \frac{mE[V_O^{R2}]}{Q} \quad . \quad (48)$$

\square

6.7 Total expected distance and cost

The individual components developed in the previous subsections are now combined to approximate the total expected distance traveled by a fleet of m vehicles:

$$D_T^{TG} = D_T^{R1} + D_T^{RE} + D_T^{OF1} + D_T^{R2} + D_T^{OF2} \quad (49)$$

Note that D_T^{TG} is a function of only two decision variables, m and r_T . It is therefore simple to determine an optimal value for r_T numerically for any given m .

In most applications, the expected total system cost per period given by (2) is dominated by the per vehicle component, $c_v m$. Thus, most systems will be designed to use the smallest possible vehicle fleet. A reasonable method to determine fleet size m^{TG} under the TGS scheme is to ensure that the total fleet capacity provides α standard deviations of buffer over the expected demand of all customers,

$$m^{TG} = \left\lceil \frac{\lambda\pi R^2 + \alpha\sqrt{\gamma\lambda\pi R^2}}{Q} \right\rceil. \quad (50)$$

7 Model validation via simulation

Expression (49) is based on a number of simplifications. To verify its accuracy for test problems, we developed a simulation in the MATLAB programming environment to model operations under the TGS scheme. The simulation includes detailed coded decision algorithms for all components of the TGS scheme for the idealized service region with a centrally-located depot. These algorithmic components could be easily adapted for design and control of real-world systems with arbitrary service regions and depot locations.

Given δ , μ , and σ^2 specifying a problem, λ and γ are determined from (4) and (5). Next, given R , Q , and α , the vehicle fleet size m^{TG} is then determined from (50). For each of n_S operating period replications, the system simulates customer locations using a spatially-homogeneous Poisson process with rate δ , and generates a random load size $q_i = \max\{0, V\}$ for each customer, where V is a normal random variable with mean μ and variance σ^2 . Phase 1, 2, and 3 vehicle tours are then determined using decision algorithms similar to those described in the previous sections; these methods are presented in detail in Erera (2000).

<i>parameter</i>		<i>value</i>	<i>units</i>
Region radius	R	10	mi
Customer density	δ	4.0	1/mi ²
Mean demand	μ	5	units
Variance demand	σ^2	2	units ²
Vehicle capacity	Q	75	units
Alpha	α	3	

Table 1: Parameter Data for Example Problem

7.1 Results for an example problem

Consider an example problem with parameters given in Table 1. This is a large-scale problem with moderate variation. The expected number of customers to be served each period is 1257, and each vehicle can transport on average the loads of 15 customers. Applying (50) yields a fleet size of $m^{TG} = 92$. Figure 8 provides a graph of D_T^{TG} as a function of r_T for m^{TG} fixed at 92. To amplify the variation, the minimum ordinate shown (1200) is chosen to be close to the minimum expected line-haul distance $D_{LH} = \frac{4m^{TG}R}{3}$, a lower bound on the expected miles traveled which assumes that each vehicle only travels to the average radius and back without any lateral deviations. For this problem, the optimal threshold radius value is $r_T^* = 5.6$, and the total expected distance traveled by the fleet given by (49) is $D_T^{TG*} = 1991$ miles.

We now use the simulation to validate the approximation model for this example problem. For each value of r_T in $\{1.0, 1.2, 1.4, \dots, 8.8, 9.0\}$, we generated $n_S = 100$ replications. Figure 8 plots the sample mean total distance from the simulated scenarios, \tilde{D}_T^{TG} , in relation to approximation function D_T^{TG} . Clearly, the shape of the means locus is quite similar to that of D_T^{TG} . Further, the simulated means always fall below the predicted curve; thus, the approximation is conservative.

Simulation results for all values of r_T are given in Table 2, where s_D is the sample standard deviation of the simulated total distance. Note that s_D never exceeds 105 miles; at the predicted optimal radius 5.6, the simulated mean was 1955 miles and $s_D = 40$ miles. Thus, the results were grouped tightly about the mean values, showing that for this example the implemented decision algorithms were quite robust. The relative errors between D_T^{TG} and \tilde{D}_T^{TG} further show that the approximation is quite accurate, and that accuracy improves close to r_T^* , the radius value that matters most. At $r_T^* = 5.6$, the relative error is 1.9%. The error increases to 3.1% at $r_T = 9$ and to 7.2% at $r_T = 1.0$.

The optimal radius determined via simulation is 4.8 which does not coincide with the result of the approximation model. However, the mean simulated distance at this radius is 1941 miles, which is only 0.7% smaller than the distance at r_T^* . The results in the table further indicate that the expected distance will be close to the optimum for a range of values of r_T close to r_T^* . This property should be useful in practice.

Level-of-service performance can also be measured from the simulation. One of the primary benefits of the phase 2 capacity reallocation is that most customers will be served by the vehicle initially assigned, and not require an overflow tour. At the optimal threshold radius r_T^* , an average of 16.35 customers required service by either a region 1 or region 2 overflow tour; this value is 1.3% of the expected number of customers. Thus, the strategy delivers a very high service level in this example.

r_T	\widetilde{D}_T^{TG}	s_D	r_T	\widetilde{D}_T^{TG}	s_D	r_T	\widetilde{D}_T^{TG}	s_D
1.0	2195	105	3.8	1995	60	6.6	1977	45
1.2	2167	83	4.0	1985	58	6.8	1979	51
1.4	2155	101	4.2	1976	56	7.0	1984	49
1.6	2146	98	4.4	1969	57	7.2	1999	47
1.8	2111	92	4.6	1952	48	7.4	1992	41
2.0	2100	90	4.8	1941	44	7.6	2003	45
2.2	2087	100	5.0	1951	47	7.8	2010	38
2.4	2065	90	5.2	1954	43	8.0	2021	38
2.6	2055	87	5.4	1953	48	8.2	2020	40
2.8	2046	74	5.6	1955	39	8.4	2025	43
3.0	2029	73	5.8	1959	52	8.6	2027	42
3.2	2021	52	6.0	1971	44	8.8	2031	45
3.4	2015	64	6.2	1964	42	9.0	2033	46
3.6	1999	64	6.4	1964	36			

Table 2: Example load-constrained problem: simulation results

8 Comparisons

To quantify expected savings for a range of problems, we now compare near-optimally configured systems under both the detour-to-depot and TGS schemes.

8.1 Methodology

Table 3 summarizes the parameter settings for 20 sample scenarios. The vehicle capacity $Q = 75$ units and idealized service region radius $R = 10$ miles remain constant over all scenarios. The last column in the table is the square root of the variance to mean ratio for the given parameters.

For this comparative analysis, we configure both the detour-to-depot and TGS schemes so that they provide equal customer service levels, measured by the total expected customer demand served by an overflow tour or following a detour to the depot. Under the detour-to-depot scheme, this quantity is the sum of the expected demand overflow for each routing zone. Under TGS, it is difficult to determine this expectation exactly and therefore we estimate it from simulation results. Let n_{OC} be the sample mean number of region 1 and region 2 customers served by an overflow tour. Equating expected overflow demand allows us to determine the area A_Z^{DD} of the detour-to-depot overflow zones using the following relation:

$$\mu n_{OC} = \pi R^2 \sqrt{\frac{\gamma\lambda}{A_Z^{DD}}} \Psi \left(\frac{Q - \lambda A_Z^{DD}}{\sqrt{\gamma\lambda A_Z^{DD}}} \right). \quad (51)$$

where $\Psi(x) = \phi(x) - x\Phi(-x)$. To perform the comparative analysis then, we employ the following steps for each scenario:

Scenario	δ	n	μ	σ^2	$\sqrt{\frac{\tau}{\lambda}}$
1	4.0	1257	5.0	2.0	0.52
2	2.0	628	5.0	2.0	0.73
3	6.0	1885	5.0	2.0	0.42
4	4.0	1257	2.5	2.0	0.57
5	2.0	628	2.5	2.0	0.81
6	6.0	1885	2.5	2.0	0.47
7	4.0	1257	7.5	2.0	0.51
8	2.0	628	7.5	2.0	0.72
9	6.0	1885	7.5	2.0	0.42
10	4.0	1257	5.0	25.0	0.71
11	2.0	628	5.0	25.0	1.00
12	6.0	1885	5.0	25.0	0.58
13	4.0	1257	5.0	12.0	0.61
14	2.0	628	5.0	12.0	0.86
15	6.0	1885	5.0	12.0	0.50
16	0.75	236	5.0	2.0	1.20
17	0.75	236	5.0	12.0	1.40
18	0.75	236	5.0	25.0	1.63
19	0.25	79	10.0	12.0	2.12
20	0.25	79	10.0	25.0	2.24

Table 3: Data for Comparative Analysis Scenarios

1. Determine m^{TG} using (50) with $\alpha = 3$.
2. Determine optimal threshold r_T^* by minimizing (49) with $m = m^{TG}$; D_T^{TG*} is the resulting minimum expected distance.
3. Simulate TGS scheme with $m = m^{TG}$, $r_T = r_T^*$, and $n_S = 100$, yielding the sample mean expected distance \tilde{D}_T^{TG*} , and n_{OC} .
4. Use (51) to determine A_Z^{DD} .
5. Using $A_Z = A_Z^{DD}$, determine m^{DD} from (14) and \underline{D}_T^{DD*} from (20).

8.2 Results

Table 4 presents comparative results for the sample scenarios. The results in the table indicate that the TGS scheme provides large savings over detour-to-depot. In each sample problem, the TGS scheme requires fewer vehicles than detour-to-depot. Furthermore, in most problems, TGS additionally requires significantly less expected travel distance. Note that for some problems with fewer customers and fewer vehicles, TGS sometimes requires more expected travel distance.

Comparing the columns for D_T^{TG*} and \tilde{D}_T^{TG*} shows that the approximation model is again accurate near the optimum for problems with both large and small numbers of customers. Although it is usually conservative, in scenarios 10–12 and scenario 15 the approximation model underestimates required travel distance, primarily due to the approximation of required region 1 and region 2 overflow distances.

Scen	r_T^*	m^{TG}	D_T^{TG*}	\tilde{D}_T^{TG*}	m^{DD}	\underline{D}_T^{DD*}	\underline{m}	\underline{D}_{LH}	D_T^{VRP}	\tilde{p}_O
1	5.6	92	1991	1955	120	2234	84	1117	1630	1.31 %
2	5.5	48	1138	1115	63	1249	42	559	921	0.84 %
3	5.6	135	2802	2789	175	3181	126	1675	2304	1.60 %
4	5.0	46	1259	1202	56	1303	42	559	1071	0.85 %
5	4.1	24	762	720	28	749	21	279	642	0.99 %
6	5.2	68	1710	1666	83	1810	63	838	1466	0.98 %
7	5.8	137	2770	2742	180	3192	126	1676	2189	2.52 %
8	5.9	71	1533	1491	96	1745	63	838	1200	1.69 %
9	5.8	202	3964	3960	267	4638	188	2513	3142	2.68 %
10	6.2	94	2086	2089	110	2248	84	1117	1630	5.59 %
11	6.2	49	1198	1199	59	1254	42	559	921	3.97 %
12	6.1	138	2932	2960	161	3221	126	1675	2304	6.16 %
13	5.9	93	2037	2025	117	2255	84	1117	1630	2.75 %
14	6.0	48	1163	1148	61	1252	42	559	921	2.01 %
15	5.9	137	2870	2882	168	3208	126	1675	2304	3.43 %
16	4.9	19	545	508	23	544	16	209	432	1.11 %
17	4.8	20	564	527	24	557	16	209	432	1.68 %
18	4.6	21	586	552	23	561	16	209	432	3.06 %
19	6.6	15	388	358	19	395	11	140	268	1.25 %
20	7.2	15	392	365	19	395	11	140	268	1.66 %

Table 4: Comparative Results for Twenty Scenarios

The column labeled \tilde{p}_O presents the mean percentage of customers requiring an overflow tour measured from the TGS simulation; in scenarios 10–12, this value ranges from 4-6% while in most other scenarios this percentage is less than 2%. Furthermore, since the values of \tilde{p}_O are generally quite small, the assumption that overflow demands become known in the TGS approximation model does not appear to affect results significantly.

Any operating scheme for a vehicle routing system with uncertainty will require a larger fleet size and more expected travel distance than the optimal routing solution for an equivalent system

with no uncertainty. Table 4 includes the following deterministic benchmark values:

$$\begin{aligned} \underline{m} &= \frac{\mu\delta\pi R^2}{Q} \\ \underline{D}_{LH} &= \frac{4R\underline{m}}{3} \\ D_T^{\text{VRP}} &= \underline{D}_{LH} + \delta^{1/2} \sqrt{\frac{2}{3}} \pi R^2 \end{aligned}$$

For a routing system with exactly $\delta\pi R^2$ customers each with lot size demand μ , quantity \underline{m} is the minimum number of required vehicles of capacity Q . If these customers were uniformly scattered over the service region, quantity \underline{D}_{LH} is the minimum expected radial distance of travel required by all tours. Finally, D_T^{VRP} is an approximation of the total expected distance required by such tours and is an approximate lower bound on the expected distance required by any scheme for the problem with uncertainty (Daganzo, 1984a).

Table 5 presents conservative estimates of the fleet size and travel distance benefits of the TGS scheme relative to the detour-to-depot scheme for the sample scenarios. Using \underline{m} as a deterministic baseline, column two in this table estimates the percentage reduction in the fleet size penalty due to uncertainty when using TGS instead of detour-to-depot. Similarly, column 3 uses D_T^{VRP} as a baseline to estimate the percentage reduction in the expected travel distance penalty.

The fleet size penalty reductions range from 29% to 82% for the sample scenarios, indicating that large risk-pooling benefits are clearly attainable at equivalent service levels. Furthermore, significant distance penalty savings arise in most problems. TGS provides the most significant savings for scenario 9, where the fleet penalty was reduced by 82% and the travel distance penalty by 45%. This scenario features a very large number of customers and vehicles, and small individual customer demand variances. In such scenarios, savings are relatively large since customer locations are known when region 2 customers are assigned to tours; since variance is also small, demand is nearly deterministic for phase 2. However, scenarios 10-12 show that large savings are also possible when individual customer demand variances are relatively high. Finally, note that TGS provides significant savings even in problems 19 and 20, which feature the fewest vehicles and customers.

9 Conclusions

The threshold global-sharing scheme for load-constrained vehicle routing problems with uncertain customer locations and demands provides significant fleet size and expected travel cost savings over the detour-to-depot scheme. Benefits are obtained by pooling the capacity of the entire vehicle fleet in a way that leads to cost-efficient vehicle operations that are controlled via a small number of real-time decisions.

Although this paper develops results for an idealized problem setting, the TGS scheme should be relatively easy to implement for problems with irregularly-shaped service areas, non-uniform travel networks, and non-homogeneous customer location and demand processes. Configuration of

Scen	$\frac{m^{DD} - m^{TG}}{m^{DD} - m}$	$\frac{D_T^{DD*} - D_T^{TG*}}{D_T^{DD*} - D_T^{VRP}}$
1	77.78%	40.23%
2	71.43%	33.84%
3	81.63%	43.22%
4	71.43%	18.97%
5	57.14%	-12.15%
6	75.00%	29.07%
7	79.63%	42.07%
8	75.76%	38.90%
9	82.28%	45.05%
10	61.54%	26.21%
11	58.82%	16.82%
12	65.71%	31.52%
13	72.73%	34.88%
14	68.42%	26.89%
15	73.81%	37.39%
16	57.14%	-0.89%
17	50.00%	-5.60%
18	28.57%	-19.38%
19	50.00%	5.51%
20	50.00%	2.36%

Table 5: Fleet Size and Expected Distance Penalty Reductions for Sample Scenarios

a TGS scheme specifies both a geographic boundary partitioning potential customer sites into either a pre-assigned region further from the depot or a dynamically-assigned region closer to the depot, and also decision algorithms for vehicle tour generation. For real-world problems, it is reasonable to determine a near-optimal boundary location either via the approximation model presented herein or via simulation.

Acknowledgements

This research was supported partially by the Dwight D. Eisenhower Graduate Transportation Fellowship program, and the University of California Transportation Center.

Appendix: Normality of Conditional Gaussian Process

Suppose $Y_R(\theta)$ is a stochastic process with independent increments, where $Y_R(\theta)$ is normal for all θ . Recall that the mean and variance of $Y_R(\theta)$ are:

$$\begin{aligned} E[Y_R(\theta)] &= \frac{\theta m}{2\pi} \mu_R \\ \text{var}(Y_R(\theta)) &= \frac{\theta m}{2\pi} \sigma_R^2 \end{aligned}$$

Clearly, the joint distribution $\{Y_R(\theta), Y_R(2\pi - \theta)\}$ is multivariate normal with the following covariance matrix:

$$\begin{bmatrix} \frac{\theta m}{2\pi} \sigma_R^2 & 0 \\ 0 & \frac{(2\pi - \theta)m}{2\pi} \sigma_R^2 \end{bmatrix}$$

Since any linear transformation of a multivariate normal random vector remains multivariate normal, $\{Y_R(\theta), Y_R(2\pi)\}$ is multivariate normal with mean vector $\{\frac{\theta m}{2\pi} \mu_R, m \mu_R\}$ and covariance:

$$\begin{bmatrix} \frac{\theta m}{2\pi} \sigma_R^2 & \frac{\theta m}{2\pi} \sigma_R^2 \\ \frac{\theta m}{2\pi} \sigma_R^2 & m \sigma_R^2 \end{bmatrix}$$

The correlation coefficient ρ is $\sqrt{\frac{\theta}{2\pi}}$. It follows now from a property of the multivariate normal distribution that the conditional distribution $Y_R(\theta)|Y_R(2\pi)$ is normal with the following mean and variance:

$$\begin{aligned} E[Y_R(\theta)|Y_R(2\pi)] &= \frac{\theta m}{2\pi} \mu_R + \frac{\theta}{2\pi} (Y_R(2\pi) - m \mu_R) = \frac{Y_R(2\pi) \theta}{2\pi} \\ \text{var}(Y_R(\theta)|Y_R(2\pi)) &= \frac{\theta m}{2\pi} \sigma_R^2 (1 - \rho^2) = m \sigma_R^2 \frac{\theta}{2\pi} \left(1 - \frac{\theta}{2\pi}\right) \end{aligned}$$

References

- D.J. Bertsimas. *Probabilistic combinatorial optimization problems*. PhD thesis, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 1988.
- D.J. Bertsimas. A vehicle routing problem with stochastic demand. *Operations Research*, 40: 574–585, 1992.
- D.J. Bertsimas and L. Howell. Further results on the probabilistic traveling salesman problem. *European Journal of Operational Research*, 65:68–95, 1993.
- C.F. Daganzo. 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, 1984a.
- C.F. Daganzo. The length of tours in zones of different shapes. *Transportation Research, Part B*, 18B:135–146, 1984b.

- M. Dror and P. Trudeau. Stochastic vehicle routing with a modified savings algorithm. *European Journal of Operational Research*, 23:228–235, 1986.
- A.L. Erera. *Design of Large-scale Logistics Systems for Uncertain Environments*. PhD thesis, Department of Industrial Engineering and Operations Research, University of California-Berkeley, Berkeley, CA, 2000.
- M. Gendreau, G. Laporte, and R. Séguin. An exact algorithm for the vehicle routing problem with stochastic demands and customers. *Transportation Science*, 29:143–155, 1995.
- M. Gendreau, G. Laporte, and R. Séguin. A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. *Operations Research*, 44:469–477, 1996.
- P. Jaillet. A priori solution of a traveling salesman problem in which a random subset of the customers are visited. *Operations Research*, 36:929–936, 1988.
- P. Jaillet. Analysis of probabilistic combinatorial optimization problems in euclidean spaces. *Mathematics of Operations Research*, 18:51–70, 1993.
- G. Laporte and F.V. Louveaux. The integer l -shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13:133–142, 1993.
- G. Laporte, F.V. Louveaux, and H. Mercure. Models and exact solutions for a class of stochastic location-routing problems. *European Journal of Operational Research*, 39:71–78, 1989.
- G. Laporte and Y. Nobert. A cutting planes algorithm for the m -salesman problem. *Journal of the Operational Research Society*, 31:1017–1023, 1980.
- W. Stewart and B. Golden. Stochastic vehicle routing: a comprehensive approach. *European Journal of Operational Research*, 14:371–385, 1983.
- B. Verweij, S. Ahmed, A. Kleywegt, G. Nemhauser, and A. Shapiro. The sample average approximation method applied to stochastic routing problems: A computational study. *Computational Optimization and Applications*, 24:289–333, 2003.

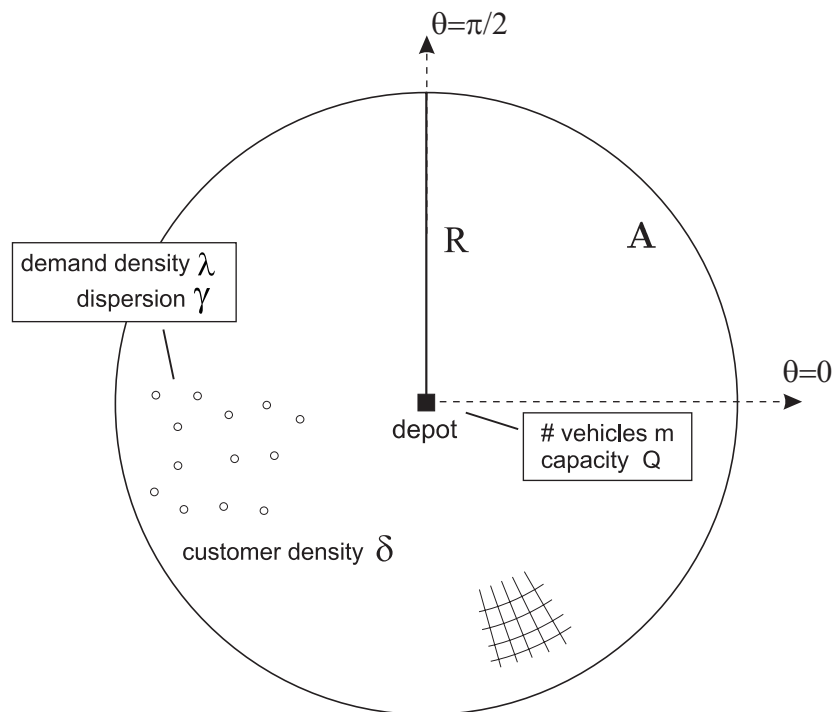


Figure 1: Idealized central depot problem

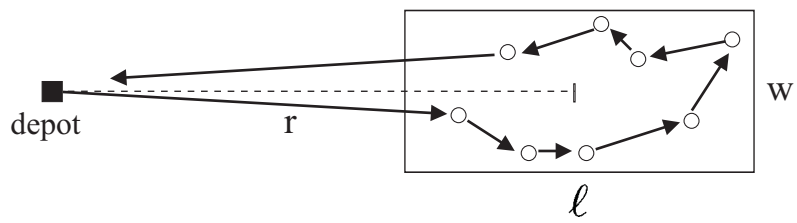


Figure 2: Rectangular routing zone

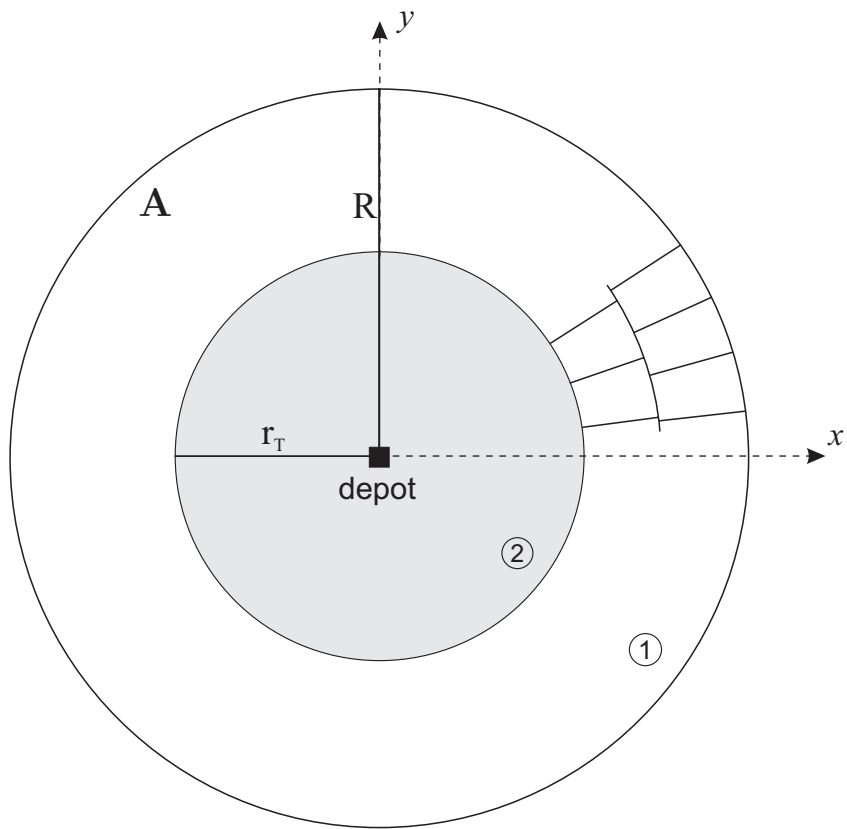


Figure 3: Threshold global-sharing strategy

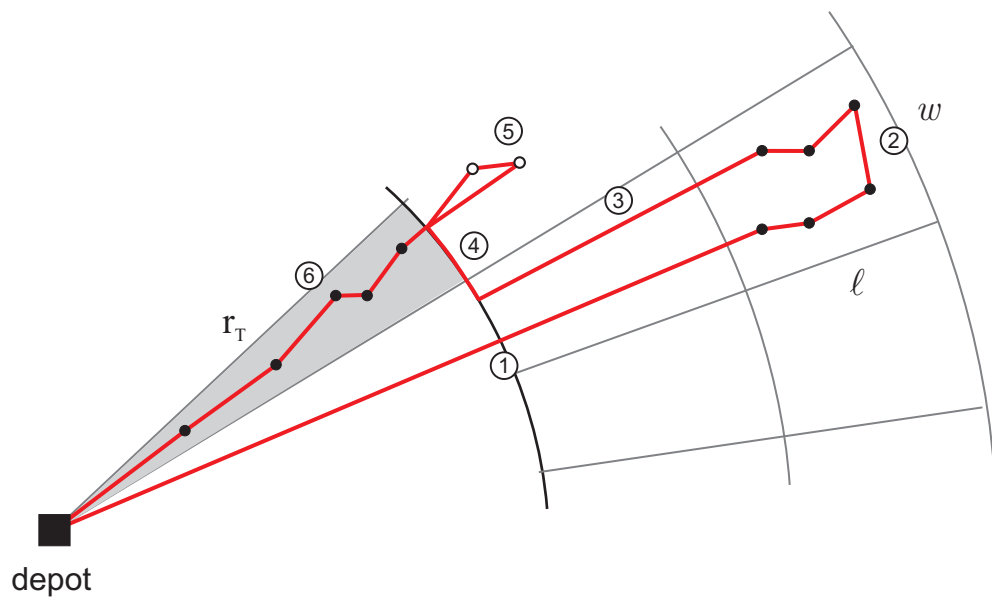


Figure 4: Sample vehicle tour for threshold global-sharing strategy

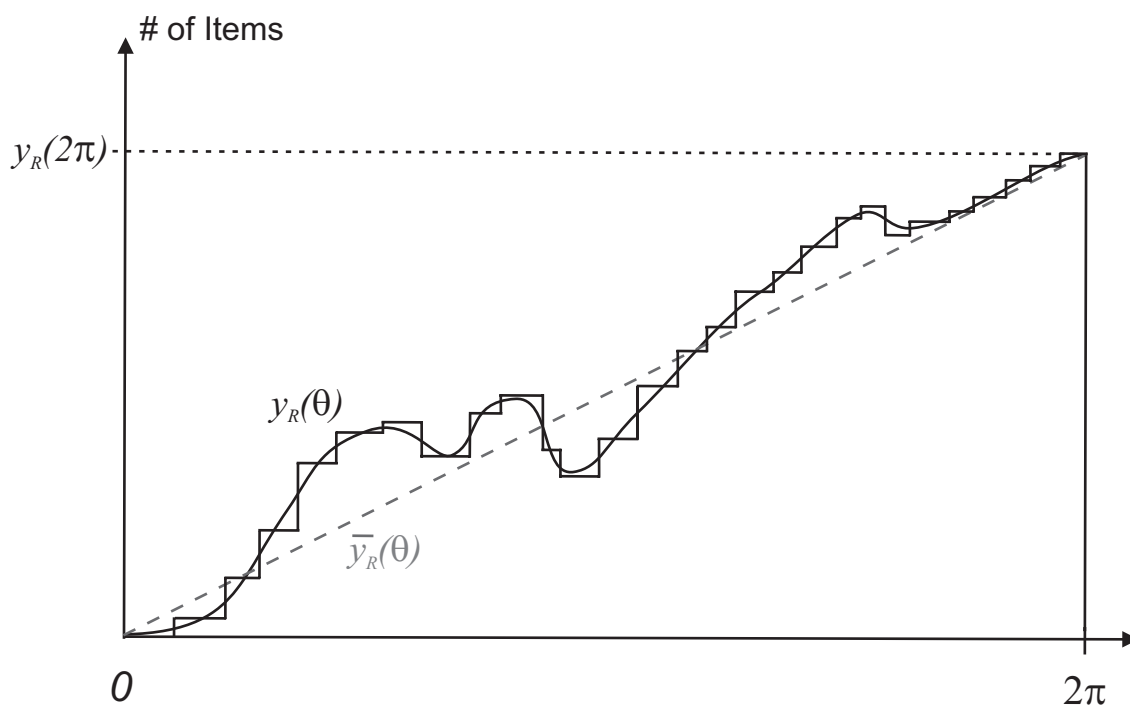


Figure 5: Cumulative remaining capacity curves

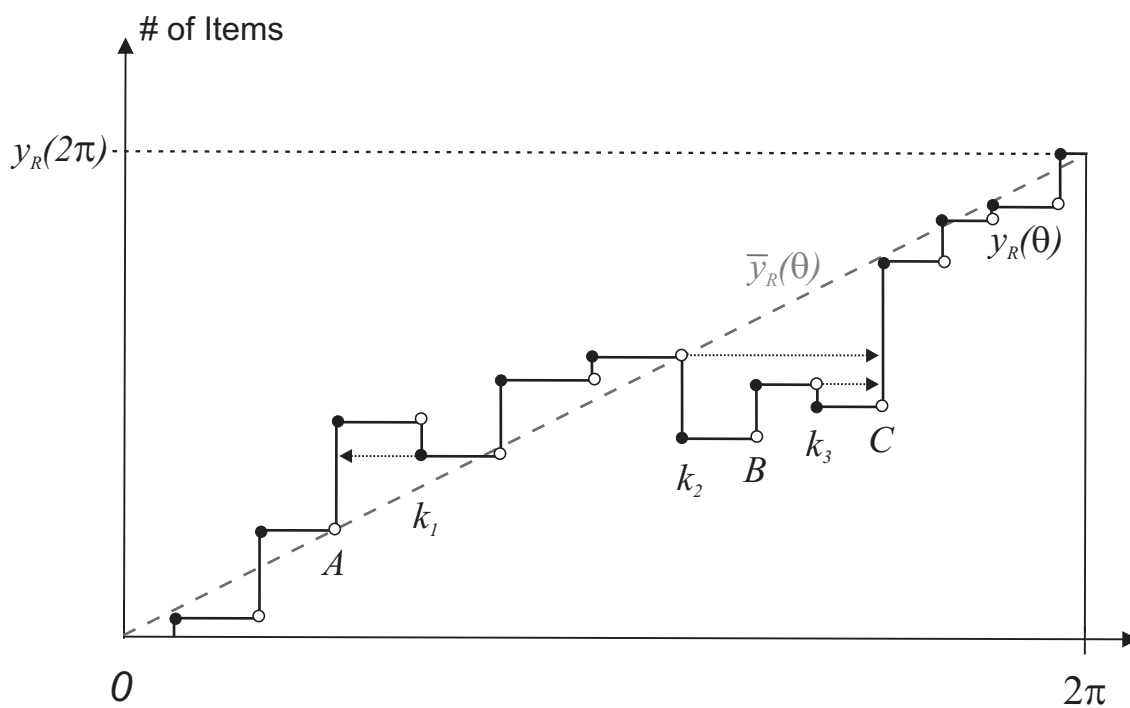


Figure 6: Region 1 overflow customer reassignment algorithm

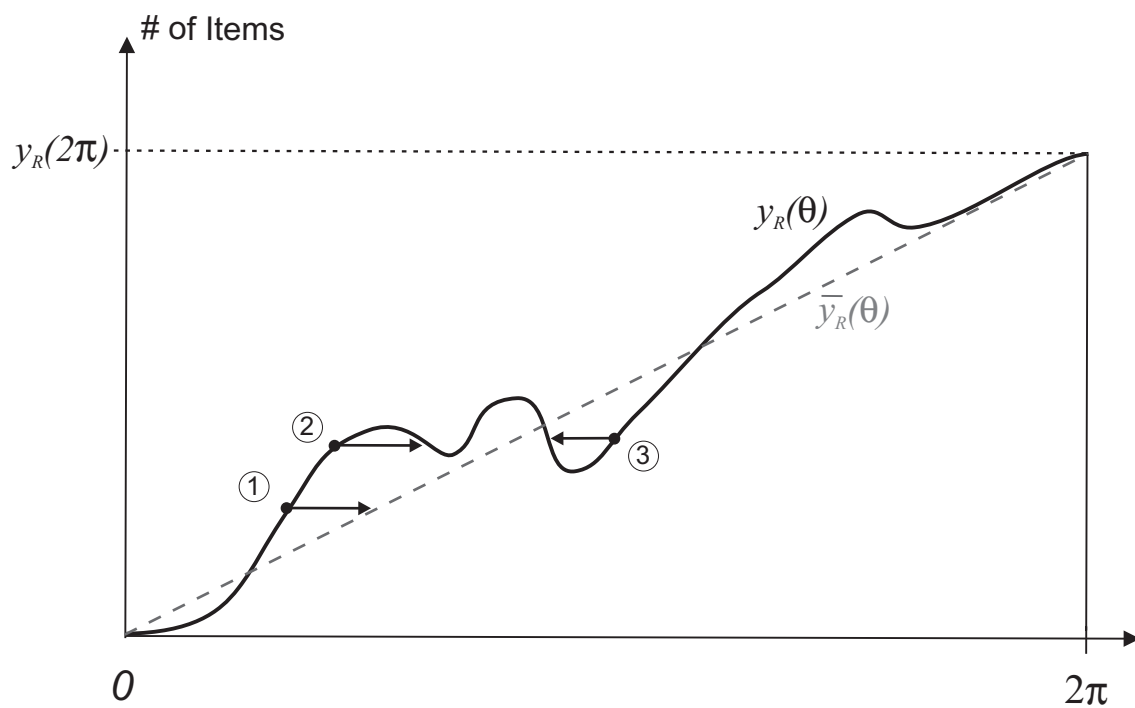


Figure 7: Calculation of repositioning distance

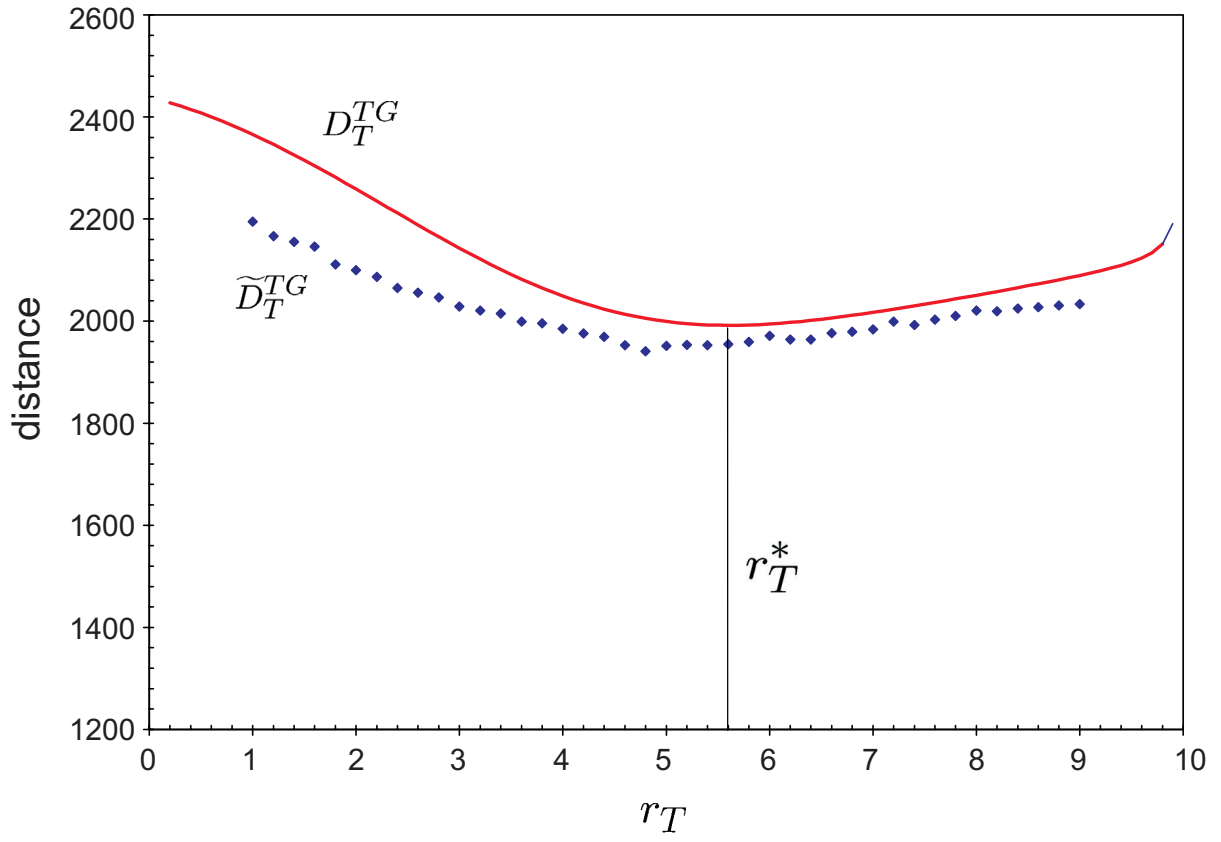


Figure 8: Example load-constrained problem: Predicted and simulated distance vs. threshold radius