
Challenges and Advances in A Priori Routing

Ann Melissa Campbell and Barrett W. Thomas

Tippie College of of Business
Department of Management Sciences
University of Iowa
Iowa City, Iowa 52242-1994
ann-campbell@uiowa.edu, barrett-thomas@uiowa.edu

Summary. An a priori route is a route which specifies an ordering of all possible customers that a particular driver may need to visit. The driver may then skip those customers on the route who do not receive a delivery. Despite the prevalence of a priori routing, construction of these routes still presents considerable challenges. Exact methods are limited to small problem sizes, and even heuristic methods are intractable in the face of real-world-sized instances. In this chapter, we will review some of the ideas that have emerged in recent years to help solve these larger instances. We focus on the probabilistic traveling salesman problem and the recently introduced probabilistic traveling salesman problem with deadlines and discuss how objective-function approximations can reduce computation time without significantly impacting solution quality. We will also present several open research questions in a priori routing.

Key words: Stochastic routing; a priori routing.

1 Introduction

For many delivery companies, only a subset of their customers require a pickup or delivery each day. Information may be not available far enough in advance to create optimal schedules each day for those customers that do require a visit or the cost to acquire sufficient computational power to find such solutions may be prohibitive. Companies have long used a priori routes to help overcome these difficulties. An a priori, or pre-planned, route is a route which specifies an ordering of all possible customers that a particular driver may need to visit. The driver may then skip those customers on the route who do not receive a delivery. A priori routes are used routinely in the package express industry to sequence the many potential stops in each driver's assigned territory. These routes create a regularity of service that can be beneficial for both the customers and the drivers. Customers will be served at roughly the same time each day they require service, and the drivers can become very familiar

with their routes. If there is time available on the day of service, starting from an a priori tour can be useful, too, as a starting point for reoptimization.

The best known a priori routing problem is the probabilistic traveling salesman problem (PTSP). Formally, the PTSP is the problem of finding a minimum expected cost tour through a set of customers $N = \{i \mid 1, \dots, n\}$ with probabilities $P = \{p_i \mid 1, \dots, n\}$ of requiring service on any given day. The travel time between any two customers i and j is given by d_{ij} , where $d_{ij} = d_{ji}$.

Despite the prevalence of a priori routing, construction of these routes still presents considerable challenges. Exact methods are limited to small problem sizes, and even heuristic methods are intractable in the face of real-world-sized instances. In this chapter, we will review some of the ideas that have emerged in recent years to help solve these larger instances. These ideas are focused on approximating the computationally expensive PTSP objective function using various techniques.

Even though the PTSP is a challenging problem by itself, many delivery companies are faced with a more complicated routing problem due to increasing just-in-time business practices. These just-in-time considerations usually come in the form of delivery deadlines. The most common example of these services is next-day package delivery featured by United Parcel Service (UPS) and FedEx. Next-day delivery providers usually offer a choice of deadlines such as 10 am, noon, or 3 pm. Such time-definite services have grown from just 4% of the parcel delivery market in 1977 to over 60% in 2002 [59]. The market for all time-definite cargo was expected to grow by 7.6% in 2006 alone [25], and the growth is expected to continue.

In the case of a priori routing with deadlines, the problem is known as the probabilistic traveling salesman problem with deadlines (PTSPD). The problem definition for the PTSPD is the same as the PTSP except that associated with each customer $i \in N$ is a known deadline l_i . The PTSPD can alternately be considered a version of the PTSP with time windows, but the opening time of all windows is set to zero. Unfortunately, due to how the deadlines impact the problem, the same techniques used to help solve large instances of the PTSP are often not applicable for the PTSPD. In this chapter, we will also review new approximation techniques that have proven successful for the PTSPD and discuss instance characteristics that influence the performance of these approaches.

This chapter is structured as follows. Section 2 provides a basic literature review of a priori routing, where Section 3 and Section 4 delve into solution approaches for the PTSP and PTSPD, respectively. Section 5 presents several open research questions in a priori routing.

2 Literature Review

One of the first appearances of a priori routing in the literature is in [2], where the authors develop a priori tours for use in meals-on-wheels routing. Jaillet [41] formally introduces the PTSP and demonstrates some interesting properties of optimal tours including the fact that such a tour may intersect itself. Jaillet [42] provides a formulation for the expected value of a tour and bounds the relationship between optimal PTSP and TSP solutions. Berman and Simchi-Levi [7] focus on instances of the PTSP with heterogeneous probabilities, where most of Jaillet’s results involve homogeneous probabilities. They establish a lower bound for such instances and explain how to combine this bound with a branch-and-bound algorithm to find an optimal a priori tour. Bowler et al. [17] offers additional characterizations of the problem.

Related to the PTSP is the Noisy Euclidean Traveling Salesman Problem (NTSP). This problem was introduced in [20] and further studied in [45]. The NTSP is based on the idea that some neighborhoods within the service area will have a higher probability of requiring service than others, and these densities can be exploited to create an a priori route or “trajectory”. In this way, each node in the resulting TSP can be thought of as being sampled from a probability distribution, where in the PTSP the customer locations are known with certainty.

While the literature contains research into many constrained versions of the TSP, there is limited research into constrained versions of the PTSP. The best known of the constrained versions is the stochastic vehicle routing problem (SVRP). The SVRP requires the consideration of vehicle capacity in the formation of the tours, and rather than customer presence, customer demand is usually the stochastic element of the problem. The first mention of this problem can be found in [58]. Bertsimas [8] introduces an analytical framework and bounds for the SVRP. Other work can be divided into chance-constrained and recourse model formulations (see [26, 27] and [16] for an overview of the two types of formulations). Stewart and Golden [55], Laporte et al. [43], and Bastian and Rinnooy Kan [3] provide chance-constrained formulations and show how the problems can be transformed into deterministic problems. Dror et al. [32], Dror [31], and Gendreau et al. [35] present stochastic programming solutions to various recourse models for the SVRP. Dror and Trudeau [33], Bramel et al. [18], Bertsimas et al. [9], Savelsbergh and Goetschalckx [54], Gendreau et al. [37], and Yang et al. [62] offer various heuristics for the SVRP. Gendreau et al. [37] provides comparisons of CPU times between their exact solution approaches in [35] and their heuristics in [37]. For different scenarios involving 11 stochastic customers, they demonstrate that exact solution approaches may require solution times a thousandfold greater than the solution times of the proposed TABUSTOCH heuristic. Larger instances where all customers are stochastic are not evaluated.

Campbell and Thomas [23] introduce the PTSPD, providing two recourse models and a chance-constrained model for the problem. In addition, compu-

tational experiments demonstrate situations in which it is important to model the problem stochastically versus situations in which deterministic models are sufficient.

The authors know of only a few other papers that address routing under uncertainty with time constraints. These papers consider time constraints in the context of stochastic travel times rather than in the context of random customers as is discussed in this chapter. Teng et al. [57] apply the L-shaped algorithm to the time-constrained traveling salesman problem (TCTSP) with stochastic travel and service times. In the TCTSP, the time constraint is on the length of the tour, which contrasts with the PTSPD where the time constraints control when individual customers can be visited. Wong et al. [61] introduce a 2-stage stochastic integer program with recourse for a problem in which travel times are stochastic, all customers must be visited, and each customer has an associated time window.

The value of a priori versus daily optimized or variable routes has been explored in the context of the SVRP. Haughton [39, 40] introduce metrics for determining the value of reoptimization versus a priori routes. Waters [60], Benton and Rosetti [5], and Savelsbergh and Goetschalckx [54] discuss circumstances in which a priori routes can be cost-competitive alternatives to reoptimization. For example, Benton and Rosetti [5] experiment with different customer realization probabilities and route rescheduling costs. Their experiments reveal that reoptimization is preferred when route scheduling costs are low and few customers are realized. Interestingly, they also comment that there are “a great deal of hidden costs associated with variable routes” and suggest this is why a priori routes are often preferred. These hidden costs can include consistency considerations and management overhead as well as customer relationships associated with a driver visiting a customer at relatively the same time everyday. A detailed discussion of consistency in route design can be found in [63] and [64] which introduce a driver learning model to capture the importance of a driver consistently working in a particular geographic area.

3 Solution Approaches for the PTSP

In this section, we formulate the PTSP, provide background on solution methods for the PTSP, and highlight two techniques for approximating the objective function.

3.1 Problem Formulation

In a solution to the PTSP, all of the customers are sequenced on one tour. On the day of service, when all demands are known, the customers that have been realized can be visited in the sequence defined by this a priori tour. Solution methods for the PTSP focus on minimizing the expected cost of these final

tours. The expected cost associated with a particular sequence of customers $1 \dots n$ can be evaluated by Equation 1 [8]:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i p_j d_{i,j} \prod_{k=i+1}^{j-1} (1-p_k) + \sum_{i=2}^n \sum_{j=1}^{i-1} p_i p_j d_{i,j} \prod_{k=i+1}^n (1-p_k) \prod_{l=1}^{j-1} (1-p_l). \quad (1)$$

The first part of the equation represents the expected cost associated with using each arc (i, j) in a *forward* direction while the second part is the expected cost associated with using each arc in the *reverse* direction to complete the tour. The expected cost of an arc is based on the probability that the customers at both end points of the arc are realized, the probability that the none of the customers in between these on the tour are realized, and the length of the arc.

3.2 Background

An exact approach for the PTSP was introduced by Laporte et al. [44], but computational tests indicate success only with instances of 50 customers or less. Consequently, much of the PTSP literature focuses on heuristic approaches. The authors found that among the instances studied, it was much harder to solve instances with low individual probability values.

Rossi and Gavioli [53] discuss how to modify construction heuristics for the TSP specifically to solve the PTSP. Their heuristics are based on Clarke and Wright and nearest neighbor techniques and do not include any local improvement. The expected costs of the resulting solutions are compared with those found using basic TSP heuristics. Based on their computational experiments, the authors conclude that it is important to use solution techniques specifically developed for the PTSP if the number of customers is greater than 50 and the probability of each customer requiring a visit is less than 60%.

Bertsimas et al. [11] discuss space-filling curve and iterative heuristics. Bertsimas and Howell [10] and Chervi [28] explore the use of TSP heuristics for solving the PTSP and propose algorithms for the PTSP based on constructing an initial solution using the space-filling curve heuristic [1] followed by local search. Variations of the 2-OPT and 1-Shift techniques developed for the TSP in [46] are introduced that compute the change in objective in an expected value sense. The equations presented by Bertsimas and Howell [10] and Chervi [28] have been shown to have small errors, corrected in [15] and [12], but even with these small errors, the authors are able to show improvement based on expected value becomes more important as n becomes large. These key instances, though, prove to be very difficult for them to solve. For example, [9] reports runtimes of half an hour to a full hour of CPU time for their simplest improvement heuristics to converge with 50-customer PTSPs. Bertsimas and Howell [10] also report that expected value based local improvement is particularly important when probability values are significantly less than 1. This confirms the results established by Rossi and Gavioli [53].

Beraldi et al. [6] extend the efficient evaluation methods of Bertsimas and Howell [10] and Chervi [28] to the probabilistic pickup and delivery traveling salesman problem.

Various metaheuristic approaches have also been applied to the PTSP. Bianchi et al. [13, 14] introduce an ant-colony optimization approach. They demonstrate that the ant-colony approach outperforms a radial-sort heuristic and a random best heuristic for the PTSP. Building upon work presented in [51], Rosenow [52] develops a genetic-algorithm approach which uses crossover operators first proposed by Grefenstette et al. [38] for the traveling salesman problem. The tests are limited and show only that the genetic algorithm outperforms a branch-and-bound scheme for the PTSP. Bowler et al. [17] uses the PTSP as a testbed to demonstrate the effectiveness of stochastic annealing. As a result of their tests on the PTSP, the authors hypothesize that a priori tours perform at most 14% worse than reoptimization strategies for the PTSP.

3.3 Approximation Approaches

The above research on solution techniques for the PTSP concludes that it is particularly important to solve instances of the PTSP using an expected value approach when the size of the instances are large and when probability values are not close to 1. This represents most instances of the PTSP that would be solved in practice, such as those at package delivery companies. Even with sophisticated metaheuristic approaches and significant increases in computing power, such large instances are still considered intractable because they involve repeated evaluations of Equation 1. This objective function is much more time consuming to evaluate than with the TSP. The PTSP objective function includes the cost of every arc that could potentially be used in the solution, leading to its $O(n^3)$ complexity rather than the TSP's $O(n)$ complexity. Thus, evaluating the cost of a small change in a PTSP route, such as in a heuristic search approach, is also potentially computationally expensive. In the TSP, the change in cost can be computed based on the neighbors directly impacted by the change. Due to the probabilistic nature of the objective function, computing a small change in an a priori route can easily be as expensive as evaluating the cost of the full objective function. Some of the more promising ideas that have emerged in recent years for solving the PTSP involve approximating the objective function and using this approximation in various search methods. These ideas differ from the existing solution approaches that rely on exact computation of expected solution values, even when they are obtained using heuristics. We will consider two types of approximation here: one where the number of terms in the original objective function is truncated and one where customers are aggregated.

Truncation and Approximation Function

The first of these approximation ideas for the PTSP was presented in detail in [56]. The objective function considered is slightly different than Equation 1 because the authors assume the tour begins at the depot (customer 0) which is present with a probability value of 1. The authors also do not consider the cost of a return trip to the depot which removes the second part of Equation 1.

The approximation is based on truncating the calculations used in computing the cost of adding a customer to the tour. For example, suppose the cost to add a customer j to the end of an arbitrary partial tour is initially computed by:

$$p_{j-1}p_jd_{j-1,j} + p_{j-2}p_jd_{j-2,j}(1-p_{j-1}) + \cdots + p_0p_jd_{0,j}(1-p_1)\cdots(1-p_i). \quad (2)$$

Equation 2 could be replaced by:

$$p_ip_jd_{i,j} + p_{i-1}p_jd_{i-1,j}(1-p_i) + \cdots + p_{i-K+1}p_jd_{i-K+1,j}(1-p_{i-K+2})\cdots(1-p_i) \quad (3)$$

where $K \in (0, i)$ is a truncation parameter. Thus, K terms at most are used in the equation that evaluates the cost of adding j . The idea is that the cost of adding j is most impacted by the customers closest to j on the tour. Customers more than K stops ahead of j , for example, will be unlikely to directly precede j on the resulting tour, so the cost of using that direct arc is highly discounted in the cost computation and can reasonably be truncated.

The idea of truncating expected cost expressions was introduced prior to [56] in [36]. In [36], the authors develop approximations for the vehicle routing problem with stochastic customers and stochastic demand. The part of their approximation related to the existence of stochastic customers is quite similar to the simple truncation proposed above.

Due to the truncation, Equation 3 will always underestimate the true costs. Tang and Miller-Hooks [56] refine this approximation approach by proposing the use of a function $f(K)$ to estimate this underestimation, changing and generalizing Equation 3 to:

$$\sum_{i=j-K}^{n-1} p_id_{i,j} \prod_{k=i+1}^{j-1} (1-p_k) + f(K)p_j. \quad (4)$$

Such an expression, after defining $f(K)$, can clearly be used to speed up a construction heuristic for the PTSP. Equation 4 can also be used in approximating the full function evaluation, enabling truncation to be used with various improvement approaches.

The amount of speedup and the quality of such an approach is dependent on the choice of K and the approximation function $f(K)$. Equation 4 will be more accurate the higher K is, but it will also be more expensive to evaluate. Tang and Miller-Hooks [56] propose choosing K based on a set

of trial experiments. They propose a series of computationally efficient functions to use for $f(K)$ involving coefficients that are tuned to the particular problem instances. For many choices of K and $f(K)$, the evaluation of the objective becomes $O(n)$, which is the same complexity as the objective of the deterministic TSP.

To overcome the sensitivity to the choice of K , Tang and Miller-Hooks [56] propose a progressive enhancement idea where K can be increased over the course of the improvement procedure. As the solution improves and less improvements are possible with the current choice of K , K is increased and $f(K)$ is improved to make the approximation more accurate. This gives the power of the speedup without a loss in quality of the final solution. It also makes the solution method less sensitive to a poor initial choice of K .

Tang and Miller-Hooks [56] experiment with datasets involving 50, 75, and 100 customers. The computational experiments reveal that it is possible to tackle large sized PTSP instances with significantly reduced computational effort but not reduced solution quality. Using a 2-OPT improvement scheme, for example, instances with 100 customers that originally took over 2100 seconds on average to converge, take on average 17.5 seconds with approximation. For these instances, the solutions found with approximation were on average 0.5% better than those found using exact objective evaluations.

We note that both Branke and Guntch [19] and Liu [47] experiment with truncation-based approximate function evaluations within their ant-colony and scatter-search metaheuristics, respectively. Branke and Guntch [19] consider instances as large as 1379 customers and demonstrate that an approximate function evaluation embedded in ant-colony optimization reduces computation time by almost 30% relative to a full evaluation. Interestingly, they found better speedup when using the approximation with heuristic search methods other than ant colony optimization. On instances with up to 100 customers, results in [47] indicate that approximate function evaluation embedded in scatter search is capable of finding good solutions while reducing computation time by as much as 86%. Unfortunately, direct comparisons between [56], [19], and [47] do not exist.

Approximation by Aggregation

The second idea we examine in approximating the PTSP is to aggregate customers into regions. By using regions in place of individual customers in Equation 1, there is an obvious opportunity for speedup in evaluating the objective function. This idea is examined in detail in [21]. Aggregation in this context refers to grouping customers together and then representing them by one point spatially and with a single probability value. In aggregating customers, there are many choices in terms of how to divide customers into regions and how many regions to create.

Customer aggregation is common in the literature, but primarily in the context of location problems (see [29, 34], for examples) and not in a routing

context. There has been little analysis, theoretical or computational, on how aggregation can be used in solving routing problems or how aggregation can impact solution quality, especially in a probabilistic context. In [4], for example, the authors aggregate deliveries by postal codes before designing routes, but there is no discussion of how this aggregation impacts solution quality and speed.

If aggregation is used in the context of the PTSP, the a priori tour becomes a tour through regions rather than customers. This requires a modification of Equation 1. If the n customers are each assigned to one of r regions, a new probability value is required for each region as well as new distance costs. The probability associated with each region S will need to reflect all of the customers assigned to this region. Since a tour will travel to each region only once, Campbell [21] proposes defining p_S as the probability that region S will require a visit. Computation of p_S is then:

$$p_S = 1 - \prod_{i \in S} (1 - p_i). \quad (5)$$

Equation 5 is the probability that region S will have at least one realized demand given that customer orders are independent events. Next, to compute Euclidean distances between regions, a spatial location is needed to represent each region. Campbell [21] proposes computing the centroid of the customers in a region, where the weights on the customers are based on their individual probabilities. These centroid coordinates can be used to compute the Euclidean distance between each pair of regions. Equation 1 can then be replaced by the aggregated a priori expected value equation found in Equation 6:

$$\sum_{S=1}^{r-1} \sum_{T=S+1}^r p_S p_T d_{S,T} \prod_{U=S+1}^{T-1} (1 - p_U) + \sum_{S=2}^r \sum_{T=1}^{S-1} p_S p_T d_{S,T} \prod_{U=S+1}^r (1 - p_U) \prod_{V=1}^{T-1} (1 - p_V) \quad (6)$$

where $d_{S,T}$ represents the distance between regions S and T . Equation 6 resembles Equation 1, but the number of terms here can be several orders of magnitude smaller than Equation 1 because of aggregation.

Note, unlike with truncation, this approximation does not necessarily underestimate the full objective function. Because in most cases it will, Campbell [21] proposes a series of functions that can be added to Equation 6 similar to the $f(K)$ functions in [56].

As stated earlier, the primary literature on aggregation comes from location theory, and there are a variety of ways suggested to group customers together based on distance. Yet with the PTSP, probability offers an important additional consideration. Grouping customers strictly based on location may lead to regions, for example, with very different probabilities of requiring a visit. Likewise grouping customers so that each region is equally likely to require a visit may create regions of very different size and shape. The study in [21] uses simple distance and probability-based aggregation approaches to

build insight into what creates a successful aggregation scheme for PTSP problems.

The distance-based aggregation scheme is a grid-based approach, such as those discussed in many location papers [50]. For a given parameter g , the customer service area will be divided evenly into g segments along the x axis and g segments along the y axis to create a total of g^2 regions of equal area. If customers are uniformly distributed, each region should have roughly the same number of customers. The potential downfall of such an approach lies in the fact that not all customers are evenly distributed over the service area in real world applications.

The second form of aggregation is to divide the customer service area into regions of roughly equal probability. This may help remedy some of the possible negative issues with grids since customers should be fairly evenly distributed among the regions, but now the regions will clearly not be of the same size. To divide potential customers into regions based on probability, Campbell [21] defines a parameter *maxp* that represents the maximum likelihood of requiring a visit in a region. Starting from an initial region that includes all of the customers, regions are repeatedly divided into smaller regions until all regions have probability less than or equal to *maxp*. The approach is similar to Voronoi Diagram/Delaunay Triangulation ideas [30]. The shape of the current region is used to guide the division process in order to preserve some of the advantages of the grid approach and keep the regions from being extremely tall or wide which would distort distance calculations.

Computational experiments involve datasets ranging from 100-1000 customers and various customer probabilities. Following the structure of [10], initial solutions were constructed using a space-filling curve heuristic and several improvement schemes were tested, including TSP techniques such as 2-OPT and 1-Shift and expected value versions of 2-OPT and 1-Shift. Using aggregation, search procedures for even the 1000 customer datasets were able to converge within the 120 second time limit, and the final solutions were usually better than those found without aggregation. Both uniformly and clustered datasets were used to see how the geographical distribution of customers impacts the performance of the aggregation approaches. The results indicate that, as expected, using grid versus probability based aggregation makes little impact on uniform datasets, but can have significant impact when datasets are clustered. Campbell [21] found that quite coarse levels of aggregation can lead to good objective value estimates, but aggregation needs to become finer as the customer probabilities increase. The proposed general rule is to divide customers into regions such that the total expected demand in a region is no more than 0.5 deliveries in order to achieve an estimate within 90% of the full objective value. The experiments also indicate that there are many research opportunities concerning the solution of the PTSP when data is not uniformly distributed. Most of the previous publications, and thus the conclusions about solving the PTSP, were based on uniformly distributed datasets,

but Campbell [21] found the relative performance of various heuristics to be very different for clustered datasets.

4 Solution Approaches for the PTSPD

In this section, we formulate the PTSPD, discuss why different approximation techniques are needed than those for the PTSP, and introduce two approximation techniques appropriate for the PTSPD.

4.1 Problem Formulation

In a solution to the PTSPD, all of the customers are sequenced on one tour, just like the PTSP. Because of the deadlines, the PTSPD is modeled with a depot (customer 0), and the tour departs the depot at time 0. As in [56], the depot exists with a probability value of 1, but, unlike [56], the model includes a return to the depot in the cost evaluation. There are several choices on how the deadlines will impact the costs in the model. We will follow the same structure as [22], introduced in [23], where the vehicle is allowed to visit each customer i after the delivery deadline has passed, but incurs a per-unit-time penalty, λ_i , for doing so. If T_i is the latest time that customer i can be reached and $g(i, t)$ is the probability that customer i is reached at time t , then the expected cost of a tour is:

$$\begin{aligned} & \sum_{j=1}^n p_j d_{0j} \prod_{k=1}^{j-1} (1 - p_k) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i p_j d_{ij} \prod_{k=i+1}^{j-1} (1 - p_k) + \sum_{i=1}^n p_i d_{i0} \prod_{k=i+1}^n (1 - p_k) \\ & + \sum_{i=1}^n p_i \sum_{t=l_i+1}^{T_i} \lambda_i g(i, t) (t - l_i). \end{aligned}$$

The probability $g(i, t)$ can be computed recursively using the following equations. When $t < d_{0i}$, $g(i, t)$ values will always be zero, since arrival cannot occur any earlier than with a direct trip from the depot. When $t = d_{0i}$, arrival at t can occur if no prior customers are realized:

$$g(i, t) = \prod_{k=1}^{i-1} (1 - p_k). \quad (7)$$

When $t > d_{0i}$, arrival at i is based on all of the possible preceding customers and their possible departure times:

$$g(i, t) = \sum_{h=1, t > (d_{hi} + d_{0h})}^{i-1} p_h g(h, t - d_{hi}) \prod_{k=h+1}^{i-1} (1 - p_k). \quad (8)$$

The complexity of the function evaluation is dominated by the computation of the g values and is $O(n^2 \max_i \{T_i\})$.

4.2 Approximation Approaches

As indicated in Section 2, the PTSPD was only recently introduced in [23]. While Campbell and Thomas [23] demonstrate the expected savings possible from modeling a time-constrained a priori routing problem probabilistically, they also show that, like with the PTSP, the probabilistic model requires significant additional computation time to solve. Given the recent developments for the PTSP, the obvious solution is to use one of the approximation approaches from Section 3. Unfortunately, simply implementing the truncation and aggregation ideas proposed for the PTSP with the PTSPD is unlikely to be successful for two reasons. First, the computational experiments in [23] reveal, in line with the worst case complexity analysis, that the run time for solving the PTSPD is dominated by the g values. Thus, truncating only the distance portion of the objective function will make little difference in the run-times. We will discuss in Section 4.2 how Campbell and Thomas [22] propose truncating the calculation of the g terms. Second, the aggregation schemes discussed for the PTSP no longer make sense when the customers have different deadlines. Since the expense of the g terms is due to the potentially high T_i values, Campbell and Thomas [22] propose discretizing time into larger units to reduce the computational burden in a way that behaves similar to customer aggregation. This discretization idea is referred to as temporal aggregation in Section 4.2. It is important to note that these approximations methods are not specific to the PTSPD, but, rather, they have application in any discrete-time problem in which the timing of events is stochastic.

Truncation Approximation

In the PTSPD, even a small change in the tour involving position i will impact the expected arrival times, and thus expected penalties, at all customers succeeding i on the route. Like with the PTSP, the change in penalty will be “felt” the strongest by the customers that are served just after i on the tour. Thus, Campbell and Thomas [22] proposes evaluating the change in the penalty portion of the objective associated with a local search move by considering only the q nearest neighbors to each customer.

The truncation approximation in [22] involves new equations for computing the g values. We will refer to the new g values by g^* . A value for $g^*(i, t)$ will be computed when $t = d_{0i}$ only if $i \leq q$. In other words, for $t = d_{0i}$ and $i > q$, we will set $g^*(i, t) = 0$. If $i \leq q$, Equation 7 is used. Next we modify Equation 8 to only consider the closest q customers to i :

$$g^*(i, t) = \sum_{h=\max(i-q, 1), t > (d_{hi} + d_{0h})}^{i-1} p_h g^*(h, t - d_{hi}) \prod_{k=h+1}^{i-1} (1 - p_k). \quad (9)$$

With this approximation, no $g^*(i, t)$ calculation requires more than $O(q)$ complexity, assuming that the products are stored and computed ahead of time.

Now the full penalty cost can be computed in $O(nq \max_i \{T_i\})$ time rather than $O(n^2 \max_i \{T_i\})$. Depending on the relative size of q and n , Equation 9 can lead to a significant speedup in the calculations.

The choice of q is important. Since this approximation scheme is based on the idea that the truncated terms will have a value close to zero due to all of the multiplied $(1 - p_k)$ probabilities involved, the initial choice of q should be based on the probabilities involved in a particular instance or dataset. For example, when the average probability values are near 0.1, the initial q can be set higher than when probabilities values average near 0.9. As with the progressive enhancement idea from Section 3.3, the choice of q can be incremented based on the progress of the local search procedure, overcoming the potential difficulty of parameter setting.

Computational experiments include instances with 40, 60, and 100 customers, varied probabilities, varied penalties, and varied deadlines. The approximation is embedded within a steepest descent local search algorithm with a 1-Shift neighborhood. The results indicate that, in most cases, truncation returns quality solutions more quickly than a similar search heuristic in which the cost of neighboring solutions is computed without approximation. For example, on 60-customer instances in which all customers have a low probability of being realized, truncation reduces runtime by as much as 72% compared to the direct computation. Tests show that truncation has the most value when customer probabilities are low and deadlines are tight. The truncation approach performs poorly in terms of runtime in the case where all customers have a high probability of requiring service.

Temporal Aggregation

If penalties are assessed based on the number of minutes that a delivery is late, it is necessary for accuracy to compute the $g(i, t)$ values with t indices representing minutes. If there are a large number of customers or if the travel times between some customers are quite long, the T_i values can easily become exponential in n .

In a local search scheme, it is typical to choose the change to the current solution that makes the largest improvement in the objective value. In this context, we can think of changes that reduce the lateness at customers by hours rather than minutes as the type of improvements we would want to look for first. This is the idea behind the temporal aggregation scheme proposed in [22]. Instead of making the t values represent minutes, or whatever the final time discretization that is required for the penalty calculation, larger time discretizations are used and are gradually refined until the final time discretization is reached.

Temporal aggregation has been applied in economic models [48] and in integer programming [49]. Most applications of temporal aggregation in integer programming are based on increasing the size of time periods for which decisions are made in an attempt to reduce the number of decision variables.

One key decision here is size of the time units that will be used in evaluating the penalty function. The selected time discretization should be large enough to gain computational advantage in the penalty calculation, but small enough such that penalty improvements can be found. If the original time discretization is in minutes, the larger discretization does not need to be hours but could be, for example, 3 minutes, 40 minutes, or 180 minutes. The choice of time discretization will vary and will clearly have a relationship to the customer dataset being considered.

For the temporal aggregation scheme proposed in [22], the units for the T_i , l_i , and d values will all need to be changed to reflect the new level of discretization. These new values will be referred to by T_i^* , l_i^* , and d^* , respectively. Since the l_i^* values are used as indices in the g functions, they must obtain integer values. Thus, a simple transformation using a particular time discretization v is to round each $\frac{l_i}{v}$ to its nearest integer to obtain l_i^* . Similarly, the same process is repeated to create the new distance values d^* , and these new distances are used to compute the T_i^* values. Note that Equation 7 does not change, just the value of d_{0i} to d_{0i}^* . Equation 8 becomes:

$$g^*(i, t) = \sum_{h=1, t > (d_{hi}^* + d_{0h}^*)}^{i-1} p_h g(h, t - d_{hi}^*) \prod_{k=h+1}^{i-1} (1 - p_k).$$

Campbell and Thomas [22] offer a general suggestion for the choice of the largest time discretization. Since the complexity of evaluating the objective function is potentially non-polynomial in n due to the T values, the recommendation is to use time units of size v where $v = \frac{\max_i \{T_i\}}{n}$. The largest T_i^* value can be is $\frac{\max_i \{T_i\}}{v}$ which is now n , making a function evaluation possible in $O(n^3)$ time.

The final step in designing a temporal aggregation scheme is to decide if and when the level of discretization should be changed. This choice can be based on the progress of the local search, as in a progressive enhancement procedure. For example, the user can begin with a choice of v that makes the function polynomial to evaluate and reduce v when the local search converges.

Using the same experiments as with the truncation approximation, computational results for the temporal approximation reveal that temporal approximation provides the same quality solutions as truncation. Like the truncation approach, temporal aggregation performs poorly in terms of runtime on instances in which the probability of all customers requiring service is high. Temporal aggregation offers better runtimes relative to truncation on 40-customer datasets. However, temporal aggregation requires much longer runtimes on the 100-customer datasets.

5 Open Questions

There are many remaining questions. It is not clear what are the best search methods in which these approximation ideas should be embedded and how the approximation parameters should be tuned in the solution process for these problems. It would also be interesting to determine if there would be benefits from using some of these approximation ideas in conjunction with each other. It is also not clear if the answers to these questions will be different depending on the geographical distribution of the customers and/or the distribution of the customer deadlines. Further, it would be interesting to see if the approximation techniques can be used in conjunction with vehicle capacity, as in the stochastic vehicle routing problem [36] and the stochastic vehicle routing problem with deadlines [24].

Although this chapter focuses on approximation techniques for a priori routing problems, it is also important to note that there are many variations of the PTSP that are not well-studied, but are applicable in real-world applications. These variants offer important areas of future study, especially if the approximation approaches presented here aid empirical study. One variant of interest is the case of travel distances which are not Euclidean distances but are road network distances. This variant is motivated by the fact that most real-world applications of the PTSP involve traveling on road networks. A second area of future research involves problems in which the probabilities that customers require service are not necessarily independent. These problems have been ignored due to the technical difficulties, but in many applications, these probabilities are often correlated. Third, the obvious extension to the PTSPD is to consider the situation where customers have delivery time windows and not just delivery deadlines. Known as the PTSP with time windows, this variant requires consideration of the possibility that vehicles must wait before the opening of window. This possibility greatly complicates the problem formulation and also raises questions about whether it should be possible to make deliveries early, but incur a penalty for doing so. Fourth, related to the PTSPD is the problem where only a subset of customers have a deadline, and this deadline is the same for those customers. This variant reflects the difficulties that some delivery providers have in offering an “express” service and might be more amenable to more specialized solution techniques.

Acknowledgments

The authors want to thank the referees for their helpful suggestions. This work was partially supported by the National Science Foundation through grant number 0237726(Campbell).

References

- [1] J. Bartholdi and L. Platzman, “An $O(N \log N)$ Planar Traveling Salesman Heuristic Based on Spacefilling Curves,” *Operations Research Letters* 1, 121–125 (1982).
- [2] J. J. Bartholdi, L. K. Platzman, R. L. Collins, and W. H. Warden, “A Minimal Technology Routing System for Meals on Wheels,” *Interfaces* 13, 1–8 (1983).
- [3] C. Bastian and A. H. G. Rinnooy Kan, “The Stochastic Vehicle Routing Problem Revisited,” *European Journal of Operational Research* 56, 407–412 (1992).
- [4] J. Beasley and N. Christofides, “Vehicle routing with a sparse feasibility graph,” *European Journal of Operational Research* 98, 499–511 (1997).
- [5] W. C. Benton and M. D. Rosetti, “The Vehicle Scheduling Problem with Intermittent Customer Demands,” *Computers and Operations Research* 19, 521–531 (1992).
- [6] P. Beraldi, G. Ghiani, G. Laporte, and R. Musmanno, “Efficient Neighborhood Search for the Probabilistic Pickup and Delivery Travelling Salesman Problem,” *Networks* 45, 195–198 (2005).
- [7] O. Berman and D. Simchi-Levi, “Finding Optimal A Priori Tour and Location of Traveling Salesman with Nonhomogenous Customers,” *Transportation Science* 22, 148–154 (1988).
- [8] D. J. Bertsimas, *Probabilistic Combinatorial Optimizations Problems*, Ph.D. thesis, Massachusetts Institute of Technology (1988).
- [9] D. J. Bertsimas, P. Chervi, and M. Peterson, “Computational Approaches to Stochastic Vehicle Routing Problems,” *Transportation Science* 29, 342–352 (1995).
- [10] D. J. Bertsimas and L. H. Howell, “Further Results on the Probabilistic Traveling Salesman Problem,” *European Journal of Operational Research* 65, 68–95 (1993).
- [11] D. J. Bertsimas, P. Jaillet, and A. R. Odoni, “A Priori Optimization,” *Operations Research* 38, 1019–1033 (1990).
- [12] L. Bianchi and A. M. Campbell, “Extension of the 2- p -opt and 1-shift algorithms to the Heterogeneous Probabilistic Traveling Salesman Problem,” *European Journal of Operational Research* 176, 131–144 (2007).
- [13] L. Bianchi, L. M. Gambardella, and M. Dorigo, “An Ant Colony Optimization Approach to the Probabilistic Traveling Salesman Problem,” in *Proceedings of the 7th International Conference on Parallel Problem Solving from Nature*, G. Goos, J. Hartmanis, and J. van Leeuwen (eds), volume 2439/2002 of *Lecture Notes in Computer Science*, 883–892, Springer, Berlin, 2002.
- [14] L. Bianchi, L. M. Gambardella, and M. Dorigo, “Solving the Homogeneous Probabilistic Traveling Salesman Problem by the ACO Metaheuristic,” in *Proceedings of ANTS 2002: Third International Workshop*,

- M. Dorigo, G. Di Caro, and M. Sampels (eds), volume 2463/2002 of *Lecture Notes in Computer Science*, 176–187, Springer, Berlin, 2002.
- [15] L. Bianchi, J. Knowles, and N. Bowler, “Local Search for the Probabilistic Traveling Salesman Problem: Correction to the 2-p-opt and 1-shift algorithms,” *European Journal of Operational Research* 162, 206–219 (2005).
 - [16] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer-Verlag, New York (1997).
 - [17] N. E. Bowler, T. M. A. Fink, and R. C. Ball, “Characterization of the probabilistic traveling salesman problem,” *Physical Review E* 68, 036703 (2003).
 - [18] J. Bramel, E. G. Coffman, P. W. Shor, and D. Simchi-Levi, “Probabilistic Analysis of the Capacitated Vehicle Routing Problem with Unsplit Demands,” *Operations Research* 340, 1095–1106 (1992).
 - [19] J. Branke and M. Guntsch, “Solving the Probabilistic TSP with Ant Colony Optimization,” *Journal of Mathematical Modelling and Algorithms* 3, 403–425 (2004).
 - [20] M. Braun and J. Buhmann, “The noisy Euclidean traveling salesman problem and learning,” in *Advances in Neural Information Processing Systems*, T. Dietterich, S. Becker, and Z. Ghahramani (eds), volume 14, 251–258, MIT Press (2002).
 - [21] A. M. Campbell, “Aggregation for the Probabilistic Traveling Salesman Problem,” *Computers & Operations Research* 33, 2703–2724 (2006).
 - [22] A. M. Campbell and B. W. Thomas, “Computational Approaches for the Probabilistic Traveling Salesman Problem with Deadlines,” (2007), submitted to *Computers and Operations Research*.
 - [23] A. M. Campbell and B. W. Thomas, “The Probabilistic Traveling Salesman Problem with Deadlines,” (2007), forthcoming in *Transportation Science*.
 - [24] A. M. Campbell and B. W. Thomas, “The Stochastic Vehicle Routing Problem with Deadlines,” (2007), working Paper.
 - [25] B. Carey, “Expedited Grows on the Surface,” *Traffic World* 1 (2006).
 - [26] A. Charnes and W. W. Cooper, “Chance-Constrained Programming,” *Management Science* 6, 73–79 (1959).
 - [27] A. Charnes and W. W. Cooper, “Deterministic Equivalents for Optimizing and Satisficing Under Chance Constraints,” *Operations Research* 11, 18–39 (1963).
 - [28] P. Chervi, *A Computational Approach to Probabilistic Vehicle Routing Problems*, Master’s thesis, Massachusetts Institute of Technology (1988).
 - [29] M. Daskin, A. Haghani, M. Khanal, and C. Malandraki, “Aggregation effects in maximum covering models,” *Annals of Operations Research* 18, 115–139 (1989).
 - [30] M. de Berg, O. Schwarzkopf, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, Springer-Verlag (2000).

- [31] M. Dror, "Modeling Vehicle Routing with Uncertain Demands as Stochastic Programs: Properties of the Corresponding Solution," *European Journal of Operational Research* 64, 432–441 (1993).
- [32] M. Dror, G. Laporte, and P. Trudeau, "Vehicle Routing with Stochastic Demands: Properties and Solution Frameworks," *Transportation Science* 23, 166–176 (1989).
- [33] M. Dror and P. Trudeau, "Stochastic Vehicle Routing with Modified Savings Algorithm," *European Journal of Operational Research* 23, 228–235 (1986).
- [34] R. L. Francis and T. J. Lowe, "On Worst-Case Aggregation Analysis for Network Location Problems," *Annals of Operations Research* 40, 229–246 (1992).
- [35] 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).
- [36] M. Gendreau, G. Laporte, and R. Séguin, "Stochastic Vehicle Routing," *European Journal of Operational Research* 88, 3–12 (1996).
- [37] 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).
- [38] J. Grefenstette, R. Gopal, B. Rosmaita, , and D. Van Gucht, "Genetic algorithms for the traveling salesman problem," in *Proceedings of the First International Conference on Genetic Algorithms*, J. Grefenstette (ed), Lawrence Erlbaum Associates, Hillsdale, New York, 1985.
- [39] M. A. Haughton, "Quantifying the Benefits of Route Reoptimisation Under Stochastic Customer Demand," *Journal of the Operational Research Society* 51, 320–332 (2000).
- [40] M. A. Haughton, "Route Reoptimization's Impact on Delivery Efficiency," *Transportation Research - Part E* 38, 53–63 (2002).
- [41] P. Jaillet, *Probabilistic Traveling Salesman Problems*, Ph.D. thesis, Massachusetts Institute of Technology (1985).
- [42] P. Jaillet, "A Priori Solution of the Traveling Salesman Problem in which a Random Subset of Customers are Visited," *Operations Research* 36, 929–936 (1988).
- [43] 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).
- [44] G. Laporte, F. V. Louveaux, and H. Mercure, "A Priori Optimization of the Probabilistic Traveling Salesman Problem," *Operations Research* 42, 543–549 (1994).
- [45] F. Li, B. Golden, and E. Wasil, "The Noisy Euclidean Traveling Salesman Problem: A Computational Analysis," in *Perspectives in Operations Research: Papers in Honor of Saul Gass' 80th Birthday*, F. Alt, M. Fu, and B. Golden (eds), 247–270, Springer (2006).

- [46] S. Lin, "Computer Solution of the Traveling Salesman Problem," *Bell System Technical Journal* 44, 2245–2269 (1965).
- [47] Y.-H. Liu, "A Scatter Search Based Approach with Approximate Evaluation for the Heterogeneous Probabilistic Traveling Salesman Problem," in *Proceedings of the 2006 IEEE Congress on Evolutionary Computation*, 1603–1609, 2006.
- [48] J. Mercenier and P. Michel, "Discrete-time Finite Horizon Approximation of Infinite Horizon Optimization Problems with Steady-state Variance," *Econometrica* 62, 635–656 (1994).
- [49] A. M. Newman and M. Kuchta, "Using Aggregation to Optimize Long-term Production Planning at an Underground Mine," *European Journal of Operational Research* 176, 1205–1218 (2007).
- [50] M. Rayco, R. Francis, and A. Tamir, "A p-center grid-positioning aggregation procedure," *Computers and Operations Research* 26, 1113–1124 (1999).
- [51] S. Rosenow, "A heuristic for the probabilistic TSP," in *Operations Research Proceedings 1996*, H. Schwarze (ed), Springer Verlag (1997).
- [52] S. Rosenow, "Comparison of an exact branch-and-bound and an approximative evolutionary algorithm for the probabilistic traveling salesman problem," (1998), working paper, available at <http://www2.hsu-hh.de/uebe/paper-eng1-SOR98.pdf>.
- [53] F. Rossi and I. Gavioli, "Aspects of Heuristic Methods in the Probabilistic Traveling Salesman Problem," in *Advanced School on Stochastics in Combinatorial Optimization*, 214–227, World Scientific (1987).
- [54] M. W. P. Savelsbergh and M. Goetschalckx, "A Comparison of the Efficiency of Fixed Versus Variable Vehicle Routes," *Journal of Business Logistics* 46, 474–490 (1995).
- [55] W. R. Stewart and B. L. Golden, "Stochastic Vehicle Routing: A Comprehensive Approach," *European Journal of Operational Research* 14, 371–385 (1983).
- [56] H. Tang and E. Miller-Hooks, "Approximate Procedures for the Probabilistic Traveling Salesman Problem," *Transportation Research Record* 1882, 27–36 (2004).
- [57] S. Y. Teng, H. L. Ong, and H. C. Huang, "An Integer L-Shaped Algorithm for the Time-Constrained Traveling Salesman Problem with Stochastic Travel Times and Service Times," *Asia-Pacific Journal of Operational Research* 21, 241–257 (2004).
- [58] F. Tillman, "The Multiple Terminal Delivery Problem with Probabilistic Demands," *Transportation Science* 3, 192–204 (1969).
- [59] United Parcel Service, "About UPS," http://www.corporate-ir.net/ireye/ir_site.zhtml?ticker=UPS&script=2100&layout=7 (2002), accessed on November 30, 2006.
- [60] C. D. J. Waters, "Vehicle Scheduling Problems with Uncertainty and Omitted Customers," *Journal of the Operational Research Society* 40, 1099–1108 (1989).

- [61] J. C. F. Wong, J. M. Y. Leung, and C. H. Cheng, “On a Vehicle Routing Problem with Time Windows and Stochastic Travel Times: Models, Algorithms, and Heuristics,” Technical Report SEEM2003-03, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, 2003.
- [62] W.-H. Yang, K. Mather, and R. H. Ballou, “Stochastic Vehicle Routing Problem with Restocking,” *Transportation Science* 34, 99–112 (2000).
- [63] H. Zhong, *Territory Planning and Vehicle Dispatching with Stochastic Customers and Demand*, Ph.D. thesis, University of Southern California (2001).
- [64] H. Zhong, R. W. Hall, and M. Dessouky, “Territory Planning and Driver Learning in Vehicle Dispatching,” *Transportation Science* (to appear).