European Journal of Operational Research 196 (2009) 509–515

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Production, Manufacturing and Logistics

An approximate dynamic programming approach for the vehicle routing problem with stochastic demands

Clara Novoa [a,*], Robert Storer [b]

[a] Texas State University, Ingram School of Engineering, Industrial Engineering Program, San Marcos, TX 78666, USA
[b] Industrial and Systems Engineering Department, Lehigh University, Bethlehem, PA 18015, USA

ARTICLE INFO

Article history:
Received 19 September 2007
Accepted 18 March 2008
Available online 25 March 2008

Keywords:
Transportation
Stochastic vehicle routing
Approximate dynamic programming

ABSTRACT

This paper examines approximate dynamic programming algorithms for the single-vehicle routing problem with stochastic demands from a *dynamic* or reoptimization perspective. The methods extend the rollout algorithm by implementing different base sequences (i.e. *a priori* solutions), look-ahead policies, and pruning schemes. The paper also considers computing the cost-to-go with Monte Carlo simulation in addition to direct approaches. The best new method found is a two-step lookahead rollout started with a stochastic base sequence. The routing cost is about 4.8% less than the one-step rollout algorithm started with a deterministic sequence. Results also show that Monte Carlo cost-to-go estimation reduces computation time 65% in large instances with little or no loss in solution quality. Moreover, the paper compares results to the perfect information case from solving exact a posteriori solutions for sampled vehicle routing problems. The confidence interval for the overall mean difference is (3.56%, 4.11%).

1. Introduction

The classical, deterministic, vehicle routing problem (VRP) seeks minimum cost routes from a depot to a geographically dispersed customers' set having known demands. In this problem, all vehicles start and end their route at the depot, all customer demands are satisfied by exactly one vehicle, and vehicle capacities are not exceeded. The VRP has received extensive attention in the literature. Bertsimas and Simchi-Levi (1996) and Toth and Vigo (2002) review exact approaches, algorithms, and relaxations for the VRP.

Stochastic vehicle routing problems (SVRP's) result when one or more VRP elements are random variables. Random elements might be the customers' set, the travel times, or the customers' demands. Gendreau et al. (1996) summarize the literature on various SVRP's. Dror et al. (1989) indicate that optimal solution properties for the VRP do not hold for the SVRP. Further, Gendreau et al. (1995, 1999), Laporte et al. (2002), Ichoua et al. (2006) show that VRP's combining stochastic, integer, and in some cases dynamic elements (i.e. elements varying over time) call for complex solution methodologies.

This paper studies the Vehicle Routing Problem with Stochastic Demands (VRPSD). In this problem, customers' demands follow known probability distributions and a customer's actual demand is only revealed when the vehicle arrives at the customer location. The VRPSD's goal is to minimize total expected route cost. VRPSD's occur in practice when delivering petroleum products, industrial gases (Chepuri and Homem-De-Mello, 2005), and home heating oil (Dror et al., 1985). Other VRPSD's arise delivering products to cities under emergency (Dessouky et al., 2005), hospitals, restaurants, vending machines (Yang et al., 2000), and bank branches. Random demands are also present when collecting money (Laporte et al., 1989), packages (Marković et al., 2005), sludge, and recycled materials from banks, homes, and industrial plants.

Most VRPSD research assumes an *a priori* solution approach (Bertsimas, 1992; Teodorovic and Pavkovic, 1992; Gendreau et al., 1995; Savelsbergh and Goetschalckx, 1995; Hjorring and Holt, 1999; Laporte et al., 2002; Bianchi et al., 2004; Novoa et al., 2006). In the first stage, complete *a priori* routes are designed before any actual demands become known. In the second stage, routes are followed, demands are revealed, and extra trips to the depot for replenishment are performed if a customer's demand exceeds current vehicle capacity. The routes order is not changed. The objective is to find a route sequence that minimizes total expected cost from original distance traveled and from extra trips to and from the depot. To minimize costs further, Bertsimas et al. (1995) and Yang et al. (2000) design *a priori* routes that may prescribe returns to the depot before vehicle capacity is depleted (i.e. proactive returns).

This paper assumes a *dynamic* solution approach that models the problem in multiple stages. Other authors (Dror et al., 1989; Bertsimas, 1992; Dror, 1993; Secomandi, 2001; Laporte et al., 2002; Secomandi and Margot, in review) call this approach "reoptimization". In this approach, routing decisions occur concurrently

* Corresponding author. Tel.: +1 512 245 4343; fax: +1 512 245 7771.
E-mail addresses: cn17@txstate.edu (C. Novoa), rhs2@lehigh.edu (R. Storer).

0377-2217/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2008.03.023

with service and are based on the most current system state. The system state updates every time the vehicle arrives at a location and observes demand. There is no planned route and decisions at each stage are which customer to visit next and whether or not send the vehicle to the depot for replenishment to minimize expected routing costs.

Powell et al. (1995) mention that stochastic and dynamic models are key for designing decision support systems that respond to changing conditions often observed in practical applications. Further, Psaraftis (1995) indicates that *dynamic* approaches respond to the need for efficient real-time logistic and that they are implementable due to advances in communications technologies such as wireless phones and global positioning systems (GPS) which facilitate interaction between drivers and dispatchers. Bastian and Rinnooy Kan (1992) and Psaraftis (1995) mention that the *dynamic* or reoptimization approach is computationally challenging but results in flexible routes that may reduce total routing costs. Erera et al. (2007) favor fix routes from *a priori* approaches since they may decrease management costs, increase drivers performance and achieve service regularity. Nevertheless, Savelsbergh and Goetschalckx (1995) present instances with a 10% increase in transportation cost from using fixed routes instead of reoptimization.

This paper examines the rollout algorithm as an efficient heuristic method for solving the dynamic single VRPSD in real-time. The rollout algorithm, originally proposed by Bertsekas and Tsitsiklis (1996), Bertsekas et al. (1997), and Bertsekas (2000, 2001), overcomes the curse of dimensionality in dynamic programming (DP). The only previous computational approaches applying rollout to the VRPSD are Secomandi (1998, 2000, 2001), Novoa (2005), and Secomandi and Margot (in review). Secomandi's contribution in regard to rollout methods is the development of a one-step rollout algorithm.

Our paper has three main contributions. The first is the development of a two-step rollout algorithm that provides 1.6% cheaper solutions than the one-step rollout algorithm. The second is the use of Monte Carlo simulation (MCS) for computing the updated base sequences expected cost as an alternative to the exact computation in Secomandi (1998, 2000, 2001). We demonstrate that MCS may reduce the total computational time by about 65% for large instances. The third contribution is the development of improved base sequences and pruning schemes leading to a cost reduction of about 4% over previous methods. The best rollout method evidences the benefits of linking *a priori* and *dynamic* approaches for VRPSD.

Paper organization is as follows. Section 2 reviews literature on *dynamic* approaches. Section 3 describes the problem. Section 4 presents the dynamic programming formulation. Section 5 describes the proposed rollout algorithms and the Monte Carlo simulation approach. Section 6 contains numerical results and Section 7 concludes the paper.

2. Literature review

There are few papers on *dynamic* approaches to the VRPSD relative to those studying the *a priori* approach. The earliest contributions are theoretical models in Dror et al. (1989) and Dror (1993) that model the dynamic VRPSD as a Markov Decision process. However, these papers do not provide any computational results. Secomandi (1998, 2000, 2001, 2003) is the first author that provides a computationally tractable heuristic.

Secomandi (2000) solves the VRPSD using two approximate DP algorithms: optimistic approximate policy iteration (OAPI) and one-step rollout algorithm (ORA). OAPI is a neuro-dynamic programming (NDP) algorithm (Bertsekas and Tsitsiklis, 1996) that approximates the optimal cost-to-go functions for all system states as linear (or nonlinear) functions of pre-selected problem features (i.e. vehicle location, vehicle capacity, etc.) with coefficients estimated by least squares. ORA is a simplified NDP algorithm that resembles the one-step policy iteration method. Given a properly selected *a priori* base sequence, ORA sequentially improves it. The ORA in Secomandi (2000) performs better than OAPI. For small instances, (less than 10 customers), Secomandi (1998, 2001) finds that ORA is less than 3% away from the optimum computed in Secomandi (1998). For problems with up to 60 customers, Secomandi (2001) finds that ORA produces average improvements from 1% to 4% over *a priori* solutions.

More recently, Secomandi and Margot (in review) solve VRPSD's with up to 100 customers using a partial-reoptimization approach. Under this framework, optimal policies are computed on selected subsets of the state set using backward DP. To be able to compute these optimal policies, it is assumed that if a customer demand exceeds current vehicle capacity (i.e. a route fails) and the customer has not been completely served, the vehicle travels to the depot and return to this customer. To find the desired subsets, the authors use two methods for partitioning the initial customer sequence: disjoint and overlapping. The partial-reoptimization approach is less than 1.5% away from the optimum in problems with 10 and 15 customers. Disjoint and overlapping methods produce average improvements of 1.71% and 2.24% over the ORA in Secomandi (2001) for instances with 5–100 customers.

In a general context, efficient approximate DP methods for solving dynamic problems are discussed in Bertsekas (2001), Bertsekas and Tsitsiklis (1996), Sutton and Barto (1998), and Powell (2007). Examples of simulation-based methods for adaptively computing value function (i.e. cost-to-go) approximations are Kleywegt et al. (2002) for solving stochastic inventory routing problems, Topaloglu and Powell (2006) and Topaloglu and Powell (2007) for solving vehicle fleet management problems. Other contributions to neuro DP using least square approximations or a rollout approach are Van-Roy et al. (1998) for solving two-echelon inventory problems, Secomandi (2003) for solving a traveling salesman problem with stochastic travel times, and Bertsekas and Castanon (1999) for solving stochastic scheduling problems. Furthermore, de Farias and Van-Roy (2003) use linear programming (LP) to compute coefficients on the cost-to-go approximating function and Adelman (2004) uses dual-prices to approximate the cost-to-go on instances of the stochastic inventory routing problem.

3. Problem description

A single-vehicle with fixed capacity Q departs from a depot to perform only deliveries (or only pick-ups) at different customer locations. Node 0 represents the depot and $I = 1, 2, \ldots, n$ represents the customers' set. Distances between customers i and j, denoted by $d(i,j)$, are known, symmetric, and satisfy the triangle inequality. Travel costs are proportional to distances.

In the computational study, customers' demands follow known discrete distributions. They are assumed statistically independent but may come from different families and may have different parameters. Let $D_i, i = 1, 2, \ldots, n$ be random variables that describe the demand for each customer and p_i be its probability distribution, $p_i(r) = Pr\{D_i = r\}, r = 0, 1, \ldots, R \leqslant Q$. Exact demand for a customer is revealed only when the vehicle arrives at the customer for the first time. Because of random demands, vehicle routes will fail when a customer's demand exceeds remaining vehicle capacity. This research assumes split deliveries only when failures occur. In such a case, the vehicle may partially satisfy a customer demand, then complete it in another non necessarily immediate visit. Upon route failure, the vehicle must return to the depot to restore the vehicle capacity to Q. The problem objective is to find a routing