Computers & Operations Research 50 (2014) 1–13

Contents lists available at ScienceDirect

Computers & Operations Research

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

Paired cooperative reoptimization strategy for the vehicle routing problem with stochastic demands

Lin Zhu [a,b], Louis-Martin Rousseau [a,c], Walter Rei [a,d,*], Bo Li [b]

[a] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), Canada
[b] College of Management and Economics, Tianjin University, Tianjin, 300072, China
[c] Mathematics and Industrial Engineering, Ecole Polytechnique de Montreal, Canada
[d] Management Science School, UQAM, Canada

ARTICLE INFO

Available online 13 April 2014

Keywords:
Stochastic vehicle routing
Dynamic programming
Reoptimization
Heuristic

ABSTRACT

In this paper, we develop a paired cooperative reoptimization (PCR) strategy to solve the vehicle routing problem with stochastic demands (VRPSD). The strategy can realize reoptimization policy under cooperation between a pair of vehicles, and it can be applied in the multivehicle situation. The PCR repeatedly triggers communication and partitioning to update the vehicle assignments given real-time customer demands. We present a bilevel Markov decision process to model the coordination of a pair of vehicles under the PCR strategy. We also propose a heuristic that dynamically alters the visiting sequence and the vehicle assignment given updated information. We compare our approach with a recent cooperation strategy in the literature. The results reveal that our PCR strategy performs better, with a cost saving of around 20–30%. Moreover, embedding communication can save an average of 1.22%, and applying our partitioning method rather than an alternative can save an average of 3.96%.

1. Introduction

Vehicle routing problems (VRPs) involve designing a set of minimal-cost routes to meet customer demands under a group of operational constraints (see, e.g., [1–3]). In classical VRPs, the demands are assumed to be known with certainty, and all the relevant information to compute the routes is available in advance. However, in practice, customer demands and several other aspects are often stochastic. Solving the problem deterministically by replacing the stochastic parameters with their expected values does not give good solutions [4]. This justifies the development of stochastic models that can construct solutions with regard to the observed informational flow (i.e., when and how the values associated with the stochastic parameters become known).

In this paper, we consider the VRP with stochastic demands (VRPSD) in which the demand is known only when the vehicle arrives at the customer location. It has many real-world applications, such as local-deposit delivery and collection from bank branches [5], home oil delivery [6], beer distribution, and garbage collection [7].

In the VRPSD, a vehicle may reach a customer location without sufficient residual capacity to fulfill the demand, leading to a route *failure*, in which case a *recourse* action is necessary. Various recourse actions are possible: (i) replenishing the vehicle at the depot; (ii) scheduling a different vehicle to visit the customer where the failure occurred; or (iii) skipping the customer altogether (in this case a penalty is incurred). We consider (i), i.e., a driver performs a *replenishment* trip to the depot when a failure occurs.

Different modeling approaches have been developed to deal with the uncertain demands. These modeling approaches depend on the way both the routing and replenishment decisions are made, either *static* or *dynamic* [8]. For static approaches, stochastic programming with recourse (SPR) is often used [9]. It is a two-stage approach that minimizes the total cost of the planned routes and the expected recourse actions (e.g., [10]). Dynamic approaches, which apply a *reoptimization* policy [11,12], use a Markov decision process (MDP) to model the real-time decisions, given the available vehicle capacity and the set of unvisited customers (e.g., [9,13–15,8,16,12]).

Given the recent technological advances, reoptimization policies are now a viable strategy to decrease routing costs in the VRPSD context. However, efficiently solving the MDP models is challenging given the large numbers of actions, stages, and states involved. Therefore, most studies assume that a single vehicle is available.

Secomandi and Margot [8] observe that the existing literature on the VRPSD with reoptimization is scant, focusing on heuristic methods for the single-vehicle situation (e.g., [17–19,14,15,8,12]). To the best of our knowledge, only one study considers the

* Corresponding author.
E-mail addresses: zhulin.sife@gmail.com (L. Zhu), louis-martin.rousseau@polymtl.ca (L.-M. Rousseau), rei.walter@uqam.ca (W. Rei).

http://dx.doi.org/10.1016/j.cor.2014.03.027
0305-0548/

multivehicle case: Goodson et al. [16] propose a roll-out algorithm, real-time information is used, and the customers are dynamically assigned to different vehicles when the demands are revealed.

The solution of the MDP model for the VRPSD in the multi-vehicle context is challenging. Dynamic routing and replenishment decisions are necessary, and the assignment of customers to vehicles should also be performed dynamically. We propose the use of two general concepts that have proved to be efficient for the VRPSD: partial reoptimization of the routes and paired-vehicle cooperation.

The partial reoptimization technique was proposed by Secomandi and Margot [8] for the single-vehicle VRPSD. It computes optimal policies locally for subsets of states, to be used for the dynamic routing and replenishment when the demand is revealed. The paired-vehicle cooperation is based on the paired locally coordinated (PLC) scheme [20]. The PLC forms pairs of vehicles and shares customers within each pair, giving a solution in which each customer is dynamically served by a vehicle or its partner.

We focus on developing a cooperation strategy, the *paired cooperative reoptimization* (PCR) strategy, for a single pair of vehicles. We can then solve the multivehicle problem by clustering the customers into groups and serving the customers in each group with a pair of vehicles, as suggested by Ak and Erera [20]. The PCR strategy is based on the partial reoptimization technique and adds *communication* between the two vehicles. Via effective communication, the customers are dynamically chosen to be served by one of the two vehicles when the updated information becomes available.

This paper's main contribution is the development of the PCR recourse strategy, which is formulated as a bilevel MDP. This strategy enables a pair of vehicles to dynamically serve a set of customers under a reoptimization policy. We propose a heuristic that relies on both partial reoptimization and real-time communication to dynamically construct the routes performed by the pair of vehicles. We compare the performance of PCR strategy through a numerical study that shows the benefits, in terms of the total travel cost, with respect to other recourse strategies.

The remainder of this paper is organized as follows. In Section 2, we present our assumptions and discuss the general paired reoptimization problem. In Section 3, we give the definition of the PCR, and in Section 4 we discuss the bilevel MDP. In Section 5, we present our heuristic. Finally, in Section 6, we give the results of the computational study. We compare our algorithm with the PLC approach and describe two experiments that illustrate the cooperation of the PCR.

2. Problem definition

2.1. Notation and assumptions

In this paper, a single pair of vehicles cooperate to serve a set of customer demands. We use notation similar to that of Secomandi and Margot [8]. Given a complete network, let the set of nodes be $\{0, 1, \ldots, N\}$, with N a positive integer. Node 0 denotes the depot and $C = \{1, \ldots, N\}$ is the set of customers. The distances $d(i,j)$ between any two nodes i and j are known, symmetric, and satisfy the triangle inequality: $d(i,j) \le d(i,l) + d(l,j)$, with l an additional node. Two vehicles with the same capacity Q, denoted as $t1$ and $t2$, are initially located at the depot and must eventually return there. Let ξ_i, $i = 1, 2, \ldots, N$ be the discrete random variable that describes the demand of customer i. Its probability mass function is $p_i(e) = \Pr\{\xi_i = e\}$, $e = 0, 1, \ldots, E \le Q$, and $p_i(e) = 0$, $e = E+1, \ldots, Q$, with E a nonnegative integer. The customer demands ξ_i are independent of the vehicle routing/replenishment policy. The realization of ξ_i becomes known when the vehicle arrives at customer location i. The total depot capacity is at least $N \cdot E$, so that all the customers can be served.

We assume that each customer can be served by only one vehicle. Moreover, split deliveries are allowed, i.e., when a failure occurs, the vehicle delivers its existing load to the customer, then returns to the depot to reload, and subsequently completes the interrupted delivery.

The vehicles can communicate to dynamically modify their routes, and the locations, available capacities, and unvisited customers are visible to both of the vehicles. The information is shared under three assumptions. First, we ignore the time spent on loading and unloading and on planning (the vehicle assignments and the next customer to visit). Second, the vehicles are not permitted to have idle time. Third, the vehicles travel at the same speed. Therefore, at any given time, each vehicle's location and status (e.g., en route or replenishing) can be found by calculating the total distance traveled.

2.2. Formulation of general paired reoptimization problem

We describe the general problem with reference to the MDP formulations for the single-vehicle situation [12,8] and the multi-vehicle case [16].

The paired-vehicle problem is a special case of the multivehicle problem, and it can be stated as follows. When a vehicle finishes serving its current customer, a new customer will be assigned, and the vehicle must decide whether to visit the new customer directly or via the depot. The new customer is chosen from the set of unassigned customers. The vehicles must coordinate their efforts by considering the influence of each decision on the other vehicle and on the future cost.

The decisions occur when a vehicle completes an assignment, not when it arrives at a new customer and observes the demand. The next location is always a customer location and not the depot.

We formulate the problem as an MDP with stages in the set $\Omega' = \{0, 1, 2, \ldots, K'\}$. Each stage $k \in \Omega' \backslash \{0\}$ starts as a vehicle finishes its current assignment. The two vehicles may complete their assignments simultaneously and trigger the next stage together. K' is the final stage that occurs when no customer is unassigned. Let the state space for the process be Ψ'. For each stage $k \in \{0, 1, 2, \ldots, K'\}$, we characterize the corresponding state as $x_k = (l_1, l_2, q_1, q_2, R_k(l_1, l_2))$. Here, l_1 and l_2 are the customer locations where the two vehicles completed their last assignments, q_1 and q_2 are the available capacities after those assignments, and $R_k(l_1, l_2)$ is the set of remaining customers at stage k. The initial system state is $x_0 = (0, 0, Q, Q, C)$ and the final system state is $x_{K'} = (l_1, l_2, q_1, q_2, \phi)$. For example, suppose the current state is $x_k = (l_1, l_2, q_1, q_2, R_k(l_1, l_2))$, and the vehicles will next serve customers j_1 and j_2. Suppose that vehicle $t1$ finishes serving customer j_1 and triggers the next stage $k+1$, while vehicle $t2$ is either en route to customer j_2 or replenishing so as to meet j_2's demand. In this case, the state updates to $x_{k+1} = (j_1, l_2, q_1', q_2, R_{k+1}(j_1, l_2))$, where q_1' is the residual capacity of vehicle $t1$ that is $q_1' = q_1 - e$ (e being the realization of the demand for customer j_1).

Given state x_k, action (a_1^k, a_2^k) assigns the two vehicles to the next customer locations. Let $t \in \{1, 2\}$ represent the vehicles $t1$ and $t2$, and $z_k \subseteq \{1, 2\}$ be the set of vehicles that trigger stage k by completing their current assignments. Clearly, $z_k \ne \phi$, and $z_0 = \{1, 2\}$ indicates that both vehicles start to serve new assignments at the beginning. In addition, let the vehicles in $\{\{1, 2\} \backslash z_k\}$, which have not completed their assignments, continue on their planned routes. The set of actions available for state x_k, $k \in \Omega'$, is then

$$A(x_k) = \{(a_1^k, a_2^k) | a_t^k = j^{(1)} \text{ or } j^{(0)}, \quad \forall j \in R_{k-1}(l_1, l_2) \backslash \{a_1^{k-1}, a_2^{k-1}\} (k \ge 1)$$

$$\forall t \in z_k; \quad a_t^k = a_t^{k-1}, \ \forall t \in \{\{1, 2\} \backslash z_k\}; a_1^k \ne a_2^k\} \quad (1)$$