
Volume 217, Issue 2, 1 March 2012
ISSN 0377-2217
ELSEVIER
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
Available online at www.sciencedirect.com
SciVerse ScienceDirect
www.elsevier.com/locate/ejor

European Journal of Operational Research 217 (2012) 324–332

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing and Logistics

Single vehicle routing problems with a predefined customer sequence, compartmentalized load and stochastic demands

D.G. Pandelis [a], E.G. Kyriakidis [b,*], T.D. Dimitrakos [c]

[a] *Department of Mechanical Engineering, University of Thessaly, Athinon Avenue, Pedion Areos, 38334 Volos, Greece*
[b] *Department of Statistics, Athens University of Economics and Business, Patission 76, 10434 Athens, Greece*
[c] *Department of Mathematics, University of the Aegean, Karlovassi, 83200 Samos, Greece*

ARTICLE INFO

Article history:
Received 15 October 2010
Accepted 15 September 2011
Available online 22 September 2011

Keywords:
Logistics
Dynamic programming
Stochastic demands
Multiple-product delivery

ABSTRACT

We consider the problem of finding the optimal routing of a single vehicle that delivers K different products to N customers according to a particular customer order. The demands of the customers for each product are assumed to be random variables with known distributions. Each product type is stored in its dedicated compartment in the vehicle. Using a suitable dynamic programming algorithm we find the policy that satisfies the demands of the customers with the minimum total expected cost. We also prove that this policy has a specific threshold-type structure. Furthermore, we investigate a corresponding infinite-time horizon problem in which the service of the customers does not stop when the last customer has been serviced but it continues indefinitely with the same customer order. It is assumed that the demands of the customers at different tours have the same distributions. It is shown that the discounted-cost optimal policy and the average-cost optimal policy have the same threshold-type structure as the optimal policy in the original problem. The theoretical results are illustrated by numerical examples.

1. Introduction

One of the most widely studied problems in combinatorial optimization is the Vehicle Routing Problem (VRP). It was introduced by Dantzig and Ramser [6] and it is considered as a generalization of the classical travelling salesman problem. Surveys of the most important results that have been derived on this significant topic during the last fifty years are given by Toth and Vigo [18], Simchi-Levi et al. [15] and Liong et al. [9]. The VRP can be described as follows. Assume that we manage a fleet of vehicles that originate from one or several depots, and deliver goods to N geographically dispersed customers. Each vehicle starts its route from a depot, visits a subset of the customers, satisfies their demands and, finally, returns to the depot. The VRP is the problem of determining the optimal delivery route for each vehicle. The cost structure includes the costs of travelling from one customer to another and the costs of travelling from a customer back to a depot for restocking. The usual optimization criterion is the minimization of the total travel cost. A different interpretation of the VRP is the determination of the optimal collection route for each vehicle if the vehicles collect expired products instead of delivering new products. Collection of household waste, delivery of goods to supermarkets, petrol delivery to petrol stations, stocking of vending machines, supply of baked foods at food stores can be considered as applications of the VRP. Three important variants of the VRP that have been studied extensively are (i) the capacitated VRP, (ii) the VRP with time-windows and (iii) the VRP with pick-up and delivery. In the first problem all vehicles are identical with finite capacity, in the second problem the vehicles deliver goods to the customers during specific time-windows and in the third problem at each customer a pickup of expired products is performed before the delivery of new products. It is also possible to consider stochastic versions of the VRP if we assume that one or more parameters of the problem are random variables. For example, the number of customers to be serviced, the demands of the customers and the travel times can be random variables. The VRP and its variants are NP-hard problems and, therefore, many researchers have developed heuristics and metaheuristics (tabu search, simulated annealing, genetic algorithms, ant colony optimization) that reach a "good solution" (see e.g. [1,3,4,13]). Exact algorithms (for example branch-and-bound, branch-and-cut, branch-and-price methods) have also been applied to the VRP that give the optimal solution (see e.g. [2,5,7,10]).

In the present paper we consider a simple interesting variant of the VRP that was introduced by Tatarakis and Minis [16]. In this variant it is assumed that a single vehicle starts its route from a depot and delivers K different products to N customers according to a predefined customer sequence $1 \to 2 \to \cdots \to N$. The vehicle is divided into K sections and each section is suitable for one type of product only. The demand of the customer $j \in \{1,\ldots,N\}$ for

* Corresponding author.
E-mail addresses: d_pandelis@mie.uth.gr (D.G. Pandelis), ekyriak@aueb.gr, kyriak@aegean.gr (E.G. Kyriakidis), dimitheo@aegean.gr (T.D. Dimitrakos).

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.09.027

product $i \in \{1,\ldots,K\}$ is a discrete random variable. The demand of each customer for a specific product cannot exceed the capacity of the corresponding section of the vehicle. The objective is to find the policy that minimizes the expected total travel cost. Tatarakis and Minis [16] developed a suitable dynamic programming algorithm for this problem and, for $K = 2$, proved that the optimal policy has the following threshold-type structure. Denote z_i, $i = 1,2$, as the load of product $i \in \{1,2\}$ carried by the vehicle after the demands of customer $j \in \{1,\ldots,N-1\}$ have been satisfied. There exists a critical number $s_j(z_1) \geqslant 0$ such that it is optimal for the vehicle to proceed to the next customer if and only if $z_2 \geqslant s_j(z_1)$; otherwise, it is optimal to return to the depot for stock replenishment, and then resume the route. Note that the problem studied in [16] is a generalization of the problem introduced by Yang et al. [20] in which it is assumed that the vehicle delivers to the customers only one product, i.e. $K = 1$. Yang et al. proved that for each customer $j \in \{1,\ldots,N-1\}$ there exists a critical number s_j such that the optimal decision, after serving customer j, is to continue to customer $j + 1$ if the remaining quantity in the vehicle is greater than or equal to s_j, or return to the depot for stock replenishment if it is less than s_j. Kyriakidis and Dimitrakos [8] proved an analogous result if the demands of the customers are continuous random variables and developed a suitable dynamic programming algorithm for the determination of the critical numbers s_j, $j = 1,\ldots,N-1$. Tsirimbas et al. [19] studied the case of multiple-product deliveries when the demand of each customer for product $i \in \{1,\ldots,K\}$ is not a random variable but a constant number. They also assumed that the vehicle visits each customer only once and they designed a suitable dynamic programming algorithm for the determination of the optimal policy. It is noteworthy that Mendoza et al. [11] investigated recently a problem in which a fleet of identical multi-compartment vehicles serves N customers whose demands for each product $i \in \{1,\ldots,K\}$ is stochastic. In this work it is assumed that the fleet size is unlimited. It is not assumed that the customers are served according to a particular order. Mendoza et al. proposed a memetic algorithm for the determination of the optimal routing strategy. This algorithm couples genetic operators and local search procedures.

The main contribution in the present paper is that the structure of the optimal policy for the problem studied in [16] is proved for any positive integer K. Our proof is much simpler and more elegant than that given in [16] for $K = 2$. We also study a corresponding infinite time-horizon problem in which the service of the customers does not stop when the demands of the last customer N are satisfied but it continues indefinitely with the same customer order. It is assumed that, when the vehicle completes a tour, the demands of the customers for the products are renewed for the next tour and follow the same distributions. Using well-known results of the Markov decision theory we prove that the policy that minimizes the total expected discounted cost and the policy that minimizes the long-run expected average cost per unit time have the same threshold-type structure as the optimal policy in the initial finite-horizon problem. Furthermore, if we assume that the demand of the customer $j \in \{1,\ldots,N\}$ for the product $i \in \{1,\ldots,K\}$ is a continuous random variable, it is possible to show that the optimal policy has again the same threshold-type structure for the finite-horizon problem and for the discounted cost infinite-horizon problem.

The rest of the paper is organized as follows. In Section 2 we formulate the finite-horizon compartmentalized load problem and we develop two alternative dynamic programming algorithms that compute the optimal policy. The structure of the optimal policy is proved. In Section 3 we present the infinite-horizon problem and we prove the structure of the discounted-cost optimal policy and the structure of the average-cost optimal policy. We give numerical results for the computation of the average-cost optimal policy. In Section 4 we consider the case of stochastic continuous demands and we prove the structure of the finite-horizon optimal policy and the structure of the infinite-horizon discounted-cost optimal policy. In the last section the conclusions of the paper are summarized.

2. The finite-horizon problem

We consider a set of nodes $V = \{0,1,\ldots,N\}$ with node 0 denoting the depot and nodes $1,\ldots,N$ corresponding to customers. The customers are serviced in the order $1,2,\ldots,N$ by a vehicle consisting of K compartments with capacities $Q_1,Q_2,\ldots,Q_K$. Each compartment is suitable for one type of product only. There are K different products to be delivered to the customers. The vehicle starts its route from the depot loaded with Q_i items of product $i \in \{1,\ldots,K\}$ and after servicing all customers it returns to the depot. The demand of customer j, $j = 1,2,\ldots,N$, for product i, $i = 1,2,\ldots,K$, is a discrete random variable ξ_i^j; we assume that the joint probability distribution of each customer's demands is known. The actual demand of each customer becomes known upon the vehicle's arrival at the customer's site. We assume that the demand of each customer for a specific product cannot exceed the capacity of this product's compartment, that is, $\max_{j=1,2,\ldots,N} \xi_i^j \leqslant Q_i$, for $i \in \{1,\ldots,K\}$. When the vehicle visits customer j for the first time it satisfies as much demand as possible. If part of the demand for product $i \in \{1,\ldots,K\}$ is not satisfied, the vehicle goes to the depot, fills its compartments with loads $Q_1,\ldots,Q_K$ and returns to the customer to satisfy the demand. After satisfying the demand of the last customer, the vehicle returns to the depot. We denote by $c_{j,j+1}$, $j = 1,2,\ldots,N-1$, the travel cost between customers j and $j + 1$, and by c_{j0}, $j = 1,2,\ldots,N$, the travel cost between customer j and the depot. We naturally assume that these costs satisfy the triangle inequality, i.e.,

$$c_{i,i+1} \leqslant c_{i0} + c_{0,i+1}, \quad i = 1,\ldots,N-1.$$

The road network is depicted in Fig. 1.

Let $z_i \in \{0,\ldots,Q_i\}$, $i = 1,2,\ldots,K$, be the load of product i carried by the vehicle after the demand for some customer has been satisfied. Then, the vehicle either (i) proceeds directly to the next customer or (ii) goes to the depot, restocks with loads Q_i, $i = 1,2,\ldots,K$, of products $1,2,\ldots,K$, and then visits the next customer. Our objective is to determine a vehicle routing policy that minimizes the expected cost covered during a visit cycle. As mentioned in [19] a practical application of this problem could be the so called ex-van sales. In ex-van sales, the driver of the vehicle acts as a salesman. She visits her customers (supermarkets, kiosks, etc.) typically according to a predefined sequence. The demands of each

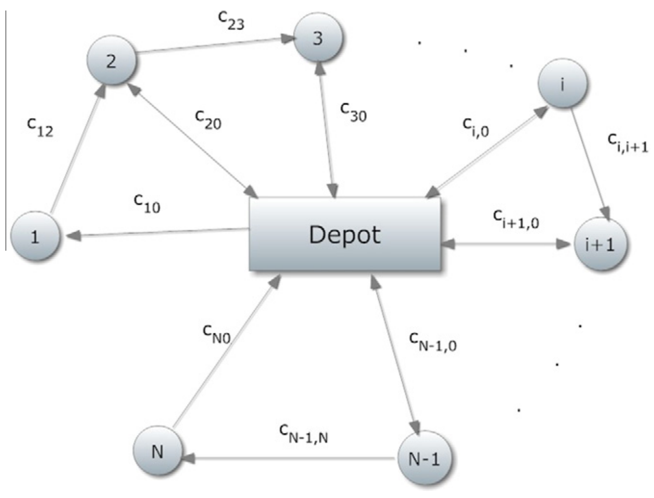


Fig. 1. The road network.

customer for the products are not known in advance but they are revealed upon arrival. If a customer's demand for a product exceeds the quantity that is loaded in the vehicle, the driver has to go to the depot for replenishment. We define vectors $\bar{z}=(z_1, z_2,\ldots,z_K)$, $\overline{Q}=(Q_1,Q_2,\ldots,Q_K)$, $\bar{\xi}^j=\left(\xi_1^j,\xi_2^j,\ldots,\xi_K^j\right)$, and denote by $P^j(\bar{\xi})=\Pr(\bar{\xi}^j=\bar{\xi})$, $\bar{\xi}=(\xi_1,\ldots,\xi_K)\in S=\{0,\ldots,Q_1\}\times\cdots\times\{0,\ldots,Q_K\}$, the joint probability mass function for the demand of customer $j\in\{1,\ldots,N\}$.

2.1. The optimal routing strategy

Let $f_j(\bar{z})$ be the minimum expected future cost when the load of product i carried by the vehicle after having satisfied demand for customer j is equal to z_i, and $\Xi_{\bar{z}}=\{\bar{\xi}:\xi_i\leqslant z_i, 1\leqslant i\leqslant K\}$. Then, an optimal routing strategy can be determined by the following dynamic programming equations (Algorithm 1). For $j=1,2,\ldots,N-1$ we have

$$f_j(\bar{z})=\min\{H_j(\bar{z}),A_j\},\quad \bar{z}\in\{0,\ldots,Q_1\}\times\cdots\times\{0,\ldots,Q_K\},\tag{1}$$

where

$$H_j(\bar{z})=c_{j,j+1}+\sum_{\bar{\xi}\in\Xi_{\bar{z}}}f_{j+1}(\bar{z}-\bar{\xi})P^{j+1}(\bar{\xi})+\sum_{\bar{\xi}\notin\Xi_{\bar{z}}}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi}),\tag{2}$$

$$A_j=c_{j0}+c_{j+1,0}+\sum_{\bar{\xi}\in S}f_{j+1}(\overline{Q}-\bar{\xi})P^{j+1}(\bar{\xi}).\tag{3}$$

In the boundary we have

$$f_N(\bar{z})=c_{N0},\quad \bar{z}\in S\tag{4}$$

and the minimum expected cost during a cycle is

$$f_0=c_{10}+\sum_{\bar{\xi}\in S}f_1(\overline{Q}-\bar{\xi})P^1(\bar{\xi}).$$

The left-hand-side term in the curly brackets in (1) corresponds to the action of proceeding to the next customer and the right-hand-side term corresponds to the action of going to the depot for restocking. In Eq. (2) the quantity $(\bar{z}-\bar{\xi})^-$ is defined as $(\min(z_1-\xi_1,0),\ldots,\min(z_K-\xi_K,0))$. Before we proceed to prove the main result, we need the following lemma.

Lemma 1. *For $j=1,2,\ldots,N$ we have $f_j(\bar{z})\leqslant 2c_{j,0}+f_j(\overline{Q})$.*

Proof. For $j=N$ the result is trivial. For $j<N$ we have from Eq. (1)

$$f_j(\bar{z})\leqslant A_j=c_{j,0}+c_{j+1,0}+\sum_{\bar{\xi}\in S}f_{j+1}(\overline{Q}-\bar{\xi})P^{j+1}(\bar{\xi}).\tag{5}$$

When the vehicle is fully loaded it is obviously optimal to proceed to the next customer. Therefore,

$$f_j(\overline{Q})=c_{j,j+1}+\sum_{\bar{\xi}\in S}f_{j+1}(\overline{Q}-\bar{\xi})P^{j+1}(\bar{\xi}).\tag{6}$$

The result follows from Eqs. (5) and (6), and the triangle inequality. □

The following theorem characterizes the optimal vehicle routing strategy after the demand of customer j has been satisfied.

Theorem 1

(i) *For given $z_i\in\{0,\ldots,Q_i\}$, $i=1,\ldots,K-1$, there exists an integer $s_j(z_1,\ldots,z_{K-1})\in\{0,\ldots,Q_K+1\}$ such that it is optimal for the vehicle to proceed to customer $j+1$ if $z_K\geqslant s_j(z_1,z_2,\ldots,z_{K-1})$.*
(ii) *$s_j(z_1,z_2,\ldots,z_{K-1})$ is non-increasing in each of its arguments.*

Proof. In view of (1), for both parts of the theorem it suffices to show that $H_j(\bar{z})$ is non-increasing in its arguments. We will also need to prove that $f_j(\bar{z})$ is non-increasing in its arguments. The proof is by induction on j. First, the induction base is established by $f_N(\bar{z})$ being non-increasing (Eq. (4)). Then, assuming that $f_{j+1}(\bar{z})$ is non-increasing, we will show that $H_j(\bar{z})$ and $f_j(\bar{z})$ are non-increasing. Once we show that $H_j(\bar{z})$ is non-increasing, the monotonicity of $f_j(\bar{z})$ follows directly from Eq. (1). Therefore, we focus on $H_j(\bar{z})$.

Let $\bar{e}_i$ be a vector of size K whose components are equal to 0 except the i-th component that is equal to one. Because of symmetry, to prove that $H_j(\bar{z})$ is non-increasing in each of its arguments, it suffices to show that $H_j(\bar{z})\geqslant H_j(\bar{z}+\bar{e}_i)$. We have from Eq. (2)

$$\begin{aligned}H_j(\bar{z})&=c_{j,j+1}+\sum_{\bar{\xi}\in\Xi_{\bar{z}}}f_{j+1}(\bar{z}-\bar{\xi})P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}\notin\Xi_{\bar{z}}:\xi_i\leqslant z_i}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}\in\Xi_{\bar{z}+\bar{e}_i}:\xi_i=z_i+1}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}\notin\Xi_{\bar{z}+\bar{e}_i}:\xi_i=z_i+1}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}:\xi_i>z_i+1}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi})\end{aligned}\tag{7}$$

and

$$\begin{aligned}H_j(\bar{z}+\bar{e}_i)&=c_{j,j+1}+\sum_{\bar{\xi}\in\Xi_{\bar{z}}}f_{j+1}(\bar{z}+\bar{e}_i-\bar{\xi})P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}\in\Xi_{\bar{z}+\bar{e}_i}:\xi_i=z_i+1}f_{j+1}(\bar{z}+\bar{e}_i-\bar{\xi})P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}\notin\Xi_{\bar{z}+\bar{e}_i}:\xi_i\leqslant z_i}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}+\bar{e}_i-\bar{\xi})^-)]P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}\notin\Xi_{\bar{z}+\bar{e}_i}:\xi_i=z_i+1}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}+\bar{e}_i-\bar{\xi})^-)]P^{j+1}(\bar{\xi})\\&\quad+\sum_{\bar{\xi}:\xi_i>z_i+1}[2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}+\bar{e}_i-\bar{\xi})^-)]P^{j+1}(\bar{\xi}).\end{aligned}\tag{8}$$

We compare $H_j(\bar{z})$ and $H_j(\bar{z}+\bar{e}_i)$ term by term. The first, fourth, and fifth sums in Eq. (7) are at least equal to the respective sums in Eq. (8) by the induction hypothesis. The second sum in Eq. (7) is at least equal to the third sum in Eq. (8), also by the induction hypothesis. It remains to compare the third sum in Eq. (7) with the second sum in Eq. (8). We have

$$f_{j+1}(\bar{z}+\bar{e}_i-\bar{\xi})\leqslant 2c_{j+1,0}+f_{j+1}(\overline{Q})\leqslant 2c_{j+1,0}+f_{j+1}(\overline{Q}+(\bar{z}-\bar{\xi})^-),$$

where the first inequality follows from Lemma 1 and the second from the induction hypothesis. □

Remark 1. If $s_j(z_1,\ldots,z_{K-1})=Q_K+1$, it is optimal for the vehicle to return to the depot for all $z_K\in\{0,\ldots,Q_K\}$.

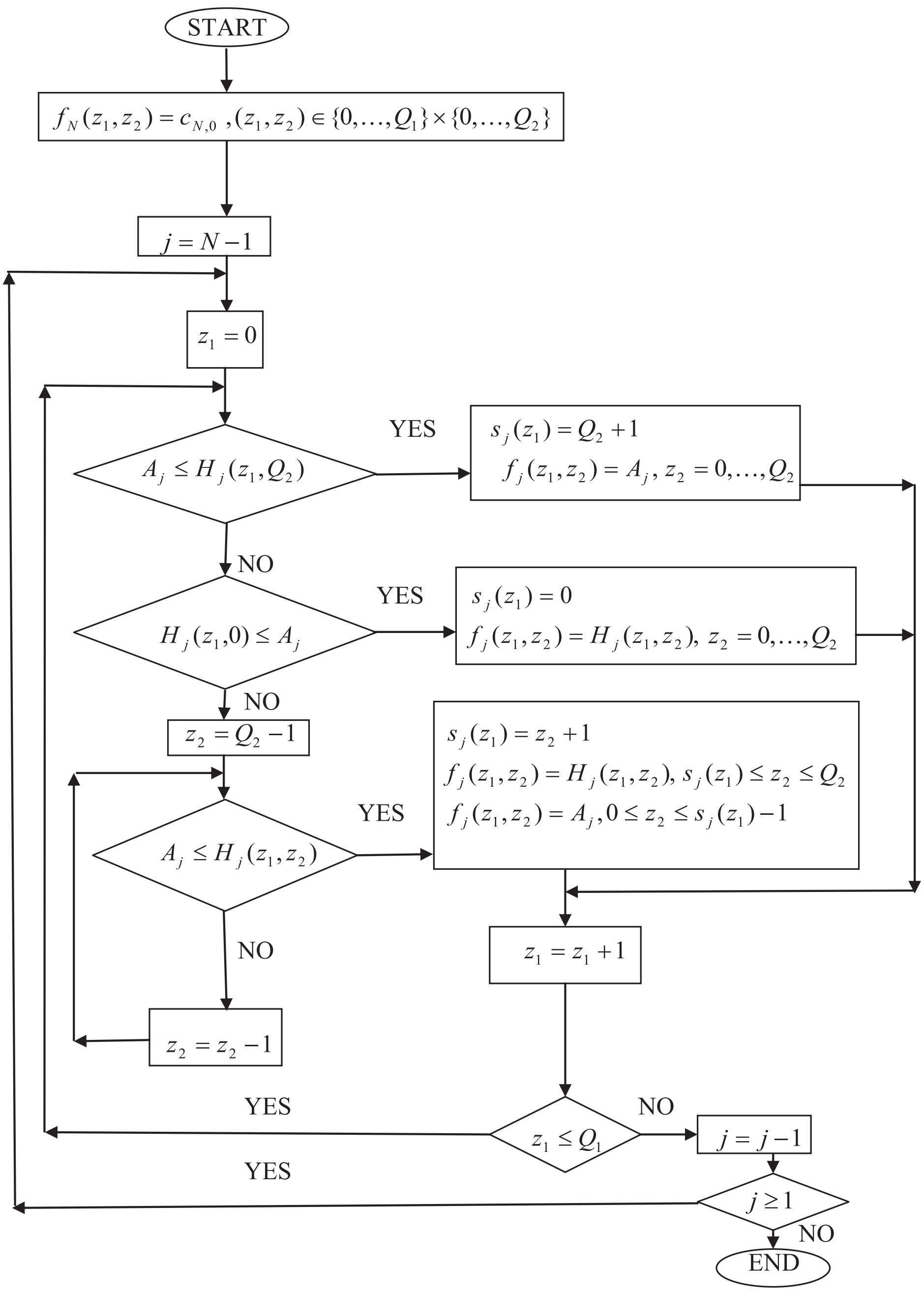
Remark 2. Tatarakis and Minis [16] proved Part (i) of the above result for $K=2$. Their proof involves complicated algebraic expressions and cannot be extended for $K>2$. They also assumed that for $j\in\{1,\ldots,N\}$ the random variables $\xi_1^j,\ldots,\xi_K^j$ are independent. In our proof we did not impose this assumption.

2.2. An improved special-purpose dynamic programming algorithm

In view of the above theorem for $K=2$ the optimal policy (i.e. the critical numbers $s_j(z_1)$ for $j=1,\ldots,N-1$, $z_1=0,\ldots,Q_1$)

can be found by the following algorithm that is presented as a flowchart.

Algorithm 2. Algorithm for the determination of $s_j(z_1)$, $z_1 = 0,\ldots,Q_1$



The above algorithm is faster than the initial dynamic programming algorithm based on Eqs. (1)–(4) since it does not calculate for $j = 1,\ldots,N-1$ and $z_1 = 0,\ldots,Q_1$ the quantities $H_j(z_1,z_2)$, $0 \leqslant z_2 < s_j(z_1) - 1$.

2.3. Numerical results

As illustration we present the following example. Suppose that $N = 10$ and $Q_1 = Q_2 = 5$. The travel costs between customers j and $j+1$, $j = 1,\ldots,9$, are given by: $c_{12} = 27$, $c_{23} = 18$, $c_{34} = 27$, $c_{45} = 22$, $c_{56} = 24$, $c_{67} = 25$, $c_{78} = 23$, $c_{89} = 22$ and $c_{9,10} = 25$. The travel costs between customers j, $j = 1,\ldots,10$, and the depot are given by: $c_{10} = 24$, $c_{20} = 22$, $c_{30} = 23$, $c_{40} = 25$, $c_{50} = 22$, $c_{60} = 20$, $c_{70} = 21$, $c_{80} = 20$, $c_{90} = 19$ and $c_{10,0} = 24$. Note that these costs satisfy the triangle inequality. We assume that for each customer $j = 1,\ldots,10$ the demands ξ_1^j and ξ_2^j for products 1 and 2 are independent and follow the discrete uniform distribution in the set $\{0,\ldots,Q_1\}$ and the binomial distribution $B(Q_2,p)$, respectively. This means that for each customer $j = 1,\ldots,10$, $\Pr\left(\xi_1^j = x\right) = (Q_1+1)^{-1}$, $x = 0,\ldots,Q_1$, and $\Pr\left(\xi_2^j = y\right) = \binom{Q_2}{y} p^y (1-p)^{Q_2-y}$, $y = 0,\ldots,Q_2$, $p \in (0,1)$. In Table 1 below we present the critical numbers $s_j(z_1)$, $z_1 = 0,\ldots,Q_1$,

Table 1
The critical numbers $s_j(z_1)$ of the optimal policy for each value of $p \in \{0.1,\ldots,0.9\}$.

Customer j	$s_j(0)$	$s_j(1)$	$s_j(2)$	$s_j(3)$	$s_j(4)$	$s_j(5)$
1	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	2,3,3,3,3,3,4,5,5	0,1,2,2,2,3,3,4,5	0,1,1,2,2,3,3,4,5
2	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	2,2,3,3,3,3,4,4,5	0,1,1,2,2,3,3,4,5	0,0,1,1,2,2,3,4,5	0,0,1,1,2,2,3,4,5
3	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	2,3,3,3,3,3,4,4,5	1,1,2,2,3,3,4,4,5	0,1,1,2,2,3,3,4,5
4	6,6,6,6,6,6,6,6,6	6,6,6,6,6 6,6,6,6	3,3,3,3,3,4,4,5,5	0,1,2,2,2,3,3,4,5	0,1,1,2,2,3,3,4,5	0,0,1,1,2,2,3,4,5
5	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	1,2,3,3,3,4,4,5,5	0,1,2,2,3,3,4,4,5	0,1,1,2,2,3,3,4,5
6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	4,4,4,4,4,4,4,5,5	1,1,2,2,3,3,4,4,5	0,1,2,2,3,3,4,4,5
7	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	1,2,3,3,3,3,4,5,5	0,1,2,2,3,3,4,4,5	0,1,1,2,2,3,3,4,5
8	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	1,2,3,3,3,4,4,5,5	0,1,2,2,2,3,4,4,5	0,1,1,2,2,3,3,4,5
9	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	6,6,6,6,6,6,6,6,6	2,2,3,4,4,5,5,5,5	1,2,2,3,3,4,4,5,5	1,1,2,2,3,3,4,4,5

Table 2
The values of f_0, h_0, C for each value of p.

p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
f_0	365.68	367.13	371.98	380.94	384.45	389.33	394.66	398.20	403.99
h_0	429.61	430.65	433.58	444.01	451.19	472.20	500.03	543.34	591.86
C	440	440	440	440	440	440	440	440	440

that correspond to the optimal policy for each customer j, $j = 1,\ldots,N-1$ and for each value of $p \in \{0.1, 0.2,\ldots,0.9\}$. In each cell of the table, the first critical number corresponds to $p = 0.1$, the second critical number corresponds to $p = 0.2,\ldots,$ the ninth critical number corresponds to $p = 0.9$.

Note that Part (ii) of Theorem 1 is confirmed numerically since, for fixed p, $s_j(z_1)$ is non-increasing in $z_1 \in \{0,\ldots,5\}$ for each customer $j \in \{1,\ldots,9\}$. We also observe that, for each customer $j \in \{1,\ldots,N-1\}$, $s_j(z_1)$ is non-decreasing as p increases. This is intuitively reasonable since, as p increases, the expected value of the demand for product 2 increases and, therefore, the action of returning to the depot for restocking becomes more favourable.

In Table 2 above we give for each $p \in \{0.1,\ldots,0.9\}$, (i) the minimum total expected cost $f_0 = c_{10} + \sum_{(\xi_1,\xi_2)\in S} f_1(Q_1 - \xi_1, Q_2 - \xi_2)P^1(\xi_1,\xi_2)$, (ii) the total expected cost h_0 that we would have if the vehicle after serving customer $j \in \{1,\ldots,N-1\}$ always proceeds directly to customer $j+1$ and (iii) the total cost C that we would have if the vehicle, after serving each customer $j \in \{1,\ldots,N-1\}$ always returns to the depot for stock replenishment before proceeding to customer $j+1$. Note that C is equal to $2\sum_{i=1}^{10} c_{0i}$ and does not depend on p. The quantity h_0 can be computed from the following equations

$$h_0 = c_{10} + \sum_{(\xi_1,\xi_2)\in S} \widetilde{H}_1(Q_1 - \xi_1, Q_2 - \xi_2)P^1(\xi_1,\xi_2), \text{ where}$$

$$\widetilde{H}_j(z_1,z_2) = c_{j,j+1} + \sum_{(\xi_1,\xi_2)\in \Xi_{(z_1,z_2)}} \widetilde{H}_{j+1}(z_1 - \xi_1, z_2 - \xi_2)P^{j+1}(\xi_1,\xi_2) + \sum_{(\xi_1,\xi_2)\notin \Xi_{(z_1,z_2)}} \Big[2c_{j+1,0} + \widetilde{H}_{j+1}(Q_1 + (z_1 - \xi_1)^-, Q_2 + (z_2 - \xi_2)^-)\Big]P^{j+1}(\xi_1,\xi_2), \quad (z_1,z_2) \in S, j = 1,\ldots,N-1,$$

$$\widetilde{H}_N(z_1,z_2) = c_{N0}, \quad (z_1,z_2) \in S.$$

From the above table we see that, for all values of $p \in \{0.1,\ldots,0.9\}$, f_0 is considerably smaller than h_0 and C. We also observe that f_0 and h_0 increase as p increases. Note that for $p \leqslant 0.3$ ($p > 0.3$) the policy that chooses always the action of proceeding to the next customer is better (worse) than the policy that always chooses the action of returning to the depot for restocking.

Consider in our example the optimal routing strategy for e.g. $p = 0.5$. The critical numbers $s_j(z_1)$, $z_1 = 0,\ldots,Q_1$, for $j = 1,\ldots,N-1$ that correspond to this strategy are given by the fifth numbers in the cells of Table 1. From Table 2 we see that the total expected cost of this strategy is equal to 384.45. If we perturb slightly this strategy by increasing the critical numbers by one unit (if $s_j(z_1) < Q_1$) or by decreasing them by one unit (if $s_j(z_1) > 0$) we find that the total expected cost becomes 387.21 and 387.89, respectively. Therefore we can conclude that slight changes of the optimal critical numbers do not cause a significant increase of the total expected cost.

We implemented the initial dynamic programming algorithm based on Eqs. (1)–(4) and the special-purpose dynamic programming algorithm by running the corresponding Matlab programs on a personal computer equipped with an Intel® Core™ 2 Duo, 2.5 GHz processor and 4 GB of RAM. We will compare the efficiencies of these algorithms by considering again the same example with $p = 0.5$ but with variable $Q = Q_1 = Q_2$. In Fig. 2 we present the graphs that show, as Q varies, the variation of the computation times (expressed in seconds) required by the initial dynamic programming algorithm (Algorithm 1) and by the special-purpose algorithm Algorithm 2.

From the figure below we observe that, as Q increases, the computation times for both algorithms increase non-linearly. The computation time required by the special-purpose algorithm is

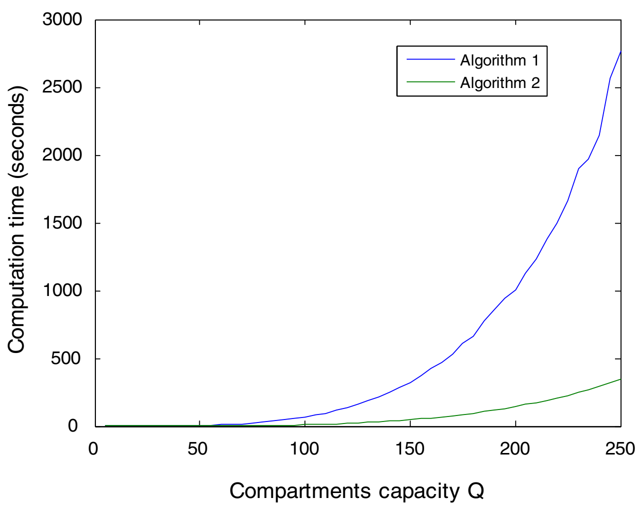


Fig. 2. The computation times of the algorithms as Q varies.

considerably smaller than the computation time required by our initial algorithm, especially for high values of Q.

3. The infinite-horizon problem

We modify the problem that we introduced in the previous section by considering an infinite-time horizon problem in which the service of the customers does not stop when the last customer N is serviced but it continues indefinitely with the same customer order. This means that after the service of customer N, the vehicle services again customer 1, customer 2, and so on. Let c_{N1} denote the travel cost from customer N to customer 1. The road network is depicted in Fig. 3.

The demands of the customers for all products are renewed at successive tours of the vehicle. We assume that the demand ξ_i^j of customer $j \in \{1,\ldots,N\}$ for product $i \in \{1,\ldots,K\}$ is a discrete random variable with the same distribution at each tour. As in the initial problem, the vehicle driver after the completion of the service of a customer has to choose among two decisions: (i) to proceed to the next customer and (ii) to go to the depot for restocking and then resume the route. It can be assumed that the driver selects her decisions at equidistant time epochs (e.g. every hour) after the completion of the service of the customers. This means that if, for example, the service of the 4th customer has been completed and the decision (i.e. proceed to the 5th customer or return to the depot for replenishment and then proceed to the 5th customer) is selected at 1 pm then the next decision (i.e. proceed to the 6th customer or return to the depot for replenishment and then proceed to the 6th customer) is selected at 2 pm. Although we impose this assumption in order to apply well-known results from the theory of Markov decision processes, there are situations in which this assumption may hold, as the so-called ex-van sales. As mentioned in Tsirimpas et al. [19], in ex-van sales the driver of the vehicle acts as a salesman. She visits her customers (supermarkets, stores, kiosks, etc.) in an area typically according to a predefined sequence. The demands of each customer for new products are not known in advance but they are revealed upon arrival. If the products must be consumed in a short period (e.g. milk, fruits, vegetables), it seems reasonable that the supply of the customers with new products continues with the same customer order for a long time horizon. It can be assumed that the driver selects her decisions at equidistant time epochs (e.g. every hour). Another example could be the routing of a self propelled vehicle in a manufacturing shop that transfers discrete parts to workcenters in a predefined sequence (see Rembold et al. [12]). Note that in addition to the main pathway connecting the workcenters, there are spurs connecting each workcenter with the material warehouse, allowing the return and the reloading of the vehicle. It can be assumed that this process continues for a long time without interruption and the decisions for the routing of the vehicle are selected after the completion of the service at the workcenters at equidistant time epochs. A third application of this system could be the situation in which a vehicle collects at equidistant time epochs waste from specific points according to a predefined order. When the vehicle is filled up with waste it must go to a specific area to empty it and then return to its route.

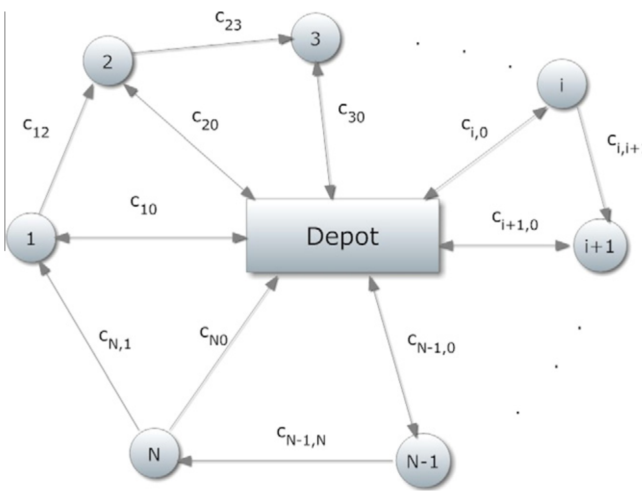


Fig. 3. The road network for the infinite-horizon problem.

3.1. The structure of the discounted-cost optimal policy and of the average-cost optimal policy

The usual optimization criteria in the infinite-horizon setting are the minimization of the expected total discounted cost and the minimization of the expected long-run average cost per unit time. Using well-known results of Markov decision processes (see Chapter 6 in Ross [14]) we will see that, under any one of these criteria, the optimal policy has the same structure as the optimal policy in the finite-horizon problem. The state space of the system is

$$I = \{(j,\bar{z}) : j = 1,\ldots,N, \bar{z} \in \{0,\ldots,Q_1\} \times \cdots \times \{0,\ldots,Q_K\}\}.$$

The state $(j,\bar{z})$ represents the situation at which the vehicle has completed the service of customer j and the remaining load of product $i \in \{1,\ldots,K\}$ is z_i. Let $V_n^\alpha(j,\bar{z}),\ 1 \leqslant j \leqslant N,\ 0 < \alpha < 1$, be the minimum n-step expected discounted cost if the initial state is $(j,\bar{z}) \in I$. The use of the discount factor $\alpha \in (0,1)$ can be explained by the economic idea that a cost to be incurred in the future is discounted in today's money, and thus we discount such costs at a rate α per unit of time. The quantity $V_n^\alpha(j,\bar{z})$ satisfies the following dynamic programming equation for $n = 1,2,\ldots$

$$V_n^\alpha(j,\bar{z}) = \min\left\{H_n^\alpha(j,\bar{z}), c_{j0} + c_{0,j+1} + \alpha\sum_{\bar{\xi}\in S} V_{n-1}^\alpha(j+1,\overline{Q}-\bar{\xi})P^{j+1}(\bar{\xi})\right\},$$

where,

$$H_n^\alpha(j,\bar{z}) = c_{j,j+1} + \alpha\sum_{\bar{\xi}\in\Xi_{\bar{z}}} V_{n-1}^\alpha(j+1,\bar{z}-\bar{\xi})P^{j+1}(\bar{\xi}) + \sum_{\bar{\xi}\notin\Xi_{\bar{z}}}[2c_{j+1,0} + \alpha V_{n-1}^\alpha(j+1,\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi}),$$

$$V_0^\alpha(j,\bar{z}) = 0.$$

In the above equation we assume that $N + 1$ is equal to 1 since the next customer of customer N is customer 1. It can be shown by induction on n that $V_n^\alpha(j,\bar{z})$ and $H_n^\alpha(j,\bar{z})$ are non-increasing in z_i, $i = 1,\ldots,K$, in the same way as we showed the monotonicity of $f_j(\bar{z})$ and $H_j(\bar{z})$ in the proof of Theorem 1. Let $V^\alpha(j,\bar{z})$ be the minimum expected α-discounted total cost if $(j,\bar{z})$ is the initial state of the system. This quantity is finite since $\alpha \in (0,1)$ and the state space is finite. It satisfies the following optimality equation:

$$V^\alpha(j,\bar{z}) = \min\left\{H^\alpha(j,\bar{z}), c_{j0} + c_{0,j+1} + \alpha\sum_{\bar{\xi}\in S} V^\alpha(j+1,\overline{Q}-\bar{\xi})P^{j+1}(\bar{\xi})\right\},$$

where,

$$H^\alpha(j,\bar{z}) = c_{j,j+1} + \alpha\sum_{\bar{\xi}\in\Xi_{\bar{z}}} V^\alpha(j+1,\bar{z}-\bar{\xi})P^{j+1}(\bar{\xi}) + \sum_{\bar{\xi}\notin\Xi_{\bar{z}}}[2c_{j+1,0} + \alpha V^\alpha(j+1,\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi}).$$

The left-hand side term in the curly brackets of the optimality equation corresponds to the decision of going directly to the next customer while the other term corresponds to the decision of returning to the depot for restocking. It is well-known (see Corollary 6.6 in [14]) that, as $n \to \infty$, $V_n^\alpha(j,\bar{z}) \to V^\alpha(j,\bar{z})$ and $H_n^\alpha(j,\bar{z}) \to H^\alpha(j,\bar{z})$. Hence, $V^\alpha(j,\bar{z})$ and $H^\alpha(j,\bar{z})$ are non-increasing in z_i, $i = 1,\ldots,K$. This implies that the α-discounted cost optimal policy has the threshold-type structure described in Theorem 1.

We focus now on the minimization of the expected average cost. First we note that the state $(1,\bar{0})$ is accessible from any other state under any stationary policy. From Corollary 6.20 in Ross [14] it follows that there exist numbers g and $h(j,\bar{z}), (j,\bar{z}) \in I$, such that

$$h(j,\bar{z}) = \min\left\{H(j,\bar{z}), c_{j0} + c_{0,j+1} - g + \sum_{\bar{\xi}\in S} h(j+1,\overline{Q}-\bar{\xi})P^{j+1}(\bar{\xi})\right\}, \quad (j,\bar{z}) \in I, \tag{9}$$

where,

$$H(j,\bar{z}) = c_{j,j+1} - g + \sum_{\bar{\xi}\in\Xi_{\bar{z}}} h(j+1,\bar{z}-\bar{\xi})P^{j+1}(\bar{\xi}) + \sum_{\bar{\xi}\notin\Xi_{\bar{z}}} [2c_{j+1,0} + h(j+1,\overline{Q}+(\bar{z}-\bar{\xi})^-)]P^{j+1}(\bar{\xi}).$$

Eq. (9) is known as average-cost optimality equation. The number g is the minimum average cost. There also exists a sequence $\alpha_n \to 1$ (see Theorem 6.18 in [14]) such that

$$h(j,\bar{z}) = \lim_{n\to\infty}[V^{\alpha_n}(j,\bar{z}) - V^{\alpha_n}(1,\bar{0})], \quad (j,\bar{z}) \in I. \tag{10}$$

The monotonicity of $H^{\alpha_n}(j,\bar{z})$ with respect to z_i, $i = 1,\ldots,K$, implies that the expression

$$\alpha_n \sum_{\bar{\xi}\in\Xi_{\bar{z}}} [V^{\alpha_n}(j+1,\bar{z}-\bar{\xi}) - V^{\alpha_n}(1,\bar{0})]P^{j+1}(\bar{\xi}) + \sum_{\bar{\xi}\notin\Xi_{\bar{z}}} [2c_{j+1,0} + \alpha_n[V^{\alpha_n}(j+1,\overline{Q}+(\bar{z}-\bar{\xi})^-) - V^{\alpha_n}(1,\bar{0})]]P^{j+1}(\bar{\xi})$$

is non-increasing in z_i, $i = 1,\ldots,K$. This property is valid if we take the limit as $n \to \infty$. In view of (10) it is deduced that $H(j,\bar{z})$ is non-increasing in z_i, $i = 1,\ldots,K$. From (9) it follows that the average-cost optimal policy has the same threshold structure as the finite-horizon optimal policy and the discounted-cost optimal policy.

3.2. Computation of the average-cost optimal policy

The average-cost optimal policy can be found numerically by the value-iteration algorithm, the policy-iteration algorithm and the linear programming formulation. We refer to Chapter 3 of Tijms's [17] book for a detailed description of these algorithms. To implement these algorithms we must specify the one-step transition probabilities and the one-step expected transition costs. Let $a \in \{0,1\}$ be the action that is selected when the system at a decision epoch is at state $(j,\bar{z}) \in I$. We assume that the action $a = 0$ means that the vehicle goes to the next customer while the action $a = 1$ means that it goes to the depot for stock replenishment. Let $p_{(j,\bar{z})(j+1,\bar{w})}(a)$ be the probability that the state at the next decision epoch will be the state $(j+1,\bar{w})$ if the present state is $(j,\bar{z})$ and the action a is selected, and let $C((j,\bar{z}),a)$ be the corresponding expected cost. Taking into account the structure of the model it is possible to specify these quantities. We give below when $K = 2$ some of these quantities:

$$p_{(j,z_1,z_2)(j+1,w_1,w_2)}(0) = P\left(\xi_1^{j+1} = z_1 - w_1, \xi_2^{j+1} = z_2 - w_2\right), \quad 0 < z_1 < Q_1,\ 0 < z_2 < Q_2,\ 0 \leqslant w_1 < z_1,\ 0 \leqslant w_2 < z_2,$$

$$p_{(j,z_1,z_2)(j+1,z_1,w_2)}(0) = P\left(\xi_1^{j+1} = Q_1, \xi_2^{j+1} = z_2 + Q_2 - w_2\right), \quad 0 < z_1 < Q_1,\ 0 < z_2 < Q_2,\ z_2 < w_2 < Q_2,$$

$$p_{(j,z_1,z_2)(j+1,z_1,z_2)}(0) = P\left(\xi_1^{j+1} = 0, \xi_2^{j+1} = 0\right) + P\left(\xi_1^{j+1} = Q_1, \xi_2^{j+1} = Q_2\right), \quad 0 < z_1 < Q_1,\ 0 < z_2 < Q_2,$$

$$p_{(j,z_1,z_2)(j+1,Q_1,w_2)}(0) = P\left(\xi_1^{j+1} \leqslant z_1, \xi_2^{j+1} = z_2 + Q_2 - w_2\right), \quad 0 < z_1 < Q_1,\ 0 < z_2 < Q_2,\ z_2 < w_2 < Q_2,$$

$$p_{(j,0,Q_2)(j+1,w_1,Q_2)}(0) = P\left(\xi_1^{j+1} = Q_1 - w_1\right), \quad 0 < w_1 \leqslant Q_1,$$

$$p_{(j,0,Q_2)(j+1,w_1,w_2)}(0) = 0, 0 < w_1 \leqslant Q_1, \quad 0 \leqslant w_2 < Q_2,$$

$$p_{(j,z_1,z_2)(j+1,w_1,w_2)}(1) = P\left(\xi_1^{j+1} = Q_1 - w_1, \xi_2^{j+1} = Q_2 - w_2\right), \quad 0 \leqslant z_1 \leqslant Q_1,\ 0 \leqslant z_2 \leqslant Q_2,\ 0 \leqslant w_1 \leqslant Q_1,\ 0 \leqslant w_2 \leqslant Q_2,$$

$$C((j,z_1,z_2),0) = c_{j,j+1} + 2c_{j+1,0}\left[1 - P\left(\xi_1^{j+1} \leqslant z_1, \xi_2^{j+1} \leqslant z_2\right)\right], \quad 0 \leqslant z_1 \leqslant Q_1,\ 0 \leqslant z_2 \leqslant Q_2,$$

$$C((j,z_1,z_2),1) = c_{j,0} + c_{0,j+1}, \quad 0 \leqslant z_1 \leqslant Q_1,\ 0 \leqslant z_2 \leqslant Q_2.$$

As illustration we present the following example. We suppose that $N = 10$, $K = 2$ and $Q_1 = Q_2 = 10$. The travel costs between customers j and $j+1$, $j = 1,\ldots,10$, are given by: $c_{12} = 17$, $c_{23} = 18$, $c_{34} = 20$, $c_{45} = 19$, $c_{56} = 19$, $c_{67} = 13$, $c_{78} = 10$, $c_{89} = 15$, $c_{9,10} = 19$ and $c_{10,1} = 16$. The travel costs between customer j, $j = 1,\ldots,10$, and the depot are given by: $c_{10} = 10$, $c_{20} = 12$, $c_{30} = 13$, $c_{40} = 15$, $c_{50} = 17$, $c_{60} = 12$, $c_{70} = 17$, $c_{80} = 20$, $c_{90} = 18$ and $c_{10,0} = 13$. The demands of the customers for products 1 and 2 are independent discrete uniformly distributed random variables, such that for each customer $j = 1,\ldots,N$, $P\left(\xi_1^j = x\right) = (Q_1+1)^{-1}$, $x = 0,\ldots,Q_1$, and

Table 3
The critical numbers of the optimal policy.

Customer j	$s_j(0)$	$s_j(1)$	$s_j(2)$	$s_j(3)$	$s_j(4)$	$s_j(5)$	$s_j(6)$	$s_j(7)$	$s_j(8)$	$s_j(9)$	$s_j(10)$
1	11	11	11	11	11	11	11	11	10	9	8
2	11	11	11	11	11	11	11	10	9	8	7
3	11	11	11	11	11	11	11	10	8	8	7
4	11	11	11	11	11	11	9	8	7	6	6
5	11	11	11	11	11	8	7	6	5	4	4
6	11	11	11	11	8	6	5	5	4	4	4
7	11	11	8	6	4	4	3	3	2	2	2
8	11	11	11	9	7	6	5	4	4	3	3
9	11	11	11	11	11	10	8	7	6	6	5
10	11	11	11	11	11	11	11	9	8	7	7

$P\left(\xi_2^j = y\right) = (Q_2+1)^{-1},\ y = 0,\dots,Q_2$. We implemented the standard policy-iteration algorithm for the determination of the optimal average-cost policy. After four iterations this algorithm converges to the optimal stationary policy with average cost equal to 215.01. The structure of the optimal policy, as expected, is of threshold-type described in Theorem 1. In Table 3 above we present the critical numbers $s_j(z_1)$, $z_1 = 0,\dots,Q_1$, that correspond to the optimal policy for each customer j, $j = 1,\dots,N$.

The computation time of the corresponding Matlab program required by the above algorithm is 12.84 seconds. The standard value-iteration algorithm does not converge in this example. This is due to the periodicity (with period N) of all states of the system under any stationary policy. This problem can be circumvented by a perturbation of the one-step transition probabilities so that a transition from a state to itself with non-zero probability is allowed. Specifically, we take the following new one-step probabilities

$$\tilde{p}_{(j,z_1,z_2)(j+1,w_1,w_2)}(a) = \tau p_{(j,z_1,z_2)(j+1,w_1,w_2)},\ \tilde{p}_{(j,z_1,z_2)(j,z_1,z_2)}(a) = 1-\tau,$$

where τ is a constant such that $0 < \tau < 1$. A reasonable choice for the value of τ is 0.5. The perturbed model has the same average-cost optimal policy as the original model (see [17, p. 209]). We implemented the value-iteration algorithm in the perturbed model and, as expected, we obtained the above optimal policy. The required computation time is 6.16 seconds which is considerably smaller than the computation time required by the policy-iteration algorithm.

4. The problem with stochastic continuous demands

We modify the problem that we introduced in Section 2 by assuming that the demand of customer $j \in \{1,\dots,N\}$ for product $i \in \{1,\dots,K\}$ is a continuous random variable ξ_i^j. The joint probability density function $g^j(\bar{\xi})$, $\bar{\xi} = (\xi_1,\dots,\xi_K) \in \widetilde{S} = [0,Q_1]\times\cdots\times[0,Q_K]$ of $\left(\xi_1^j,\dots,\xi_K^j\right)$ is known. A realistic interpretation of this model could be the situation in which the vehicle delivers K different kinds of petrol to a sequence of petrol stations. The vehicle has K compartments and each compartment is suitable for a particular kind of petrol. Note that Kyriakidis and Dimitrakos [8] investigated this problem when $K = 1$. Using the notation of Section 2, the optimal routing strategy in the modified problem can be determined by the following dynamic programming equations. For $j = 1,\dots,N-1$ we have:

$$f_j(\bar{z}) = \min\{H_j(\bar{z}), A_j\},\quad \bar{z} = (z_1,\dots,z_K) \in [0,Q_1]\times\cdots\times[0,Q_K],$$

where,

$$\begin{aligned} H_j(\bar{z}) = c_{j,j+1} &+ \int_{\bar{\xi}\in\Xi_{\bar{z}}} f_{j+1}(\bar{z}-\bar{\xi})g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\notin\Xi_{\bar{z}}} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi}, \end{aligned}$$

$$A_j = c_{j0} + c_{j+1,0} + \int_{\bar{\xi}\in\widetilde{S}} f_{j+1}(\overline{Q}-\bar{\xi})g^{j+1}(\bar{\xi})d\bar{\xi}.$$

In the boundary we have

$$f_N(\bar{z}) = c_{N0},\quad \bar{z} \in [0,Q_1]\times\cdots\times[0,Q_K]$$

and the minimum total expected cost is

$$f_0 = c_{10} + \int_{\bar{\xi}\in\widetilde{S}} f_1(\overline{Q}-\bar{\xi})g^1(\bar{\xi})d\bar{\xi}.$$

The structure of the finite-horizon optimal policy is the same as in the case of discrete demands since Lemma 1 and Theorem 1 are valid in this modified problem. The proof of Lemma 1 is identical to the corresponding proof of Section 2. The proof of Theorem 1 is similar to that of Section 2 with the following difference. Instead of showing that $H_j(\bar{z}) \geqslant H_j(\bar{z}+\bar{e}_i)$ in the modified model it suffices to show that $H_j(\bar{z}) \geqslant H_j(\bar{z}^*)$, where $\bar{z} = (z_1,\dots,z_i,\dots,z_K)$ and $\bar{z}^* = (z_1,\dots,z_i+y,\dots,z_K)$ with $y > 0$ such that $z_i + y \leqslant Q_i$. It can be seen that

$$\begin{aligned} H_j(\bar{z}) = c_{j,j+1} &+ \int_{\bar{\xi}\in\Xi_{\bar{z}}} f_{j+1}(\bar{z}-\bar{\xi})g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\notin\Xi_{\bar{z}}:\xi_i\leqslant z_i} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\in\Xi_{\bar{z}^*}:z_i<\xi_i\leqslant z_i+y} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\notin\Xi_{\bar{z}^*}:z_i<\xi_i\leqslant z_i+y} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\in\widetilde{S}:\xi_i>z_i+y} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi}, \end{aligned}$$

$$\begin{aligned} H_j(\bar{z}^*) = c_{j,j+1} &+ \int_{\bar{\xi}\in\Xi_{\bar{z}}} f_{j+1}(\bar{z}^*-\bar{\xi})g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\in\Xi_{\bar{z}^*}:z_i<\xi_i\leqslant z_i+y} f_{j+1}(\bar{z}^*-\bar{\xi})g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\notin\Xi_{\bar{z}^*}:\xi_i\leqslant z_i} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}^*-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\notin\Xi_{\bar{z}^*}:z_i<\xi_i\leqslant z_i+y} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}^*-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi} \\ &+ \int_{\bar{\xi}\in\widetilde{S}:\xi_i>z_i+y} [2c_{j+1,0} + f_{j+1}(\overline{Q} + (\bar{z}^*-\bar{\xi})^-)]g^{j+1}(\bar{\xi})d\bar{\xi} \end{aligned}$$

By comparing $H_j(\bar{z})$ and $H_j(\bar{z}^*)$ term by term as in the proof of Theorem 1 we deduce that $H_j(\bar{z}) \geqslant H_j(\bar{z}^*)$.

It is also possible to consider the corresponding infinite-horizon problems with stochastic continuous demands. The state space of the system becomes:

$$I = \{(i,\bar{z}) : i = 1,\dots,N, \bar{z} = (z_1,\dots,z_K) \in [0,Q_1]\times\cdots\times[0,Q_K]\}.$$

Using standard results of Markov decision theory (see Chapter 6 in [14]) it is possible to prove in the same way as in Section 3 that the infinite-horizon α- discounted cost optimal policy has the same threshold-type structure as the corresponding finite-horizon problem. It seems intuitively reasonable that the average-cost optimal policy has the same structure. However a rigorous proof seems to be difficult due to the fact that the state space I is continuous, since it is not easy to prove that the minimum expected α- discounted total cost is equicontinuous (see p.150 in [14]).

5. Conclusions

In this paper we investigated a simple variant of the classical capacitated vehicle routing problem. We assumed that a single vehicle starts its route from a depot and delivers K different products to N customers according to a particular order. The vehicle is divided into K sections. Each product is stored in its own section in the vehicle. The demands of the customers for each product were assumed to be discrete or continuous random variables with known distributions. A suitable dynamic programming algorithm was proposed for the determination of the policy that minimizes the total expected cost for the service of all customers. It was proved that the optimal policy divides the set of all possible loads carried by the vehicle into two disjoint subsets. If after servicing a customer the loads carried by the vehicle belong to the first subset then it is optimal to proceed to the next customer, while if these loads belong to the second subset it is optimal to return to depot for restocking and then go to the next customer.

We point out that the development of the dynamic programming algorithm relies heavily on the fact that only one vehicle is available and on the assumption that the vehicle visits the customers according to a predefined sequence.

We also studied a corresponding infinite-horizon problem in which the service of the customers is not completed when the last customer's demands are satisfied but it continues periodically with the same customer order. The times between the decision epochs are constant. The demands of the customers for each product are renewed in each cycle and follow the same distribution. We prove that the total discounted-cost optimal policy and the average-cost optimal policy have the same threshold-type structure as in the finite-horizon case.

A subject for future research could be the investigation of a more general problem in which the vehicle delivers new products to each customer and collects expired or useless products, and the demands for both kinds of products are random.

Acknowledgment

The authors thank the referees whose useful comments improved the content of the present work.

References

[1] B.M. Baker, M.A. Ayechew, A genetic algorithm for the vehicle routing problem, Computers and Operations Research 30 (2003) 787–800.
[2] R. Baldacci, E.A. Hadjiconstantinou, A. Mingozzi, An exact algorithm for the capacitated vehicle routing problem based on a two-commodity network flow formulation, Operations Research 52 (2004) 723–738.
[3] B. Bullnheimer, R.F. Hartl, C. Strauss, An improved ant system algorithm for the vehicle routing problem, Annals of Operations Research 89 (1999) 319–328.
[4] W. Chiang, R.A. Russell, Simulated annealing metaheuristics for the vehicle routing problem with time windows, Annals of Operations Research 63 (1996) 3–27.
[5] C.H. Christiansen, J. Lysgaard, A branch-and-price algorithm for the capacitated vehicle routing problem with stochastic demands, Operations Research Letters 35 (2007) 773–781.
[6] G. Dantzig, R. Ramster, The truck dispatching problem, Management Science 6 (1959) 80–91.
[7] R. Fukasawa, H. Longo, J. Lysgaard, M. Poggi de Aragao, M. Reis, E. Uchoa, R.F. Werneck, Robust branch-and-cut-and-price for the capacitated vehicle routing problem, Mathematical Programming A 106 (2006) 491–511.
[8] E.G. Kyriakidis, T.D. Dimitrakos, Single vehicle routing problem with a predefined customer sequence and stochastic continuous demands, Mathematical Scientist 33 (2008) 148–152.
[9] C.Y. Liong, I. Wan Rosmanira, O. Khairuddin, M. Zirour, Vehicle routing problem: Models and solutions, Journal of Quality Measurement and Analysis 4 (2008) 205–218.
[10] J. Lysgaard, A. Letchford, R.W. Eglese, A new branch-and-cut algorithm for the capacitated vehicle routing problem, Mathematical Programming 100 (2004) 423–445.
[11] J.E. Mendoza, B. Castanier, C. Gueret, A.L. Medaglia, N. Velasco, A memetic algorithm for the multi-compartment vehicle routing problem with stochastic demands, Computers and Operations Research 37 (2010) 1886–1898.
[12] U. Rembold, C. Blume, R. Dillmann, Computer Integrated Manufacturing Technology and Systems, Marcel Dekker, New York, 1985.
[13] Y. Rochat, R.E. Taillard, Probabilistic diversification and intensification in local search for vehicle routing, Journal of Heuristics 1 (1995) 147–167.
[14] S.M. Ross, Applied Probability Models with Optimization Applications, Dover, New York, 1992.
[15] D. Simchi-Levi, X. Chen, J. Bramel, The Logic of Logistics, second ed., Theory, Algorithms and Applications for Logistics and Supply Chain Management, Springer, New York, 2005.
[16] A. Tatarakis, I. Minis, Stochastic single vehicle routing with a predefined customer sequence and multiple depot returns, European Journal of Operational Research 197 (2009) 557–571.
[17] H.C. Tijms, Stochastic Models: An Algorithmic Approach, Wiley, New York, 1994.
[18] P. Toth, D. Vigo (Eds.), The Vehicle Routing Problem, Siam, Philadelphia, PA, 2002.
[19] P. Tsirimpas, A. Tatarakis, I. Minis, E.G. Kyriakidis, Single vehicle routing with a predefined customer sequence and multiple depot returns, European Journal of Operational Research 187 (2008) 483–495.
[20] W.-H. Yang, K. Mathur, R.H. Ballou, Stochastic vehicle routing problem with restocking, Transportation Science 34 (2000) 99–112.