Scheduling in a queuing system with impatience and setup costs

Alain Jean-Marie, Emmanuel Hyon

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Scheduling in a queuing system with impatience and setup costs

Alain Jean-Marie — Emmanuel Hyon

N° 6881 — version 2

initial version March 2009 — revised version February 2010

Thème COM

Rapport de recherche

ISSN 0249-6399   ISRN INRIA/RR--6881--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SOPHIA ANTIPOLIS - MÉDITERRANÉE

Scheduling in a queuing system with impatience and setup costs

Alain Jean-Marie*, Emmanuel Hyon†

Thème COM — Systèmes communicants
Équipe-Projet Maestro

Rapport de recherche n° 6881 — version 2 — initial version March 2009 — revised version February 2010 — 18 pages

Abstract: We consider a single server queue in discrete time, in which customers must be served before some limit sojourn time of geometrical distribution. A customer who is not served before this limit leaves the system: it is impatient. The fact of serving customers and the fact of losing them due to impatience induce costs. The purpose is to decide when to serve the customers so as to minimize costs. We use a Markov Decision Process with infinite horizon and discounted cost. We establish the structural properties of the stochastic dynamic programming operator, and we deduce that the optimal policy is of threshold type. In addition, thanks to a pathwise comparison analysis of two threshold policies, we are able to compute explicitly the optimal value of this threshold according to the parameters of problem.

Key-words: Scheduling, queuing system, impatience, deadline, optimal control, Markov decision processes

* INRIA and LIRMM, CNRS/Université Montpellier 2, 161 Rue Ada, F-34392 Montpellier, ajm@lirmm.fr.
† Université Paris Ouest Nanterre la Défense, LIP6, UPMC 4 place Jussieu, F-75252 Paris Cedex, Emmanuel.Hyon@u-paris10.fr.

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Politique de service optimal dans une file d'attente en temps discret avec impatiences

Résumé : Nous considérons un modèle d'une file d'attente à un serveur en temps discret, dans laquelle les clients doivent être servis avant une durée de séjour limite, de distribution géométrique. Un client qui n'est pas servi avant cette limite quitte le système: il est impatient. Le fait de servir les clients et le fait de perdre des clients par impatience induisent des coûts. Il s'agit de décider de façon optimale quand servir les clients. Nous utilisons un processus de décision Markovien à horizon infini et à coûts actualisés. Nous établissons les propriétés structurelles de l'opérateur de programmation dynamique stochastique, et nous déduisons que la politique optimale est à seuil. Par ailleurs, grâce à une analyse trajectorielle comparative de deux politiques à seuil, nous calculons explicitement la valeur optimale de ce seuil en fonction des paramètres du problème.

Mots-clés : Ordonnancement, file d'attente, impatience, échéance, contrôle optimal, processus de décision markovien

1 Introduction

In this paper we are interested in the optimal control of a queuing system with impatient customers (or, equivalently said, customers with deadlines). The set-up of customer services, the storage of the customers in the queue as well as their departure from the queue due to impatience (called "losses" in the remainder of this paper) induce some costs and it has to be decided when to begin the service in order to minimize these costs. This is a genuine tradeoff problem, since instantaneous costs associated with the decisions serve/not serve are not constant: there is no reason *a priori* why the decision should be the same whatever the state of the system.

Controlled queuing models, deterministic as well as stochastic, have been largely studied in the literature since their application fields are numerous: networking (see [1] and references therein), resources allocation (see [7] and references therein) to quote just a few. Nevertheless most of these works do not consider impatient customers but rather losses due to overflow. Yet, the phenomenon of impatience, associated with deadlines or "timeouts", has become non negligible in several fields: cellular communication networks [16, 3], call center [10], yield management or reservations problems (see [14] for discrete-time finite-horizon problems), real-time systems etc.

The literature features papers on the performance evaluation of queues with impatience, but relatively few on optimal control of such queues. One branch, represented by [4, 18], is concerned with finding the optimal scheduling algorithm so as to minimize deadline misses. Another direction is to consider the optimal routing between several queues [12, 8], still in order to minimize average deadline overrun. In the present paper, our focus is on the influence of set-up costs: we consider that serving customers has a cost, which adds up to the costs of missing deadlines and holding customers. On the other hand, the question of the order of service is not relevant to us, since we assume that deadlines have a memoryless distribution.

Our longer-term objective is to solve the same problem but with batch service. The problem of optimally controlling a batch server in a queue (without impatience) has been addressed in [5] and [13] (see also the references therein). Its resolution is based on the Markov Decision Process (MDP) formalism, and goes through establishing some structural properties of the value function and the dynamic programming operator. This then allows to deduce some properties of the optimal policy, which in turns implies that the solution is a threshold (or *control limit*) policy. Unhappily, it appears that extending the techniques developed in [13] to queues with impatience is not straightforward. Indeed, it has been noted in [9] (quoted in [8]) that impatience tends to destroy the structural properties that are commonly used for proving the optimality of threshold policies. The importance of structural properties in the study of admission control policy and optimality proofs has been underlined in [6]. In this paper, we show that structural properties exist although losses under given conditions on the value function and present the solution for service batches of unit size. For this purpose, we use some tools which, in our opinion, will be useful for solving more complex cases.

More precisely, in this work, we adapt the framework of structural analysis of Markov Decision Processes, as described for instance in [15]. We adopt the infinite-horizon, discounted cost criterion. We establish the structural properties of the stochastic dynamic programming operator and we show that the optimal policy is a threshold policy. Furthermore, using a sample path comparison between policies with different values of the threshold, we explicitly compute the threshold value as a function of the parameters, and we conclude that the optimal policy is actually "always serve" or "never serve", based on a simple criterion derived from the cost parameters. We discuss in our conclusion some problems encountered with general batch sizes.

This paper is organized as follows: Section 2 deals with the model, while Section 3 establishes the structural properties and 4 focuses on the effective computations of the threshold.

2 Model

We consider a discrete time (or slotted) model, where the slot is the time unit. Customers are assumed to arrive at the beginning of each slot.[1] They are stored in an infinite buffer in which they wait for to be admitted in the server to be processed. This admission decision is made by a controller. The beginning of a new service induces a cost (for example a setup cost). Holding customers in the buffer also induces a cost. The service duration is assumed to be equal to the duration of a slot.

Customers are impatient: while they are in the buffer, they can leave spontaneously the system with fixed probability α, independently from the past and from each other. On the other hand, customers admitted in service are not impatient anymore. Each time a customer leaves the buffer, this induces a cost.

2.1 System dynamics

We proceed with introducing some additional notation, and formulating the optimal control problem in the framework of Markov Decision Processes, using the notation of Puterman [15].

Assume that slots are numbered from 0 and denote with A_n the number of arrivals at the beginning of slot n. The sequence $\{A_n\}_{n\in\mathbb{N}}$ is assumed to be an i.i.d. sequence of random variables. With the usual abuse of notation, we denote generically this common distribution with A. We furthermore assume that A is of mean λ, so that the arrival process is of intensity λ. Examples include Poisson-distributed arrivals ($\mathbb{P}(A = k) = \frac{\lambda^k}{k!}e^{-\lambda}$), Bernoulli arrivals ($\mathbb{P}(A = 1) = \lambda$ and $\mathbb{P}(A = 0) = 1 - \lambda$) as well as geometric arrivals ($\mathbb{P}(A = k) = (1 - \mu)\mu^k$, $\mu = \lambda/(1 + \lambda)$). Our results do not depend on the specific distribution, however.

We call $x_n \in \mathbb{N}$ the number of waiting customers, counted *just after* the arrivals at slot n. The admission decision of the controller takes place just after arriving customers have been taken into account. The set of decisions, or action space, is denoted with $\mathcal{Q} = \{0, 1\}$, where $q_n \in \mathcal{Q}$ is the number of customers admitted into service at slot n: 0 if no customer is admitted and 1 otherwise. We assume that the controller may choose $q_n = 1$ even if $x_n = 0$, which has of course no effect. The number of customers remaining in the buffer just after the decision is then $y_n = (x_n - q_n)^+$, with $x^+ = \max(0, x)$.

During a slot, losses can occur because customers become impatient and leave. It is assumed that each customer has a constant probability $\alpha \in [0, 1]$ of leaving in each slot, independently from the past and from other customers. This is equivalent to assuming that the patience of each customer is geometrically distributed on $\mathbb{N}$ with parameter α. [2] For notational convenience, we introduce the stochastic operators $I(y)$ and $S(y)$ which count, respectively, the number of customers lost (impatient) and remaining (survivors), out of y present at the beginning of a slot. Conditioned on the value of $y_n = y$, the number of losses during one slot, $I(y)$, is a Binomial random variable with mean αy; likewise, $S(y)$ is a Binomial random variable with mean $\overline{\alpha}y$, where $\overline{\alpha} = 1 - \alpha$. This is a consequence of the independence and memoryless assumptions made on the impatience. Observe that for a given slot, $I(y)$ and $S(y)$ are correlated since $I(y) + S(y) = y$. This will not cause problems in the analysis.

With this notation, we can state the central representation of the dynamics of the system. The evolution of the state from slot n to slot $n + 1$ is given by the recurrence equation:

$$x_{n+1} \;=\; R(x_n, q_n) \;:=\; S\left((x_n - q_n)^+\right) + A_{n+1}\,, \tag{1}$$

[1] As usual in this type of models, customers arriving during a slot are supposed to be notified to the system only at the beginning of the next slot.

[2] This is the equivalent for discrete time to the exponential distribution of deadlines in continuous time models.

whereas the number of customers lost in slot n is equal to $I((x_n - q_n)^+)$.

Stochastic Monotonicity Properties. We conclude by recalling useful properties of the stochastic operators we have introduced. From basic properties of the binomial distribution, it is possible to write the equality between distributions

$$S(x+1) \ =_d \ S(x) \ + \ B(1) \ , \tag{2}$$

with $B(1)$ a Bernoulli random variable of parameter $\overline{\alpha}$. The random variables in the right-hand side being independent. For a possible generalization based on this property, one can see the conclusion.

Recall the following notion of stochastic ordering of random variables, and the property of the Binomial family.

Definition 1 (Stochastic order [17]). *We say that a random variable X is stochastically greater than a random variable Y (denoted by $X \geq_{st} Y$) if $\mathbb{P}(X \leq t) \leq \mathbb{P}(Y \leq t)$, or, equivalently, if $\mathbb{E}w(X) \geq \mathbb{E}w(Y)$ for every nondecreasing function w.*

Proposition 2. *For any $x \geq y \in \mathbb{N}$ we have $S(x) \geq_{st} S(y)$. If $X \geq_{st} Y$, then $S(X) \geq_{st} S(Y)$.*

2.2 Elements of the Markov Decision Process

2.2.1 Transition probabilities

The dynamics of the controlled process are characterized by the probabilities to move in state z, given that the state is y and the action is $q \in \mathcal{Q}$: $\mathbb{P}\left(z|(y,q)\right)$. Formally

$$\mathbb{P}\left(z|(y,q)\right) \ = \ \mathbb{P}\left(x_{n+1} = z | x_n = y, q_n = q\right) \ . \tag{3}$$

These probabilities do not depend on n: the transition probabilities are homogeneous in time. Given the recurrence (1), and conditioning on the number of survivors, this can be expressed as:

$$\mathbb{P}\left(z|(y,q)\right) \ = \ \sum_{i=0}^{\min(z,(y-q)^+)} \mathbb{P}\big(S\left((y-q)^+\right) = i\big)\mathbb{P}\big(A = z - i\big).$$

2.2.2 Rewards/Costs

The costs associated with decisions and transitions are the following. First, there is a setup cost which is incurred at the decision epoch when the controller chooses to admit one customer into service. If the choice is to keep the customer in the queue no cost is incurred. We denote by c_B this setup cost. It is assumed that $c_B > 0$. Second, there is a cost associated to each customer leaving the queue due to impatience. This cost, at slot n, is $c_L I(y_n)$, where c_L is the cost of a single loss. Finally, there is an holding cost c_H per remaining customer. We assume that it applies to all customers present after the service admission decision, so that the cost for slot n is $c_H y_n$.

In order to fit in the framework of [15], we have to express the immediate cost (or negative reward) at each slot as a deterministic function of the current state and the current decision. Conditioned on the number of customers present $x_n = x$ and given that the decision is $q_n = q$, the *average* cost due to losses (see Section 2.1) is $c_L\mathbb{E}(I(x-q)^+) = c_L\alpha(x-q)^+$.

It is therefore indeed possible to express the cost incurred by taking decision q when the state is x, as the function of (x, q):

$$c(x,q) \;=\; q\,c_B \;+\; (c_L\,\alpha + c_H)\,(x-q)^+ \;=\; q\,c_B \;+ c_C\,(x-q)^+\;, \tag{4}$$

where $c_C = \alpha c_L + c_H$ is the *per-capita* cost for customers. Observe that this cost function is *not* bounded, unless $c_C = 0$.

2.2.3 Decision policies

We call policy a sequence of decision rules $\pi = (d_0, d_1, \ldots)$, each decision rule mapping some information set to some action. While the most general set of policies is that of history-dependent randomized policies, Π^{HR}, the classical results on discounted, infinite-horizon, time-homogeneous Markovian optimal control allow us to concentrate on *Markov Deterministic Policies*. Such policies are characterized by a single, deterministic decision rule which maps the current state to a decision. We denote with Π the set of such policies (denoted as D^{MD} in [15]).

We consider a discounted cost criterion and the discount factor is denoted by θ. We make this choice in order to avoid the complexities associated with the average cost criterion. Nevertheless a brief explanation for the case $\theta \to 1$ can be seen in conclusion. Under each policy π, the evolution of the system generates a random sequence of states x_n and decisions q_n. We define the value function of policy π by the total expected discounted cost:

$$v_\theta^\pi(x) \;=\; \mathbb{E}_x^\pi\left[\sum_{n=0}^{\infty} \theta^n\, c(x_n, q_n)\right]\;,$$

where $x_0 = x$. Our aim is to find the optimal policy $\pi^* \in \Pi^{HR}$ such that

$$\forall x \in \mathbb{N}\,, \quad v_\theta^{\pi^*}(x) = v_\theta^*(x) = \min_{\pi \in \Pi^{HR}} v_\theta^\pi(x).$$

The following operators T, L and L_d, acting on functions v, will be useful in the analysis:

$$(Tv)(x,q) \;=\; c(x,q) \;+\; \theta \sum_{y \in \mathbb{N}} \mathbb{P}\left(y|(x,q)\right) v(y) \;=\; c(x,q) \;+\; \theta\, \mathbb{E}\left[v\left(R(x,q)\right)\right]\;, \tag{5}$$

$$(L_d v)(x) \;=\; (Tv)(x, d(x))\;.$$

and

$$(Lv)(x) \;=\; \min_{d \in \Pi} (L_d v)(x)\;.$$

With this notation, the dynamic programming equation is $v_\theta = Lv_\theta$ and finding the solution to this fixed point problem solves the dynamic programming problem.

3 Structural properties of the optimal policy

In this part we study the structural properties of value functions in order to get qualitative results on the optimal policy. Specifically, we prove that the optimal policy is of threshold type.

The framework is that of property propagation through the Dynamic Programming operator. It consists in three steps: first exhibit some structural properties of the operator Tv under special conditions on v (this implies some qualitative results on π^*). Then show that the properties of v are kept by the operator Tv. At last, check that these properties are kept when passing to the limit. A *structure theorem* then allows to ensure that there exists an optimal policy and states, at the same time, that this optimal policy can be chosen in the set of structured policies.

For easier reference, we first recall the methodological framework and the results we will need from the literature. Next, we prove that the dynamic programming operator for our problem propagates monotonicity and convexity. Finally, we deduce the desired property for the optimal policy.

3.1 Structured policies

We say that a policy is a *structured* policy if it has a special form (for example increasing, decreasing...). Together with the notion of structured policies, comes a notion of *structured* value functions. Both notions are adapted to each other. Let V^σ be the set of structured value functions and $\mathcal{D}^\sigma$ the set of structured decision rules. A structured policy Π^σ is a sequence of structured decision rules.

The following theorem indicates which properties have to be conserved (or propagated) by the dynamic programming operator. Two sets of properties are required in this theorem: properties (6)–(8) are related with the existence of solutions under unboundedness of the cost function; conditions $0)$ to $iii)$ are the structural conditions *per se*.

Theorem 3 ([15], Theorem 6.11.3)**.** *Assume that the following properties hold: there exists a positive function on the state space, w, such that:*

$$\sup_{(x,q)} \frac{|c(x,q)|}{w(x)} \; < \; +\infty \; , \tag{6}$$

$$\sup_{(x,q)} \frac{1}{w(x)} \sum_y \mathbb{P}(y|x,q) w(y) \; < \; +\infty \; , \tag{7}$$

and for every μ, $0 \le \mu < 1$, there exists η, $0 \le \eta < 1$ and some integer J, such that, for every J-uple of Markov Deterministic decision rules $\pi = (d_1, \ldots, d_J)$, and every x,

$$\mu^J \sum_y P_\pi(y|x) w(y) \; \le \; \eta w(x) \; , \tag{8}$$

where P_π denotes the J-step transition matrix under policy π.

Let V_w be the set of functions on the state space which have a finite w-weighted supremum norm (i.e. $\sup_x |v(x)/w(x)| < +\infty$) and $V^\sigma \subset V_w$. Assume that:

0. for each $v \in V_w$, there exists a deterministic Markov decision rule d such that $Lv = L_d v$.

If, furthermore,

i. $v \in V^\sigma$ implies $Tv \in V^\sigma$,

ii. $v \in V^\sigma$ implies there exists a decision d such that $d \in \mathcal{D}^\sigma \cap \arg\min_d L_d v$,

iii. V^σ is a closed subset of the set of value functions by simple convergence.

Then, there exists an optimal stationary policy $(d^*)^\infty$ in Π^σ with $d^* \in \arg\min_d L_d v$.

We introduce now the key submodularity property and its consequences.

Definition 4 (Submodularity [6]). *A real-valued function g defined on two partially ordered sets $\mathcal{X} \times \mathcal{Q}$ is called submodular (or subadditive) if it has monotone decreasing differences. That is, if it verifies the inequality*

$$g(x^+, q^+) - g(x^-, q^+) \le g(x^+, q^-) - g(x^-, q^-),$$

for any $x^+ \ge x^- \in \mathcal{X}$ and any $q^+ \ge q^- \in \mathcal{Q}$.

Proposition 5 ([15], Lemma 4.7.6). *Let g be a real valued function defined on $\mathbb{N} \times \{0, 1\}$. If the difference function $x \mapsto g(x, 1) - g(x, 0)$ is nonincreasing, then the function g is submodular.*

These properties allow to deduce a structural property of the optimal control

Proposition 6 (Monotone optimal control [15], Lemma 4.7.1). *Let $g(x, q)$ be a submodular function on two partially ordered sets such that, $\min_x g(x, q)$ exists for any x. If $g(x, q)$ is submodular then the control function*

$$q(x) = \min\{\arg\min_{q \in \mathcal{Q}} (g(x, q))\} ,$$

is increasing i.e. $x^+ \ge x^-$ yields $q(x^+) \ge q(x^-)$.

3.2 Structural properties of the dynamic programming operator

In this part we establish structural results of the dynamic programming operator for our system: propagation of monotonicity, submodularity and convexity.

Lemma 7 (Propagation of monotonicity). *Let $\tilde{v}$ be the function defined by $\tilde{v}(x) = \min_q Tv(x, q)$ for any $x \in \mathbb{N}$. Then $\tilde{v}$ is nondecreasing in x if v is nondecreasing in x.*

Proof. The definition of Tv in (5) involves two terms given in Eqs. (4) and (1). We show first that the lump costs $c(x, q)$ are nondecreasing for a given decision q. Indeed, from Equation (4) the cost is either equal to $(x - 1)c_C + c_B$ or xc_C which are nondecreasing in x.

From Proposition 2, it follows that $S\left((x + 1 - q)^+\right) \ge_{st} S\left((x - q)^+\right)$. Therefore we have that $R(x + 1, q) \ge_{st} R(x, q)$. Recall (Definition 1) that if $X \ge_{st} Y$, then $\mathbb{E}w(X) \ge \mathbb{E}w(Y)$ for every nondecreasing function w. As a consequence, the function $Tv(x, q)$ is the sum of two increasing functions of x for every q. The minimum over q is therefore also increasing. □

Lemma 8 (Submodularity). *For any nondecreasing convex function v, the function $Tv(x, q)$ is submodular on $\mathbb{N} \times \mathcal{Q}$.*

Proof. We have the decomposition:

$$\Delta_q Tv(x) \stackrel{\triangle}{=} Tv(x, 1) - Tv(x, 0) \ = \ c(x, 1) - c(x, 0) + \theta \Delta_q \hat{T}v(x) \ , \tag{9}$$

where:

$$\begin{aligned} \Delta_q \hat{T}v(x) \ &= \ \sum_{y \in \mathbb{N}} \mathbb{P}\big(y|(x, 1)\big)v(y) \ - \ \sum_{y \in \mathbb{N}} \mathbb{P}\big(y|(x, 0)\big)v(y) \\ &= \ \mathbb{E}v(S((x - 1)^+) + A) \ - \ \mathbb{E}v(S(x) + A) \ . \end{aligned} \tag{10}$$

First of all, observe that the difference $c(x, 1) - c(x, 0) = c_B - c_C$ does not depend on $x \ge 1$. On the other hand, for $x = 0$, we simply have: $\Delta_q \hat{T}v(0) = 0$ and $c(0, 1) - c(0, 0) = c_B$ thus $\Delta_q Tv(0) = c_B$.

We then prove the nonincreasingness of $x \mapsto \Delta_q \hat{T}v(x)$ for any $x > 0$. In that case, we use the stochastic decomposition (2), written as $S(x) = S(1) + S(x-1)$, in (10) to get:

$$\begin{aligned} \Delta_q \hat{T}v(x) &= \mathbb{E}v(S(x-1)+A) - \mathbb{E}v(S(x-1)+S(1)+A) \\ &= \sum_{a,s} \mathbb{P}(A=a, S(1)=s) \left(\mathbb{E}v(S(x-1)+a) - \mathbb{E}v(S(x-1)+s+a)\right) \\ &= -\sum_{a,s} \mathbb{P}(A=a, S(1)=s)\, \mathbb{E}\left[u_{a,s}(S(x-1))\right] , \end{aligned} \tag{11}$$

where we have defined: $u_{a,s}(y) \stackrel{\triangle}{=} v(y+s+a) - v(y+a)$. Since v is increasing and convex, the function $u_{a,s}(y)$ is nonnegative and increasing for all nonnegative values of a and s. The stochastic increasingness of the $S(x)$ (Proposition 2), implies (see Definition 1) that $\mathbb{E}u_{a,s}(S(x)) \geq \mathbb{E}u_{a,s}(S(x-1))$, for all $x \geq 1$ and all s, $a \geq 0$. This last inequality is conserved by convex combinations. As a result, the expression (11) is a nonincreasing function of $x > 0$. It is also negative, so that when $x = 1$:

$$\Delta_q Tv(1) = c_B - c_C + \theta \Delta_q \hat{T}v(1) \leq c_B = \Delta_q Tv(0) .$$

The function is therefore nonincreasing at $x = 0$ as well.

We conclude that $\Delta_q Tv(x) = Tv(x,1) - Tv(x,0)$ is indeed nonincreasing in x and then, by Proposition 5, $Tv(x,q)$ is submodular. □

Lemma 9 (Propagation of convexity)**.** *Let $\tilde{v}$ be the function defined by $\tilde{v}(x) = \min_q Tv(x,q)$ for any $x \in \mathbb{N}$. Then $\tilde{v}$ is nondecreasing convex in x if v is nondecreasing convex in x.*

Proof. The nondecreasingness comes from Lemma 7. We just have to prove the convexity.

Assume then that v is nondecreasing convex. Define $q^*_y \stackrel{\triangle}{=} \arg\min_q Tv(x,q)$, and let $\Delta_x \tilde{v}(y)$ be defined as $\Delta_x \tilde{v}(y) \stackrel{\triangle}{=} \tilde{v}(y+1) - \tilde{v}(y) = Tv(y+1, q^*_{y+1}) - Tv(y, q^*_y)$. We shall prove that the function $\Delta_x \tilde{v}(y)$ is nondecreasing, which is equivalent to the convexity of $\tilde{v}$.

As proved in Lemma 8, the function $Tv(y,q)$ is submodular. As a consequence of Proposition 6, the function $y \mapsto q^*_y$ is therefore nondecreasing. Henceforth, the couple (q^*_{y+1}, q^*_y) can take only one of the three values $(0,0)$, $(1,0)$ and $(1,1)$. These cases correspond to, respectively, $y \leq \bar{y} - 2$, $y = \bar{y} - 1$ and $y \geq \bar{y}$, where $\bar{y}$ is the integer number, potentially infinite, such that $\Delta_q Tv(\bar{y} - 1) > 0$ and $\Delta_q Tv(\bar{y}) \leq 0$. Observe that $\bar{y} \geq 1$ because $\Delta_q Tv(0) = c_B > 0$.

The case $y = \bar{y} - 1$ is the simplest: in that case $\Delta_x \tilde{v}(\bar{y} - 1) = c_B$. For $y \leq \bar{y} - 2$, since $q_{y+1} = q_y = 0$, we have:

$$\Delta_x \tilde{v}(y) = c_C + \theta \left[\mathbb{E}v(S(y+1)+A) - \mathbb{E}v(S(y)+A)\right] . \tag{12}$$

If on the other hand $y \geq \bar{y}$, since $q_{y+1} = q_y = 1$, we have:

$$\Delta_x \tilde{v}(y) = c_C + \theta \left[\mathbb{E}v(S(y)+A) - \mathbb{E}v(S(y-1)+A)\right] . \tag{13}$$

The same stochastic increasingness arguments as in the proof of Lemma 8 allow to conclude that $\Delta_x \tilde{v}(y)$ is increasing in both cases, that is, for $y \leq \bar{y} - 2$ and $y \geq \bar{y}$. There remains to show that $\Delta_x \tilde{v}(\bar{y} - 2) \leq \Delta_x \tilde{v}(\bar{y} - 1) (= c_B) \leq \Delta_x \tilde{v}(\bar{y})$. Using (9), (10) and (12), we conclude that:

$$c_B - \Delta_x \tilde{v}(\bar{y} - 2) = \Delta_q Tv(\bar{y} - 1) .$$

Likewise

$$\Delta_x \tilde{v}(\bar{y}) - c_B = -\Delta_q Tv(\bar{y}) .$$

By definition of $\bar{y}$, both these quantities are nonnegative. Therefore $\Delta_x \tilde{v}(y)$ is nondecreasing and thus the convexity propagates. □

3.3 Structural properties of the optimal policy

We present now deterministic policies which are called either threshold policies in [1] or control limit policies in [5].

One calls threshold policy a Markovian deterministic policy such that

$$q(x) = \begin{cases} q_1 & \text{if } x < \nu \\ q_2 & \text{if } x \geq \nu \end{cases}$$

where q_1 and q_2 are in $\mathcal{Q}$ and ν is called the threshold. In other words, if the system is in a state under the threshold then it is optimal to perform q_1 and q_2 is optimal elsewhere.

Theorem 10. *The optimal policy is increasing in x (it is a monotone control) and is a threshold policy.*

Proof. We apply Theorem 3 with V^σ the set of nondecreasing convex functions, and $\mathcal{D}^\sigma$ the set of monotone controls.

Let us first check that the preliminary conditions (6)–(8) are satisfied. We choose as weighing function: $w(x) = C + c_C x$, for some constant $C > 0$ to be determined in the course of the proof. Such a function satisfies (6) because the ratio in the right-hand side of (6) is an homographic and bounded function.

Firstly, we have, using the notation introduced earlier,

$$\begin{aligned} \sum_y \mathbb{P}(y|(x,q))w(y) &= \mathbb{E}(w(A + S((x-q)^+))) \;=\; C \;+\; c_C\, \mathbb{E}(A + S((x-q)^+)) \\ &= C \;+\; c_C\,(\lambda + \overline{\alpha}((x-q)^+)) \;\leq\; C \;+\; c_C\lambda + \overline{\alpha}(c_C x)\,. \end{aligned}$$

It follows that $\sup_q \sum_y \mathbb{P}(y|(x,q))w(y)/w(x)$ is also bounded by an homographic and bounded function of x.

Finally, consider any sequence of Markov decision rules d_j. Since w is increasing, we always have: $\mathbb{E}(w(A + S((x - d_j(x))^+))) \leq \mathbb{E}(w(A + S(x)))$. Let us call $\widetilde{R}(x) = R(x, 0)$ and denote with $\widetilde{R}^{(n)}$ the n-th iteration of this operator. It follows that

$$\begin{aligned} \sum_y \mathbb{P}_\pi(y|x,q)w(y) &\leq \mathbb{E}(w(\widetilde{R}^{(J)}(x))) \;=\; C \;+\; c_C\,\mathbb{E}(\widetilde{R}^{(J)}(x)) \\ &= C \;+\; c_C\,(\lambda + \bar{\alpha}\mathbb{E}(\widetilde{R}^{(J-1)}(x))) \;=\; C \;+\; c_C\lambda\frac{1-\overline{\alpha}^J}{1-\overline{\alpha}} \;+\; c_C\overline{\alpha}^J x\,. \end{aligned}$$

Assume first that $\alpha > 0$. If one chooses the constant $C = \lambda c_C/\alpha$, then:

$$\sum_y \mathbb{P}_\pi(y|x,q)w(y) \;\leq\; C \;+\; c_C\lambda\frac{1}{\alpha} \;+\; \overline{\alpha}^J(c_C x) \;=\; 2C \;+\; \overline{\alpha}^J(c_C x)\,.$$

Next, for any $\mu < 1$, there exists some J such that $2\mu^J$ is less than 1. For this J and every sequence of J decision rules, and taking $\eta \in (2\mu^J, 1)$

$$\mu^J \sum_y \mathbb{P}_\pi(y|x,q)w(y) \;\leq\; 2\mu^J C \;+\; \mu^J\overline{\alpha}^J(c_C x) \;\leq\; \eta w(x)\,.$$

Consider now the case $\alpha = 0$. Then

$$\sum_y \mathbb{P}_\pi(y|x,q)w(y) \;\leq\; C \;+\; c_C\lambda J \;+\; c_C x\,.$$

Whatever the constant C, for any $\mu < 1$, there exists an integer J such that $\mu^J(C + \lambda c_C J)/C < 1$. It is therefore possible to find a constant $\eta < 1$ as above. Finally, Property (8) holds in all cases.

In order to apply Theorem 3, it is next required that condition 0) hold. Here, Theorem 6.2.10 of [15] implies that this is the case, as a consequence of the finiteness of the action space.

We proceed to check that the three structural conditions of Theorem 3 are satisfied. Lemma 9 insures that the value function issued from operator T is still nondecreasing convex if the input value function is nondecreasing convex. Thus *i)* is checked. Moreover, Lemmas 7 and 8 combined with Proposition 6 prove that the optimal control is nondecreasing as soon as the value function is nondecreasing convex. This shows *ii)*. At last, the point-wise convergence of a sequence of nondecreasing convex functions is convex, therefore *iii)* holds.

Therefore, there exists an optimal policy which is a monotone control. Given that the action space has two elements, this is actually a threshold policy. $\square$

As a corollary, we also obtain that the value function is nondecreasing and convex.

4 The optimal threshold

We are now interested by computing the threshold values accordingly the parameters values. This means determining the customer load of the system from which it becomes more efficient to serve customers than to do nothing. An infinite threshold means that it is never optimal to accept customers. The approach used here is to compute an expression for the value function of a general threshold policy, then use a sample path analysis (see *e.g.* [11]) in order to establish the monotonicity of this value function.

Theorem 11. *Let ψ be the number defined by*

$$\psi \;=\; c_B \;-\; \frac{c_C}{1-\overline{\alpha}\theta} \;.$$

Then,

i. If $\psi > 0$, the optimal threshold is $\nu = +\infty$.

ii. If $\psi < 0$, the optimal threshold is $\nu = 1$.

iii. If $\psi = 0$, any threshold $\nu \geq 1$ gives the same value.

The proof needs a preliminary Lemma. Define first the operator R_ν by

$$R_\nu(x) \;=\; A \;+\; S\big((x - 1_{x\geq\nu})^+\big) \;, \tag{14}$$

and denote with $R_\nu^{(n)}$ the n-th iteration of the operator R_ν.

Lemma 12. *Let $\Phi_\nu(x,\theta)$ be the function defined by*

$$\Phi_\nu(x,\theta) = \sum_{n=0}^{\infty} \theta^n \mathbb{P}(R_\nu^{(n)}(x) \geq \nu) \;.$$

Then the function $\Phi_\nu(x,\theta)$ is positive, increasing with respect to x for every fixed ν and decreasing with respect to ν for fixed x.

The Proof of Lemma 12 is postponed in appendix.

Proof of Theorem 11. When the policy is a threshold policy with threshold ν, the evolution of the system is: $x_n = R_\nu(x_{n-1}) = R_\nu^{(n)}(x_0)$. Then, according to Eq. (14), we have:

$$\mathbb{E}R_\nu(x) \;=\; \mathbb{E}A \;+\; \overline{\alpha}\mathbb{E}\left([x-1_{x\geq\nu}]^+\right) \;=\; \mathbb{E}A \;+\; \overline{\alpha}\big(\mathbb{E}X - \mathbb{P}(X\geq\nu)\big)\;, \tag{15}$$

since $x - 1_{x\geq\nu}$ is always nonnegative, for every $x \geq 0$ and $\nu \geq 1$.

Next, if the state is x, the cost incurred by the control policy becomes:

$$c(x) \;=\; c_B\, 1_{x\geq\nu} \;+\; c_L\, I\big(x-1_{x\geq\nu}\big) \;+\; c_H\,\left(x-1_{x\geq\nu}\right)\,,$$

since the acceptance of a customer occurs only when the queue is not empty, implying that the threshold is strictly greater than 0. Hence, if X is a random variable, we have the expectation:

$$\begin{aligned}\mathbb{E}c(X) \;&=\; c_B\,\mathbb{P}(X\geq\nu) \;+\; \alpha c_L\,\mathbb{E}(X-1_{X\geq\nu}) \;+\; c_H\,\mathbb{E}(X-1_{X\geq\nu})\\ &=\; c_B\,\mathbb{P}(X\geq\nu) \;+\; c_C\,\mathbb{E}(X-1_{X\geq\nu})\\ &=\; c_B\,\mathbb{P}(X\geq\nu) \;+\; c_C\,\left(\mathbb{E}(X)-\mathbb{P}(X\geq\nu)\right)\\ &=\; (c_B-c_C)\,\mathbb{P}(X\geq\nu) \;+\; c_C\,\mathbb{E}(X)\,.\end{aligned}$$

Using these properties, the value function, when the threshold policy is with threshold ν, at the initial state x can be computed as follows:

$$\begin{aligned}V_\nu(x) &= \mathbb{E}_x\left(\sum_{n=0}^{\infty}\theta^n c(x_n)\right) = \sum_{n=0}^{\infty}\theta^n\mathbb{E}_x\big(c(x_n)\big)\\ &= \sum_{n=0}^{\infty}\theta^n\left((c_B-c_C)\,\mathbb{P}(R_\nu^{(n)}(x)\geq\nu) \;+\; c_C\,\mathbb{E}(R_\nu^{(n)}(x))\right)\,.\end{aligned}$$

With, using Equation (15),

$$\begin{aligned}\sum_{n=0}^{\infty}\theta^n\mathbb{E}(R_\nu^{(n)}(x)) &= x \;+\; \sum_{n=1}^{\infty}\theta^n\mathbb{E}(R_\nu^{(n)}(x))\\ &= x \;+\; \sum_{n=1}^{\infty}\theta^n\left(\mathbb{E}(A)+\overline{\alpha}\mathbb{E}(R_\nu^{(n-1)}(x))-\overline{\alpha}\mathbb{P}(R_\nu^{(n-1)}(x)\geq\nu)\right)\\ &= x \;+\; \frac{\mathbb{E}(A)\theta}{1-\theta} + \theta\overline{\alpha}\sum_{n=0}^{\infty}\theta^n\mathbb{E}(R_\nu^{(n)}(x)) - \theta\overline{\alpha}\Phi_\nu(x,\theta)\,.\end{aligned}$$

Consequently, we get

$$\begin{aligned}\sum_{n=0}^{\infty}\theta^n\mathbb{E}(R_\nu^{(n)}(x)) &= \frac{1}{1-\theta\overline{\alpha}}\left(x \;+\; \frac{\theta\mathbb{E}(A)}{1-\theta} \;-\; \overline{\alpha}\theta\Phi_\nu(x,\theta)\right)\,,\\ V_\nu(x) &= \frac{c_C}{1-\theta\overline{\alpha}}\left(x \;+\; \frac{\theta\mathbb{E}(A)}{1-\theta}\right) \;+\; \left(c_B \;-\; \frac{c_C}{1-\overline{\alpha}\theta}\right)\Phi_\nu(x,\theta)\,.\end{aligned}$$

The dependence on ν is concentrated in the function Φ_ν, but, by Lemma 12, we know that $\Phi_\nu(x,\theta)$ is positive, increasing with respect to x for every fixed ν and decreasing with respect to ν for fixed x. Since we want for all x, $\arg\min_\nu V_\nu(x)$, if $\psi > 0$ then the optimal value of $\nu = \infty$ while if $\psi < 0$ the optimal value is $\nu = 1$. If $\psi = 0$, V_ν does not depend on ν and all values give the same result. $\square$

5 Discussion and Extensions

5.1 Interpretations and Comparative Statics

An interpretation of Theorem 11 is as follows. Given the choice between serving *immediately or never* the first customer in line, the controller faces an immediate cost c_B for service, and a future, discounted cost for letting the customer remain in the system until it gets impatient. This cost turns out to be $c_C/(1-\overline{\alpha}\theta)$. Assume that the controller chooses not to serve this customer. It has then no reason for serving any other customer because they are all symmetric, given the memoryless deadline distribution. At the next slot, for every customer present the same dilemma would occur between two *average* costs, still because of the Markovian nature of the impatience process. Instead, in a discounted case the cost of postponing an decision has to be taken into account. The monotonicity result of Lemma 8 show this cost has either a negligible long term cost ($\nu = +\infty$) or not ($\nu = 1$). Therefore, the optimal decision can be made on a customer-per-customer basis. This leads to a policy where customers are served as soon as they arrive ($\nu = 1$) or never ($\nu = +\infty$).

According to this interpretation, it is not too surprising that the optimal policy does not depend on λ, although the value function does. It may be for instance that the rate of arrivals is larger than the service capacity. If the impatience process is active ($\alpha > 0$), the queue always remain stable, and average holding costs remain bounded over time. If $\alpha = 0$, the queue may build up, and holding costs increase over time. However, the presence of the discount factor makes the value function finite.

Next, the value of ψ defined in Theorem 11 is: increasing with respect to c_B, decreasing with respect to c_L and c_H; the optimal threshold ν^* varies in the same direction as ψ (from 1 to $+\infty$), which is conform to intuition. It can be seen also that ψ and ν^* are decreasing with respect to θ. Finally, ν^* is increasing with respect to α if $c_L \leq \theta c_H/(1-\theta)$, and decreasing in the converse case.

The extreme cases for α also have a reasonable explanation. When $\alpha = 1$, every customer not served in one slot leaves the system before the next slot. It is therefore optimal to serve one customer among the recently arrived ones (corresponding to a threshold of one) if the setup cost c_B is smaller than the per-customer cost $c_C = c_L + c_H$. This is consistent with Theorem 11 since $\psi = c_B - (c_L + c_H)$ in this case. At the opposite, when $\alpha = 0$, customers never leave spontaneously the system. There is no loss cost incurred and $\psi = c_B > c_H/(1-\theta)$ in Theorem 11. If the holding cost is zero, the threshold is therefore infinite: it is optimal never to serve any customer. If we had allowed c_B to be negative, the threshold would be equal to 1.

5.2 Methodological issues

It could be objected to the present paper that since the optimal policy is very simple (serve or no serve), there is probably a simpler way to prove it than to use the machinery of property propagation. This feeling is reinforced by the interpretations of the previous paragraph. We argue in this section that this criticism is true only to some extent: in our opinion, the direct proofs may not be substantially more compact, and they have less potential for generalization.

Two alternate proof techniques are candidate: direct verification and coupling.

Direct verification. It is indeed possible to prove Theorem 11 by computing the value function for the optimal policy (always serve if $\psi < 0$, never serve if $\psi > 0$) and check that they solve Bellman's equation $Lv = v$. Observe first that this approach still necessitates a proof that this equation has a unique solution: one still needs to check that conditions (6)–(8) hold, as in the proof of Theorem 10.

Computing the candidate value function is what we have done in Section 4. Proving that it solves Bellman's equation turns out to be indeed simple in the case $\psi > 0$ with the function

$$V_\infty(x) \;=\; \frac{c_C}{1-\theta\overline{\alpha}}\left(x \;+\; \frac{\theta\mathbb{E}(A)}{1-\theta}\right)\;.$$

For the case $\psi(x) < 0$, the value function has no such simple closed form, and proving that it solves Bellman's equation requires establishing additional monotonicity properties similar to Lemma 15 as well as convexity properties. In the end, the proof does not appear to be substantially shorter.

Coupling arguments. The simplicity of the decision rule suggests that simple comparison arguments should be sufficient. For instance, if the optimal policy is "always serve", show that not serving a customer is sub-optimal as compared to serving it. Such reasoning is usually made formal through coupling arguments: both trajectories are constructed on the same probability space.

The most natural coupling is perhaps the one we use in Appendix 7. With this coupling, it is not possible to show that serving one customer systematically results in a gain or a loss. This can be seen through a counterexample 13 and also through the fact that the criterion on ψ involves *average* costs, and not only instantaneous costs. Another reason is that we use a discounted cost, for which the cost of some action depends on the time at which it is taken. Interchange arguments are usually used in the context of average or total costs.

Example 13. *We assume that the parameters are such that $c_B < \frac{c_C}{1-\bar{\alpha}\theta}$ (implying the optimality of the "always serve" policy) and such that $c_B > c_H + c_L$. These assumptions are satisfied for a wide range of parameters since $\bar{\alpha}$ and θ are smaller than one and since that $c_H + c_L = c_C + \bar{\alpha}c_L$. We however show on a simple sample of a path that this is not better to serve the customer than to let it in the system.*

Assume there is a set of x customers waiting. By coupling arguments let us assume that there is only one customer expected to leave the system at the next step and that this dedicated customer is this one for which a decision has to be taken. Using Eq. (5, the costs induced by the choice to admit the customer is $c_B + (x-1)c_H + \theta\mathbb{E}(v(x-1))$ while the costs induced by the other choice is $c_L + xc_H + \theta\mathbb{E}(v(x-1))$. The difference between these two costs is equal to $c_B - (c_L + c_H)$ which is non negative by assumption. So, it would be better to let the customer in the system.

The coupling technique which corresponds directly to the interpretation of Section 5.1 consists in coupling residual impatience times of present customers in two trajectories corresponding to two distinct controls. This is possible *a priori* thanks to the memoryless property of the Geometric distribution. Using this technique requires the introduction of a formal apparatus, additional notation and a rigorous proof of the fact that relevant distributions are the same in the coupled systems (see *e.g.* [11]).

We have therefore chosen not to pursue this line of proof.

5.3 Extensions

We discuss here some extensions of the present model for which the principal result is preserved. As we have mentioned in the introduction, we envision other natural extensions (multiple servers, finite capacity, several impatience classes, non-unit service) but those do not fit yet in the framework we have developed here.

The extension of the results to the average cost criterion does not seem to be difficult, in the sense that letting $\theta = 1$ in Theorem 11 does not cause problems: the value of ψ is just defined as

$c_B - c_L - c_H/\alpha$, with an obvious interpretation. The needed increasingness property of Lemma 12 holds for the limit: $\lim_{\theta\to 1}(1-\theta)\Phi_\nu(x,\theta)$, so that Tauberian arguments are likely to be applicable.

Finally, the results obtained here extend to the case where customers do not simply disappear due to impatience, but are replaced with a random number B of new customers, in the manner of Galton-Watson branching processes. In that situation, the parameter α in the cost function (4) is to be interpreted as $\mathbb{P}(B=0)$. The parameter $\overline{\alpha}$ is the branching factor $\mathbb{E}B$ and does not coincide anymore with $(1-\alpha)$. The stochastic increasingness property (2) still holds, with $B(1) \equiv B$. The preliminary conditions for Theorem 3 (see the proof of Theorem 10) hold as long as $\overline{\alpha} = \mathbb{E}B < 1$, which corresponds to the stability of the process without service. The coupling argument of Lemma 15 can be extended as well to a branching process.

6 Conclusion

In this paper we have shown that the optimal control of service in a single queue with impatience is a threshold policy and we give the closed form of the value of this threshold. If the framework used could seem to be usual, its application here requires some additional concepts which do not appear in previous work. For example, here the monotonicity of the control requires a convex value function contrarily to the usual cases where only monotonicity of the value function is required (see [15]). This is due to the random departures due to impatience, which completely modify the dynamical behavior.

Because of this, the extension of the problem to the case where the server may serve more than one customer at a time, does not work in a straightforward manner. This may be seen on the simulation below, where for a case of batch of size larger than one, the dynamical operator does not have the submodular property of Lemma 8. The trajectory comparison result in Lemma 15 does not hold either.

Example 14. *Consider the model presented in Section 2, except that the server can serve batches of size up to $B = 5$. The evolution of the system* (1) *then becomes*

$$x_{n+1} \;=\; R(x_n, q_n) \;:=\; S\left((x_n - q_n B)^+\right) + A_{n+1}\;,$$

where $q_n \in \{0,1\}$ as before. The parameters are $\theta = 0.9$, $c_B = 1$, $c_L = 0.5$, $c_H = 0$, $\alpha = 0.2$, $\lambda = 1$ and the distribution of arrivals is Poisson. The sequence $Tv(x,1) - Tv(x,0)$ is non monotone. Indeed its values obtained by simulations (taking a size of the queue of 100*) for respectively x from* 0 *to* 8 *are:* 1.0, 0.709, 0.438, 0.192, −0.029, −0.249, −0.203, −0.197 *and* −0.224*. If the function Tv were submodular, this sequence would be nonincreasing (Proposition 5); this is not the case.*

On the other hand, no experimental evidence has contradicted, so far, the possibility that the optimal control still be of threshold type. Similar issues are addressed in [19] for the dual problem of admission control of batches. The challenge of further research on the topic will therefore be to find the appropriate properties that can be propagated by the dynamic programming operator in this case. This is why such a work is a first step basis. First, by showing that structural properties can be found even though it is not the submodularity for more complex cases. And equally by showing that the propagation of properties of the value function is necessary but not sufficient. Indeed, notions of "K-convexity" used in contexts which look similar (see [2] and [13]) do not work, because they would result in a submodular function $Tv(x,q)$, which is contradicted by Example 1 above.

References

[1] E. Altman. *Handbook of Markov Decision Processes Methods and Applications*, chapter Applications of Markov Decision Processes in Communication Networks : a survey. Kluwer, 2001.

[2] E. Altman and G. Koole. On submodular value functions and complex dynamic programming. *Stochastic Models*, 14:1051–1072, 1998.

[3] J. Artalejo and V. Pla. On the impact of customer balking, impatience and retrials in telecommunication systems. *Computers and Mathematics with Applications*, 57(2):217–229, 2009.

[4] P. P. Bhattacharya and A. Ephremides. Optimal scheduling with strict deadlines. *IEEE Trans. Automatic Control*, 34(7):721–728, July 1989.

[5] R. K. Deb and R. F. Serfozo. Optimal control of batch service queues. *Advances in Applied Probability*, 5(2):340–361, 1973.

[6] P. Glasserman and D. Yao. *Monotone Structure in Discrete-Event Systems*. Wiley, 1994.

[7] A. J. Kleywegt and J. D. Papastavrou. The dynamic and stochastic knapsack problem. *Operations Research*, 46:17–35, 1998.

[8] Y. L. Kocaga and A. R. Ward. Admission control for a multi-server queue with abandonment. Manuscript, July 2009.

[9] G. Koole. Monotonicity in markov reward and decision chains: Theory and applications. *Foundation and Trends in Stochastic Systems*, 1(1), 2006.

[10] G. Koole and A. Mandelbaum. Queueing models of call centers: An introduction. *Annals of Operations Research*, 113:41–59, 2002.

[11] Z. Liu, P. Nain, and D. Towsley. Sample path methods in the control of queues. *Queueing Systems*, 21:293–335, 1995.

[12] A. Movaghar. Optimal control of parallel queues with impatient customers. *Performance Evaluation*, 60:327–343, 2005.

[13] K. P. Papadaki and W. B. Powell. Exploiting structure in adaptative dynamic programming algorithms for a stochastic batch service problem. *European Journal of Operational Research*, 142:108–127, 2002.

[14] J. D. Papastavrou, S. Rajagopalan, and A. J. Kleywegt. The dynamic and stochastic knapsack problem with deadlines. *Management Science*, 42(12):1706–1718, 1996.

[15] M. Puterman. *Markov Decision Processes Discrete Stochastic Dynamic Programming*. Wiley, 2005.

[16] G. C. Resende and P. Pardalos. *Handbook of optimization in telecommunications*. Springer Science & Buisness Media, 2006.

[17] D. Stoyan. *Comparison Methods for Queues and Other Stochastic Models*. Wiley, 1983.

[18] D. Towsley and S. S. Panwar. On the optimality of minimum laxity and earliest deadline scheduling for real-time multiprocessors. In *Proceedings of IEEE EUROMICRO-90 Real Time Workshop*, pages 17–24, June 1990.

[19] E. B. Çil, E. L. Örmeci, and F. Karaesmen. Structural results on a batch acceptance problem for capacitated queues. *Mathematical Methods of Operations Research*, 66:263–274, 2007.

7 Appendix: Proof of Lemma 12

The proof of Lemma 12 requires some preliminary results and notations. Hence, we denote with π_ν the threshold policy with threshold ν. Let us denote $O_\nu^{(n)}$ the set of customers present in the system at slot n under policy π_ν. This set is evaluated at the beginning of the nth slot just before the decision. Customers are leaving this set due to either losses or service. We shall compare the sequences $O_\nu^{(n)}$ and $O_{\nu+1}^{(n)}$. We assume that the two stochastic systems are *coupled* through their arrival sequences, and the impatience of each customer. In other words, it is assumed that, at each slot, the same amount of customers arrive and also that the customers which are present simultaneously in both queues will leave simultaneously because of impatience.

It may however happen that at some slot, one customer is expected to leave under one of the policies, whereas it has been already serviced under the other policy and is not present in the corresponding queue. In this case nothing else occurs for this customer.

Lemma 15. *For every trajectory, we have*

$$O_{\nu+1}^{(n)} = \begin{cases} O_\nu^{(n)} \\ O_\nu^{(n)} \cup \{j_n\} \end{cases}$$

where j_n is the customer of smaller index in $O_{\nu+1}^{(n)}$, and the union is disjoint in the second case.

Proof. The proof is by induction. Obviously since only arrivals have occurred at time 0 we have $O_{\nu+1}^{(0)} = O_\nu^{(0)}$. Assume the result is true up to slot n. Then two cases may occur:

Case $O_{\nu+1}^{(n)} = O_\nu^{(n)}$. Three possibilities have to be considered.

If $\#O_\nu^{(n)} < \nu$, then nothing happens. Since arrivals and losses are identical it follows $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)}$.

If $\#O_\nu^{(n)} = \nu$, then a customer is served under π_ν and not under $\pi_{\nu+1}$. If we denote by $\bar{O}_\nu$ the intermediate set of customers which are present after decision under π_ν, we have $\bar{O}_\nu = O_\nu^{(n)} \setminus \{j_n\}$ and $\bar{O}_{\nu+1} = O_{\nu+1}^{(n)} = O_\nu^{(n)}$. Again two cases may occur, depending on the impatience of customer $\{j_n\}$. Either $\{j_n\}$ leaves the queue, yielding $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)}$. Indeed, because of the coupling, arrivals and departures of customers who are not $\{j_n\}$, are the same in both sets. Or $\{j_n\}$ remains, thus $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)} \cup \{j_n\}$ for the same reason.

If $\#O_\nu^{(n)} > \nu$, then one customer is served in both systems and this customer is the same. Consequently, $\bar{O}_\nu = \bar{O}_{\nu+1}$ and then $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)}$ once coupled arrivals and impatiences have been taken into account.

Case $O_{\nu+1}^{(n)} = O_\nu^{(n)} \cup \{j_n\}$. Two possibilities have to be considered.

Either $\#O_\nu^{(n)} < \nu$ and $\#O_{\nu+1}^{(n)} < \nu+1$. No service occurs. If $\{j_n\}$ leaves the queue with the other impatient customers then $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)}$. Otherwise, still because of the coupling, $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)} \cup \{j_n\}$.

Or else, $\#O_\nu^{(n)} \geq \nu$ and $\#O_{\nu+1}^{(n)} \geq \nu+1$, so that service occurs in both queues. Then, since j_n is the first customer in the queue under π_ν, we get $\bar{O}_\nu = O_\nu^{(n)} \setminus \{j_n\}$ and $\bar{O}_{\nu+1} = O_\nu^{(n)}$. If j_n belong to the set of customers that get impatient, then $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)}$. Otherwise, we have $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)} \cup \{j_n\}$.

Our claim is therefore proved. □

We can now prove the lemma.

Proof of Lemma 12. It is sufficient to prove that each function $\mathbb{P}(R_\nu^{(n)}(x) \geq \nu)$ enjoys the claimed properties. These functions are obviously positive. They are increasing with respect to x because the operator R_ν is stochastically increasing: $R_\nu(x+1) \geq_{st} R_\nu(x)$. There remains to prove the non-increasingness with respect to ν for every fixed x. This amounts to show that

$$\mathbb{P}\left(R_{\nu+1}^{(n)}(x) \geq \nu+1\right) \;\leq\; \mathbb{P}\left(R_\nu^{(n)}(x) \geq \nu\right) \; . \tag{16}$$

Observe that $R_\nu^{(n)}(x) = \#O_\nu^{(n)}$. According to Lemma 15, two cases may occur. When $O_{\nu+1}^{(n)} = O_\nu^{(n)}$ then $R_\nu^{(n)}(x) = R_{\nu+1}^{(n)}(x)$. Hence, if $R_{\nu+1}^{(n)}(x) \geq \nu+1$ then $R_\nu^{(n)}(x) \geq \nu+1$ which implies $R_\nu^{(n)}(x) \geq \nu$. When $O_{\nu+1}^{(n+1)} = O_\nu^{(n+1)} \cup \{j_n\}$ then $R_\nu^{(n)}(x) + 1 = R_{\nu+1}^{(n)}(x)$. Hence the events $R_{\nu+1}^{(n)}(x) \geq \nu+1$ and $R_\nu^{(n)}(x) \geq \nu$ are the same. So (16) is satisfied. □

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399