Optimal Static Priority Rules for Stochastic Scheduling with Impatience

Laboratoire G-SCOP
Grenoble Institute of Technology

A. Salch, J.-P. Gayon*, P. Lemaire

Abstract

We consider a stochastic scheduling problem with impatience to the end of service or impatience to the beginning of service. The impatience of a job can be seen as a stochastic due date. Processing times and due dates are random variables. Jobs are processed on a single machine with the objective to minimize the expected weighted number of tardy jobs in the class of static list scheduling policies. We derive optimal schedules when processing times and due dates follow different probability distributions.

Keywords: Scheduling Algorithms, Stochastic Systems

1. Introduction

In many service systems (e.g. hotlines), customers make requests and have to wait until their requests are met. If the service is not provided quickly enough, customers can decide to quit (or renege) the system. This is what happens when customers are subject to impatience. In the context of scheduling, the impatience of a customer can be modelled as a due date, because the service has to be provided before the customer becomes impatient. This due date is a random variable because one can not know *a priori* how long a customer will be ready to wait for a certain service.

Systems which deal with impatience can be separated in two categories: problems with Impatience to the End of Service (IES), traditionally considered in the scheduling literature, or with Impatience to the Beginning of Service (IBS). In problems with IES, a job has to end before its due date to be on time. This case is more consistent with production systems where a demand is satisfied when the production of the ordered item ends.

In problems with IBS, a penalty cost is incurred only when a job begins after its due date. No penalty is incurred otherwise. In particular, if the due date occurs during the process, the process can last for a long period of time without any cost. This second case is consistent with systems providing services to the customer such as call centers. In these systems, a customer is satisfied when she/he reaches the hotline. After that, one can consider that she/he will not hang up.

A difference must also be made between impatience and abandonment. A customer is impatient when she/he considers to have been waiting for too long (incurring a cost), but nevertheless remains in the queue; if the customer actually leaves the system, then it is an abandonment. Consequently, in systems with impatience, late jobs are processed but in systems with abandonment, they are not.

In this paper we study the problem of scheduling jobs on a single machine in order to minimize the expected weighted number of late jobs with IBS or IES and no abandonment. The optimal policy is searched among the class of static list scheduling policies. In this class of policies, a schedule is built at time zero and it cannot be changed thereafter.

2. Literature review

Relatively many papers consider abandonment (or reneging) in queueing systems with a single class of jobs (see e.g. [1, 2, 3, 4, 5]). Several papers investigate the scheduling of different classes of jobs in a system with abandonment in order to minimize the (weighted) number of late jobs. These papers consider strict priority rules. In a strict priority rule, jobs are ordered according to their priority and whenever a job of higher priority arises, the current job is preempted to the benefit of this new job. Atar et al. [6] prove that a strict priority rule is asymptotically optimal in an overloaded system. Panwar et al. [7] characterize an optimal policy when all durations are known at the arrival of jobs. Down et al. [8] consider a problem with two classes of customers, Poisson arrivals and all durations being exponential. They provide sufficient conditions for a strict priority rule to be optimal. Jang [9] and Jang and

*Corresponding author: 46, Avenue Félix Vialet 38031 Grenoble Cedex, France, Tel: (+33) 4 76 57 47 46, Fax: (+33) 4 76 57 46 95
Email address: `jean-philippe.gayon@grenoble-inp.fr` (J.-P. Gayon)

Klein [10] propose a heuristic for the problem of scheduling n different jobs with stochastic processing times and deterministic due dates. Seo et al. [11] give near optimal schedules when processing times are normal random variables and with a common due date.

Some other papers consider the scheduling of jobs with impatience costs but without abandonment. Argon et al. [12] study a scheduling problem with IES where all jobs are available at time zero, the objective being to minimize the expected number of late jobs in the class of dynamic policies. When only two classes of customers enter the system, the authors give conditions under which a strict priority rule is optimal. Pinedo [13] studies a stochastic IES scheduling problem on a single machine with the objective to minimize the expected weighted number of late jobs in the class of static list scheduling policies. For the particular case where processing times follow independent exponential distributions and due dates follow general distribution function, he proves that processing the jobs in non-increasing order of the ratio of their weights times their mean processing times, the so-called $c\mu$ rule, is optimal. Boxma and Forst [14] consider the same problem for some other probability distributions. In all the results mentioned above, either processing times or due dates are identically distributed.

In this paper, we extend the results of [13] and [14]. First, we consider situations where both processing times and due dates are not identically distributed. Second, we extend their results to IBS situations. This IBS counterpart has not been studied in the literature, to the best of our knowledge.

The remainder of this paper is organized as follows. Section 3 formulates the problem and introduces notations. Section 4 shows a summary of our results and Section 5 gives optimal static priority rules and their proofs.

3. Problem description

We consider a scheduling problem where n jobs have to be processed on a single machine. All jobs are available at time zero. A job j has a processing time X_j with mean $1/\mu_j$, a due date D_j with mean $1/\gamma_j$ and a deterministic weight w_j. We assume that all random variables are independent. A random variable Y that has a cumulative density function (c.d.f.) F_Y, is noted $Y \sim F_Y$. The probability density function (p.d.f.) of Y is noted f_Y. Especially, $Y \sim \exp(\gamma)$ means that Y is exponentially distributed with mean $1/\gamma$. A family Y_j of independent and identically distributed (i.i.d.) random variables with c.d.f. F_D will be denoted by $D_j \sim F_D$, without specifying the index of the distribution.

In the IES case, a job j is late if the end of its execution, C_j, occurs after its due date, D_j (i.e. when $C_j > D_j$). The value of the objective function for a schedule S is $C(S) = E(\sum w_j U_j)$, where U_j is assigned the value 1 if $C_j > D_j$, and 0 otherwise.

The IBS problem is similar to the IES problem except that a job j is late if the starting time of its execution, S_j, occurs after its due date, D_j (i.e. when $S_j > D_j$). The value of the objective function for a schedule S is $\tilde{C}(S) = E(\sum w_j \tilde{U}_j)$, where $\tilde{U}_j$ is assigned the value 1 if $S_j > D_j$, and 0 otherwise.

For both IES and IBS problems, our objective is to sequence the jobs in order to minimize the expected weighted number of late jobs in the class of static list scheduling policies. Since we are looking for a static policy, we are not allowed to change the schedule after time zero. Hence if a job is already late, it has to be processed in the predefined order, possibly before a job still on time. One can remark that there always exists an optimal solution without idle time. Consequently, S_j coincides with C_{j-1}, the completion time of job $j-1$.

Note that the deterministic IBS problem can be polynomially reduced to the deterministic IES problem by changing the due date d_j of the instance of the IBS problem to $d'_j = d_j + p_j$; the reverse is also true. In a stochastic setting, however, the sum of two random variables from the same class of distribution does not necessarily remain in this class of distribution. Hence, when the theoretical results depend on the selected distributions, the IBS problem can not be reduced to the IES problem, and vice-versa (see [15] for details).

4. Summary of results and comments

Table 1 summarizes the results of the literature together with our contributions for the IES and the IBS problems. Our results are proved in the next section.

In Problem 1, neither the processing times nor the due dates are identically distributed. To the best of our knowledge, either processing times or due dates are always assumed identically distributed in the literature. For exponential distributions, we show that if jobs can be simultaneously sequenced by non-decreasing γ_j and by non-increasing $w_j \gamma_j \mu_j$, then such a sequence is optimal for the IES problem. For the IBS problem, the additional condition that the weights w_j are non-increasing is needed. None of these assumptions can be relaxed.

In Problem 2, if jobs can be simultaneously sequenced by non-decreasing stochastic order of their processing times and by non-increasing order of their weights, then such a sequence is optimal. Suppose that X and Y are two stochastic variables, X is said to be stochastically smaller than Y ($X \leq_{st} Y$) if and only if $\mathbb{P}(X > t) \leq \mathbb{P}(Y > t)$ for all t. However, such a sequence does not necessarily exist, because two c.d.f. cannot always be compared with regard to stochastic order. The same scheduling rule is also optimal for the problem with IES.

In Problem 3 there are only two homogeneous classes of jobs in the system (n_1 jobs of class 1, n_2 jobs of class 2). For the IBS problem, we show that the optimal policy

♯	Problem	IES	IBS
1	$X_j \sim \exp(\mu_j)$, $D_j \sim \exp(\gamma_j)$	$I : \gamma_j \nearrow$ and $w_j\gamma_j\mu_j \searrow$	$II : \gamma_j \nearrow$, $w_j\gamma_j\mu_j \searrow$ and $w_j \searrow$
2	$X_j \sim F_{X_j}$, $D_j \sim F_D$	$X_j \nearrow_{st}$ and $w_j \searrow$ (*)	$X_j \nearrow_{st}$ and $w_j \searrow$
3	$X_j \sim F_X$, $D_j \sim \exp(\gamma_j)$, 2 classes of jobs	Threshold policy	Threshold policy
4	$X_j \sim F_{X_j}$, $D_j \sim \exp(\gamma)$	$\beta_j \searrow$ [14]	$\beta'_j \searrow$
5	$X_j \sim F_X$, $D_j \sim F_D$	$w_j \searrow$ [14]	$w_j \searrow$

(*) The special case $w_j = w$ was proved by [14]

Table 1: Optimal static list scheduling policies

is a threshold policy and we derive a closed form formula for the optimal threshold. Below this threshold, priority is given to one of the classes and above this threshold, priority is given to the other class. A threshold policy is also optimal for the problem with IBS.

In Problem 4, the optimal IES schedule is to process jobs in non-increasing order of $\beta_j = w_j/(1/\mathcal{L}\{f_{X_j}\}(\gamma)-1)$ and in non-increasing order of $\beta'_j = w_j/(1-\mathcal{L}\{f_{X_j}\}(\gamma))$ for the IBS problem, where $\mathcal{L}\{f\}(\gamma) = \int_{t=0}^{+\infty} f(t)\exp(-\gamma t)dt$ is the Laplace transform of f in γ.

For Problem 5, IBS and IES can be reduced one to the other and the same schedule is optimal in both cases.

In the next section, we detail the proofs of the results for Problem 1 to Problem 5 with IBS and for Problem 1 with IES. We do not provide proofs of the additional results for the IES problems, since the techniques that are used are similar (see technical report [15]).

5. Optimal static priority rules

First we provide a property which gives an analytic formula to compute the cost of a schedule for both kinds of impatience when processing times and due dates are exponentially distributed.

Property 1. *Consider a scheduling problem with n jobs. Job j has a weight w_j, an exponentially distributed processing time with mean $1/\mu_j$ and an exponentially distributed due date with mean $1/\gamma_j$. Then the expected weighted number of late jobs for schedule $S = \{1, 2, \dots n\}$ is*

$$C(S) = \sum_{j=1}^{n} w_j \left(1 - \prod_{k=1}^{j} \frac{\mu_k}{\mu_k+\gamma_j}\right) \text{ for IES problems and}$$
$$\tilde{C}(S) = \sum_{j=2}^{n} w_j \left(1 - \prod_{k=1}^{j-1} \frac{\mu_k}{\mu_k+\gamma_j}\right) \text{ for IBS problems.}$$

Proof. Since there is no idle time on the machine and jobs are performed even if they are tardy, the value of the objective function of schedule S is

$$C(S) = \sum_{j=1}^{n} w_j \mathbb{P}(C_j > D_j) \text{ for IES problems and} \quad (1)$$
$$\tilde{C}(S) = \sum_{j=2}^{n} w_j \mathbb{P}(X_1 + \ldots + X_{j-1} > D_j) \text{ for IBS problems.} \quad (2)$$

From Equation (1),

$$C(S) = \sum_{j=1}^{n} w_j \int_{t=0}^{+\infty} (1 - F_{C_j}(t)) f_{D_j}(t) dt = \sum_{j=1}^{n} w_j\gamma_j \int_{t=0}^{+\infty} (1 - F_{C_j}(t)) e^{-\gamma_j t} dt.$$

Using $\mathcal{L}\{\mu\exp(-\mu t)\}(\gamma) = \mu/(\mu+\gamma)$, the Laplace transform of an exponential p.d.f., leads to

$$C(S) = \sum_{j=1}^{n} w_j\gamma_j \mathcal{L}\{1 - F_{C_j}\}(\gamma_j) = \sum_{j=1}^{n} w_j \left(1 - \prod_{k=1}^{j} \frac{\mu_k}{\mu_k+\gamma_j}\right).$$

For these simplifications, we used the derivation property ($\mathcal{L}\{f\}(\gamma) = \gamma\mathcal{L}\{F\}(\gamma) - f(0)^-$) and the convolution property ($\mathcal{L}\{f_{X_i} * f_{X_j}\}(\gamma) = \mathcal{L}\{f_{X_i}\}(\gamma)\mathcal{L}\{f_{X_j}\}(\gamma)$) of the Laplace transforms [16].

Adapting this result for problems with IBS is trivial, up to some modifications in the indices of sums and products. This ends the proof of Property 1. □

Preliminaries

In order to prove the theorems, we use a pairwise interchange argument between two adjacent jobs. The two schedules

$$S : \{1, 2, \ldots s, i, u, s+3, \ldots n\} \quad \text{and}$$
$$S' : \{1, 2, \ldots s, u, i, s+3, \ldots n\}$$

differ only by the two jobs i and u which are swapped. These two jobs are in position $s+1$ and $s+2$, depending on the schedule under consideration. Since all jobs are processed, the end of execution of job s and the beginning of execution of job $s+3$, occur at the same time in both schedules.

IES case

The starting time of job i in schedule S, denoted by $Z = \sum_{j=1}^{s} X_j$, is equal to the starting time of job u in schedule S'. Then, from Equation (1),

$$C(S) = \sum_{j=1}^{s} w_j \mathbb{P}(C_j > D_j) + \sum_{j=s+3}^{n} w_j \mathbb{P}(C_j > D_j) + w_i \mathbb{P}(Z + X_i > D_i) + w_u \mathbb{P}(Z + X_i + X_u > D_u).$$

The same separation can be made on the cost of schedule S' and making the difference between these costs gives

$$\begin{aligned} &C(S') - C(S) \\ &= w_u \mathbb{P}(Z + X_u > D_u) + w_i \mathbb{P}(Z + X_u + X_i > D_i) \\ &\quad - w_i \mathbb{P}(Z + X_i > D_i) - w_u \mathbb{P}(Z + X_i + X_u > D_u). \end{aligned} \tag{3}$$

Schedule S performs better than schedule S' iff this difference is positive. However, as stated in [14], this formula "is too general to allow useful comments", and further assumptions are required. In [14], it was chosen to use independent and identically distributed processing times and/or due dates, to drasticaly simplify Equation (3). We do not assume i.i.d. variables, and thus require the two new lemmas below.

When considering exponential due dates, we can further simplify Equation (3).

Lemma 1. *When $D_j \sim \exp(\gamma_j)$ and $X_j \sim F_{X_j}$, we have*

$$\begin{aligned} &C(S') - C(S) \\ &= w_i \mathbb{P}(Z + X_i \le D_i) \mathbb{P}(X_u > D_i) \\ &\quad - w_u \mathbb{P}(Z + X_u \le D_u) \mathbb{P}(X_i > D_u) \\ &= w_i \mathcal{L}\{f_Z\}(\gamma_i) \mathcal{L}\{f_{X_i}\}(\gamma_i) \left[1 - \mathcal{L}\{f_{X_u}\}(\gamma_i)\right] \\ &\quad - w_u \mathcal{L}\{f_Z\}(\gamma_u) \mathcal{L}\{f_{X_u}\}(\gamma_u) \left[1 - \mathcal{L}\{f_{X_i}\}(\gamma_u)\right] \end{aligned}$$

Proof. We have

$$\begin{aligned} &\mathbb{P}(Z + X_u + X_i > D_i) \\ &= \mathbb{P}(Z + X_u + X_i > D_i \mid Z + X_i > D_i) \mathbb{P}(Z + X_i > D_i) \\ &\quad + \mathbb{P}(Z + X_u + X_i > D_i \mid Z + X_i \le D_i) \mathbb{P}(Z + X_i \le D_i) \end{aligned}$$

The memoryless property of D_i implies that

$$\mathbb{P}(Z + X_u + X_i > D_i \mid Z + X_i \le D_i) = \mathbb{P}(X_u > D_i).$$

Consequently,

$$\mathbb{P}(Z + X_u + X_i > D_i) = \mathbb{P}(Z + X_i > D_i) + \mathbb{P}(X_u > D_i) \mathbb{P}(Z + X_i \le D_i).$$

Applying this formula to $\mathbb{P}(Z + X_u + X_i > D_i)$ and $\mathbb{P}(Z + X_i + X_u > D_u)$ in Equation (3) explains the first equality of Lemma 1. If D is exponentially distributed with mean $1/\gamma$ then $\mathbb{P}(Z \le D) = \int_{t=0}^{+\infty} F_Z(t) \gamma \exp(-\gamma t) dt = \mathcal{L}\{f_Z\}(\gamma)$ which results in the second equality of the lemma using the convolution property of the Laplace transform. This ends the proof. □

IBS case

From Equation (2), the difference between the costs of schedules S and S' is

$$\begin{aligned} \tilde{C}(S') - \tilde{C}(S) = &w_u \mathbb{P}(Z > D_u) + w_i \mathbb{P}(Z + X_u > D_i) \\ &- w_i \mathbb{P}(Z > D_i) - w_u \mathbb{P}(Z + X_i > D_u). \end{aligned} \tag{4}$$

Again, considering exponential due dates leads to a simpler formula:

Lemma 2. *When $D_j \sim \exp(\gamma_j)$ and $X_j \sim F_{X_j}$, we have*

$$\begin{aligned} \tilde{C}(S') - \tilde{C}(S) = &w_i \mathbb{P}(Z \le D_i) \mathbb{P}(X_u > D_i) \\ &- w_u \mathbb{P}(Z \le D_u) \mathbb{P}(X_i > D_u) \\ = &w_i \mathcal{L}\{f_Z\}(\gamma_i) \left[1 - \mathcal{L}\{f_{X_u}\}(\gamma_i)\right] \\ &- w_u \mathcal{L}\{f_Z\}(\gamma_u) \left[1 - \mathcal{L}\{f_{X_i}\}(\gamma_u)\right]. \end{aligned}$$

Proof. Similar to that of Lemma 1. □

5.1. Problem 1: $X_j \sim \exp(\mu_j)$, $D_j \sim \exp(\gamma_j)$

Theorem 1. *Assume job j has a weight w_j, an exponentially distributed processing time with mean $1/\mu_j$ and an exponentially distributed due date with mean $1/\gamma_j$.*

If jobs can be ordered such that

$$I: \begin{cases} i) & \gamma_1 \le \gamma_2 \le \cdots \le \gamma_n \text{ and} \\ ii) & w_1\gamma_1\mu_1 \ge w_2\gamma_2\mu_2 \ge \cdots \ge w_n\gamma_n\mu_n \end{cases}$$

then scheduling jobs in non-decreasing order of the indices is optimal for problems with IES.

If jobs can be ordered such that

$$II: \begin{cases} i) & \gamma_1 \le \gamma_2 \le \cdots \le \gamma_n, \\ ii) & w_1\gamma_1\mu_1 \ge w_2\gamma_2\mu_2 \ge \cdots \ge w_n\gamma_n\mu_n \text{ and} \\ iii) & w_1 \ge w_2 \ge \cdots \ge w_n, \end{cases}$$

then scheduling jobs in non-decreasing order of the indices is optimal for problems with IBS.

Proof.
IES case

From Lemma 1, when all distributions are exponential, we have $C(S') - C(S) \ge 0$ iff

$$\frac{w_u\gamma_u}{\mu_i + \gamma_u} \frac{\mu_u}{\mu_u + \gamma_u} \mathcal{L}\{f_Z\}(\gamma_u) \le \frac{w_i\gamma_i}{\mu_u + \gamma_i} \frac{\mu_i}{\mu_i + \gamma_i} \mathcal{L}\{f_Z\}(\gamma_i). \tag{5}$$

As the Laplace transform is a decreasing function, $\gamma_i \le \gamma_u$ implies that $\mathcal{L}\{f_Z\}(\gamma_i) \ge \mathcal{L}\{f_Z\}(\gamma_u)$.

Moreover, $i)$ also implies that $(\mu_i+\gamma_u)(\mu_u+\gamma_u) \geq (\mu_u+\gamma_i)(\mu_i+\gamma_i)$ and then

$$\frac{w_u\gamma_u}{\mu_i+\gamma_u}\frac{\mu_u}{\mu_u+\gamma_u} \leq \frac{w_i\gamma_i}{\mu_u+\gamma_i}\frac{\mu_i}{\mu_i+\gamma_i} \text{ because of } ii).$$

Hence, with the set of assumptions I of Theorem 1, Inequality (5) holds for every pair of jobs and therefore it is optimal to schedule jobs in non-decreasing order of their indices. This ends the proof of Theorem 1 for IES.

IBS case

From Lemma 2, when all distributions are exponential, we have $\tilde{C}(S') - \tilde{C}(S) \geq 0$ iff

$$\frac{w_u\gamma_u}{\mu_i+\gamma_u}\mathcal{L}\{f_Z\}(\gamma_u) \leq \frac{w_i\gamma_i}{\mu_u+\gamma_i}\mathcal{L}\{f_Z\}(\gamma_i). \tag{6}$$

As before, with the set of assumptions II of Theorem 1, Inequality (6) holds for every pair of jobs and therefore it is optimal to schedule jobs in non-decreasing order of their indices. This ends the proof of Theorem 1 for IBS. □

Compared to the IES case, the IBS case requires the extra condition that the weights are non-increasing. To understand why, consider an instance with two jobs and equal due dates ($\gamma_1 = \gamma_2 = \gamma$). When γ is very large, the two jobs have very little chance to be completed on time. In the IBS problem, priority should be given to the job with the largest weight, in order to minimize the short term cost (no cost is paid for the job scheduled in first). This is not the case in the IES problem.

None of the assumptions of Theorem 1 can be relaxed. Table 2 provides counter-examples when one condition is relaxed. For example, the second row of Table 2 indicates that if condition i) is relaxed, then the instance described in the second column is a counter-example. Property 1 has been used to compute the costs.

Relaxed condition	Counter-examples for IES			
		w_i	μ_i	γ_i
i) $\gamma_j \nearrow$	Job 1	1	1	2
	Job 2	1	1	1
ii) $w_j\gamma_j\mu_j \searrow$	Job 1	1	1	1
	Job 2	1	2	2

Relaxed condition	Counter-examples for IBS			
		w_i	μ_i	γ_i
i) $\gamma_j \nearrow$	Job 1	1.5	0.5	2.5
	Job 2	1	0.5	3
	Job 3	0.5	0.5	0.5
ii) $w_j\gamma_j\mu_j \searrow$	Job 1	2	1	1
	Job 2	1	2	3
iii) $w_j \searrow$	Job 1	1	3	3
	Job 2	2	1	4

Table 2: Counter-examples when one condition is relaxed in Theorem 1

5.2. Problem 2: $X_j \sim F_{X_j}, D_j \sim F_D$

Theorem 2. *Assume that the processing times are independent ($X_j \sim F_{X_j}$) and that the due dates are i.i.d. ($D_j \sim F_D$). If jobs can be simultaneously sequenced 1) by non-decreasing stochastic order of their processing times and 2) by non-increasing order of their weights, then such a sequence is optimal for the IBS problem.*

Proof. Given that the due dates are i.i.d. ($D_j \sim F_D$) and using cumulative and density probability functions, Equation (4) can be simplified to

$$\begin{aligned}&\tilde{C}(S') - \tilde{C}(S)\\ &= \int_{t=0}^{+\infty} \big[(w_i - w_u)F_Z(t) - w_iF_{Z+X_u}(t)\\ &\qquad + w_uF_{Z+X_i}(t)\big] f_D(t)dt.\end{aligned}$$

Now suppose that $X_i \leq_{st} X_u$, i.e. $\mathbb{P}(X_i \leq t) \geq \mathbb{P}(X_u \leq t)$ $\forall t \geq 0$. Then for all $t \geq 0$ and $0 \leq z \leq t$, $\mathbb{P}(X_i \leq t - z \mid Z = z) \geq \mathbb{P}(X_u \leq t-z \mid Z = z)$ because D and all X_j are independent. By using the formula of conditional probability we obtain $\mathbb{P}(Z + X_i \leq t) = \mathbb{P}(X_i \leq t - z \mid Z = z)\mathbb{P}(Z = z) \geq \mathbb{P}(X_u \leq t - z \mid Z = z)\mathbb{P}(Z = z) = \mathbb{P}(Z + X_u \leq t)$. As a direct consequence, $F_{Z+X_i}(t) \geq F_{Z+X_u}(t)$ $\forall t \geq 0$. Then,

$$\begin{aligned}&\tilde{C}(S') - \tilde{C}(S)\\ &\geq \int_{t=0}^{+\infty} \big[(w_i - w_u)(F_Z(t) - F_{Z+X_u}(t))\big] f_D(t)dt.\end{aligned}$$

Given that all possible values for X_u are non-negative, we have the following inequality: $F_{Z+X_u}(t) = \mathbb{P}(Z + X_u \leq t) \leq \mathbb{P}(X_u \leq t) = F_{X_u}(t)$ for all $t \geq 0$. Then, all the terms in the integral are positive. This ends the proof of Theorem 2. □

5.3. Problem 3: $X_j \sim F_X, D_j \sim \exp(\gamma_j)$, 2 classes of jobs

Theorem 3. *Let the processing times be i.i.d. and follow a general distribution function F_X. Assume that there are n_1 jobs from class 1 (with weight w_1 and impatience rate γ_1) and n_2 jobs from class 2 (with weight w_2 and impatience rate γ_2). The total number of jobs is $n = n_1 + n_2$. Let $\alpha_j = \mathcal{L}\{f_X\}(\gamma_j)$ for $j \in \{1, 2\}$ and assume without loss of generality that $\alpha_1 \geq \alpha_2$. Then, it is optimal to give priority to class 1 in any position s such that*

$$s \geq t_{IBS} = \ln\left(\frac{w_2(1-\alpha_2)}{w_1(1-\alpha_1)}\right)\left(\ln\left(\frac{\alpha_1}{\alpha_2}\right)\right)^{-1}.$$

Proof. Starting again from Lemma 2 with i.i.d. processing times

$$\begin{aligned}&\tilde{C}(S') - \tilde{C}(S)\\ &= w_i\left[\mathcal{L}\{f_X\}(\gamma_i)\right]^s(1 - \mathcal{L}\{f_X\}(\gamma_i))\\ &\quad - w_u\left[\mathcal{L}\{f_X\}(\gamma_u)\right]^s(1 - \mathcal{L}\{f_X\}(\gamma_u)).\end{aligned} \tag{7}$$

This inequality describes that it is preferable to process job i before job u when looking for which job to process in position $s+1$. If it is always preferable to process job i before job u, i.e. the inequality is valid for all s, then i should be processed before u in an optimal schedule.

When there are only two classes of jobs in the system, the quantity $(\tilde{C}(S') - \tilde{C}(S))$ is non-negative in Equation (7) if $w_1\alpha_1^s(1-\alpha_1) \geq w_2\alpha_2^s(1-\alpha_2)$.

When $\alpha_1 = \alpha_2$, $(\tilde{C}(S') - \tilde{C}(S))$ is non-negative if $w_1 \geq w_2$, independently of position s. It is therefore optimal to always give priority to class 1 when $w_1 \geq w_2$ and $\alpha_1 = \alpha_2$.

When $\alpha_1 > \alpha_2$, it is optimal to give priority to class 1 when

$$s \geq t_{IBS} = \ln\left(\frac{w_2(1-\alpha_2)}{w_1(1-\alpha_1)}\right)\left(\ln\left(\frac{\alpha_1}{\alpha_2}\right)\right)^{-1}.$$

Hence it is optimal to give priority to class 1 in position s when $s \geq t_{IBS}$. If $t_{IBS} \leq 0$, priority is always given to class 1. This occurs for example when w_1 is very large. If $t_{IBS} > n$, it means that priority is always given to class 2. This occurs for example when w_2 is very large. □

The same proof can be used to prove the optimality of a threshold policy for the problem with IES. In this case it is optimal to give priority to class 1 in position s when $s \geq t_{IES} = t_{IBS} - 1$. (See [15] for a proof.)

5.4. Problem 4: $X_j \sim F_{X_j}, D_j \sim \exp(\gamma)$

Theorem 4. *Let the processing times be independent random variables ($X_j \sim F_j$) and the due dates be i.i.d., exponentially distributed with mean $1/\gamma$. An optimal IBS schedule is to process jobs in non-increasing order of $\beta'_j = w_j/(1-\mathcal{L}\{f_{X_j}\}(\gamma))$.*

Proof. Starting from Lemma 2 with i.i.d. and exponential due dates, $\tilde{C}(S') - \tilde{C}(S) \geq 0$ iff $w_i[1-\mathcal{L}\{f_{X_u}\}(\gamma) \geq w_u[1-\mathcal{L}\{f_{X_i}\}(\gamma)]$, which ends the proof of the theorem. □

5.5. Problem 5: $X_j \sim F_X$, $D_j \sim F_D$

Theorem 5. *Let the processing times be i.i.d. random variables ($X_j \sim F_X$) and the due dates be also i.i.d. random variables ($D_j \sim F_D$). An optimal IBS static list scheduling policy is to process the jobs in non-increasing order of their weights.*

Proof. The optimal IES schedule is to process the jobs in non-increasing order of their weights [14]. Here one can modify the due dates from the IBS problem to $D'_j = D_j + X_j$ resulting in an instance of an IES problem. The due dates are still i.i.d. and since the IES policy does not depend on the probability distributions of the due dates, the same optimal scheduling rule holds for the IBS problem. □

References

[1] S. Zeltyn, Call centers with impatient customers: exact analysis and many-server asymptotics of the $M/M/n+G$ queue, Ph.D. thesis, Israel Institute of Technology (2004).

[2] A. Movaghar, On queueing with customer impatience until the end of service, Stochastic Models 22 (1) (2006) 149–173. doi:10.1080/15326340500481804.

[3] A. Movaghar, On queueing with customer impatience until the beginning of service, Queueing Systems 29 (2-4) (1998) 337–350.

[4] A. Ward, S. Kumar, Asymptotically optimal admission control of a queue with impatient customers, Mathematics of Operations Research 33 (1) (2008) 167–202.

[5] S. Benjaafar, J.-P. Gayon, S. Tepex, Optimal control of a production-inventory system with customer impatience, Operations Research Letters 38 (4) (2010) 267–272.

[6] R. Atar, C. Giat, N. Shimkin, The $c\mu/\theta$ rule for many-server queues with abandonment, Operations Research 58 (2010) 1427–1439. doi:10.1287/opre.1100.0826.

[7] S. Panwar, D. Towsley, J. Wolf, Optimal scheduling policies for a class of queues with customer deadlines to the beginning of service, Journal of the ACM 35 (4) (1988) 832–844.

[8] D. Down, G. Koole, M. Lewis, Dynamic control of a single-server system with abandonments, Queueing Systems 67 (2011) 63–90.

[9] W. Jang, Dynamic scheduling of stochastic jobs on a single machine, European Journal of Operational Research 138 (3) (2002) 518–530.

[10] W. Jang, C. Klein, Minimizing the expected number of tardy jobs when processing times are normally distributed, Operations Research Letters 30 (2) (2002) 100–106.

[11] D. Seo, C. Klein, W. Jang, Single machine stochastic scheduling to minimize the expected number of tardy jobs using mathematical programming models, Computers and Industrial Engineering 48 (2) (2005) 153–161.

[12] N. Argon, S. Ziya, R. Righter, Scheduling impatient jobs in a clearing system with insights on patient triage in mass casualty incidents, Probability in the Engineering and Informational Sciences 22 (2008) 301–332.

[13] M. Pinedo, Stochastic scheduling with release dates and due dates, Operations Research 31 (3) (1983) 559–572.

[14] O. Boxma, F. Forst, Minimizing the expected weighted number of tardy jobs in stochastic flow shops, Operations Research Letters 5 (3) (1986) 119–126.

[15] A. Salch, J.-P. Gayon, P. Lemaire, Stochastic scheduling with impatience, Tech. rep., available online (2011). URL `http://hal.archives-ouvertes.fr/hal-00641681/fr/`

[16] P. Kuhfittig, Introduction to the Laplace Transform, PLENUM Publishing Co., N.Y., 1978.