Les cahiers
Leibniz

Stochastic Scheduling with Impatience to the Beginning of Service

Alexandre Salch, Jean-Philippe Gayon, Pierre Lemaire

Laboratoire G-SCOP
46 av. Félix Viallet, 38000 GRENOBLE, France
ISSN : 1298-020X

n° 194

November 2011

Site internet : http://www.g-scop.inpg.fr/CahiersLeibniz/

Stochastic Scheduling with Impatience to the Beginning of Service

Alexandre Salch, Jean-Philippe Gayon* and Pierre Lemaire

Laboratoire G-SCOP

Grenoble Institute of Technology

46, Avenue Félix Vialet

38031 Grenoble Cedex, France

* Corresponding author

E-mail: jean-philippe.gayon@grenoble-inp.fr

Tel: (+33) 4 76 57 47 46, Fax: (+33) 4 76 57 46 95

Stochastic Scheduling with Impatience to the Beginning of Service

July 28, 2011

Abstract

In this paper we discuss a stochastic scheduling problem with impatience to the beginning of service. The impatience of a job can be seen as a due date and a job is considered to be on time if it begins to be processed before its 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. We also study the relation between the problem with impatience to the beginning of service and the problem with impatience to the end of service. Finally, we provide some additional results for the problem with impatience to the end of service.

Keywords: Stochastic scheduling, Impatience

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 modeled 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 Beginning of Service (IBS), considered here, or with Impatience to the End of Service (IES), traditionally considered in the scheduling literature. In problems with IBS, jobs which are in process are already considered to be on time. A penalty cost is incurred only when a job begins after its due date. The process could last for a long period of time without additional cost. This problem is consistent with systems providing services to the customer such as call centers. In these systems, a customer is satisfied when he reaches the hotline. After that, one can consider that he will not hang up. In problems with IES, a job has to end before its due date to be on time. This second case is more consistent with production systems where a demand is satisfied when the production of the ordered item ends.

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 and no abandonment. That is all jobs have to be processed, even if they are already tardy. The optimal policy is searched among the class of static list scheduling policies. If the optimal policy is searched among the class of dynamic policies, decisions can be made at any time, taking advantage of all available information. Consequently it would be clearly sub-optimal to process jobs which already passed their due dates before jobs still on time. When the optimal policy is searched among the class of static policies, a static list scheduling policy is built at time zero and this list can not be changed thereafter. As a consequence, a job that passed its due date, or equivalently, a job being impatient, may be processed before jobs that are still on time.

The remainder of this paper is organized as follows. In Section 2 we present a literature review on scheduling stochastic jobs with impatience. In Section 3 we formulate the problem and introduce notations. In Section 4 we study the relation between the IES and the IBS problems. In Section 5 we summarize our main results on the IBS problem. Section 6 details the proofs of these results. Section 7 presents conclusions and avenues for research.

2 Literature review

Several papers consider impatience in systems with a single class of jobs and reneging. Zeltyn and Mandelbaum [17] propose results on performance evaluation for IBS problems, so does

Movaghar for both IES [12] and IBS [11] problems. Other authors study the optimal control of these systems. Ward and Kumar [16] give an optimal admission control in heavy traffic for problems with IBS. Benjaafar et al. [3] study a make-to-stock system with IES. In these papers, all jobs have the same characteristics (processing time, impatience,...).

Several papers investigate the scheduling of different classes of jobs in a system with abandonment (or reneging). Atar et al. [2] prove that a strict priority rule is asymptotically optimal in an overloaded system with IBS. Panwar et al. [13] characterize an optimal policy when all durations are known at the arrival of jobs. Down et al. [5] 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 [8] and Jang and Klein [9] propose a heuristic for the problem of scheduling n different jobs with stochastic processing times and deterministic due dates. Seo et al. [15] 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 without abandonment. Argon et al. [1] 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 [14] 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 [4] consider the same problem for some other probability distributions. We will later detail the results of [14] and [4] (see Table 1, Section 5).

In this paper, we investigate the IBS counter-part of the problems investigated by [14] and [4]. This IBS counterpart has not been studied in the literature, to the best of our knowledge. More precisely, we address the problem of scheduling jobs with IBS and without abandonment, in the class of static list scheduling policies. We derive the optimal schedule when processing times and due dates follow different probability distributions. We also study the relation between the IBS problem and the IES problem. Moreover we provide some additional results for the IES problem.

3 Problem description

We first describe in details the IBS problem before presenting quickly the IES problem. 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 and a due date D_j, that are independent random variables, and a deterministic weight w_j. 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 will be 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.

The objective is to sequence the jobs in order to minimize the expected weighted number of late jobs. A job j is said to be 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 $C(S) = E(\sum w_j \tilde{U}_j)$, where $\tilde{U}_j$ is assigned the value 1 if $S_j > D_j$, and 0 otherwise. We investigate the conditions of optimality of schedules among 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. That is why we only consider such schedules. Consequently, S_j coincides with C_{j-1}, the completion time of job $j-1$. The deterministic version of the IBS problem with due dates d_j can be noted $1 \mid d_j \mid \sum w_j \tilde{U}_j$ by adapting Graham's notation [7].

The IES problem is similar to the IBS problem except that 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 $E(\sum w_j U_j)$, where U_j is assigned the value 1 if $C_j > D_j$, and 0 otherwise. The deterministic version of the IBS problem with due dates d_j can be noted $1 \mid d_j \mid \sum w_j U_j$, with Graham's notation.

4 Relation between IES and IBS problems

We remind that the definition of tardiness depends on whether IES or IBS is assumed. In this section, we investigate the relation between IES and IBS problems.

4.1 Deterministic due dates and processing times

We show here that our deterministic IBS problem can be polynomialy reduced to the deterministic IES problem within a polynomial time. Let s_j and c_j be the begining and the end of job j. We assume that we have an oracle which solves the IES problem. Then changing the due date d_j of the instance of the IBS problem to d'_j as follow $d'_j = d_j + p_j$ allows us to solve the modified instance of the IBS problem with the oracle. The reciprocal is shown with the same kind of transformation of an instance of the IES problem ($d'_j = d_j - p_j$). Both transforms (from IES to IBS and the other way around) are polynomial reductions, showing that both problems belong to the same class of computational complexity [6]. But even if IES and IBS problems can be reduced from one to the other, we show in Appendix A that an optimal IBS schedule can perform arbitrarily bad for the IES problem and vice-versa. The value of the objective function of an optimal schedule of an IES problem can nevertheless be used as an upper bound for the IBS problem on the same instance. This holds because a job that is on time for an IES problem is also on time for an IBS problem.

4.2 Stochastic due dates and/or processing times

Again, we assume that we have an oracle which solves the stochastic IES problem. We modify the random variables of the due dates D_j of an instance of the IBS problem to $D'_j = D_j + X_j$. This way we obtain an instance of an IES problem. However the sum of two random variables from the same class of distribution does not necessarily remain in this class of distribution. For example, the sum of two random variables that follow an exponential distribution does not follow an exponential distribution. If the optimal schedule takes advantage of properties of D to be tractable (e.g., the absence of memory of an exponential distribution), it will not be possible to use the same argument replacing D by $D + X$. Since, in a stochastic context, results in the literature depend on the class of distributions under consideration (see Table 1 in next section), there is no reason that it would always be possible to solve to optimality this IES instance with the same oracle. Consequently, modifying the distribution of a stochastic problem implies that prior results can not always be applied as it is the case in the previous section, unless the original and modified random due dates have the same properties.

We now provide an example where results from the stochastic IES problem can be extended to the stochastic IBS problem.

Theorem 1. *When the processing times are i.i.d. random variables ($X_j \sim F_X$) and due dates* *are also i.i.d. random variables ($D_j \sim F_D$), an optimal IBS static list scheduling policy that minimizes the expected weighted number of late jobs is to process the jobs in non-increasing order of their weights.*

Intuitively, since the processing times and due dates of all jobs are the same, jobs are processed according to the only parameter which differs from one job to the other: their weights.

Proof. Boxma and Forst [4] prove that the optimal IES schedule is to process the jobs in non-increasing order of their weights for the IES problem. Here we can modify the due dates from the IBS problem by $D'_j = D_j + X_j$ 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. □

5 Summary of results

Table 1 summarizes the results of the literature and our contributions (in bold) for the IBS and the IES problem. Our results are proved in the next sections.

$\sharp$	Problem	IES	IBS
1	$X_j \sim \exp(\mu_j)$, $D_j \sim F_D$	$w_j\mu_j \searrow$ [4]	Counter-example **(App. B)**
2	$X_j \sim F_X$, $D_j \sim F_D$	$w_j \searrow$ [4]	$w_j \searrow$ **(Sec. 4.2)**
3	$X_j \sim F_{X_j}$, $D_j \sim F_D$	$X_j \nearrow_{st}$ and $w_j = w$ [4] $X_j \nearrow_{st}$ and $w_j \searrow$ **(Sec. D.2)**	$X_j \nearrow_{st}$ and $w_j \searrow$ **(Sec. 6.2)**
4	$X_j \sim F_{X_j}$, $D_j \sim \exp(\gamma)$	$\beta_j \searrow$ [4]	$\beta'_j \searrow$ **(Sec. 6.3)**
5	$X_j \sim F_X$, $D_j \sim \exp(\gamma_j)$	$\delta_j(s) \searrow$ [4]	$\delta'_j(s) \searrow$ **(Sec. 6.4)**
6	$X_j \sim F_X$, $D_j \sim \exp(\gamma_j)$, 2 classes of jobs	Threshold policy **(Sec. D.3)**	Threshold policy **(Sec. 6.5)**

Table 1: Optimal static list scheduling policies

In Problem 1, the processing time of job j follows an exponential distribution with mean $1/\mu_j$ and the due dates are i.i.d. random variables. An optimal IES schedule is to process the jobs in non-increasing order of $w_j\mu_j$. We are not able to characterize the optimal list scheduling policy for the IBS counterpart. However we provide a counter-example in Appendix B that shows that processing jobs in non-increasing order of $w_j\mu_j$ is not optimal (see Appendix B). To gain some insight on this problem, one can notice that for the IES problem, whatever the rate of the due dates is, the optimal policy serves the job with highest $w_j\mu_j$ first, even if the rate of the due dates is large compared to the other parameters of the system. In such a case, the IES scheduling rule is clearly not optimal for the IBS problem since there is no chance to process more than one job on time. Hence the job with the highest weight should be processed first. If an optimal static list scheduling policy exists, it certainly depends on the distribution of the due dates.

In Problem 2, the processing times and the due dates are respectively i.i.d. The optimal IBS and IES schedules are to process the jobs in non-increasing order of their weights, as stated in Theorem 1. Remind that in this case, IES and IBS problems can be reduced to each other (see Section 4.2).

In Problem 3, the processing times are independent random variables and the due dates are i.i.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. 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. can not always be compared with regard to stochastic order.

In Problem 4, the processing times follow independent random variables and due dates are i.i.d. and exponential with mean $1/\gamma$. Then the optimal IBS schedule is to process jobs in non-increasing order of $\beta'_j = w_j(1-\mathcal{L}\{f_{X_j}\}(\gamma))$ and in non-increasing order of $\beta_j = w_j/(1/\mathcal{L}\{f_{X_j}\}(\gamma)-1)$ for the IES problem, where $\mathcal{L}\{f\}(s) = \int_{t=0}^{+\infty} e^{-st} f(t)\, dt$ is the Laplace transform of a function f in s.

In Problem 5, the processing times follow i.i.d. random variables and the due date of a job j is exponential with mean $1/\gamma_j$. Then an optimal IBS schedule is to process jobs in non-increasing order of $\delta'_j(s) = w_j\left[\mathcal{L}\{f_X\}(\gamma_j)\right]^s(1-\mathcal{L}\{f_X\}(\gamma_j))$ for all $s \in \{0, 1, \ldots, n-2\}$, and in non-increasing order of $\delta_j(s) = w_j\left[\mathcal{L}\{f_X\}(\gamma_j)\right]^{s+1}(1-\mathcal{L}\{f_X\}(\gamma_j))$ for all s for the IES problem, where s represents the position of the last job before those we interchange. As in Problem 3, it may not be possible to order the set of jobs under these relations.

Problem 6 is a special case of Problem 5 with only two homogeneous classes of jobs in the system (n_1 jobs of class 1, n_2 jobs of class 2). For this problem, we show that the optimal policy is a threshold policy. Below a certain threshold, priority is given priority to one of the classes and above this threshold, priority is given to the other class.

In the next section, we detail the proofs of our results for the IBS problem. In Appendix D, we detail the proofs of our additional results for the IES problem.

6 Optimal IBS Static Priority Rules

6.1 Preliminaries

In order to prove the optimality of the IBS scheduling rules summarized in Table 1, we use a pairwise interchange argument between two adjacent jobs. The two schedules

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

differ only by the two adjacent 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 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

$$C(S) = \sum_{j=2}^{n} w_j \mathbb{P}(X_1 + \ldots + X_{j-1} > D_j).$$

Let us introduce the random variable $Z_s = \sum_{j=1}^{s} X_j$, then the costs of the two schedules are

$$\begin{aligned} C(S) &= \sum_{j=2}^{s} w_j \mathbb{P}(X_1 + \ldots + X_{j-1} > D_j) \\ &+ w_i \mathbb{P}(Z_s > D_i) + w_u \mathbb{P}(Z_s + X_i > D_u) \\ &+ \sum_{j=s+3}^{n} w_j \mathbb{P}(X_1 + \ldots + X_i + X_u + \ldots + X_{j-1} > D_j) \quad \text{and} \\ C(S') &= \sum_{j=2}^{s} w_j \mathbb{P}(X_1 + \ldots + X_{j-1} > D_j) \\ &+ w_u \mathbb{P}(Z_s > D_u) + w_i \mathbb{P}(Z_s + X_u > D_i) \\ &+ \sum_{j=s+3}^{n} w_j \mathbb{P}(X_1 + \ldots + X_u + X_i + \ldots + X_{j-1} > D_j). \end{aligned}$$

By making the difference between the two schedules, we have

$$\begin{aligned} C(S') - C(S) = & \; w_u \mathbb{P}(Z_s > D_u) + w_i \mathbb{P}(Z_s + X_u > D_i) \\ & - w_i \mathbb{P}(Z_s > D_i) - w_u \mathbb{P}(Z_s + X_i > D_u) \end{aligned}$$

and using cumulative and density probability functions leads to

$$C(S') - C(S) = w_i \int_{t=0}^{+\infty} (F_{Z_s}(t) - F_{Z_s+X_u}(t)) f_{D_i}(t)\, dt \qquad (1)$$
$$- w_u \int_{t=0}^{+\infty} (F_{Z_s}(t) - F_{Z_s+X_i}(t)) f_{D_u}(t)\, dt.$$

When this difference is positive, schedule S performs better than schedule S' and one can take advantage of the value of this difference to obtain conditions on a list scheduling policy. Equation (1) will be further simplified using the particular distribution functions considered in the next problems.

6.2 Problem 3: $X_j \sim F_{X_j}, D_j \sim F_D$

Theorem 2. *Consider independent processing times ($X_j \sim F_{X_j}$) and i.i.d. due dates ($D_j \sim F_D$). 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 problem of minimizing the expected weighted number of late jobs with IBS.*

Proof. Given that the due dates are i.i.d. ($D_i \sim F_D$), Equation (1) can be simplified as follow

$$C(S') - C(S) = \int_{t=0}^{+\infty} \left[(w_i - w_u)F_{Z_s}(t) - w_i F_{Z_s+X_u}(t) + w_u F_{Z_s+X_i}(t)\right] f_D(t) dt.$$

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 $0 \leq z \leq t$, $\mathbb{P}(X_i \leq t - z \mid Z_s = z) \geq \mathbb{P}(X_u \leq t - z \mid Z_s = z)$ because D and all X_j are independent. By using the formula of conditional probability we obtain $\mathbb{P}(Z_s + X_i \leq t) = \mathbb{P}(X_i \leq t - z \mid Z_s = z)\mathbb{P}(Z_s = z) \geq \mathbb{P}(X_u \leq t - z \mid Z_s = z)\mathbb{P}(Z_s = z) = \mathbb{P}(Z_s + X_u \leq t)$. As a direct consequence, $F_{Z_s+X_i}(t) \geq F_{Z_s+X_u}(t)\ \forall t \geq 0$. This relation implies a lower bound on the difference between the objective functions,

$$C(S') - C(S) \geq \int_{t=0}^{+\infty} \left[(w_i - w_u)F_{Z_s}(t) - (w_i - w_u)F_{Z_s+X_u}(t)\right] f_D(t) dt$$
$$\geq \int_{t=0}^{+\infty} \left[(w_i - w_u)(F_{Z_s}(t) - F_{Z_s+X_u}(t))\right] f_D(t) dt.$$

Given that all possible values for X_u are non-negative, we have the following inequality: $F_{Z_s+X_u}(t) = \mathbb{P}(Z_s + X_u \leq t) \leq \mathbb{P}(X_u \leq t) = F_{X_u}(t)$ for all $t \geq 0$. And finally

$$C(S') - C(S) \geq (w_i - w_u) \int_{t=0}^{+\infty} f_D(t) dt = w_i - w_u.$$

So we can conclude that the difference is non-negative if $w_i \geq w_u$. If the jobs can be ordered such that their processing times are non-decreasing in the sense of stochastic order and if among this order, the weights are non-increasing, then this schedule is optimal. □

For example, when $X_j \sim \exp(\mu_j)$, $D_j \sim F_D$ and $w_j = w$, the optimal IBS schedule is to process the jobs in non-increasing order of μ_j, since $X_i \leq_{st} X_u$ if and only if $\mu_i \geq \mu_j$.

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

Theorem 3. *In the case of independent processing times ($X_j \sim F_j$) and i.i.d. due dates, exponentially distributed with mean $1/\gamma$, an optimal schedule is to process jobs in non-increasing order of $\beta'_j = w_j/(1 - \mathcal{L}\{f_{X_j}\}(\gamma))$.*

Remember that when $D \sim \exp(\gamma)$, $\mathbb{P}(X > D) = \int_{t=0}^{+\infty}(1 - F_X(t))\gamma \exp^{-\gamma t} dt$, which can be rewritten $\mathbb{P}(X > D) = 1 - \mathcal{L}\{f_X\}(\gamma)$. That is to say, the denominator of β'_j represents the probability that job j ends after a realization of the random variables defining the due dates. Consequently a job with a large processing time with respect to the due dates should be processed at the end of the schedule, in order not to penalize other jobs, with respect to a correction factor which is the weight of the job.

Proof. In this proof we will use the convolution and the derivative properties of Laplace transforms that can be found in almost any reference on Laplace transforms, as for example in [10]. Starting from Equation (1) and replacing $f_{D_i}(t)$ by $\gamma e^{-\gamma t}$ results in

$$\begin{aligned} & C(S') - C(S) \\ & = \gamma \int_{t=0}^{+\infty} e^{-\gamma t} \Big[(w_i - w_u)F_{Z_s}(t) - w_i F_{Z_s+X_u}(t) + w_u F_{Z_s+X_i}(t)\Big]\, dt. \end{aligned}$$

This is the expression of a Laplace transform and the difference can be further simplified to

$$\begin{aligned} & C(S') - C(S) \\ & = \gamma \mathcal{L}\{(w_i - w_u)F_{Z_s} - w_i F_{Z_s+X_u} + w_u F_{Z_s+X_i}\}(\gamma). \end{aligned}$$

The c.d.f. of all processing times must be equal to 0 in $t = 0$, so we can use the derivation property of Laplace transforms, which leads to

$$\begin{aligned} & C(S') - C(S) \\ & = \mathcal{L}\{(w_i - w_u)f_{Z_s} - w_i f_{Z_s+X_u} + w_u f_{Z_s+X_i}\}(\gamma) \end{aligned}$$

and we can apply the convolution property of Laplace transforms to obtain the final expression

$$\begin{aligned} & C(S') - C(S) \\ & = \left(\prod_{k=1}^{s} \mathcal{L}\{f_{X_k}\}(\gamma)\right) \Big[(w_i - w_u) - w_i \mathcal{L}\{f_{X_u}\}(\gamma) + w_u \mathcal{L}\{f_{X_i}\}(\gamma)\Big]. \end{aligned}$$

The difference is non-negative if and only if $\frac{w_i}{1-\mathcal{L}\{f_{X_i}\}(\gamma)} \geq \frac{w_u}{1-\mathcal{L}\{f_{X_u}\}(\gamma)}$. Which ends the proof of the theorem. $\square$

For example, when $X_j \sim \exp(\mu_j)$ and $D_j \sim \exp(\gamma)$, the optimal IBS schedule is to process the jobs in non-increasing order of $w_j(\mu_j + \gamma)$, using the Laplace transform of an exponential p.d.f.: $\mathcal{L}\{f(t) = \mu e^{-\mu t}\}(\gamma) = \mu/(\mu + \gamma)$.

6.4 Problem 5: $X_j \sim F_X, D_j \sim \exp(\gamma_j)$

Theorem 4. *If the processing times are i.i.d. and follow a general distribution function F_X and the due date of job j is exponential with mean $1/\gamma_j$ and if it is possible to sequence the jobs in non-increasing order of $\delta'_j(s) = w_j\left[\mathcal{L}\{f_X\}(\gamma_j)\right]^s(1 - \mathcal{L}\{f_X\}(\gamma_j))$ for all $s \in \{0, 1, \ldots, n-2\}$, then such a sequence is optimal for the IBS problem.*

When $D_j \sim \exp(\gamma_j)$ then $\mathbb{P}(X < D_j) = \mathcal{L}\{f_X\}(\gamma_j)$. s represents the position of the last job before j (j is processed in position $s+1$). Consequently, $\left[\mathcal{L}\{f_X\}(\gamma_j)\right]^s = \left[\mathbb{P}(X < D_j)\right]^s$ is the probability that job j is still not late at the end of the execution of job s. Multiplying this quantity by $1 - \mathcal{L}\{f_X\}(\gamma_j) = \mathbb{P}(X > D_j)$ will represent the probability that job j abandons just during the execution of the job in position $s+1$. Consequently $\delta'_j(s)$ represents a just-in-time factor for job j in position s. The greater this factor, the closest job j is from its optimal position in the schedule. And if the jobs can be sequenced such that all of them are in their optimal positions, then this sequence is optimal.

Proof. Starting again from Equation (1) and replacing $f_{D_j}(t)$ by $\gamma_j e^{-\gamma_j t}$ leads to the difference

$$\begin{aligned} &C(S') - C(S) \\ &= w_i\gamma_i\mathcal{L}\{F_{Z_s} - F_{Z_s+X}\}(\gamma_i) - w_u\gamma_u\mathcal{L}\{F_{Z_s} - F_{Z_s+X}\}(\gamma_u). \end{aligned}$$

Now we use the integer $s \in \{0, 1, \ldots, n-2\}$, which is the position of the job just before those we interchange and the difference becomes

$$\begin{aligned} &C(S') - C(S) \\ &= w_i\left[\mathcal{L}\{f_X\}(\gamma_i)\right]^s(1 - \mathcal{L}\{f_X\}(\gamma_i)) \\ &\quad - w_u\left[\mathcal{L}\{f_X\}(\gamma_u)\right]^s(1 - \mathcal{L}\{f_X\}(\gamma_u)). \end{aligned} \tag{2}$$

If $w_i\left[\mathcal{L}\{f_X\}(\gamma_i)\right]^s(1 - \mathcal{L}\{f_X\}(\gamma_i)) \geq w_u\left[\mathcal{L}\{f_X\}(\gamma_u)\right]^s(1 - \mathcal{L}\{f_X\}(\gamma_u))$ then this difference is non-negative. 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. We shall say that i dominates u: $i \succ u$. If the jobs of the instance can be ordered by this dominance criterion (up to a change of indexes, all jobs verify $i \succ j$ if and only if $i < j$), then processing jobs in this order is optimal. $\square$

Let us consider for example an instance with three jobs to process. The weights are all equal to 1, processing times are exponential with mean 1 and impatiences are exponential with mean $\gamma_1 = 3/7$, $\gamma_2 = 1/4$ and $\gamma_3 = 1/9$. For this instance, the optimal schedule is to process the jobs in the order of their indexes, since in this case, we have indeed $1 \succ 2 \succ 3$ (see Appendix C for details).

6.5 Problem 6: $X_j \sim F_X, D_j \sim \exp(\gamma_j)$, 2 classes of jobs

In what follows, we consider a special case of Problem 5 where there is only two types of jobs. Even for this very simple case, the optimal policy is not trivial as we shall see.

Assume that there are n_1 jobs from class 1 (with weight w_1 and impatient rate γ_1) and n_2 jobs from class 2 (with weight w_2 and impatient rate γ_2). The total number of jobs is $n = n_1+n_2$. All processing times are i.i.d. ($X_j \sim F_X$). Let $\alpha_j = \mathcal{L}\{f_X\}(\gamma_j)$ for $j \in \{1, 2\}$ and assume that $\alpha_1 \geq \alpha_2$, without loss of generality.

The quantity $(C(S') - C(S))$ is non-negative in Equation (2) if

$$w_1\alpha_1^s(1 - \alpha_1) \geq w_2\alpha_2^s(1 - \alpha_2)$$

When $\alpha_1 = \alpha_2$, $(C(S') - C(S))$ 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

$$s \geq t_{IBS} = \ln\left(\frac{w_2(1 - \alpha_2)}{w_1(1 - \alpha_1)}\right)\left(\ln\left(\frac{\alpha_1}{\alpha_2}\right)\right)^{-1} \tag{3}$$

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.

To illustrate how to use formula (3), consider an instance with $\alpha_1 > \alpha_2$ and $n_1 = 4$ jobs of class 1 and $n_2 = 6$ jobs of class 2. If $t_{IBS} = -2.5$, the optimal static schedule is $(1, 1, 1, 1, 2, 2, 2, 2, 2, 2)$. If $t_{IBS} = 3.2$, the optimal static schedule is $(2, 2, 2, 1, 1, 1, 1, 2, 2, 2)$. If $t_{IBS} = 100.3$, the optimal static schedule is $(2, 2, 2, 2, 2, 2, 1, 1, 1, 1)$

A similar analysis can be made for the IES counterpart of this problem (see Appendix D.3).

7 Conclusion and future research

In this paper we study a stochastic scheduling problem with IBS. The impatience of jobs is modelled by stochastic due dates, but unlike the traditional literature on scheduling with due dates, a job is considered to be on time if its execution begins before its due date. The optimal schedule is searched among the class of static list scheduling policies. All decisions have to be taken at time zero and jobs have to be processed in the predefined order, whether they are tardy or not. We show that the deterministic IBS problem reduces to the deterministic IES problems. However, the stochastic IBS problem and the stochastic IES problem can not be reduced to each other. We derive the optimal schedule for the IBS problem for different processing time and due date distributions. Finally, we provide some additional results for the problem with IES.

Future research could focus on extending our results for more general processing time and due date distributions. In particular, we provide results when either processing times or due dates distributions are i.i.d. but not when they both follow different probability distributions. It could also be interesting to consider a problem mixing IES and IBS. Another avenue for research is to try to extend some of our results to the dynamic problem of scheduling with impatience and reneging, because none of the references we found deals with the dynamic problem with IBS, except for Atar et al.[2] who consider an asymptotic regime. Maybe some of our insights are still valid and may lead to some interesting heuristics for the dynamic problem and for which performance analysis should be done. But finding optimal results might not be achievable keeping the same model, regarding the quite restrictive results found by Down and al. [5] on the IBS counterpart of this problem.

References

[1] N.T. Argon, S. Ziya, and 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*, 22:301–332, 2008.

[2] R. Atar, C. Giat, and N. Shimkin. The $c\mu/\theta$ rule for many-server queues with abandonment. *Operations Research*, 58:1427–1439, 2010.

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

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

[5] D.G. Down, G. Koole, and M.E. Lewis. Dynamic control of a single server system with abandonments. *To appear in: Queueing Systems*.

[6] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[7] R.L. Graham, E.L. Lawler, J.K. Lenstra, and A.H.G. Rimooy Kan. Optimization and approximation in deterministic sequencing and scheduling : a survey. *Annals of Discrete Mathematics*, 5:287–326, 1979.

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

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

[10] P.K.F. Kuhfittig. *Introduction to the Laplace Transform*. PLENUM Publishing Co., N.Y., 1978.

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

[12] A. Movaghar. On queueing with customer impatience until the end of service. *Stochastic Models*, 22(1):149–173, 2006.

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

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

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

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

[17] S. Zeltyn and A. Mandelbaum. *Call centers with impatient customers: exact analysis and many-server asymptotics of the $M/M/n+G$ queue*. PhD thesis, Israel Institute of Technology, 2004.

A Arbitrarily bad schedules

We first provide an instance where there are n jobs to be executed. All of them are available at time zero, have unit processing times and weights ($p_j = 1$ and $w_j = 1$), and for all $j \in \{1, 2, \ldots, n\}$, $d_j = j - 0.5$. The optimal IBS solution is to schedule the jobs in increasing order of the indexes and all of them are on time. Using the same schedule for the IES problem leads to a cost equal to n, since all jobs are late. The optimal IES schedule is the same schedule except for job 1 which is the last to be processed, leading to a cost equal to 1 (job 1 is the only one to be late). The ratio of the costs, n, goes to infinity with the number of jobs in the instance. Consequently we can state that an IBS optimal schedule can perform arbitrarily bad for the IES problem.

We now provide an instance with only two jobs such that $d_1 < p_1$, $d_1 < p_2$, $d_2 > p_1$, $d_2 > p_2$ and $d_2 < p_1 + p_2$. The optimal IES schedule is to process job 2 and then job 1 leading to a cost $C_{IES} = w_1$, since only job 2 can end on time. The optimal IBS schedule is to process job 1 then job 2, in the IBS problem both jobs end on time. But in the IES problem, processing the jobs in this order implies that both are late. Hence, the cost for this schedule is $C_{IBS} = w_1 + w_2$. The rate between the two costs for these schedule $C_{IBS}/C_{IES} = 1 + w_1/w_2$ goes to infinity with w_1/w_2. This little example shows that an IES optimal schedule can perform arbitrarily bad for the IBS problem too.

B Counter-example for Problem 1

Let us consider for example a problem with two jobs and where $X_j \sim \exp(\mu_j)$ and $D_j \sim \exp(\gamma)$. The parameters of the instance are $w_1 = 1$, $\mu_1 = 3$, $w_2 = 2$, $\mu_2 = 1$ and $\gamma = 2$. For the IES problem, it is optimal to process job 1 first because $w_1\mu_1 > w_2\mu_2$. But if we evaluate both possible schedules for the IBS problem, we obtain $C(\{1,2\}) = w_2\gamma/(\gamma + \mu_1)$ and $C(\{2,1\}) = w_1\gamma/(\gamma + \mu_2)$. Replacing the parameters by their values, we have $C(\{1,2\}) = 4/5 > C(\{2,1\}) = 2/3$. The optimal IBS schedule is to process job 2 first. This simple counter-example proves that processing jobs in non-increasing order of $w_j\mu_j$ is not optimal for the IBS counterpart of the problem.

C Details of the example in Section 6.4

As in the example, we consider an instance with three jobs to process. The weights are all equal to 1, processing-times are exponential with mean 1 and impatiences are exponential with mean $\gamma_1 = 3/7$, $\gamma_2 = 1/4$ and $\gamma_3 = 1/9$.

Let us prove that $1 \succ 2 \succ 3$. The criterion to evaluate is $[\mathcal{L}\{f_X\}(\gamma_i)]^s\,(1-\mathcal{L}\{f_X\}(\gamma_i))$. Since $f_X(t) = e^{-t}$, we have $\mathcal{L}\{f_X\}(\gamma_j) = \frac{1}{1+\gamma_j}$. For $s = 0$, $\delta_1'(0) = 3/7$, $\delta_2'(0) = 1/4$ and $\delta_3'(0) = 1/9$: the order holds. For $s = 1$ $\delta_1'(1) = 0.3$, $\delta_2'(1) = 0.2$ and $\delta_3'(1) = 0.1$: the order holds again. We proved that the order holds for all s, consequently $1 \succ 2 \succ 3$ is proved as well.

We showed that this instance respects the assumptions of the theorem, now we will evaluate the objective function on all possible schedule with these three tasks. The following expression of the objective function is obtained using the same arguments as for Equation (1) $C(S) = \sum_{i=2}^{n} w_i \int_{t=0}^{+\infty} (1 - F_{Z_{i-1}}(t)) f_{D_i}(t)dt$. Replacing $f_{D_i}(t)$ by $\gamma_i e^{-\gamma_i t}$ leads to

$$\begin{aligned}
C(S) &= \sum_{i=2}^{n} w_i \gamma_i \int_{t=0}^{+\infty} (1 - F_{Z_{i-1}}(t)) e^{-\gamma_i t} dt \\
&= \sum_{i=2}^{n} w_i \left(1 - \mathcal{L}\{f_{Z_{i-1}}\}(\gamma_i)\right) \\
&= \sum_{i=2}^{n} w_i \left(1 - \prod_{k=1}^{i-1} \mathcal{L}\{f_X\}(\gamma_i)\right) \\
&= \sum_{i=2}^{n} w_i \left(1 - [\mathcal{L}\{f_X\}(\gamma_i)]^{i-1}\right).
\end{aligned}$$

Replacing $\mathcal{L}\{f_X\}(\gamma_i)$ by $\frac{1}{1+\gamma_i}$ and w_i by 1, we obtain

$$C(S) = \sum_{i=2}^{n} \left(1 - \left(\frac{1}{1+\gamma_i}\right)^{i-1}\right).$$

With this formula we can compute the value of the objective function for the 6 possible schedules and we obtain the following values:

S	$C(S)$
$\{1,2,3\}$	0.39
$\{1,3,2\}$	0.46
$\{2,1,3\}$	0.49
$\{2,3,1\}$	0.61
$\{3,1,2\}$	0.66
$\{3,2,1\}$	0.71

Table 2: Value of the objective function on all possible schedules

The conclusion is that the first sequence, processing the tasks in the order of the indexes, is indeed optimal.

D Complementary results for the IES problem

D.1 Preliminaries

In order to prove the optimality of IES schedules for Problems 3 and 6, we again use a pairwise interchange argument between schedule

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

The objective function of schedule S is

$$C_{IES}(S) = \sum_{j=1}^{n} w_j \mathbb{P}(X_1 + \ldots + X_j > D_j).$$

Let us introduce the random variable $Z_s = \sum_{j=1}^{s} X_j$, then the objective functions of the two schedules are

$$\begin{aligned}
C_{IES}(S) = &\sum_{j=1}^{s} w_j \mathbb{P}(X_1 + \ldots + X_j > D_j) \\
&+ w_i \mathbb{P}(Z_s + X_i > D_i) + w_u \mathbb{P}(Z_s + X_i + X_u > D_u) \\
&+ \sum_{j=s+3}^{n} w_j \mathbb{P}(X_1 + \ldots + X_i + X_u + \ldots + X_j > D_j) \quad \text{and} \\
C_{IES}(S') = &\sum_{j=1}^{s} w_j \mathbb{P}(X_1 + \ldots + X_j > D_j) \\
&+ w_u \mathbb{P}(Z_s + X_u > D_u) + w_i \mathbb{P}(Z_s + X_u + X_i > D_i) \\
&+ \sum_{j=2}^{s+3} w_j \mathbb{P}(X_1 + \ldots + X_u + X_i + \ldots + X_j > D_j).
\end{aligned}$$

By making the difference between the two schedules, we have

$$\begin{aligned}
C_{IES}(S') - C_{IES}(S) = &\; w_u \mathbb{P}(Z_s + X_u > D_u) + w_i \mathbb{P}(Z_s + X_u + X_i > D_i) \\
&- w_i \mathbb{P}(Z_s + X_i > D_i) - w_u \mathbb{P}(Z_s + X_i + X_u > D_u)
\end{aligned}$$

and using distribution and density probability functions leads to the difference

$$\begin{aligned}
C_{IES}(S') - C_{IES}(S) = &\; w_i \int_{t=0}^{+\infty} (F_{Z_s+X_i}(t) - F_{Z_s+X_i+X_u}(t)) f_{D_i}(t)\, dt \\
&- w_u \int_{t=0}^{+\infty} (F_{Z_s+X_u}(t) - F_{Z_s+X_i+X_u}(t)) f_{D_u}(t)\, dt.
\end{aligned} \tag{4}$$

D.2 Problem 3 with IES: $X_j \sim F_{X_j}, D_j \sim F_D$

Theorem 5. *If the processing time of job j is a random variable with general distribution function F_{X_j} and due dates are i.i.d. with general distribution function 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 problem of minimizing the expected weighted number of late jobs with IES.*

Proof. Given that the due-dates are i.i.d. ($D_i \sim F_D$), Equation (4) can be simplified as

$$C_{IES}(S') - C_{IES}(S) = \int_{t=0}^{+\infty} \left[w_i F_{Z_s+X_i}(t) - w_u F_{Z_s+X_u}(t) - (w_i - w_u) F_{Z_s+X_i+X_u}(t) \right] f_D(t) dt.$$

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 $0 \leq z \leq t$, $\mathbb{P}(X_i \leq t - z \mid Z_s = z) \geq \mathbb{P}(X_u \leq t - z \mid Z_s = z)$ because D and all X_j are independent. By using the formula of conditional probability we obtain $\mathbb{P}(Z_s + X_i \leq t) = \mathbb{P}(X_i \leq t - z \mid Z_s = z)\mathbb{P}(Z_s = z) \geq \mathbb{P}(X_u \leq t - z \mid Z_s = z)\mathbb{P}(Z_s = z) = \mathbb{P}(Z_s + X_u \leq t)$. As a direct consequence, $F_{Z_s+X_i}(t) \geq F_{Z_s+X_u}(t)\ \forall t \geq 0$. This relation implies a lower bound on the difference between the objective functions,

$$\begin{aligned} C_{IES}(S') - C_{IES}(S) &\geq \int_{t=0}^{+\infty} \left[(w_i - w_u)F_{Z_s+X_u}(t) - (w_i - w_u)F_{Z_s+X_i+X_u}(t)\right] f_D(t)dt \\ &\geq \int_{t=0}^{+\infty} \left[(w_i - w_u)(F_{Z_s+X_u}(t) - F_{Z_s+X_i+X_u}(t))\right] f_D(t)dt. \end{aligned}$$

Given that all possible values for X_i are non-negative, the following inequality holds: $F_{Z_s+X_i+X_u}(t) = \mathbb{P}(Z_s + X_i + X_u \leq t) \leq \mathbb{P}(Z_s + X_u \leq t) = F_{Z_s+X_u}(t)$ for all $t \geq 0$. And finally

$$C_{IES}(S') - C_{IES}(S) \geq (w_i - w_u) \int_{t=0}^{+\infty} f_D(t)dt = w_i - w_u,$$

so we can conclude that the difference is non-negative if $w_i \geq w_u$. If the tasks can be ordered such that their processing-times are increasing in the sense of stochastic order and if among this order, the weights are decreasing, then this schedule is optimal. □

D.3 Problem 6 with IES: $X_j \sim F_X, D_j \sim \exp(\gamma_j)$, 2 classes of jobs

We consider the same setting as in Section 6.5 except that impatience is of IES type. In order to find an optimal schedule, we start from Equation (4) (Section D.1) replacing $f_{D_j}(t)$ by $\gamma_j \exp(-\gamma_j t)$ the difference becomes

$$\begin{aligned} C_{IES}(S') - C_{IES}(S) &= w_i \int_{t=0}^{+\infty} (F_{Z_s+X_i}(t) - F_{Z_s+X_i+X_u}(t))\gamma_i e^{-\gamma_i t}\, dt \\ &\quad - w_u \int_{t=0}^{+\infty} (F_{Z_s+X_u}(t) - F_{Z_s+X_i+X_u}(t))\gamma_u e^{-\gamma_u t}\, dt \\ &= w_i \left(\mathcal{L}\{f_{Z_s+X_i}\}(\gamma_i) - \mathcal{L}\{f_{Z_s+X_i+X_u}\}(\gamma_i)\right) \\ &\quad - w_u \left(\mathcal{L}\{f_{Z_s+X_u}\}(\gamma_u) - \mathcal{L}\{f_{Z_s+X_i+X_u}\}(\gamma_u)\right). \end{aligned}$$

And taking advantage of the fact that processing-times are i.i.d. and of the convolution property of Laplace transforms leads to

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

One can notice that it is quite the same expression as Equation (2), except for the exponent which is here equal to $s + 1$. Consequently, the same arguments can be used as in Section

6.5 and the optimal schedule is a threshold policy for which the threshold can be computed $t_{IES} = \ln\left(\frac{w_2(1-\alpha_2)}{w_1(1-\alpha_1)}\right)\left(\ln\left(\frac{\alpha_1}{\alpha_2}\right)\right)^{-1}$.

Without surprise, the two thresholds verify $t_{IBS} = t_{IES} + 1$. The optimal position of job j in the IBS problem arises one position after the optimal position in the IES problem, because the duration of the execution of job j is not taken into account in the IBS problem.

Les cahiers Leibniz ont pour vocation la diffusion des rapports de recherche, des séminaires ou des projets de publication sur des problèmes liés au mathématiques discrètes.

Responsables de la publication : Nadia Brauner et András Sebő
ISSN : 1298-020X -