Minimizing the expected weighted number of tardy jobs with non-identically distributed processing times and due dates

Frank G. Forst
Loyola University Chicago

Abstract

In this paper we are concerned with finding a job sequence which minimizes the expected weighted number of tardy jobs on one machine. Three sufficient optimality condition are derived when both the job processing times and the job due dates are independent, non-identically distributed random variables. We then derive more specific optimality conditions for the special case of normally distributed job processing times. A numerical example is also provided.

Key words: stochastic job sequencing, one machine, due dates, tardy jobs, normal probability distribution
INTRODUCTION

We consider a single machine and a set of n jobs, $J = (1, 2, \ldots, n)$, to be processed. All the jobs are available at time zero. Both the job processing times $X_k$ and the job due dates $D_k$ ($k = 1, 2, \ldots, n$) are independent, non-identically distributed random variables. The $X_k$ and $D_k$ are also independent. Each job $k$ has a constant weight $w_k > 0$. Let $C_k$ represent the completion time of job $k$ on the single machine. When $C_k > D_k$ the decision-maker incurs the penalty $w_k$, and when $C_k \leq D_k$ the decision-maker incurs no penalty for job $k$.

The objective is to obtain a job sequences $S = (j_1, j_2, \ldots, j_n)$, which is a permutation of $J$, that minimizes the expected weighted number of tardy jobs (i.e., jobs that are completed after their due date) $E[N(S)]$. Job preemption is not allowed, and the jobs must be sequenced before processing of the jobs begins. Thus we only consider static job sequences in this paper. Define $U_k = 1$ when job $k$ is tardy and $U_k = 0$ when job $k$ is not tardy. Then the objective may be written as

$$\text{Minimize } E(N(S)) = E \left[ \sum_{k=1}^{n} w_k U_k(S) \right] \quad (1.1)$$

SUMMARY OF RELATED LITERATURE

Karp (1972) has investigated the complexity of the deterministic version of the single-machine job tardiness problem. He has proven that the problem of minimizing the weighted number of tardy jobs is binary NP-hard, which means that an efficient polynomial-bounded algorithm probably does not exist. Several results have appeared in the literature for the stochastic problem of minimizing the expected weighted number of tardy jobs on one machine.

Pinedo (1983) analyzed the case in which the job processing times are independent, exponentially distributed random variables with rates $\lambda_k$ ($k = 1, \ldots, n$) and the due dates are independent, identically distributed (i.i.d.) random variables. Pinedo proved that the optimal sequence is obtained by sequencing the jobs in decreasing order of $\lambda_k w_k$.

Boxma and Forst (1986) demonstrated that in the case of i.i.d. random due dates and i.i.d. random processing times, an optimal job sequence is obtained by sequencing the jobs in decreasing order of the weights. When all due dates are i.i.d. random variables and all weights are identical, if the jobs can be sequenced according to an increasing stochastic ordering, this ordering yields an optimal job sequence. Random variable $X$ is called stochastically smaller than random variable $Y$, written $X \leq_{st} Y$, if $\Pr(X > t) \leq \Pr(Y > t)$ for all $t$. Finally, Boxma and Forst (1986) derived a simple, optimal job sequencing rule for the case when the job processing times are independent random variables and the due dates follow i.i.d. exponential distributions.

De, Ghosh, and Wells (1991) analyzed the case where the job processing times
are independent random variables and the jobs have a common, exponentially
distributed due date. They derived two sufficient conditions for the existence of a job
sequence which stochastically minimizes the weighted number of tardy jobs. They
also determined a simple sequencing rule which stochastically minimizes the number
of tardy jobs.

The case of normally distributed processing times has also been addressed in the
job sequencing literature.

Sarin, Erel, and (1991) studied the single machine case in which the job
processing times are independent, normally distributed random variables, the
variances of the processing times are proportional to their means, and the jobs have a
common, constant due date. For this situation, they proved that a necessary, but not
sufficient, condition for a job sequence to be optimal is that it have either a W-shape
or V-shape with respect to the mean processing times.

Soroush and Fredendal (1994) analyzed the single machine problem with
earliness and tardiness costs where the job processing times are independent,
normally distributed random variables having distinct means and variances, while the
due dates are distinct and constant. They proved that the total expected earliness and
tardiness cost of a job sequence is a nonlinear, increasing function of the expected
values and standard deviations of the job completion times for the jobs in that
sequence. The authors then derived lower and upper bounds for the total expected
E/T cost of a job sequence. Next they developed three heuristics to provide three
candidates for the optimal job sequence. The authors conducted fives-, six-, and
seven-job experiments which demonstrated that one of the three candidates was often,
but not always, the optimal job sequence. Thus, none of the three heuristics was
always successful in identifying the optimal job sequence.

Cai and Zhou (1997) sought to identify a job sequence which minimizes the
expected weighted combination of earliness, tardiness, and flow time costs when the
job processing rimes are independent and normally distributed random variables with
variances that are proportionional to the means. The job due dates are random variables
following a common probability distribution. The authors show that an optimal job
sequence must be V-shaped with respect to the mean processing times.

Jang (2002) studied the problem of minimizing the expected number of tardy jobs
when jobs having normally distributed processing times and deterministic due dates
arrive randomly. Jang develops a heuristic, based on a myopically optimal solution,
from which he obtains a simple and robust dynamic scheduling policy. But it is not
always optimal.

Recently Seo, Klein, and Jang (2005) analyzed the problem of minimizing the
expected number of tardy jobs have normally distributed processing times and a
common, constant due date. They transformed the original stochastic problem into an
equivalent non-linear integer programming model, Model A1. The authors then
compared the performance of three other models (A2, A3, and A4) with that of Model
A1. In experiments involving 20 to 50 jobs, Model A2 generated solutions with average
errors ranging from 2.2 percent to 3.8 percent compared to Model A1’s optimal solutions,
but with significant computational time savings.

To our knowledge, the case in which both the job processing times and the due
dates are non-identically distributed random variables, and the objective is to find a job
sequence which minimizes the expected weighted number of tardy jobs, has not yet been addressed in the literature. This case is considered in the present paper.

RESULTS

The strategy is to analyze the effect on the expected weighted number of tardy jobs when two adjacent jobs are interchanged. Consider the following two job sequences:

\[ S: j_1, j_2, \ldots, j_{i-1}, j_i, j_{i+1}, \ldots, j_n \]

\[ S*: j_1, j_2, \ldots, j_{i-1}, j_{i+2}, j_i, j_{i+1}, \ldots, j_n, \]

Where \( S* \) is obtained from \( S \) by interchanging jobs \( j_i \) and \( j_{i+1} \) for some \( 1 \leq i \leq n-1 \).

Now we derive sufficient conditions for job sequence \( S \) to be preferred to job sequence \( S* \); that is, for \( E(N(S)) \leq E(N(S*)) \).

THEOREM 3.1 \( E(N(S)) \leq E(N(S*)) \); that is, job \( j_i \) should precede job \( j_{i+1} \), if

\[ w_i \leq w_{i+1}, \quad (3.1) \]
\[ D_i \leq s D_{i+1}, \quad (3.2) \]
\[ X_i - D_i \leq s X_{i+1} - D_{i+1}. \quad (3.3) \]

These three sufficient conditions are transitive.

PROOF. Since all the job processing times and due dates are independent, the expected weighted numbers of tardy jobs in the sets \( (j_1, \ldots, j_{i-1}) \) and \( (j_{i+2}, \ldots, j_n) \) are the same for job sequences \( S \) and \( S* \). Thus we need only consider the expected weighted number of tardy jobs in the set \( (j_i, j_{i+1}) \) for \( S \) and \( S* \). Let \( C*_{i} \) and \( C*_{i+1} \) represent the completion times of job \( j_i \) and job \( j_{i+1} \), respectively, for sequence \( S* \). Then

\[ E(N(S)) - E(N(S*)) = w_i \Pr(C_i > D_i) + w_{i+1} \Pr(C_{i+1} > D_{i+1}) - w_{i+1} \Pr(C*_{i+1} > D_{i+1}) - w_i \Pr(C*_{i} > D_i). \]

Thus \( E(N(S)) - E(N(S*)) \leq 0 \) if and only if

\[ w_i \Pr(C_i > D_i) - w_{i+1} \Pr(C_{i+1} > D_{i+1}) \leq w_{i+1} \Pr(C*_{i+1} > D_{i+1}) - w_i \Pr(C*_{i} > D_i). \quad (3.4) \]

Two sufficient conditions for (3.4) to hold are

\[ w_i \leq w_{i+1}, \quad (3.1) \]
\[ \Pr(C*_{i} > D_i) - \Pr(C_{i+1} > D_{i+1}) \geq \Pr(C_{i+1} > D_{i+1}) - \Pr(C*_{i} > D_i). \quad (3.5) \]

Writing the job completion times in terms of the job processing times, and letting \( X=X_1+X_2+\ldots+X_{i-1} \), we may express inequality (3.5) as

\[ \Pr(X+X_i+X_{i+1} > D_i) - \Pr(X+X_{i+1} > D_{i+1}) \geq \Pr(X+X_{i+1} > D_{i+1}) - \Pr(X+X_i > D_i). \quad (3.6) \]

Since the job processing times are independent of the job due dates, two sufficient conditions for inequality (3.6) to hold are

\[ D_i \leq s D_{i+1}, \quad (3.2) \]
\[ \Pr(X+X_i > 0) \leq \Pr(X+X_{i+1} > 0). \quad (3.7) \]

Now a sufficient condition for (3.7) to hold is

\[ X_i - D_i \leq s X_{i+1} - D_{i+1}. \quad (3.3) \]

Thus conditions (3.1), (3.2), and (3.3) are sufficient for \( E(N(S)) \leq E(N(S*)) \).

Condition (3.1) is transitive since \( w_i \geq w_{i+1} \) and \( w_{i+1} \geq w_{i+2} \) imply \( w_i \geq w_{i+2} \) for three adjacent jobs \( j_i, j_{i+1}, \) and \( j_{i+2} \). By similar argument, conditions (3.2) and (3.3) are also transitive.

Q.E.D.
Theorem 3.1 implies that “fast” jobs with high penalty costs and tight due dates should be processed first. If for all pairs of jobs \( j_i \) and \( j_{i+1} \) the three conditions of Theorem 3.1 or their contrapositives hold, we can determine a complete and optimal job sequence that minimizes the expected weighted number of tardy jobs. However, an optimal job sequence need not satisfy conditions (3.1), (3.2), and (3.3) for all pairs of jobs.

Theorem 3.1 also implies that when the due dates are independent, identically distributed random variables and all the weights are the same, an optimal sequence is obtained if the jobs can be sequenced in increasing stochastic order of their processing times. This result was first proven by Boxma and Forst (1986). Note that if the due dates are independent, exponentially distributed random variables with rates \( \lambda_k \) \((k=1, 2, \ldots, n)\), then condition (3.2) may be written \( \lambda_i \geq \lambda_{i+1} \).

Now, we analyze the special case in which the job processing times are independent, normally distributed random variables. Normally distributed job processing times are justified in practice when each job consists of many elementary tasks, each of which have random processing times.

To largely avoid the problem of negative values for random variables, we will assume \( \mu > 3\sigma \).

**COROLLARY 3.1.** Suppose the job processing times \( X_k \) are independent, normally distributed random variables with mean \( \mu_k \) and common standard deviation \( \sigma \), for \( 1 \leq k \leq n \). Then job \( j_i \) should precede job \( j_{i+1} \) if

\[
\begin{align*}
    w_i &\geq w_{i+1} \quad (3.1) \\
    D_i &\leq_{st} D_{i+1}, \text{ and} \quad (3.2) \\
    \mu_i - D_i &\leq_{st} \mu_{i+1} - D_{i+1} \quad (3.3)
\end{align*}
\]

These three sufficient conditions are transitive.

**PROOF.** The proof of Theorem 3.1 indicates that job \( j_i \) should precede job \( j_{i+1} \) if

\[
\begin{align*}
    w_i &\geq w_{i+1} \quad (3.1) \\
    D_i &\leq_{st} D_{i+1}, \text{ and} \quad (3.2) \\
    \Pr(X_1+\ldots+X_{i-1}+X_i>X_i+\ldots+X_{i+1}+X_{i+1}>D_i) &\leq \Pr(X_1+\ldots+X_{i-1}+X_i+X_{i+1}>D_i+1). \quad (3.7)
\end{align*}
\]

Since the job processing times are independent and normally distributed, \((X_1+\ldots+X_{i-1}+X_i)\) also follows a normal distribution with parameters \((\mu_1+\ldots+\mu_i+\sigma\sqrt{i})\). Similarly, \((X_1+\ldots+X_{i+1}+X_{i+1})\) follows a normal distribution with parameters \((u_1+\ldots+\mu_{i-1}+\sigma\sqrt{i})\).

Then inequality (3.7) holds if and only if

\[
D_i - (\mu_1+\ldots+\mu_{i-1}+\mu_i) \geq_{st} D_{i+1} - (\mu_1+\ldots+\mu_{i-1}+\mu_{i+1}) \quad (3.9)
\]

Now inequality (3.9) holds whenever

\[
D_i - \mu_i \geq_{st} D_{i+1} - \mu_{i+1}; \text{ that is, } \mu_i - D_i \leq_{st} \mu_{i+1} - D_{i+1}.
\]

Thus conditions (3.1), (3.2), and (3.3) are sufficient for job \( j_i \) to precede job \( j_{i+1} \); that is, for \( E(N(S)) \leq E(N(S^*)) \).

The transitivity of conditions (3.1) and (3.2) is readily established. Condition (3.3) is transitive because \( \mu_i - D_i \leq_{st} \mu_{i+1} - D_{i+1} \) and \( \mu_{i+1} - D_{i+1} \leq_{st} \mu_{i+2} - D_{i+2} \) imply
\[ \mu_i - d_i \leq \mu_{i+2} - d_{i+2} \] for three adjacent jobs \( j_i, j_{i+1}, \) and \( j_{i+2} \). Q.E.D.

**COROLLARY 3.2** Suppose the job processing times \( X_k \) are independent normal with parameters \((\mu_k, \sigma)\), and the due dates are distinct constants \( d_k \), for \( 1 \leq k \leq n \). Then job \( j_i \) should precede job \( j_{i+1} \) if

\[
\begin{align*}
 w_i & \geq w_{i+1} \\
 d_i & \leq d_{i+1}, \text{ and} \\
 \mu_i - d_i & \leq \mu_{i+1} - d_{i+1}
\end{align*}
\]  

These sufficient conditions are transitive.

**PROOF.** Since the due dates are all constants \( d_k \), condition (3.10) follows immediately from condition (3.2) of Corollary 3.1, while condition (3.11) follows directly from condition (3.8) of Corollary 3.1. Transitivity is readily established. Q.E.D.

**A NUMERICAL EXAMPLE**

A small numerical example involving five jobs to illustrate Corollary 3.2 will now be discussed. In this example, the job processing times \( X_k \) are independent, normally distributed random variables with means \( \mu_k \) and common standard deviation \( \sigma \), for \( 1 \leq k \leq n \). The due dates are distinct constants \( d_k \). Let \( \sigma = 3 \).

<table>
<thead>
<tr>
<th>Job</th>
<th>( w_k )</th>
<th>( \mu_k )</th>
<th>( d_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>35</td>
<td>30</td>
</tr>
</tbody>
</table>

Note that \( \mu_k > 3\sigma = 3(3) = 9 \) for all \( k \).

Also note \( w_3 > w_2 > w_4 > w_1 > w_5 \), \( d_3 < d_2 < d_4 < d_1 < d_5 \), and
\[
\begin{align*}
 \mu_3 - d_3 & = 0 < \mu_2 - d_2 = 2 \leq \mu_4 - d_4 = 2 < \mu_1 - d_1 = 3 < \mu_5 - d_5 = 5.
\end{align*}
\]

Thus by Corollary 3.2, an optimal job sequence \( S = (3,2,4,1,5) \).

**FUTURE RESEARCH**

The case is currently being analyzed in which both the mean job processing times \( \mu_k \) and the standard deviations \( \sigma_k \) are distinct. The plan is to study the special case in which the due dates \( D_k \) are independent, exponentially distributed random variables.
REFERENCES


