Operations Research Letters 10 (1991) 467–472
North-Holland

Proof of a conjecture of Schrage about the completion time variance problem

Nicholas G. Hall

Faculty of Management Sciences, The Ohio State University, Columbus, Ohio 43210-1399, USA

Wieslaw Kubiak

Faculty of Business Administration, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1B 3X5

The problem of minimizing the completion time variance of n jobs on a single machine has been studied by several authors. We prove the correctness of a 1975 conjecture due to L. Schrage about the position of the third longest job in an optimal schedule.

scheduling; completion time variance

1. Introduction and formulation

The problem of scheduling n jobs nonpreemptively on a single machine, so as to minimize the variance of job completion times, was first proposed by Merten and Muller [3]. They motivate this problem as a way of modeling file organization decisions in which it is important to provide uniform response times to users. This objective is particularly important for on-line systems. Schrage [4] describes conditions which are necessary for optimality, as well as providing the conjecture which is discussed below.

Eilon and Chowdhury [1] prove the important property that an optimal schedule must be V-shaped. That is, the jobs must be ordered by nonincreasing processing time if placed before the shortest job, and by nondecreasing processing time if placed after the shortest job. Some heuristic approaches are also proposed and tested in that paper. Kanet [2] motivates the completion time variance problem as being applicable to any service or manufacturing setting where it is desirable to provide jobs or customers with approximately the same treatment. Two ways of dealing with the apparent difficulty of this performance measure are described, namely a proxy measure using total absolute differences in completion times, and a computationally effective heuristic. Vani and Raghavachari [5], by using job interchange arguments, are able to provide a list of possible optimal schedules for small problems. For example, when $n = 7$, there are only five schedules which need to be considered. A heuristic approach, competitive with those of Eilon and Chowdhury [1] and of Kanet [2], is also described and tested.

In order to formalize the problem considered here, we begin with several definitions. Let

p_j = integer processing time of job j, $j = 1, \dots, n$,

C_j = completion time of job j in schedule σ, $j = 1, \dots, n$,

$\overline{C} = \sum_{j=1}^{n} C_j/n$,

Π = the set of all possible schedules.

The problem which we consider is the minimization of completion time variance, or

$$\text{(CTV)} \qquad \min_{\sigma \in \Pi} z(\sigma) = \sum_{j=1}^{n} \left(C_j - \overline{C}\right)^2/n.$$

2. Schrage's conjecture and preliminary results

We assume throughout that the jobs are numbered so that $p_1 \leq \cdots \leq p_n$. A well known conjecture by Schrage [4] is that there exists, for every instance of CTV, an optimal schedule of the form $(n, n-2, n-3, \ldots, n-1)$. Kanet [2] provides a counterexample in the form of an 8 job instance, for which no optimal schedule is of this form. A weaker form of Schrage's conjecture which has remained open, however, is that for every instance there is an optimal schedule of the form $(n, n-2, \ldots, n-1)$. Vani and Raghavachari [5] prove that this weaker form of Schrage's conjecture is correct for any instance of CTV with $n \leq 18$. Below, we prove that this result holds for any problem size. We begin with some preliminary results.

Lemma 1.

$$\sum_{C_j \leq \bar{C}} |C_j - \bar{C}| = \sum_{C_j > \bar{C}} |C_j - \bar{C}|.$$

Proof. Follows directly from the definition of $\bar{C}$. □

Lemma 2. *For any instance of* CTV, *there exists an optimal schedule of the form* $(n, \ldots, n-1)$.

Proof. Schrage [4] proves that, in any optimal schedule for CTV, the longest job is scheduled first. He also proves that the order of the last $n-1$ jobs can be reversed without changing the variance. From Eilon and Chowdhury [1], every optimal schedule is V-shaped. The combination of these last two results completes the proof. □

Lemma 3 (Vani and Raghavachari [5]). *Let* $f(d) = \sum_{j=1}^{n} (C_j - d)^2$ *denote the sum of squared deviations around any point in time d in the schedule. Then* $f(d)$ *is minimized at* $d = \bar{C}$.

Proof.

$$f(d) = \sum_{j=1}^{n} (C_j - d)^2 = \sum_{j=1}^{n} C_j^2 - 2d \sum_{j=1}^{n} C_j + nd^2 \Rightarrow f'(d) = -2 \sum_{j=1}^{n} C_j + 2nd = 0,$$

or $d = \sum_{j=1}^{n} C_j / n = \bar{C}$ at a minimum point (where $f''(d) > 0$). □

3. Proof of the conjecture

Theorem 1. *For any instance of* CTV, *there exists an optimal schedule of the form* $(n, n-2, \ldots, n-1)$.

Proof. From Lemma 2 and the V-shaped property of Eilon and Chowdhury [1], we need only show that the assumption that job $n-2$ immediately precedes job $n-1$ in *every* optimal schedule for some instance I with $n \geq 2$ leads to a contradiction. Let us consider an optimal schedule for I with mean flow time $\bar{C}$. Let $p_n = \alpha_0$, $p_{n-1} = \beta_1$, and $p_{n-2} = \beta_2$ denote the processing times of the three longest jobs, respectively. Let $\alpha_1 < \beta_2$ denote the processing time of the job scheduled immediately after job n. This change of notation will simplify the derivation of our results. Where no ambiguity will arise, we may refer to a job by its processing time.

We consider several cases in which the positions of the jobs are changed, and show that the resulting sum of squared deviations around some point in the schedule is less than or equal to the previous sum of squared deviations around $\bar{C}$. It then follows from Lemma 3 that the sum of squared deviations around the new $\bar{C}$ has not increased. We assume for now that job α_1 finishes at or before $\bar{C}$, and that job β_2 starts at or after $\bar{C}$.

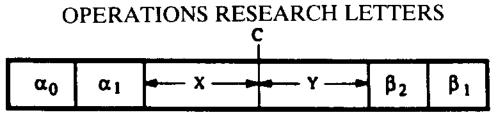

Fig. 1. Schedule for Theorem 1, Case 1

Case 1. $2X \leq 4Y + \alpha_1 + 2\beta_1 + \beta_2$. We assume that jobs are scheduled as shown in Figure 1.

Interchanging jobs α_1 and β_2, and keeping all the remaining jobs in their original positions, results in a difference between the new cost with respect to $\overline{C} + \beta_2 - \alpha_1$, and the previous cost with respect to $\overline{C}$, given by

$$\begin{aligned}\Delta &= (X+\beta_2)^2 + (Y+\alpha_1)^2 + (Y+\alpha_1+\beta_1)^2 - (X+\alpha_1)^2 - (Y+\beta_2)^2 - (Y+\beta_1+\beta_2)^2 \\ &= 2X(\beta_2-\alpha_1) + 4Y(\alpha_1-\beta_2) + (\alpha_1^2-\beta_2^2) + 2\beta_1(\alpha_1-\beta_2) \\ &= (\beta_2-\alpha_1)[2X - 4Y - \alpha_1 - 2\beta_1 - \beta_2] \leq 0,\end{aligned}$$

thus the existence of another solution with equal or lower cost provides a contradiction.

Case 2. $2X > 4Y + \alpha_1 + 2\beta_1 + \beta_2$. We consider Case 2 under 4 subcases (2A, 2B, 2C and 2D). First, however, consider the schedule as shown in Figure 2.

Let $N_E = |\{j \mid C_j < \overline{C}\}|$ and $N_L = |\{j \mid C_j > \overline{C}\}|$ denote the number of 'early' and 'late' jobs, respectively. If there is a job j with $C_j = \overline{C}$, then job j does not belong to either N_E or N_L. In Figure 2, we let X_i denote the time between the completion of the i-th *earliest* job finishing before $\overline{C}$, and $\overline{C}$, $i = 1, \ldots, N_E$, and we let Y_i denote the time between $\overline{C}$ and the completion of the i-th *latest* job finishing after $\overline{C}$, $i = 1, \ldots, N_L$. Note that, since $X > \beta_1 > \alpha_1$ in Figure 1, and $\alpha_1 \geq \alpha_2$ by the V-shaped property, the third job in the schedule must finish at or before $\overline{C}$, thus $N_E \geq 3$. Also, $N_L \geq 3$ from Lemma 1 and the condition for Case 2. Let $\alpha_i = X_i - X_{i+1}$, $i = 1, \ldots, N_E$, and $\beta_i = Y_i - Y_{i+1}$, $i = 1, \ldots, N_L$, and define $X_{N_E+1} = 0$, and $Y_{N_L+1} = 0$. Note that, interpreting Figure 2 in terms of Figure 1, $X = X_2$ and $Y = Y_3$. We begin by showing that, in an optimal schedule, there exists an index i, $2 \leq i \leq \min\{N_E, N_L\}$, such that *either* $\alpha_i > \beta_i$ *and* $X_i > Y_i$, *or* $X_i \leq Y_i$. Assuming that no such index i exists, then $\alpha_j \leq \beta_j$ and $X_j > Y_j$, $j = 2, \ldots, \min\{N_E, N_L\}$. Now, using the condition for Case 2, we have

$$\begin{aligned}&2X_2 > 4Y_3 + \alpha_1 + 2\beta_1 + \beta_2 \\ &\quad \Leftrightarrow X_1 + X_2 > 3Y_3 + 2\alpha_1 + 2\beta_1 + Y_2 \qquad \text{(from the definitions in Figure 2)} \\ &\quad \Rightarrow X_1 + X_2 > 2Y_3 + Y_2 + 2\alpha_1 + (Y_3 + \beta_1 + \beta_2) \qquad (\text{since } \beta_1 \geq \beta_2) \\ &\quad \Rightarrow (X_1 + X_2) - (Y_1 + Y_2) > 2Y + 2\alpha_1, \qquad \text{(again from Figure 2).} \qquad (1)\end{aligned}$$

Also, since $X_2 = \sum_{i=2}^{N_E} \alpha_i > \sum_{i=2}^{N_L} \beta_i = Y_2$ in any optimal schedule, and $\alpha_i \leq \beta_i$, $i = 2, \ldots, \min\{N_E, N_L\}$, it follows that $N_E > N_L$. Thus, since $X_1 + X_2 > Y_1 + Y_2$ and $X_j > Y_j$, $j = 3, \ldots, \min\{N_E, N_L\}$, we have $\sum_{i=1}^{N_E} X_i > \sum_{i=1}^{N_L} Y_i$, which contradicts Lemma 1, and is therefore an impossibility.

Thus, let $i^* = \min\{i \mid 2 \leq i \leq \min(N_E, N_L), \alpha_i > \beta_i$ *and* $X_i > Y_i$, *or* $X_i \leq Y_i\}$. Since $2X_2 = 2X > 4Y_3 + \alpha_1 + 2\beta_1 + \beta_2 > 2Y_2 + \alpha_1 + \beta_1$, we have $X_2 > Y_2$. Since $\alpha_2 \leq \alpha_1 < \beta_2$, it follows that $i^* \geq 3$.

Case 2A. $X_i^* \leq Y_i^*$. This implies that $X_{i-1}^* - \alpha_{i-1}^* \leq Y_{i-1}^* - \beta_{i-1}^*$, in which case *either* $X_{i+1}^* > Y_{i-1}^*$ and $\alpha_{i-1}^* \leq \beta_{i-1}^*$ which is algebraically impossible, *or* $X_{i-1}^* > Y_{i-1}^*$ *and* $\alpha_{i-1}^* > \beta_{i-1}^*$, *or* $X_{i-1}^* \leq Y_{i-1}^*$. In either of these last two cases, i^* was not minimal, a contradiction.

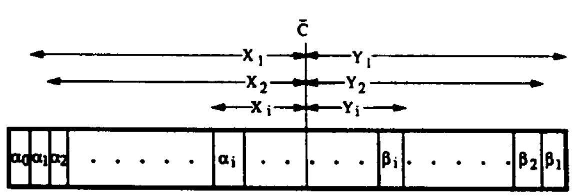

Fig. 2. Schedule for Theorem 1, Cases 2A–C

The remaining cases thus have $\alpha_i^* > \beta_i^*$ and $X_i^* > Y_i^*$. Let $\gamma = \alpha_i^* - \beta_i^*$. We say that job j is 'split' if its processing begins before $\bar{C}$ and ends after $\bar{C}$, i.e. $C_j - p_j < \bar{C} < C_j$.

Case 2B. $\alpha_i^* > \beta_i^*$, $X_i^* > Y_i^*$, and α_i^*, β_i^* are not split. Interchanging jobs α_i^* and β_i^*, and keeping all the remaining jobs in their original positions, results in a difference in the new cost with respect to $\bar{C} - \gamma$, and the previous cost with respect to $\bar{C}$, given by

$$\begin{aligned}\Delta &= \sum_{i=1}^{i^*} (x_i - \gamma)^2 + \sum_{i=1}^{i^*} (Y_i + \gamma)^2 - \sum_{i=1}^{i^*} X_i^2 - \sum_{i=1}^{i^*} Y_i^2 \\ &= 2i^*\gamma^2 - 2\gamma\left(\sum_{i=1}^{i^*} X_i - \sum_{i=1}^{i^*} Y_i \right) \\ &= 2\gamma\left[\sum_{i=1}^{i^*} (Y_i - X_i) + i^*\gamma \right].\end{aligned}$$

We need only show that for any value of $i^* \geq 3$, $\sum_{i=1}^{i^*} (X_i - Y_i) \geq i^*\gamma$. Firstly, since $\alpha_i \leq \beta_i$, $i = 1, \ldots, i^* - 1$, and from the definitions following Figure 2, we have

$$X_3 - Y_3 \leq X_4 - Y_4 \leq \cdots \leq X_{i^*-1} - Y_{i^*-1}. \tag{2}$$

Secondly, using the condition for Case 2,

$$\begin{aligned}&2X_3 + 2\alpha_2 > 2Y_3 + 2Y + \alpha_1 + 2\beta_1 + \beta_2 \\ &\quad \Rightarrow X_3 - Y_3 > Y - \alpha_2 + \tfrac{1}{2}\alpha_1 + \beta_1 + \tfrac{1}{2}\beta_2 \Rightarrow X_3 - Y_3 > Y + \alpha_1.\end{aligned} \tag{3}$$

Then from (1)–(3), and the condition that $X_i^* > Y_i^*$,

$$\sum_{i=1}^{i^*} (X_i - Y_i) > (Y + \alpha_1)[2 + 1 + (i^* - 4)] = (Y + \alpha_1)(i^* - 1). \tag{4}$$

Now from the definitions in Figure 2, if $i^* \geq 4$, then $Y > \beta_3 \geq \alpha_3 \geq \alpha_i^*$, thus from (4),

$$\sum_{i=1}^{i^*} (X_i - Y_i) > 2\alpha_{i^*}(i^* - 1) > i^*\alpha_{i^*}. \tag{5}$$

Alternatively, if $i^* = 3$, from (1) and (3),

$$\sum_{i=1}^{3} (X_i - Y_i) > 3\alpha_1 \geq i^*\alpha_{i^*}. \tag{6}$$

Thus, combining (5) and (6), we have

$$\sum_{i=1}^{i^*} (X_i - Y_i) > i^*\alpha_{i^*} \quad \text{for } i^* \geq 3. \tag{7}$$

Then from (7), and the fact that $\alpha_{i^*} > \gamma$, we have $\Delta < 0$, which provides the necessary contradiction.

Case 2C. $\alpha_i^* > \beta_i^*$, $X_i^* > Y_i^*$ and β_i^* is split. Here $N_E \geq N_L$, and thus, since $X_1 + X_2 > Y_1 + Y_2$ and $X_i > Y_i$, $i = 3, \ldots, N_L$, it follows that $\sum_{i=1}^{N_E} X_i > \sum_{i=1}^{N_L} Y_i$, therefore from Lemma 1 this situation is an impossibility.

Case 2D. $\alpha_i^* > \beta_i^*$, $X_i^* > Y_i^*$, β_i^* is not split and α_i^* is split. The schedule is shown in Figure 3, where the definitions of X_i and Y_i, $i = 1, \ldots, i^* - 1$, remain as in Figure 2.

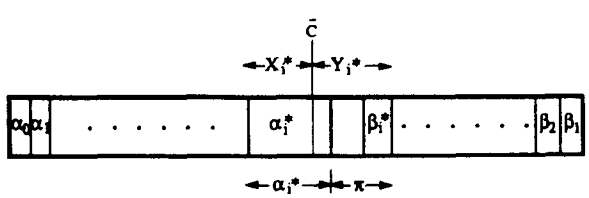

Fig. 3. Schedule for Theorem 1, Case 2D

Let π denote the total time between the completion of job α_i^* and that of job β_i^*, and let $\omega = X_i^* - \pi$. Since $X_i^* > Y_i^* \geq \pi$, clearly $\omega > 0$. Consider the following changes. We schedule jobs β_i, $i = i^*, \ldots, N_L - 1$, between $\overline{C} - \pi - \omega$ and $\overline{C} - \omega$ in nonincreasing order of their processing times, and job α_i^* between $\overline{C} - \omega$ and $\overline{C} + \alpha_i^* - \omega$. Note that jobs β_i, $i = i^*, \ldots, N_L - 1$, finish no further from $\overline{C} - \omega$ than they did from $\overline{C}$ before the interchange. We keep all the remaining jobs in their original positions with respect to the other jobs. Thus the resulting difference between the new cost with respect to $\overline{C} - \omega$, and the previous cost with respect to $\overline{C}$, is given by

$$\begin{aligned}\Delta &\leq \sum_{i=1}^{i^*} (X_i - \omega)^2 + \sum_{i=1}^{i^*} (Y_i + \omega)^2 - \sum_{i=1}^{i^*} X_i^2 - \sum_{i=1}^{i^*} Y_i^2 \\ &= 2\omega \sum_{i=1}^{i^*} (Y_i - X_i + \omega).\end{aligned}$$

From the definitions in Figure 3, $X_i^* = \alpha_i^* \Rightarrow \omega = \alpha_i^* - \pi < \alpha_i^*$. Then from (7), $\Sigma_{i=1}^{i^*} (X_i - Y_i) > i^* \alpha_{i^*} > i^* \omega \Rightarrow \Delta < 0$, contradiction.

So far, we have assumed throughout that job α_1 finishes at or before $\overline{C}$, and that job β_2 starts at or after $\overline{C}$. From Lemma 1, job α_1 must not finish after $\overline{C}$. Thus it only remains to consider the case when job β_2 is split. Let $\lambda = Y_2 - \alpha_1$. If $\lambda \leq 0$, then interchanging jobs α_1 and β_2, and keeping all the remaining jobs in their original positions, results in equal or lower cost for every job with respect to the original $\overline{C}$. Thus, from Lemma 3, total cost with respect to the new $\overline{C}$ is not increased. If $\lambda > 0$, then a similar interchange results in a difference in the new cost with respect to the new beginning of job α_1, (at $\overline{C} + \lambda$), and the previous cost with respect to $\overline{C}$, given by

$$\begin{aligned}\Delta &\leq (X_1 + \lambda)^2 + (Y_1 - \lambda)^2 + \alpha_1^2 - X_1^2 - Y_1^2 - Y_2^2 \\ &= \lambda(2X_1 + \lambda) - \lambda(2Y_1 - \lambda) - \lambda(\alpha_1 + Y_2) \\ &= \lambda(2X_1 + \lambda - 2Y_1 - 2\alpha_1).\end{aligned}$$

From Lemma 1, we know that $Y_1 + (Y_1 - \beta_1) \geq X_1 + (X_1 - \alpha_1) \Rightarrow 2Y_1 \geq 2X_1 + \beta_1 - \alpha_1 > 2X_1 + \lambda$. Thus $\Delta < 0$, a contradiction. □

4. Conclusions and potential applications

From the results of this paper, it is possible to fix the positions of the three longest jobs when seeking to minimize completion time variance. This property may be used to reduce the amount of enumeration required to solve the problem optimally. For special cases where the problem can be decomposed into small subsets of jobs, our results may provide an efficient algorithm. Finally, the authors are currently investigating the performance of heuristic procedures which implement similar ideas for instances of the general completion time variance problem. The recognition version of that problem is apparently open as to ordinary NP-completeness.

Acknowledgement

This research was supported in part by NSERC of Canada under Grant OGP0105675.

References

1 S. Eilon and I.G. Chowdhury, "Minimising waiting time variance in the single machine problem", *Management Sci.* **23**, 567–575 (1977).

2 J.J. Kanet, "Minimizing variation of flow time in single machine systems", *Management Sci.* **27**, 1453–1459 (1981).
3 A.G. Merten and M.E. Muller, "Variance minimization in single machine sequencing problems", *Management Sci.* **18**, 518–528 (1972).
4 L. Schrage, "Minimizing the time-in-system variance for a finite jobset", *Management Sci.* **21**, 540–543 (1975).
5 V. Vani and M. Raghavachari, "Deterministic and random single machine sequencing with variance minimization", *Oper. Res.* **35**, 111–120 (1987).