Scheduling to minimize the coefficient of variation

Prabuddha De[a], Jay B. Ghosh[b,*], Charles E. Wells[a]

[a] *Department of MIS and Decision Sciences, University of Dayton, Dayton, USA*
[b] *Faculty of Business Administration, Bilkent University, Bilkent 06533, Ankara, Turkey*

Abstract

In this paper, we address the problem of uninterruptedly scheduling a set of independent jobs that are ready at time zero with the objective of minimizing the coefficient of variation (CV) of their completion times. We first show that, for high processing time values of the longest job, a variance (V) minimizing schedule also minimizes CV. Using this equivalence, we next demonstrate the invalidity of an earlier conjecture about the structure of a CV-optimal schedule and proceed to establish the NP-hardness of the **CV** problem. Finally, drawing from our prior work on the **V** problem, we provide a pseudo-polynomial dynamic programming algorithm for the solution of the **CV** problem.

Keywords: Machine scheduling; Algorithms; Complexity

1. Introduction

Consider the situation where a set of n $(n \geqslant 1)$ independent jobs is ready at time zero for processing on a single machine which is continuously available. Assume that the preemption of a job and the insertion of machine idle times are not allowed. Assume further, without loss of generality, that associated with each job j, $j = 1, \ldots, n$, there is a distinct processing time p_j and that the jobs are indexed such that $p_1 > p_2 > \cdots > p_n$. Let σ be a sequence of the n jobs, and note that a sequence is sufficient to specify a schedule in the present context. Now, letting $[k]$ designate the index of the job in position k in a sequence, define

$$C_{[i]}(\sigma) = \sum_{1 \leqslant k \leqslant i} p_{[k]}(\sigma),$$

$$\bar{C}(\sigma) = (1/n) \sum_{1 \leqslant i \leqslant n} C_{[i]}(\sigma),$$

$$V(\sigma) = (1/n) \sum_{1 \leqslant i \leqslant n} [C_{[i]}(\sigma) - \bar{C}(\sigma)]^2,$$

$$CV(\sigma) = [V(\sigma)]^{1/2} / \bar{C}(\sigma).$$

A scheduling objective that has received much attention over the last two decades is that of identifying a job sequence which minimizes V or the completion time variance (call it the **V** problem); see Vani and Raghavachari [1] and De et al. [2] for a review of the pertinent literature. Recently, Raghavachari and Zammouri [3] have proposed the minimization of CV or the coefficient of variation as an alternate objective (call it the **CV** problem). The rationale for using this objective is that

* Corresponding author. Fax: + 90 312 266 4958; e-mail: ghosh@bilkent.edu.tr.

CV is a unit-less measure of variation and also that its use may be appropriate in some scheduling situations. The **CV** problem is particularly relevant if one is looking for a high mean time in system ($\bar{C}$) but a low time-in-system variance (V); consider, for example, the scheduling of chemical processes which on termination yield toxic wastes that are usually stored for future disposal. The **CV** problem has recently been studied by Mittenthal et al. [4] as well.

In the sequel, we present several new and interesting results for the **CV** problem. First, we establish an equivalence between the **V** and **CV** problems for high processing time values of the longest job. We then exploit this equivalence to demonstrate that an earlier conjecture about the structure of CV-optimal sequences, as stated in a remark in [3, p. 307] is incorrect, and establish further that the **CV** problem, like its **V** counterpart, is NP-hard. However, we go on to show that the **CV** problem is solvable in pseudo-polynomial time through the extension of a dynamic program for the **V** problem given in [2]. We conclude with a summary of our work.

2. New results

It has already been shown [3] that a CV-optimal sequence, like a V-optimal sequence, is $\bigvee$-shaped, i.e., the jobs preceding the shortest job (job n) in this sequence appear in the longest-processing-time-first order and those following it appear in the shortest-processing-time-first order. It has been shown further [3] that the longest job (job 1) in a CV-optimal sequence necessarily appears in the first position (as it does in a V-optimal sequence also).

We now show that, for high values of the parameter p_1 beyond a certain point, a V-optimal sequence in fact becomes CV-optimal. Let I be the base problem instance with parameters n and p_j, $j = 1, \ldots, n$, and Φ be the set of all sequences which have the largest $\bar{C}$ among the V-optimal sequences for I. Further, let Θ be the set of all $\bigvee$-shaped sequences which have the longest job in the first position and have $\bar{C}$ that are at least as large as $\bar{C}(\sigma)$, $\sigma \in \Phi$. From the conditions for V- and CV-

optimality, viz., $\bigvee$-shape and longest-job-first [2, 3], and the definitions of Φ and Θ, it follows that Φ is a subset of Θ. Noting further that a CV-optimal sequence cannot have $\bar{C}$ smaller than $\bar{C}(\sigma)$, $\sigma \in \Phi$ (because otherwise a V-optimal sequence in Φ will have a smaller CV and thus refute its optimality), it also follows that the set of all CV-optimal sequences for I, call it Ω, is contained in Θ as well.

Example. To illustrate the above defined sets, we consider a 5-job problem instance (I), where $p_1 = 10$, $p_2 = 9$, $p_3 = 8$, $p_4 = 7$ and $p_5 = 1$. Let $(i_1, \ldots, i_5)$ represent a sequence where i_k denotes the index of the job in the kth position. Notice that, in this instance, $\Theta = \{(1, 2, 3, 4, 5), (1, 2, 3, 5, 4), (1, 2, 4, 5, 3), (1, 3, 4, 5, 2)\}$, whereas $\Phi = \{(1, 3, 4, 5, 2)\}$ and $\Omega = \{(1, 2, 4, 5, 3)\}$. Of the V-optimal sequences, $(1, 2, 5, 4, 3)$ and $(1, 3, 4, 5, 2)$, the latter has the larger $\bar{C} (= 22.8)$ and thus constitutes Φ; recall that the cited $\bar{C}$ value is critical in the determination of Θ. Clearly, as expected, Φ and Ω are both subsets of Θ.

Returning to the task at hand, we now introduce a new problem instance I' such that $p'_1 = p_1 + \delta$ for $\delta \geqslant 0$, and $p'_j = p_j$, $j = 2, \ldots, n$. From the expressions for $\bar{C}$ and V in terms of the positional processing times as given in Schrage [5], we can see that $V'(\sigma) = V(\sigma)$ and $\bar{C}'(\sigma) = \bar{C}(\sigma) + \delta$ for σ in Φ, Θ and Ω. If we define Φ', Θ' and Ω' for I' in the same manner as we have defined Φ, Θ and Ω for I, then it becomes clear from the preceding observation that $\Phi' = \Phi$ and $\Theta' = \Theta$. Using the same arguments used earlier for Ω, it also becomes clear that Ω' is a subset of Θ' (and, therefore, of Θ).

Example. In the example introduced earlier, we let $\delta = 240$ and create an instance I' from I such that $p'_1 = 250$ and $p'_j = p_j$ for $j = 2, \ldots, 5$. Note that the sets Φ' and Θ' are, respectively, identical to the sets Φ and Θ, and that $\Omega' = \{(1, 3, 4, 5, 2)\}$ is again a subset of Θ' as expected.

We now state the result which shows that a member of Ω' (the complete set of the CV-optimal sequences for I') comes from the chosen subset of the V-optimal sequences for I', Φ' (and, therefore, from the corresponding set for I, Φ) and thus establishes the desired equivalence between the **V** and **CV** problems.

Results 1. There is a δ such that σ_Φ in Φ is also in Ω'

To avoid confusion and excessive notation, we choose to work primarily with Φ and Θ. First, we let $\Psi = \Theta - \Phi$. If Ψ is empty (as in the case for $n \leqslant 3$), σ_Φ in Φ is obviously also in Ω and the above result is trivially proved at $\delta = 0$. For the foregoing, we thus assume Ψ to be non-empty.

For all $\sigma_\Phi \in \Phi$ and $\sigma_\Psi \in \Psi$, first note that

$$V'(\sigma_\Phi) = V(\sigma_\Phi) < P^2, \quad V'(\sigma_\Psi) \geqslant V'(\sigma_\Phi) + (1/n)^3,$$

$$\bar{C}(\sigma_\Phi) > \delta, \quad \bar{C}(\sigma_\Psi) < P + \delta,$$

where $P = \sum_{1 \leqslant j \leqslant n} p_j$. Next, let $\delta = 3n^3 P^3$, and note that

$$V'(\sigma_\Phi)[\bar{C}'(\sigma_\Psi)]^2 - V'(\sigma_\Psi)[\bar{C}'(\sigma_\Phi)]^2$$

$$< V'(\sigma_\Phi)[P + \delta]^2 - [V'(\sigma_\Phi) + (1/n)^3]\delta^2$$

$$= V'(\sigma_\Phi)[P^2 + 2\delta P] - (1/n)^3\delta^2$$

$$< P^2[P^2 + 6n^3 P^4] - 9n^3 P^6$$

$$< 0.$$

$$\Rightarrow CV'(\sigma_\Phi) < CV'(\sigma_\Psi).$$

It follows that $CV'(\sigma_\Phi) \leqslant CV'(\sigma_\Theta)$ for all $\sigma_\Theta \in \Theta_,$. and that σ_Φ is CV-optimal for I' and thus in Ω'. This completes our proof.

Example. In our 5-job problem, $\sigma_\Phi = (1, 3, 4, 5, 2)$ is V-optimal and $\sigma_\Psi = (1, 2, 4, 5, 3)$ is CV-optimal for I. However, for I' (with $\delta = 240$), σ_Φ becomes both V-optimal and CV-optimal.

We now state a negative result about the conjecture [3, p. 307] that a CV-optimal sequence is always of the form $(1, 2, \ldots)$.

Result 2. A CV-optimal sequence is *not* always of the form $(1, 2, \ldots)$.

It is sufficient to show that there exists a problem instance I with a V-optimal sequence $\sigma_\Phi = (1, 3, \ldots)$; because of Result 1, we are then able to create another instance I' such that σ_Φ is CV-optimal as well. We have already shown the existence of exactly such an instance through our 5-job numerical example. Based on results in [5], it may be noted that, for 5-job problem instances in general,

the sequence $(1, 3, 4, 5, 2)$ can be taken to be σ_Φ as long as $3(p_2 - p_3) < (p_4 - p_5)$.

Result 1 also provides a convenient basis for proving that the **CV** problem is NP-hard. The proof is made possible by a recent result, due to Kubiak [6], that the related **V** problem is NP-hard.

Result 3. The **CV** problem is NP-hard.

Consider the NP-complete **V** decision problem: given an instance I, is there a sequence σ such that $V(\sigma) \leqslant v$? Assume, without loss of generality, that $v < P^2$ and that $n^3 v$ is an integer. From instance I, construct an instance I' by using $\delta = 3n^3 P^3$ as before. Obviously, this is accomplished in time and space that are both polynomial in the length of I. We now show that the answer to the **V** decision problem is "yes" if and only if the answer to the **CV** decision problem – viz., given I', is there a sequence σ such that $CV'(\sigma) \leqslant v^{1/2}/\delta$? – is "yes". Since the **CV** decision problem is easily seen to be in NP, showing the above establishes that it is NP-complete and thus that its corresponding optimization problem is NP-hard.

First, assume that the answer to the **V** decision problem is "yes", i.e., there is a sequence σ such that $V(\sigma) \leqslant v$. Clearly,

$$CV'(\sigma) = [V'(\sigma)]^{1/2}/\bar{C}'(\sigma)$$

$$= [V(\sigma)]^{1/2}/[\bar{C}(\sigma) + \delta]$$

$$< v^{1/2}/\delta.$$

This indicates that the answer to the **CV** decision problem is "yes" as well.

Next, assume that the answer to the **CV** decision problem is "yes", i.e., there is a sequence σ such that $CV'(\sigma) \leqslant v^{1/2}/\delta$. This implies that

$$V(\sigma) = V'(\sigma)$$

$$\leqslant (v/\delta^2)[\bar{C}'(\sigma)]^2$$

$$= (v/\delta^2)[\bar{C}(\sigma) + \delta]^2$$

$$= v + 2v\bar{C}(\sigma)/\delta + v[\bar{C}(\sigma)]^2/\delta^2.$$

Substituting $\delta = 3n^2 P^3$, $v < P^2$ and $\bar{C}(\sigma) < P$, and noting that both $n^3 V$ and $n^3 v$ are integers, it follows

that

$$V(\sigma) < v + (2/3)(1/n)^3 + (1/9)(1/n)^6(1/P)^2$$

$$< v + (1/n)^3.$$

$$\Rightarrow V(\sigma) \leqslant v.$$

Thus, the answer to the **V** decision problem is also "yes". This, in effect, completes the proof of Result 3.

Though the **CV** problem is NP-hard, we now show that it can be solved in pseudo-polynomial time; the problem is, therefore, NP-hard only in the ordinary sense. The solution that we provide is basically an extension of the dynamic programming algorithm for the **V** problem given in [2]. The algorithm recursively builds upon $\bigvee$-shaped partial sequences to arrive at the $\bigvee$-shaped CV-optimal complete sequence. Let σ_k and σ'_k be two $\bigvee$-shaped partial sequences formed of the shortest k jobs (jobs $n - k + 1$ through n). The main result related to their completion follows.

Result 4. If $\bar{C}(\sigma_k) = \bar{C}(\sigma'_k)$ and $V(\sigma_k) < V(\sigma'_k)$, then σ'_k cannot yield a CV-optimal sequence upon completion.

Let σ' be the $\bigvee$-shaped sequence obtained from the optimal completion of σ'_k. Assume that σ_k is completed identically to obtain the $\bigvee$-shaped sequence σ. Under the stated conditions, it has been shown [2] that $\bar{C}(\sigma) = \bar{C}(\sigma')$ and $V(\sigma) < V(\sigma')$. It immediately follows that $CV(\sigma) < CV(\sigma')$, proving Result 4.

Let stage k of the dynamic program (**DP**) correspond to the scheduling of the kth shortest job (job $n - k + 1$) to form a k-job partial sequence σ_k. Define, for stage k, $f_k(t) = \min\{kV(\sigma_k)\}$, where the minimum is taken over all k-job partial sequences σ_k with $k \bar{C}(\sigma_k) = t$. Also, define $P_k = \sum_{n-k+1 \leqslant j \leqslant n} p_j$. The **DP** recursion is as follows:

$$f_k(t) = \min\{f_{k-1}(t - kp_{n-k+1}) + (t - kp_{n-k+1})^2/[k(k-1)],$$

$$f_{k-1}(t - kP_k) + (t - kP_k)^2/[k(k-1)]\},$$

for $t = 0, \ldots, kP_k$ and $k = 2, \ldots, n$. The initial conditions are given by $f_1(t) = 0$ for $t = p_n$ and $f_1(t) = \infty$ for $t = 0, \ldots, p_n - 1$. Also,

$f_k(t) = \infty$ for $t < 0$ or $t > kP_k$ at $k = 1, \ldots, n - 1$. The optimal solution to the **CV** problem is found at the end of the recursion by computing $g^* = \min_{0 \leqslant t \leqslant nP}\{[nf_n(t)]^{1/2}/t\}$.

Result 5. The **DP** algorithm solves the **CV** problem in $O(n^2P)$ time and space.

The main part of the **DP** algorithm is essentially a recursive implementation of the enumerative algorithm given in [2]; it implicitly enumerates over all $\bigvee$-shaped partial sequences that are potentially optimal (cf. Result 4), and is therefore correct. Its time and space requirements at any stage k are bounded by kP_k. The overall complexity of the algorithm, both in terms of time and space, is thus $O(n^2P)$.

3. Summary and conclusion

With regard to the **CV** problem, we have shown its equivalence to the **V** problem under certain parameter values, proved the invalidity of an earlier conjecture about the structure of CV-optimal sequences, established its NP-hardness, and provided a pseudo-polynomial dynamic programming algorithm for its effective solution. Practically speaking, the problem has thus far been solved using heuristic methods only [3, 4]. Given our computational experience with the **DP** algorithm [2], the present results show that problem instances with up to 100 jobs can be solved exactly in a reasonable amount of time.

References

[1] Vani, V. and Raghavachari, M., 1987. Deterministic and random single machine sequencing with variance minimization, Oper. Res., 35: 111–120.

[2] De P., Ghosh, J.B. and Wells C.E., 1992. On the minimiz-ation of completion time variance with a bicriteria exten-sion, Oper. Res., 40: 1148–1155.

[3] Raghavachari, M. and Zammouri, M., 1993. Single machine scheduling with coefficient of variation minimization. Eur. Oper. Res., 62: 302–310.

[4] Mittenthal, J., Raghavachari, MJ. and Rana, A.I., 1993. A hybrid simulated annealing approach for single machine scheduling problems with non-regular penalty functions. Comput. Oper. Res., 20: 103–111.

[5] Schrage, L., 1975. Minimizing the time-in-system variance for a finite jobset. Management Science, 21: 540–543.

[6] Kubiak W., 1993. Completion time minimization on a single machine is difficult. Oper. Res. Lett., 14: 49–59.