Local search for stochastic parallel machine scheduling: improving performance by estimating the makespan

Guido Passage, Marjan van den Akker, Han Hoogeveen
Department of Information and Computing Sciences
Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands
g.j.p.n.passage@gmail.com, J.M.vandenAkker@uu.nl, J.A.Hoogeveen@uu.nl

We consider the following scheduling problem. There are n jobs that have to be carried out by m identical machines; processing job j takes time p_j, which is a stochastic variable. Job j cannot start before the release date r_j, and there are some *minimum delay precedence constraints*: job j must wait q_{ij} time units after job i has completed, before it can start. We want to find a feasible schedule with minimum expected makespan, that is, the completion time of the last job. To specify a schedule, we need to determine for each machine which jobs it executes and in which order this is done.

We present a local search approach. The main challenge here is that the expected makespan cannot be computed exactly and hence must be estimated in each iteration of the local search. In [1] we use discrete event simulation, but this requires a sufficient number of simulation runs to get a reliable estimate. To speed this up we have investigated several algorithms to estimate the makespan of a new candidate solution. Below we describe the algorithm that performed best.

Given a schedule, we compute the start time S_j of job j by starting it as soon as possible. Hence, S_j is equal to the maximum of a number of stochastic variables, which are: r_j; the values $S_i + p_i + q_{ij}$ for all predecessors i of j; and the completion time of the machine predecessor of j. Denoting these stochastic variables by $D_1, \ldots, D_l$, we find that $S_j = \max\{D_1, \ldots, D_l\}$. Next, we use [2], who describe how to find $E[X]$ and $Var[X]$ of $X = \max\{Y_1, Y_2\}$, where Y_1 and Y_2 are normally distributed stochastic variables with correlation coefficient ρ.Pretending that all D_k are normally distributed, we estimate $E[S_j]$ and $Var[S_j]$ by iteratively computing $X_k = \max\{X_{k-1}, D_k\}$, where $X_1 = D_1$. Here the variables D_k are considered in the following order: first all relations with jobs i executed on the same machine as j, starting with the machine predecessor of j, then all other predecessor relations, and then r_j. For stochastic variables corresponding to jobs on different machines we use $\rho = 0$; for the other ones we estimate ρ.

Computational experiments reveal that combining local search with the above algorithm finds solutions that are just as good as local search with estimating the makespan using 300 simulation runs per iteration, in a fraction of the time. This suggests that estimation outperforms simulation in local search for stochastic parallel machine scheduling problems.

References

[1] van den Akker, J.M., van Blokland, K. ,and Hoogeveen, J.A. Finding robust solutions for the stochastic job shop scheduling problem by including simulation in local search. In *Experimental Algorithms – SEA 2013*, Vol. 7933 of Lecture Notes in Computer Science, 402–413, 2013. Springer Berlin Heidelberg.

[2] Nadarajah. S; Kotz. S.: Exact distribution of the max/min of two Gaussian random variables. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 16*, 210-212, 2008.