UNIVERSITY OF TWENTE.

Scheduling with Uncertain Processing Times

Marc Uetz
m.uetz@utwente.nl

Berkeley, 2016

This Talk

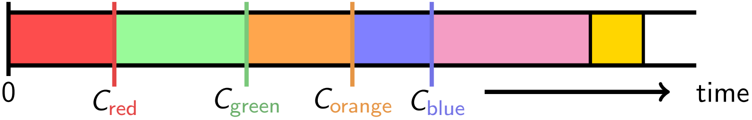
Mixture of

1. (mini) **introductory lecture**, stochastic scheduling
2. along with some **recent results** (unrelated machine scheduling) [Skutella, Sviridenko & U. 2016]

Single Machine Scheduling

Given: n jobs j with weights $w_j > 0$, processing times $p_j \in \mathbb{Z}_{>0}$

Task: sequence jobs on 1 machine; at most one job at a time;



Objective: minimize $\sum_j w_j C_j$ where $C_j = j$'s completion time;

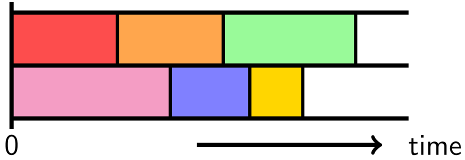
Theorem (Smith 1956)

Smith's rule, sequencing jobs in order $w_j/p_j \searrow$ is optimal □

Identical Parallel Machine Scheduling

Given: n jobs as above; m identical parallel machines

Task: schedule each job on any one machine; minimize $\sum_j w_j C_j$



Theorem

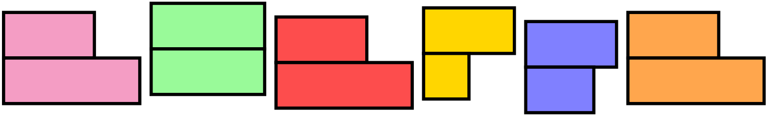
Problem is strongly NP-hard [Garey & Johnson, Problem SS13]

Smith's rule: tight 1.21-approximation [Kawaguchi & Kyan, 1986]

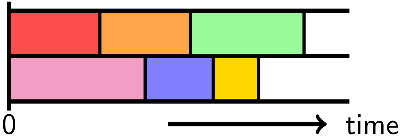
There exists a PTAS [Skutella & Woeginger, 2000]

Unrelated Machine Scheduling

Given: m machines, machine-dependent processing times p_{ij}



Task: schedule each job on one machine; minimize $\sum_j w_j C_j$



Theorem

Problem is APX-hard [Hoogeveen et al., 2002]

Exists $(\frac{3}{2} - c)$*-approximation* [Bansal, Srinivasan, Svensson, 2016]

Main Result

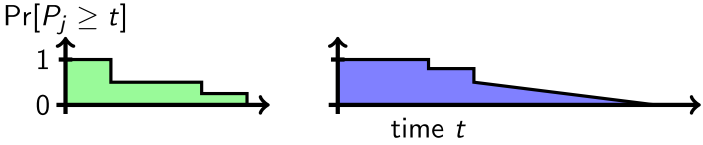
Theorem (Skutella, Sviridenko, U. 2016)

Stochastic unrelated machine scheduling *has a* $(\frac{3+\Delta}{2})$*-approximation.*

$\Delta =$ bounds the (squared) coeff. of variation of processing times

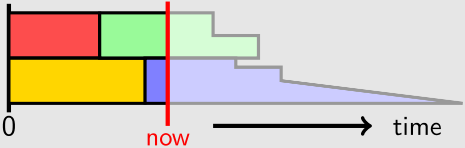
Stochastic Scheduling

processing time = (independent) random variables P_j (or P_{ij}); all known to us



Solution: Non-anticipatory scheduling policy Π

Decisions based on information **up to now** and a priori knowledge about P_j (or P_{ij}); no further information about the future.



Optimality

On instance I with policy Π

$\Pi(I) :=$ cost of policy Π on I, is a **random variable**

Definition (Optimal Policy)

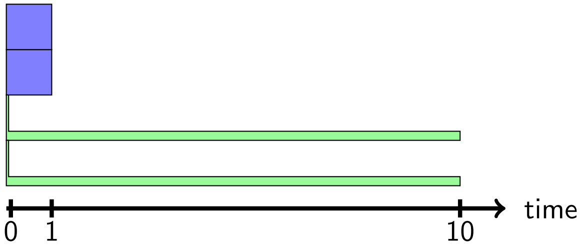
Call Π^{OPT} optimal if it achieves

$$\inf\{\ \mathbb{E}[\Pi(I)] \mid \Pi \text{ non-anticipatory policy }\}$$

Existence follows from [Möhring, Radermacher, Weiss 1985]

An Example

$n = 4$ jobs, unit weights



blue jobs: $P_j = 1$

green jobs: $P_j = \begin{cases} 0 & \text{probability } 4/5 \\ 10 & \text{probability } 1/5 \end{cases}$ (note $\mathbb{E}[P_j] = 2$)

Stochastic World

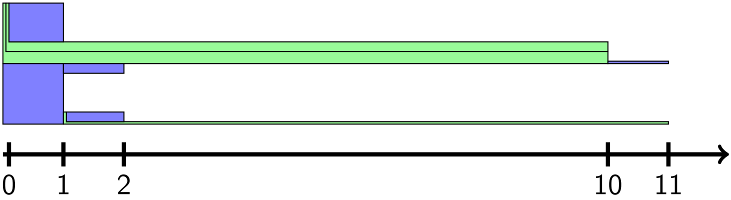
Tradeoff: better to delay

large $\mathbb{E}[P_j]$ or **large Pr**$(P_j$ **"large"**$)$ (heavy tail) ?

Claim

Unique 2-machine optimal policy: green,blue → green → blue
[with $\Pi(I) = \mathbb{E}[\sum_j C_j] = 6.92$].

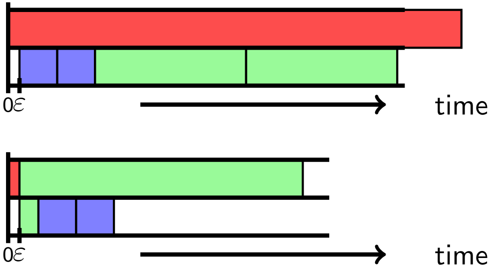
Just work it out □



Stochastic World: Deliberate Idleness

Theorem (U. 2003)

There are instances where only optimal policy **deliberately leaves machines idle**.



Approximation Algorithms

Optimal policies

- intuitively complex, exponential size decision tree; definitely NP(APX)-hard, ...
- only computing $\mathbb{E}[\Pi(I)]$ can be #P-hard [Hagstrom, 1988]

Definition (Approximation)

Policy Π has **performance guarantee** $\alpha \geq 1$, if for all instances I

$$\mathbb{E}[\Pi(I)] \;\leq\; \alpha\; \mathbb{E}[\Pi^{\mathrm{OPT}}(I)]$$

Our adversary is non-anticipatory, too!

Approximation Algorithms

Möhring, Schulz & U. [JACM, 1999]
First LP-based approximation algorithms
e.g.: Smith's rule has performance guarantee ($\frac{3+\Delta}{2}$).

Skutella & U. [SICOMP, 2005]
Extension to problems w. precedence constraints.

Megow, U. & Vredeveld [MOR, 2006] as well as Chou et al. [2006], Schulz [2008]
Stochastic jobs that arrive online.

All for **identical** machines; use LP lower bound on Π^{OPT}

LP Relaxation Identical Machines

Core ingredient: stochastic version of **load inequalities** [Möhring, Schulz & U., 1999]

$$\sum_{j\in S}\mathbb{E}[P_j]\,\mathbb{E}[C_j^\Pi] \;\geq\; \frac{1}{2m}\Bigl(\sum_{j\in S}\mathbb{E}[P_j]\Bigr)^2 + \frac{1}{2}\sum_{j\in S}\mathbb{E}[P_j]^2 - \frac{m-1}{2m}\sum_{j\in S}\mathbb{V}\mathrm{ar}[P_j] \quad \forall \text{ subsets } S$$

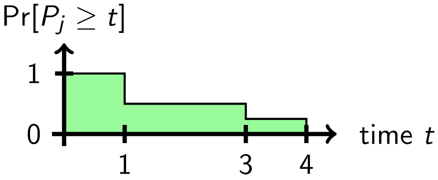
Generalizes LPs used earlier [Wolsey, 1985; Queyranne, 1993 & 1995; Hall, Schulz, Shmoys & Wein 1997]

But: doesn't generalize to **unrelated machines**

Structuring The Input

Theorem

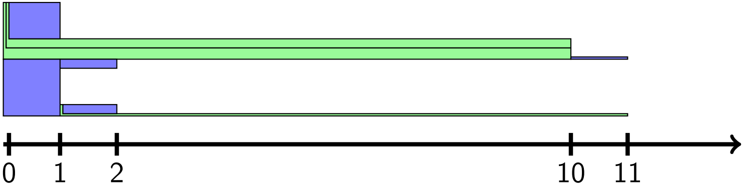
At a cost of $(1+\varepsilon)$, may assume w.l.o.g. input is integer valued.



Time-Indexed LP Relaxation: Intuition

Instance I and non-anticipatory policy Π, define

$$x_{ijt} := \Pr[\Pi \text{ starts job } j \text{ on machine } i \text{ at time } t \in \mathbb{Z}_{\geq 0}]$$



second blue job, $j = 4$, has

$$x_{1,4,0} = 16/25$$
$$x_{2,4,1} = 8/25$$
$$x_{1,4,10} = 1/25$$

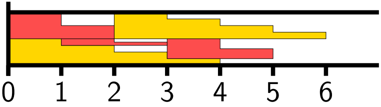
Time-Indexed LP Relaxation

Instance I and non-anticipatory policy Π

$$x_{ijt} := \Pr[\Pi \text{ starts job } j \text{ on machine } i \text{ at time } t \in \mathbb{Z}_{\geq 0}]$$

Properties of x_{ijt} (Π non-anticipatory!):

- $\mathbb{E}[C_j] = \sum_{i,t}\big(t + \mathbb{E}[P_{ij}]\big)\, x_{ijt}$
- $\sum_{i,t} x_{ijt} = 1$ for all jobs j
- $\Pr\big[i \text{ processes } j \text{ in } [s, s+1]\big] = \sum_{t=0}^{s} x_{ijt} \Pr[P_{ij} > s - t]$

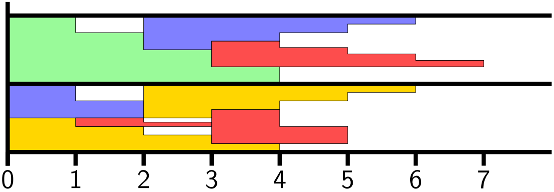


- $\displaystyle\sum_{j} \sum_{t=0}^{s} x_{ijt} \Pr[P_{ij} > s - t] \leq 1$ for each machine i and time s

Time-Indexed LP Relaxation

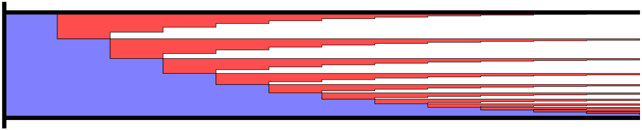
$$
\begin{aligned}
\min \quad & \sum_{i,j,t} w_j \left(t + \mathbb{E}[P_{ij}]\right) x_{ijt} \\
\text{s.t.} \quad & \sum_{i,t} x_{ijt} = 1 && \text{jobs } j, \\
& \sum_j \sum_{t=0}^{s} x_{ijt} \Pr[P_{ij} > s - t] \leq 1 && \text{machines } i, \text{ times } s, \\
& x_{ijt} \geq 0 && \text{jobs } j, \text{ machines } i, \text{ times } t.
\end{aligned}
$$

Example:

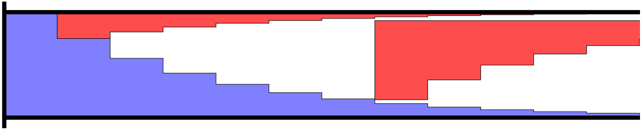


Technical Detail: Infinite LP Solution?

Two identical jobs with exponentially distributed processing times:



But: There are feasible LP solutions that are finite, e.g.



Theorem

∃ *finite optimal LP solution; LP can be solved efficiently (FPTAS).*

LP-Based Scheduling Policy

Algorithm

1. find an optimal (or approximate) LP solution (x_{ijt});
2. assign each job j independently at random to a machine i with
$$\Pr\big[j \text{ assigned to } i\big] = \sum_t x_{ijt} \quad ;$$
3. apply Smith's rule on each machine (that's optimal!);

Theorem (Skutella, Sviridenko & U. 2014)

This algorithm is a $(\frac{3+\Delta}{2})$*-approximation.*

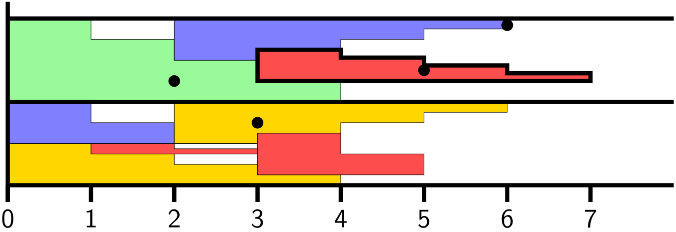
$$\Delta \geq \mathbb{CV}^2[P_{ij}] := \frac{\mathbb{V}\mathrm{ar}[P_{ij}]}{\mathbb{E}^2[P_{ij}]} \qquad \text{for all } P_{ij}$$

Proof of Performance Ratio

Idea: Analyze more complicated and provably worse algorithm:

1. find an optimal (or approximate) LP solution (x_{ijt});
2. for each job j
   a) choose pair (i, t) independently at random with probability x_{ijt};
   b) choose $r \in \mathbb{Z}_{\geq 0}$ indep. at random with probability $\frac{\Pr[P_{ij} > r]}{\mathbb{E}[P_{ij}]}$;
   c) set the *tentative start time* of j to $s := t + r$;
3. on each machine, sequence jobs by incr. tentative start times;

Example:



Proof of Performance Ratio

Key Lemma

Total exp. processing before job $j \to (i,s) \le$ tent. start time $s + \frac{1}{2}$

Which yields

$$\mathbb{E}[C_j] \le \sum_i \sum_{s \in \mathbb{Z}_{\ge 0}} \left(s + \tfrac{1}{2} + \mathbb{E}[P_{ij}]\right) \mathrm{Pr}\left[j \to (i,s)\right]$$

$$\le \cdots \le \left(\frac{3+\Delta}{2}\right) C_j^{\mathrm{LP}}$$

Lower Bounds

Performance bounds are tight ...

Theorem

Any *"fixed assignment" policy can have optimality gap* $\Delta/2$. [identical machines].

Theorem

The time-indexed LP can have optimality gap $\Delta/2$. [1 machine].

Final Remarks

To improve, **only adaptivity** is not enough

[Smith's rule, even though adaptive, can be as bad as $\Omega(\sqrt{\Delta})$]

open problems are

1. const. approximation (indep. of Δ)?

   Im, Moseley & Pruhs [STACS 2015] get

   $$O(\log^2 n + m \log n) - \text{approximation}$$

   (identical machines); balancing $\mathbb{E}[P_j]$ vs. $\Pr[P_j \text{ “large”}]$

2. nontrivial hardness / bounds on approximation ?

Beating $\Omega(\Delta)$: Im, Moseley, Pruhs 2015

Which jobs first? we want to **exclude** jobs with...

- $\Pr[P_j = \text{"large"}]$ maximal
- $\mathbb{E}[P_j]$ maximal

if ALG schedules k jobs A, OPT schedules k jobs A^*, then

Core Lemma

$$\begin{aligned} & \Pr[\text{ALG}(A) \text{ blocks all machines beyond } \tau] \\ \le\ & \Pr[\text{OPT}(A^*) \text{ blocks all machines beyond } \tau/\alpha] \end{aligned}$$

$\alpha = \Theta(m \log n)$

Some Reading

- Approximation in Stochastic Scheduling:...
  J ACM **46**, 1999
- When Greediness Fails: Examples from Stochastic Scheduling
  OR Lett **31**, 2003
- Stochastic Machine Scheduling with Precedence Constraints
  SIAM Comp **34**, 2005
- Models and Algorithms for Stochastic Online Scheduling
  MOR **31**, 2006
- Unrelated Machine Scheduling with Stochastic Processing Times
  MOR **41**, 2016

[w/ Megow, Möhring, Schulz, Skutella, Sviridenko, Vredeveld]

- Stochastic Scheduling of Heavy-Tailed Jobs
  Im, Moseley & Pruhs, *STACS*, 2015