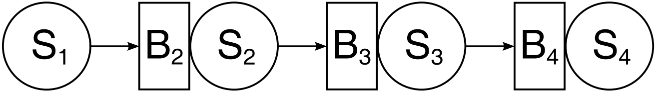
The Bounding Discrete Phase-Type Method

J.S. Tancrez and P. Semal
Université Catholique de Louvain

A.A. Markov Anniversary Meeting • Charleston, S.C. • June 12-14, 2006

Production Systems

1. Models of production systems have always been essential.
   - Design phases : equipments, layout, synchronization.
   - Operational level : load, sequencing, staffing.
2. Here, we focus on **production lines** with stochastic service times. The main difficulty lies in their productivity losses due to blocking and starving.
3. We take **non-restrictive assumptions** :
   - General finite service time distributions,
   - Finite buffer sizes,
   - Saturation : infinite arrival and infinite demand,
   - Blocking after service.



Models for Production Lines

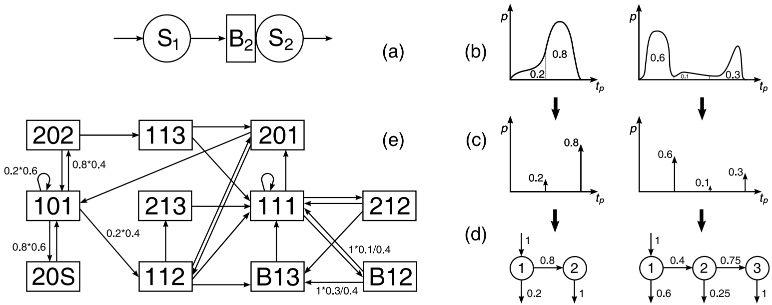
1. **Exact analytical models** allow direct understanding but can be used for simple production lines only :
   - Closed-form models are available for very simple configurations.
   - State models use Markov Chains. The main difficulty lies in the explosion of the state space size.
   - Holding time models construct recursive relationships.
2. At the other end of the continuum, **simulation** is very general. Its weakness mainly lies in its development cost.
3. **Approximate analytical models** are in between : the development cost keeps low but the result are uncertain.
   - Decomposition solve more but much easier subproblems.
   - Expansion adds the concept of artificial node.

The Bounding Discrete Phase-Type Method

1. The distributions are discretized, using a same time step τ, in order to describe the evolution of the production system by a **Markov chain**.
2. Our originality lies in the discretization method :
   - We concentrate the probability mass in $[k\tau, (k+1)\tau]$...
   - ... at the end of this interval, $(k+1)\tau$.
3. Aiming to :
   - Keep an intelligible link with the original distributions,
   - Prove bounds on the productivity.
4. We call it the **"grouping at the end" discretization**.

The BDPH Method

Trivial Example



1. Transient and steady-state performances can be computed.
2. The method is approximate but offers theoretical control.
3. As it is a state model, the weakness lies in the explosion of the state space. Other methods will be considered in the future.

Production Run

1. We consider a random infinite real time production run r. To construct a run, we only need a sequence of random processing times :

$$\{I^r(W_{i,k})\} \stackrel{\Delta}{\longmapsto} r.$$

where $W_{i,k}$ denotes the job k at station i.

2. If these times are discretized, we get $\overline{r}$:

$$\{I^r(W_{i,k})\} \stackrel{\text{disc.}}{\longmapsto} \{\overline{I^r(W_{i,k})}\} \stackrel{\Delta}{\longmapsto} \overline{r}.$$

When using the "grouping at the end" discretization :

$$\overline{I^r(W_{i,k})} \triangleq \left\lceil \frac{I^r(W_{i,k})}{\tau} \right\rceil \tau.$$

It leads to the following interesting property :

$$\overline{I^r(W_{i,k})} - \tau \leq I^r(W_{i,k}) \leq \overline{I^r(W_{i,k})}.$$

Structural Properties

Lemma (Structural properties of a production line)

Given a production line including a buffer of size b_i before each station i, the jobs verify the following inequalities, $\forall\, r, i, k$:

$$
\begin{aligned}
t^r_{start}\left(W_{i,k}\right) &\geq t^r_{end}\left(W_{i-1,k}\right) && (1)\\
&\geq t^r_{end}\left(W_{i,k-1}\right) && (2)\\
&\geq t^r_{end}\left(W_{\substack{i+1, & k-b_{i+1}-2\\ i+2, & k-b_{i+1}-b_{i+2}-3\\ i+3, & k-b_{i+1}-b_{i+2}-b_{i+3}-4\\ \vdots & \vdots}}\right) && (3)
\end{aligned}
$$

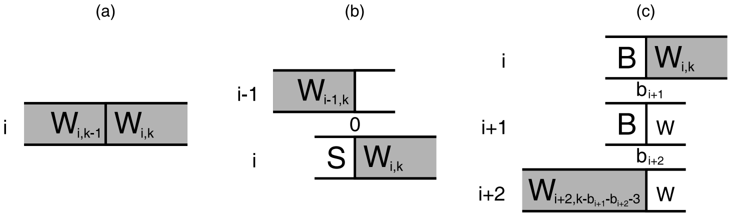
with at least one equality.

Structural Properties

Informal Proof

1. A job k can only be started on station i if :
(1) its processing on the previous station $i-1$ is ended,
(2) the processing of the job $k-1$ in station i is ended and
(3) this previous job $k-1$ is not blocked in station i.

2. There is no reason to wait once all these conditions are satisfied. The start moment $t^r_{start}(W_{i,k})$ is thus given by the maximum of the right hand sides.



Productivity and Critical Path

1. We focus on the time needed to produce p units in a run r, given by $t^r_{end}(W_{m,p})$. Only the p-part of a run is relevant :

$$\{ l^r(W_{i,k}) \mid 1 \leq i \leq m, 1 \leq k \leq p\} \stackrel{\Delta}{\longmapsto} r_p.$$

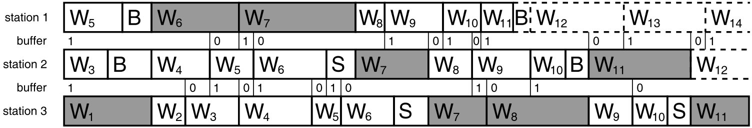
2. In order to state $l(r_p) = t^r_{end}(W_{m,p})$ in function of the processing times, we introduce the concept of **critical path** of r_p, $cp(r_p)$. It is defined as the sequence of jobs that covers r_p.

$$l(r_p) = l(cp(r_p)) = \sum_{W_{i,k} \in cp(r_p)} l^r(W_{i,k}).$$

3. By the Lemma, every run r_p has at least one critical path.
4. As the inequalities are valid for any run, the absence of overlap is independent of the considered run, unlike the absence of gap.

Example of Critical Path

1. The critical path is build starting with the last job and looking which job end, in this run, has triggered its start, i.e. which inequality of the Lemma is satisfied at equality.
2. The notion of critical part is best illustrated on a Gantt Chart :



3. The concept of critical path allows to relate the intelligible transformation of the job lengths and its effect on the global length of a production run.

Founding Result

Lemma

The time an m–station line takes to produce p units in a random real time production run r can be bounded as follows :

$$l(\bar{r}_p) - \tau(p+m-1) \leq l(r_p) \leq l(\bar{r}_p).$$

The **proof** is based on $\overline{I^r(W_{i,k})} - \tau \leq I^r(W_{i,k}) \leq \overline{I^r(W_{i,k})}$. It summarized in the two following equations :

$$l(r_p) = \sum_{j \in cp(r_p)} l^r(j) \leq \sum_{j \in cp(r_p)} \overline{l^r(j)} \leq \sum_{j \in cp(\bar{r}_p)} \overline{l^r(j)} = l(\bar{r}_p),$$

$$l(\bar{r}_p) - \tau|cp(\bar{r}_p)| = \sum_{j \in cp(\bar{r}_p)} (\overline{l^r(j)} - \tau) \leq \sum_{j \in cp(\bar{r}_p)} l^r(j) \leq \sum_{j \in cp(r_p)} l^r(j).$$

Bounds on the Throughput

Only average performances are really useful. We consider the mean time T_P necessary to reach a given production P.

$$T_P = \int f(r_P) l(r_P) dr_P.$$

Theorem

The mean time T_P an m–station line takes to produce P units can be bounded on the basis of the information computed by the BDPH method. If $\overline{T}_P$ is the mean time to produce P units using the "grouping at the end" discretized times, we have :

$$\overline{T}_P - \tau(P + m - 1) \leq T_P \leq \overline{T}_P.$$

The **proof** relies on Lemma and on the fact that the probabilities of r_P and $\bar{r}_P$ are the same since they are derived from the same run r.

Bounds on the Throughput (II)

If we are interested in a fixed time instead of a fixed production, bounds can quite easily be derived from the previous Theorem.

Theorem

The mean production P_T produced by an m–station line during a fixed time T can be bounded on the basis of the informations computed by the BDPH method. If $\overline{P}_T$ is the mean production during time T using the "grouping at the end" discretized times, and $\overline{P}^*$ is the mean production during discrete time $\overline{T}^*$, where chosen minimal such that $\overline{T}^* - \tau(\overline{P}^* + m - 1) \geq T$, we have :

$$\overline{P}_T \leq P_T \leq \overline{P}^*.$$

Bounds on the Steady-State Productivity

The average time between the completion of two units, in steady-state, is called the cycle time c where $c = lim_{P \to \infty} T_P / P$.

Corollary (BDPH bounds for a production line)

When it measures the productivity of a production line in steady–state, the BDPH method is pessimistic, i.e. the cycle time is overvalued. Moreover, the error is smaller than the discretization time step :

$$\overline{c} - \tau \leq c \leq \overline{c}.$$

1. Our method allows to bound the productivity.
2. These bounds become tighter and converge when the discretization step is decreased.
3. The accuracy is directly related to the discretization step.

Approximations of the Steady-State Productivity

Approximation

The cycle time of a line can be approximated by :

$$c \approx \overline{c} - \frac{\tau}{2}.$$

Approximation

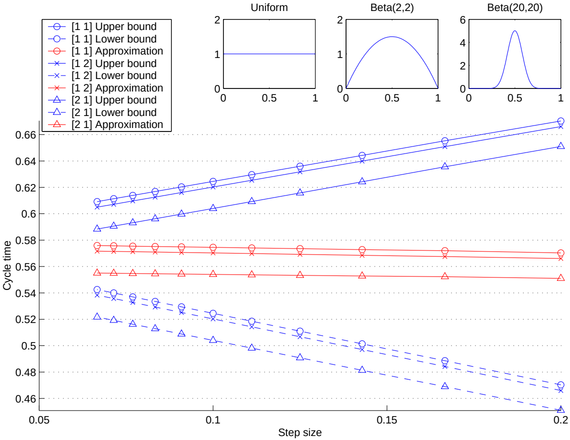
The cycle time of a line can be approximated, $\forall\, i$, by :

$$c \approx \overline{c} - e_W(i),$$

with $e_W(i) = E[\overline{I^r(W_{i,k})}] - E[I^r(W_{i,k})]$, *the discretization bias on the service distribution of station* i.

Example

We consider a simple three stations production line and compute the bounds and the approximation given by the BDPH Method.



Summary

1. We introduced a new method, called the BDPH method. It discretizes the distributions and then model the evolution of the line by a Markov Chain.
2. Our originality lies in the "grouping at the end" discretization.
3. The BDPH method computes upper and lower bounds on the productivity of a production line. These bounds converge and their accuracy is related to the discretization step.
4. Approximations of the productivity, and of other performances, can also be computed.
5. The critical path allowed the main results to be proved.
6. The weakness lies in the explosion of the Markov Chain size.

Future Work

1. Investigate how other performances of interest can be bounded (buffer utilization, average job flow time).
2. Derive more subtle approximations, using the critical path.
3. Apply our approach to more complex systems, by stating their structural properties and using the concept of critical path.
4. Implement better solution methods to solve the generated discrete time Markov Chain.
5. Speed the resolution up using approximation methods, like the decomposition method.