Scheduling in a random environment: stability and asymptotic optimality

Martin Erausquin

UPV/EHU, BCAM

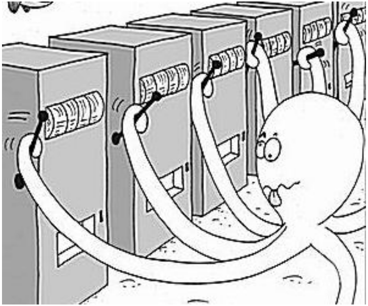
October 26, 2011

- Ayesta U., Erausquin M., Jacko P., *A Modeling Framework for Optimizing the Flow-Level Scheduling with Time-Varying Channels*, Performance Evaluation 67 (2010)
- Ayesta U., Erausquin M., Jonckheere M., Verloop I.M., *Scheduling in a random environment: stability and asymptotic optimality*, submitted.

Motivation: wireless application

- In each time slot, the base station selects a flow (user, job) to serve
- Channel conditions vary due to fading and interference effects
- In each channel condition, the probability of completing the job in one time slot is different

Multi Armed Bandit Problem



- A sequential decision problem where at each time slot the agent must choose one of K available options
- Depending on the chosen action, the agent receives a payoff at the end of the time slot
- Goal: maximize the present value of the future payoffs, choosing the right sequence of actions

Model description

- Time is slotted
- K jobs, $k \in \mathcal{K} = \{1, 2, ..., K\}$
- $\mathcal{N}_k = \{1, 2, ..., N_k\}$, set of possible states for job k
- $\forall n \in \mathcal{N}_k$, $q_{k,n}$: probability for job k of being at state n
- $\forall n \in \mathcal{N}_k$, $\mu_{k,n}$: departure probability for job k, if served, when it is at state n
- $0 \leq \mu_{k,1} \leq \mu_{k,2} \leq \cdots \leq \mu_{k,N_k} \leq 1$
- c_k: holding cost of job k per slot waiting for service
- Independence in the state evolution history, independence between different job's current states

MDP formulation

- $A = \{0, 1\}$: Action space. Action 0 means 'not serving', action 1 means 'serving'
- The expected one-period reward earned by job k at state n, depending on the action, will be given by

$$R^1_{k,n} = -c_k(1 - \mu_{k,n}) \qquad R^0_{k,n} = -c_k$$

- $X_k(\cdot)$: state process of job k
- $a_k(\cdot)$: action process of job k
- Objective:

$$\max_{\pi \in \Pi_{\mathbf{X},\mathbf{a}}} \mathbb{E}^{\pi} \left[\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k \in \mathcal{K}} R^{a_k(t)}_{k, X_k(t)} \right] \tag{P}$$

$$\text{s. t.} \sum_{k \in \mathcal{K}} a_k(t) = 1, \text{ for all } t \in \mathcal{T}$$

Relaxations (Whittle, 1988)

- Problem (P) can not be solved, neither analytically nor numerically
- We relax the constraint:

$$\sum_{k \in \mathcal{K}} a_k(t) = 1, \text{ for all } t \in \mathcal{T} \Rightarrow \mathbb{E}^{\boldsymbol{\pi}} \left[\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k \in \mathcal{K}} a_k(t) \right] = 1$$

- We get the following relaxed problem

$$\max_{\boldsymbol{\pi} \in \Pi_{\mathbf{X},\mathbf{a}}} \mathbb{E}^{\boldsymbol{\pi}} \left[\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k \in \mathcal{K}} R_{k, X_k(t)}^{a_k(t)} \right] \qquad \text{(RP)}$$

$$\text{s. t. } \mathbb{E}^{\boldsymbol{\pi}} \left[\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{k \in \mathcal{K}} a_k(t) \right] = 1$$

Solution: Potential Improvement rule

- The relaxed problem can be solved analytically
- It can be approached using Lagrangian methods

$$\max_{\boldsymbol{\pi}\in\Pi_{\mathbf{X},\mathbf{a}}} \mathbb{E}^{\boldsymbol{\pi}}\left[\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\left(\sum_{k\in\mathcal{K}} R_{k,X_k(t)}^{a_k(t)} - \nu\sum_{k\in\mathcal{K}} a_k(t)\right)\right]$$

- We decompose this problem in K subproblems:

$$\max_{\widetilde{\pi}_k\in\Pi_{\mathbf{X},a_k}} \mathbb{E}^{\widetilde{\pi}_k}\left[\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\left(R_{k,X_k(t)}^{a_k(t)} - \nu a_k(t)\right)\right] \qquad \text{(SRP)}$$

- We solve K subproblems, and we obtain the joint optimal policy for the relaxed problem combining them

Solution: Potential Improvement rule

- The next theorem gives the optimal solution for the relaxed problem:

Theorem

Let

$$\nu_{k,n} = \frac{c_k \mu_{k,n}}{\sum_{m>n} q_{k,m}(\mu_{k,m} - \mu_{k,n})} \text{ for } n \neq N_k, \qquad \nu_{k,N_k} = \infty$$

Then:

- *If* $\nu \leq \nu_{k,n}$*, it is optimal to serve job* k *if it is in state* $n \in \mathcal{N}_k$*;*
- *If* $\nu \geq \nu_{k,n}$*, it is optimal not to serve job* k *if it is in state* $n \in \mathcal{N}_k$*;*

- We construct a feasible policy for the original problem using this theorem
- **Potential Improvement rule:** gives service at time t to job $k^*(t)$ such that:

$$k^*(t) := \arg\max_{k \in \mathcal{K}} \nu_{k,X_k(t)}$$

Scheduling disciplines

- Potential Improvement index:

$$\nu_{k,n}^{\mathrm{PI}} = \frac{c_k \mu_{k,n}}{\sum_{m>n} q_{k,m}(\mu_{k,m} - \mu_{k,n})} \text{ for } n \neq N_k \qquad \nu_{k,N_k}^{\mathrm{PI}} = \infty$$

- $c\mu$ index:

$$\nu_{k,n}^{c\mu} := c_k \mu_{k,n} \text{ for } n \in \mathcal{N}_k$$

- Score Based index (T.Bonald, 2004):

$$\nu_{k,n}^{\mathrm{SB}} := \sum_{m=1}^{n} q_{k,m}, \text{ for } n \in \mathcal{N}_k$$

- Relatively Best index (Qualcomm 3G standard, 2000):

$$\nu_{k,n}^{\mathrm{RB}} := \frac{c_k \mu_{k,n}}{\sum_{m=1}^{N_k} q_{k,m} \mu_{k,m}}, \text{ for } n \in \mathcal{N}_k$$

Tie Breaking rules

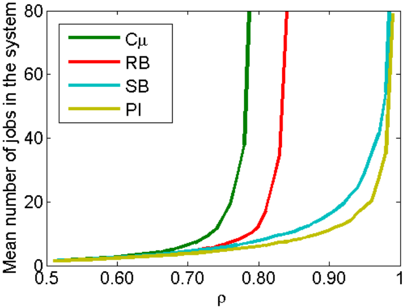
- Under some schedulers, users in their best possible state have the same index value:

$$\nu_{k,N_k}^{\mathrm{PI}} = \infty \text{ for all } k \qquad \nu_{k,n}^{\mathrm{SB}} = 1 \text{ for all } k$$

- Ties broken at random under SB
- Ties broken following the $c\mu$-rule under PI

Numerical simulation

- Two classes, $k \in \mathcal{K} = \{1, 2\}$
- At each time slot, a new class-k job arrives with probability λ_k
- λ_2 fixed
- $c_1 = c_2 = 1$
- We move λ_1 such that $\varrho = \frac{\lambda_1}{\mu_{1,N_1}} + \frac{\lambda_2}{\mu_{2,N_2}}$ varies from 0.5 to 1



How good is Potential Improvement rule?

- Questions:
  - Stability
  - Optimal in some sense?
- K classes, $k \in \mathcal{K} = \{1, 2, ..., K\}$, with arrival rate λ_k
- $X_k(t)$: number of class-k jobs at time t. $X(t) = (X_1(t), ..., X_K(t))$.
- The fluid limit of the process $X(t)$ is defined as follows:

$$Y^r(t) = \frac{X(rt)}{r}$$

  when r goes to ∞, starting with $X_k(0) = r$, for all k
- We characterize the maximum stability condition, the set of policies that achieve maximum stability, and the set of policies that are asymptotically fluid optimal

Classes of policies

Definition

A policy belongs to the family of **BR policies** if, whenever there is a job in its best state, this job is preferred over any other job which is not in its best possible state

Definition

A policy belongs to the family of **BRP policies** if it is a BR policy and it uses the $c\mu$-rule as the tie-breaking rule

- RB, SB and PI are BR policies
- PI is a BRP policy

Asymptotic drifts

- Drift: $\delta(x) = \mathbb{E}(X(1) - x | X(0) = x)$
- If the drift vector $\delta(x)$ has uniform limits, we can define the asymptotic drifts, $\delta^{\mathcal{U}} : \mathbb{N}^{\mathcal{U}} \to \mathbb{R}^K$,

$$\delta^{\mathcal{U}}(x_{\mathcal{U}}) := \lim_{x_k \to \infty,\ k \in \mathcal{U}^c} \delta(x)$$

- Let $X^{\mathcal{U}}$ denote the $|\mathcal{U}|$-dimensional stochastic process corresponding to the original process seeing an infinite number of users of class $k \in \mathcal{U}^c$ and let $\pi^{\mathcal{U}}$ denote its stationary measure
- Averaged drift vectors are

$$\tilde{\delta}^{\mathcal{U}} = \sum_{x \in \mathbb{N}^{\mathcal{U}}} \delta^{\mathcal{U}}(x) \pi^{\mathcal{U}}(x)$$

Strong fluid limit

Theorem

For a given policy f inducing a partially increasing vector field with uniform limits drift, we have:

$$\lim_{r\to\infty} \mathbb{P}\big(\sup_{0\le s\le t} |Y^{f,r}(s) - y^f(s)| \ge \epsilon\big) = 0, \quad \text{for all } \epsilon > 0,$$

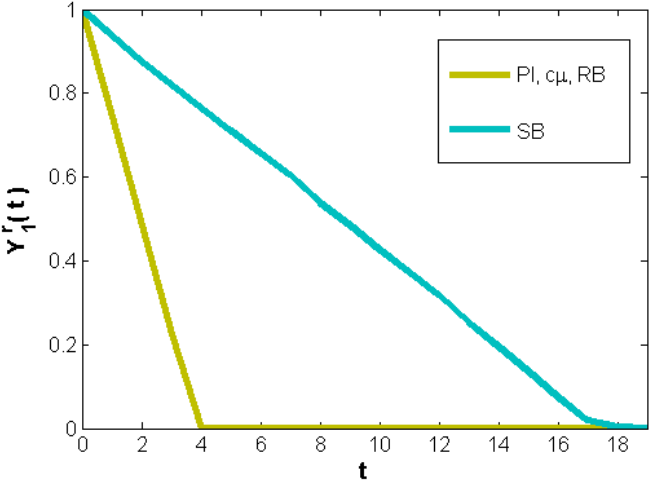
with $y^f(t)$ a piece-wise linear function such that (let $\mathcal{U}_0 = \emptyset$ and $T_0 = 0$):

$$\frac{\mathrm{d}y_k^f(t)}{\mathrm{d}t} = \tilde{\delta}_k^{f,\mathcal{U}_l}, \quad t \in [T_l^f, T_{l+1}^f],$$

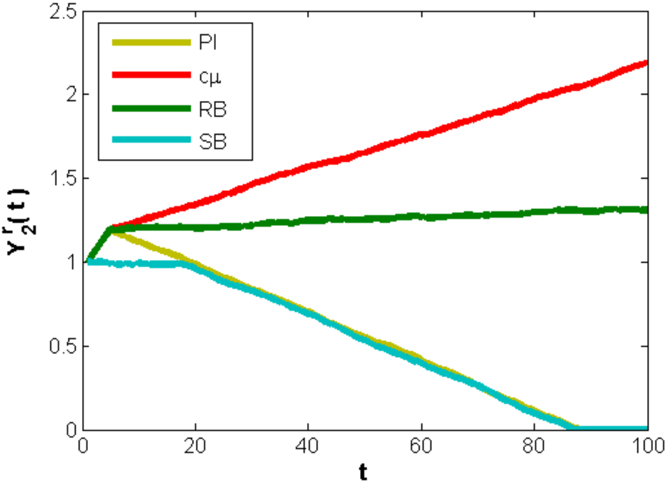
$$\text{with } T_{l+1}^f = T_l^f + \min_{k\in\mathcal{U}_l^c, \tilde{\delta}_k^{\mathcal{U}_l}<0} \frac{y_k^f(T_l^f)}{-\tilde{\delta}_k^{f,\mathcal{U}_l}},$$

$$\text{and } \mathcal{U}_{l+1} = \mathcal{U}_l \cup \arg\min\{\frac{y_k^f(T_l^f)}{-\tilde{\delta}_k^{f,\mathcal{U}_l}}, k \in \mathcal{U}_l^c\}.$$

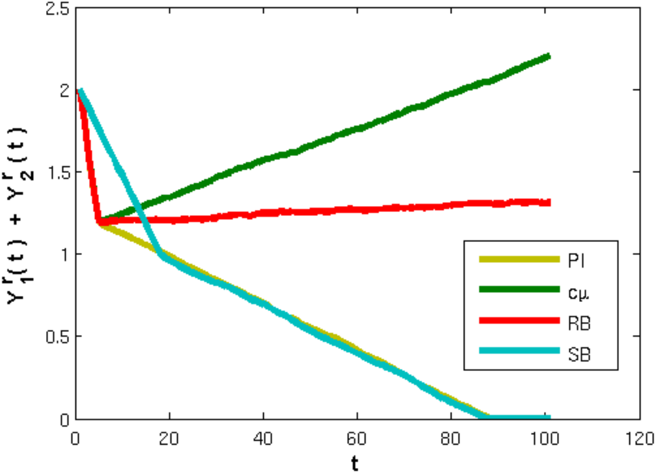
Strong fluid limit



Strong fluid limit



Strong fluid limit



Stability

Theorem

A policy f inducing a partially increasing drift with uniform limits achieves stability if $T_l^f < \infty$ for all l, where $T_l^f = \inf\{t \geq 0 : y_l^f(t) = 0\}$

- This theorem allows us to calculate numerically the stability region for a large class of policies

Theorem (Proposed (but not proven) in S.Aalto, P.Lassila (2010))

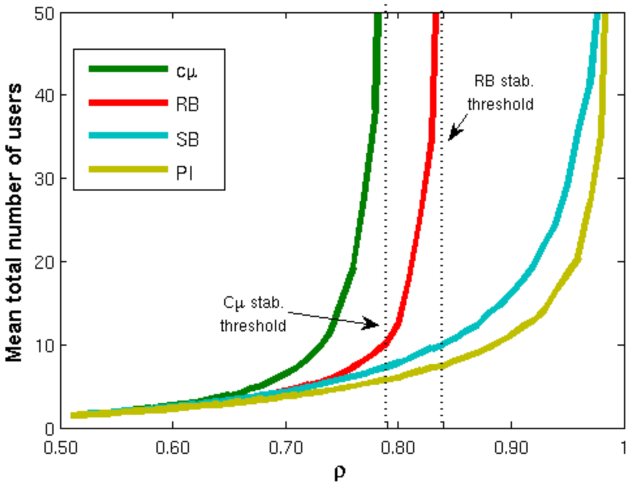
1. *The maximum stability region is*

$$\sum_{k=1}^{K} \frac{\lambda_k}{\mu_{k,N_k}} \leq 1$$

2. *Any BR policy has maximum stability region*

- In particular, PI and SB are maximum stable

Stability



cμ
RB
SB
PI
RB stab.
threshold
Cμ stab.
threshold
Mean total number of users
ρ
0
10
20
30
40
50
0.50
0.60
0.70
0.80
0.90
1

Asymptotic fluid optimality

- Deterministic fluid control model

$$\min \sum_{k=1}^{K} c_k x_k(t), \text{ for all } t \geq 0$$

$$x_k(t) = x_k(0) + \lambda_k t - \sum_{n=1}^{N_k} \mu_{k,n} \int_0^t u_{k,n}(v)\mathrm{d}v,$$

$$x_k(t) \geq 0, \; k = 1, \ldots, K$$

such that for all $v \geq 0$,

$$\sum_{k=1}^{K} \sum_{n=1}^{N_k} u_{k,n}(v) \leq 1, \quad u_{k,n}(v) \geq 0, \text{ for all } k, n$$

Lemma

Assume $\mu_{1,N_1} \geq \mu_{2,N_2} \geq \ldots \geq \mu_{K,N_K}$*. The fluid control that solves the fluid control problem is as follows: Let* $l = \arg\min\{k : x_k(t) > 0\}$*. Then*

$$u^*_{k,N_k}(t) = \frac{\lambda_k}{\mu_{k,N_k}}, \quad \textit{for} \quad k < l, \quad u^*_{l,N_l}(t) = 1 - \sum_{i=1}^{l-1} \frac{\lambda_i}{\mu_{i,N_i}}$$

Lemma

For any policy f *we have*

$$\liminf_{r\to\infty} \sum_{k=1}^{K} Y_k^{f,r}(t) \geq \sum_{k=1}^{K} x_k^*(t)$$

Definition

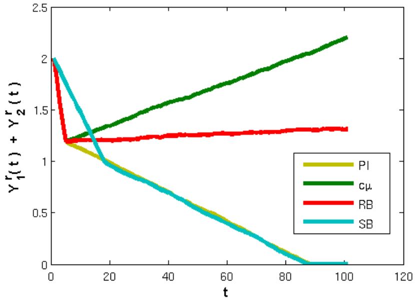
A policy f is said to be **asymptotically fluid optimal** when the lower bound is obtained for all $t \geq 0$

Asymptotic fluid optimality

Theorem

Any BRP policy is asymptotically fluid optimal

- PI is asymptotically fluid optimal



Conclusions and future work

- Main conclusions:
  - Whittle's index rule (PI) consistently outperforms all the other policies (or is equivalent to the best one)
  - BR policies achieve maximum stability region. In particular, PI is maximum stable
  - We can calculate numerically the stability region for a wide class of policies
  - BRP policies are asymptotically fluid optimal. PI is the only known scheduling rule which is asymptotically fluid optimal

  Future work:
  - Can we construct policies which outperform PI?
  - Can we say something about the optimal tie-breaking rule?
  - Include correlations in the channel state evolution