Scheduling of Users with Markovian Time-Varying Service Rates

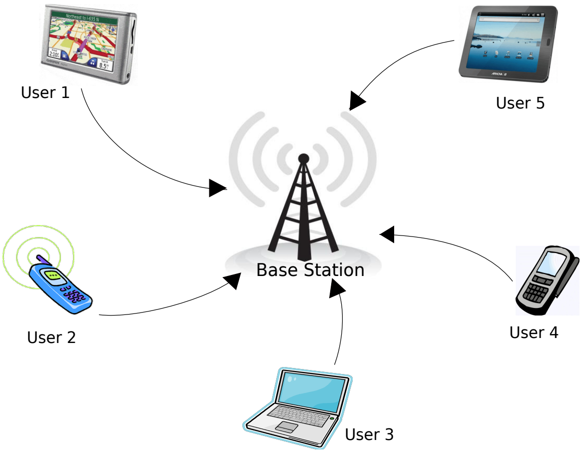
Fabio Cecchi

Joint work with Peter Jacko, Lancaster University and BCAM

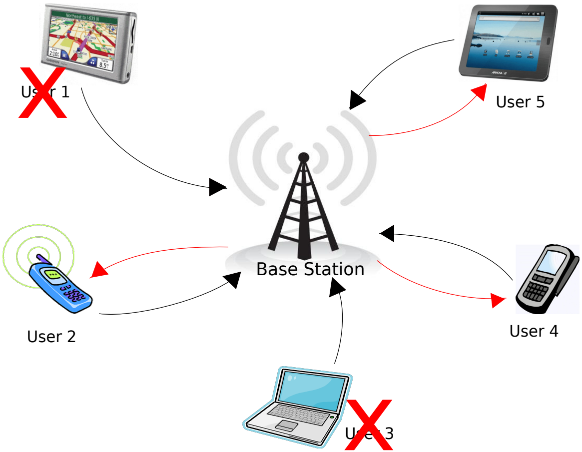
University of Pisa and BCAM

19 June 2013

Problem description



Problem description



Background

Objectives

- minimization of the expected time-average waiting cost;
- minimization of the expected time-average number of uncompleted jobs;
- maximization of some time-average fairness function across users.

Background

Objectives

- minimization of the expected time-average waiting cost;
- minimization of the expected time-average number of uncompleted jobs;
- maximization of some time-average fairness function across users.

Time varying Channel Quality Condition

- GOOD: Exploit the variation $\rightarrow$ *Be opportunistic*
- BAD: Difficult in analysis.

Background

Objectives

- minimization of the expected time-average waiting cost;
- minimization of the expected time-average number of uncompleted jobs;
- maximization of some time-average fairness function across users.

Time varying Channel Quality Condition

- GOOD: Exploit the variation $\rightarrow$ *Be opportunistic*
- BAD: Difficult in analysis.

History

- *MaxRate*: Myopically opportunistic [Knopp, Humblet 1995];
- *Proportionally Fair*: Fairly opportunistic [Chaponniere et al. 2002];
- *Best Condition Schedulers*: Smartly opportunistic [Ayesta et al. 2010].

What's new?

Markovian evolution of the channel quality conditions

$$\Downarrow$$

More realistic than IID evolution!;
(A step towards autoregressive distribution)

What's new?

Markovian evolution of the channel quality conditions

$$\Downarrow$$

More realistic than IID evolution!;
(A step towards autoregressive distribution)

Objective: Minimization of the expected time-average waiting cost;

What could be the structure of an optimal scheduler policy?

Approach

Approach:

1. *Multi-armed restless bandit problem formulation [Whittle 1988]*
    - Markov Decision Process (MDP) with infinite constraints ;

Approach

Approach:

1. *Multi-armed restless bandit problem formulation [Whittle 1988]*
   - Markov Decision Process (MDP) with infinite constraints ;
2. *Relaxation*
   1. Whittle Relaxation;
      - MDP with one constraint;
   2. Lagrangian Relaxation;
      - MDP without constraints;
   3. Decomposition into independent single user subproblem;
      - K one armed restless bandit problem without constraints;

Approach

Approach:

1. *Multi-armed restless bandit problem formulation [Whittle 1988]*
   - Markov Decision Process (MDP) with infinite constraints ;
2. *Relaxation*
   1. Whittle Relaxation;
      - MDP with one constraint;
   2. Lagrangian Relaxation;
      - MDP without constraints;
   3. Decomposition into independent single user subproblem;
      - K one armed restless bandit problem without constraints;
3. *Recomposition into a feasible heuristic solution.*

MDP Formulation

At every decision time an action must be chosen for every user in system:

Action space: $a \in \mathcal{A}_k = \{0, 1\}$

MDP Formulation

$$(\mathcal{N}_k, (\mathbf{W}_k^a)_{a \in \mathcal{A}_k}, (\mathbf{R}_k^a)_{a \in \mathcal{A}_k}, (\mathbf{P}_k^a)_{a \in \mathcal{A}_k})$$

MDP Formulation

At every decision time an action must be chosen for every user in system:

Action space: $a \in \mathcal{A}_k = \{0, 1\}$

MDP Formulation

$$(\mathcal{N}_k, (\mathbf{W}_k^a)_{a\in\mathcal{A}_k}, (\mathbf{R}_k^a)_{a\in\mathcal{A}_k}, (\mathbf{P}_k^a)_{a\in\mathcal{A}_k})$$

- *State Space*: $\mathcal{N}_k = \{0\} \cup \{1, ..., N_k\}$;

MDP Formulation

At every decision time an action must be chosen for every user in system:
Action space: $a \in \mathcal{A}_k = \{0, 1\}$

MDP Formulation

$$(\mathcal{N}_k, (\mathbf{W}_k^a)_{a \in \mathcal{A}_k}, (\mathbf{R}_k^a)_{a \in \mathcal{A}_k}, (\mathbf{P}_k^a)_{a \in \mathcal{A}_k})$$

- *State Space*: $\mathcal{N}_k = \{0\} \cup \{1, ..., N_k\}$;
- $\mathbf{W}_k^a := \left(W_{k,n}^a\right)_{n \in \mathcal{N}_k}$ is the one period expected *Work* required,

$$W_{k,n}^1 = 1, \qquad W_{k,n}^0 = 0;$$

MDP Formulation

At every decision time an action must be chosen for every user in system:

Action space: $a \in \mathcal{A}_k = \{0, 1\}$

MDP Formulation

$$(\mathcal{N}_k, (\mathbf{W}_k^a)_{a\in\mathcal{A}_k}, (\mathbf{R}_k^a)_{a\in\mathcal{A}_k}, (\mathbf{P}_k^a)_{a\in\mathcal{A}_k})$$

- *State Space*: $\mathcal{N}_k = \{0\} \cup \{1, ..., N_k\}$;
- $\mathbf{W}_k^a := \left(W_{k,n}^a\right)_{n\in\mathcal{N}_k}$ is the one period expected *Work* required,

$$W_{k,n}^1 = 1, \qquad W_{k,n}^0 = 0;$$

- $\mathbf{R}_k^a := \left(R_{k,n}^a\right)_{n\in\mathcal{N}_k}$ is the one period expected *Reward* earned,

$$R_{k,0}^a = 0, \qquad R_{k,n}^0 = -c_k, \qquad R_{k,n}^1 = -c_k(1 - \mu_{k,n});$$

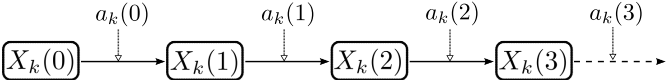
- $\mathbf{P}_k^a := \left(P_{k,n,m}^a\right)_{n,m\in\mathcal{N}_k}$ is the one period *State Transition Probability Matrix*,

$P_{k,n,m}^a$	(n,m)	$(n,0)$	$(0,m)$	$(0,0)$
$a=0$	$q_{k,n,m}$	0	0	1
$a=1$	$(1-\mu_{k_n})q_{k,n,m}$	$\mu_{k,n}$	0	1

- $\mathbf{P}_k^a := \left(P_{k,n,m}^a\right)_{n,m\in\mathcal{N}_k}$ is the one period *State Transition Probability Matrix*,

$P_{k,n,m}^a$	(n,m)	$(n,0)$	$(0,m)$	$(0,0)$
$a=0$	$q_{k,n,m}$	0	0	1
$a=1$	$(1-\mu_{k_n})q_{k,n,m}$	$\mu_{k,n}$	0	1

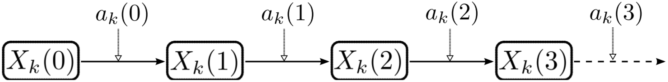
The user k evolution is captured by the state and action processes $(X_k(t), a_k(t))$:



- $\mathbf{P}_k^a := \left(P_{k,n,m}^a\right)_{n,m\in\mathcal{N}_k}$ is the one period *State Transition Probability Matrix*,

$P_{k,n,m}^a$	(n,m)	$(n,0)$	$(0,m)$	$(0,0)$
$a=0$	$q_{k,n,m}$	0	0	1
$a=1$	$(1-\mu_{k_n})q_{k,n,m}$	$\mu_{k,n}$	0	1

The user k evolution is captured by the state and action processes $(X_k(t), a_k(t))$:



Problem P_k

$$\max_{\tilde{\pi}_k\in\Pi_{\mathbf{X}_k,\mathbf{a}_k}} \mathbb{E}_0^{\tilde{\pi}_k}\left[\sum_{t=0}^{\infty}\beta^t R_{k,X_k(t)}^{a_k(t)} - \nu W_{k,X_k(t)}^{a_k(t)}\right]. \quad (1)$$

How to solve this one armed restless bandit problem?

$$\max_{\tilde{\pi}\in\Pi_{\mathbf{x},\mathbf{a}}} \mathbb{E}_0^{\tilde{\pi}}\left[\sum_{t=0}^{\infty}\beta^t R_{X(t)}^{a(t)} - \nu W_{X(t)}^{a(t)}\right]$$

How to solve this one armed restless bandit problem?

$$\max_{\tilde{\pi}\in\Pi_{\mathbf{x},\mathbf{a}}} \mathbb{E}_0^{\tilde{\pi}}\left[\sum_{t=0}^{\infty}\beta^t R_{X(t)}^{a(t)} - \nu W_{X(t)}^{a(t)}\right]$$

Definition

Threshold policy: $\forall \nu \in \mathbb{R}$ there exists a threshold state $n(\nu) \in \mathcal{N}$ such that

- it is optimal to serve the user in state n if $n \geq n(\nu)$,
- it is optimal not to serve the user in state n if $n < n(\nu)$.

How to solve this one armed restless bandit problem?

$$\max_{\tilde{\pi}\in\Pi_{\mathbf{x},\mathbf{a}}} \mathbb{E}_0^{\tilde{\pi}} \left[\sum_{t=0}^{\infty} \beta^t R_{X(t)}^{a(t)} - \nu W_{X(t)}^{a(t)} \right]$$

Definition

Threshold policy: $\forall \nu \in \mathbb{R}$ there exists a threshold state $n(\nu) \in \mathcal{N}$ such that

- it is optimal to serve the user in state n if $n \geq n(\nu)$,
- it is optimal not to serve the user in state n if $n < n(\nu)$.

Definition

Index policy: The problem (1) is indexable if there exist values $\nu_n^* \in \mathbb{R} \cup \{-\infty, \infty\} \quad \forall n \in \mathcal{N}$ such that

- it is optimal to serve the user in state n if $\nu_n^* \geq \nu$,
- it is optimal not to serve the user in state n if $\nu_n^* \leq \nu$.

Such values ν_n^* are called the (Whittle) index values, and define an optimal index policy for the problem.

Conjecture about indexability

Conjecture

The single user subproblem is indexable.

Conjecture about indexability

Conjecture

The single user subproblem is indexable.

- *Proved in particular cases*:
  1. IID evolution of channel quality conditions [Ayesta et al. 2010];
  2. Markovian evolution of channel quality conditions, with 2 states [Jacko 2011];
- *Massive numerical experiments*: Index are computed through the $\mathcal{AG}$-algorithm [Niño-Mora 2007]. It provides a certificate of indexability for a specified instance and the indexability test never failed.

Index values - General case

Highest index

Theorem

Under Conjecture (1), the index value $\nu_N = \frac{c\mu_N}{1-\beta}$ *and we have that* $\nu_N \geq \nu_n \ \forall n \in \mathcal{N}$, *i.e.* $\nu_N = \infty$ *under time average criterion.*

Index values - General case

Highest index

Theorem

Under Conjecture (1), the index value $\nu_N = \frac{c\mu_N}{1-\beta}$ *and we have that* $\nu_N \geq \nu_n \ \forall n \in \mathcal{N}$, *i.e.* $\nu_N = \infty$ *under time average criterion.*

2nd Highest index

Conjecture

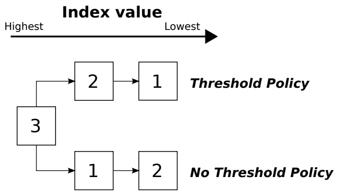
Under Conjecture (1), the second highest index value under time average criterion is of (every) state n which satisfies

$$n \in \underset{m \in \mathcal{N}' \setminus \{N\}}{\arg\max} \left\{ \frac{c\mu_m}{q_m^*(\mu_N - \mu_m)} \right\},$$

and the index value of such state(s) is the corresponding maximum.

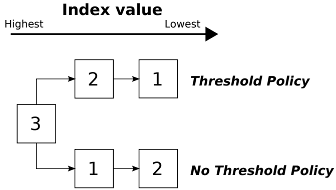
3 Channel Quality Conditions

The Conjecture about the 2nd highest index is proved.
An expression for the lowest index is also available.



3 Channel Quality Conditions

The Conjecture about the 2nd highest index is proved.
An expression for the lowest index is also available.



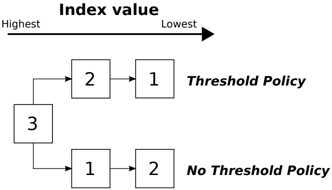
Example

$$\mathbf{Q} = \begin{pmatrix} 0.9 & 0.08 & 0.02 \\ 0.05 & 0.15 & 0.8 \\ 0.03 & 0.07 & 0.9 \end{pmatrix}, \quad \mu = \begin{pmatrix} 0.10 \\ 0.12 \\ 0.20 \end{pmatrix},$$

$$c = 1, \quad \beta = 0.999$$

3 Channel Quality Conditions

The Conjecture about the 2nd highest index is proved.
An expression for the lowest index is also available.



Example

$$\mathbf{Q} = \begin{pmatrix} 0.9 & 0.08 & 0.02 \\ 0.05 & 0.15 & 0.8 \\ 0.03 & 0.07 & 0.9 \end{pmatrix}, \quad \mu = \begin{pmatrix} 0.10 \\ 0.12 \\ 0.20 \end{pmatrix},$$

$$c = 1, \quad \beta = 0.999$$

In this case the indices obtained through the $\mathcal{AG}$-algorithm are

$$\nu_3^* = 200 \quad \nu_1^* = 3.372 \quad \nu_2^* = 2.089$$

Approximation through Steady State distribution

Theorem

Let us fix a bound M such that $\mu_3 \le M \le 1$. Then we have that the index value of state 1 and 2,

$$\nu_1^* = \frac{c\mu_1 + \mathcal{O}\left(M^2\right)}{\sum_{m=2,3} q_m^{SS}(\mu_m - \mu_1) + \mathcal{O}\left(M^2\right)}, \quad \nu_2^* = \frac{c\mu_2 + \mathcal{O}\left(M^2\right)}{q_3^{SS}(\mu_3 - \mu_2) + \mathcal{O}\left(M^2\right)}$$

M	Absolute Error	Relative Error
1	0.3880	14.08%
0.5	0.1854	7.424%
0.3	0.1273	4.498%
0.1	0.0399	1.571%
0.05	0.0237	0.828%
0.01	0.0051	0.176%
0.001	0.0005	0.017%

Table: Mean absolute and relative errors of the approximation of ν_1^*.

- Absolute errors are linear in M, hyperbolic in the job size;
- Mp3 5MB $\rightarrow M \sim 0.0003$.

Proposed Scheduler

Scheduling problem

PI* scheduler

Serve the C users with the highest index at the beginning of every slot of time.

Proposed Scheduler

Scheduling problem

PI* scheduler

Serve the C users with the highest index at the beginning of every slot of time.

Possible generalization

- PI^{SS}, where
$$\nu^*_{k,n} = \frac{c_k \mu_{k,n}}{\sum_{m>n} q^{\text{SS}}_{k,m}(\mu_{k,m} - \mu_{k,n})}.$$
- $\text{PI}^{\mathcal{AG}}$, where the indices are computed through the $\mathcal{AG}$-algorithm.

Conclusion

Simulations:

The PI^* scheduler has been tested in various scenarios and compared with the other schedulers existing in literature: the results are definitely interesting.

Conclusion

Simulations:

The PI^* scheduler has been tested in various scenarios and compared with the other schedulers existing in literature: the results are definitely interesting.

Open questions:

Threshold policies

- Sufficient conditions;
- Properties of the PI^{SS} scheduler:
  - Fairness;
  - N-state generalization.

Model extensions

- Lack of knowledge of some parameters;
- Possibility of abandonments;
- Multiple base stations.

Thank you for your attention!

Grazie per l'attenzione!

Numerical Simulation

Other schedulers

- the $\mathbf{c}\boldsymbol{\mu}$ rule, i.e. $\nu^{c\mu}_{k,n} = c_k \mu_{k,n}$, [Knopp, Humblet 1995];
- the **Relatively Best** rule, i.e. $\nu^{RB}_{k,n} = \frac{c_k \mu_{k,n}}{\sum_{m=1}^{N_k} q^{SS}_{k,m} \mu_{k,m}}$, [Chaponniere et al. 2002];
- the **Proportionally Best** rule, i.e. $\nu^{PB}_{k,n} = \frac{c_k \mu_{k,n}}{\mu_{k,N_k}}$, [Aalto et al. 2010];
- the **Score Based** rule, i.e. $\nu^{SB}_{k,n} = c_k \sum_{m=1}^{n} q^{SS}_{k,m}$ [Bonald 2004];

Traffic intensity

$$\varrho_k := \frac{\lambda_k}{\mu_{k,N_k}} \qquad \varrho := \sum_{k \in K} \varrho_k$$

.

- Simulations time range: 5 to 15 minutes.
- Time slot $\sim$ 0.167msec.

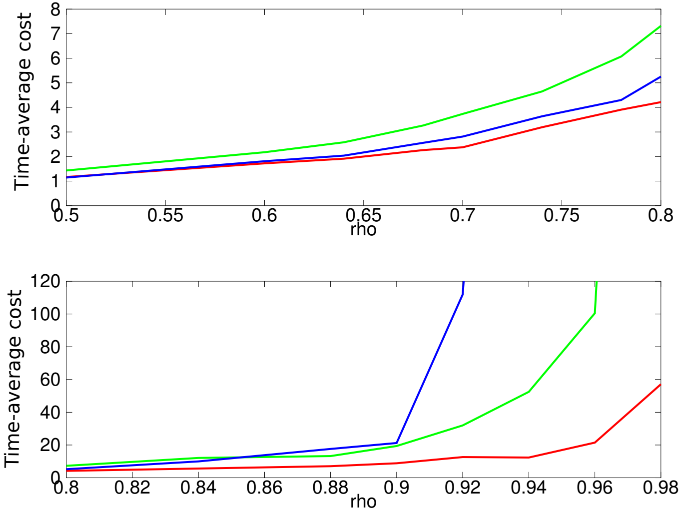
Scenario 1

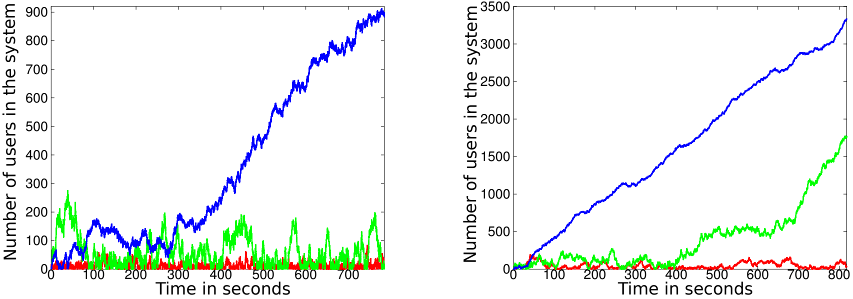
Example

	Job	Service rate	c	Q
1	0.5MB	$\begin{pmatrix} 53.76 \\ 67.2 \\ 80.64 \end{pmatrix}$ Mb/sec	1	$\begin{pmatrix} 0.4 & 0.21 & 0.39 \\ 0.48 & 0.5 & 0.02 \\ 0.26 & 0.3 & 0.44 \end{pmatrix}$
2	0.5MB	$\begin{pmatrix} 4.2 \\ 26.88 \\ 33.6 \end{pmatrix}$ Mb/sec	1	$\begin{pmatrix} 0.34 & 0.35 & 0.31 \\ 0.27 & 0.45 & 0.28 \\ 0.45 & 0.15 & 0.4 \end{pmatrix}$

- In this scenario the PI*, the PB and the SB rules lead to the same policy;

Figure: PI*,SB,PB (red), RB (green), $c\mu$ (blue).





(a) $\varrho = 0.94$.

(b) $\varrho = 0.98$.

Figure: PI*,SB,PB (red), RB (green), $c\mu$ (blue).

Scenario 2

Example

	Job	*Service rate*	*c*	**Q**
1	0.5MB	$\begin{pmatrix} 26.88 \\ 53.76 \\ 80.64 \end{pmatrix}$ Mb/sec	10	$\begin{pmatrix} 0.38 & 0.20 & 0.42 \\ 0.43 & 0.19 & 0.38 \\ 0.48 & 0.27 & 0.25 \end{pmatrix}$
2	0.5MB	$\begin{pmatrix} 26.88 \\ 53.76 \\ 80.64 \end{pmatrix}$ Mb/sec	1	$\begin{pmatrix} 0.38 & 0.20 & 0.42 \\ 0.43 & 0.19 & 0.38 \\ 0.48 & 0.27 & 0.25 \end{pmatrix}$

- In this scenario the PB, the RB, the SB and the $c\mu$ rules lead to the same policy;

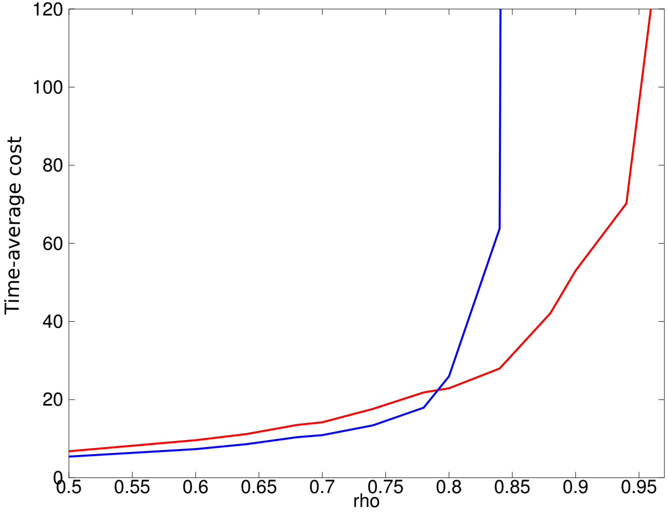


Figure: PI* (red), $c\mu$,SB,PB,RB (blue).

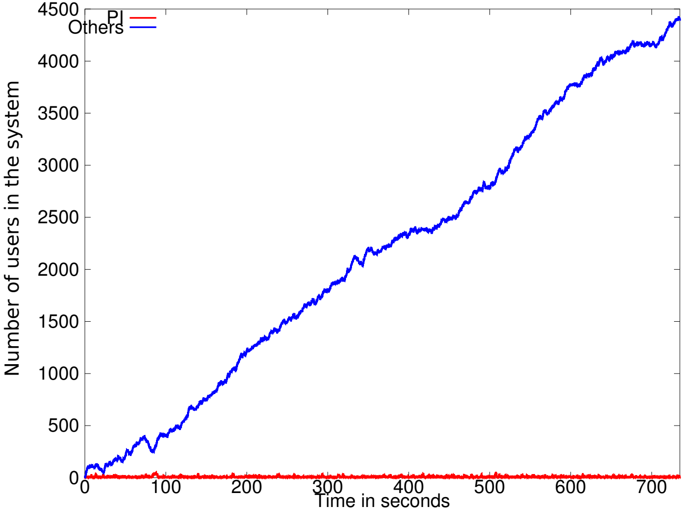
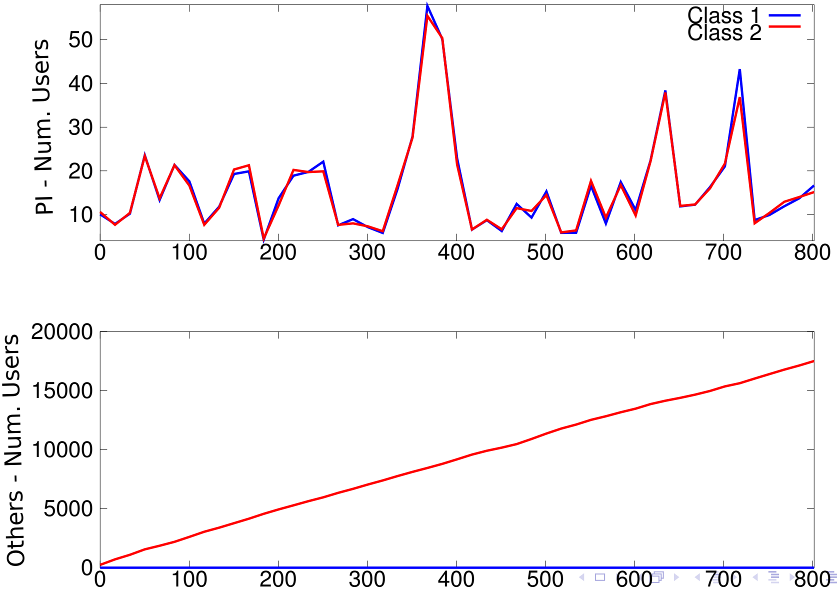


Figure: $\varrho = 0.88$. PI* (red), $c\mu$,SB,PB,RB (blue)


Class 1
Class 2
PI - Num. Users
50
40
30
20
10
0
100
200
300
400
500
600
700
800
Others - Num. Users
20000
15000
10000
5000
0

Sufficient conditions for the Threshold Policy optimality

Theorem

If Conjecture (1) holds for the problem with $N = 3$ *and* $\beta = 1$*, then we have that* $q_{13} \geq q_{23}$ *implies that* $\nu_2^* > \nu_1^*$*, i.e., the problem is solvable by threshold policies.*

Theorem

Let us denote by $\Delta := \min\{\mu_3 - \mu_2, \mu_2 - \mu_1\}$ *and* $1 > M \geq \mu_3$ *. If Conjecture (1) holds for the problem with* $N = 3$ *and* $\beta = 1$*, then we have that* $\Delta \geq \frac{M^2}{3(1-M)} =: \varepsilon$ *implies that* $\nu_2^* > \nu_1^*$*, i.e., the problem is solvable by threshold policies.*

M	1	0.5	0.3	0.1	0.05	0.01
ε	$+\infty$	0.16667	0.04286	0.00370	0.00088	0.00003