Contents lists available at ScienceDirect
Performance Evaluation
journal homepage: www.elsevier.com/locate/peva

A modeling framework for optimizing the flow-level scheduling with time-varying channels[☆]

Urtzi Ayesta [a,b], Martin Erausquin [c], Peter Jacko [a,*]

[a] BCAM — Basque Center for Applied Mathematics, 48170 Bilbao, Spain
[b] IKERBASQUE — Basque Foundation for Science, 48011 Bilbao, Spain
[c] UPV/EHU — University of the Basque Country, Bilbao, Spain

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Multi-armed restless bandit problem
Markov Decision Process
Stochastic dynamic programming
Opportunistic scheduling
Optimal scheduling
$c\mu$-rule
Index policy

ABSTRACT

We introduce a comprehensive modeling framework for the problem of scheduling a finite number of finite-length jobs where the available service rate is time-varying. The main motivation comes from wireless data networks where the service rate of each user varies randomly due to fading. We employ recent advances on the restless bandit problem that allow us to obtain an opportunistic scheduling rule for the system without arrivals. When the objective is to minimize the mean number of users in the system or to minimize the mean waiting time, we obtain a priority-based policy which we call the "Potential Improvement" (PI) rule, since the priority index equals the ratio between the current available service rate and the expected potential improvement of the service rate. We also show that for certain objective functions, the index rule takes the form of known opportunistic scheduling rules like "Relatively Best" (RB) or "Proportionally Best" (PB). Thus our model provides a formal justification for the deployment of opportunistic scheduling rules in order to improve the flow-level performance in the presence of time-varying capacities. We further analyze the performance of the PI rule in the presence of randomly arriving users. When the service rates are constant, PI is equivalent to the $c\mu$-rule, which is known to be optimal with any distribution of arrivals. Using a recent characterization for the stability region of flow-level scheduling rules under random arrivals, we show that PI achieves the maximum stability region. We perform numerical experiments in a wide range of scenarios and compare the performance of PI with other popular disciplines like RB, PB, Score-Based (SB) and the $c\mu$-rule. Our results show that RB, PB, SB or the $c\mu$-rule might outperform the others depending on the scenario, but regardless of this, the performance of PI is always superior or equivalent to the best of these opportunistic rules.

1. Introduction

In this paper we are interested in studying the problem of optimal flow-level scheduling in time-varying systems. The main motivation comes from a downlink wireless data network where the download rate of each user fluctuates over time due to fading. Scheduling in such a time-varying system has received a lot of attention in recent years. Broadly speaking, researchers have explored this problem by studying packet-level and flow-level models. In packet-level formulations it is

☆ Research partly supported by grant MEC08/85 of the Ministry of Education (Spanish Government) and grant PI2010-2 of the Department of Education and Research (Basque Government).

* Corresponding author.
E-mail addresses: ayesta@bcamath.org (U. Ayesta), martin.erausquin@ehu.es (M. Erausquin), jacko@bcamath.org, peter.jacko@gmail.com (P. Jacko).

0166-5316/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.peva.2010.08.015

typically assumed that the scheduler can keep track of the number of packets in the queue and the objective is typically to minimize the queue length of the various users/classes. We refer for example to [1–7] for this line of research. Other researchers have analyzed the performance of scheduling disciplines at the flow level in order to better capture the performance perceived by users, see for example [8–12]. In this line of work the objective becomes to minimize the number of jobs in the system. For a review on flow-level modeling we refer to [13,14].

The Proportionally Fair scheduler (with its variant which we call the Relatively Best scheduler) is one of the most known scheduling disciplines in this time-varying context (see for example [15,9,2,11]). The Proportionally Fair (PF) scheduler is implemented in several time-slotted networks and at every time slot serves the user with the highest ratio of the current service rate and the obtained time-average throughput, while the Relatively Best (RB) scheduler serves the user with the highest ratio of the current service rate and the time-average expected service rate, and is broadly analyzed in the literature. It was shown in [11] that the two schedulers are roughly equivalent under certain system conditions. [15] showed that the Proportionally Fair scheduler maximizes the aggregate logarithmic utility of obtained throughput in a fixed population of permanent users.

Flow-level models with time-varying service rates are extremely difficult to solve for optimality. There are several important works (see for example [11]) that show that so-called "opportunistic schedulers" perform well, but to the best of our knowledge there is no mathematical model that shows that opportunistic schedulers are (even asymptotically) optimal in the context of users with finite-length jobs. A rule is called opportunistic (or channel-aware) if it takes advantage of the channel fluctuations by serving a user whose channel condition is "good" in some sense with respect to its own statistical behavior. The main objective of this paper is to develop a mathematical model that shows that opportunistic schedulers are optimal for the flow-level scheduling of users in a time-varying system.

The general problem statement is as follows. We wish to schedule a finite number of users. The available service rate of each user in a slot is an independent and identically distributed (i.i.d.) user-dependent random variable. The common resource can be preemptively allocated to at most one user in every time slot and the selected user may complete the service and depart from the system with a probability that depends on the current channel condition. Every user in the system incurs a cost per slot (which may depend on the class and on whether the user is selected or not) and the objective is to minimize the expected total discounted or undiscounted cost. We formulate the problem as a Markov Decision Process (MDP). The state of each user is a two dimensional random variable. One dimension captures whether service has been completed or not, and the other one is the current available service rate of the user. A crucial aspect of our model is that the cost of a user k depends on a parameter γ_k which allows us to model a wide variety of cost functions.

This formulation can be seen as a generalization of the job sequencing problem with geometrically distributed service times solved in [16]. It falls within the class known as *multi-armed restless bandit problems*, an optimization problem extremely difficult and PSPACE-hard even in its deterministic variant [17], and is typically attributed to suffer from the *curse of dimensionality*. The term "restless" refers to the fact that the state of all users in the system varies in time regardless of whether they are served or not. Restless problems can be solved analytically only in a few cases (typically with largely restricted dynamics), and this explains to some extent why the results on scheduling in time-varying scenarios are so scarce in queueing theory.

In order to tackle our restless bandit formulation we follow the approach of Whittle [18] that allows us to derive a nearly-optimal heuristic scheduling rule by computing an index policy as described in [19,20]. The main idea is to relax the sample path constraint (that imposes that only one user is served at a time) by letting the average discounted number of users served in a slot to be one. This relaxation simplifies significantly the problem since it allows to decompose it into single-user subproblems. The optimal policy of the relaxed formulation becomes now of index type (as in the multi-armed classic bandit problem), that is, under certain conditions, we can calculate for each user a certain index value called *price* (that depends only on the current state and generalizes the Gittins index) such that the optimal scheduler serves the users in every slot with an actual price higher than a threshold (the value of the threshold ensures that the average discounted number of users served in a slot is precisely one). The optimal policy for the relaxed problem need not be feasible for the original problem, but allows us to construct a heuristic index policy for the original problem by serving the user with the currently highest price. This heuristic rule is feasible (only one user is served at a time) and is typically reported to have an extremely good performance [19]. In addition, it was shown in [21] that index policies approach optimality as the number of users grows to infinity (if certain additional assumptions hold).

As our main contribution we design a scheduling rule as an index policy for the problem at hand which depends on the vector of parameters γ_k. In the important case $\gamma_k = 1$ for all users k then every user in the system incurs a holding cost per slot while present in the system, thus, the objective becomes to minimize the average discounted holding cost. Then, the time-average (i.e., undiscounted) limit of the index rule has a very appealing form and is equal to the ratio between the current available service rate multiplied by the holding cost and the expected potential improvement of the service rate. We call this (opportunistic) index policy the Potential Improvement (PI) rule. If the available rate for each user is constant, PI is equivalent to the $c\mu$-rule, which is known to be optimal in this setting. When the γ_k equals the inverse of the expected service rate of user k and the discount factor is not too large, then the index policy approximately selects the user with highest ratio between the instantaneous rate and the expected rate, the policy RB defined above. When γ_k equals the inverse of the maximum transmission rate of user k and the discount factor is not too large, then the index policy selects approximately

Table 1
The transmission rates and the corresponding channel condition probabilities in the CDMA 1xEV-DO wireless network, as reported in [24], and values used in the simulation study for the two classes in Scenarios 1 & 2.

Transmission rate (kb/s)	38.4	76.8	102.6	153.6	204.8	307.2	614.4	921.6	1228.8	1843.2	2457.6
Probabilities in CDMA	0.00	0.01	0.04	0.08	0.15	0.24	0.18	0.09	0.12	0.05	0.04
Probabilities for class 1	0	0	0.05	0	0.23	0	0.42	0	0.21	0	0.09
Probabilities for class 2	0	0	0.15	0	0.33	0	0.52	0	0	0	0

the user with the highest ratio between the instantaneous rate and the maximum rate, a policy known as Proportionally Best (PB) [22]. Finally, when $\gamma_k = 0$, then the problem is to determine the policy that maximizes the flow-level throughput, and we show that if the discount factor is not too large, then the discounted variant of the PI rule is the index solution for this case. However, our results suggest that opportunistic rules may not be appropriate in some situations in the time-average case.

We perform numerical experiments to assess the performance of PI in the presence of randomly arriving users. For a wide range of parameter values we compare the performance of PI with RB [11], PB [22], Score-Based (SB) [9] and the $c\mu$-rule. As performance criterion we take the mean number of users in the system. Our results show that depending on the scenario and on the load in the system, RB, PB, SB or the $c\mu$-rule might outperform the others, but interestingly the performance of PI is always comparable or superior to the best of the opportunistic rules. Our simulations indicate that in many scenarios PI outperforms the other policies in a stochastic ordering sense, and not only in the mean. We also investigate stability issues. The results of [22,23] imply that PI is stable under the maximum stability region, while SB and PB are such only if these rules are independent of the holding costs. Our simulations (in which the holding costs are set to 1) show that indeed depending on the scenario, RB or $c\mu$ fail to achieve maximum stability, whereas PI and SB always do.

The rest of the paper is organized as follows. In Section 2 we present the problem description. Section 3 formulates the problem as a Markov Decision Process. Section 4 contains the main contribution of this paper, that is, the analytical resolution of the relaxed problem and the heuristic opportunistic rule for the original stochastic optimization formulation. In Section 5 we discuss the PI scheduling rule and its relation to alternative known rules. Section 6 presents several properties of PI in systems with random arrivals and Section 7 presents the numerical simulations. For the sake of readability, proofs are postponed to the Appendix.

2. Problem description

In this section we consider the discrete-time job sequencing problem (without arrivals), in which we allow for time-varying service rate due to the variation in the users' wireless channel quality. Consider $K - 1$ jobs waiting for service at a base station that can serve one job at a time by transmitting a data flow through a dedicated channel to the corresponding user. There is a 1: 1: 1 correspondence between jobs, users, and channels, so we use these terms interchangeably, and we also use the expression *job-channel-user triple*. Let $c_k > 0$ be the holding cost per period incurred for user k waiting while the job transmission is not completed. Suppose that channel k can take quality conditions from a non-empty finite set $\mathcal{N}'_k := \{1, 2, \ldots, N_k\}$ so that condition $n \in \mathcal{N}'_k$ happens in a certain period with probability $q_{k,n}$, having $\sum_{n\in\mathcal{N}'_k} q_{k,n} = 1$. The transmission quality condition of each channel evolves randomly and independently of the other channels, of the decisions of the base station, and of the channel quality evolution history. Thus, the probabilities $q_{k,n}$ can be seen as the steady-state distribution of the channel conditions.

Further, under channel condition n, the probability that the service of job k is completed within one period if being transmitted is $\mu_{k,n}$. Without loss of generality we assume that the channel condition labels are ordered so that $0 \le \mu_{k,1} \le \mu_{k,2} \le \cdots \le \mu_{k,N_k} \le 1$. To ensure that eventually all users leave the system we assume that for every user k, $q_{k,n}\mu_{k,n} \neq 0$ for some channel condition $n \in \mathcal{N}'_k$. Thus, the job-channel-user triples are independent of each other and compete for the bandwidth of the base station.

If the base station is allocated to a user whose job has already been completed, then no transmission occurs. We incorporate the following parameter to make the problem extensively flexible to incorporate additional conditions or options. It is allowed to allocate the base station resources to an alternative task (such as battery recharging or service maintenance), for which we obtain an *alternative-task reward* κ per period. For instance, the role of this alternative task with a positive κ could be to turn off the base station allocation when all the users have very bad channel conditions. On the other hand, by setting this parameter to a negative value we may force the base station to be non-idling (whenever there are waiting jobs), or it can be simply set to zero narrowing the focus to the traditional problem.

The server is assumed to be preemptive (i.e., the service of a job or the alternative task can be interrupted at the beginning of any time period even if not completed). Thus, the base station decides at the beginning of every period to which user (or to the alternative task) it should be allocated during that period.

Remark 1 (*Modeling of a Wireless Data Network*). Assume a time-slotted system (with slot duration denoted by $t_c = 1.67$ ms) as the CDMA 1xEV-DO [24]. The available service rate of each user fluctuates due to fading, and as a consequence, it varies from one slot to another in an independent fashion. In the particular case of the CDMA 1xEV-DO system the service

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015

rates that are available are presented in Table 1 (the probabilities are only informative). Let the service requirement (in bits) of user k be a geometrically distributed random variable denoted by B_k, and let $\mathbb{E}[B_k]$ denote its expectation. Let Δ denote the amount of bits transferred in one slot under the current channel condition. Note that in practice Δ will vary from slot to slot depending on the channel condition. Then the probability that a user leaves the system is approximately $\mathbb{P}(b \le B_k \le b+\Delta | B_k > b) \approx \Delta/\mathbb{E}[B_k]$, which does not depend on the attained service b (memoryless property of geometric distribution). This expression becomes asymptotically exact as the ratio $\Delta/\mathbb{E}[B_k]$ becomes small, that is, when the mean service requirement (in bits) of a user is much larger than the amount of bits that can be served in one slot. Let $s_{k,n}$ denote the service rate (in bits per second) of user k while in channel condition n, and let $q_{k,n}$ be probability that the channel is in condition n in a slot. We can then conclude that in a wireless data network with geometric service requirements the departure probability of a class-k user under channel condition n can be approximated by

$$\mu_{k,n} := \frac{s_{k,n} \cdot t_c}{\mathbb{E}[B_k]}. \quad (1)$$

2.1. γ-parameterized objective

In order to provide a comprehensive modeling framework for opportunistic scheduling rules, we introduce a parameterized objective function. We suppose that there is a parameter $\gamma_k \ge 0$ associated with each job k, see definitions in Section 3. The joint goal is to minimize the expected aggregate γ-parameterized holding cost minus the alternative-task reward, over an infinite horizon.

The problem with $\gamma_k = 1$ for all k corresponds to minimization of the mean waiting time and minimization of the mean number of jobs in the system. Interestingly, $\gamma_k = 0$ for all k leads to throughput maximization. Moreover, PB and RB rules can be approximately recovered as prices (i.e., index-based solutions) to the problem for other values of γ. We leave the more detailed discussion to Section 5.

3. MDP formulation

In this section we present an MDP formulation of the problem under the general γ-parameterized objective. The setting fits the *multi-armed restless bandit problem* [18,25] and in this paper we follow the framework of its generalization called the *dynamic and stochastic resource capacity allocation problem* introduced in [26].

Consider the time slotted into epochs $t \in \mathcal{T} := \{0, 1, 2, \ldots\}$ at which decisions can be made. The time epoch t corresponds to the beginning of the time period t. Suppose that at $t = 0$ there are $K - 1 \ge 1$ users awaiting transmission from the base station that at each epoch chooses (at most) one of the users to which it transmits a data stream through a user-dedicated wireless channel. If no user is chosen, then the base station is allocated to the alternative task, i.e., there are K competing options, labeled by $k \in \mathcal{K}$. Thus, the base station is allocated to exactly one option at a time.

3.1. Jobs, channels, and users

Every user $k = 1, 2, \ldots, K - 1$ can be allocated either zero or full capacity of the base station. We denote by $\mathcal{A} := \{0, 1\}$ the *action space*, i.e., the set of allowable levels of capacity allocation. Here, action 0 means allocating zero capacity (i.e., "not serving"), and action 1 means allocating full capacity (i.e., "serving"). This action space is the same for every user k.

Each job-channel-user triple k is defined independently of other jobs, channels and users as the tuple $\left(\mathcal{N}_k, \left(\boldsymbol{W}_k^a\right)_{a\in\mathcal{A}}, \left(\boldsymbol{R}_k^a\right)_{a\in\mathcal{A}}, \left(\boldsymbol{P}_k^a\right)_{a\in\mathcal{A}}\right)$, where

- $\mathcal{N}_k := \{0\} \cup \mathcal{N}_k'$ is the *state space*, where state 0 represents a job already completed, and $\mathcal{N}_k' := \{1, \ldots, N_k\}$ is the set of possible quality conditions of channel k provided the job is uncompleted;
- $\boldsymbol{W}_k^a := \left(W_{k,n}^a\right)_{n\in\mathcal{N}_k}$, where $W_{k,n}^a$ is the (expected) one-period capacity consumption, or *work* required by user k at state n if action a is decided at the beginning of a period; in particular, for any $n \in \mathcal{N}_k$,

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

- $\boldsymbol{R}_k^a := \left(R_{k,n}^a\right)_{n\in\mathcal{N}_k}$, where $R_{k,n}^a$ is the expected one-period *reward* earned by user k at state n if action a is decided at the beginning of a period; in particular, for any $n \in \mathcal{N}_k'$,

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

- $\boldsymbol{P}_k^a := \left(p_{k,n,m}^a\right)_{n,m\in\mathcal{N}_k}$ is the user-k stationary one-period *state-transition probability matrix* if action a is decided at the beginning of a period, i.e., $p_{k,n,m}^a$ is the probability of moving to state m from state n under action a; in particular, denoting by $\widetilde{\mu}_{k,n} := 1 - \mu_{k,n}$,

$$\boldsymbol{P}_k^1 := \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ N_k \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & \cdots & N_k \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ \mu_{k,1} & \widetilde{\mu}_{k,1} q_{k,1} & \cdots & \widetilde{\mu}_{k,1} q_{k,N_k} \\ \mu_{k,2} & \widetilde{\mu}_{k,2} q_{k,1} & \cdots & \widetilde{\mu}_{k,2} q_{k,N_k} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{k,N_k} & \widetilde{\mu}_{k,N_k} q_{k,1} & \cdots & \widetilde{\mu}_{k,N_k} q_{k,N_k} \end{pmatrix} \end{array}, \quad \boldsymbol{P}_k^0 := \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ N_k \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & \cdots & N_k \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & q_{k,1} & q_{k,2} & \cdots & q_{k,N_k} \\ 0 & q_{k,1} & q_{k,2} & \cdots & q_{k,N_k} \\ & & & \ddots & \\ 0 & q_{k,1} & q_{k,2} & \ddots & q_{k,N_k} \\ 0 & q_{k,1} & q_{k,2} & \cdots & q_{k,N_k} \end{pmatrix} \end{array}.$$

The dynamics of user k is thus captured by the *state process* $X_k(\cdot)$ and the *action process* $a_k(\cdot)$, which correspond to state $X_k(t) \in \mathcal{N}_k$ and action $a_k(t) \in \mathcal{A}$ at all time epochs $t \in \mathcal{T}$. As a result of deciding action $a_k(t)$ in state $X_k(t)$ at time epoch t, the user k consumes the allocated capacity, earns the reward, and evolves its state for the time epoch $t + 1$.

We have the same action space $\mathcal{A}$ available at every state, which assures a technically useful property that $\boldsymbol{W}_k^a, \boldsymbol{R}_k^a, \boldsymbol{P}_k^a$ are defined in the same dimensions under any $a \in \mathcal{A}$. Note that state 0 is absorbing.

3.2. Alternative task

We model the alternative task as a *static user* with a single state 0 and with reward κ if served. i.e., such a user $k = K$ is defined by $\mathcal{N}_K := \{0\}$, $W_{K,0}^a := a$, $R_{K,0}^a := \kappa a$, $p_{K,0,0}^a := 1$ for all $a \in \mathcal{A}$.

3.3. A unified optimization criterion

Before describing the problem we first define an averaging operator that will allow us to discuss the infinite-horizon problem under the traditional β-discounted criterion and the time-average criterion in parallel. Let $\Pi_{X,a}$ be the set of all the policies that for each time epoch t decide (possibly *randomized*) action $a(t)$ based only on the state-process history $X(0), X(1), \dots, X(t)$ and on the action-process history $a(0), a(1), \dots, a(t-1)$ (i.e., *non-anticipative*). Let $\mathbb{E}_\tau^\pi$ denote the expectation over the state process $X(\cdot)$ and over the action process $a(\cdot)$, conditioned on the state-process history $X(0), X(1), \dots, X(\tau)$ and on policy π.

Consider any expected one-period quantity $Q_{X(t)}^{a(t)}$ that depends on state $X(t)$ and on action $a(t)$ at any time epoch t. For any policy $\pi \in \Pi_{X,a}$, any initial time epoch $\tau \in \mathcal{T}$, and any *discount factor* $0 \le \beta \le 1$ we define the infinite-horizon *β-average quantity*[1]

$$\mathbb{B}_\tau^\pi \left[Q_{X(\cdot)}^{a(\cdot)}, \beta, \infty \right] := \lim_{T \to \infty} \frac{\sum_{t=\tau}^{T-1} \beta^{t-\tau} \mathbb{E}_\tau^\pi \left[Q_{X(t)}^{a(t)} \right]}{\sum_{t=\tau}^{T-1} \beta^{t-\tau}}. \tag{2}$$

The β-average quantity recovers the traditionally considered quantities in the following three cases:

- *expected time-average quantity* when $\beta = 1$;
- *expected total β-discounted quantity*, scaled by constant $1 - \beta$, when $0 < \beta < 1$;
- *myopic quantity* when $\beta = 0$.

Thus, when $\beta = 1$, the problem is formulated under the *time-average criterion*, whereas when $0 < \beta < 1$ the problem is considered under the *β-discounted criterion*. The remaining case when $\beta = 0$ reduces to a static problem and hence is considered in order to define a *myopic policy*. In the following we consider the discount factor β to be fixed and the horizon to be infinite, therefore we omit them in the notation and write briefly $\mathbb{B}_\tau^\pi \left[Q_{X(\cdot)}^{a(\cdot)} \right]$.

3.4. Optimization problem

Now we describe in more detail the problem we consider and formulate it below. Let $\Pi_{\boldsymbol{X},\boldsymbol{a}}$ be the space of randomized and non-anticipative policies depending on the joint state-process $\boldsymbol{X}(\cdot) := (X_k(\cdot))_{k \in \mathcal{K}}$ and deciding the joint action-process $\boldsymbol{a}(\cdot) := (a_k(\cdot))_{k \in \mathcal{K}}$, i.e., $\Pi_{\boldsymbol{X},\boldsymbol{a}}$ is the *joint policy space*.

[1] For definiteness, we consider $\beta^0 = 1$ for $\beta = 0$.

For any discount factor β, the problem is to find a joint policy $\boldsymbol{\pi}$ maximizing the *objective* given by the β-average aggregate reward starting from the initial time epoch 0 subject to the family of *sample path* allocation constraints, i.e.,

$$\max_{\boldsymbol{\pi}\in\Pi_{\boldsymbol{X},\boldsymbol{a}}} \mathbb{B}_0^{\boldsymbol{\pi}}\left[\sum_{k\in\mathcal{K}} R_{k,X_k(\cdot)}^{a_k(\cdot)}\right] \tag{P}$$

$$\text{subject to } \mathbb{E}_t^{\boldsymbol{\pi}}\left[\sum_{k\in\mathcal{K}} a_k(t)\right] = 1, \quad \text{for all } t \in \mathcal{T}.$$

Note that the constraint could equivalently be expressed by restricting $\Pi_{\boldsymbol{X},\boldsymbol{a}}$ to policies satisfying $\sum_{k\in\mathcal{K}} a_k(t) = 1$ for any possible joint state–process history $\boldsymbol{X}(0), \boldsymbol{X}(1), \ldots, \boldsymbol{X}(t)$, for all $t \in \mathcal{T}$.

4. Solution

Problem (P) can be relaxed as proposed in [18], which is further approached by the Lagrangian methods and can be decomposed into a parameterized (bi-criteria) optimization problem as we describe next (for more details see [26]). Notice that any joint policy $\boldsymbol{\pi} \in \Pi_{\boldsymbol{X},\boldsymbol{a}}$ defines a set of single-user policies $\widetilde{\pi}_k$ for all $k \in \mathcal{K}$, where $\widetilde{\pi}_k$ is a randomized and non-anticipative policy depending on the *joint* state-process $\boldsymbol{X}(\cdot)$ and deciding the *user*-k action-process $a_k(\cdot)$. We will write $\widetilde{\pi}_k \in \Pi_{\boldsymbol{X},a_k}$. We will therefore study the user-k subproblem

$$\max_{\widetilde{\pi}_k\in\Pi_{\boldsymbol{X},a_k}} \mathbb{B}_0^{\widetilde{\pi}_k}\left[R_{k,X_k(\cdot)}^{a_k(\cdot)} - \nu W_{k,X_k(\cdot)}^{a_k(\cdot)}\right]. \tag{3}$$

The main idea of our approach is to identify a set of optimal policies $\widetilde{\pi}_k^*$ for (3) for each $k \in \mathcal{K}$, and using them to construct a joint policy $\boldsymbol{\pi}$, feasible though not necessarily optimal for problem (P).

4.1. Optimal solution to single-user subproblem via prices

Problem (3) falls into the framework of *restless bandits* and can be optimally solved by assigning a set of prices $\nu_{k,n}$ to each state $n \in \mathcal{N}_k$ under the so-called *indexability* condition [19,20].

Let us denote for channel condition $n \in \mathcal{N}_k'$ of user $k \le K-1$,

$$\nu_{k,n} := \frac{c_k\mu_{k,n}\left(1-\beta+\beta\gamma_k\right)}{(1-\beta)+\beta\sum\limits_{m>n} q_{k,m}(\mu_{k,m}-\mu_{k,n})}, \quad \nu_{k,0} := 0, \tag{4}$$

and for the alternative task $k = K$, $\nu_{K,0} := \kappa$. The following are the main results of this paper.

Proposition 1 (*Optimality of Threshold Policies*)**.**

(i) *If $\nu > 0$, then there exists $n \in \mathcal{N}_k \cup \{-1\}$ such that threshold policy transmitting in channel conditions $\mathcal{S}_{N-n} := \{m \in \mathcal{N}_k : m > n\}$ and not transmitting otherwise is optimal for problem (3).*

(ii) *If $\gamma_k \ge \mu_{k,N}$ or $\beta \le 1/(1+\mu_{k,N}-\gamma_k)$ holds for user $k \le K-1$, then for every real-valued ν there exists $n \in \mathcal{N}_k \cup \{-1\}$ such that threshold policy transmitting in channel conditions $\mathcal{S}_{N-n} := \{m \in \mathcal{N}_k : m > n\}$ and not transmitting otherwise is optimal for problem (3).*

Proposition 2 (*Indexability*)**.** *Suppose that $\gamma_k \ge \mu_{k,N}$ or $\beta \le 1/(1+\mu_{k,N}-\gamma_k)$ holds for user $k \le K-1$. The following holds for problem (3):*

(i) *if $\nu \le \nu_{k,n}$, then it is optimal to transmit under user's k channel condition $n \in \mathcal{N}_k'$;*
(ii) *if $\nu \ge \nu_{k,n}$, then it is optimal not to transmit under user's k channel condition $n \in \mathcal{N}_k'$;*
(iii) *if $\nu \le \nu_{k,0}$, then it is optimal to transmit when the job k is already completed (i.e., when $n = 0$);*
(iv) *if $\nu \ge \nu_{k,0}$, then it is optimal not to transmit when the job k is already completed (i.e., when $n = 0$);*
(v) *if $\nu \le \nu_{K,0}$, then it is optimal to allocate the base station to the alternative task K;*
(vi) *if $\nu \ge \nu_{K,0}$, then it is optimal not to allocate the base station to the alternative task K.*

4.2. Scheduling rule for original problem

Since the original problem requires to allocate the base station to exactly one option (to one of the users or to the alternative task), then at any time epoch t we propose to allocate the base station to the option $k^*(t)$ with the highest actual price, i.e.,

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

Notice that under the time-average criterion (i.e., $\beta = 1$, and therefore $\gamma_k \geq \mu_{k,N}$ is required), users with their channel in the best condition have an infinite price. Therefore, in addition we propose the following tie-breaking rule: If at least one of the users' channel is in its best state (i.e., $X_k(t) = N_k$ for some k), then allocate the base station to the user

$$k^*(t) := \arg\max_{k\in\mathcal{K}} \lim_{\beta\to 1} (1-\beta)\nu_{k,X_k(t)}.$$

Such a tie-breaking would therefore be based on the second term of the Laurent expansion of $\nu_{k,n}$ at $\beta = 1$. This tie-breaking may itself have ties; these are resolved arbitrarily.

5. γ-parameterized objective and relationships with scheduling rules

Notice in (4) that for user k under channel condition n we can write $\nu_{k,n} = c_k\mu_{k,n}f_{k,n}$, where

$$f_{k,n} := \frac{(1-\beta) + \beta\gamma_k}{(1-\beta) + \beta\sum_{m>n} q_{k,m}(\mu_{k,m} - \mu_{k,n})}.$$

A quick inspection reveals that $\gamma_k \geq \mu_{k,N}$ is a sufficient condition for $f_{k,n} \geq 1$, which implies $\nu_{k,n} \geq c_k\mu_{k,n}$. On the other hand, under $\beta = 0$ we have $f_{k,n} = 1$, i.e., the above-defined scheduling rule is the *myopic* $c\mu$-rule, in which μ is the instantaneous job completion probability under the actual channel conditions, which is well-known to be optimal in that case. When there is a single available channel condition (i.e., $N_k = 1$) for every user k and if γ_k is the same for all k, then the above-defined scheduling rule is equivalent to the $c\mu$-rule (only scaled by a constant for a given β), which is optimal in that case [16].

5.1. Potential improvement rule

Problem (P) with $\gamma_k = 1$ for all $k \leq K-1$ corresponds to the problem of minimization of the β-average holding costs, since $R^a_{k,n}$ equals the expected holding cost paid at the end of the current slot if action a is decided at the beginning of the slot. In the special case of $c_k = 1$ for all k, problem (P) is equivalent to both minimization of the β-average number of users in the system and minimization of the β-average waiting time.

Under the time-average criterion ($\beta = 1$), the prices in (4) simplify to what we call the *Potential Improvement* (PI) index:

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

and the tie-breaking quantity is

$$\lim_{\beta\to 1}(1-\beta)\nu^{\mathrm{PI}}_{k,n} = 0 \quad \text{for } n \neq N_k, \qquad \lim_{\beta\to 1}(1-\beta)\nu^{\mathrm{PI}}_{k,N_k} = c_k\mu_{k,N_k}.$$

Thus, PI rule results in giving absolute priority to users whose actual channel's quality is the best possible, i.e., the base station can serve a user with non-best channel quality only if there is no user with best channel quality. The absolute priority to the best channel quality users is the most distinguishing property of PI from the RB (PF) rule. Further, if there are several users with absolute priority, PI prescribes to serve users according to the classic $c\mu$-rule, where μ is the instantaneous job completion probability (under the best channel condition). Finally, when no channel achieves its best quality, PI allocates the bandwidth to the user with the highest ratio of the actual service rate with respect to the expected potential improvement of the service rate.

Remark 2 (*Application of PI in a Wireless Data Network*)**.** In view of (1), the straightforward implementation of PI in a wireless network will be by using the index

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

with arbitrary tie-breaking except if more than one user is in its best channel condition, in which case we select any one of those with highest value $c_k s_{k,N_k}/\mathbb{E}[B_k]$. We note that the PI index only depends on the channel statistics, and that it depends on the expected service requirement $\mathbb{E}[B_k]$ only for tie-breaking. If one uses PI with randomized tie-breaking or if all the jobs have the same expected length, then PI would belong to the family of opportunistic schedulers like, for example, RB, PB discussed in the next subsection.

5.2. Relatively best and proportionally best rules

For a given discount factor β, suppose that the job is long enough so that $\sum_{m>1} q_{k,m}(\mu_{k,m} - \mu_{k,1}) \ll (1-\beta)/\beta$ and suppose that $(1-\beta)/\beta \ll \gamma_k$. Then we have $f_{k,n} \approx \gamma_k\beta/(1-\beta)$, so that the scheduling rule defined above is approximately the rule selecting the job with the highest $\nu_{k,n} = c_k\mu_{k,n}\gamma_k$. In the following we show that Relatively Best (RB) and Proportionally Best (PB) rules can be recovered in this way in our modeling framework.

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015

When $\gamma_k = 1/\overline{\mu}_k$ for all k, where $\overline{\mu}_k := \sum_m q_m \mu_{k,m}$, then the scheduling rule recovers the well-studied RB rule, which (with $c_k = 1$ for all k) approximates the PF rule implemented in practice. When $\gamma_k = 1/\mu_{k,N}$ for all k, then the scheduling rule recovers the PB rule introduced in [22].

We remark that we have not been able to find a convincing interpretation of the objective function that gives rise to RB and PB rules. RB optimizes the "natural metrics" (which are the mean number of customers and mean waiting time, obtained when $\gamma_k = 1$) only if the jobs are so short that they can be completed within a slot even under the worst channel condition (i.e., $\overline{\mu}_k = 1$). Note that this assumption in fact leads to a no-channel variation case, in which RB is equivalent to $c\mu$-rule and therefore optimal.

PB optimizes the natural metrics only if the jobs are rather short and can be completed within a slot under the best channel condition (i.e., $\mu_{k,N} = 1$). In this case, however, the values of the discount factor must be moderate in order to satisfy the approximation conditions from the beginning of this subsection. Hence the time-average case is not achievable with this approximation and we are not able to draw any conclusions about the time-average performance of the PB rule.

5.3. On priority-based rule for throughput maximization

Notice that the case with $\gamma_k = 0$ for all k corresponds to maximizing the β-average throughput, weighted by the holding costs c_k. The results from the previous section thus apply if $\beta \le 1/(1+\mu_{k,N})$ for all $k \le K-1$, which is system-independent only if $\beta \le 1/2$. Under these conditions, the scheduling rule given by the prices in (4) is the same as the rule under the natural metrics, since it is only scaled by the constant $1-\beta$. However, we were unable to extend the result to the time-average case, and we believe that there might be cases in which threshold policies are not a sufficient class for optimality. Notice that if $\beta = 1$, then the price for the best channel is $c\mu$, while being 0 for all the other channels. This suggests (though does not prove) that there is no opportunistic rule that could maximize throughput even in an asymptotic sense.

6. Performance of PI in systems with random arrivals

In this section we investigate the performance of the PI rule in the presence of random arrivals. We assume that users are grouped in K classes. The channel of a class-k user can be in N_k conditions, and condition $n \in \mathcal{N}_k$ happens with probability $q_{k,n}$. Users of the same class are thus characterized by having the same channel statistics, but it still holds that the channel condition of every user evolves independently of other channels. Let λ_k be the probability that a new class-k user arrives during a slot to the system; thus, several users are allowed to arrive during the same slot provided they belong to different classes.

The optimal scheduling rule could be calculated numerically when 2 jobs are in the system and no new jobs arrive. The numerical experiments (not reported here) indicate that in this case the suboptimality gap of PI is rather small. Unfortunately the curse of dimensionality makes it impossible to compute the optimal policy for a larger number of users, or for the case of randomly arriving users. Thus we cannot draw any conclusion on how far from the true optimal policy PI is.

Existing literature gives strong evidence to support the claim that the optimal policy to schedule a fixed number of users, may also perform very well in the presence of new arrivals, particularly if the arrival process is Bernoulli or Poisson. In fact it has been shown in a wide variety of models that the optimal scheduling policy with a fixed number of users is also optimal in the case of arrivals (see for example [27] and [28, Theorem 3.28] for the M/G/1 queue, [29–31] for the $c\mu$-rule, [32] for a single server queue with feedback and [33,28,34] for the multi-armed bandit problem). This gives hope that the PI rule should perform well with arrivals. The rest of the paper is devoted to investigating this issue.

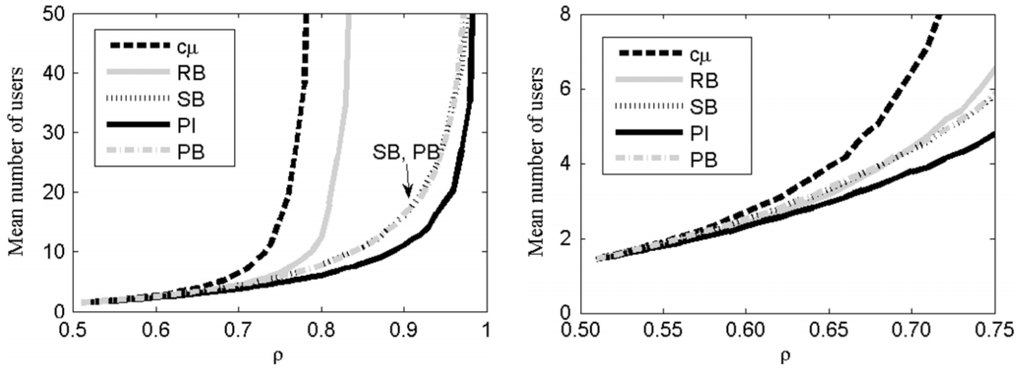
6.1. Stability of PI

Let $\varrho_k := \lambda_k/\mu_{k,N_k}$ and let $\varrho := \sum_k \varrho_k$. Then it is known under the assumption of a time-scale separation (see [11]) that the maximum stability region for our model is $\varrho < 1$, that is, no scheduling rule is stable if this inequality is not satisfied. By stability we understand that the Markov chain is positive recurrent. It is known that a sufficient condition for a policy to achieve the maximum stability region is that any user in its best channel condition will be preferred over any user in a non-best channel condition (see, for example [22,23]). As we explained in Section 5 PI gives absolute priority to users whose channel's actual quality is the best possible, and thus, PI is stable under the maximum stability region.

We note that no other rule in its full generality achieves the maximum stability region and so each of them fails to be stable in some stabilizable systems. PB and SB satisfy the above-mentioned condition only if $c_k = c$ for all k. The $c\mu$-rule satisfies it only if $c_k\mu_{k,N_k} = c$ for all k. In an important result, [11] showed that for the so-called "symmetric" system (see below) the necessary stability condition is also sufficient for RB (and PF). Our numerical simulations nevertheless indicate that in asymmetric scenarios (even if $c_k = 1$ for all k), the stability region under RB and $c\mu$ is strictly smaller than the maximal one, whereas PI always provides maximum stability.

6.2. Flow-level performance of PI in symmetric system

Suppose that all users have the same set of possible channel conditions $\mathcal{N}' := \{1, \ldots, N\}$ and $c_k = 1$, for all k. Let Y denote a discrete random variable such that $\mathbb{P}(Y = y_n) = q_n$, $n \in \mathcal{N}'$. Assume that, as before, in every slot a class-k channel



(a) When $0.51 \le \varrho \le 0.99$.

(b) Zoomed in when $0.51 \le \varrho \le 0.75$.

Fig. 1. Mean number of users in the system in Scenario 1.

is in condition n with probability $q_{k,n} = q_n$ (independently of other users) and that $\mu_{k,n} = r_k y_n$, where r_k is a class dependent constant and it holds $\mu_{k,n} < 1, \forall(k, n)$. In [11] this system is referred to as "symmetric", because it models the situation in which the relative fluctuations of the various users around r_k are identical. As shown in [11] (see also [14]), in the symmetric system RB behaves roughly as PF. We will show now that under this condition, PI and RB are equivalent.

In this system the PI index becomes

$$\nu_{k,n} = \frac{y_n}{\sum\limits_{m>n} q_m(y_m - y_n)} \quad \text{for } n \neq N_k, \qquad \nu_{k,N_k} = \infty, \tag{5}$$

and it does not depend anymore on k. It then holds that a class-k user in channel condition n will be preferred over a class-j user in channel condition l if and only if $\nu_{k,n} > \nu_{j,l}$, i.e., if and only if $n > l$. Thus the channel condition characterizes completely the policy (independently of the class the user belongs to).

RB chooses the user with the highest $\frac{\mu_{k,n}}{\sum_m q_m \mu_{k,m}} = \frac{y_n}{\sum_m q_m y_m}$, and thus a user in state n will be preferred over a user in state l if and only if $n > l$, that is, RB and PI are the same policy. This in particular implies that the analysis carried out in [11, Section 3] to investigate the performance of RB with randomly arriving users can be readily applied to PI. From this analysis we conclude that the performance of PI at the flow level can be characterized by a state-dependent Processor Sharing queue.

7. Simulation study in systems with random arrivals

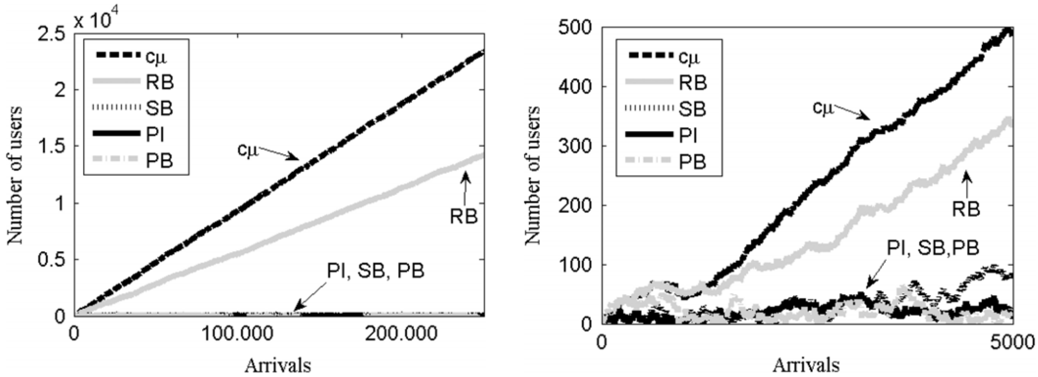
In this section we report numerical simulations to compare the performance of PI with other popular scheduling rules proposed in the literature. Time is slotted. In every slot a new class-k user arrives to the system with probability λ_k. At the beginning of each time slot, we simulate the current transmission rate for the users and then depending on the rule we select which user to serve. This in turn determines the probability that the user departs from the system during the slot. We fix the following parameters for all our simulations: $\kappa = 0$ and $c_k := 1$, for all k; with these values the objective is to minimize the mean number of users in the system. The statistics of the number of users are collected by invoking the GASTA property (Geometric Arrivals See Time Averages) of Geometric arrivals [35].

We consider the policies PI, RB, PB, SB, and $c\mu$-rule. All these policies are priority-based, i.e., they serve the user with the highest index in the system. The expression for the indices PI, RB and PB were explained in Section 5. The remaining indices are $\nu_{k,n}^{c\mu} := c_k \mu_{k,n}$ and $\nu_{k,n}^{\text{SB}} := c_k \sum_{m=1}^{n} q_{k,m}$, for $n \in \mathcal{N}_k$. We simulate RB, and not PF, because simulating PF would be very time consuming as it requires to keep track of the history of the process. We have a similar situation with SB. In [9] SB calculates the score depending on the past measurements of the channel fluctuations. This scheme is very time-consuming to simulate, and we have therefore used this alternative definition that makes the process Markovian. Under the assumption that the channel quality evolves over time in an i.i.d. fashion, we expect both definitions to be very close to each other. In addition to these rules, we also simulate the policy that is optimal for the simpler scheduling problem in which there is only one user of each class and no new users arrive. This rule is obtained by solving numerically the dynamic programming equation and we denote it as the NA rule (No Arrivals rule). We emphasize that the curse of dimensionality makes it unfeasible to numerically find an optimal solution in much more complex scenarios. If there are more than one user with the highest index value in a certain moment, then in RB, SB and PB we allocate the slot to a randomly chosen user, as proposed in [36,9,22] respectively.

Our main conclusions are the following:

(i) Depending on the value of the parameters, RB, SB, PB or $c\mu$ might outperform the others.
(ii) PI consistently outperforms all the other policies (or is equivalent to the best one) with respect to the mean, and in many cases even stochastically.
(iii) The stability region of RB, $c\mu$ and NA is strictly smaller than the maximal one ($\varrho < 1$).

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015



(a) For 230,000 user arrivals.

(b) Zoomed for 5000 user arrivals.

Fig. 2. Sample path of the number of users in the system in Scenario 1 when $\varrho = 0.95$.

Numerical experiments in the symmetric system (see Section 6.2) show that PI, SB, PB and RB are equivalent, outperforming NA and $c\mu$ rules, while all these rules are stable in the maximum stability region. Next we describe in more detail the results we have obtained for a variety of asymmetric scenarios.

7.1. Scenario 1

In the first family of simulations we consider a set of five transmission rates for users of class 1 and a set of three transmission rates for users of class 2. These transmission rates are adopted from the CDMA 1xEV-DO system and the channel condition probabilities are given in Table 1. We set $\lambda_2 := 0.005$, and choose 102.57 kb as the expected length of both class-1 and class-2 jobs. In view of (1) we have $\varrho_2 = 0.5$. We consider a range of values for class-1 arrival probability λ_1 such that ϱ varies from 0.51 to 0.99.

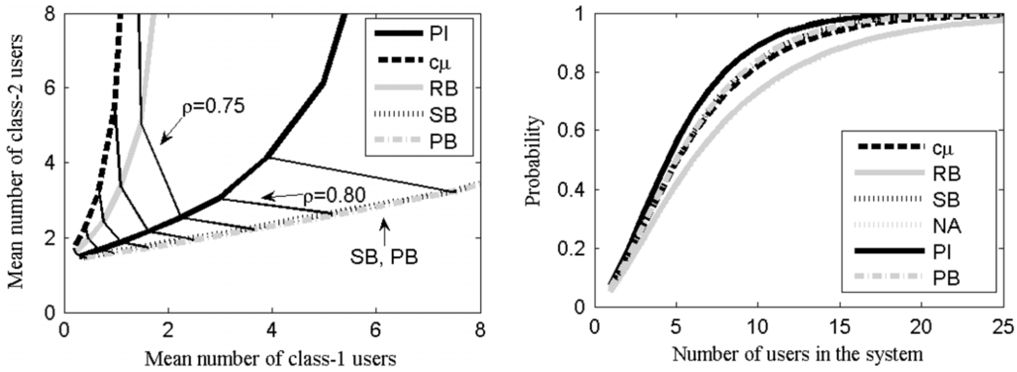
In Fig. 1(a) we plot the mean number of users in the system versus ϱ under different rules. NA and RB rules are equivalent in this setting, i.e., they both take the same scheduling decision (and therefore NA is omitted in the figures). Although SB and PB are not completely equivalent (they take different decisions in some of the cases when no users are in their best states), they show a very similar performance. We observe that instability of the $c\mu$-rule arises with $\varrho \geq 0.79$ and that of RB and NA with $\varrho \geq 0.84$. SB, PB and PI are stable in the whole range of ϱ, nevertheless we note that the performance of PI is always superior to that of SB and PB. In Fig. 1(b) we show the same data zoomed in when $0.51 \leq \varrho \leq 0.75$. All the rules show similar performance when $\varrho = 0.55$, but we can observe that the mean number of users in the system under PI becomes significantly smaller than under other rules as ϱ grows. On the other hand, the $c\mu$ rule becomes increasing worse as ϱ grows. Fig. 2 depicts the sample path evolution of the number of users in the system when $\varrho = 0.95$ and indicates that RB (and therefore also NA), and $c\mu$ are unstable, whereas the system is stable under PI, SB and PB.

For better understanding of the performance, we further plot in Fig. 3(a) the indifference map (well-known in the microeconomic theory) for this scenario, where the points at the same indifference curve (or isoline) correspond to the same value of ϱ. When the objective is to minimize the aggregate number of users in the system, then the minimizer among the plotted rules always coincides with PI. This is because of the shape of the indifference curves and the fact that PI keeps the number of users of different classes fairly balanced. Notice that $c\mu$ and RB tend to keep a small number of class-1 users, causing the number of class-2 users to skyrocket as ϱ increases. Conversely, SB and PB maintain a small number of users of class-2, while class 1 grows steadily.

7.2. Scenario 2

We consider the same transmission rates and channel condition probabilities as in Scenario 1. We set the expected length of class-2 jobs to 102.57 kb and let $\lambda_1 = \lambda_2 := 0.005$ for both classes. With these values we obtain $\varrho_2 = 0.5$. We vary the expected length of class 1 users so that ϱ varies between 0.51 and 0.99. Note that changing the expected job length modifies the departure probabilities (see (1)), which in turn implies that $c\mu$ and NA might differ for different values of ϱ.

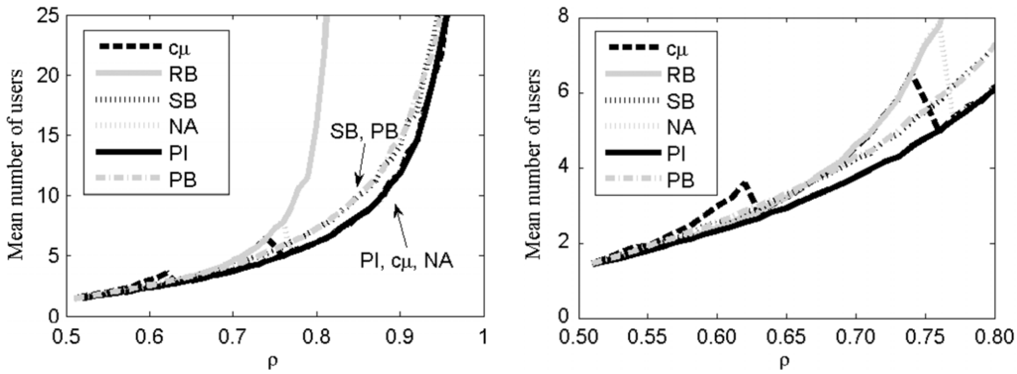
Fig. 4(a) plots the mean number of users in the system versus ϱ. We can observe that RB is unstable for $\varrho \geq 0.83$. Again, SB and PB show a very similar performance although they are not equivalent. On the other hand, PI, SB, NA and $c\mu$ rules remain stable for all values of $\varrho \leq 0.99$. In Fig. 4(b) we zoom in on the data when $0.51 \leq \varrho \leq 0.80$ and can observe that PI dominates other rules increasingly more as ϱ grows, whereas RB and NA become increasingly more dominated. However, notice in Fig. 4(a) that on the interval $0.77 \leq \varrho \leq 0.99$ the performance of PI, $c\mu$ and NA is quite similar and slightly outperforming the SB and PB rules. For $\varrho \leq 0.76$ the performance of NA deteriorates similarly to the unstable RB, but then there is an abrupt improvement in the performance. This is caused by the change in the priorities given by NA; indeed, for



(a) Indifference map for Scenario 1.

(b) Distribution functions of scheduling rules.

Fig. 3. Further comparisons of opportunistic rules.



(a) When $0.51 \leq \varrho \leq 0.99$.

(b) When $0.51 \leq \varrho \leq 0.80$.

Fig. 4. Mean number of users in the system in Scenario 2.

$\varrho \geq 0.77$ any user in his best channel is given priority over all the users that are in a non-best channel condition, while for $\varrho \leq 0.76$ this is not true.

7.3. Other scenarios

We have studied more scenarios, changing available transmission rates and the corresponding channel condition probabilities. In general, we have obtained similar qualitative conclusions as in scenarios 1 and 2: PI, PB and SB rules are stable in all the scenarios whenever $\varrho < 1$, showing that PI performs strictly better than PB and SB almost in every simulation. On the other hand, RB, NA and $c\mu$ rules have more irregular performance and show instability problems in some scenarios.

7.4. Stochastic ordering

In this section, our goal is to illustrate that, in fact, the number of users under different rules are often ordered in a stochastic sense. Recall that two random variables X and Y are stochastically ordered if and only if $\mathbb{P}(X \leq z) \geq \mathbb{P}(Y \leq z)$, $\forall z$. In particular, the stochastic ordering implies that $\mathbb{E}[X] \leq \mathbb{E}[Y]$. In Fig. 3(b) we represent the distribution functions for a particular example (in Scenario 2 when $\varrho = 0.75$), but we emphasize that we have obtained similar pictures in almost all our simulations. Our simulations strongly suggest that PI is stochastically better than the other rules.

8. Conclusion

Optimizing the flow-level performance in a system with time-varying service rates, random arrivals, and asymmetric channel conditions remains a challenging task. We have proposed a comprehensive modeling framework for the problem without arrivals and have exploited it to derive the Potential Improvement rule, which can be implemented in practice not more costly than other popular opportunistic schedulers. Numerical results indicate that PI is superior or comparable to the best of the existing scheduling rules under several scenarios. It would be desirable to investigate the performance of PI in other scenarios, for example, by dropping the assumption that the channel quality is not a correlated process. It would

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015

also be interesting to evaluate the performance of PI with user mobility and to compare the performance of PI with PF in real-world systems. The practical issue of efficient learning of the parameters of a particular user based on the history is another important topic that deserves more research attention.

Appendix. Work-reward analysis and proofs

We will next focus on the case $\beta < 1$, i.e., the problem under the discounted criterion. Problem (3) is a standard stationary MDP problem, for which it is well known that there is an optimal policy which is deterministic (i.e., non-randomized), stationary (i.e., Markovian), and independent of the initial state [37, Chapter 6]. In particular, this implies that there exists an optimal policy which only depends on the user-k state-process $X_k(\cdot)$. Indeed, policy $\widetilde{\pi}_k \in \Pi_{\boldsymbol{X},a_k}$ that depends on the joint state-process $\boldsymbol{X}(\cdot)$ can be seen as a randomized policy, since the user-l state-process $X_l(\cdot)$ for $l \neq k$ is not influenced by the user-k action-process $a_k(\cdot)$ prescribed by $\widetilde{\pi}_k$.

Therefore, in order to find an optimal policy to problem (3) it is enough to concentrate on stationary policies $\pi_k \in \Pi_{X_k,a_k}$. Every such policy can be represented in terms of a *serving set* $\mathcal{S} \subseteq \mathcal{N}_k$, which prescribes to allocate the base station's service (i.e., to serve) whenever the user is in state $n \in \mathcal{S}$ and not to serve whenever the user is in state $n \notin \mathcal{S}$. Thus, an optimal policy to problem (3) can be obtained by solving

$$\max_{\mathcal{S}\subseteq\mathcal{N}_k} \mathbb{B}_0^{\mathcal{S}}\left[R_{k,X_k(\cdot)}^{a_k(\cdot)}\right] - \nu\mathbb{B}_0^{\mathcal{S}}\left[W_{k,X_k(\cdot)}^{a_k(\cdot)}\right]. \tag{A.1}$$

Notice that (A.1) is a parametric (bi-objective) optimization problem and every policy (i.e., serving set) $\mathcal{S}$ is associated with a bi-dimensional point $\mathbb{B}_0^{\mathcal{S}}\left[W_{k,X_k(\cdot)}^{a_k(\cdot)}\right], \mathbb{B}_0^{\mathcal{S}}\left[R_{k,X_k(\cdot)}^{a_k(\cdot)}\right]$. If depicted in a plane with works on the x-axis and rewards on the y-axis, then the optimal policies to (A.1) lie on the upper boundary of such a region, since the parameter ν gives the slope of the supporting hyperplane (a line in this case) defining an optimum point (i.e., an optimal policy).

We will next analyze scaled quantities $\mathbb{B}_0^{\mathcal{S}}\left[R_{k,X_k(\cdot)}^{a_k(\cdot)}\right]/(1-\beta)$, writing briefly $\mathbb{R}_n^{\mathcal{S}}$ if the initial state $X_k(0) = n \in \mathcal{N}_k$. Analogously, we will write briefly $\mathbb{W}_n^{\mathcal{S}}$ and we denote the value function under policy $\mathcal{S}$ by $\mathbb{V}_n^{\mathcal{S}} := \mathbb{R}_n^{\mathcal{S}} - \nu\mathbb{W}_n^{\mathcal{S}}$. These scaled quantities correspond to the usual quantities under the β-discounted criterion. An optimal solution in terms of prices for $\beta = 1$, i.e., under the time-average criterion, is obtained in the limit. In the rest of this section we will omit the user subscript $k \le K-1$ to simplify the notation.

A.1. Proof of Proposition 1

Let us denote the optimal value function by $\mathbb{V}_n^*$. An inspection of the Bellman equation leads to the following lemma.

Lemma 1. (i) *Suppose that $\nu > 0$ holds. Then, if it is optimal to transmit in state $n \in \mathcal{N}\setminus\{N\}$, then it is optimal to transmit in state $n+1$.*
(ii) *Suppose that $\nu \le 0$ and $\gamma \ge \mu_N$ holds. Then it is optimal to transmit in any state $n \in \mathcal{N}$.*
(iii) *Suppose that $\nu \le 0$ and $\gamma < \mu_N$ and $\beta \le 1/(1+\mu_N-\gamma)$ holds. Then it is optimal to transmit in any state $n \in \mathcal{N}$.*

Proof. The Bellman equation for state $n \in \mathcal{N}$ is $\mathbb{V}_n^* = \max_{a\in\mathcal{A}}\left\{R_n^a - \nu W_n^a + \beta\sum_{m\in\mathcal{N}} p_{n,m}^a \mathbb{V}_m^*\right\}$. After plugging the definitions of the action-dependent parameters for state $n \in \mathcal{N}\setminus\{0\}$, we obtain

$$\begin{aligned}\mathbb{V}_n^* &= \max\left\{R_n^1 - \nu W_n^1 + \beta\sum_{m\in\mathcal{N}'}(1-\mu_n)q_m\mathbb{V}_m^* + \beta\mu_n\mathbb{V}_0^*\,;\,R_n^0 - \nu W_n^0 + \beta\sum_{m\in\mathcal{N}'}q_m\mathbb{V}_m^*\right\}\\ &= -c\gamma + \beta\sum_{m\in\mathcal{N}'}q_m\mathbb{V}_m^* + \max\left\{-\nu + \mu_n\left(c + \beta\mathbb{V}_0^* - \beta\sum_{m\in\mathcal{N}'}q_m\mathbb{V}_m^*\right);0\right\},\end{aligned}$$

where the first term in the curly brackets corresponds to action 1 and the second one to action 0. Transmitting (i.e., action 1) is optimal in state $n \in \mathcal{N}\setminus\{0\}$, if the first term is greater than or equal to the second term. Let us denote by $Z := c + \beta\mathbb{V}_0^* - \beta\sum_{m\in\mathcal{N}'}q_m\mathbb{V}_m^*$.

(i) If $\nu > 0$ and transmitting is optimal in state $n \in \mathcal{N}\setminus\{0,N\}$, then $\mu_n Z \ge \nu > 0$. Since $\mu_{n+1} \ge \mu_n$, we also have $-\nu + \mu_{n+1}Z \ge 0$, that is, transmitting is optimal in state $n+1$. Further, using the Bellman equation it is straightforward to obtain that $\mathbb{V}_0^* = 0$ because action 1 is not optimal in state 0, and therefore the statement holds under $\nu > 0$ for all $n \in \mathcal{N}\setminus\{N\}$.
(ii) If $\nu \le 0$, then we proceed as follows. First, using the Bellman equation it is straightforward to obtain that $\mathbb{V}_0^* = -\nu/(1-\beta)$ because action 1 is optimal in state 0 and thus $-\nu$ is obtained in every period forever. Notice that the one-period net reward, $R_n^a - \nu W_n^a$, is for any state $n \in \mathcal{N}$ and any action $a \in \mathcal{A}$ upperbounded by $-\nu$ whenever $\gamma \ge \mu_N$. Hence $\mathbb{V}_m^* \le -\nu/(1-\beta) = \mathbb{V}_0^*$ for any $m \in \mathcal{N}'$, and therefore (using $c > 0$) also $Z > 0$, and finally, for any state $n \in \mathcal{N}\setminus\{0\}$, $-\nu + \mu_n Z > 0$. That is, transmitting is optimal in any state $n \in \mathcal{N}$.

(iii) Similarly to (ii), the one-period net reward, $R_n^a - \nu W_n^a$, is for any state $n \in \mathcal{N}$ and any action $a \in \mathcal{A}$ upperbounded by $-c(\gamma - \mu_N) - \nu$ whenever $\gamma < \mu_N$. Hence $\mathbb{V}_m^* \leq (-c(\gamma - \mu_N) - \nu)/(1-\beta) = -c(\gamma - \mu_N)/(1-\beta) + \mathbb{V}_0^*$ for any $m \in \mathcal{N}'$, and therefore also

$$Z \geq \frac{c}{1-\beta}\left[1 - \beta + \beta(\gamma - \mu_N)\right].$$

The right-hand side is non-negative due to the assumption $\beta \leq 1/(1 + \mu_N - \gamma)$. So, similarly to (ii), transmitting is optimal in any state $n \in \mathcal{N}$. □

Note that analogously to Lemma 1(iii), one could also prove a sharper condition for β involving parameter ν, which is, however, not useful for further discussion in this paper. Finally, Lemma 1 establishes optimality of threshold policies in Proposition 1.

A.2. Proof of Proposition 2

In order to prove the existence of optimal prices ν_n in terms of properties (i) and (ii) of Proposition 2, we will establish validity of a sufficient condition called *LP-indexability* introduced in [38, Definition 5.3], which will be stated after defining some necessary concepts. The analysis in the following paragraphs also shows how to evaluate such prices provided they exist. An immediate result are the *balance equations* given in the following lemma.

Lemma 2. *For all states $n \in \mathcal{N}$ we have $\mathbb{W}_n^{\mathcal{N}} = 1/(1-\beta)$, and under any policy $0 \notin \mathcal{S}$ we have*

$$\mathbb{R}_n^{\mathcal{S}} = \begin{cases} -c(\gamma - \mu_n) + (1-\mu_n)\beta \sum_{m \in \mathcal{N}'} q_m \mathbb{R}_m^{\mathcal{S}}, & \text{if } 0 \neq n \in \mathcal{S}, \\ -c\gamma + \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{R}_m^{\mathcal{S}}, & \text{if } 0 \neq n \notin \mathcal{S}, \\ 0, & \text{if } n = 0. \end{cases}$$

$$\mathbb{W}_n^{\mathcal{S}} = \begin{cases} 1 + (1-\mu_n)\beta \sum_{m \in \mathcal{N}'} q_m \mathbb{W}_m^{\mathcal{S}}, & \text{if } 0 \neq n \in \mathcal{S}, \\ \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{W}_m^{\mathcal{S}}, & \text{if } 0 \neq n \notin \mathcal{S}, \\ 0, & \text{if } n = 0. \end{cases}$$

Proof. Directly from the definition of β-average reward and work, respectively, we have

$$\mathbb{R}_n^{\mathcal{S}} = R_n^{n \in \mathcal{S}} + \beta \sum_{m \in \mathcal{N}} p_{n,m}^{n \in \mathcal{S}} \mathbb{R}_m^{\mathcal{S}},$$

$$\mathbb{W}_n^{\mathcal{S}} = W_n^{n \in \mathcal{S}} + \beta \sum_{m \in \mathcal{N}} p_{n,m}^{n \in \mathcal{S}} \mathbb{W}_m^{\mathcal{S}},$$

where $n \in \mathcal{S}$ equals 1 if true and 0 otherwise. Substituting the values of $R_n^{n \in \mathcal{S}}$, $W_n^{n \in \mathcal{S}}$ and $p_{n,m}^{n \in \mathcal{S}}$ given in the definition of the job-channel-user triple, and simplifying, results in the above characterization. □

If price ν_n for $n \in \mathcal{N}'$ with the desired properties (i) and (ii) (and price ν_0 for $n = 0$ with the desired properties (iii) and (iv)) stated in Proposition 2 exists, then both transmitting and not transmitting is optimal if $\nu = \nu_n$. This means that there is a policy, say $\mathcal{S}^*$, such that both including state n in $\mathcal{S}^*$ and not including it lead to the same objective value, i.e.,

$$\mathbb{R}_n^{\mathcal{S}^* \cup \{n\}} - \nu_n \mathbb{W}_n^{\mathcal{S}^* \cup \{n\}} = \mathbb{R}_n^{\mathcal{S}^* \setminus \{n\}} - \nu_n \mathbb{W}_n^{\mathcal{S}^* \setminus \{n\}}.$$

A straightforward consequence of this and of the balance equation is that changing the action only in the initial period must also lead to the same objective values, i.e.,

$$\mathbb{R}_n^{\langle 0, \mathcal{S}^* \rangle} - \nu_n \mathbb{W}_n^{\langle 0, \mathcal{S}^* \rangle} = \mathbb{R}_n^{\langle 1, \mathcal{S}^* \rangle} - \nu_n \mathbb{W}_n^{\langle 1, \mathcal{S}^* \rangle},$$

where policy $\langle a, \mathcal{S}^* \rangle$ is the policy that employs action a in the initial period and then proceeds according to $\mathcal{S}^*$. Then, whenever $\mathbb{W}_n^{\langle 1, \mathcal{S}^* \rangle} - \mathbb{W}_n^{\langle 0, \mathcal{S}^* \rangle} \neq 0$, we have

$$\nu_n = \frac{\mathbb{R}_n^{\langle 1, \mathcal{S}^* \rangle} - \mathbb{R}_n^{\langle 0, \mathcal{S}^* \rangle}}{\mathbb{W}_n^{\langle 1, \mathcal{S}^* \rangle} - \mathbb{W}_n^{\langle 0, \mathcal{S}^* \rangle}}. \tag{A.2}$$

We will therefore study ν_n under all policies $\mathcal{S}$, defined as

$$\nu_n^{\mathcal{S}} := \frac{\mathbb{R}_n^{\langle 1, \mathcal{S} \rangle} - \mathbb{R}_n^{\langle 0, \mathcal{S} \rangle}}{\mathbb{W}_n^{\langle 1, \mathcal{S} \rangle} - \mathbb{W}_n^{\langle 0, \mathcal{S} \rangle}}. \tag{A.3}$$

From the balance equations we can obtain the following characterization of these quantities.

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015

Lemma 3. *For any state $n \in \mathcal{N}'$ under any policy $0 \notin \mathcal{S}$ we have*

$$\nu_n^{\mathcal{S}} = \mu_n \frac{c - \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{R}_m^{\mathcal{S}}}{1 - \mu_n \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{W}_m^{\mathcal{S}}}, \quad \nu_0^{\mathcal{S}} = 0. \tag{A.4}$$

Proof. Using the characterization of $\mathbb{R}_n^{\langle a, \mathcal{S}\rangle}$ for $a \in \mathcal{A}$ from Lemma 2, we obtain that

$$\mathbb{R}_n^{\langle 1, \mathcal{S}\rangle} - \mathbb{R}_n^{\langle 0, \mathcal{S}\rangle} = \mu_n \left(c - \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{R}_m^{\mathcal{S}} \right).$$

Similarly, we obtain that

$$\mathbb{W}_n^{\langle 1, \mathcal{S}\rangle} - \mathbb{W}_n^{\langle 0, \mathcal{S}\rangle} = 1 - \mu_n \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{W}_m^{\mathcal{S}}. \tag{A.5}$$

Then, the definition in (A.3) yields the stated characterization. Finally, $\nu_0^{\mathcal{S}} = 0$ is obtained trivially. □

Since our objective is to find the optimal prices ν_n in terms of properties (i) and (ii) of Proposition 2, we postulate that

$$\text{For all } n \in \mathcal{N} : \nu_n = \nu_n^{\mathcal{S}_{N-n}} \quad \text{for } \mathcal{S}_{N-n} := \{m \in \mathcal{N} : m > n\}. \tag{A.6}$$

We will verify this postulate using a sufficient condition LP-indexability [38, Definition 5.3], which in our problem can be simplified to the following.

Definition 1. Problem (3) is LP-indexable with prices ν_n given in (A.6), if the following conditions hold:

(i) $\mathbb{W}_n^{\langle 1, \emptyset\rangle} - \mathbb{W}_n^{\langle 0, \emptyset\rangle} > 0$ and $\mathbb{W}_n^{\langle 1, \mathcal{N}\rangle} - \mathbb{W}_n^{\langle 0, \mathcal{N}\rangle} > 0$ for all $n \in \mathcal{N}$;
(ii) $\mathbb{W}_n^{\langle 1, \mathcal{S}_{N-n}\rangle} - \mathbb{W}_n^{\langle 0, \mathcal{S}_{N-n}\rangle} > 0$ and $\mathbb{W}_{n+1}^{\langle 1, \mathcal{S}_{N-n}\rangle} - \mathbb{W}_{n+1}^{\langle 0, \mathcal{S}_{N-n}\rangle} > 0$ for each $n \in \mathcal{N} \setminus \{N\}$;
(iii) For every real-valued ν there exists $n \in \mathcal{N} \cup \{-1\}$ such that the serving set $\mathcal{S}_{N-n}$ is optimal.

We will first need to characterize the above quantities under $\mathcal{S}_{N-n}$ for any $n \in \mathcal{N}$. According to the balance equations in Lemma 2, we have

$$\mathbb{W}_l^{\mathcal{S}_{N-n}} = 1 + (1 - \mu_l)\beta \sum_{m \in \mathcal{N}'} q_m \mathbb{W}_m^{\mathcal{S}_{N-n}}, \quad \text{if } l \in \mathcal{S}_{N-n},$$
$$\mathbb{W}_l^{\mathcal{S}_{N-n}} = \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{W}_m^{\mathcal{S}_{N-n}}, \quad \text{if } l \notin \mathcal{S}_{N-n}.$$

Denoting by

$$\overline{\mathbb{W}}^{\mathcal{S}_{N-n}} := \beta \sum_{m \in \mathcal{N}'} q_m \mathbb{W}_m^{\mathcal{S}_{N-n}},$$

we can solve such a system of linear equations obtaining

$$\overline{\mathbb{W}}^{\mathcal{S}_{N-n}} = \frac{\beta \sum_{m \in \mathcal{S}_{N-n}} q_m}{1 - \beta + \beta \sum_{m \in \mathcal{S}_{N-n}} q_m \mu_m}. \tag{A.7}$$

Analogously, we can obtain

$$\overline{\mathbb{R}}^{\mathcal{S}_{N-n}} = \frac{c \left(-\beta\gamma + \beta \sum_{m \in \mathcal{S}_{N-n}} q_m \mu_m \right)}{1 - \beta + \beta \sum_{m \in \mathcal{S}_{N-n}} q_m \mu_m}.$$

Plugging these expressions into (A.4), we have

$$\nu_l^{\mathcal{S}_{N-n}} = \frac{c\mu_l (1 - \beta + \beta\gamma)}{1 - \beta + \beta \sum_{m \in \mathcal{S}_{N-n}} q_m (\mu_m - \mu_l)}. \tag{A.8}$$

Next we establish that the LP-indexability holds under a mild condition.

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015

Lemma 4. *Suppose that $\gamma \geq \mu_N$ or $\beta \leq 1/(1+\mu_N-\gamma)$ holds. Problem (3) is LP-indexable with prices ν_n given in (A.6).*

Proof. (i) It is straightforward to obtain from Lemma 2 that $\mathbb{W}_n^{\langle 1,\emptyset\rangle} - \mathbb{W}_n^{\langle 0,\emptyset\rangle} = \mathbb{W}_n^{\langle 1,\mathcal{N}\rangle} - \mathbb{W}_n^{\langle 0,\mathcal{N}\rangle} = 1$ for all $n \in \mathcal{N}$.

(ii) Using (A.5) and (A.7) we obtain

$$\mathbb{W}_l^{\langle 1,\mathcal{S}\rangle} - \mathbb{W}_l^{\langle 0,\mathcal{S}\rangle} = 1 - \frac{\beta \sum\limits_{m\in\mathcal{S}_{N-n}} q_m\mu_l}{1-\beta+\beta \sum\limits_{m\in\mathcal{S}_{N-n}} q_m\mu_m},$$

which is nonincreasing in l due to the monotonicity assumption upon μ_l's. Hence, it is enough to prove $\mathbb{W}_{n+1}^{\langle 1,\mathcal{S}_{N-n}\rangle} - \mathbb{W}_{n+1}^{\langle 0,\mathcal{S}_{N-n}\rangle} > 0$. By the same assumption we have $\mu_{n+1} \leq \mu_m$ for all $m \in \mathcal{S}_{N-n}$, which together with $\beta < 1$ implies the result.

(iii) This is established in Proposition 1. □

Notice that expression (4) with $l := n$ is equivalent to (A.8). Therefore, since LP-indexability is a sufficient condition for the properties (i) and (ii) (and (iii) and (iv)) of Proposition 2, we conclude its proof. Validity of the properties (v) and (vi) of Proposition 2 can be derived analogously.

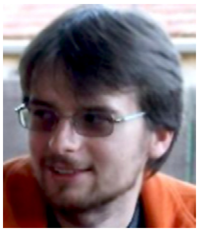
References

[1] L. Tassiulas, A. Ephremides, Dynamic server allocation to parallel queues with randomly varying connectivity, IEEE Transactions on Information Theory 39 (2) (1993) 466–478.
[2] M. Andrews, Instability of the proportional fair scheduling algorithm for HDR, IEEE Transactions on Wireless Communications 3 (5) (2004) 1422–1426.
[3] A. Stolyar, Maxweight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic, The Annals of Applied Probability 14 (1) (2004) 1–53.
[4] A. Eryilmaz, R. Srikant, J. Perkins, Stable scheduling policies for fading wireless channels, IEEE/ACM Transactions on Networking 13 (2) (2005) 411–424.
[5] M. Neely, Order optimal delay for opportunistic scheduling in multi-user wireless uplinks and downlinks, IEEE/ACM Transactions on Networking 16 (5) (2008) 1188–1199.
[6] M. Andrews, K. Kumaran, K. Ramanan, S. Stolyar, R. Vijayakumar, P. Whiting, Scheduling in a queueing system with asynchronously varying service rates, Probability in the Engineering and Informational Sciences 18 (2) (2004) 191–217.
[7] B. Sadiq, S.-J. Baek, G. de Veciana, Delay-optimal opportunistic scheduling and approximations: the Log rule, in: Proceedings of IEEE INFOCOM, 2009.
[8] M. Hu, J. Zhang, J. Sadowsky, Traffic aided opportunistic scheduling for wireless networks: algorithms and performance bounds, Computer Networks 46 (4) (2004) 505–518.
[9] T. Bonald, A score-based opportunistic scheduler for fading radio channels, in: Proceedings of European Wireless, 2004, pp. 283–292.
[10] D. Park, H. Seo, H. Kwon, B. Lee, Wireless packet scheduling based on the cumulative distribution function of user transmission rates, IEEE Transactions on Communications 53 (11) (2005) 1919–1929.
[11] S. Borst, User-level performance of channel-aware scheduling algorithms in wireless data networks, IEEE/ACM Transactions on Networking 13 (3) (2005) 636–647.
[12] P. Lassila, S. Aalto, Combining opportunistic and size-based scheduling in wireless systems, in: Proceedings of ACM MSWiM, 2008, pp. 323–332.
[13] X. Liu, E. Chong, N. Shroff, Optimal opportunistic scheduling in wireless networks, in: Proceedings of VTC, 2003.
[14] S. Borst, T. Bonald, N. Hegde, M. Jonckheere, A. Proutiere, Flow-level performance and capacity of wireless networks with user mobility, Queueing Systems 63 (1–4) (2009) 131–164.
[15] H. Kushner, P. Whiting, Convergence of proportional-fair sharing algorithms under general conditions, IEEE Transactions on Wireless Communications 3 (2004) 1250–1259.
[16] D.R. Cox, W.L. Smith, Queues, Methuen & Co., 1961
[17] C. Papadimitriou, J. Tsitsiklis, The complexity of optimal queueing network, Mathematics of Operations Research 24 (2) (1999) 293–305.
[18] P. Whittle, Restless bandits: activity allocation in a changing world, Journal of Applied Probability 25 (1988) 287–298.
[19] J. Niño-Mora, Dynamic priority allocation via restless bandit marginal productivity indices, Top 15 (2) (2007) 161–198.
[20] P. Jacko, Dynamic Priority Allocation in Restless Bandit Models, Lambert Academic Publishing, 2010.
[21] R. Weber, G. Weiss, On an index policy for restless bandits, Journal of Applied Probability 27 (1990) 637–648.
[22] S. Aalto, P. Lassila, Flow-level stability and performance of channel-aware priority-based schedulers, in: NGI 2010, 6th EURO-NF Conference on Next Generation Internet, Paris, 2010.
[23] U. Ayesta, M. Erausquin, M. Jonckheere, I.M. Verloop, Opportunistic scheduling in wireless systems: stability and asymptotic optimality (2010) (submitted for publication).
[24] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, A. Viterbi, CDMA/HDR: a bandwidth-efficient high-speed wireless data service for nomadic users, IEEE Communications Magazine 38 (7) (2000) 70–77.
[25] J. Niño-Mora, Restless bandits, partial conservation laws and indexability, Advances in Applied Probability 33 (1) (2001) 76–98.
[26] P. Jacko, Adaptive greedy rules for dynamic and stochastic resource capacity allocation problems, Medium for Econometric Applications 17 (4) (2009) 10–16. Available online at: http://www.met-online.nl.
[27] K. Sevcik, Scheduling for minimum total loss using service time distributions, Journal of the ACM 21 (1974) 66–75.
[28] J. Gittins, Multi-Armed Bandit Allocation Indices, Wiley, Chichester, 1989.
[29] W. Smith, Various optimizers for single stage production, Naval Research Logistics Quarterly 3 (1956) 59–66.
[30] D. Fife, Scheduling with random arrivals and linear loss functions, Management Science 11 (3) (1965) 429–437.
[31] C. Buyukkoc, P. Varaya, J. Walrand, The $c\mu$ rule revisited, Advances in Applied Probability 17 (1985) 237–238.
[32] I. Meilijson, G. Weiss, Multiple feedback at a single server station, Stochastic Processes and their Applications 5 (1977) 195–205.
[33] J. Gittins, D. Jones, A dynamic allocation index for the sequential design of experiments, in: J. Gani (Ed.), Progress in Statistics, North-Holland, 1974, pp. 241–266.
[34] P. Whittle, Arm-acquiring bandits, Annals of Probability 9 (2) (1981) 284–292.
[35] A. Makowski, B. Melamed, W. Whitt, On averages seen by arrivals in discrete time, in: Proceedings of IEEE Conference on Decision and Control, 1989, pp. 1084–1086.
[36] F. Berggren, R. Jäntti, Asymptotically fair transmission scheduling over fading channels, IEEE Transactions on Wireless Communications 3 (1) (2004) 326–336.
[37] M.L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, John Wiley & Sons, 2005.
[38] J. Niño-Mora, Characterization and computation of restless bandit marginal productivity indices, in: Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools, ICST, Brussels, Belgium, October 2007.

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015

Urtzi Ayesta is an IKERBASQUE researcher at BCAM, the Basque Center for Applied Mathematics, Bilbao, Spain. Previously he has been a CNRS researcher working at LAAS, Toulouse, France and an ERCIM postdoc fellow at CWI, Amsterdam, the Netherlands. He received the Ph.D. degree in Computer Science from Université de Nice-Sophia Antipolis (France). His Ph.D. research work was carried out at the research laboratories of INRIA (MAESTRO team) and France Telecom R&D. Urtzi Ayesta holds a M.Sc. degree in Electrical Engineering from Columbia University (US) and a Diplome in Telecomunication Engineering from Nafarroako Unibertsitate Publikoa-Universidad Publica de Navarra (Spain).

Martin Erausquin obtained his B.Sc. in Mathematics from the University of the Basque Country (UPV/EHU) in 2007 and he studied the "Master in Banking and Quantitative Finance" between 2007 and 2010, at the UPV/EHU. As of September 2010, Martin is a Ph.D. student at UPV/EHU under the supervision of Urtzi Ayesta (BCAM) and Eva Ferreira (UPV/EHU).



Peter Jacko is currently with the BCAM — Basque Center for Applied Mathematics, Spain. He obtained his Ph.D. in Business Administration and Quantitative Methods (2009) and D.E.A. in Statistics and Operations Research (2006) from the Universidad Carlos III de Madrid, Spain. He obtained his Mgr. (2003) and Bc. (2002) degrees in Mathematics from the Univerzita P. J. Šafárika v Košiciach, Slovakia. E-mail: jacko@bcamath.org.

Please cite this article in press as: U. Ayesta, et al., A modeling framework for optimizing the flow-level scheduling with time-varying channels, Performance Evaluation (2010), doi:10.1016/j.peva.2010.08.015