
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Aalto, Samuli; Lassila, Pasi; Taboada, Ianire

Indexability of an opportunistic scheduling problem with partial channel information

Published in:

Proceedings of the 12th EAI International Conference on Performance Evaluation Methodologies and Tools, VALUETOOLS 2019

DOI:

[10.1145/3306309.3306324](https://doi.org/10.1145/3306309.3306324)

Published: 12/03/2019

Document Version

Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:

Aalto, S., Lassila, P., & Taboada, I. (2019). Indexability of an opportunistic scheduling problem with partial channel information. In *Proceedings of the 12th EAI International Conference on Performance Evaluation Methodologies and Tools, VALUETOOLS 2019* (pp. 95-102). ACM. <https://doi.org/10.1145/3306309.3306324>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Indexability of an opportunistic scheduling problem with partial channel information

Samuli Aalto
Aalto University, Finland
samuli.aalto@aalto.fi

Pasi Lassila
Aalto University, Finland
pasi.lassila@aalto.fi

Ianire Taboada
University of the Basque Country,
Spain
ianire.taboada@ehu.eus

ABSTRACT

Opportunistic scheduling in wireless cellular systems utilizes random channel quality variations in time by favoring the users with good channel conditions. However, the success of such schedulers is heavily depending on the accuracy of the available information on the channel states of users. In this paper, we consider the opportunistic scheduling problem of downlink data traffic with partial channel information, where the target is to minimize the flow-level holding costs. In earlier works, the Whittle index approach has successfully been utilized to develop near-optimal scheduling policies for the corresponding problems. Using the same approach, we complement and extend the results found thus far. More specifically said, our novel contributions are (i) finding sufficient conditions under which the problem with partial channel information is provably indexable, and (ii) deriving an explicit formula for the corresponding Whittle index. In addition, we evaluate the performance of the derived Whittle index policies and compare them with some greedy policies by numerical simulations.

CCS CONCEPTS

• **Networks** → **Network resources allocation; Network performance evaluation;**

KEYWORDS

Wireless cellular networks, Opportunistic scheduling, Partial channel information, Whittle index

ACM Reference Format:

Samuli Aalto, Pasi Lassila, and Ianire Taboada. 2019. Indexability of an opportunistic scheduling problem with partial channel information. In *12th EAI International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS 2019)*, March 12–15, 2019, Palma, Spain. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3306309.3306324>

1 INTRODUCTION

The channel quality in wireless cellular systems is inherently randomly varying in time. Opportunistic schedulers utilize such variations by favoring the users with good channel conditions [5, 8, 15,

17, 31]. However, their success is heavily depending on the accuracy of the available information on the channel states of users.

In this paper, we consider the opportunistic scheduling problem of downlink data traffic with *partial channel information*, where the target is to minimize the flow-level holding costs. In line with many related papers [4, 14, 20, 22, 23, 25], we assume that the scheduler estimates the channel state based on ARQ-type feedback from the *scheduled* users. Thus, from the users that are not scheduled in the time slot, the scheduler does not get any new channel state information. The wireless channel dynamics are described by discrete-time Markov processes, and the traffic consists of elastic flows corresponding to downloading of files using TCP.

Such a flow-level opportunistic scheduling problem with IID or Markovian channel dynamics has earlier been studied under the assumption of *exact channel information* [1–3, 6, 7, 10–13, 16, 28, 30, 32]. Near-optimal scheduling policies have successfully been developed by applying the *Whittle index approach*, which was originally developed for the restless bandits problems [34]. In this approach, the original (mathematically intractable) problem is first relaxed by allowing the scheduler to choose multiple users in the same time slot and only requiring that the time slot is given on average for at most one user, which makes the problem more tractable by decomposing it to separate user-related subproblems. While the problem becomes separable, it is, however, not clear, a priori, whether the problem is indexable, or not. If the problem is indexable, the resulting Whittle index policy is known to be asymptotically optimal, at least under certain technical conditions [33].

Under the more realistic assumption of *partial channel information*, the main emphasis has been on the opportunistic scheduling problem where the target is to maximize the average or discounted packet-level throughput of a fixed user population [4, 18–26, 35]. With partial information, the problem can be formulated in the framework of Partially Observable Markov Decision Processes. Even here, the Whittle index approach has been used in many papers to develop near-optimal scheduling policies [19–26].

The main difference between the maximization of packet-level throughput and the minimization of flow-level holding costs arises from the fact that in the previous problem a user stays in the system forever, while in the latter one, a user leaves the system as soon as the service of the corresponding flow has been completed. So the cost structure is different and, therefore, the solutions, as well. In particular, the optimality and near-optimality results for packet-level models found in [4, 18–26, 35] are not applicable in our flow-level problem.

The flow-level opportunistic scheduling problem with partial channel information has, thus far, only been considered in [14, 29]. In [14], Jacko and Villar assume ARQ-type feedback and two-state

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

VALUETOOLS 2019, March 12–15, 2019, Palma, Spain

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6596-3/19/03...\$15.00

<https://doi.org/10.1145/3306309.3306324>

positively autocorrelated Markovian channels. By applying the Whittle index approach, the scheduling problem is decomposed into separate user-related subproblems. However, they are not able to prove but just *conjecture* that the subproblems are indexable. Under this conjecture and some additional assumptions concerning the model parameters (that will be discussed in Section 2), they establish the optimality of threshold type policies for the subproblems and also present a formula for the corresponding Whittle index. However, the formula for the Whittle index given in [14] is not fully explicit. In [29], Taboada et al. study a slightly different model related to periodic feedback relying on the Whittle index approach. They allow multi-state Markovian channels but present only an approximate analysis, where the indexability property is again conjectured.

Our contributions in the present paper are as follows: We complement and extend the results given in [14] by

- (i) finding sufficient conditions under which the flow-level opportunistic scheduling problem with partial channel information is provably indexable, which was only conjectured in [14];
- (ii) deriving an explicit formula for the corresponding Whittle index, which is surprisingly simple when compared to the less explicit formula given in [14].

Moreover, we evaluate the performance of the derived Whittle index policies and compare them with greedy policies by numerical simulations. The results show that the Whittle index policies perform well in all our scenarios, independent of the degree of positive correlation in channel dynamics.

The flow-level opportunistic scheduling problem with partial channel information is described in Section 2 and the mathematically more tractable relaxed version of the problem in Section 3. In Section 4, we introduce a class of threshold policies and characterize the optimal policy for the relaxed problem. In Section 5, the relaxed problem is shown to be indexable, at least in a certain part of the parameter space, and the corresponding Whittle index is derived explicitly. In Section 6, we define the Whittle index policy for the original opportunistic scheduling problem, and its performance is compared to some greedy policies by numerical simulations in Section 7. Section 8 concludes the paper.

2 SCHEDULING PROBLEM WITH PARTIAL CHANNEL INFORMATION

We consider the following opportunistic scheduling problem with partial channel information in discrete time. The same problem was earlier introduced and partially solved in [14].

At time 0 (i.e., in the beginning of the first time slot), there are K users in the area of a wireless base station, each with a file (a.k.a. flow) to be downloaded. For any time slot $t \in \{0, 1, \dots\}$, the scheduler in the base station chooses (at most) M of the flows for service until all the flows have been downloaded. Denote $A_k^\pi(t) = 1$ if flow k is chosen in time slot t , where π refers to the scheduling policy used; otherwise $A_k^\pi(t) = 0$. Thus, for any policy π and time slot t , we have the constraint

$$\sum_{k=1}^K A_k^\pi(t) \leq M, \quad (1)$$

where $M \in \{1, 2, \dots\}$.

We assume that, for any user k , the wireless channel is described by a two-state discrete-time Markov chain $N_k(t) \in \{1, 2\}$. Let $q_{k,n,m} > 0$ denote the transition probability from channel state n to state m for user k . We assume that these Markov chains are independent of each other and, furthermore, *positively autocorrelated*, i.e.,

$$\rho_k = q_{k,2,2} - q_{k,1,2} > 0. \quad (2)$$

In addition, let $q_{k,n}$ denote the steady-state probability to be in channel state n ,

$$q_{k,1} = \frac{q_{k,2,1}}{q_{k,2,1} + q_{k,1,2}}, \quad q_{k,2} = \frac{q_{k,1,2}}{q_{k,2,1} + q_{k,1,2}}. \quad (3)$$

Finally, let $\mu_{k,n}$ denote the probability that flow k is completed in a time slot where it is scheduled and its channel state is *known* to be n . Without loss of generality, we may assume that

$$\mu_{k,2} \geq \mu_{k,1} > 0, \quad (4)$$

i.e., channel state 2 is better than state 1.

The scheduling decisions are, however, not based on full information but just partial information on the channel states of the users. We assume that any user scheduled in time slot t sends back its true channel state at the end of this time slot. This assumption leads to a Partially Observable Markov Decision Process (POMDP).

In the POMDP formulation, the state of user k is given by the *belief state* $X_k^\pi(t)$, which denotes the current probability to be in the good channel state 2, unless the flow has already been completed, which is denoted by a special symbol '*'. The set of possible belief states for each user k is countably infinite:

$$\mathcal{S}_k = \{x_{k,2,j} : j \in \{0, 1, \dots\}\} \cup \{q_{k,2}\} \cup \{x_{k,1,j} : j \in \{0, 1, \dots\}\} \cup \{*\},$$

where $x_{k,2,0} = q_{k,2,2}$, $x_{k,1,0} = q_{k,1,2}$, and

$$x_{k,n,j+1} = x_{k,n,j}q_{k,2,2} + (1 - x_{k,n,j})q_{k,1,2}. \quad (5)$$

Note that user k with an uncompleted flow is in belief state $x_{k,n,j}$ if and only if this user was last time scheduled for $j + 1$ slots ago and its channel state was then known to be n . When the flow of user k is completed, it will enter the absorbing state '*'.¹

It is easy to show by induction that

$$x_{k,2,j} = q_{k,1,2} \frac{1 - \rho_k^{j+1}}{1 - \rho_k} + \rho_k^{j+1}, \quad x_{k,1,j} = q_{k,1,2} \frac{1 - \rho_k^{j+1}}{1 - \rho_k}. \quad (6)$$

Now it follows from our assumption (2) that

$$x_{k,2,0} > x_{k,2,1} > \dots > q_{k,2} > \dots > x_{k,1,1} > x_{k,1,0}. \quad (7)$$

Recall also that belief state $q_{k,2}$ refers to the equilibrium state of channel k .

Let $\mu_k(x)$ denote the probability that flow k is completed in a time slot where it is scheduled and its belief state is $x \in \mathcal{S}_k \setminus \{*\}$. We adopt the same *rate adaptation* scheme as in [14]: If flow k is scheduled in belief state x for which $x > \mu_{k,1}/\mu_{k,2}$, the base station transmits with the best rate. In such a case, the flow is completed with probability $x\mu_{k,2}$. But if $x \leq \mu_{k,1}/\mu_{k,2}$, the base station transmits with the worst rate so that the flow is completed with

¹ This is a characteristic property of flow-level opportunistic scheduling models, which the corresponding packet-level models lack.

probability $\mu_{k,1}$. In other words, the flow completion probability if scheduled in belief state x is given by

$$\mu_k(x) = \max\{x\mu_{k,2}, \mu_{k,1}\}. \quad (8)$$

In addition, let $\mu_k(x, n)$ denote the probability that flow k is completed in a time slot where it is scheduled, its belief state is $x \in \mathcal{S}_k \setminus \{*\}$, and the channel state is n . According to our rate adaptation scheme, we have

$$\mu_k(x, 2) = \begin{cases} \mu_{k,2}, & \text{if } x > \mu_{k,1}/\mu_{k,2}, \\ \mu_{k,1}, & \text{otherwise;} \end{cases}$$

$$\mu_k(x, 1) = \begin{cases} 0, & \text{if } x > \mu_{k,1}/\mu_{k,2}, \\ \mu_{k,1}, & \text{otherwise.} \end{cases}$$

Note that $\mu_k(x) = x\mu_k(x, 2) + (1-x)\mu_k(x, 1)$.

For any user k , holding costs are accumulated at rate $h_k > 0$ until the whole file is downloaded and the flow is completed. Thus, the objective function in our scheduling problem is given by

$$E \left[\sum_{t=0}^{\infty} \sum_{k=1}^K h_k 1_{\{X_k^\pi(t) \neq *\}} \right]. \quad (9)$$

The aim is to find the optimal scheduling policy π that minimizes the expected total holding costs (9) subject to the strict capacity constraint (1) for all t and assuming that the scheduling decisions at each time slot t are based on the belief states $X_k^\pi(t)$ of users.

3 RELAXED OPTIMIZATION PROBLEM

The original optimization problem described in the previous section belongs to the class of *restless bandit* problems, which are known to be mathematically intractable. Jacko and Villar [14] followed Whittle's approach [34] to develop near-optimal heuristic solutions for the problem: They modified the original problem by replacing the strict capacity constraint (1) by an averaged one and approached the relaxed problem by the Lagrangian methods. In this paper, we use the same approach, which results in the following (separate) sub-problems for each user k : Find the optimal policy π that minimizes the objective function

$$f_k^\pi + \nu g_k^\pi, \quad (10)$$

where ν can be interpreted as the unit price of work, f_k^π as the expected total holding costs of user k , and g_k^π as the expected total amount of work needed for user k ,

$$f_k^\pi = E \left[\sum_{t=0}^{\infty} h_k 1_{\{X_k^\pi(t) \neq *\}} \right], \quad g_k^\pi = E \left[\sum_{t=0}^{\infty} A_k^\pi(t) \right].$$

Note that, if $\nu < 0$, there is no need to postpone the scheduling decision but user k should be scheduled in each time slot until its completion. Thus, from this on, we assume that $\nu \geq 0$.

The separable subproblems of the Lagrangian version of the relaxed scheduling problem are now considered in the context of partially observed Markov decision processes. The possible actions $a \in \mathcal{A} = \{0, 1\}$ are "to schedule" ($a = 1$) and "not to schedule" ($a = 0$). For any belief state $x \in \mathcal{S}_k \setminus \{*\}$, both actions $a \in \{0, 1\}$ are allowed, while for $x = *$, only action $a = 0$ is possible.

Let $p_k(y|x, a) \geq 0$ denote the transition probability from belief state $x \in \mathcal{S}_k$ to belief state $y \in \mathcal{S}_k$ after action $a \in \mathcal{A}$. It follows

from the previous discussion that the non-zero transition probabilities are as follows:

$$\begin{aligned} p_k(x_{k,n,j+1}|x_{k,n,j}, 0) &= 1, \quad n \in \{1, 2\}, j \in \{0, 1, \dots\}, \\ p_k(q_{k,2}|q_{k,2}, 0) &= 1, \\ p_k(x_{k,2,0}|x, 1) &= x(1 - \mu_k(x, 2)), \quad x \in \mathcal{S}_k \setminus \{*\}, \\ p_k(x_{k,1,0}|x, 1) &= (1-x)(1 - \mu_k(x, 1)), \quad x \in \mathcal{S}_k \setminus \{*\}, \\ p_k(*|x, 1) &= \mu_k(x), \quad x \in \mathcal{S}_k \setminus \{*\}, \\ p_k(*|*, 0) &= 1. \end{aligned} \quad (11)$$

Note that $*$ is an absorbing state for any policy.

Finally, let $c_k(x, a)$ denote the immediate cost in state $x \in \mathcal{S}_k$ after action $a \in \mathcal{A}$. In our model,

$$\begin{aligned} c_k(x, 0) &= h_k, \quad x \in \mathcal{S}_k \setminus \{*\}, \\ c_k(x, 1) &= h_k + \nu, \quad x \in \mathcal{S}_k \setminus \{*\}, \\ c_k(*, 0) &= 0. \end{aligned} \quad (12)$$

Since the state space \mathcal{S}_k is discrete, the action space \mathcal{A} is finite, and the immediate costs are bounded, there is a stationary deterministic policy π_k^* (described by a function $a^{\pi_k^*}(x)$ from the state space \mathcal{S}_k to the action space \mathcal{A}) that minimizes the expected total costs (10) [27].

Finally, let $V_k(x; \nu)$ denote the *value function* for state $x \in \mathcal{S}_k$ related to the minimization of the expected total costs (10) with Lagrangian parameter ν . The corresponding *optimality equations* in this model read as follows:

$$\begin{aligned} V_k(*; \nu) &= 0, \\ V_k(q_{k,2}; \nu) &= h_k + \min\{V_k(q_{k,2}; \nu), \\ &\quad \nu + q_{k,2}(1 - \mu_k(q_{k,2}, 2))V_k(x_{k,2,0}; \nu) + \\ &\quad (1 - q_{k,2})(1 - \mu_k(q_{k,2}, 1))V_k(x_{k,1,0}; \nu)\}, \\ V_k(x_{k,n,j}; \nu) &= h_k + \min\{V_k(x_{k,n,j+1}; \nu), \\ &\quad \nu + x_{k,n,j}(1 - \mu_k(x_{k,n,j}, 2))V_k(x_{k,2,0}; \nu) + \\ &\quad (1 - x_{k,n,j})(1 - \mu_k(x_{k,n,j}, 1))V_k(x_{k,1,0}; \nu)\}. \end{aligned} \quad (13)$$

In addition, let $V_k^\pi(x; \nu)$ denote the corresponding value function for policy π . The policy π_k^* for which

$$V_k^{\pi_k^*}(x; \nu) = V_k(x; \nu)$$

for all $x \in \mathcal{S}_k$ is said to be ν -optimal for user k .

4 CHARACTERIZATION OF THE OPTIMAL POLICY FOR THE RELAXED PROBLEM

In this section, we characterize the ν -optimal policy π_k^* for the relaxed problem (10) in a certain part of the parameter space related to user k . We start with a small lemma (the proof of which is elementary and, thus, omitted here). Thereafter, we present and prove four auxiliary results (Propositions 1-4) needed for the characterization given in Theorem 1.

Note that, in this section, we leave out the user index k to simplify notations. Thus, for example, the value function $V_k(x; \nu)$ is briefly referred to by $V(x; \nu)$ and so on.

LEMMA 1. *If $a \leq b$ and $0 < q \leq p < 1$, then*

$$pa + (1-p)b \leq qa + (1-q)b.$$

PROPOSITION 1. *$V(x; \nu) < \infty$ for all $x \in \mathcal{S}$ and $V(x_{2,0}; \nu) \leq V(x_{1,0}; \nu)$.*

PROOF. 1° Assume first that $x_{2,0} \leq \mu_1/\mu_2$. It follows from (8) and (7) that under this assumption the base station transmits with the worst rate μ_1 in any belief state $x \in \mathcal{S} \setminus \{*\}$ so that

$$\mu(x) = \mu(x, 2) = \mu(x, 1) = \mu_1.$$

Thus, there is no reason to postpone the decisions to schedule but it is optimal to schedule the user in each time slot. In addition, since the completion trials in each time slot are independent, the value function for any $x \in \mathcal{S} \setminus \{*\}$ is equal to

$$V(x; v) = \frac{h + v}{\mu_1} < \infty, \quad (14)$$

which is also easy to verify from the optimality equations (13) in this case. In particular, we have $V(x_{2,0}; v) - V(x_{1,0}; v) = 0$.

2° Assume now that $x_{2,0} > \mu_1/\mu_2$. It follows from (8) and (7) that under this assumption the flow completion probability $\mu(x)$ in any belief state x is upper [lower] bounded by $x_{2,0}\mu_2$ [μ_1]. In addition, since the completion trials in each time slot are independent, the value function for any $x \in \mathcal{S} \setminus \{*\}$ is lower and upper bounded by

$$\frac{h + v}{x_{2,0}\mu_2} \leq V(x; v) \leq \frac{h + v}{\mu_1} < \infty. \quad (15)$$

Since $x_{2,0} > \mu_1/\mu_2$, we also have

$$\mu(x_{2,0}) = x_{2,0}\mu_2, \quad \mu(x_{2,0}, 2) = \mu_2, \quad \mu(x_{2,0}, 1) = 0.$$

Thus, from the optimality equation (13), we get

$$V(x_{2,0}; v) \leq h + v + x_{2,0}(1 - \mu_2)V(x_{2,0}; v) + (1 - x_{2,0})V(x_{1,0}; v)$$

implying, together with the lower bound of (15), that

$$(1 - x_{2,0})(V(x_{2,0}; v) - V(x_{1,0}; v)) \leq h + v - x_{2,0}\mu_2 V(x_{2,0}; v) \leq 0.$$

Thus, $V(x_{2,0}; v) - V(x_{1,0}; v) \leq 0$, which completes the proof. \square

PROPOSITION 2. If π is v -optimal and $x_{2,j+1} > \mu_1/\mu_2$, then

$$a^\pi(x_{2,j+1}) = 1 \quad \Rightarrow \quad a^\pi(x_{2,j}) = 1.$$

PROOF. Assume that π is v -optimal. Let $j \in \{0, 1, \dots\}$, and assume that $x_{2,j+1} > \mu_1/\mu_2$ and $a^\pi(x_{2,j+1}) = 1$. By (7), we deduce that $x_{2,j} > \mu_1/\mu_2$. Now if $a^\pi(x_{2,j}) = 0$, then it follows from optimality equations (13), Proposition 1, and Lemma 1 that

$$\begin{aligned} V(x_{2,j+1}; v) &\leq v + x_{2,j}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{2,j})V(x_{1,0}; v) \\ &\leq v + x_{2,j+1}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{2,j+1})V(x_{1,0}; v) \\ &< h + v + x_{2,j+1}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{2,j+1})V(x_{1,0}; v) = V(x_{2,j+1}; v), \end{aligned}$$

which is a contradiction. Thus, we have $a^\pi(x_{2,j}) = 1$. \square

PROPOSITION 3. If π is v -optimal and $x_{1,j} > \mu_1/\mu_2$, then

$$a^\pi(x_{1,j}) = 1 \quad \Rightarrow \quad a^\pi(x_{1,j+1}) = 1.$$

PROOF. Assume that π is v -optimal. Let $j \in \{0, 1, \dots\}$, and assume that $x_{1,j} > \mu_1/\mu_2$ and $a^\pi(x_{1,j}) = 1$. By (7), we deduce that $x_{1,k} > \mu_1/\mu_2$ for any $k > j$. Now it follows from optimality equations (13) that

$$\begin{aligned} V(x_{1,j}; v) &= h + v + x_{1,j}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,j})V(x_{1,0}; v) \end{aligned}$$

and, for any $k > j$,

$$\begin{aligned} &h + v + x_{1,j}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,j})V(x_{1,0}; v) \\ &\leq (k - j + 1)h + v + x_{1,k}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,k})V(x_{1,0}; v). \end{aligned}$$

Thus,

$$(k - j)h + (x_{1,k} - x_{1,j})((1 - \mu_2)V(x_{2,0}; v) - V(x_{1,0}; v)) \geq 0.$$

In addition, from (6), it is easy to verify that, for any $k > j$,

$$x_{1,k} - x_{1,j} > x_{1,k+1} - x_{1,j+1} > 0.$$

By combining this with Proposition 1, we deduce that

$$\begin{aligned} &(k - j + 1)h + v + x_{1,k+1}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,k+1})V(x_{1,0}; v) \\ &\quad - h - v - x_{1,j+1}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad - (1 - x_{1,j+1})V(x_{1,0}; v) \\ &= (k - j)h \\ &\quad + (x_{1,k+1} - x_{1,j+1})((1 - \mu_2)V(x_{2,0}; v) - V(x_{1,0}; v)) \\ &> (k - j)h \\ &\quad + (x_{1,k} - x_{1,j})((1 - \mu_2)V(x_{2,0}; v) - V(x_{1,0}; v)) \geq 0. \end{aligned}$$

Since this is true for any $k > j$, it follows from optimality equations (13) that

$$\begin{aligned} V(x_{1,j+1}; v) &= h + v + x_{1,j+1}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,j+1})V(x_{1,0}; v) \end{aligned}$$

so that $a^\pi(x_{1,j+1}) = 1$. \square

PROPOSITION 4. If policy π is v -optimal, then

- (i) $a^\pi(x_{2,j}) = 1$ for infinitely many j ,
- (ii) $a^\pi(q_2) = 1$, and
- (iii) $a^\pi(x_{1,j}) = 1$ for infinitely many j .

PROOF. It is easy to see from the optimality equations (13) that if any of the three condition is violated, $V^\pi(x; v) = \infty$ for some belief state $x \in \mathcal{S} \setminus \{*\}$, which contradicts Proposition 1. \square

Now we are ready to characterize the v -optimal policy π_k^* for the relaxed problem related to user k . We will first introduce a family of scheduling policies (so called *threshold policies*), and then reveal the cases when these policies are optimal, which is the main result of this section. As seen from the previous propositions, parameter μ_1/μ_2 has an important role here.

DEFINITION 1. Let $\pi_{1,j}$, where $j \in \{0, 1, \dots\}$, refer to the policy π such that

$$a^\pi(x) = \begin{cases} 0, & \text{for } x \in \{x_{1,0}, \dots, x_{1,j-1}\} \cup \{*\}, \\ 1, & \text{otherwise.} \end{cases} \quad (16)$$

In addition, let Π_1 denote the whole family of such policies

$$\Pi_1 = \{\pi_{1,j} \mid j \in \{0, 1, \dots\}\}, \quad (17)$$

which are called *threshold policies*.

Note that, for policy $\pi_{1,0}$, the user is scheduled in any belief state $x \in \mathcal{S} \setminus \{*\}$.

THEOREM 1. *If $x_{1,0} > \mu_1/\mu_2$, then there is $\pi_{1,j} \in \Pi_1$ such that $\pi_{1,j}$ is v -optimal.*

PROOF. Assume that $x_{1,0} > \mu_1/\mu_2$, and let π denote the v -optimal policy in this case.

By (7), we deduce that $x_{2,j} > \mu_1/\mu_2$ for any $j \in \{0, 1, \dots\}$. Now it follows from Propositions 2 and 4(i) that $a^\pi(x_{2,j}) = 1$ for all $j \in \{0, 1, \dots\}$.

By (7), we also deduce that $x_{1,j} > \mu_1/\mu_2$ for any $j \in \{0, 1, \dots\}$. Now it follows from Propositions 3 and 4(iii) that there is $j_1 \in \{0, 1, \dots\}$ such that $a^\pi(x_{1,j}) = 0$ for all $j \in \{0, \dots, j_1 - 1\}$ and $a^\pi(x_{1,j}) = 1$ for all $j \in \{j_1, j_1 + 1, \dots\}$. In addition, since $a^\pi(q_2) = 1$ by Proposition 4(ii), the policy π_{1,j_1} is v -optimal. \square

We note that the optimality of threshold policies in [14, Proposition 4] was based on a conjecture that the problem itself is indexable, while Theorem 1 above does not require any such conjecture.

5 WHITTLE INDEX WITH PARTIAL CHANNEL INFORMATION

In this section, we continue the characterization of the optimal solution of the relaxed problem (10) related to user k . More precisely said, we show that the relaxed optimization problem is indexable under the assumption specified in Theorem 1 above, which was only conjectured but not proved in [14]. In addition, we derive expressions for the corresponding Whittle index values, which are more explicit than those given in [14, Theorem 4.1].

We start by defining the indexability property, after which we present and prove three auxiliary results (Propositions 5-7) needed for the main results of this section given in Theorem 2. Note that, after the definition of indexability, we again leave out the user index k to simplify notations in this section.

DEFINITION 2. *The relaxed optimization problem (10) related to user k is indexable if, for any belief state $x \in \mathcal{S}_k \setminus \{*\}$, there exists $W_k(x) \in [-\infty, \infty]$ such that*

- (i) *decision $a = 1$ (to schedule user k) is optimal in state x if $v \leq W_k(x)$;*
- (ii) *decision $a = 0$ (not to schedule user k) is optimal in state x if $v \geq W_k(x)$.*

Assume now that

$$x_{1,0} > \mu_1/\mu_2. \quad (18)$$

which is the same case as explicitly considered in [14]. It follows from (7) and (18) that, in this case, the base station transmits with the best rate in any belief state $x \in \mathcal{S} \setminus \{*\}$:

$$\mu(x) = x\mu_2, \quad \mu(x, 2) = \mu_2, \quad \mu(x, 1) = 0.$$

The optimality equations (13) read now as follows:

$$\begin{aligned} V(*) &= 0, \\ V(q_2) &= h + \min\{V(q_2; v), \\ &\quad v + q_2(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - q_2)V(x_{1,0}; v)\}, \\ V(x_{n,j}; v) &= h + \min\{V(x_{n,j+1}; v), \\ &\quad v + x_{n,j}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{n,j})V(x_{1,0}; v)\}. \end{aligned} \quad (19)$$

For any $j \in \{0, 1, \dots\}$, let us define

$$v_{1,j} = h \left(\frac{x_{1,j}}{x_{1,j+1} - x_{1,j}} - j - 1 \right). \quad (20)$$

It follows from (6) that

$$v_{1,j} = h \left(\frac{1 - \rho^{j+1}}{(1 - \rho)\rho^{j+1}} - j - 1 \right). \quad (21)$$

Below we show that these values are positive and increasing unboundedly with respect to j .

PROPOSITION 5. *For any $j \in \{0, 1, \dots\}$,*

$$0 < v_{1,j} < v_{1,j+1}. \quad (22)$$

In addition, $\lim_{j \rightarrow \infty} v_{1,j} = \infty$.

PROOF. By (21),

$$v_{1,0} = h \left(\frac{1}{\rho} - 1 \right) > 0,$$

since $0 < \rho < 1$. Let then $j \in \{0, 1, \dots\}$. Now, again by (21), we have

$$v_{1,j+1} - v_{1,j} = h \left(\frac{1}{\rho^{j+2}} - 1 \right),$$

which is strictly positive and an increasing function of j , since $0 < \rho < 1$. The claims follow from these facts. \square

PROPOSITION 6. *Assume (18). If policy $\pi_{1,0}$ is v -optimal, then $v \leq v_{1,0}$.*

PROOF. Assume that policy $\pi_{1,0}$ is v -optimal, and denote briefly $\pi = \pi_{1,0}$. Now it follows from optimality equations (19) that

$$\begin{aligned} V(x_{1,0}; v) &= h + v + x_{1,0}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,0})V(x_{1,0}; v), \\ V(x_{2,0}; v) &= h + v + x_{2,0}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{2,0})V(x_{1,0}; v), \end{aligned}$$

which gives

$$\begin{aligned} V(x_{1,0}; v) &= (h + v) \frac{1 - (x_{2,0} - x_{1,0})(1 - \mu_2)}{x_{1,0}\mu_2}, \\ V(x_{2,0}; v) &= (h + v) \frac{1 - x_{2,0} + x_{1,0}}{x_{1,0}\mu_2}. \end{aligned} \quad (23)$$

Similarly, by optimality equations (19),

$$\begin{aligned} V(x_{1,1}; v) &= h + v + x_{1,1}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,1})V(x_{1,0}; v). \end{aligned}$$

On the other hand, since $a^\pi(x_{1,0}) = 1$, we have, again by optimality equations (19), that

$$\begin{aligned} V(x_{1,1}; v) &\geq v + x_{1,0}(1 - \mu_2)V(x_{2,0}; v) \\ &\quad + (1 - x_{1,0})V(x_{1,0}; v). \end{aligned}$$

Thus,

$$V(x_{1,0}; v) \leq V(x_{2,0}; v)(1 - \mu_2) + \frac{h}{x_{1,1} - x_{1,0}}.$$

By combining this with (23), we finally deduce that

$$v \leq h \left(\frac{x_{1,0}}{x_{1,1} - x_{1,0}} - 1 \right) = v_{1,0},$$

which completes the proof. \square

PROPOSITION 7. *Assume (18), and let $j \in \{1, 2, \dots\}$. If policy $\pi_{1,j}$ is v -optimal, then*

$$v_{1,j-1} \leq v \leq v_{1,j}.$$

PROOF. Assume that policy $\pi_{1,j}$ is ν -optimal, and denote briefly $\pi = \pi_{1,j}$. First, for any state $\{x_{1,0}, \dots, x_{1,j-1}\}$, we have, by optimality equations (19),

$$V(x_{1,i}; \nu) = h + V(x_{1,i+1}; \nu)$$

implying that

$$V(x_{1,0}; \nu) = jh + V(x_{1,j}; \nu)$$

In addition, it follows from optimality equations (19) that

$$\begin{aligned} V(x_{1,j}; \nu) &= h + \nu + x_{1,j}(1 - \mu_2)V(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,j})V(x_{1,0}; \nu), \\ V(x_{2,0}; \nu) &= h + \nu + x_{2,0}(1 - \mu_2)V(x_{2,0}; \nu) \\ &\quad + (1 - x_{2,0})V(x_{1,0}; \nu). \end{aligned} \quad (24)$$

Thus,

$$\begin{aligned} V(x_{1,0}; \nu) &= \frac{(h+\nu)(1-(x_{2,0}-x_{1,j})(1-\mu_2))}{x_{1,j}\mu_2} \\ &\quad + \frac{jh(1-x_{2,0}(1-\mu_2))}{x_{1,j}\mu_2}, \\ V(x_{1,j}; \nu) &= \frac{(h+\nu)(1-(x_{2,0}-x_{1,j})(1-\mu_2))}{x_{1,j}\mu_2} \\ &\quad + \frac{jh(1-x_{2,0}(1-\mu_2)-x_{1,j}\mu_2)}{x_{1,j}\mu_2}, \\ V(x_{2,0}; \nu) &= \frac{(h+\nu)(1-x_{2,0}+x_{1,0})+jh(1-x_{2,0})}{x_{1,j}\mu_2}. \end{aligned} \quad (25)$$

Now, since $a^\pi(x_{1,j-1}) = 0$, we have, by optimality equations (19), that

$$\begin{aligned} V(x_{1,j}; \nu) &\leq \nu + x_{1,j-1}(1 - \mu_2)V(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,j-1})V(x_{1,0}; \nu), \end{aligned}$$

which implies, by (24), that

$$V(x_{2,0}; \nu)(1 - \mu_2) + \frac{h}{x_{1,j}-x_{1,j-1}} \leq V(x_{1,0}; \nu).$$

By combining this with (25), we deduce that

$$\nu \geq h \left(\frac{x_{1,j-1}}{x_{1,j}-x_{1,j-1}} - j \right) = v_{1,j-1}.$$

On the other hand, since $a^\pi(x_{1,j}) = 1$, we have, by optimality equations (19) and, that

$$\begin{aligned} V(x_{1,j+1}; \nu) &\geq \nu + x_{1,j}(1 - \mu_2)V(x_{2,0}; \nu) \\ &\quad + (1 - x_{1,j})V(x_{1,0}; \nu). \end{aligned}$$

In addition, again by optimality equations (19),

$$\begin{aligned} V(x_{1,j+1}; \nu) &= h + \nu + x_{j+1,1}(1 - \mu_2)V(x_{2,0}; \nu) \\ &\quad + (1 - x_{j+1,1})V(x_{1,0}; \nu). \end{aligned}$$

Thus,

$$V(x_{1,0}; \nu) \leq V(x_{2,0}; \nu)(1 - \mu_2) + \frac{h}{x_{1,j+1}-x_{1,j}}.$$

By combining this with (25), we finally deduce that

$$\nu \leq h \left(\frac{x_{1,j}}{x_{1,j+1}-x_{1,j}} - j - 1 \right) = v_{1,j},$$

which completes the proof. \square

THEOREM 2. *Under assumption (18), the relaxed optimization problem (10) is indexable, and the corresponding Whittle index $W(x)$ is given by*

$$\begin{aligned} W(x_{2,j}) &= \infty, & j \in \{0, 1, \dots\}, \\ W(q_2) &= \infty, \\ W(x_{1,j}) &= v_{1,j}, & j \in \{0, 1, \dots\}, \end{aligned} \quad (26)$$

where $v_{1,j}$ is given in (21).

PROOF. The results follow straightforwardly from Theorem 1 and Propositions 5-7. \square

We note that, under assumption (18), the Whittle index values do *not* depend on the ‘‘rate’’ parameters μ_1 and μ_2 at all but only on the ‘‘channel dynamics’’ parameters $q_{1,2}$ and $q_{2,2}$ together with the holding cost weight h and the time $j+1$ from the previous channel state update, as seen from (26) and (21), since $\rho = q_{2,2} - q_{1,2}$ by (2). However, this fact is not clearly visible from the implicit Whittle index formula given in [14, Theorem 4.1] under the same assumption (18).

6 WHITTLE INDEX POLICY

Now we are ready to define the Whittle index policy for the original opportunistic scheduling problem described in Section 2, where it is strictly required that (at most) M users are scheduled at each time slot t , see (1).

Assume that there are K users whose channel dynamics satisfy condition (18). Let $x_k(t)$ denote the belief state of user k in the beginning of time slot t . The initial belief state $x_k(0)$ is assumed to be the equilibrium state of the corresponding two-state Markov process, $x_k(0) = q_{k,2}$, given by (3). At each time slot t , the *Whittle index policy* schedules the $\min\{K, M\}$ users with the highest Whittle index values $W_k(x_k(t))$, where index $W_k(x_k(t))$ is given by (26). All ties are here broken randomly. Let us use acronym *WhR* for this policy.

We also define a slightly modified version of the Whittle index policy, where the difference comes from the tie-breaking rule. Let us use acronym *WhM* for the Whittle index policy that breaks the ties related to an infinite Whittle index value by applying the rule that is based on a myopic index $M_k(x_k(t))$ defined by

$$M_k(x) = h_k \max\{x\mu_{k,2}, \mu_{k,1}\}. \quad (27)$$

All the other ties are broken randomly.

For comparison, we define two other index policies (cf. [14]). The *myopic index policy* (*Mpc*) schedules greedily the $\min\{K, M\}$ users with the highest myopic index values $M_k(x_k(t))$ given in (27), thus maximizing the short-term utilization. The other one, called the *belief state index policy* (*Blf*) schedules the $\min\{K, M\}$ users with the highest belief state values $x_k(t)$. For both of these policies, all ties are broken randomly.

If all users are homogeneous, *Mpc* is throughput-optimal for the corresponding packet-level scheduling problem with partial channel information [4, 18, 35]. In this case, both *Blf* and *WhM* make, in fact, identical decisions with *Mpc* being, thus, also optimal. For our flow-level problem with heterogeneous users, *Mpc* becomes optimal in the limit $\rho_k \rightarrow 0$ for all k (as already noticed in [14]), since in this limit the channels become IID in time, and the optimal scheduling policy is known to be the $c\mu$ -rule [9], which is equivalent with the *Mpc* policy in this case.

Note that, due to mathematical tractability, the development of the Whittle index policies in our flow-level opportunistic scheduling problem (as in all similar flow-level problems [1–3, 6, 10–12, 14, 28–30, 32]) is based on a transient setting with a finite number of jobs in the beginning and no new arrivals later on. Still, as also done in many of the papers mentioned above, the developed Whittle index policies can be applied in the dynamic setting with new flows

arriving randomly. Next we evaluate numerically the performance of our Whittle index policies in such a dynamic setting.

7 NUMERICAL RESULTS

In this section, we illustrate the properties of the Whittle index policies, WhR and WhM, and compare them against Mpc and Blf. In addition, as a reference scheduler providing insight on the value of *exact* channel state information we use the well-known *PF* scheduler [8], which applies an index rule, where the index is the instantaneous channel rate divided by the current throughput. Another extreme is provided by the *Rnd* policy, which does not use any channel information but the scheduled flow is selected randomly. The rate adaptation is still greedily optimized as in other policies. The performance of the policies using partial channel information (WhR, WhM, Mpc, Blf) should lie in between Rnd and PF.

We implemented a discrete event simulator to simulate a dynamic setting with two classes of flows and $M = 1$. In the simulations, flows with a random size arrive according to a Poisson process with rate λ_k in each class k , and each class has its own channel dynamics. The simulation proceeds in discrete time from one time slot to the next one, and the scheduling is decided at the beginning of the time slot, as in actual cellular systems. During the time slot, the channel rate of each flow is constant and changes according to the considered 2-state Markov chain between the successive time slots. Let $r_{k,n}$ denote the channel rate (Mbit/s) of class- k flows in channel state $n \in \{1, 2\}$. In LTE the time slot length is 1 ms, but in our simulations, the time slot length was $\Delta = 10$ ms in order to keep simulation times manageable. The results of each simulation with a given load consists of 400 000 flow arrivals. The original flow size is exponentially distributed with mean $E[B] = 5$ Mbit for both classes k . Thus, the corresponding completion probabilities in our POMDP model are $\mu_{k,n} = r_{k,n}\Delta/E[B]$, $n = 1, 2$. This kind of mapping of the parameters between the model and a flow-level discrete event simulation has been used in many similar studies, see, e.g. [6, 10, 14].

The load in the system is defined as $\gamma = \lambda_1 E[B]/r_{1,2} + \lambda_2 E[B]/r_{2,2}$, and the maximal stability condition (with exact channel information) is $\gamma < 1$ [16]. In our simulation scenarios, we study the performance as a function of the load γ fixing $\lambda = \lambda_1 = \lambda_2$. Thus, for a given load γ , we have

$$\lambda = \frac{\gamma}{E[B](1/r_{1,2} + 1/r_{2,2})}.$$

We considered three scenarios, where the parameters, to be defined below, satisfy condition (18), i.e., $q_{k,1,2} > \mu_{k,1}/\mu_{k,2}$ for both classes k . All results are presented in Figure 1, where each panel shows the mean total number of flows in the system for different policies relative to the corresponding performance of our WhR policy as a function of the load γ .

In Scenario 1, we consider a setting where the rates in both classes are the same ($r_{1,1} = r_{2,1} = 0.01$ Mbit/s, $r_{1,2} = r_{2,2} = 50$ Mbit/s), but the channel dynamics are asymmetric ($q_{1,1,2} = 0.06$, $q_{1,2,2} = 0.96$, $q_{2,1,2} = 0.06$, $q_{2,2,2} = 0.86$), and thus have relatively high positive correlation with $\rho_1 = 0.9$ and $\rho_2 = 0.8$. The results are shown in Figure 1 (left panel). Rnd performs very poorly and it has been left out of the figure, which highlights the importance of the belief state information with highly positively correlated channels.

On the other hand, the performance of PF gives a very loose lower bound, showing that exact information gives by far superior results. On the other hand, all policies utilizing belief state information perform quite well with WhM being somewhat better than the rest. Note that since the rates are symmetric in both classes, Mpc and Blf are identical policies in this scenario.

In Scenario 2, we change the maximum rate parameters to be highly asymmetric by reducing the maximum rate of class-2 flows to $r_{2,2} = 3$ Mbit/s. In this way, under the Mpc policy, the system behaves, in fact, as a strict priority system so that class-1 flows always have absolute priority over class-2 flows independent of the belief state. The results are depicted in Figure 1 (middle panel). The Rnd and PF policies behave similarly as in Scenario 1. However, Mpc is indeed now showing clearly worse performance: because of the $c\mu$ -rule it is favoring too heavily the class-1 flows with very high $r_{1,2}$. The other policies using belief state information (WhM, WhR, Blf) all perform roughly equally well.

Finally, in Scenario 3, we reduce the degree of positive correlation significantly by choosing $\rho_1 = 0.4$ ($q_{1,1,2} = 0.3$ and $q_{1,2,2} = 0.7$) and $\rho_2 = 0.1$ ($q_{2,1,2} = 0.6$ and $q_{2,2,2} = 0.7$). Also, the rates are asymmetric between the classes so that $r_{1,1} = r_{2,1} = 8.4$ Mbit/s, but the maximum rates are $r_{1,2} = 33.6$ Mbit/s and $r_{2,2} = 16.8$ Mbit/s. This scenario should favor the Mpc policy as it is known to be optimal in the IID setting. Also, the Rnd policy should benefit as the belief state information used by the other policies is not in this context as useful as in the case with higher positive correlation. The results are shown in Figure 1 (right panel). Indeed, as predicted, Rnd is performing better than in the previous scenarios. Also, Mpc gives the best performance among the policies utilizing the belief state. Both Whittle index policies (WhM, WhR) are also performing well, but Blf is clearly worse than all these.

Overall, the Whittle index policies (WhM, WhR) perform well in all these scenarios, independent of the degree of positive correlation in channel dynamics of the flows. This is in contrast with the properties of the more heuristic Mpc and Blf policies, which may change with the correlations.

8 CONCLUSIONS

We considered flow-level opportunistic scheduling problem of down-link data traffic with partial channel information, where the channel dynamics are described by two-state positively autocorrelated Markov processes and the scheduler estimates the channel state based on ARQ-type feedback from the scheduled users.

Since the original problem belongs to the class of restless bandit problems, which are known to be mathematically intractable, the Whittle index approach has been utilized in earlier studies to develop near-optimal scheduling policies for the corresponding problems. We used the same approach and managed to extend the results found thus far: We were able to find sufficient conditions under which the problem with partial channel information is provably indexable, derived an explicit formula for the corresponding Whittle index, and proved that the optimal policy for the relaxed optimization problem is of threshold type in this case.

In our numerical simulations with dynamic flow-level traffic, the derived Whittle index policies performed well, independent of the degree of positive correlation in channel dynamics, which was not

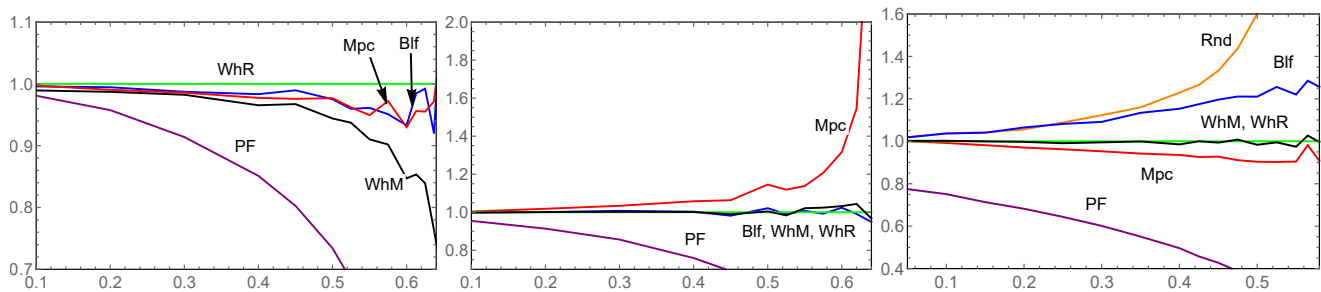


Figure 1: Mean total number of flows in the system for different policies relative to WhR policy as a function of the load for Scenario 1 (left panel), Scenario 2 (middle panel) and Scenario 3 (right panel).

the case with the more heuristic policies that we studied in our simulations.

As further work, it would be interesting to find out whether this flow-level opportunistic scheduling problem with partial channel information, where the channel dynamics are described by two-state positively autocorrelated Markov processes, is indexable for any other parameter values. From the practical point of view, the assumption that the channels are positively autocorrelated may be the most interesting. However, from the theoretical point of view, it would also be nice to find out whether the indexability property may be proved for negatively autocorrelated channels. Even harder would still be to consider the corresponding problems with more than two channel states.

REFERENCES

- [1] S. Aalto, P. Lassila, and P. Osti. Whittle index approach to size-aware scheduling with time-varying channels. In *Proc. of ACM Sigmetrics*, pages 57–69, June 2015.
- [2] S. Aalto, P. Lassila, and P. Osti. Whittle index approach to size-aware scheduling for time-varying channels with multiple states. *Queueing Systems*, 83:195–225, 2016.
- [3] S. Aalto, P. Lassila, and P. Osti. Opportunistic scheduling with flow size information for Markovian time-varying channels. *Performance Evaluation*, 112:27–52, 2017.
- [4] S. Ahmad, M. Liu, T. Javadi, Q. Zhao, and B. Krishnamachari. Optimality of myopic sensing in multichannel opportunistic access. *IEEE Transactions on Information Theory*, 55:4040–4050, 2009.
- [5] A. Asadi and V. Mancuso. A survey on opportunistic scheduling in wireless communications. *IEEE Communications Surveys & Tutorials*, 15:1671–1688, 2013.
- [6] U. Ayesta, M. Eraisquin, and P. Jacko. A modeling framework for optimizing the flow-level scheduling with time-varying channels. *Performance Evaluation*, 67:1014–1029, 2010.
- [7] U. Ayesta, M. Eraisquin, M. Jonckheere, and I. Verloop. Scheduling in a random environment: Stability and asymptotic optimality. *IEEE/ACM Transactions on Networking*, 21:258–271, 2013.
- [8] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi. CDMA/HDR: A bandwidth efficient high speed wireless data service for nomadic users. *IEEE Communications Magazine*, 38(7):70–77, 2000.
- [9] C. Buyukkok, P. Varaiya, and J. Walrand. The $c\mu$ rule revisited. *Advances in Applied Probability*, 17:237–238, 1985.
- [10] F. Cecchi and P. Jacko. Scheduling of users with Markovian time-varying transmission rates. In *Proc. of ACM Sigmetrics*, pages 129–140, June 2013.
- [11] F. Cecchi and P. Jacko. Nearly-optimal scheduling of users with Markovian time-varying transmission rates. *Performance Evaluation*, 99-100:16–36, 2016.
- [12] P. Jacko. Value of information in optimal flow-level scheduling of users with Markovian time-varying channels. *Performance Evaluation*, 68:1022–1036, 2011.
- [13] P. Jacko, E. Morozov, L. Potakhina, and I. Verloop. Maximal flow-level stability of best-rate schedulers in heterogeneous wireless systems. *Transactions on Emerging Telecommunications Technologies*, 28(1), 2017.
- [14] P. Jacko and S. Villar. Opportunistic schedulers for optimal scheduling of flows in wireless systems with ARQ feedback. In *Proc. of ITC 24*, September 2012.
- [15] A. Jalali, R. Padovani, and R. Pankaj. Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system. In *Proc. of IEEE VTC 2000 Spring*, pages 1854–1858, May 2000.
- [16] J. Kim, B. Kim, J. Kim, and Y. Bae. Stability of flow-level scheduling with Markovian time-varying channels. *Performance Evaluation*, 70:148–159, 2013.
- [17] R. Knopp and P. Humblet. Information capacity and power control in single-cell multiuser communications. In *Proc. of IEEE ICC*, pages 331–335, June 1995.
- [18] G. Koole, Z. Liu, and R. Righter. Optimal transmission policies for noisy channels. *Operations Research*, 49:892–899, 2001.
- [19] M. Larrañaga, M. Assaad, A. Destounis, and G. Paschos. Asymptotically optimal pilot allocation over Markovian fading channels. *IEEE Transactions on Information Theory*, 64:5395–5418, 2018.
- [20] J. Le Ny, M. Dahleh, and E. Feron. Multi-UAV dynamic routing with partial observations using restless bandit allocation indices. In *Proc. of ACC*, pages 4220–4225, June 2008.
- [21] K. Liu, R. Weber, and Q. Zhao. Indexability and Whittle index for restless bandit problems involving reset processes. In *Proc. of IEEE CDC-ECC*, pages 7690–7696, December 2011.
- [22] K. Liu and Q. Zhao. Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access. *IEEE Transactions on Information Theory*, 56:5547–5567, 2010.
- [23] J. Niño-Mora. An index policy for dynamic fading-channel allocation to heterogeneous mobile users with partial observations. In *Proc. of NGI*, April 2008.
- [24] W. Ouyang, A. Eryilmaz, and N. Shroff. Asymptotically optimal downlink scheduling over Markovian fading channels. In *Proc. of IEEE Infocom*, pages 1224–1232, March 2012.
- [25] W. Ouyang, A. Eryilmaz, and N. Shroff. Low-complexity optimal scheduling over time-correlated fading channels with ARQ feedback. *IEEE Transactions on Mobile Computing*, 15:2275–2289, 2016.
- [26] W. Ouyang, S. Murugesan, A. Eryilmaz, and N. Shroff. Exploiting channel memory for joint estimation and scheduling in downlink networks. In *Proc. of IEEE Infocom*, pages 3056–3064, April 2011.
- [27] M. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2005.
- [28] I. Taboada, P. Jacko, U. Ayesta, and F. Liberal. Opportunistic scheduling of flows with general size distribution in wireless time-varying channels. In *Proc. of ITC 26*, September 2014.
- [29] I. Taboada, F. Liberal, J. Fajardo, and B. Blanco. An index rule proposal for scheduling in mobile broadband networks with limited channel feedback. *Performance Evaluation*, 117:130–142, 2017.
- [30] I. Taboada, F. Liberal, and P. Jacko. An opportunistic and non-anticipating size-aware scheduling proposal for mean holding cost minimization in time-varying channels. *Performance Evaluation*, 79:90–103, 2014.
- [31] P. Viswanath, D. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory*, 48:1277–1294, 2002.
- [32] K. Wang, J. Yu, L. Chen, and M. Win. Opportunistic scheduling revisited using restless bandits: Indexability and index policy. In *Proc. of IEEE Globecom*, December 2017.
- [33] R. Weber and G. Weiss. On an index policy for restless bandits. *Journal of Applied Probability*, 27:637–648, 1990.
- [34] P. Whittle. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25A:287–298, 1988.
- [35] Q. Zhao, B. Krishnamachari, and K. Liu. On myopic sensing for multichannel opportunistic access: Structure, optimality, and performance. *IEEE Transactions on Wireless Communications*, 7:5431–5440, 2008.