Lassila, Pasi; Aalto, Samuli

Near-optimal uplink scheduling for age-energy tradeoff in wireless systems

Published in:
2021 33th International Teletraffic Congress (ITC-33)

Published: 01/01/2021

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

Please cite the original version:
Lassila, P., & Aalto, S. (2021). Near-optimal uplink scheduling for age-energy tradeoff in wireless systems. In *2021 33th International Teletraffic Congress (ITC-33)* IEEE. https://ieeexplore.ieee.org/document/9625609

Near-optimal uplink scheduling for age-energy tradeoff in wireless systems

Pasi Lassila
Department of Communications and Networking
Aalto University, Finland

Samuli Aalto
Department of Communications and Networking
Aalto University, Finland

***Abstract*—We study the optimal age-energy tradeoff arising in the context uplink scheduling of multiple heterogeneous devices that are transmitting time-sensitive update packets through the base station in a cellular network. In the model, new updates arrive stochastically at the devices and they are subject to losses when attempting to transmit. Also, associated with the update transmission are related energy costs. The problem is to develop policies for minimizing the weighted sum of the total mean age and energy costs. The Whittle index approach is applied to obtain a near-optimal policy. In particular, we prove the indexability of the problem and explicitly derive the associated Whittle indices. Our numerical experiments demonstrate the superiority of the resulting Energy-aware Whittle index policy compared with other heuristics.**

I. Introduction

Age-of-Information (AoI) has in the recent years become an important performance metric for time-sensitive data collected by various devices and transmitted over a network to servers, where the data is processed and utilized. The interest has been strongly motivated by recent efforts in the standardization of 5G cellular systems to provide support for machine-to-machine type communications with extremely low latency requirements.

Essentially, AoI measures the freshness of the information updates at the receiver and it is defined as the time difference between the current time and the time that the previous update was received. AoI is a stochastic process since the time that the previous update was received is characterized by the output process of the network over which the update message was sent. In the seminal paper [1], the authors analyzed the AoI process as a general stochastic process, where update transmissions require a random service time and are served by a single server FIFO queue, as well as derived the mean AoI for elementary queues, such as M/M/1 and M/D/1. The main observation from the analysis is that to minimize mean AoI one needs to balance the frequency of the generation of updates (inter-arrival times) with the likelihood of observing a long queue yielding a strictly positive value for the optimal load, which is very different to mean delay performance that would be minimized by having load zero or to throughput performance that is maximized by load arbitrarily close to one. Since then the research on AoI has flourished and by now a large body of research exists, see [2] for a very recent and in-depth survey. The analysis of optimization of AoI has been extended to alternative age metrics (even non-linear age costs), optimization of AoI in more complex queueing systems with various update packet replacement policies and wireless networks under different physical and MAC layer assumptions, also sometimes including energy constraints.

In this paper, we consider a set of devices (sources) that are attempting to transmit their update packets in the uplink direction to a base station. The task of the base station is to schedule the devices in each time slot and we allow the base station to schedule multiple devices per time slot. The devices are bufferless and new updates arrive randomly in each time slot. When a device is scheduled the update is successful with a given probability. Associated with each uplink transmission is additionally an energy cost at the device. If the energy cost is high enough it could be better from the device point of view to still postpone the update transmission attempt, which gives rise to a tradeoff between the age of the devices and the energy consumption to perform the updates. The base station is assumed to be aware of the age of each device. Based on the age information together with other system parameters, including the energy costs of the devices, the objective is to determine the optimal scheduling policy that minimizes the weighted sum of the mean AoI and the energy costs. Note that in the downlink direction this tradeoff does not exist since the base station is always on, but in the uplink each device can save energy by not transmitting.

The age-energy tradeoff considered in our problem differs from the one studied in the context of information age with energy harvesting devices, where the problem setup involves a single device with a finite battery supply which is being charged by a harvesting process creating a stochastic supply of energy available for transmitting information updates. The tradeoff then involves selecting when to make transmission attempts subject to the energy constraint to minimize AoI. This has been analyzed assuming an error-free wireless channel in [3], [4], and the optimal policy has been shown to have a certain threshold structure. Similar results for noisy channels have been recently given in [5].

The problem that we consider can be formulated as a Markov Decision Process (MDP) problem. However, explicitly solving the optimal policy is infeasible and to overcome this we apply the celebrated Whittle index approach, originally proposed in [6] for restless multiarmed bandit problems (RMAB). Here the idea is to relax the original strict scheduling constraint that must be satisfied in every time slot such that it is only required to hold on average. By using Lagrangian techniques, the original problem involving joint optimization of the age process for all devices is decomposed into an MDP, where each device is considered separately, thus reducing the complexity. A technical assumption in the approach is that the problem must be so-called indexable. However, a priori it is not necessarily known whether a given problem is indexable or not and it needs to be verified and this is the technical challenge in the approach. The Whittle index approach has been used in many application scenarios recently, e.g., the opportunistic scheduling problem [7], [8], [9] and the job dispatching problem [10], [11], and shown numerically to perform very well. Also, it is known that, if the problem is indexable, the Whittle index policy is asymptotically optimal [12], [13], at least under certain technical conditions.

As our main theoretical contribution, we show that our considered problem is indexable and we derive the explicit form of the Whittle index. In our numerical simulation experiments, we demonstrate that our Energy-aware Whittle index policy performs significantly better than any other heuristic policy. Also, we consider the impact of scheduling multiple devices in a time slot. By simulations we also illustrate that the Energy-aware Whittle index policy systematically provides better performance the higher the number of devices that can be scheduled. Based on the index values, the base station can estimate whether it is actually advantageous for a device to transmit, and, if not, the base station can allocate that transmission opportunity for use by other traffic in the system. This is in contrast with all the other heuristic policies which always select exactly the given number of devices and there is for any combination of other parameters an optimal number of devices to be scheduled for minimizing the performance. Thus, the Energy-aware Whittle index is robust against the precise selection of the number of devices to be scheduled.

The Whittle index approach has been applied to age-optimal scheduling with multiple devices, but without any energy aspects, recently in a few papers [14], [15], [16], [17], [18]. A similar model to ours, including the bufferless assumption, has been analyzed in [15], see also the conference version [14]. The authors prove the indexability and the explicit form of the Whittle index is derived. Similarly, [16] also presents an index policy, but the age model is slightly different than in [14], [15]. Finally, the authors in [18] (original conference paper [17]) relax the no buffer assumption resulting in a two-dimensional age process with exact knowledge of age at sender and scheduler. Indexability is proven and the index values are derived, thus generalizing the results of [15]. Our model differs from all these works in that we include the age-energy tradeoff. Additionally, we study the impact of scheduling simultaneously multiple devices, which is not considered in any of the previous works. Also, technically our proof of indexability uses a different approach.

The paper is organized as follows. Section II presents the system model and the original MDP formulation. The Whittle index approach and the decomposition of the MDP problem in given in Section III. The proof of indexability is given in Section IV. The policies to be used in numerical examples, including the Energy-aware Whittle index policy, are presented in Section V. Numerical examples are in Section VI and the conclusions can be found in Section VII.

II. System model and the original MDP problem

A. System model

Consider a wireless base station with K devices transmitting time-sensitive information updates in the uplink direction from the devices towards the base station. The devices are indexed by $k = 1, \ldots, K$. Time is slotted with n denoting the n:th time slot, i.e., $n \in \{0, 1, \ldots\}$. In each time slot, a fresh information update is generated at device k with probability λ_k, which is independent from everything else in the system. Also, for simplicity, we assume that the device does not have a buffer, where an update message could be saved when a new one is not available in the current slot.

The devices try to send their information updates to the receiver through the base station as quickly as possible in order to have the most recent information available at the receiver. To measure the timeliness of the information from each device, the base station keeps track of the time since the previous update was observed at the base station, which is called the age of information for device k and it is denoted by $A_k(n)$. Initially, we assume that at time $n = 0$ the age of each device is set to 1, i.e., $A_k(0) = 1, k = 1, \ldots, K$.

Device k can only make a transmission attempt in time slot n if there is an update available, which happens with probability λ_k independently at any slot n. Thus, we assume (for tractability) that the devices do not have a buffer where to store the most recent update. The devices are additionally energy-constrained and given that device k makes a transmission attempt in time slot n, this will always incur an energy cost E_k, which represents the amounf of energy (in Joules) spent by the device for the transmission during a time slot. However, the transmission attempt will be successful with probability p_k and with probability $(1 - p_k)$ the transmission attempt fails due to poor channel conditions.

Device k can not independently decide to transmit, but needs a scheduling grant from the base station. Thus, the problem is for the base station to schedule the devices in order to optimize the costs incurred by the age of the devices and the energy costs. In our case, this can be represented as an MDP problem.

B. MDP formulation for optimal age-energy tradeoff

We assume that the base station makes a decision $D_k(n)$, for all k, at the beginning of slot n and that the decision is instantaneously available at the transmitters, which then implement the actions. However, the base station at the instant of making the scheduling decisions is not aware of whether an update is currently available at each of the devices. It can only utilize age information of the devices at the beginning of the slot $A_k(n)$ and the other parameters, i.e., energy costs E_k, arrival rates λ_k and success probabilities p_k. We define that $D_k(n) = 1$, if the base station decides to schedule device

k in slot n. Correspondingly, $D_k(n) = 0$, if the base station decides not to schedule. Altogether, in one time slot the base station can schedule a maximum of M devices, and thus we have the constraint

$$\sum_{k=1}^{K} D_k(n) \leq M, \quad \forall n. \tag{1}$$

The decisions $D_k(n), \forall(k,n)$, constitute the policy π and affect the age process $A_k(n)$ of each device. To highlight this dependence we from now on explicitly include π in the decision variables, $D_k^\pi(n)$ and age variables $A_k^\pi(n)$.

The state of the system in slot n is described by $\{A_1^\pi(n), \ldots, A_K^\pi(n)\}$, i.e., the vector of all ages, whose components evolve in the following manner. Consider device k in time slot n with age $A_k^\pi(n) = a$. If $D_k^\pi(n) = 1$, the device is allowed to transmit and if there is an update available and the transmission is successful in time slot n, which happens with joint probability $\lambda_k p_k$, the age of device k is initialized to 1 at the base station in time slot $n+1$, i.e., $A_k^\pi(n+1) = 1$. With probability $(1-\lambda_k p_k)$ the scheduled transmission attempt fails and the age increases by one, i.e., $A_k^\pi(n+1) = a+1$. Also, if $D_k^\pi(n) = 0$, then naturally $A_k^\pi(n+1) = a+1$ with probability 1. To minimize age $A_k^\pi(n)$, one should schedule as often as possible, but this decision involves also an energy cost E_k with probability λ_k, i.e., if an update is available at device k. Thus, there is an inherent tradeoff between minimizing the age and the energy costs.

To represent the age-energy tradeoff, as the cost for each device k in time slot n we consider the weighted sum of the instantaneous age and the expected energy cost, i.e., $A_k^\pi(n) + D_k^\pi(n)\lambda_k w_k E_k$, where $w_k \geq 0$ is a weighting factor that converts the energy cost (in Joules) into time slots (dimensionless unit), i.e., the unit of w_k is [1/J]. For generality the weight parameter w_k is allowed here to depend on the device index. The overall objective is then to find a policy π^* that minimizes the long run average costs (weighted sum of mean age and energy costs),

$$\lim_{N\to\infty} E\left[\frac{1}{N}\sum_{n=0}^{N-1}\sum_{k=1}^{K}(A_k^\pi(n) + D_k^\pi(n)\lambda_k w_k E_k)\right], \tag{2}$$

subject to the scheduling constraint (1). This is clearly an MDP problem. However, due to the multidimensional and unbounded nature of the age process, obtaining the optimal solution to the problem is not feasible. As an approximate solution to the problem, we will next consider the so-called Whittle index approach.

III. Relaxed optimization problem

The problem (2) belongs in the class of restless multi-armed bandit problems (RMAB). The innovative idea developed originally by Whittle [6] for overcoming the inherent intractability of such problems was to relax the strict scheduling constraint (1), which must be satisfied for every time slot n, by requiring it to hold only on average. This allows to combine the original objective function (2) and the constraint into one, the so-called Lagrangian function. Furthermore, the overall problem involving K devices can be decoupled into K separate sub-problems, one for each device, involving determining the policy π_k^* that minimizes

$$\lim_{N\to\infty} E\left[\frac{1}{N}\sum_{n=0}^{N-1}(A_k^\pi(n) + D_k^\pi(n)\lambda_k w_k E_k) + \nu D_k^\pi(n)\right], \tag{3}$$

where ν is a real-valued variable referred to as the Lagrange multiplier. The variable ν can also be interpreted as the unit price of activity (scheduling in this case). Minimizing (3) can again be characterized as an MDP, but with a significantly simpler structure than in our original problem. Utilizing the solution to this allows us later on to define an index policy that can then be used as the near optimal policy for solving the original problem.

Note that our formulation of the constraint (1) does not require it to be satisfied as a strict equality, unlike in the original formulation in [6]. However, allowing the constraint to be defined as an inequality has been analyzed, e.g., in [13]. We will later see how the inequality affects the optimal solution.

Next let us define in more detail the MDP formulation associated with minimizing (3). The problem is now one-dimensional, where the state of a bandit is given by $a \in \mathcal{A}$, where $\mathcal{A} = \{1, 2, \ldots\}$ is the state space of the age of device k, and let $d \in \mathcal{D}$, where $\mathcal{D} = \{0, 1\}$ is the decision space.

We denote by $c_k(a, d; \nu)$ the mean cost in state a with decision d and a given value of the price of activity (scheduling) ν for device k. In our case, this is given by

$$c_k(a, d; \nu) = a + d(\lambda_k w_k E_k + \nu). \tag{4}$$

Also, let $q_k(a'|a, d)$ denote the conditional one-step transition probabilities for making a transition to state a' given that in the previous slot the state was a and that decision d was taken. The non-zero transition probabilities $q_k(a'|a, d)$ in our case are clearly, for any $a \geq 1$,

$$\begin{aligned} &q_k(a+1|a, 0) = 1, \\ &q_k(1|a, 1) = \lambda_k p_k, \\ &q_k(a+1|a, 1) = 1 - \lambda_k p_k. \end{aligned}$$

Since the state space is discrete, the decision space finite and the cost rate (4) is linear with respect to the state a there is a stationary deterministic policy π_k^*, which minimizes the long run average costs (3) [19]. Furthermore, the optimal policy π_k^* is characterized by the set of optimality equations defined for each state $a \in \mathcal{A}$

$$v_k(a;\nu) = \min_{d\in\{0,1\}}\Bigg\{c_k(a, d; \nu) - \bar{c}_k(\nu) + \sum_{a'\in\mathcal{A}} q_k(a'|a, d) v_k(a';\nu)\Bigg\}, \tag{5}$$

where $\bar{c}_k(\nu)$ is the minimum expected cost under the optimal policy π_k^* and $v_k(a;\nu)$ is the so-called value function referring to the expected difference in the total cost when starting the process from state a and in equilibrium under the optimal policy π_k^*. In our case, (5) becomes

$$\begin{aligned} v_k(a;\nu) = \min\{& a - \bar{c}_k(\nu) + v_k(a+1;\nu), \\ & a + \lambda_k w_k E_k + \nu - \bar{c}_k(\nu) + \lambda_k p_k v_k(1;\nu) \\ & + (1-\lambda_k p_k) v_k(a+1;\nu)\}, \end{aligned} \tag{6}$$

where the terms inside the minimum correspond to mean costs for the decisions not to schedule ($d = 0$) and to schedule ($d = 1$), respectively.

Let us denote by $\Delta_k(a;\nu)$ the difference of the value functions in state a and state 1,

$$\Delta_k(a;\nu) = v_k(a;\nu) - v_k(1;\nu).$$

From (6), we clearly get the following condition for optimal decisions in each state $a \in \mathcal{A}$,

(i) If $\lambda_k w_k E_k + \nu - \lambda_k p_k \Delta_k(a+1;\nu) < 0$,
it is optimal to schedule ($d = 1$) device k.
(ii) If $\lambda_k w_k E_k + \nu - \lambda_k p_k \Delta_k(a+1;\nu) > 0$,
it is optimal not to schedule ($d = 0$) device k. (7)
(iii) If $\lambda_k w_k E_k + \nu - \lambda_k p_k \Delta_k(a+1;\nu) = 0$,
both decisions are equally optimal for device k.

We refer to (7) as the optimality condition in state a and it will be utilized often in our proof of the indexability of our problem.

IV. The Whittle index

In the Whittle index approach, the main technical challenge is that one needs to prove that the problem is indexable [6], which is not a priori obvious in our case. In this section, we prove the indexability of our problem and explicitly derive the corresponding Whittle index values.

The optimal solution to the relaxed problem with the objective function (3) can be considered now as an unconstrained problem for any given value of the price of activity $\nu \in [-\infty, \infty]$. The solution is characterized by the Whittle index values, when the problem is indexable, which is defined as follows.

Definition 1: The optimization problem with objective function (3) is *indexable* if, for any $a \in \mathcal{A}$, there exists $\nu_k^*(a) \in [-\infty, \infty]$ such that

(i) decision 1 (to schedule device k) is optimal in state a if $\nu_k^*(a) \geq \nu$;
(ii) decision 0 (not to schedule device k) is optimal in state a if $\nu_k^*(a) \leq \nu$.

The value $\nu_k^*(a)$ is referred to as the Whittle index for state a. Also, note that it follows from the definition that if $\nu_k^*(a) = \nu$ both actions are equally good. In other words, the Whittle index for state a, $\nu_k^*(a)$, is precisely the value of the Lagrange parameter (price of scheduling) ν, which makes the decisions 0 (not to schedule) and 1 (to schedule) equally optimal (or good) in state a. Note that, as shown in [13], if for some state a the optimal $\nu = \nu_k^*(a) < 0$ then that device is never scheduled in the optimal solution since our constraint (1) does not have to be satisfied as an equality.

Next we proceed to proving the indexability property. The analysis only concerns a given device k in isolation, and hence we leave out the explicit dependence on k from the notation.

A. Analysis of threshold policy $\pi(m)$

Our proofs rely on the properties of the following threshold policy $\pi(m)$. The policy $\pi(m)$ is defined such that the device is scheduled (i.e., decision $d = 1$) in any state $a \geq m \geq 1$, and not scheduled in the states $a \in \{1, \ldots, m-1\}$, respectively.

Let us denote by $p(a,m)$ the steady state probability for the age process under policy $\pi(m)$. It is straightforward to derive the steady state probabilities from the global balance equations, which gives

$$p(a,m) = \frac{\lambda p}{(m-1)\lambda p + 1} \cdot \begin{cases} 1, & \text{if } a < m, \\ (1-\lambda p)^{a-m}, & \text{if } a \geq m. \end{cases}$$

Correspondingly, the cost rate in any given state is given by (4) and hence the mean cost $\bar{c}(m)$ with policy $\pi(m)$ equals

$$\begin{aligned} \bar{c}(m) &= \sum_{a=1}^{m-1} c(a,0;\nu)\, p(a,m) + \sum_{a=m}^{\infty} c(a,1;\nu)\, p(a,m) \\ &= \frac{m}{2} + \frac{1}{\lambda p} + \frac{w\lambda E + \nu - m/2}{1 + (m-1)\lambda p}. \end{aligned}$$

In particular, the value function differences $\Delta^{\pi(m)}(a;\nu)$ for state a under policy $\pi(m)$ are key in the optimality condition (7) and are utilized in the indexability proof. These are given in the proposition below.

Proposition 1: The value function difference in a given state a and value ν for policy $\pi(m)$, $\Delta^{\pi(m)}(a;\nu)$, is given by

$$\begin{aligned} \Delta^{\pi(m)}(a;\nu) &= v^{\pi(m)}(a;\nu) - v^{\pi(m)}(1;\nu) \\ &= \begin{cases} 0, & \text{if } a = 1, \\ f_0(a,m,\nu), & \text{if } a = 2, \ldots, m-1, \\ f_1(a,m,\nu), & \text{if } a \geq m, \end{cases} \end{aligned} \tag{8}$$

where

$$\begin{aligned} f_0(a,m,\nu) = &\frac{1}{(m-1)\lambda p + 1} \times \\ &\Big(\frac{a-1}{\lambda p} - \frac{1}{2}(a-1)(a-2(m-1)) \\ &+ \frac{m-1}{2}(a-1)(m-a)\lambda p + (a-1)(w\lambda E + \nu)\Big), \end{aligned}$$

and

$$\begin{aligned} f_1(a,m,\nu) = &\frac{1}{(m-1)\lambda p + 1} \times \\ &\Big(\frac{a-1}{\lambda p} + (m-1)a - \frac{1}{2}(m^2 + m - 2) \\ &+ (m-1)(w\lambda E + \nu)\Big). \end{aligned}$$

Proof. The value functions for each state and given policy $\pi(m)$, $v^{\pi(m)}(a;\nu)$, are defined, up to an arbitrary additive constant that can be chosen as $v^{\pi(m)}(1;\nu)$, by a set of linear equations, one for each state, i.e., the so-called Howard equations [19]. In our case, the Howard equations read as follows:

$$\begin{aligned} v^{\pi(m)}(a;\nu) &= a - \bar{c}(m) \\ &\quad + v^{\pi(m)}(a+1;\nu), && \text{if } 1 \le a \le m-1, \\ v^{\pi(m)}(a;\nu) &= a + w\lambda p + \nu - \bar{c}(m) \\ &\quad + \lambda p v^{\pi(m)}(1;\nu) \\ &\quad + (1-\lambda p) v^{\pi(m)}(a+1;\nu), && \text{if } a \ge m. \end{aligned}$$

The two equations are related to states, where the device is not scheduled and scheduled, respectively. Note that when $m = 1$, device is always scheduled and the first equation is redundant.

It is easy to check that, for any m, the value functions as defined in (8) satisfy the above Howard equations in any state $1 \le a < m-1$ (device is not scheduled in state a and next state $a+1$), state $a = m-1$ (device is not scheduled in state $m-1$ but is scheduled in state m) and state $a \ge m$ (device is always scheduled), which completes the proof. □

B. Proof of indexability

Theorem 1: The relaxed MDP problem with objective function (3) is indexable and the Whittle index in state a, $\nu^*(a)$, is given by

$$\nu^*(a) = a + \frac{\lambda p}{2} a(a-1) - w\lambda E. \tag{9}$$

Proof. In the proof, we cover the whole interval $-\infty < \nu < \infty$. The first interval spans $-\infty < \nu \le \nu^*(1)$, while the remaining part of the range of ν values is split into subintervals $\nu^*(a-1) \le \nu \le \nu^*(a)$. Since $\nu^*(a)$ is an increasing function of a, the whole range $-\infty < \nu < \infty$ will be covered by the proof.

1° Assume that $-\infty < \nu \le \nu^*(1)$. Below we show that the threshold policy $\pi(1)$, which always schedules the device for any state $a \ge 1$, is optimal.

From (8), we get that

$$\Delta^{\pi(1)}(a;\nu) = \frac{a-1}{\lambda p}.$$

Consider state $a = 1$. By the optimality condition (7), $\pi(1)$ is optimal in state 1 if and only if

$$\begin{aligned} & w\lambda E + \nu - \lambda p \Delta^{\pi(1)}(2;\nu) \le 0 \\ \Leftrightarrow & w\lambda E + \nu - \lambda p \frac{1}{\lambda p} \le 0 \\ \Leftrightarrow & \nu \le 1 - w\lambda E = \nu^*(1). \end{aligned}$$

Next we consider the states $a \in \{2, 3, \ldots\}$. From the assumption $\nu \le \nu^*(1)$, it follows

$$\begin{aligned} & w\lambda E + \nu - \lambda p \Delta^{\pi(1)}(a+1;\nu) \\ \le\ & w\lambda E + (1 - w\lambda E) - \lambda p \frac{a}{\lambda p} \\ =\ & 1 - a < 0, \end{aligned}$$

since $a \ge 2$. We conclude by the optimality condition (7) that if $\nu \le \nu^*(1)$ then $\pi(1)$ is optimal for any state a, and in state 1 both decisions are equally optimal if $\nu = \nu^*(1)$, which is also the Whittle index of state 1.

2° Now assume that $m \ge 2$ and $\nu^*(m-1) \le \nu \le \nu^*(m)$. Below we show that the threshold policy $\pi(m)$ is optimal in the considered generic interval. The proof is done in four steps.

Step 1: Let the state $a \in \{1, \ldots, m-2\}$, i.e., the set of states, where policy $\pi(m)$ does not schedule the device. Note that this step is redundant for $m = 2$ and thus we may assume $m \ge 3$. Since $\nu \ge \nu^*(m-1)$, it follows by using (8), where $f_0(a, m, \nu)$ is an increasing function of ν, and (9) that

$$\begin{aligned} & w\lambda E + \nu - \lambda p \Delta^{\pi(m)}(a+1;\nu) \\ \ge\ & w\lambda E + \nu^*(m-1) - \lambda p f_0(a+1, m, \nu^*(m-1)) \\ =\ & \frac{1}{2}(m - (a+1))(\lambda p(m - (a+2)) + 2) \ge 1 > 0, \end{aligned}$$

since $a \le m-2$. Thus, by the optimality condition (7) policy $\pi(m)$ is optimal in states $a \in \{1, \ldots, m-2\}$ when $\nu \ge \nu^*(m-1)$.

Step 2: Now consider the state $a = m-1$. According to policy $\pi(m)$ the device is not scheduled in state $a = m-1$, but it is scheduled in state $a = m$. By the optimality condition (7), $\pi(m)$ is optimal in state $a = m-1$ if and only if

$$\begin{aligned} & w\lambda E + \nu - \lambda p \Delta^{\pi(m)}(m;\nu) \ge 0 \\ \Leftrightarrow & w\lambda E + \nu - \lambda p f_1(m, m, \nu) \ge 0 \\ \Leftrightarrow & \nu \ge \nu^*(m-1). \end{aligned}$$

The last inequality above can be verified readily by using the expressions from (8) and (9). We conclude by the optimality condition (7) that if $\nu > \nu^*(m-1)$ then $\pi(m)$ is optimal for state $a = m-1$ and both decisions are equally optimal if $\nu = \nu^*(m-1)$, which is also the Whittle index of state $a = m-1$.

Step 3: Now consider the state $a = m$. According to policy $\pi(m)$ the device is scheduled in state $a = m$ and also in state $a = m+1$. By the optimality condition (7), $\pi(m)$ is optimal in state $a = m$ if and only if

$$\begin{aligned} & w\lambda E + \nu - \lambda p \Delta^{\pi(m)}(m+1;\nu) \le 0 \\ \Leftrightarrow & w\lambda E + \nu - \lambda p f_1(m+1, m, \nu) \le 0 \\ \Leftrightarrow & \nu \le \nu^*(m). \end{aligned}$$

The last inequality above can be verified readily by using the expressions from (8) and (9). We conclude by the optimality condition (7) that if $\nu < \nu^*(m)$ then $\pi(m)$ is optimal for state $a = m$ and both decisions are equally optimal if $\nu = \nu^*(m)$, which is also the Whittle index of state $a = m$.

Step 4: Let the state $a \in \{m+1, \ldots\}$, i.e., the set of states, where policy $\pi(m)$ always schedules the device. Since $\nu \le \nu^*(m)$ it follows by using (8) and (9)

$$\begin{aligned} & w\lambda E + \nu - \lambda p \Delta^{\pi(m)}(a+1;\nu) \\ \le\ & w\lambda E + \nu^*(m) - \lambda p f_1(a+1, m, \nu^*(m)) \\ =\ & m - a \le -1 < 0, \end{aligned}$$

since $a \ge m+1$. Thus, by the optimality condition (7) policy $\pi(m)$ is optimal in states $a \in \{m+1, \ldots\}$ when $\nu \le \nu^*(m)$.

By steps 1 – 4 above we have shown that policy $\pi(m)$ is optimal in the interval $\nu^*(m-1) \leq \nu \leq \nu^*(m)$ for any state a. Furthermore, if $\nu = \nu^*(m-1)$ (or $\nu = \nu^*(m)$, respectively) both decisions are equally optimal in state $a = m-1$ (or state $a = m$, respectively), which are also the corresponding Whittle indices. □

V. The Whittle index policy and other heuristics

In this section, we define the Whittle index policy and other heuristic index policies that will then be applied in the numerical examples. Consider now the original problem to minimize the objective function (2) with K devices and the scheduler (i.e., the base station) must select in each time slot at most M devices to be scheduled, according to the constraint (1). Here we again bring back into the notation the explicit dependence on the device index k.

Energy-aware Whittle index policy: For a device k in state a, we define the index

$$\nu_k^{\mathrm{EW}}(a) = \nu_k^*(a) = a + \frac{\lambda_k p_k}{2} a(a-1) - w_k \lambda_k E_k,$$

as given in Theorem 1. The *energy-aware Whittle index* policy is defined as the policy that selects at most M devices with the largest Whittle index values $\nu_k^{\mathrm{EW}}(a)$ such that $\nu_k^{\mathrm{EW}}(a) > 0$. In case of any ties, they are broken randomly.

Max-age policy: As a simple heuristic, we define the non-energy aware policy, which simply attempts to stabilize the age process of all devices. Thus, the index of device k is simply

$$\nu_k^{\mathrm{Max}}(a) = a.$$

The *Max-Age* policy is defined as the policy that selects exactly M devices with the largest ages, i.e., the index values $\nu_k^{\mathrm{MA}}(a)$. In case of any ties, they are broken randomly.

Myopic policy: In the Myopic policy, the idea is to only consider the change in the immediate mean age and energy costs between the two decisions. If the device is scheduled the immediate mean costs are simply $\lambda_k p_k \cdot 1 + (1 - \lambda_k p_k) \cdot (a+1) + w_k \lambda_k E_k$. If the device is not scheduled the mean cost is only $(a+1)$. Thus, for device k the index is defined as the difference in the costs for the two decisions,

$$\nu_k^{\mathrm{Myo}}(a) = -\lambda_k p_k a + w_k \lambda_k E_k.$$

The *Myopic* policy is defined as the policy that greedily selects M devices with the smallest index values $\nu_k^{\mathrm{Myo}}(a)$.

All above policies are dynamic, depending on the instantaneous state of the devices. As a static randomized policy we can consider the *Random* policy that schedules randomly M devices (out of K) in each time slot. Thus, the probability that an arbitrary device is scheduled in a time slot equals M/K, independent of its own state and the states of the other devices, and each device behaves stochastically independently. The age of a device will be initialized to 1 in each time slot with probability $\lambda_k p_k (M/K)$ and otherwise it increases by one. The mean age is $K/(\lambda_k p_k M)$ and the mean energy cost is $(M/K)\lambda_k E_k$. The total mean cost is then given by

$$\bar{c}^{\mathrm{Rnd}}(M) = \sum_{k=1}^{K} \frac{K}{\lambda_k p_k M} + w_k \frac{M}{K} \lambda_k E_k, \tag{10}$$

where we have highlighted explicitly the dependence of the policy on parameter M.

The Random policy can be easily optimized with respect to M. The unconstrained minimum cost (assuming M is continuous valued) is achieved by the value M^*, which equals

$$M^* = K \sqrt{\frac{\sum_k \frac{1}{\lambda_k p_k}}{\sum_k w_k \lambda_k E_k}}.$$

Since M is integer valued and constrained such that $M \in \{1, \ldots, K\}$, the final minimized cost is obtained as

$$\bar{c}^{\mathrm{OptRnd}} = \tag{11}$$
$$\begin{cases} \min\{\bar{c}^{\mathrm{Rnd}}(\lfloor M^* \rfloor), \bar{c}^{\mathrm{Rnd}}(\lceil M^* \rceil)\}, & \text{if } 1 \leq M^* \leq K, \\ \min\{\bar{c}^{\mathrm{Rnd}}(1), \bar{c}^{\mathrm{Rnd}}(K)\}, & \text{otherwise.} \end{cases}$$

The performance of the other polices depends on M, as well. We will explore this later in the numerical examples.

VI. Numerical examples

In this section, we compare the performance of the policies by simulating the system for different parameter values. For each simulation run, the system has been simulated for at least $2 \cdot 10^5$ time slots.

To simplify the parameterization of the simulations we consider a scenario, where there are two classes of devices with a given number of devices in each class, denoted by n_1 and n_2. Similarly, the other parameters are also given per class. Within each class the devices then have identical parameters.

A. Performance of different policies

First we consider a scenario with a small number of devices with $n_1 = 1$ and $n_2 = 2$ (i.e., $K = 3$ devices), where we only schedule one user in a time slot, i.e., $M = 1$. The classes have asymmetric parameters in order to highlight the differences in the performance. In Scenario 1, the parameters are selected such that energy costs will be relatively high compared with the age costs. For the class 1 device the following parameters are used: $p_1 = 0.5, w_1 = 1, E_1 = 50$. For class-2 devices the following parameters are used: $\lambda_2 = 0.5, p_2 = 0.4, w_2 = 5, E_2 = 100$. Clearly, for class-2 devices energy costs are significantly greater than for class 1. We then simulate the system for an increasing value of λ_1. In Scenario 2, all other parameters are kept the same, except we decrease the impact of the energy cost by one order of magnitude by changing the weights to $w_1 = 0.1$ and $w_2 = 0.5$. The results for the total mean cost in the system as a function of λ_1 for the different policies are shown in Figure 1, where upper panel corresponds to Scenario 1 and lower panel to Scenario 2. In the figures, the curves for different policies are labeled as follows: Random policy (Rnd), Max-age policy (Max), Myopic policy (Myo) and Energy-aware Whittle index policy (EW).

In Scenario 1 (upper panel), the Myopic policy performs very poorly with the mean cost exceeding 1000 across all values of λ_1 and therefore its results are not shown. Closer inspection of the results reveals that the Myopic policy gives a too great emphasis on the energy costs and as a result the

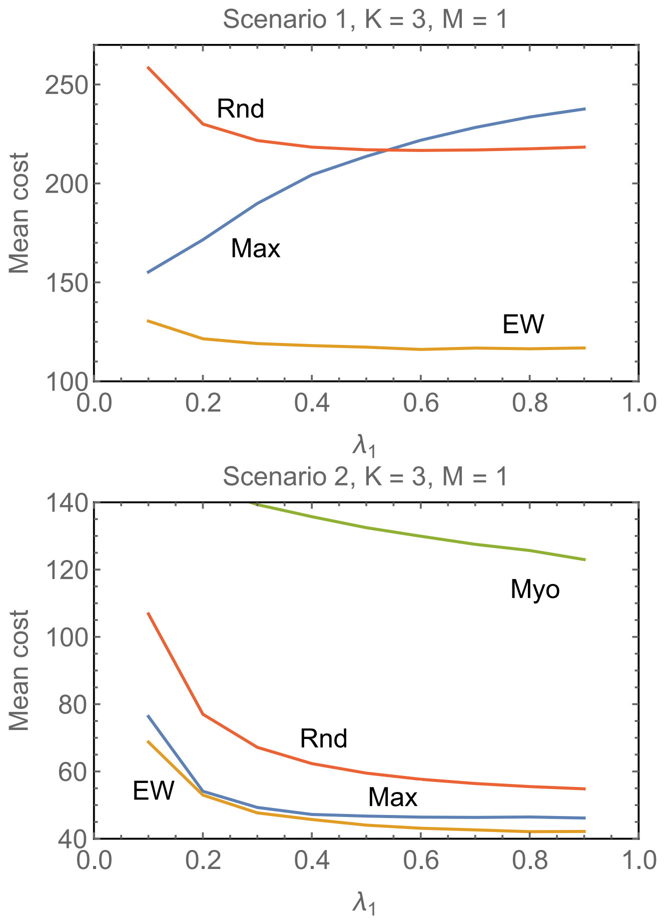

Fig. 1. Total mean cost in a small system with $K = 3$ devices and $M = 1$ for the different policies (Rnd = Random policy, Myo = Myopic policy, EW = Energy-aware Whittle index policy) as a function of λ_1 in Scenario 1 (upper panel) and Scenario 2 (lower panel).

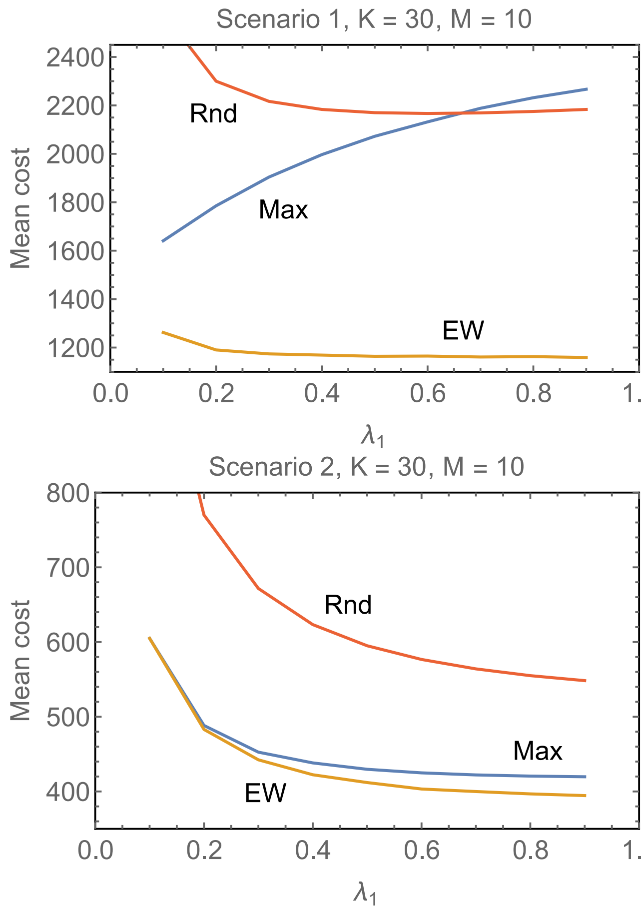

Fig. 2. Total mean cost in a large system with $K = 30$ devices and $M = 10$ for the different policies (Rnd = Random policy, Myo = Myopic policy, EW = Energy-aware Whittle index policy) as a function of λ_1 in Scenario 1 (upper panel) and Scenario 2 (lower panel).

age of class-2 devices are very high compared with the other policies. Random policy performs typically worse than Max-age and Energy-aware Whittle index policies and the total cost first decreases and then practically flattens or slowly increases as a function of λ_1. This behavior can be inferred from the expression of the total cost for the Random policy (10). In our case, as a function of λ_1 the age of class 1 device decreases but it also increases the energy consumption, while the costs from class 2 devices are unaffected. However, with our given parameter values the decrease in the age is first outweighing the increase in energy cost, until the increasing energy costs start to dominate. Max-age policy performs somewhat differently. By inspecting the detailed results from simulations, the reason is that the Max-age policy aims to equalize the ages of all devices across both classes. Therefore, with increasing λ_1 the age of class 1 device decreases, but it simultaneously also decreases the age of class-2 devices, which have an even greater energy cost than the class 1 device. As the energy costs anyway have a very high weight in this scenario, this causes the total costs of Max-age policy to increase as a function of λ_1 so that for small values of λ_1 it is performing better than Random policy but for larger values it performs even worse than Random policy. Notably, the Energy-aware Whittle index policy is clearly outperforming all other policies.

In Scenario 2, where impact of energy costs have been reduced significantly, costs of all policies are decreasing sharper as a function of λ_1, i.e., effects as was seen in Scenario 1 about the counter acting impacts of increasing energy costs are not seen here. Also in this case, Myopic policy still gives clearly the worst performance, but it is relatively better than in Scenario 1 due to the decreased weight of energy costs and it can be reasonably shown in the figure. Max-age policy is clearly performing now uniformly better than Random policy and at times even almost as good as the best performing Energy-aware Whittle index policy.

Next, we study a larger system and increase the number of devices to $n_1 = 10$ and $n_2 = 20$, i.e., there are $K = 30$ devices. Similarly, we increase the number of devices to be scheduled in a time slot to $M = 10$. The same two scenarios, Scenario 1 (high energy weight) and Scenario 2 (low energy weight), are considered as earlier. The results for the total mean cost in the system as a function of λ_1 for the different policies are shown in Figure 2, where upper panel corresponds to Scenario 1 and lower panel to Scenario 2. It can be observed that the results are very similar to the ones in Figure 1, only the scale is higher due to the larger number of devices. Due to the poor performance of Myopic policy, its results are not shown in the figure.

B. *Impact of optimizing M*

The number of devices to schedule M has a great impact on the achievable costs of the policies. As was seen earlier, with the Random policy the optimal M can be even analytically determined. However, for the other policies we need to study this through simulations.

We consider Scenario 1 with $K = 30$ devices with $\lambda_1 = 0.2$ and $\lambda_1 = 0.9$, see Figure 3. For the Energy-aware Whittle index policy the cost appears to monotonously decrease as a function of M. This is natural as the policy inherently decides, whether it is beneficial from the total cost point of view to schedule or not, i.e., if the Whittle index is positive or negative valued. Thus, the parameter M is a soft constraint in the system. The scheduler does not have to schedule exactly M devices, but it selects up to M, if it is cost effective at that time. From the system point of view this is a nice property since M can be set to a relatively high value and whatever out of M remains unutilized it can be allocated to other users in the system.

This is in contrast with all the other policies, which have a distinct value for M that minimizes the cost and any deviation from that will result in inferior performance. Moreover, the minimizing value depends on the parameters, as seen in Figure 3 with $\lambda_1 = 0.2$ (upper panel) and $\lambda_1 = 0.9$ (lower panel), which may be difficult to estimate. In these policies, there is no natural way for the scheduler to decide whether it is beneficial to schedule or not. Note that for the Random policy, the minimum point on the curve corresponds to the costs of the optimized Random policy given by (11).

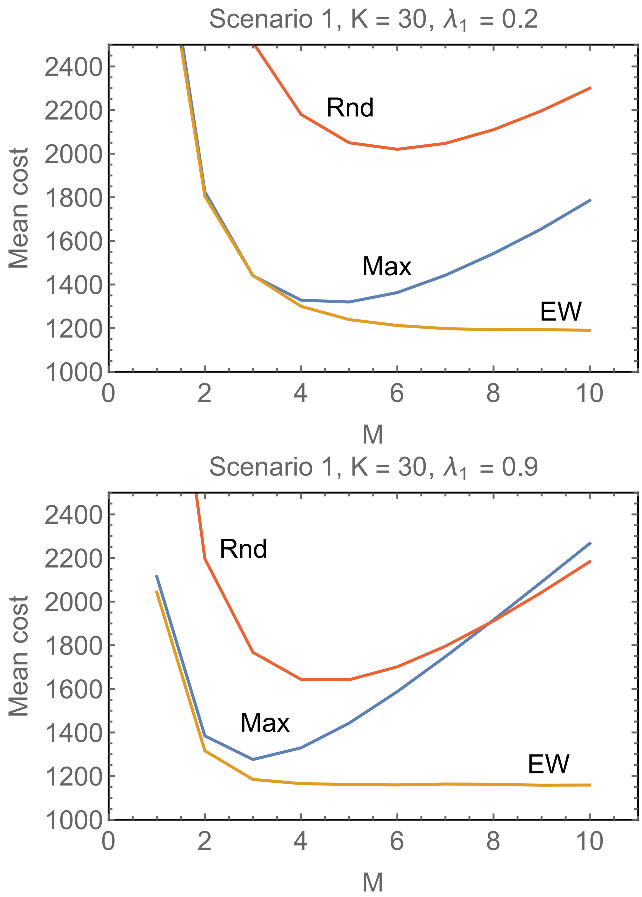

Fig. 3. Total mean cost in a large system with $K = 30$ devices and $M = 10$ for the different policies (Rnd = Random policy, Myo = Myopic policy, EW = Energy-aware Whittle index policy) as a function of $M \in \{1, \dots, 10\}$ for $\lambda_1 = 0.3$ (upper panel) $\lambda_1 = 0.9$ (lower panel) in Scenario 1.

VII. Conclusions

We considered the optimal age-energy tradeoff arising in the context uplink scheduling of multiple devices that are transmitting time-sensitive update packets through the base station in a cellular network. In the model, new updates arrive stochastically at the devices and they are subject to losses when attempting to transmit. Also, associated with the update transmission are related energy costs. As a measure of the information freshness at the base station, we consider the age of update packets. The problem is to develop policies for minimizing the weighted sum of the total mean age and energy costs when devices having heterogeneous parameters.

The problem is an instance of RMABs and we applied the Whittle index approach to obtain a near-optimal policy. We proved the indexability of the problem and explicitly solved the associated Whittle indices. In our numerical experiments, we demonstrated the superiority of the resulting Energy-aware Whittle index policy compared with the other heuristics (Random policy, Max-age policy, Myopic policy). In practical settings, multiple devices can be scheduled in a time slot and the Whittle index policy was illustrated to behave in a robust manner without a need to optimize precisely the number of devices to schedule, since the Whittle index value is also a direct measure of the usefulness of scheduling a device.

A possible extension of the work here is to consider the same model but assuming that the scheduler has exact knowledge, whether a fresh update is available or not, thus extending the state description from mere knowledge of the instantaneous age to also being aware of the realization of the Bernoulli random variable characterizing the generation of a new update. Additionally, in our model, the bufferless assumption at the device is done for tractability reasons. It is worth exploring how this assumption could be relaxed in the analysis, but it will lead to a two-dimensional process, which significantly complicates the analysis.

References

[1] S. Kaul, R. Yates, and M. Gruteser, “Real-time status: How often should one update?” in *IEEE INFOCOM*, 2012, pp. 2731–2735.

[2] R. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, “Age of information: An introduction and survey,” *ArXiv*, vol. abs/2007.08564, 2020.

[3] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 193–204, 2018.

[4] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies," *IEEE Transactions on Information Theory*, vol. 66, no. 1, pp. 534–556, 2020.

[5] Y. Gu, H. Chen, Y. Zhou, Y. Li, and B. Vucetic, "Timely status update in internet of things monitoring systems: An age-energy tradeoff," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5324–5335, 2019.

[6] P. Whittle, "Restless bandits: activity allocation in a changing world," *Journal of Applied Probability*, vol. 25A, pp. 287–298, 1988.

[7] P. Jacko, "Value of information in optimal flow-level scheduling of users with Markovian time-varying channels," *Performance Evaluation*, vol. 68, pp. 1022–1036, 2011.

[8] S. Aalto, P. Lassila, and P. Osti, "Whittle index approach to size-aware scheduling for time-varying channels with multiple states," *Queueing Systems*, vol. 83, pp. 195–225, 2016.

[9] K. Wang, J. Yu, L. Chen, P. Zhou, X. Ge, and M. Z. Win, "Opportunistic scheduling revisited using restless bandits: Indexability and index policy," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4997–5010, 2019.

[10] J. Niño-Mora, "Dynamic allocation indices for restless projects and queueing admission control: A polyhedral approach," *Mathematical Programming*, vol. 93, pp. 361–413, 2002.

[11] S. Aalto and P. Lassila, "Near-optimal dispatching policy for energy-aware server clusters," *Performance Evaluation*, vol. 135, 2019.

[12] M. Larrañaga, U. Ayesta, and I. Verloop, "Asymptotically optimal index policies for an abandonment queue with convex holding cost," *Queueing Systems*, vol. 81, pp. 99–169, 2015.

[13] I. M. Verloop, "Asymptotically optimal priority policies for indexable and nonindexable restless bandits," *Annals of Applied Probability*, vol. 26, no. 4, pp. 1947–1995, 08 2016.

[14] Y. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 2634–2638.

[15] Y. P. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Transactions on Mobile Computing*, vol. 19, no. 12, pp. 2903–2915, 2020.

[16] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2637–2650, 2018.

[17] Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Can decentralized status update achieve universally near-optimal age-of-information in wireless multiaccess channels?" in *30th International Teletraffic Congress (ITC 30)*, vol. 01, 2018, pp. 144–152.

[18] J. Sun, Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Closed-form Whittle's index-enabled random access for timely status update," *IEEE Transactions on Communications*, vol. 68, pp. 1538–1551, 2020.

[19] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2005.