Optimal Sampling for State Change Detection with Application to the Control of Sleep Mode

A. P. Azad*, S. Alouf*, E. Altman*, V. Borkar† and G. Paschos+
* Maestro group, INRIA, 2004 Route des Lucioles, F-06902 Sophia Antipolis
†School of Technology, TIFR, Mumbai
+ CERTH, Thessaloniki, Greece
{aazad,salouf,altman}@sophia.inria.fr, borkar@tifr.res.in, gpasxos@ee.upatras.gr

***Abstract*— This work considers systems with inactivity periods that have an unknown duration. We study the question of scheduling "waking up" instants in which a server can check whether the inactivity period is over. There is a cost proportional to the delay from the moment the inactivity period ends until the server discovers it, a (small) running cost while the server is away and also a cost for waking up. As an application to the problem, we consider the energy management in WiMax where inactive mobiles reduce their energy consumption by entering a sleep mode. Various standards exist which impose specific waking-up scheduling policies at wireless devices. We check these and identify optimal policies under various statistical assumptions. We show that periodic fixed vacation durations are optimal and derive the optimal period. We show that this structure does not hold for other inactivity distributions but manage to obtain some suboptimal solutions which perform strictly better than the periodic ones. We finally obtain structural properties for optimal policies for the case of arbitrary distribution of inactivity periods.**

I. Introduction

Mobile terminals using contemporary radios can benefit greatly by shutting off the transceiver whenever there is no scheduled activity. Nevertheless, if the attention of the mobile is suddenly required, the mobile will be shut off and therefore unavailable. The longer the shut off (vacation) periods, the longer the expected response delay. Therefore, one can identify the inherent tradeoff of energy management; increase vacation length to improve energy saving or decrease vacation length to reduce delays.

Past approaches have considered incoming/outgoing traffic, the effect of setup time, or even the queueing implications in the analysis. Concerning the arrival process, it has been assumed to be Poisson, having a hyper-Erlang distribution or a hyper-exponential distribution. In all cases, it does not depend on the energy management scheme. As for delay, it is the average packet delay in the system that is considered.

Our work departs from the existing models in two aspects. First, rather than an exogeneous independent arrival process, we have in mind elastic arrival processes in which (i) a "think time" or "off period" begins when the activity of the server ends, and (ii) the duration of the "on period" does not depend on the wake up delay, defined as the time that elapses between the instant a request is issued and the instant at which the request service actually begins. Both assumptions are appropriate to interactive applications such as web browsing. As a result, the measure for delay is taken to be the wake up delay. We shall be using a hyper-exponential distribution for the off period. The motivation comes from works that provide evidence of heavy-tailed off period distributions on the Internet and on the World Wide Web with a Pareto type distribution in [1], [2]. It is well-known that heavy-tailed distributed random variables (rvs) can be well approximated by hyper-exponential distributions [3], [4]. Once the performance of the proposed standard mechanism is known, one is interested to optimize the degrees of freedom available (customizable variables) in order to achieve the desired balance between delay and energy saving. Recent work in the literature focuses on heuristic adaptive algorithms, see [5], [6], [7].

We shall thus investigate in this paper optimal energy management system under one of the following assumptions on the off time distribution: (i) *Exponential distribution*; (ii) *Hyper-exponential distribution*; (iii) *General distribution*. Our contributions are as follows: (1) Our problem formulation allows us to obtain optimal policies which minimize the delay cost as well as energy cost at the same time. We use DP approach which allows to obtain optimal vacation size at each wake up instant. (2) For exponential off times, we show that equal sized vacation policies are optimal and we derive it. (3) For hyper-exponential policies we derive interesting structural properties. We show that optimal policy are bounded. Asymptotically, the optimal policy converges to fixed policy corresponding to the smallest parameter irrespective of the initial state. These policies can be computed numerically using value iteration. (4) For any general off time distribution, we show that optimal polices are bounded. (5) We propose suboptimal policies using policy iteration which perform strictly better than optimal policies within a class of vacation and are simpler to compute. We show numerically the performance of such suboptimal solutions using one stage and two stage policy iteration. (6) We compare the proposed policy with that of standard protocol under various statistical assumptions.

In the rest of the paper, Sect. II outlines our system model and introduces the cost function. Section III investigates different strategies that could be considered in the power save mode and derives the corresponding optimal control. Section IV presents relevant characteristic of the optimal policy whereas Sect. V tackles the problem of finding the optimal policy under the worst case process of arrivals. Numerical results and a comparative study of the different optimal strategies and of the IEEE 802.16e standard are reported in Sect. VI, before concluding the paper in Sect. VII.

II. System Model

We consider a system with repeated vacations. As long as there are no customers, the server goes on vacation. We are interested in finding the optimal policy, so that at any start of vacation, the length of this vacation is optimal.

This system models a mobile device that turns off its radio antenna while inactive to save energy. A vacation is then the time during which the mobile is sleeping. At the end of a vacation, the mobile needs to turn on the radio to check for packets (customers).

Let X denote the number of vacations in an idle period. X is a discrete random variable (rv) taking values in $\mathbb{N}^*$. Let τ denote the time length between the start of the first vacation and the arrival of a customer; this time is referred to as the "off time". τ is a rv whose probability density function is $f_\tau(t), t \geq 0$.

The duration of the kth vacation is a rv denoted B_k, for $k \in \mathbb{N}^*$. For analytical tractability, we consider vacations $\{B_k\}_{k\in\mathbb{N}^*}$ that are mutually independent rvs. The time at the end of the kth sleep interval is a rv denoted T_k, for $k \in \mathbb{N}^*$. We denote T_0 as the time at the beginning of the first vacation; by convention $T_0 = 0$. We naturally have $T_k = T_{k-1} + B_k = \sum_{i=1}^k B_i$.

We will use the following notation $\mathcal{Y}^*(s) := \mathbb{E}[\exp(-sY)]$ to denote the Laplace-Stieltjes transform of a generic rv Y evaluated at s. Hence, we can readily write $\mathcal{T}_k^*(s) = \prod_{i=1}^k \mathcal{B}_i^*(s)$. Observe that the time at the end of a generic idle period is simply T_X. Hence, since the arrival time of the first customer during the idle period is τ, its service will be delayed for $T_X - \tau$ units of time.

The energy consumed by a mobile while *listening* to the channel and checking for customers is denoted E_L. This is actually a penalty paid at the end of each vacation. The *power* consumed by a mobile in a sleep state is denoted E_S. The energy consumed by a mobile during vacation B_k is then equal to $E_L + E_S B_k$, and that consumed during a generic idle period is equal to $E_L X + E_S T_X$.

We are interested in minimizing the cost of the power save mode, which is seen as a weighted sum of the energy consumed during the power save mode and the *extra* delay incurred on the traffic by a sleeping mobile. Let V be this cost, which is written as follows

$$V := \bar{\epsilon}\,\mathbb{E}[T_X - \tau] + \epsilon\,(E_L \mathbb{E}[X] + E_S \mathbb{E}[T_X]) \qquad (1)$$

where ϵ is a *normalized weight* that takes value between 0 and 1. Let $\bar{\epsilon} = 1 - \epsilon$.

The following can be easily derived

$$\begin{aligned}\mathbb{E}[T_X - \tau] &= \mathbb{E}_X[\mathbb{E}[T_X - \tau | X]] \\ &= \sum_{k=1}^{\infty} \mathbb{E}[(T_k - \tau)\mathbb{1}\{T_{k-1} < \tau \leq T_k\}];\end{aligned}$$

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} kP(X = k), \quad \mathbb{E}[T_X] = \sum_{k=1}^{\infty} \mathbb{E}[T_k]P(X = k);$$

where, for $k \in \mathbb{N}^*$,

$$P(X = k) = P(T_{k-1} < \tau \leq T_k). \qquad (2)$$

The cost can then be rewritten

$$\begin{aligned}V &= \sum_{k=1}^{\infty} \Big\{ \bar{\epsilon}\,\mathbb{E}[(T_k - \tau)\mathbb{1}\{T_{k-1} < \tau \leq T_k\}] \\ &\qquad + \epsilon P(X = k)\,(E_L k + E_S \mathbb{E}[T_k]) \Big\}. \qquad (3)\end{aligned}$$

We first observe that

$$\begin{aligned}&\mathbb{E}[(T_k - \tau)\mathbb{1}\{T_{k-1} < \tau \leq T_k\}] \\ &\quad = \mathbb{E}[(T_k - \tau)\mathbb{1}\{\tau \leq T_k\}\,|\tau > T_{k-1}]P(\tau > T_{k-1}).\end{aligned}$$

Hence, and using (2), one can rewrite (3) as follows

$$\begin{aligned}V &= \bar{\epsilon} \sum_{k=1}^{\infty} P(\tau > T_{k-1})\mathbb{E}[(T_k - \tau)\mathbb{1}\{\tau \leq T_k\}\,|\tau > T_{k-1}] \\ &\quad + \epsilon \sum_{k=1}^{\infty} P(T_{k-1} < \tau \leq T_k) \sum_{i=1}^{k} (E_L + E_S \mathbb{E}[B_i])\,.\end{aligned}$$

Rearranging the terms in the summations, we obtain

$$\begin{aligned}V &= \sum_{k=0}^{\infty} P(\tau > T_k) \Big\{ \bar{\epsilon}\,\mathbb{E}[(T_{k+1} - \tau)\mathbb{1}\{\tau \leq T_{k+1}\}\,|\tau > T_k] \\ &\quad + \epsilon\,(E_L + E_S \mathbb{E}[B_{k+1}]) \Big\}. \qquad (4)\end{aligned}$$

The term between braces is actually the portion of the cost that is due to the $k+1$st vacation solely. The first part corresponds to the cost due to extra delay and the second part corresponds to the cost due to energy consumption.

We will next formulate a parametric optimization problem using (4), which allows us to obtain optimal vacation sizes for given class of vacations. The classes that we have considered are: (i) exponentially distributed vacations; (ii) equally sized vacations (periodic pattern); (iii) general vacations that follow a scaled version of a known distribution; and (iv) general discrete vacations. Due to space limitation, we will present solely the case of exponentially distributed vacations as its results will be needed subsequently in Sect. IV. The results for the other cases can be found in [8].

III. Parametric Optimization

In this section, we focus on minimizing (3). We assume that τ is hyper-exponentially distributed with n phases and parameters $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and $\boldsymbol{q} = (q_1, \ldots, q_n)$. In other words, we have

$$f_\tau(t) = \sum_{i=1}^{n} q_i \lambda_i \exp(-\lambda_i t), \quad \sum_{i=1}^{n} q_i = 1. \qquad (5)$$

Recall that τ represents the time that elapses since the beginning of the save mode (i.e., beginning of the first vacation) until the arrival of the first packet (i.e., customer). Therefore, τ is the conditional residual inter-arrival time. Observe that when $n = 1$, τ will be exponentially distributed with parameter $\lambda = \lambda_1$. This case is equivalent to having a Poisson arrival process with rate λ, thanks to the memoryless property.

We will now compute the elements of (3). The distribution of X is given as

$$P(X=k) = P(\tau > T_{k-1}) - P(\tau > T_k) = \sum_{i=1}^{n} q_i \mathcal{T}_{k-1}^*(\lambda_i)\left(1-\mathcal{B}_k^*(\lambda_i)\right).$$

Also, we can compute

$$\mathbb{E}[(T_k-\tau)\mathbb{1}\{T_{k-1}<\tau\le T_k\}] = \sum_{i=1}^{n} q_i \mathcal{T}_{k-1}^*(\lambda_i)\left(\mathbb{E}[B_k] - \frac{1-\mathcal{B}_k^*(\lambda_i)}{\lambda_i}\right).$$

After some calculus, the cost simplifies to (let $\eta := \bar\epsilon + \epsilon E_S$)

$$V = -\bar\epsilon\,\mathbb{E}[\tau] + \sum_{k=0}^{\infty}\sum_{i=1}^{n} q_i \mathcal{T}_k^*(\lambda_i)\left(\epsilon E_L + \eta\mathbb{E}[B_{k+1}]\right), \quad (6)$$

where $\mathbb{E}[\tau] = \sum_{i=1}^{n} q_i/\lambda_i$ is the expectation of τ.

Assuming now that all vacations are identically distributed, i.e., the control is static, and letting B be a generic rv having same distribution as any of the vacations, then (6) can be rewritten

$$V = -\bar\epsilon\,\mathbb{E}[\tau] + (\epsilon E_L + \eta\mathbb{E}[B])\sum_{i=1}^{n}\frac{q_i}{1-\mathcal{B}^*(\lambda_i)}. \quad (7)$$

If B is exponentially distributed with mean b, i.e., $\mathcal{B}^*(s) = 1/(1+bs)$. The performance is optimized by controlling the mean vacation size b. The cost, denoted by $V_e(b)$ with the subscript standing for "exponential", simplifies to

$$V_e(b) = -\bar\epsilon\,\mathbb{E}[\tau] + (\epsilon E_L + \eta b)\sum_{i=1}^{n} q_i\left(1+\frac{1}{b\lambda_i}\right), = \epsilon\left(E_S + \frac{E_L}{b}\right)\mathbb{E}[\tau] + (\epsilon E_L + \eta b). \quad (8)$$

Proposition 3.1: The cost $V_e(b)$ is a convex function having a minimum at

$$b_e^\star = \sqrt{\frac{\epsilon E_L \mathbb{E}[\tau]}{\eta}} = \sqrt{\frac{\epsilon E_L\mathbb{E}[\tau]}{\bar\epsilon + \epsilon E_S}}. \quad (9)$$

The minimal cost is

$$V_e(b_e^\star) = \epsilon(E_S\mathbb{E}[\tau] + E_L) + 2\sqrt{\epsilon\eta E_L\mathbb{E}[\tau]} \quad (10)$$

Proof: Let us compute the first and second derivative of the cost. We find

$$V_e'(b) = \eta - \frac{\epsilon E_L\mathbb{E}[\tau]}{b^2}; \quad V_e''(b) = 2\frac{\epsilon E_L\mathbb{E}[\tau]}{b^3}.$$

Clearly, $V_e''(b) \ge 0$ for any $b > 0$, hence $V_e(b)$ is a convex function. The derivative $V_e'(b)$ has a root at $b_e^\star$ as given in (9), which yields a minimum in the cost $V_e(b)$ at $b_e^\star$. Substituting the optimal $b_e^\star$ in (8) we obtain the minimal cost (10). $\diamondsuit$

The optimal expected vacation duration and the minimal cost are obtained in closed-form.

Remark 3.1: The optimal mean vacation (9) and minimal cost (10) stand valid for any distribution of off times τ for a known mean of off time $\mathbb{E}[\tau]$.

This derives directly from (3). Observe in (3) that the term $\mathbb{E}[(T_k-\tau)\mathbb{1}\{T_{k-1}<\tau\le T_k\}]$ is exponentially distributed due to memoryless property of vacations. Hence it is valid for any distribution of off times τ for a known $\mathbb{E}[\tau]$.

IV. Dynamic Programming

Dynamic programming (DP) is a well-known tool which allows to compute the optimal decision policy to be taken at each intermediate observation point, taking into account the whole lifetime of the system. Considering our system model, we want to identify the optimal sleep strategy where decisions are taken at each intermediate wake-up instance. Hence, a DP approach is a natural candidate for determining the optimal policy.

We introduce the following DP

$$V_k^\star(t_k) = \min_{b_{k+1}\ge 0}\Big\{\mathbb{E}[c(t_k, b_{k+1})] + P(\tau_{t_k} > b_{k+1})V_{k+1}^\star(t_{k+1})\Big\}. \quad (11)$$

Here, $V_k^*(t_k)$ represents the optimal cost at time t_k where the argument t_k denotes the state of the system at time t_k. The term $P(\tau_{t_k} > b_{k+1})$ represents the transition probability at t_k. Term $c(t_k, b_{k+1})$ denotes the cost to go from stage t_k when the control(vacation) taken is b_{k+1}. In generic notation the per stage cost is represented as

$$c(t,b) = \bar\epsilon\,\mathbb{E}[(b-\tau_t)\mathbb{1}\{\tau_t \le b\}] + \epsilon(E_L + E_S b) \quad (12)$$

Hence we can see that given a terminal cost $V_N^*(t_N)$ one can obtain the total optimal cost recursively using DP equation (11). We can see that in the DP problem, each stage is characterized by the distribution of the residual off time. The state of the system in sleep mode can then by described by the distribution of τ_t.

Next, we derive some interesting properties of the off times' hyper-exponential distribution. In the rest of this section, three cases will be considered following the distribution of the off time. We start with DP solutions for hyper-exponential distribution of off time with n phases and derive some structural property, then present the results found under the assumption of a Poisson arrival process (exponential off time). Last, the case of general off times is considered. We then discuss suboptimal solutions through DP.

A. Hyper-Exponential Off Time

Before we derive some properties of hyper-exponential off time, recall the its definition from (5).

1) Distribution of the Residual Off Time: Recall that τ_t denoted the residual time of τ at time t, given that $\tau > t$. We compute its tail as follows

$$P(\tau_t > a) = P(\tau > t+a \,|\, \tau > t) = \frac{P(\tau > t+a)}{P(\tau > t)} = \sum_{i=1}^{n} g_i(q,t)\exp(-\lambda_i a) \quad (13)$$

where

$$g_i(q,t) := \frac{q_i\exp(-\lambda_i t)}{\sum_{j=1}^{n} q_j\exp(-\lambda_j t)}, \quad i = 1,\ldots,n. \quad (14)$$

We denote $g(\boldsymbol{q}, t)$ as the n-tuple of functions $g_i(q,t)$, $i = 1,\ldots,n$. Observe that $g(\boldsymbol{q}, 0) = \boldsymbol{q}$. The function g transforms the distribution $\boldsymbol{q}$ into another distribution $\boldsymbol{q}'$ such that $\sum_{j=1}^{n} q_j' = 1$ and $q_j' > 0$.

Equation (13) is nothing but the tail of a hyper-exponential rv having n phases and parameters $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and $g(\boldsymbol{q}, t)$. Except for the probabilities of the n phases, the off time τ and its residual time τ_t have the same distribution and same parameter $\boldsymbol{\lambda}$. As time goes on, the residual time keeps its distribution but updates its probabilities, through the transform function g. It can be shown that

$$g_i(q, b_1 + b_2) = g_i\Big(g_i(q, b_1), b_2\Big). \tag{15}$$

In other words, the function g is such that the result of the transformation after $b_1 + b_2$ units of time is the same as that of a first transformation after b_1 units of time, followed by a second transformation after b_2 units of time.

To simplify the notation, we will drop the subscript of the residual off time τ_t, and instead, we will add as argument the current probability distribution (which is transformed over time through the function g). For instance, if at some point in time, the residual off time has the probability distribution $\boldsymbol{q}'$, then we will use the notation $\tau(\boldsymbol{q}')$.

The results above can be extended to account for a random passed time T. We have

$$P(\tau > T + a \,|\, \tau > T) \;\; = \;\; \sum_{i=1}^{n} g_i(q, T) \exp(-\lambda_i a)$$

where

$$g_i(q, T) := \frac{q_i \mathcal{T}^*(\lambda_i)}{\sum_{j=1}^{n} q_j \mathcal{T}^*(\lambda_j)} = \frac{q_i \mathcal{T}^*(\lambda_i)}{P(\tau > T)}. \tag{16}$$

There is an abuse of notation in the definition of $g_i(q, T)$, as this function depends on the *distribution* of T and not on the rv T itself. The function $g_i(q, T)$ is *not* a rv. Observe that (14) is a particular case of (16) where time is deterministic. Equations (14) and (16) will be referred to as "update rules". Below we briefly discuss property induced from the asymptotic property $g(.)$.

Define the composition $g^m(q, T) = g(g^{m-1}(q, T))$ where $g^1(q, T)$ is the vector whose ith element is given in (16). Assume without loss of generality that $\lambda_1 \le \ldots \le \lambda_n$. Let $e(i)$ be the n-dimensional vector whose ith element is 1 and all other elements are zero.

Lemma 4.1: Fix q and let $I(q)$ be the smallest j for which $q_j > 0$. The following limit holds:

$$\lim_{m \to \infty} g^m(q, T) = e(I(q)).$$

Proof: Let $\alpha(i) := \frac{\mathcal{T}^*(\lambda_i)}{\mathcal{T}^*(\lambda_{I(q)})}$. Then

$$g_i(q, T) = \frac{q_i \mathcal{T}^*(\lambda_i)}{\sum_{j=I(q)}^{n} q_j \mathcal{T}^*(\lambda_j)} = \frac{q_i \alpha_i}{\sum_{j=I(q)}^{n} q_j \alpha_j}$$

$$g_i^{m+1}(q, T) = \frac{q_i \alpha_i^m}{\sum_{j=I(q)}^{n} q_j \alpha_j^m}$$

this implies the lemma. ♢

Assume that the set of available actions is bounded below by $\underline{b} > 0$. Then for all $i > I(q)$ we have $\alpha_i \le \exp(-\underline{b}\Delta)$ where $\Delta := \lambda_{I(q)+1} - \lambda_{I(q)}$. Thus

$$\sum_{i > I(q)} g_i^{m+1}(q, T) \le \Delta^m (1 - q_{I(q)}).$$

Lemma 4.2: For any q we have

$$\lim_{q' \to q} V(q') = V(q)$$

Proof: Refer to [8] for a detailed proof. ♢

Lemma 4.1 states that the off time distribution translates its mass towards the phase with the smallest λ, and converges asymptotically irrespective of the initial distribution. This suggests that there exists a threshold after which the optimal policy is the one that corresponds to the optimal policy for state $I(q)$. Lemma 4.2 states that as the state converges, the value also converges to the value at the converged state.

B. DP Solution for Hyper-Exponential Off Time

Below we formulate the optimization problem as an MDP where the state space is taken to be the simplex of dimension n, i.e. the set of probability measures over the set $\{1, 2, ..., n\}$. At each stage, the residual off time sees its probability distribution being updated. Let $\boldsymbol{q}^0$ denote the probability distribution of the *total* off time. It is then the probability distribution of the residual off time at time 0. Thanks to the property (15), the probability distribution of the residual off time at stage $k+1$, i.e., at time t_k, is $\boldsymbol{q} = g(\boldsymbol{q}^0, t_k)$. Henceforth, there is a one to one relation between the stage and the current probability distribution of the residual off time. Without loss of optimality, either of them can be the state in the MDP [9, Sect. 5.4].

The system state is denoted $\boldsymbol{q}$ and represents the *current* probability distribution of the residual off time. The initial state is $\boldsymbol{q}^0$. We assume that the controller can choose any time b (a constant or a rv) until he wakes up. The transition probabilities are simply

$$P_{\boldsymbol{q}, b, \boldsymbol{q}'} = \mathbb{1}\{\boldsymbol{q}' = g(\boldsymbol{q}, b)\}\,.$$

We are faced with an MDP with a Borel action space and a state space where the state space is the set of probability vectors $\boldsymbol{q}$. Note however that starting from a given $\boldsymbol{q}$, there is a countable set Q of $\boldsymbol{q}$'s so that only states within Q can be reached from $\boldsymbol{q}$. Therefore we may restrict the state space to the countable set Q. We can again use [10] to conclude that we may restrict to policies that choose at each state a non-randomized decision b, and the decision depends only on the current state (and need not depend on the previous history).

We next show that there is some $\bar{b}$ such that we may restrict all actions to the compact interval $[0, \bar{b}]$ without loss of optimality.

Consider the policy w that takes always a constant vacation of length of one unit. It is easily seen that the total expected cost when using this policy is bounded by

$$\bar{\epsilon} + \epsilon \left(1 + \sup_i \frac{1}{\lambda_i}\right)(E_L + E_S)$$

Here the term $\bar{\epsilon}$ is an upper bound on the expected waiting cost; the term $(1 + \sup_i \frac{1}{\lambda_i})$ is an upperbound on both the expected number of vacations till detection of the end of the off period as well as on the expected time till that instant.

We conclude that

Lemma 4.3: (i) For all $\boldsymbol{q}$, $V(\boldsymbol{q}) \le \bar{b}$ where $\bar{b} = \bar{\epsilon} + \epsilon(1 + \sup_i \frac{1}{\lambda_i})(E_L + E_S)$.

(ii) Without loss of optimality, one may restrict to policies that take only actions within $[0, \widetilde{b}]$ where

$$\widetilde{b} = \frac{1}{\bar{\epsilon}}\{\bar{b} + 1 + 1/(\min_i \lambda_i)\}$$

To prove the second part of the Lemma, let u be an ϵ-optimal Markov policy that does not use randomization, where $\epsilon \in (0,1)$. If $u_i > \widetilde{b}$ for some i then the expected immediate cost at step i is itself larger than the total expected cost that would be incurred under the policy w plus 1:

$$\mathbb{E}\Big[\Big(b - \tau(\boldsymbol{q})\Big)\mathbb{1}\{\tau(\boldsymbol{q}) \le b\}\Big] > \bar{b} + 1.$$

Thus by switching from time i onwards to w, the expected cost strictly decreases by at least 1 unit so that u cannot be ϵ-optimal.

We conclude that the MDP can be viewed as one with a countable state space, compact action space, discrete time, and non-negative costs (known as "negative dynamic programming"). Using [11] we then conclude:

(i) The optimal value (minimal cost) is given by the minimal solution of the following DP:

$$V(\boldsymbol{q}) = \min_{b\ge 0}\Big\{\bar{\epsilon}\mathbb{E}\Big[\Big(b - \tau(\boldsymbol{q})\Big)\mathbb{1}\{\tau(\boldsymbol{q}) \le b\}\Big] + \epsilon(E_L + bE_S) + P\Big(\tau(\boldsymbol{q}) > b\Big)V\Big(g(\boldsymbol{q}, b)\Big)\Big\}. \quad (17)$$

(ii) Let $B(\boldsymbol{q})$ denote the set of all b's that minimize the right hand side of (17) for a given $\boldsymbol{q}$. Then any policy that chooses at state $\boldsymbol{q}$ some $b \in B(\boldsymbol{q})$ is optimal.

The value iteration can be used as an iterative method to compute $V(\boldsymbol{q})$. Starting with $V_0 = 0$ we write

$$V_{k+1}(\boldsymbol{q}) = \min_{b\ge 0}\Big\{\bar{\epsilon}\mathbb{E}\Big[\Big(b - \tau(\boldsymbol{q})\Big)\mathbb{1}\{\tau(\boldsymbol{q}) \le b\}\Big] + \epsilon(E_L + bE_S) + P\Big(\tau(\boldsymbol{q}) > b\Big)V_k\Big(g(\boldsymbol{q}, b)\Big)\Big\}.$$

Then $V(\boldsymbol{q}) = \lim_{k\to\infty} V_k(\boldsymbol{q})$, see [9]. This iteration is to be performed for every possible state $\boldsymbol{q}$. Notice that $g(\boldsymbol{q}, b)$ is the moving state converges asymptotically to $e(I(q))$ (4.1). To complete the value iteration, we need to compute the terms in the above DP equations. We elaborate on the equality found in (6), and write, for a fixed b

$$\mathbb{E}\Big[\Big(b - \tau(\boldsymbol{q})\Big)\mathbb{1}\{\tau(\boldsymbol{q}) \le b\}\Big] = b - \sum_{i=1}^{n} q_i \frac{1 - \exp(-\lambda_i b)}{\lambda_i}.$$

1) Exponential Off Time : When arrivals form a Poisson process with rate λ, the off time τ will be exponentially distributed with parameter λ which is nothing but a special case of hyper-exponential off time by fixing $n = 1$. It is well known that the residual time τ_t will have the same distribution as τ whatever t is. From (14) we can see, $g(q, t) = q$ for any t . This distribution is characterized solely by the rate λ. In other words, as time goes on, the state of the system is always represented by the parameter λ. Henceforth, the DP involves a single state, denoted λ.

We are faced with a Markov Decision Process (MDP), a single state λ, a Borel action space (the positive real numbers) and discrete time. Note that the sleep durations are not discrete. However, decisions are taken at discrete embedded times: the kth decision is taken at the end of the $k-1$st vacation. Therefore, we are dealing with a discrete time MDP. This is so called "negative" dynamic programming [11]. It follows from [10] that we can restrict to stationary policies (that depend only on the state) and that do not require randomization. Since there is only one state (at which decisions are taken) this implies that one can restrict to vacation sizes that have fixed size and that are the same each time a decision has to be taken. In other words, the optimal sleep policy is the constant one.

Hence the optimal value is given by the minimization of the following MDP :

$$V^\star(\lambda) = \min_{b\ge 0}\Big\{\bar{\epsilon}\mathbb{E}\Big[\Big(b - \tau(q)\Big)\mathbb{1}\{\tau(q) \le b\}\Big] + \epsilon(E_L + bE_S) + P\Big(\tau(q) > b\Big)V^\star\Big(\lambda\Big)\Big\}. \quad (18)$$

Proposition 4.1: The optimal vacation size for exponential off time is given by a

$$b_c^\star = -\frac{1}{\lambda}\left(\zeta + W_{-1}(-e^{-\zeta})\right) \quad (19)$$

$$\text{with } \zeta := \frac{\lambda\epsilon E_L}{\eta} + 1,$$

where $W_{-1}(-e^{-\zeta})$ denotes the branch of the Lambert W function[1] that is real-valued on the interval $[-\exp(-1), 0]$ and always below -1. The minimal cost is

$$V^\star(\lambda) = -\frac{1}{\lambda}\left(\bar{\epsilon} + \eta W_{-1}(-e^{-\zeta})\right). \quad (20)$$

Proof: From (18) we can express

$$V(\lambda) = \frac{\bar{\epsilon}\mathbb{E}\Big[\Big(b - \tau(q)\Big)\mathbb{1}\{\tau(q) \le b\}\Big] + \epsilon(E_L + bE_S)}{1 - P\Big(\tau(q) > b\Big)} \quad (21)$$

Substituting $\mathbb{E}\Big[\Big(b - \tau(q)\Big)\mathbb{1}\{\tau(q) \le b\}\Big] = \frac{\lambda b - 1 + \exp(-\lambda b)}{\lambda}$ and $P\Big(\tau(q) > b\Big) = \exp(-\lambda b)$ in the above equation and differentiating w.r.t. b we obtain

$$V'(b) = \eta\left\{\frac{1 - \exp(-\lambda b)(\zeta + \lambda b)}{(1 - \exp(-\lambda b))^2}\right\}.$$

Here we change the argument to b because we are trying to find b^*. The extremum of $V(b)$, denoted $b_c^\star$, must verify $V'(b_c^\star) = 0$. In other words, we must have

$$\begin{aligned} 1 - \exp(-\lambda b_c^\star)(\zeta + \lambda b_c^\star) &= 0 \\ \Leftrightarrow \quad \exp(-\zeta - \lambda b_c^\star)(-\zeta - \lambda b_c^\star) &= -\exp(-\zeta). \end{aligned}$$

The last expression is of the form $y\exp(y) = x$ with $y = -\zeta - \lambda b_c^\star$ and $x = -\exp(-\zeta)$. The solution y is the Lambert W function [12], denoted W, at the point x. Hence,

$$-\zeta - \lambda b_c^\star = W(-\exp(-\zeta)).$$

Since $\zeta \ge 1$, we have $-\exp(-1) \le -\exp(-\zeta) < 0$. Therefore, we need $W(-\exp(-\zeta))$ to be real-valued in $[-\exp(-1), 0[$. Also, given that $\zeta + \lambda b_c^\star \ge 1$, we need

[1]The Lambert W function, satisfies $W(x)\exp(W(x)) = x$. As $y\exp(y) = x$ has an infinite number of solutions y for each (non-zero) value of x, the function $W(x)$ has an infinite number of branches.

$W(-\exp(-\zeta))$ to be always negative and smaller than -1. Both conditions are satisfied by the branch numbered -1. Hence, $-\zeta - \lambda b_c^\star = W_{-1}(-\exp(-\zeta))$ and (19) is readily found. Replacing (19) in (21) and appropriate substitution and using the relation $\exp(y) = x/y$, one can derive (20).

Similarly we proceed for the second order conditions to determine if $b_c^\star$ yields minimum cost. This concludes the proof. Refer to [8] for a detailed proof. ♢

C. General Distribution of Off Time

Unlike what was done earlier, we no longer consider that the system state can be represented by the parameters of the distribution of the residual off time. Indeed, when the off time has a general distribution, one cannot expect any more that the residual off time will keep the same distribution over time, updating only its parameters, as was the case for the hyper-exponential distribution.

We therefore consider henceforth as system state the instant t at which a vacation is to start. We use again τ_t to denote the residual value of τ at time t (i.e. $\tau - t$) given that it is larger than t.

As a state space, we consider the set of non-negative real numbers. An action b is the duration of the next vacation. We shall assume that b can take value in a finite set. The set of t reachable (with positive probability) by some policy is countable. We can thus assume without loss of generality that the state space is discrete. Then the following holds:

Proposition 4.2:
(i) There exists an optimal deterministic stationary policy.
(ii) Let $V^0 := 0$, $V^{k+1} := \mathcal{L}V^k$, where

$$\mathcal{L}V(t) := \min_b \{c(t,b) + P(\tau_t > b)V(t+b)\}$$

where $c(t,b)$ has been defined in (12). Then V^k converges monotonically to the optimal value $V^\star$.
(iii) $V^\star$ is the smallest nonnegative solution of $V^\star = \mathcal{L}V^\star$. A stationary policy that chooses at state t an action that achieves the minimum of $\mathcal{L}V^\star$ is optimal.

Proof: (i) follows from [11, Thm 7.3.6], and (ii) from [11, Thm 7.3.10]. Part (iii) is due to [11, Thm 7.3.3]. ♢

Note that V^k expresses optimal cost for the problem of minimizing the total cost over a horizon of k steps.

Proposition 4.3: Assume that τ_t converges in distribution to some limit $\widehat{\tau}$. Define

Then $$v(b) := \frac{\widehat{c}(b)}{1 - P(\widehat{\tau} > b)}.$$

(i) $\lim_{t\to\infty} V^\star(t) = \min_b v(b)$.
(ii) Assume that there is a unique b that achieves the minimum of $v(b)$ and denote it by $\widehat{b}$. Then there is some stationary optimal policy $b(t)$ such that for all t large enough, $b(t)$ equals $\widehat{b}$.

Proof: Refer to [8] for a detailed proof. ♢

We have shown in this subsection, that for a general off time, it is enough to consider deterministic policies to achieve optimal performance. Also, if the residual off time distribution converges in time then the optimal policy converges to the constant policy and in fact becomes constant after finite time (under the appropriate conditions). This can be shown to be the case with the hyper-exponential distribution. Indeed, its residual time converges in distribution to an exponential distribution, having as parameter the smallest among the rates of the hyper-exponential distribution.

D. Suboptimal policies through dynamic programming

The DP approach has the merit of providing optimal policy whereas the parametric optimization approach of Sect. III is simpler to compute but is restricted to obtaining solutions within a class of vacation policies. In this section, we propose a suboptimal solution approach using policy iteration for a few stages. For the rest of the stages, we fix the policy to identical class of vacations and use parametric optimization.

In the simple case of one stage policy iteration, the vacations have the form $\boldsymbol{q} = (B_0, B_1, B_2, ...)(q)$. Let U_1^{exp} denote the case where $\{B_i\}_{i\geq 1}$ are i.i.d. exponentially distributed rvs with parameter $1/b'$. We can use DP to compute the optimal policy within U_1^{exp}, which is given by

$$V_1^*(\boldsymbol{q}) = \min_{b\geq 0}\Big\{\bar{\epsilon}\mathbb{E}\Big[\Big(b-\tau(\boldsymbol{q})\Big)\mathbb{1}\{\tau(\boldsymbol{q})\leq b\}\Big] + \epsilon(E_L + bE_S) + P\Big(\tau(\boldsymbol{q}) > b\Big)V_e^*(g(\boldsymbol{q},b))\Big\} \quad (22)$$

where $V_e^*(g(\boldsymbol{q},b))$ is equivalent to $V_e^*(b')$ in (10), and depends only on the state $g(\boldsymbol{q},b)$; b'^* is obtained from (9).

If we add the constraint that the first vacation should be exponentially distributed with the same distribution as B_1, i.e. $b = b'$, then we will be back to the problem of finding an optimal exponentially distributed vacation with state-independent mean. Since we do not impose this restriction, the resulting policy in U_1^{exp} will do strictly better.
Similarly, we can consider the class of vacation policies U_1^{det} of the form $\boldsymbol{q} = (\bar{b}, b', b', b', ...)(q)$ where b' is a constant. Then (22) holds with $V_e(b)$ replaced with $V_c(b)$ which corresponds to deterministic equal vacations (cf. [8]).

This suboptimal method for one stage policy iteration can be extended to more stages. Instances of the two stage policy U_2 are provided in Sect VI. As the number of stages of the policy iteration increases, the suboptimal solution converges to the optimal solution obtained from (17).

V. Worst Case Performance

We consider in this section the case where the off time is exponentially distributed with an unknown parameter. When the distribution of the parameter is known (Bayesian framework) the problem reduces to the study of the hyper-exponentially distributed off time. In practice there could be many situations when the statistical distribution of the off time is unknown or hard to estimate. In such non-Bayesian frameworks, we can conduct a a worst-case analysis: optimize the performance under the worst case choice of the unknown parameter. We assume that this parameter lies within the interval $[\lambda_a, \lambda_b]$. The worst case is identified as follows

$$\lambda_w := \arg\max_{\lambda\in[\lambda_a,\lambda_b]}\min_{\{B_k\},k\in\mathbb{N}^*} V \quad (23)$$

We obtain the λ_w for the following cases:

Optimal vacation policy (exponential off time) : The minimal cost under this policy has been derived in Sect. IV-B.1. It is expressed as follows

$$V^{\star}(\lambda) = \frac{-\bar{\epsilon} - \eta W_{-1}\left(-\exp\left(-1 - \frac{\lambda\epsilon E_L}{\eta}\right)\right)}{\lambda}.$$

We have studied this function using the mathematics software tool, Maple 11. We found the following: $V_c^{\star}(\lambda)$ is a monotonic function, decreasing with λ; $\lim_{\lambda\to+\infty} V_c^{\star}(\lambda) = \epsilon E_L$; and $\lim_{\lambda\to 0} V_c^{\star}(\lambda) = +\infty$. Evidently, $\lambda_{w,c} = \arg\max_{\lambda\in[\lambda_a,\lambda_b]} V_c^{\star}(\lambda) = \lambda_a$, just like what was found with the exponential vacation policy.

VI. Numerical Investigation

In this section we show some numerical results of our model. We study power saving scheme performance of the IEEE 802.16e standard, which allows mobile terminals to go on sleep mode when there are no packets to serve. Our model is powerful enough to analyze power save mode for large range of wireless systems. Consider the system composed of a base station and a wireless mobile terminal. The wireless mobile terminal can be seen as a server which goes on repeated vacations until a customer arrives. When this happens, the mobile node completes the current vacation before serving the customer (the packet). This captures the fact that the mobile terminal goes to sleep by turning off the radio as long as there are no packets destined to it.

In practice, the mobile needs to check for any packet arrived while it was off. So at the end of each sleep duration, it will awaken and listen to the channel for any notification message from the base station. The mobile needs to switch on the radio every time it wakes up and then listen to the channel for a (small) fixed-duration listening period. Hence, each "wake up" costs E_L amount of energy consumption. The rate of energy consumption during sleep, E_S, is considerably less than that during listening.

The cost V, defined in (1), captures the main performance measures: energy consumed during the sleep duration and extra delay incurred due to the sleep mode. The cost V is a weighted sum of both metrics. From (1), it comes that a large value of ϵ makes V more sensitive to the energy consumption than to the extra delay, whereas a small ϵ gives more weight to the delay.

We compare the optimal policy performance with IEEE 802.16e standard protocol. Policies are compared considering as performance metrics the optimal weighted cost and the corresponding optimal sleep duration. We define the improvement ratio, denoted I, as follows:

$$I := \frac{V_{\text{Std}} - V_{\text{Optimal}}}{V_{\text{Std}}}. \qquad (24)$$

To obtain V_{Std}, we substitute the following parameters of IEEE 802.16e [13] in (4). More precisely, we substitute the parameters for type I power saving classes. There, the size of a sleep window (i.e., a vacation) is doubled over time until a maximum permissible sleep window, denoted $b_{\max}$. The size of the kth vacation is then

$$b_k = b_1 2^{\min\{k-1,l\}}, \quad k \in \mathbb{N}^*$$

where $l := \log_2(b_{\max}/b_1)$. We also have

$$t_k = b_1\left(\frac{1}{2^{\min\{k,l\}}-1} + 2^l(k-l)\mathbb{1}\{k>l\}\right), \quad k \in \mathbb{N}^*.$$

We can thus express cost for standard protocol as

$$V_{\text{Std}} = -\bar{\epsilon}\,\mathbb{E}[\tau] + \sum_{k=0}^{\infty}\sum_{i=1}^{n} q_i e^{-\lambda_i t_k}\left(\epsilon E_L + \eta b_1 2^{\min\{k,l\}}\right) \quad (25)$$

The physical parameters are set to the following values: $E_L = 10$, and $E_S = 1$. The parameters of the standard protocol are $b_1 = 2$ and $l = 10$.

A. *Exponential Off Time*

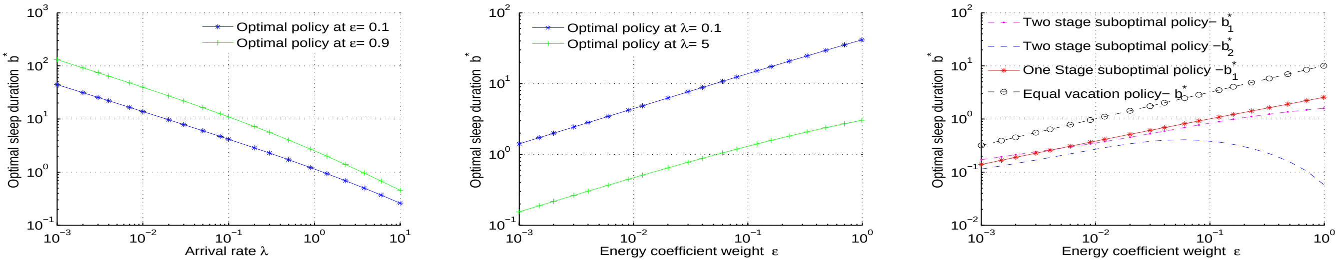
The performance metrics at hand depend on the rate λ and on the normalized weight ϵ. In the following evaluation, we will alternatively vary one of the parameters and fix the other.

The optimal sleep duration reported graphically in Fig. 1(a) against the arrival rate λ for two values of ϵ, 0.1 and 0.9. The initial sleep window for standard protocol is fixed to 2. The weight ϵ equal to 0.1 mimics the situation when energy consumption is given higher priority over delay, while ϵ equal to 0.9 mimics the opposite situation. Observe that optimal sleep window decrease as the arrival rate increases. The physical explanation for that is that a larger arrival rate forces the server to be available after shorter breaks, otherwise the cost is too high.

Figure 2(a) depicts the optimal and standard costs obtained against the rate λ and for the same values of ϵ. The optimal costs correspond to the optimal parameters plotted in Fig. 1(a). The most relevant observation on Fig. 2(a) is that the optimal policy yields better performance than the standard policy. This observation depicts the consistency of optimality.

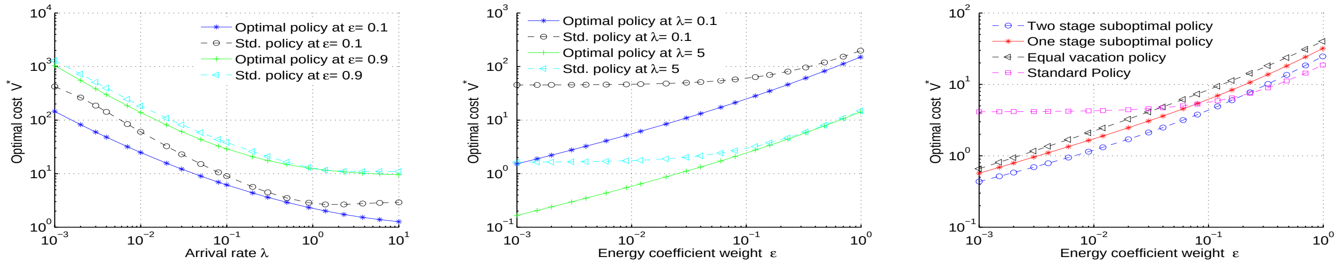
The monotonic decreasing trend of the optimal costs that has been shown analytically in Sect. V is confirmed by the numerical computations. Observe also how the cost decreases asymptotically to ϵE_L (1 for $\epsilon = 0.1$ and 9 for $\epsilon = 0.9$) as the rate λ increases. The same trend is observed for the cost of the standard policy. As λ decreases, the increase in both optimal costs is due to the increase of the optimal sleep duration (or of its expectation), while for standard policy it is due to the extra (useless and costly) listening.

We next fix λ to either 0.1 or 5, and vary ϵ. Intuitively, a smaller value of ϵ makes the extra delay more penalizing, enforcing then smaller optimal sleep durations (constant and optimal policy). This is observed in Fig. 1(b). As mentioned earlier, smaller optimal sleep durations yield smaller optimal costs. Thus, the optimal costs increase as ϵ increase, as can be observed in Fig. 2(b). For $\epsilon < 0.1$, the cost of the standard policy is fairly insensitive. This is because the standard has been designed to favor delay over energy: the first sleep duration is small ($b_1 = 2$) and it takes a while before the sleep duration becomes penalizing in terms of delay. This is confirmed by the sudden increase in cost as $\epsilon \geq 1$ (notice the logarithmic scale): when energy consumption costs start to have more weight, the standard policy's performance degrades.



(a) Exponential off time (b) Exponential off time (c) Hyper-Exponential off time

Fig. 1. Optimal expected sleep duration of various vacation policies vs λ and ϵ.



(a) Exponential off time (b) Exponential off time (c) Hyper-Exponential off time

Fig. 2. Optimal expected cost for various vacation policies vs λ and ϵ.

B. Hyper-Exponential Off Time

In this section we numerically investigate the suboptimal solutions and compare them with that of standard protocol and the i.i.d. exponentially distributed vacations policy. We consider the following parameters, initial distribution $\boldsymbol{q} = \{0.1, 0.3, 0.6\}$ with parameter $\boldsymbol{\lambda} = \{0.2, 3, 10\}$. We graphically depict cost in Fig. 2(c) and vacation sizes in Fig. 1(c) against ϵ. The suboptimal solutions are evaluated using (22), equal mean vacation policy using (10). The first stage optimal vacation size is denoted by b_1^*, in case of two stage policy iteration the second stage optimal vacation size is denoted b_2^*. Notice that the suboptimal policies performs strictly better than equal mean vacation policy. As expected two stage suboptimal policy performs better than that of one stage. As we vary ϵ, we observe a similar behaviour as that of exponential off time (cf. Fig. 2(b)). Interestingly, for large value of ϵ, standard policy outperforms all the other policies. This can be explained as large ϵ makes the system more delay sensitive where standard protocol is well suited than that of exponential protocols. However one can expect that n-stage policy iteration could perform even better than that of standard. Due to space limitation we present only the exponential vacation policy and its suboptimal policy; refer to [8] for the study of the deterministic policy also.

VII. Concluding Remarks

We have introduced a model for the control of vacations for optimizing energy saving in wireless networks taking into account the tradeoff between energy consumption and delays. Previous models studied in the literature have considered an exogeneous arrival process, where as we considered an on-off model in which the off duration begins when the server leaves on vacation and where the duration of the on period does not depend on when it starts. We note that our model can handle more general scenarios. Our model could be used for other scenarios as well; indeed, one can even consider exogeneous arrival processes by replacing the "off" period (that starts at the end of a busy period) by the *residual inter-arrival time* at the end of the busy period (the computation of its distribution is left for future work).

References

[1] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson, "Self-similarity through high variability: Statistical analysis of ethernet lan traffic at the source level," in *Proc. of ACM SIGCOMM, Cambridge, MA*, vol. 25, 1995, pp. 110–113.

[2] M. Crovella and A. Bestavros, "Self-similarity in world wide web traffic-evidence and possible causes," in *Proc. of ACM Sigmetrics, Philadelphia, PE*, 1996, pp. 160–169.

[3] A. Riska, V. Diev, and E. Smirni, "Efficient fitting of long-tailed data sets into hyperexponential distributions," in *Proc. of IEEE GLOBECOM*, vol. 3, November 2002, pp. 2513–2517.

[4] A. Feldmann and W. Whitt, "Fitting mixtures of exponentials to long-tail distributions to analyze network performance models," *Performance Evaluation*, vol. 31, no. 8, pp. 963–976, August 1998.

[5] N.-H. Lee and S. Bahk, "MAC sleep mode control considering downlink traffic pattern and mobility," in *Proc. of IEEE VTC 2005-Spring, Stockholm, Sweden*, vol. 3, May 2005, pp. 3102–3106.

[6] J. Xiao, S. Zou, B. Ren, and S. Cheng, "An enhanced energy saving mechanism in ieee 802.16e," in *Proc. of IEEE GLOBECOM 2006*, November 2006, pp. 1–5.

[7] D. G. Jeong and W. S. Jeon, "Performance of adaptive sleep period control for wireless communications systems," *IEEE Trans. on Wireless Communications*, vol. 5, pp. 3012–3016, November 2006.

[8] A. P. Azad, S. Alouf, E. Altman, V. Borkar, and G. Paschos, "Optimal sampling for state change detection with application to the control of sleep mode," http://www-sop.inria.fr/members/Sara.Alouf/.

[9] D. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 1996, vol. I.

[10] E. Feinberg, "On stationary strategies in borel dynamic programming," *Math. of Operations Research*, vol. 17, no. 2, pp. 392–397, May 1992.

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

[12] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.

[13] I. S. 802.16e 2005, "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands," 2005.