Battery-Aware Selective Transmitters in Energy-Harvesting Sensor Networks: Optimal Solution and Stochastic Dual Approximation

Jesús Fernández-Bes, Carlos III University of Madrid, Leganes, 28911, Madrid, SPAIN
Antonio G. Marques, King Juan Carlos University, Fuenlabrada, 28943, Madrid, SPAIN
Jesús Cid-Sueiro, Carlos III University of Madrid, Leganes, 28911, Madrid, SPAIN

Abstract

Energy-harvesting devices alleviate the problem of sensor nodes being powered by finite-capacity batteries. Since harvested energy is limited and stochastic, development of energy-efficient schemes to operate the network is well motivated. We investigate a scenario where, at every time instant, sensors have to decide whether to transmit or discard a message based on: its importance, the transmission energy cost, and their battery level. First, we formulate the problem as a sequential decision optimization and rely on dynamic programming to solve it. Second, we use dual stochastic methods to develop a suboptimal scheme that entails much lower computational complexity. This scheme leverages a relation between the battery level and the Lagrange multiplier associated with a long-term energy-conservation constraint. This finding is general and can be used in setups different that the considered here.

1 Introduction

A challenge for the deployment and practical application of Wireless Sensor Networks (WSNs) is that, in most scenarios, sensor nodes are powered by batteries. This limits their operational lifetime and renders energy management one of the most critical tasks. As periodic replacement of the sensor batteries is typically difficult, expensive, and even infeasible; a more practical alternative is to equip the nodes with energy harvesting devices, which are able to collect ambient energy from the environment; see, e.g., [1]. But, even when nodes are capable to harvest energy, the availability of ambient energy is usually limited and stochastic. Hence, energy-efficient operation policies are still critical to achieve a good network performance.

In this work, we design energy-efficient policies for WSNs implementing *selective* communications strategies [2], also referred to as *censoring* strategies [3]. Such strategies assume that nodes are able to evaluate/quantify the importance of the messages and use it to make a decision about transmitting or discarding them. To carry out the design, we consider a setup where, at every time instant, energy-harvesting sensors have to decide whether to transmit or discard a message based on: a) the importance/reward of the current message; b) the transmission cost; and c) the energy available at their batteries.

The contribution is threefold. First, the battery-aware decision problem is formulated as a sequential decision problem and Dynamic Programming (DP) tools are used to obtain its optimal solution. The problem is casted as a DP because current decisions (i.e., transmitting or discarding a message) change the amount of energy stored in the batteries and, therefore, have an impact not only on the current state, but also on future states. As a consequence current decisions have to take into account the (expected) impact on the future costs/rewards. Second, even in its simplest forms, DPs are difficult to solve and approximate solutions are almost always required [4]. For this reason, we leverage dual stochastic methods to develop a suboptimal scheme that entails much lower computational complexity and achieves almost optimal performance. The third scheme relies on an stochastic approximation to estimate the value of the dual variable using a biased and scaled version of the battery level. This fact does not depend on the specific problem investigated in this paper, so that it can be exploited in designs and system setups very different from the ones considered here.

In our prior work, a sequential decision framework is also used to optimize the *aggregate* importance of the messages transmitted by the sensors [2] and those that successfully delivered to the sink in a multihop network [5], but nodes were not able to harvest energy. Other sequential decision methods for energy-harvesting sensors are proposed in [6] and [7], which are based on the optimization of a long-term average reward and a very simplistic battery recharge model. In our paper we replace the long-term average by a discounted aggregate reward (which is more robust to non-stationarities), we generalize the recharge battery model to arbitrary charge distributions, and we apply a stochastic approximation method that relaxes short-term energy constraints using long-term restrictions, taking advantage of the links between battery levels and stochastic multipliers. Additionally, related DP and Stochastic Dual Approximation has also been used to deal with harvesting constraints in general wireless communication problems, e.g., [8]. Finally, Stochastic estimations of dual variables have been proposed in the context of wireless communication networks [9]. Indeed, there are several works relating such stochastic estimates to the length of queues where packets are buffered before transmission [10, 11, 12]. However, to the best of our knowledge, this is the first use of a similar strategy to establish a relationship between the stochastic dual estimates and the battery levels.

2 System model

In our model, each node in the network receives a sequence of requests to transmit messages (generated by itself or received from its neighbors) graded with different importance levels. This importance is a generic metric that depends on the application and/or quality of service requirements and accounts for the significance, priority, relevance or utility of each message [2]. At each *decision epoch* k, the node makes a transmission decision based on the available state information, so that a long-term reward is maximized. We consider k as an *event index*, so that the duration of every slot is random. In this section, we introduce notation and describe the model for the state dynamics, the decisions, and the reward.

The *state* of a node comprises two variables: b_k, the battery level at epoch k, and x_k, the importance of the message to be sent at epoch k. Consequently, the overall state of the node at epoch k is $\mathbf{s}_k = (b_k, x_k)$.

At epoch k, the node must *decide* wether to send the message or discard it. Let $d_k = 1$ if the decision (action) is to send the message and zero otherwise. A forwarding policy $\pi = \{d_k\}_{k=0}^{\infty}$ at a given node is a sequence of decision rules, which are functions of the state vector; i.e.,

$$d_k = d_k(\mathbf{s}_k) = d_k(b_k, x_k). \tag{1}$$

Through the manuscript, different decision rules are designed, and each will be denoted with a specific superindex. The (statistical) models for processes b_k and x_k are described next. Since our focus is on the energy management, we assume that x_k is a statistically independent sequence, and independent of b_{k-n} or d_{k-n}, for any $n > 0$. The model for b_k is more elaborated, because its value depends on the taken (past) actions. Let c_k denote the energy consumed when $d_k = 1$; i.e., the energy *cost* of transmitting the message k. Moreover, let $e_{in,k}$ be the amount of energy (if any) harvested by the node at epoch k. We will assume that both $e_{in,k}$ and c_k are random (the latter due to, for example, communications over fading channels). Then, b_k can be written recursively as

$$b_{k+1} = \Big[b_k - d_k c_k + e_{in,k}\Big]_0^B. \tag{2}$$

where the projection $[\cdot]_0^B = \max(0, \min(\cdot, B))$ guarantees that the energy stored in the battery neither is negative nor exceeds its maximum capacity B.

Our objective is maximizing the aggregated importance of the messages transmitted by the nodes. To formulate the reward, let first $q_k \in \{0, 1\}$ denote the *success index*, which is one if the transmission is successful and zero otherwise. The meaning of the success index may depend on the application scenario, for example, q_k can be set to 1 if the neighboring node forwards the transmitted message [5], alternatively, q_k can be set to 1 only if the message arrives to the sink. The *instantaneous reward* is given by $r_k = d_k q_k x_k$. In words, the node receives a positive reward x_k if it decides to transmit the message ($d_k = 1$) and the transmission is successful ($q_k = 1$). Otherwise, the reward is 0. Note that the exact value of q_k may not be known at instant k. Hence, transmission policies will be designed so that the *expected aggregate reward*

$$\mathbb{E}\left\{\sum_{k=0}^{\infty} \gamma^k d_k q_k x_k\right\} \tag{3}$$

is maximized. The discount factor $0 < \gamma < 1$ weights the rewards at different times exponentially, so that more focus is placed on early rewards [4], and facilitates the development of stationary decision rules. The expectation accounts not only for uncertainties at future time instants but also for those at time k (such as the value of q_k). The next sections develop algorithms aimed to maximize (3).

3 Optimal sequential decision

The optimal decision sequence is obtained as the solution to the following problem ("s. to" stands for "subject to")

$$\max_{\{d_k\}_{k=0}^{\infty}} \quad \mathbb{E}\Big\{\sum_{k=0}^{\infty} \gamma^k d_k q_k x_k\Big\} \tag{4a}$$

$$\text{s. to} \quad d_k \le u(b_k - c_k + e_{in,k}),\ \forall k \tag{4b}$$

$$d_k \in \{0, 1\},\ \forall k, \tag{4c}$$

where $u(\cdot)$ stands for the Heaviside step function. Constraint (4b) guarantees that transmissions take place only if there is enough energy in the battery. The constraint implicitly assumes that energy is harvested before taking the transmission decision.

The above optimization can be casted as a DP problem, in particular, a Markov Decision Process (MDP) [4], where, according to the model of section 2, its transition probabilities are defined as

$$p(s_{k+1}|s_k, d_k) = p(b_{k+1}|b_k, d_k)p(x_{k+1}) \tag{5}$$

This equation implies that current decisions have an impact not only on current rewards, but also on future rewards. As a result, decisions at different times are coupled and must be jointly optimized. Due to the formulation of the objective in (4a), the problem is a DP with *infinite horizon* (because all future time instants are considered) and *discounted cost* (with discount factor γ) [4]. The presence of γ guarantees that a stationary solution for the problem exists, meaning that the optimal action at any k depends only on the state at such time, i.e., that $d_k(\mathbf{s}_k) = d(\mathbf{s}_k)$. This is not only convenient from a practical perspective, but also facilitates the theoretical analysis [4].

To solve the DP, we will use Bellman's principle of optimality, stating that the optimal decision satisfies

$$d_k^{DP*} = \arg\max_d \mathbb{E}\left\{d q_k x_k + \gamma V(\mathbf{s}_{k+1})|\mathbf{s}_k\right\}, \quad \text{s. to (4b), (4c);} \tag{6}$$

where * is used to denote the optimal value of a variable, and $V(\cdot)$ represents the value function, which must satisfy the following recursive condition

$$V(\mathbf{s}_k) = \mathbb{E}\{x_k q_k d_k^{DP*} + \gamma V(\mathbf{s}_{k+1})\,|\mathbf{s}_k, d_k = d_k^{DP*}\} \tag{7}$$

Conditions (6) and (7), together with the simplifying assumptions described in previous sections, can be used to obtain an expression for the optimal decision. It can be shown (due to space limitations, proofs are omitted through the manuscript) that the optimal decision has the form of a threshold rule,

$$d_k^{DP*} = u\left(R_k - \mu^{DP}(b_k)\right). \tag{8}$$

where R_k is the (expected) instantaneous reward associated with the transmission at epoch k and it is defined as

$$R_k = \mathbb{E}\{u(b_k - c_k + e_{in,k})q_k | x_k, b_k\} x_k. \tag{9}$$

Moreover, the threshold μ^{DP} can be calculated as

$$\begin{aligned}\mu^{DP}(b_k) = \gamma(&\mathbb{E}\{V([b_k + e_{in,k}]_0^B)|b_k\} \\ &- \mathbb{E}\{V([b_k - c_k + e_{in,k}]_0^B)|b_k\})\end{aligned} \tag{10}$$

In words, the threshold is set to the (expected) *marginal* future reward if the transmission at epoch k is not performed and the corresponding energy is saved [cf. (10)].

Even if full statistical knowledge of the state variables available, the expression for $V(\cdot|b_k)$ cannot be found analytically, so that neither the value function nor the transmission threshold can be found in closed form. However, because our problem was posed as an infinite horizon DP, the value function is stationary, so that iterative numerical algorithms can be used to obtain $V(\cdot|b_k)$. Value iteration and policy iteration are the two classical algorithms used to obtain the optimal solution of a DP [4]. Note that the output of those algorithms is a function, so that they are computationally expensive. A typical approach is to sample the domain of the value function, estimate the function at the points in the sampling grid iteratively, and interpolate the values outside the grid. The aim on the next sections is to rely on convex optimization, dual decomposition and stochastic approximation to derive low-complexity suboptimal solutions for (4). Moreover, such an approach will reveal meaningful links between the stochastic estimates of the dual variables and the battery levels.

4 Stochastic dual approximation

The main difficulty in dealing with (4) is that current and future decisions are coupled and must be jointly optimized. Mathematically, this is due to the constraint (4b) and the recursive definition of b_k in (2). To account for (4b) and (2) in an alternative manner, we define $e_{out,k}(d_k) = d_k c_k$, so that the dependence of the energy consumption on d_k is explicit. Then, we formulate the constraint $e_{out,k}(d_k) \leq b_k$, which has to hold for each and every k. By using (2), the constraint can be re-written as

$$\textstyle\sum_{k=0}^{k'} e_{out,k}(d_k) \leq \sum_{k=0}^{k'} e_{in,k};\ \forall k', \tag{11}$$

where b_0 has been subsumed into $e_{in,0}$. Although (11) guarantees that (4b) is satisfied, the problem resulting from replacing (4b) with (11) would still require solving for all time instants jointly.

An alternative to facilitate the solution is to relax (11) and reformulate it as a single long-term constraint. By doing so, we have

$$\max_{\{d_k\}_{k=0}^{\infty}} \quad \mathbb{E}\Big\{\sum_{k=0}^{\infty} \gamma^k d_k q_k x_k\Big\} \tag{12a}$$

$$\text{s. to} \quad \mathbb{E}\Big\{\sum_{k=0}^{\infty} \gamma^k e_{out,k}(d_k)\Big\} \leq \mathbb{E}\Big\{\sum_{k=0}^{\infty} \gamma^k e_{in,k}\Big\} \tag{12b}$$

$$d_k \in \{0,1\} \tag{12c}$$

Note that (12b) is a single constraint, while (4b) was a set with infinite number of constraints (one per epoch). Constraint (12b) guarantees a balance between the average energy spent and the average energy harvested. In other words, (12b) is an (exponentially-weighted) average energy-conservation constraint. The latter is similar to the concept of Energy Neutral Operation, first presented in [13], and used later in a number of works. From a mathematical point of view, the motivation for (12) is that, except for the Boolean constraints in (12c), it is a convex (hence, tractable) problem. Specifically, it will be shown soon that after dualizing (12b), the optimal decision can be found separately for each epoch k.

Once the motivation to work with (12) is clear, the next step is to find its optimal solution. Let λ denote the Lagrange multiplier associated with the constraint (12b). Using the fact that $e_{out,k}(d_k) = c_k d_k$, and with L accounting for all terms that do not depend on d_k, the Lagrangian of (12) can be written as

$$\begin{aligned}\mathcal{L}(\{d_k\}_{k=0}^{\infty}, \lambda) = L& \\ + \mathbb{E}\Big\{\sum_{k=0}^{\infty} \gamma^k d_k \bar{q}_k x_k - \gamma^k \bar{c}_k d_k \lambda\Big\}&\end{aligned} \tag{13}$$

where $\bar{q}_k = \mathbb{E}\{q_k|x_k\}$ and $\bar{c}_k = \mathbb{E}\{c_k|x_k\}$ as q_k and c_k can be unobserved at the decision instant. Upon defining $\mathcal{L}_k(d_k, \lambda) = \mathbb{E}\{q_k x_k d_k - \lambda c_k d_k\}$, it is clear that (13) can be decomposed across time, so that the optimization for each d_k can be performed separately; see, e.g., [10, 9]. Hence, the optimum decision at each k is

$$d_k^{DR*}(\lambda) = \arg\max_{d\in\{0,1\}} \mathcal{L}_k(d, \lambda) = u(\bar{q}_k x_k - \bar{c}_k \lambda) \tag{14}$$

where ^{DR} stands for "Dual Relaxation". For the second equality we have leveraged that $\mathcal{L}_k(d, \lambda)$ is linear in d.

It can be shown that (12) has zero-duality gap, so that the Lagrangian relaxation is optimal [14]. Hence, the only remaining step to get the optimal solution to (12) is to find λ^*, which is the value optimizing the dual function associated with (12), and substitute $\lambda = \lambda^*$ into (14). Provided that the state distribution is known, a classical dual subgradient method can be used to carry out this task. This entails low computational burden and exhibits guaranteed convergence [14].

The main drawback associated with (14) is that it can give rise to decisions that are not feasible for (4). Specifically, there may be epochs k for which $u(\bar{q}_k x_k - \bar{c}_k \lambda^*) = 1$ but $u(b_k + e_{in,k} - \bar{c}_k) = 0$. To render d_k^{DR*} feasible, we consider the following alternative to (14)

$$d_k^{DF*} = d^{DF*}(b_k, \mathbf{x}_k, \lambda^*) := u(R_k - \bar{c}_k \lambda^*), \qquad (15)$$

where ^{DF} stands for "Dual Feasible". The iterative algorithm to compute λ^* can be updated to account for this change. Upon defining the threshold $\mu_k^{DF} = \bar{c}_k \lambda^*$, it is clear that the decision in (15) is made by comparing the instantaneous reward with μ^{DF}. The threshold in this case does not depend on the specific value of the battery b_k and leverages the expected transmission cost $\bar{c}_k$ and the long-term price of satisfying the energy-harvesting constraint represented by λ^*.

4.1 Batteries as stochastic multipliers

In the context of resource allocation for wireless networks, algorithms that do not search the optimum value of the multipliers associated with average constraints, but an on-line (time-variant) estimate of them have been proposed; see, e.g., [12]. The main motivation is threefold: i) the stochastic schemes require lower computational complexity, ii) they are robust to non-stationarities, and iii) the state distribution does not need to be known. Experimental results will show that in addition to these advantages, the stochastic schemes designed in this section: iv) are a better approximation to (8) than (15), and v) reveal existing links between Lagrange multipliers and battery levels, which can provide insights on the behavior of the schemes, and can be further exploited in future designs.

Let λ_k denote the stochastic estimate (approximation) of λ^* at epoch k. With η denoting a small *constant* stepsize, we will focus on estimates of the form

$$\lambda_{k+1} = \left[\lambda_k + \eta(e_{out}(d_k^{DF*}) - e_{in,k})\right]_0^\infty \qquad (16)$$

where the update corresponds to the stochastic estimation of the subgradient of the dual function at epoch k [9, 12]. The next step is to substitute (16) into the optimal decision rule in (15). This yields

$$d_k^{SD*} = d^{DF*}(b_k, x_k, \lambda_k) = u(R_k - \bar{c}_k \lambda_k), \qquad (17)$$

where ^{SD} stands for "Stochastic Dual". Interestingly, we observe how the stochastic decision rule implements a time-varying threshold $\bar{c}_k \lambda_k$. Following steps similar to those in [9, 12], we can show that these schemes are strictly feasible and entail a small loss of optimality (proportional to η and the variance of the stochastic update).

Moreover, the stochastic estimates of the multipliers can be related to the sensor batteries. Specifically, suppose that λ_0 and η are chosen so that up to epoch k', the projections in (2) and (16) are not needed, it holds then that $\lambda_{k'} - \lambda_0 = \eta(b_0 - b_{k'})$. In words, *the stochastic multiplier λ_k can be viewed as a scaled and offset version of the battery b_k*. This in turn implies that although b_k was not considered as an explicit state variable in (12), it naturally emerged as a stochastic estimate of the Lagrange multiplier λ which accounts for the average energy conservation constraint in the node [cf. (12b)]. This basically means that both variables play a related role, which is to account for the tradeoff between energy consumed and energy harvested.

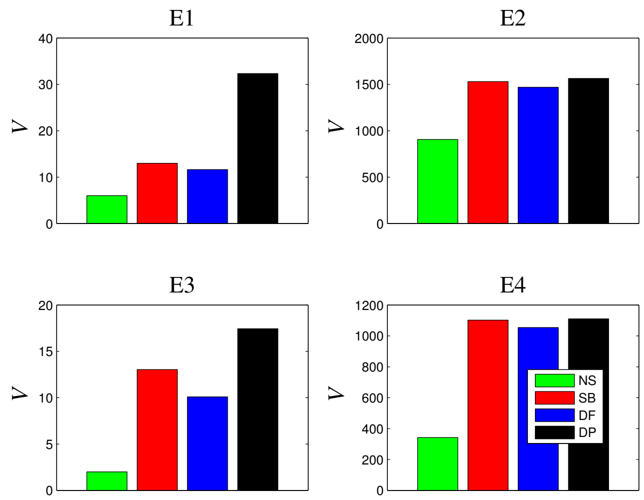
Based on this relation, an alternative stochastic estimator for the multiplier is proposed $\tilde{\lambda}_k := [\lambda_0 - \eta b_k]_0^\infty$, whose *asymptotic* feasibility and convergence properties are similar to those for (16). When $\lambda_k = \tilde{\lambda}_k$ is substituted into (17), we have that

$$d_k^{SB*} = d^{DF*}(b_k, x_k, \tilde{\lambda}_k) = u(R_k - \bar{c}_k[\lambda_0 - \eta b_k]_0^\infty) \qquad (18)$$

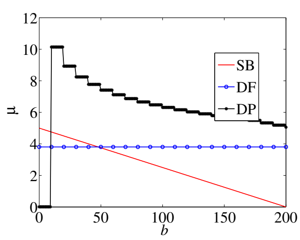
where ^{SB} stands for "Stochastic Batteries". It holds that for the stochastic dual schemes, the node decides by comparing the instantaneous reward R_k with the threshold $\mu^{SB}(b_k, \mathbf{x}_k) = \bar{c}_k[\lambda_0 - \eta b_k]_0^\infty$, which does depend on the battery level. A quick "sanity check" shows that the variation of μ^{SB} with respect to b_k is reasonable. Low batteries levels render the value of $\tilde{\lambda}_k$ high, so that unless $\bar{c}_k$ is small, μ_k^{SB} will be high and transmissions will be unlikely. Differently, if b_k is high, the value of $\tilde{\lambda}_k$ will be small, so that unless c_k is very high or $\bar{q}_k x_k$ too small, transmission will occur. Constants λ_0 and η will have an impact on the performance of the algorithm and their specific values should be set based on the system setup (finite battery capacity, messages importance, etc.). For example, the values of λ_0 and η can be set so that when $b_k = B$, almost all (x_k, c_k) combinations, say 95%, give rise to $d_k = 1$.

5 Numerical Simulations

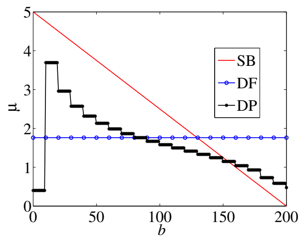
Three feasible decision schemes have been developed d^{DP*}, d^{DF*}, and d^{SB*}. Each of them implements the decision by comparing the instantaneous reward $R(b_k, x_k)$ with a different threshold. Due to space limitations, only a couple of representative numerical examples are presented. Additional material will be available in the journal version of this conference paper. In all the experiments, the importance of the messages is generated according to a exponential distribution with mean $m_x = 2$. The battery size is 200 for experiments E1 and E2, and $B = 1000$ for experiments E3 and E4. The transmission cost is $c = 10$ for E1 and E2, and $c = 20$ for E3 and E4. Energy refill is $e_{in} = 30$ with probability p_H, and $e_{in} = 0$ with probability $1 - p_H$. For E1 and E3 p_H is set to 0.001, while for E2, E4 $p_H = 0.15$. The presented results are averaged across 100 simulations. We compare the performance of four schemes: (NS) a nonselective approach; (SB) the stochastic dual solution in (18); (DF) the non-stochastic dual solution of (15), where λ^* obtained off-line using a Montecarlo approach; and (DP) the optimal scheme, which uses DP and statistical information about x. Results plotted in Fig. 1.a show that: a) both SB and DF yield a performance (reward) close to that of DP, which is a much more complex method, in most simulated cases and outperform clearly NS; and b) in most cases the stochastic approximation SB (which is computationally simpler and does not require statistical knowledge) outperforms DF. We can gain additional intuition analyzing the forwarding policy itself. Consequently, we plot the thresholds of the three schemes for a case where SB performs worse than DP (Fig. 1.b) and for one case where it performs similarly (Fig. 1.c). The plots demonstrate that 1) the threshold that SB implements is more general than the one for DF (first order vs. zero order approximation) and 2) when SB entails a big performance loss, it is because the values of η and λ_0 are grossly suboptimal. This suggests that if η and λ_0 are "wisely" selected based on prior information, SB will perform close to DP and outperform DF. Space limits prevent us from elaborating on this issue and is left as future work.



(a) Long-term importance V for different scenarios and policies.



(b) μ for E1.



(c) μ for E2.

Figure 1 Experimental results for one-hop network.

6 Conclusion

We considered an energy-harvesting WSN where nodes decide to send or discard a message based on its importance, the energy cost associated with the transmission, and the available sensor's battery. The problem was formulated as a discounted infinite-horizon DP and three different solutions were proposed. The first one satisfied the Bellman's principle of optimality and involved the computation of the *stationary* value function associated to the DP. To alleviate the computational burden associated with the DP, a convex reformulation was also considered. The basic idea was to introduce a constraint guaranteeing that the long-term consumed energy was equal to the long-term harvested energy. Such a constraint was dualized and the resultant problem was solved using dual decomposition, with the dual variable accounting for the cost of energy transmission. The third scheme relied on stochastic approximation to estimate the value of the dual variable. In turns out that a scaled and biased version of the battery level can serve as such a stochastic estimate.

7 Acknowledgement

This work was supported by the Spanish MECD FPU programme, Spanish MINECO grants Nos. TEC2011-22480, PRI-PIBIN-2011-1266 and EU FP7 ICT-2011-9 TUCAN3G.

8 References

[1] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Comput.*, vol. 4, no. 1, pp. 18–27, 2005.

[2] R. Arroyo-Valles, A. G. Marques, and J. Cid-Sueiro, "Optimal selective transmission under energy constraints in sensor networks," *IEEE Trans. on Mobile Comput.*, vol. 8, no. 11, pp. 1524–1538, Nov. 2009.

[3] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Energy-efficient detection in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 693–702, Apr. 2005.

[4] D.P. Bertsekas, *Dynamic programming and optimal control*, vol. 1, Athena Scientific Belmont, MA, 2 edition, 1995.

[5] R. Arroyo-Valles, A.G. Marques, and J. Cid-Sueiro, "Optimal selective forwarding for energy saving in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 164 –175, 2011.

[6] N. Michelusi, K. Stamatiou, and M. Zorzi, "On optimal transmission policies for energy harvesting devices," in *Inform. Theory and Applicat. Workshop (ITA), 2012.* IEEE, 2012, pp. 249–254.

[7] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 547 –551, feb. 2009.

[8] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4808–4818, 2012.

[9] A. Ribeiro, "Ergodic stochastic optimization algorithms for wireless communication and networking," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6369 – 6386, Dec. 2010.

[10] L. Georgiadis, M.J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends in Netw.*, vol. 1, no. 1, pp. 1–144, 2006.

[11] L. Huang and M.J. Neely, "Delay reduction via lagrange multipliers in stochastic network optimization," *IEEE Trans. Autom. Control*, vol. 56, no. 4, pp. 842–857, 2011.

[12] A. G. Marques, L. M. Lopez-Ramos, G. B. Giannakis, J. Ramos, and A. Caamano, "Optimal cross-layer resource allocation in cellular networks using channel and queue state information," *IEEE Trans. Vehic. Tech.*, vol. 61, no. 6, pp. 2789 – 2807, Jul. 2012.

[13] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embed. Comput. Syst.*, vol. 6, no. 4, Sept. 2007.

[14] D.P. Bertsekas, *Nonlinear programming*, Athena Scientific, 1999.