

Transactions Papers

Energy Efficient Transmission Scheduling for Delay Constrained Wireless Networks

Pavan Nuggehalli, *Member, IEEE*, Vikram Srinivasan, *Member, IEEE*,
and Ramesh R. Rao, *Senior Member, IEEE*

Abstract—In this paper, we address the problem of energy efficient packet scheduling in a wireless environment. We consider a wireless transmitter which is limited by its finite battery resource. Our objective is to design a transmission schedule that maximizes battery lifetime subject to some delay constraints. To achieve this, we exploit two previously unconnected ideas: (i) Channel coding can be used to conserve energy by transmitting at reduced power levels over longer durations; (ii) Electro-chemical mechanisms in batteries allow them to recover energy during idle periods. While the first idea favors extending transmission durations, the second idea requires the transmitter to be idle to allow for recovery. In other words, bursty packet transmissions interspersed with idle periods extend battery life. Therefore, a strategy which is based entirely on either one or the other idea is not optimal. We provide a framework to merge the two ideas. We consider two kinds of delay constraints, one a deadline constraint and the other an average delay constraint and show that energy aware scheduling strategies for both these scenarios can result in significant energy savings.

Index Terms—Scheduling, power control, wireless LAN.

I. INTRODUCTION

WIRELESS networks hold the promise of ubiquitous access to information. This promise has led to the widespread deployment of wireless based voice and data services exemplified by cellular networks and wireless local area networks. In the last few years, there has also been a concerted effort to harness wireless technology for such diverse applications as disaster relief, home entertainment and inventory management.

Manuscript received September 5, 2002; revised September 13, 2003; accepted January 10, 2005. The associate editor coordinating the review of this paper and approving it for publication was A. Sheik.

P. Nuggehalli is with the Centre for Electronics Design and Technology, Indian Institute of Science, Bangalore, 560012 (e-mail: pavan@cedt.iisc.ernet.in).

V. Srinivasan is with the Department of Electrical and Computer Engineering, National University of Singapore, 117576 (e-mail: elevs@nus.edu.sg).

R.R. Rao is with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093 USA (e-mail: rao@cw.cucsd.edu).

This work was supported by the Center for Wireless Communications, UC San Diego. An earlier version of this paper appeared in INFOCOM 2002.

Digital Object Identifier 10.1109/TWC.2006.03010

The full potential of wireless networks is limited by their dependence on energy sources with finite capacity. For example, cell phones and laptops need to be recharged every now and then from power outlets. In sensor networks deployed in inhospitable terrain, it is often impossible to recharge dead batteries and the network is rendered useless when the nodes run out of energy. It follows that the greater the energy capacity of the battery, the longer the run time of a wireless device. On the other hand, the shrinking form factor of wireless devices dictates that batteries be small in size and light weight thus limiting their energy capacity. This problem is exacerbated by the fact that advances in battery technology have not kept pace with progress in digital integration technology. Unlike transistor density which has been doubling every 18 months to the beat of Moore's law, battery energy density has improved only by a factor of 2 to 4 in the past 30 years and is expected to improve by only 20% in the next four years [4].

In the last few years, a number of research efforts have focused on energy conserving network protocols and architectures for enhancing device lifetimes. These include low-power digital and analog design as well as protocol enhancements at every layer of the protocol stack. See [5] and [6] for a current review of some of these approaches.

In this paper, we deal with the problem of designing an energy aware transmission schedule for a wireless node subject to some delay constraints. We do so by combining two previously unconnected concepts, one involving the use of channel coding [1] and the other which utilizes battery recovery [2]. In [1], it was shown that instead of transmitting packets at constant power for a fixed duration, channel coding can be used to conserve energy by transmitting at reduced power levels over longer durations. This idea favors extending transmission durations for as long as possible. In [2], it was observed that electro-chemical mechanisms in batteries allow them to recover energy during idle periods. This idea requires the transmitter to be idle to allow for recovery. In other words, the device saves energy by transmitting in bursts and recovering in the idle periods. Clearly the two energy conserving ideas above are at variance with each other. Thus, a strategy which is based entirely on either one or the other

idea is not optimal. Note that delay constraints prohibit the use of arbitrarily large transmission or idle times. In this paper we address the following question: can these two concepts be amicably merged to devise a delay constrained energy efficient transmission schedule?

In [1], they observe that, in a variety of channel coding situations, the energy required to transmit a bit is a strictly convex and decreasing function of its transmission time. Using this insight, they propose an optimal off-line lazy scheduling policy. The policy is off-line in the sense that the arrival times of all the packets in a bounded interval $(0, T)$ are known *a priori*. The aim of the policy is to schedule all departures before the deadline T . They contend that, for energy consumption to be minimized, a transmitter should attempt to transmit a packet for as long as possible subject to the deadline constraint, T . Motivated by the off-line analysis, they propose an online algorithm which they show to be near optimal. They then consider the problem of scheduling packets over an infinite time horizon with an average queuing delay constraint. They favorably compare a certain heuristically derived scheduling policy with a constant service policy that results in the same average delay. However they leave open the problem of devising an optimal scheduling policy for an arbitrary average delay constraint.

In [2], [3], [14], the electro-chemical behavior of batteries has been studied in the context of wireless devices. Batteries deliver power due to the electro-chemical reactions that occur between the electrodes and the active material near the electrodes. This leads to a depletion of active mass near the electrode which sets up a concentration gradient in the interior of the battery. Consequently, when the battery is idle, active mass diffuses towards the electrodes replenishing lost charge. This is called the *Recovery effect*. A stochastic model for charge recovery has been proposed which tracks the evolution of the energy state of a battery. This model can be used to track the amount of energy restored during an idle period.

The experimental justification for using the battery recovery mechanism for improved energy performance can be found in results reported in [7], where the capacity, energy and delivered power of Lithium ion cells has been studied under different high-power discharge profiles. These results show that batteries operating in discontinuous mode can deliver twice as much energy as when discharged using a constant current. For example, an E-tech 250 mA H battery delivered 2.78 KJ of energy when pulsed using a 20% duty cycle. In contrast it delivered only 1.3619 KJ of energy under constant current discharge. In both cases, the current level was 8 times the rated current. The authors of [7] also identify the discharge regimes where the stochastic model proposed by [2] is valid.

In this paper, we consider a single wireless device transmitting packets over a wireless channel. We assume that the device has sole access over the channel and there are no collisions or other types of contention for the channel from other wireless devices. For example, it could be a cell phone transmitting to a base station using its unique assigned channel. We also assume that for each power level, the transmitter can adaptively transmit with capacity achieving codes at that power level. We consider two kinds of delay constraints, a deadline constraint and an average delay con-

straint. The deadline constraint could model, for instance, a battery operated device that is periodically recharged. It is desirable that the battery not die before a scheduled recharge. The average delay constraint can be thought of as a Quality of Service (QoS) guarantee. In this case, it is of interest to consider strategies which can extend battery lifetime and still provide the required service. Our model is general and can be applied to any underlying physical layer scheme including FDMA, TDMA and spread spectrum systems.

We first tackle the deadline constraint problem by adapting the optimal off line scheduling policy in [1] to a discrete-time scenario. We then assume a simple battery recovery model and derive the optimal transmission strategy. Our results show that energy savings of more than 50% are feasible by accounting for energy recovery.

Next, the average delay case is considered. We assume that, packets arrive at a transmitter's queue according to a Poisson process. Packets are served on a first come, first serve basis. In order to solve the open problem posed in [1], viz. the design of an energy optimal scheduling policy for an average delay constraint, we use Constrained Dynamic Programming (CDP) [10], [13], to devise a transmission strategy using only the channel coding concept and ignoring battery dynamics. The resulting transmission policy is called the *Non-Battery Aware* (NBA) policy. Then, a stochastic model of battery behavior based on [2] is used. This allows the transmitter to conserve energy by scheduling idle periods after a transmission. However, incorporating a realistic model of energy recovery along with the channel coding concept in the CDP framework is computationally prohibitive. We resolve this by making simplifying assumptions about battery recovery. This leads to policies which may be sub-optimal. We call these the *Battery Aware* (BA) policies. We then construct a simulation model with a realistic battery component. For a given average delay constraint, the BA and NBA policies are evaluated against a *Constant Service* (CS) policy using our simulation model. With tight delay constraints, the NBA policy can extend lifetime by as much as 40% over a CS policy. With the BA policy, this figure can improve up to 90% for low arrival rates. As the delay constraint is relaxed, the CS policy is seen to do almost as well as the NBA policy. However, the gains from the BA policy become more substantial for low arrival rates.

In Section II, we address the problem of optimal scheduling in discrete-time with a deadline constraint and devise policies with and without battery recovery. In Section III, we formulate the delay constrained optimization problem and consider a more realistic battery model. In Section IV, we discuss some results. Finally we provide conclusions and directions for further research in Section V.

II. OPTIMAL OFF-LINE SCHEDULE WITH DEADLINE CONSTRAINT

In this section, we first recast the off line scheduling scheme proposed in [1] in a discrete-time framework. We start out by showing how channel coding can conserve energy in an AWGN channel at the expense of lower rates (higher transmission durations). Let $e(t)$ be the energy required to

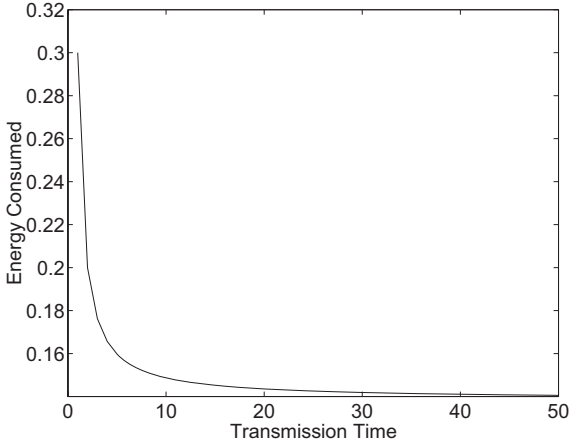


Fig. 1. Energy versus Time, derived from equation (3).

transmit a bit for a duration t . Consider the case of optimal channel coding for the AWGN case. Assume each channel use or transmission lasts for one unit of time. From Shannon's channel capacity theorem, we have

$$C = \frac{1}{2} \log_2 \left(1 + \frac{P}{N} \right) \text{ bits/transmission} \quad (1)$$

where C is the optimal capacity, given an average power constraint P and noise power N . If one assumes the use of a capacity achieving code, the number of transmissions required to transmit 1 bit is given by

$$s = \frac{1}{C} \quad (2)$$

Assuming one transmission takes 1 unit of time, the energy required to transmit 1 bit is then given by

$$\begin{aligned} e(t) &= tP \\ &= tN(2^{\frac{2}{t}} - 1) \end{aligned} \quad (3)$$

This function is a non-negative and strictly decreasing convex function (See Fig. 1). Other channel coding situations lead to a similar result. It turns out that the strictly convexity of $e(t)$ can be exploited to devise an optimal schedule which is independent of the specific nature of $e(t)$ [1].

We consider a slotted system, with each slot of unit duration. Assume that a transmitter needs to send M packets in time $[0, T]$. In other words, T is the deadline constraint. The arrival times of the packets are given by t_i , $1 \leq i \leq M$. Arriving packets are placed in a buffer where they wait to be transmitted. In our discrete time model, the t_i 's can take only integer values. The inter-arrival times are given by d_i . Define $d_M = T - t_M$, and let $\vec{\tau} = [\tau_1, \tau_2, \dots, \tau_M]$ be the transmission durations of the M packets. The τ_i 's are integers. All packets must be transmitted before T . A transmission schedule is said to be feasible if

- 1) For $1 \leq k < M$, $\sum_{i=1}^k \tau_i \geq \sum_{i=1}^k d_i$
- 2) $\sum_{i=1}^M \tau_i \leq T$

The above conditions imply that no packet is transmitted before it arrives and all packets meet the deadline constraint T .

A transmission schedule is said to be *optimal*, if it consumes the least amount of energy over all other feasible transmission schedules. Let $w(\vec{\tau})$ be the energy consumed using schedule $\vec{\tau}$. Then

$$w(\vec{\tau}) = \sum_{k=1}^M e(\tau_k) \quad (4)$$

Note that, for a schedule to be optimal, $\sum_{i=1}^M \tau_i = T$. Otherwise, one can increase the transmission time of the last packet. This would result in a feasible schedule consuming less energy than the optimal one, which is a contradiction.

We now describe the optimal scheduling policy. The proof of optimality is very similar to that in [1] and is omitted for the sake of brevity. Let $k_0 = 0$ and define

$$\begin{aligned} m_1 &= \max_{k \in \{1, \dots, M\}} \left\{ \frac{1}{k} \sum_{i=1}^k d_i \right\} \\ k_1 &= \max \left\{ k : \frac{1}{k} \sum_{i=1}^k d_i = m_1 \right\} \end{aligned}$$

For $j \geq 1$, we define

$$\begin{aligned} m_{j+1} &= \max_{k \in \{1, \dots, M-k_j\}} \left\{ \frac{1}{k} \sum_{i=1}^k d_{k_j+i} \right\} \\ k_{j+1} &= k_j + \max \left\{ k : \frac{1}{k} \sum_{i=1}^k d_{k_j+i} = m_{j+1} \right\} \end{aligned}$$

Suppose m_j is not an integer. Then it will be of the form

$$\begin{aligned} m_j &= m'_j + \frac{r_j}{k_j - k_{j-1}} \quad m'_j \in \mathcal{N} \text{ and} \\ 1 &\leq r_j \leq k_j - k_{j-1} \end{aligned}$$

Here \mathcal{N} is the set of positive integers.

Theorem 1: Suppose energy spent is a strictly convex and decreasing function of transmission time. Then, the optimal schedule $\vec{\tau}^*$ is given as follows:

Suppose $k_{j-1} < i \leq k_j$. Then

$$\tau_i^* = \begin{cases} m_j & m_j \in \mathcal{N} \\ m'_j + 1 & k_{j-1} + 1 \leq i \leq k_{j-1} + r_j \\ m'_j & k_{j-1} + r_j + 1 \leq i \leq k_j \end{cases} \quad (5)$$

Proof: See [1] for details. \blacksquare

The basic intuition is to equalize transmission times as far as possible without violating causality and deadline constraints. [1] also proposes some extensions to the online case, when the arrival epochs are unknown. However, this is still an open problem and we omit further discussion. We show next how the formalism presented so far can be extended when energy recovery is taken into account.

A. Energy Recovery

In this section, we consider a simple model for energy recovery to illustrate the need to account for this phenomenon. In later sections, we will consider more realistic models of

battery behavior. Here we assume that the battery recovery process is a strictly increasing concave function of idle time. This assumption is a good first approximation as borne out in the experimental figures reported in [9]. In order to prolong battery lifetime, we devise transmission schedules wherein, each packet transmission is followed by an idle duration. Experimental results in [7] show that such an on off discharge process can prolong battery lifetime significantly. Let $v(\tau)$ be the energy recovered in time τ . In order to describe a transmission schedule, we need to state both the actual time spent in transmission and the time spent in recovery. Let τ_i, s_i and $\tau_i - s_i$ be the total transmission time, actual transmission time and recovery time for packet i , respectively. Note that given τ_i , one can find an optimal division of τ_i into s_i and $\tau_i - s_i$ such that the least amount of energy is expended. Thus, for an optimal scheduling scheme, only the τ_i 's need to be specified. Let $z(\tau_i)$ be the optimal energy spent when the total transmission time is τ_i . We have,

$$\begin{aligned} z(\tau_i) &= \inf_{0 \leq x \leq \tau_i} e(x) - v(\tau_i - x) \\ s_i &= \arg \min_{0 \leq x \leq \tau_i} e(x) - v(\tau_i - x) \end{aligned} \quad (6)$$

In order to treat this problem using the framework developed in Section II, we use the following fact.

Lemma 1: Let f_1, \dots, f_2 be strictly convex functions on \mathcal{R}^n , and let

$$f(x) = \inf\{f_1(x_1) + \dots + f_m(x_m) | x_i \in \mathcal{R}^n, x_1 + \dots + x_m = x\}$$

Then f is a strictly convex function on \mathcal{R}^n .

Proof: See [17]. ■

Using Lemma 1, we see that $z(\tau_i)$ is strictly convex and decreasing in τ_i . Hence, $z(\tau)$ plays the exact role of $e(\tau)$ in the previous section. Thus the same theory holds and the optimal sum of transmission and idle times, τ_i , can be derived as before. The actual transmission and recovery times can then be obtained by performing the minimization in (6).

B. Results

We assume that 10 arrivals occur and need to be scheduled before a deadline $T = 59$. The first arrival occurs in the first slot at $t = 0$. The arrival times of the packets are 0, 4, 11, 20, 27, 30, 31, 35, 44, 49. The inter-arrival times can be derived from the above information. We assume an AWGN channel with a noise power of 0.1. We consider a naive scheduling scheme where the transmission durations are equated to the inter-arrival times. The total energy consumed (without recovery) was found to be 1.7215 by using (3). For the optimal scheme developed in Section II, the energy consumed without energy recovery is equal to 1.563, an improvement of around 10%. We now factor in energy recovery using (6). The energy recovered in time τ is assumed to be

$$v(\tau) = \alpha(1 - e^{-\tau}) \quad (7)$$

The energy recovery model considered here is very simple and serves mainly to illustrate the gains that can accrue by remaining idle. This is a concave increasing function as

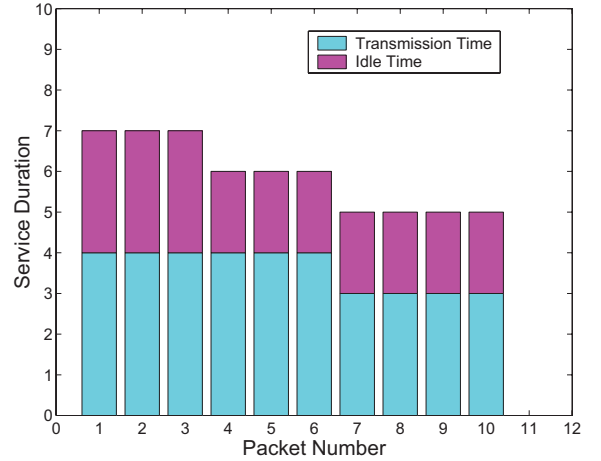


Fig. 2. Optimal scheduling policy. This policy includes energy recovery. The darker bars represent the sleep duration after packet transmission.

required and is similar in form to the experimental results shown in [9]. We took the parameter α to be 0.1 to fit the curve in [9]. With the above assumption, the energy consumed was found to be 0.7981, an improvement of around 54% over the naive scheme. The optimal scheme is shown in Fig. 2.

It can be seen that the optimal policy involves recovering for some time after transmitting. For example, the first packet is transmitted in 4 slots. The transmitter is then idle for 3 slots before serving the second packet. If energy recovery were not taken into account, the first packet would be transmitted over 7 slots.

III. DELAY CONSTRAINED SCHEDULING

In this section we consider the problem of optimal scheduling given an average delay constraint. We initially ignore battery dynamics and assume that there is no energy recovery. We use the channel coding concept to devise an optimal transmission strategy for an average delay constraint using CDP [10], [11]. This solves the open problem in [1]. We then incorporate energy recovery in the model and provide some heuristic methods to devise energy efficient transmission schemes. These schemes are evaluated via simulation.

A. Model Description

We assume a discrete-time system. Packets arrive according to a Poisson process with rate λ . All the arrivals in a slot are assumed to occur at the beginning of the slot. At the end of a transmission, the transmitter schedules the Head of Line (HOL) packet if the buffer is non-empty. This transmission duration depends on the buffer occupancy. Intuitively, if the buffer occupancy is low, the transmitter can schedule longer service periods and still meet the delay constraint. If the buffer is crowded, the transmitter will attempt to rapidly serve packets. An average delay constraint can be recast as an average "number in system" constraint using Little's law. In the formulation below, we find it convenient to use the average number of packets in the system as the constraint.

In order to describe the system in terms of a discrete-time Markov chain, we need to track two variables, viz., the buffer occupancy and the state of the service process. We model our system as a two dimensional process $\{X_n, S_n\}$, where X_n is the buffer occupancy at time n and S_n is the state of service at time n . It represents the time left to complete the present transmission. For example, the system is in state $\{5, 3\}$ when the buffer occupancy (including the packet in service) is 5 and the time remaining for completion of service is 3 slots. The state $\{5, 0\}$ represents a decision epoch. The server has just completed service and has to decide on the transmission duration of the HOL packet. In order to apply the results of Section III-B, we need to make the state space and action space finite. We do so by restricting the buffer to be of length B . By using the method of approximating sequences, we solve for the infinite buffer case. The action space $A = \{0, 1, \dots, T\}$ is simply the number of slots used for service. Note that restricting T to be finite is not a serious limitation because delay constraints and stability requirements preclude high values of T . When service is *in progress*, the only valid action is the null action denoted by 0. We assume that energy costs are incurred only at the beginning of transmission. For a state $\{x, s\}$ and action a , the energy cost, $C(x, s, a)$ is given as (3),

$$C(x, s, a) = \begin{cases} 0 & s \neq 0, \quad \forall a \in A \\ aN(2^{\frac{2}{a}} - 1) & s = 0, \quad a \in A \setminus \{0\} \end{cases} \quad (8)$$

The delay cost represented by buffer occupancy is given as

$$D(x, s, a) = x \quad (9)$$

Let $P_{\{x,s\}\{x',s'\}}(a)$ be the probability of transiting from state $\{x, s\}$ to state $\{x', s'\}$ when action a is taken. Let $p_q, 0 \leq q \leq B-1$ be the probability of q arrivals in one slot and $p_B(k)$ be the probability of $B-k$ or more arrivals in a slot. We have

$$p_q = e^{-\lambda} \frac{\lambda^q}{q!} \\ p_B(k) = 1 - \sum_{i=0}^{B-k-1} p_i \quad (10)$$

The transition probabilities are given as follows

$$\begin{aligned} P_{\{0,0\}\{q,0\}}(0) &= \begin{cases} p_q & 0 \leq q < B \\ p_B(0) & q = B \end{cases} \\ P_{\{k,1\}\{k+q-1,0\}}(0) &= \begin{cases} p_q & 0 \leq q < B-k+1 \\ p_B(k-1) & q = B-k+1 \end{cases} \\ P_{\{k,j\}\{k+q,j-1\}}(0) &= \begin{cases} p_q & 0 \leq q < B-k \\ p_B(k) & q = B-k \end{cases} \\ P_{\{k,0\}\{k+q-1,0\}}(1) &= \begin{cases} p_q & 0 \leq q < B-k+1 \\ p_B(k-1) & q = B-k+1 \end{cases} \\ P_{\{k,0\}\{k+q,a-1\}}(a) &= \begin{cases} p_q & 0 \leq q < B-k \\ p_B(k) & a \in \{2, \dots, T\} \\ p_B(k) & q = B-k \end{cases} \end{aligned} \quad (11)$$

We note here that by restricting the buffer length to B , we are not *dropping* packets but are transferring the excess

probability to the states with buffer occupancy B . This is merely a mathematical trick to solve for the infinite buffer case where there are no packet losses. See [10] for further details.

B. Constrained Dynamic Programming

Our discussion in this section closely follows the one in [10]. Consider a stochastic process in discrete time. We assume that this process can take at most a countable number of states denoted by S . For each state $s \in S$, a finite set of actions represented by the set $A(s)$ are possible. We associate two non-negative costs with each state-action pair, $C(s, a)$ and $D(s, a)$, $a \in A(s)$. Positivity is not required as long as the costs are bounded below. Let π be some scheduling policy. Let X_n and a_n be the state and action chosen at time n respectively. Then we define the following metrics

$$\begin{aligned} G_\pi(s) &= \limsup_{n \rightarrow \infty} \frac{1}{n} E_\pi \left(\sum_{i=0}^{n-1} C(X_i, a_i) | X_0 = s \right) \\ K_\pi(s) &= \limsup_{n \rightarrow \infty} \frac{1}{n} E_\pi \left(\sum_{i=0}^{n-1} D(X_i, a_i) | X_0 = s \right) \end{aligned} \quad (12)$$

Assume there is a constraint that K_π cannot exceed K . A policy π said to be feasible if $K_\pi(s) \leq K, \forall s \in S$. Let Δ be the set of all policies. The constrained optimization problem is to determine a policy π^* such that

$$G_{\pi^*}(s) = \inf_{\pi \in \Delta} G_\pi(s), \quad \forall s \in S$$

Subject to:

$$K_\pi(s) \leq K \quad \forall s \in S \quad (13)$$

The solution approach in [10] is to transform a CDP problem into an unconstrained dynamic programming problem which can then be solved using well known techniques such as the Value Iteration Algorithm which assume a finite state space.

In order to solve (13), the method of Lagrange multipliers is employed. Let $\lambda \in [0, \infty)$. The Lagrangian, for a policy π is defined as follows.

$$\begin{aligned} J_\pi(s, \lambda) &= \\ \limsup_{n \rightarrow \infty} \frac{1}{n} E_\pi \left(\sum_{i=0}^{n-1} \{C(X_i, a_i) + \lambda D(X_i, a_i)\} | X_0 = s \right) \\ &\leq G_\pi(s) + \lambda K_\pi(s), \quad \forall s \in S \end{aligned} \quad (14)$$

We define

$$J(s, \lambda) = \inf_{\pi} J_\pi(s, \lambda) \quad (15)$$

We say that a policy π is λ -optimal if $J_\pi(s, \lambda) = J(s, \lambda), \forall s \in S$. We state, without proof, the following Lemma. Details can be found in [10].

Lemma 2: If there exists $\lambda > 0$ and a λ -optimal policy π such that $G_\pi(s) < \infty$ and $K_\pi(s) = K, \forall s \in S$, then π solves the constrained optimization problem (13).

We state now the key result of this section from [10].

Theorem 2: If S is finite and $J(s, \lambda) \equiv J(\lambda), \forall \lambda > 0$ and $\forall s \in S$, then the following is true

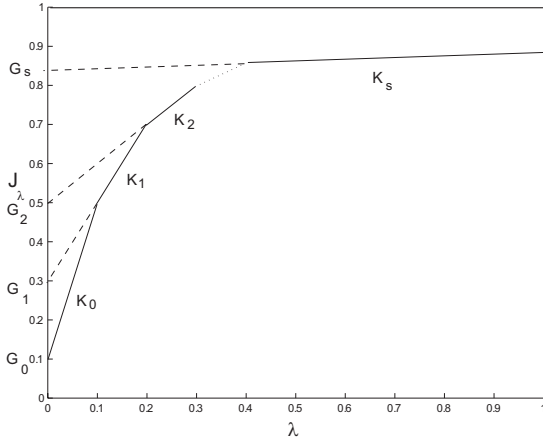


Fig. 3. A sample plot of J_λ versus λ . It is a piecewise linear increasing concave function with decreasing slope and a finite number of breakpoints. The K 's denote the slopes of various segments and the G 's are the respective intercepts.

- 1) $J(\lambda)$ is an increasing, continuous, and concave function consisting of finitely many linear segments each with a non-negative intercept. See Fig. 3.
- 2) The unique breakpoints $0 = \lambda_0 < \lambda_1 < \dots < \lambda_b$ define the $b + 1$ intervals $[\lambda_0, \lambda_1], \dots, [\lambda_{b-1}, \lambda_b], [\lambda_b, \infty)$, on each of which $J(\lambda)$ is linear.
- 3) Let $J(\lambda) = G_j + \lambda K_j$, for $\lambda \in [\lambda_j, \lambda_{j+1}]$. Then $K_0 > K_1 > \dots > K_b$ and $G_0 < G_1 < \dots < G_b$.
- 4) Assume that a stationary policy e is λ^* optimal for some $\lambda^* \in (\lambda_j, \lambda_{j+1})$, then $G_e(s) = G_j$ and $K_e(s) = K_j, \forall s \in S$. Hence e is optimal for $\lambda \in [\lambda_j, \lambda_{j+1}]$ and is said to be *essential* on $[\lambda_j, \lambda_{j+1}]$.

We now provide some intuition for Theorem 2. Since the action and state spaces are finite, there are only a finite number of policies. For any policy e , $J_e(s, \lambda)$ is a straight line with slope $K_e(s)$, intersecting the abscissa at $G_e(s)$. A little thought shows that minimizing over these straight lines (or policies) results in a curve as in Fig. 3. This is the basis of the proof for the first three properties of Theorem 2. To see why the last property is true, assume that the policy e is λ^* optimal for some $\lambda^* \in (\lambda_j, \lambda_{j+1})$. This implies that $J_e(i, \lambda^*)$ is equal to $J(\lambda^*)$ at λ^* . Since the curve is linear in $[\lambda_j, \lambda_{j+1}]$, we have that $J(\lambda)$ is the same as $J_e(i, \lambda)$ in this interval.

We now describe the use of Lemma 2 and Theorem 2 to solve the problem given by (13). If the constraint $K = K_j$ for some j , then we are done. For, using part (3) of Theorem 2 and the Value Iteration Algorithm [11], a stationary policy e , *essential* in $\lambda^* \in (\lambda_j, \lambda_{j+1})$ can be found. This policy, by Lemma (2) solves problem (13). Suppose $K_j < K < K_{j+1}$. Consider a policy which randomly combines two policies, one *essential* in the interval $[\lambda_j, \lambda_{j+1}]$ and the other *essential* in the interval $[\lambda_{j+1}, \lambda_{j+2}]$. The randomizing probability p is chosen such that $pK_j + (1-p)K_{j+1} = K$. Again by Lemma (2), this policy solves problem (13).

The above methodology can be used to derive the energy optimal scheduling policy for an arbitrary average delay constraint. This is the NBA (Non Battery Aware) policy.

C. Energy Recovery

We now incorporate energy recovery as well in our model to derive the BA (Battery Aware) policy. Various electrochemical models have been proposed for different kinds of batteries. In [18], for example, a Lithium insertion cell has been considered. [14] provides experimental results for Lithium coin cell batteries. These kind of batteries are often used in mobile devices. In [2], [3], a stochastic recovery model which approximates the electro-chemical behavior in a battery has been proposed. The evolution of battery state has been modeled as a stochastic process. Our battery model has four parameters, L , M , E_R and β . The charge delivered under a constant discharge called the rated current until the battery dies is called the *nominal capacity* denoted by M . The total charge which the battery can potentially deliver is called the *theoretical capacity*, L . The model we have used is described as follows. The state of the battery represents the energy state of the battery. The energy state is the amount of active charge material near the electrodes of the battery. The battery starts off in state M . When a packet is transmitted, the battery transits to a reduced energy state determined by (3). Energy is recovered when the transmitter is idle. In an idle slot, the battery recovers E_R units of energy with probability β , independent of state. The battery is considered dead when the energy state is 0 or if L units of charge have been consumed. It is clear that tracking the evolution of our model would require knowledge of the energy state of the battery as well as the amount of energy consumed. Thus incorporating energy recovery in the CDP framework results in an explosion of state space which makes the problem computationally intractable. To resolve this, we propose a coarse model of energy recovery to devise a scheduling policy which is possibly sub-optimal. In our coarse model, for a value of the recovery probability, β , we assume that the battery recovers $E_R * \beta$ units of energy in every idle slot. Our coarse model has only two parameters, E_R and β . We assume that the battery never dies.

The scheduling policy determines the duration of transmission of the HOL packet and the recovery duration at the end of transmission, for a given backlog. The state of the system is represented by the triplet $\{x, s, r\}$, where $\{x, s\}$ represents the backlog and state of service as defined before and r represents the remaining idle duration. For example, the state $\{5, 3, 4\}$ means that the backlog is 5, there are 3 slots remaining for end of service and once the service is over, the transmitter will sleep for 4 slots. The state $\{5, 0, 4\}$ means that the transmitter is idle and will resume service after 4 slots. The transition probabilities can be derived in a similar fashion as before and are omitted for the sake of brevity.

D. Discussion

Using the constrained dynamic programming methodology, we were able to obtain an optimal NBA scheduling policy and a possibly sub-optimal BA policy for a given arrival rate and delay constraint. In order to meaningfully evaluate the performance of the above policies, we compare them with the most energy efficient CS policy that satisfies the delay constraint. In our problem setting, these CS policies can employ only a finite number of service durations. Hence

arbitrary delay constraints cannot be satisfied exactly using only CS policies. These policies will not be optimal by Lemma 2. From the Pollaczek-Khintchine formula [12], we can compute the average delay (W_S) for a CS policy of duration a as follows

$$W_S = a + \frac{\lambda a^2}{2(1 - \lambda a)} \quad (16)$$

This can be used to determine the appropriate CS policy to compare performance with.

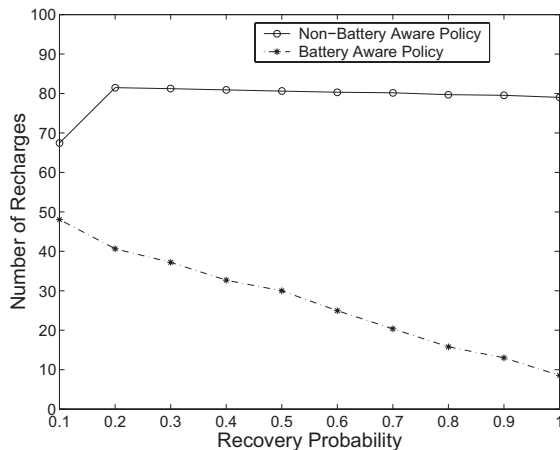


Fig. 4. Arrival Rate = 0.03, Average Delay Constraint = 1.5. Each curve shows the number of recharges required per 100 CS policy recharges. The BA policy is seen to outperform the NBA policy even for low values of recovery probability.

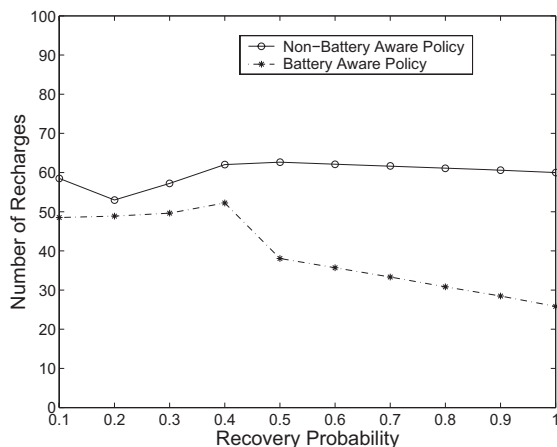


Fig. 5. Arrival Rate = 0.07, Average Delay Constraint = 1.5. Each curve shows the number of recharges required per 100 CS policy recharges. For stringent delay constraints, the NBA policy does much better than the CS policy.

IV. RESULTS

We consider an AWGN channel with a noise power of 0.1. To account for battery dynamics, we use the battery model developed before. We set E_R to be one-seventh the energy

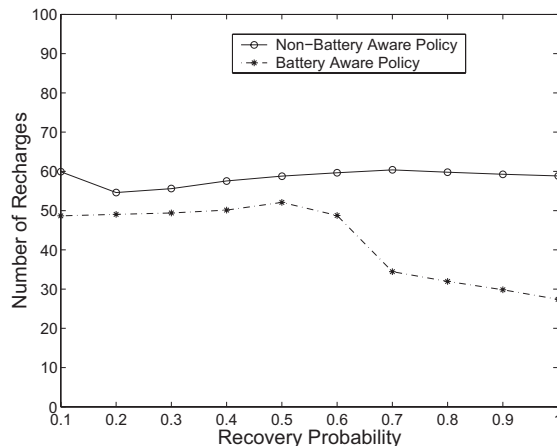


Fig. 6. Arrival Rate = 0.01, Average Delay Constraint = 1.5. Each curve shows the number of recharges required per 100 CS policy recharges. As arrival rate increases, the BA policy outperforms the NBA policy for high values of β .

spent in transmitting a packet in one slot ($\epsilon(1)/7$). This is motivated by the fact that systems like GSM employ a duty cycle of 1 : 8. Also in [9], it has been shown that a battery recovers completely in 22 ms after subjecting it to a discharge for 3 ms. We assume that a battery starts out with its energy level equal to the nominal capacity ($M = 50$). The theoretical capacity is assumed to be 10 times the nominal capacity ($L = 500$). Every time the battery dies, we assume it is recharged instantaneously. This allows us to compare the performance of different policies. The optimal policy would require the smallest number of recharges.

We consider different arrival rates and delay constraints. For a given arrival rate and delay constraint, we consider three scenarios. The first scenario is for a CS policy which best meets the delay requirement. We then compute the NBA and BA scheduling policies. Using our simulation model, we compare these three policies for different battery models by varying the recovery probability, β .

In Figs. 4, 5 and 6, we have considered three different arrival rates for a fixed stringent delay constraint. Each curve in these plots represents the number of recharges that would be required for every 100 recharges with the appropriate CS scheme. We observe that the NBA policy can result in energy savings of up to 40% over the CS policy. The BA policy does much better, saving as much as 90% over the CS policy. However, we see that as the arrival rate increases, higher values of the recovery probability (β) are required for significant gain over the NBA policy. For lax delay constraints, Figs 5 and 7 reveal that the NBA policy does not do much better than the CS policy. For example, with a delay constraint of 1.5 and an arrival rate of 0.07, the NBA policy requires only 60 recharges for every 100 recharges with the CS policy. If the delay constraint is relaxed to 5, the NBA policy requires 90 recharges, a gain of only 10%. On the other hand, the performance of the BA policy improves as the delay constraint is relaxed. For a recovery probability of 0.7, and delay constraints of 1.5 and 5, 35 and 15 recharges were respectively required for every 100 recharges with the

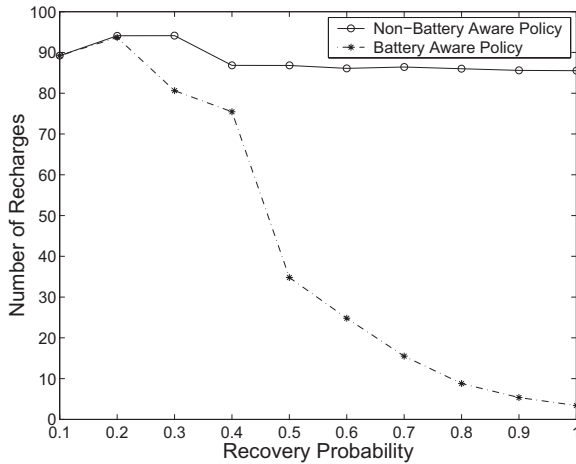


Fig. 7. Arrival Rate = 0.07, Average Delay Constraint = 5.0. Each curve shows the number of recharges required per 100 CS policy recharges. For relaxed delay constraints, the CS policy and NBA policy are comparable. The BA policy becomes more efficient as the recovery probability increases.

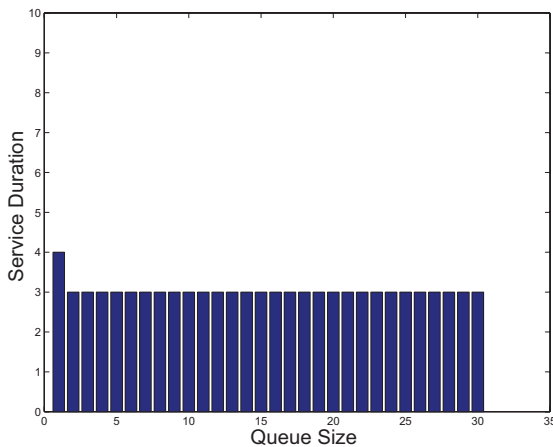


Fig. 8. NBA Policy, Arrival Rate = 0.07, Average Delay Constraint = 5.

CS policy.

The trend in the results above can be explained as follows. The energy consumption curve in (3) has a large slope in the beginning and quickly flattens out as can be seen in Fig. 1. For low arrival rates, the NBA policy attempts to transmit each packet over very long durations without recovering. However, the energy saved does not change much after a certain point and the transmitter would have been better off by being idle and recovering energy.

In Fig. 8, we plot the NBA policy for an arrival rate of 0.07 and an average delay constraint of 5 slots. In Fig. 9, we plot the corresponding BA policy. For low values of backlog, the transmitter schedules long recovery periods. On the other hand, for the NBA policy, the transmitter schedules almost uniformly long service periods. This prevents recovery and the marginal gains from transmitting for longer periods are not substantial.

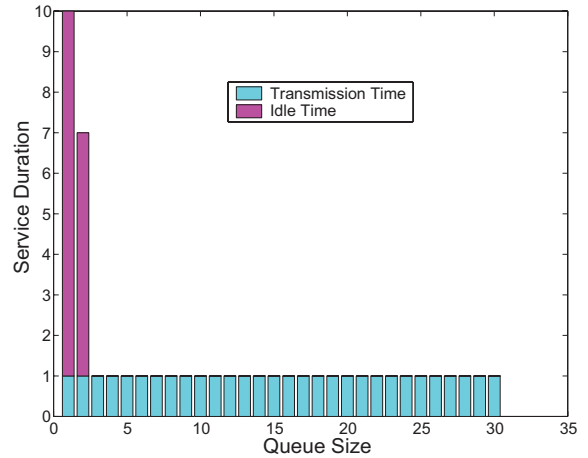


Fig. 9. BA Policy, Arrival Rate = 0.07, Average Delay Constraint = 5. The darker bars represent the sleep durations after packet transmission.

V. CONCLUSIONS

In this paper, we addressed the problem of energy efficient packet scheduling for a wireless transmitter powered by a battery. We merged together two previously unconnected concepts: (i) Channel coding can be used to conserve energy by transmitting at reduced power levels over longer durations and (ii) Electro-chemical mechanisms in batteries allow them to recover energy during idle periods. Delay constraints prohibit the use of arbitrarily large transmission or idle times. Hence, we considered the problem of devising energy optimal scheduling policies subject to delay constraints. Two kinds of delay constraints were considered, one a deadline constraint and the other an average delay constraint. For the deadline constraint, an optimal off line scheduling policy in discrete-time was devised assuming a simple model of energy recovery. We realized gains of up to 50% over a naive scheduling strategy.

Then, using the framework of Constrained Dynamic Programming and ignoring battery dynamics, we were able to solve the problem of devising an energy optimal schedule for an average delay constraint. Next, we incorporated energy recovery mechanisms in our model. However, realistic battery models led to computationally prohibitive solution methods. To address this, we used simplified battery models to devise scheduling policies that exploit the gains due to channel coding and battery recovery. We used simulation to compare the performance of our policies with a CS policy. With tight delay constraints, the NBA policy saved as much as 40% over a CS policy. With the BA policy, this figure improves up to 90% for low arrival rates. As the delay constraint was relaxed, the CS policy was seen to do almost as well as the NBA policy. However, the gains from the BA policy became more substantial for low arrival rates.

We can identify several directions for further research. While recent experimental evidence [7] indicates that battery recovery mechanisms can be exploited to draw more energy from high performance batteries, battery and circuit behavior need to be better understood to create accurate analytical

models. We believe that the Poisson arrival model is, in some sense, a lower bound on achievable performance. Since real life traffic is more predictable, more gains appear feasible. Lastly, it will be interesting to incorporate more realistic channel models and coding mechanisms.

REFERENCES

- [1] B. Prabhakar, E. Uysal, and A. El Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proc. Infocom*, Apr. 2001.
- [2] C. F. Chiasserini and R. R. Rao, "Energy efficient battery management," *IEEE J. Select. Areas Commun.*, vol. 19, no. 7, pp. 1235-1245, July 2001.
- [3] C. F. Chiasserini and R. R. Rao, "Improving battery performance by using traffic shaping techniques," *IEEE J. Select. Areas Commun.*, vol. 19, no. 7, pp. 1385-1394, July 2001.
- [4] R. Powers, "Advances and trends in primary and small secondary batteries," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 9, no. 4, pp. 32-36, Apr. 1994.
- [5] *IEEE Wireless Commun. Mag.*, vol. 9, no. 4, Aug. 2002.
- [6] L. Bennini, A. Bogliolo, and G. De Micheli, "A survey of design techniques for system level dynamic power management," *IEEE Trans. VLSI Syst.*, vol. 8, no. 3, pp. 299-316, June 2000.
- [7] M. B. Patel, D. F. Kimball, and R. R. Rao, "Efficient utilization of Lithium cell battery at high discharge rate for mobile devices and mobile ad hoc networks," To be submitted. [Online] Available at <http://fieng9.ucsd.edu/~m4patel/research.htm>
- [8] A. K. Salkintzis and C. Chamzas, "An in-band power-saving protocol for mobile data networks," *IEEE Trans. Commun.*, vol. 46, pp. 1194-1205, Sep. 1998.
- [9] R. M. LaFollette, "Design and performance of high specific power, pulsed discharge, bipolar lead acid batteries," in *Proc. 10th Annual Battery Conference on Applications and Advances*, Jan. 1995.
- [10] L. I. Sennott, "Computing average optimal constrained policies in stochastic dynamic programming," *Probability in the Engineering and Information Sciences*, vol. 15, pp. 103-133, Feb. 2001.
- [11] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. New York: John Wiley & Sons, 1999.
- [12] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
- [13] E. Altman, *Constrained Markov Decision Processes*. London, UK: Chapman & Hall/CRC, 1999.
- [14] S. Park, A. Savvides, and M. B. Srivastava, "Battery capacity measurement and analysis using Lithium coin cell battery," in *Proc. of ISLPED*, Aug. 2001.
- [15] I. Chlamtac, C. Petrioli, and J. Redi, "Energy-conserving access protocols for identification networks," *IEEE/ACM Trans. Networking*, vol. 7, pp. 51-59, Feb. 1999.
- [16] L. E. Larson, "Radio frequency integrated circuit technology for low-power wireless communications," *IEEE Pers. Commun.*, vol. 5, no. 3, pp.11-19, June 1998.

- [17] R. T. Rockafellar, *Convex Analysis*. Princeton: Princeton University Press, 1970.
- [18] M. Doyle, T. F. Fuller, and J. S. Newman, "Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell," *J. Electrochem. Soc.*, vol. 140, pp. 1526-1533, June 1993.



Pavan Nuggehalli (S'98-M'03) received the M.Sc (Engg.) degree in electrical engineering from the Indian Institute of Science, Bangalore, India, in 1998, and the Ph.D. degree in electrical and computer engineering from the University of California at San Diego, La Jolla, in 2003.

He is currently an Assistant Professor with the Centre for Electronics Design and Technology, Indian Institute of Science, Bangalore. His research interests lie in performance analysis of wireless ad hoc and sensor networks.



Vikram Srinivasan (S'98-M'03) received the B.Sc degree in physics from the University of Madras, Chennai, India in 1994, the M.E. degree in electrical communications engineering from the Indian Institute of Science, Bangalore, India, in 1998, and the Ph.D. degree in electrical and computer engineering from the University of California at San Diego, La Jolla, in 2003.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, the National University of Singapore, Singapore. His

research interests include resource allocation and performance analysis of communication protocols for wireless networks.



Ramesh R. Rao (M'85-SM'90) received the B.E. degree (with honors) in Electronics and Communications from the University of Madras, Chennai, India, in 1980 and the M.S. and Ph.D. degrees from the University of Maryland, College Park, Maryland, in 1982 and 1984, respectively.

Since 1984, he has been on the faculty of the department of Electrical and Computer Engineering at the University of California, San Diego, where he is currently Professor and Director of the San Diego Division of the California Institute of Telecommunications and Information Technology. Prior to this appointment, he served as the Director of the UCSD Center for Wireless Communications (CWC) and was the Vice Chair of Instructional Affairs in the Department of Electrical and Computer Engineering. In April 2004, he was named Qualcomm Endowed Chair in Telecommunications and Information Technology. He served as the Editor of the Information Theory Society Newsletter from 1993 to 1995 and is the founding Web Editor of the Information Theory Society website. His research interests include architectures, protocols and performance analysis of wireless, wire line and photonic networks for integrated multi-media services.

Dr. Rao was elected to the IEEE Information Theory Society Board of Governors from 1997 to 1999 and from 2000 to 2002.