Optimal Adaptive Modulation Policy For QoS Constrained Wireless Networks With Renewable Energy Sources
Abhishek Agarwal and Aditya K. Jagannatham

Abstract
In this paper we derive the optimal policy to minimize the average number of dropped packets for a delay constrained wireless node with a renewable energy source. The proposed framework employs adaptive modulation for transmission of the optimal number of packets towards satisfying the Quality of Service (QoS) constraints. This framework is formulated as a Markov Decision Process (MDP) with a state space consisting of the joint channel state, energy stored in the renewable battery, and data queue length. The optimal policy derived from the above MDP yields the appropriate modulation, and hence the number of packets to be transmitted for long term packet drop minimization. Further, we demonstrate that the optimal policy is monotone with respect to each of the joint state components. Simulation results are presented to validate the performance and monotone structure of the proposed optimal policy.

Index Terms

I. INTRODUCTION
Modern wireless communication networks are rapidly progressing towards the transmission of delay-sensitive rich multimedia content over packet switched networks. In such scenarios satisfying the associated QoS constraints with respect to delay, jitter, and throughput under erratic channel conditions requires large amounts of energy. Hence, in energy limited battery...
powered systems, it is critical to optimally allocate energy for QoS satisfaction. In this scenario, to ease the system energy demands, renewable energy sources (such as solar, wind etc.) can be utilized to enhance the battery and hence performance of mobile nodes. The random charging nature of such renewable sources calls for efficient packet scheduling policies for performance maximization.

Towards this end, we propose a framework for optimal packet scheduling in such delay constrained wireless networks which employ renewable energy resources. This is formulated as a MDP consisting of the wireless channel state, renewable battery energy, and data queue length at the transmitting wireless node. The above framework is employed to derive an adaptive modulation based optimal policy to efficiently vary the rate and power of transmission while satisfying a maximum bit-error rate (BER) constraint on the transmitted data. Such a policy aims to minimize the long term packet drop percentage through optimal selection of the transmit power and data rate parameters in accordance with the dynamically evolving wireless channel environment. Further, we demonstrate that the proposed policy has the desirable monotone properties with respect to the key channel, battery, and queue state parameters. This leads to the design of a computationally feasible scheme for packet scheduling suitable for wireless nodes with limited processing capabilities.

Recently, resource allocation for wireless nodes with renewable energy sources has attracted significant research interest. In [1] the authors present a scheme towards long term capacity maximization through efficient transmit power allocation. However, the authors therein assume infinite backlogged data at the transmitter and continuously adaptable transmission rate as a function of the transmit power. In this paper we study the problem in a more practical context in which the buffer length is finite and the data transmission is constrained to a finite class of adaptive M-ary Quadrature Amplitude Modulation (M-QAM) constellations. Other related studies such as [2], [3] have been carried out for average power constrained systems. For instance, in [2] the authors consider a model similar to ours employing adaptive modulation to vary the data rate while satisfying maximum BER constraints. However, since the work therein does not consider renewable energy sources it does not effectively capture the balance between the power consumed and packets dropped. Hence, the battery inclusive joint state MDP in our formulation captures the trade-off between the packet drop rate and the power consumed through the Markovian state transitions. Thus, our optimal policy naturally leads to efficient long term power utilization.
This paper is organized as follows. Section II provides an overview of the system and formulates the MDP for the above delay constrained wireless transmission scenario. Section III presents the central results that establish the monotonicity properties of the optimal policy for the MDP. Simulation results are presented in Section IV and Section V concludes the paper.

II. SYSTEM MODEL AND MDP FORMULATION

We consider a renewable energy source based wireless mobile user (MU) with a finite data buffer. The energy recharge of the renewable battery in each time slot is modeled as a random process to capture the variation in the charging phenomenon. The wireless channel between the MU and the base station (BS) is assumed to be Rayleigh Fading. The MU employs adaptive modulation to vary its data rate with the channel conditions such that the maximum BER is constrained by the QoS restrictions. This requires the MU to transmit a given number of bits (constellation size of modulation scheme) at a minimum power. Hence, the transmit power required for a QAM constellation of order $2^a$ is given from [4] as,

$$P(a, ||h||^2) = \frac{\Gamma (2^a - 1)}{||h||^2},$$

(1)

where $\Gamma$ is a constant dependent on the noise power and the required BER and $h$ is the Rayleigh fading channel coefficient. Each time slot comprises of $N$ symbol transmissions of the adaptively chosen constellation belonging to the class of M-QAM constellations. Hence, without loss of generality we consider a model normalized by $N$.

We formulate this paradigm as an infinite horizon discounted Markov Decision Process with $s_n$, the state at time $n$ defined as,

$$s_n = (g_n, e_n, b_n) \in \mathcal{S},$$

where $g_n = ||h_n||^2$, $e_n$ is the battery available at time $n$, $b_n$ is the data queue length, and $\mathcal{S}$ denotes the state space. The optimal stationary policy for the above MDP gives the appropriate choice of the adaptive QAM constellation size $a_n$ at time instant $n$ towards long term packet drop minimization depending upon the state $s_n$. The action space at time $n$, $\mathcal{A}(s_n) = \{0, 1, \ldots, a_n\}$ is restricted by the amount of energy available and the bits present in the buffer. Hence, $P(a_n, g_n) \leq e_n$ and $a_n \leq b_n$. Due to the strict delay constraint the untransmitted data is rendered redundant after one time epoch. Hence, the the reward at time $n$ can be defined as,

$$r_n (s_n, a_n) = -(b_n - a_n),$$

(2)
which essentially imposes a penalty proportional to the number of packets lost. The MU objective of long term average reward maximization is captured by the average discounted reward function defined as,

\[ R^\pi(s_n) = r(s_n, \pi(s_n)) + \beta \sum_{s_{n+1} \in S} q(s_{n+1}|s_n, a_n) R^\pi(s_{n+1}), \]

where \( \pi : S \rightarrow \bigcup_{s \in S} A(s) \) is the MU adaptive modulation policy, \( q(s_{n+1}|s_n, a_n) \) is the transition probability and \( \beta \) is the discount factor. The transition from one state to another can be defined as,

\[ s_{n+1} = (g_{n+1}, e_n - P(a_n, g_n) + \xi_n, b_{n+1}), \]

where \( g_{n+1} \) is the uncorrelated exponentially distributed channel gain, \( \xi_n \) is an independent random variable denoting the amount of energy recharged, and \( b_{n+1} \) is an independent data arrival random variable.

Let \( g_n \in \{0,1,\ldots,H_m\}, e_n \in \{0,1,\ldots,E_m\}, \) and \( b_n \in \{0,1,\ldots,L\} \) denote the size of the state space, where \( L \) denotes the data buffer size and \( E_m \) the number of energy units in the battery, and \( H_m \) the size of the channel state space. Let the maximum constellation size of the M-QAM be \( 2^M \). Then, as shown in Section IV, even for small parameter values such as \( M = 4, E_m = 7, L = 9 \), the size of the state space is \( |S| \approx 5 \times 8^2 \times 10 = 3200 \), which renders the standard value and policy iteration based techniques for optimal policy computation impractical. Hence, in the next section we prove key properties of the above MDP, for arbitrary distributions of the data arrival and energy recharge processes, such that the computation of an optimal policy becomes tractable.

### III. Monotonicity Properties of the Optimal Policy

In this section we employ the convexity properties of \( P(a, g) \) in (1) to derive the key monotonicity properties of the optimal policy for the above MDP.

#### A. Concavity of Average Reward in \( e \)

Observe that the optimal value function must be non-decreasing in \( e \), since increasing \( e \) can only increase the action space. Secondly, the function \( P(a, g) \) is convex in \( a \). Thus we have,

\[ \alpha P(a^1,g) + (1-\alpha) P(a^2,g) \geq P(\bar{a},g), \]
where $0 \leq \alpha \leq 1$ and $\bar{a} = \alpha a^1 + (1 - \alpha) a^2$. We now prove the following result

**Lemma 1.** Let $\bar{e}$ be defined as, $\bar{e} = \alpha e^1 + (1 - \alpha) e^2$. It then follows that,

$$\alpha R (g, e^1, b) + (1 - \alpha) R (g, e^2, b) \leq R (g, \bar{e}, b).$$

**Proof:** We prove the above result using induction. For ease of illustration, we replace $R_k (g, e, b)$ by $R_k (e)$ in this section. It has been demonstrated in [5] that optimal value function $R (s)$ using value iteration is,

$$R_{k+1} (s) = \max_a \{r (s, a) + \beta \times \mathbb{E}_{g', \xi, b'} [R_k (s - P (g, a) + \xi)]\}.$$

Under the inductive assumption, let $R_k (s)$ be concave in $e$. Thus, to prove the concavity of $R (s)$ it is necessary to demonstrate the concavity of $R_{k+1} (e)$, which is described as,

$$\alpha R_{k+1} (e^1) + (1 - \alpha) R_{k+1} (e^2) \leq R_{k+1} (\bar{e}).$$

Let the optimal action at states $(g, e^i, b)$ be given by $a^i$, $i \in 1, 2$. Thus, $R_{k+1} (e^i)$ is given as,

$$r \left( b, a^i \right) + \beta \times \mathbb{E}_{g', \xi, b'} [R_k (e^i - P (g, a^i) + \xi)],$$

where $i = 1, 2$. Also, from (2) we have,

$$\alpha r \left( b, a^1 \right) + (1 - \alpha) r \left( b, a^2 \right) = r \left( b, \bar{a} \right),$$

where $\bar{a} = \alpha a^1 + (1 - \alpha) a^2$. Thus, from the concavity of $R_k$, we have,

$$\alpha r \left( b, a^1 \right) + (1 - \alpha) r \left( b, a^2 \right) \leq \alpha r \left( b, \bar{a} \right) + (1 - \alpha) r \left( b, \bar{a} \right) + \beta \mathbb{E} [R_k (\bar{e} - P (g, \bar{a}) + \xi)] \leq R_k (\bar{e} - P (g, \bar{a}) + \xi),$$

where the last inequality follows from the convexity of $P (g, a)$ and the non-decreasing nature of $R (g, e, b)$ in $e$. Thus from (4) and (5) we have,

$$\alpha R_{k+1} (e^1) + (1 - \alpha) R_{k+1} (e^2) \leq r (b, \bar{a}) + \beta \mathbb{E} [R_k (\bar{e} - P (g, \bar{a}) + \xi)] \leq r (b, a^*) + \beta \mathbb{E} [R_k (\bar{e} - P (g, a^*) + \xi)] = R_{k+1} (g, \bar{e}, b),$$

where the first inequality follows from the linearity of the expectation operator and the subsequent inequality follows from the fact that $a^*$ is the optimal action for $R_{k+1} (g, \bar{e}, b)$. 

We now prove the monotonicity of the optimal policy with respect to the channel gain $g$. 

August 23, 2012
B. Monotonicity of the optimal policy in \( g \)

Let the optimal policy in state \( s^i = (g^i, e, b) \) be \( a^i, i \in \{1, 2\} \). Consider the value function, 
\[
R((g, e, b), a) = r(b, a) + \beta E[R(g', e - P(g, a) + \xi, b')].
\]  
(6)

Thus, for \( g^1 < g^2 \) and \( a^1 > a^2 \), we have \( |A(s^2)| \geq |A(s^1)| \). Hence, we have,
\[
R(s^1, a^1) \geq R(s^1, a^2),
\]
(7)
\[
R(s^2, a^2) \geq R(s^2, a^1),
\]
(8)

Now, adding inequalities (7) and (8) and employing the relation in (6) we have,
\[
E_R(p_{1,1}) + E_R(p_{2,2}) \geq E_R(p_{1,2}) + E_R(p_{2,1}),
\]
(9)
where \( p_{i,j} = e - P(a^i, g^j), 1 \leq i, j \leq 2 \) and \( E_R(p) = E_{g',\xi,\beta'}[R(g', p + \xi, b')]. \) From the definition of \( P(a, g) \) in (1) it follows that,
\[
P(a^1, g^1) - P(a^2, g^1) > P(a^1, g^2) - P(a^2, g^2)
\]
\[
\implies p_{1,1} + p_{2,2} < p_{1,2} + p_{2,1},
\]
(10)

Further, we have either \( p_{1,1} \leq p_{1,2} \leq p_{2,1} \leq p_{2,2} \) or \( p_{1,1} \leq p_{2,1} \leq p_{1,2} \leq p_{2,2} \). Thus, from the concavity and monotonicity of \( R(b, g, e) \) in \( e \) and the above inequality we have,
\[
E_R(p_{1,1}) + E_R(p_{2,2}) \leq E_R(p_{1,2}) + E_R(p_{2,1}),
\]
which contradicts (9). Thus, given \( g^1 < g^2 \) we must have \( a^1 \leq a^2 \), which completes the proof.

C. Monotonicity of the optimal policy in \( e \)

Employing an argument similar to the one in section III-B above, we prove that the the optimal policy is also monotone in the battery energy \( e \). Let the optimal policy in state \( s^i = (g, e^i, b) \) be \( a^i, i \in \{1, 2\} \). Let \( e^1 < e^2 \) and \( a^1 > a^2 \). Adding the relations (7) and (8) and following a procedure similar to the proof illustrated above, we have,
\[
E_R(p_{2,2}) + E_R(p_{1,1}) \geq E_R(p_{2,1}) + E_R(p_{1,2}),
\]
(11)
where \( p_{i,j} = E_R(e^i - P(a^j, g)) i, j \in \{1, 2\} \). Also from the definition of \( P(a, g) \) in (1) we have,
\[
p_{2,2} \geq p_{2,1} \geq p_{1,2} \geq p_{1,1},
\]
Employing the concavity property of $R(g,e,b)$ in $e$ and the linearity of the expectation operator, it follows from the above equations that,

$$E_R(p_{2,2}) + E_R(p_{1,1}) \leq E_R(p_{2,1}) + E_R(p_{1,2}),$$

which contradicts (11). Thus, given $e^1 < e^2$ it follows that $a^1 \leq a^2$.

**D. Monotonicity of the optimal policy in $b$**

In this section we demonstrate that the optimal policy is also monotone in the data queue length $b$. The optimal value function for a given action in state $s = (g,e,b)$ is,

$$R(s,a) = r(b,a) + \beta \times E[R(g',e-P(g,a) + \xi,b')],$$

(12)

where $a \in \mathcal{A}(s)$. Consider the states $s^i = (g,e,b^i)$ for $i = \{1,2\}$ such that $b^1 < b^2$. Let the optimal actions in $s^i$ be $a^i$, $i \in \{1,2\}$. Thus, since the action space is determined by $P(g,e)$ and the queue length $b$, we have $|\mathcal{A}(s^1)| \leq |\mathcal{A}(s^2)|$. From (12) it can be observed that the second term in $R(s^1)$ and $R(s^2)$ is identical for all $a$. The first term $r(b,a)$ differs only by a constant, $(b^1 - b^2)$. Thus if $a^2 \in \mathcal{A}(s^1)$ we must have $a^1 = a^2$. Otherwise if $a^2 \in \mathcal{A}(s^2) - \mathcal{A}(s^1)$ we necessarily have $a^2 > a^1$. Hence, the optimal action is monotone in $b$.

We now show that if $a^2 \in \mathcal{A}(s^2) - \mathcal{A}(s^1)$ we must have $a^1 = \max \mathcal{A}(s^1)$. The first term in (12) is linearly increasing in $a$. The second term, $T_2(a)$, is decreasing. Further, we have,

$$[e - P(a - 1, g)] + [e - P(a + 1, g)] \leq 2 \times [e - P(a, g)]$$

and $a - 1 \leq a \leq a + 1$. Thus, employing the concavity result in III-A we have,

$$T_2(a - 1) - T_2(a) \leq T_2(a) - T_2(a + 1).$$

(13)

Hence, $T_2(a)$ is concave decreasing in $a$. Therefore, given $a^2 \in \mathcal{A}(s^2) - \mathcal{A}(s^a)$, $a^1 = \max \mathcal{A}(s^1)$.

**IV. Simulation Results**

In our simulation setup for the above delay constrained wireless scheduling scenario, we consider M-QAM constellations with maximum modulation size $2^M = 2^4$. The buffer length considered is $L = 7$ and the maximum number of energy units is $E_m = 9$. The data arrival process $b_n$ is assumed to be Poisson distributed with mean $\lambda$. The probability of data loss due to buffer overflow is neglected. Hence, we consider a finite support $\{0,1, \ldots, L\}$ for $b_n$. The
discount factor $\beta = 0.9$. The energy renewal process $\xi_n$ is independent exponentially distributed with support $\{0, 1, \ldots, \xi_{\text{max}}\}$, $\xi_{\text{max}} = 7$, and mean $\mu$. The total battery energy is divided into $E_m$ equal energy units. The maximum energy parameter $E_{\text{max}}$ is set such that the outage probability for transmitting a single bit at the maximum power level is 0.1.

For a practical wireless system scenario, the energy and channel states have to be quantized. Hence, the associated energy function $P(a, g)$ is adapted to satisfy the QoS constraints as,

$$P(a, g) = \left\lceil \frac{\Gamma(2^a - 1)}{g} \times \frac{E_m}{E_{\text{max}}} \right\rceil,$$

where $\lceil x \rceil$ denotes the ceiling function. For the transmission of $a$ bits the channel gain can be divided into $E_m + 1$ intervals according to the cost required in each interval. Thus,

$$H_m + 1 \leq (E_m + 1) \times (M + 1).$$

The actual value of $H_m$ will be slightly lower than $(E_m + 1) \times (M + 1)$ due to the overlapping nature of the interval boundaries. Hence, the state space is limited as,

$$|\mathcal{S}| \leq (M + 1) \times (E_m + 1)^2 \times (L + 1).$$

As illustrated in section II, even for moderate values of the MDP parameters $M, E_m, L$ the size of the state space becomes intractably large. Hence, the monotone properties derived in section III are crucial for the computation of the optimal policy. In Fig. 1, 2, 3 (left) we plot the optimal policy of the above MDP as a function of channel gain $g$, battery energy $e$, and data queue length $b$ parameters respectively. From the plots we observe that the optimal policy for the quantized MDP is monotone in the above MDP joint state components. The stronger properties for the optimal policy proved in III-D can also be readily seen from these plots. Thus, one can employ the monotonicity properties of the optimal policy proved above to significantly reduce the computations required to arrive at the optimal adaptive modulation scheduling policy.

In Fig. 1, 2, 3 (right) we plot the performance of the optimal policy, in terms of the long term packet drop rate, $R(g, e, b)$ in (3), with respect to the state components $g, e, b$. The plots for different values of the parameters $M$ and $\lambda$ show the variation of $R(g, e, b)$ with these parameters. As expected, the drop rate increases with $\lambda$, the data arrival rate. Also, as the maximum constellation size $M$ increases, the drop rate decreases since the action space becomes larger. Finally, when both $M$ and $\lambda$ are increased, the variation in $R(g, e, b)$ is small since the increase in the data arrival rate counters the effect of the increase in the action space.
V. CONCLUSION

In this paper we formulate the problem of variable rate, variable power data transmission for a QoS constrained wireless user with renewable battery energy, as a Markov Decision Process. It has been demonstrated that the high dimensionality of the problem makes it prohibitively complex to employ straightforward techniques such as value and policy iteration for optimal policy computation. Thus, employing the convexity properties of the power function, for the above transmission scenario, we have derived key monotonicity properties of the optimal transmission
Figure 3. Optimal number of transmitted bits vs. data queue length $b$ (left) at $M = 4, \lambda = 4$. Packet Drop Rate vs. queue length $b$ (right) at $g = 12$ and $e = 8$. The mean of the random variables $b_n$ and $\xi_n$ are $\lambda = 4$ and $\mu = 2.953$, respectively.

policy for the above MDP. These properties significantly reduce the computation required for the optimal policy for long term packet drop minimization.

REFERENCES


