Optimal Adaptive Modulation 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) which minimizes the long term packet drop rate. Further, we demonstrate that the optimal policy is monotone with respect to each of the joint state components of the MDP. Simulation results are presented to validate the monotone structure and superior performance of the proposed policy with respect to the existing schemes.**

***Index Terms*—Markov decision process, adaptive modulation, renewable energy source, quality of service.**

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 Quality of Service (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 this delay and QoS constrained wireless scenario which employs renewable energy resources. This problem is formulated as an Markov Decision Process (MDP) with a maximum bit-error rate (BER) QoS constraint. The state space of the MDP consists of the wireless channel state, renewable battery energy, and data queue length at the transmitting wireless node. The optimal policy of the MDP aims to minimize the long term packet drop percentage through optimal selection of the transmit power and data rate parameters. Further, we demonstrate that the proposed policy has the desirable monotone properties with respect to each state component, leading to a computationally feasible scheme for packet scheduling suitable for nodes with limited processing capabilities.

Manuscript received September 20, 2012. The associate editor coordinating the review of this letter and approving it for publication was D. Niyato.

The authors are with the Department of EE, IIT Kanpur, 208016, India (e-mail: adityaj@iitk.ac.in).

Digital Object Identifier 10.1109/WCL.2012.112012.120696

Recently, resource allocation for renewable energy wireless nodes 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. Other works such as [2], [3], [4], [5] consider similar scheduling problems for single and multiple users without renewable energy sources. The authors in [2], [3] consider data scheduling with the aim of minimizing the average transmit power while limiting the average delay assuming that an instantaneous error-free feedback channel is available. The scheme in [5] uses multiple packet transmissions to counter channel fading. A significant drawback of the scheme therein is the absence of delay constraints rendering the scheme unsuitable for multimedia transmission in practical scenarios. Other related studies such as [6], [7] have been carried out for average power constrained systems without renewable energy sources.

In this paper we study the data scheduling problem in a practical context with finite buffer length and the data transmission constrained to finite adaptive M-ary Quadrature Amplitude Modulation (M-QAM) constellations. The joint state MDP captures the trade-off between the packet drop rate and the power consumed through the Markovian state transitions. Our novel low complexity monotone policy describes an optimal data scheduling for delay and QoS constrained systems with renewable energy sources for long term efficient power utilization. The proposed policy minimizes the average packet drop rate employing insights into the convexity properties of the system transmit power cost function. Results demonstrate the superior performance of the proposed scheme compared to the existing optimal policy for renewable energy sources derived in [1].

This paper is organized as follows. Section II provides an overview of the system and formulates the MDP for the above 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 is modeled as a random process independent across different time slots to capture the variation in the charging phenomenon, similar to [1]. 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 [8] as,

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

where h is the Rayleigh fading channel coefficient, $\Gamma = \frac{N_0}{3}\left[Q^{-1}\left(P_e/4\right)\right]^2$ is a constant dependent on the noise power spectral density N_0 and the QoS based maximum tolerable BER constraint P_e. 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. The action space at time n, $\mathcal{A}(s_n) = \{0, 1, \ldots, a_{max}(s_n)\}$ is restricted by the amount of energy available and the bits present in the buffer. Hence, $P\left(a_{max}(s_n), g_n\right) \leq e_n$ and $a_{max}(s_n) \leq b_n$. 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,

$$r_n\left(b_n, a_n\right) = -\left(b_n - a_n\right), \tag{2}$$

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,

$$R^{\pi}\left(s_n\right) = r\left(s_n, \pi\left(s_n\right)\right) + \beta \sum_{s_{n+1} \in \mathcal{S}} q\left(s_{n+1} | s_n, a_n\right) R^{\pi}\left(s_{n+1}\right), \tag{3}$$

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

$$s_{n+1} = \left(g_{n+1}, e_n - P\left(a_n, g_n\right) + \xi_n, b_{n+1}\right),$$

where g_{n+1} is the uncorrelated exponentially distributed channel gain, ξ_n is an independent random variable denoting the amount of energy recharged, b_{n+1} is an independent data arrival random variable, and a_n is the action chosen at time n. The transition probability $q(s_{n+1}|s_n, a_n)$ can be computed from the probability densities of the random variables g_{n+1}, ξ_n, and b_{n+1} as follows,

$$\begin{aligned} q(s_{n+1}|s_n, a_n) &= q(g_{n+1}) \times q(e_{n+1}|e_n, g_n, a_n) \times q(b_{n+1}) \\ &= q(g_{n+1}) \times q(\xi_n = e_{n+1} - e_n + P(a_n, g_n)) \\ &\quad \times q(b_{n+1}). \end{aligned} \tag{4}$$

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. Here H_m denotes the size of the channel state space, E_m the number of energy units in the battery, and L the data buffer size. Let the maximum constellation size of the M-QAM be 2^M. The action space is limited by maximum energy that can be used in one time epoch, E_{max}. As shown in Section IV, even for small parameter values such as $M = 4$, $E_m = 27$, $E_{max} = 9$, $L = M = 4$, the size of the state space is $|S| \approx 6000$, 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. We denote the channel, energy, and buffer state components by $g, e,$ and b , respectively by dropping the time index. The action is denoted by the variable a.

A. Concavity of Average Reward in energy state, e

The action space is limited by the amount of battery energy, $P(a_{max}(s), g) \leq e$, and the data queue, $a_{max}(s) \leq b$. Hence, increasing the battery energy increases the action space i.e. if $e_1 \leq e_2$ and $\mathcal{A}_1, \mathcal{A}_2$ are the action spaces associated with the states (g, e_1, b) and (g, e_2, b), respectively, then $\mathcal{A}_2 \supseteq \mathcal{A}_1$. Thus, the optimal value of the function $R\left(g, e, b\right)$ must be non-decreasing in e i.e. $R\left(g, e_1, b\right) \leq R\left(g, e_2, b\right)$. Secondly, it can be readily seen from (1) that the function $P(a, g)$ is convex in a. Thus, for two different actions a^1, and a^2 we have,

$$\alpha P\left(a^1, g\right) + (1 - \alpha) P\left(a^2, g\right) \geq P\left(\bar{a}, g\right), \tag{5}$$

where $0 \leq \alpha \leq 1$ and $\bar{a} = \alpha a^1 + (1 - \alpha) a^2$. We now prove the concavity of $R\left(g, e, b\right)$ in e in the following lemma.

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

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

Proof: We prove the above result using induction. For ease of illustration, we replace $R_k(s) = R_k(g, e, b)$ by $R_k(e)$ in this section. In [9], it has been demonstrated that optimal value function $R\left(s\right)$ using value iteration is,

$$R_{k+1}(e) = \max_a \left\{r(b, a) + \beta \times \mathrm{E}_{g', \xi, b'}\left[R_k\left(e - P\left(g, a\right) + \xi\right)\right]\right\}. \tag{6}$$

Under the inductive assumption, let $R_k\left(s\right)$ be concave in e. Now, to prove the lemma above we have to demonstrate,

$$\alpha R_{k+1}\left(e^1\right) + (1 - \alpha) R_{k+1}\left(e^2\right) \leq R_{k+1}\left(\bar{e}\right). \tag{7}$$

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

$$R_{k+1}\left(e^i\right) = r\left(b, a^i\right) + \beta \times \mathrm{E}_{g', \xi, b'}\left[R_k\left(e^i - P\left(g, a^i\right) + \xi\right)\right], \tag{8}$$

where $i = 1, 2$. From the concavity of R_k we have,

$$\begin{aligned} &\alpha R_k \left(e^1 - P\left(g, a^1\right) + \xi\right) + (1-\alpha) R_k \left(e^2 - P\left(g, a^2\right) + \xi\right) \\ &\leq R_k \left(\bar{e} - \alpha P\left(g, a^1\right) - (1-\alpha) P\left(g, a^2\right) + \xi\right) \\ &\leq R_k \left(\bar{e} - P\left(g, \bar{a}\right) + \xi\right), \end{aligned} \tag{9}$$

where the last inequality follows from (5) and the non-decreasing nature of $R(g, e, b)$ in e. Thus, employing (9) and the linearity of the stage reward $r(b, a)$ in a, we have,

$$\begin{aligned} &\alpha R_{k+1}(e^1) + (1-\alpha) R_{k+1}(e^2) \\ &\leq r(b, \bar{a}) + \beta \times \mathrm{E}[R_k(\bar{e} - P(g, \bar{a}) + \xi)] \\ &\leq R_{k+1}(g, \bar{e}, b), \end{aligned}$$

where the last inequality follows from the fact that the best action for R_{k+1} in (6) cannot be worse than $\bar{a}$, thus proving (7). ■

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

B. Monotonicity of the optimal policy in channel state, g

Let the optimal policy in state $s^i = \left(g^i, e, b\right)$ be a^i, $i \in 1, 2$. We need to demostrate $g_1 < g_2 \implies a_1 \leq a_2$. Assume instead, $g^1 < g^2$ and $a^1 > a^2$. Consider the argument of the maximization in (6) for action a, defined for state $s = (g, e, b)$ as,

$$R(s, a) = r(b, a) + \beta \times \mathrm{E}[R(g', e - P(g, a) + \xi, b')]. \tag{10}$$

We have,

$$R\left(s^1, a^1\right) \geq R\left(s^1, a^2\right), \tag{11}$$

$$R\left(s^2, a^2\right) \geq R\left(s^2, a^1\right). \tag{12}$$

Adding inequalities (11) and (12) and employing the relation in (10) we have,

$$E_R\left(p_{1,1}\right) + E_R\left(p_{2,2}\right) \quad \geq \quad E_R\left(p_{1,2}\right) + E_R\left(p_{2,1}\right), \tag{13}$$

where $p_{i,j} = e - P\left(a^i, g^j\right)$, $1 \leq i, j \leq 2$ and $E_R(p) = \mathrm{E}_{g',\xi,b'}\left[R\left(g', p + \xi, b'\right)\right]$. From the definition of $P(a, g)$ in (1) it can be seen that,

$$p_{1,1} + p_{2,2} < p_{1,2} + p_{2,1}, \tag{14}$$

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 (14) 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 (13). Thus, given $g^1 < g^2$ we must have $a^1 \leq a^2$, which completes the proof.

Using similar arguments we can prove the monotonicity of the optimal policy in the energy state e. This proof is omitted for the sake of brevity. We now prove the monotonicity of the optimal policy in b.

C. Monotonicity of the optimal policy in buffer state, b

Consider the states $s^i = \left(g, e, b^i\right)$ for $i = \{1, 2\}$ such that $b^1 < b^2$. Let the optimal actions in s^i be a^i, $i \in \{1, 2\}$. One can show that $b_1 < b_2 \implies \left|\mathcal{A}\left(s^1\right)\right| \subseteq \left|\mathcal{A}\left(s^2\right)\right|$. From (10) it can be observed that the second term in $R\left(s^1, a\right)$ and $R\left(s^2, a\right)$ is identical for all a. The first term $r(b, a)$ differs only by a constant, $\left(b^1 - b^2\right)$. Thus if $a^2 \in \mathcal{A}\left(s^1\right)$ we must have $a^1 = a^2$. Otherwise if $a^2 \in \mathcal{A}\left(s^2\right) - \mathcal{A}\left(s^1\right)$ we necessarily have $a^2 > a^1$. Hence, the optimal action is monotone in b.

Further, we can show that if $a^2 \in \mathcal{A}\left(s^2\right) - \mathcal{A}\left(s^1\right)$ we must have $a^1 = \max \mathcal{A}\left(s^1\right)$. The first term in (10) is linearly increasing in a. From the convexity of $P(a, g)$ and the concavity of $R(g, e, b)$ in e it can be readily seen that the second term is concave decreasing in a. Therefore, if $a^2 \in \mathcal{A}\left(s^2\right) - \mathcal{A}\left(s^1\right)$, restricting the action space to $\mathcal{A}_1$, it follows that $a^1 = \max \mathcal{A}\left(s^1\right)$.

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 $L = M = 4$. The maximum number of energy units is $E_m = 27$. We assume $\mathrm{E}[g] = \mathrm{E}[||h||^2] = 1$ and $\Gamma = 0.19$. The action space is restricted to use a maximum of $E_{max} = 9$ energy units in one time epoch. The data arrival process b_n is assumed to be Poisson distributed with mean λ. 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 $\beta = 0.85$. The energy renewal process ξ_n is considered to be an independent process similar to [1] with an exponential distribution with support $\{0, 1, \ldots, \xi_{max}\}$, $\xi_{max} = 9$, and mean μ.

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,

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

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,

$$H_m + 1 \leq E_{max}(M+1) + 1. \tag{15}$$

The actual value of H_m will be slightly lower than $E_{max}(M + 1) + 1$ due to the overlapping nature of the interval boundaries. Hence, the size of the state space $|S| = (E_m + 1) \times (H_m + 1) \times (L + 1)$ can be approximated using (15) as follows,

$$\begin{aligned} |\mathcal{S}| &\leq (E_m + 1) \times (E_{max}(M+1) + 1) \times (L+1) \\ &= (27+1) \times (9 \times 5 + 1) \times (4+1) = 6440 \end{aligned} \tag{16}$$

Thus, even for moderate values of the MDP parameters M, E_m, L the size of the state space is intractably large. Hence, the monotone properties derived in section III are crucial for the computation of the optimal policy. In Fig. 1(a), 2(a), 3(a) we plot the optimal policy of the above MDP as a function of channel gain g, battery energy e, and data

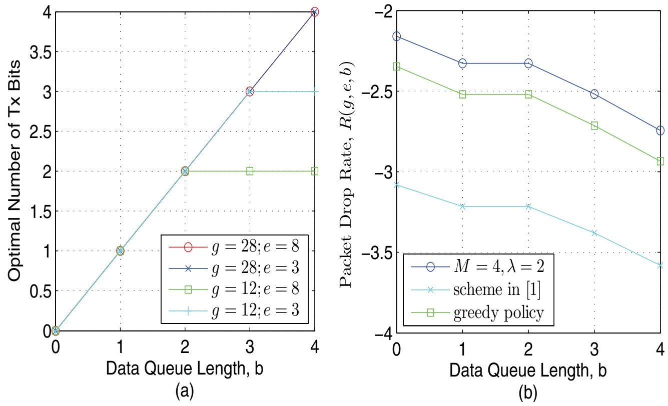

Fig. 1. Optimal number of transmitted bits vs. channel g (left) at $M = 4$ and Packet Drop Rate vs. channel g (right) at $b = 4$ and $e = 8$. The mean of the random variables b_n and ξ_n are $\lambda = 2$ and $\mu = 0.992$, respectively.

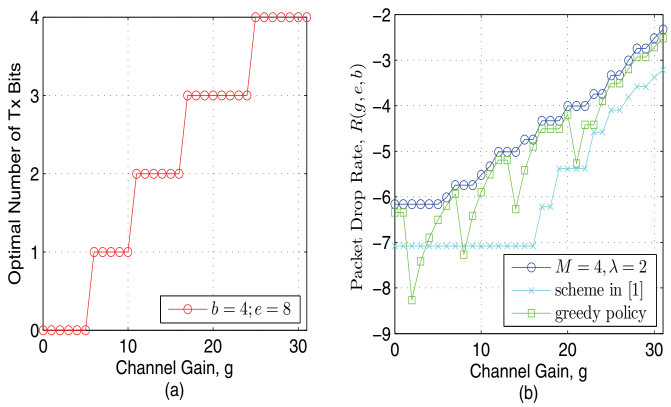

Fig. 2. Optimal number of transmitted bits vs. battery energy e (left) at $M = 4$ and Packet Drop Rate vs. battery energy e (right) at $b = 4$ and $g = 23$. The mean of the random variables b_n and ξ_n are $\lambda = 2$ and $\mu = 0.992$, respectively.

queue length b 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-C can also be readily seen from these plots. Thus, one can employ the monotonicity policy iteration algorithm in [9] (section 6.11.2) to compute the optimal policy. This significantly reduces the computations required to arrive at the optimal adaptive modulation scheduling policy since the size of the optimal action set monotonically decreases with each iteration of the algorithm.

In Fig. 1(b), 2(b), 3(b) 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, and b. This policy is compared with the sub-optimal policy computed using the approach in [1] and the greedy policy which uses maximum allowed battery energy at each time epoch. As expected, the optimal policy surpasses both the policies. Note that for a difference of one unit in $R(g, e, b)$ the policies will drop an average of $1 - \beta = 0.15$ more packets at every decision epoch. Thus, both sub-optimal policies would drop large number of packets even during small time intervals.

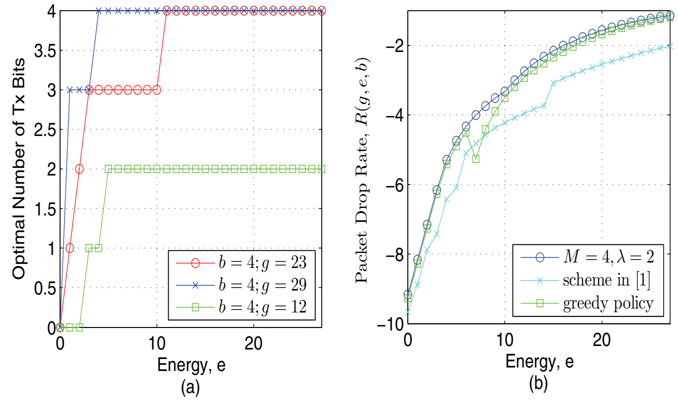

Fig. 3. Optimal number of transmitted bits vs. data queue length b (left) at $M = 4$. Packet Drop Rate vs. queue length b (right) at $g = 28$ and $e = 8$. The mean of the random variables b_n and ξ_n are $\lambda = 2$ and $\mu = 0.992$, respectively.

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 policy iteration for optimal policy computation. Thus, employing the convexity properties of the transmit power function we have derived key monotonicity properties of the optimal transmission policy for the above MDP. These properties significantly reduce the computation required for the optimal policy for long term packet drop minimization. Simulation results have been presented to illustrate the significantly lower long term packet drop rate of the proposed scheme compared to the existing ones.

References

[1] A. Sinha and P. Chaporkar, "Optimal power allocation for a renewable energy source," in *Proc. 2012 National Conference on Communications*, pp. 1–5.

[2] M. H. Ngo and V. Krishnamurthy, "On optimality of monotone channel-aware transmission policies: a constrained Markov decision process approach," in *Proc. 2007 International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 621–624.

[3] N. Salodkar, A. Karandikar, and V. Borkar, "A stable online algorithm for energy-efficient multiuser scheduling," *IEEE Trans. Mobile Comput.*, vol. 9, no. 10, pp. 1391–1406, Oct. 2010.

[4] Z. Ye, A. A. Abouzeid, and J. Ai, "Optimal stochastic policies for distributed data aggregation in wireless sensor networks," *IEEE/ACM Trans. Networking*, vol. 17, no. 5, pp. 1494–1507, Oct. 2009.

[5] B. Xiong, W. Yan, C. Lin, and F. Ren, "Cross-layer optimal policies for stochastic reliable transmission in wireless sensor networks," in *Proc. 2010 International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, pp. 944–951.

[6] A. T. Hoang and M. Motani, "Buffer and channel adaptive modulation for transmission over fading channels," in *Proc. 2003 IEEE International Conference on Communications*, vol. 4, pp. 2748–2752.

[7] A. Hoang and M. Motani, "Buffer and channel adaptive transmission over fading channels with imperfect channel state information," in *Proc. 2004 IEEE Wireless Communications and Networking Conference*, vol. 3, pp. 1891–1896.

[8] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[9] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1st edition. John Wiley & Sons, Inc., 1994.