Optimal and Suboptimal Scheduling over Time Varying Flat Fading Channels

Dejan V. Djonin, Ashok K. Karmokar, and Vijay K. Bhargava

Department of Electrical and Computer Engineering, 2356 Main Mall,
University of British Columbia, Vancouver, BC, V6T 1Z4, Canada.
E-mail: {ddjonin, ashokk, vijayb}@ece.ubc.ca

***Abstract*— This paper explores optimal and suboptimal packet schedulers for time-varying flat fading channels that trade-off between minimization of the average delay and the average transmitted power. Both uncorrelated and correlated block fading channels are investigated. Extending a previous work, we formulate the trade-off as a unconstrained Markov decision processes and find the stationary deterministic optimal policy using both relative value iteration and policy iteration algorithm. As well, we present constrained Markov decision processes formulation of the problem and linear programming algorithm to solve it and show that optimal schedulers are randomized in this case. In order to alleviate the computational complexity needed to determine the optimal scheduling policy we propose a suboptimal log-scheduling policy that has performance close to that of the optimal scheduler. The proposed policy is also robust to different channel models. It is demonstrated that log-policy is favorable to the water-filling policy for very slow fading channels.**

I. Introduction

Modern and future wireless networks will support a numerous services with a wide range of quality of service (QoS), e.g., delay, rate, bit error rate (BER) etc. requirements. The time-varying nature of the channel poses a challenging task of delivering wide variety of services. Various adapting techniques are developed to compensate for the channel's time variation or fading. These methods include adjusting the transmission power, changing the constellation size and coding rate, and varying the spreading gain in code division multiple access (CDMA) based systems (e.g., [1] and references therein). In a wireless network, mobile devices usually rely on a battery with a limited amount of energy. So, wise minimization of the transmission power can lead to more efficient utilization of battery energy and hence longer battery life of the mobile devices. We consider the situation depicted in Fig. 2 where packet arrives from some higher layer application and is placed into a finite transmission buffer. A scheduler removes some of the packets from the buffer periodically, encodes them, and transmits the encoded packets over a fading channel. We assume that the current channel state and the buffer occupancy are available at the transmitter. By delaying packet in the buffer, the transmission power can be saved, but different users may have different QoS (i.e. delay) requirements, and too much delay can result in buffer overflows and packet dropping. In this paper, we study the tradeoff between transmission power and communication delay. We extend the results of [1], where the optimal tradeoff between the average power and the average delay has been studied over a discrete-time, block-fading channel. In our work, we consider the tradeoff over a more realistic correlated fading channel model, namely, finite state Markov channel (FSMC) model. We also consider practical finite-size buffer instead of infinite size buffer. We formulate the optimal power/delay problem as a constrained Markov decision processes (MDP) problem as well and give a linear programming (LP) solution of the problem. We focus on schedulers which do not drop the packets. Sometimes, in practice, the statistical characterization of the channel is not known in advance or computational resources at the transmitter or the receiver are limited. For those cases, we propose a simple suboptimal scheduler called *log-scheduler*.

The remaining part of the paper is organized as follows. In Section II, we describe the channel model for our work. The system model is described in Section III. In Section IV, we formulate the tradeoff between the average power and average delay as an unconstrained MDP (UMDP) and constrained MDP (CMDP) problem. We give a numerical example and some results for the optimal power/delay problem. In Section V, we analyze a simple suboptimal scheduler. We also give simulation results for this scheduler and compare the performance of suboptimal scheduler with that of optimal and mixed scheduler. We conclude with Section VI.

II. FSMC for the Wireless Fading Channel

In a multipath propagation environment, the received signal envelope usually has the Rayleigh distribution. A slowly varying flat fading Rayleigh channel can be represented as a FSMC model as shown in Fig. 1. By partitioning the range of the received channel power gains into a finite number of intervals, a finite state channel model for the Rayleigh fading channel can be built. FSMC model is very useful and popular block fading channel model that considers the correlation of the fades between blocks. Let h be the received instantaneous

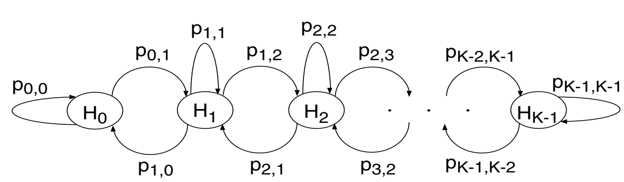

Fig. 1. Finite State Markov Model for the Rayleigh Fading

channel power gain which is proportional to the square of the signal envelope. The pdf of h is exponentially distributed and can be written as

$$p(h) = \frac{1}{h_0} \exp\left(-\frac{h}{h_0}\right) \text{ for } h \geq 0 \tag{1}$$

 0-7803-8533-0/04/$20.00 (c) 2004 IEEE

where $h_0 = \mathbb{E}\{h\}$ is the average channel power gain.

Let the received gains be partitioned into K states and $\mathcal{H} = \{H_0\ H_1\ \cdots\ H_{K-1}\}$ denote the state space of the FSMC. Suppose, $\overrightarrow{\Gamma} = [\Gamma_0\ \Gamma_1\ \cdots\ \Gamma_K]^T$ are the received gain thresholds in increasing order with $\Gamma_0 = 0$ and $\Gamma_K = \infty$. Then the Rayleigh fading channel is said to be in state $H_k,\ k = 0, 1, \cdots, K-1$, if the received gain is in the interval $[\Gamma_k, \Gamma_{k+1})$. Let $p_k,\ i = 0, 1, \cdots, K-1$ denote the steady state probabilities associated with states $H_k,\ k = 0, 1, \cdots, K-1$. We can express p_k as follows,

$$p_k = \int_{\Gamma_k}^{\Gamma_{k+1}} p(h)dh = \exp\left(-\frac{\Gamma_k}{h_0}\right) - \exp\left(-\frac{\Gamma_{k+1}}{h_0}\right). \quad (2)$$

The choice of the gain thresholds is flexible and the partitioning can be done in many ways. We choose the equal probability method for dividing gains as in [2], which is computationally simple and fairly accurate, for our simulations.

We assume that the Rayleigh fading channel is slow enough that the received gain remains at a certain level for the time duration of a block. Furthermore, the channel states associated with consecutive blocks are assumed to be neighboring states. Now, let the transmission rate of system be R_p packets per second, so the average packets per second transmitted during which the channel is in state H_k is $R_p^k = R_p p_k$. Based on the FSMC model proposed in [2], the crossover probabilities can be written as,

$$p_{k,k+1} \approx \frac{N(\Gamma_{k+1})}{R_p^k},\ k = 0, 1, \cdots, K-2 \quad (3)$$

and $$p_{k,k-1} \approx \frac{N(\Gamma_k)}{R_p^k},\ k = 1, 2, \cdots, K-1 \quad (4)$$

where, $N(\Gamma_k),\ k = 1, 2, \cdots, K-1$ is the expected number of times per second the received channel gain passes downward across the corresponding threshold Γ_k and is given by,

$$N(\Gamma_k) = \sqrt{\frac{2\pi\Gamma_k}{h_0}} f_m \exp\left(-\frac{\Gamma_k}{h_0}\right). \quad (5)$$

Here, $f_m = v/\lambda$ is the maximum Doppler frequency, where v is the speed of the mobile station and λ is the wavelength of the radio wave. The transition probability of staying in the same state can be found as: $p_{0,0} = 1 - p_{0,1},\ p_{K-1,K-1} = 1 - p_{K-1,K-2}$ and $p_{k,k} = 1 - p_{k,k-1} - p_{k,k+1},\ k = 1, 2, \cdots, K-2$. We use the average gain for each state for calculating the immediate cost of that state. The average gain for state $H_k,\ k = 0, 1, \cdots, K-1$ are found as,

$$\tilde{h}_k = \frac{\int_{\Gamma_k}^{\Gamma_{k+1}} hp(h)dh}{\int_{\Gamma_k}^{\Gamma_{k+1}} p(h)dh} = h_0 + \frac{\Gamma_k \exp\left(\frac{\Gamma_{k+1}-\Gamma_k}{h_0}\right) - \Gamma_{k+1}}{\exp\left(\frac{\Gamma_{k+1}-\Gamma_k}{h_0}\right) - 1}. \quad (6)$$

III. System Model

We consider a single user system with an finite transmission buffer as shown in the Fig. 2 and a discrete-time model of the buffer where a time slot corresponds to a single block of M channel uses of a block-fading FSMC. Data arrives from the higher layer application and is placed into a finite transmission buffer of length L. Incoming data, as well as data stored in the buffer and data being sent in one block is packetized, where the size of each packet is G bits. Let A_n denote the number of packets arriving at the buffer input between time $(n-1)$ and n. Suppose $\{A_n\}$ form an ergodic Markov chain with state space $\mathcal{A} \subset \mathbb{R}^+$. We assume that $\{A_n\}$ is independent of the channel fading and noise processes. The average packet arrival rate in steady state can be given by $\bar{A} = \mathbb{E}\{A_n\}$. Assume that at the beginning of the nth block the transmitter takes U_n packets from the buffer and maps these into a rate GU_n/M codeword which will be transmitted over the next M channel uses. We will assume fixed-length/variable-rate codewords, i.e., all codewords are sent over the same number of channel uses, but the number of possible codewords can vary. All the packet transmitted in a block experience the same channel gain.

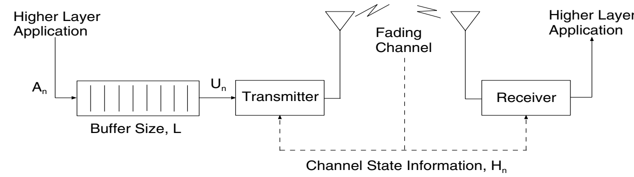

Fig. 2. Schematic of System Model

Let $\mathcal{B} = \{B_0\ B_1\ \cdots\ B_L\}$ denotes state space of the buffer's packet occupancy. The dynamics of the buffer in terms of packet occupancy are then given by

$$B_{n+1} = \min\{B_n - U_n + A_{n+1}, B_L\}. \quad (7)$$

We assume that the transmitter can choose U_n based on the buffer state B_n, the channel gain H_n, and the source state A_n. All packets that arrive in time-slot n can be transmitted only in time-slot $(n+1)$ or later. A natural constraint on U_n is that $0 \le U_n \le B_n$ for all n i.e. no transmission when the buffer is empty and we cannot transmit more packets than available in the queue.

Let $P(h, u)$ be the required transmission power during the block when the channel gain is h and the transmitter chooses to transmit u packets and can be given by,

$$P(h, u) = \frac{\sigma^2}{h}\left(2^{Gu/M} - 1\right). \quad (8)$$

For all h with $|h| > 0$, $P(h, u)$ is an increasing and strictly convex function of $u \ge 0$. The average total delay experienced by a packet in the system in Fig. 2 is the sum of the delay in the buffer plus the time from when a packet leaves the buffer until it is decoded. The average packet delay is related to the average buffer occupancy via Little's theorem as follows,

$$\bar{D} = \frac{1}{\bar{A}}\mathbb{E}\{B_n\}. \quad (9)$$

Let $\mathcal{U}$ be the set of all possible transmission rate with U elements, i.e., $\mathcal{U} = \{u^1, u^2, \cdots, u^U\}$. Also, let μ be a stationary policy such that $\mu : \mathcal{B} \times \mathcal{H} \times \mathcal{A} \mapsto \mathcal{U}$. The expected long-term average power and average delay with policy μ are respectively,

$$\bar{P}(\mu)=\lim_{m\to\infty}\frac{1}{m}\sum_{n=1}^{m}\mathbb{E}\left\{P\left(H_n,\mu(B_n,H_n,A_n)\right)\right\} \quad (10)$$

$$\text{and} \quad \bar{D}(\mu)=\lim_{m\to\infty}\frac{1}{m}\sum_{n=1}^{m}\frac{\mathbb{E}\left\{B_n\right\}}{\bar{A}}. \quad (11)$$

Here, the expectation operator is over the random fading gains.

IV. Optimal Scheduling

A. Unconstrained MDP Formulation

We are interested in two objectives, minimizing the average delay and minimizing the average power. Since the two goals are conflicting, we are interested in the trade-off between minimizing the average power and the average delay. We consider minimizing a weighted combination of the two criteria and formulate the problem as an average cost UMDP problem with state space $\mathcal{S}=\mathcal{B}\times\mathcal{H}\times\mathcal{A}$. Therefore, for $\beta>0$, we seek to find the policy μ which minimizes

$$J(\mu)=\lim_{m\to\infty}\frac{1}{m}\sum_{n=1}^{m}\mathbb{E}\left\{P\left(H_n,\mu(B_n,H_n,A_n)\right)+\beta\frac{B_n}{\bar{A}}\right\}. \quad (12)$$

The weighting factor β indicates the relative importance of the average buffer delay over the average power; smaller values of β correspond to placing less importance on delay and vice versa. At each time slot, the transmitter chooses a control action, namely, the transmission rate, and incurs a per-stage cost of $\left(P\left(H_n,\mu(B_n,H_n,A_n)\right)+\beta\frac{B_n}{\bar{A}}\right)$. For constant incoming traffic $\bar{A}$, the problem at hand forms at weakly communicating MDP and there always exists a stationary policy μ which is optimal and satisfies Bellman's equation for the average cost problem. Such MDP problems can be solved via dynamic programming techniques, e.g., relative value iteration, or policy iteration algorithm given in [3]. Let, for a given value of β, μ^* be an optimal policy and $\bar{P}(\mu^*)$ and $\bar{D}(\mu^*)$ be the corresponding average power and average delay, as given in (10) and (11). Here, $\bar{P}(\mu^*)$must be the minimum average power such that the average delay is less than $\bar{D}(\mu^*)$. By varying β and finding the deterministic optimal policy the corresponding average power and average delay provide different points on the optimum power/delay curve. Points above optimal power/delay curve are achievable with certain scheduling, while there exists no scheduler that can have power/delay performance below that curve.

B. Constrained MDP Formulation

In this section, we formulate the problem as CMDP, where we bound the average delay to a specific value and seek to find the optimal stationary policy μ^* that minimizes the average power as,

$$\min_{\mu} J_P(\mu)=\lim_{m\to\infty}\frac{1}{m}\sum_{n=1}^{m}\mathbb{E}\left\{P\left(H_n,\mu_n(B_n,H_n,A_n)\right)\right\}$$
$$\text{subject to: } \lim_{m\to\infty}\frac{1}{m}\sum_{n=1}^{m}\frac{\mathbb{E}\left\{B_n\right\}}{\bar{A}}<\tilde{D} \quad (13)$$

where, $\tilde{D}$ is maximum tolerable average delay. The problem can also be formulated by bounding the average power to a specific value and finding the optimal policy that minimizes the average delay. The above CMDP problem can be solved using equivalent LP methodology as presented below [3]. It can be shown that there is an one to one correspondence between feasible (optimal) solution of the LP and the feasible (optimal) solution of CMDP [4]. Let $\nu(s,u)$ represents the "steady-state" probability that the process is in state $s\in\mathcal{S}$ and action $u\in\mathcal{U}$ is applied. We seek to find the optimal control policy which is represented in terms of probability distribution ν over $\mathcal{S}\times\mathcal{U}$ by solving the linear program,

$$\begin{aligned}
&\min_{\nu} \sum_{s\in\mathcal{S},u\in\mathcal{U}} c(s,u)\nu(s,u)\\
\text{subject to:}\quad &\sum_{s\in\mathcal{S},u\in\mathcal{U}} d(s)\nu(s,u)\le\tilde{D}\\
&\sum_{u\in\mathcal{U}}\nu(t,u)=\sum_{s\in\mathcal{S},u\in\mathcal{U}}\nu(s,u)p_{st}(u),\ \forall t\in\mathcal{S}\\
&\sum_{s\in\mathcal{S},u\in\mathcal{U}}\nu(s,u)=1\\
&\nu(s,u)\ge 0,\quad \forall s\in\mathcal{S},\forall u\in\mathcal{U}
\end{aligned} \quad (14)$$

where $c(s,u)=P(h,u)$ and is given by (8), $d(s)$ for the composite state $s\in\mathcal{S}$ returns the number of packets in the queue divided by average arrival rate $\bar{A}$. The optimal randomized scheduler μ^* for the CMDP problem is uniquely characterized with probability $\theta_{\mu^*(s)}(u)$ of applying policy $u\in\mathcal{U}$ in state $s\in\mathcal{S}$. As opposed to the general randomized scheduler, in the case of deterministic scheduler, the probabilities $\theta_{\mu^*(s)}(u)$ take on only values 0 or 1. This formulation provides a simple tool for determination of the optimal randomized scheduling policies. Suppose there exists a optimal solution ν^* to the LP problem. Then there exists an optimal policy μ^* for the CMDP problem, where μ^* satisfies

$$\theta_{\mu^*(s)}(u)=\frac{\nu^*(s,u)}{\sum_{u'\in\mathcal{U}}\nu^*(s,u')} \text{ if} \sum_{u'\in\mathcal{U}}\nu^*(s,u')>0. \quad (15)$$

If $\sum_{u'\in\mathcal{U}}\nu^*(s,u')=0$ for some $s\in\mathcal{S}$, an action that drives the system to $\mathcal{S}_{\nu^*}=\{s\in\mathcal{S}:\sum_{u'\in\mathcal{U}}\nu^*(s,u')>0\}$ is chosen in each state [3]. It is seen from the simulations that the number of states within the recurrent class where randomized policy is applied is not greater than one. This resembles the property of recurrent constrained MDP as shown in Theorem 4.4 of [4]. Provided that the dimensionality of the problem is not too large, linear programs like the previous one can be easily solved using standard mathematical packages (e.g., MATLAB Optimization Toolbox).

C. Results

In this section, we show the results on the delay/power trade-off of optimal schedulers under different number of actions and fading rates, namely f_m = 50Hz, 100Hz, 500Hz and i.i.d. The model under investigation has constant incoming traffic of $\bar{A} = 2$ packets per block. Each block has $M = 3G$ channel uses. The average code rate for this case is $2/3$. Noise power $\sigma^2 = 1mW$ and average channel power gain $h_0 = 0.1$. The packet rate $R_p = 10^4$ packets/sec. In the simulations, the buffer size, $L = 100$ packets while the number of channel states $K = 5$. In order to avoid overflows and dropped packets in the buffer, the immediate cost function has been modified. In our simulations we have applied very high (theoretically infinite) penalty cost in the case when the application of a certain policy in a certain state would lead to buffer overflows and dropped packets. The problem of avoiding the possibility of sending more packets than the buffer currently holds is addressed in a similar manner. Fig. 3 demonstrates the

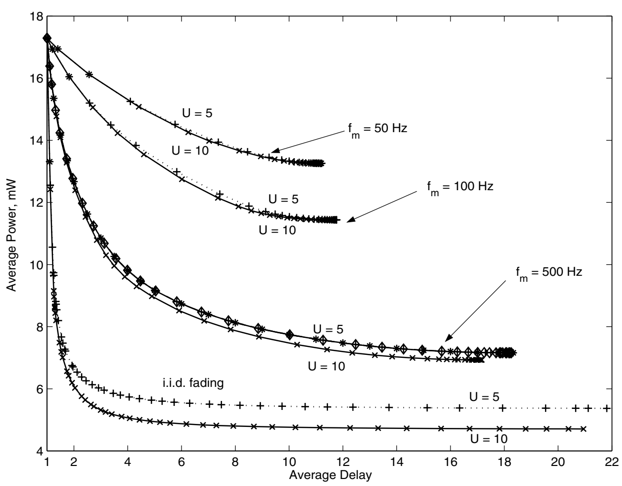

Fig. 3. Influence of fading rates and number of actions on power/delay tradeoffs.

influence of the Doppler frequency of the correlated fading channel as well as the number of policies on the optimal power/delay curve. The figure also shows that relative value iteration algorithm (+ marked curve for $f_m = 500$) and policy iteration algorithm ($\diamond$ marked curve) for unconstrained formulation and linear programming solution for constrained formulation (* marked curve) give the same plot. It can be seen that for very small delays, the minimal necessary powers for all analyzed cases approach the same value. This case corresponds to the delay-limited capacity of the channel [5].[1] The minimal necessary power decreases with the increase of the delay and it will eventually reach the level predicted by the water-filling power allocation policy (see e.g. [1]). It is also seen that there exists a maximum delay (e.g., for $f_m = 500Hz$ and $U = 5$, maximum average delay=18.3323 slots) after that the increase in allowable delay doesn't decrease the power. The rate of decrease of the necessary power is decreasing with the decrease of the Doppler frequency. Note that the decrease in the number of available rate policies from $U = 10$ to $U = 5$ does not significantly degrade the optimal power/delay curve.

[1] Note that the theoretical delay-limited capacity for Rayleigh fading channel is 0 and necessary power for unit delay would be infinite. However, the finite power in our case comes from the fact that we used approximate Rayleigh fading channel with only 5 states.

V. Suboptimal Scheduling

In certain cases of practical interest it is not possible to use the optimal scheduling policy explained in the previous section. This may occur when computational resources at the transmitter or the receiver are limited or when the statistical characterization of the channel is not known in advance. We present here a simple approximate policy that we will call *log-scheduling*. This policy does not depend on the statistics of the fading channel. As can be seen from (8), the power is increasing exponentially with the transmission rate. The purpose of the log-scheduling is to limit the transmitted power and make it linearly proportional to the buffer occupancy. The second guideline that influences this choice of scheduling is our goal to have the policy that ensures equal powers distributed across all channel states. This power allocation strategy allows us to balance the powers in different channel states for the same buffer occupancies. Furthermore the short-term average power is balanced even for relatively slow fading rates when there is not enough mixing between the states of the FSMC model. The previous guidelines can be accommodated with the following policy

$$U_n(B_n, H_n) = \lfloor \log_2(\tau B_n \tilde{h}_n) \rfloor \tag{16}$$

where, $\lfloor x \rfloor$ denotes the floor function of x. The coefficient τ determines how aggressive the power allocation policy is going to be. Larger value of τ assigns higher rates for all states and the policy resembles channel inversion power control. Smaller value of τ implies that weaker states will choose smaller transmission rates while stronger states will be assigned higher transmission rates and the policy resembles water-filling power control. Substituting the previous equation in (8) it can be noticed that the power is linearly proportional to the buffer occupancy for all channel states. This ensures that there is no excessive power dissipation for any channel or buffer state. In practical implementations the scheduling policy of (16) has to be slightly modified. For constant incoming traffic the scheduling policy can be as follows

$$\begin{aligned} & U_n(B_n, H_n) \qquad (17) \\ = \;& \max\left(\min\left(B_n, \lfloor \log_2(\tau B_n \tilde{h}_n) \rfloor, u^U\right), B_n - (L - \bar{A})\right) \end{aligned}$$

where the inner minimization operation ensures that the number of the transmitted bits at time slot n is not greater than the buffer occupancy B_n or the highest possible rate. The outer maximization ensures that there are no buffer overflows for the constant incoming traffic $\bar{A}$. In case of burst sources, $\bar{A}$ should be substituted with the largest number of incoming packets that can arrive within one block. The simulation results for log-scheduling are given in Fig. 4. It can be

 0-7803-8533-0/04/$20.00 (c) 2004 IEEE

observed that power/delay curve of log-schedulers follows the same decreasing pattern as the respective curves of optimal schedulers. However the log-schedulers suffers from power loss of 10% - 20% for longer delays.

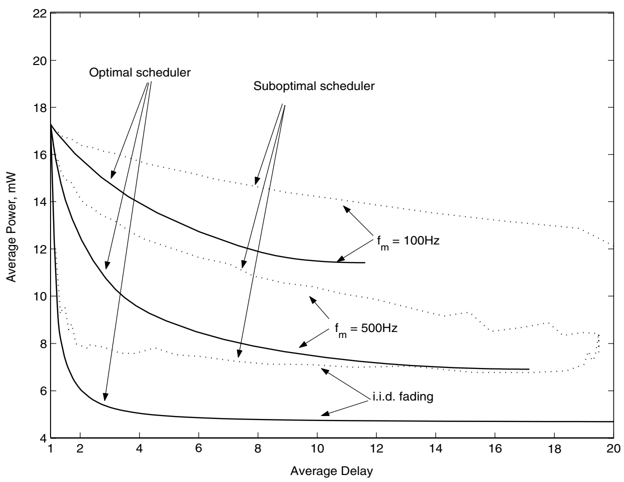

Fig. 4. Comparison of Optimal vs. Log-scheduling

We have also compared the optimal and sub-optimal schedulers with the information-theoretical results of channel inversion and water-filling power control policy. Let deterministic policy μ_C be defined as $U(B_n, H_n) = \min(B_n, u^U)$ i.e. maximum possible number of packets are always sent from the buffer irrespective of the channel state. This policy resembles the channel inversion power allocation policy [6]. Furthermore, let the deterministic policy μ_W be defined as

$$U_n(B_n, H_n) = \max\left(\min\left(B_n, U_W(\tilde{h}_n)\right), B_n - (L - \bar{A})\right)$$

where $U_W(\tilde{h}_n)$ is the opportunistic rate allocation policy dependent only on the current channel state that corresponds to the water-filling power allocation policy for the given fading statistics and capacity [6]. Maximization and the minimization in (18) modify the opportunistic rate allocation $U_W(\tilde{h})$ so as to ensure that there are no dropped packets in the buffer and the number of packets sent from the buffer does not exceed its occupancy. We now form the *mixed scheduling* policy μ_M, by randomizing the previous two policies. This policy μ_M applies scheduling according to the policy μ_W with probability α and scheduling according to the policy μ_C with probability $(1-\alpha)$, for some $\alpha \in [0, 1]$. The influence of the coefficient α on the power/delay curve for mixed scheduling is shown in Fig. 5. For $\alpha = 0$ mixed scheduling is equivalent to deterministic channel inversion power control (with unit delay), while for $\alpha = 1$ mixed scheduling is equivalent to water-filling power control. It can be easily observed that this mixed policy behaves worse than log-scheduling for almost all values of α. As the value of α increases, the average power first start increasing and then after crossing some value of α, it again start decreasing. We may explain this fact in the following way: when the fading gain is in very low state of the channel state space $\mathcal{H}$, the water filling allocation allows the packet to accumulate in the buffer, but channel inversion allocation sends these packet with very high cost. It is also seen that the average power required decreases as the fading speed increases. This is due to the fact that as the fading speed increases, there is less chance of fading state to remain in the bad state for a long time and consequently less chance to accumulate the packet in the buffer for a long time. This prevents the average buffer size to become very big for slow fading. It is interesting that for slower fading rates and $\alpha = 1$ i.e. if pure water-filling is applied, the necessary power is larger for this policy than for log-scheduler. Therefore, we can conclude that straightforward application of water-filling power allocation is not suitable for practical systems where buffer dynamics as well as delay is taken into account.

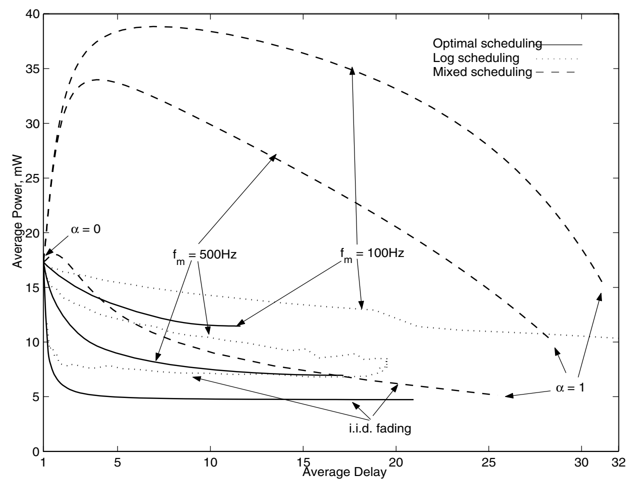

Fig. 5. Comparison of Optimal vs. Policy mixing

VI. Conclusions

We have presented unconstrained and constrained MDP formulations that can produce optimal power and rate allocation policies that minimize the average transmission power and average delay over a flat fading channel for given constant incoming traffic and given channel fade statistics. Suboptimal rate and power allocation policy is also proposed and it is shown that it has performance close to that of the optimal schedulers especially for the smaller delays.

References

[1] R. A. Berry, and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inform. Theory*, vol. 48, No. 5, pp. 1135-1149, May 2002.

[2] H. S. Wang, and N. Moayeri, "Finite-state Markov channel- a useful model for radio communication channels," *IEEE Trans. on Veh. Technol.*, vol. 44, No. 1, pp. 163-171, Feb. 1995.

[3] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: John Wiley & Sons, 1994.

[4] E. Altman, *Constrained Markov Decision Processes: Stochastic Modeling*. London: Chapman and Hall/CRC, 1999.

[5] S. Hanly, and D. Tse, "Multi-access fading channels:delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, No. 7, pp. 2816-2831, Nov. 1998.

[6] A. J. Goldsmith, and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, No. 6, pp. 1986-1992, Nov. 1997.