# Delay-aware Power Adaptation for Incremental Redundancy Hybrid ARQ over Fading Channels with Memory

Ashok K. Karmokar, Dejan V. Djonin, and Vijay K. Bhargava
Department of Electrical and Computer Engineering, 2356 Main Mall
University of British Columbia, Vancouver, BC, V6T 1Z4, Canada.
E-mails: {ashokk,ddjonin,vijayb}@ece.ubc.ca

***Abstract*— We investigate transmit power adaptation strategy for IR-HARQ scheme over correlated fading channels. In order to jointly analyze physical and link layers, the transmitter model incorporates a finite-size buffer that receives randomly varying traffic from a higher layer application. It is assumed that channel variations can be modeled with first order Markov chain. We show that the optimal power adaptation law under buffer delay and packet overflow constraints can be obtained using the framework of semi-Markov decision processes. The optimal policy is computed using the linear programming approach. Simulation results show that substantial power savings can be achieved if the transmission delay requirements are relaxed.**

## I. Introduction

Incremental redundancy hybrid automatic repeat request (IR-HARQ) is an integral part of EDGE standard and is also proposed as a part of 3G evolution cellular system standards, such as W-CDMA high speed downlink packet access (HS-DPA) for high speed reliable packet data communications. IR-HARQ employs forward error correction (FEC) technique in physical layer as well as automatic repeat request (ARQ) technique in data link layer to cope with the time-varying fading channels, and to guarantee both the high reliability and the high throughput. In this scheme, information packets are first transmitted with no or few parity bits for error detection and correction. Incremental redundancy bits are transmitted upon retransmission request. The receiver combines the transmitted and retransmitted bits together to form a more powerful error correction code to recover the information. Rate compatible punctured convolutional (RCPC) codes proposed in [1] are particularly useful for IR-HARQ systems. RCPC codes are constructed from a single rate $1/N$ convolutional code, wherein a family of higher rate codes is formulated by puncturing successively greater numbers of code symbols. These codes have practical utility in that the system requires a single rate $1/N$ convolutional encoder and a Viterbi decoder. In [1], a truncated IR-HARQ scheme with RCPC is analyzed over an AWGN channel and an ideally interleaved Rayleigh fading channel, assuming independent decoding attempts.

Packet errors occur in wireless channels due to various propagation artifacts, such as multi-path and fading are correlated. To deal with correlated errors, a first order Markov model with a finite number of states is considered in [2]. This model is widely used for packet communications due to combination of its analytical tractability and accuracy. Generalized type-II HARQ using RCPC which combines the IR-HARQ strategy of Hagenauer with the code-combining ARQ strategy of Chase is analyzed in [3]. In [4], an IR-HARQ with selective combining is investigated over a finite state Markov channel (FSMC) model. Coding rate adaptation for type-I HARQ system over a partially observable correlated Rayleigh fading channel has been studied in [5]. In [6], a type-II HARQ system with a finite size receiver buffer is analyzed over a two-state Markov channel using rate 1/2 convolutional code and truncated HARQ.

We adapt the transmission power of an IR-HARQ system based on both the channel state and the buffer state to minimize three goals: transmission power, delay and overflow. Due to dynamic nature of optimization criteria, the problem falls under the purview of *stochastic dynamic programming* methods. Further, due to stochastic nature of the duration of successive decision-epochs and dependence of the costs on the decision-epoch duration, the optimization problem is formulated as a semi-Markov decision process (SMDP) problem. The optimal solution is found by converting the problem into an equivalent auxiliary discrete-time Markov decision process (DT-MDP) problem with the linear programming solution technique. To our best knowledge, this is the first work that analyzes SMDP-based cross-layer power adaptation law under latency and overflow constraints for IR-HARQ systems.[1]

The remaining part of the paper is organized as follows. In Section II, we describe the system model including incoming traffic, buffer and channel models used in our paper. The observation probability for RCPC scheme is also described in this section. We explain the formulation of the cross-layer adaptation problem as SMDP in Section III. In Section IV, the equivalent discrete-time formulation of the SMDP problem and its solution techniques are given. We give simulation results in Section V and conclude the paper in Section VI.

## II. IR-HARQ Modeling

Consider a type-II IR-HARQ system using RCPC with single-transmit single-receive antenna in Fig. 1. The transmitter is equipped with a finite buffer that can accommodate $B$ packets. Discrete-time representation of relevant variables is adopted and duration of a discrete time-slot is $T_B$. Unless

[0]This work is supported jointly by the Natural Sciences and Engineering Research Council (NSERC) of Canada under a strategic project grant, UBC Graduate Fellowship, and NSERC Postdoctoral Fellowship.

[1]The methodology described in this paper can equally be applied to any HARQ scheme with packet (e.g. Chase) combining at the receiver.



1-4244-0355-3/06/$20.00 (c) 2006 IEEE

*This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.*

otherwise specified, we use superscript $k$ to denote the value of a particular variable at $k^{\text{th}}$ time-slot. We assume that the feedback channel is noiseless and the channel condition is known at the transmitter perfectly. Let $a^k$ denotes the number of incoming packets at the buffer in time-slot $k$. We assume that the incoming traffic is non-constant, independent and identically distributed (i.i.d). Let $P^a(a_i)$ be the probability of $a_i$ packet arrivals and $A$ be the maximum number of packet arrivals per time-slot. In particular, for Poisson distributed traffic, the probability of $a_i$ arrivals in time-slot $k$ can be given by $P^a(a^k = a_i) = \exp(-\lambda T_B)\frac{(\lambda T_B)^i}{i!},\ i = 0, \cdots, A$, where $P^a(a^k = a_A) \to 0$.

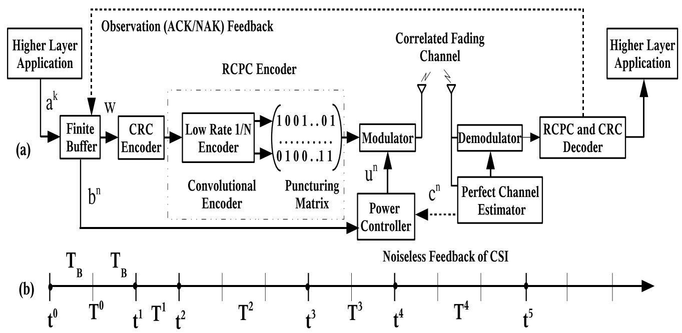


Fig. 1. (a) System diagram of the incremental redundancy type-II hybrid ARQ system and (b) Typical sample path for a SMDP.

Assume that at the start of decision-epoch $n \in \mathbb{Z}^* = \{0, 1, \cdots\}$, the scheduler chooses a particular action $u^n$ to transmit $w$ packets from the buffer. The buffer occupancy $b^n$ and channel condition $c^n$ in decision-epoch $n$ determines the choice of action $u^n$. Each decision-epoch consists of 1 upto $R+1$ time-slots, where $R$ is the number of retransmissions. The duration of the decision-epoch is a random variable and depends on the decoding results. The control action is taken at the start of a decision-epoch and the same action is continued until the end of that decision-epoch. We denote decision-epoch by superscript $n$. Note that there is no transmission when the buffer has insufficient number of packets. Let $\mathcal{B} = \{b_0, b_1, \cdots, b_B\}$ denote the buffer state space in terms of packet occupancy, where $b_i$ corresponds to $i$ packets in the buffer. The buffer dynamics with $a^n = a^k$ is given by,

$$b^{n+1} = b^n - \alpha w + a^k + \cdots + a^{k+r} \tag{1}$$

where multiplier $\alpha$ has value of 1 for positive acknowledgment (ACK) and 0 for negative acknowledgment (NAK) feedback in the last retransmission. In (1), $a^n = a^k$ is to be assumed and $r$ is the random variable that characterizes the number of retransmission. Note that the ACK and NAK feedback from the receiver that reflect the decoding result constitute the observation for the IR-HARQ scheme.

### A. Rayleigh Fading Finite State Markov Channel

The wireless channel in the analyzed IR-HARQ system is assumed to be flat fading obeying Rayleigh distribution. The probability density function of power gain for Rayleigh fading channel is described with exponential distribution,

$$p(\gamma) = \frac{1}{\bar{\gamma}}\exp(-\frac{\gamma}{\bar{\gamma}}), \text{ for } \gamma \geq 0, \tag{2}$$

where $\bar{\gamma} = \mathbb{E}\{\gamma\}$ is the average received channel power gain. A finite state Markov channel is described with a set of channel states $\mathcal{C} = \{c_1, \cdots, c_C\}$ and a matrix of transition probabilities among states $P^c = [P_{c_i,c_j},\ 1 \leq i,j \leq C]$, where $C$ is the number of finite non-overlapping channel states and $P_{c_i,c_j}$ is the probability of transition from state $c_i$ to state $c_j$. The partitioning of channel states can be done in several ways, e.g., in equal probability method, the stationary probability of all states are assumed to be equal. That is, $P_i = 1/C,\ \forall c_i \in \mathcal{C}$ [2]. Let $\Gamma = \{\gamma_0, \cdots, \gamma_C\}$ denotes the set of gain partition of the channel, where $\gamma_0 = 0$, $\gamma_i < \gamma_{i+1}$, and $\gamma_C = \infty$. The cross-over transition probabilities can be expressed in terms of block rate, $R_B$ as,

$$P_{c_i,c_{i+1}} = \frac{N_i}{R_B^i}, \text{ and } P_{c_i,c_{i-1}} = \frac{N_{i-1}}{R_B^i},\ \forall c_i \in \mathcal{C}, \tag{3}$$

where, $N_i = \sqrt{\frac{2\pi\gamma_i}{\bar{\gamma}}} f_m \exp\left(-\frac{\gamma_i}{\bar{\gamma}}\right)$ is the level crossing rate at the threshold $\gamma_i$ and $R_B^i = R_B P_i = P_i/T_B$ is the average number of block per second in channel state $c_i$. Maximum Doppler frequency $f_m = v/\lambda_{rw}$, where $v$ and $\lambda_{rw}$ are the mobile station's speed and wavelength of the radio wave. When the fading rate of the channel is slow, the non-adjacent transition probabilities can be assumed to be zero, i.e., $P_{c_i,c_j} = 0$ if $|j-i| > 1$. Self transition probabilities can be found using the property that sum of all outgoing transition probabilities is equal to 1. In a special case when $C = 2$, the model can be described as Gilbert-Elliot channel containing a "bad" state (state $c_1$) and "good" state (state $c_2$). The steady-state probabilities of such two state channel can be given by $P_i = \int_{\gamma_{i-1}}^{\gamma_i} p(\gamma)d\gamma = \frac{1-P_{c_{3-i},c_{3-i}}}{2-(P_{c_1,c_1}+P_{c_2,c_2})},\ i = 1, 2$.

### B. RCPC Code and Observation Probability

Let $C_1 > C_2 > \cdots > C_M$ denote the $M$ rates offered by a family of RCPC codes which are obtained from a low rate $C_M = 1/N$ code. Parity check bits $m_{crc}$ for error detection and tail bits $m_{tb}$ to properly terminate the encoder memory and decoder trellis are appended with $m_{ib} = wG$ information bits that correspond to $w$ packets taken from the buffer, where $G$ is the size of each incoming data packet in bits. Total $m_T = m_{ib} + m_{crc} + m_{tb}$ bits are encoded with the original mother code of rate $C_M$ encoder. In the first transmission, code bits in the starting code $C_1$ are sent to the receiver and a Viterbi decoder is used for error correction followed by cyclic redundancy check (CRC) error detection. If errors are detected, the receiver sends a NAK. The incremental redundancy bits yielding code $C_2$ from code $C_1$, which were deleted by puncturing process, are then transmitted and decoding is performed using code $C_2$ by combining the first and second sets of transmitted bits. This process is continued until decoding process results in no errors being detected or the maximum number of retransmissions is reached. In either



This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

case, the buffer occupancy is updated and the whole process is started in the next decision-epoch. As in [4], we assume that the first transmission and each subsequent retransmission is of equal duration $T_B$. The upper bound for the first error event probability of the Viterbi decoding algorithm with code $C_r$, given channel states $\{c^1, \cdots, c^r\}$, is expressed as [4],

$$P_E(C_r|c^1,\cdots,c^r) \leq \sum_{d=d_{free}^{(r)}}^{\infty} \underbrace{\sum_{d_1}\cdots\sum_{d_r}}_{d_1+\cdots+d_r=d} a_d^{(r)}(d_1,\cdots,d_r) \times P_d(d_1,\cdots,d_r|c^1,\cdots,c^r). \quad (4)$$

In (4), $d_{free}^{(r)}$ is the free distance of code $C_r$, and $a_d^{(r)}(d_1,\cdots,d_r)$ is the number of paths of weight $d=(d_1+\cdots+d_r)$, where $d_1$ is the distance contribution of code $C_1$, $d_2$ is the distance contribution of the added bits to $C_1$ yielding code $C_2$ and so on. The term $P_d(d_1,\cdots,d_r|c^1,\cdots,c^r)$ is the probability that a wrong path at distance $d$ from the correct path is selected and, for hard decision decoding, it can be given by,

$$P_d(d_1,\cdots,d_r|c^1,\cdots,c^r) = \sum_{e_1}\binom{d_1}{e_1}\epsilon_1^{e_1}(1-\epsilon_1)^{(d_1-e_1)} \times\cdots\times\sum_{e_r}\binom{d_r}{e_r}\epsilon_r^{e_r}(1-\epsilon_r)^{(d_r-e_r)} \quad (5)$$

where $\epsilon_i$, $i=1,\cdots,r$ is the bit error rate (BER) for channel state $c^i$. For BPSK transmission, it can be calculated approximately as, $\epsilon_i \approx \frac{1}{2}\text{erfc}(\sqrt{\frac{\tilde{\gamma}_i P_t}{\sigma^2}})$, where $\tilde{\gamma}_i$ is the average channel gain of state $c^i$ and can be obtained as $\tilde{\gamma}_i = \frac{1}{P_i}\int_{\gamma_{i-1}}^{\gamma_i}\gamma p(\gamma)d\gamma$. In (5), $e_i$ is the number of bits received in error among $d_i$ bits in their specific positions transmitted in channel state $c^i$. The $d_i$ bit locations are determined by the specific path through the decoding trellis. The summation of $e_i$ satisfy the inequality,

$$\sum_{i=1}^{r} e_i \geq \sum_{i=1}^{r}\frac{d_i}{2}. \quad (6)$$

Note that, like [4], our scheme is also holds for soft decision decoding. The probability of decoding failure (NAK probability) with code $C_r$ when action $u$ is taken can be given by,

$$\begin{aligned} P_N(c^1,\cdots,c^r,u) &= 1 - P_A(c^1,\cdots,c^r,u) \\ &= 1-(1-P_E(C_r|c^1,\cdots,c^r))^l. \end{aligned} \quad (7)$$

In (7), $l$ is the number of stages in a trellis for decoding $(m_{ib}+m_{crc}+m_{tb})$ bits with code $C_r = \frac{V}{V+L_r}$ and $l = \frac{m_{ib}+m_{crc}+m_{tb}}{V}$, where $L_r$ can be varied from 1 to $(N-1)V$ and $V$ is the puncturing period [4]. The term $P_A(c^1,\cdots,c^r,u)$ denotes the probability of decoding success (ACK probability) with code $C_r$ and action $u$. Here, we assumed that all errors can be detected by the error detection code. Note that for the case of $C_1$, the probability of decoding failure or NAK probability, $P_N(c^1,u)$, is given by [4],

$$P_N(c^1,u) = 1-(1-\epsilon_{c^1})^{m_{ib}+m_{crc}+m_{tb}}. \quad (8)$$

## III. SMDP Formulation of the Scheduling Problem

We have discussed in Section II that the time between successive control choices is variable and depends on the current state and the choice of action. The cost per decision-epoch depends on the time required for transition from one state to the next. Therefore, the problem at hand forms a semi-Markov decision process problem ([7], Section 5.3). The SMDP problem can be modeled through a tuple $\{\mathcal{S},\mathcal{U},\mathcal{F},\mathcal{Q},\mathcal{G}\}$. $\mathcal{S}=\{s_1,\cdots,s_S\}$ is system state space that is composed of buffer state and channel state. $\mathcal{U}=\{u_1,\cdots,u_U\}$ is finite action space, where each action corresponds to a power level of the transmitter. We denote by $\mathcal{U}_s \subset \mathcal{U}$ those actions that are available at state $s \in \mathcal{S}$. The choice of actions in a state is determined by a policy. In general, the policy $\pi$ in the policy space $\Pi$ can be described as $\pi = \{\mu^1,\mu^2,\cdots\}$, where action $u^n = \mu^n(s^n)$ is applied at decision-epoch $n$. A stationary policy does not vary with decision-epoch $n$ and can be denoted as $\pi=\{\mu,\mu,\cdots\}$. $\mathcal{F}$ is the set of distribution function of sojourn times. Sojourn time for decision-epoch $n$, $T^n = t^{n+1}-t^n$ represents the time spent in a particular state before moving to next state, where $t^n$ is the time of occurrence of the start of the $n^{th}$ decision-epoch (see Fig. 1(b)). $\mathcal{Q}$ is the set of transition distributions $Q_{s_i,s_j}(\tau,u_i)$, where $Q_{s_i,s_j}(\tau,u_i)$ represents the probability of moving from state $s_i$ to state $s_j$ at or before time $\tau$ if action $u_i$ is chosen. $\mathcal{G}$ is the set of cost matrices; thus, $G(s,u)$ is the cost associated with state-action pair $(s,u)$ for particular objective.

We consider the average cost criterion for scheduling packets in incremental redundancy schemes, where our goal is to minimize average cost for stationary policy $\pi$ over a set of policies $\Pi$ and it is given by the following equation,

$$G^\pi = \limsup_{m\mapsto\infty}\frac{1}{\mathbb{E}_s^\pi\{T_m\}}\mathbb{E}_s^\pi\left\{\int_0^{T_m} g(s(t),u(t))dt\right\}, \quad (9)$$

where, $s(t)=s^n$ and $u(t)=u^n$ for $t^n \leq t < t^{n+1}$, and $T_m$ is the completion time of the $m^{th}$ transition. The expectation operator $\mathbb{E}_s^\pi$ is the conditional expectation when the probability measure is determined by the policy $\pi$, and the conditioning event is $\{s^0=s\}$. It can be shown that the problem at hand has a unichain structure. A unichain MDP has a single recurrent class, where for every pair of states there exists a stationary policy under which one state is accessible from other states, and possibly a empty set of transient states.

### A. SMDP Kernel and Transition Probability

Mathematically, the transition distributions for a given state-action pair $(s^n,u^n)$ can be given by,

$$Q_{s^n,s^{n+1}}(\tau,u^n) = P\{t^{n+1}-t^n \leq \tau, s^{n+1}|s^n,u^n\},\ \forall s^{n+1}\in\mathcal{S}. \quad (10)$$

Let $p(\tau,T_f)$ denotes a step function having value of 1 if $T_f < \tau$ and 0 elsewhere. In practical systems, truncated HARQ with limited number of retransmissions have been found to give good trade-off between throughput and delay, i.e., it increases the throughput and minimizes the packet delay in the buffer [6]. It is also intuitively clear that rather than



This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

retransmitting large number of times, it is better to try limited number of times and if still decoding fails, then transmit again with higher transmission power. Therefore, without loss of generality and to avoid cumbersome expressions, we assume that the maximum number of retransmissions is two. Therefore, the transition distributions function can be expressed as,

$$Q_{s^n,s^{n+1}}(\tau,u^n) = p(\tau,T_B)P^c_A(c^{n+1}|c^n,u^n)\tilde{P}^{1,A}_{b^n,b^{n+1}}(u^n)$$
$$+p(\tau,2T_B)P^c_{N,A}(c^{n+1}|c^n,u^n)\tilde{P}^{2,A}_{b^n,b^{n+1}}(u^n)$$
$$+p(\tau,3T_B)P^c_{N,N,A}(c^{n+1}|c^n,u^n)\tilde{P}^{3,A}_{b^n,b^{n+1}}(u^n)$$
$$+p(\tau,3T_B)P^c_{N,N,N}(c^{n+1}|c^n,u^n)\tilde{P}^{3,N}_{b^n,b^{n+1}}(u^n). \quad (11)$$

In (11), the term $P^c_{N,N,A}(c^{n+1}|c^n,u^n)$, for example, is the probability of switching to state $c^{n+1}$ from state $c^n$ for action $u^n$, which causes NAK in the first transmission and in the first retransmission and causes ACK in the second retransmission. It can be expressed as,

$$P^c_{N,N,A}(c^{n+1}|c^n,u^n) = \sum_{c^{k+2}}\sum_{c^{k+1}} P_N(c^n,u^n)P_N(c^n,c^{k+1},u^n)$$
$$\times P_A(c^n,c^{k+1},c^{k+2},u^n)P_{c^n,c^{k+1}}P_{c^{k+1},c^{k+2}}P_{c^{k+2},c^{n+1}}, \quad (12)$$

where, $c^{k+1}$ and $c^{k+2}$ are the intermediate channel states. Other transition probabilities for specific observations can be obtained similarly. In (11), the probability of incoming traffic $a^n$ that causes buffer occupancy $b^{n+1} \in \mathcal{B}$ in the next time-slot for a particular state-action pair $(s^n,u^n)$ can be given by,

$$\tilde{P}^{l,A}_{b^n,b^{n+1}}(u^n) = \sum_{a^n\in\mathcal{A}} \delta\{b^{n+1}-(b^n-w+a^n)\}P^{la}(a^n)$$
$$\text{and } \tilde{P}^{3,N}_{b^n,b^{n+1}}(u^n) = \sum_{a^n\in\mathcal{A}} \delta\{b^{n+1}-(b^n+a^n)\}P^{3a}(a^n)$$

where $P^{la}(a^n)$ is the probability of $a^n \in \mathcal{A} = \{0,1,\cdots,l\times A\}$ packet arrivals in $l \in \{1,2,3\}$ time-slots and $\delta(x)$ is a delta function whose value is 1 at $x=0$ and zero otherwise.

The expected value of the transition time corresponding to state-action pair $(s^n,u^n)$ can be given by,

$$\bar{\tau}_{s^n}(u^n) = \sum_{s^{n+1}\in\mathcal{S}} \int_0^\infty \tau Q_{s^n,s^{n+1}}(d\tau,u^n). \quad (13)$$

As a consequence of choosing particular action $u^n$ in a particular state $s^n$, the system moves to a new state $s^{n+1}$ with probability given by transition probability. The transition probabilities can be specified by transition distributions via,

$$p_{s^n,s^{n+1}}(u^n) = \lim_{\tau\mapsto\infty} Q_{s^n,s^{n+1}}(\tau,u^n). \quad (14)$$

### B. Costs Associated with Different Objectives

In this paper, our objective is to minimize three parameters that guarantee the QoS requirements for the IR-HARQ problem: power, buffer delay and packet overflow. Minimizing transmission power is of particular importance for wireless devices that usually operate with battery of limited energy. Since the power cost doesn't depend on future system state, therefore for state-action pair $(s^n,u^n)$, it can be written as,

$$G^P(s^n,u^n) = g^P(s^n,u^n)\bar{\tau}_{s^n}(u^n), \quad (15)$$

where $g^P(s^n,u^n)$ is the immediate power cost per time-slot and is simply equal to transmission power.

Since different traffic have different delay sensitivity, delay is another parameter that quantifies QoS requirements in wireless networks. The delay cost for IR-HARQ problem depends on present state as well as next state. Therefore, we take $g^D(s^n,u^n,s^{n+1}) = (b^n+b^{n+1}-2)/2\lambda,\ \forall c^n$ as the immediate delay cost per time-slot. The reason behind such choice is based on the approximation that the buffer occupancy changes linearly between the successive decision-epochs. The expected delay cost for IR-HARQ system with two retransmission for composite state $s^n = [b^n,c^n]$ and action $u^n$ can be given by,

$$G^D(s^n,u^n) = \sum_{b^{n+1}\in\mathcal{B}}\sum_{c^{n+1}\in\mathcal{C}} \left(\frac{b^n+b^{n+1}-2}{2\lambda}\right)T_B$$
$$[P^c_A(c^{n+1}|c^n,u^n)\tilde{P}^{1,A}_{b^n,b^{n+1}}(u^n) + 2P^c_{N,A}(c^{n+1}|c^n,u^n)$$
$$\times\tilde{P}^{2,A}_{b^n,b^{n+1}}(u^n) + 3P^c_{N,N,A}(c^{n+1}|c^n,u^n)\tilde{P}^{3,A}_{b^n,b^{n+1}}(u^n)$$
$$+3P^c_{N,N,N}(c^{n+1}|c^n,u^n)\tilde{P}^{3,N}_{b^n,b^{n+1}}(u^n)]. \quad (16)$$

Buffer overflow is an important QoS requirements when the incoming traffic is bursty. Overflow occurs if the buffer has less vacancy than the number of incoming packets. When the buffer has $B-A+1$ or more packets, then $0,1,\cdots$, up to a maximum of $3A$ packets can be dropped due to overflow. The expected overflow cost for above buffer states is given by,

$$G^O(s^n,u^n) = \sum_{c^{n+1}\in\mathcal{C}} T_B[P^c_A(c^{n+1}|c^n,u^n)\sum_{a=1+r^n}^{A} P^a(a)(a-r^n)$$
$$+P^c_{N,A}(c^{n+1}|c^n,u^n)\sum_{a=1+r^n}^{2A} P^{2a}(a)(a-r^n)$$
$$+P^c_{N,N,A}(c^{n+1}|c^n,u^n)\sum_{a=1+r^n}^{3A} P^{3a}(a)(a-r^n)$$
$$+P^c_{N,N,N}(c^{n+1}|c^n,u^n)\sum_{a=1+r^n}^{3A} P^{3a}(a)(a-r^n+w)], \quad (17)$$

where $r^n$ is the maximum number of packets that can be accommodated without overflow and equals $B+w-b^n$. Note that when $B-2A+1 \le b^n \le B-A$, the number of packets that can be dropped due to overflow is $0,1,2,\cdots$ up to a maximum of $2A$. When $B-3A+1 \le b^n \le B-2A$, the number of packets that can be dropped due to overflow is $0,1,2,\cdots$ up to a maximum of $A$. The expected overflow cost for these buffer occupancies can be given similarly. For all other buffer occupancies, no packet overflow occurs, therefore the cost is zero.

## IV. Equivalent Discrete-time MDP Formulation

Semi-Markov average cost problem formulated in Section III is equivalent to an auxiliary discrete-time average cost problem. It can also be shown that the auxiliary discrete-time average cost problem and the semi-Markov average cost problem have the same probabilistic structure [7]. Therefore, dynamic programming for DT-MDP can be applied to the auxiliary problem in order to solve the semi-Markov problem. Let us assume that $p_{s_i,s_i}(u_i) < 1$ for all $s_i \in \mathcal{S}$ and $u_i \in \mathcal{U}_{s_i}$, and $\gamma$ be any scalar satisfying following inequality,

$$0 < \gamma < \frac{\bar{\tau}_{s_i}(u_i)}{1 - p_{s_i,s_i}(u_i)}. \quad (18)$$

The transition probability and the expected average cost per stage for auxiliary discrete-time problem can be obtained respectively from [7],

$$\tilde{p}_{s_i,s_j}(u_i) = \begin{cases} \frac{\gamma p_{s_i,s_j}(u_i)}{\bar{\tau}_{s_i}(u_i)}, & \text{if } s_j \neq s_i; \\ 1 - \frac{\gamma(1-p_{s_i,s_i}(u_i))}{\bar{\tau}_{s_i}(u_i)}, & \text{if } s_j = s_i. \end{cases} \quad (19)$$

$$\text{and } \tilde{G}(s_i, u_i) = \frac{G(s_i, u_i)}{\bar{\tau}_{s_i}(u_i)}. \quad (20)$$

The auxiliary DT-MDP problem can be solved either using

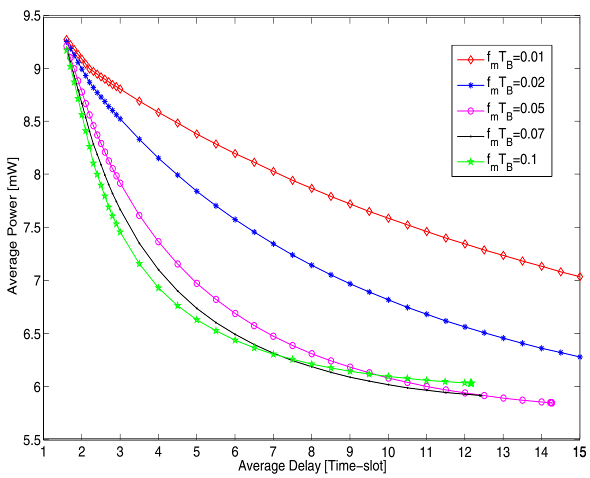


Fig. 2. Tradeoff between transmitter power and buffer delay for constant incoming traffic. Packet overflow bound $P_{\text{of}} = 10^{-3}$.

unconstrained formulation technique, where a Lagrangian sum of the costs are minimized or constrained formulation technique, where cost for one objective is minimized keeping other costs (called constraints costs) below some specified bounds. In this paper, we consider constrained Markov decision process (CMDP) formulation that can easily incorporate more than one constraints and be solved using LP method. The CMDP problem can be expressed by the following equations, where our objective is to find optimal policy $\pi^*$ over a set of all stationary policies $\Pi$ that satisfy,

$$\min_{\pi\in\Pi} \limsup_{m\mapsto\infty} \frac{1}{m}\mathbb{E}_\pi\left[\sum_{n=1}^{m} \tilde{G}^P(s^n, \pi(s^n))\right] \quad (21)$$

$$\text{s.t.: } \limsup_{m\mapsto\infty} \frac{1}{m}\mathbb{E}_\pi\left[\sum_{n=1}^{m} \tilde{G}^D(s^n, \pi(s^n))\right] \leq D \quad (22)$$

$$\text{and } \limsup_{m\mapsto\infty} \frac{1}{m}\mathbb{E}_\pi\left[\sum_{n=1}^{m} \tilde{G}^O(s^n, \pi(s^n))\right] \leq P_{\text{of}}, \quad (23)$$

where, the objective is to minimize expected long-term power cost with constraints on expected long-term delay and overflow costs. Nonnegative constants $D$ and $P_{\text{of}}$ are the maximum allowable long-term average delay and packet overflow. The choice of these parameters depends on the QoS requirements of a particular user. Due to the ergodic nature of the buffer and channel state processes the problem has unichain structure [7]. Let $\nu(s, u)$ represents the "steady-state" probability that the process is in state $s$ and action $u$ is applied. We seek to find the control policy which is represented in terms of probability distribution $\nu$ over $\mathcal{S} \times \mathcal{U}$. The optimal policy $\nu^*$ can be obtained by solving the LP [8]:

$$\min_{\nu} \sum_{s_i\in\mathcal{S},u_i\in\mathcal{U}_{s_i}} \tilde{G}^P(s_i, u_i)\nu(s_i, u_i) \quad (24)$$

$$\text{s.t.: } \sum_{s_i\in\mathcal{S},u_i\in\mathcal{U}_{s_i}} \tilde{G}^D(s_i, u_i)\nu(s_i, u_i) \leq D$$

$$\sum_{s_i\in\mathcal{S},u_i\in\mathcal{U}_{s_i}} \tilde{G}^O(s_i, u_i)\nu(s_i, u_i) \leq P_{\text{of}}$$

$$\sum_{u_i\in\mathcal{U}_{s_i}} \nu(s_j, u_i) = \sum_{s_i\in\mathcal{S},u_i\in\mathcal{U}_{s_i}} \nu(s_i, u_i)p_{s_i,s_j}(u_i), \ \forall s_j \in \mathcal{S}$$

$$\sum_{s_i\in\mathcal{S},u_i\in\mathcal{U}_{s_i}} \nu(s_i, u_i) = 1; \ \nu(s_i, u_i) \geq 0, \ \forall s_i \in \mathcal{S}, \forall u_i \in \mathcal{U}_{s_i}.$$

There exists an optimal randomized stationary policy $\mu^*$ for the CMDP problem if there exists an optimal solution $\nu^*$ to the LP problem. The probability of applying policy $u \in \mathcal{U}_s$ in state $s \in \mathcal{S}$ satisfies,

$$\theta_{\mu^*(s_i)}(u_i) = \frac{\nu^*(s_i, u_i)}{\sum_{u_j\in\mathcal{U}_{s_i}} \nu^*(s_i, u_j)} \text{ if } \sum_{u_j\in\mathcal{U}_{s_i}} \nu^*(s_i, u_j) > 0.$$

## V. Simulation Results

In this section, we show simulation results for the following set of data: number of channel states $C = 2$, buffer size $B = 100$ packets, average channel gain $\bar{\gamma} = 1$, noise power $\sigma^2 = 1$ mW, number of actions $U = 3$, set of power levels $P_t$ corresponding to three actions is $P_s = \{6.4, 15, 30\}$ mW, maximum number of retransmissions $R = 2$, block rate $R_B = 10^4$ blocks/sec, number of packets taken for transmission in each decision-epoch $w = 2$ and average arrival rate $\lambda = 1$ packet per time-slot. The set of rates of the family of RCPC codes at the transmitter is $\{1, 1/2, 1/3\}$, which are generated from a rate $1/3$ code with memory $m_{tb} = 4$ [4]. In Fig. 2, we show the tradeoff between the transmission power and the buffer delay for constant incoming traffic of 1 packet/time-slot. It is seen from the figure that the power decreases as delay increases and the rate of decrease of power is more for faster fading channels in the lower delay regions. But in the higher delay regions, the rate of decrease of power is more for slower fading channels. For smaller delays, flexibility of storing packets is limited. Therefore, when the fading rate is slow, the scheduler has to choose higher transmission power to send the packets. For larger delays, the scheduler has more flexibility to store packets in bad channel states when the fading is slow and can allow more fluctuation in



This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

the buffer. Also, for slower fading the channel states in a decision-epoch are more predictable, therefore the decoding success is increased. Hence, in these regions the faster fading channels needs more power. The power/delay trade-off for Poisson traffic with maximum number of packet arrival $A = 7$ is shown in Fig. 3. The figure shows the same trend as the constant traffic, but the power needed to maintain the same delay is more for Poisson traffic. In Fig. 4, we give a comparison between the constant and Poisson distributed incoming traffic for same average packet arrival rate of 1 packet/time-slot. It is seen from the figure that the difference between transmission power for constant and Poisson traffic with same delay increases as the fading correlation increases. The packet overflows as a function of delay for constant and Poisson traffic is shown in Fig. 5. Note that to maintain same overflow, Poisson traffic needs more delay, which is expected.

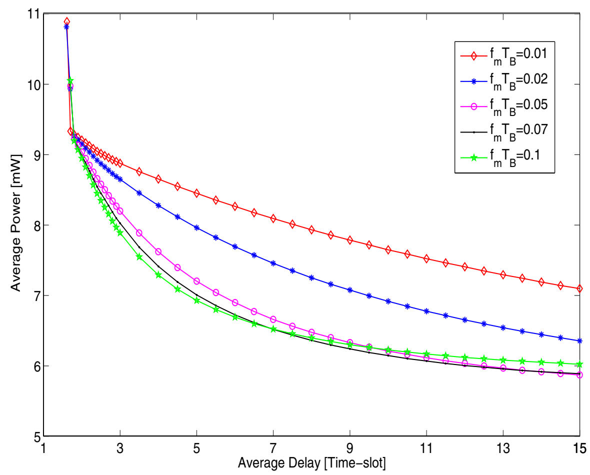


Fig. 3. Tradeoff between transmitter power and buffer delay for Poisson incoming traffic. Packet overflow bound $P_{\text{of}} = 10^{-3}$.

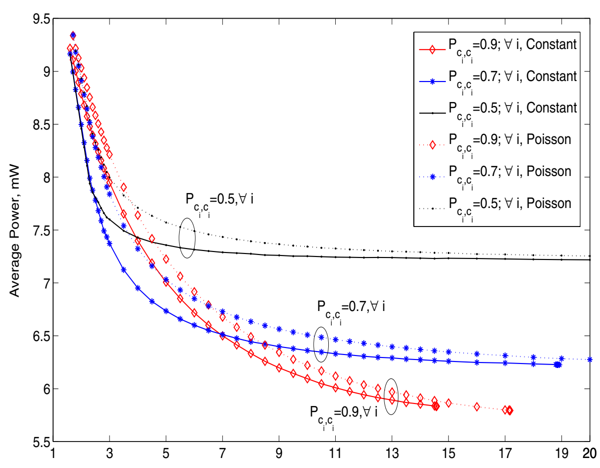


Fig. 4. Comparison between constant and Poisson traffic in terms of transmit power for different channel memories. Packet overflow bound $P_{\text{of}} = 10^{-4}$.

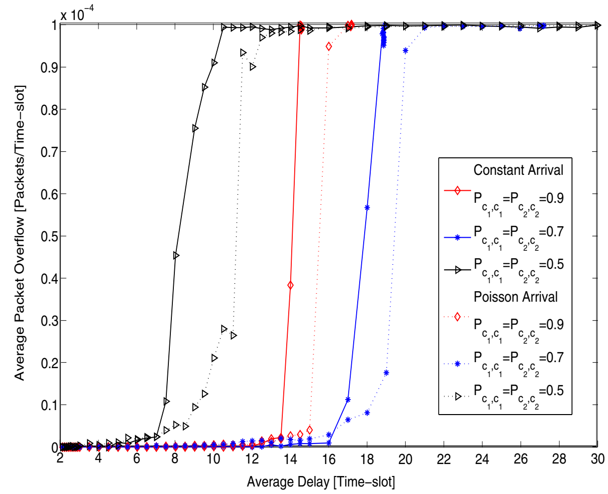


Fig. 5. Variation of packet overflow with buffer delay for Poisson incoming traffic. Packet overflow bound $P_{\text{of}} = 10^{-4}$.

## VI. Conclusions

Semi-Markov decision process framework has been utilized to calculate the optimal transmit power adaptation policy of an IR-HARQ system.We have derived the SMDP kernel, transition probability and costs associated with different objectives to be optimized. The SMDP problem has then been converted into an auxiliary discrete-time CMDP problem and its solution is obtained by linear programming methods. Simulation results have been given to examine the influence of randomly varying channel and traffic parameters on the performance of the optimal power allocation policy. It has been shown that by employing optimal power allocation, faster fading channel performs better under stringent delay constraint, while the situation is the opposite if the delay constraint is relaxed.

## References

[1] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC) and their applications," *IEEE Trans. Commun.*, vol. 36, pp. 389-400, April 1988.

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

[3] S. Kallel, and D. Haccoun, "Generalized type II hybrid ARQ scheme using punctured convolutional coding," *IEEE Trans. Commun.*, vol. 38, pp. 1938-1946, Nov. 1990.

[4] Qinqing Zhang, and S. A. Kassam, "Hybrid ARQ with selective combining for fading channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 867-880, May 1999.

[5] A. K. Karmokar, and V. K. Bhargava, "Coding rate adaptation for hybrid ARQ systems over time varying fading channels with partially observable state", in *Proc. of ICC'05*, May. 16-20, 2005, vol. 4, pp. 2797-2801.

[6] L. Lugand, D. J. Costello, Jr., and R. H. Deng, "Parity retransmission hybrid ARQ using rate 1/2 convolutional codes on a nonstationary channel," *IEEE Trans. Commun.*, vol. 37, pp. 755-765, Jul. 1989.

[7] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Second Edition, vol. II. Belmont, Massachusetts: Athena Scientific, 2001.

[8] E. A. Feinberg, and A. Shwartz, *Handbook of Markov Decision Processes: Models and Applications*. Boston: Kluwer Academic, 2001.



*This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.*