

# Optimal Control of Service Rate and Reliability in Scheduling over Wireless Channels

Aditya Karnik

Global General Motors R&D  
India Science Lab, GM Tech Center (India)  
Bangalore 560066, India  
Email: aditya.karnik@gm.com

Catherine Rosenberg

Dept. of Electrical and Computer Engineering  
University of Waterloo  
Waterloo, Ontario, Canada N2L 3G1  
Email: cath@ece.uwaterloo.ca

**Abstract**—This work investigates a problem of scheduling tasks on a single server in which the rate at which the server can complete its tasks is not fixed but can be varied- a *higher service rate* resulting in correspondingly *less reliable* service. This is a prevalent trade-off in many practical systems including wireless communication systems. Introducing the idea of mode of service, we formulate the rate-reliability problem as a Markov control problem. We obtain existence results and show that the optimal policy has a threshold mode. The optimal policy is also shown to provide average delay guarantee while maximizing goodput. We provide numerical results supporting our analysis and showing advantage of the optimal policy over *fixed rate* policies.

## I. INTRODUCTION

The problem of scheduling tasks on a single server so as to provide certain performance guarantees to its users arises in numerous areas, most notably in communication, manufacturing and computer systems. For instance, in a communication system the server may be a link and the task may be to transmit packets for different flows. There is a vast literature on this problem addressing different variations. This work investigates a case in which the rate at which the server can complete its tasks is not fixed but can be varied- a *higher service rate* resulting in correspondingly *less reliable* service. Though this is a prevalent trade-off in many practical systems, our motivation comes from wireless communication systems; hence they form the focus of the paper.

In a communication system, the rate-reliability trade-off arises from the fact that for a fixed received power or equivalently, signal-to-noise ratio (SNR) at the receiver<sup>1</sup>, increasing the data transmission rate increases the bit error rate (BER) [1]. For example in M-ary PSK single antenna uncoded system, a high transmission rate pertains to a bigger constellation size and for a fixed SNR, results in a higher BER in comparison to a lower transmission rate (smaller constellation size). Figure 1 roughly depicts this trade-off. Obviously, an overall increase in the transmission power changes the trade-off curve by

Part of this work was done when A. Karnik was at University of Waterloo, Waterloo, Canada.

<sup>1</sup>A constant SNR may result from transmission at a fixed power level, under time-invariant channel gains, or from power control if the channel gains are varying with time.

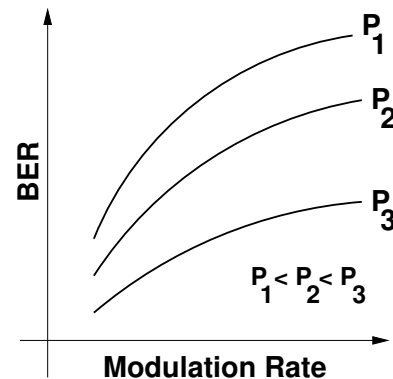


Fig. 1. Rate-Reliability trade-off in communication systems.

yielding a lower BER for a given rate (see Figure 1). However, transmission power cannot be scaled arbitrarily owing to regulatory constraints. The rate-BER trade-off is not typically exploited in existing wireless systems since they are operated under a peak power constraint and a given BER requirement. We believe that by exploiting it, the wireless systems can be much more efficient and flexible. Operating a system this way might be more complicated than operating with fixed BER, however, operating the latter is not necessarily simple since if the channel characteristics are varying with time, so must the physical layer parameters (i.e., transmit power, modulation, etc.) to maintain the fixed BER. Data traffic usually requires long term performance guarantees such as goodput (i.e., the average rate at which packets are successfully received at the receiver) and/or average delay, and, therefore, is not sensitive to small time-scale changes in service reliability as long as their performance (QoS) objectives are met. This is, in particular, true for multimedia (voice, video) sources which are sensitive to delay but not to the loss of some packets (QoS typically imposes certain goodput and delay requirements; hence retransmitting packets on transmission failures is futile). Thus, by trading rate and BER, a wireless system can potentially handle more traffic without compromising the long term quality of transmission. Henceforth we will be concerned only with multimedia traffic.

This raises a basic (and obvious) question: how should

the system be operated under such a rate-reliability trade-off so that the performance guarantees required by the data traffic (and assumed to be feasible) are met? There are two ways. In the first, at a relatively slower time-scale, the server can adapt the rate of transmission to the input rate, thereby, guaranteeing high goodput if the input traffic rate is low while giving gracefully degrading goodput performance with increasing traffic rate. This is unlike a fixed BER system, in which the maximum transmission rate is essentially fixed. This adaptive system can potentially support higher traffic rates (though with low reliability). In the second, the server can adapt, at a much finer scale, to the stochastic variations in the instantaneous load- when the buffers have high occupancy, the BER can be traded with the transmission rate and when the buffer occupancy is low the server may switch to a lower rate yielding high reliability. Thus, by adaptively switching between rate-reliability regimes based on buffer occupancy the server can provide delay guarantees to multimedia sources while improving the total goodput. Note that an on-line version of such a scheme would obviate the need to explicitly adapt to the traffic rate since it will adapt transparently to the time-varying traffic conditions. The other compelling factor for the second strategy is the time-varying nature of the wireless channel gains. As the channel varies from ‘good’ to ‘bad’, the BER can be lowered by lowering the transmission rate if the delay requirements permit. In fact, owing to differences in channel quality for different users, the server can adapt transmission rates individually for each user and do so opportunistically, thereby, striking a balance between efficiency and robustness.

In this paper, we develop a scheme that trades rate for reliability and vice versa in response to the buffer occupancy and channel conditions depending on the performance guarantees; thus the server operates at rates corresponding to different BER values. This differs from the adaptive modulation schemes (e.g., [2], [3], [4]) in which the transmission parameters are changed only in response to changes in the channel quality so as to maintain a fixed SNR (or BER). Our scheme may, thus, be seen as a cross-layer optimization scheme where the two layers involved are the physical and link layer. While we are not concerned with the practicality of such a scheme in this paper, we believe that it directly impacts the following two scenarios of practical interest.

1) **Adaptive burst profiles in single antenna systems:**

The data transmission rate in single antenna transceivers is determined by the physical layer parameters, namely, the modulation-coding schemes. Many wireless systems, in particular IEEE 802.16 standard for broadband wireless access, provide an adaptive physical layer. The radio link control is typically capable of adaptive burst profiling in which the modulation and coding schemes can be adjusted individually to each user on a frame-by-frame basis. Though mechanisms to monitor and control the burst profiles have been provided in IEEE 802.16 PMP, the question of how this control is executed is left open.

2) **Scheduling over MIMO antennas:** MIMO systems provide an additional dimension of varying the transmission rate by manipulating multiple antennas. Thus, a packet may be transmitted over multiple antennas simultaneously resulting in low transmission rate but high reliability, or multiple packets may be transmitted simultaneously, say, one on each antenna yielding high transmission rate but low reliability. The former may be identified as diversity gain while the latter multiplexing gain. Diversity-multiplexing trade-off has received wide attention in recent years. However, studies have been purely from the physical layer perspective and the question of trading rate-reliability or equivalently multiplexing-diversity to guarantee traffic performance is yet to be resolved.

Our aim is to formulate and study the rate-reliability problem analytically. In this paper, we consider a single buffer at the server in which arriving packets are queued till served. This models single user MIMO systems or an IEEE 802.16 subscriber station employing only one buffer on its wireless uplink. We assume a finite number of possible service rates to capture finite modulation-coding schemes and combinations of multiple antennas. Assuming that the server serves packets in frames/slots, this leads to a notion of ‘mode’ of transmission; each mode specifies the maximum number of packet transmissions in a given time period and a probability distribution for their successful transmission.

The rate-reliability problem is formulated as a Markov control problem: the server, at the beginning of each slot, needs to determine the mode of service based on its buffer occupancy so as to maximize some reward function. The positive part of the reward corresponds to the number of successfully served packets (i.e., a measure of goodput) whereas the negative part corresponds to the holding cost (i.e., a cost for delaying packets). We provide existence and structural results for the optimal policy. In addition, we show that the optimal policy can provide average delay guarantees (assuming they are feasible) while maximizing the server goodput. Numerical results using an equivalent linear programming formulation are also provided; in particular we obtain the goodput-delay performance frontier. While the problem formulation is motivated by the two systems discussed above, in this paper we do not quantify gains of the adaptive scheme in their specific contexts. Also, to keep the analysis simple we assume single channel state (hence adaptation is with respect to only queue-length). These and other possible extensions (discussed in Section VI) will be undertaken in a longer version of the paper.

A remark is in order. The problem of controlling the service rate of server(s) is not novel. Different versions of it have been studied, for example,  $M/G/1$  queues in [5], tandem queues in [6] and  $M/M/1$  queues in [7]. To our knowledge, the idea of exploiting the rate-reliability trade-off and the corresponding formulation to maximize goodput subject to delay seems to be new.

The paper is organized as follows. In Section II we discuss the model and formulate the problem. Existence and struc-

tural results are proved in Section III. The delay-constrained problem is considered in Section IV. Numerical results are provided in Section V. We conclude in Section VI.

## II. MODEL AND FORMULATION

Consider a system consisting of a server with an infinite capacity buffer, in which exogenously arriving packets are queued till they are served, after which they depart the system. Service is first-in-first-out. Packets may belong to one user or to multiple users of the same class. Time is assumed to be slotted with slot (alternatively, frame)  $k$  denoting a time period of  $[k, k + 1)$ ,  $k = 0, 1, \dots$ . Packets arriving in slot  $k$  are accumulated and considered for service only in slots  $(k + 1)$ -onwards.  $v(k)$  denotes the number of packet arrivals in slot  $k$ .  $\{v(k), k \geq 0\}$  is assumed to be an i.i.d sequence with  $P_v(\cdot)$  denoting the common p.m.f.  $P_v(0) > 0$  and  $\lambda := E[v]$ . Thus,  $\lambda$  denotes the average rate of the input traffic per slot. The server can operate in  $M + 1$  modes indexed  $m = 0, 1, 2, \dots, M$ . The set of modes is denoted by  $\mathcal{M}$ . Mode  $m$  specifies the maximum number of packets that can be served in a slot,  $\eta_m$ , and a set of conditional distributions,  $P_m(\cdot|j)$  for each  $j \leq \eta_m$ ;  $P_m(l|j)$  gives the probability that  $l \leq j$  packets are successfully served given that  $j$  were served in mode  $m$ . In this paper, we assume the  $P_m(\cdot|j)$ 's to be Binomial with probability  $p_m$ . In other words,  $p_m$  is the probability that a packet is successfully served in mode,  $m$ , and successful transmissions of packets served are independent events. It follows that mode  $m$  is characterized by  $(\eta_m, p_m)$ . Without loss of generality, we assume that  $0 = \eta_0 < \eta_1 < \dots < \eta_M$  and  $1 = p_0 > p_1 > \dots > p_M$ . Thus, the modes are ordered in increasing  $\eta_m$ 's and we say  $m_2$  is faster/larger than  $m_1$  ( $m_1 \leq m_2$ ) if  $\eta_{m_1} \leq \eta_{m_2}$ . Let  $x(k)$  denote the queue-length at the beginning of slot  $k$ ;  $x(k) \in Z_+$ , the set of nonzero integers.  $\{x(k), k \geq 0\}$  constitutes the observable state process of the system. We denote by  $\mathcal{X}$  the state space and by  $x$  a generic element of  $\mathcal{X}$ .

The scheduling decision in slot  $k = 0, 1, \dots$ , taken at the beginning of the slot, is to determine the mode of service,  $m(k)$ , after observing  $x(k)$ . If  $x(k) = x$  the number of packets transmitted is  $\min\{\eta_m, x\}$ . Since our focus is on the multimedia traffic, we assume that the packets that are lost are not retransmitted. We denote by  $a$  a scheduling decision, i.e.,  $a = m \in \mathcal{M}$ .  $m_a$  denotes the mode of transmission associated with a particular action  $a$ .  $\mathcal{A}$  denotes the set of actions  $a$ . The set of feasible actions in state  $x$  is  $\mathcal{A}(x) = \mathcal{M}$ . Denote by  $Q(x'|x, a)$  the probability of transition to  $x'$  given  $x$  and action  $a \in \mathcal{A}(x)$ . Then  $Q(x'|x, a) = P_v(x' - (x - \eta_{m_a})^+)$ , where  $(u)^+$  denotes  $\max\{u, 0\}$ . The expected one stage reward in state  $x$  and action  $a \in \mathcal{A}(x)$ ,  $r(x, a)$ , is  $\min\{\eta_{m_a}, x\}p_{m_a} - E_v[h((x - \eta_{m_a})^+ + v)]$ . Note that  $\min\{\eta_{m_a}, x\}p_{m_a}$  is the average number of successfully served packets and  $h(\cdot)$  denotes the cost for holding packets in the buffer. It is assumed that once scheduled,  $\min\{\eta_{m_a}, x\}$  packets are removed from the buffer and do not incur any cost in that slot. For reasons that will be clear in Section IV, we

assume a linear holding cost. Thus,

$$r(x, a) = \min\{\eta_{m_a}, x\}p_{m_a} - \beta(x - \eta_{m_a})^+ \quad (1)$$

for some  $\beta > 0$  which controls the trade-off between goodput and buffer occupancy. Note that the positive part in the expected one-stage reward (related to goodput) favors 'robust' (i.e., having higher value of  $\eta_m p_m$ ) modes whereas the negative part (corresponding to buffer occupancy) favors 'fast' modes.

A discrete time Markov control problem is now specified by  $(\mathcal{X}, \mathcal{A}, \{\mathcal{A}(x)|x \in \mathcal{X}\}, Q, r)$ . Denote by  $\Pi$  the set of feasible policies<sup>2</sup> and by  $E_x^\pi$  the expectation with respect to probability measure determined by policy  $\pi \in \Pi$  and the initial distribution  $P(x(0) = x) = 1$ . The objective is to find a scheduling policy, if it exists, which maximizes the time-average reward, i.e.,

$$\max_{\pi \in \Pi} \liminf_{n \rightarrow \infty} \frac{1}{n} E_x^\pi \left[ \sum_{k=0}^{n-1} r(x(k), a(k)) \right] \quad (2)$$

## III. EXISTENCE AND STRUCTURAL RESULTS

Let

$$J^\pi(x) := \liminf_{n \rightarrow \infty} \frac{1}{n} E_x^\pi \left[ \sum_{k=0}^{n-1} r(x(k), a(k)) \right]$$

for  $\pi \in \Pi$  and

$$J^*(x) := \max_{\pi \in \Pi} J^\pi(x)$$

Recall that a stationary deterministic policy  $\pi$  is a function from  $\mathcal{X}$  to  $\mathcal{A}$ . Thus,  $\pi(x)$  denotes the optimal action in state  $x$ . Since  $\mathcal{X}$  is countable,  $\pi$  can essentially be treated as an infinite-dimensional vector so that  $\pi(x)$  also denotes the  $x^{\text{th}}$  co-ordinate of  $\pi$  for  $x = 1, 2, \dots$ . Let  $\Pi_{SD}$  denote the subset of stationary deterministic policies in  $\Pi$ .

*Proposition 3.1:* For any arrival rate  $\lambda$  within the stability region of the system, i.e.,  $\lambda < \eta_M$ ,

- 1) there exists an optimal policy,  $\pi^*$ , in  $\Pi_{SD}$ .
- 2)  $J^*(x)$  is a constant for all  $x \in \mathcal{X}$  and is obtained as limit.
- 3) the average cost optimality equation holds.
- 4) any limit point of  $\alpha$ -discounted reward optimal policies is average reward optimal.

*Proof:* Note that  $r(x, a) \leq \max_m \eta_m p_m$ . For any  $\pi \in \Pi_{SD}$ ,  $\{x(k), k \geq 0\}$  is a Markov chain. It is aperiodic since  $P_v(0) > 0$ . Consider policy  $\pi^M$  under which the server operates only the mode  $M$ . Then under  $\pi^M$ ,  $\{x(k), k \geq 0\}$  is irreducible and positive recurrent. Thus,  $\pi^M$  is a 0 standard policy. Further, for each  $u > 0$ , the set  $\{x|r(x, a) \geq -u \text{ for some } a\}$  is finite. The proposition now follows from [8] (Theorem 7.2.3 and Theorem 7.5.6). ■

We now show that the optimal policy has a 'threshold mode', denoted by  $m^*$ . Modes faster than  $m^*$  are not employed. Moreover, for large enough queue-lengths, the optimal

<sup>2</sup>A feasible policy,  $\pi$ , is a sequence  $\{\pi_k\}_{k \geq 0}$  where  $\pi_k$  is the conditional probability distributions on  $\mathcal{A}$  given the history (sequence of states and actions) up to slot  $k$ .

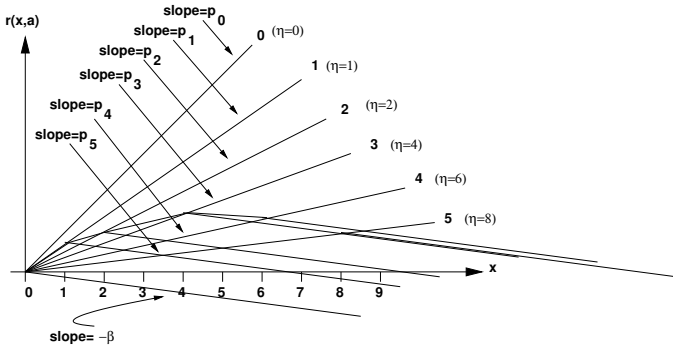


Fig. 2. Variation of the expected one stage reward  $r(x, a)$  with  $x$  when the server can operate in 6 modes. The bold curve shows  $V_{\alpha,1}^*(x)$ . The threshold mode  $m_{\alpha,1}^* = 4$ .

mode of transmission is the same,  $m^*$ . Formally, the result is stated as follows.

**Proposition 3.2:** There exists an  $m^* \in \mathcal{M}$  such that  $m_{\pi^*(x)} \leq m^*$  for all  $x \in \mathcal{X}$ . Moreover,  $m_{\pi^*(x)} = m^* \Rightarrow m_{\pi^*(x+1)} = m^*$ .

Let for  $0 \leq \alpha < 1$ ,  $V_{\alpha,n}^*(x)$  (resp.  $V_{\alpha}^*(x)$ ) denote the value function of the  $\alpha$ -discounted  $n$ -horizon (resp. infinite horizon) reward problem.  $\pi_{\alpha,n}^*$  and  $\pi_{\alpha}^*$  denote the respective optimal stationary deterministic policies. We define a class of functions called *eventually decreasing functions* (Definition 3.1) and show that for each  $n$ ,  $V_{\alpha,n}^*(x)$  is such a function (Proposition 3.3). This guarantees that  $\pi_{\alpha,n}^*$  has the structure given in Proposition 3.2. We then use the vanishing discount approach to show that  $\pi_{\alpha}^*$  has the same structure.

**Definition 3.1:**  $g(\cdot)$  is eventually linearly decreasing (*e.l.d.*) if there exists an  $\hat{x}$  and  $b \geq 0$  such that  $g(x) = -b(x - \hat{x}) + g(\hat{x})$  for  $x \geq \hat{x}$ .

Denote the class of *e.l.d.* functions on  $Z_+$  by  $G$ . Let  $g_k \in G$  for  $k = 1, 2, \dots$ . Denote by  $\hat{x}_k$  and  $b_k$  the ‘threshold’ and the ‘slope’ corresponding to  $g_k$ .

- Proposition 3.3:** 1)  $\max_{1 \leq k \leq K} g_k(x) \in G$ . Moreover, let  $k(x) := \arg \max_{1 \leq k \leq K} g_k(x)$ . Then there exists an  $\hat{x}$  such that  $x \geq \hat{x} \Rightarrow k(x) = k(\hat{x})$ .  
 2) For  $c_k \geq 0$ ,  $\sum_{k=1}^K c_k g_k(x) \in G$  with  $\hat{x} \leq \max_k \hat{x}_k$ .  
 3) If  $\max_k \hat{x}_k \leq X < \infty$  then  $g = \lim_k g_k \Rightarrow g \in G$ .

*Proof:*

- 1) It is easy to see that  $\max_{1 \leq k \leq K} g_k(x)$  eventually decreases linearly with  $b = \min_{1 \leq k \leq K} b_k$ . This also means that for large enough  $x$ ,  $k(x) = \arg \min_k b_k$ .
- 2) For  $x \geq \max_{1 \leq k \leq K} \hat{x}_k$ ,  $\sum_{k=1}^K c_k g_k(x) = \sum_{k=1}^K c_k (-b_k(x - \hat{x}_k) + g_k(\hat{x}_k))$ . Thus,  $\sum_{k=1}^K c_k g_k(x)$  is *e.l.d.* with  $\hat{x} \leq \max_k \hat{x}_k$  and  $b = -\sum_{k=1}^K c_k b_k$ .
- 3) The bound on  $\hat{x}_k$  implies that each  $g_k$  decreases linearly for  $x \geq X$ . Thus, for  $x \geq X$ ,  $g_k(x) = -b_k(x - X) + g_k(X)$ . Since the limit of converging linear functions is linear it follows that  $g$  is *e.l.d.* with  $\hat{x} \leq X$ . ■

**Proposition 3.4:**  $V_{\alpha,n}^*(x)$  is *e.l.d.* for each  $n \geq 0$ . There exists  $m_{\alpha,n}^* \in \mathcal{M}$  such that  $m_{\pi_{\alpha,n}^*(x)} \leq m_{\alpha,n}^*$  for all  $x \in \mathcal{X}$ , and  $m_{\pi_{\alpha,n}^*(x)} = m_{\alpha,n}^* \Rightarrow m_{\pi_{\alpha,n}^*(x+1)} = m_{\alpha,n}^*$ .

*Proof:* Our proof is by induction. Terminal cost being 0,  $V_{\alpha,0}^*$  is trivially *e.l.d.* To get an insight consider  $n = 1$ .  $r(x, m)$  is *e.l.d.* for each  $m$ . Therefore, by Proposition 3.3,  $V_{\alpha,1}^* = \max_m r(x, m)$  is *e.l.d.* Let

$$\mathcal{M}_1 := \{m | \eta_m p_m - (\eta_{\bar{m}} - \eta_m) \beta \geq \eta_{\bar{m}} p_{\bar{m}} \text{ for all } \bar{m} > m\}$$

and  $\hat{m} := \min \mathcal{M}_1$ . Then for  $x \geq \hat{m}$ ,  $m_{\pi_{\alpha,1}^*(x)} = \hat{m}$  and for  $x < \hat{m}$ ,  $m_{\pi_{\alpha,1}^*(x)} < \hat{m}$  (since  $m_{\pi_{\alpha,1}^*(x)} \in \mathcal{M}(x)$  for  $x \leq \eta_M$ , where,  $M(x) = \min\{m | x \leq \eta_m\}$  and  $\mathcal{M}(x) = \{m \leq M(x)\}$ ). Thus,  $V_{\alpha,1}^*$  is *e.l.d.* with  $\hat{x}_1 = \eta_{\hat{m}}$ ,  $b_1 = \beta$ , and  $m_{\alpha,1}^* = \hat{m}$ . The proposition, therefore, holds for  $n = 1$ .

Assume now that it holds for  $k = n - 1, n - 2, \dots, 0$  for  $n > 2$ . Also assume that  $b_k = \beta$  for each  $k$ .  $V_{\alpha,n}^*(\cdot)$  solves the following optimality equation.

$$V_{\alpha,n}^*(x) = \max_m (r(x, m) + \alpha E_v V_{\alpha,n-1}^*((x - \eta_m)^+ + v))$$

Fix  $m$  and let

$$V_{\alpha,n}^m(x) := r(x, m) + \alpha E_v V_{\alpha,n-1}^*((x - \eta_m)^+ + v)$$

Note that  $V_{\alpha,n-1}^*(x)$  *e.l.d.* with  $\hat{x} = \hat{x}_{n-1}$  implies that  $V_{\alpha,n-1}^*((x - \eta_m)^+ + v)$  is *e.l.d.* with  $\hat{x} = (\hat{x}_{n-1} - v)^+ + \eta_m$ . Thus, if  $g_k(x) := \sum_{v=0}^k P(v) V_{\alpha,n-1}^*((x - \eta_m)^+ + v)$  then  $g_k \uparrow E_v V_{\alpha,n-1}^*((x - \eta_m)^+ + v)$  with  $\hat{x}_k \leq \hat{x}_{n-1} + \eta_m$  for all  $k$  (Proposition 3.3 (2)). It now follows from Proposition 3.3 (3) that  $E_v V_{\alpha,n-1}^*((x - \eta_m)^+ + v)$  is *e.l.d.* Hence  $V_{\alpha,n}^m(x)$  is *e.l.d.* (Proposition 3.3 (2)). Since  $r(x, m)$  is *e.l.d.* with  $b = \beta$  and by assumption  $V_{\alpha,n-1}^*$  has  $b_{n-1} = \beta$ , it is clear that  $V_{\alpha,n}^m$  has  $b_n^m = \beta$  implying that the slopes of  $V_{\alpha,n}^m$  for each  $m$  are aligned. Note that  $\hat{x}_n^m = \hat{x}_{n-1} + \eta_m$  and, being *e.l.d.*,  $V_{\alpha,n-1}^*(\hat{x}_{n-1} + v) = -\beta v + V_{\alpha,n-1}^*(\hat{x}_{n-1})$ . Hence

$$\begin{aligned} V_{\alpha,n}^m(\hat{x}_{n-1} + \eta_m) &= \eta_m p_m - \beta(\hat{x}_{n-1} + \eta_m - \eta_m) + \\ &\quad \alpha \sum_v P(v) (V_{\alpha,n-1}^*(\hat{x}_{n-1}) - \beta v) \end{aligned}$$

Define  $y_n^m := \eta_m p_m - \beta \hat{x}_{n-1} + \alpha (V_{\alpha,n-1}^*(\hat{x}_{n-1}) - \beta \lambda)$  and let

$$\mathcal{M}_n := \{m | y_n^m - y_n^{\bar{m}} \geq (\eta_{\bar{m}} - \eta_m) \beta \text{ for all } \bar{m} > m\}$$

and  $\hat{m} := \min \mathcal{M}_n$ .

Since  $V_{\alpha,n}^m(x)$  is *e.l.d.* for each  $m$ , it follows from Proposition 3.3 (1) that  $V_{\alpha,n}^*(\cdot)$  is *e.l.d.* Hence, for  $x \geq \hat{x}_n$ ,  $m_{\pi_{\alpha,n}^*(x)} = m_{\pi_{\alpha,n}^*(\hat{x})}$ . Since the slopes of  $V_{\alpha,n}^m$  for each  $m$  are aligned,  $\hat{m}$  denotes the mode which eventually ‘dominates’ (hence  $b_n = \beta$  and  $\hat{x}_n = \hat{x}_{n-1} + \eta_{\hat{m}}$ ). ■

*Proof of Proposition 3.2:* Recall that  $r(x, a) \leq \max_m \eta_m p_m$ . Hence, the problem can be modified so that the expected one-stage reward equals  $(r(x, a) - \max_m \eta_m p_m)$  in state  $x$  and action  $a$  without changing the previous results. In this case,  $V_{\alpha,n}^*(x)$  is non-increasing in  $n$  and converges to  $V_{\alpha}^*(x)$  as  $n \rightarrow \infty$ . Therefore,  $V_{\alpha}^*$  is *e.l.d.* ( $\hat{x}_{\alpha}$  and  $b_{\alpha} = \beta$ ). Any limit point of sequence  $\pi_{\alpha,n}^*(x)$ ,  $n \geq 1$  is  $\alpha$ -discount optimal for the infinite horizon [8] (Proposition 4.3.1). Consider subsequence  $k_n$  such that  $\pi_{\alpha,k_n}^*(x) \rightarrow \pi_{\alpha}^*(x)$ . From the properties of  $\pi_{\alpha,n}^*(x)$  proved above, it is clear that  $\pi_{\alpha}^*(x)$

retains those properties. Now consider sequence  $\alpha_n$  converging to 1. Let  $k_n$  be a subsequence such that  $\pi_{\alpha_{k_n}}^* \rightarrow \pi^*$ . Proposition 3.1 (4) implies that  $\pi^*$  is average reward optimal. Since each  $\pi_{\alpha_{k_n}}^*$  has the property stated in Proposition 3.4, the proposition follows. ■

#### IV. DELAY GUARANTEES: CONSTRAINED OPTIMIZATION

Though the above formulation in itself provides a reasonable way to operate a rate-reliability adaptive scheme, the main reason behind its importance is that its optimal policy,  $\pi^*$ , can provide *delay guarantee* while maximizing the system goodput. Proving this result is undertaken in this section.

Consider the problem of maximizing goodput subject to the constraint that the average delay can be at most  $\bar{d}$ . Using Little's Law, the average delay constraint can equivalently be stated in terms of a constraint on the average queue-length (say,  $\bar{x}$ ). Thus, the problem is

$$\max_{\pi} \liminf_{n \rightarrow \infty} \frac{1}{n} E_x^{\pi} \left[ \sum_{k=0}^{n-1} s(k) \right] \quad (3)$$

where,  $s(k)$  denotes the expected number of successful transmissions in slot  $k$  and

$$\pi \in \{d | \limsup_{n \rightarrow \infty} \frac{1}{n} E_x^d \left[ \sum_{k=0}^{n-1} x(k) \right] \leq \bar{x}\} \quad (4)$$

Taking  $\beta$  to be a Lagrange multiplier, (2) may now be seen as a 'relaxation' of (3) (establishing this was precisely the reason behind choosing a linear holding cost in (1)).

We show below that the average reward optimal policy for the relaxed problem,  $\pi^*$ , is also optimal for the constrained problem (3). The result is stated formally as follows.

*Proposition 4.1:* Given a delay bound, equivalently a bound on the expected queue-length  $\bar{x}$ , let the set of feasible policies (4) be non-empty. Then there exists  $\beta_{\bar{x}} > 0$  such that the average reward optimal policy for (2) is optimal for (3).

Since the investigation in this section is centered on  $\beta$  we make the notation explicit in  $\beta$ . Thus the quantities defined for (2) will now be denoted as  $r_{\beta}(x, a)$ ,  $J_{\beta}^{\pi}(x)$ ,  $J_{\beta}^*(x)$  and  $\pi_{\beta}^*$ . Proposition 3.1 proves existence of an optimal policy,  $\pi_{\beta}^*$ , for the problem  $\max_{\pi} J_{\beta}^{\pi}(x)$  for every  $\beta > 0$ . The following is true in addition. We will denote by  $\{x_{\beta}(k), k \geq 0\}$  the queue-length process under  $\pi_{\beta}^*$ .

*Lemma 4.1:*  $\{x_{\beta}(k), k \geq 0\}$  is irreducible positive recurrent Markov chain.

*Proof:* Consequence of the facts that  $P_v(0) > 0$ ,  $\pi_{\beta}^*$  is work-conserving and [8] (Theorem 7.5.6). ■

Thus, (3) and (4) both are obtained as limits. Let

$$\bar{x}_{\beta} := \lim_{n \rightarrow \infty} \frac{1}{n} E_x^d \left[ \sum_{k=0}^{n-1} x_{\beta}(k) \right]$$

Clearly,  $\bar{x}_{\beta} = E_{\mu_{\beta}}[x_{\beta}]$ , where the expectation is with respect to the stationary distribution,  $\mu_{\beta}$ .

*Lemma 4.2:*  $\beta_1 < \beta_2 \Rightarrow m_{\pi_{\beta_1}^*}(x) \leq m_{\pi_{\beta_2}^*}(x)$ .

Lemma 4.2 basically formalizes the intuitive idea that with less penalty for delay, the optimal policy chooses modes which are

efficient from the point of view of goodput (see Figure 2). It is proved by showing super-additivity of  $\alpha$ -discounted value function and using the fact that any limit point of  $\alpha$ -discounted reward optimal policies is average reward optimal (Proposition 3.1). We omit the proof here.

*Proposition 4.2:*  $\beta \rightarrow \bar{x}_{\beta}$  is continuous over  $[\delta, \infty)$  for any  $\delta > 0$ .

*Proof:* Lemma 4.2 implies that for  $\beta_1 < \beta_2$ , transition law  $Q_{\beta_1}$  stochastically dominates  $Q_{\beta_2}$ . It follows that  $\mu_{\beta_1}$  stochastically dominates  $\mu_{\beta_2}$ . Therefore, the family of measures indexed by  $\beta \in [\delta, \infty)$  is tight and hence  $\beta \rightarrow \mu_{\beta}(x)$  is continuous for each  $x \in \mathcal{X}$  [9] (Theorem 6.1.5). Since  $\mu_{\delta-\epsilon}$ , for  $\epsilon$  small enough, stochastically dominates  $\mu_{\beta}$  for  $\beta \in [\delta, \infty)$ , stationary queue-lengths  $x_{\beta}$  are uniformly integrable. The proposition, thus, follows from [10] (Theorem 5.4). ■

*Proof of Proposition 4.1:* Proposition 4.2 proves that, given  $\bar{x}$ , there exists  $\beta_{\bar{x}}$  for which  $\pi_{\beta_{\bar{x}}}^*$  yields an average queue-length of  $\bar{x}$ . The proposition now follows from the following three conditions on  $\pi_{\beta_{\bar{x}}}^*$  [11]: (i) it yields (3) and (4) as limits (Lemma 4.1) (ii) it meets the constraint (4) with equality, and (iii) it is average reward optimal for (2) for  $\beta = \beta_{\bar{x}}$ . ■

#### A. Summary of Results

Given a server which can operate in multiple modes with different reliability levels and which serves an infinite buffer to which packets arrive in a given (i.i.d.) arrival process of rate  $\lambda$ , we have shown the following. For every  $\beta > 0$  in (1),

- 1) If the queue is stable, then there exists an optimal stationary deterministic policy which maximizes (2); stationary deterministic means that the policy specifies the mode to employ for each value of the queue-length (Proposition 3.1).
- 2) The optimal policy is such that no mode faster than a threshold mode (which depends on  $\beta$ ) is used and for all queue-lengths greater than a threshold queue-length (which depends on  $\beta$ ), the optimal mode of transmission is the threshold mode (Proposition 3.2).
- 3) The optimal policy achieves the maximum goodput among all policies which do not exceed its average delay (this is an alternate form of Proposition 4.1).

These results imply that the goodput-delay performance frontier (the set  $\Omega := \{(d, S^*(d)) | d \geq 1\}$  where  $S^*(d)$  denotes the maximum achievable goodput under the average delay constraint  $d$ ) is determined by the set of optimal policies for (2) indexed by  $\beta$ .

Note that Proposition 4.1 guarantees only the existence of  $\beta_{\bar{x}}$  for a given queue-length (or equivalent delay) constraint. Explicit computation of  $\beta_{\bar{x}}$  is an important issue and will be taken up in a longer version of this paper.

## V. NUMERICAL RESULTS

Since the state space is countable we use the approximating sequence method [8] for numerical computation. Let  $J^N$  and  $\pi^{*N}$  denote the average reward value function and optimal

$\beta$	$m^*$
0.001	4
0.01	4
0.05	5
0.1	6
0.5	7
1.0	8
2.0	10

TABLE I

VARIATION OF THRESHOLD MODES WITH PENALTY FOR DELAY ( $\beta$ ).  $\lambda = 3$  PACKETS/SLOT.

stationary policy for the truncated state space  $\{x \in \mathcal{X} | x \leq N\}$ . Then the following holds.

*Proposition 5.1:*  $J^* = \lim_N J^N$  and any limit point of  $\pi^{*N}$  is an average reward optimal policy.

We compute  $J^N$  and  $\pi^{*N}$  using a linear programming (LP) formulation [12] for the following scenario. The buffer size is taken to be 250 packets (thus,  $N = 250$ ) and the number of packets arriving in the buffer per slot is distributed binomially; thus,  $v \sim B(n, q)$  with  $n = 10$  and  $\lambda = nq$ . The server can operate in 10 modes with  $(\eta, p)$  equaling  $(1, 0.9999)$ ,  $(2, 0.99)$ ,  $(3, 0.95)$ ,  $(4, 0.9)$ ,  $(5, 0.8)$ ,  $(6, 0.7)$ ,  $(7, 0.6)$ ,  $(8, 0.4)$ ,  $(9, 0.1)$  and  $(10, 0.05)$ .

Note that in the following the phrase ‘adaptive scheme’ refers to the general scheme of adapting transmission modes according to queue-length with the objective of (2) and not necessarily to a particular optimal policy.

Table I shows the variation of the threshold mode with  $\beta$  when the arrival rate  $\lambda = 3$  packets/slot. As  $\beta$  increases, the delay component in the reward becomes more prominent and pushes for faster modes.

Figure 3 shows a comparison of the adaptive scheme with two fixed mode schemes: one in which the server operates only in mode 4 with a per packet reliability (i.e., probability of success) 0.9 (referred to as fixed mode 4 scheme) and the other, called fixed mode 6 scheme, in which the only mode of service is 6 with per packet reliability 0.7.  $\beta$  is set to 0.01 for the adaptive scheme. Figure 3 clearly shows the superiority of the adaptive scheme in terms of the reward rate (average weighted sum of goodput and delay). Note that in Figure 3 the optimal adaptive policy for each value of the arrival rate is possibly different (see the earlier remark on ‘adaptive scheme’). In fact, with increasing arrival rate, the optimal policy is required to employ more and more ‘faster’ modes to keep the delay cost low (the threshold mode correspondingly becomes higher). This may result in decreasing goodput and optimal reward rate for high arrival rates (in Figure 3 this happens for arrival rates greater than 6 packets/slot).

An advantage of the adaptive scheme is that fixed mode 4 and fixed mode 6 schemes are stable only for  $\lambda < 4$  and  $\lambda < 6$  respectively (hence, in Figure 3 the results corresponding to them have been shown only up to arrival rates 3 packets/slot and 5 packets/slot respectively) whereas the adaptive scheme can support arrival rates up to 10 packets/slot.

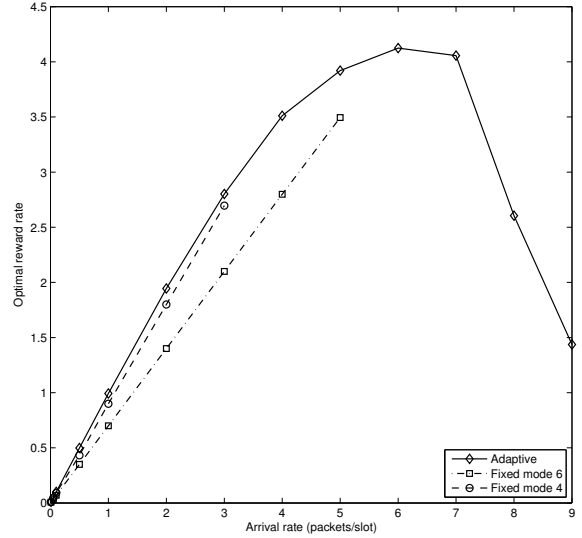


Fig. 3. Variation of the optimal average reward with arrival rate for the adaptive scheme with  $\beta = 0.01$  and two fixed mode schemes (4 and 6).

Next, to construct the goodput-delay frontier, we compute optimal policies for different values of  $\beta$  (using the LP formulation) and simulate them to calculate the corresponding values of goodput and average delay (see Section IV-A). Shown in Figure 4 is the performance frontier for  $\lambda = 3$  packets/slot. We also show comparative performance of three fixed mode schemes. Fixed mode 4 and fixed mode 6 schemes have been introduced above. In Fixed mode 5 scheme the

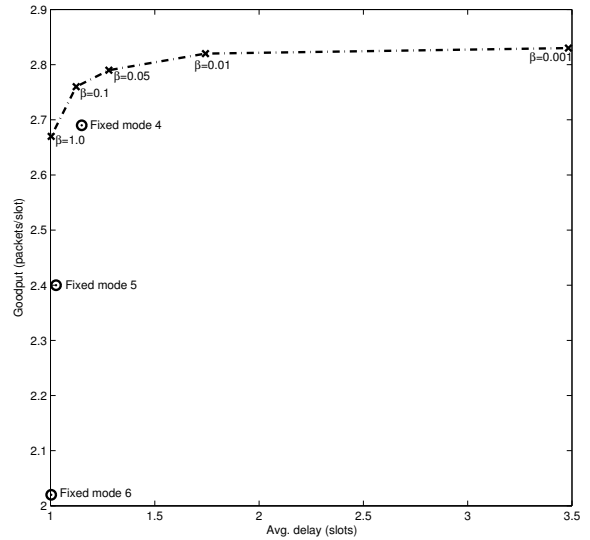


Fig. 4. The goodput-delay frontier for  $\lambda = 3$  packets/slot obtained by varying  $\beta$  for the adaptive scheme. Also shown is performance of three fixed mode schemes (4, 5 and 6) for comparison.

server operates only mode 5, thereby, always transmitting 5 packets (less if queue-length warrants it) with reliability 0.8. Observe that while fixed mode 5 and fixed mode 6 schemes achieve an average delay close to 1, their respective goodputs are roughly only 89% and 74% of that of an adaptive policy that achieves the same delay ( $\beta = 1$ ). On the other hand, fixed mode 4 scheme is close to the frontier, away by roughly 3% in goodput from an optimal policy having the same delay ( $\beta = 0.1$ ). However, if the server is forced to use only a fixed mode scheme then an average delay constraint stricter than 1.15 means that the server must use fixed mode 5 scheme and lose roughly 11% in goodput.

## VI. CONCLUSION

Rate-reliability is a prevalent and important trade-off in wireless communication system. Using adaptive burst profiling and scheduling over MIMO antennas as motivating scenarios, we formulated the rate-reliability problem as a Markov control problem and obtained structural results for the optimal policy. We also showed that the optimal policy achieves the maximum goodput among all policies which do not exceed its average delay. Therefore, the goodput-delay performance frontier is determined by the set of optimal adaptive policies. We provided supporting numerical results and computed a goodput-delay frontier.

The current formulation can be extended directly to account for the following.

- **Time-varying channel:** The formulation implicitly assumes single channel state. With multiple such states, each mode is specified by the probability of success in each of these states.
- **Markovian arrival process:** for realistically modeling the multimedia traffic.
- **Time-varying slot allocation:** Instead of a full frame/slot

for transmission, the server is allocated a time-varying fraction of the frame-time for transmission. Such allocation are typical in IEEE 802.16 PMP mode.

The linear programming approach leads to an implementable table-lookup policy. It requires the statistics of the arrival process explicitly. Development of an adaptive, learning-based scheme could be an interesting research direction.

## VII. ACKNOWLEDGMENT

The authors wish to thank Aravind Iyer for his helpful comments and suggestions.

## REFERENCES

- [1] J. Proakis. *Digital Communication* McGraw-Hill, 2000.
- [2] B. Vucetic. An adaptive coding scheme for time-varying channels. *IEEE Trans. on Communications*, vol. 39, pp. 653-663, May 1991.
- [3] H. Matsuoka, S. Sampei, N. Morinaga, Y. Kamio. Adaptive Modulation System with Variable Coding Rate Concatenated Code for High Quality Multi-Media Communication Systems. *Proceedings of the IEEE VTC'96*, pp. 487-491, April 1996.
- [4] S.-G. Chua and A.J. Goldsmith. Adaptive coded modulation for fading channels. *IEEE Trans. on Communications*, pp. 595-602, May 1998.
- [5] O. Hernandez-Lerma and S. Marcus. Adaptive Control of Service in Queueing Systems. *Systems & Control Letters*, vol. 3, pp. 283-289, Nov 1983.
- [6] R. Weber and S. Stidham. Optimal Control of Service Rates in Networks of Queues. *Adv. in Applied Probability*, vol. 19, pp. 202-218, 1987.
- [7] J. George and J. M. Harrison. Dynamic Control of a Queue with Adjustable Service Rate. *Operations Research*, vol. 49(5), pp. 720-731, Sep 2001.
- [8] L. Sennott. *Stochastic Dynamic Programming and the Control of Queueing Systems*. Wiley Series in Probability and Statistics, 1999.
- [9] G. Fayolle, V. Malyshev and M. Menshikov. *Topics in the Constructive Theory of Countable Markov Chains*. Cambridge University Press, 1995.
- [10] P. Billingsley. *Convergence of Probability Measures*. New York: Wiley, 1968.
- [11] D. Ma, A. Makowski and A. Shwartz. Estimation and Optimal Control for Constrained Markov chains. *Proceedings of the CDC*, pp. 994-999, December 1986.
- [12] D. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.