

Optimal Buffer Scheduling over a Fading Channel with an Average Power Constraint: The Single User Case

Munish Goyal, Anurag Kumar and Vinod Sharma

Department of Electrical Comm. Engg., Indian Institute of Science, Bangalore, India.

{munish, anurag, vinod}@ece.iisc.ernet.in

Abstract — We consider an optimal power and rate scheduling problem for a single user transmitting to a base station on a fading wireless link with the objective of providing transmission delay guarantees. The base station acts as a controller which, depending upon the transmitter buffer lengths and the signal power to interference ratio (SIR) on the uplink pilot channel, allocates transmission rate and power to the user. We provide structural results for an average cost optimal policy under a long run average transmitter power constraint. We obtain a closed form expression relating the optimal policy when the SIR is the best, to the optimal policy for any other SIR value.

I. INTRODUCTION

In communication systems, many fundamental problems involve the management and allocation of resources. In a wired network, the crucial resources are the transmission data rates available on the link. Techniques such as flow control, routing and admission control are all centered around allocating these resources. We consider a resource allocation problem that arises in mobile wireless communication systems. Several challenging and exciting analytical problems arise because of the unique feature of a wireless link. The first is the time varying nature of the multipath channel, and the second is the limited battery power available at a typical wireless handset. Thus it is desirable to allocate the transmission rates to the user such that the energy used to transmit the information is minimized while keeping the errors under control. This would be sufficient if one is interested only in store and forward data traffic that can handle arbitrary large delays. For transmitting telephone quality interactive packet voice and other real time traffic, however, the transmission scheduling mechanism should take care of quality of service (QoS) objectives such as mean delay, delay jitter and the throughput. Thus there is a need for optimal allocation of wireless resources which provides such QoS guarantees subject to the above said error and energy constraints.

Various methods for allocating transmission resources are part of most third generation cellular standards. They include adjusting the transmission power, changing the coding rate and varying the spreading gain in a CDMA based system. There are some restriction on how much they can vary. In this work we assume that the transmitter can transmit at any arbitrary rate and power level. The system model is given in Fig. 1 and is explained below. We assume a slotted system where the higher layer presents the data, that arrives over a slot, to the link layer at the end of each slot. The link layer is assumed

to have an infinite capacity buffer to hold the data. The buffer length information is communicated to the receiver/controller at the end of each slot. We assume that the channel gain and the interference from other users transmission remains fixed over a slot and varies independently from slot to slot. For a small time slot shown as shaded in Fig. 1, the user transmits pilot bits at a fixed power level which we refer to as a pilot channel. The receiver estimates the signal to interference ratio (SIR) on the pilot channel. We assume that the estimates are perfect. Depending on the SIR estimates and the buffer length information, the receiver evaluates the optimal transmission rate and power for the current slot and communicates it back to the transmitter. The transmitter removes that much amount of data from the buffer and encodes it at the allocated rate. All this exchange of information and the encoding is assumed to be completed within the time slot shown as shaded in the Fig. 1. After this the transmitter starts to transmit the encoded data.

Goldsmith and Varaiya [2] are probably the first to obtain the optimal resource allocation policy for a single link fading wireless channel. There emphasis was on the optimal physical layer performance, while ignoring the network layer performance such as delay guarantees. In recent work, Berry and Gallager [1] have considered a problem similar to ours. They obtained structural result exhibiting a tradeoff between the network layer and physical layer performance, i.e. the optimal power and delay. They show that the optimal power vs the optimal delay curve is convex, and as the average power available for transmission increases, the mean delay decreases. They also provide some structural results for the optimal policy that achieves any point on the power delay curve. In this work, we improve upon the results obtained in [1]. We prove the existence of stationary average optimal policy, and give a closed form expression for the optimal policy for any SIR value in terms of the optimal policy when the SIR is one, i.e, the best SIR.

II. PROBLEM FORMULATION

We consider a fluid model for our analysis and will later show how it can be used for a packetized data model. Let, for $n \in \{0, 1, 2, \dots\}$, $s[n]$ be the amount of fluid in the buffer at the n^{th} decision epoch; let $\gamma[n]$ be the SIR and $a[n]$ be the amount of fluid arriving in the n^{th} slot (i.e., the interval $[n\tau, (n+1)\tau]$). We assume the arrival and the SIR process to be iid; let $F(a)$ and $G(\gamma)$ be the respective distribution function. According to our model, if in a slot n , the user transmits a signal $y_s[n]$, then the receiver gets $y_r[n] = h[n]y_s[n] + \zeta[n]$,

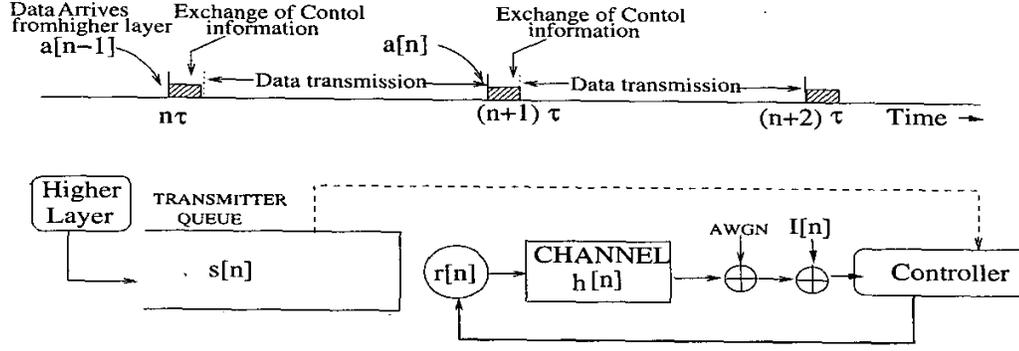


Figure 1: System model

where $\zeta[n]$ constitutes the additive white Gaussian noise and the other users' interference signal. Let σ^2 be the ambient noise power. The SIR comprises of the channel gain ($h \in (0, 1)$), ambient noise power and the interference ($I \in [0, \infty)$) from other user transmissions, i.e., $\gamma = \frac{\sigma^2 h}{\sigma^2 + I}$. Thus it follows that $\gamma \in (0, 1)$. Let the state of the system be represented as $x[n] := (s[n], \gamma[n])$. At the n^{th} decision instant, the controller decides upon the amount of fluid $r[n]$ to be transmitted in the current slot depending on the entire history of state evolution, i.e., $x[k]$ for $k = \{1, 2, \dots, n\}$. Naturally $r[n] \leq s[n]$, for all n . Now the buffer state evolution equation can be written as $s[n+1] = s[n] - r[n] + a[n]$. We assume N channel uses in a slot obtained from the Nyquist theorem. In state x , the energy cost for serving r units of fluid in a slot is $P(x, r) = \frac{\sigma^2}{\gamma} (e^{\theta r} - 1)$, where $\theta = \frac{2 \ln(2)}{N}$. The controller's objective is to obtain the optimal $r(\cdot)$ that minimizes the long run average amount of data in the buffer

$$\limsup_n \frac{1}{n} E \sum_{k=0}^n s[k],$$

subject to the average power constraint,

$$\limsup_n \frac{1}{n} E \sum_{k=0}^n P(x[k], r[k]) \leq \bar{P}.$$

We can restate the above objective as a constrained Markov decision problem. Consider a controlled Markov chain $\{X[n], n \in \{0, 1, 2, \dots\}\}$ with state space $\mathcal{X} = \mathcal{R}^+ \times (0, 1]$. The set of feasible actions in state $x = (s, \gamma)$ is $[0, s]$. Define a set of feasible state-action pair \mathcal{K} as $\{(x, a) | x = (s, \gamma) \in \mathcal{X}, a \in [0, s]\}$. The transition kernel on \mathcal{X} given \mathcal{K} is denoted by Q , where

$$Q(y \in (S', \Gamma') \subset \mathcal{X} | (x, r)) = \int_{S' - s + r} dF(y) \int_{\Gamma'} dG(\gamma).$$

A policy π generates at time n an action $R[n]$ depending upon the entire history of the process, i.e., at decision instant $n \in \{0, 1, 2, \dots\}$, π_n is a mapping from $\mathcal{K}^n \times \mathcal{X}$ to $[0, s[n]]$. Let Π be the space of all such policies. A stationary policy $f \in \Pi$ is a measurable mapping from \mathcal{X} to $[0, s]$. For a policy $\pi \in \Pi$, define $B^\pi = \limsup_n \frac{1}{n} E^\pi \sum_{k=0}^n S[k]$ and $K^\pi =$

$\limsup_n \frac{1}{n} E^\pi \sum_{k=0}^n p(X[k], R[k])$. Given $\bar{P} > 0$, denote by $\Pi_{\bar{P}}$ the set of all admissible control policies $\pi \in \Pi$ which satisfy the constraint $K^\pi \leq \bar{P}$. Thus the controller objective can be restated as a constrained control problem (CP) defined as,

$$(CP) : \text{Minimize } B^\pi \text{ subject to } \pi \in \Pi_{\bar{P}} \quad (1)$$

III. ANALYSIS

We convert the constrained problem (CP) into a family of unconstrained problems (UP_β) through a Lagrangian approach and use the results stated in Theorem III.1 to obtain the solution of (CP). The problem (UP_β) is shown to be a standard Markov decision problem with an average cost criterion. We define a corresponding discounted cost MDP with discount factor α . We study the average cost problem as a limit of these discounted cost problems when the discount factor α goes to one.

Definitions:

- Define a mapping $c_\beta : \mathcal{K} \rightarrow \mathcal{R}^+$ by

$$c_\beta(x, r) = s + \beta \frac{\sigma^2}{\gamma} (\exp^{\theta r} - 1)$$

- Define a Lagrangian functional for any policy $\pi \in \Pi$ as

$$J_\beta^\pi = \limsup_n \frac{1}{n} E^\pi \sum_{k=1}^n c_\beta(X[k], R[k])$$

- The unconstrained problem (UP_β), for $\beta > 0$ is,

$$(UP_\beta) : \text{Minimize } J_\beta^\pi \text{ subject to } \pi \in \Pi \quad (2)$$

- A stationary optimal rate allocation policy is $r(x)$, i.e., when the state is x , the optimal transmission rate is $r(x)$.
- Define $u(x) = s - r(x)$, i.e., $u(x)$ is the optimal amount of data that should be left unserved when the state is x
- The total expected discounted cost

$$V_\alpha(x) = \min_{\pi \in \Pi} E_x^\pi \left[\sum_{k=0}^{\infty} \alpha^k c_\beta(X[k], R[k]) \right]$$

- Define $\eta = E\left[\frac{e^{\theta A}}{\Gamma}\right]$
- Given $\epsilon > 0$, define Q as a solution to

$$E\left(Q - \frac{1}{\theta} \ln\left(\frac{1}{\Gamma}\right)\right)^+ > E(A) + \epsilon$$

Theorem III.1 [3] *Let, for some $\beta > 0$, $\pi^* \in \Pi$ be the policy that solves the unconstrained problem (UP $_{\beta}$). If π^* yields the expressions B^{π^*} and K^{π^*} as limits and $K^{\pi^*} = \bar{P}$, then the policy π^* necessarily solves the constrained problem (CP).*

Proof: See [3].

The discounted cost optimality equation (DCOE) that solves for the discounted cost problem stated above is,

$$V_{\alpha}(x) = \min_{0 \leq r \leq s} \{c_{\beta}(x, r) + \alpha \int_{\mathcal{X}} V_{\alpha}(y) Q(dy|(x, r))\} \quad (3)$$

However, from the definition of Q it follows that $Q(dy|(x, r)) = Q(dy|u)$. Thus we can rewrite the DCOE (Equation 3) in terms of u as,

$$V_{\alpha}(x) = \min_{u \in [0, s]} \left\{s + \frac{\beta \sigma^2}{\gamma} (e^{\theta(s-u)} - 1) + \alpha H_{\alpha}(u)\right\}, \quad (4)$$

where $H_{\alpha}(u)$ is given by,

$$H_{\alpha}(u) := \int_0^{\infty} \int_0^1 V_{\alpha}(u+a, \gamma) dF(a) dG(\gamma) \quad (5)$$

Theorem III.2 *Let $u_{\alpha_n}(s, 1)$ be the minimizer of the right hand side of the following equation,*

$$V_{\alpha_n}(s, 1) = \min_{u \in [0, s]} \left\{s + \beta \sigma^2 (e^{\theta(s-u)} - 1) + \alpha_n H_{\alpha_n}(u)\right\}.$$

- There exists a stationary average cost optimal policy $u(x)$.
- Given any sequence of discount factor converging to one, there exists a subsequence $\{\alpha_n\}$ such that for any $s \in S$, the average optimal policy $u(s, 1) = \lim_n u_{\alpha_n}(s, 1)$, i.e., the choice of subsequence does not depend upon the choice of s . The policy $u(s, 1)$ is monotonic nondecreasing and Lipschitz with parameter one
- Given any $x = (s, \gamma) \in \mathcal{X}$, the average optimal policy $u(x)$ representing the amount of data not served when in state x , is

$$u(x) = \min\left(s, u\left(s + \frac{1}{\theta} \ln\left(\frac{1}{\gamma}\right), 1\right)\right)$$

- If $\eta < \infty$, then $\frac{s}{2} - \frac{1}{2\theta} \ln(\kappa_2) \leq u_1(s) \leq s - \frac{1}{\theta} \ln(\kappa_1)$, where κ_1 and κ_2 are constants
- If $Q < \infty$, then there exists a Lagrange multiplier $\beta > 0$ such that the optimal policy for (UP $_{\beta}$) is also optimal for the constrained problem (CP)

Remark A nearly optimal policy for the packetized model is obtained via the randomization between two fluid optimal policies.

Sketch of the Proof:

- In order to prove the existence of a stationary average optimal policy, we make use of a result obtained in [4]. We show that the model satisfies the hypothesis (W) of the Theorem 3.8 [4] thus proving the existence of stationary average optimal policy and that the average optimal policy $u(x)$ can be obtained as a limit of discount optimal policies $u_{\alpha}(x)$. Further, $u(x)$ is *limit discount optimal* in the sense that, for any $x \in \mathcal{X}$ and given any sequence of discount factors converging to one, there exist a subsequence $\{\alpha_m\}$ of discount factors and a sequence $x_m \rightarrow x$ such that $u(x) = \lim_{m \rightarrow \infty} u_{\alpha_m}(x_m)$. The subsequence depends upon the choice of x
- To prove this we use the properties of discounted optimal policies generated as a solution to Equation 4. Firstly, it is easy to show that the function $V_{\alpha}(s, \gamma)$ is convex and increasing in s ; thus $H_{\alpha}(u)$ is convex, increasing and the problem is a constrained convex programming problem. Thus using standard techniques for solving a constrained convex programming problem, we obtained the following result. For each $\alpha \in (0, 1)$,

- $u_{\alpha}(x) = 0$ for $\{x \in \mathcal{X} : \frac{\beta \theta \sigma^2}{\alpha \gamma} e^{\theta s} \leq H'_{\alpha}(0)\}$.
- $u_{\alpha}(x) = s$ for $\{x \in \mathcal{X} : H'_{\alpha}(s) \leq \frac{\beta \theta \sigma^2}{\alpha \gamma}\}$
- Else $u_{\alpha}(x)$ is the solution of $\frac{\sigma^2}{\gamma} e^{\theta s} = e^{\theta u} \frac{\alpha H'_{\alpha}(u)}{\beta \theta}$.

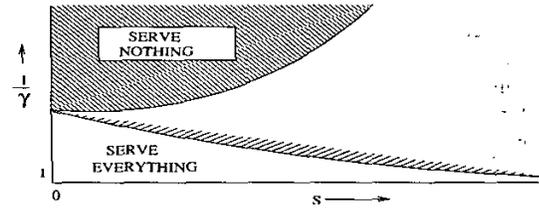


Figure 2: Characterization of the optimal policy for the discounted cost MDP

Observations: For the discounted cost problem,

- It is optimal not to serve anything when the SIR estimated is low.
- When the SIR estimated at the receiver is high, it is rather optimal to serve everything till a larger value of the buffer size.
- In the low SIR region, the amount of data served increases with the buffer size since the delay cost exceeds the power cost for transmitting at low SIR.

Now we show that the discount optimal policy for any SIR γ can be calculated by knowing the optimal policy

when the SIR is 1. Consider two states x_1 and x_2 . It follows from the previous result that except for the case when $u_\alpha(x) = s$, the controls $u_\alpha(x_1)$ and $u_\alpha(x_2)$ are the same if $\frac{\sigma^2}{\gamma_1} e^{\theta s_1} = \frac{\sigma^2}{\gamma_2} e^{\theta s_2}$. Thus we can compute the optimal policy $u_\alpha(x)$ for any $x \in \mathcal{X}$ by knowing the optimal policy for the case when γ is fixed to one and only s is allowed to vary. In order to compute $u_\alpha(s, \gamma)$, we first obtain s_1 such that $e^{\theta s_1} = e^{\theta s} \frac{1}{\gamma}$ and the optimal policy $u_\alpha(s, \gamma)$ can be written in terms of $u_\alpha(\cdot, 1)$ as,

$$u_\alpha(s, \gamma) = \min \left\{ s, u_\alpha \left(s + \frac{1}{\theta} \ln \left(\frac{1}{\gamma} \right), 1 \right) \right\}. \quad (6)$$

Now by simple interchange argument and using the convexity of $H_\alpha(u)$, we can show that both $r_\alpha(s, 1)$ and $u_\alpha(s, 1)$ are monotonic nondecreasing. But since they should sum to s , the policy $u_\alpha(s, 1)$ is monotonic nondecreasing and Lipschitz with parameter one. Note that this implies globally equi-Lipschitz property for any sequence of functions $\{u_{\alpha_n}(s, 1)\}_{n=1}^\infty$. Using this property we can strengthen the result of Theorem 3.8 [4]. The new result is, given $s \in S$ and any sequence of discount factors converging to one, there exists a subsequence $\{a_n\}$ such that $u(s, 1) = \lim_n u_{\alpha_n}(s, 1)$. Now it follows from the monotonic nondecreasing nature of $u_\alpha(s, 1)$ that the optimal policy $u(s, 1)$ is monotonic nondecreasing and Lipschitz with parameter one. Again using these properties of $u(s, 1)$ and globally equi-Lipschitz property of discount optimal policies, we can further strengthen the result in a following way. Given any sequence of discount factors converging to 1, there exists a subsequence $\{a_n\}$ such that for any $s \in S$, the average cost optimal policy

$$u(s, 1) = \lim_n u_{\alpha_n}(s, 1) \quad (7)$$

The choice of subsequence does not depend upon the choice of s . The proof essentially uses the Cantor diagonalization procedure to obtain a subsequence that works for a dense subset of the state space and then uses the globally equi-Lipschitz property to show that the subsequence works for the whole state space.

- (iii). This follows by passing through the limit in Equation 6. Note that this follows only because of our stronger result of Theorem III.2(II).
- (iv). We first prove bounds for $H'_\alpha(u)$. Consider a feasible policy that serves everything, i.e., $u_\alpha(x) = 0$ for all $x \in \mathcal{X}$. For this policy we have,

$$V_\alpha(x) \leq s + \frac{\beta \sigma^2}{\gamma} (e^{\theta s} - 1) + \alpha H_\alpha(0)$$

Also as $H_\alpha(u)$ is increasing, we have $V_\alpha(x) \geq s + \alpha H_\alpha(0)$ independent to the choice of the policy. Thus $H_\alpha(u) \leq u + E[A] + \eta \beta \sigma^2 e^{\theta u} + \alpha H_\alpha(0)$ and $H_\alpha(u) \geq u + E[A] + \alpha H_\alpha(0)$. Now since $H_\alpha(u)$ is convex, we get $H'_\alpha(u) \leq 1 + \theta \beta \eta e^{\theta u}$.

Now we show that $H'_\alpha(u) \geq 1$. When the buffer is empty, the optimal policy is $u_\alpha(0, \gamma) = 0$ for all γ . Thus

$V_\alpha(0, \gamma) = \alpha H_\alpha(0)$. Also $V_\alpha(s, \gamma) \geq s + \alpha H_\alpha(0)$. Thus we have

$$V'_\alpha(0, \gamma) = \lim_{\epsilon \rightarrow 0} \frac{V_\alpha(\epsilon, \gamma) - V_\alpha(0, \gamma)}{\epsilon} \geq 1$$

Thus $V'_\alpha(s, \gamma) \geq 1$ for all (s, γ) . Using the dominated convergence theorem we can take the differentiation inside of the integral sign in Equation 5. Thus $H'_\alpha(u) \geq 1$. Now from the solution to the convex problem, the bounds for $H'_\alpha(u)$ and the convergence result of Theorem III.2(II), the result follows.

- (v). We need to show the existence of a $\beta > 0$ for which the hypothesis of the Theorem III.1 holds. We consider the instantaneous cost function $c_\beta(x, r)$. As the parameter β is increased, power becomes more expensive and hence the average amount of fluid in the buffer increases. It has earlier been shown [1] that as the delay increases, the power required decreases and the power-delay tradeoff curve is convex. Thus, each point on the curve can be obtained with a particular choice of β , i.e., there exists a $\beta_0 > 0$ such that $K'' = \bar{P}$. But we have this result only in the lim sup sense. To show that the lim sup and the lim inf are equal, we prove that the controlled chain is ergodic under the policy obtained for this choice of β_0 .

We use the negative drift argument to obtain sufficient conditions for the ergodicity of the controlled chain $\{X[n]\}$. Given $\epsilon > 0$, it is sufficient to find a compact set $C \subset X$ and a $k > 0$ such that for $(s, \mathbf{y}) \in C$ following holds,

$$E \left((S_{n+k} - S_n) | (X[n] = (s, \gamma), u(s, \gamma)) \right) < -\frac{\epsilon}{2} \quad (8)$$

Lemma 111.1 *There exists a $\beta_0 > 0$ for which the controlled chain is ergodic.*

Sketch of the Proof: We first show that for Equation 8 to be satisfied for some compact set, it is sufficient to find an $s_0 < \infty$ such that $r(s_0, 1) > Q$. Then using the discounted cost value iteration algorithm we show that for discount factor α , there exist an $s_0(\alpha) < \infty$ such that for all $s > s_0(\alpha)$

$$r_\alpha(s, 1) > \frac{1}{\theta} \ln \left(\frac{\alpha}{(1 - \alpha)\beta\theta\sigma^2} \right) - \epsilon$$

Now using the convergence result of Theorem III.2(II), we get an $s_0 < \infty$ satisfying $r(s_0, 1) > Q$. Hence done.

REFERENCES

- [1] Randall A. Berry, R. G. Gallager, "Communication over Fading Channels with Delay Constraints," *IEEE Transaction on Information Theory*, May 2002.
- [2] A. J. Goldsmith, P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Transaction on Information Theory*, Nov 1997.
- [3] D I. Ma, A.M.Makowski, A.Schwartz, "Estimation and Optimal Control for Constrained Markov Chains," *IEEE Conf. on Decision and Control*, Dec 1986.
- [4] Manfred Schäl, "Average Optimality in Dynamic Programming with General State Space," *Mathematics of Operations Research*, Feb 1993.