Stochastic Control for Energy Efficient Resource Allocation in Wireless Networks

Abhijeet Bhorkar
(01D07014)

under the guidance of
Prof. A. Karandikar

Indian Institute of Technology Bombay,
Powai, Mumbai-76

July 05, 2006

Outline

Motivation

- Wireless LANs, ad-Hoc networks, sensor networks
  limited battery life, bandwidth
- Maintain acceptable QoS metric
  rate, delay, fairness
- Efficient utilization of limited resources
- Time varying wireless channel conditions
- Exploit channel variations
  - opportunistic scheduling :
    Schedule user with the best channel condition
- Can we perform power control to exploit channel variations?
  Joint opportunistic and energy efficient control

Motivation

- Wireless LANs, ad-Hoc networks, sensor networks
  limited battery life, bandwidth
- Maintain acceptable QoS metric
  rate, delay, fairness
- Efficient utilization of limited resources
- Time varying wireless channel conditions
- Exploit channel variations
  - opportunistic scheduling :
    Schedule user with the best channel condition
- Can we perform power control to exploit channel variations?
  Joint opportunistic and energy efficient control

Motivation

- Wireless LANs, ad-Hoc networks, sensor networks
  limited battery life, bandwidth
- Maintain acceptable QoS metric
  rate, delay, fairness
- Efficient utilization of limited resources
- Time varying wireless channel conditions
- Exploit channel variations
  - opportunistic scheduling :
    Schedule user with the best channel condition
- Can we perform power control to exploit channel variations?
  Joint opportunistic and energy efficient control

Scenario

Power optimal scheduling

- Minimum rate guarantee (Multi-user)
- Fairness guarantee (Multi-user)
- Average delay guarantee with finite buffer (point to point link)

Preliminaries(1)

Shannon's capacity:

$$P = \frac{N_0 W}{x}\left(e^{u/W} - 1\right),$$

N_0	:	Spectral density of AWGN channel
W	:	Spectrum bandwidth
u	:	Transmission rate

Stochastic approximation:

$$\lambda(n+1) = \lambda(n) + \alpha(n)\left(H(\lambda(n)) + M(n+1)\right) \quad (1)$$

$$\mathbf{E}[h(\lambda, x)] = H(\lambda)$$

$$\text{Martingale } M(n+1) = h(\lambda(n), x(n)) - H(\lambda(n))$$

If step sizes satisfy , $\sum_{n=0}^{\infty} \alpha(n) = \infty, \quad \sum_{n=0}^{\infty} \alpha(n)^2 < \infty$

then (1) tracks the Ordinary Differential equation (ODE),

$$\dot{\lambda}(t) = H(\lambda(t)) \quad (2)$$

System Model

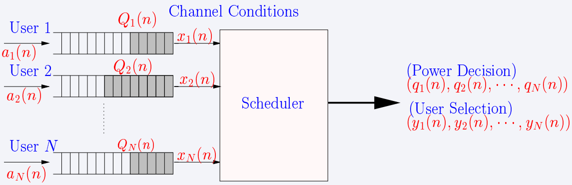


Figure: Single hop system model

- Slotted single-hop TDMA system
- Uplink scheduling
- Perfect channel state information
- Channel process ergodic (i.i.d. or Markovian)

Problem Formulation

- Minimize average power

$$\min \quad \limsup_{M\to\infty} \frac{1}{M}\sum_{n=1}^{M} q(n),$$

- Subject to average rate constraints C_i

$$\begin{aligned} \limsup_{M\to\infty} \frac{1}{M}\sum_{n=1}^{M} U_i(q_i(n),x_i(n)) &\geq C_i \quad \forall i, \\ q(n) &\geq 0, \\ \sum_{i=1}^{N} y_i(n) &\leq 1 \quad \forall n \end{aligned} \qquad (3)$$

  - U is information theoretic rate and is concave differentiable function of x_i, q_i

    $U = \log(1 + x_i q_i)$
  - $x = (x_1, x_2, \cdots, x_N)$
  - $y = (y_1, y_2, \cdots, y_N)$

Problem Formulation

- Minimize average power

$$\min \quad \limsup_{M\to\infty} \frac{1}{M}\sum_{n=1}^{M} q(n),$$

- Subject to average rate constraints C_i

$$\begin{aligned} \limsup_{M\to\infty} \frac{1}{M}\sum_{n=1}^{M} U_i(q_i(n),x_i(n))) &\geq C_i \quad \forall i, \\ q(n) &\geq 0, \\ \sum_{i=1}^{N} y_i(n) &\leq 1 \quad \forall n \end{aligned} \qquad (3)$$

  - U is information theoretic rate and is concave differentiable function of x_i, q_i

    $U = \log(1+x_iq_i)$
  - $x = (x_1, x_2, \cdots, x_N)$
  - $y = (y_1, y_2, \cdots, y_N)$

Multiuser Optimal Solution

Proposition

Optimal Policy for multiple users is to select k^{th} user and transmit with power $q*$

Sketch of Proof

- Use ergodicity to convert optimization problem (3) in continuous domain
- Minimize Lagrangian of (3) w.r.t. q first, then w.r.t. y
- Optimal power for single user,
  $q_i^* = \left(\lambda_i - \frac{1}{x_i}\right)^+$,where λ_i is the Lagrange multiplier
- Minimizing w.r.t. y, we get,

$$k = \arg\min_i \left(q_i^* - \lambda_i \left[\log(1 + q_i^* x_i) - C_i\right]\right)$$

▸ Details

Online Algorithm (1)

- After minimizing over the primal variables, optimal value of Lagrangian is,

$$F(\lambda) = [E(\min_i(q_i^* - \lambda_i \log(1 + q_i^* x_i(n)) - C_i)]$$

- $F(\lambda)$ strictly concave $\rightarrow$ unique maximum
- Need to find saddle point

- Consider for example $f(x)$ is continuous differentiable
- Gradient ascent scheme for maximizing f is,

$$x_{n+1} = x_n + \alpha_n \dot{f}(x_n)$$

- It tracks the differential equation,

$$\dot{x}(t) = \frac{df}{dx} = \dot{f}$$

Online Algorithm (2)

- Estimate λ_i online

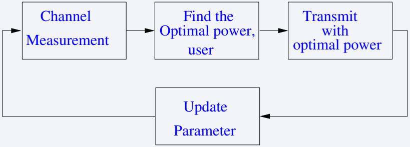


Figure: Block diagram for on-line policy

- Update Equation

$$\lambda_i(n+1) = \{\lambda_i(n) - \alpha(n)[\underbrace{y_i(n)\log\left(1+\left(\lambda_i(n)-\frac{1}{x_i(n)}\right)^+ x_i(n)\right)}_{h_i(\lambda)}] - C_i\}^+ \quad \forall i, \tag{4}$$

Optimality and Stability of Update Equation

- Iterations converge to differential inclusion almost surely to ,

$$\dot{\lambda}(t) = h(\lambda(t))$$

  and thus to a supergradient ascent scheme

$$\dot{\lambda}(t) \in \partial F(\lambda(t))$$

  ∂F supergradient of F
- Stability
  - Boundedness of λ_i using projection method or linear stochastic approximation method

Simulations

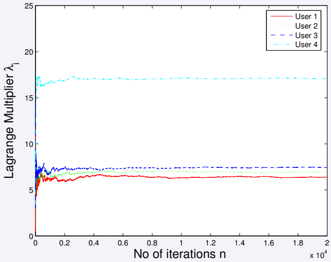


Figure: Convergence for Markovian channel $\alpha = 0.3$, $\mathbf{C} = (0.6, 0.8, 0.7, 0.2)$, $\lambda(0)$ = $(1,1,1,1)$ and $\sigma = (1,1,0.9,0.3)$

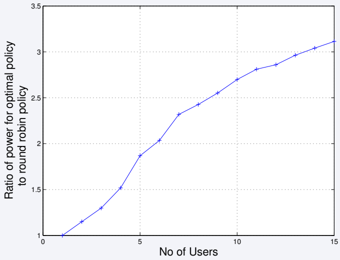


Figure: Gain of the optimal policy over variable power round robin policy, C=0.6, $\gamma = 0.7$

Temporal Fairness

- Time as resource
- Different users receive different time resource
  - Long term fair: Proportional share on long run
    - Starvation
    - HOL blocking
  - Short term fair: Proportional share on finite window size M

Temporal Fairness

- Time as resource
- Different users receive different time resource
  - Long term fair: Proportional share on long run
    - Starvation
    - HOL blocking
  - Short term fair: Proportional share on finite window size M

Temporal Fair scheduler

- Long term temporal fair
  - ϕ_i temporal share to user i on infinite run

$$\liminf_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} \mathbf{E}\, y_i(n) \geq \phi_i \quad \forall i.$$

  - Modified update equation

$$\begin{aligned}\lambda_i(n+1) &= \Big[\lambda_i(n) - a(n)\Big[y_i(n)\log\Big(\\ &\quad 1+\Big(\lambda_i - \frac{1}{x_i(n)}\Big)^+ x_i(n)\Big)\Big] - C_i\Big]^+ \\ \lambda_i'(n+1) &= [\lambda_i'(n) - a(n)(y_i(n) - \phi_i)]^+ \quad \forall i\end{aligned}$$

- Short term temporal fair
  - ϕ_i temporal share on finite window of M
  - Heuristic: Elimination policy
    Remove the user if its share is exhausted

Simulations

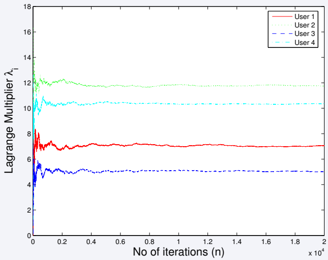


Figure: Trajectory of $\lambda_i(n)$ $\phi = (0.3, 0.4, 0.2, 0.1)$

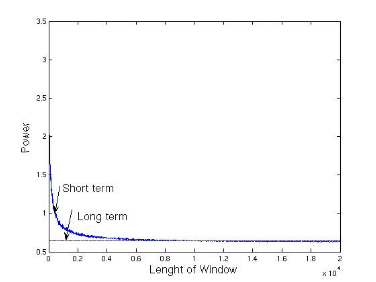


Figure: Power required for the short term and long term fairness

Short term Thoughput

Preliminaries(2)

Average cost Markov Decision Process:

$$\text{Path wise Average Cost} = \limsup_{N\to\infty} \mathbf{E}\left[\frac{1}{N}\sum_{n=0}^{N-1} c(s_n, u_n)\right]$$

$$\phi^* + h(s) = \min_{u\in U}\{c(s,u) + P(\bar{s}|s,u)h(\bar{s})\}$$

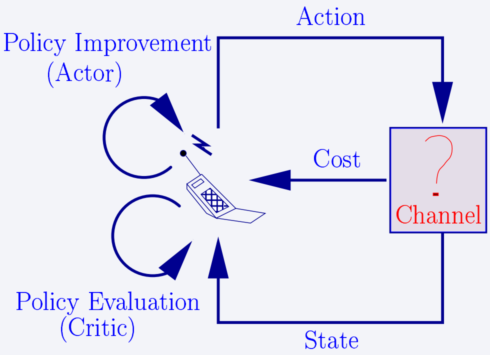
Bellman equation with unique $h(s)$ exists iff,

$$h(s) = \min_{u\in U}\left\{c(s,u) + \mathbf{E}h(\bar{s}) - h(s^0)\right\} \tag{5}$$

where,

- $s \in S$: State of the process
- $u \in U$: Action taken
- c : Immediate cost
- $\bar{s}$: Next state
- h : Difference value function
- ϕ : Average cost
- $s^0 \in S$: Some arbitrary state
- $P(\bar{s}|s,u)$:Transition matrix (kernel)

Reinforcement Learning



Learn how to take action in each state to minimize the cost

System Model and Problem Formulation

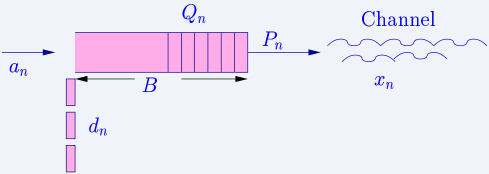


Figure: Transmitter of a point to point link

Optimization problem

$$\text{Minimize } P = \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} P_n, \quad (6)$$

$$\text{Subject to } D - \bar{D} \leq 0 \quad \text{Delay constraint} \quad (7)$$

$$\varepsilon - \bar{\varepsilon} \leq 0^a \quad \text{Drop probability constraint}$$

[a] We denote the constraint by ‘¬’ .

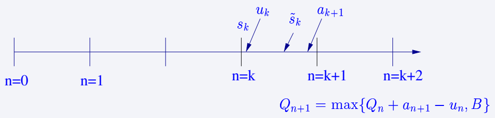


Figure: System Model

By Little's law (7) is,

$$Q_{avg} - \bar{a}\bar{D} \leq 0$$

where,

$$\begin{aligned}
\varepsilon &= \limsup_{N\to\infty} \frac{\sum_{n=0}^{N} \overbrace{\max(0, Q_n - u_n + a_{n+1} - B)}^{d_n}}{\sum_{n=0}^{N} a_n} \\
\bar{a} &= a_{avg}\,(1-\varepsilon) \quad \text{(Effective arrival rate)} \\
Q_{avg} &= \limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N} Q_n, \\
a_{avg} &= \limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N} a_n
\end{aligned}$$

Discrete State Space Markov Decision Formulation : Post Decision

The immediate cost of the constrained MDP is given by,

$$\begin{aligned} c_n &= P_n + \lambda_1 (Constraint_1) + \lambda_2 (Constraint_2) \\ Constraint_1 &: (Q_n - \bar{D}(a_{n+1} - d_n)) \\ Constraint_2 &: (d_n - \bar{\varepsilon} a_{n+1}) \end{aligned}$$

c_n is convex function of u_n and s_n. Hence no duality gap.

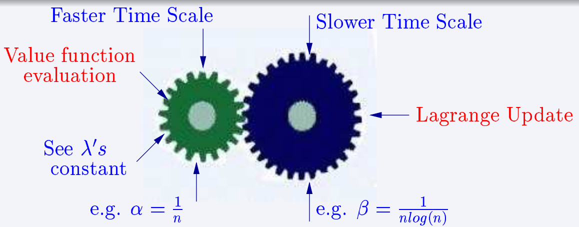
- Post decision state $\tilde{s}$:State after decision is taken

Bellman Equation (5) using the post-decision state is given as,

$$\tilde{h}_{n+1}(\tilde{s}) = \underbrace{\mathbf{E}\left[\min_{u \in U}\right.}_{\text{Interchange of min and E operator}} \left.\left\{ c(s, \lambda, u) + \tilde{h}_{n+1}(\tilde{s}|\tilde{s}_n) - \tilde{h}(\tilde{s}^0) \right\}\right]$$

Interchange of $\min$ and $\mathbf{E}$ operator

Two time scale update



Post learning: Faster time scale

$$\tilde{h}_{n+1}(\tilde{s}) = \tilde{h}_n(\tilde{s}) + \alpha(\nu(\tilde{s},n)) I\{\tilde{s}_n = \tilde{s}\} \left[\min_{u \in U} \left\{ c(s,\lambda_n,u) + \tilde{h}_n(\tilde{\tilde{s}}) \right\} - \tilde{h}_n(\tilde{s}_0) - \tilde{h}_n((\tilde{s}) \right] \quad \forall \tilde{s}$$

Lagrangian update: Slower time scale

$$\lambda_{1_{n+1}} = \Gamma_1[\lambda_{1_n} + \beta(n)(Q_n - \bar{D}(a_n - d_n))]$$
$$\lambda_{2_{n+1}} = \Gamma_2[\lambda_{2_n} + \beta(n)(d_n - \bar{\varepsilon} a_n)]$$

Γ_1, Γ_2 are projection operators.

Simulations

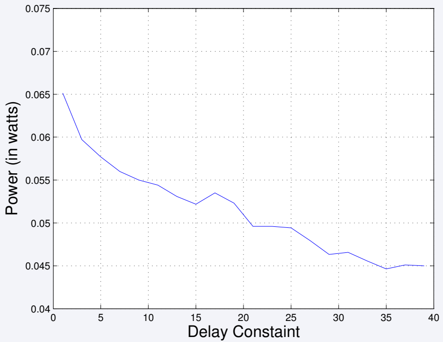
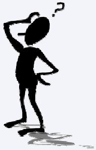


Figure: Power delay curve with finite state space

Continuous State Space Formulation

How to deal with large buffer and continuous channel?

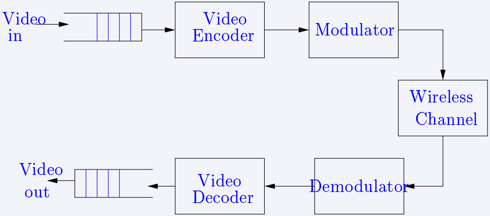


Approximate the value function:

$$h(x) = \sum_{i=1}^{K} f_i(x) r_i,$$

where,
$f = [f_1, f_2, \cdots f_K]$:Feature vectors e.g. [1, Q, x, Qx]
$r = [r_1, r_2, \cdots r_K]$: Weights Details

Joint Source and Channel Coding



Device an online algorithm to find:

- Quantization q_k per macro block (MB)
- Transmission rate u_k for each MB

and minimize transmission energy subject to distortion and absolute delay constraint per MB

Finite Horizon MDP Optimization Model

Figure: k^{th} MB transmission at nth slot

- M :No of MB in Frame
- δ_k: Delay of k^{th} MB
- D_k : Distortion of k^{th} MB

$$\min_{u_k,q_k} \mathbf{E} \sum_{k=0}^{M} \left\{ \sum_{l=n}^{L_k+n-1} P(x_l,u_k)T_c \,|x_n \right\}$$

$$\text{such that } \frac{1}{M}\mathbf{E}\sum_k D_{k=0}^{M-1} \leq \frac{1}{M} D_{max} \quad (8)$$

$$\delta_k \leq T_{max}, \forall k, \quad (9)$$

Formulate as finite Horizon MDP with immediate cost:

$$c_n = E_n + \lambda D_n$$

.

Finite horizon to Infinite Horizon

- Learning algorithms specifically for infinite horizon
- Join horizon M to horizon 0 to get infinite cycle
- Modified learning for finite horizon

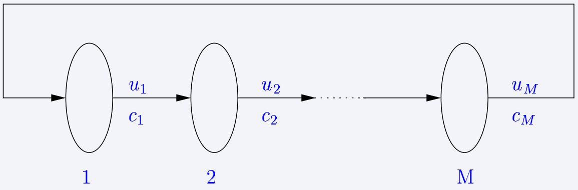


Figure: Finite horizon to infinite horizon

Learn Online Reinforcement-Learning ▸ Q Learning

Simulation

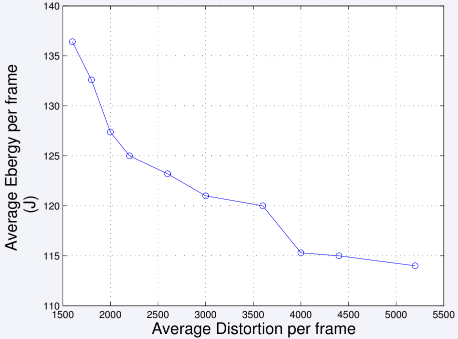


Figure: Power-Distortion curve

Summary

- Considered joint opportunistic and power optimal solution for minimum rate guarantee
- Proposed power optimal temporal fair scheduling
- Proposed average delay constrained power optimal scheme using MDP formulation.
- Proved convergence of the online policies
- Investigated the issue of power optimal video transmission over wireless

Future Research Direction

- More practical, discrete rate scheduler needs to be designed for minimum rate guarantee
- Definition of fairness in fading is an open issue
- Convergence of function approximation algorithms using multiple policies is unresolved
- Power optimal variable packet length online scheduling algorithm for video transmission can be considered

Thank you

Using ergodicity from (3),

$$\min \int \nu(dx_1, \cdots dx_N) \sum_{y \in A} \int_{[0,\infty)} p_1(dq|y,x) p_2(y|x) q,$$

subject to
$$\int \nu(dx_1 \cdots dx_N) \sum_{y \in A} \int_{[0,\infty)} p_1(dq|y,x) p_2(y|x) \log(1 + q y_i x_i) \geq C_i \quad \forall i,$$
$$q \geq 0. \qquad (10)$$

The Lagrange function associated with (10) is,

$$f(p_1, p_2, \lambda) \triangleq \int \nu(dx_1 \cdots dx_N) \sum_{y \in A} \int_{[0,\infty)} p_1(dq|y,x) p_2(y|x) \left(q - \sum_i \lambda_i \left[\log(1 + q y_i x_i) - C_i \right] \right). \qquad (11)$$

- ν Joint distribution of channel
- p_1, p_2 conditional distributions

Minimizing w.r.t. p_1, p_2,

$$F(\lambda) = \min_{p_1, p_2} f(p_1, p_2, \lambda)$$

Back

Short Term Thoughput Fair

- Thoughput: MTC_i for window of slot M, each slot duration T
- State :$\{\mathbf{r(n)}, \mathbf{x}(n)\}$
  - $r(n)$ Residual thoughput at slot n
  - $x(n)$ Channel state vector
  - $u(n)$ Transmission vector at slot n
- DP formulation

$$\begin{aligned} \mathbf{V}(n, \mathbf{r(n)}, \mathbf{x}(n)) &= \min\left(q(n) + \bar{\mathbf{V}}(n+1, \mathbf{r}(n+1))\right), \\ &= \min_{y_i(n), q(n)} \sum_i^N \left(\frac{1}{x_i}\left(e^{(u_i(n) y_i(n))} - 1\right) + \bar{\mathbf{V}}(n+1, \mathbf{r}(n+1))\right), \\ \bar{\mathbf{V}}(M+1, \mathbf{r}(M+1)) &= \infty. \end{aligned} \quad (12)$$

- Heuristic : Elimination policy

Back

Function Approximation

Least Square Policy Iteration:

Minimize the difference between actual and estimated approximated value function by,

$$
\begin{aligned}
\bar{r}_n &= \arg\min_r \sum_{m=0}^{n} \Bigg\{ \Big(f(s_m)' r - f(s_m)' r_n - \sum_{k=m}^{n} \\
&\qquad (\alpha\Lambda)^{k-m} d_n(s_m, s_{m+1}) \Big)^2 \Bigg\} \\
r_{n+1} &= r_n + \beta_n(\bar{r}_n - r_n) && (13) \\
d_n(s_m, s_{m+1}) &= c(s_m, s_{m+1}) - \phi_n + (f(s_{m+1}) - f(s_m) r_n, \forall k, n && (14) \\
\phi_{n+1} &= \phi_n + \gamma_n (c(s_n, s_{n+1}) - \phi_n) && (15)
\end{aligned}
$$

- Existence of fixed point proved
- Convergence not proved

Back

Solution using Q Learning

$$
\begin{aligned}
Q_{n+1}(s,u) &= Q_n(s,u) + \alpha_n(s,u)e_n \\
\phi_{n+1} &= \phi_n + \beta_n e'_n, \\
e_n &= \begin{cases} c_n - \phi_n + max_b Q_n(y_n,b) - Q_n(s,u) \\ \text{if } (s,u) = (s_n,u_n), x_n \in S_i, i < N \\ c_n - \phi_n \\ \text{if } (s,u) = (s_n,u_n), x_n \in S_N \\ 0 \text{ otherwise} \end{cases}, \\
e'_n &= \begin{cases} c_n - \phi_n + max_b Q_n(y_n,b) - Q_n(s,u) \\ \text{if } s = s_n, s_n \in S_i, i < N \\ c_n - \phi_n \\ \text{if } s = s_n, s_n \in S_N \\ 0 \text{ otherwise} \end{cases},
\end{aligned} \tag{16}
$$

Back