Adaptive Distributed Algorithms for Optimal Random Access Channels

Yichuan Hu & Alejandro Ribeiro
Dept. of Electrical and Systems Engineering
University of Pennsylvania

(yichuan, aribeiro) @seas.upenn.edu
http://fling.seas.upenn.edu/~yichuan/wiki/
http://alliance.seas.upenn.edu/~aribeiro/wiki/

This work is supported by ARO P-57920-NS and NSF CAREER CCF-0952867.

Penn Seminar on Communications and Networking, Nov. 11, 2010
Random access channels in presence of fading

- Random access channel
  - Terminals make independent transmission decisions
  - Simultaneous transmissions result in lost packets
- Minimal coordination yet reasonable resource utilization (36%)

- In presence of fading
  - Adapt transmission probability
  - Adapt transmitted power
- Only own channel is known
- Only own channel needs to be known
- **Opportunistic** random access
  - Transmit only when channel exceeds *threshold* [Qin-Berry '06]
  - *Optimal strategy* for concave rate functions [Yu-Giannakis '06]

- Threshold rules have been extensively studied
  - [Adireddy-Tong '05], [Bai-Zhang '06], [Xue, et al '07], ...

- Require access to channel probability distribution (pdf)

- Suboptimal adaptive rules have been proposed
  - [Al Harthi-Borst '09], [Salodkar-Karandikar '09]

- Optimal adaptive algorithm operating without knowledge of channel pdf
Random access in the presence of fading

- $n$ terminals transmit to common AP
- $h_i(t)$ Channel from $i$ to AP at time $t$
- At time $t$ terminals determine
  - Schedule $q_i(t) = Q_i(h_i(t)) \in \{0, 1\}$
  - Transmitted power $p_i(t) = P_i(h_i(t))$
- At time $t$, terminal $i$ transmits with power $p_i(t)$ when $q_i(t) = 1$
- If transmission is scheduled, communication rate is determined by SNR
- Write generically as $C_i(h_i(t)p_i(t))$
  - Capacity achieving codes $C_i(h_i(t)p_i(t)) = B \log \left(1 + \frac{h_i(t)p_i(t)}{BN_0}\right)$
  - AMC $C_i(h_i(t)p_i(t)) = \sum_{m=1}^{M} \tau_m \mathbb{I} \left( \eta_m \leq \frac{h_i(t)p_i(t)}{BN_0} \leq \eta_{m+1} \right)$
- Function $C_i(h_i(t)p_i(t))$ not necessarily concave. Discontinuities allowed
Information is delivered from terminal $i$ to the AP if and only if

- $\Rightarrow$ Terminal $i$ is scheduled $q_i(t) = 1$
- $\Rightarrow$ All other terminals are not scheduled $q_j(t) = 0$ for all $j \neq i$

Information delivered by $i$ at time $t$ is therefore

$$r_i(t) = q_i(t) \prod_{j=1,j \neq i}^{n} [1 - q_j(t)] C_i(h_i(t)p_i(t))$$

Long term average rate then given by

$$r_i := \lim_{t \to \infty} \frac{1}{t} \sum_{u=1}^{t} r_i(u) = \lim_{t \to \infty} \frac{1}{t} \sum_{u=1}^{t} \left[ q_i(u) \prod_{j=1,j \neq i}^{n} [1 - q_j(u)] C_i(h_i(u)p_i(u)) \right]$$
From long term averages to expected values

- Assuming (requiring) the use of ergodic schedules and power allocations
  \[ r_i = \mathbb{E}_h \left[ Q_i(h_i) C_i(h_i P_i(h_i)) \prod_{j=1, j \neq i}^n [1 - Q_j(h_j)] \right] \]

- Recall \( q_i(t) = Q_i(h_i(t)) \) and \( p_i(t) = P_i(h_i(t)) \) (which presupposes ergodicity)

- Invoking independence of schedules of different terminals
  \[ r_i = \mathbb{E}_{h_i} \left[ Q_i(h_i) C_i(h_i P_i(h_i)) \prod_{j=1, j \neq i}^n [1 - \mathbb{E}_{h_j} [Q_j(h_j)]] \right] \]

- Likewise, average power consumption of terminal \( i \) is
  \[ p_i := \lim_{t \to \infty} \frac{1}{t} \sum_{u=1}^t q_i(u)p_i(u) = \mathbb{E}_{h_i} \left[ Q_i(h_i) P_i(h_i) \right] \]
Introduce proportional fair (PF) utility $U(r) = \sum_{i=1}^{n} w_i \log(r_i)$

Tends to equalize rates of different terminals

Choose schedules and power allocations that maximize PF utility

$$P = \max U(r)$$

subject to

$$(r_i) = \mathbb{E}_{h_i} \left[ Q_i(h_i) C_i(h_i P_i(h_i)) \right] \prod_{j=1, j \neq i}^{n} \left[ 1 - \mathbb{E}_{h_j} [Q_j(h_j)] \right]$$

$${\mathbb{E}}_{h_i} \left[ Q_i(h_i) P_i(h_i) \right] \leq p_{i}^{avg}$$

$${Q}_i(h_i) \in Q, \quad {P}_i(h_i) \in \mathcal{P}_i,$$

$p_{i}^{avg} = \text{average power constraint. } Q = \text{set of functions with image } \{0, 1\}.$

$\mathcal{P}_i = \text{instantaneous power constraints}$
Challenges

(1) Optimization problem is not convex
⇒ Schedules are binary. Rate function not necessarily concave.

(2) Infinite dimensional schedules and power allocations are functions of $h_i$

(3) Design algorithm without knowledge of channel’s probability distribution

(4) Schedules are required to be independent across terminals
⇒ Solution method has to guarantee independence of schedules

(5) Schedule and power allocation of terminal $i$ is a function of $h_i$ only

- As always, problem is nicer in dual domain (convex and finite dimensional)
- Can obviously overcome (1) and (2) on dual domain. Can also solve (3).
- But duality gap is strictly positive
A simple, yet effective, reformulation

- Substitute rate expression into PF utility (log of prod = sum of logs)

\[
U(r) = \sum_{i=1}^{n} w_i \left[ \log \mathbb{E}_{h_i} \left[ Q_i(h_i) C_i(h_i P_i(h_i)) \right] + \sum_{j=1, j \neq i}^{n} \log \left( 1 - \mathbb{E}_{h_j} \left[ Q_j(h_j) \right] \right) \right]
\]

- Reordering terms it’s easy to see that

\[
U(r) = \sum_{i=1}^{n} \left[ w_i \log \left( \mathbb{E}_{h_i} \left[ Q_i(h_i) C_i(h_i P_i(h_i)) \right] \right) + \tilde{w}_i \log \left( 1 - \mathbb{E}_{h_i} \left[ Q_i(h_i) \right] \right) \right]
\]

- Which can be written as a sum of local “utilities” defined as

\[
U_i = w_i \log \left( \mathbb{E}_{h_i} \left[ Q_i(h_i) C_i(h_i P_i(h_i)) \right] \right) + \tilde{w}_i \log \left( 1 - \mathbb{E}_{h_i} \left[ Q_i(h_i) \right] \right)
\]

- A term that accounts for contribution to i’s rate
- A term that accounts for interference to other terminals
Equivalent problem formulation

► Original problem is then equivalent to set of problems

\[ P_i = \max \ w_i \log x_i + \tilde{w}_i \log(1 - y_i) \]

s.t. \[ x_i \leq \mathbb{E}_{h_i} \left[ Q_i(h_i) C_i(h_i P_i(h_i)) \right] \]

\[ y_i \geq \mathbb{E}_{h_i} \left[ Q_i(h_i) \right] \]

\[ \mathbb{E}_{h_i} \left[ Q_i(h_i) P_i(h_i) \right] \leq p_{\text{avg}}^i \]

\[ x_i \geq 0, \quad 0 \leq y_i \leq 1, \quad Q_i(h_i) \in \mathcal{Q}, \quad P_i(h_i) \in \mathcal{P}_i \]

► It is a set of per terminal problems

\[ \Rightarrow \] Solutions depend on local channel, i.e., \( h_i \) only, cf. Issue (5)

► Solutions of different terminals are statistically independent, cf. Issue (4)

► Plus, these problems do have null duality gap [Ribeiro-Giannakis ’10]

► Solves (1) non-convexity and (2) infinite-dimensionality

► Also solves (3) ignorance of channel’s pdf stochastic subgradient descent
Lagrangian and descent directions of dual function

- Descent direction of dual function subgradient
  - Given dual variables find primal variables that maximize Lagrangian
  - Evaluate constraint violation at Lagrangian maximizers

- Introduce multipliers $\lambda_1$, $\lambda_2$, and $\lambda_3$. Define Lagrangian

$$
\mathcal{L}_i(x_i, P_i(h_i), \lambda_i) = w_i \log x_i + \tilde{w}_i \log (1 - y_i) + \lambda_1 \left[ \mathbb{E}_{h_i} [Q_i(h_i)C_i(h_iP_i(h_i))] - x_i \right] \\
+ \lambda_2 \left[ y_i - \mathbb{E}_{h_i} [Q_i(h_i)] \right] + \lambda_3 \left[ p_i^{avg} - \mathbb{E}_{h_i} [Q_i(h_i)P_i(h_i)] \right]
$$

- Reorder Lagrangian to group by primal variables

$$
\mathcal{L}_i(x_i, P_i(h_i), \lambda_i) = \lambda_3 p_i^{avg} + \left[ w_i \log x_i - \lambda_1 x_i \right] + \left[ \tilde{w}_i \log (1 - y_i) + \lambda_2 y_i \right] \\
+ \mathbb{E}_{h_i} [Q_i(h_i)\left[ \lambda_1 C_i(h_iP_i(h_i)) - \lambda_2 - \lambda_3 P_i(h_i) \right] ]
$$

- Maximization to respect to primals is easy (separable)
  - But there is an infinite number of them
Lagrangian maximization with respect to $x_i$ and $y_i$ is simple

\[ x_i(t) = \arg\max_{x \geq 0} \{ w_i \log x_i - \lambda_{i1}(t)x \} = \frac{w_i}{\lambda_{i1}(t)} \]

\[ y_i(t) = \arg\max_{0 \leq y \leq 1} \{ \tilde{w}_i \log_i(1 - y_i) + \lambda_{i2}(t)y \} = \left[ 1 - \frac{\tilde{w}_i}{\lambda_{i2}(t)} \right]^+ \]

With respect to $Q_i(h_i)$ and $P_i(h_i)$ consider current value of channel $h_i(t)$

Compute maximizers for current channel only

\[ \{q_i(t), p_i(t)\} = \arg\max_{q_i \in \{0,1\}, p_i \in [0,p_i^{\text{inst}}]} \{ q_i \left[ \lambda_{i1}(t)C_i(h_i(t)p_i) - \lambda_{i2}(t) - \lambda_{i3}(t)p_i \right] \} \]

Non-convex maximization, yet easy to do (only two variables involved)

Policy Transmit with power $p_i(t)$ if and only if $q_i(t) = 1$

$x_i(t)$ and $y_i(t)$ are auxiliary variables
Stochastic subgradient descent (II)

- Stochastic subgradients defined instantaneous constraint violations

\[ s_{i1}(t) = q_i(t)C_i(h_i(t)p_i(t)) - x_i(t), \]
\[ s_{i2}(t) = y_i(t) - q_i(t), \]
\[ s_{i3}(t) = p_{i}^{\text{avg}} - q_i(t)p_i(t). \]

- Expected value of stochastic subgradients is constraint violation
- Which is known to be a descent direction of dual function
- Algorithm completed with (stochastic) descent in dual domain \((l = 1, 2, 3)\)

\[ \lambda_{il}(t + 1) = [\lambda_{il}(t) - \epsilon s_{il}(t)]^+ \]

- Algorithm does not require pdf of \(h_i\)
Initialize Lagrangian multipliers $\lambda_i(0)$; 

for $t = 0, 1, 2, \cdots$ do 

Compute primal variables: 

$$x_i(t) = \frac{w_i}{\lambda_{i1}(t)};$$

$$y_i(t) = \left[1 - \frac{\tilde{w}_i}{\lambda_{i2}(t)}\right]^+;$$

$$p_i(t) = \operatorname{argmax}_{p_i \in [0, p_i^\text{inst}]} \{\lambda_{i1}(t)C_i(h_i(t)p_i) - \lambda_{i2}(t) - \lambda_{i3}(t)p_i\};$$

$$q_i(t) = H(\lambda_{i1}(t)C_i(h_i(t)p_i(t)) - \lambda_{i2}(t) - \lambda_{i3}(t)p_i(t));$$

if $q_i(t) = 1$ then 

Transmit with power $p_i(t)$;

end

Compute stochastic subgradients:

$$s_{i1}(t) = q_i(t)C_i(h_i(t)p_i(t)) - x_i(t);$$

$$s_{i2}(t) = y_i(t) - q_i(t);$$

$$s_{i3}(t) = p_i^\text{avg} - q_i(t)p_i(t);$$

Update dual variables: 

$$\lambda_{il}(t + 1) = [\lambda_{il}(t) - \epsilon s_{il}(t)]^+, \quad \text{for } l = 1, 2, 3;$$

end
Almost sure ergodic optimality

Hypotheses

(H1) Bounded second moment \( \hat{S}^2 \geq \mathbb{E} \left[ \| \hat{s}(t) \|^2 \mid \lambda(t) \right] \)

(H2) Strictly feasible \( x_0 \in \mathcal{X} \) and \( p_0(h) \) exist

(H3) Channel realizations are independent identically distributed

Theorem

- Average power constraint is almost surely satisfied

\[
\lim_{t \to \infty} \frac{1}{t} \sum_{u=1}^{t} q_i(u)p_i(u) \leq p_i^{\text{avg}} \quad \text{a.s.,}
\]

- The utility of the ergodic limit of the transmission rates almost surely converges to a value within \( \epsilon/2 \sum_{i=1}^{n} \hat{S}_i^2 \) of optimality,

\[
P - U(r) := P - \sum_{i=1}^{n} w_i \log \left( \lim_{t \to \infty} \frac{1}{t} \sum_{u=1}^{t} r_i(u) \right) \leq \frac{\epsilon}{2} \sum_{i=1}^{n} \hat{S}_i^2
\]
Does the algorithm provide intuition on structural properties of the optimal operating point?

**Theorem**

- The optimal scheduling function $Q_i^*(h_i)$ for random access channels is a *threshold rule*. I.e., there exists a constant $h_0$ such that

$$Q_i^*(h_i) = H(h_i - h_0).$$

Moreover, the threshold $h_0$ is completely determined by optimal dual variables $\lambda_{il}^*$ ($l = 1, 2, 3$).

**Implication**

- All we need is the optimal dual variables. Instead of learning the channel pdf (possible infinite dimensional), we only need to know three dual variables $\lambda_{il}^*$ ($l = 1, 2, 3$).
Simulations – Setup

- Random access with \( n = 20 \) terminals communicating with common AP
- Randomly placed in square with side \( L = 100m \)
- Rayleigh fading
- Exponential pathloss law for average channel
  - exponent 2, coefficient \( 10^{-6} \)
- Channel bandwidth \( B = 1\text{Hz} \) (rates in \( \text{b/s/Hz} \))
- Noise level \( N_0 = 10^{-10}\text{W} \)
- All PF utility weights set to \( w_i = 1 \)

- Instantaneous power constraint \( p_i^{\text{inst}} = 100\text{mW} \). Average power \( p_i^{\text{avg}} = 100\text{mW} \)
- AMC with four modes
  - Rates of \( \tau_1 = 1\text{b/s/Hz} \), \( \tau_2 = 2\text{b/s/Hz} \), \( \tau_3 = 3\text{b/s/Hz} \), and \( \tau_4 = 4\text{b/s/Hz} \)
  - Transitions at received SNRs \( \eta_1 = 1 \), \( \eta_2 = 4 \), \( \eta_3 = 8 \), and \( \eta_4 = 16 \)
Simulations – Convergence

- Convergence in about 200 iterations ($\epsilon = 0.1$)
Simulations – Transmission probabilities

- Transmission probabilities of one run compared with optimal probs
- All terminals transmit with equal probability (PF utility)
Rates exhibit variation due to differing average channels
No power is allocated when channel realizations are bad. Terminal 3 uses only the AMC mode with the lowest rate $\tau_1 = 1 \text{ bits/s/Hz}$, while Terminal 13 uses two modes with rates $\tau_2 = 2 \text{ bits/s/Hz}$ and $\tau_3 = 3 \text{ bits/s/Hz}$.

(a) Terminal 3

(b) Terminal 13
Conclusions

- Introduced adaptive random access protocol
  - Adapts schedules and transmitted power to channel realizations
  - Operates without knowledge of channel’s distribution
  - Maximizes proportional fair utility

- Reformulate as per terminal subproblems
- Implement stochastic subgradient descent in dual domain
- Pursue similar ideas in general ad-hoc networks