Delay-Optimal Cross-Layer Design for Wireless Systems

Vincent Lau
Dept of ECE
Hong Kong University of Science and Technology
May 2009
Outline

- Introduction and Motivation
- Survey of Existing Approaches
- Example: Distributive Delay-Optimal Control for Uplink OFDMA via Localized Stochastic Learning and Auction Game
- Convergence Analysis
- Asymptotic Optimality
- Conclusion
Introduction and Motivation

Why delay performance is important?

- "WHAT??!! He is stuck in the air!! (!$*(&#%*!()"
- "You must be kidding me! Buffering at such an important moment!!??"

Fact 1:
Real-life applications are delay-sensitive
Introduction and Motivations

Multiple delay-sensitive applications running at different devices:

- Keep track of a game
- Play multi-player game
- Keep talking to some friends

**Fact II:** Different users and applications have heterogeneous delay requirements
Related Works

**OFDMA Joint Power and Subband Design for PHY Performance**

- [Yu’02], [Hoo’04],[Seong’06], etc.
  - Selects the strongest user per subband
  - Time-Frequency Water-filling Power Allocation
  - Assuming knowledge of **perfect CSIT**.

- [Lau’05], [Wong’09], [Brah’07] etc.
  - Robust Power and Subband Control with **limited feedback** or outdated CSIT (packet errors).

**Remark:**
Only adapt based on **CSIT**, ignoring queue states and optimize **PHY layer performance only** (throughput or PFS)

**Conclusion:** Very important to make use of both **(channel state info) CSI** and **(queue state info) QSI** for delay sensitive applications
Introduction and Motivations

**Challenges to incorporate QSI and CSI in adaptation**

**Challenge 1:** Requires both **Information theory** (modeling of the PHY dynamics) & **Queueing theory** (modeling of the delay/buffer dynamics)

**Challenge 2:** Brute-force approach cannot lead to any viable solution

When Shannon meets Kleinrock...
Existing Approaches to deal with Delay-Optimal Control

Various approaches dealing with delay problems

**Approach I: Stability Region and Lyapunov Drift** [Berry’02], [Neely’07], etc.
- Discuss **stability region** of point-to-point SISO and multiuser SISO.
- Also considered asymptotically delay-optimal control policy based on “Lyapunov Drift”
- The authors obtained interesting tradeoff results as well as insight into the structure of the **optimal control policy at large delay regime**.

**Remark:**
This approach allows simple control policy with design insights but the control will be good only for **asymptotically large delay regime**.

[Diagram showing buffer partitioning and state transitions with arrows indicating changes in buffer state s when S<1/v or S>1/v]
Various approaches dealing with delay problems

Approach II [Yeh’01PhD], [Yeh’03ISIT]
- Symmetric and homogeneous users in multi-access fading channels
- Using stochastic majorization theory, the authors showed that the longest queue highest possible rate (LQHPR) policy is delay-optimal
Related Works

Various approaches dealing with delay problems

**Approach III**: [Hui’07], [Tang’07], etc.

To convert the delay constraint into average rate constraint using tail probability at large delay regime (large derivation theory) and solve the optimization problem using information theoretical formulation based on the rate constraint.

**Remark:**
While this approach allows potentially simple solution, the control policy will be a function of CSIT only and such control will be good only for large delay regime.

**Note:**
In general, the delay-optimal power and precoder adaptation should be a function of both the CSI and the QSI.
Related Works

Various approaches dealing with delay problems

**Approach IV** : [Bertsekas’87]
The problem of finding the optimal control policy (to minimize delay) is casted into a **Markov Decision Problem (MDP)** or a stochastic control problem.

**Remark:**
- Unfortunately, it is well-known that there is no easy solution to MDPs in general.
- Brute-force **value iteration** and **policy iteration** are very complex and time-consuming.
- The curse of **dimensionality**!!
Technical Challenges To be Solved

**Challenge 1:**
A systematic approach for low complexity delay-optimal control policy in general delay regime.

**Challenge 2:**
Curse of Dimensionality: Exponential Complexity due to coupling among multiple delay-sensitive heterogeneous users.

**Challenge 3:**
Structure of the delay-optimal policy.

**Challenge 4:**
Distributive Implementation (Function of Local CSI and Local QSI, e.g. uplink OFDMA)
Example:
Distributive Delay-Optimal Control for Uplink OFDMA via Localized Stochastic Learning and Auction Game
Uplink OFDMA System Model

QSI from the K mobiles

Q_k, ..., Q_k, ..., Q_1

Resource Allocation Controller

CSI from the K mobiles

SC alloc. policy $\Omega_S$

P alloc. policy $\Omega_P$

$\lambda_1$ → $\mu_1(Q)$ → 1

$\lambda_k$ → $\mu_k(Q)$ → k

$\lambda_K$ → $\mu_K(Q)$ → K
**OFDMA PHY Model**

Subcarrier & Power Allocation

\[ R_k = \sum_{n=1}^{N_F} s_{k,n} I (X_{k,n}; Y_{k,n} | H_{k,n}) = \sum_{n=1}^{N_F} s_{k,n} \log \left( 1 + p_{k,n} |H_{k,n}|^2 \right) \]

\[ Y_{k,n} = H_{k,n} X_{k,n} + Z_{k,n} \]
Source Model and System States

- G-MAP Packets
- YouTube Packets
- PHY Frames
- Packet Arrivals
- MAC State
- PHY State
- Power & Subband Allocation
- QSI $Q = \{Q_1, Q_2\}$
- CSI $H_{N_r \times N_t}$

Channel is quasi-static in a slot
i.i.d. between slots
OFDMA Queue Dynamics

- Time domain partitioned into scheduling slots
  - CSI $H(t)$ remains quasi-static within a slot and is i.i.d. between slots
  - Packet arrival $A(t) = (A_1(t), \ldots, A_K(t))$ where $A_k(t)$ i.i.d. according to a general distribution $P(A)$.
  - $N_k(t)$ denotes the random packet size, i.i.d.
  - $Q_k(t)$ denotes the number of packets waiting in the $k$-th buffer at the $t$-th slot.

$$Q_k(t + 1) = \min\{[Q_k(t) - R_k(t)\tau / N_k(t)]^+ + A_k(t), N_Q\}$$

- Global System State (CSI, QSI)
  - $\chi(t) = (H(t), Q(t))$

Total number of bits Transmitted in the $t$-th slot
Stationary Power and Subband Allocation Control Policy

- A mapping \( \Omega = (\Omega_p, \Omega_s) \) from the system state \( \mathcal{X} \) to a power and subband allocation actions.

\[
\Omega_p(\mathcal{X}) = \{p_{k,n}\}, \quad \Omega_s(\mathcal{X}) = \{s_{k,n}\}
\]

- Power Constraint

\[
\sum_{n=1}^{N_F} E[p_{k,n}] \leq P_k \quad \forall k \in \{1, K\}, \quad p_{k,n} \geq 0
\]

- Subband Allocation Constraint

\[
\sum_{k=1}^{K} s_{k,n} = 1 \quad \forall n \in \{1, N_F\}
\]

- Packet Drop Rate Constraint

\[
\Pr[Q_k = N_Q] \leq P_k^d \quad \forall k \in \{1, K\}
\]
OFDMA Delay-Optimal Formulation

Definitions: Average Delay, Power and Packet Drop Constraints under a control policy $\Omega$

$$T_k(\Omega) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[Q_k(t)] = \mathbb{E}_{\pi_x}[Q_k] \ \forall k \in \{1, K\}$$

$$\overline{P}_k(\Omega) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E} \left[ \sum_{n} p_{k,n}(t) \right] = \mathbb{E}_{\pi_x} \left[ \sum_{n} p_{k,n} \right] \leq P_k \ \forall k \in \{1, K\}$$

$$\overline{P}_k^d(\Omega) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E} \left[ 1 \{Q_k(t) = N_Q\} \right] = \mathbb{E}_{\pi_x} \left[ 1 \{Q_k = N_Q\} \right] \leq P_k^d \ \forall k \in \{1, K\}$$

$\mathbb{E}_{\pi_x}$ denotes expectation w.r.t. the underlying measure $\pi$

Little’s Law: average no. of packets=average arrival rate * average delay

the average delay (in terms of seconds) $\propto$ the average queue length
Problem Formulation: Find the optimal control policy that minimizes the Pareto Optimal delay boundary.

\[ d(\chi, \{ p, s \}) = \sum_k \beta_k Q_k \]

\[ \min_{\Omega} J^V_\beta = \sum_{k=1}^K \beta_k T_k(\Omega) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E} \left[ d(\chi(t), \Omega(\chi(t))) \right] \]

subject to the power and packet drop rate constraints.

Solution: Markov Decision Problem (MDP)

Key Idea: Divide-and-Conquer

To break a large problem (optimization over the whole policy space) into smaller subproblems (optimization over a control action at a stage).
Overview of Markov Decision Problem Formulation

- Specification of an Infinite Horizon Markov Decision Problem

  - Decisions are made at points of time – decision epochs
  - System state and Control Action Space:
    - At the t-th decision epoch, the system occupies a state $S_t$
    - The controller observes the current state and applies an action $A_t$
  - Per-stage Reward & Transition Probability
    - By choosing action $A_t$ the system receives a reward $R(S_t, A_t)$
    - The system state at the next epoch is determined by a transition probability kernel $\Pr(S_{t+1} \mid S_t, A_t)$
  - Stationary Control Policy:
    - The set of actions for all system state realizations $A_t = \pi(S_t)$
  - The Optimization Problem:
    - Average Reward
      - $R^* = \max_{\pi} \lim_{T \to \infty} \frac{1}{T} \mathbb{E} \left[ \sum_{t=1}^{T} R(S_t, A_t) \right]$
Overview of Markov Decision Problem Formulation

Solution of a Markov Decision Problem

Key Criterion: Bellman’s Equation
Under some technical conditions, the optimal value of the problem is given by the solution of the Bellman’s Equation.

Optimal average reward

\[ \bar{R}^* = \theta \]

Optimal policy (Fixed Point Problem on Functional Space)

\[ \pi^* = \arg \max_{A_i} \left\{ r(S_i, A_i) + \sum_{S_j} Pr(S_j | S_i, A_i) V(S_j) \right\} \]
Constrained Markov Decision Problem Formulation

CMDP Formulation: Find the optimal control policy $\Omega$ that minimizes

$$
\min_{\Omega} J^\Omega_\beta = \sum_{k=1}^K \beta_k T_k(\Omega) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}\left[d(\chi(t), \Omega(\chi(t)))\right]
$$

subject to the power and packet drop rate constraints

Lagrangian approach to the Constrained MDP:

$$
\min_{\Omega} \ L^\Omega_\beta(\gamma) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}\left[g(\chi(t), \Omega(\chi(t)))\right]
$$

$$
g(\gamma, \chi, \Omega(\chi)) = \sum_k \left(\beta_k Q_k + \gamma_k^k(\sum_n P_{k,n} - P_k) + \gamma_k^k(1[Q_k = N_Q] - P_k^d)\right)
$$
Optimal Solution

- **Infinite Horizon Average Reward MDP**

  Given a stationary control policy $\Omega$, the random process $\{\chi(t), g(\chi(t), \Omega(\chi(t)))\}$ evolves like a Markov Chain with transition kernel:

  $$\Pr[\chi(t+1)|\chi(t), \Omega(\chi(t))] = \Pr[H(t+1)] \Pr[Q(t+1)|\chi(t), \Omega(\chi(t))]$$

  Solution is given by the “Bellman Equation”

  $$\theta + V(\chi^i) = \min_{u(\chi^i)} \left[ g(\chi^i, u(\chi^i)) + \sum_{\chi^j} \Pr[\chi^j|\chi^i, u(\chi^i)]V(\chi^j) \right]$$

  “Potential function” (contribution of the state $i$ to the average reward)

  “Optimal Value” $\theta = J_\beta^* = \inf_\Omega J_\beta^\Omega$

  $$(N_Q + 1)^K$$ Equations and unknowns
Centralized Solution?

Obtain knowledge of global QSI from K users (Uplink)?

Heavy signaling loading to deliver these QSI from mobiles to BS

Must have distributive solution!

Optimal Solution – Online Learning

How to determine the potential function?

- Brute-Force solution of the Bellman Equation? (Value Iteration):
  - Too complicated, exponential complexity and memory requirement

- Online stochastic learning!
  - Iteratively estimate potential function based on real time
    observation of CSI and QSI - online value iteration

Distributive Solution:

Per-user Potential and LMs Initialization

Online Policy Improvement Based on Per-subband Auction

Online Per-user Potential and LMs Update [Local CSI, Local QSI]

Termination
Decentralized Solution (I)

Online Per-user Primal-Dual Potential Learning Algorithm via Stochastic Approximation

\[ \tilde{V}_{l+1}^k(Q_k^i) = \begin{cases} \tilde{V}_l^k(Q_k^i) + \epsilon^v_l \left[ (g_k(\gamma_l^k, \chi_k(l)) + \tilde{V}_l^k(Q_k(l + 1))) \right. \\
- (g_k(\gamma_l^k, \chi_k^l) + \tilde{V}_l^k(Q_k^l) - \tilde{V}_l^k(Q_k^i)) - \tilde{V}_l^k(Q_k^i) \left. \right] \text{ if } Q_k^i = Q_k(l) \\
\tilde{V}_l^k(Q_k^i) \text{ if } Q_k^i \neq Q_k(l) \end{cases} \]

\[ \tilde{\gamma}_{l+1}^k = \Gamma(\tilde{\gamma}_l^k + \epsilon_l^\gamma \left( \sum_{k,n} P_{k,n} - P_k \right)) \]

\[ \gamma_{l+1}^k = \Gamma(\gamma_l^k + \epsilon_l^\gamma \left[ Q_k(l) - N_Q \right] - P_k^i) \]

Remark (Comparison to the deterministic NUM): Iterative updates are performed within the CSI coherence time. This limits the number of iterations and the performance.

Proposed online algorithm: Iterative updates evolve in the same time scale as the CSI and QSI, allowing it to converge to a better solution (no longer limited by the coherence time of CSI).

Both the per-user potential and 2 LMs are updated simultaneously.

New Observation at the beginning of the (l+1)-th slot

\[ \sum_l \epsilon^v_l = \infty, \epsilon^v_l \geq 0, \epsilon^v_l \to 0, \sum_l \epsilon^\gamma_l = \infty, \epsilon^\gamma_l \geq 0, \epsilon^\gamma_l \to 0 \]

\[ \sum_l \left( (\epsilon^v_l)^2 + 2(\epsilon^\gamma_l)^2 \right) < \infty, \frac{\epsilon^v_l}{\epsilon^\gamma_l} \to 0 \]
Decentralized Solution (II)

- **Per-stage auction with K bidders (MSs) and one auctioneer (BS)**

- **Low complexity Scalarized Per-Subband Auction**

  - **Bidding:** Each user submits a bid \( \hat{X}_{k,n} \)
  
  - **Subband allocation:**
    \[
    s_{k,n}(H_n, Q^i) = \begin{cases} 
    1, & \text{if } k = k^*_n \text{ and } \hat{X}_{k^*_n,n} > 0 \\
    0, & \text{otherwise } \end{cases}
    \]

  - **Power allocation:**
    \[
    p_{k,n}(H_n, Q^i) = s_{k,n}(H_n, Q^i) \left( \sum_{k=1}^{K} \frac{\partial V^m(Q_k)}{\partial Q_k} - \frac{1}{|H_{k,n}|^2} \right)^+ 
    \]

  - **Charging:**
    \[
    C_{k,n} = s_{k^*_n,n} \hat{X}_{k^*_n,n} 1\{k = k^*_n\} \hat{K}_{k^*_n,n} = \arg\max_{k \neq k^*_n} \hat{X}_{k,n} 
    \]

- **Lemma:** The per-stage social optimal scalarized bid

  Water-level depends on QSI (via potential function)

  \[
  \delta \tilde{V}^k(Q_k^i) = \tilde{V}(Q_1^i, \cdots, Q_k^i, \cdots, Q_K^i) - \tilde{V}(Q_1^i, \cdots, [Q_k^i - 1]^+, \cdots, Q_K^i) 
  \]

  \[
  \hat{X}_{k,n}^* = \arg\max_{k \neq k^*_n} \hat{X}_{k,n} 
  \]

  \[
  \frac{1}{|H_{k,n}|^2} \left( \sum_{k=1}^{K} \frac{\partial V^m(Q_k)}{\partial Q_k} - \frac{1}{|H_{k,n}|^2} \right)^+ 
  \]
Decentralized Solution

Theorem (Convergence of online per-user learning) Under some mild conditions, the distributive learning converges almost surely.

Remark (Comparison to conventional stochastic learning)

Conventional SL: (1) for unconstrained MDP only or LM for CMDP are determined offline by simulation; (2) designed for centralized solution with control action determined entirely from the potential update → Convergence Proof based on standard “contraction Mapping” and Fixed-Point Theorem argument.

Proposed SL: (1) simultaneous update of LM and the potential function; (2) control action is determined by all the users’ potential via per-stage auction → per-user potential update is NOT a contraction mapping & standard proof does not apply.

Bellman equation.
Numerical Results

Average Delay per user vs SNR

The number of users $K = 2$, the buffer size $N_Q = 10$, the mean packet size $\overline{N}_k = 305.2$ Kbyte/pck, the average arrival rate $\lambda_k = 20$ pck/s

Huge gain in delay performance compared with Modified-Largest Weighted Delay First (M-LWDF), which is the queue length weighted throughput maximization.

Close-to-optimal performance even for small number of users.
Numerical Results

Average Delay per user vs No. of users

The buffer size $N_Q = 10$, the mean packet size $\overline{N}_k = 78.125$ Kbyte/pck

the average arrival rate $\lambda_k = 20$ pck/s, the queue weight $\beta_k = 1$ at a transmit SNR = 10dB.

The distributive solution has huge gain in delay performance compared with 3 Baselines.
Numerical Results

Illustration of convergence property:
Potential function vs. the scheduling slot index (K=10)
Conclusion

Distributive Implementation via Decentralized Stochastic Learning and Auction Game

Online Per-user Learning:
Simultaneous update of LMs and Potentials.
Almost sure convergence

Optimal Strategy for the Auction Game:
Delay-Optimal Power Control: Multi-Level Water-Filling
(QSI → water level; CSI → instantaneous allocation)
Delay-Optimal Subband Allocation: User selection based on (QSI,CSI)

Asymptotically Global Optimal for large K


Thank you!

Questions are Welcomed!

Vincent Lau - eeknlau@ee.ust.hk

http://www.ee.ust.hk/~eeknlau