Multimedia Communications
and Systems Laboratory

http://medianetlab.ee.ucla.edu

A Unified Framework for Delay-Sensitive Communications

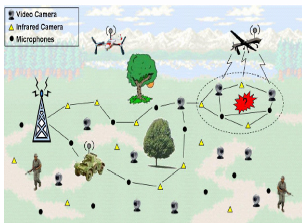
Fangwen Fu

fwfu@ee.ucla.edu

Advisor: Prof. Mihaela van der Schaar

UCLA

Motivation

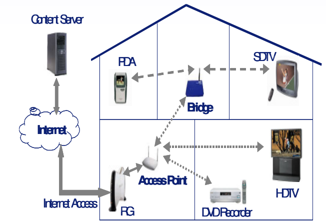


Sensor networks

VOIP

Wireless video phone

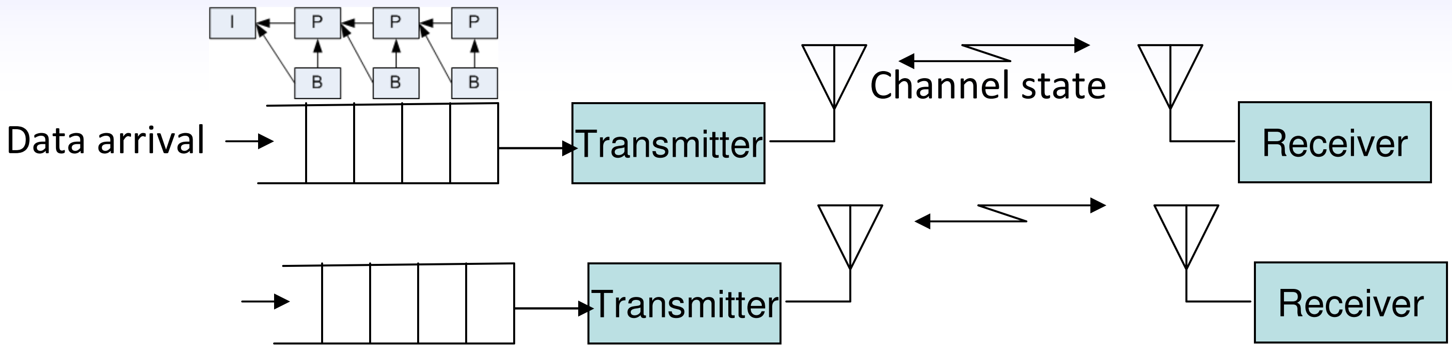
Video conference



In-home streaming

- Delay sensitive multimedia applications are booming over a variety of time-varying networks (e.g. sensor networks, WiMax, Wireless LAN, etc.)
- Existing dynamic distributed network environments cannot provide adequate support for delay-sensitive multimedia applications
- This problem has been investigated for a decade, but we still do not have efficient solutions for it.

Challenges



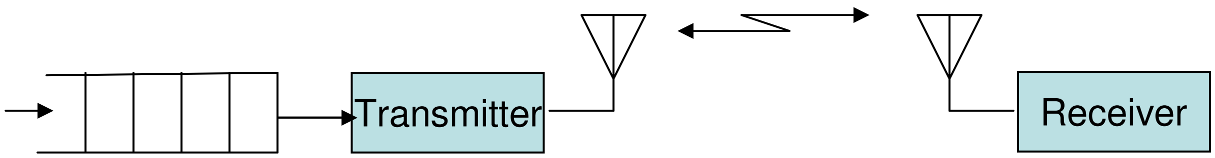
- *Challenge 1*: Unknown time-varying environments
  - Time-varying data arrivals and channel conditions
  - Lack of statistic knowledge of dynamics

- *Challenge 2:* Heterogeneity in the data to transmit (e.g. media data)
  - Different delay deadlines, importance, and dependencies

- *Challenge 3*: Coupling in multi-user transmission
  - Mutual impact due to dynamically sharing of the same network resources (e.g. bandwidth, transmission opportunities) by multiple users

Existing solutions-1

- Minimize average delay for *homogeneous traffic* in point-to-point communications

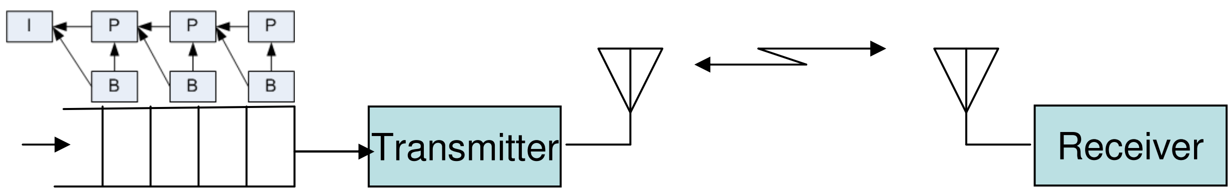


- Information theory [Shannon and beyond] –*Challenge 1*
  - Water-filling algorithms
  - Maximize the throughput without delay constraints

- Control theory – *Challenge 1*
  - Markov decision process (MDP) formulation [Berry 2002, Borkar 2007, Krishnamurthy 2006]
    - Statistic knowledge of the underlying dynamics is required
  - Online learning [Krishnamurthy 2007, Borkar 2008]
    - Slow convergence and large memory requirement
  - Stability-constrained optimization for single-user transmission [Tassiulas 1992,2006, Neely 2006, Kumar 1995, Stolyar 2003]
    - Queue is stable, but delay performance is suboptimal (for low delay applications)

Existing solutions-2

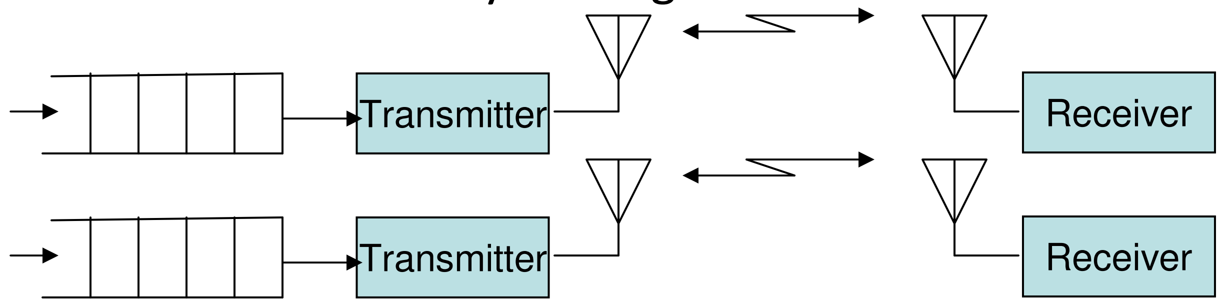
- Maximize quality of delay-sensitive applications with *heterogeneous traffic*



- Multimedia communication theory –*Challenge 2*
  - Cross-layer optimization [van der Schaar 2001, 2003, 2005, Katsaggelos 2002]
    - Observes and then optimizes (i.e. myopic optimization)
  - Rate distortion optimization (RaDiO) [Chou, 2001, Frossard 2006, Girod 2006, Ortega 2009]
    - Explicitly considers importance, delay deadlines and dependencies of packets
    - Linear transmission cost (e.g. not suitable for energy-constrained transmission)
    - No learning ability in unknown environments
  - *Both solutions only explore the heterogeneity in the media data, but do not explore the network dynamics (e.g. time-varying channel conditions) and resource constraints.*

Existing solutions-3

- Multi-user transmission by sharing network resources



- Network optimization theory
  - Network utility maximization [Chiang 2007, Katsaggelos 2008] –*Challenge 3*
    - Uses static utility function without considering the network dynamics
    - No delay guarantee
    - No learning ability in unknown environments
  - Stability-constrained optimization for multi-user transmission [Tassiulas 1992, 2006, Neely 2006, 2007, Kumar 1995, Stolyar 2003] - *Challenges 1 and 3*
    - Queue is stable, but delay performance is suboptimal (for low-delay applications)
    - Does not consider heterogeneous media data

A unified foresighted optimization framework

Challenges	Solutions
dynamic systems	Foresighted optimization framework
Unknown dynamics	Online learning
Learning efficiency Heterogeneity Multi-user coupling	Separation principles

Current utility

State-value function

$$\max_y \{ u(s,y) \quad + \quad \mathbf{E}_w V(f(s,y,w)) \}$$

Queue length
Channel condition
Heterogeneity

State: s

State: $s' = f(s,y,w)$

Action: y

Dynamics: w

Current time slot

Next time slot

Key accomplishments

	Previous state-of-art methods	Improvements
Energy-efficient data transmission*	Stability constrained optimization [Neely 2006]	**Reduce the delay by 70%** (at low delay region)
Wireless video transmission	Rate-distortion optimization [Chou 2001]	**Improve up to 5dB in video quality**
Multi-user video transmission	Network utility maximization [Chiang 2007]	**Improve 1~3dB in video quality**

*minimize the average delay

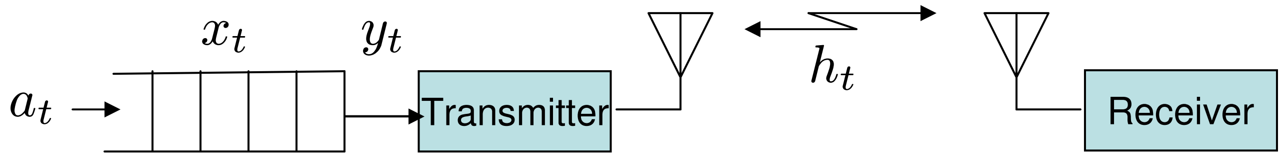
Roadmap

- Separation principle 1 (improving learning efficiency)
  - Post-decision state-based formulation
  - Structure-aware online learning with adaptive approximation
- Separation principle 2 (Separating the foresighted decision for heterogeneous media data transmission)
  - Context-based state
  - Priority-based scheduling
- Separation principle 3 (decomposing multi-user coupling)
  - Multi-user Markov decision process formulation
  - Post-decision state value function decomposition

Roadmap

- Separation principle 1 (improving learning efficiency)
  - Post-decision state-based formulation
  - Structure-aware online learning with adaptive approximation
- Separation principle 2 (Separating the foresighted decision for heterogeneous media data transmission)
  - Context-based state
  - Priority-based scheduling
- Separation principle 3 (decomposing multi-user coupling)
  - Multi-user Markov decision process formulation
  - Post-decision state value function decomposition

Energy-efficient data transmission

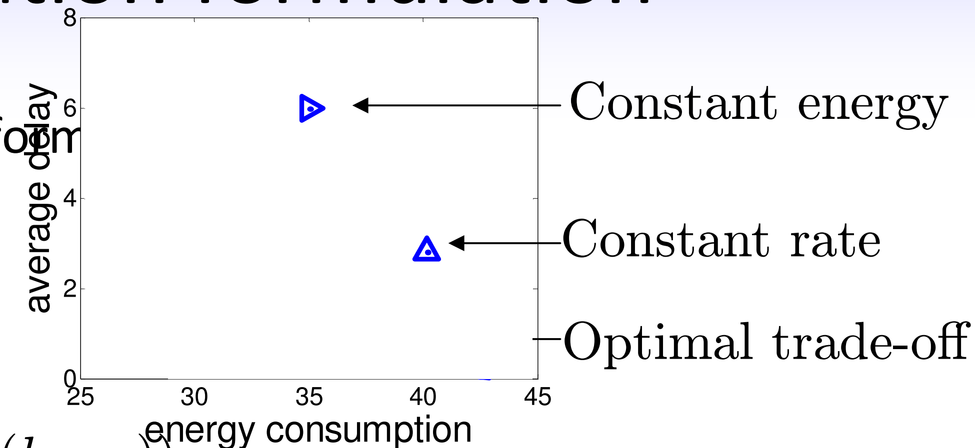


- Point-to-point time-slotted communication system
- System variables
  - Backlog (queue length): x_t
  - Channel state: h_t Finite state Markov chain (e.g. Rayleigh fading)
  - Data arrival process: a_t: i.i.d.
- Decision at each time slot
  - Amount of data to transmit (transmission rate): $y_t, 0 \leq y_t \leq x_t$
  - Energy consumption: $\rho(h_t, y_t)$, convex in y_t, e.g. $\rho_t(h_t, y_t) = \sigma^2 \frac{(2^{y_t} - 1)}{h_t}$.

What is the optimal (queueing) delay and energy trade-off?

Foresighted optimization formulation

- Foresighted optimization (MDP) form
  - State: (x_t, h_t)
  - Action: y_t
  - Policy: $\pi : (x_t, h_t) \to y_t$
  - Utility function:



$$u(x_t, h_t, y_t) = -(x_t - y_t + \lambda \rho(h_t, y_t)).$$

- Objective (optimize the trade-off between delay and energy consumption)

$$\max_{\pi} \mathbf{E} \sum_{t=0}^{\infty} \alpha^t \{u(x_t, h_t, \pi(x_t, h_t))\} \quad \alpha \in [0, 1) \text{ is discount factor.}$$

  - State value function: $V(x_t, h_t) = \max_{\pi} \mathbf{E} \sum_{k=t}^{\infty} \alpha^{(k-t)} \{u(x_k, h_k, \pi(x_k, h_k))\}$
- Bellman's equations

$$V(x, h) = \max_{\pi} \{u(x, h, \pi(x, h)) + \alpha \mathbf{E}_{a, h'|h} V(x - \pi(x, h) + a, h')\}$$

  - Policy iteration

Challenges for solving the Bellman's equations

Bellman's equation:

$$V(x,h) = \max_{\pi}\{u(x,h,\pi(x,h)) + \alpha \mathbf{E}_{a,h'|h} V(x-\pi(x,h)+a,h')\}$$

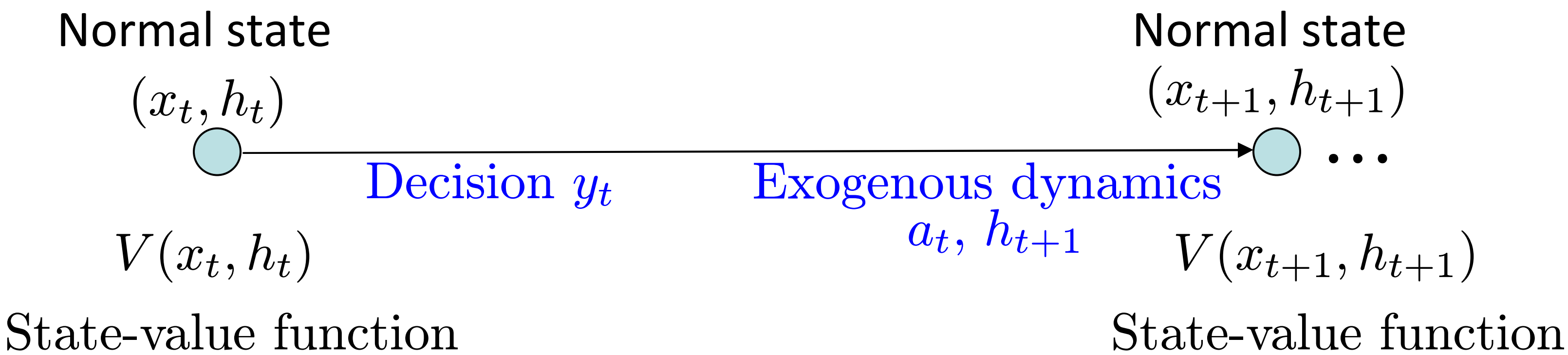
- Lack of statistical knowledge of the underlying dynamics
  - Unknown traffic characteristics
  - Unknown channel (network) dynamics

- Coupling between the maximization and expectation

- Curses of dimensionality
  - Large state space
    - Intractable due to large memory and heavy computation requirements

Conventional online learning methods

- Decision and dynamics



- Foresighted optimization

$$V(x,h) = \max_{0 \le y \le x} \{u\underbrace{(x,h,y) + \alpha \mathbf{E}_{a,h'|h} V(x-y+a,h')}_{Q(x,h,y)}\}$$

- Online learning
  - Learn Q-function (Q-learning): $Q(x,h,y)$

Low convergence, high space complexity

Our approach- separation via post-decision state

Normal state (x_t, h_t) $\xrightarrow{\text{Decision } y_t}$ Post-decision state $(x_t - y_t, h_t)$ $\xrightarrow{\text{Exogenous dynamics } a_t,\ h_{t+1}}$ Normal state (x_{t+1}, h_{t+1}) $\cdots$

$V(x_t, h_t) \leftarrow U(\tilde{x}_t, h_t) \leftarrow V(x_{t+1}, h_{t+1})$

State-value function — Post-decision state-value function — State-value function

Foresighted decision

$$V(x,h) = \max_y \{u(x,h,y) + \alpha U(x-y,h)\}$$

Expectation over dynamics

$$U(x,h) = \mathbf{E}_{a,h'|h} V(x+a, h')$$

Post-decision state separates foresighted decision from dynamics.

$$\begin{tikzcd} \pi, V \arrow[r, bend left, "\text{Expectation over dynamics}"] & U \arrow[l, bend left, "\text{Foresighted decision}"] \end{tikzcd}$$

Post-decision state-based online learning

$$U(x,h) = \boxed{\mathbf{E}_{a,h'|h}} V(x+a,h')$$
$$V(x,h) = \max_{y}\{u(x,h,y) + \alpha U(x-y,h)\}$$

- Online learning

$$\boxed{U_t(x,h_{t-1}) = (1-\beta_t)U_{t-1}(x,h_{t-1}) + \beta_t V_t(x,h_t)} \quad \text{e.g. } \beta_t = 1/t$$

Online update   Time-average

$$\begin{tikzcd} \pi, V \arrow[rr, bend left, "\text{Online update}"] & & U \arrow[ll, bend left, "\text{Foresighted decision}"] \end{tikzcd}$$

$$V_t(x,h_t) = \max_{y\in\mathcal{Y}}\{u(x,h_t,y) + \alpha U_{t-1}(x-y,h_t)\}$$

Theorem:
Online adaptation converges to the optimal solution when $t \to \infty$

Expectation is independent of backlog $x \to$ batch update (fast convergence).

Batch update incurs high complexity. ☹

Structural properties of optimal solution

$$\boxed{\begin{aligned} U(x,h) &= \mathbf{E}_{a,h'|h} V(x+a,h') \\ V(x,h) &= \max_y \{u(x,h,y) + \alpha U(x-y,h)\} \end{aligned}}$$

- Structural properties of optimal solution
  - Assumption: $u(x,h,y)$ is jointly concave and supermodular* in (x,y)

    .

$\pi(x,h)$ is monotonic in x

$$\begin{tikzcd} \pi, V \arrow[rr, bend left, "\text{Expectation over dynamics}"] & & U \arrow[ll, bend left, "\text{Foresighted decision}"] \end{tikzcd}$$

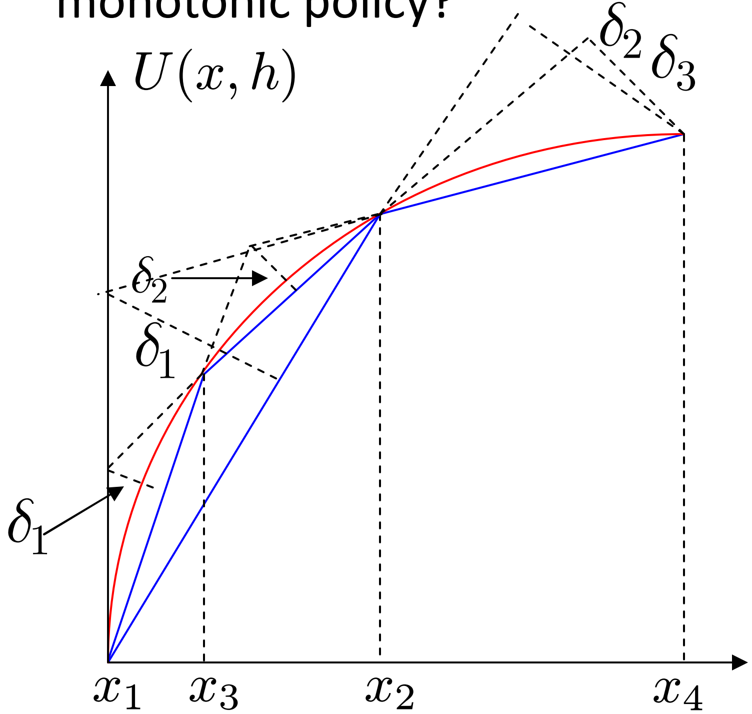
$U(x,h)$ is concave in x

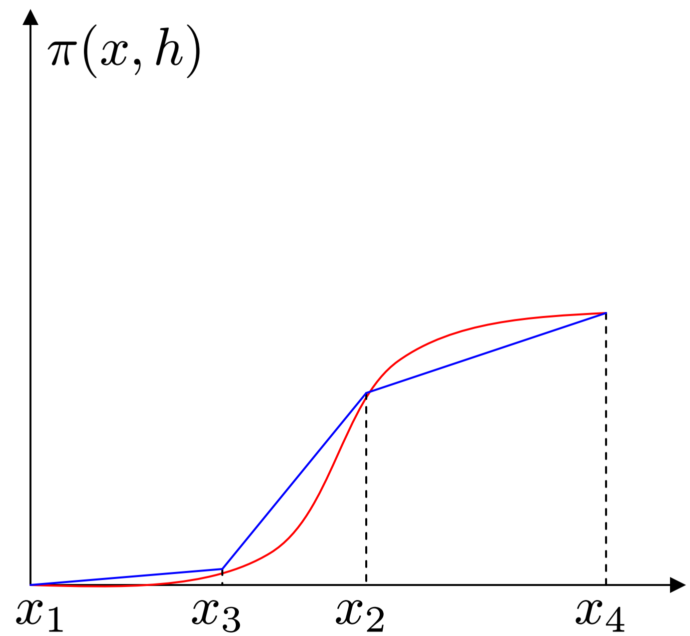
How can we utilize these structural properties in online learning?

$* \; u(x',h,y') - u(x',h,y) \geq u(x,h,y') - u(x,h,y) \text{ if } x' \geq x, y' \geq y$

Piece-wise linear approximation

- How to compactly represent post-decision state-value function and monotonic policy?



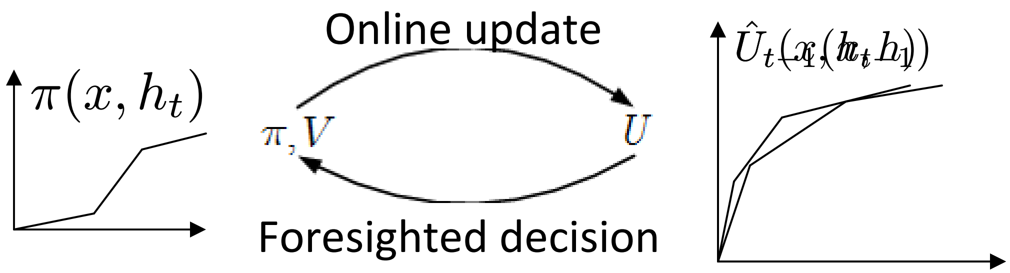


Adaptive approximation operator (A_δ)

For each channel state *h*, we approximate the post-decision state-value function such that $\min_{n} \max_{i=1,\cdots,n} \delta_i \le \delta$ (threshold).

Online learning with adaptive approximation

$$\hat{U}_t(x, h_{t-1}) = \underset{\delta}{A}\{(1-\beta_t)\hat{U}_{t-1}(x, h_{t-1}) + \beta_t V_t(x, h_t)\}$$



$$V_t(x, h_t) = \max_{y \in \mathcal{Y}}\{u(x, h_t, y) + \alpha \hat{U}_{t-1}(x-y, h_t)\}$$

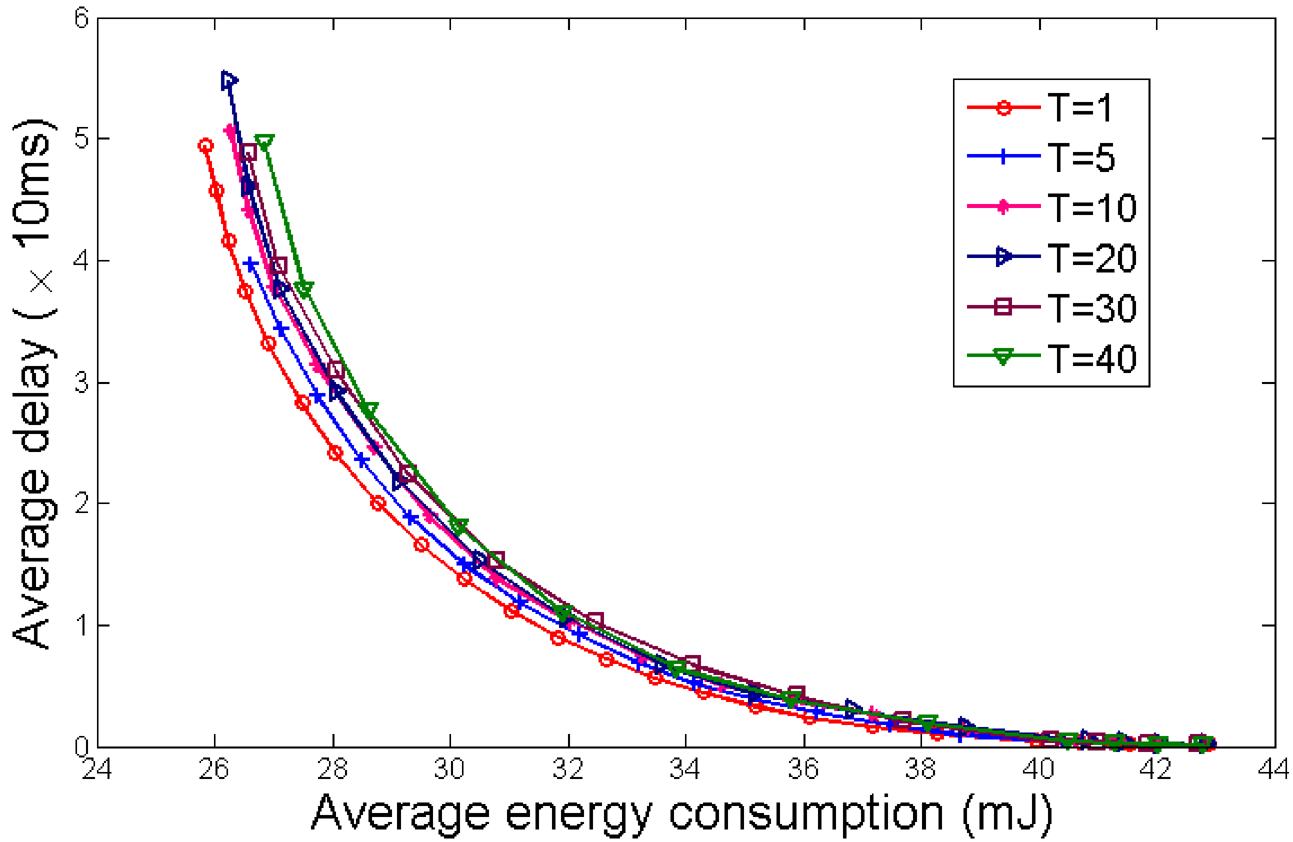
Theorem: Online learning with adaptive approximation converges to an ε-optimal solution, where $\varepsilon = \frac{\delta}{1-\alpha}$

Variant: Update $U(x, h)$ and $\pi(x, h)$ every T time slots

Performance of learning with approximation

Rayleigh fading channel

Average channel gain $\frac{h^2}{\sigma^2} = 0.14$ #channel state=8 $\alpha = 0.95$



Comparison with stability-constrained optimization

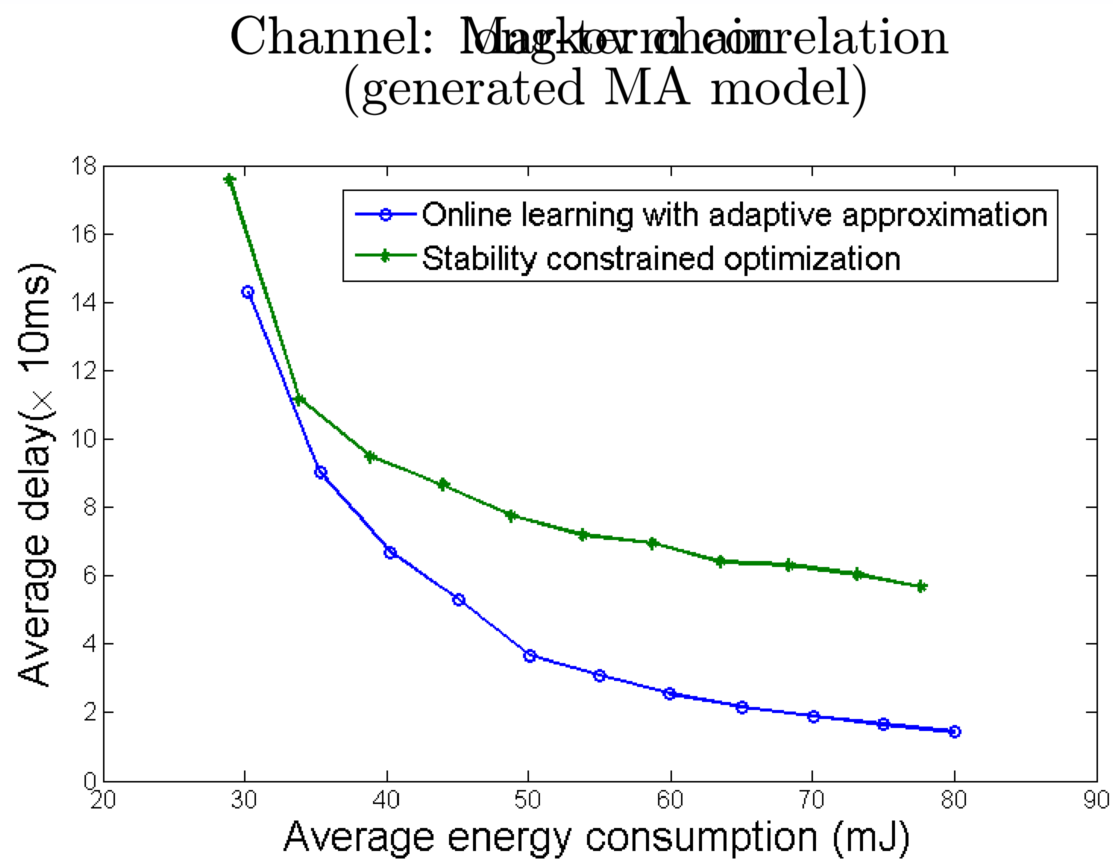
- Stability-constrained optimization [Neely, 2006]
  - Minimize the trade-off between Lyapunov drift and energy consumption

$$\min \lambda\rho(h_t, y_t) + \underbrace{(x_t - y_t)^2 - x_t^2}_{\text{Lyapunov drift}}$$

$$\max_{y_t \in \mathcal{Y}} \underbrace{-(x_t - y_t + \lambda\rho(h_t, y_t))}_{\substack{\text{Utility function} \\ u(x_t, h_t, y_t)}} + \underbrace{x_t - y_t - (x_t - y_t)^2}_{\substack{\text{Post-decision state value functon} \\ U(x_t - y_t, h_t)}} + x_t^2$$

  - Do not consider the effect of the utility function on post-decision state value function
  - Do not consider the time-correlation of the channel states
  - Only ensure queue stability, but result in poor delay performance

Comparison to stability-constrained optimization

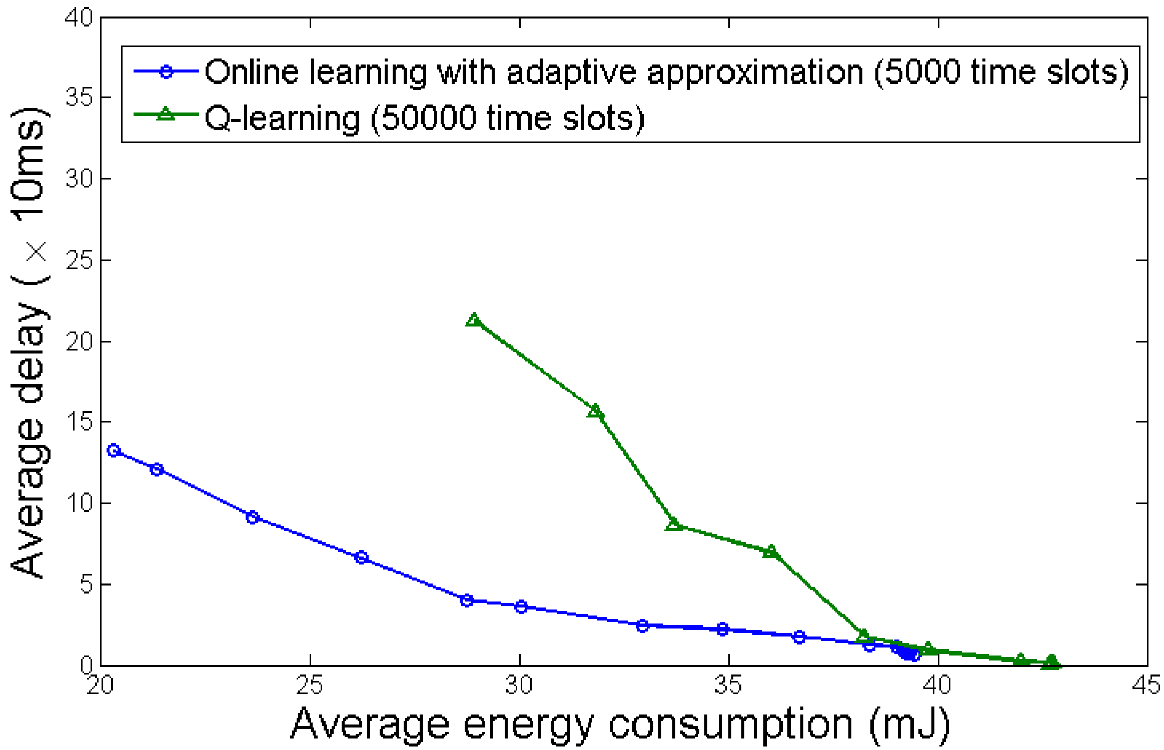


Stability constrained optimization
- Minimize Lyapunov drift ≠ Minimize delay

Our proposed solution
- Minimize queue size = Minimize delay

Comparison to Q-learning

- Markovian Rayleigh fading channel
- Q-learning: update the state-value function one state at each time slot (learn over 50000 time slots)
- Online learning with adaptive approximation: T=10, learn over 5000 time slots

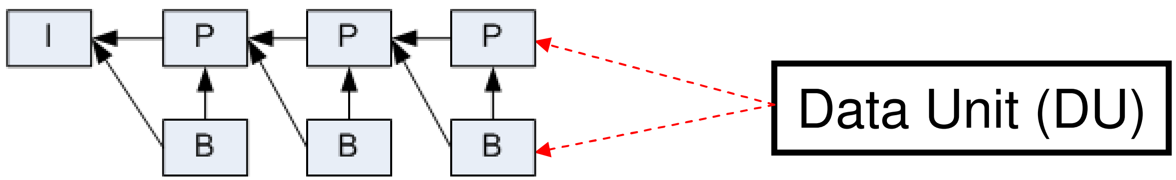


Roadmap

- Separation principle 1 (improving learning efficiency)
  - Post-decision state-based formulation
  - Structure-aware online learning with adaptive approximation
- Separation principle 2 (Separating the foresighted decision for heterogeneous media data transmission)
  - Context-based state
  - Priority-based scheduling
- Separation principle 3 (decomposing multi-user coupling)
  - Multi-user Markov decision process formulation
  - Post-decision state value function decomposition

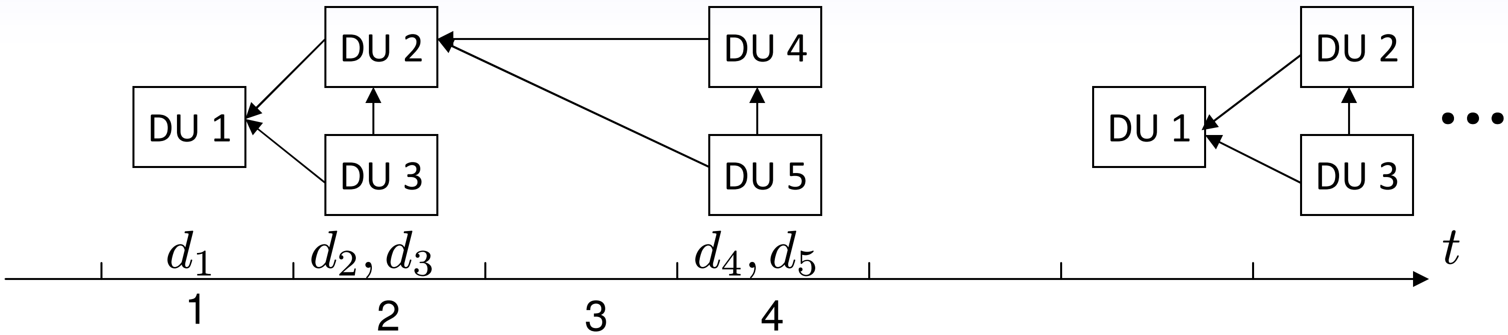
Heterogeneous media data

Media data representation:



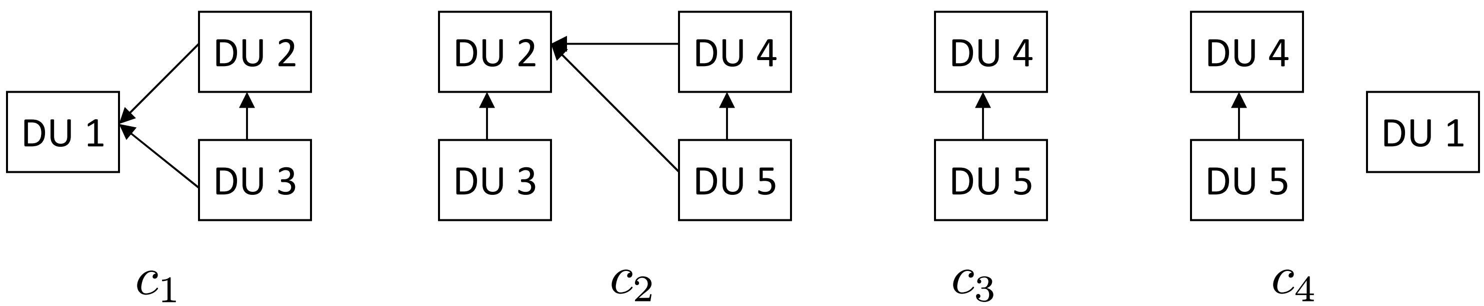
- Each DU has the following attributes:
  - Arrival time: time at which the DU is ready for processing: t^i
  - Delay deadline: d^i
  - Size : L^i in packets
  - Distortion impact: q^i per packet
  - Interdependency between DUs: *expressed by Directed Acyclic Graph (DAG)*

Context



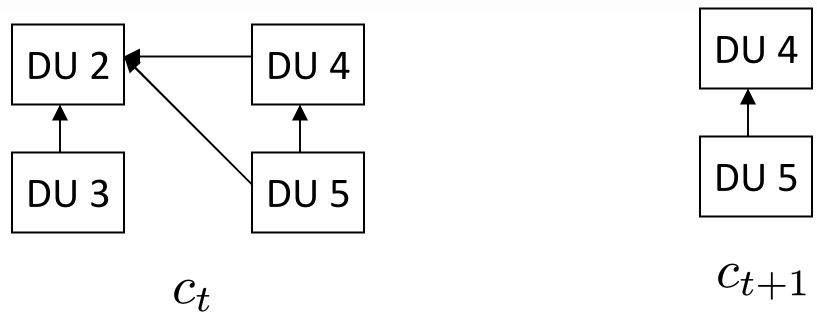
- Fixed GOP (i.e. group of DUs) structure
- Context (c_t) at each time slot t
  - Include the DUs whose deadlines are within a time window W

e.g. $W = 3$



- Context transition is deterministic

Foresighted optimization



State: $(c_t, \mathbf{x}_t, h_t)$ $\mathbf{x}_t = (x_t^2, x_t^3, x_t^4, x_t^5)$

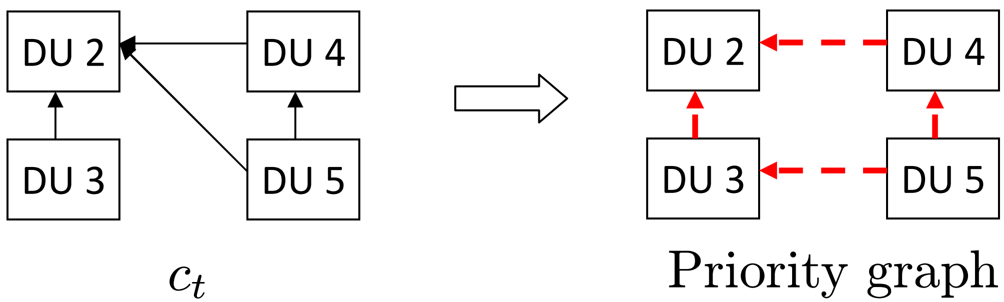
- Multi-DU Foresighted decision

$$\max_{y_t^i, i \in c_t} \left\{ \underbrace{\sum_{i \in c_t} u_i(x_t^i, h_t^i, y_t^i)}_{\text{Current utility}} + \underbrace{\alpha U(c_t, \mathbf{x}_t - \mathbf{y}_t, h_t)}_{\text{Post-decision state-value function}} \right\}$$

  - Which DU should be transmitted first?
  - How much data should be transmitted for each DU?

Priority-based scheduling

- Prioritization
  - Based on distortion impacts, delay deadlines and dependencies



Separate foresighted decision across DUs

- Priority-based scheduling
  - If there is only one DU with the highest priority, transmit the data in this DU by solving the foresighted optimization;
  - If there are multiple DUs that have same priorities, solve the foresighted optimization for each DU, transmit the data from the DU with highest long-term utility.

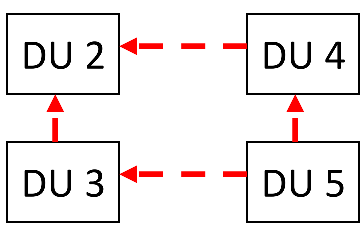
Single-DU foresighted decision:

$$V_t^i = \max_{y_t^i \in \mathcal{Y}(h_t)} \{\tilde{u}_i(x_t^i, h_t, \sum_{j \lhd i} y_t^{j*}, y_t^i) + \alpha U_i(c_t, x_t^i - y_t^i, h_t)\}$$

$j \lhd i$: DU j has higher priority than DU i.

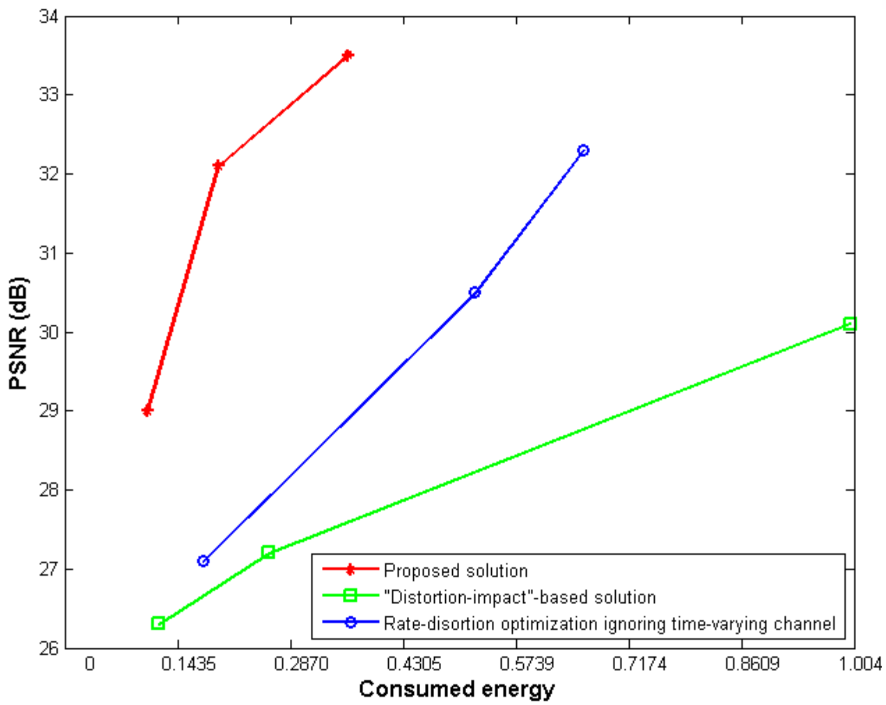
One dimensional concave function given c_t and h_t .

It can be updated using the proposed online learning.

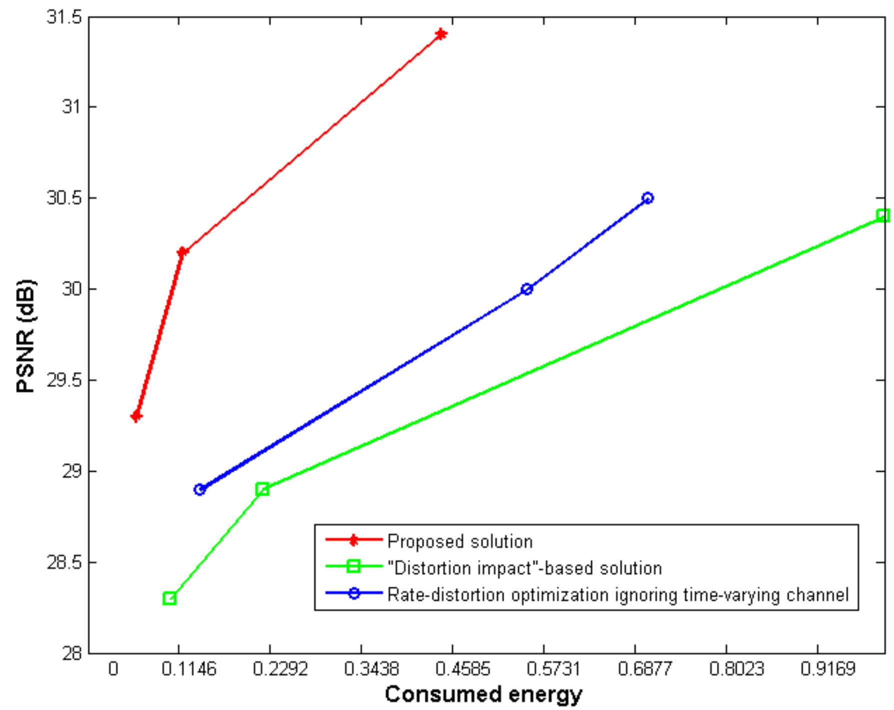


Multi-DU foresighted decision → Multiple single-DU foresighted decision

Simulation results for single-user transmission



Foreman



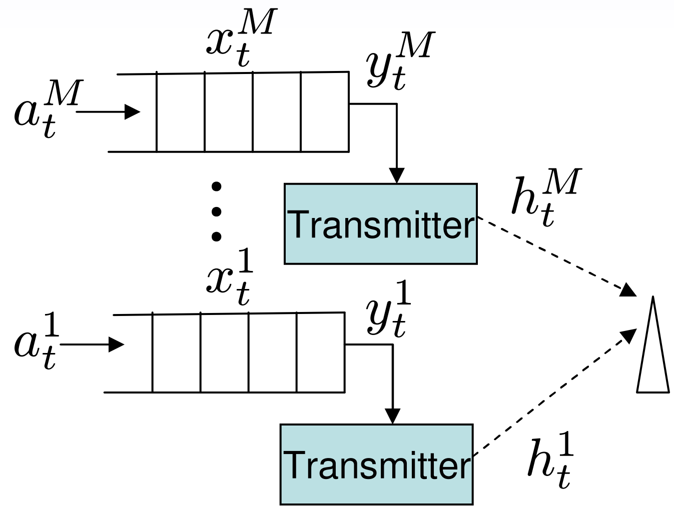
Coastguard

Channel: Rayleigh fading, modeled as 8-state Markov chain

Roadmap

- Separation principle 1 (improving learning efficiency)
  - Post-decision state-based formulation
  - Structure-aware online learning with adaptive approximation
- Separation principle 2 (Separating the foresighted decision for heterogeneous media data transmission)
  - Context-based state
  - Priority-based scheduling
- Separation principle 3 (decomposing multi-user coupling)
  - Multi-user Markov decision process formulation
  - Post-decision state value function decomposition

Delay-sensitive multi-access communications

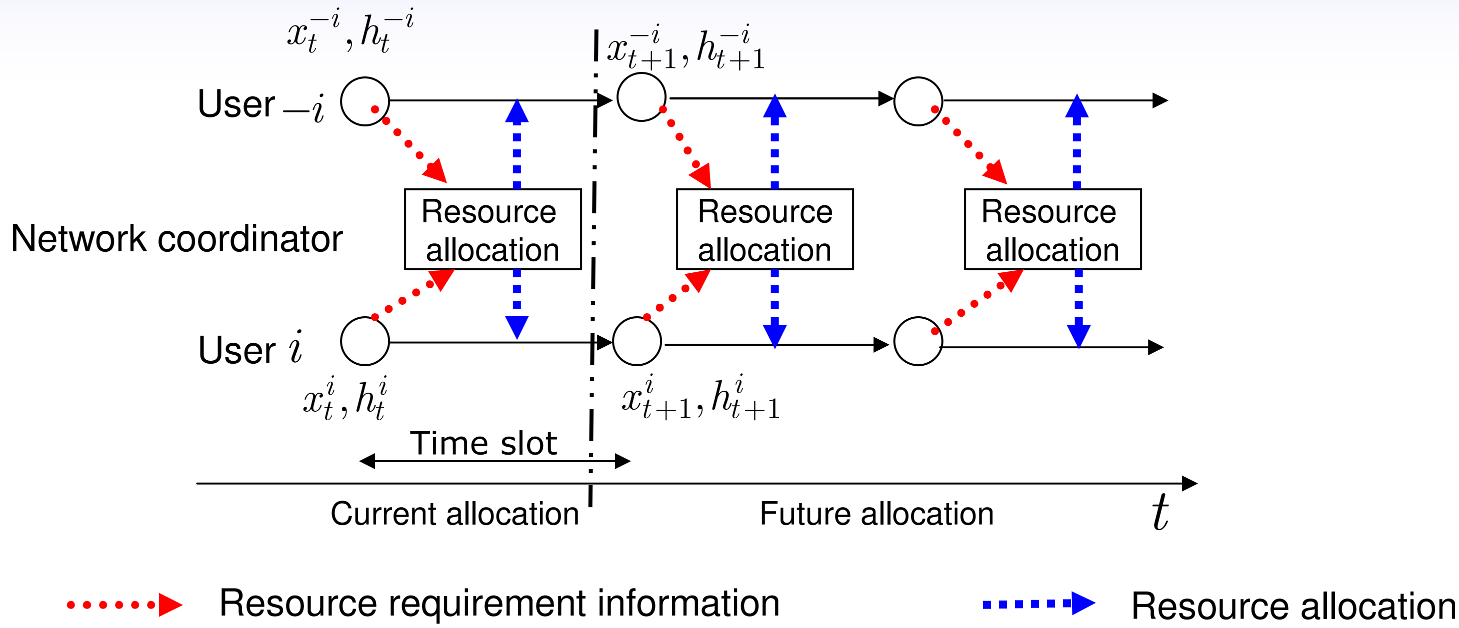


$$\max_{\mathbf{y}_t, \forall t} \mathbf{E} \sum_{t=0}^{\infty} \alpha^t \sum_{i=1}^{M} u_i(x_t^i, h_t^i, y_t^i)$$

$$\text{s.t. } [y_t^1, \cdots, y_t^M] \in \Pi(\mathbf{h}_t), \forall t \geq 0$$

Resource constraint (e.g. transmission time constraint in TDMA)

Foresighted optimization formulation



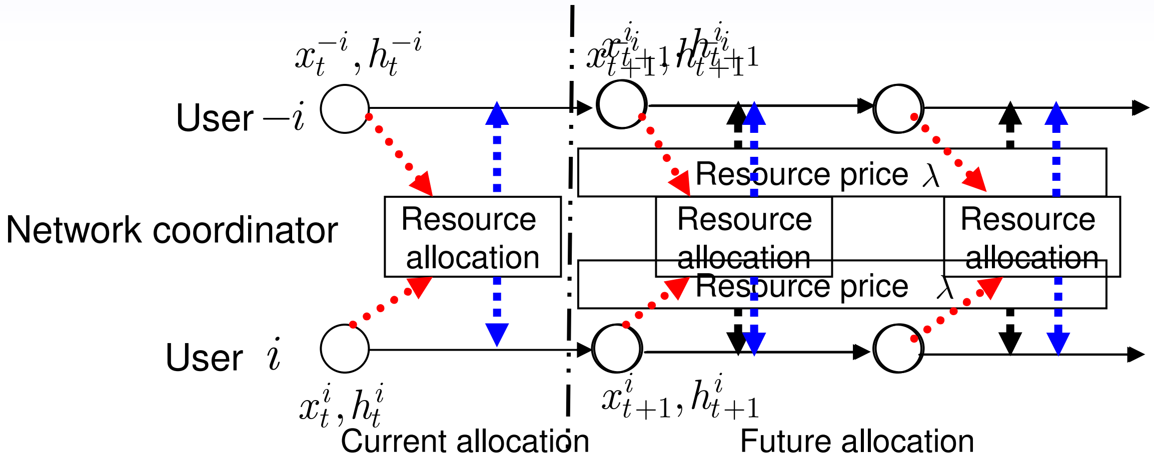
- Formulate as Multi-user MDP (MUMDP) and perform foresighted decision

Depends on all users state

$$V(\mathbf{x}_t, \mathbf{h}_t) = \max_{\mathbf{y}_t \in \Pi(\mathbf{h}_t)} \{ \sum_{i=1}^{M} u_i(x_t^i, h_t^i, y_t^i) + \alpha \boxed{U(\mathbf{x}_t - \mathbf{y}_t, \mathbf{h}_t)} \}$$

Our goal: decouple the post-decision state value function across users

Decomposition of post-decision state-value function



- Relax the resource constraints (e.g. TDMA-like access)

$$\sum_{i=1}^{M} \frac{y_k^i}{R(h_k^i)} \leq 1, \forall k = t+1, \cdots \quad \Rightarrow \quad \sum_{k=t+1}^{\infty} \alpha^k \sum_{i=1}^{M} \frac{y_k^i}{R(h_k^i)} \leq \frac{1}{1-\alpha}$$

Access time

- Introduce scalar resource price λ , and compute post-decision state-value function $U_i^{\lambda}(x_t^i, h_t^i)$ individually based on single-user MDP.

$$U_i^{\lambda}(x^i, h^i) = \max_{0 \leq y^i \leq x^i} \{u_i(x^i, h^i, y^i) - \lambda y^i / R(h^i) + \alpha U_i^{\lambda}(x^i - y^i, h^i)\}$$

- Upper bound

Resource allocation

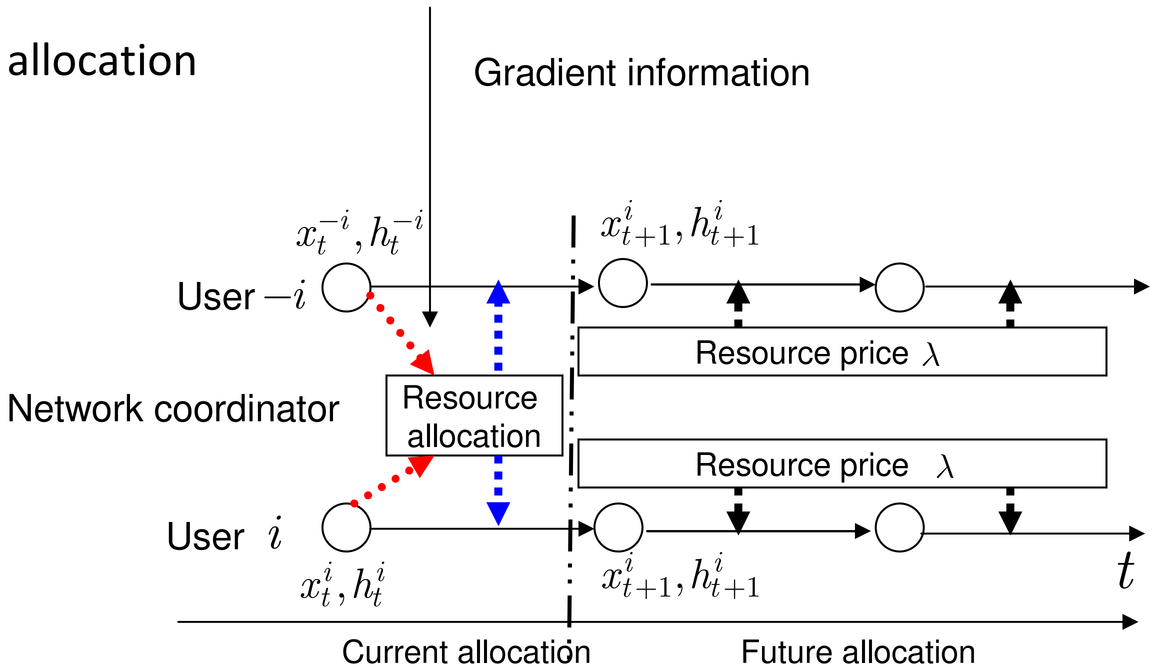
- Post-decision state value function decomposition

$$U(\mathbf{x}_t, h_t) \approx \sum_{i=1}^{M} U_i^{\lambda}(x_t^i, h_t^i)$$

- Resource allocation

$$\max_{\mathbf{y}_t \in \Pi(\mathbf{h}_t)} \sum_{i=1}^{M} \underbrace{\{u_i(x_t^i, h_t^i, y_t^i) + \alpha U_i^{\lambda}(x_t^i - y_t^i, h_t^i)\}}$$

  - Gradient-based allocation
- Lower bound



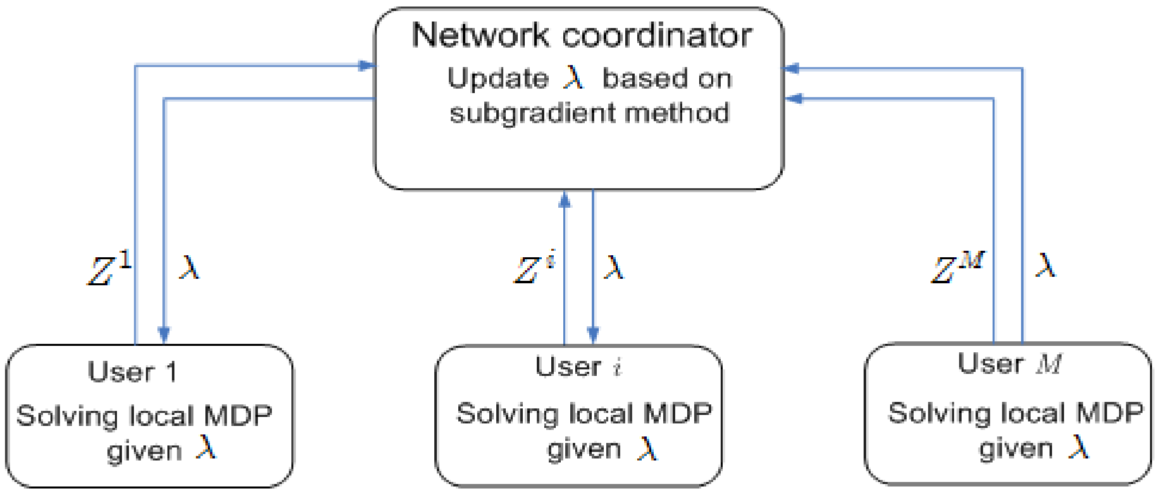
Resource price update

- Subgradient method to update resource price

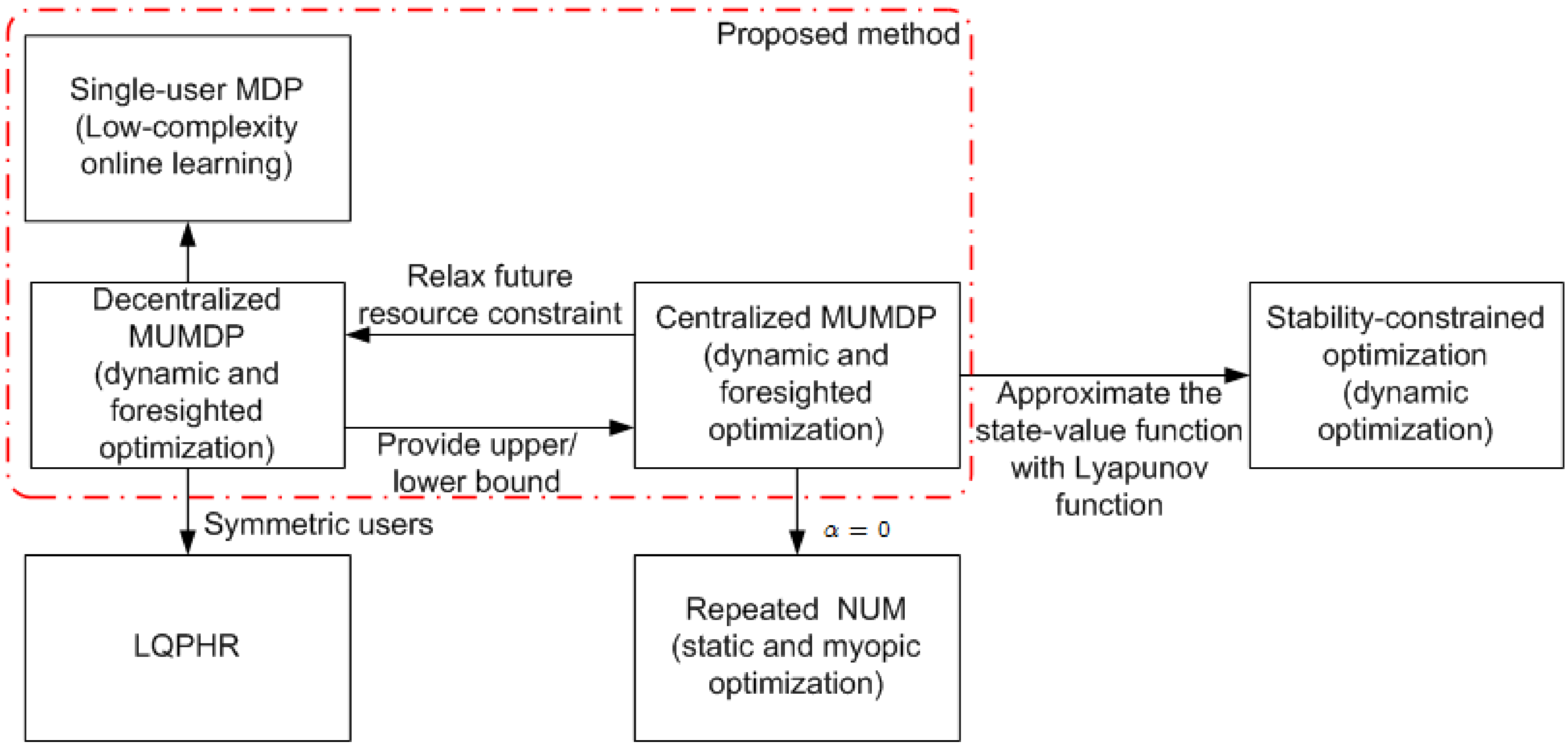
The resource price is updated by

$$\lambda^{k+1} = [\lambda^k + \beta^k \underbrace{(\sum_{i=1}^{M} Z^i - \frac{1}{1-\alpha})}_{\text{subgradient}}]^+$$

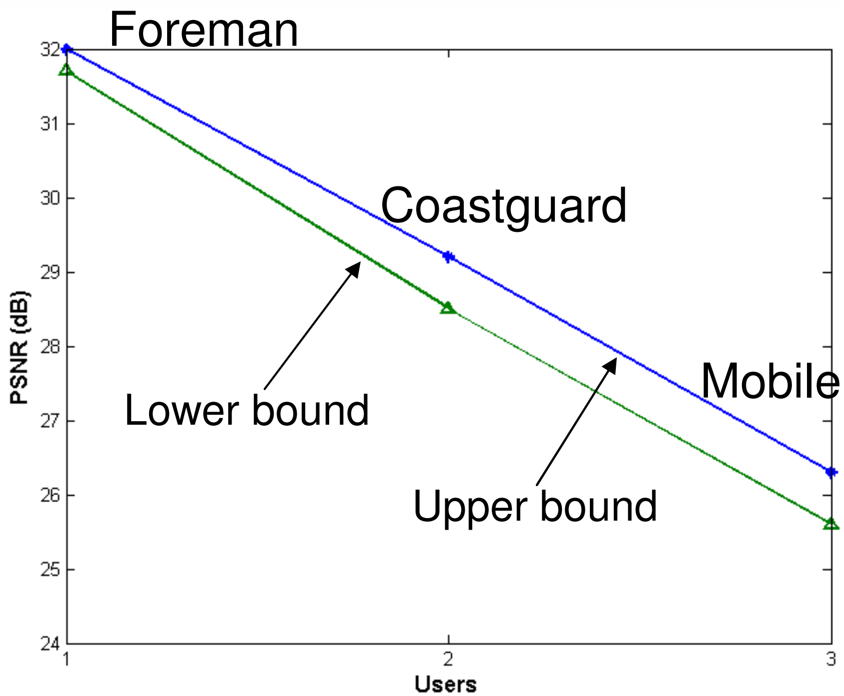
where Z^i is the expected consumed resource by user i and is individually computed by user i.



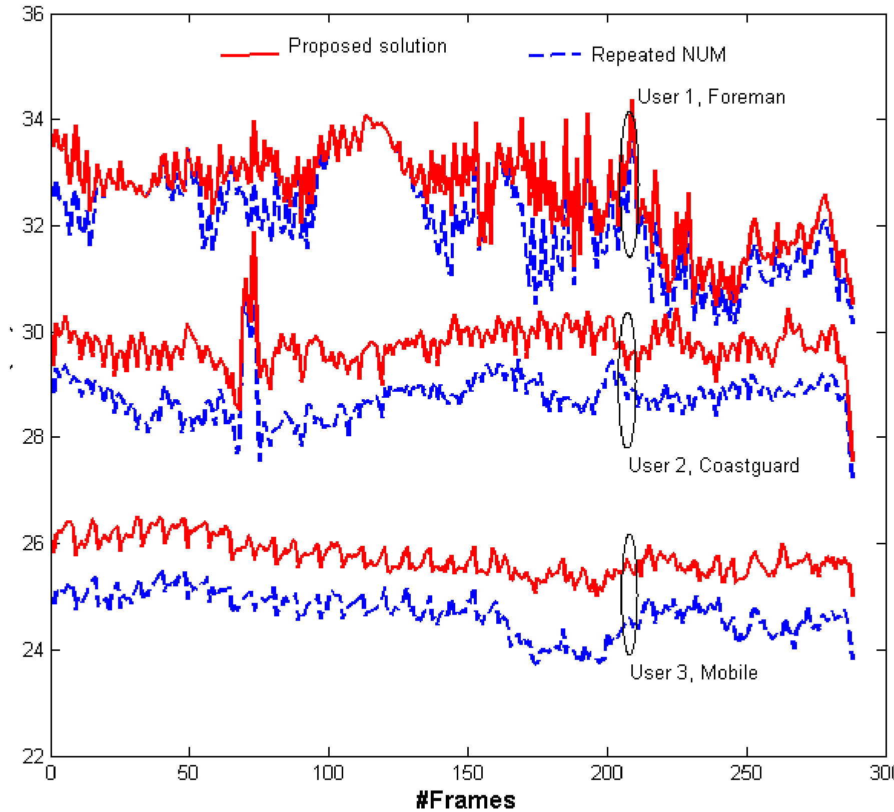
Relationship of different solutions



Simulation results for multi-user transmission



Users experienced with average channel conditions of 28dB



1. Each user uses multiple queues to represent video data;
2. Markov chain model for Rayleigh fading channel
3. TDMA-type channel access

Other applications developed in our lab

- Cross-layer optimization via layer separation [Fu 2009, Zhang 2010]
  - Each layer performs dynamic optimization individually
  - Message exchange across layers
- Media-TCP [Shiang 2010]
  - Context-based congestion control
- Dynamic voltage scaling for video decoding [Mastronarde 2009]
  - Post-decision state-based formulation
  - Context-based scheduling
- Wireless video network with cooperation [Mastronarde 2010]
  - Structure-aware online learning

Summary: separation principle 1

- Foresighted optimization framework

$$\max_y \{u(s,y) \quad + \quad \mathbf{E}_w V(f(s,y,w))\}$$

Current time slot

Next time slot

$$\max_y u(s,y) + U(g(s,y))$$

Post-decision state

$$\tilde{s} = g(s,y)$$

$$U(\tilde{s}) = \mathbf{E}_w V(g'(\tilde{s},w))$$

Current time slot

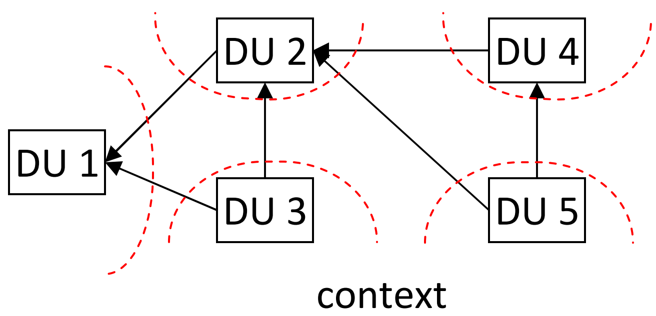
Next time slot

- *Separation principle 1*
  - Post-decision state-based foresighted optimization formulation: separation between foresighted decision and dynamics
  - Structure-aware online learning
    - Low complexity, fast convergence and achieving ε-optimal solutions

Summary: separation principle 2

- Foresighted optimization framework

$$\max_{\mathbf{y}} u(\mathbf{s},\mathbf{y}) + U(g(\mathbf{s},\mathbf{y}))$$



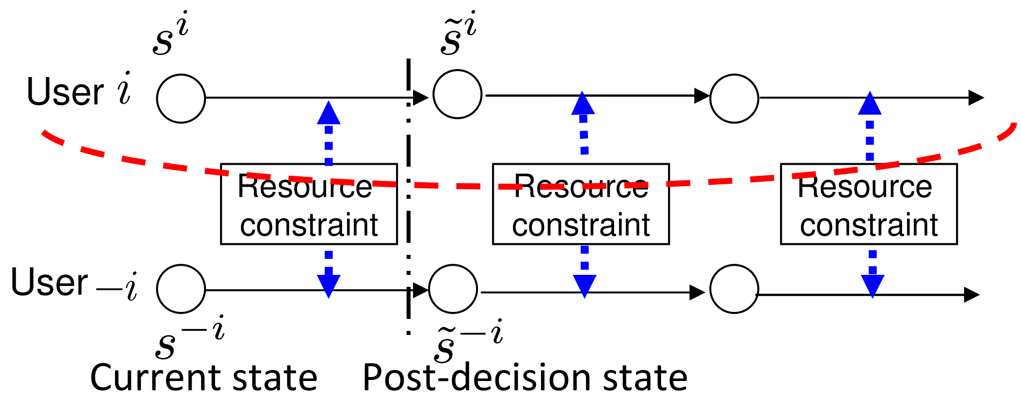
DU: data unit

- *Separation principle 2*
  - Context-based state to capture heterogeneity in data units at each time slot
  - Priority graph-based scheduling: separation across data units

Summary: separation principle 3

- Foresighted optimization framework

$$\max_{\mathbf{y}\in\Pi} \sum_{i=1}^{M} u^i(s^i, y^i) + U(g(\mathbf{s},\mathbf{y}))$$



- *Separation principle 3*
  - Decomposition of post-decision state value function: separation across users

Future research

- Extend the unified framework to
  - Multi-hop delay-sensitive data transmission
  - Non-collaborative multi-user data transmission
  - Energy-efficient parallel data processing in media systems

Related Journal Publications

[Fu10a] **Fangwen Fu,** Mihaela van der Schaar, "Structural solutions for cross-layer optimization of wireless multimedia transmission," In submission.

[Fu10b] Fangwen Fu, Mihaela van der Schaar, "Structure-aware stochastic control for transmission scheduling" in submission.

[Fu10c] Fangwen Fu, Mihaela van der Schaar, "A Systematic Framework for Dynamically Optimizing Multi-User Video Transmission," *IEEE J. Sel. Areas Commun*., vol. 28, no. 3, pp. 308-320, Apr. 2010.

[Fu10d] Fangwen Fu, Mihaela van der Schaar, "Decomposition Principles and Online Learning in Cross-Layer Optimization for Delay-Sensitive Applications", *IEEE Trans. Signal Process.*, vol 58, no. 3, pp. 1401-1415, Feb. 2010.

[Fu09a] Mihaela van der Schaar and **Fangwen Fu**, "Spectrum Access Games and Strategic Learning in Cognitive Radio Networks for Delay-Critical Applications," *Proc. of IEEE, Special issue on Cognitive Radio*, vol. 97, no. 4, pp. 720-740, Apr. 2009.

[Fu09b] Yu Zhang, **Fangwen Fu**, Mihaela van der Schaar, "On-line Learning and Optimization for Wireless Video Transmission," *IEEE Transactions on Signal Processing*, accepted, 2009.

[Fu09c] Fangwen Fu, Mihaela van der Schaar, "A New Systematic Framework for Autonomous Cross-Layer Optimization," *IEEE Trans. Veh. Tech*., vol. 58, no. 4, pp. 1887-1903, May, 2009.

[Fu09d] Fangwen Fu, Mihaela van der Schaar, "Learning to Compete for Resources in Wireless Stochastic Games," *IEEE Trans. Veh. Tech.*, vol. 58, no. 4, pp. 1904-1919, May 2009.

Acknowledgements

- PhD committee: Professor Mihaela van der Schaar, Lixia Zhang, Jason Speyer, Lieven Vandenberghe, and Gregory J. Pottie

- Labmates: Brian Foo, Hyunggon Park, Nick Mastronarde, Brian Foo, Xiaolin Tong, Yi Su, Yu Zhang, Shaolei Ren, Jaeok Park, Khoa Tran Phan, Zhichu Lin, and Yuanzhang Xiao

- Intern mentors: Dr. Deepak Turaga, Dr. Olivier Verscheure, and Dr. Ulas Kozat

- Collaborators: Dr. Tudor Stoenescu, Dr. Ulrich Berthold, Dr. Ahmad Fattahi

- Family: my wife, parents, sister and brother