FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Distributed Learning, Estimation and Control In the Routing Game

Walid Krichene **Alexandre Bayen**

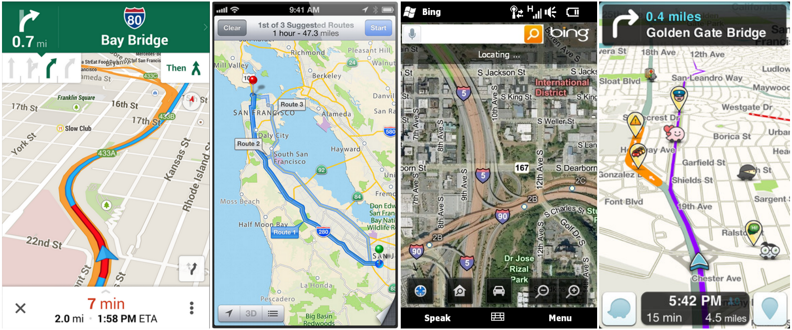
Dept. of Electrical Engineering & Computer Sciences,
UC Berkeley, CA, USA

Learning dynamics in the routing game

- Routing games model congestion on networks.
- Nash equilibrium quantifies efficiency of network in steady state.

System does not operate at equilibrium. Beyond equilibria, we need to understand decision dynamics (learning).

- A realistic model for decision dynamics is essential for prediction, control.



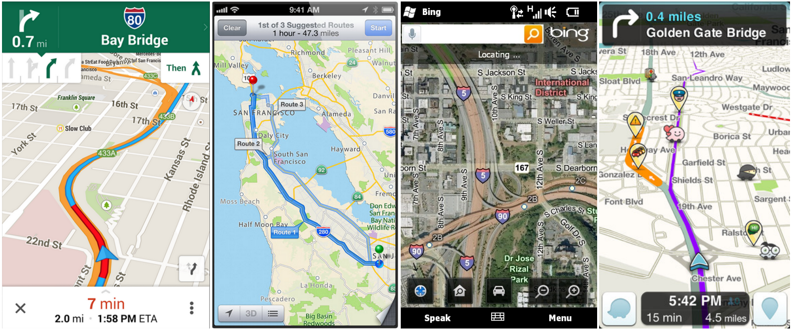
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Learning dynamics in the routing game

- Routing games model congestion on networks.
- Nash equilibrium quantifies efficiency of network in steady state.

System does not operate at equilibrium. Beyond equilibria, we need to understand decision dynamics (learning).

- A realistic model for decision dynamics is essential for prediction, control.



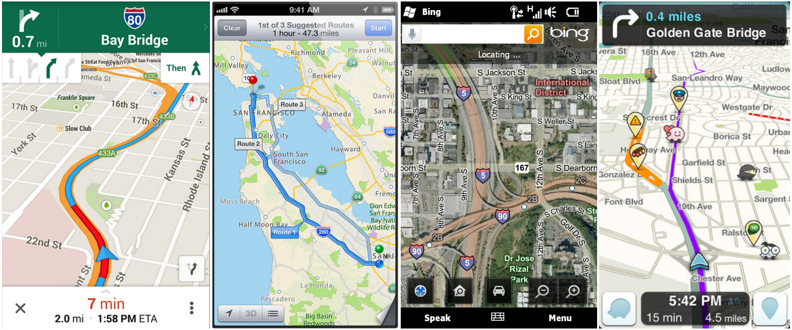
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Learning dynamics in the routing game

- Routing games model congestion on networks.
- Nash equilibrium quantifies efficiency of network in steady state.

System does not operate at equilibrium. Beyond equilibria, we need to understand decision dynamics (learning).

- A realistic model for decision dynamics is essential for prediction, control.



FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Desiderata

Learning dynamics should be

- Realistic in terms of information requirements, computational complexity.
- Consistent with the full information Nash equilibrium.

$$x^{(t)} \to \mathcal{X}^{\star}$$

- Robust to stochastic perturbations, e.g. observation noise.

Desiderata

Learning dynamics should be

- Realistic in terms of information requirements, computational complexity.
- Consistent with the full information Nash equilibrium.

$$x^{(t)} \to \mathcal{X}^{\star}$$

- Robust to stochastic perturbations, e.g. observation noise.

Desiderata

Learning dynamics should be

- Realistic in terms of information requirements, computational complexity.
- Consistent with the full information Nash equilibrium.

$$x^{(t)} \to \mathcal{X}^{\star}$$

- Robust to stochastic perturbations, e.g. observation noise.

Outline

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Outline

Online learning in the routing game

- Player drives from source to destination node
- Chooses path from $\mathcal{A}_k$
- Mass of players on each edge determines cost on that edge.

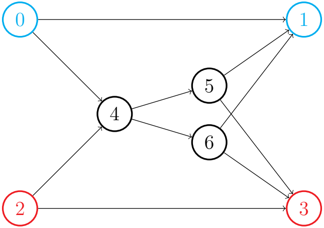


Figure: Routing game

Online learning in the routing game

- Player drives from source to destination node
- Chooses path from $\mathcal{A}_k$
- Mass of players on each edge determines cost on that edge.

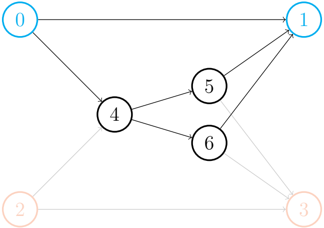


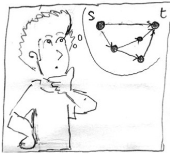
Figure: Routing game

Online learning in the routing game

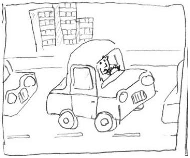
Online Learning Model

for $t \in \mathbb{N}$ **do**
  Play $p \sim x^{(t)}_{\mathcal{A}_k}$
  Discover $\ell^{(t)}_{\mathcal{A}_k}$
  Update $x^{(t+1)}_{\mathcal{A}_k} = u_k\left(x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k}\right)$
end for

$x^{(t)}_{\mathcal{A}_1} \in \Delta^{\mathcal{A}_1}$

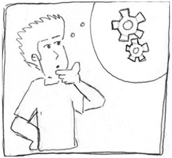


Sample $p \sim x^{(t)}_{\mathcal{A}_1}$



Discover $\ell^{(t)}_{\mathcal{A}_1}$

Update $x^{(t+1)}_{\mathcal{A}_1}$



Main problem

Define class of dynamics $\mathcal{C}$ such that

$$u_k \in \mathcal{C}\ \forall k \Rightarrow x^{(t)} \to \mathcal{X}^\star$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Online learning in the routing game

Online Learning Model

for $t \in \mathbb{N}$ **do**
  Play $p \sim x^{(t)}_{\mathcal{A}_k}$
  Discover $\ell^{(t)}_{\mathcal{A}_k}$
  Update $x^{(t+1)}_{\mathcal{A}_k} = u_k\left(x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k}\right)$
end for

$x^{(t)}_{\mathcal{A}_1} \in \Delta^{\mathcal{A}_1}$ | Sample $p \sim x^{(t)}_{\mathcal{A}_1}$ | Discover $\ell^{(t)}_{\mathcal{A}_1}$ | Update $x^{(t+1)}_{\mathcal{A}_1}$

Main problem

Define class of dynamics $\mathcal{C}$ such that

$$u_k \in \mathcal{C}\ \forall k \Rightarrow x^{(t)} \to \mathcal{X}^\star$$

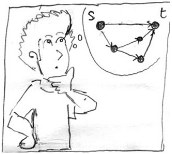
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Online learning in the routing game

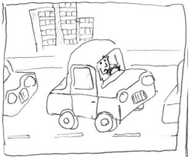
Online Learning Model

for $t \in \mathbb{N}$ **do**
  Play $p \sim x^{(t)}_{\mathcal{A}_k}$
  Discover $\ell^{(t)}_{\mathcal{A}_k}$
  Update $x^{(t+1)}_{\mathcal{A}_k} = u_k\left(x^{(t)}_{\mathcal{A}_k}, \ell^{(t)}_{\mathcal{A}_k}\right)$
end for

$x^{(t)}_{\mathcal{A}_1} \in \Delta^{\mathcal{A}_1}$

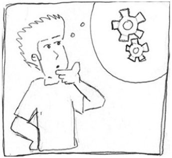


Sample $p \sim x^{(t)}_{\mathcal{A}_1}$



Discover $\ell^{(t)}_{\mathcal{A}_1}$

Update $x^{(t+1)}_{\mathcal{A}_1}$



Main problem

Define class of dynamics $\mathcal{C}$ such that

$$u_k \in \mathcal{C} \ \forall k \Rightarrow x^{(t)} \to \mathcal{X}^\star$$

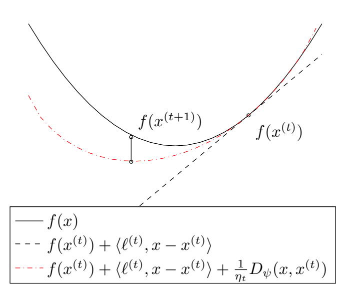
Stochastic convex optimization

Idea:

- The set of Nash equilibria is $\arg\min_{x\in\mathcal{X}} f(x)$ (the Rosenthal potential).
- View the learning dynamics as a distributed algorithm to minimize f.

Algorithm 1 MD Method

for $t \in \mathbb{N}$ **do**
  observe $\ell^{(t)} \in \partial f(x^{(t)})$
  $x^{(t+1)} = \arg\min_{x\in\mathcal{X}} \left\langle \ell^{(t)}, x \right\rangle + \frac{1}{\eta_t} D_\psi(x, x^{(t)})$
end for



- η_t: learning rate
- D_ψ: Bregman divergence

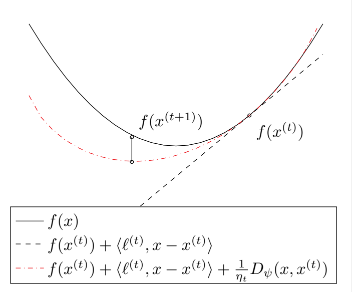
Stochastic convex optimization

Idea:

- The set of Nash equilibria is $\arg\min_{x \in \mathcal{X}} f(x)$ (the Rosenthal potential).
- View the learning dynamics as a distributed algorithm to minimize f.

Algorithm 2 MD Method

1: **for** $t \in \mathbb{N}$ **do**
2: observe $\ell_{\mathcal{A}_k}^{(t)} \in \partial_{\mathcal{A}_k} f(x^{(t)})$
3: $x_{\mathcal{A}_k}^{(t+1)} = \arg\min_{x \in \mathcal{X}_{\mathcal{A}_k}} \left\langle \ell_{\mathcal{A}_k}^{(t)}, x \right\rangle + \frac{1}{\eta_t^k} D_{\psi_k}(x, x_{\mathcal{A}_k}^{(t)})$
4: **end for**



- η_t: learning rate
- D_ψ: Bregman divergence

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Stochastic convex optimization

Idea:

- The set of Nash equilibria is $\arg\min_{x\in\mathcal{X}} f(x)$ (the Rosenthal potential).
- View the learning dynamics as a distributed algorithm to minimize f.

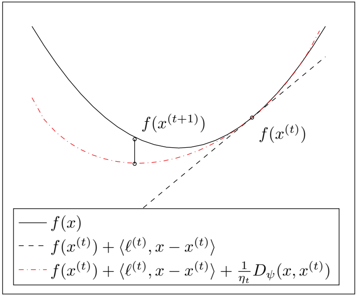
Algorithm 2 SMD Method

1: **for** $t \in \mathbb{N}$ **do**

2: observe $\hat{\ell}^{(t)}_{\mathcal{A}_k}$ with $\mathbb{E}\left[\hat{\ell}^{(t)}_{\mathcal{A}_k} | \mathcal{F}_{t-1}\right] \in \partial_{\mathcal{A}_k} f(x^{(t)})$

3: $x^{(t+1)}_{\mathcal{A}_k} = \arg\min_{x\in\mathcal{X}_{\mathcal{A}_k}} \left\langle \hat{\ell}^{(t)}_{\mathcal{A}_k}, x \right\rangle + \frac{1}{\eta^k_t} D_{\psi_k}(x, x^{(t)}_{\mathcal{A}_k})$

4: **end for**



- η_t: learning rate
- D_ψ: Bregman divergence

Convergence

Convergence of Distributed Stochastic Mirror Descent

For $\eta_t^k = \frac{\theta_k}{t^{\alpha_k}}, \ \alpha_k \in (0,1)$,

$$\mathbb{E}\left[f(x^{(t)})\right] - f^\star = \mathcal{O}\left(\sum_k \frac{\log t}{t^{\min(\alpha_k, 1-\alpha_k)}}\right)$$

[1] Syrine Krichene, Walid Krichene, Roy Dong, and Alexandre Bayen. Convergence of heterogeneous distributed learning in stochastic routing games.
In *53rd Allerton Conference on Communication, Control and Computing*, 2015

Convergence

Convergence of Distributed Stochastic Mirror Descent

For $\eta_t^k = \frac{\theta_k}{t^{\alpha_k}}$, $\alpha_k \in (0,1)$,

$$\mathbb{E}\left[f(x^{(t)})\right] - f^\star = \mathcal{O}\left(\sum_k \frac{\log t}{t^{\min(\alpha_k, 1-\alpha_k)}}\right)$$

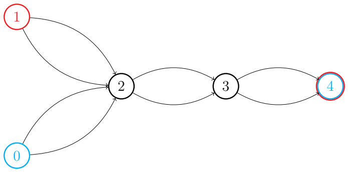


Figure: Example network with 2 populations.

- Centered Gaussian noise on edges.
- Population 1: Hedge with $\eta_t^1 = t^{-.3}$
- Population 2: Hedge with $\eta_t^2 = t^{-.4}$

Convergence

Convergence of Distributed Stochastic Mirror Descent

For $\eta_t^k = \frac{\theta_k}{t^{\alpha_k}}, \ \alpha_k \in (0,1)$,

$$\mathbb{E}\left[f(x^{(t)})\right] - f^\star = \mathcal{O}\left(\sum_k \frac{\log t}{t^{\min(\alpha_k, 1-\alpha_k)}}\right)$$

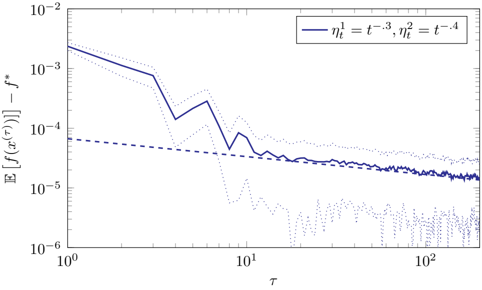


Figure: Potential values.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Convergence

Convergence of Distributed Stochastic Mirror Descent

For $\eta_t^k = \frac{\theta_k}{t^{\alpha_k}}$, $\alpha_k \in (0,1)$,

$$\mathbb{E}\left[f(x^{(t)})\right] - f^\star = \mathcal{O}\left(\sum_k \frac{\log t}{t^{\min(\alpha_k, 1-\alpha_k)}}\right)$$

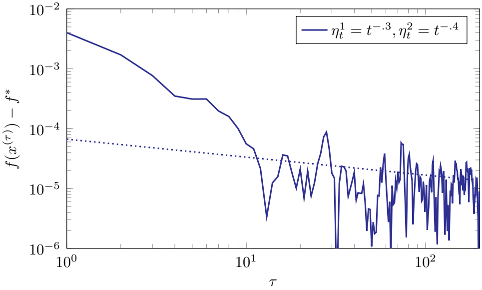


Figure: Potential values.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Outline

1 Distributed learning model

2 Estimation of Learning Rates

3 Optimal Control

A routing experiment

- Interface for the routing game.
- Used to collect sequence of decisions $\bar{x}^{(t)}$.

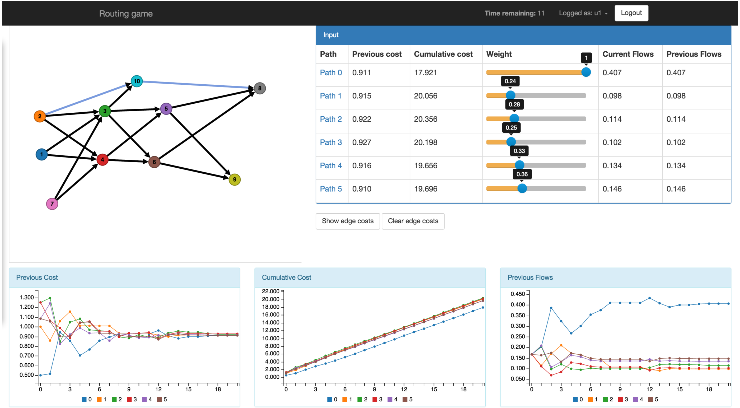


Figure: Interface for the routing game experiment.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

A routing experiment

- Interface for the routing game.
- Used to collect sequence of decisions $\bar{x}^{(t)}$.

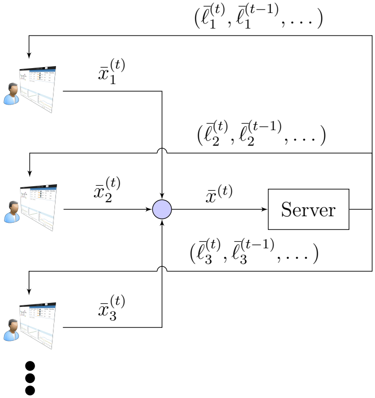


Figure: Interface for the routing game experiment.

Estimation of learning dynamics

- We observe a sequence of player decisions $(\bar{x}^{(t)})$ and losses $(\bar{\ell}^{(t)})$.
- Can we fit a model of player dynamics?

Mirror descent model

Estimate the learning rate in the mirror descent model

$$x^{(t+1)}(\eta) = \arg\min_{x \in \Delta^{\mathcal{A}_k}} \left\langle \bar{\ell}^{(t)}, x \right\rangle + \frac{1}{\eta} D_{KL}(x, \bar{x}^{(t)})$$

Then $d(\eta) = D_{KL}(\bar{x}^{(t+1)}, x^{(t+1)}(\eta))$ is a convex function. Can minimize it to estimate $\eta_k^{(t)}$.

[2]Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game.
In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Estimation of learning dynamics

- We observe a sequence of player decisions $(\bar{x}^{(t)})$ and losses $(\bar{\ell}^{(t)})$.
- Can we fit a model of player dynamics?

Mirror descent model

Estimate the learning rate in the mirror descent model

$$x^{(t+1)}(\eta) = \arg\min_{x \in \Delta^{\mathcal{A}_k}} \left\langle \bar{\ell}^{(t)}, x \right\rangle + \frac{1}{\eta} D_{KL}(x, \bar{x}^{(t)})$$

Then $d(\eta) = D_{KL}(\bar{x}^{(t+1)}, x^{(t+1)}(\eta))$ is a convex function. Can minimize it to estimate $\eta_k^{(t)}$.

---

[2]Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game.
In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015

Estimation of learning dynamics

- We observe a sequence of player decisions $(\bar{x}^{(t)})$ and losses $(\bar{\ell}^{(t)})$.
- Can we fit a model of player dynamics?

Mirror descent model

Estimate the learning rate in the mirror descent model

$$x^{(t+1)}(\eta) = \arg\min_{x \in \Delta^{\mathcal{A}_k}} \left\langle \bar{\ell}^{(t)}, x \right\rangle + \frac{1}{\eta} D_{KL}(x, \bar{x}^{(t)})$$

Then $d(\eta) = D_{KL}(\bar{x}^{(t+1)}, x^{(t+1)}(\eta))$ is a convex function. Can minimize it to estimate $\eta_k^{(t)}$.

---

[2]Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game.
In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015

Preliminary results

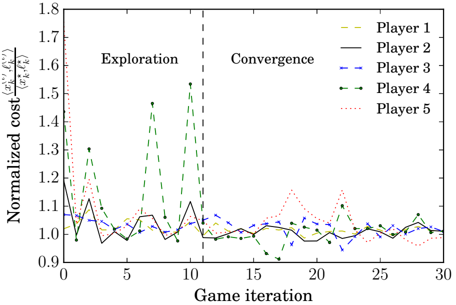


Figure: Costs of each player (normalized by the equilibrium cost)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Preliminary results

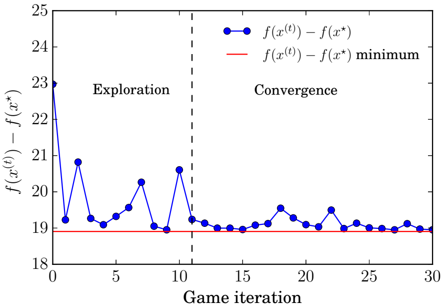


Figure: Potential function $f(x^{(t)}) - f^{\star}$.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Preliminary results

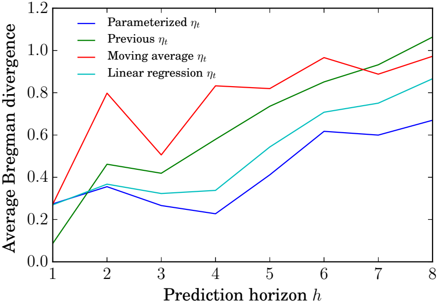


Figure: Average KL divergence between predicted distributions and actual distributions, as a function of the prediction horizon h.

Outline

1. Distributed learning model
2. Estimation of Learning Rates
3. Optimal Control

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Optimal routing with learning dynamics

Assumptions

- A central authority has control over a fraction of traffic: $u^{(t)}$
- Remaining traffic follows learning dynamics: $x^{(t)}$

Optimal routing under selfish learning constraints

$$\begin{aligned} \text{minimize}_{u^{(1:T)},x^{(1:T)}} \quad & \sum_{t=1}^{T} J(x^{(t)}, u^{(t)}) \\ \text{subject to} \quad & x^{(t+1)} = u(x^{(t)} + u^{(t)}, \ell(x^{(t)} + u^{(t)})) \end{aligned}$$

Solution methods:

- Greedy method: Approximate the problem with a sequence of convex problems.
- Mirror descent with the adjoint method.

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Optimal routing with learning dynamics

Assumptions

- A central authority has control over a fraction of traffic: $u^{(t)}$
- Remaining traffic follows learning dynamics: $x^{(t)}$

Optimal routing under selfish learning constraints

$$\begin{aligned} \text{minimize}_{u^{(1:T)}, x^{(1:T)}} \quad & \sum_{t=1}^{T} J(x^{(t)}, u^{(t)}) \\ \text{subject to} \quad & x^{(t+1)} = u(x^{(t)} + u^{(t)}, \ell(x^{(t)} + u^{(t)})) \end{aligned}$$

Solution methods:

- Greedy method: Approximate the problem with a sequence of convex problems.
- Mirror descent with the adjoint method.

Application to the L.A. highway network

- Simplified model of the L.A. highway network.
- Cost functions uses the B.P.R. function, calibrated using the work of [4].

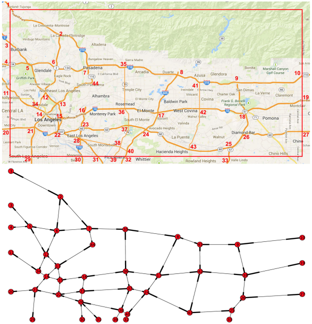


Figure: Los Angeles highway network.

---

[4]J. Thai, R. Hariss, and A. Bayen. A multi-convex approach to latency inference and control in traffic equilibria from sparse data.
In *American Control Conference (ACC), 2015*, pages 689–695, July 2015

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Application to the L.A. highway network

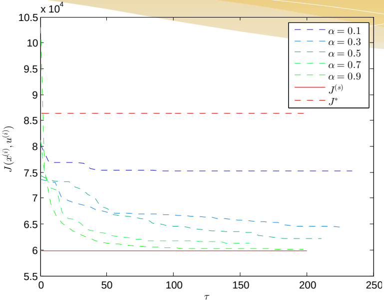


Figure: Average delay without control (dashed), with full control (solid), and different values of α.

---

[3]Milena Suarez, Walid Krichene, and Alexandre Bayen. On optimal routing under selfish learning dynamics.
Transactions on Control of Network Systems, in review, 2015

Summary and ongoing work

- A model of online learning as coupled sequential decision problems.
- Design / analysis of learning dynamics using stochastic optimization.
- Estimation of player dynamics, optimal control under learning.
- Can be applied to model predictive control.

Thank you!

eecs.berkeley.edu/∼walid/

Summary and ongoing work

- A model of online learning as coupled sequential decision problems.
- Design / analysis of learning dynamics using stochastic optimization.
- Estimation of player dynamics, optimal control under learning.
- Can be applied to model predictive control.

Thank you!

`eecs.berkeley.edu/~walid/`

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

References I

[1] Syrine Krichene, Walid Krichene, Roy Dong, and Alexandre Bayen. Convergence of heterogeneous distributed learning in stochastic routing games. In *53rd Allerton Conference on Communication, Control and Computing*, 2015.

[2] Kiet Lam, Walid Krichene, and Alexandre M. Bayen. Estimation of learning dynamics in the routing game. In *International Conference on Cyber-Physical Systems (ICCPS), in review.*, 2015.

[3] Milena Suarez, Walid Krichene, and Alexandre Bayen. On optimal routing under selfish learning dynamics. *Transactions on Control of Network Systems, in review*, 2015.

[4] J. Thai, R. Hariss, and A. Bayen. A multi-convex approach to latency inference and control in traffic equilibria from sparse data. In *American Control Conference (ACC), 2015*, pages 689–695, July 2015.