FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# Distributed learning dynamics convergence in routing games

Alex Bayen, EECS / CEE

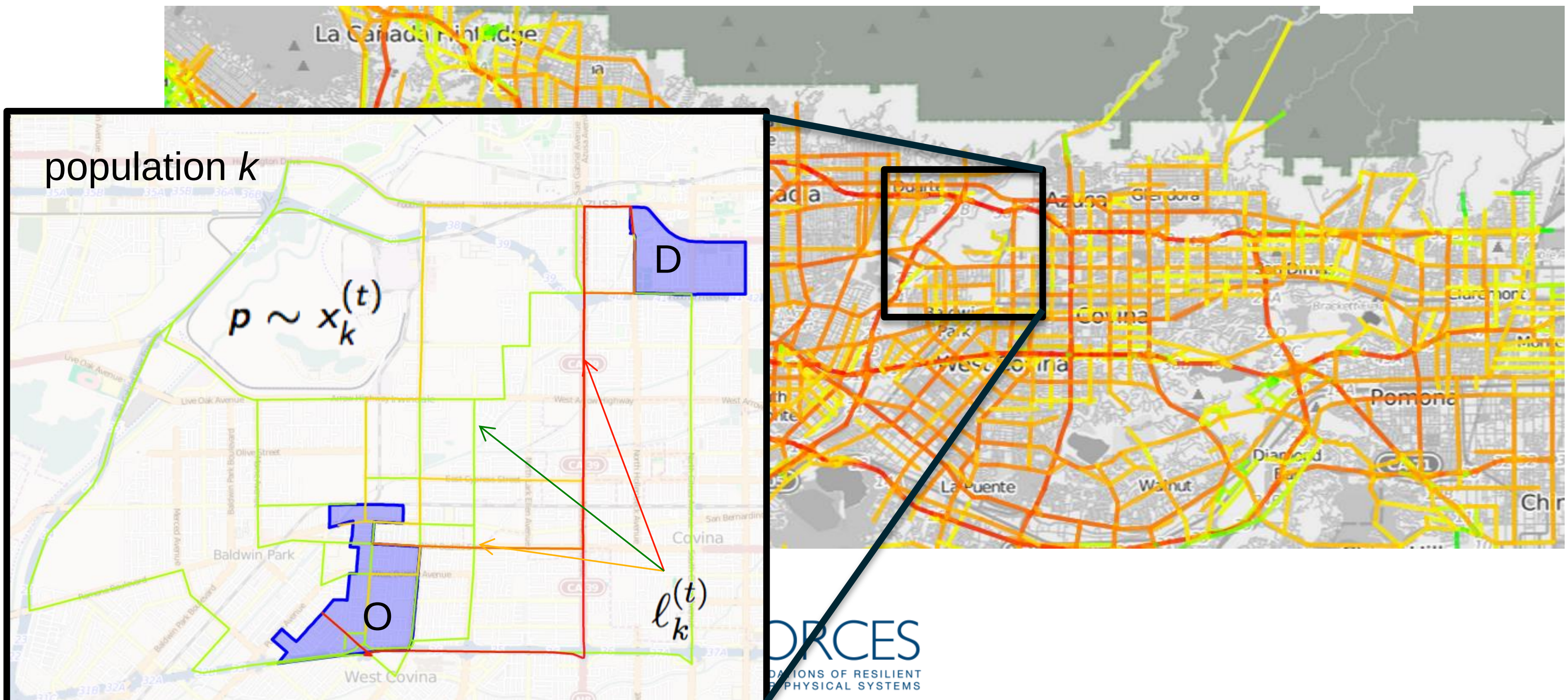
with Walid Krichene and Jerome Thai

Berkeley UNIVERSITY OF CALIFORNIA
Massachusetts Institute of Technology
NSF
MICHIGAN
VANDERBILT UNIVERSITY

# Problem formulation

Distributed learning dynamics in routing games

- Each player routes population $k$ according to distribution $p \sim x_k^{(t)}$ (corresponding to one OD pair)
- At each iteration, the population $k$ discovers their outcome $\ell_k^{(t)}$



# Problem formulation

Distributed learning dynamics in routing games

- Each player routes population *k* according to distribution $p \sim x_k^{(t)}$ (corresponding to one OD pair)
- At each iteration, the population *k* discovers their outcome $\ell_k^{(t)}$
- The routing of population *k* at the next step is subsequently updated according to the following law $x_k^{(t+1)} = u_k\left(x_k^{(t)}, \ell_k^{(t)}\right)$

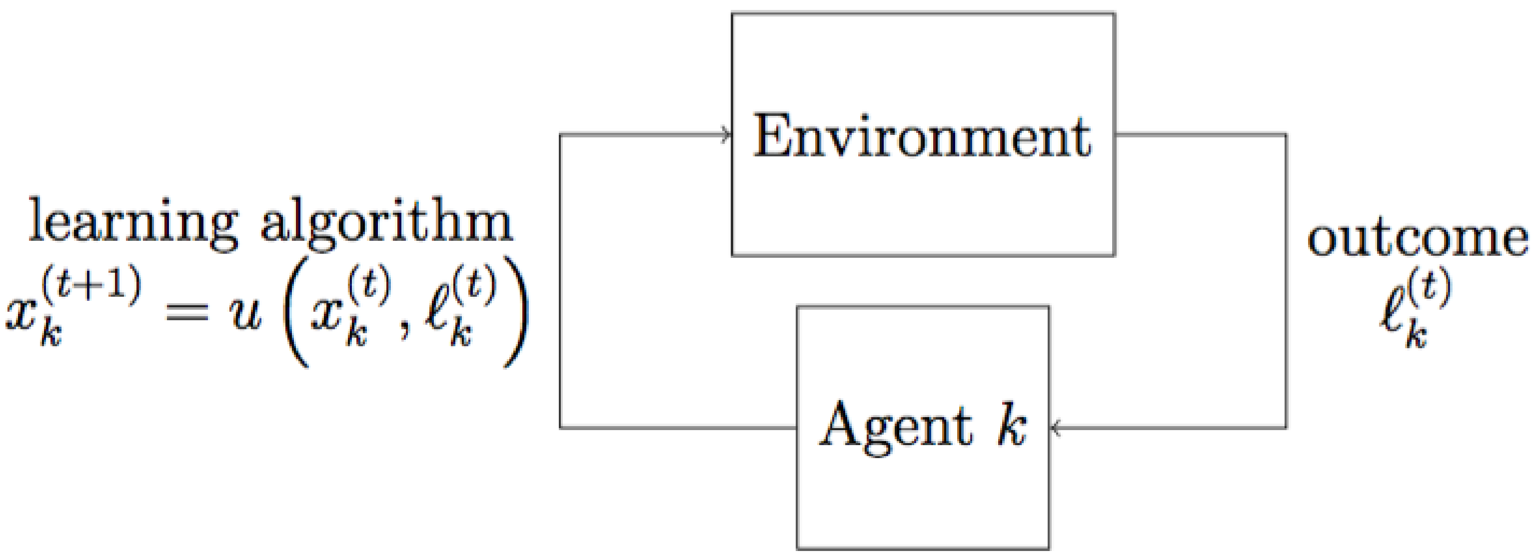
**Online Learning Model**

**for** $t \in \mathbb{N}$ **do**
    Play $p \sim x_k^{(t)}$
    Discover $\ell_k^{(t)}$
    Update $x_k^{(t+1)} = u_k\left(x_k^{(t)}, \ell_k^{(t)}\right)$
**end for**

# Sequential decision problem

Distributed learning dynamics in routing games

- Each player routes population *k* according to distribution $p \sim x_k^{(t)}$ (corresponding to one OD pair)
- At each iteration, the population *k* discovers their outcome $\ell_k^{(t)}$
- The routing of population *k* at the next step is subsequently updated according to the following law $x_k^{(t+1)} = u_k\left(x_k^{(t)}, \ell_k^{(t)}\right)$



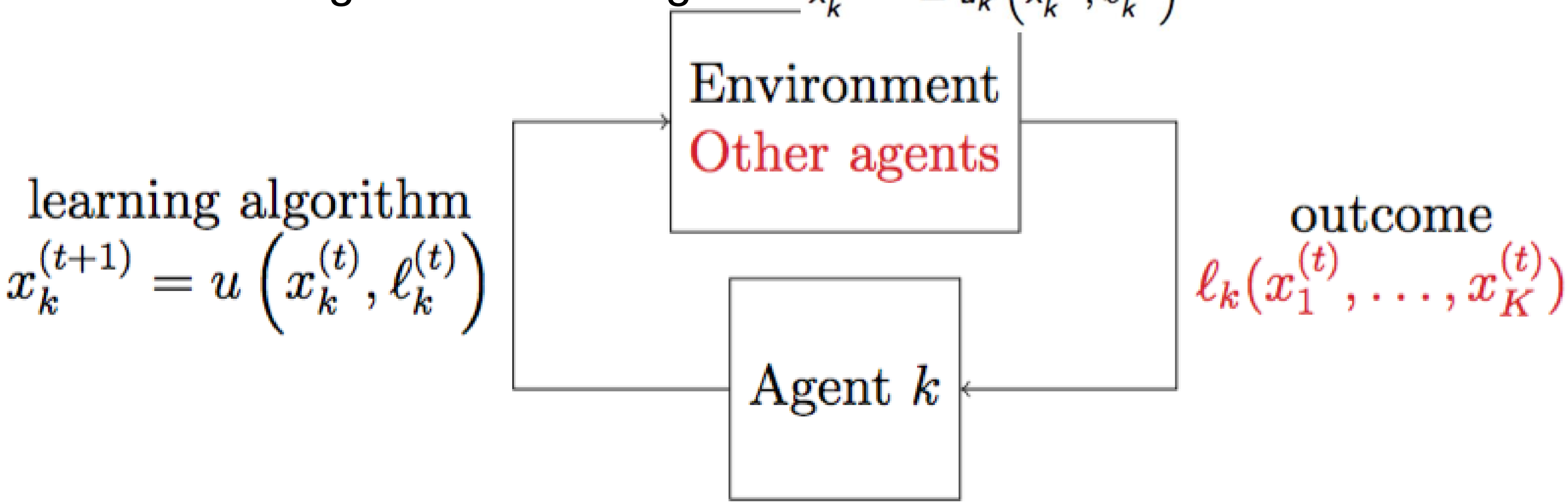
# Coupled sequential decision problem

Distributed learning dynamics in routing games

- Each player routes population *k* according to distribution $p \sim x_k^{(t)}$ (corresponding to one OD pair)

- At each iteration, the population *k* discovers their outcome $\ell_k^{(t)}$

- The routing of population *k* at the next step is subsequently updated according to the following law $x_k^{(t+1)} = u_k\left(x_k^{(t)}, \ell_k^{(t)}\right)$



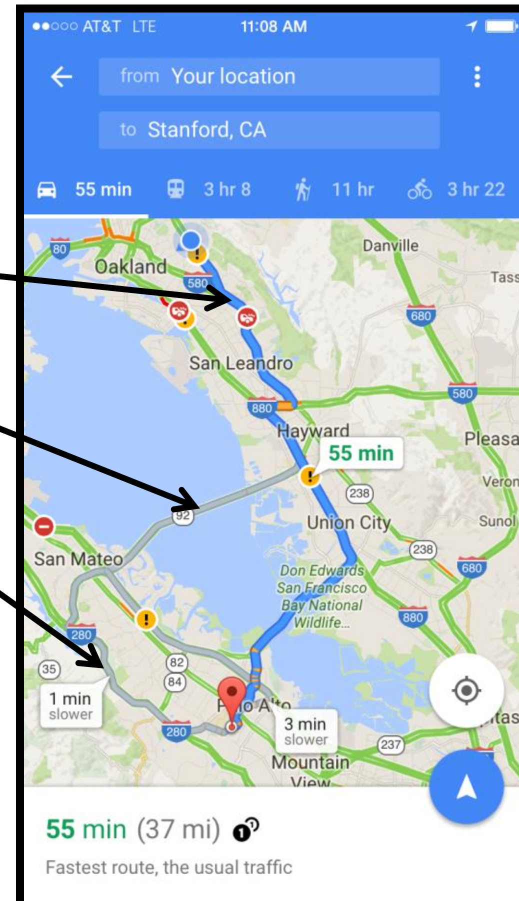
# Coupled sequential decision problem

This also represents the process of apps (companies) routing users

- Each gives shortest path (given previous information)
- Previous information is mostly statistical (experience from previous day and some statistical forecast)

All paths proposed are nearly equal:

- Shortest path (55mins)
- Third shortest path (58 mins)
- Second shortest path (56 mins)

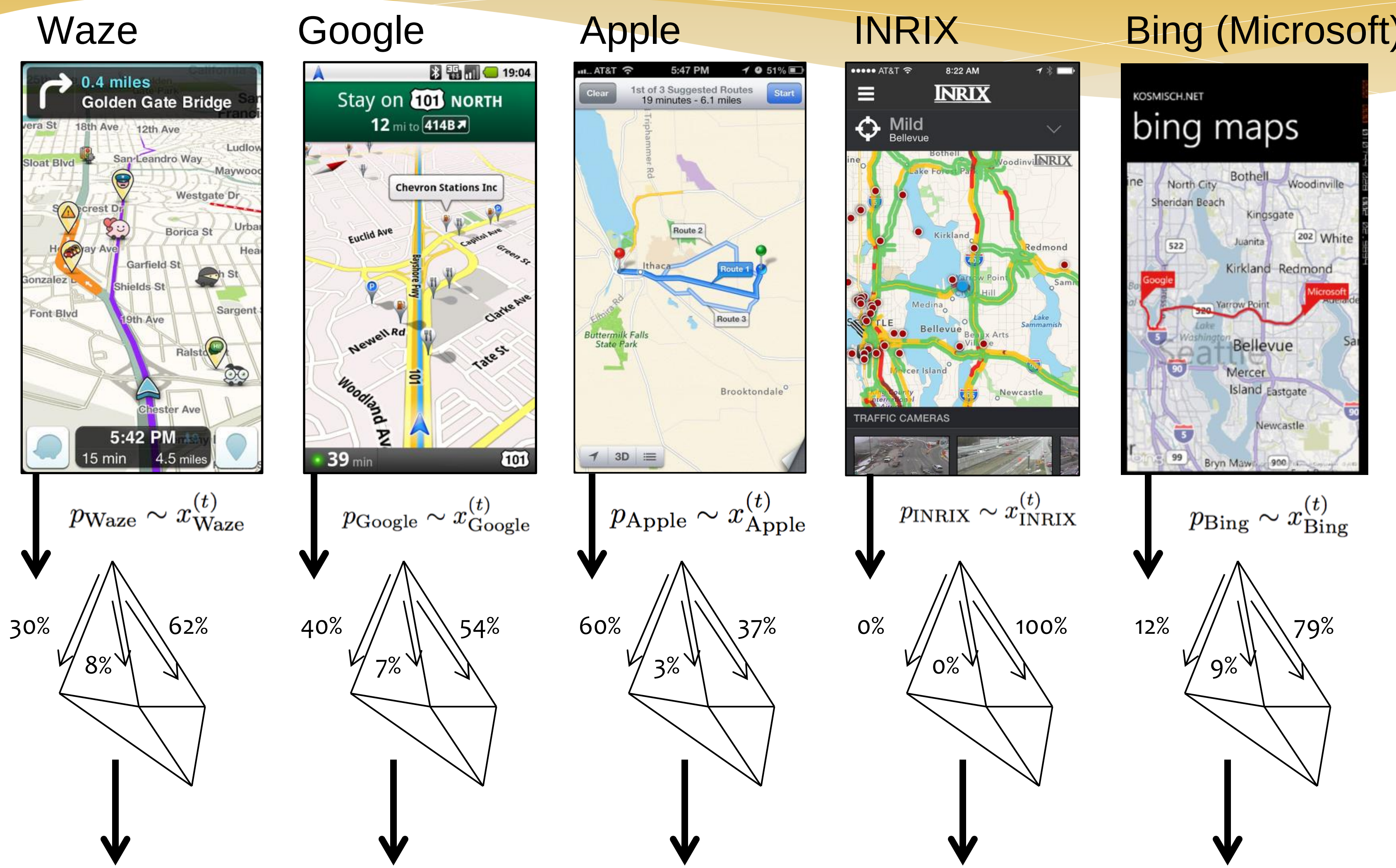



Routing does in general not depend on

- Forecast of the network loading using demand data (incomplete today)
- Forecast of the network using potential impact of routing (i.e. routed users) on the network
- Knowledge of what competitors of the app are doing (in the present case, Apple, INRIX, 511, etc.)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# Coupled sequential decision problem




All users of each company "equal" by standards of the company i.e. same (shortest) travel time according to the company, "essentially" Nash.

# Distributed learning in games

Non equilibrium situations

- Equilibria: good description of system efficiency at steady-sate.
- But systems rarely operate at equilibrium, hence
  - A prescriptive model: How do we drive system to eq.
  - A descriptive model: How would players behave in the game.

Goals of the work

- Define algorithm classes for which we can prove convergence
- Robustness to stochastic perturbations.
- Heterogeneous learning: different agents use different algorithms
- Convergence rates.

Related work

- Discrete time: Hannan consistency (Hannan 1957), Hedge algorithm for two-player games (Freund 1999), regret based algorithms: (Hart 2001), online learning in games (Cesa 2006)
- Continuous time: Potential games under dynamics with positive correlation condition (Sandholm 2009), replicator dynamics in evolutionary game theory (Weibull 1997), no-regret dynamics for two player games (Hart 2001)

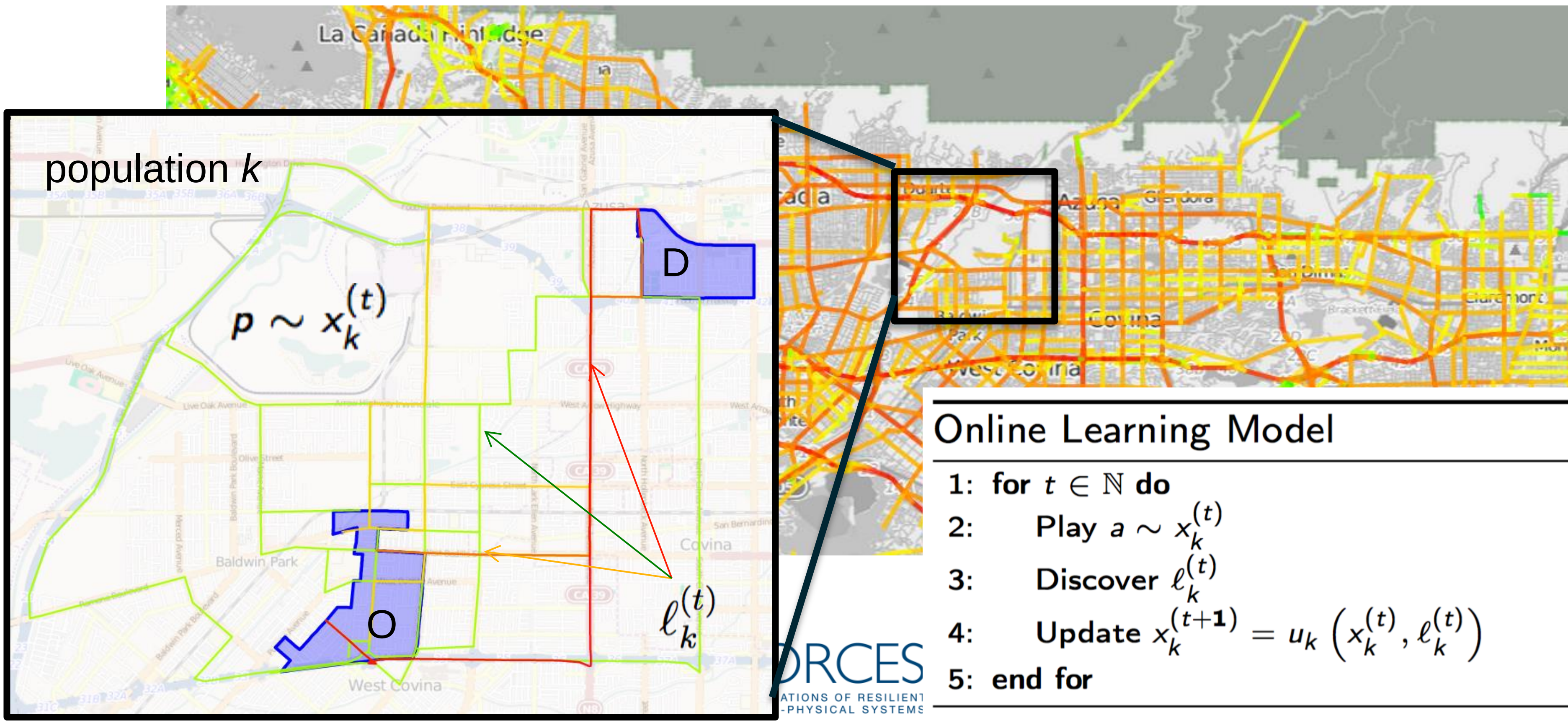
# Problem formulation

**Main problem**

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

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

Important question: what is $\mathcal{X}^\star$?




**Online Learning Model**

1: **for** $t \in \mathbb{N}$ **do**
2: Play $a \sim x_k^{(t)}$
3: Discover $\ell_k^{(t)}$
4: Update $x_k^{(t+1)} = u_k\left(x_k^{(t)}, \ell_k^{(t)}\right)$
5: **end for**

# Nash equilibrium

Write

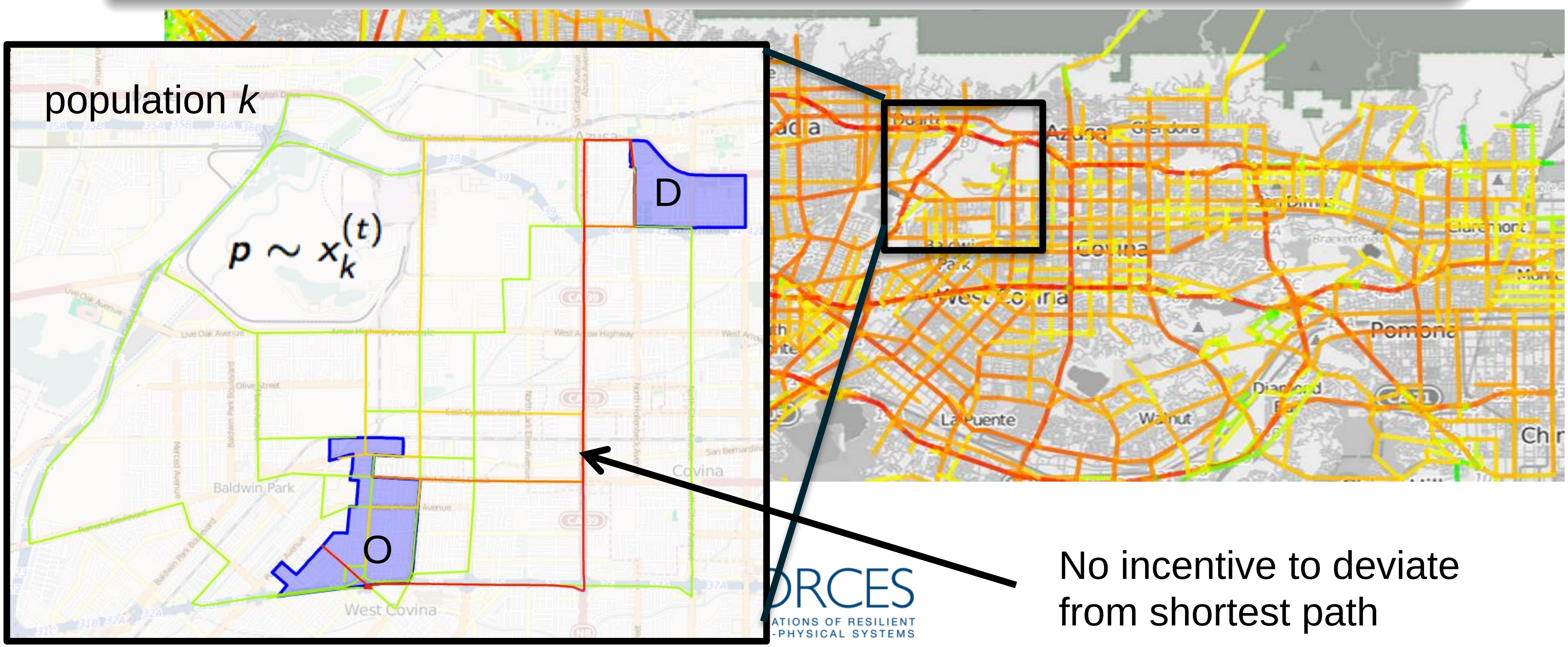
$x = (x_1, \dots, x_K) \in \Delta^{\mathcal{A}_1} \times \cdots \times \Delta^{\mathcal{A}_K}$

$\ell(x) = (\ell_1(x), \dots, \ell_K(x))$

**Nash equilibria $\mathcal{X}^\star$**

$x^\star$ is a Nash equilibrium if for all $k$, paths in the support of $x_k^\star$ have minimal loss.

$$\forall x,\ \langle \ell(x^\star), x - x^\star \rangle \geq 0$$



# Equilibrium of a game: tangential condition

Write
$x = (x_1, \dots, x_K) \in \Delta^{\mathcal{A}_1} \times \cdots \times \Delta^{\mathcal{A}_K}$
$\ell(x) = (\ell_1(x), \dots, \ell_K(x))$

**Nash equilibria $\mathcal{X}^\star$**

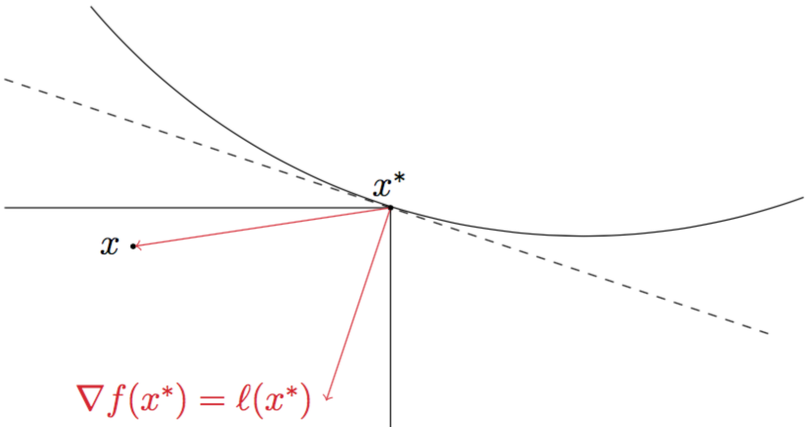
$x^\star$ is a Nash equilibrium if for all $k$, paths in the support of $x_k^\star$ have minimal loss.

$$\forall x, \ \langle \ell(x^\star), x - x^\star \rangle \geq 0$$

**Rosenthal potential**

$\exists f$ convex such that $\nabla f(x) = \ell(x)$.

| Nash condition | $\Leftrightarrow$ | first order optimality |
|---|---|---|
| $\forall x, \ \langle \ell(x^\star), x - x^\star \rangle \geq 0$ | | $\forall x, \ \langle \nabla f(x^\star), x - x^\star \rangle \geq 0$ |



# Equilibrium of a game: tangential condition

Write
$x = (x_1, \ldots, x_K) \in \Delta^{\mathcal{A}_1} \times \cdots \times \Delta^{\mathcal{A}_K}$
$\ell(x) = (\ell_1(x), \ldots, \ell_K(x))$

**Nash equilibria $\mathcal{X}^\star$**

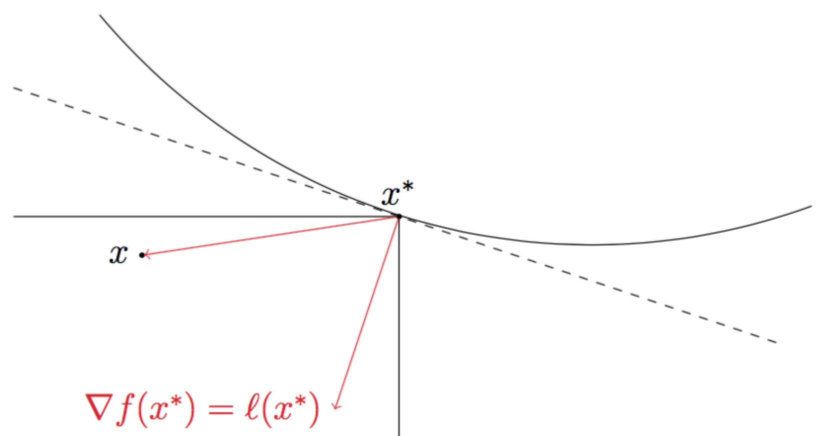
$x^\star$ is a Nash equilibrium if for all $k$, paths in the support of $x_k^\star$ have minimal loss.

$$\forall x, \ \langle \ell(x^\star), x - x^\star \rangle \geq 0$$

**Rosenthal potential**

$\exists f$ convex such that $\nabla f(x) = \ell(x)$.

| Nash condition | $\Leftrightarrow$ | first order optimality |
|---|---|---|
| $\forall x, \ \langle \ell(x^\star), x - x^\star \rangle \geq 0$ | | $\forall x, \ \langle \nabla f(x^\star), x - x^\star \rangle \geq 0$ |



# Approach 1: regret analysis

## Interpretation of the regret and the convergence

- Cumulative regret models the comparison of playing over time the best strategy possible (without changing it), and comparing it to the strategy obtained by the game.
- In the case of sublinear regret, the game converges on average towards a Nash equilibrium
- Good for optimization purposes
- Bad for operational purposes (no guarantee on what the outcome of the game is)

Cumulative regret

$$R_k^{(t)} = \sup_{x_k \in \Delta^{\mathcal{A}_k}} \sum_{\tau \leq t} \left\langle x_k^{(t)} - x_k, \ell_k(x^{(t)}) \right\rangle$$

"Online" optimality condition. Sublinear if $\limsup_t \frac{R_k^{(t)}}{t} \leq 0$.

**Convergence of averages**

$$\left[\forall k, R_k^{(t)} \text{ is sublinear}\right] \Rightarrow \bar{x}^{(t)} \to \mathcal{X}^\star$$

$\bar{x}^{(t)} = \frac{1}{t} \sum_{\tau=1}^{t} x^{(\tau)}$. proof

# Application to the routing game

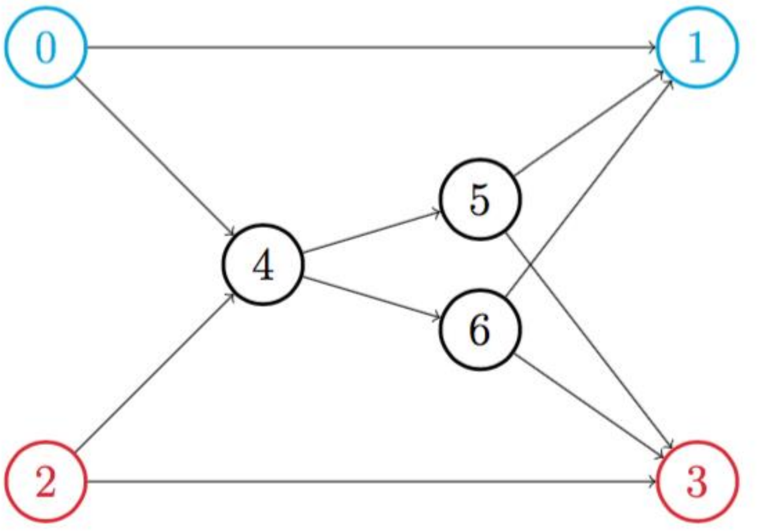
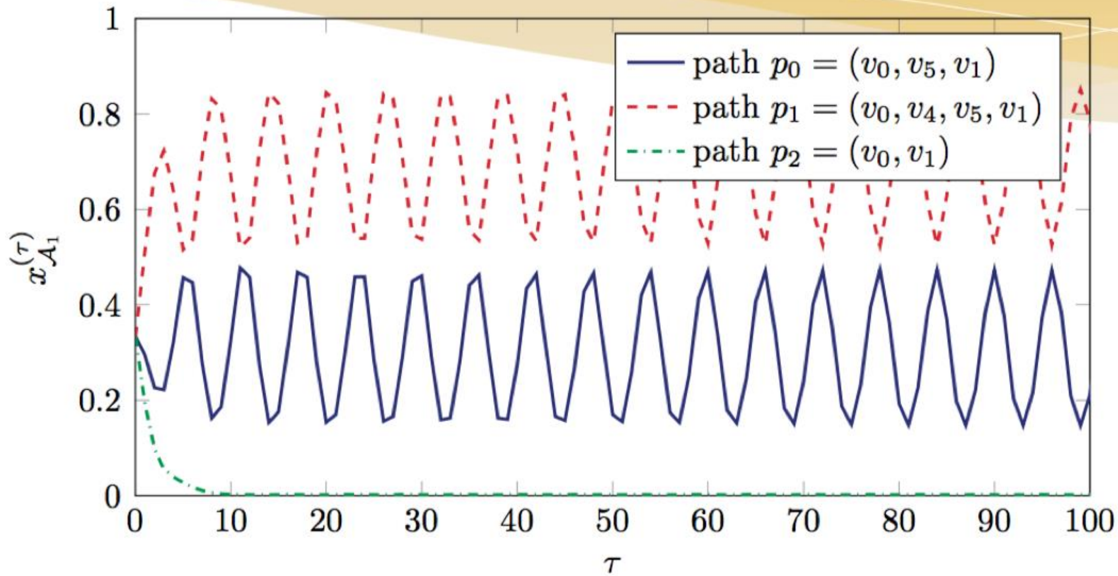



Figure: Example with strongly convex potential.

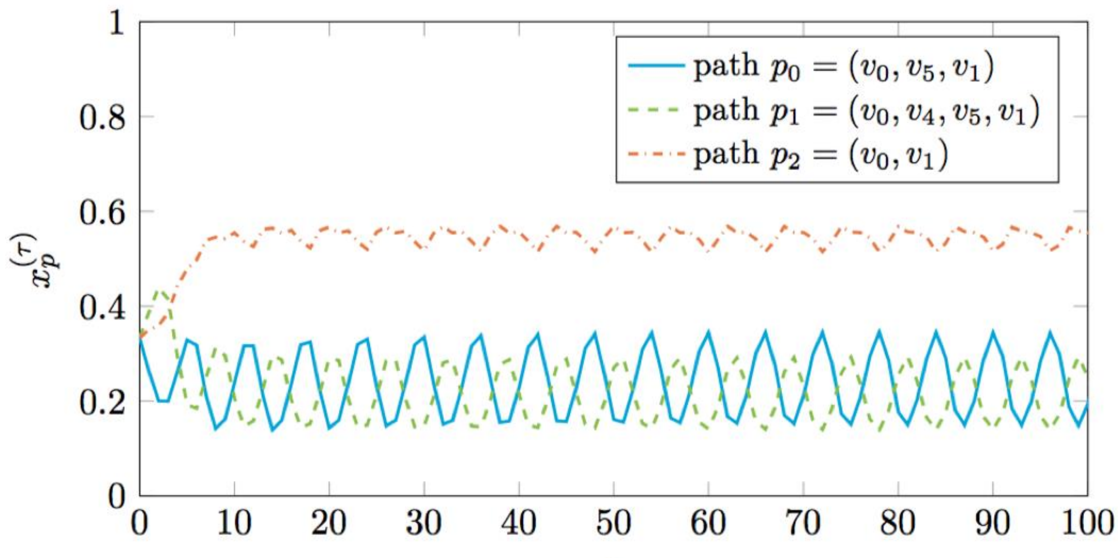
- Population 1: Hedge with $\eta_t^1 = t^{-1}$
- Population 2: Hedge with $\eta_t^2 = t^{-1}$

# Convergence on average

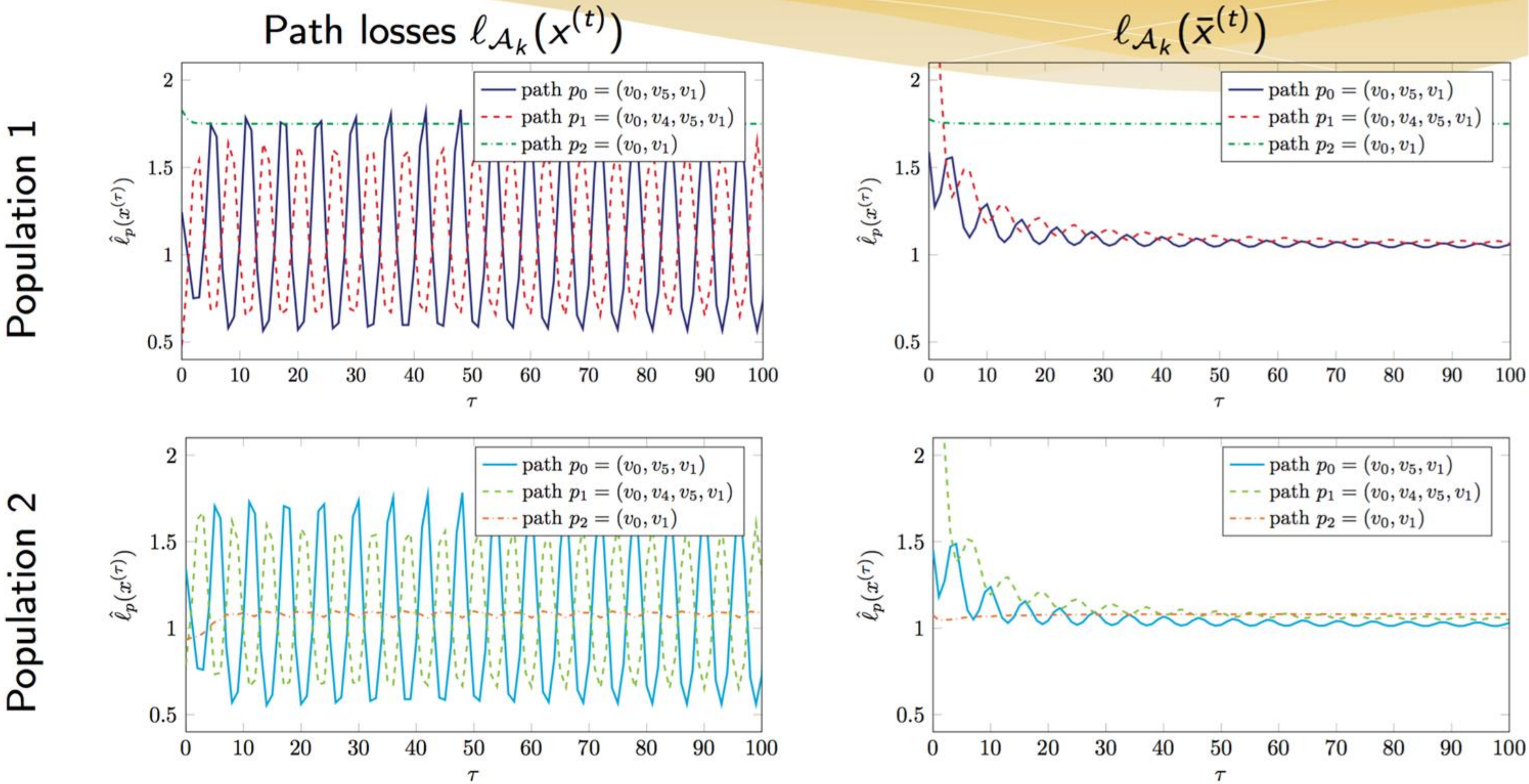
Population 1




Population 2



# Convergence on average



# Approach 2: stochastic approximation

**Idea:**

- View the learning dynamics as a discretization of an ODE
- Study the convergence of the ODE
- Relate the convergence of the discrete algorithm to the convergence of the ODE

In Hedge $x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t \ell_a^{(t)}}$, take $\eta_t \to 0$. ▸ More on Hedge

**Replicator equation [25]**

$$\forall a \in \mathcal{A}_k, \frac{dx_a}{dt} = x_a \left(\langle \ell(x), x \rangle - \ell_a(x)\right)$$

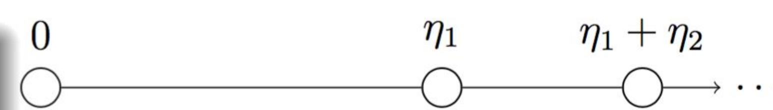



Figure: Underlying continuous time

**Definitions:**

- $\eta_t$ Discretization (in time)
- $x_a$ Distribution of flow along one arc
- $\mathcal{A}_k$ Set of arcs for population $k$

Weibull, *Evolutionary Game Theory*, 1997

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# AREP: approximate replicator dynamics

Idea:

- View the learning dynamics as a discretization of an ODE
- Study the convergence of the ODE
- Relate the convergence of the discrete algorithm to the convergence of the ODE

In Hedge $x_a^{(t+1)} \propto x_a^{(t)} e^{-\eta_t \ell_a^{(t)}}$, take $\eta_t \to 0$. ▸ More on Hedge

**Replicator equation [25]**

$$\forall a \in \mathcal{A}_k, \frac{dx_a}{dt} = x_a \left( \langle \ell(x), x \rangle - \ell_a(x) \right)$$

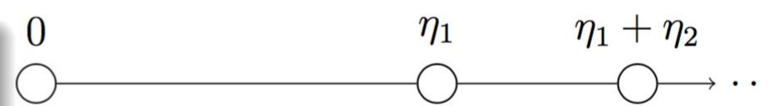



Figure: Underlying continuous time

**Discretization of the continuous-time replicator dynamics**

$$\frac{x_a^{(t+1)} - x_a^{(t)}}{\eta_t} = x_a^{(t)} \left( \left\langle \ell(x^{(t)}), x^{(t)} \right\rangle - \ell_a(x^{(t)}) \right) + U_a^{(t+1)}$$

Benaim, *Dynamics of stochastic approximation algorithms*, 1999

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# AREP: approximate replicator dynamics

Idea:

- View the learning dynamics as a discretization of an ODE
- Study the convergence of the ODE
- Relate the convergence of the discrete algorithm to the convergence of the ODE

$$\frac{dx_a}{dt} = x_a \left( \langle \ell(x), x \rangle - \ell_a(x) \right)$$

**Discretization of the continuous-time replicator dynamics**

$$\frac{x_a^{(t+1)} - x_a^{(t)}}{\eta_t} = x_a^{(t)} \left( \left\langle \ell(x^{(t)}), x^{(t)} \right\rangle - \ell_a(x^{(t)}) \right) + U_a^{(t+1)}$$

- $\eta_t$ discretization time steps.
- $(U^{(t)})_{t \geq 1}$ perturbations that satisfy for all $T > 0$,
  $\lim_{\tau_1 \to \infty} \max_{\tau_2 : \sum_{t=\tau_1}^{\tau_2} \eta_t < T} \left\| \sum_{t=\tau_1}^{\tau_2} \eta_t U^{(t+1)} \right\| = 0$

(a sufficient condition is that $\exists q \geq 2$: $\sup_\tau \mathbb{E} \| U^{(\tau)} \|^q < \infty$ and $\sum_\tau \eta_\tau^{1+\frac{q}{2}} < \infty$)

# AREP: approximate replicator dynamics

Idea:

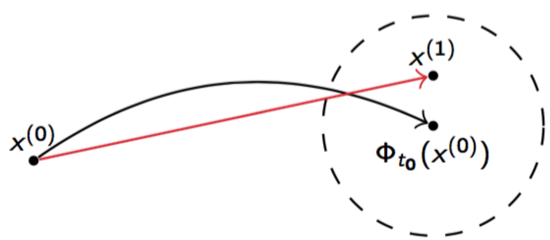
- View the learning dynamics as a discretization of an ODE
- Study the convergence of the ODE
- Relate the convergence of the discrete algorithm to the convergence of the ODE, but no convergence rates

**Theorem [13]**

In convex potential games, under AREP updates, if $\eta_t \downarrow 0$ and $\sum \eta_t = \infty$, then

$$x^{(t)} \to \mathcal{X}^\star \text{ a.s.}$$

- Affine interpolation of $x^{(t)}$ is an asymptotic pseudo trajectory of ODE.




- Use $f$ as a Lyapunov function. proof details

# AREP: approximate replicator dynamics

Idea:

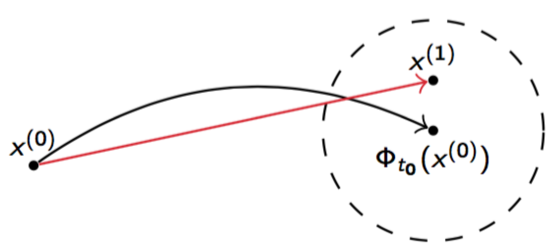
- View the learning dynamics as a discretization of an ODE
- Study the convergence of the ODE
- Relate the convergence of the discrete algorithm to the convergence of the ODE, but no convergence rates

**Theorem [13]**

In convex potential games, under AREP updates, if $\eta_t \downarrow 0$ and $\sum \eta_t = \infty$, then

$$x^{(t)} \to \mathcal{X}^\star \text{ a.s.}$$

- Affine interpolation of $x^{(t)}$ is an asymptotic pseudo trajectory of ODE.




- Use $f$ as a Lyapunov function. proof details

# AREP: approximate replicator dynamics

Idea:

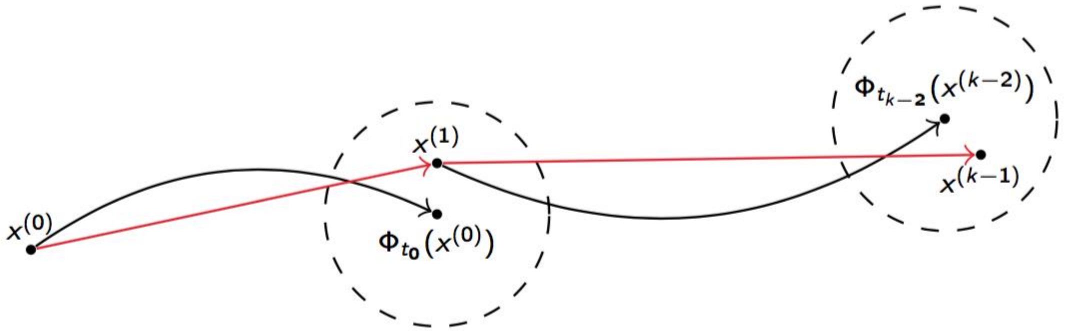
- View the learning dynamics as a discretization of an ODE
- Study the convergence of the ODE
- Relate the convergence of the discrete algorithm to the convergence of the ODE, but no convergence rates

**Theorem [13]**

In convex potential games, under AREP updates, if $\eta_t \downarrow 0$ and $\sum \eta_t = \infty$, then

$$x^{(t)} \to \mathcal{X}^\star \text{ a.s.}$$

- Affine interpolation of $x^{(t)}$ is an asymptotic pseudo trajectory of ODE.




- Use $f$ as a Lyapunov function. ▸ proof details

# AREP: approximate replicator dynamics

Idea:

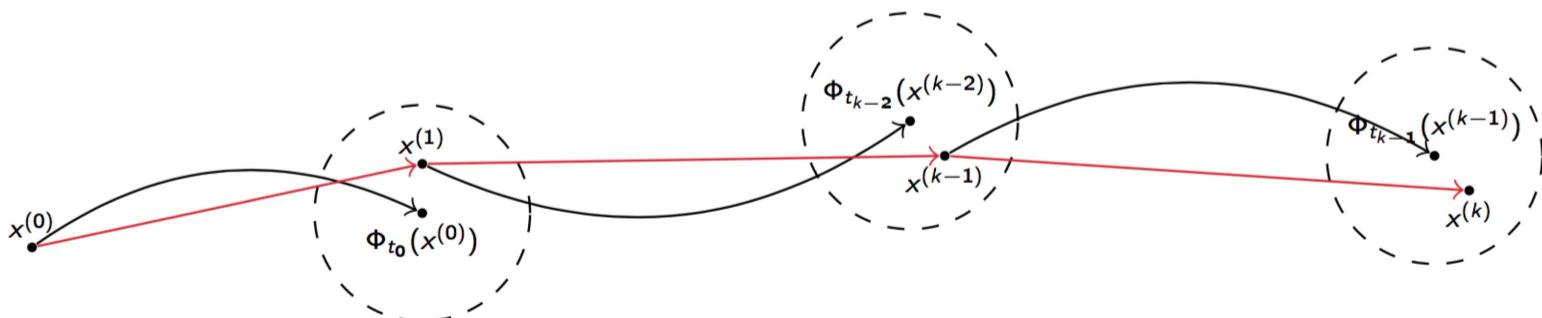
- View the learning dynamics as a discretization of an ODE
- Study the convergence of the ODE
- Relate the convergence of the discrete algorithm to the convergence of the ODE, but no convergence rates

**Theorem [13]**

In convex potential games, under AREP updates, if $\eta_t \downarrow 0$ and $\sum \eta_t = \infty$, then

$$x^{(t)} \to \mathcal{X}^\star \text{ a.s.}$$

- Affine interpolation of $x^{(t)}$ is an asymptotic pseudo trajectory of ODE.

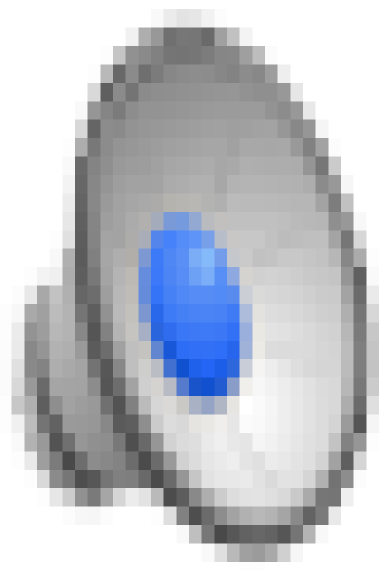



- Use $f$ as a Lyapunov function. proof details

# AREP: approximate replicator dynamics: illustration

Blue: discretized AREP

Red: continuous trajectory of replicator dynamics



# Approach 3: convex optimization

Idea:

- View the learning dynamics as a distributed algorithm to minimize the function $f$.
- Allows us to analyze convergence rates.

Here:

- Class of distributed optimization methods: stochastic mirror descent

| | | |
|---|---|---|
| minimize | $f(x)$ | convex function |
| subject to | $x \in \mathcal{X} \subset \mathbb{R}^d$ | convex, compact set |

**Bregman Divergence**

Strongly convex function $\psi$

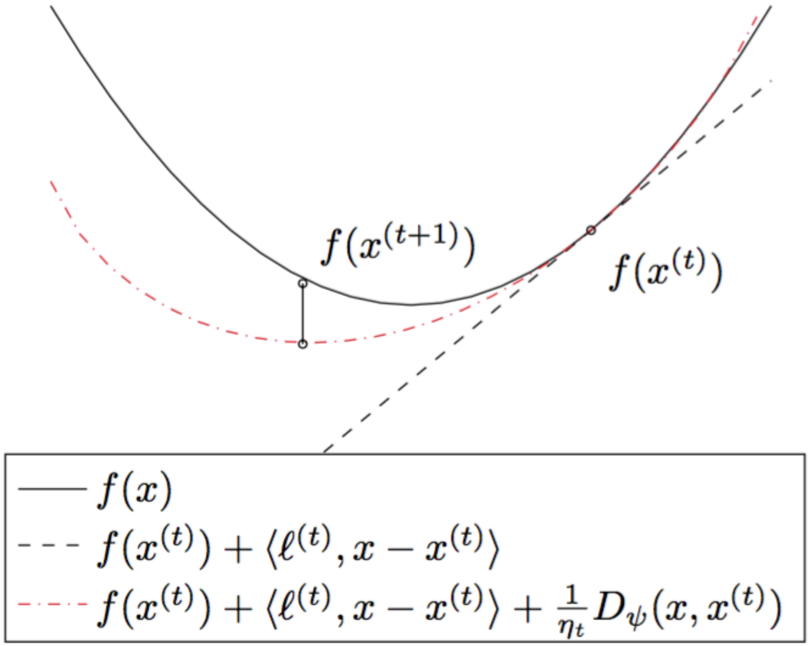
$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla \psi(y), x - y \rangle$$

# Approach 3: convex optimization

| | | |
|---|---|---|
| minimize | $f(x)$ | convex function |
| subject to | $x \in \mathcal{X} \subset \mathbb{R}^d$ | convex, compact set |

**Algorithm 2** MD Method with learning rates $(\eta_t)$

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




- $\eta_t$: learning rate
- $D_\psi$: ▸ Bregman divergence

**Bregman Divergence**

Strongly convex function $\psi$

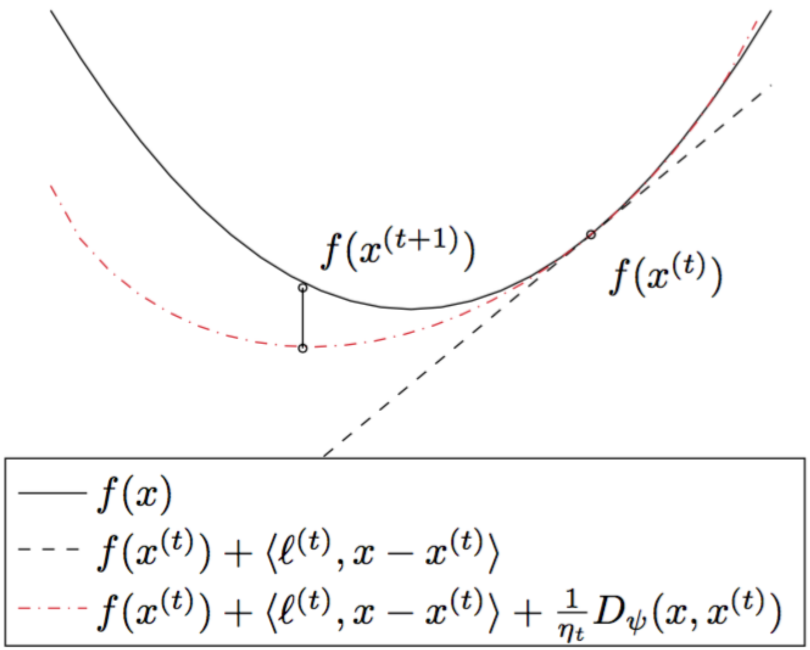
$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla \psi(y), x - y \rangle$$

# Approach 3: convex optimization

| | | |
|---|---|---|
| minimize | $f(x)$ | convex function |
| subject to | $x \in \mathcal{X} \subset \mathbb{R}^d$ | convex, compact set |

**Algorithm 2** MD Method with learning rates $(\eta_t)$

**for** $t \in \mathbb{N}$ **do**
  observe $\ell_k^{(t)} \in \partial_k f(x^{(t)})$
  $x_k^{(t+1)} = \arg\min_{x \in \mathcal{X}_k} \left\langle \ell_k^{(t)}, x \right\rangle + \frac{1}{\eta_t^k} D_{\psi_k}(x, x_k^{(t)})$
**end for**




- $\eta_t$: learning rate
- $D_\psi$: ▸ Bregman divergence

**Bregman Divergence**

Strongly convex function $\psi$

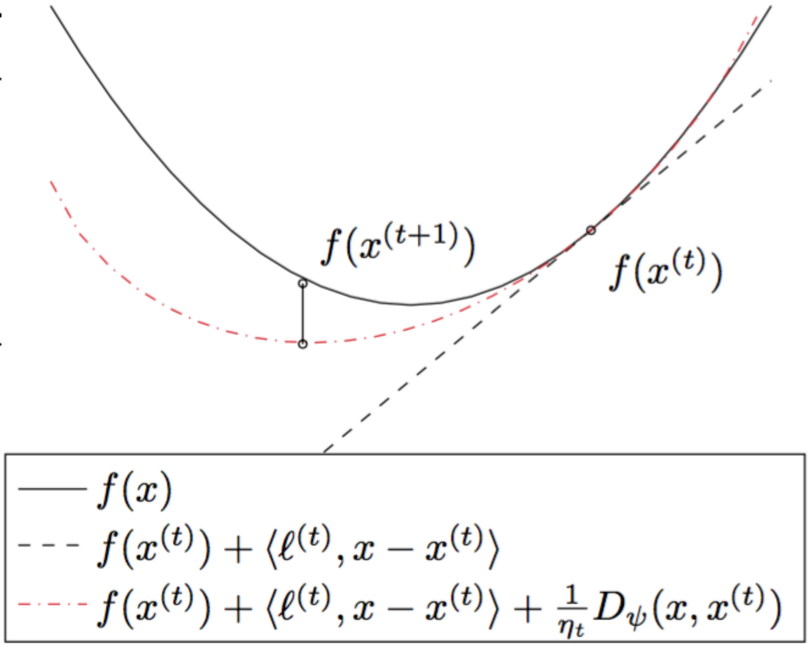
$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y \rangle$$

# Approach 3: convex optimization

| | | |
|---|---|---|
| minimize | $f(x)$ | convex function |
| subject to | $x \in \mathcal{X} \subset \mathbb{R}^d$ | convex, compact set |

**Algorithm 2** SMD Method with learning rates $(\eta_t)$

1: **for** $t \in \mathbb{N}$ **do**
2: observe $\hat{\ell}_k^{(t)}$ with $\mathbb{E}\left[\hat{\ell}_k^{(t)} | \mathcal{F}_{t-1}\right] \in \partial_k f(x^{(t)})$
3: $x_k^{(t+1)} = \arg\min_{x \in \mathcal{X}_k} \left\langle \hat{\ell}_k^{(t)}, x \right\rangle + \frac{1}{\eta_t^k} D_{\psi_k}(x, x_k^{(t)})$
4: **end for**




- $\eta_t$: learning rate
- $D_\psi$: Bregman divergence

**Bregman Divergence**

Strongly convex function $\psi$

$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y \rangle$$

# Convergence

$d_\tau = D_\psi(\mathcal{X}^\star, x^{(\tau)})$.

**Main ingredient**

$$\mathbb{E}\left[d_{\tau+1} | \mathcal{F}_{\tau-1}\right] \le d_\tau - \eta_\tau(f(x^{(\tau)}) - f^\star) + \frac{\eta_\tau^2}{2\mu}\mathbb{E}\left[\|\hat{\ell}^{(\tau)}\|_*^2 | \mathcal{F}_{\tau-1}\right]$$

From here,

- Can show a.s. convergence $x^{(t)} \to \mathcal{X}^\star$ if $\sum \eta_t = \infty$ and $\sum \eta_t^2 < \infty$
  $d_\tau$ is an almost super martingale [19], [5]

  Deterministic version:
  If $d_{\tau+1} \le d_\tau - a_\tau + b_\tau$, and $\sum b_\tau < \infty$, then $(d_\tau)$ converges.

---

[19]H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications.
*Optimizing Methods in Statistics*, 1971

[5]Léon Bottou. Online algorithms and stochastic approximations.
1998

# Convergence

- To show convergence $\mathbb{E}\left[f(x^{(t)})\right] \to f^\star$, generalize the technique of Shamir et al. [22].

**Convergence of Distributed Stochastic Mirror Descent**

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

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

Non-smooth, non-strongly convex.

More details

[22]Ohad Shamir and Tong Zhang. Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes.
In *ICML*, pages 71–79, 2013

[12] 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

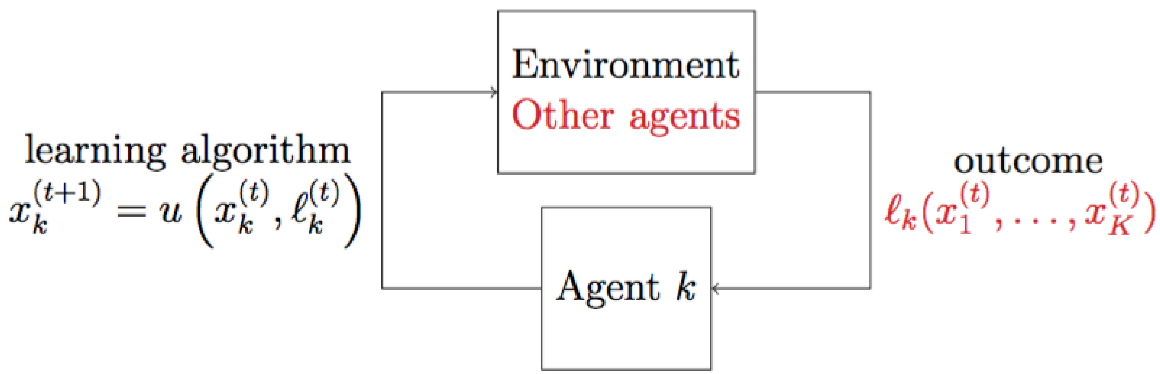
# Summary

Distributed learning dynamics in routing games

– Each player routes population $k$ according to distribution (corresponding to one OD pair) $p \sim x_k^{(t)}$

– At each iteration, the population $k$ discovers their outcome $\ell_k^{(t)}$

– The routing of population $k$ at the next step is subsequently updated according to the following law $x_k^{(t+1)} = u_k\left(x_k^{(t)}, \ell_k^{(t)}\right)$




- Regret analysis: convergence of $\bar{x}^{(t)}$
- Stochastic approximation: almost sure convergence of $x^{(t)}$
- Stochastic convex optimization: almost sure convergence, $\mathbb{E}\left[f(x^{(t)})\right] \to f^\star$, $\mathbb{E}\left[D_\psi(x^\star, x^{(t)})\right] \to 0$, convergence rates.

# Application to the routing game

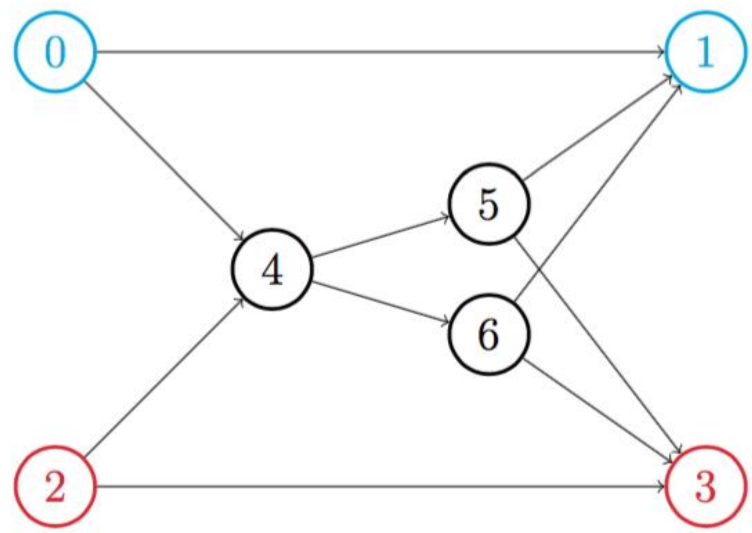
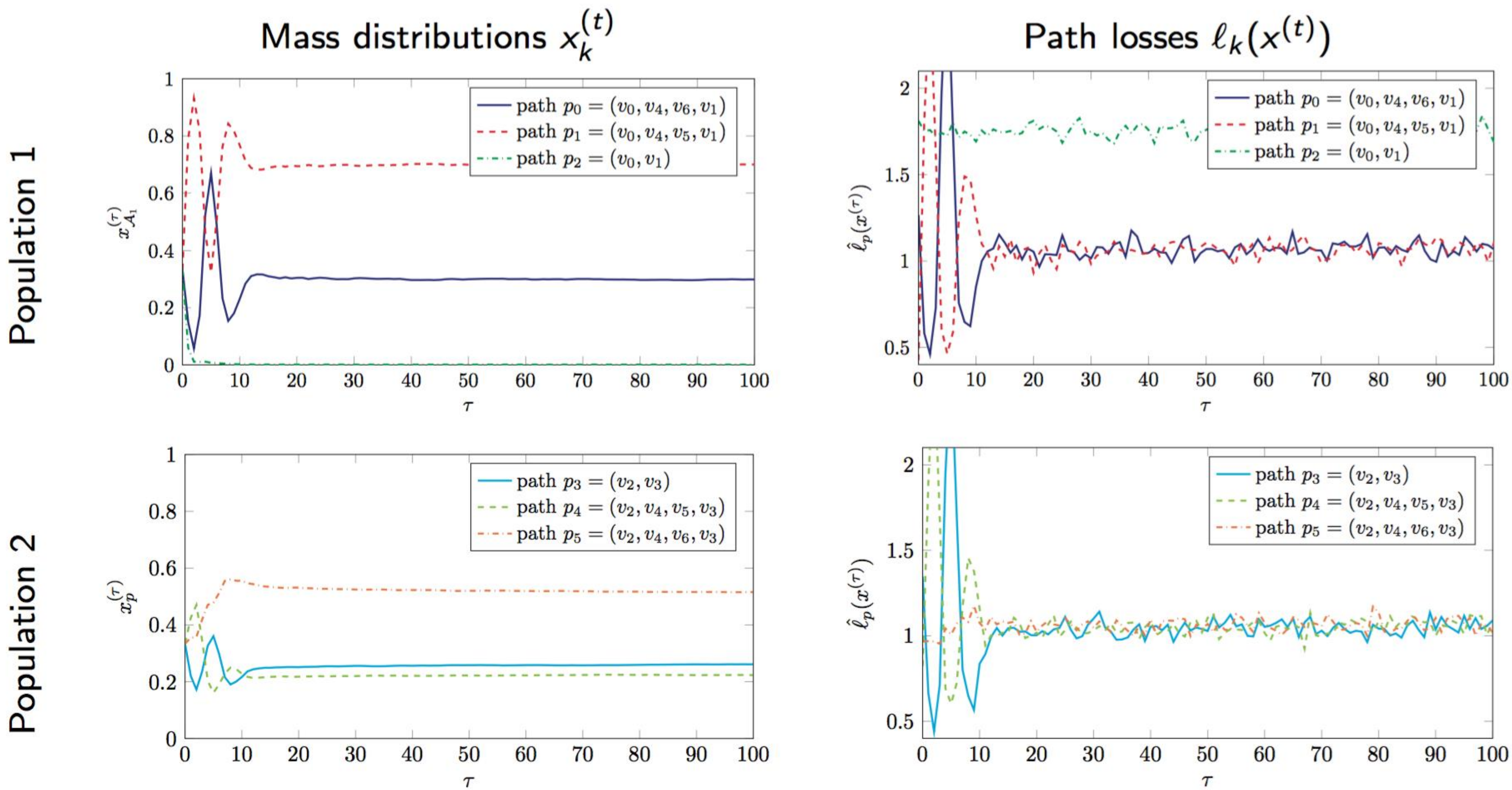



Figure: Example with strongly convex potential.

- Population 1: Hedge with $\eta_t^1 = t^{-1}$
- Population 2: Hedge with $\eta_t^2 = t^{-1}$

# Routing game with strongly convex potential



# Routing game with strongly convex potential

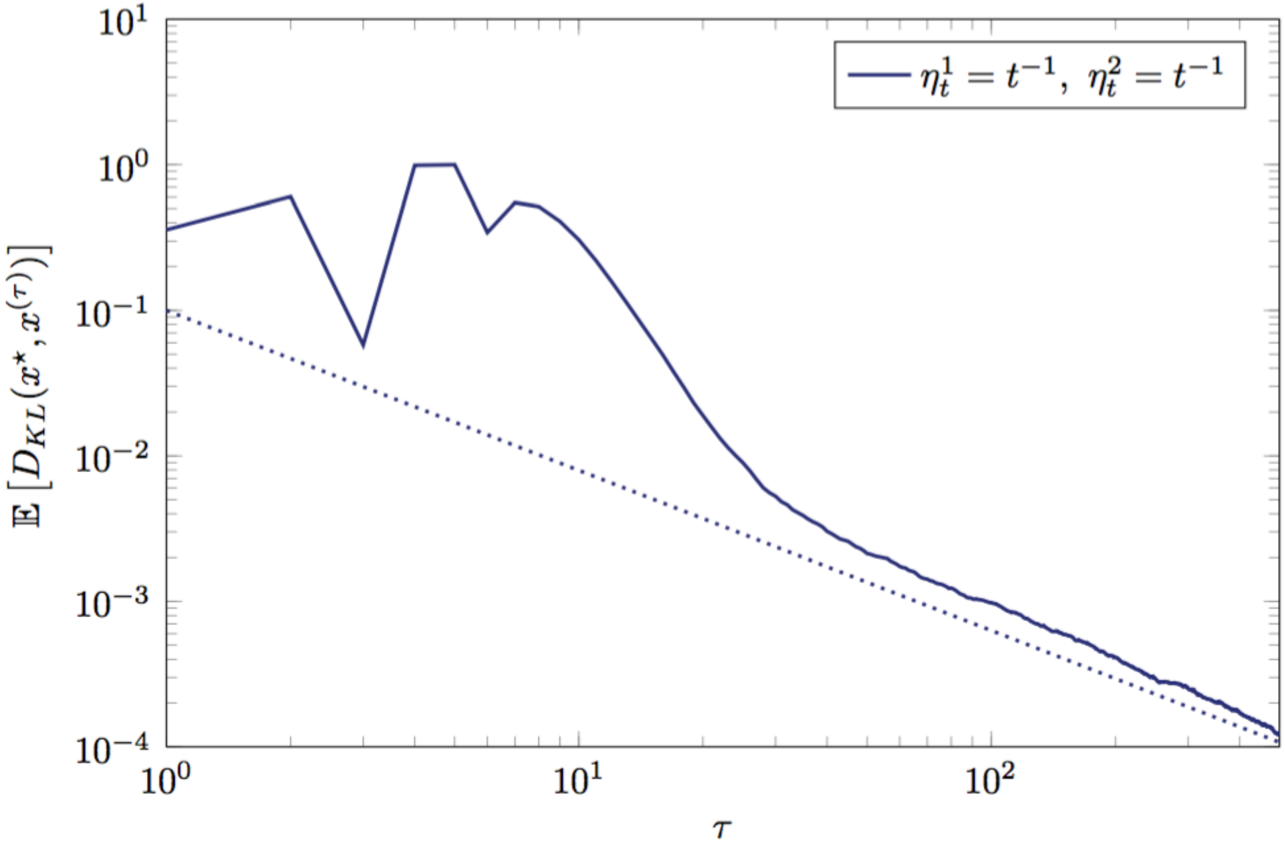


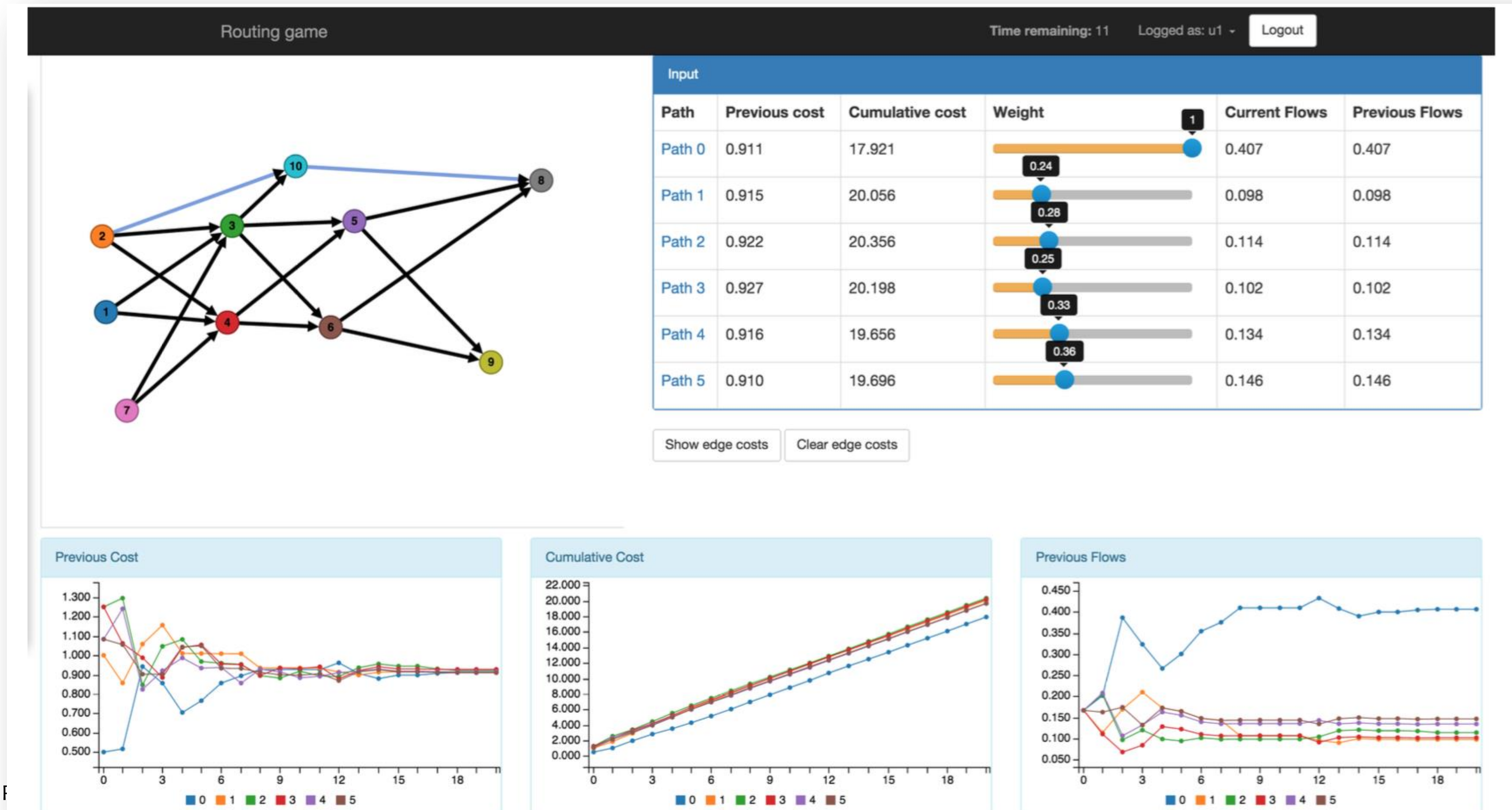

Figure: Distance to equilibrium.

For $\eta_t^k = \frac{\theta_k}{\ell_f t^{\alpha_k}}$, $\alpha_k \in (0, 1]$, $\mathbb{E}\left[D_\psi(x^\star, x^{(t)})\right] = O(\sum_k t^{-\alpha_k})$

# Practical game implementation: field experiment

Idea of the game: study non-cooperative behavior of routing applications "managers"

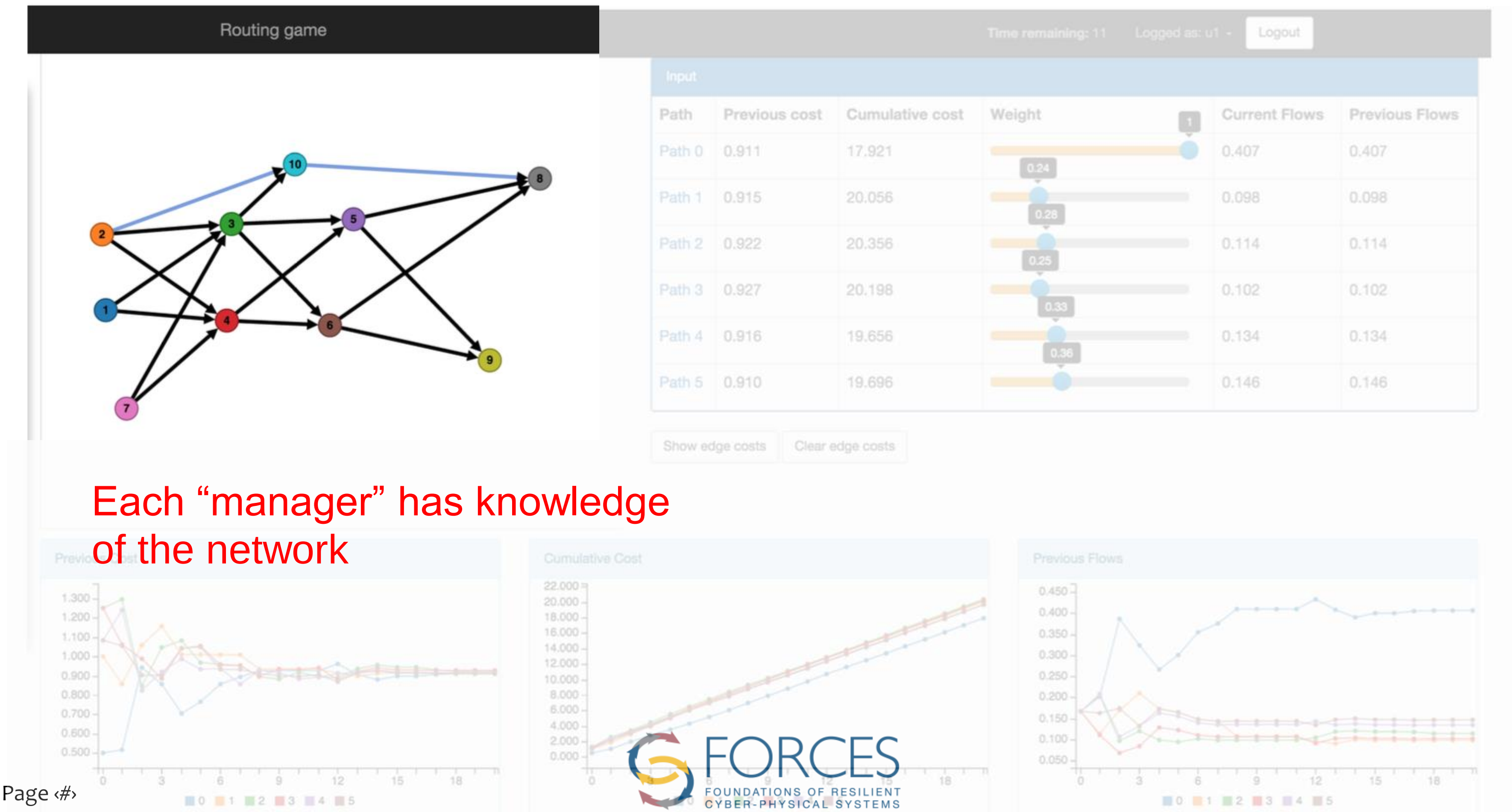
- As if Google was "playing against" Apple, INRIX etc.
- Study evolution of distribution over successive iterations



# Practical game implementation: field experiment

Idea of the game: study non-cooperative behavior of routing applications "managers"

- As if Google was "playing against" Apple, INRIX etc.
- Study evolution of distribution over successive iterations

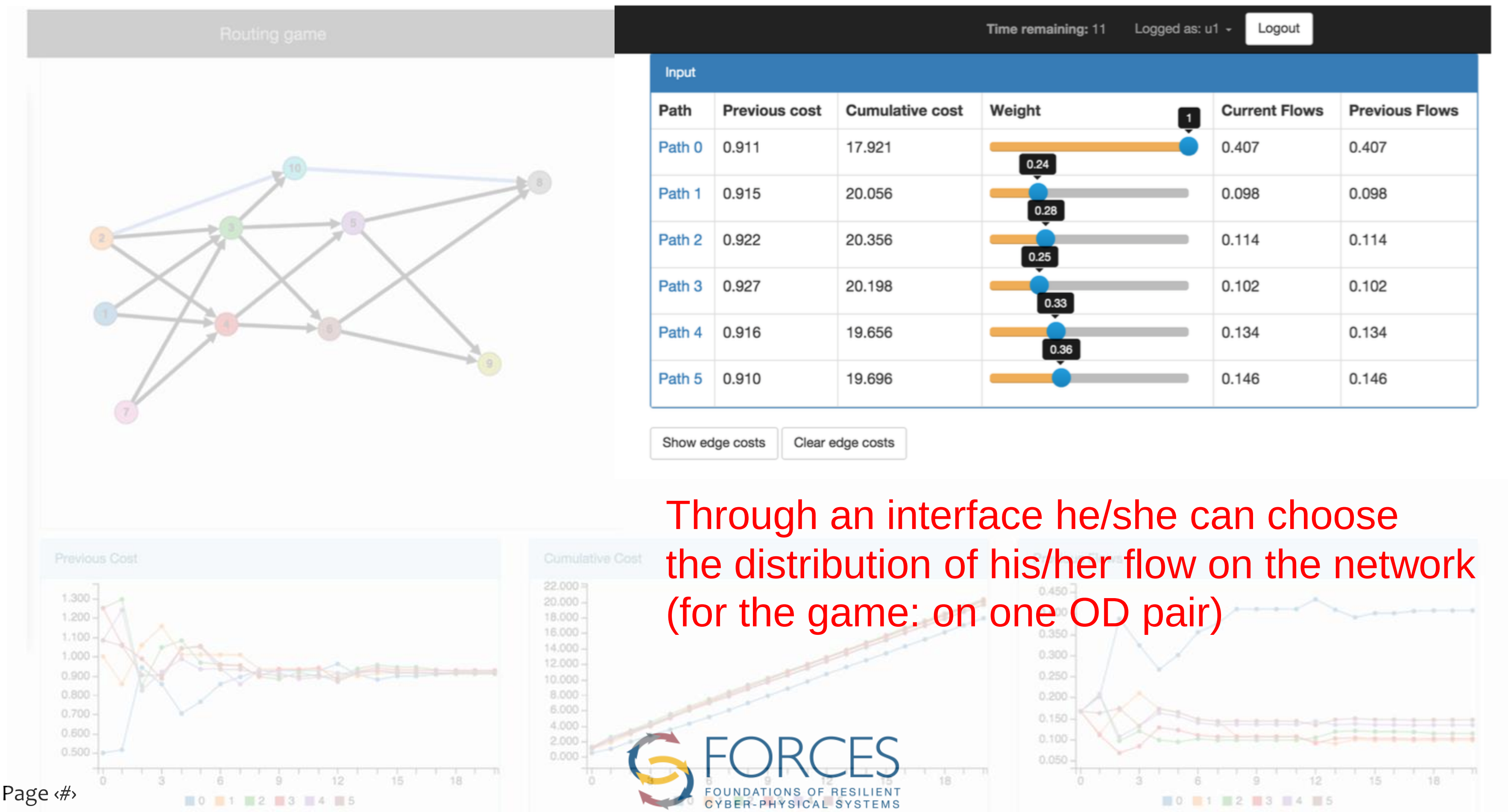



Each "manager" has knowledge of the network

# Practical game implementation: field experiment

Idea of the game: study non-cooperative behavior of routing applications "managers"

- As if Google was "playing against" Apple, INRIX etc.
- Study evolution of distribution over successive iterations



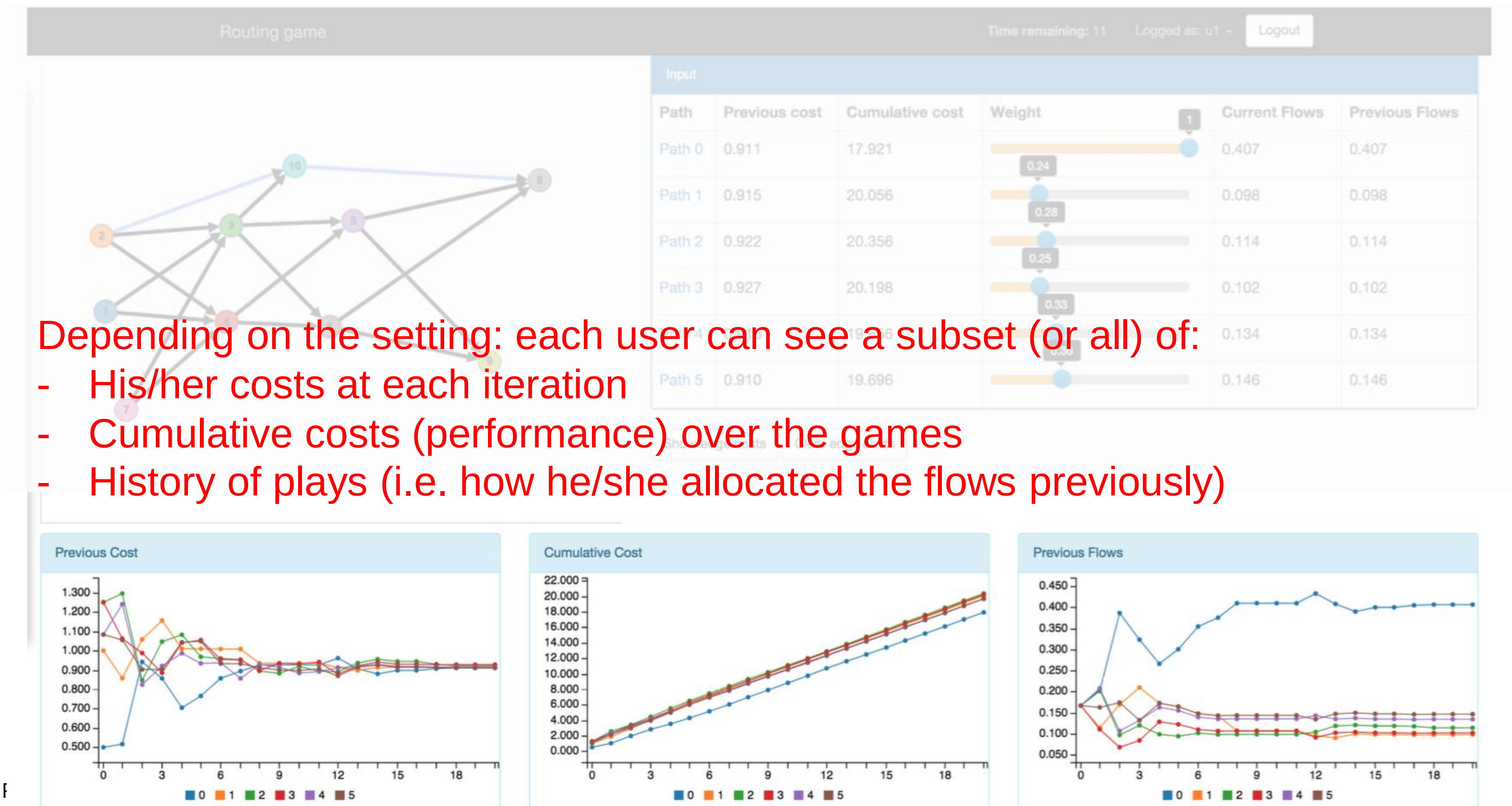

Through an interface he/she can choose the distribution of his/her flow on the network (for the game: on one OD pair)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# Practical game implementation: field experiment

Idea of the game: study non-cooperative behavior of routing applications "managers"

- As if Google was "playing against" Apple, INRIX etc.
- Study evolution of distribution over successive iterations



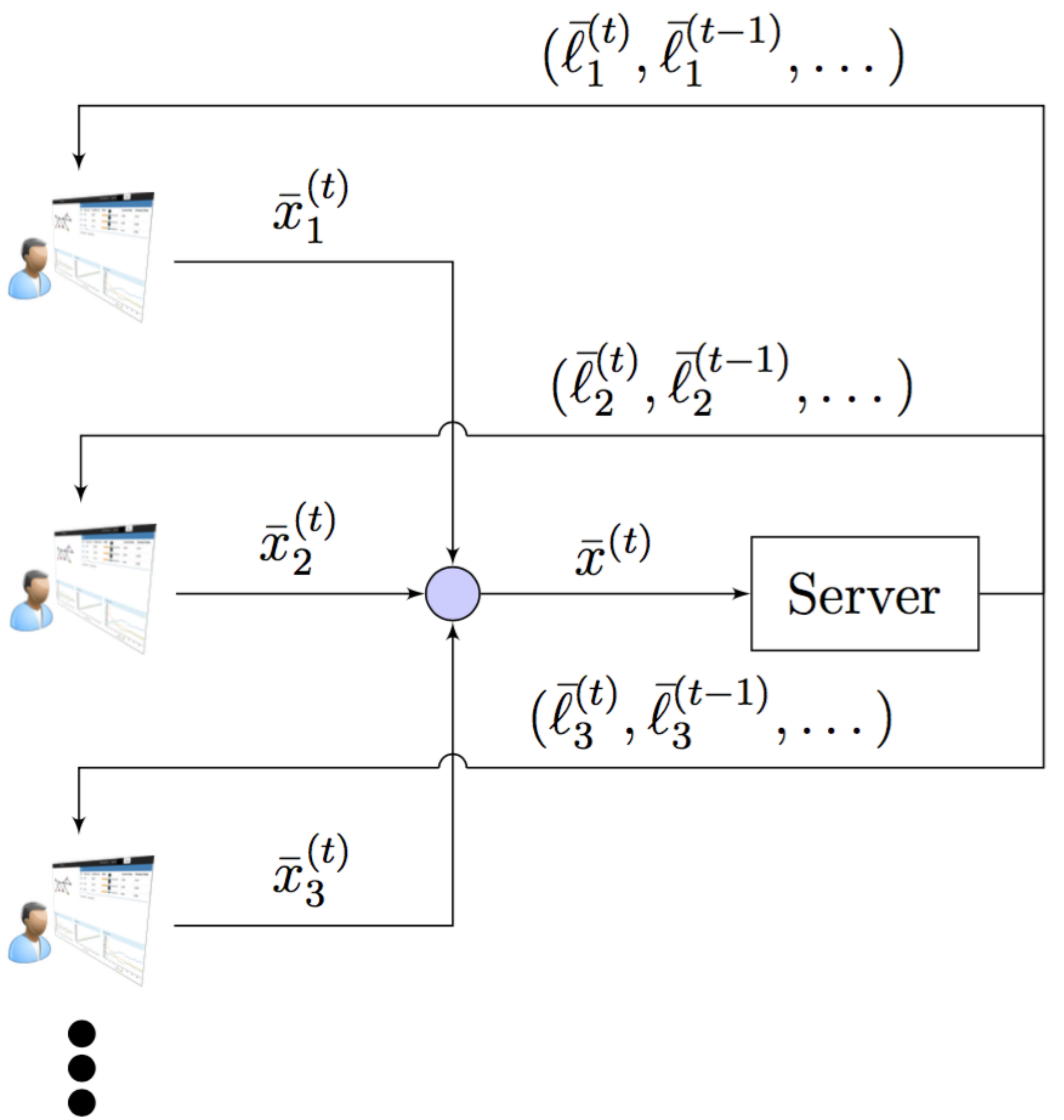

Depending on the setting: each user can see a subset (or all) of:

- His/her costs at each iteration
- Cumulative costs (performance) over the games
- History of plays (i.e. how he/she allocated the flows previously)

# Game process

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# Game process



# Learning how players learn

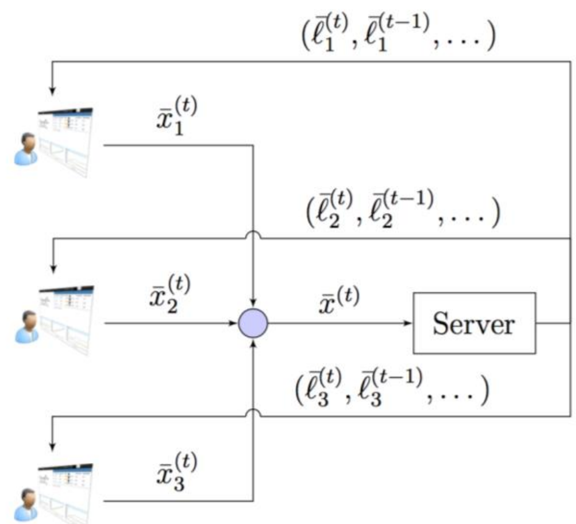
- We observe a sequence of player decisions ($\bar{x}^{(t)}$) and losses ($\bar{\ell}^{(t)}$).
- Can we fit a model of player dynamics?

## Mirror descent model

Estimate the learning rate in the mirror descent model

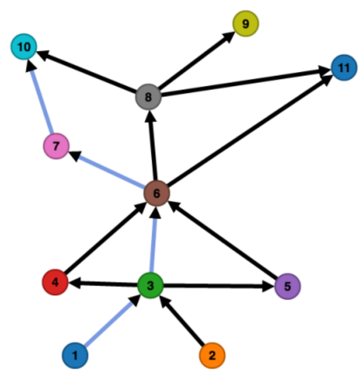
$$x^{(t+1)}(\eta) = \arg\min_{x \in \Delta^{\mathcal{A}_k}} \left\langle \bar{\ell}^{(t)}, x \right\rangle + \frac{1}{\eta} D_{KL}(x, \bar{x}^{(t)})$$

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



# Interface for one player

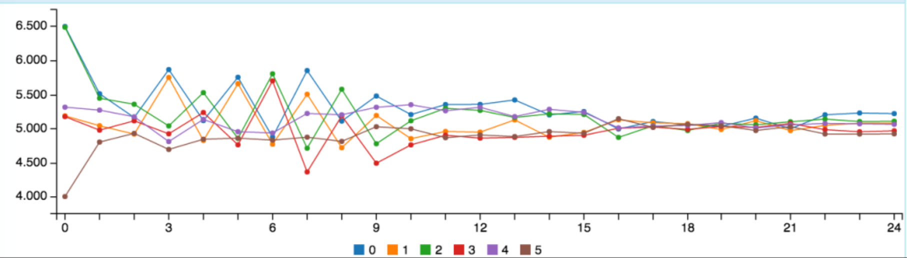
Routing game | Time remaining: 0 | Turn left: 25 | Logged as: link | Logout



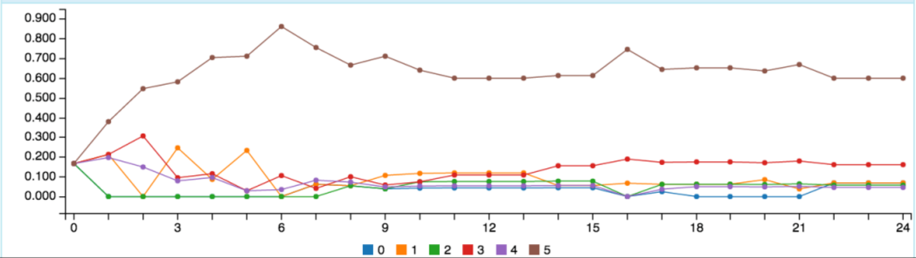

| Path | Path's previous cost | Weight | Current Flows | Previous Flows |
|---|---|---|---|---|
| Path 0 | 5.217 | 0.06 | 0.069 | 0.069 |
| Path 1 | 5.077 | 0.06 | 0.069 | 0.069 |
| Path 2 | 5.106 | 0.05 | 0.057 | 0.057 |
| Path 3 | 4.966 | 0.14 | 0.161 | 0.161 |
| Path 4 | 5.061 | 0.04 | 0.046 | 0.046 |
| Path 5 | 4.921 | 0.52 | 0.598 | 0.598 |

Show edge costs | Clear edge costs

Previous Cost




Previous Flows

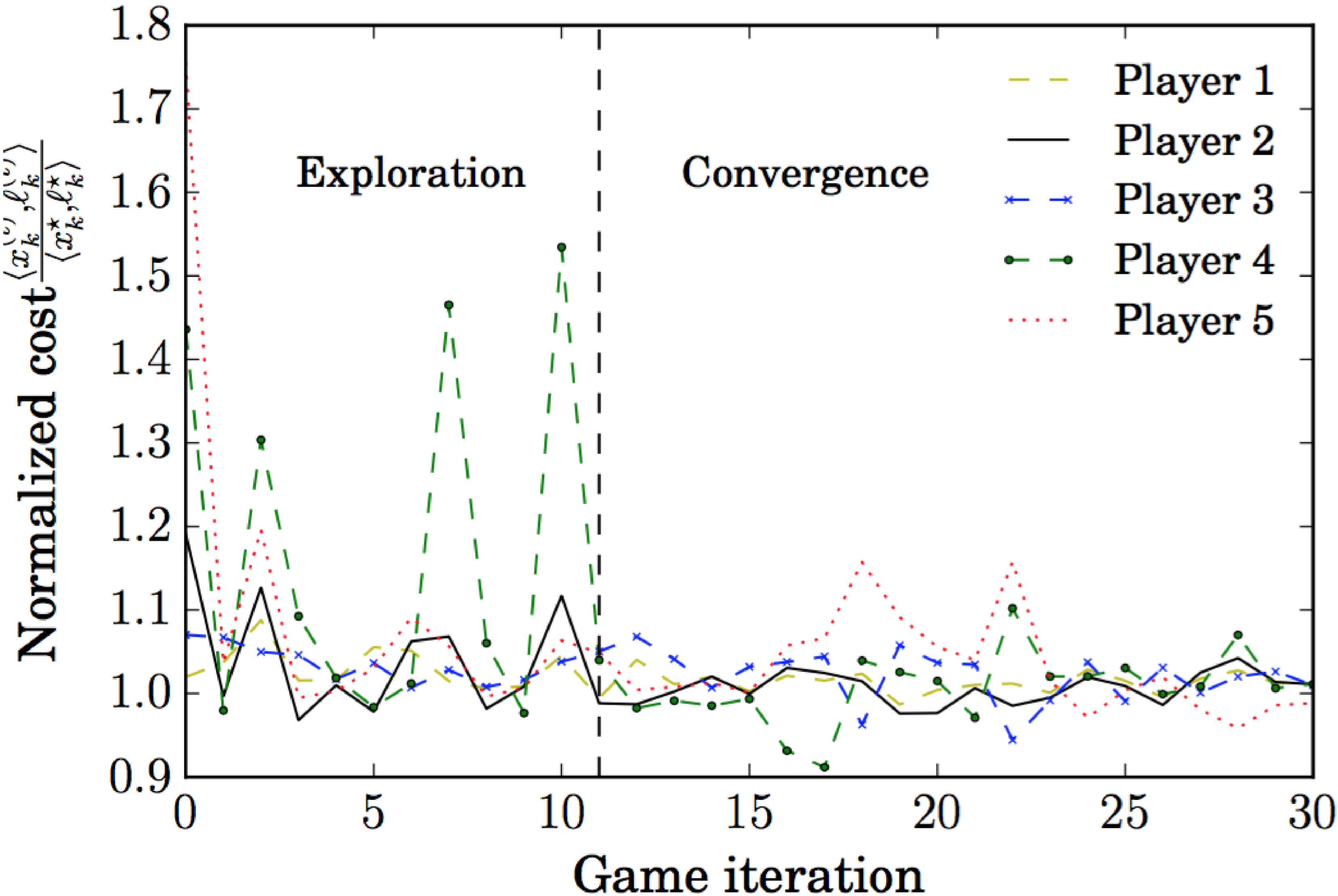



Want to work on this HIT?

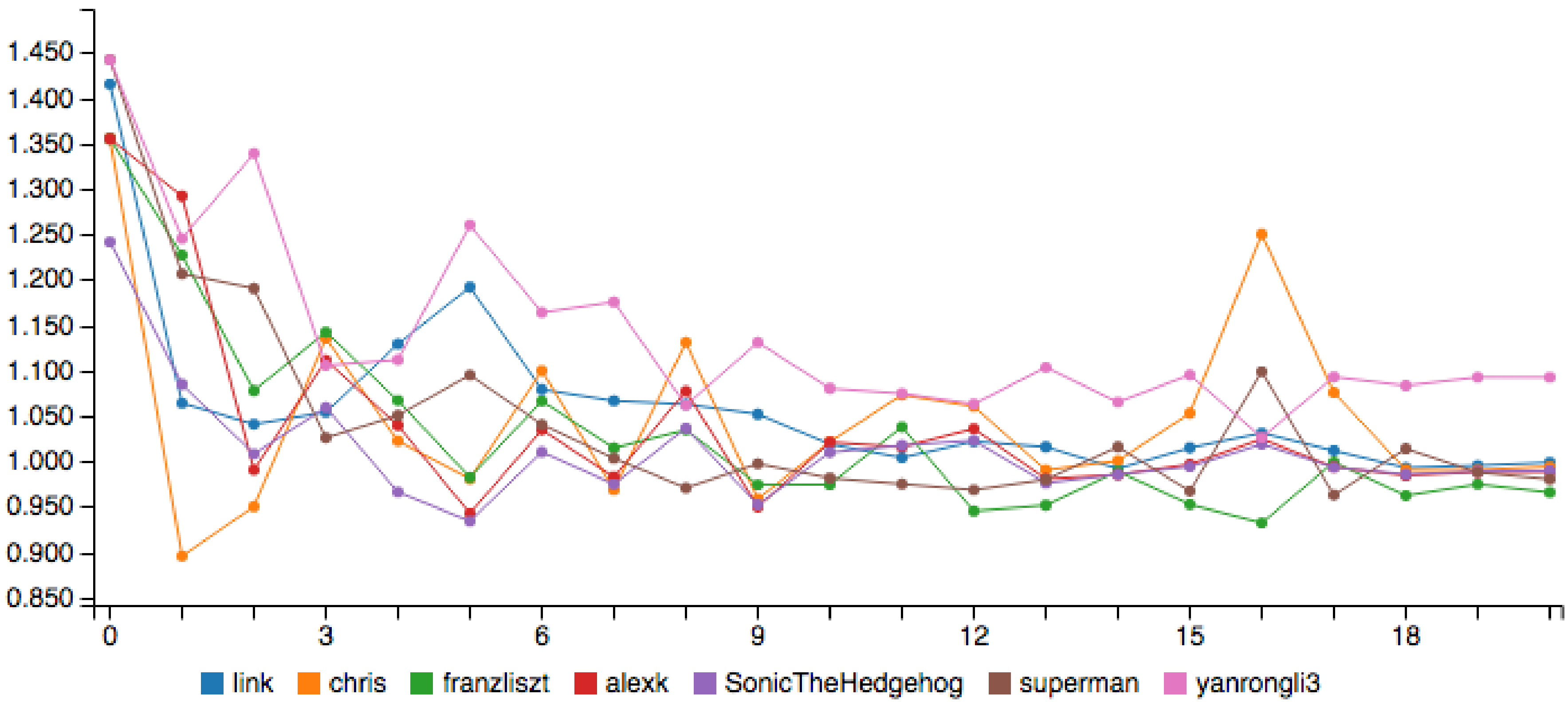
Accept HIT

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

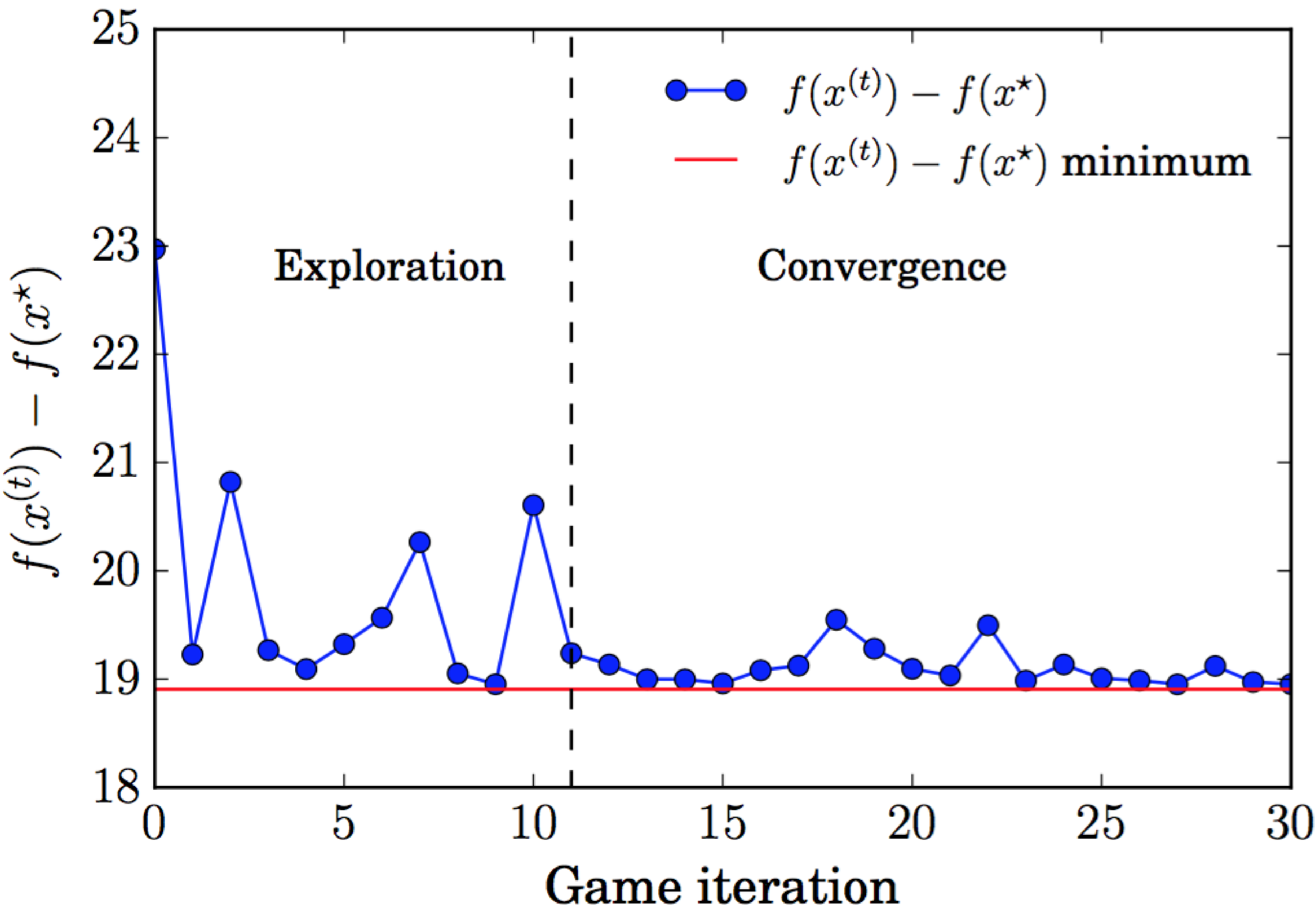
# Cost of each player (normalized by eq. cost)



# Cost of each player (normalized by eq. cost)



# Value of potential function

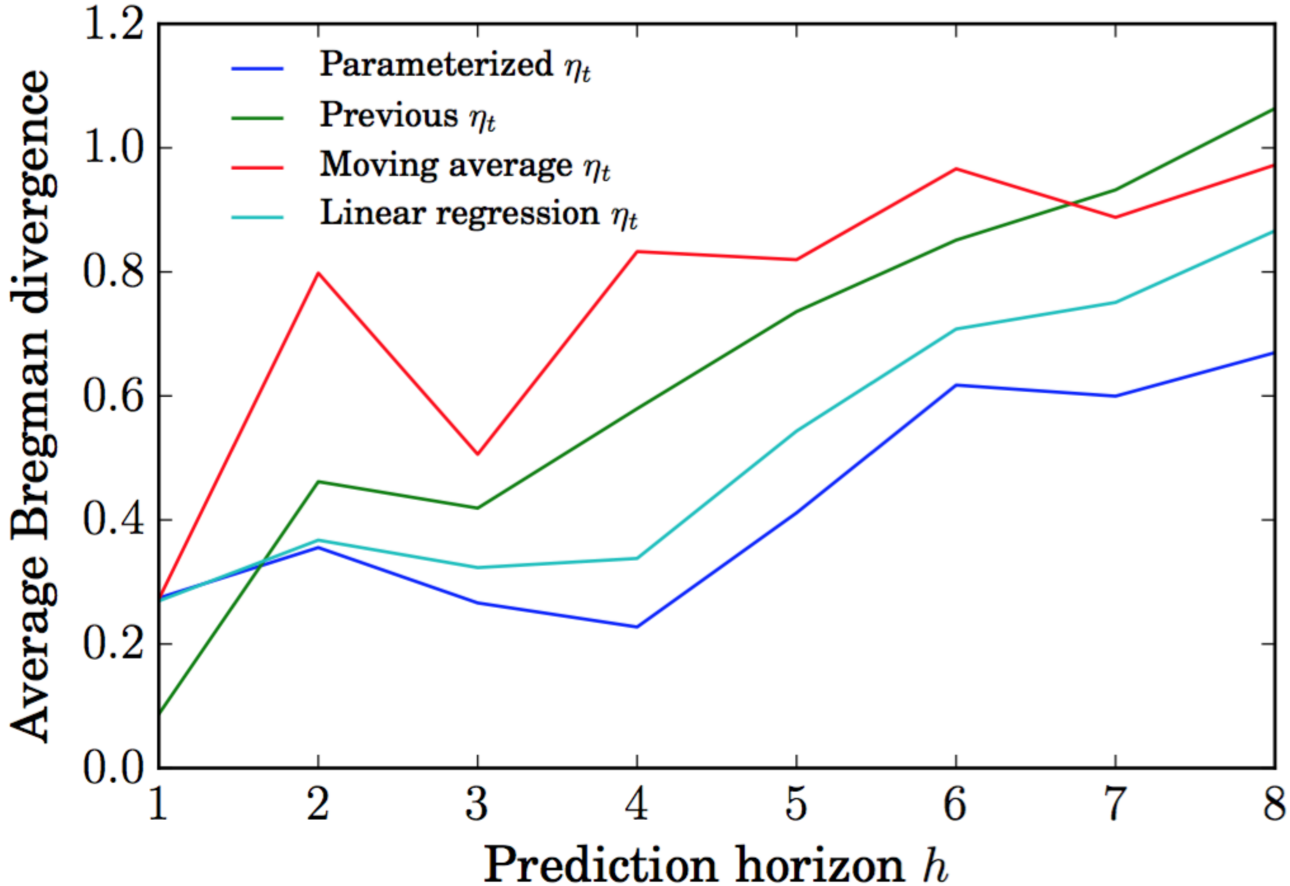


# Average of KL divergence

Average KL divergence between

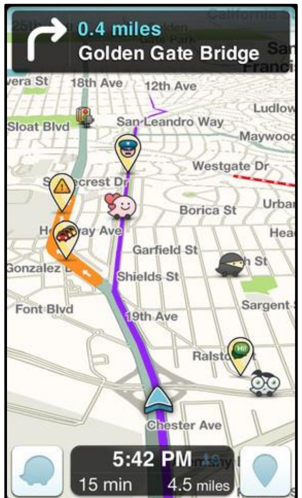
- Predicted distributions
- Actual distributions

As a function of the prediction horizon *h*

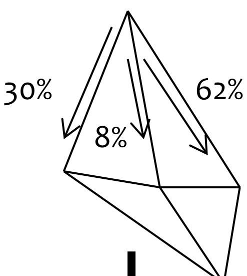


# Back to Coupled sequential decision problem

Waze

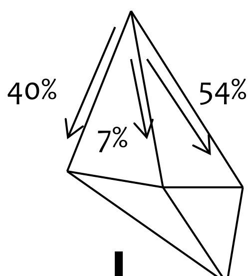



$p_{\text{Waze}} \sim x^{(t)}_{\text{Waze}}$

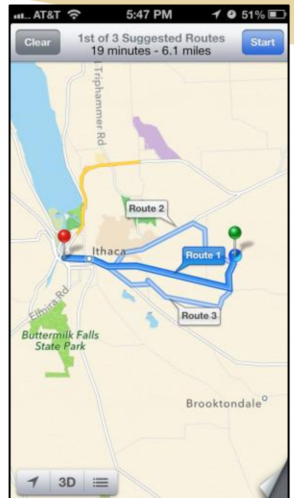



Google

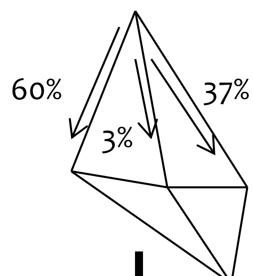

$p_{\text{Google}} \sim x^{(t)}_{\text{Google}}$



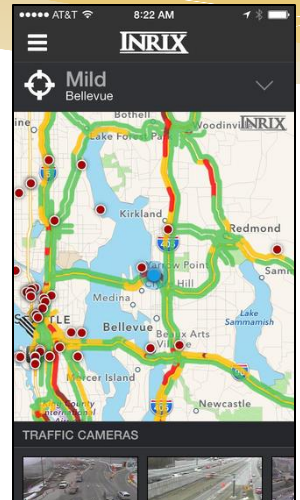

Apple



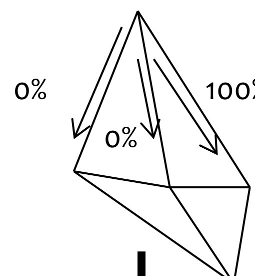

$p_{\text{Apple}} \sim x^{(t)}_{\text{Apple}}$



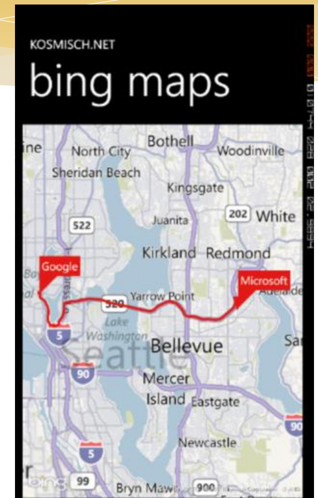

INRIX



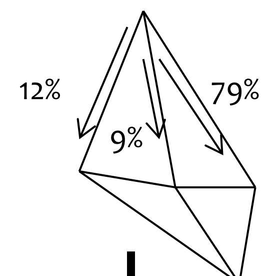

$p_{\text{INRIX}} \sim x^{(t)}_{\text{INRIX}}$




Bing (Microsoft)

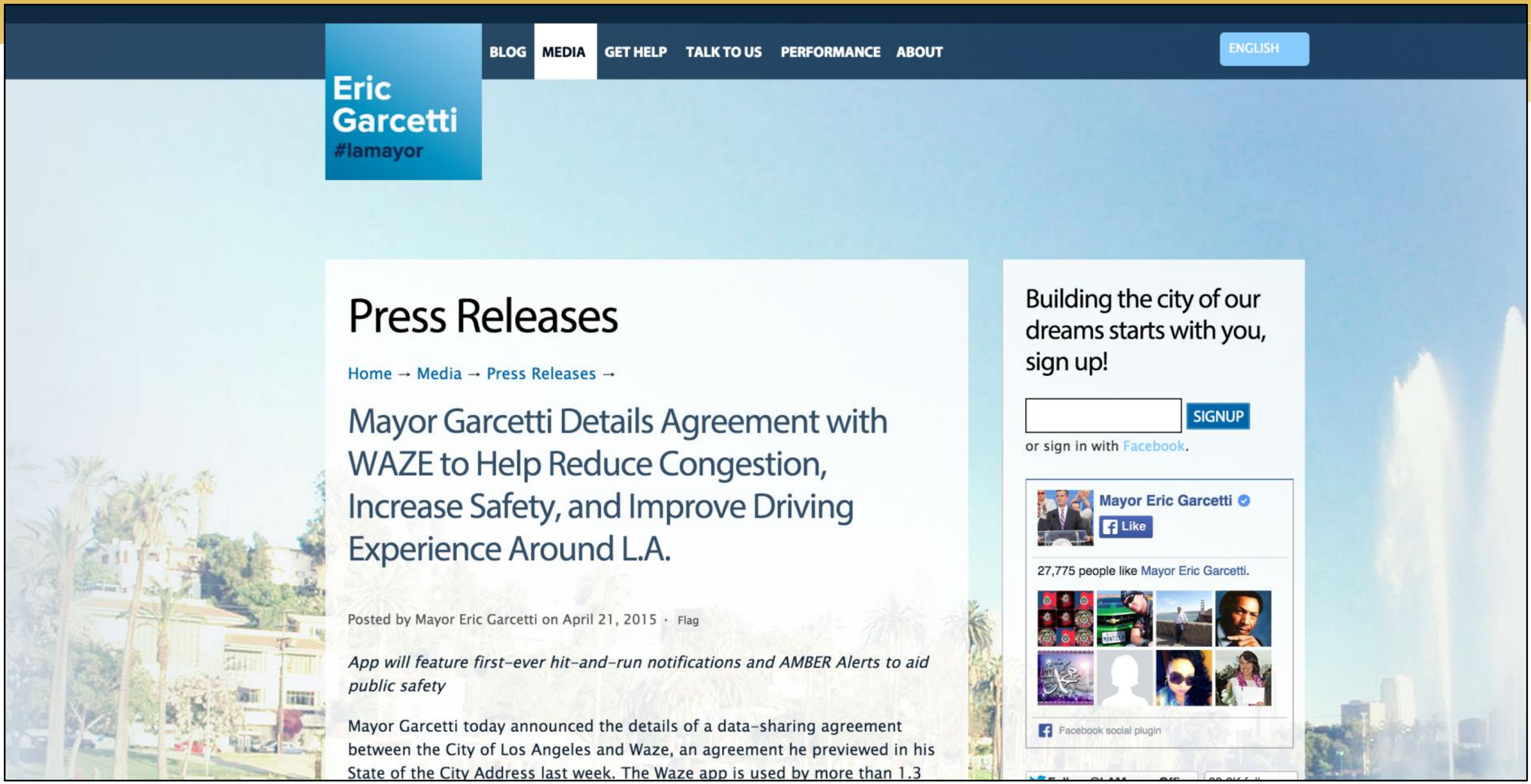



$p_{\text{Bing}} \sim x^{(t)}_{\text{Bing}}$




All users of each company “equal” by standards of the company i.e. same (shortest) travel time according to the company, “essentially” Nash.

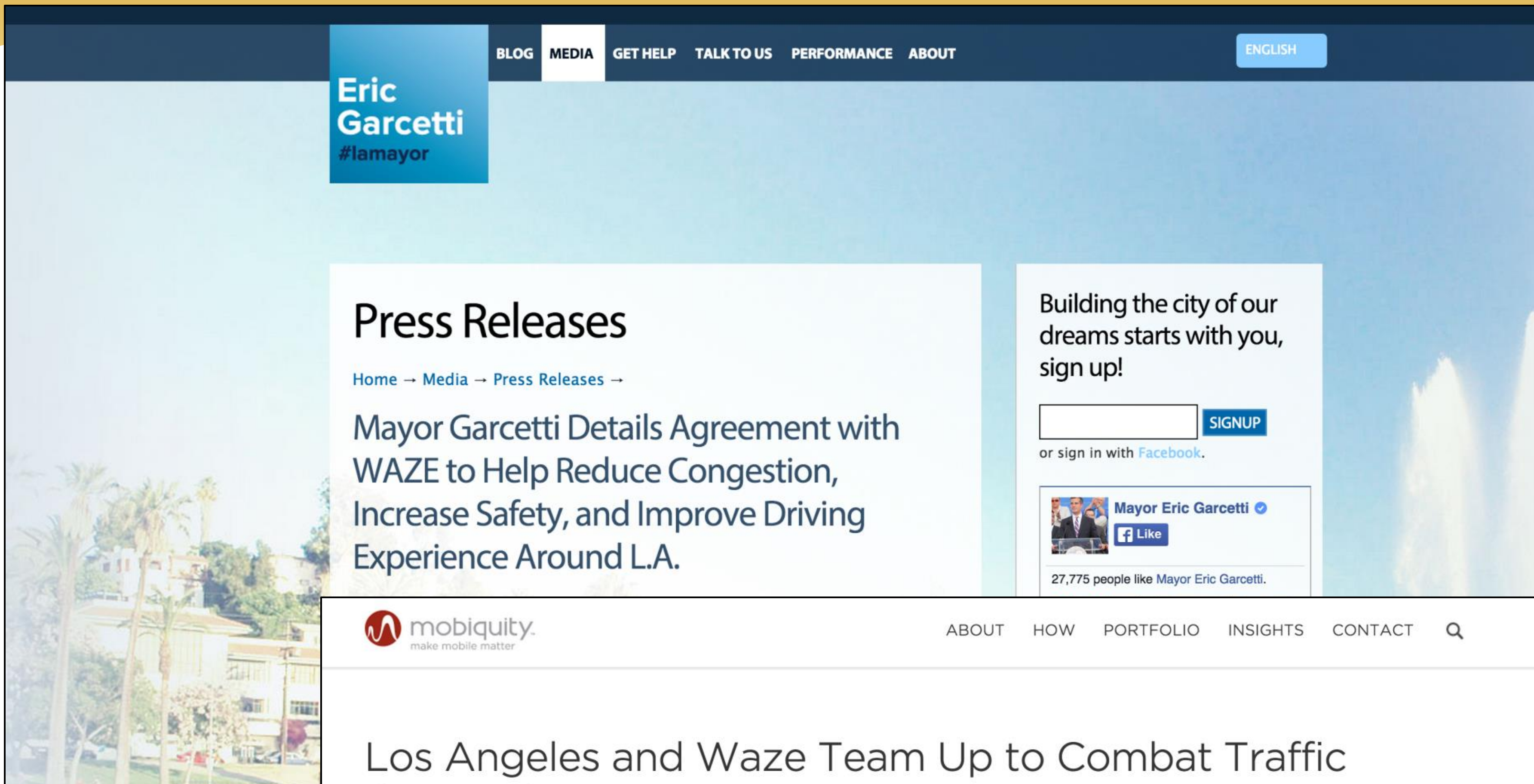
# The "LA problem" (soon in a city near you)






FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# The "LA problem" (soon in a city near you)



BLOG MEDIA GET HELP TALK TO US PERFORMANCE ABOUT ENGLISH

Eric Garcetti #lamayor

## Press Releases

Home → Media → Press Releases →

Mayor Garcetti Details Agreement with WAZE to Help Reduce Congestion, Increase Safety, and Improve Driving Experience Around L.A.

Building the city of our dreams starts with you, sign up!

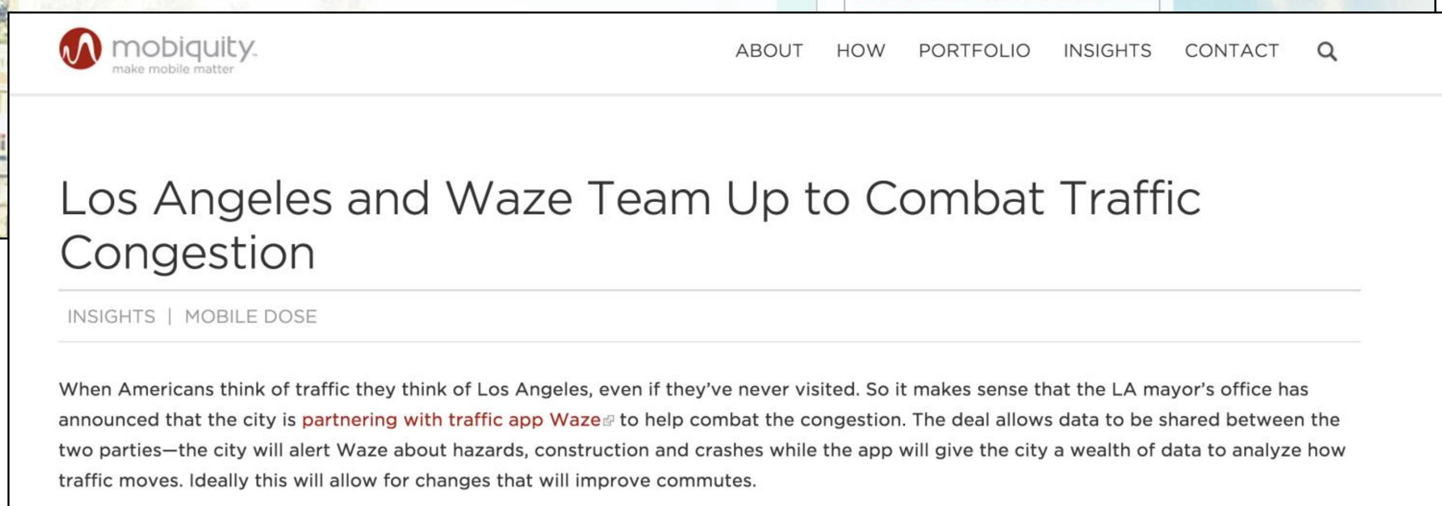
SIGNUP

or sign in with Facebook.

Mayor Eric Garcetti

Like

27,775 people like Mayor Eric Garcetti.



mobiquity make mobile matter

ABOUT HOW PORTFOLIO INSIGHTS CONTACT

## Los Angeles and Waze Team Up to Combat Traffic Congestion

INSIGHTS | MOBILE DOSE

When Americans think of traffic they think of Los Angeles, even if they've never visited. So it makes sense that the LA mayor's office has announced that the city is partnering with traffic app Waze to help combat the congestion. The deal allows data to be shared between the two parties—the city will alert Waze about hazards, construction and crashes while the app will give the city a wealth of data to analyze how traffic moves. Ideally this will allow for changes that will improve commutes.

# The "LA problem" (soon in a city near you)

≡ SECTIONS The Boston Globe SEARCH

beta Boston

**Today's top tech event**

28 Find a Startup Job You'll Love: Meet Startup Institute Boston **Details**
**More events**

GOOGLE

## Boston partners with Google's Waze app to improve traffic flow in the city


Waze logo added to a recent photo of the Southeast Expressway via BARRY CHIN/GLOBE STAFF



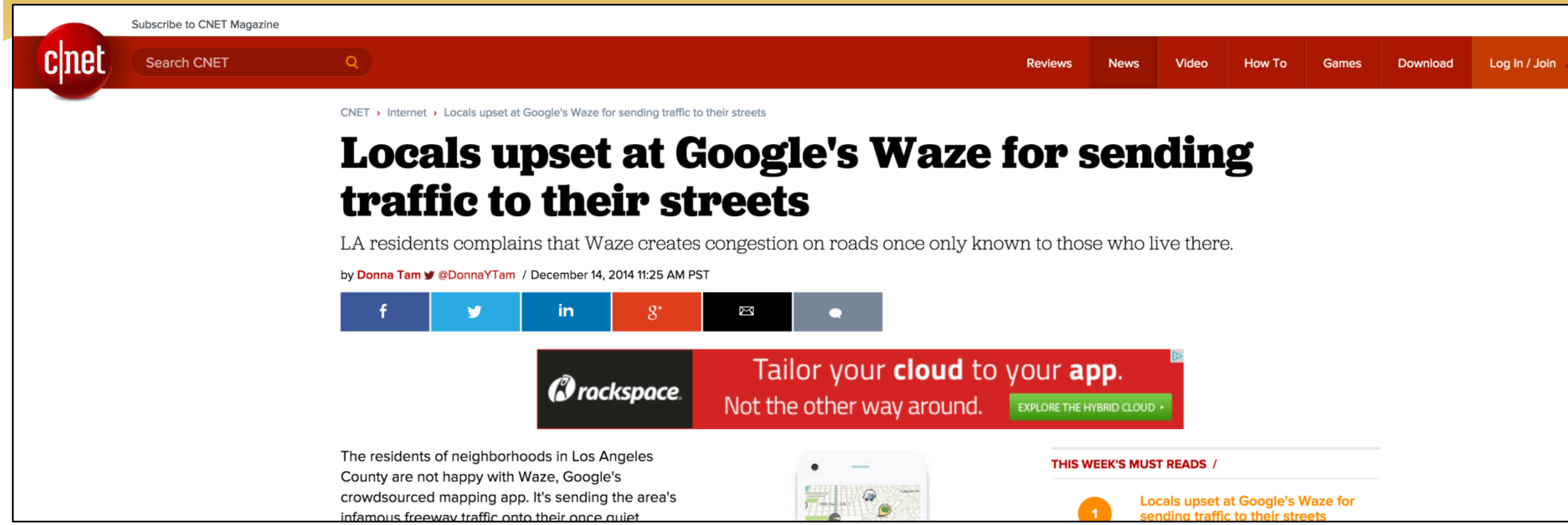
**BetaBoston in your email**

Your email address

☑ Daily

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# The "LA problem" (soon in a city near you)



Subscribe to CNET Magazine

Search CNET

Reviews | News | Video | How To | Games | Download | Log In / Join

CNET › Internet › Locals upset at Google's Waze for sending traffic to their streets

## Locals upset at Google's Waze for sending traffic to their streets

LA residents complains that Waze creates congestion on roads once only known to those who live there.

by Donna Tam @DonnaYTam / December 14, 2014 11:25 AM PST



The residents of neighborhoods in Los Angeles County are not happy with Waze, Google's crowdsourced mapping app. It's sending the area's infamous freeway traffic onto their once quiet

THIS WEEK'S MUST READS /

1 Locals upset at Google's Waze for sending traffic to their streets

# The "LA problem" (soon in a city near you)

Subscribe to CNET Magazine

cnet | Search CNET | Reviews | News | Video | How To | Games | Download | Log In / Join

CNET › Internet › Locals upset at Google's Waze for sending traffic to their streets

## Locals upset at Google's Waze for sending traffic to their streets

LA residents complains that Waze creates congestion on roads once only known to those who live there.

by Donna Tam

The residents
County are no
crowdsourced
infamous free

RATTER

## Waze Has No Concept Of The Hell That Is LA Traffic

Brittany Malooly
4/22/15 2:18pm

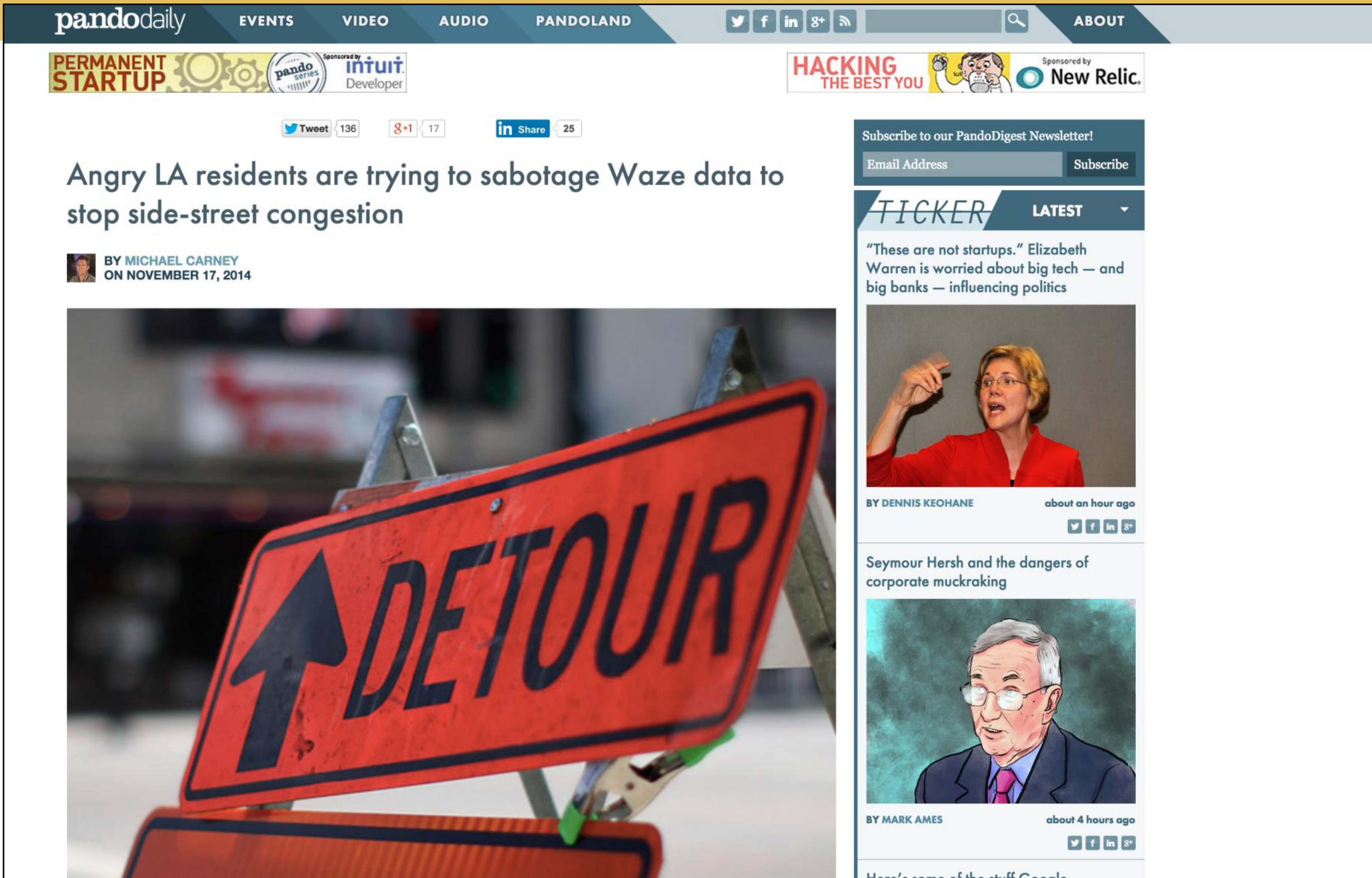
652

Waze markets itself as a hip, modern, community-based app that helps urban drivers save time and stay safe on the road, but Waze is the very same company that is repeatedly fucking over Angelenos during rush hour traffic.



Waze consistently recommends something people are referring to on Reddit as the "suicide left," which entails turning from a small side street onto a busy, multi-lane road during peak traffic hours without a stoplight. Other users also complain that the app will suggest clearing the entire road straight across. Not only do these options waste time as drivers either wait for a chance to cross or turn, but these suggestions are also dangerous.

# The "LA problem" (soon in a city near you)



pandodaily EVENTS VIDEO AUDIO PANDOLAND ABOUT

PERMANENT STARTUP pando series Sponsored by intuit Developer

HACKING THE BEST YOU Sponsored by New Relic.

Tweet 136 | g+1 17 | in Share 25

Angry LA residents are trying to sabotage Waze data to stop side-street congestion

BY MICHAEL CARNEY
ON NOVEMBER 17, 2014



Subscribe to our PandoDigest Newsletter!
Email Address Subscribe

TICKER LATEST

"These are not startups." Elizabeth Warren is worried about big tech — and big banks — influencing politics

BY DENNIS KEOHANE about an hour ago

Seymour Hersh and the dangers of corporate muckraking

BY MARK AMES about 4 hours ago

Here's some of the stuff Google

# The "LA problem" (soon in a city near you)

ABOUT STAFF CONTACT ADVERTISE FAQ PRIVACY POLICY TERMS OF SERVICE

mynewsLA.com



CRIME GOVERNMENT BUSINESS EDUCATION SPORTS HOLLYWOOD LIFE WEATHER OC Search...

LATEST NEWS La Mirada man accused in murder of his wife in 1992 arrested in Antigua

Home » Government » This Article

## 'Cut-through' traffic caused by Waze app must stop, L.A. councilman says

POSTED BY JOHN SCHREIBER ON APRIL 28, 2015 IN GOVERNMENT | 10,658 VIEWS | 2 RESPONSES

Recommend Share 120



A Los Angeles city deal with traffic app Waze may be great, but some local communities are being inundated with "cut-through" traffic that must stop, a Los Angeles City Councilman said Tuesday.

Paul Krekorian introduced a motion to help local neighborhoods, saying Waze should send drivers away from residential streets and onto major roadways as part of the company's data-sharing agreement with the city.

Mayor Eric Garcetti announced last week that the city is sharing road closure data with Waze to improve its service, and in return the city is getting live updates about traffic patterns

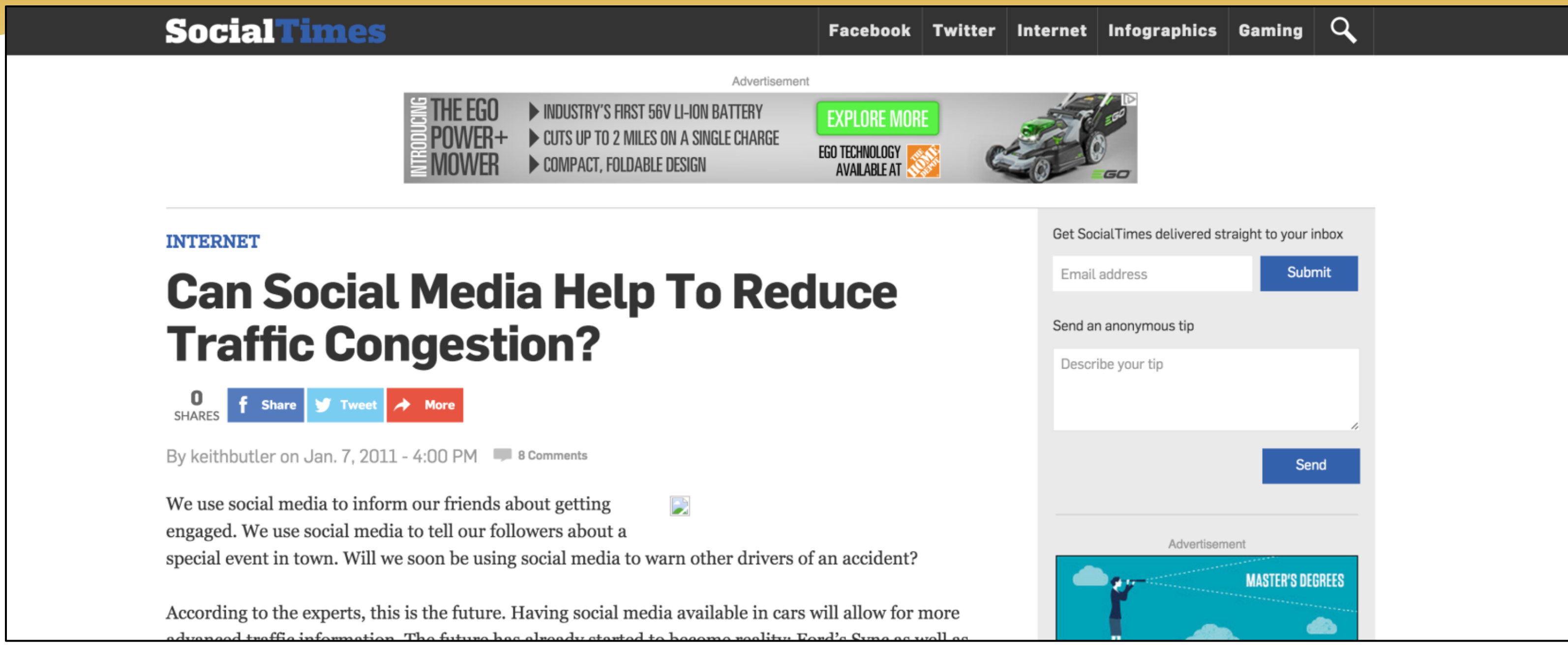
GET MYNEWSLA.COM'S FREE NEWSLETTERS

We'll send you the latest headlines every morning at 7 and every weekday afternoon at 5. Our newsletters are **free** and your email address is secure.

Email Address

SUBSCRIBE

# The one (?) million dollar question


SocialTimes | Facebook | Twitter | Internet | Infographics | Gaming



INTERNET

## Can Social Media Help To Reduce Traffic Congestion?

0 SHARES Share Tweet More

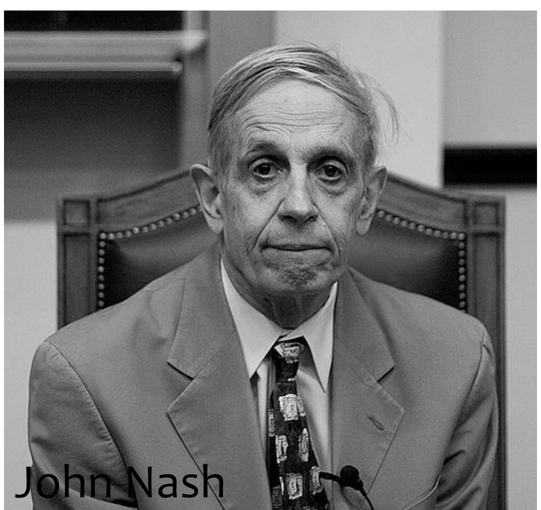
By keithbutler on Jan. 7, 2011 - 4:00 PM 8 Comments

We use social media to inform our friends about getting engaged. We use social media to tell our followers about a special event in town. Will we soon be using social media to warn other drivers of an accident?

According to the experts, this is the future. Having social media available in cars will allow for more

Get SocialTimes delivered straight to your inbox
Email address Submit
Send an anonymous tip
Describe your tip
Send




John Nash

Is steering mobility towards Nash eq. good?

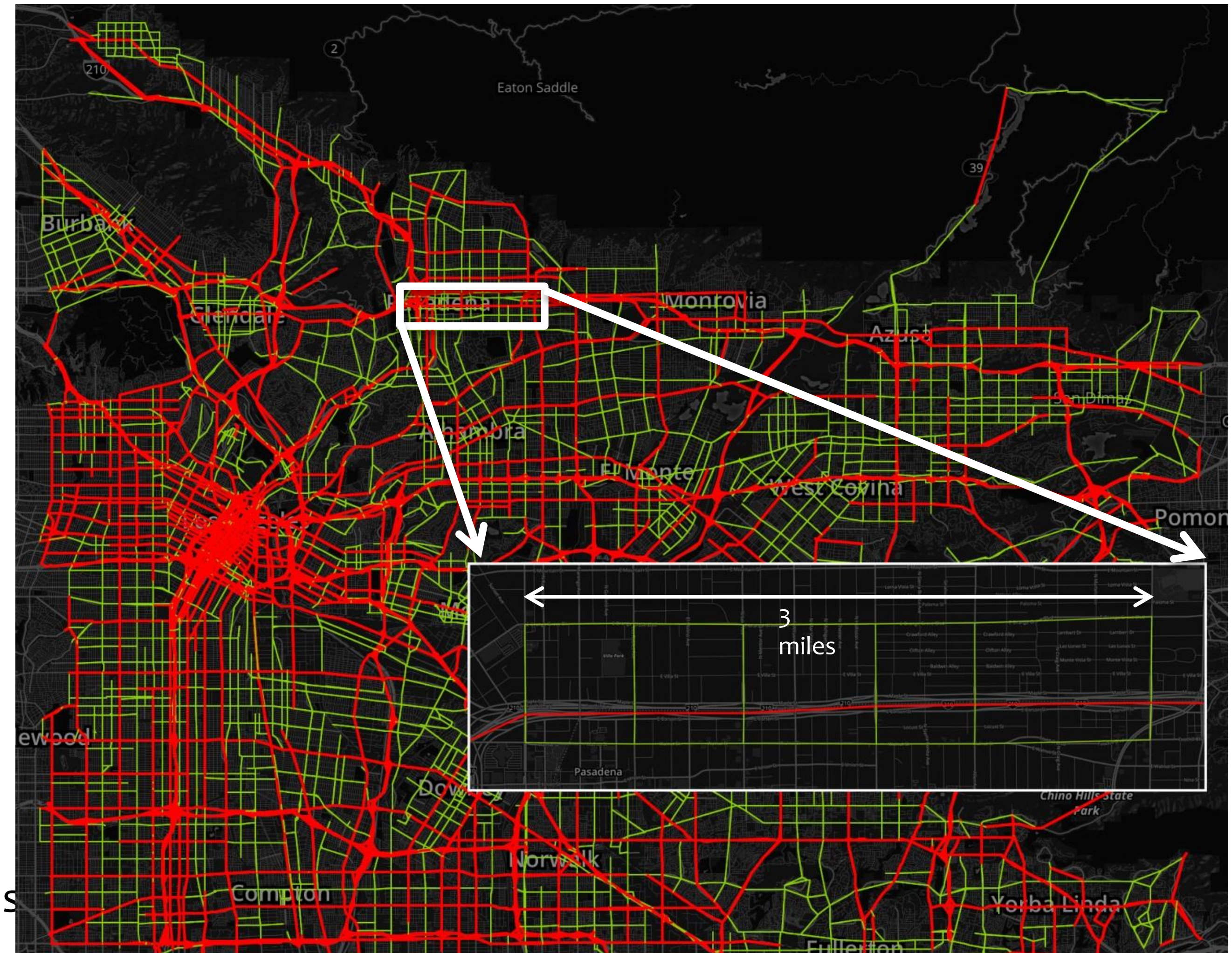
- System now could potentially be doing worse than Nash.
- Nash is obviously not as good as system optimum (hence price of anarchy, value of altruism etc.)
- How bad / good is displacing current equilibrium towards Nash, which is what apps are doing?

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# Example for 3 miles in Pasadena

Let us assume overnight, 15% of users of I210 start using Waze:

- Immediate massive reroute through Pasadena
- Travel time in Pasadena instantaneously goes up by 17%



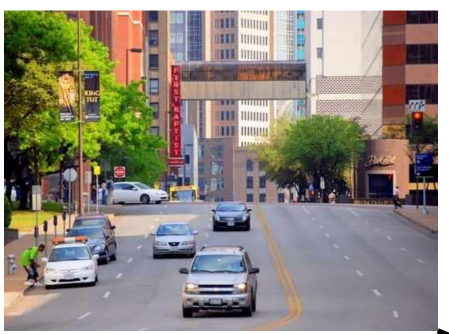

John Nas

# Example for 3 miles in Pasadena

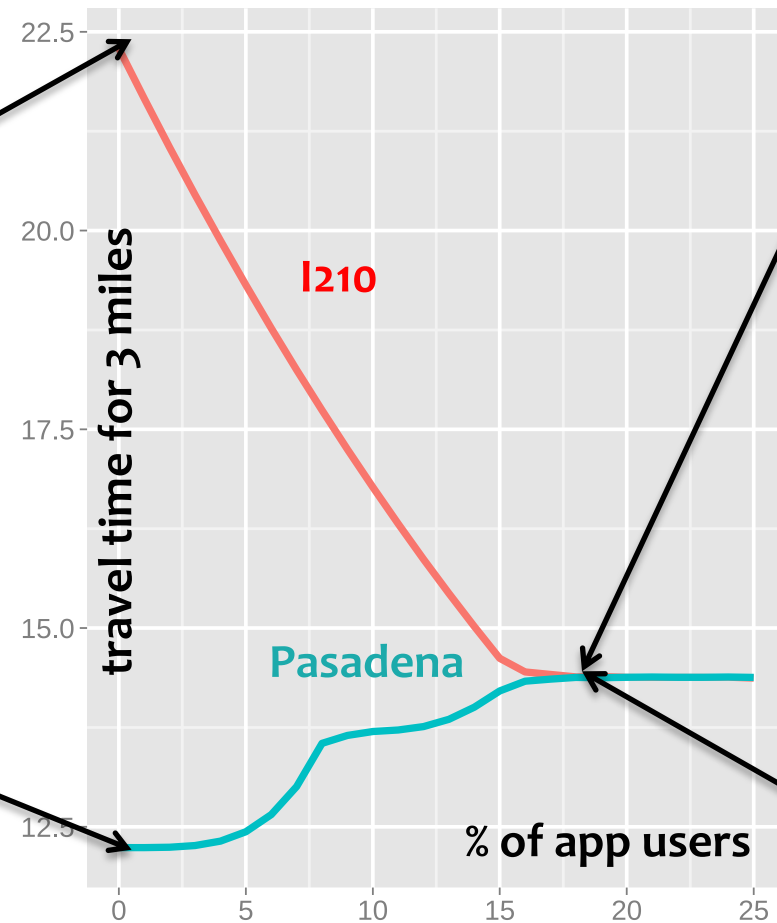
Let us assume overnight, 15% of users of I210 start using Waze:

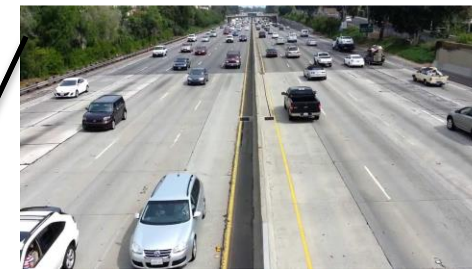
- Immediate massive reroute through Pasadena
- Travel time in Pasadena instantaneously goes up by 17%

I210 "before"



City "before"



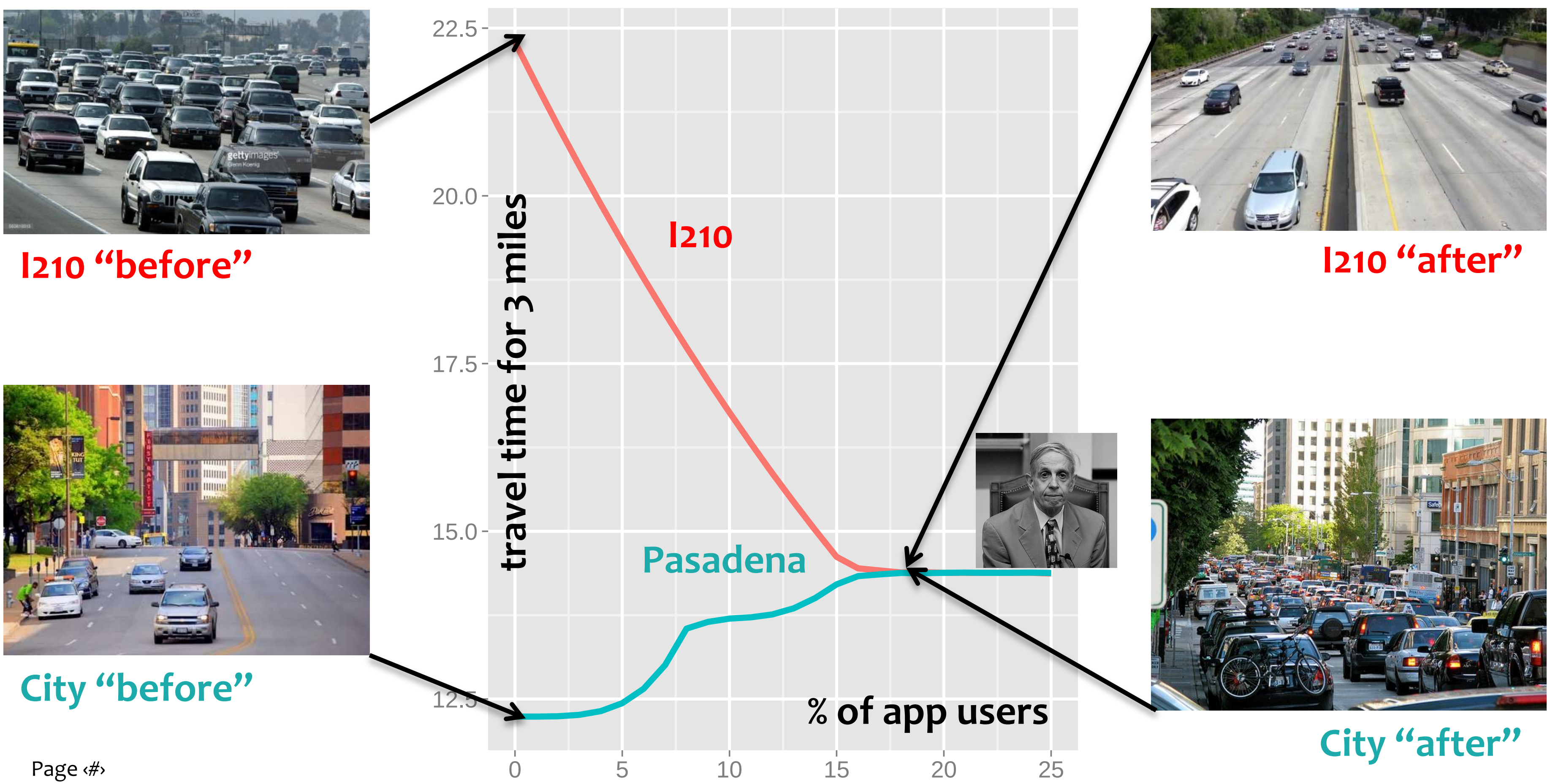
I210 "after"

City "after"

# Example for 3 miles in Pasadena

Let us assume overnight, 15% of users of I210 start using Waze:

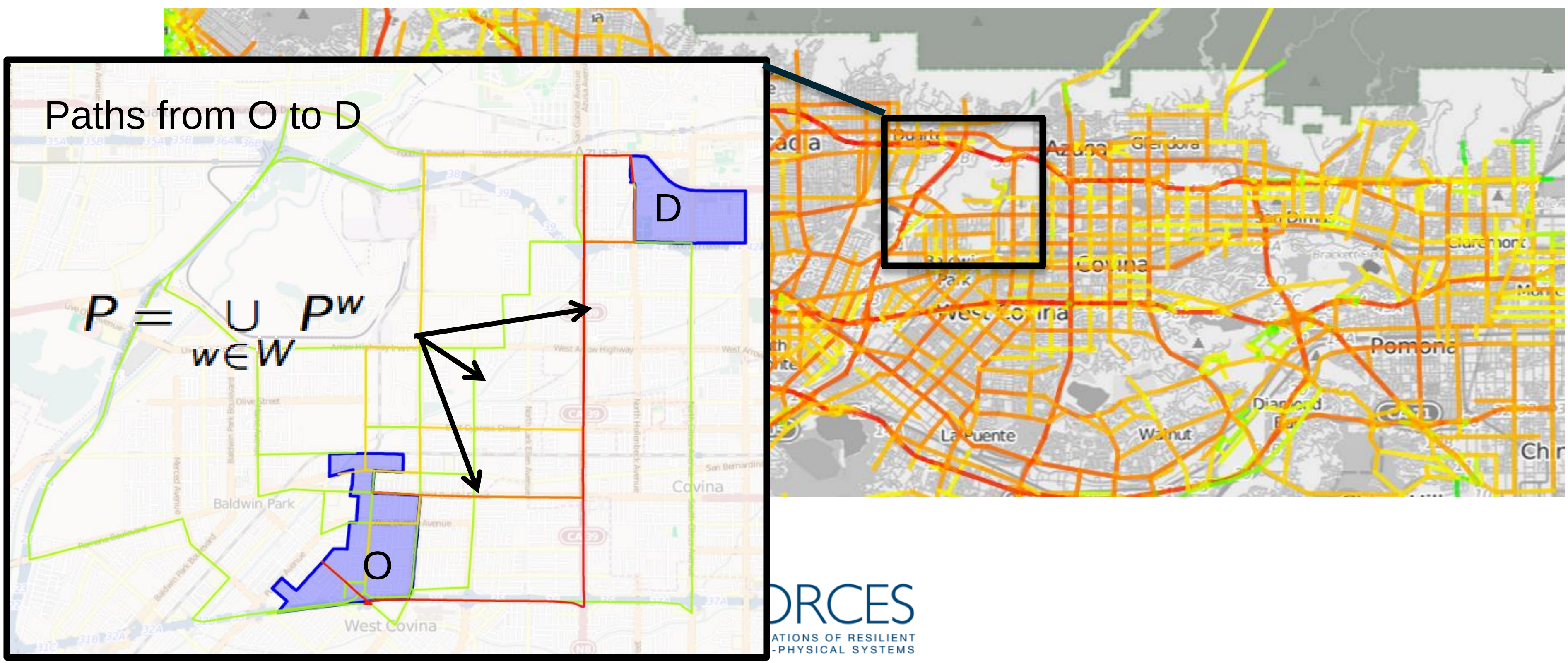
– Immediate massive reroute through Pasadena
– Travel time in Pasadena instantaneously goes up by 17%



# Graph formulation

Given a directed graph $G = (N, A)$

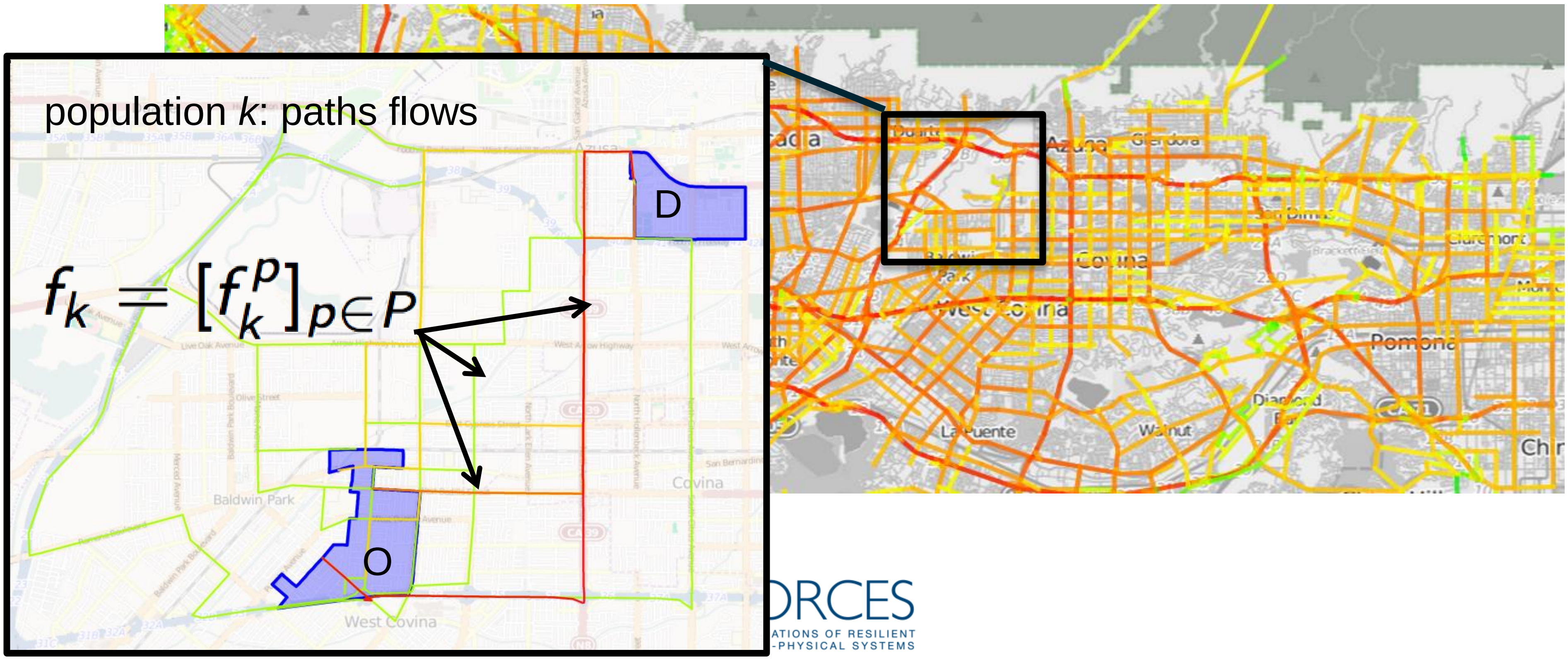
- OD pairs $w \in W \subset N^2$ with paths $P = \bigcup_{w \in W} P^w$
- Arc-path incidence matrix $\Delta = I(a \in p)$



# Graph formulation

$N$ types of users $k = 1, \cdots, N$:

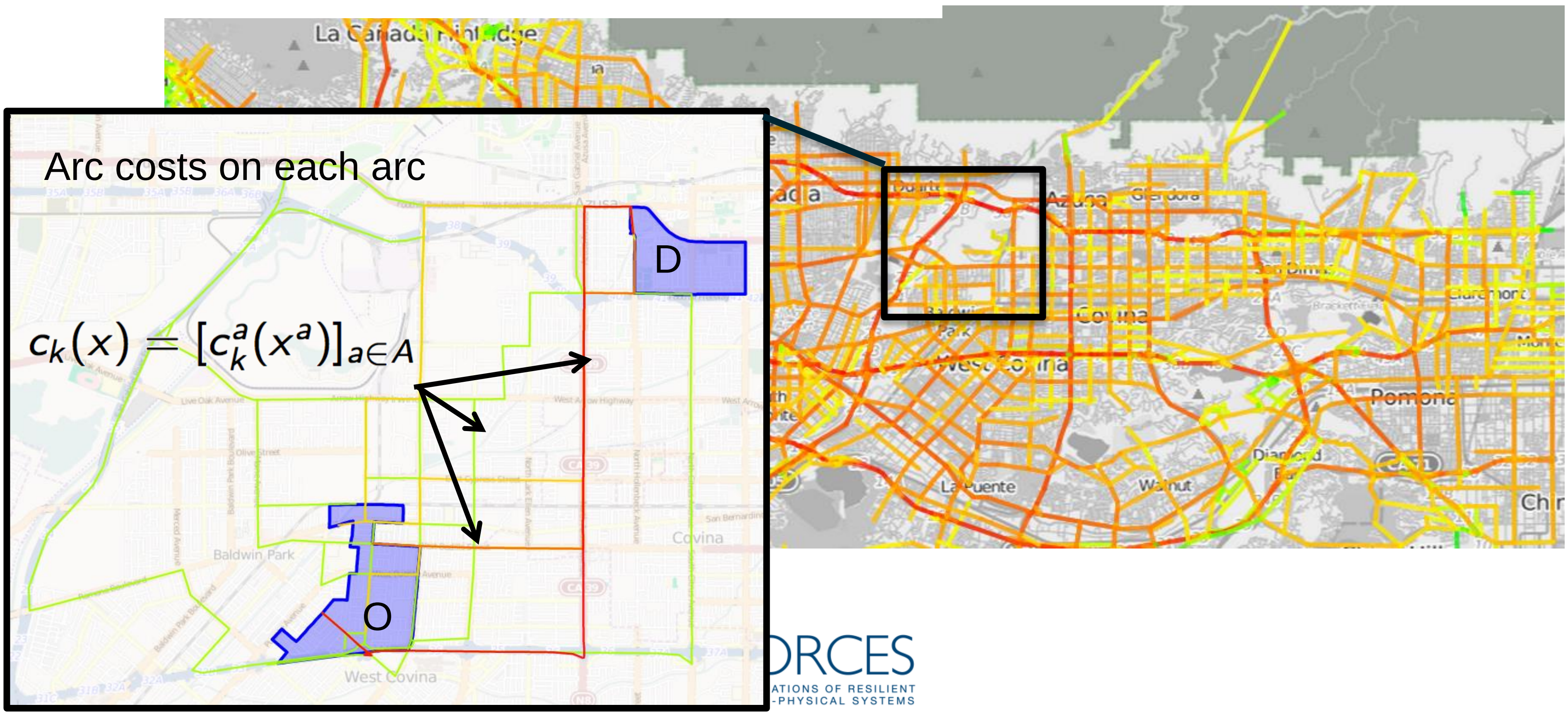
- Demand $T_k = [T_k^w]_{w \in W}$
- Path flows $f_k = [f_k^p]_{p \in P}$ and arc flows $x_k = [x_k^a]_{a \in A} = \Delta f_k$
- Total flows $f = \sum_{k=1}^{N} f_k$ and $x = [x^a]_{a \in A} = \Delta f$



# Graph formulation

$N$ types of users $k = 1, \cdots, N$:

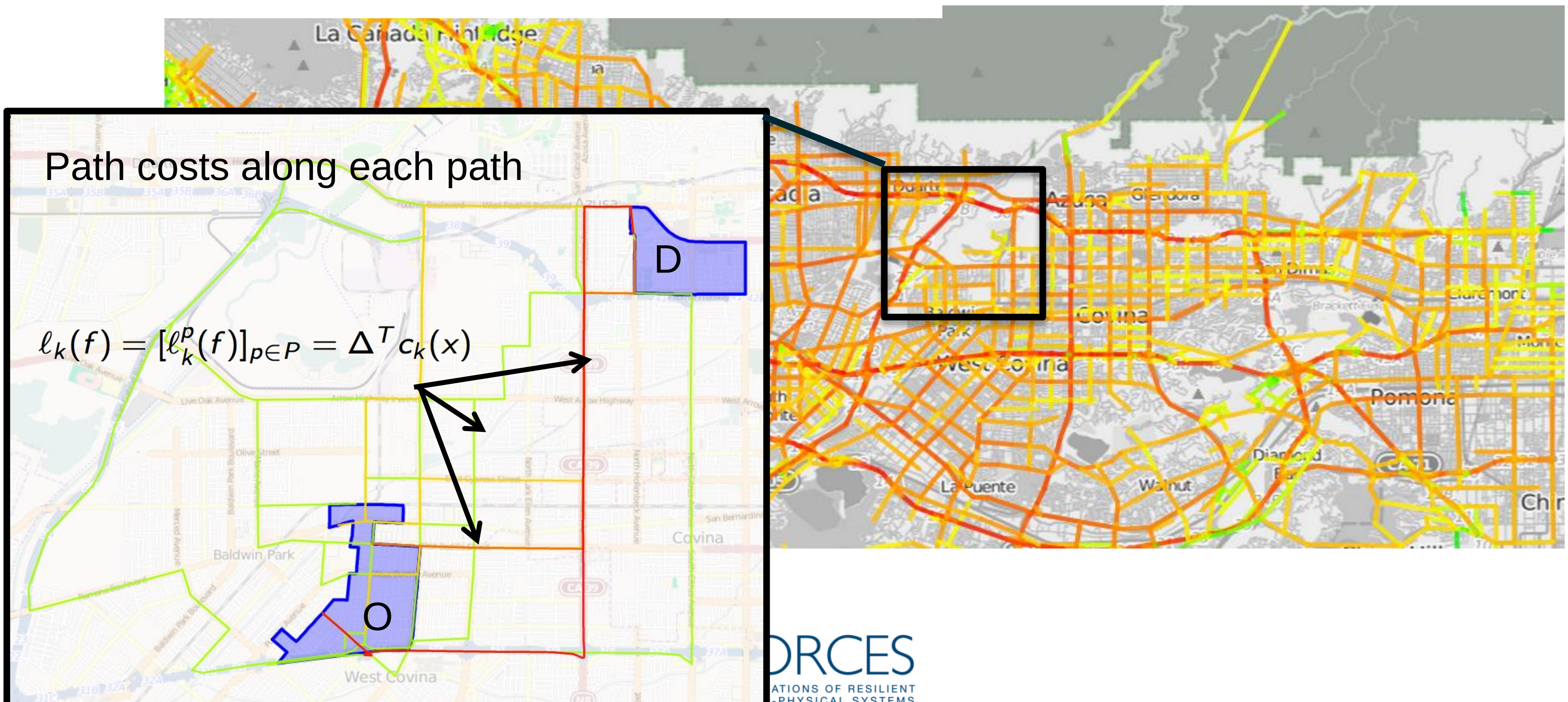
- Total flows $f = \sum_{k=1}^{N} f_k$ and $x = [x^a]_{a \in A} = \Delta f$
- Arc costs $c_k(x) = [c_k^a(x^a)]_{a \in A}$
- Path costs $\ell_k(f) = [\ell_k^p(f)]_{p \in P} = \Delta^T c_k(x)$



# Graph formulation

$N$ types of users $k = 1, \cdots, N$:

- Total flows $f = \sum_{k=1}^{N} f_k$ and $x = [x^a]_{a \in A} = \Delta f$
- Arc costs $c_k(x) = [c_k^a(x^a)]_{a \in A}$
- Path costs $\ell_k(f) = [\ell_k^p(f)]_{p \in P} = \Delta^T c_k(x)$



# Graph formulation

Given a directed graph $G = (N, A)$

- OD pairs $w \in W \subset N^2$ with paths $P = \bigcup_{w \in W} P^w$
- Arc-path incidence matrix $\Delta = I(a \in p)$

$N$ types of users $k = 1, \cdots, N$:

- Demand $T_k = [T_k^w]_{w \in W}$
- Path flows $f_k = [f_k^p]_{p \in P}$ and arc flows $x_k = [x_k^a]_{a \in A} = \Delta f_k$
- Total flows $f = \sum_{k=1}^{N} f_k$ and $x = [x^a]_{a \in A} = \Delta f$
- Arc costs $c_k(x) = [c_k^a(x^a)]_{a \in A}$
- Path costs $\ell_k(f) = [\ell_k^p(f)]_{p \in P} = \Delta^T c_k(x)$

# Traffic equilibrium (heterogeneous players)

$$C_k^w(f) := \min_{q \in P^w} \ell_k^q(f) \quad \forall\, k = 1, \cdots, N \tag{1}$$

**Definition: Nash equilibrium**

$(f_k)_{k=1,\cdots,N}$ is an eq. flow if $\forall\, w \in W$, $\forall\, p \in P^w$, $\forall\, k = 1, \cdots, N$:

$$f_k^p > 0 \quad \Longrightarrow \quad \ell_k^p(f) = C_k^w(f) \tag{2}$$

**Variational inequality**

Equivalently, for all feasible path flows $(h_k)_{k=1,\cdots,N}$

$$\sum_k \ell_k(f)^T h_k \geq \sum_k \ell_k^p(f)^T f_k \tag{3}$$

Arc flow formulation, for all feasible arc flows $(x_k)_{k=1,\cdots,N}$

$$\sum_k c_k(x)^T y_k \geq \sum_k c_k(x)^T x_k \tag{4}$$

# Coupled optimization programs

No potential exists because externality symmetry does not hold.[2]

Coupled convex potentials

$$\phi_k(x) = \sum_a \int_0^{x_a} c_k^a(u)du \quad \Longrightarrow \quad \nabla \phi_k(x) = [c_k^a(x^a)]_a \tag{6}$$

**Definition: Equilibrium as a solution to a Nash Equilibrium game**

$\{x_k^\star\}_{k=1,\cdots,N}$ is an equilibrium if and only if

$$x_k^\star \in \text{argmin}_{x_k \in X_k} \phi_k(x_k + x_{-k}^\star) \quad \forall\ k \tag{7}$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

# Coupled optimization programs

## Convergence of Gauss-Seidel best response-based algorithm to the equilibrium

At each iteration every player, given the strategies of the others, updates his own strategy by solving his convex optimization problem

$$\min_{x_k \in X_k} \phi_k(x_k + x_{-k}) \quad \forall k \tag{8}$$

Convergence is guaranteed when each $\phi_k$ is continuously differentiable and convex in $x_k$ for fixed $x_{-k}$, and the strategy sets $X_k$ are closed and convex.[a]

[a] G. Scutari, D. P. Palomar, and J-S. Pang, "Convex Optimization, Game Theory, and Variational Inequality Theory", *IEEE Signal Processing Magazine*, Vol. 35, May 2010.

**Block coordinate descent for solving the heterogeneous game**

**for** $t \in 1, 2, \cdots$ **do**
  **for** $k \in 1, 2, \cdots, N$ **do**
    $x_k^{t+1} \leftarrow \text{argmin}_{y_k \in C_k} \phi_k(x_1^{t+1} + \cdots + x_{k-1}^{t+1} + y_k + x_{k+1}^{t} + \cdots + x_N^{t})$
    $t \leftarrow t + 1$
  **end for**
**end for**

# Conclusions

## Historical perspective

- The years 2007-2012 have brought information to mobility, giving drivers the ability to achieve shortest travel time.
- The years 2010-2016 have seen the impact of these technologies on mobility patterns (changes in modality, routing, behavior)
  - Companies (apps): are "learning"
  - Users are "learning"

## Scientific contributions in a "post travel time / optimal control era"

- Under certain conditions, companies working non-cooperatively on user routing might converge to a Nash equilibrium.
- Conditions of this convergence depends on the assumptions on the model of these companies
- Practical implementations on "humans" reveals convergence to Nash equilibria

## Public policy perspective

- Today, in many regions of the world, traffic "[non]-equilibrium" is probably worse than Nash equilibrium
- Apps probably contribute to steering system towards Nash
- While Nash is probably still better than current situation globally, it redistributes congestion, leading to increased congestion in sub-urban areas