Mean field for large bike sharing systems

Christine Fricker,
Nicolas Gast (EPFL, Switzerland)
Hanene Mohamed (Universite Paris Ouest, France)

March, 22th, 2012

Motivation

- Bike sharing systems become popular
  - Since Velib in Paris 2007
  - More than 200 programs around the world
    Ex: Barcelona, Lyon, London, Montreal, Washington, Hangzou (China)
- why?
  - Good for the town (pollution, traffic jams, health);
  - Good for the citizen (not to buy, to park the bike, no risk of theft).
- BUT

Empty station

:(

Full station

:(

Good stations

:)

problematic stations

Empty station

:(

Full station

:(

Good stations

:)

$\underbrace{\hspace{10em}}_{\textbf{problematic stations}}$

problematic stations

Le Monde

Mercredi 13 Juillet 2011

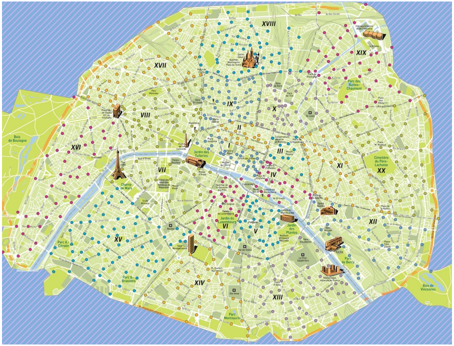
& vous

MODE DE VIE

Le Vélib' ne se trouve pas toujours à sa place

Quatre ans après son lancement, le système se heurte encore au problème des bornes vides ou engorgées

Velib : a large stochastic network



Map of Velib' stations in Paris.

- **1500** stations
- **20000** bikes

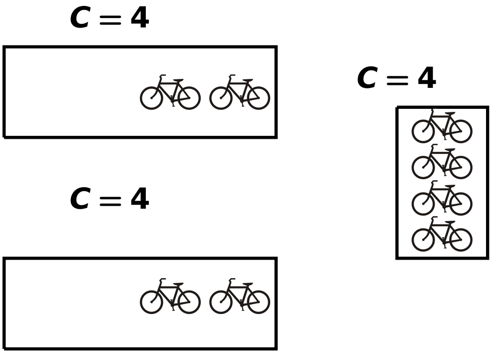
Usage:

- Take a bike from any station.
- Use it.
- Return it to a station of your choice

- Performance: low proportion of
  - empty stations
  - full stations

- huge number of nodes (stations) $\rightarrow +\infty$
  $\Rightarrow$ mean field

A simple case: homogeneous



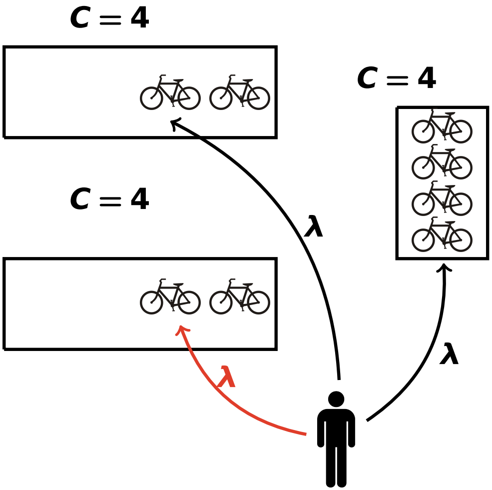
For all N stations:

- Capacity C

Extended to non-homogeneous case

- arrival rate, routing probability

A simple case: homogeneous



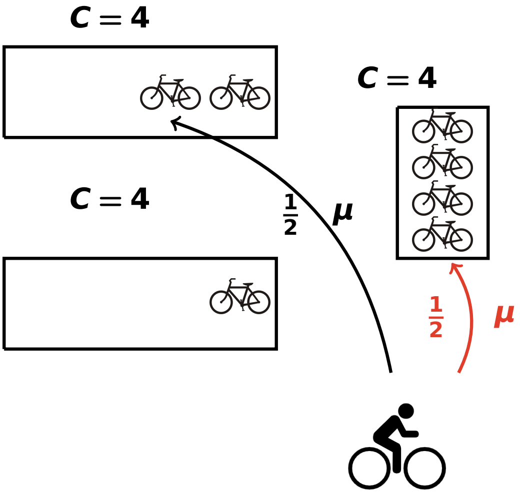
For all N stations:

- Capacity C
- Arrival rate λ

Extended to non-homogeneous case

- arrival rate, routing probability

A simple case: homogeneous



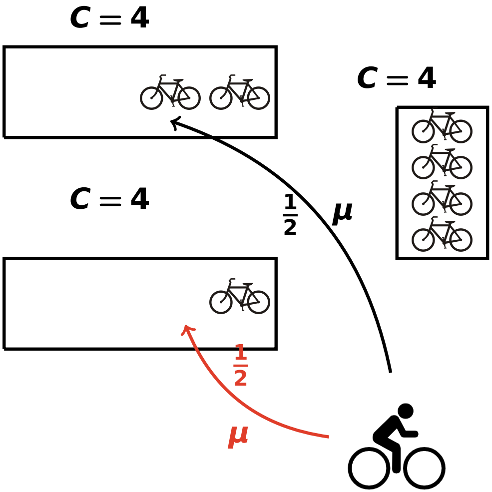
For all N stations:

- Capacity C
- Arrival rate λ
- Travel time: exponential of mean $1/\mu$
- Choose one station at random.

Extended to non-homogeneous case

- arrival rate, routing probability

A simple case: homogeneous



For all N stations:

- Capacity C
- Arrival rate λ
- Travel time: exponential of mean $1/\mu$
- Choose one station at random.
- If full, try again at random ($\approx$ local search)

Extended to non-homogeneous case

- arrival rate, routing probability

A first result

The stationary distribution of the number of bikes in a station converges to $\boldsymbol{geom(\rho)}$ on $\mathbf{0, \ldots, C}$, with $\boldsymbol{\rho}$ solution of

$$s = mean(geom(\rho)) + \frac{\lambda}{\mu}\rho$$

as $\boldsymbol{M, N}$ tends to $+\infty$,

$\boldsymbol{M/N \to s}$ the average number of bikes per station.

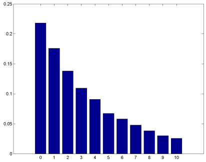
A first result

The stationary distribution of the number of bikes in a station converges to $\boldsymbol{geom(\rho)}$ on $\mathbf{0, \ldots, C}$, with $\boldsymbol{\rho}$ solution of

$$s = mean(geom(\rho)) + \frac{\lambda}{\mu}\rho$$

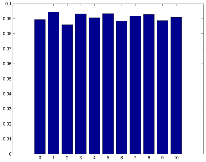
as $\boldsymbol{M, N}$ tends to $+\infty$,
$\boldsymbol{M/N \to s}$ the average number of bikes per station.

$$s < \frac{C}{2} + \frac{\lambda}{\mu}$$



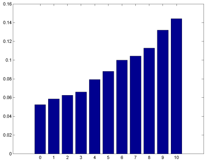
$$\rho < 1$$

$$s = \frac{C}{2} + \frac{\lambda}{\mu}$$



$$\rho = 1$$

$$s > \frac{C}{2} + \frac{\lambda}{\mu}$$

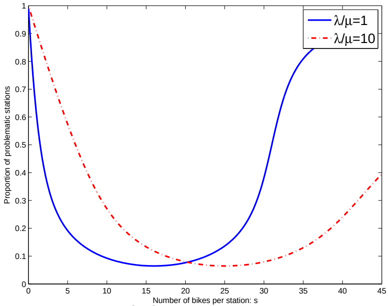


$$\rho > 1$$

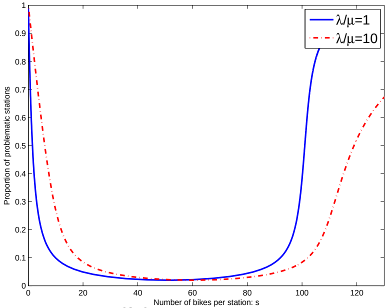
Plotting performance...

y-axis: limiting proportion of problematic stations
x-axis: number of bikes/station s.

$$\rho \mapsto (s(\rho), geom(\rho)(0) + geom(\rho)(C)).$$



(a) $C = 30$.



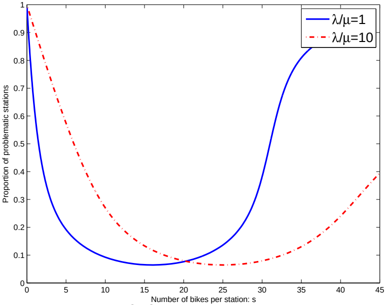
(b) $C = 100$.

- minimal $2/(C+1)$ for $s = s_{opt} \stackrel{def}{=} \lambda/\mu + C/2$
- "flat" at s_{opt}.

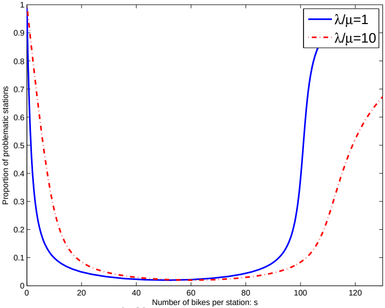
Plotting performance...

y-axis: limiting proportion of problematic stations
x-axis: number of bikes/station s.

$$\rho \mapsto (s(\rho), geom(\rho)(0) + geom(\rho)(C)).$$



(c) $C = 30$.



(d) $C = 100$.

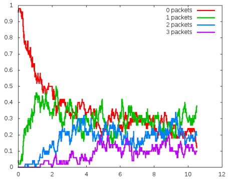
- minimal $2/(C+1)$ for $s = s_{opt} \stackrel{def}{=} \lambda/\mu + C/2$
- "flat" at s_{opt}.

Ex: for $C = 30$: at least **6.5%** of problematic stations.

Techniques: mean field limit

- Gives the steady state when $\boldsymbol{N}$ goes to infinity of

$$\boldsymbol{X_i} = \frac{\mathbf{1}}{\boldsymbol{N}} \# \{\textbf{stations with i bikes}\}$$



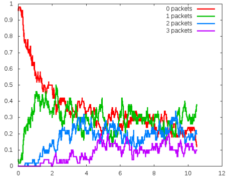
For fixed $\boldsymbol{N}$, $\boldsymbol{X_i}$ is a (reversible) Markov process

- Steady state (explicit)

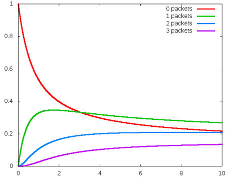
Techniques: mean field limit

- Gives the steady state when $\boldsymbol{N}$ goes to infinity of

$$X_i = \frac{1}{N}\#\{\textbf{stations with i bikes}\} \rightarrow \frac{\rho^i}{Z}$$



$N \rightarrow \infty$



For fixed $\boldsymbol{N}$, $\boldsymbol{X_i}$ is a (reversible) Markov process

- Steady state (explicit)

System described by an ODE

- The ODE has a unique equilibrium point.
- Closed-form formula.

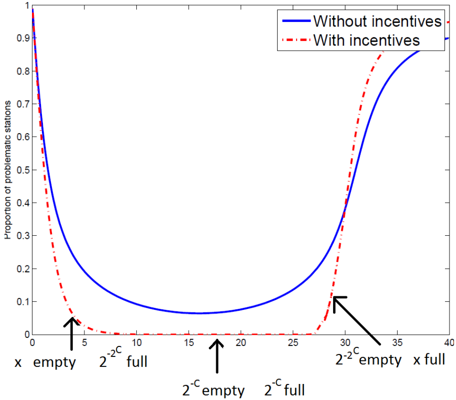
Key: proba. interpretation of the ODE

Contributions on mean field

- mean field limit: explicit derivations
- CV of invariant measures via a Lyapunov function
- extension to some heterogeneity
  idea of proof: discrete/continuous

Choice

- Rule: When returning, choose at random two stations and return to the least loaded.
  (Load balancing via power of choice)



- Improves performance from **1/*C*** to **2^{-C}**
- True even if incentives=just a proportion choose

Choice

- Rule: choose two neighbors
- Simulations (2D) show performance is $\approx \mathbf{2^{-C/2}}$.
  (recall: no geometry: $\mathbf{2^{-C}}$, no incentive: $\mathbf{1/C}$).

Conclusion

Good understanding of the homogeneous case

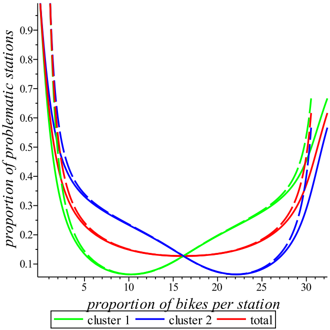
- Poor performance: $\mathbf{1/C}$ problematic stations (even in symmetric case)
- Choose among 2 neighbors helps a lot: $\mathbf{2^{-c/2}}$
- Avoid empty and full stations is useless!
- About redistribution by trucks
  - an simple efficient algorithm
  - a minimal value for truck rate of $\mathbf{1/C}$: 0!

Future work

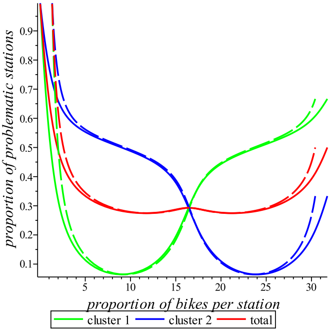
- mean field with geometry (local search)
- redistribution/incentives +heterogeneity: algorithms
- experimentation building a realistic model for Paris collaboration with Mairie de Paris

Heterogeneity

From numerical simulations to theoretical results



(e) $\boldsymbol{\lambda_1/\lambda_2 = 1.3}$



(f) $\boldsymbol{\lambda_1/\lambda_2 = 2}$

Figure: Performance of the system

A more complex system

- heterogeneity: housing vs working, uphill vs downhill,...
- geometry local search
- time dependency: daily period