Sizing, Incentives and Regulations in Bike-sharing Systems

Nicolas Gast [1]

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

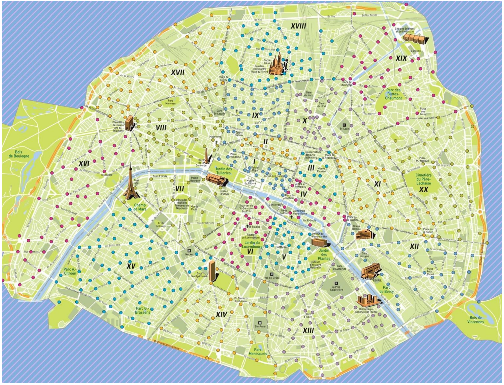
Informs Meeting 2011, Charlotte, NC

1. joint work with Christine Fricker (Inria)

Outline

A new transportation system.

- Bike sharing systems started in the 60s.
- Increasing popularity. Ex : Velib' in Paris (2007).
- $>$ 400 cities (and counting). Ex : Barcelona, Montreal, Washington.



Map of Velib' stations in Paris (France).

Example of Velib' :

- 20000 bikes
- 1500 stations.

Usage :

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

Public but different from public transportation

Many advantages :

- Good for the town (pollution, traffic jams, health) ;
- Good for the citizen (not to buy, to park the bike, no risk of theft).

However : congestions problems due to flows and random choices.

Empty station

:(

Full station

:(

Good stations

:)

problematic stations

How to manage them ?

- Sizing : number of stations ? bikes ? locations for bikes per station ?
- performance : a low number of problematic stations
  - low number of empty or full stations
- time dependent arrival rate : daily period
- heterogeneity : popular or non popular stations
  (housing and working areas, uphill and downhill stations,...)

Our approach : study the impact of random choices

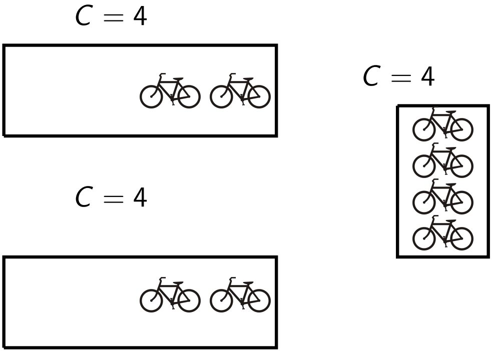
1. Qualitative behavior and quantitative impact of different factors.
2. Strategies : redistribution (trucks) and incentives (pricing).

Related work :

- Traces analysis (Borgnat et al. 10, Vogel et al. 11, Nair et al. 11]
- Redistribution based of forcast [Raviv et al. 11, Chemla et al. 09]
- Few stochastic models. In a similar context : limiting regime with infinite capacity [Malyshev Yakovlev 96, Georges Xia 10]

Outline

The simplest case : homogeneous



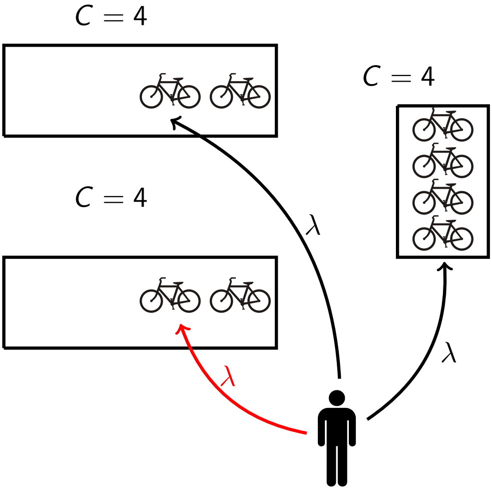
For all N stations :

- Fixed capacity C

Will be extended to non-homogeneous :

- arrival rate, routing probability

The simplest case : homogeneous



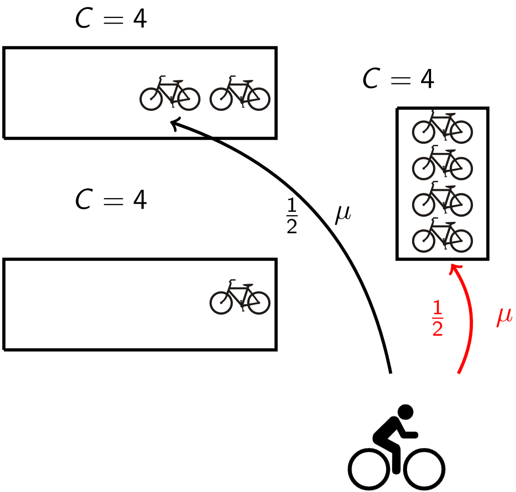
For all N stations :

- Fixed capacity C
- Arrival rate λ.

Will be extended to non-homogeneous :

- arrival rate, routing probability

The simplest case : homogeneous



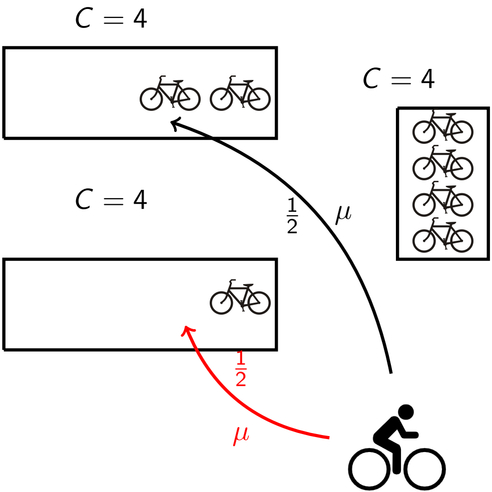
For all N stations :

- Fixed capacity C
- Arrival rate λ.
- Routing matrix : homogeneous.
- Travel time : exponential of mean $1/\mu$.

Will be extended to non-homogeneous :

- arrival rate, routing probability

The simplest case : homogeneous



For all N stations :

- Fixed capacity C
- Arrival rate λ.
- Routing matrix : homogeneous.
- Travel time : exponential of mean $1/\mu$.
- Other destination chosen if full ($\approx$ local search).

Will be extended to non-homogeneous :

- arrival rate, routing probability

A first result : distribution of stations

We focus on the distribution of occupancy in steady state.

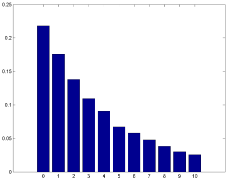
Theorem

There exists ρ, such that in steady state, as N goes to infinity :

$$x_i = \frac{1}{N}\#\{\text{stations with i bikes}\} \propto \rho^i.$$

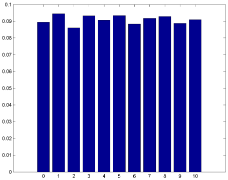
We have $\rho \le 1$ iff $s \le \frac{C}{2} + \frac{\lambda}{\mu}$ where s be the average number of bikes per stations.

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



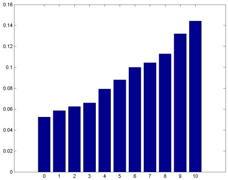
$\rho < 1$

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



$\rho = 1$

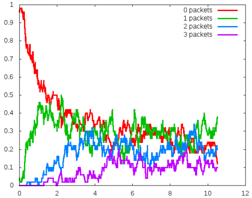
$s > \frac{C}{2} + \frac{\lambda}{\mu}$



$\rho < 1$

Proof based on mean field approximation

$$x_i = \frac{1}{N}\#\{\text{stations with i bikes}\}$$

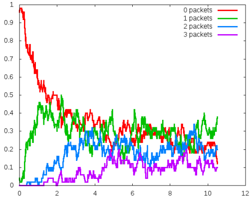


For fixed N, X_i is a complicated stochastic process

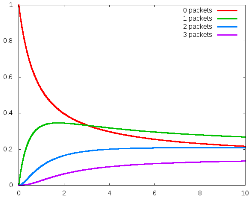
- Reversible process but steady state not explicit.

Proof based on mean field approximation

$$x_i = \frac{1}{N}\#\{\text{stations with i bikes}\} \propto \rho^i$$



$N \to \infty$



For fixed N, X_i is a complicated stochastic process

- Reversible process but steady state not explicit.

System described by an ODE

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

Use mean field approximation [Kurtz 79]

- Study the system when the number of stations N goes to infinity.

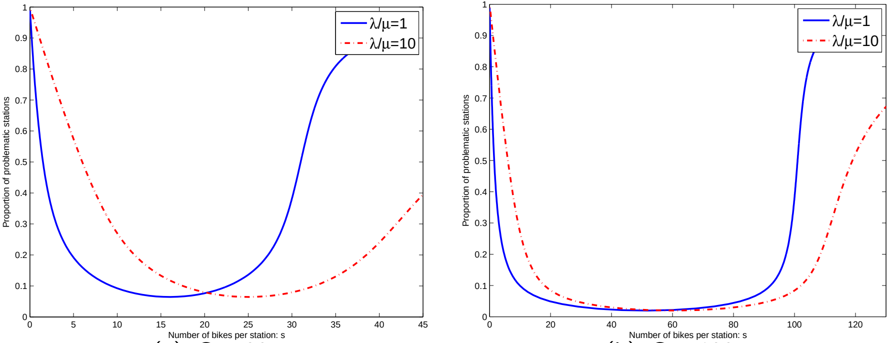
Consequences

Proportion of problematic stations (=empty+full) $x_0 + x_C$ is minimal for

$$\rho = 1 \quad \text{i.e.} \quad s = s_c \stackrel{def}{=} \lambda/\mu + C/2$$

- Prop. of problematic stations is at least $2/(C+1)$ and "flat" at s_c.

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



(a) $C = 30$. (b) $C = 100$.

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

A first rule : "two choices" rule.

Users can observe the occupation of stations.

- Rule : choose the least loaded among 2 stations close to destination to return the bike.

Paradigm known as "*the power of two choices*" :

- Used in balls and bills [Azar et al. 94]
- Supermarket model, server farm [Vvedenskaya 96, Mitzenmacher 96]

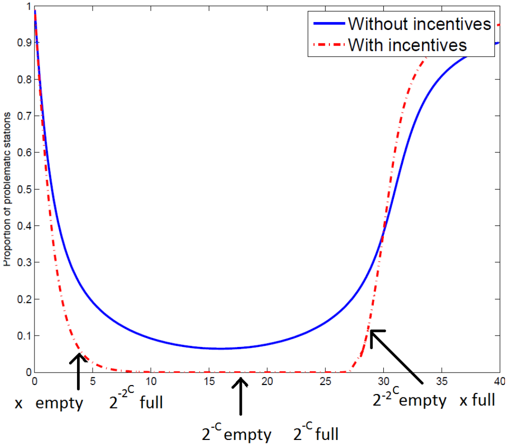
Characteristics of bike-sharing systems :

1. Finite capacity of stations.
2. Local search (choice among neighbors).

Two choices – finite capacity but no geometry

With no geometry, we can solve in close-form.

- Proof uses similar mean field argument.

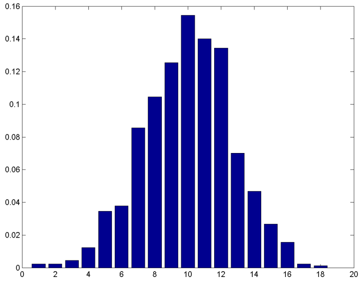


Choosing two stations at random, improves perf. from $1/C$ to 2^{-C}

Two choices – taking geometry into acount

Problem hard to solve : mean field do not apply (geometry) :(.

- Existing results for balls and bins (see [Kenthapadi et al. 06])
- Only numerical results exists for server farms (ex : [Mitzenmacher 96])



We rely on simulation

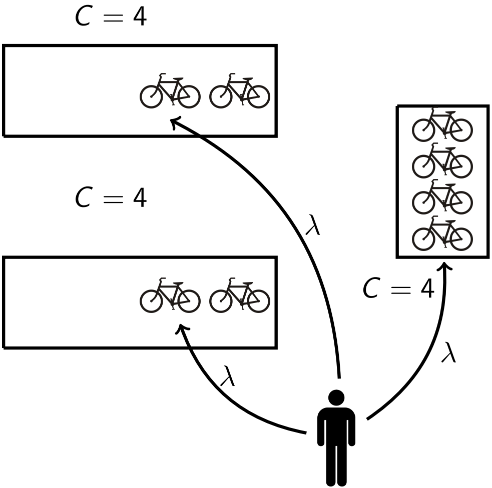
Occupancy of stations
x-axis = occupation of station.
y-axis : proportion of stations.

Recall : with no incentives, the distribution would be uniform.

Empirically :

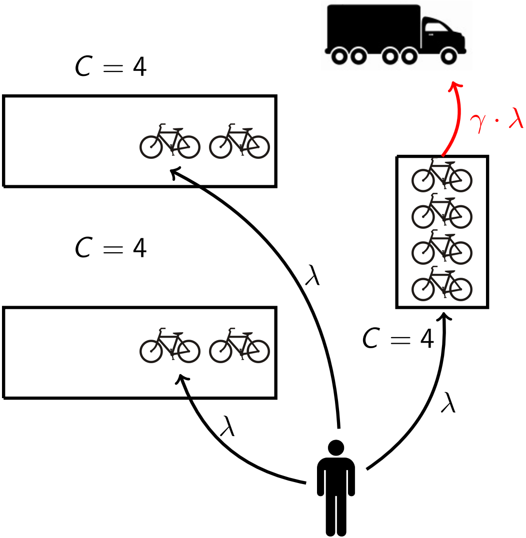
- with geometry 2D : proportion of problematic stations is $\approx 2^{-C/2}$.
  (recall : with no-geometry : 2^{-C}, with no incentive : $1/C$).

Regulation



Same model as before
with a truck

Regulation



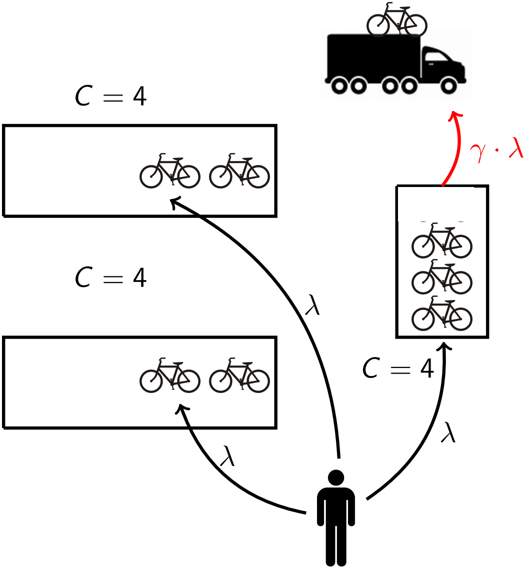
Same model as before with a truck

With rate $\gamma \cdot \lambda$:

- Take a bike from the most loaded.
- Put it in the least loaded.

Question : what should be γ ? 10%, 20%, more ?

Regulation



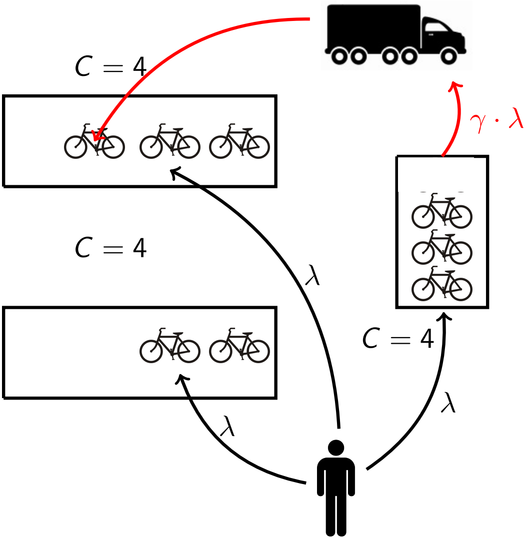
Same model as before with a truck

With rate $\gamma \cdot \lambda$:

- Take a bike from the most loaded.
- Put it in the least loaded.

Question : what should be γ ? 10%, 20%, more ?

Regulation



Same model as before
with a truck

With rate $\gamma \cdot \lambda$:

- Take a bike from the most loaded.
- Put it in the least loaded.

Question : what should be γ ? 10%, 20%, more ?

Optimal rate of regulation

Recall C is the capacity, s the fleet size and N the number of stations.

Theorem

As N goes to infinity, we have :

- *The number of problematic stations decreases as γ increases.*
- *If $\gamma > \frac{1}{2[C-(s-\lambda/\mu)]-1}$, then there is no problematic stations.*

For example : if $s = \frac{C}{2} + \frac{\lambda}{\mu}$, a regulation rate of $1/(C-1)$ suffices.

Proof. Again mean field approximation but with *discontinuous dynamics*

- The dynamical system is described by a differential inclusion

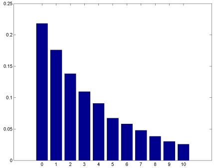
$$\dot{x} \in F(x).$$

- The DI has a unique solution. We can solve in close-form.
  See [Gast Gaujal 2010].

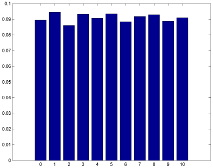
Optimal rate of regulation, illustration

Example : capacity is $C = 10$. Fleet size is 3,5 or 7 bikes/stations.

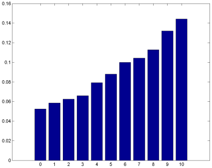
1. No regulation, $\gamma = 0$



$s = 3$

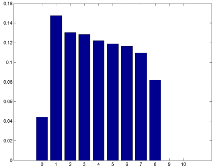


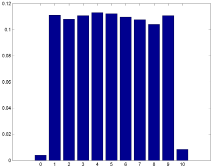
$s = 5$

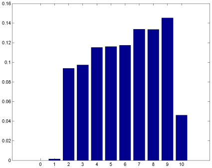


$s = 7$

2. Regulation ($\gamma = 10\%$).







x-axis = occupancy of stations, from 0 to 10.
y-axis = proportion of stations.

Conclusion on the homogeneous model

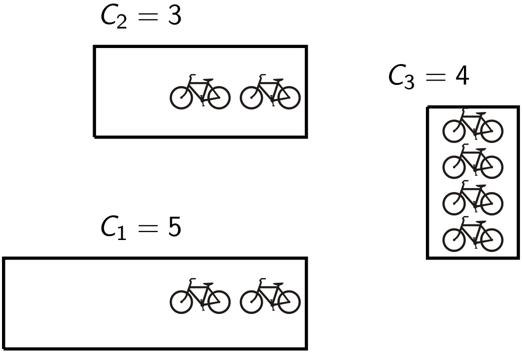
		prop. of problematic stations	ex : $N = 30$
Original model		1/C	6.5%
Two choices	(random)	2^{-C}	$10^{-9} \approx 0$
	(geom)	$2^{-C/2}$	$10^{-4.5}$
Regulation	$\gamma > \frac{1}{C-1}$	0	$\gamma = .032$

However : as mentioned before, there are some important factor :

- time dependent arrival rate : daily period
- heterogeneity : popular or non popular stations
  (housing and working areas, uphill and downhill stations,...)

Outline

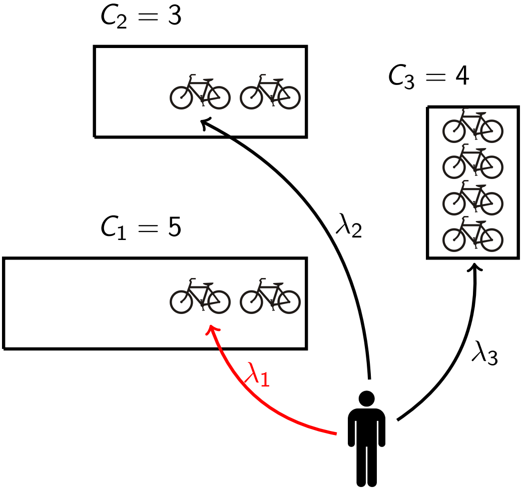
Heterogeneous model



For each station i :

- Fixed capacity C_i

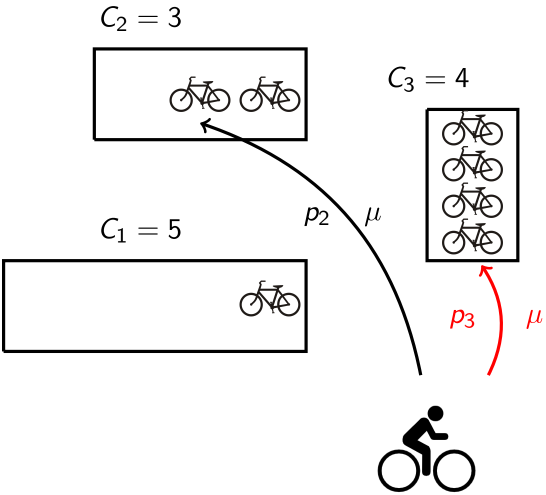
Heterogeneous model



For each station i :

- Fixed capacity C_i
- Arrival rate λ_i.

Heterogeneous model



For each station i :

- Fixed capacity C_i
- Arrival rate λ_i.
- Popularity of station P.
- Travel time : exponential of mean $1/\mu_{ij}$.
- Local search if full.

Steady state performance

There are N stations. Assume that as N goes to infinity, the popularity of the parameters $p_i = (\lambda_i, p_i)$ goes to some distribution.

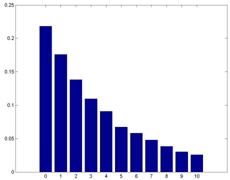
Theorem (Propagation of chaos-like result)

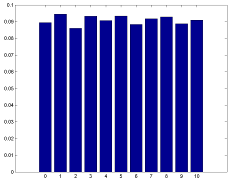
There exists a function $\rho(p)$ such that for all k, if stations $1, \dots k$ have parameter $p_1, \dots p_k$, then, as N goes to infinity :

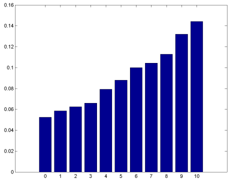
$$P(\#\{\text{bikes in stations j}\} = i_j \text{ for } j = 1..k) \propto \prod_{j=1}^{k} \rho(p_j)^{i_j}$$

Depending on popularity, stations have different behaviors :

Popular start $\rightarrow$ Popular destination







Steady-state performance : numerical example

- In general, ρ is the solution of a fixed-point equation.
- Can be plotted in closed form for particular cases.

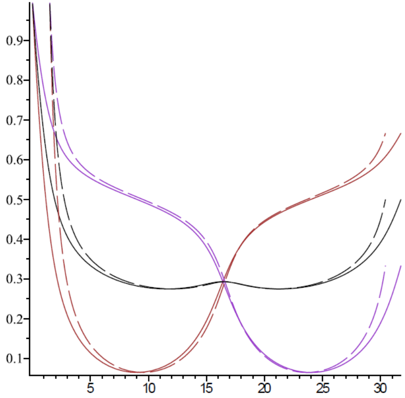


Figure: Two types of stations : popular and non-popular for arrivals : $\lambda_1/\lambda_2 = 2$.

Outline

Current and future work

Good understanding of the symmetric model

- Performance poor : $1/C$ problematic stations (even for symmetric !).
- Simple incentives helps a lot : $2^{-C/2}$.
- Optimal regulation rate is function of capacity : $1/C$.

Current and future work

- Building a realistic model of Paris (using traces).
- Analyze transient and steady-state behavior.
- Difference effect of flows vs random perturbations.
- Develop model to approximate the influence of geometry.