Stochastic Models for Bike Sharing Systems

Nicolas Gast

Inria, Grenoble, France

Journée Thématique: systèmes de véhicules en libre-service, 2019,
Nanterre, France

Outline

Outline

(From far away), the System Looks Predictable

Vélib' Data (Paris) : availability at stations + trips info from September 2013 to December 2014

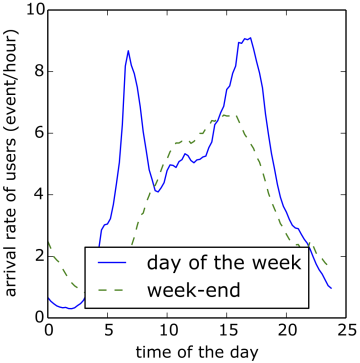
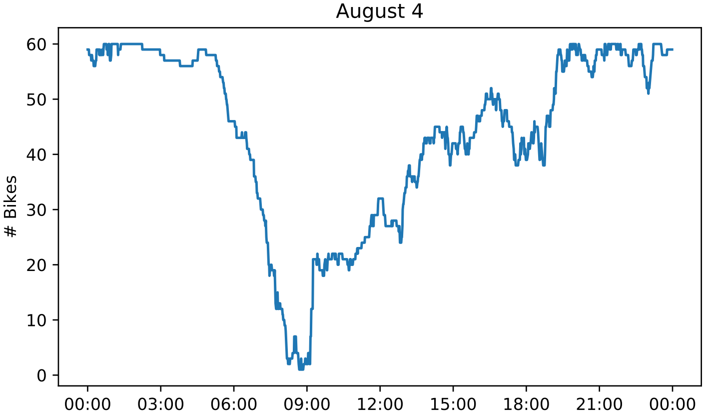
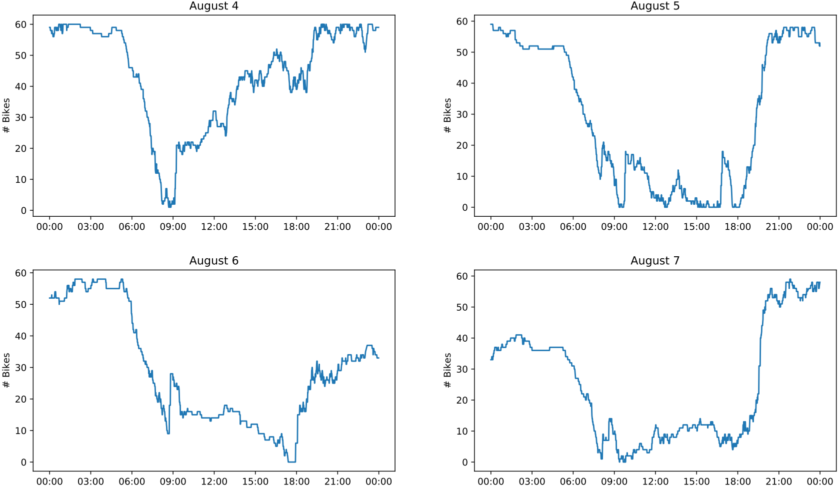


Figure: Evolution of the average departure rate from *Vélib'*stations during the day

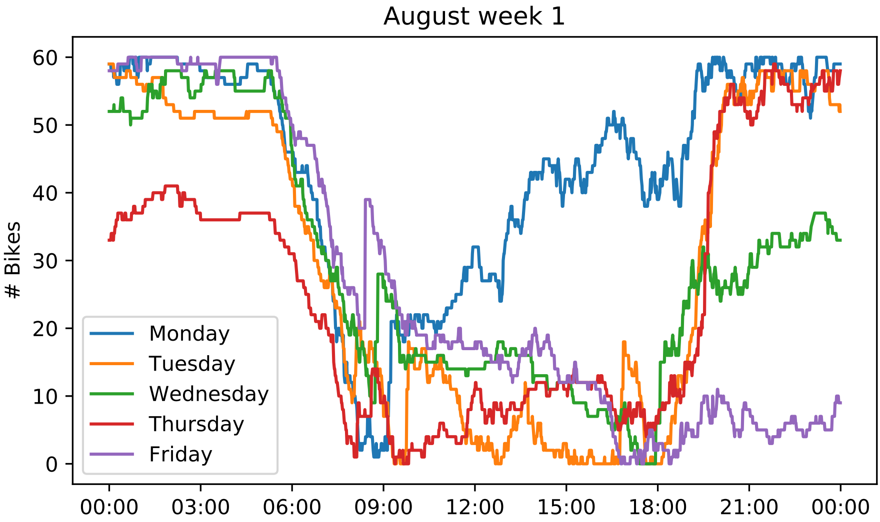
Zoom on one station (Gare de Lyon), August 2014



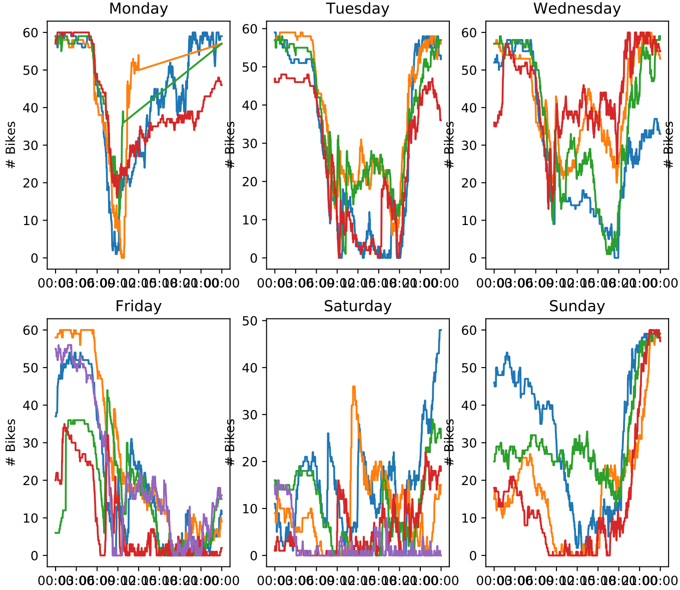
Zoom on one station (Gare de Lyon), August 2014



Zoom on one station (Gare de Lyon), August 2014



Zoom on one station (Gare de Lyon), August 2014



- Regular patterns
- Randomness

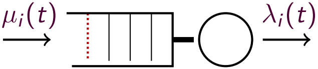
These systems can be viewed as closed queuing-networks

Our model: Demand from station i to station j: Poisson process of intensity $\lambda_{ij}(t) = \lambda_i(t)p_j(t)$.

These systems can be viewed as closed queuing-networks

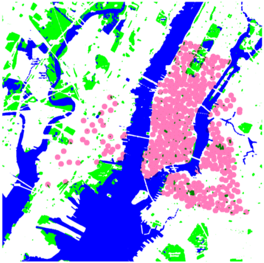
Our model: Demand from station i to station j: Poisson process of intensity $\lambda_{ij}(t) = \lambda_i(t)p_j(t)$.

As **an approximation** (valid when many stations), you can zoom on one station:



How to Build Such a Model? (stations and $\lambda_{ij}(t)$)

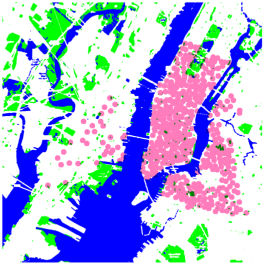
Existing BSS (e.g., NY)



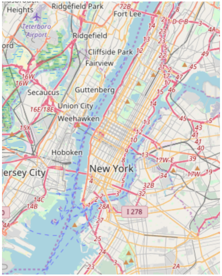
- Record traces
- "**Infer**" demand

How to Build Such a Model? (stations and $\lambda_{ij}(t)$)

Existing BSS (e.g., NY)



No Existing BSS



openstreetmap

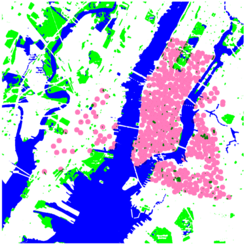
- Record traces
- "**Infer**" demand

- Where will station will be?
- Which traffic flow?

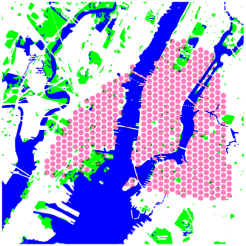
From Spatial Data to Spatial Models

Example of New-York's BSS

Real system



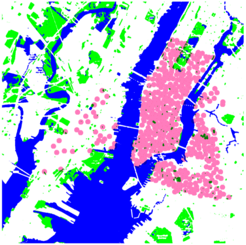
Regular



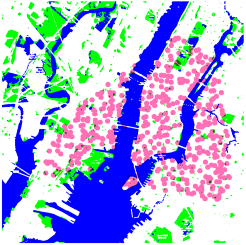
From Spatial Data to Spatial Models

Example of New-York's BSS

Real system

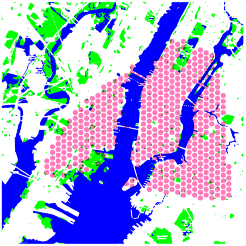


Regular (Noisy)



Poisson

Regular



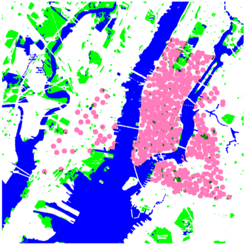
Ginibre

Rating-Weighted

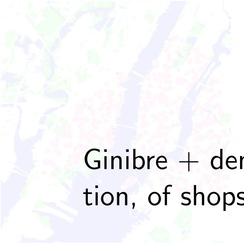
From Spatial Data to Spatial Models

Example of New-York's BSS

Real system



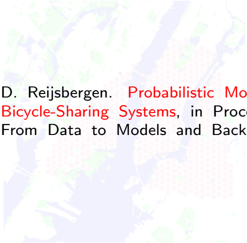
Regular (Noisy)



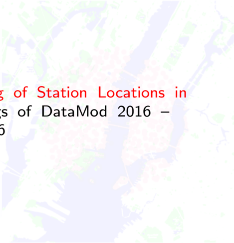
Poisson

Ginibre + densities of population, of shops, . . .

Regular



Ginibre



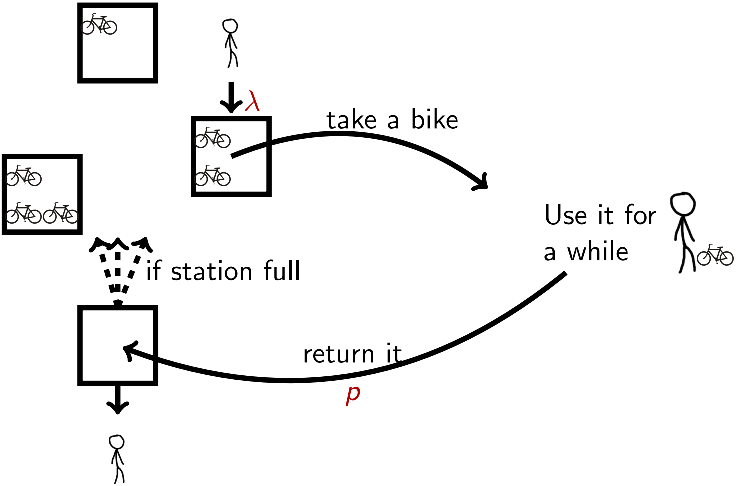
D. Reijsbergen. Probabilistic Modelling of Station Locations in Bicycle-Sharing Systems, in Proceedings of DataMod 2016 – From Data to Models and Back, 2016

Rating-Weighted

Outline

Idealized Scenario: The Homogeneous Model: $\lambda_{ij}(t) = \lambda$.

[Fricker-Gast 14]



Distribution of x_i, the fraction of station with i bikes

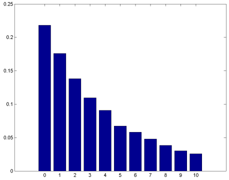
Theorem

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

$$x_i \propto \rho^i.$$

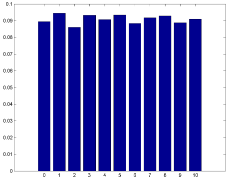
$\rho \leq 1$ *iff* $s \leq \dfrac{C}{2} + \dfrac{\lambda}{\mu}$ *where s be the average number of bikes per stations.*

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



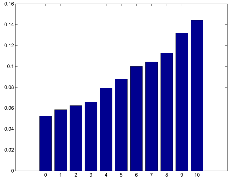
$$\rho < 1$$

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



$$\rho = 1$$

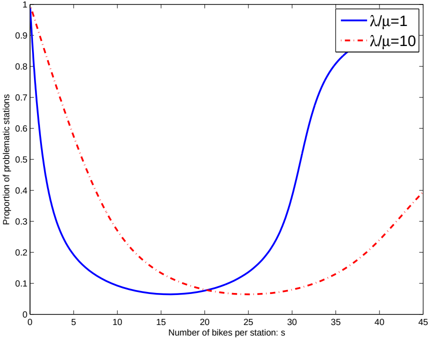
$$s > \frac{C}{2} + \frac{\lambda}{\mu}$$



$$\rho < 1$$

Consequences: optimal performance for $s \approx C/2$

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

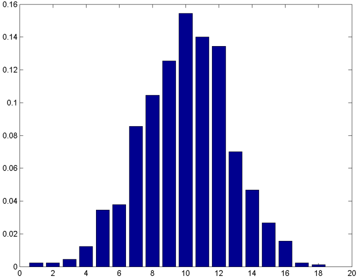


Fraction of problematic stations (=empty+full) minimal for $s = \lambda/\mu + C/2$

- Prop. of problematic stations is at least $2/(C+1)$ (6.5% for $C = 30$)

With a small help from users, everything can be better

Each users returns her bike to most empty of two neighboring stations.



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

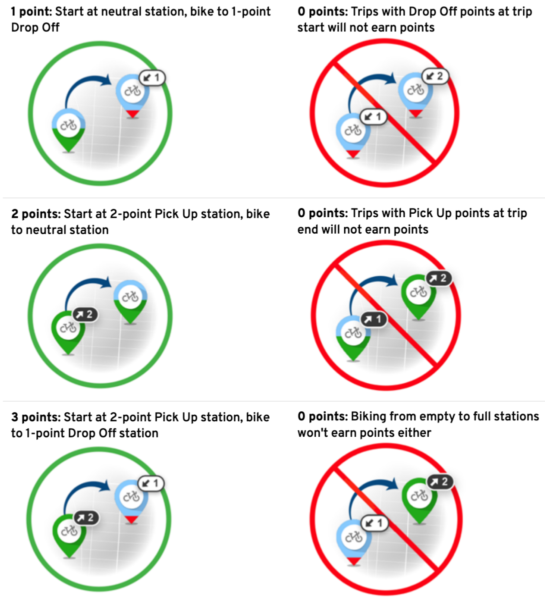
Recall: with no incentives, the distribution would be uniform.

Empirically:

- In a 2D grid, the proportion of problematic stations is about $2^{-C/2}$. (recall: without the help of users: $2/(C+1)$).

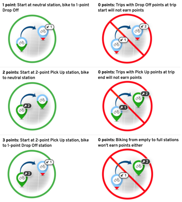
Incentives in Practice: Example of Bike Angels

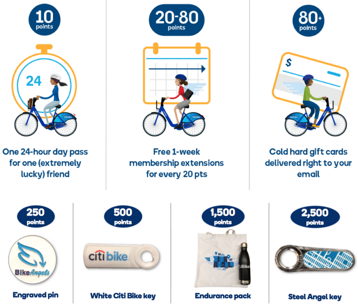
https://www.citibikenyc.com/bikeangels/rewards [Chung et al. 18]



Incentives in Practice: Example of Bike Angels

https://www.citibikenyc.com/bikeangels/rewards [Chung et al. 18]





citi bike — How it Works | Pricing | Station Map | Explore NYC

BIKE ANGELS | POINTS AND REWARDS | MONTHLY LEADERBOARD

Monthly Leaderboard

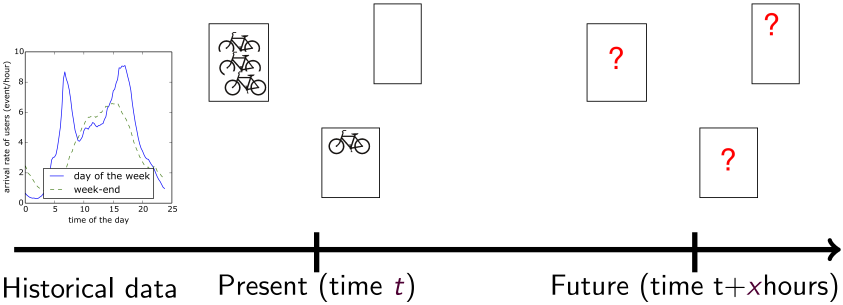
Each month, the highest-earning Angels earn a gift card bonus for helping the highest number of Citi Bike riders. Scores reset at the start of each month, and there's no limit to how many points can be earned.

1st place wins $100, 2nd place wins $75, 3rd place wins $50, 4th and 5th place win $25.

RANK	ANGEL	POINTS
1	JS610	935
2	SR013	685
3	JG855	234
4	LY382	219
5	RP544	191

Outline

Forecasting : what and why?



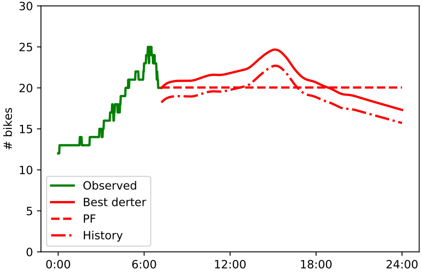
Why forecasting :

- Operator perspective (rebalancing)
- User perspective (will I find a bike?)

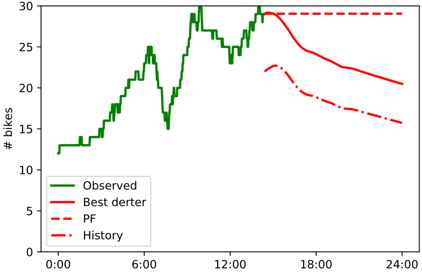
The Traditional approach is to use deterministic forecasts

Examples

- **Last-Value (LVP)** : availability at $t + h$ is equal to availability at t.
- **Historical (HP)** : average availability at this hour (based on historical observations).
- Machine learning tools (ARIMA, Bayesian network,...)



Forecast issued at 7am



Forecast issued at 3pm

Our Model shows that Forecasting Cannot be Good

Quality of forecast=**RMSE**

$$\sqrt{\mathbb{E}\left[\|X_t(t+h) - X(t+h)\|^2\right]}$$

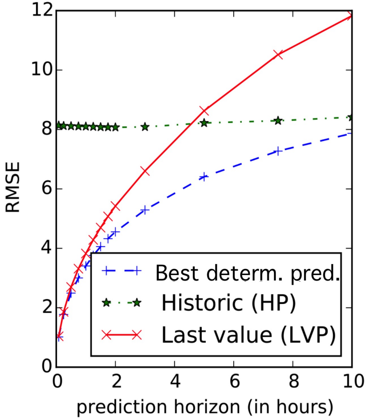


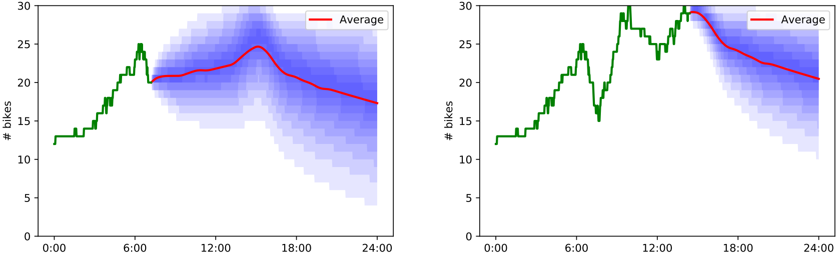
Figure: Comparison of the RMSEs for different predictors.

Lower bounds on error

- 3 bikes for $h = 30$ min
- 5 bikes for $h = 2$h.

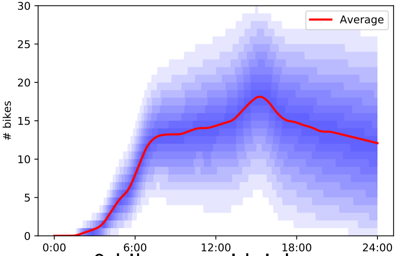
We can go beyond deterministic forecast

By using our queuing model:

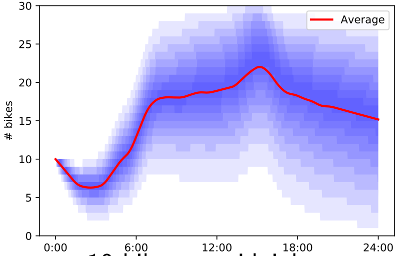


- Use in practice: "what is the probability that I will find a bike?"
- Quality can be evaluated by **scoring rules**

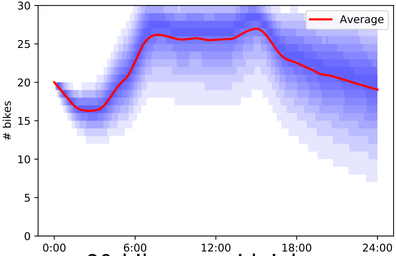
These forecasts can also be use for static redistribution
[Raviv et al 12]



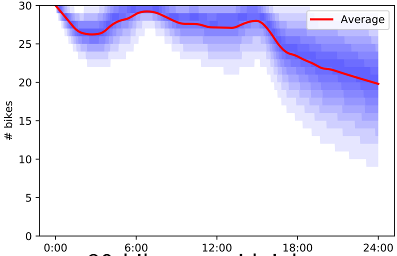
0 bikes at midnight



10 bikes at midnight



20 bikes at midnight



30 bikes at midnight

Outline

Recap

- Bike sharing systems are **probabilistic**.
- We can **build** simple stochastic models.

Applications:

- **Understand** and **design**
  - Without regulation or incentive, performance is poor.
  - System or fleet dimensioning
- System **management**
  - Forecasting
  - Repositioning

Some References

http://mescal.imag.fr/membres/nicolas.gast

nicolas.gast@inria.fr

Reijsbergen 16 *Probabilistic Modeling of Station Locations in Bicycle-Sharing Systems*, D. Reijsbergen, DataMod 2016

Fricker-Gast 14 *Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity.*, C. Fricker and N. Gast. , EJTL, 2014.

Chung et al 18 *Bike Angels: An Analysis of Citi Bike's Incentive Program*, H. Chung, D. Freund and D. Shmoys, COMPASS'18

G. et al 15 *Probabilistic forecasts of bike-sharing systems for journey planning*, N. Gast, G. Massonnet, D. Reijsbergen, and M. Tribastone, CIKM 2015

Raviv et al 14 *Static repositioning in a bike-sharing system: models and solution approaches*, T. Raviv, M. Tzur, I. Forma, EJTL 2014