Public Bike Share (PBS)

IT-driven intelligence for better bikeshare

Felisa Vázquez-Abad
CUNY Institute CoSSMO. Contact: felisav@hunter.cuny.edu
Introduction

Integrating information technology for a self-regulated PBS

Felisa Vázquez-Abad, 2016
Why the hype?
A New Way to Use Bikes

Bikes are good for sustainable cities.

Felisa Vázquez-Abad, 2016
Brief History of Public Bike Share (PBS)

• 1\textsuperscript{st} Gen (1965) Amsterdam White Bikes, many to follow.

• 2\textsuperscript{nd} Gen (1995) Copenhagen “bike library”: borrow with a deposit, many to follow.

• 3\textsuperscript{rd} Gen (1998) Rennes, Lyons: use of “smart card” technology
  • User identification
  • Information, monitoring, etc
How did Vélib achieve success?

- **Access fee:**
  - Annual memberships
  - Short term memberships

- **Usage fees:**
  - Above 30-45 minute ride
  - Geometric progression: $7, $14, etc, $441 for three hours!
Statistics: PBS is here to stay!

Spain leads the world with 132 separate bike-share programs. Italy has 104, and Germany, 43.

The world’s largest bike-sharing program is in Wuhan, China’s sixth largest city, with 9 million people and 90,000 shared bikes. In 2013, China was home to 82 bike-sharing programs, with a combined fleet of some 380,000 bicycles.

In early 2014, some 600 cities in 52 countries host advanced bike-sharing programs with a combined fleet of more than one million bikes.

The United States hosts 36 modern bike-sharing programs. With a number of new programs in the works and planned expansions of existing programs, the U.S. fleet is set to nearly double to over 37,000 publicly shared bicycles by the end of 2014.
Is it working?

Citi Bikes’ Canadian manufacturer files for bankruptcy: report

Montreal-based PBSC Urban Solutions, a nonprofit known as Bixi, has debt of almost 50 million in Canadian dollars, the Montreal Gazette reported.

NEW YORK DAILY NEWS / Tuesday, January 21, 2014, 2:26 AM
Challenges of 3rd Gen PBS

- **Customer dissatisfaction** – lack of availability
- **Financial difficulties** – need for outside funds
- **Safety issues** -- anxiety facing route and time uncertainty may yield to reckless driving.
Customer (dis)satisfaction

Dissatisfaction Index

Proportion of customers that find no bikes when and where they wish to use one.

Redistribution

Carry bikes from depot (or other nearby stations) to replenish starving stations.

Expensive to do continually. Difficult algorithmic problem.
The time limit

Unlimited number of rides, but each with a time limit,

Why do this?

- People learn: find a dock, park and get another bike
- Drivers may rush to park in time, and risk accidents
- Today 33% revenue in US PBS comes from usage fees

Significant revenue at Risk
Negative Impact on Safety
Negative Impact on Availability

Felisa Vázquez-Abad, 2015
Better data

Better use of data

Towards 4G-PBS

Felisa Vázquez-Abad, 2016
Our Vision of the Future for the 4th Gen PBS

Better data: Engage users to provide more data. Implement statistical correction of anomalies.

Better use of data: algorithms for prediction, reservations, and re-balancing. Provide users with more information.
Our Vision for the 4th Gen PBS

- Integration of data analytics (GPS, customer classes, preferences, patterns)
- Humans as users but also as sensors/agents
- Incentivize redistribution by reward systems
- Allow reservations

**Self-regulation:** use information for
- Dock/bike dynamic reallocation
- Maintenance, expansion
- Dynamic yield management
Some Examples

- Classification of stations
- Redistribution
- Reservations
- Economic model
- Pricing strategies
- Simulation
Station Classification

Finding patterns
Creating new knowledge
“Mirror” patterns as origin and as destination.
Patterns Grand Central

Can you guess what is going on?

Felisa Vázquez-Abad, 2016
Patterns Grand Central

Lack of bikes “pulls” demand on nearby stations
Unusual Patterns

Possible situations when pattern is true

- Station isolated and badly light at night
- People use subway in the morning and bike in the evening
- People use bike in the morning and other safer transport at night

Possible data anomaly

- Grand Central is known for being very popular: no bikes when origin, no docks when destination (“valet” parking)
- Historical data: trips taken
- Streaming data: actual situations and occupancy
Machine Learning ++

Training (historical data) versus learning (streaming data)

• Function for automatic detection of unusual patterns.
• Statistical corrections (use streaming and historical data) to assess real demand.
• Classify stations (neural nets for supervised learning, distance function for unsupervised learning).

Automatic detection and classification (for any city!) and classification.

Why is this useful?
Redistribution
Birth and Death Model

- Bike comes in (birth)
- Bike goes out (death)

Model is known as the "gambler’s ruin". Markov chain methods can be used to calculate expected time to depletion.

Using this model, we propose methods for optimizing **when** to call for redistribution to satisfy a given dissatisfaction rate \( \alpha \).
Day-segments: regimes and threshold controls

- **Super-starvation.** Much higher rental demand than parking. Calling the truck when empty is too late.
  - Control: call truck when there are 0 bikes.
- **Starvation.** More rental demand than parking, but we can wait after station is empty.
  - Control: call truck when empty time exceeds 0 minutes.
- **Neutral:** demands “balance” each other.
  - Control: None.
- **Filling.** More parking demand than rentals, but we can wait after all docks are full.
  - Control: call the truck when full time exceeds 0 minutes.
- **Super Filling.** Much higher parking demand than rental. Calling the truck when station is full is too late.
  - Control: call truck when there are 0 free docks.
Theoretical Results

Lemma 1. Assume that $\lambda < \mu$ and $C$ is total capacity. Let

$$\tau_C(k) = \min(s: X(t+s) = 0 | X(t) = k)$$

denote the time until starvation of the station, and $T_C(k) = \mathbb{E}[\tau_C(k)]$, then

$$T_C(k) = A \left( \left( \frac{\mu}{\lambda} \right)^k - 1 \right) + \frac{k}{\mu - \lambda}, \quad A = - \left( \frac{\lambda}{\mu} \right)^C \frac{\lambda}{(\mu - \lambda)^2}.$$ 

$\rho$: delay until truck arrives, from time when call is placed.

$L(\theta)$: expected length of cycle (time between two consecutive re-distributions).

$I(\theta)$: expected time empty during a cycle.

Lemma 2. Let $R(\delta t)$ be the residual empty time during a period of time $\delta t$ starting empty. Under the control policy for SS regime

$$L(\theta) = T_{C_0 - \theta}(C_0 - \theta) + \rho, \quad \text{and} \quad I(\theta) = \mathbb{E}[R(\rho - \tau_{C_0}(\theta))1_{\{\tau_{C_0}(\theta) < \rho\}}].$$

Under the control policy for the S regime

$$L(\theta) = T_{C_0}(C_0) + \lambda \theta T_{C_0}(1) + \theta + \rho, \quad \text{and} \quad I(\theta) = \theta + \mathbb{E}[R(\rho)].$$
Stochastic Optimization

Minimize cost of redistribution subject to a constraint of the form “at least 95% of arriving clients are satisfied”.

\[
\min_{\theta} \left( J(\theta) \overset{\text{def}}{=} \frac{\kappa}{L(\theta)} \right) \quad \text{s.t.} \quad I(\theta) \leq \alpha L(\theta)
\]

- To satisfy the constraint we need to find \( \theta^* \) such that \( I(\theta^*)/L(\theta^*) = \alpha \) (\( \alpha = 0.05 \), for example).

SS or SF: stochastic binary search
S or F: target tracking with stochastic approximation.
Target Tracking

Felisa Vázquez-Abad, 2016
Extensions

Current work: create novel algorithms to include both stochastic models and streaming data.

Use station classification for better predictions.
Reservations
Ranking and Selection
Applied probability models
Ranking and selection

\[ \mathbb{P}(X_d(t) = C_d \mid T_d(o) = t, S_d(t_0, t) = n) \approx \begin{cases} 0, & n \leq C_d - m_d \\ e^{-\lambda_d t} \sum_{k=0}^{m_d+n-C_d} \frac{(\lambda_d t)^k}{k!} m, & n + m_d > C_d \end{cases} \]
Economic Model
**Solution 2:** Economic model for pricing alternatives.

**Ongoing:** stochastic analysis of dynamic system with feedback for optimal balance (economic equilibrium)
Research Team

- **Felisa Vázquez-Abad**: stochastic optimization, probability theory, mathematical models.

- **Michael Fu**: optimization, applied probability.

- **Ted Brown**: simulations, data gathering, apps development.

- **Jason Young**: behavioral psychology, modeling human reactions and behavior

- **Research Assistants**: Silvano Bernabel, Larent Barreaud.

- **Students**: Cynthiaann Bryant, Paulina Toro Izasa, Greg Javens and Elias Morgenroth.
The End