Big Data Analytics for CPS:
A Parking Management and Congestion Reduction Case Study

Eric Mazumdar

with Lillian Ratliff, Baosen Zhang, and Shankar Sastry

Motivation

* Congestion in cities can be reduced by a more efficient use of parking infrastructure

* Inefficient use of parking is costly to urban centers:
  * Past studies have shown that 30%-50% of peak traffic in city centers is people cruising for parking[1]

* Adopting 'smart' technologies is a potentially lucrative solution:
  * Total market for smart parking in North America and Europe was $7 billion in 2014 and set to grow to $43 billion by 2025[2]

* Seattle Department of Transportation (SDOT) has asked us to come up with control schemes using data and technology to increase revenues and reduce congestion in the city center

[1]Millard-Ball, Weinberger, Hampshire (2014)
[2]Barua (2015)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Problem Statement

* Working with SDOT we aim to:
  * Gain insights from Seattle's extensive parking data
  * Build accurate data-driven models of parking to develop ways to use established parking infrastructure in urban centers more efficiently
  * Propose methods to achieve SDOT's desired level of 70%-90% occupancy.
    * Occupancy is defined as the time average of the fraction of occupied spots divided by the total number of spots[1].

$$Occupancy = \frac{1}{T}\int_0^T \frac{Occupied\ spots(t)}{Total\ spots(t)}dt$$

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

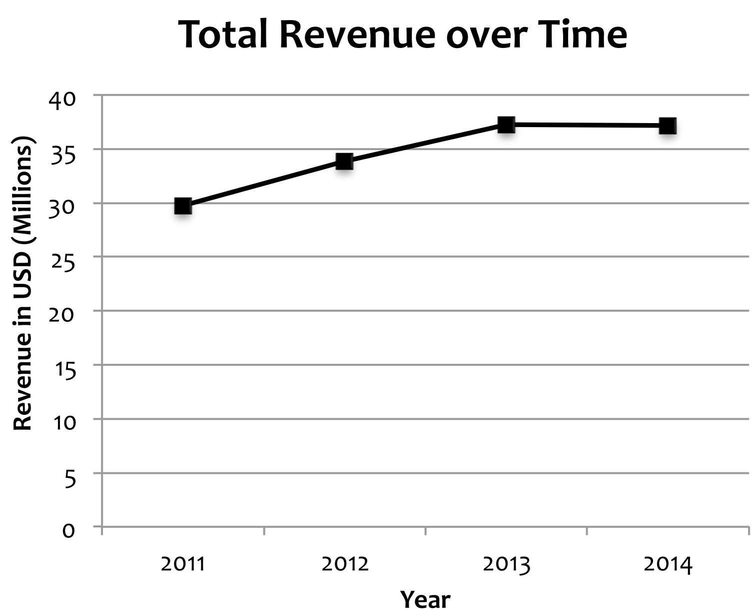
[1]SFPark: Sensor Independent Rate Adjustments (SIRA)

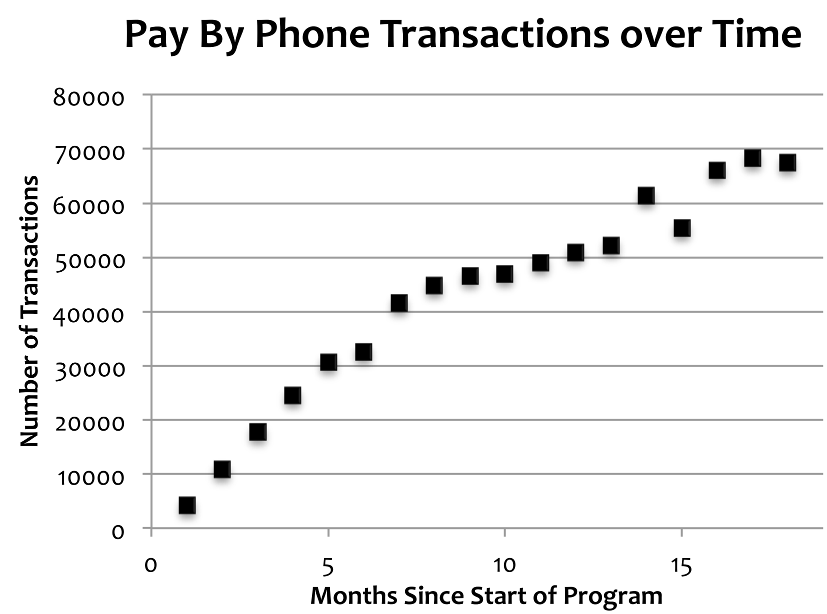
Seattle Parking Data

* We have obtained street-parking data from SDOT. It includes:
  * Geographical positions of pay-stations and the number of parking spots per blockface, where a blockface is one side of a city block
  * 5 years of Pay Station transactions (~45.5 million transactions)
    * This includes: pay-station identifier, time of transaction, the amount paid, and the duration paid for
    * Note: we only have pay-station level granularity in the data
  * 2 years of Pay-by-Phone transactions (~800,000 transactions)
    * This includes: pay-station identifier, time of transaction, the amount paid, and the duration paid for
    * Note: no user identifying information
  * Yearly manual survey of Parking Spots from years 2012-2014
    * One day per year of manually recorded parking transactions including: pay-station identifier, number of cars parked on a blockface at a given time, supply of parking spaces on a blockface at a given time

Initial Insights from SDOT Data

* Parking accounts for $37.2 million in revenue in 2013 and 2014
* Steady increase in Pay-By-Phone Transactions since its inception





FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Initial Insights from SDOT Data

- A first pass at the data was to look at the load on a blockface over time to identify any meaningful trends in the data
  - We define the load on a blockface *i* at time *k* as the number of cars parked on *i* at time *k* divided by the total number of available parking spots on *i*

$$Load_i[k] = \frac{Occupied\ spots_i[k]}{Total\ spots_i[k]}$$

  - Since we only know the arrival time of a car on a blockface *i* and an upper limit to the duration of its stay, we approximated the load on *i* over time by the following equation:

$$Load_i[k] = \frac{n_i * Load_i[k-1] + Arrivals_i[k] - Departures_i[k]}{n_i}$$

Where: - n_i is the maximum number of available parking spots on blockface *i*

-$Arrivals_i[k]$ is the number of new cars parked on blockface *i* at time *k*

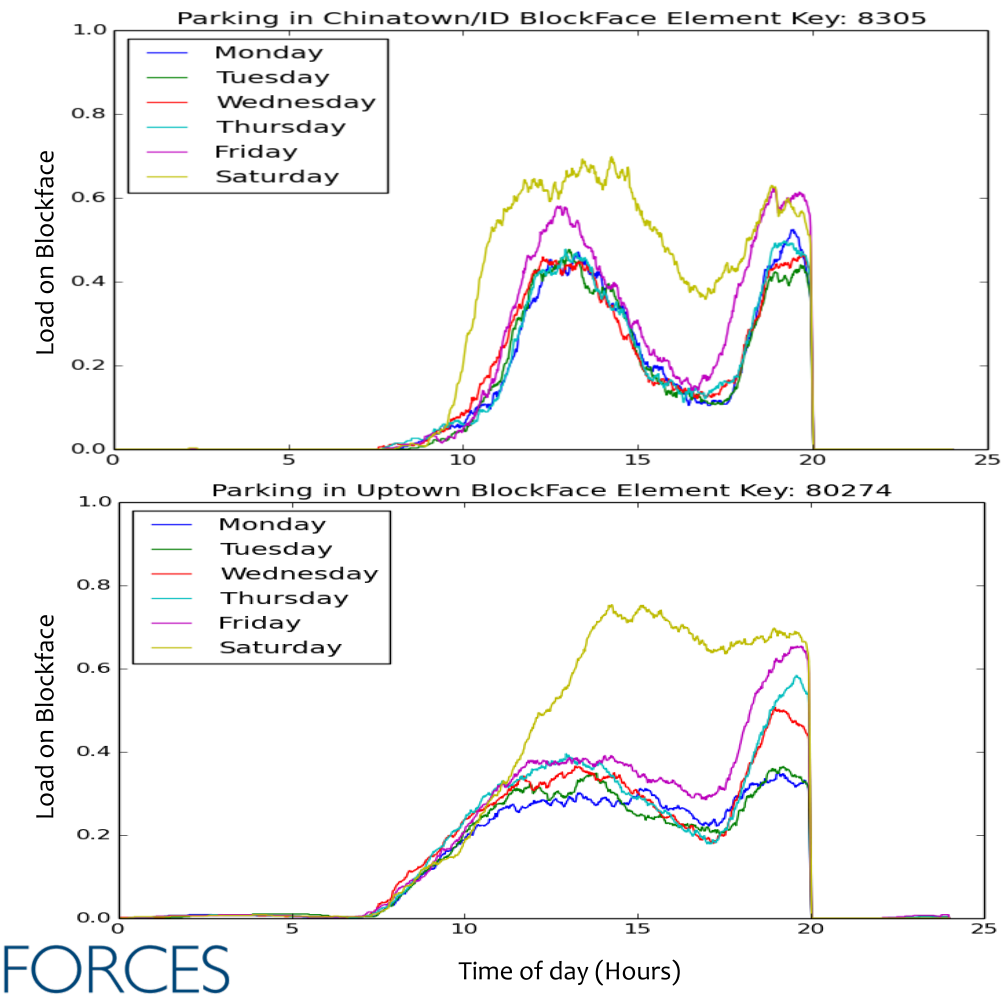
-$Departures_i[k]$ is the number of cars leaving blockface *i* at time *k*

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Initial Insights from SDOT Data

Results from load calculations:

* Parking in Seattle is well below its full capacity
* Different areas have different parking profiles
* 50% mean occupancy over all blockfaces between 10 am and 4 pm in Seattle
    * occupancy is the time average of the load



Future Work: Modeling Parking

* Building on this preliminary analysis, we aim to build a spatio-temporal model of parking in the city

* Past work in parking modeling in cities has relied on instrumenting a limited number of parking spots and measuring the total number of spaces available at a given time and the exact length of stay[1].

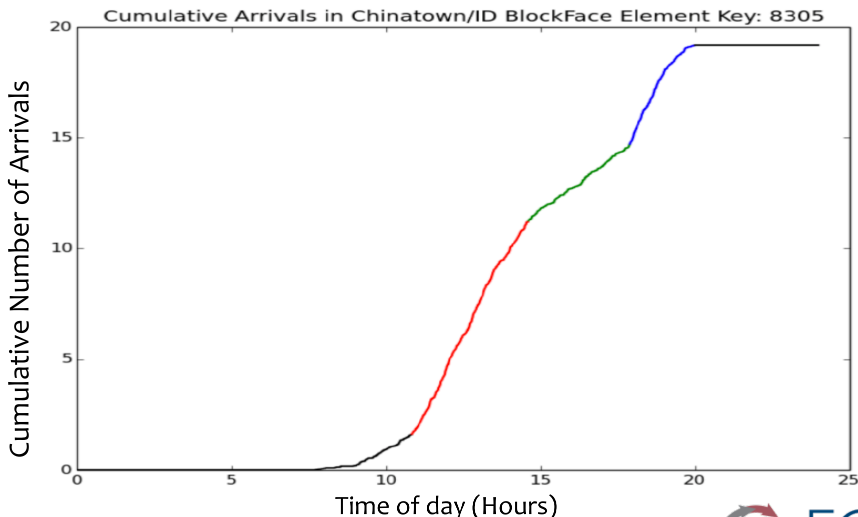
* Modeling occupancy in Seattle is challenging due to:
  * The lack of the "ground truth".
  * The lack of key data that would make construction of a deterministic model possible.
    * i.e . exact duration of parking session per transaction, number of available parking spots per blockface on a given day, the total number of cars looking for parking (not just cars that successfully parked)

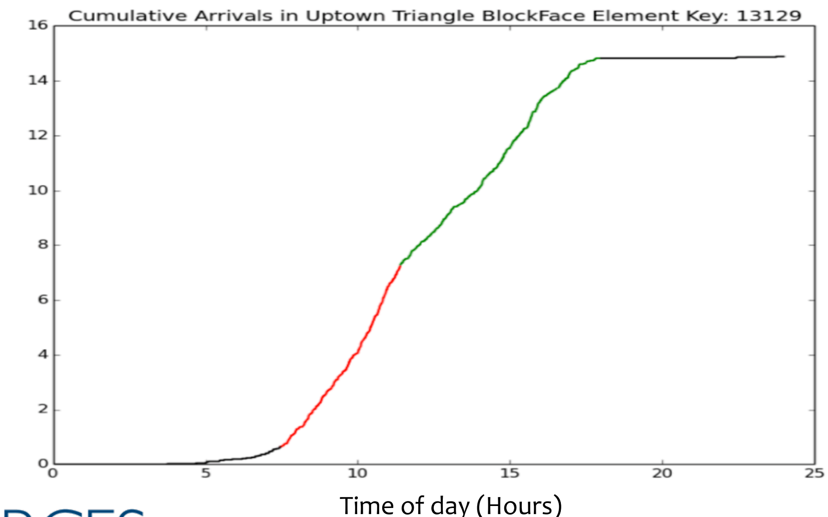
FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

[1]SFPark: Sensor Independent Rate Adjustments (SIRA)

Future Work: Modeling Parking

- Model parking on the blockface level as a M/G/n queue with arrivals modeled as a Poisson Process within time intervals
  - Each blockface *i* has:
    - parameter λ_{ik}: the rate of arrival of new cars during period *k*. This rate is constant during a period, but varies over different periods. We can estimate λ_{ik} from the data[1,2]
    - A prior distribution for the duration of a parking session drawn from the data
    - *n* servers representing the n available parking spots on the blockface





[1]Kuhl, Damerdji, Wilson (1998)
[2]Massey, Parker, Whitt (1995)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Future Work: Modeling Parking

* Benefits of modeling parking as a queue:
  * Each block is parameterized by its arrival rate, a distribution for service times, and the number of spots
  * Model is easily scaled up to larger granularities

* Design questions:
  * How to deal with new arrivals once the queue is at capacity
    * Blocking, push to neighboring queue, etc.

* Model verification:
  * SF-Park Dataset which has the "ground truth" data
  * SDOT's yearly manual survey of parking spots

Further Directions

* Use the parking model to:
  * Evaluate the difference between the public's optimal consumption of parking resources and the city's
    * Develop incentive schemes to push consumers towards equilibriums that are efficient from the perspective of the city and the consumer
  * Find optimal locations for new parking infrastructure
  * Inferring impact of inefficient parking usage on traffic congestion
  * Look at ways to maximize revenue for SDOT

Thank You

Questions?

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS