Applications through human mobility lens and around…

Vsevolod Salnikov
Background

Moscow State Lomonosov University
(2007-2012)
Math department, chair of differential geometry and application

Écoles normales supérieures - Lyon
(2011-2012)
Master (2) de Systemes Complexes (IXXI)
PhD

2012 — 2015
in the Networks Group
of Mathematical department

of UNIVERSITÉ DE NAMUR
Research interests

- Complex systems
- Networks
- Data mining
- Mobility
- High - Tech, Robotics
- Electronics
- Computer and mathematical modeling
human mobility
Data acquisition
CHECK-IN BASED

+ Broad/diverse (but biased) public
- Sampling: the person makes a check-in only (?) where he wants

«RADIO» SENSORS

+ Possibility to achieve high accuracy
- Limited spatial coverage and number of individuals involved

GPS TRACKERS

+ Very good sampling (both spatial and temporal)
- High power consumptions and limited number of individuals

SMARTPHONE TRACKING

+ Different data types
- High price and limited number of individuals

QUESTIONNAIRES

+ Possible to get more meta data
- «Sampling»
LFG 1.0: the app

- Fast way to save your train delays and thus get compensation from SNCB
- Free, but surveys a person and uploads his tracks to the server
LFG 2.0 — Late Train

Tableau de bord

De
Aalter

A
Aalst-Kerrebroeck

Heure
17:23 -> 18:26

Ajouter un retard
Mes retards

Retard de plus de 15 min:
11

Retard de plus de 30 min:
10

RÉSUMÉ

DÉTAILS

1 17/12/14 Namen -> Brussel-Luxemburg 31 min
2 17/12/14 Namen -> Brussel-Luxemburg 33 min
3 17/12/14 Namen -> Brussel-Luxemburg 30 min
4 17/12/14 Brussel-Luxemburg -> Namen 32 min
5 17/12/14 Brussel-Luxemburg -> Namen 33 min
6 17/12/14 Brussel-Luxemburg -> Namen 32 min
7 17/12/14 Brussel-Luxemburg -> Namen 32 min
8 17/12/14 Brussel-Luxemburg -> Namen 35 min
LFG: Users

[Graph showing the number of events (cum) over time from 11/01/12 to 05/01/13, with two lines representing delays and registrations.]
LFG: delays heatmap
LFG: DELAYS HEATMAP

8-9 a.m.

9-10 p.m.
LFG: How Belgium works
Estimating the price of urban transport

Deterministic formula, based on distance and duration

Tariff 1

For any hiring during Monday to Friday between 06:00 and 20:00, other than on a public holiday:

- For the first 252.4 metres or 54.2 seconds (whichever is reached first) there is a minimum charge of £2.40
- For each additional 126.2 metres or 27.1 seconds (whichever is reached first), or part thereof, if the fare is less than £17.40 then there is a charge of 20p
- Once the fare is £17.40 or greater then there is a charge of 20p for each additional 88.5 metres or 19 seconds (whichever is reached first), or part thereof
Estimating the price of **TAXI** transport

**Deterministic formula, based on distance and duration**

**Fluctuations in price are driven by fluctuations in traffic**
Estimating the price of urban transport... in a deregulated world

Deterministic formula (UberX in NYC is $2.15/mile + 40 cents/minute).

but also increased price variability, by the law of offer and demand

Prices changing in real time in a manner that is hard to predict

The knowledge and transparency on pricing in urban transport is lost.

Uber's Prices Surged in Sydney During the Hostage Crisis, and Everyone Is Furious
What about flights?

Complex pricing strategies

How The Prices Change
How the average cost of a domestic and international flight changes by the day:

- **International ticket price**
  - Lowest price: $1,004 (171 days before departure)
  - Highest price: $1,961

- **Domestic ticket price**
  - Lowest price: $402 (57 days before departure)
  - Highest price: $698

300 days before flight
What about flights?

... but price comparison tools to help users

<table>
<thead>
<tr>
<th>Flight Details</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>79,30 €</strong></td>
<td>Vol aller</td>
</tr>
<tr>
<td>17:00</td>
<td>17:10</td>
</tr>
<tr>
<td>Brussels National, Brux...</td>
<td>Gatwick, Londres</td>
</tr>
<tr>
<td>easyJet</td>
<td>1h 10'</td>
</tr>
<tr>
<td>Détails</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flight Details</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 637,06 €</strong></td>
<td>Vol aller</td>
</tr>
<tr>
<td>17:20</td>
<td>20:35</td>
</tr>
<tr>
<td>Brussels National, Brux...</td>
<td>Heathrow, Londres</td>
</tr>
<tr>
<td>4h 15'</td>
<td></td>
</tr>
<tr>
<td>Détails</td>
<td></td>
</tr>
</tbody>
</table>
How these prices compare

Williamsburg to East Village

- UberX: $15
- Old UberX: $19
- Taxi: $16

Keep in mind

These prices are only in effect for a limited time. The more you ride, the more likely we can keep them this low!

We know you may be asking yourself how this affects our partner drivers. What we’ve seen in cities across the country is that lower fares mean greater demand, lower pickup times and more trips per hour – increasing earning potential and creating better economics for drivers. What does what mean in the long run? They’ll be making more than ever!
Freedom of Information Act (United States)

From Wikipedia, the free encyclopedia

This article is about the U.S. federal law. For freedom of information in the fifty U.S. states, see Freedom of information in the United States.

The Freedom of Information Act (FOIA), 5 U.S.C. § 552, is a federal freedom of information law that allows for the full or partial disclosure of previously unreleased information and documents controlled by the United States government. The Act defines agency records subject to disclosure, outlines mandatory disclosure procedures and grants nine exemptions to the statute.[1][2] This amendment was signed into law by President Lyndon B. Johnson, despite his misgivings,[3][4] on July 4, 1966, and went into effect the following year.[5]

As indicated by its long title, FOIA was actually extracted from its original home in Section 3 of the Administrative Procedure Act (APA). Section 3 of the APA, as enacted in 1946, gave agencies broad discretion concerning the publication of governmental records. Following concerns that the provision had become more of a withholding than a disclosure mechanism, Congress amended the section in 1966 as a standalone act to implement "a general philosophy of full agency disclosure." The amendment required agencies to publish their rules of procedure in the Federal Register, 5 U.S.C. § 552(a)(1)(C), and to make available for public inspection and copying their opinions, statements of policy, interpretations, and staff manuals and instructions that are not published in the Federal Register, § 552(a)(2). In addition, § 522(a)(3) requires every agency, "upon any request for records which ... reasonably describes such records" to make such records "promptly available to any person." If an agency improperly withholds any documents, the district court has jurisdiction to order their production. Unlike the review of other

Freedom of Information Act

Long title

An Act to amend section 3 of the Administrative Procedure Act, chapter 324, of the Act of June 11, 1946 (60 Stat. 238), to clarify and protect the right of the public to information, and for other purposes.

Acronyms

FOIA

Nicknames

Public Information Act of 1966
Public Information Availability

Enacted by

the 89th United States Congress

Effective

July 5, 1967

Citations

Public law

89-487

Statutes at Large

80 Stat. 250

Codification

Acts amended

Administrative Procedure Act

Titles amended

THE NEW YORK CITY TAXI DATASET

FOILing NYC’s Taxi Trip Data

Freedom of Information Law

2013 Trip Data, 11GB, zipped!

2013 Fare Data, 7.7GB

Idea: Uber Vs Yellow Taxi Price Comparison.
THE EXPERIMENT

1. For every trip in NYC taxi dataset.
2. Record origin & destination coordinates.
3. Retrieve total fare paid.
4. Query Uber API price for the same trip.
5. Compare yellow taxi VS uber prices.
Figure 2: Distribution of prices per journey for Uber X and Yellow Taxis in New York City.
uber more expensive for short trips
Most taxi movements are within a short distance range with longer movements occurring less frequently in the data.
Yellow taxi data historic (2013)

We have not taken into account the hour of the day/week when querying the API.

Anomalous events, traffic and weather conditions can have an effect on prices.

But

Last major increase in fares was in 2012, after 8 years, +17%.

We have taken at least seasonal changes controlling for the month of Uber data collection.

That’s why we have built an app that will hopefully incorporate a less biased sophisticated model in the future.
Application Logic: Electing The Cheapest Provider

1. Given as input a pair of origin destination cells (cell size ~ 30mx30m)

2. Calculate yellow cab price averaging across all journeys that fall in the cell.

3. Query the Uber API in real time given as input the geo-coded origin and destination addresses. Uber returns [min, max] estimate.

4. Compare Uber X (cheapest Uber) with the Yellow Cab Price.
Super-interesting. Data Mining Reveals When A Yellow Taxi Is Cheaper Than Uber
bit.ly/1BQO158 Nice work, @openstreetcab!

MIT Tech Review

Data Mining Reveals When a Yellow Taxi Is Cheaper Than Uber

Computer scientists have compared a vast dataset of Yellow Taxi fares in New York City against Uber prices for the first time.
This App Tells You When Uber Is Cheaper Than A Yellow Cab

Getting around New York City is expensive as are most things in New York City.

Data Mining Reveals When a Yellow Taxi Is Cheaper Than Uber

Computer scientists have compared a vast dataset of Yellow Taxi fares in New York City against Uber prices for the first time.
OpenStreetCab

OpenStreetCab London

OpenStreetCab New York

Take Uber!

Trip Origin
Place in London

Destination
Place in London

Uber or Black

Take a Yellow cab!

Uber
£0

Black cab
£13

Saving
£5

The above prices are estimates. Paid more? Send us your feedback!

Share
Feedback

Share
Feedback

The above prices are estimates. Paid more? Send us your feedback!

← Go Back

← Go Back
And later on
> 5K only in New York (~8K downloaded the app)

> 14000 search queries generated
Reverse-engineering the Surge algorithm

Deviation from the baseline price

High frequency of surge price (almost one in four times): surge is not a rare event
Spatial impact on Surge: importance of the origin

Correlation between the time series equal to Pearson’s $r = 0.96$
Ranking areas (Surge)

Predictable!

User feedback

Instead if just the winner, can you also show the price details? (Anonymous User, New York)

Love the app..!! Would be useful to have an estimate of the price difference between taxi and uber when the recommendation is given. (Anonymous User, New York)

Please expand outside New York (Justin, U.S.A.)

From 569 Lexington Ave to Laguardia Airport. Your estimate: 38 Actual: 52 It looked like 38 was about correct for mileage and time but going to the airport they added a fee, plus tolls. (Haley, User, New York).
Your taxi fares estimate always seems high. There are a few regular journeys I inputted and did the same with quite a few jobs and every time I have come in under your taxi estimate.

3 examples were: New Kings Rd, SW6 to Grovesnor Crescent, SW1, usually £9, your app says 11 £, Belgrave Sq, SW1 to Heathrow Terminal 5 usually £55, your app says 60 £, New Kings Rd to Canary Wharf usually £38 when your app says 40 £. (Ross, Driver, London).

I find your estimated prices very inaccurate with regards to black cab. Journey from Kings St to W12 9ba costs £8 in black cab instead of £19 that you estimated. This app favors uber. (Goran, Driver, London).
Another thing I thought of is to maybe let people know that these estimates are for the Taxi day rate (rate 1) and for Uber without a surge price in effect.

For taxis there are 3 rates, rate 1 from 06:00-20:00, rate 2 from 20:00-22:00 and rate 3 from 22:00-06:00. To be honest there isn’t much difference between rates 1 and 2, but rate 3 does make it a bit more expensive.

(Ross, Driver, London).

I compared your estimates with about 10 jobs I done today and your pretty much spot on. A couple were pounds 1-2 over, a couple were under and some were on the button!
Radipole rd sw6 to 148 Harley st when I get there I will streak naked through London if my meter agrees with your £29 estimate! Will let u know, guess what I'm here and keeping my clothes on £22 in the real world!! Exactly the same as uber but twice as quick!

Mark, black cab driver
Route Tracking

trip id: 89
uber price: 9.000000
taxi price: 14.000000

Start trip
Uber  Black

price paid:
0

Stop trip

Back
Experiment

29 journeys

Pick Uber near Black Cab ranks or use Hailo!

3 days: 11 am - 11 pm

> 300 km covered

Budget: 1.5K - 2K GBP

<table>
<thead>
<tr>
<th>Provider</th>
<th>Max Abs Diff</th>
<th>Mean Abs Diff</th>
<th>Std Abs Diff</th>
<th>Max % Dev</th>
<th>Mean % Dev</th>
<th>Std % Dev</th>
<th>Pearson’s ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Cab</td>
<td>4.4</td>
<td>0.06</td>
<td>1.96</td>
<td>0.45</td>
<td>0.15</td>
<td>0.11</td>
<td>0.81</td>
</tr>
<tr>
<td>Black No feedback</td>
<td>3.5</td>
<td>-1.04</td>
<td>2.01</td>
<td>0.59</td>
<td>0.18</td>
<td>0.16</td>
<td>0.81</td>
</tr>
<tr>
<td>Uber X</td>
<td>3.31</td>
<td>0.03</td>
<td>1.08</td>
<td>0.32</td>
<td>0.10</td>
<td>0.06</td>
<td>0.83</td>
</tr>
</tbody>
</table>
Time is Money

- Time Gain [minutes]
  - Frequency
  - Median: 0.354838709677

- Price Gain [GBP]
  - Frequency
  - Median: 0.1
Urban Complexity & Performance

- Pink: Uber faster
- Black: Black Cab faster
- Yellow: Journey duration tie

*Trip Density* = \( \frac{1}{|T|} \sum_{i=1}^{|T|} \frac{P(x = lng_i, y = lat_i, r = 200m)}{\pi r^2} \)
Drivers: Black Cab VS Yellow

uses his (big) brain

Does not know where is Brooklyn!

Location Traces from OSC Users
Driver view

Users receive updates about activity at ranks in real time
1. Draw a polygon using the draw tools.
2. Add a polygon name by double clicking.
3. Click ‘Save’ when you are finished.

Admin view — Specify areas of interest
Admin view: live activity

In the 'Rank Live' section you can obtain a real time overview of the cards ranking. Use the Rank Live to help your fleet coordinate the effort and maximize business.

Live Cards Ranking

1. Feeder 12 12

2. Station 12 12

3. Drummer 12 12

4. St. Andrew 12 12

5. Parkside 12 12
## List of Drivers: Cambridge Drivers

### Logged-in users:

<table>
<thead>
<tr>
<th>UID</th>
<th>Code</th>
<th>App version</th>
<th>Manage</th>
</tr>
</thead>
<tbody>
<tr>
<td>382</td>
<td>fmj35ud7Foe2</td>
<td>AndroidUnknown</td>
<td></td>
</tr>
<tr>
<td>406</td>
<td>svlov39vvn3inf</td>
<td>AndroidUnknown</td>
<td></td>
</tr>
<tr>
<td>192</td>
<td>3ru ej0ichj38ij</td>
<td>Android_1.17</td>
<td></td>
</tr>
<tr>
<td>035</td>
<td>3ryn9e9f928</td>
<td>AndroidUnknown</td>
<td></td>
</tr>
<tr>
<td>176</td>
<td>un85 h348f9u</td>
<td>AndroidUnknown</td>
<td></td>
</tr>
<tr>
<td>015</td>
<td>fe27gve773dg</td>
<td>Android_1.18</td>
<td></td>
</tr>
<tr>
<td>267</td>
<td>38hjce913du9</td>
<td>Android_1.20</td>
<td></td>
</tr>
<tr>
<td>097</td>
<td>0qabo1ej9h91</td>
<td>AndroidUnknown</td>
<td></td>
</tr>
<tr>
<td>059</td>
<td>bknkbn2039nz</td>
<td>Android_1.18</td>
<td></td>
</tr>
<tr>
<td>312</td>
<td>2vgj201bci29jw</td>
<td>AndroidUnknown</td>
<td></td>
</tr>
<tr>
<td>306</td>
<td>e9qudh1hjd18</td>
<td>Android_1.18</td>
<td></td>
</tr>
<tr>
<td>124</td>
<td>ecjg10j1imdm</td>
<td>AndroidUnknown</td>
<td></td>
</tr>
</tbody>
</table>

Admin view — user management
THANK YOU!

Questions?

vsevolod@salnikov.be

@vsevolodNS