Can big data contribute to improve our quality of life in big, congested cities?

Marcela A. Munizaga, Mauricio Zúñiga ,
Felipe Hernandez, Jacqueline Arriagada
Universidad de Chile

Passive data sources

- GPS
- Mobile phone traces
- Ticket sales records
- Toll payment
- Smartcard fare payment
- Apps, social networks
- Sensors
- …

Oportunities

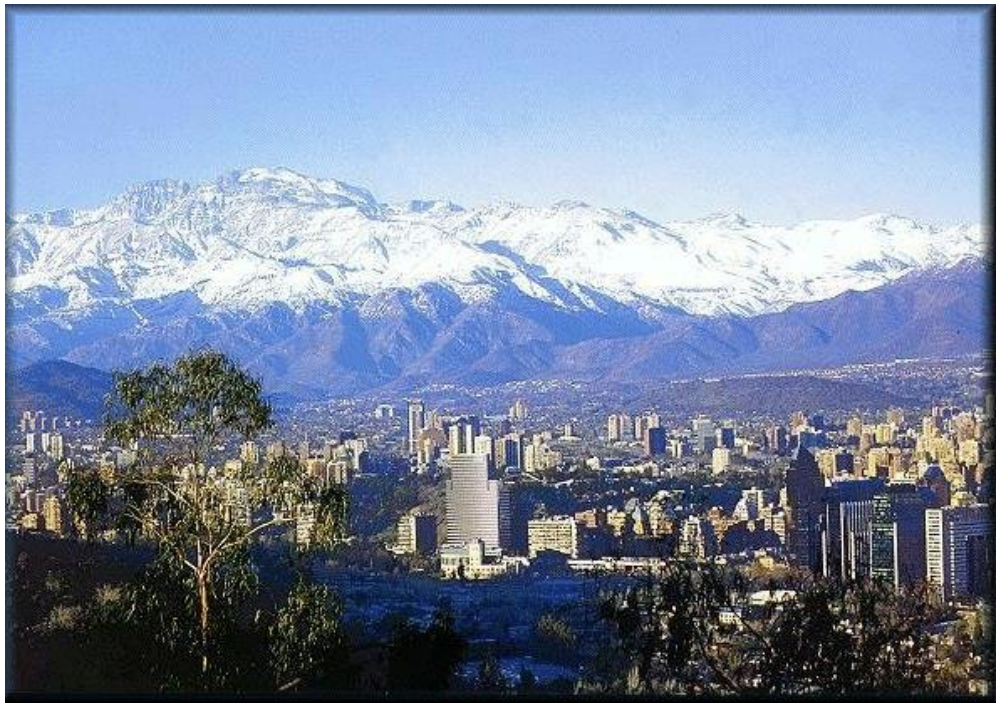
- Quantum leap in terms of availability, quality and cost of data
- Tools for planning, operation and control of transport systems
- Observe and understand traveller's behaviour
- Formulate and test hipothesis
- Formulate new policies

Challenges

- Access
- Confidentiality
- Processing
- Enrichment
- Transform this opportunity into better transport systems and better cities

Case of study: Santiago

- Santiago. Capital City of Chile:
  - Population: 6.6 million
  - 2 million households
  - Area: 1,400 km2
  - 34 Municipalities
  - 18M trips per day
    - **28% public transport**
    - 35% walk
    - 26% car



Case of study: Transantiago

- Multimodal integrated public transport system (bus, metro)
- Equipped with Smartcard and GPS
- Transactions sequence → travel structure

- ❑ 6,500 buses **all equipped with GPS**
- ❑ 10,000 bus stops
- ❑ Metro: 5 lines, 100 km, 54 trains
- ❑ **Only smartcard payment in buses** (global penetration rate 97%)
- ❑ 4.6M trips per day

Collaboration between SUBTRANS-Universidad de Chile + CONICYT (PBCT, FONDEF, ISCI)

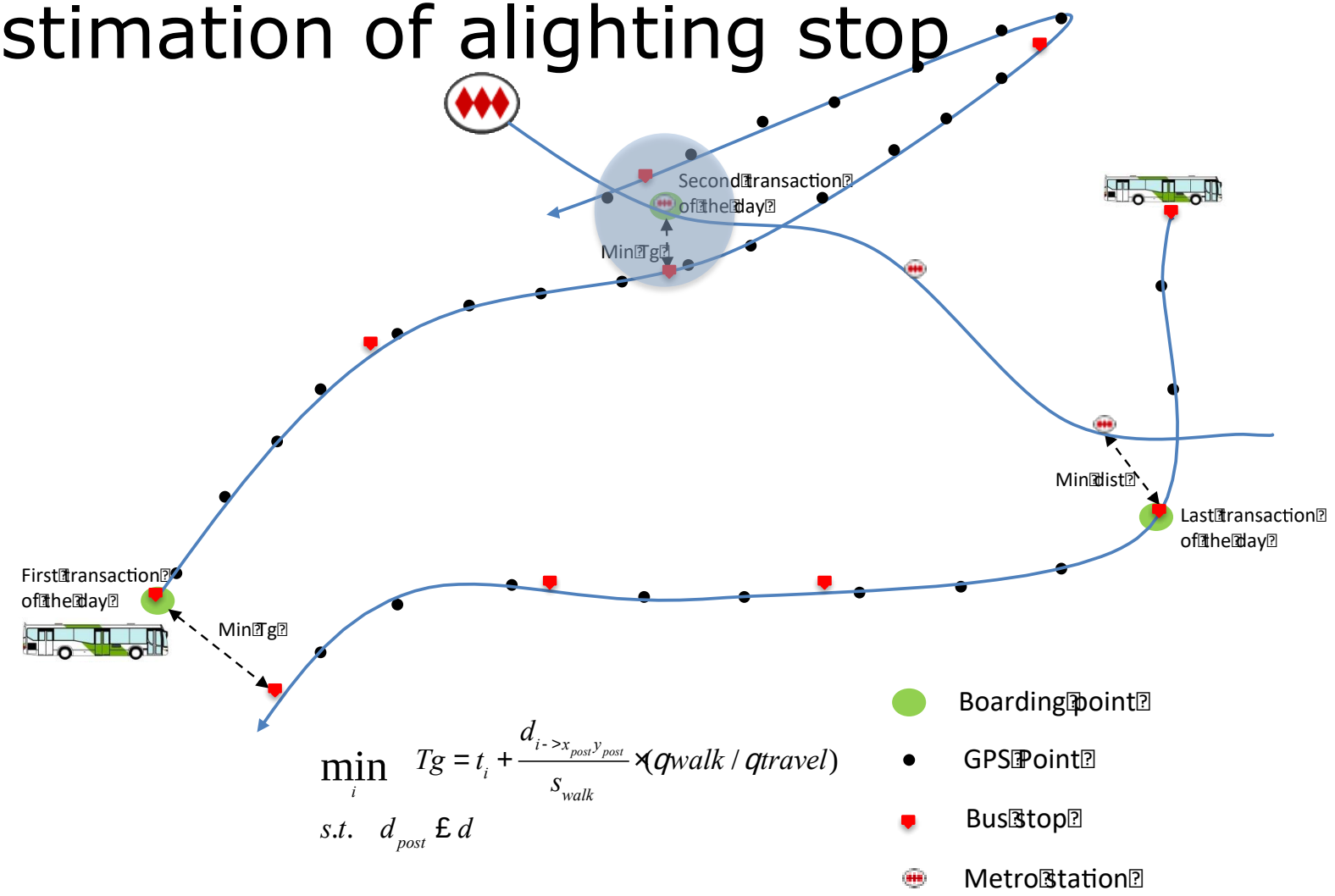
Access ✓

The Data

- Buses GPS: 1 record every 30s, 80–100 M records per week
- bip! transactions: 35-40 M records per week
- Other information:
  - Routes paths
  - Route assignments
  - Position of bus stops
  - Position of Metro stations
  - Position of bus stations

Processing

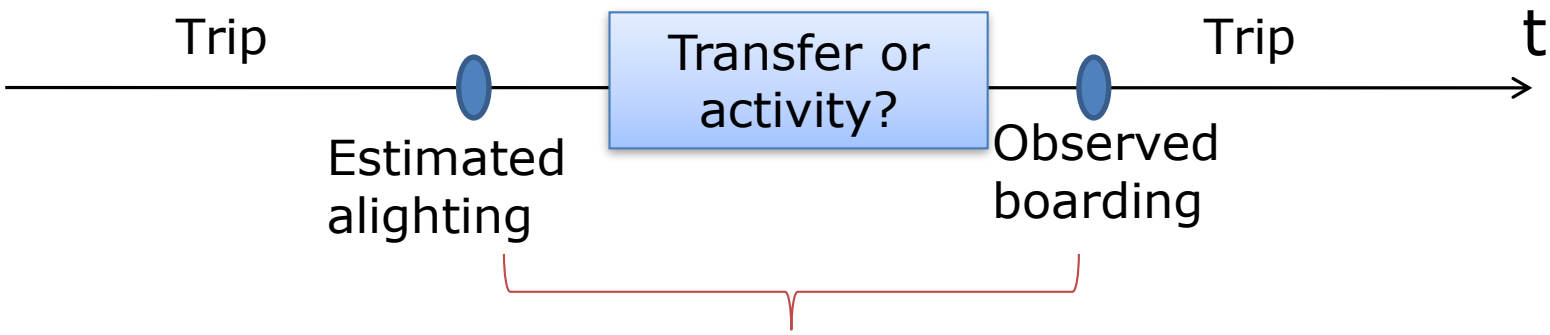
- Estimation of alighting stop



$$\min_i \quad Tg = t_i + \frac{d_{i->x_{post}y_{post}}}{s_{walk}} \times (q_{walk} / q_{travel})$$

$$s.t. \quad d_{post} \pounds d$$

Processing: Stages and Trips



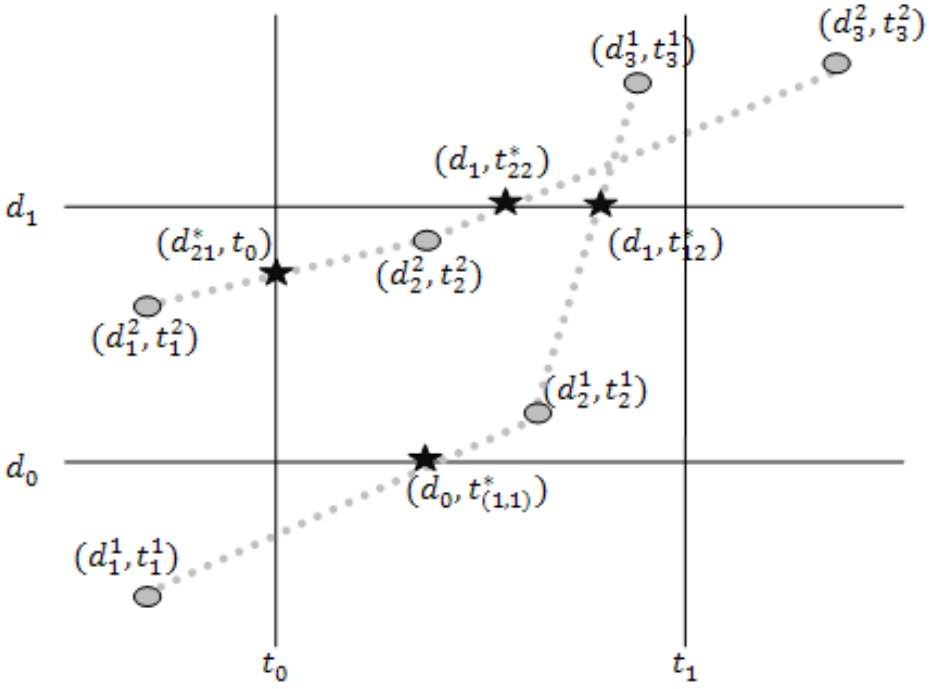
Criteria to distinguish destination from transfer

- Time elapsed
- Transaction sequence
- Frequency of PT services
- Ratio: distance on the route / Euclidean distance

Trip purpose estimation

- **Work** : Card type is adult, activity is longer than 2h, trip before is not the last of the day
- **Study** : Card type is student, activity is longer than 2h, trip before is not the last of the day
- **Home** : Trip before is the last of the day
- **Other** : Activity is between 1 min and 2h long. Trip before is not the last of the day

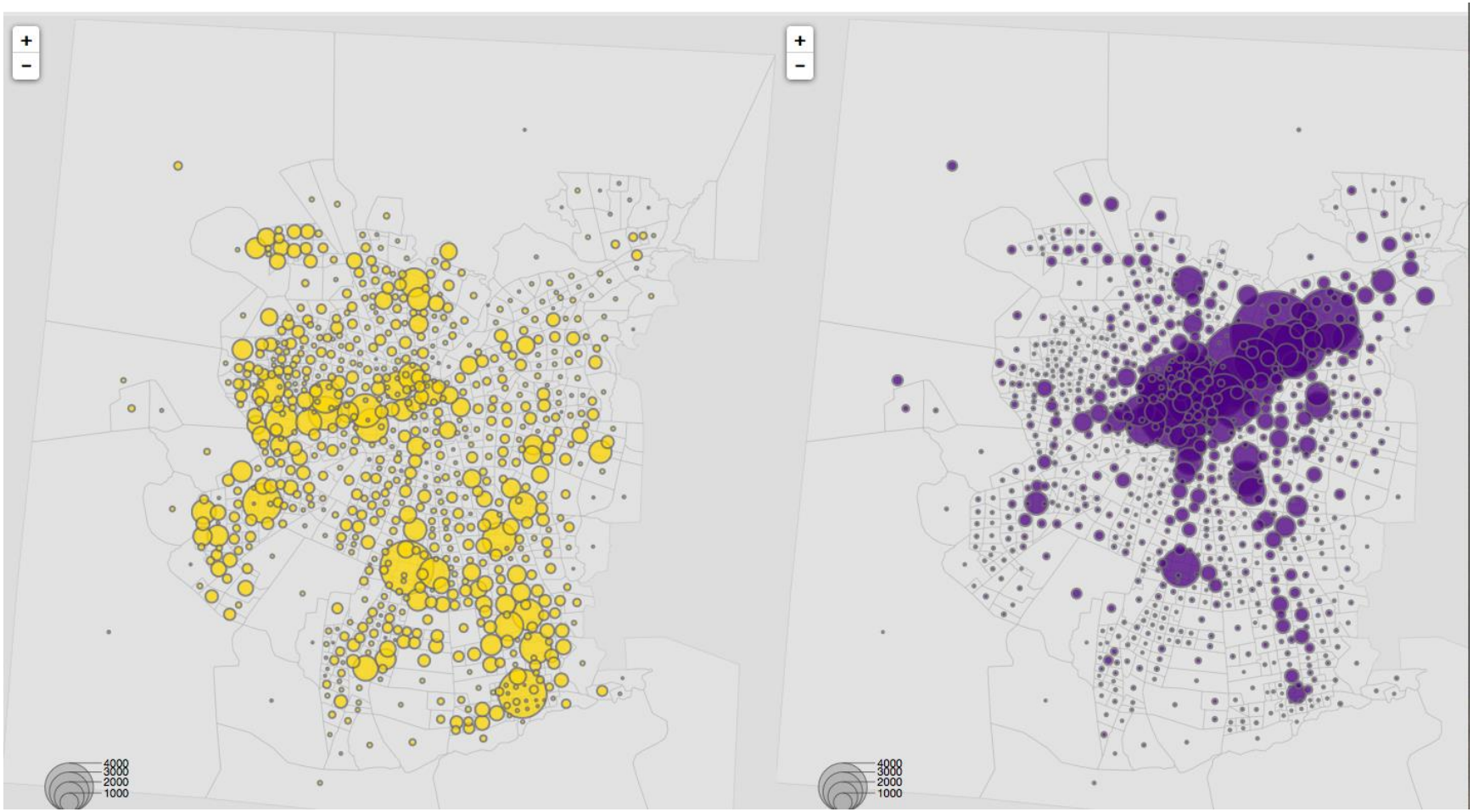
Processing: Speed estimation



$$\overline{S_g} = \frac{\sum_i D_i^g}{\sum_i T_i^g}$$

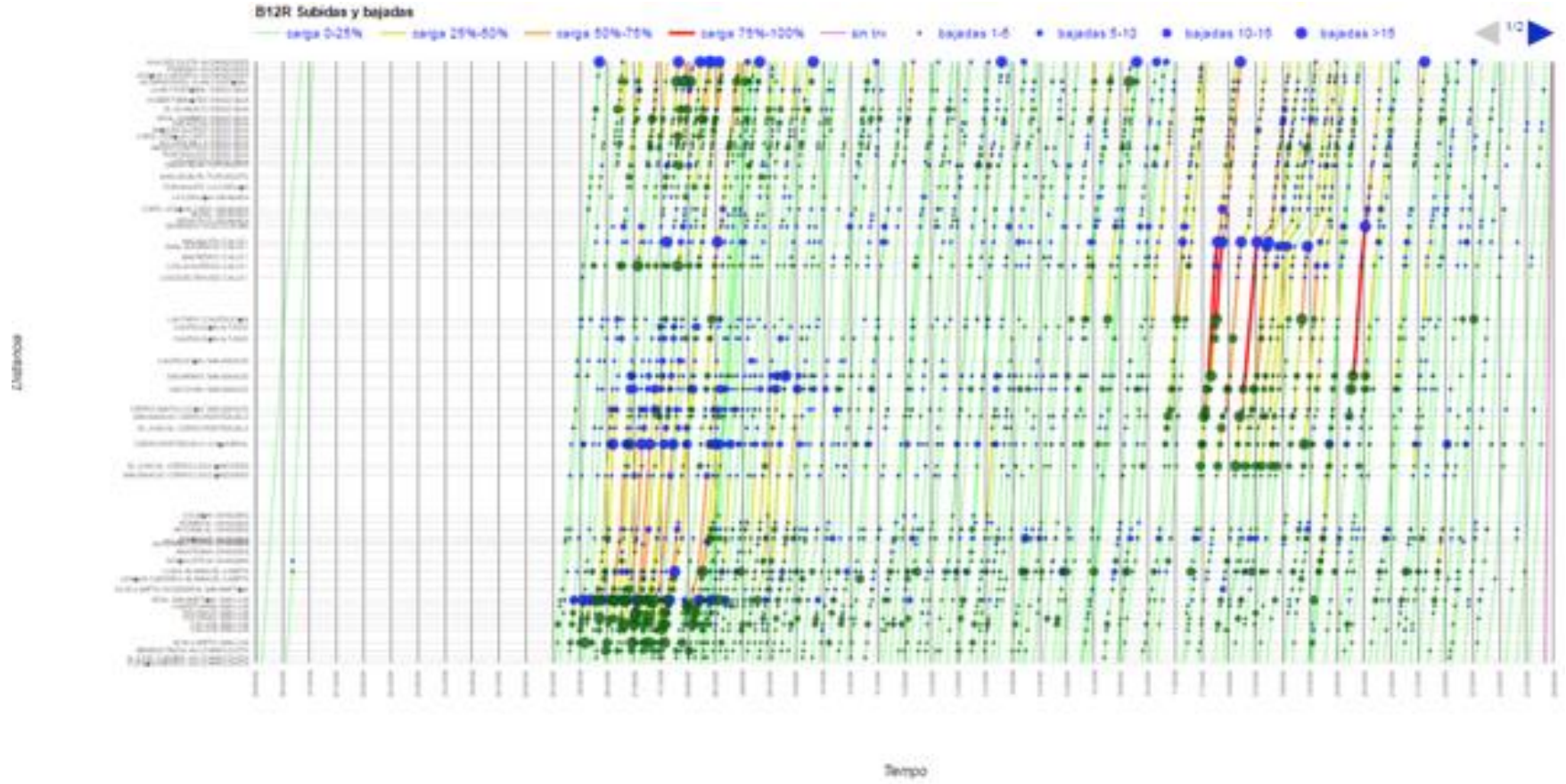
Visualization

Origin-destination matrix



Visualization

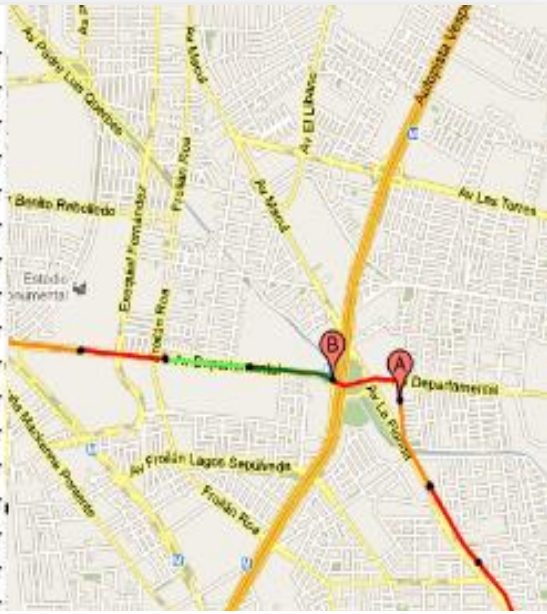
Bus Trajectories and Load Factor of One Bus Route

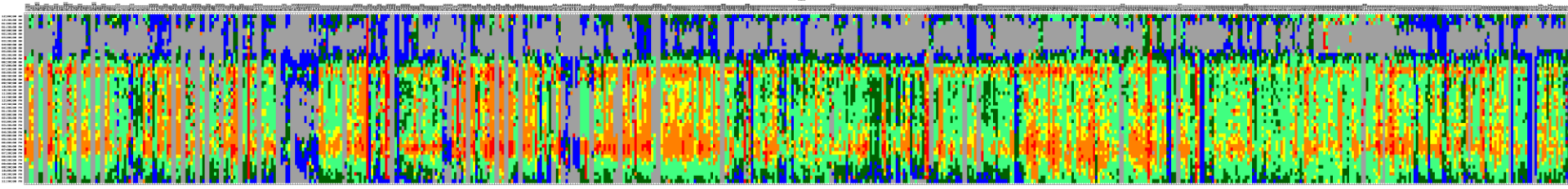


ADATRAP software that provides crucial information for decision making using GPS and smartcard) data

- Buses speed profiles
- Origin-destination flows
- Indicators

Servicio	Velocidades	Tramo	Obs.	
108R	4.06656	3	52	Ver
B08I	4.13589	14	14	Ver
102R	4.19141	17	20	Ver
B13R	4.28096	19	30	Ver
106R	4.56693	8	52	Ver
309I	4.58557	19	28	Ver
C04R	4.59294	6	14	Ver
I01I	4.62022	30	25	Ver
I12R	4.6342	25	36	Ver
B21R	4.75928	13	24	Ver
305I	4.84329	7	54	Ver
I02I	4.86831	28	36	Ver
J16I	4.95687	20	24	Ver
I07I	4.98261	15	39	Ver
213I	5.06699	33	32	Ver
B06R	5.09086	13	40	Ver
101I	5.18172	39	24	Ver

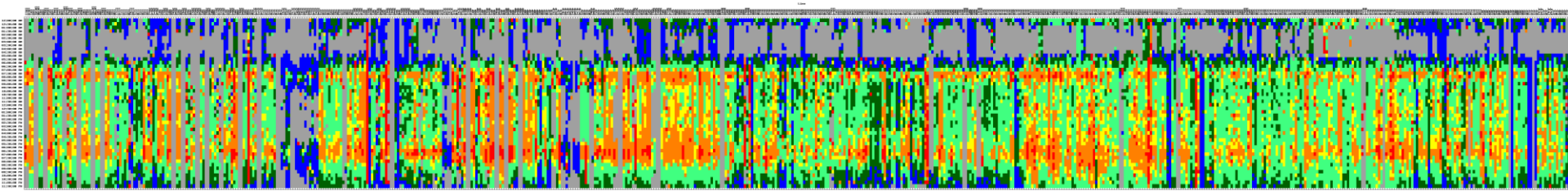




ADATRAP: main source of PT demand information

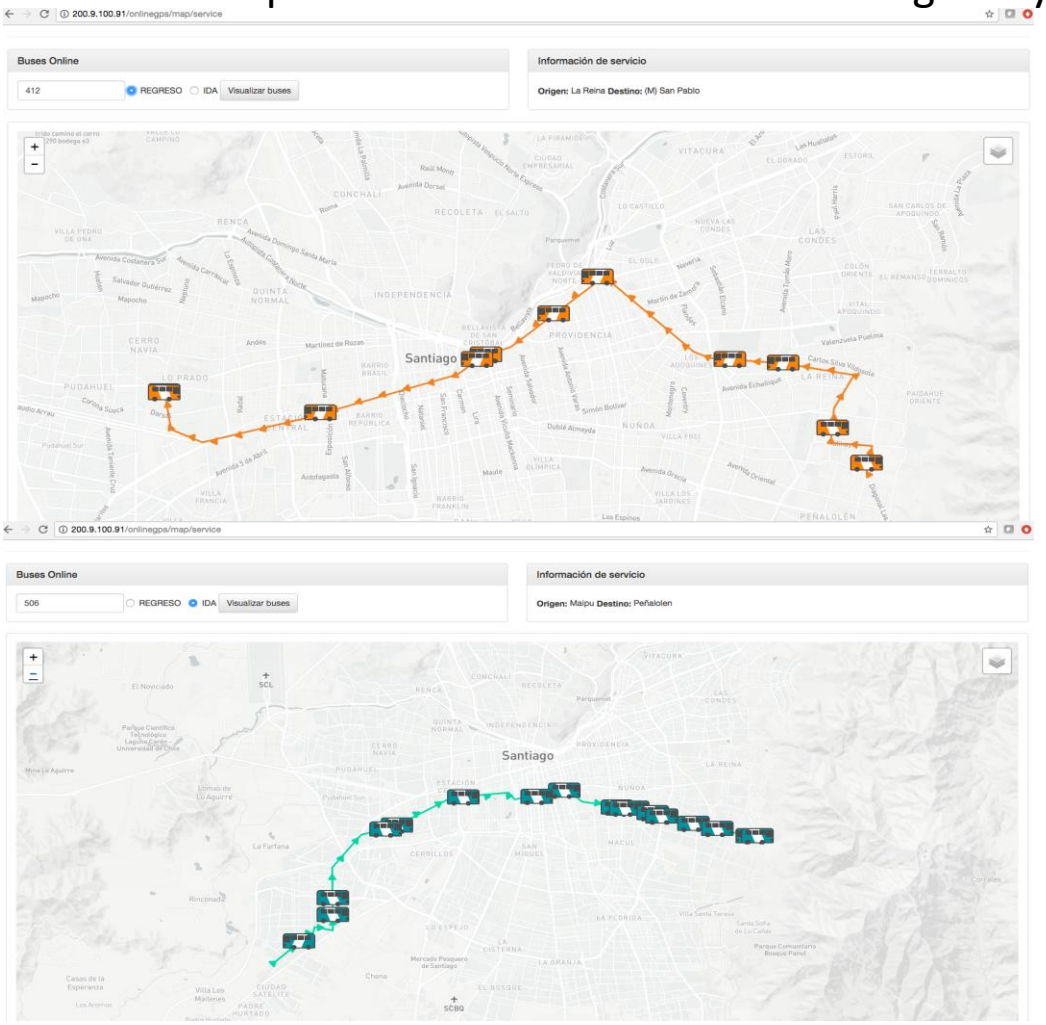
Used by: PT authorities, operators, universities, NGOs, consultants, other transport firms

- Route evaluation and modification
- Infrastructure design and prioritization
- Information campaigns design
- Monitoring level of service
- Critical point identification



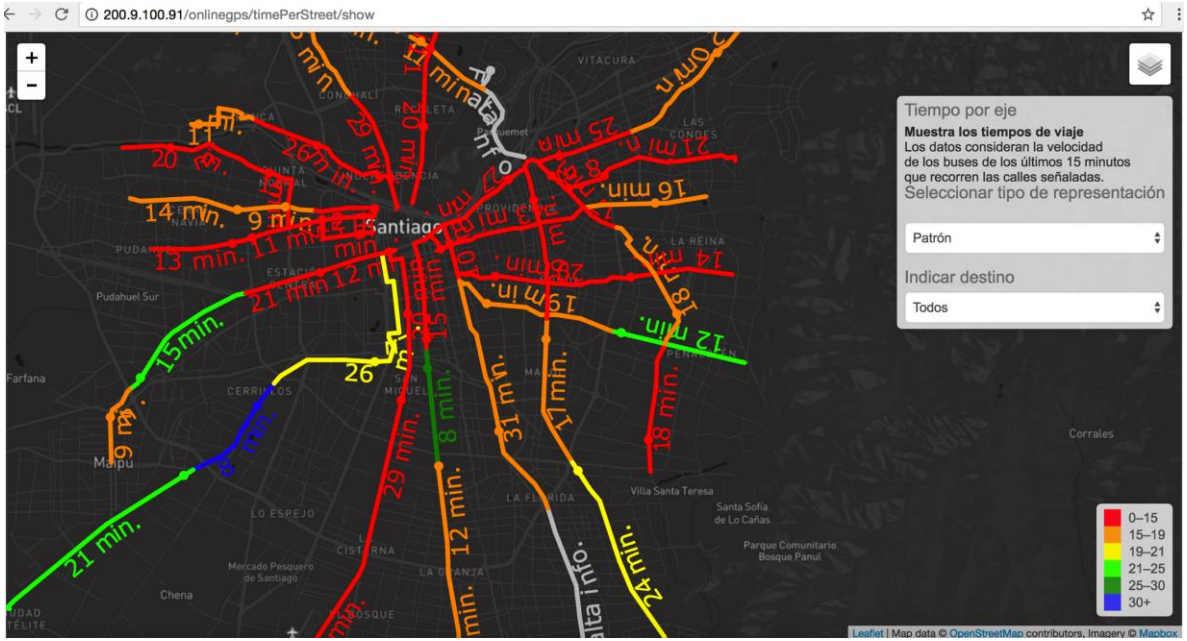
Real time monitoring

Location of buses for a particular bus route: bus bunching analysis



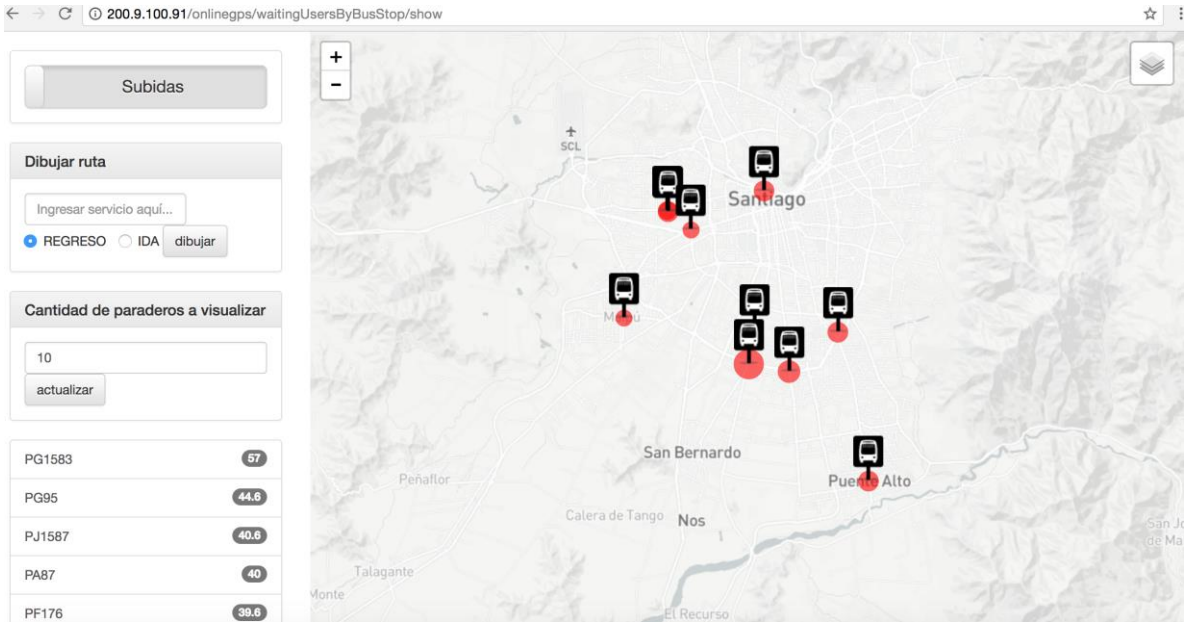
Real time monitoring

Online monitoring of speeds at street level



Estimation based on online and stored data

Estimated number of passengers waiting at a bus stop

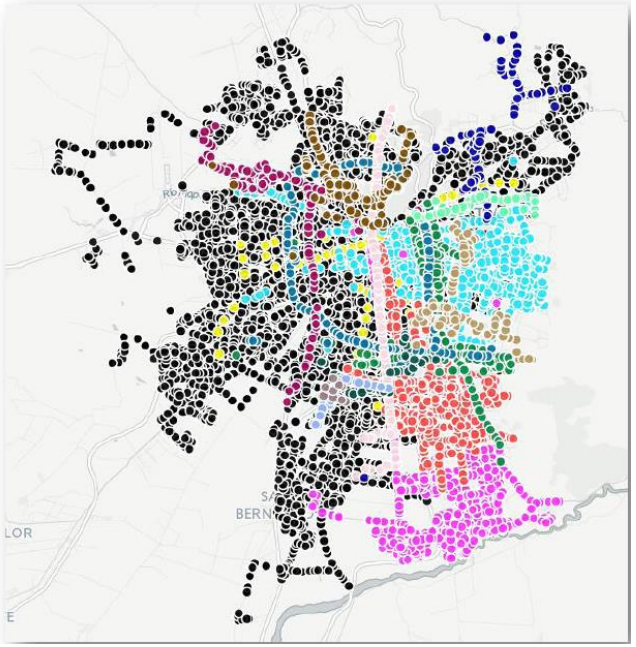


Information to users

- Transapp: Mobile phone application where users can provide and receive information of the status of the public transport system

Other developments

- Migration model
- Bus bunching model
- Fare evasion incorporation
- Spatial pattern analysis (communities)
- Environmental impacts



Conclusions

- Quantum leap on information availability and cost
- Many tools can be developed to improve planning, operation and control
- We can advance on understanding behavior and test hypothesis
- Solid grounds to formulate new (better focused) policies

Further research

- Incorporate additional information
  - Vehicle detectors, private GPS, mobile phone traces, bluetooth, online applications, surveys, twitter
- Analyze aspects of travel that were difficult to observe in the past due data limitations
  - Uncertainty, variability
  - Group behavior
  - Time/space disaggregation

- **New age for transport analysis**

Some references:

Amaya, M., Cruzat, R., Munizaga, M.A. (2017) Estimating the residence zone of frequent public transport users to make travel pattern and time use analysis. Journal of Transport Geography (in Press).

Byon, Y.J., Cortés, C.E., Young-Seon J., Martinez, F.J., Munizaga, M.A., Zúñiga, M. (2017). Developing Headway Adherence Indices and Insights for Bunching for Public Transit with GPS. International Journal of Civil Engineering, 1-12.

Cortés, C., Gibson, J., Gschwender, A., Munizaga, M.A., Zúñiga, M. (2011) Commercial bus speed diagnosis based on GPS-monitored data. Transportation Research C 19(4), 695-707.

Devillaine, F., Munizaga, M.A. y Trepanier, M. (2012) Detection activities of public transport users by analyzing smart card data. Transportation Research Record 2276, 48-55.

Gschwender, A., Munizaga, M.A. y Simonetti, C. (2016) Using Smartcard and GPS data for policy and planning: the case of Transantiago. Research in Transport Economics. Accepted for publication.

Munizaga, M.A., Palma, C. (2012) Estimation of a disaggregate multimodal public transport origin-destination matrix from passive Smart card data from Santiago, Chile. Transportation Research 24C(12), 9-18.

Munizaga, M.A., Devillaine, F., Navarrete, C. , Silva, D. (2014) Validating travel behavior estimated from smartcard data. Transportation Research 44C, 70-79.

Gibson, J. Munizaga, M.A., Tirachini, A. y Schneider, C. (2016) Estimating the Bus User Time Benefits of Implementing a Median Busway: Methodology and Case Study. Transportation Research Part A 84, 72-83.