Frederic Roulland
Xerox Research Centre Europe

Using passengers’ data to measure perceived impact of schedule deviation in public transit
Global context

• Leverage transportation operations data to plan and optimize city mobility

  • Short term challenges: Operations efficiency
    • Productivity
    • Quality of Service

  • Long term challenges: Sustainable mobility
    • Traffic congestion
    • Energy consumption, pollution
Our Data Analytics Vision

- Forecasting
- Control
- Visualization
- Modelling
- Simulation
- Auth
- Throttle
- Storage
- Process
- Portal
- Service Desk
- Privacy
- KB

Transport Operations Data
- Public Transport
- Tolling
- Parking

External Data
- Weather
- Demographics
- GIS
- Social Networks
- Sensor Networks

Value-Add Partners
- Transport Operations Data
- External Data

Analytics Platform
- Visualization
- Diagnostics
- Simulation
- Forecasting
- Modelling
- Control

Cloud

Data & Biz Intelligence Consumers
- Transportation Authorities
- Citizens
- Partners
- Local Economy Startups
- Universities

October 27, 2016
The challenges of data analytics

- Heterogeneity
  - Different providers
  - Different domains
  - Different spatio-temporal granularities
  - Different user needs

- Quality
  - Noisy
  - Partial
  - Low level

- Size
Platform conceptual architecture

- Minimal complexity and easy load distribution through highly decoupled multilayered approach
Perceived impact of schedule deviation

- How schedule deviation really impact people?
  - Waiting times
  - Lateness at final destination
  - Missed connections
- How to know this?
  - Combining schedule adherence data with passengers individual trips
  - Modeling the effect of schedule deviation over time on people behaviors

Size: Number of service trips
Color: Number of trips out of schedule

Size: Number of travellers
Color: Average waiting time
Data available

- Validators from fare collections systems
  - Very often Check-in only
- Automated Vehicle Location systems
  - Official Schedules
Passenger trip reconstruction

• Alighting inference
  Using spatio-temporal constraints on sequence of check-in to infer most probable alighting
  – Regular users are representative
  – Symmetry of daily travels

• Trip alignment
  Minimizing distance between set of observed trips from validations and set of scheduled trips
Estimated waiting cost

- Components
  - Platform waiting time
  - Budgeted waiting time
  - Stress factors

- 3 Platform waiting time scenarios
  - High frequency service
  - Connections
  - Low frequency service

Platform waiting time and budgeted waiting time depend on schedule adherence variance
Aggregating the metrics

Waiting cost

Late arrival at final destination

Missed connections

Colour: Average waiting cost
Size: Number of passengers

Colour: Average lateness (late/all)
Size: Number of (late/all) people

Colour: sum of the lateness at final destination for every trip where the connection was missed / number of connections missed
Size: Number of missed connections
Conclusion

- New usage of transportation data allows a closer understanding of mobility
- Our approach go beyond analytics model: Fully automation from data collection to decision support
- This works is an example of how it can be used to improve user experience
  - Next step is to adopt and validate such metrics in transport organization