Measure Door-to-Door Mobility Through GPS Tracks & Big Data Methodologies

Sylvain Coppéré, Stéphane Mastalerz, Félix Motot and Marwan Ghanem

TransitData 2019
Outline

- Context
- Data
- Global Process
- Detecting Door-to-Door trip
- Clustering
- Results and Performance
- Modal detection
- Perspectives and Conclusion
Context

Shared, autonomous and virtuous mobility => Need to understand how individuals travel

GPS

- True multimodal analysis
- Observed mobility
- Detailed geographical analysis
- Dynamic analysis

Technical Issues

1. Data cleansing
2. Detection of rides
3. Modal allocation
4. Calibration
5. Data visualisation
• Department of Calvados
• Number of users: 21K
• Number of gps coordinates: 67M
• Duration: 1 month (November – December)
• Frequency:
  • Coordinates are registered when a user moves 25 or 100 meters
  • Spatial check every 20 seconds
Global Process

Step 1: Raw Data
Step 2: Filtering by altitudes
Step 3: Detecting pauses
Step 4: Trip Identification
Step 5: Clustering 50m
Detecting Door-to-Door trips

- Treat GPS coordinates in temporal order
- End of OD:
  - 30 minutes in the same 300 meters zone
  - 30 minutes without any information
- Trips from one region to another
Clustering

- Regroup GPS coordinates that are at most 50 meters from one another
- Replaced by centroids
- Reduces errors
- Facilities future tasks
Results: Temporal Evolution

- Daily oscillations
- Peaks around 8 AM, 12 PM, 6 PM
- Fewer trips during the weekend
Trips between IRIS

- IRIS Unit => area with a population generally between 1,800 and 5,000.
- Most trips are inside the same IRIS
- Temporal analysis: several IRIS rarely visited during weekends
Performance

• Linear in relationship to number of users
• Average time per user
  • 1 day => 0.01 seconds
  • 1 week => 0.1 seconds
  • 1 month => 0.33 seconds
• Memory usage 200 MB
• Complexity : $O(N^2)$
Calibration

- Residence => Number of nights spent + Number of weekends
- Financial status has no relationship to application usage
- Calibration using INSEE census
- 6% penetration on average
Modal Detection

1. Filter irrelevant ODs
   - ODs with a duration longer than 3 hours
   - ODs with distances less than 300m
   - ODs with less than 10 GPS points

2. Identify Pedestrian trips
   - Distance
   - Instant Vitesse

3. Public transportation v.s Private vehicle
   - Bus station near origin and destination
   - Bus route
   - Time
Perspectives and Conclusion

• Perspectives
  • Finding the best compromise between time and precision
  • Better understanding of how people travel
  • Creating an automated process to analyse data on daily bases

• Conclusion
  • Proposed an algorithm to compute origin-destination trips from raw GPS points
  • Simple statistics to detect residences