Human Mobility Data for Resilient Urban Planning

Presenters: Nick Jones, Takahiro Yabe, Maham Khan, Sam Heroy (GFDRR)

Contributors: Oliver Lock, Satish Ukkusuri (GFDRR/Purdue), Tatiana Peralta (Transport GP), Enrico Ubaldi, Emanuele Strano (MindEarth)
Mobility data: from innovation to mainstream

- What's new with human mobility datasets?
- Insights from
  - Kathmandu (Nepal)
  - San José (Costa Rica)
  - Chennai (India)
- Seeking feedback: strengthening the contribution of these methods
Products like Urbanization Reviews and post-disaster needs assessments address abiding questions about **spatial form & economic growth**, access to **jobs**, urban **service delivery**, and **recovery after shocks**. How could new mobility datasets inform these?

Focused on the **opportunity presented by GPS datasets**, our team is developing tools & use cases to inform urban and disaster risk management using mobility data.

<table>
<thead>
<tr>
<th>Smartphone location data</th>
<th>Traditional CDR data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial resolution: ~ 15 meters</td>
<td>Spatial resolution: catchment zone of nearest cell tower (e.g. 1km +)</td>
</tr>
<tr>
<td>Temporal resolution: up to 100 locations per day</td>
<td>Temporal resolution: only when call placed</td>
</tr>
<tr>
<td>Scalability: Single channel to access data, in consistent format, under Development Data Partnership</td>
<td>Scalability: New MOU required with each country’s cell provider</td>
</tr>
<tr>
<td>Legal framework: Bank-wide MOU</td>
<td>Legal frameworks: Ad hoc approach</td>
</tr>
</tbody>
</table>

**Studies:** Mexico, India, Costa Rica, Nepal

**CoP:** Disaster Mobility Data Network

**Tools:** MobilKit
Mobility data: a single user

1. The raw data

<table>
<thead>
<tr>
<th>Date</th>
<th>Lat</th>
<th>Lng</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020-01-23 01:48:49</td>
<td>38.90041</td>
<td>-77.04408</td>
</tr>
<tr>
<td>2020-01-23 01:50:55</td>
<td>38.90112</td>
<td>-77.04574</td>
</tr>
<tr>
<td>2020-01-23 01:53:06</td>
<td>38.90188</td>
<td>-77.04762</td>
</tr>
</tbody>
</table>

2. Identify place visits

3. Count visits by geographic unit

4. Example analysis: Daily distance traveled
Mobility data: a population

Conventional work-flow: (i) check quality; (ii) filter out irrelevant users; (iii) test representativeness; (iv) analysis. GPS datasets are ‘convenient samples.’

https://medium.com/analytics-sampling-bias-bbde6560fa
# Key concepts

<table>
<thead>
<tr>
<th>Metric &amp; Concept</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radius of gyration (RoG)</td>
<td>A measure of how far an individual travels away from their home location.</td>
</tr>
<tr>
<td>Significant location</td>
<td>Locations regularly visited by the individual; typically includes home, workplace, schools, etc.</td>
</tr>
<tr>
<td>Home and workplace detection</td>
<td>Clustering algorithms are used to estimate home and work locations, linking anonymized users with demographic information.</td>
</tr>
<tr>
<td>Origin-Destination Matrix</td>
<td>A table of how many people travel between location pairs (e.g. home and work).</td>
</tr>
<tr>
<td>Stop detection</td>
<td>Algorithms used to estimate length and location of people’s visits to points of interest based on raw GPS data.</td>
</tr>
<tr>
<td>Anomaly detection</td>
<td>Identifying statistically significant change in population levels during a treatment period (e.g. a natural disaster) compared with baseline.</td>
</tr>
</tbody>
</table>
Urban planning applications

Three topics addressed today:

■ What do a city’s mobility patterns mean for its economy?
■ How are public spaces used by residents and tourists?
■ How can flooding affect urban transportation?
1. Urban connectivity and labor markets
Day vs. nighttime mobility distinguishes sleepy periphery areas (red) from day-active core areas (blue)

- Substantially less movement on weekends
More “sleepy” night-time oriented districts associated with single use, lagging employment and commercial development (e.g. the Arco Sur in red)

- Districts with more daytime use are indicative of higher employment and commercial activity
Clusters of districts with similar mobility patterns further differentiate more/less active areas

- Light green areas have the most daytime activity, followed by teal and grey (sleep)
- Daytime active areas have more commercial and amenity development
Telemetry-derived routing data reveals metropolitan inequities in commute efficiency

- For fixed distance, Arco Sur residents have slower commutes to the city center
- Better E-W than N-S connectivity in the metro
2. How do citizens use public space?
How are parks in Chennai used? When are they busiest?

- Popular spot in evening
- Generally busy during daytime; especially weekend mornings
How are parks in Chennai used? How far do people travel to get there?

**Catchment area of public amenities:**
visitors to Kamarajapuram Football Ground make longer trips on weekends.
How are parks used? Do they serve rich people, poor people or both?

Visited by the richest

Mixed wealth visitors

Visited by the poorest
How are large metropolitan parks used? What times of day are various parts of a park busy?

La Sabana Metropolitan Park was Costa Rica’s national airport prior to its redevelopment as “the lungs of San Jose” in 1977. Pictured: visit density within the park during 2021.
How are large metropolitan parks used? What times of day are various parts of a park busy?
Kathmandu’s cultural heritage sites are highly co-visited. Figure shows the frequency of co-visits on the same day between the major 9 sites. Follow-up analysis will look at economic impact of heritage sites through co-visits to ancillary industries (shops, restaurants, hotels).
3. Disaster resilience of urban transport
Cyclone Nivar (Chennai, India)
Significant decrease in active population in central business district

- Comparison with “typical” active population before the Cyclone
- Substantial reduction around George Town and Chennai Central Station
Cyclone Nivar

3-hour delay in evening peak in central business district during cyclone

- **Substantial decrease** in trips on Nov 25 (orange)
- **Slower evening peak** on Nov 25, 26, 27 compared to typical weekdays
Cyclone Nivar
Disproportionate impact on visits to hospitals during cyclone

- The cyclone had a disproportionate impact on poor people’s ability to access essential facilities.
- Visits to hospitals on day after cyclone

Hospitals (Nov 26)

Lower 10%: -31%
Upper 10%: -24%
Cyclone Nivar
Larger & longer disruptions on local stores compared to supermarkets

- More (~70%) and persistent disruption (until Nov 29th) to local stores
Quantify commute mobility reduction in different timings during the pandemic

Gradual recovery towards mid-2021 after the huge surge in cases in May 2021
Next steps and priorities
Zoning schemes assume a model of how people move (or should move) across the city

Q: Does the actual pattern of H-W trips correspond with planners’ assumptions?
Functions of mobilkit

Overview

Input
- Load data
- Localize timest.
- Dump/reload

Statistics
- Pings per user
- Active days
- Filter users

Spatial
- Tessellation
- Home-work OD
- Pop. estimation

Temporal
- Per-area profile
- Residual activity
- Land-use

Shifts
- Group areas
- By ROI
- By stat
- Displacement
- Spot deviations

AUX data
- Geo-referenced indicators
- Time-stamped events
- Tessellation (shapefile)
- Census figures
- Socio-economic

 UID | UTC  | lat  | lon  | acc |
-----|------|------|------|-----|
1322 | 16246| 45.2344 | 12.3443 | 12 |
1223 | 16274| 44.3348 | 11.2223 | 23 |
....|....|....|....|....|
2589 | 16295| 45.7655 | 13.4534 | 33 |

Output

Per-user stats/filter
- Dataset stats
- Pop. coverage

Pop. density
- Local densities
- Exposure map

OD
- Commuting patterns

Land use estimation
- Estimate biases

Displacement measure
- Spot changes in venue visit rate
Thank you!
mobilkit

github.com/mindearth/mobilkit

mobilkit.readthedocs.io

pip install mobilkit
Telemetry-derived routing data also reveals congestion points and helps to diagnose bottlenecks in the highway network

- High betweenness centrality or lack of redundancy (blue) in the road network is thought to cause traffic bottlenecks
- Congested areas (red) are colocated with high betweenness, suggesting lack of routes causes congestion and emphasizes need for outer ring road

Commuting patterns and congestion: inequities in commute congestion