Big Data and Security

Jeffrey Borowitz, PhD

Lecturer

Sam Nunn School of International Affairs

CDRs for Development

How Might Big Data Affect Development?

- Cell phone call Call Detail Records (CDRs)
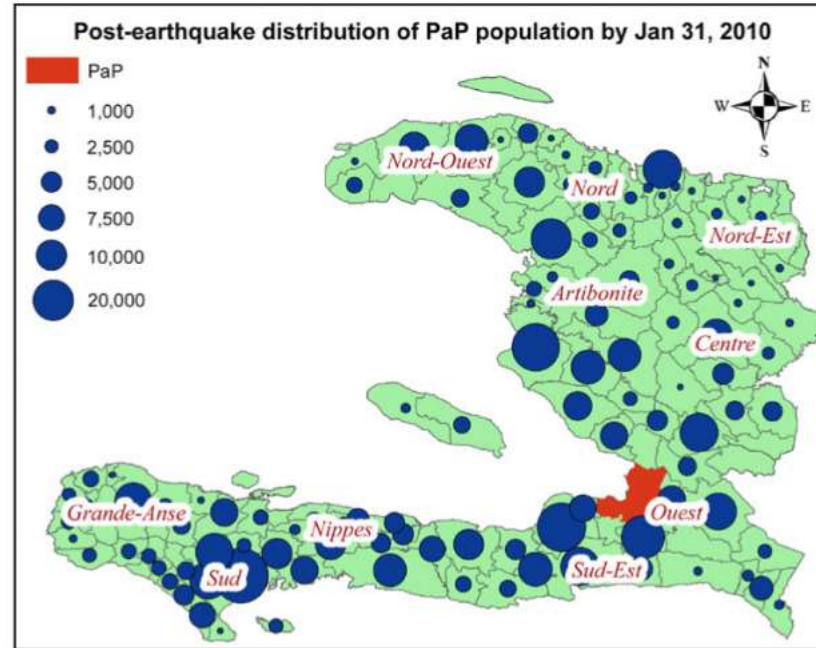- Replacing government statistics

Call Detail Records

- Contains information on who calls whom, when, where
    - Caller ID – specific to sim card
    - Caller Cell Tower Location (lat/lon)
    - Recipient ID – specific to sim card
    - Recipient Cell Tower Location
    - Call Time –
    - Call Duration – How long was the call?

Georgia Tech

Use of Call Detail Records: Disaster Relief

- Disasters cause people to leave their homes

- People also need aid and services during disasters

- But it's hard to track people down to deliver services

- Example: Haiti Earthquake
  - Unsurprisingly (?) migration is to regions of previous call activity
  - So CDR could be used to predict migration patterns in response to disasters

Georgia Tech

Haiti Migration from Port au Prince

Post-earthquake distribution of PaP population by Jan 31, 2010

Bengtsson, L., Lu, X., Thorson, A., Garfield, R., & Schreeb, J. von. (2011, August). Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti. Retrieved from https://journals.plos.org/plosmedicine/article/file?type=printable&id=10.1371/journal.pmed.1001083
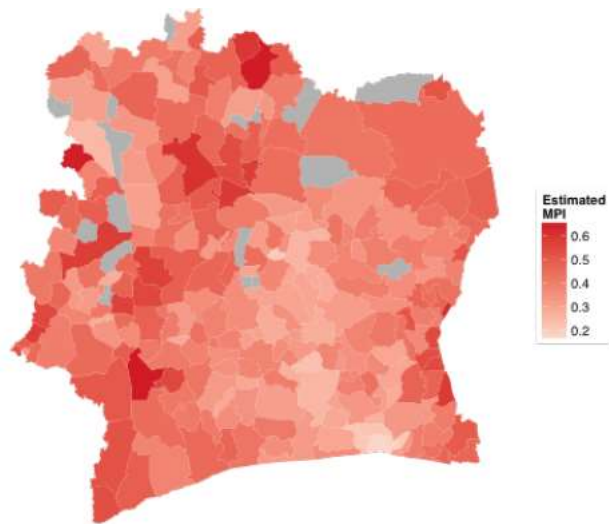
CDR as Proxies for Socioeconomic Variables

- Frias-Martinez and Virsesa (2012) combine call data with economic statistics in Latin America

- They calculate information about a few topics:
  - Mobility - spread of phone calls
  - Network - diversity and spread of people you call
  - Behavior - number of calls, length of calls, etc

- These topics can explain about 70-80% of variation in socioeconomic variables from survey data, such as:
  - Overall income
  - Education
  - Whether house has cement floor

- Benefit: could compute reasonable versions of economic indicators for less money than statistical infrastructure

Georgia Tech

CDR Proxies SES in Côte d'Ivoire

- Smith, Mashhadi, and Capra (2013)
    - Main characteristics from CDRs are link diversity within an area
    - Results match independent measurements of development index
    - But can be provided at a finer grain of detail

Smith-Clarke, C., Alcatel-Lucent, A. M., Capra, L., Christopher Smith-Clarke University College London, University College London, Licia Capra University College London, … Autodesk Research. (2014, April 1). Poverty on the cheap: estimating poverty maps using aggregated mobile communication networks. Retrieved from https://dl.acm.org/doi/pdf/10.1145/2556288.2557358

Cell Location Data via Apps

- CDRs track who calls whom – who/how you choose to call says a lot about you

- But you only understand where people are when they choose to call.

- Most phones are recording where people are most of the time in one app or another.

- SafeGraph buys and packages this data for a range of analytics related to foot traffic: https://www.safegraph.com/

Georgia
Tech

COVID

- This type of data is just being explored for research purposes

- Example: Alcott et al (2020) uses place visit rates to measure how much people are traveling/moving.

Lesson Summary

- Call Detail Records (CDRs) contain information including: who calls whom, where they are, and when the call takes place

- This information can be used in circumstances such as natural disasters to predict migration patterns

- CDR can be used as proxies for socioeconomic variables for less money than statistical infrastructure

- There's another generation of new applications coming from mobile app location tracking

Georgia Tech