

Behavioral insights for development from Mobile Network Big Data

Roshanthi Lucas Gunaratne and Sriganesh Lokanathan
LIRNEasia

10 September 2014

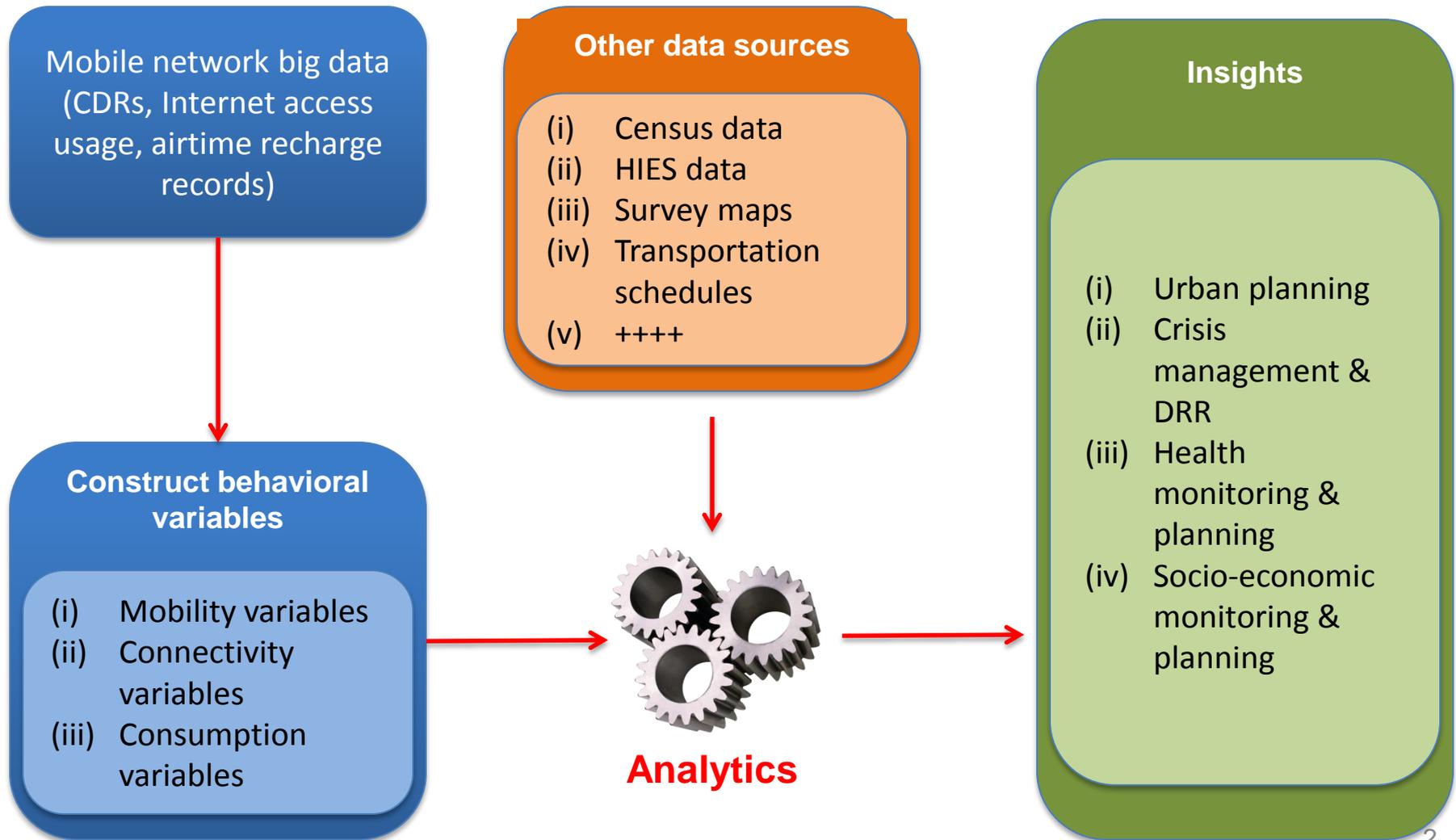


Canada

This work was carried out with the aid of a grant from the International Development Research Centre, Canada and the Department for International Development UK..



Taking a broader view: the overall process of leveraging mobile network big data for development

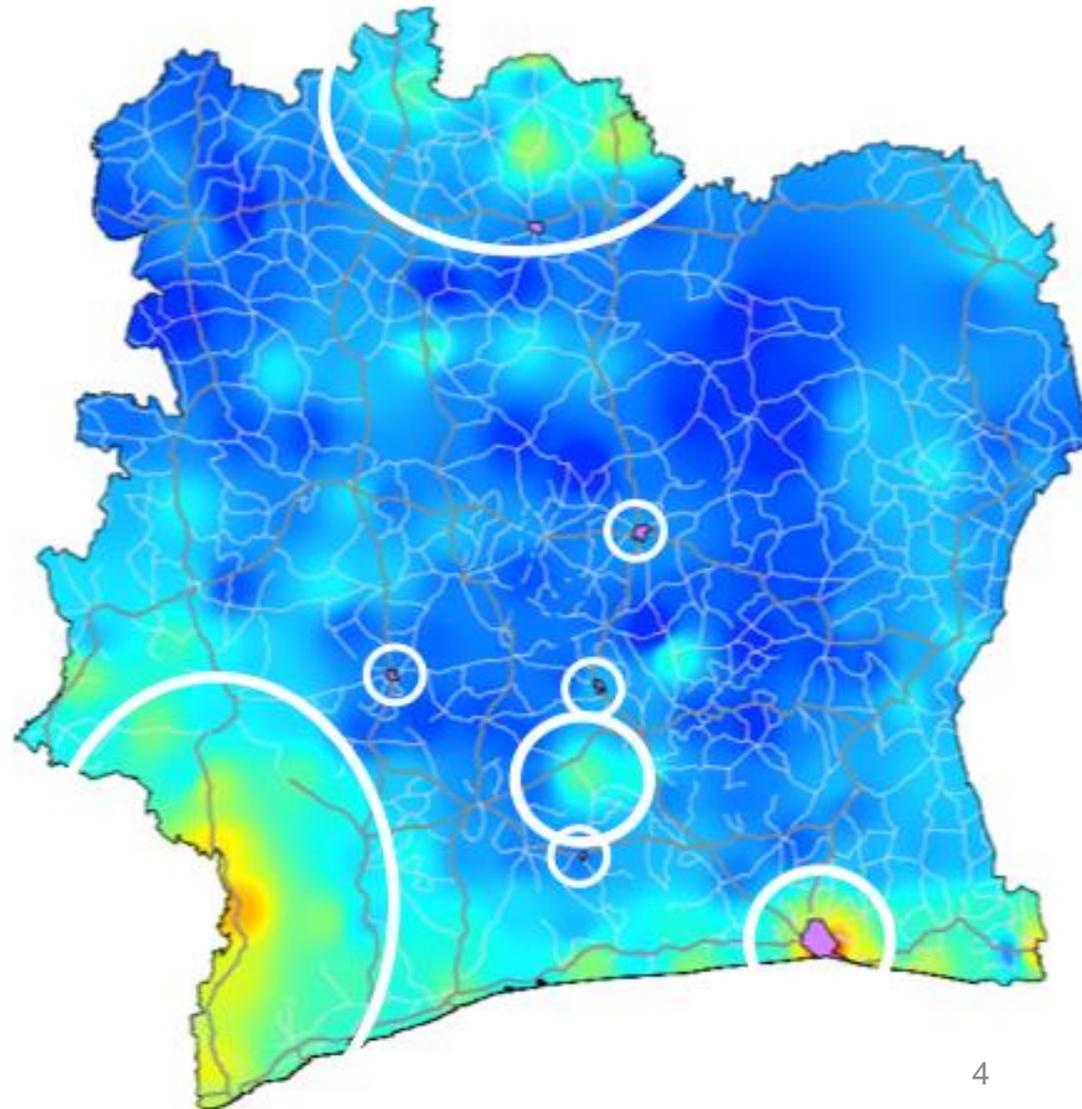
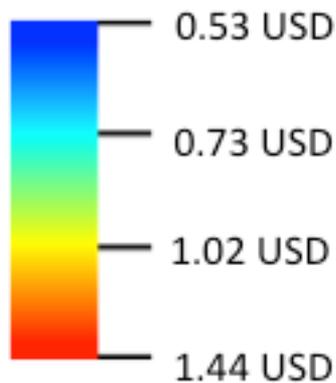


Socio-economic monitoring and planning: creation of poverty maps

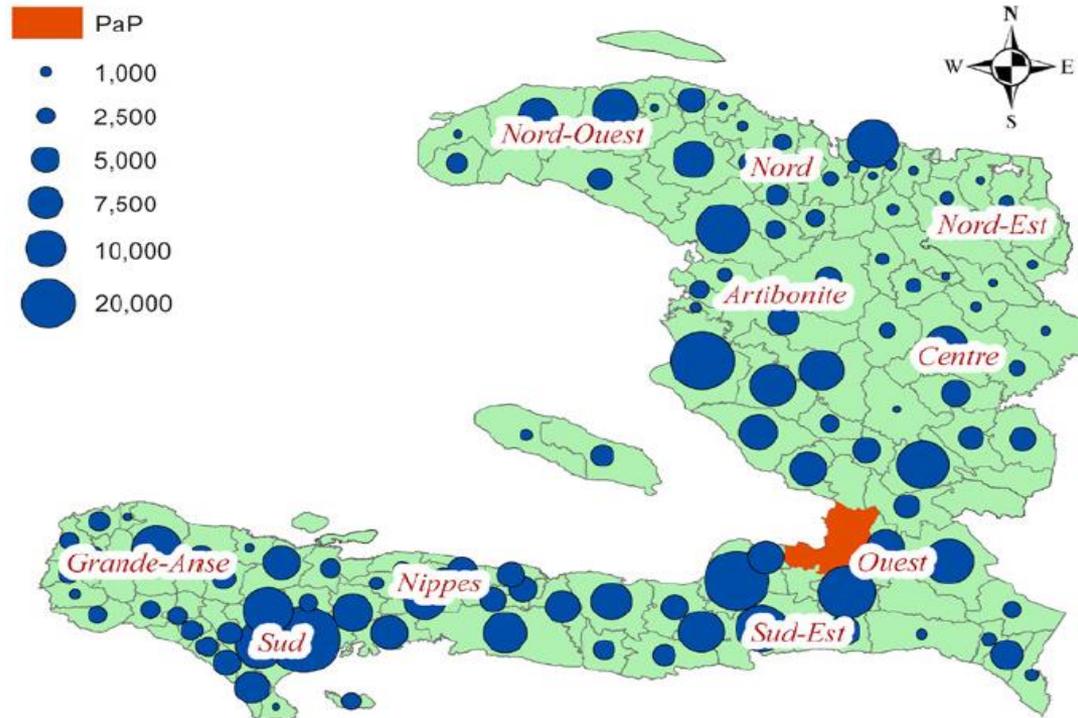
An example from Cote d'Ivoire

Source: Thoralf Gutierrez, Gautier Krings, Vincent D. Blondel, (2013).
Evaluating socio-economic state of a country analyzing airtime credit and mobile phone datasets. Available at <http://arxiv.org/abs/1309.4496>

Average airtime recharge



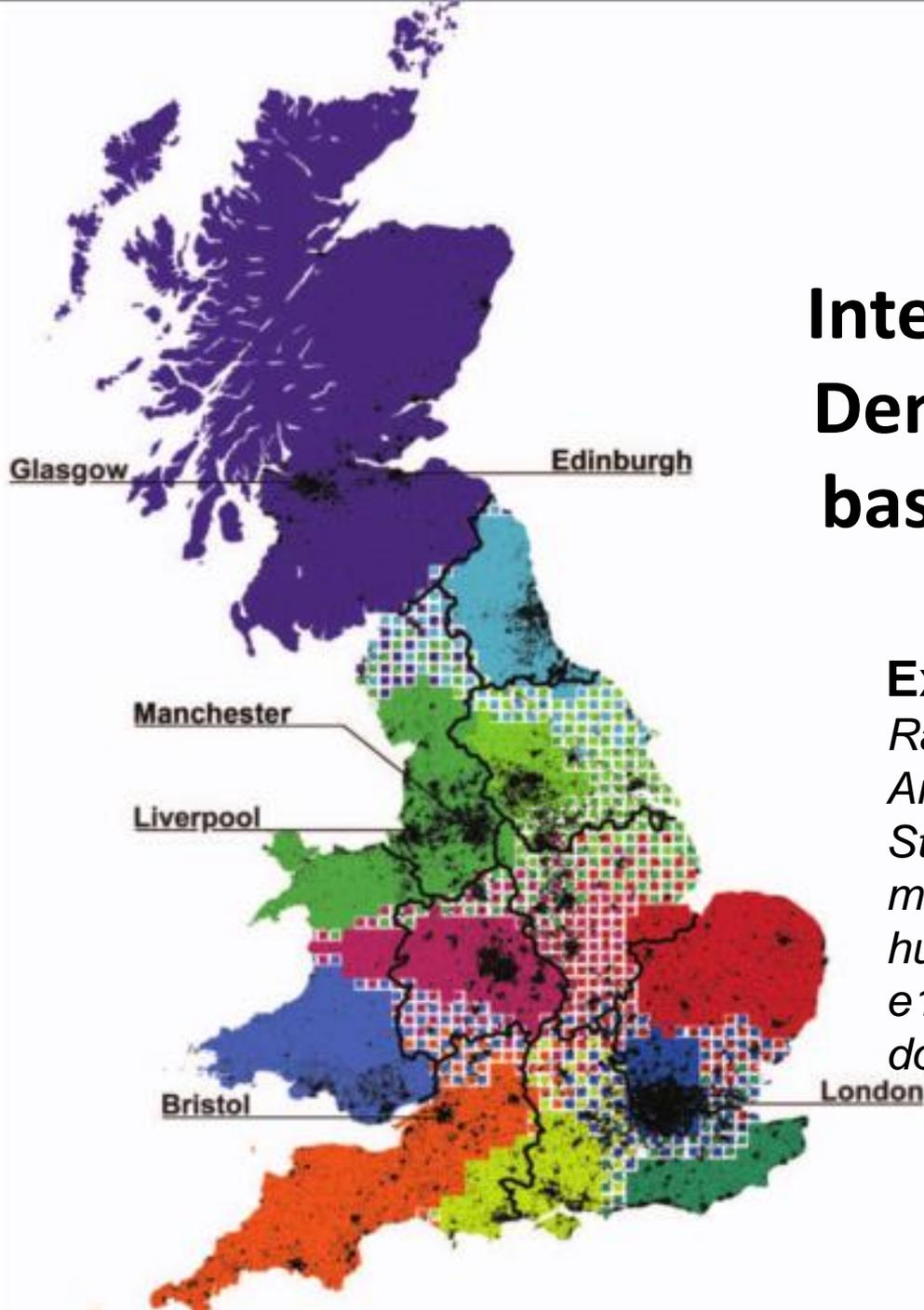
Disaster Management: Population distribution maps after a disaster event



Example from Haiti - Post-earthquake distribution of Port au Prince (Haiti) population following after 2010 earthquake

Source: Bengtsson, L., Lu, X., Thorson, A., Garfield, R., & von Schreeb, J. (2011). *Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti*. *PLoS Medicine*, 8(8), e1001083. doi:10.1371/journal.pmed.1001083

Regional/ International Integration : Demarcation of a country based on calling patterns



Example from the United Kingdom
Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Strogatz, S. H. (2010). Redrawing the map of Great Britain from a network of human interactions. PloS One, 5(12), e14248.

doi:10.1371/journal.pone.0014248

Challenges of trying to leverage big data for development

- Getting access to data
 - Most private data sources will be through ad hoc means, though that is changing
 - E.g. Orange Data for Development competition & Telecom Italia Big Data challenge
 - LIRNEasia and others working towards opening up the playing field for others
- Privacy
 - There are personal and collective privacy implications, difficulty is that:
 - Informed consent is meaningless in a big data world
 - People actually don't really know their privacy needs or how they might evolve
 - Ways to deal with the privacy issue:
 - Anonymization (with caveats)
 - Different levels of disaggregation of shared data
 - Evolution from a rights based approach to a harms based approach

Analytical challenges of working with big data

- Representativeness - volume may make sampling rate irrelevant but doesn't make data representative.
- Ground context - Knowing and understanding real world context is important, otherwise you might make false inferences
- Causation vs. correlation - Big Data analyses can **ONLY** reveal correlation **NOT** causation
- Transparency & replicability – helps catch errors and improve models
- Data provenance - tracing the pathways taken by data from originating source
- Data cleaning – verify quantitative and categorical variables are accurately coded, then remove outliers.
- Behavioral change - Once the studied subject knows the process, they will try to beat it
- The role of 'small data' for verification as well as for bootstrapping - big data is complement and not a substitute for 'small data'