11th World Telecommunication/ICT Indicators Symposium (WTIS-13)

Mexico City, México, 4-6 December 2013

Contribution to WTIS-13

Document C/21-E
6 December 2013

English

SOURCE: LIRNEasia

TITLE: Leveraging Mobile Network Big Data for Development in Sri Lanka

Leveraging Mobile Network Big Data for Development in Sri Lanka

Sriganesh Lokanathan, LIRNE*asia*

WTIS 2013

Mexico City, 6th December 2013

LIRNEasia
Pro-poor. Pro-market.

IDRC CRDI This work was carried out with the aid of a grant from the International Development Research Centre, Canada. Canada

LIRNE*asia*'s exploratory research in 2012-2014

- LIRNE*asia* has negotiated access to telecom network meta-data from multiple operators in Sri Lanka
  - Combined subscriber base of more than 50% (~10 million) of Sri Lanka's population
- Over the course of the two years, we are:
  - Conducting exploratory research on answering a *few* social science questions
  - Developing a framework with privacy and self-regulatory guidelines for the collection, use and sharing of mobile phone data.
- Technical partners:
  - Auton Lab (Carnegie Mellon University) and WSO2 will provide technical and analytical support

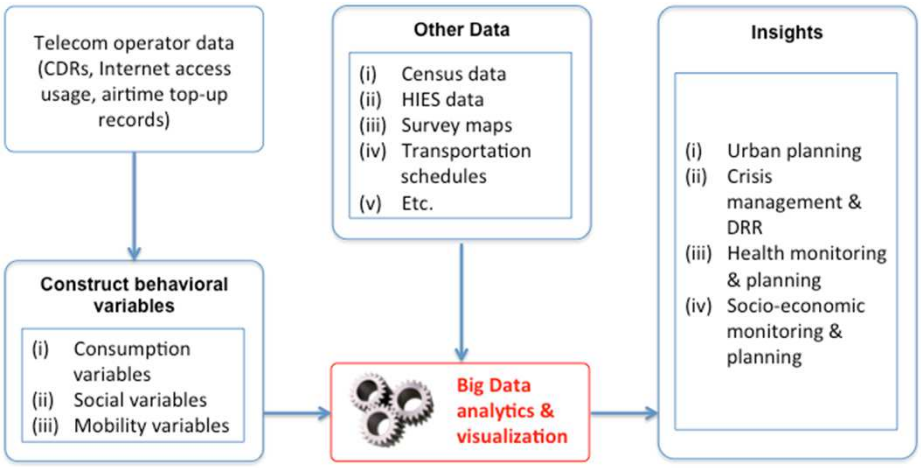
The data sets

- Multiple mobile operators in Sri Lanka have provided LIRNE*asia* access to 4 different types of meta-data:
  - Call Detail Records (CDRs)
  - SMS detail records
  - Internet access records
  - Airtime top-up records
- Data sets do not include any Personally Identifiable Information (PII).
  - All phone numbers are anonymized and
  - LIRNE*asia* does not maintain any mappings of identifiers to original phone numbers

LIRNEasia Pro-poor. Pro-market.

3

Big Data for Development in Sri Lanka: The process



LIRNEasia Pro-poor. Pro-market.

Diagram adapted from Vanessa Frias-Martinez and Enrique Frias-Martinze (2012), http://www.unglobalpulse.org/publicpolicyandcellphonedata

4

Negotiating access

- In retrospect, getting the operator CEOs to say 'yes' was the easy part.
  - Subsequently we had to have multiple meetings with different departments (Legal and regulatory affairs, Marketing, Business intelligence, Network engineers).
- Two common understandable concerns of MNOs:
  - Will my proprietary business intelligence be compromised?
  - Will the regulator have objections?
- In the end we had to sign strict NDAs with each collaborating operator
- Even after access had been negotiated, data extraction required close coordination and quite a few iterations.

LIRNEasia
Pro-poor. Pro-market.

5

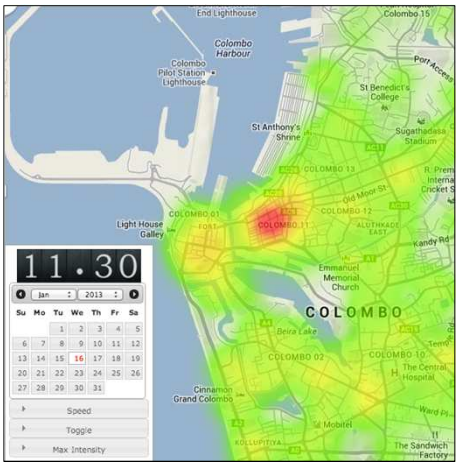
Some questions we are trying to answer via this research

- What is the extent of domicile and employment activities within different localities of Colombo?
- How do population densities of Colombo and its localities vary over time (intra-day, daily, monthly)?
- What is the extent of inhabitant and migrant flows within Colombo, and between Colombo and surrounding regions over time (intra-day, daily, monthly)? What are the corresponding sources and sinks for these flows?
- What is the temporal topology of social ties amongst the habitant population of Colombo (and its localities) and the rest of the country? How does tie strength vary over time?
- How can the above questions be answered for for different socio-economic groups

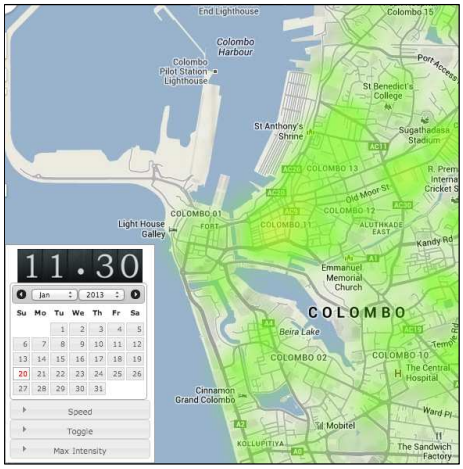
LIRNEasia
Pro-poor. Pro-market.

6

Understanding variations in population densities within cities



11:30am
Wednesday, 16th January 2013



11:30am
Sunday, 20th January 2013

Using telco big data for transportation planning

- We are working with transportation experts In Sri Lanka to test the viability of using Origin Destination (OD) matrices derived from mobile network data.
- Advantages:
  - No need for surveys; temporal snapshots at higher frequency
- Issues:
  - Is the data representative?

LIRNEasia
Pro-poor. Pro-market.

8

Complementing official statistics

- Sri Lanka Census data & Household Income and Expenditure Survey (HIES) data are important inputs to this research
  - For understanding how representative the mobile datasets may be.
  - To bootstrap poverty mapping activities we are exploring using telco big data.
- There are high synergies for close interaction
  - Statistics derived from telco data is cheaper and faster.

BUT

  - The onus is initially on us to show a viable proof of concept.

LIRNEasia Pro-poor. Pro-market.

9

Some challenges we have faced

- Attracting the right people
  - Hiring in-situ Southern computer science graduates to work on broader development focused research is hard.
  - We addressed this by forging partnerships with Universities.
- You don't appreciate the "B" in Big Data till you actually start working with it and experience the resultant infrastructure and methodological challenges.

LIRNEasia Pro-poor. Pro-market.

10

What about privacy concerns?

- As a first step we've ensured that LIRNE*asia* does not have access to any PII
- We are merging the datasets of the different operators:
  - Gives us richer insights while also addressing operator concerns.
- The broader issues are more tricky
  - What does privacy mean?
  - Regulatory rules are ambiguous at times.
    - Are simple rules that just protect PII from being shared sufficient
    - In what way can the operators leverage their own data
- Our approach has been to initiate a consultative process in Sri Lanka, leading to a set of self-regulatory guidelines
  - Still in the early stages

LIRNEasia Pro-poor. Pro-market.

11

In sum

- Mobile operator big data can be a boon for ICT4D
- Negotiating access to telco data is not easy.
  - We are documenting our process so that others can benefit from our experiences
  - Reduction of regulatory ambiguity for ICT4D efforts using telco big data will smoothen a lot of concerns.
- We cannot expect that NSOs and other producers of official data will automatically jump on the bandwagon
  - Having a working viable proof of concept is very important as a first step.

LIRNEasia Pro-poor. Pro-market.

12