

# Mobile network big data for Disaster Risk Reduction

Rohan Samarajiva

UN ESCAP Workshop on knowledge and policy gaps in disaster risk reduction  
and development planning

Bangkok, Thailand

8-9 March 2016



# Mobile networks can contribute in different ways

- Immense potential for early warning because of ubiquity of always-on devices carried on the person
  - E.g., cell broadcasting which, if based on Common Alerting Protocol, can communicate geographically targeted messages and is immune to congestion
- But focus of this presentation is on big data generated as a by-product of the operation of mobile services
  - Our focus is on Mobile Network Big Data (MNBD) because it represents all phones, smart and otherwise

# Mobile Network Big Data is most representative at this time

	Mobile SIMs/100	Internet users/100	Facebook users/100
Myanmar	50	2	12
Bangladesh	76	10	9
Pakistan	73	14	11
India	73	18	9
Sri Lanka	107	26	16
Philippines	112	40	41
Indonesia	125	17	25
Thailand	143	35	49

Source: ITU Measuring Information Society 2014; Facebook advertising portal

# Visitor Location Registry (VLR) & Call Detail Records (CDR) data

- VLR data are generated with no human intervention: every mobile device periodically signals its presence to the nearest base station
  - Massive volumes; not routinely stored
- Call Detail Records (CDRs) are meta data recording actions by customers needed for billing purposes, not content of communications
  - Records of calls
  - SMS
  - Internet access
  - Airtime recharge records
- Both VLR and CDR data provide data not only on “who called whom” type actions, but also of movement through time and space
- Data sets need not include any Personally Identifiable Information
  - LIRNEasia has obtained historical, pseudonymized MNBD from multiple mobile operators in Sri Lanka

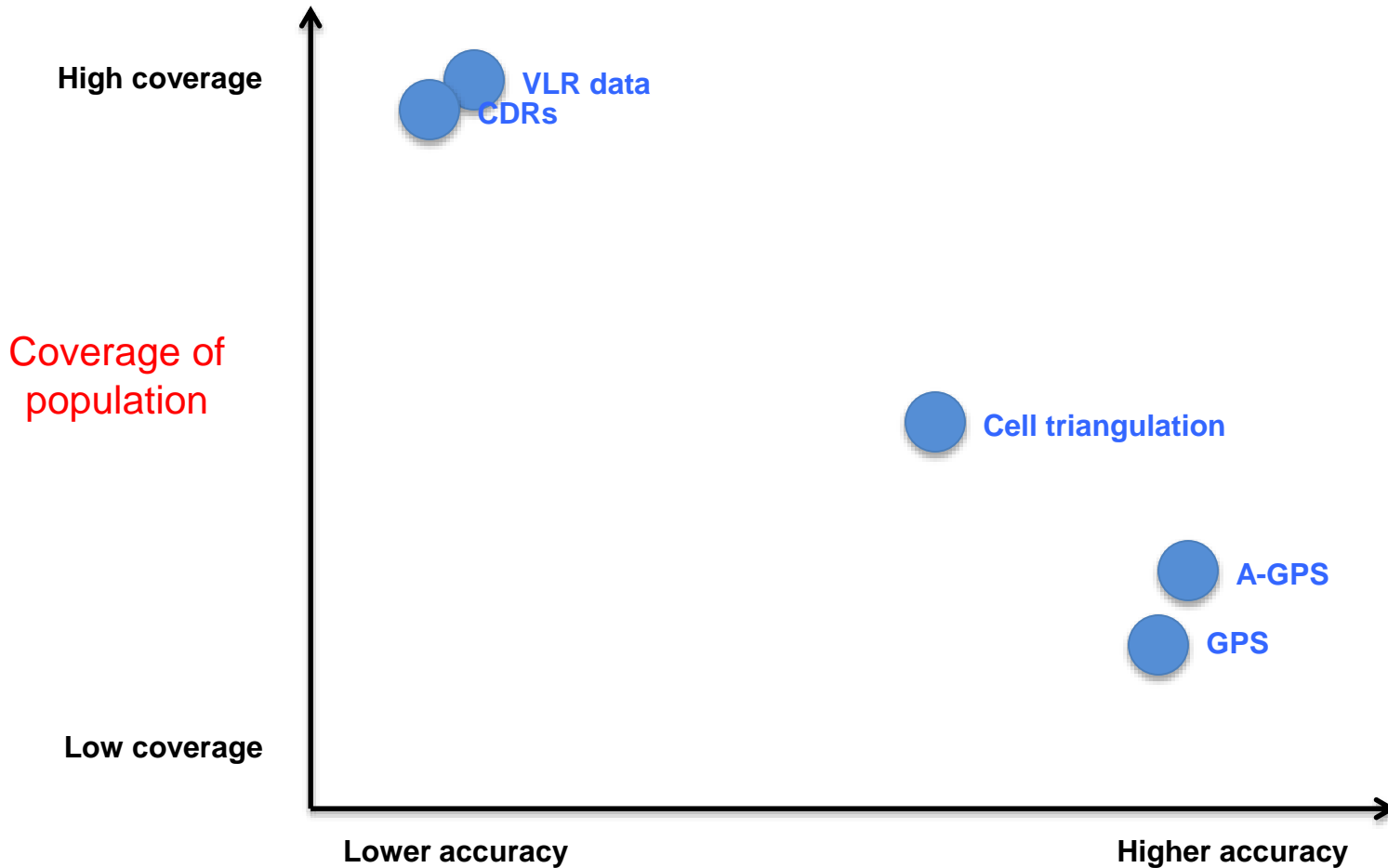
# What MNBD can do for DRR, response and relief

- It can produce high-frequency and granular reports on how people move around
  - Useful for informed urban planning, which would ideally have DRR built in
  - Special insights relevant to spread of infectious diseases
  - Normal movement patterns can be baseline for evacuation planning
    - Abnormal patterns can yield insights
  - Also useful for reconstruction planning
- Can locate congregations quickly after a disaster so relief can be supplied
- Reload data in aggregate can also yield relief & recovery related insights

# How can the location of a mobile subscriber be determined

- GPS
- Assisted GPS (A-GPS)
  - GPS + triangulation using cell towers
- Pure cell triangulation
  - Time of Arrival (ToA)
  - Angle of Arrival (AoA)
- Passive location data (i.e. Cell antenna coordinates)
  - VLR
  - CDRs, etc.

# Most handsets are not GPS enabled in developing economies



# Other limitations

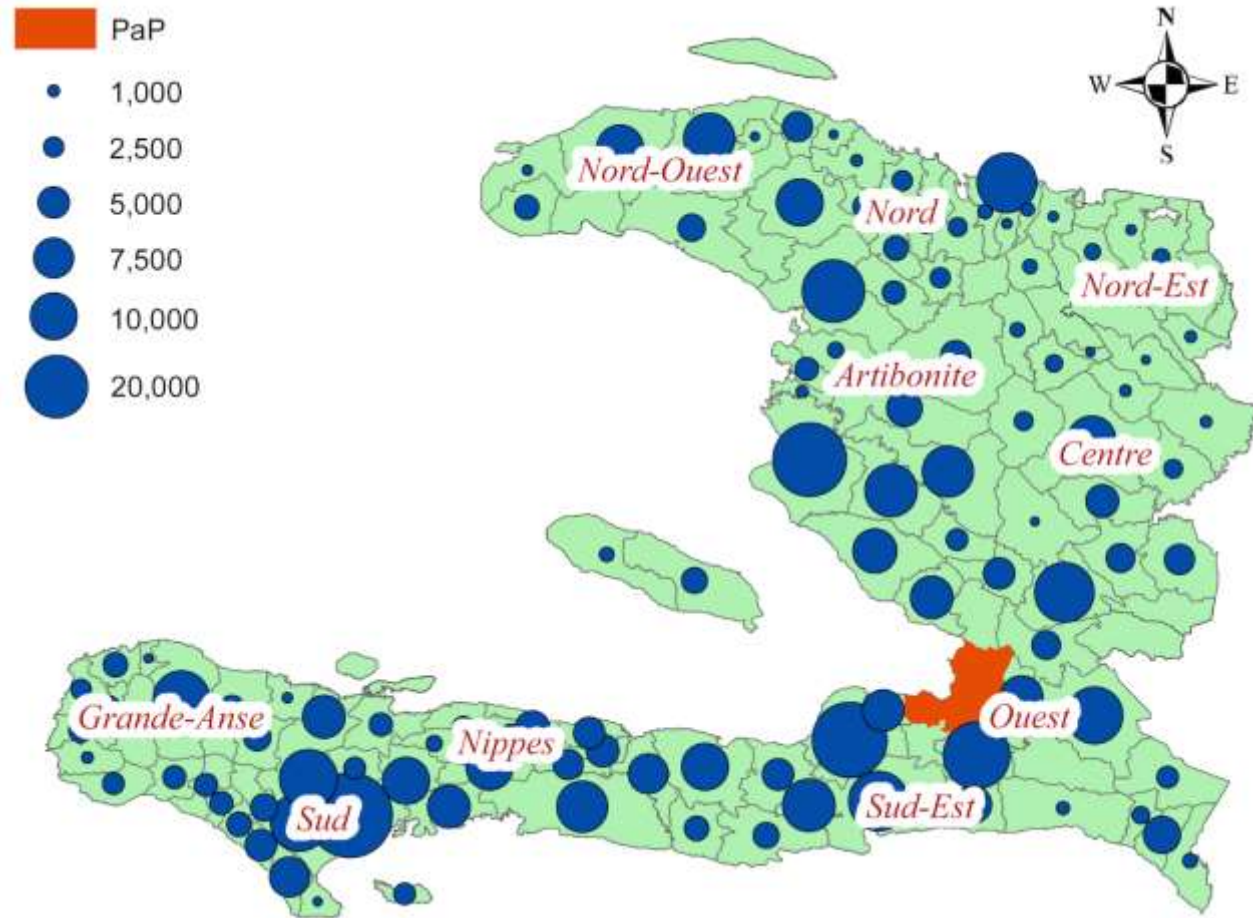
- Continuous triangulation of every phone in a network is not always feasible
- Continuous storage of VLR data is costly since data volumes are huge
- CDR data is event based → there are gaps when trying to understand individual mobility patterns



**BUT DESPITE LIMITATIONS, CDRS  
ARE USEFUL FOR DISASTER  
RESPONSE AND RECOVERY**

# Facilitating recovery efforts: POC from Haiti

Post-earthquake distribution of Port au Prince (Haiti) population after the 2010 earthquake: The figure shows the number of people estimated to have been in Port-au-Prince (PaP) on the day of the 2010 Haiti earthquake, but outside the capital 19 days later. The circles represent the numbers of people who were displaced.



**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

# MNBD for planning: Population density changes in Colombo region

Pictures depict the change in population density at a particular time relative to midnight

**Weekday**

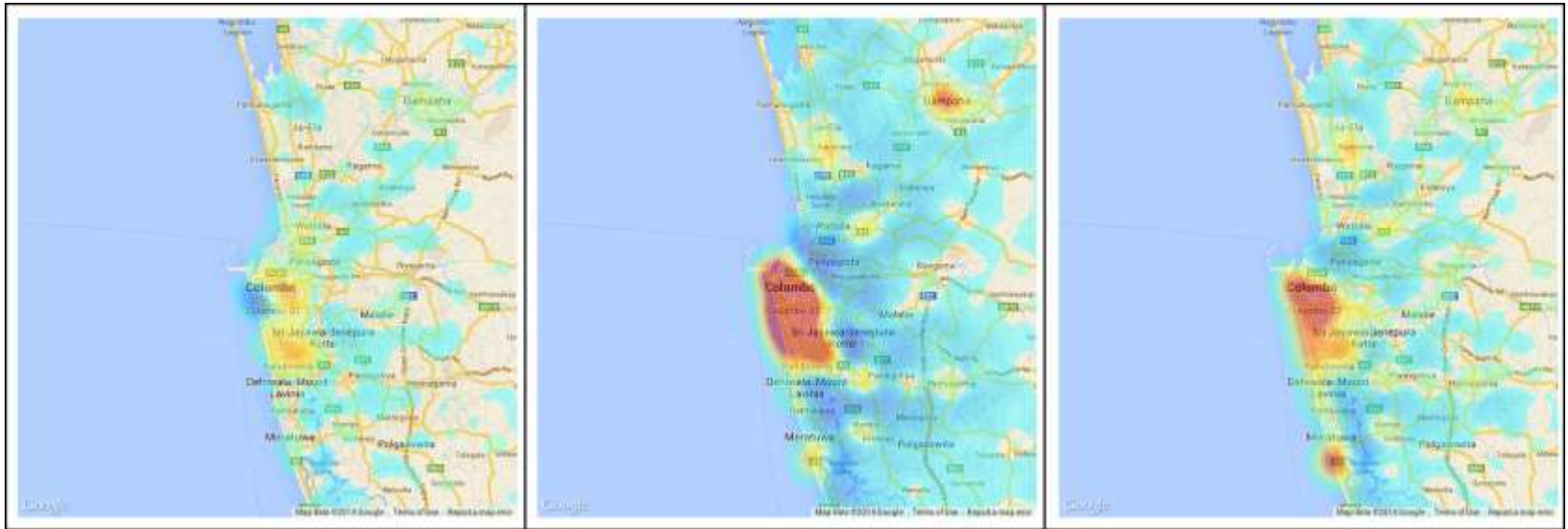


**Time 06:30**

**Time 12:30**

**Time 18:30**

**Sunday**

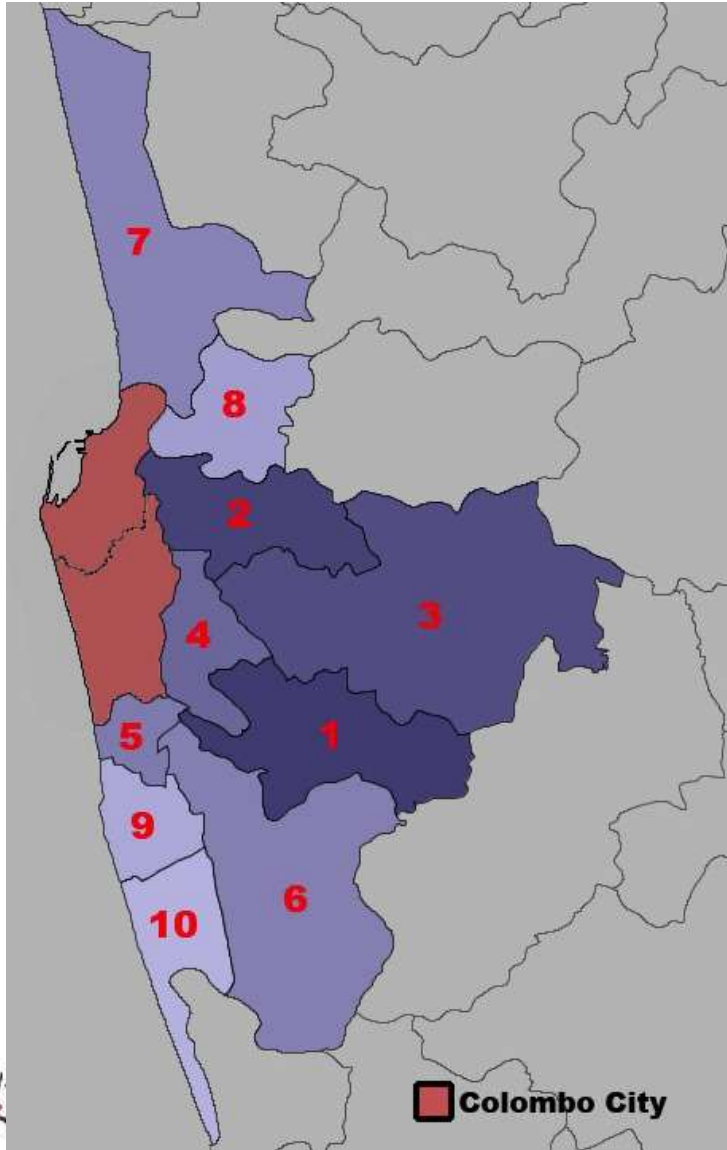


**Decrease in Density**

**Increase in Density**



# MNBD for planning: 46.9% of **Colombo City's** daytime population comes from the surrounding regions



Colombo city is made up of Colombo and Thimbirigasyaya DSDs

Home DSD	%age of Colombo's daytime population
Colombo city	53.1
1. Maharagama	3.7
2. Kolonnawa	3.5
3. Kaduwela	3.3
4. Sri Jayawardanapura Kotte	2.9
5. Dehiwala	2.6
6. Kesbewa	2.5
7. Wattala	2.5
8. Kelaniya	2.1
9. Ratmalana	2.0
10. Moratuwa	1.8

# **SPREAD OF INFECTIOUS DISEASES**

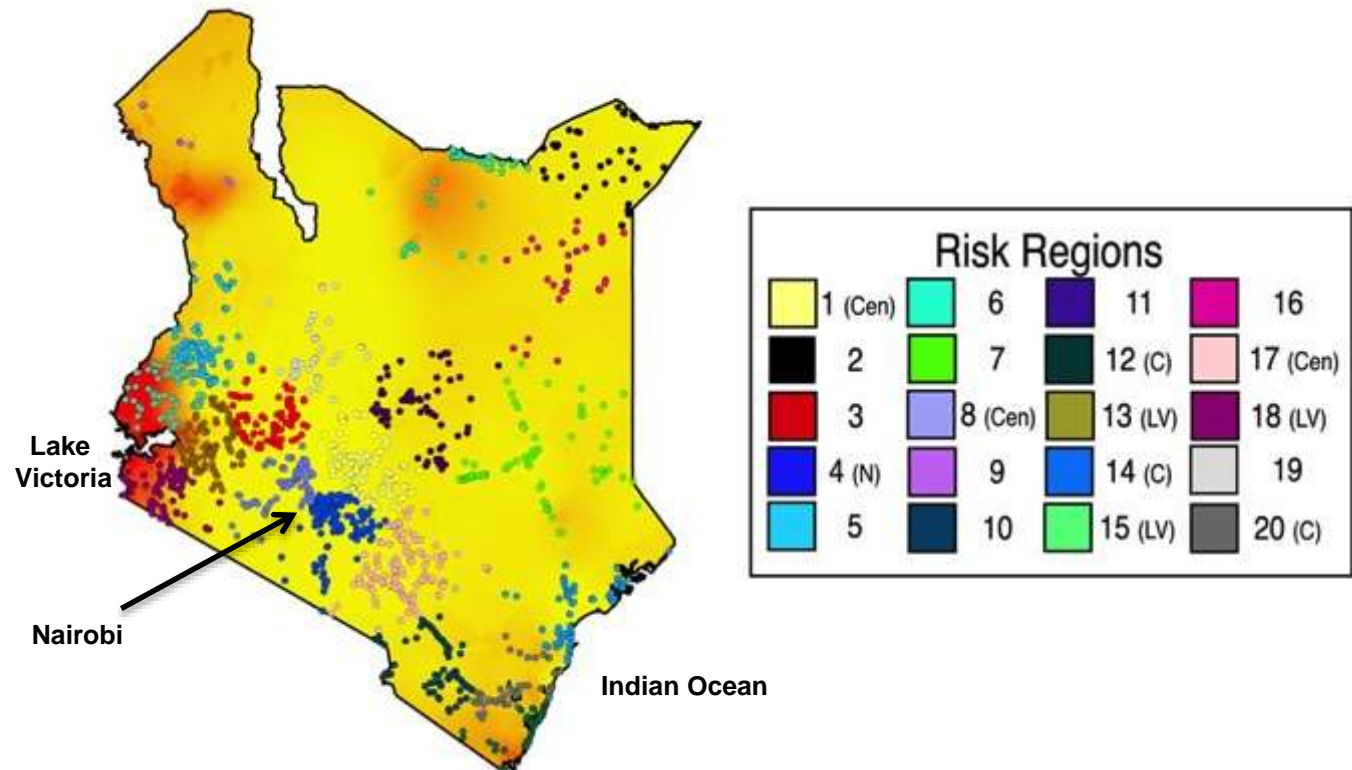
# Relevance of MNDB for tracking/predicting spread of infectious diseases

- Infectious diseases, even vector-borne diseases, are propagated by movement of people
  - Mosquitos travel less than 500 m
  - It is infected human who travel to new locations and infect mosquitos



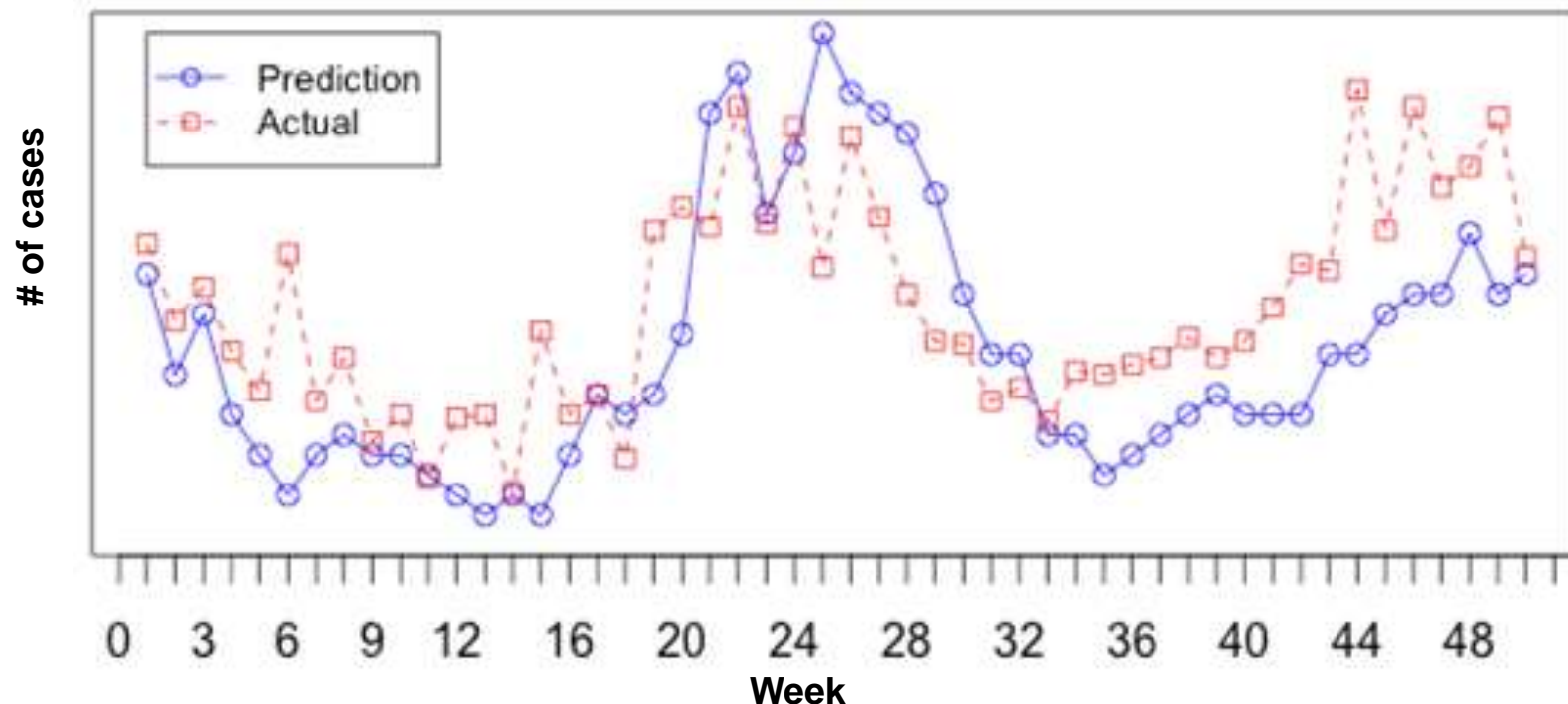
# MNBD can help track the spread of infectious diseases (e.g. malaria in Kenya)

The figure represents a visual mapping of malaria transmission routes. Each settlement was allocated to one of twenty regions by a clustering based on homogenous malaria risk and geography (the different color dots). Regions near Lake Victoria (LV), in Nairobi (Nairobi), the central areas (Cen), and along the coast (C) are labeled accordingly. The background yellow and red colors, represents areas of high malaria prevalence (in red) and those of low prevalence (in yellow)



# At LIRNEasia our initial results in predicting dengue outbreaks in Sri Lanka are promising

Comparing predicted & actual dengue outbreaks for Colombo city in 2014



- Such analyses will be very important should any cases of zika be detected in Sri Lanka



**WHERE AND WHEN DO PEOPLE CONGREGATE?**

# Why is it important?

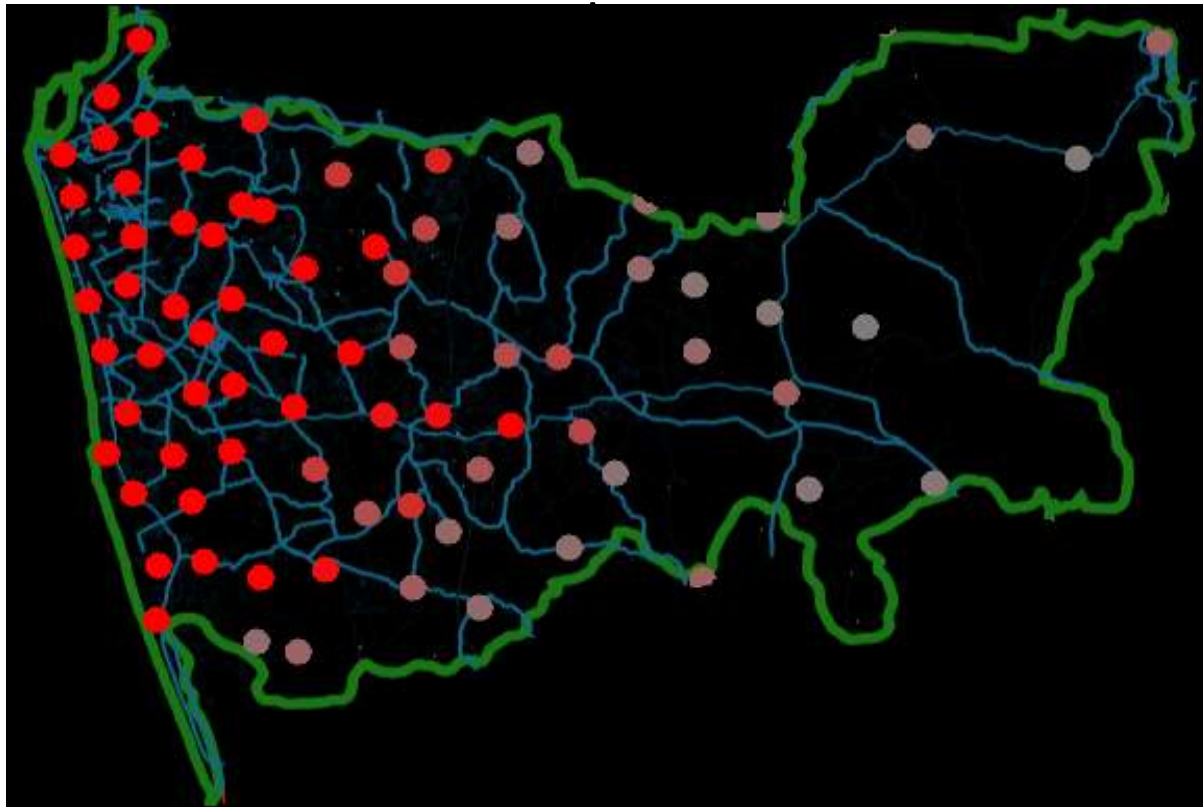
- Persons staying at a location for a significant time consume resources/goods/services from there
- Availability of facilities and resources must correspond to volume of people at a location at specific times
  - Knowing where people congregate is also useful for deciding locations of new service outlets

# Using mobile network big data to identify congregations

- We identify where individuals have spent a minimum period with reasonable confidence (Stays)
  - A *stay* is a sequence of events for a user,
    - which include instances of voice, GPRS and SMS usage
    - within a given distance of each other (1km)
    - spanning a given minimum period of time ( 10 minutes)
    - with consecutive events no more than a given time apart (1 hour)
  - A stay is represented as: **(user,latitude,longitude,start,end)**
- We aggregate stays at the population level to identify Points of Interest (POI), in this case congregations

# Where and when do people congregate in Colombo District?

During 9am – 10am on a weekday



Low  High  
Density of People

# Potential of research on congregations

- Techniques can be developed on the basis of historical MNDB
- Baseline information can be helpful for evacuation planning
- In the post-disaster relief and response phase, locating congregations can be very useful
  - However, this would require near real-time MNDB

# Challenges in using big data for DRR and DRM

- Gaining access
  - Often data belong to private organizations and are outside scope of open data policies
  - Considerable transaction costs including privacy concerns
- Data must be contextualized and analyzed by multi-disciplinary teams
  - To best of our knowledge LIRNEasia and UN Global Pulse Jakarta & Kampala are only South-based entities that work on big data in the Global South
- Data analytics skills need to be developed
- Hardware and software costs are low