

An aerial night photograph of a city, showing a dense network of lights from buildings and streets. The city is partially obscured by white clouds in the foreground. The sky is dark with a hint of orange light from the setting or rising sun.

Raising Awareness: Mobile Phone Data (MPD)

UN-CEBD Task Team
on Mobile Phone Data

Contents



1. What is the Raising Awareness course?
2. Course structure
3. Introduction to the learning materials
4. Future plans



**What is the
raising
awareness
course?**



Purpose

- Introduce the **key concepts** underpinning the use of MPD
- Highlight the **opportunities** presented by MPD
- Explain the **strengths and limitations** of MPD
- Outline how to **run a successful MPD project**

**Support you to start your journey towards using
MPD for official statistics**





Course structure



Structure

What is Mobile Phone Data?

- Different types of MPD
- How MPD are generated
- Information contained in MPD
- Why use MPD?
 - Strengths
 - Limitations

Applications of Mobile Phone Data

- 5 case studies
- Covering:
 - Transportation
 - Public health
 - Disaster response
 - Tourism
 - Information society

How to Manage Mobile Phone Data Projects

- Obtaining MPD
- Requirements for working with MPD
- Preserving individual privacy

End of course quiz and certification



Content



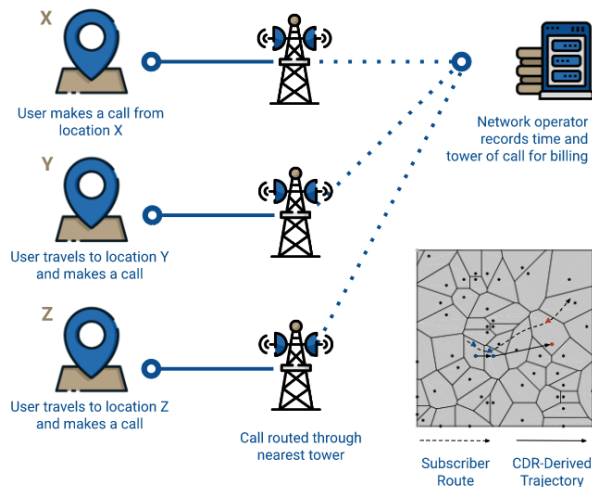
What is Mobile Phone Data (MPD)?

How are CDRs generated?

Call Detail Records (CDRs) are routinely produced by mobile network operators (MNOs) for billing purposes. Each time a subscriber makes or receives a call, sends or receives a text message, or uses mobile data (collectively referred to as “network events”), a new record is generated.

Each record contains the type of network event and the subscriber’s identifier, as well as the time of the event and the cell site it was routed through. The MNO stores these records in a database and may connect them with other information on the subscriber and their account, to inform their billing process. These data can be analyzed to better understand how people move within a country.

As CDR data record the cell tower a network event was routed through, and not the position of the subscriber’s device, the spatial resolution of the data is limited by the density of cell towers. Similarly, the temporal resolution of the data is determined by the frequency with which subscribers use their mobile devices which is affected by a range of factors. We will discuss these limitations and how they can be addressed further in the limitations section.



- What are the **different types** of MPD?
- How are MPD **generated**?
- **What information** is contained in MPD?
- What are the **strengths and limitations** of MPD?



Applications of MPD

Case Study 2 - Informing the government response to a health crisis in The Gambia

Situation:

In The Gambia, the COVID-19 onset in the mid of March 2020. As an immediate response to prevent the spread of the disease, the government imposed a social-distancing policy, a state of emergency was declared, and tourism economy collapsed due to the shut down of international borders. As these drove up unemployment, many migrants returned to their home villages, creating an urban exodus. Trade and travel within the country were reduced to the strict minimum, as authorities enforced restrictions on movement.

Problem:

- External shocks, such as environmental hazards and pandemics trigger population movements and displacements. In response, decision-makers require information about origins and destinations of migrations/relocations to inform rapid policy responses.
- Survey and administrative data have shortcomings which complicate the assessment of crisis situations. Traditional data can provide detailed information on the situation of people but are likely to be outdated during a quickly evolving crisis situation, and rapid data collection is expensive and requires time to coordinate on the ground.

Solution:

- When the COVID-19 hit The Gambia, the team could quickly produce analytics focused on the impact of COVID-19 on patterns of human mobility by utilizing existing infrastructure and code in place under the existing institutional framework among the Gambia Bureau of Statistics, Public Utilities Regulatory Authority, The World Bank, and The University of Tokyo.
- [Standardized mobility indicators](#) proposed by The World bank were used for producing mobility statistics. The team already had a forum to discuss institutional, organizational, or technical challenges through a series of workshop provided prior to the COVID-19 onset.

Outcomes:

- Mobility statistics showed that economic-lockdown measures reduced human mobility and pushed people to leave the capital city region and return to rural areas. The magnitude of impacts was greater and prolonged more in rural small villages where the poorer tend to be concentrated.
- The use case demonstrates how the existing institutional framework and infrastructure can make timely, disaggregated analysis based on CDR data available for quick decision-making.
- Along with the CDR data analysis, [a series of high-frequency phone surveys](#) were conducted to monitor the effect of COVID-19 and government intervention. CDR data provided quick, overall trends and phone surveys generated detailed insights about people later. Both results showed that the most suffered were the poorest and they were reported to have benefited most from the government cash transfer program.

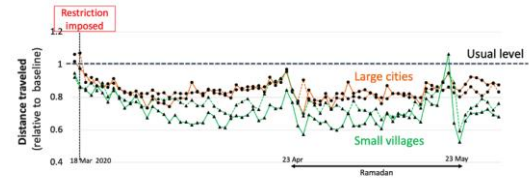


Figure. Daily median distance traveled at the district level is used as the proxy of the magnitude of mobility level. In this chart the daily median distance is presented as the relative value to the baseline period. Baseline is the average of two weeks before the COVID-19 onset. For more information, you can read [this article](#).

Source: Arai et al. (2021)

5 Case Studies

- **Commuters** in Tartu, Estonia
- Government response to a **health crisis** in The Gambia
- **Emergency response to an earthquake** in Haiti
- **Tourism statistics** in Estonia and Indonesia
- **Information Society statistics** in Indonesia and Brazil



How to Manage MPD Projects

How to secure MPD and preserve individual privacy?

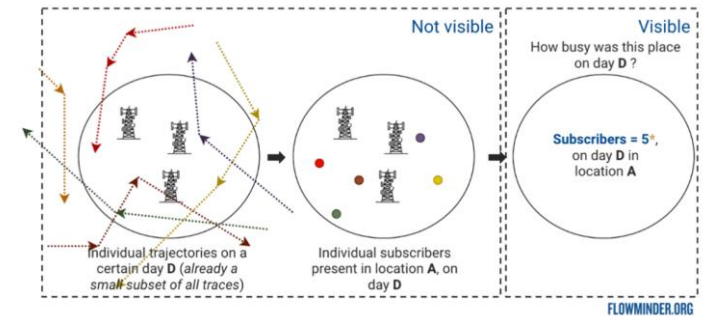
MPD data are personal data, and as such are sensitive information which can potentially have severe impacts on individuals if publicly shared without precautions. Mobile network data, including **call detail records** (CDRs), are also **commercially-sensitive** from the perspective of the **mobile network operator** (MNO). Thus **MPD data must be properly handled and have numerous robust protections**, including data anonymisation and data security.

Directly identifying information (such as a name, or a telephone number for example) linked to MPD events must be removed from the data. This process is often called **"pseudonymisation"**, where the same string of random characters is applied to the same identifier across the dataset (see the figure below). This enables us to still derive insights about subscriber mobility without storing information which can be used to directly identify an individual.

However, the **uniqueness and regularity of human mobility** means that mobile data can act like a **fingerprint** and could be used to reidentify an individual, even with the directly identifying information removed as described above.

MPD data is therefore also **aggregated** spatially (e.g. by district or region) and/or temporally (e.g. by day, month), depending on the requirements of the analytical outputs. The movements of a single individual thus **cannot be directly distinguished** from the analysis result. Once aggregated, MPD data is anonymised provided each data point in the aggregated data set represents a large enough number of subscribers. In cases where some data points represent only a small number of subscribers, the results should be **redacted** to preserve the individual privacy of subscribers.

The data analysis environment also needs to be **secure** and protect the data. The data should be **encrypted** and **stored behind a firewall**. In addition, access to the data (and permissions associated with specific algorithms or analytics) should be **carefully managed and controlled**. A platform requiring username and passwords can for example be used to select particular users, verify their identity, and grant them particular authorisations on the data. This information can also be logged, so as to have a history of data access for **auditing** in case of problems.



- How and where to **obtain access** to MPD
- How to **preserve the individual privacy** of subscribers



End Quiz

When are [call detail records](#) (CDRs) generated? When a subscriber (tick all that apply)

- a. sends/receives SMS messages
- b. uses mobile data
- c. connects to a tower
- d. turns on their device
- e. makes or receives a phone call

- 4 questions
- 75% (3 out of 4) to pass
- 3 attempts



The UN Committee of Experts on Big Data and Data Science
for Official Statistics awards this certificate to:

Thomas Smallwood

for the successful completion of the e-learning course:

Mobile Phone Data - Awareness Course

via the e-learning system on the
United Nations Global Platform

February 16, 2023

Risenga Maluleke
Chair
UN Committee of Experts on Big Data and
Data Science for Official Statistics

Stefan Schweinfest
Director
Statistics Division/UNDESA

Verification code: **ekh60b37s**



Where to find the course

Learning Hub

UN Global Platform Learning Management System

Supporting a series of e-learning courses on various statistical and Big Data topics

Available through the UN Global Platform LMS

(learning.officialstatistics.org)



Future plans



Future courses

- Next course: **Managing an MPD Project**
- More **in-depth courses** for each of the guidelines
 - Dynamic populations and censuses
 - Tourism
 - Migration
 - Information society
 - Displacement and disaster
 - Transport and commuting



Future courses

What courses would you find useful?



UN-CEBD
Task Team on
Mobile Phone Data



Partners



FLOWMINDER.ORG



<https://unstats.un.org/bigdata/>