FLOWMINDER.ORG

Enabling MNOs to produce mobility indicators to support the COVID-19 response

Resources for Mobile Network Operators

Executive Summary

Anonymised mobile operator data can support interventions against COVID-19 by improving our understanding of population movements. The data can be analysed in near real-time and provide an overview of mobility patterns across an entire country. To enable these aggregated, anonymous data to be leveraged more widely to assist with the COVID-19 response, Flowminder is assisting mobile network operators with the production of mobility indicators, whilst preserving the privacy of mobile network subscribers. These indicators can be used by governments and health workers to inform response efforts. Flowminder is providing a repository of code and analysis guidelines that will enable mobile network operators to produce these indicators. The resources will be free, and released under the open access MPL license. Flowminder will update the resources during the course of the pandemic.

Français / French: *Les données de téléphonie mobile anonymisées peuvent soutenir les interventions contre COVID-19 en améliorant notre compréhension des mouvements de population. Ces données peuvent être analysées en temps quasi réel et donner un aperçu des schémas de mobilité à travers un pays entier. Pour permettre à ces données agrégées et anonymisées d'être exploitées plus largement pour aider à la réponse contre COVID-19, Flowminder met à disponibilité des opérateurs de réseaux mobiles des ressources permettant la production d'indicateurs de mobilité, tout en préservant la confidentialité des données du réseau mobile. Ces indicateurs peuvent ensuite être utilisés par les gouvernements et les agents de santé pour guider les efforts de réponse. Flowminder fournit un répertoire de code ainsi que des manuels d'instructions et plan d'analyse qui permettront aux opérateurs de produire ces indicateurs. Les ressources seront gratuites et distribuées sous la licence d'accès libre MPL. Flowminder mettra à jour les ressources au cours de l'épidémie.*

Español / Spanish: *Los datos anónimos de los teléfonos móviles pueden respaldar las intervenciones contra COVID-19, mejorando nuestra comprensión de los movimientos de población. Los datos se pueden analizar casi en tiempo real y dar una visión general de las pautas de movilidad en un país entero. Para permitir que estos datos agregados y anónimos se utilicen más ampliamente para ayudar a la respuesta a COVID-19, Flowminder pone a disposición de los operadores de redes móviles recursos que permiten la producción de indicadores de movilidad, siempre garantizando que se preserve la confidencialidad de los suscriptores de la red móvil. Estos indicadores pueden ser utilizados por los gobiernos y los trabajadores de salud para informar los esfuerzos de respuesta. Flowminder provee un repositorio de código y pautas de análisis que permitirán a los operadores producir estos indicadores. Los recursos serán gratuitos y se distribuirán bajo la licencia de acceso abierto MPL. Flowminder actualizará los recursos durante la pandemia.*

How can mobile phone data be used to inform COVID-19 response efforts?

Analysing anonymised mobile phone data has been shown to be a valuable way of improving our understanding of the spread of many infectious diseases. It is an effective way of monitoring the effect of travel restrictions during disease outbreaks[1] and the data can also be used as an input to epidemiological models to predict how interventions will affect the spread of a disease[2] [3].

'Call Detail Records' (CDRs) and mobility estimates

The mobile phone data that are most commonly used for studying diseases in low- and middle-income countries are 'Call Detail Records' (CDRs), which can be leveraged to generate insights about the mobility behaviour of the population of mobile phone subscribers.

Estimates of the amount of time people spend away from home, the number of locations people visit in a day, and the volume of people travelling between two regions can be derived from the data. These estimates can be calculated on a daily basis, typically providing timely insights into what happened 24 to 48 hours ago. Analysis is performed in a way that guarantees that the privacy of subscribers is fully respected at all times, and that no information about any individual subscriber is disclosed. Only locational information is used, and not the content of any calls, text messages, or data sessions.

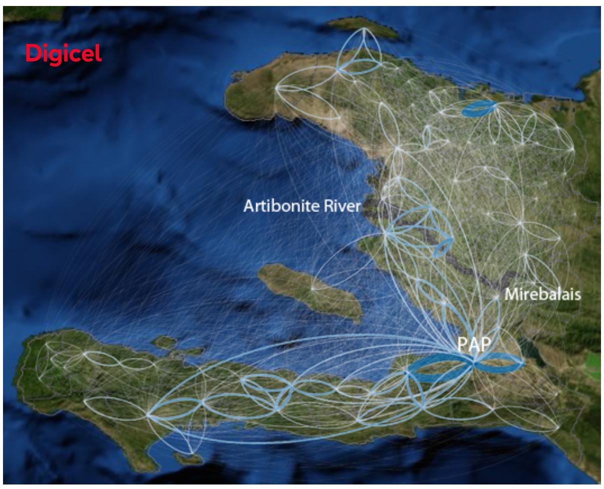
Applications

Monitoring mobility behaviour

COVID-19 spreads most rapidly when people move and interact a lot. In order to reduce the spread of COVID-19, many governments have imposed advisory or mandatory restrictions on travel and social gatherings, and enforced the closure of public places and services. Information can be derived from CDRs to indicate the degree to which these measures are being adhered to. Examples of questions from governments that can be informed by CDR analysis are:

- How much less are people travelling as a result of the restrictions?
- How much time are people still spending away from home?
- How much travel is still happening between administrative regions?
- Has the number of people visiting 'hot-spots', such as the Central Business District of a city, decreased?

Modelling disease spread



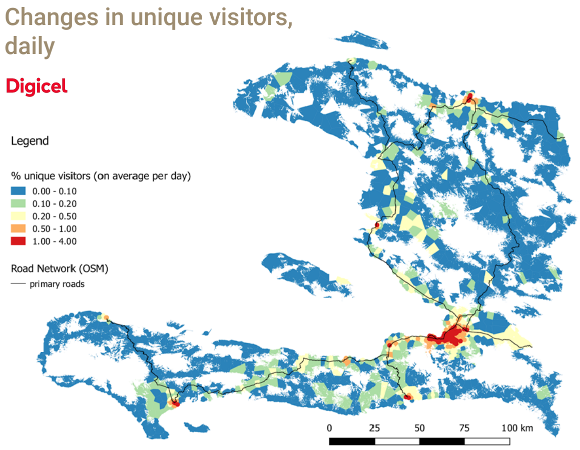
Anonymous, aggregated CDR data can also be passed on to epidemiologists and public health professionals. In countries where data on reported COVID-19 cases are considered representative of the epidemic situations, these experts can combine the aggregated CDR outputs with data in order to establish in what way, and to what extent, travel restrictions and other mobility related interventions are affecting the spread of the virus. They can also simulate how the disease may spread in a number of scenarios where different interventions have been imposed[4] .

Map: Mobility patterns in Haiti during the 2010 cholera epidemic. Thicker, bluer lines indicate larger number of travellers[5].

Identifying locations with high population mixing

CDR analysis can be used to proactively identify locations where it may be beneficial to prioritise communications and interventions. These are locations that are frequently visited by a large number of different people, which means that there is likely to be a lot of 'mixing' between different groups that otherwise would not come into contact with one another.

Average number of daily visitors as a proportion of total subscribers, Haiti, 2016



This map of Haiti shows the number of visitors to each region during a day. Places with larger numbers of visitors have larger population mixing. To target areas for specific interventions, data outputs could be provided per mobile cell tower area.

Flowminder resources for mobile network operators: supporting the production of data insights

Flowminder will assist mobile network operators to generate outputs from their CDR data that can be used by governments, public health professionals, and other relevant experts to respond to the COVID-19 outbreak.

We have initially focused on producing code and guidelines for mobile network operators that may have limited technical capacity, especially those in low- and middle-income countries. This means that the code we have provided is both simple to modify, and also should not be extremely computationally intensive to run. However, we will continually add more resources to the repository, including material that is suitable for settings where more capacity is available. In these cases, it will be possible to produce more complex outputs and analyses.

Repository of code and guidelines

We have created a repository of code and guidelines that is freely available if you wish to use your data to help inform the COVID-19 response effort. The repository currently contains code and guidance to enable your organisation to produce a set of basic, general-purpose outputs that can be used as described in the previous section. We will shortly also be adding guidelines on how to analyse these outputs, and incorporate any feedback that we receive from mobile network operators and relevant experts.

Additional support

In some instances, we will be able to provide remote support, via video-conferencing, to help you adapt the code to run on your systems. We may also be able to assist with initial analyses of the outputs. We welcome requests, but given the scale of the outbreak we may not be able to support all requests. It is also recommended that any advanced analyses be performed by epidemiology experts with specific knowledge of COVID-19, and the country context, in order to ensure that the most accurate, relevant, and valuable information is produced in a timely manner.

Privacy and confidentiality

All of the outputs produced by Flowminder's code are aggregated data, meaning that they do not contain any information about individual subscribers. Further guidance on how to ensure privacy of subscribers is included in the Github repository. Flowminder staff will be available to provide support and advice if there are any concerns in this area, or concerns regarding the possible disclosure of commercially sensitive information.

About Flowminder

Flowminder is a non-profit foundation that specialises in the analysis of anonymised mobile phone data, satellite imagery, and household survey data for humanitarian and international development purposes. Flowminder was the first organisation to respond to a large-scale infectious disease epidemic (the Haiti cholera outbreak in 2010) using mobile operator data, and Flowminder's researchers were the first to show that mobile operator data can predict the spatial spread of an infectious disease[5]. The foundation provides information to governments, non-governmental organisations, inter-governmental organisations, and researchers, and has developed sustainable partnerships with numerous mobile network operators in low- and middle-income countries. Flowminder's multidisciplinary team of experts includes epidemiological researchers, data scientists, software developers, and humanitarian practitioners.

References

[1] Peak CM, et al. Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data, *International Journal of Epidemiology,* Volume 47, Issue 5, October 2018, Pages 1562–1570, https://doi.org/10.1093/ije/dyy095

[2] Wesolowski A, et al. Commentary: Containing the Ebola Outbreak – the Potential and Challenge of Mobile Network Data. *PLOS Currents Outbreaks*. 2014 Sep 29 . Edition 1. doi: 10.1371/currents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e.

[3] Buckee CO, et al. Aggregated mobility data could help fight COVID-19. *Science* 23 Mar 2020: eabb8021, DOI: 10.1126/science.abb8021

[4] Lai S, et al. Effect of non-pharmaceutical interventions for containing the COVID-19 outbreak in China, *medRxiv* 2020.03.03.20029843; doi: https://doi.org/10.1101/2020.03.03.20029843

[5] Bengtsson L, et al. Using Mobile Phone Data to Predict the Spatial Spread of Cholera. Sci Rep 5, 8923 (2015). https://doi.org/10.1038/srep08923

Appendices (next page)

- Appendix 1 - Description of CDR outputs
- Appendix 2 - 'Call Detail Records' (CDRs) and mobility estimates
- Appendix 3 - Limitations of CDR data

To access the code repository, please visit:
github.com/Flowminder/COVID-19

Appendices

Appendix 1 – Description of CDR outputs

Flowminder's repository contains code templates to produce the outputs described in the table below. The templates are accompanied by detailed guidelines on how to use and adapt the code, and descriptions of suggested use cases for each output. Flowminder will update and add to the content of the repository in response to feedback from mobile network operators and the epidemiological community. We will additionally share further information on how the above indicators can be analysed over time and combined with each other to suit the needs of end-users. We welcome suggestions and feedback.

	Output	Description
1	Number of unique subscribers that use their phone at each location[1], each day.	This indicates how travel restrictions are affecting each location, in terms of the number of people that spend time there each day.
2	Number of 'non-residents'[2] (visitors) that use their phone at each location[1], each day.	This is similar to (1), but counts only the people at each location that do not live at that location, and therefore are 'visitors'. In cases where travel is severely restricted, the number of non-residents seen at each location should drop to close to zero.
3	Number of unique subscribers that use their phone at each location[1], each week.	This can be combined with (1) to identify locations where a lot of 'mixing' may occur between different groups of people: A location may experience a high number of visitors each day, with the same people visiting each day. Then there may be locations which receive fewer visitors each day, but because different people visit each day, the number of distinct visitors seen over a week is higher. This would indicate that the second location potentially experiences a higher degree of mixing than the first location.
4	Number of 'non-residents'[2] (visitors) that use their phone at each location[1], each week.	Similar to (3), but excluding residents from the count in order to establish the degree of 'outside traffic' from 'visitors'.
5	Number of unique subscribers moving between each pair of locations[1] each day.	This indicates how 'connected' each pair of regions is, by counting how many subscribers are seen at each pair of locations within a day.

[1] A 'location' refers to an area such as an administrative region, that is specified by defined boundaries.

[2] Each subscriber is assigned to a 'home location' based on the region where they most frequently use their phone for the last time each day. A subscriber is counted as a 'non-resident' when they use their phone in a region that is not their home location.

Appendix 2 - ‘Call Detail Records’ (CDRs) and mobility estimates

The mobile phone data that are most commonly used for studying diseases in low- and middle-income countries are ‘Call Detail Records’ (CDRs). CDRs are generated each time a mobile phone subscriber makes or receives a call, sends or receives a SMS, or uses mobile data. Each record includes an anonymous identifier of the subscriber, a timestamp, and the cell tower that the transaction was routed through. The location of the cell tower can be assumed to be a good proxy for the location of the subscriber. A CDR dataset therefore contains a history of each subscriber’s movements, which can be leveraged to generate insights about the mobility behaviour of the population of mobile phone subscribers.

Appendix 3 - Limitations of CDR data

CDR data from a mobile network operator only contains information about the people who use a SIM card from that operator. This subset of people is unlikely to be perfectly representative of the entire population of a country because not everyone uses a mobile phone, and not all mobile phone users are subscribed to a single mobile network operator. For example, young children and elderly people in many countries do not use a mobile phone, and in low-income countries, individuals in the lowest socioeconomic strata may not own a mobile phone. Additionally, the geographic, demographic, and socioeconomic distribution of subscribers to each mobile network operator is often not representative of the full population.

In spite of its limitations, CDR data can still provide a good indicator of any changes in a population’s mobility behaviour because a significant proportion of the population are included in the dataset. Whilst further work is needed to understand the extent to which CDR data are biased, and the effect that has on an analysis, previous studies have shown that mobility estimates are not strongly affected by the biases in a dataset[1] [2]. CDR data also has the significant advantage that it can be analysed to generate near ‘real-time’ insights, with a typical lag of between just a few hours to a few days due to processing times.

[1] Bengtsson L, et al. Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti, *PLoS Medicine,* August 2011 | Volume 8 | Issue 8 | e100108

[2] Wesolowski A, et al, 2013. The impact of biases in mobile phone ownership on estimates of human mobility. *J R Soc Interface,* 10: 20120986. http://dx.doi.org/10.1098/rsif.2012.0986

To access the code repository, please visit:

github.com/Flowminder/COVID-19