CAMBRIDGE
UNIVERSITY PRESS

TRANSLATIONAL ARTICLE

Analysis of call detail records to inform the COVID-19 response in Ghana—opportunities and challenges

Tracey Li[1,*] , Rachel Bowers[2], Omar Seidu[2], Gloria Akoto-Bamfo[2], David Bessah[2], Victor Owusu[2] and Laurent Smeets[2]

[1]Flowminder Foundation, Southampton, United Kingdom
[2]Ghana Statistical Service, Accra, Ghana
*Corresponding author. E-mail: tracey.li@flowminder.org

Abstract

Telecommunications data are being explored by many countries as a new source of data that can be incorporated into their national statistical systems. In particular, "mobile positioning data" are increasingly being used to study population movements and population distributions. However, the legal, ethical, and technical complexities of working with this type of data often pose many barriers, which can prevent the data from being used at the times when it is most urgently needed. We demonstrate how having a robust public–private partnership framework, a privacy-preserving technical setup, and a communications strategy already in place, prior to an emergency, can enable governments to harness the advantages of telecommunications data at the times when it is most valuable. However, even once these foundations are in place, the challenges of competing priorities, managing expectations, and maintaining communication with data consumers during a pandemic mean that the potential of the data is not automatically translated into direct impact. This highlights the importance of sensitisation exercises, targeted at potential data users, to make clear the potential and limitations of the data, as well as the importance of being able to maintain direct communication with data users. The views expressed in this work belong solely to the authors and should not be interpreted as the views of their institutions.

Policy Significance Statement

In recent years, the Ghana Statistical Service (GSS) has made efforts toward incorporating telecommunications data into the National Statistical System. The onset of the COVID-19 pandemic and the associated restrictions on traditional in-person surveying methods presented an opportunity to leverage the infrastructure of an existing project to inform the Government's response to the pandemic, particularly regarding nonpharmaceutical interventions. The analysis of telecommunications data provided timely and policy-relevant insights about population movements. Two rapid assessments of the effectiveness of Government-mandated travel restrictions were produced and shared with the Ministry of Health and the Office of the President. This resulted in increased awareness of new data sources within the Government, and the inclusion of the GSS in additional policy spaces.

CrossMark

1. Background

A formal partnership between Ghana Statistical Service (GSS), Vodafone Ghana, and Flowminder has been in place since the start of 2019. The partnership was set up with the aim of building capacity within Ghana to enable aggregated, anonymized call detail records (CDRs) to be incorporated into the production of official data. The project was funded by the Vodafone Foundation and the William and Flora Hewlett Foundation.

Prior to the onset of the COVID-19 pandemic at the start of 2020, the project had already progressed significantly. The following key activities had been completed: (a) The partners had developed the necessary legal and partnership frameworks to enable the project to proceed. This included engaging with the Ghanaian Data Protection Commission to ensure compliance with local data protection legislations, as well as ensuring compliance with the European Union General Data Protection Regulation (EU GDPR) legislation. (b) The technical infrastructure required to securely access and analyze the CDR data in an ethical, privacy-preserving, and legally compliant manner had been set up. This enabled Flowminder staff to analyze the data, via a secure remote connection, so that all data were processed, physically, within Vodafone Ghana's premises. (c) The project team at GSS had started a program of training to acquire the skills to enable them to access and analyze the aggregated, anonymized CDR data, and to understand the potential use cases and limitations of CDR data. (d) Engagement with the ministries, departments, and agencies (MDAs) of the Ghanaian Government, that could be potential consumers of the project outputs, had been ongoing. This included a showcase event in December 2019, where the project was presented to current and prospective stakeholders, and the media. This event, and related public communications, served to disseminate the aims and outputs of the project. The communications also ensured transparency about how the data were being used.

By the start of 2020, the partners were therefore in a position to start identifying use cases, and then to produce and deliver analytic outputs to the Ghanaian Government.

2. COVID-19 Response

2.1. Initial engagement with the Government

In March 2020, 1 week after the first cases of COVID-19 were detected in Ghana, the GSS approached the Minister of Planning to propose how CDR data could inform the Government's response to the pandemic. The Minister for Planning immediately approached the Minister for Health and Minister for Communication to convene a meeting with the GSS team, while he stepped back from further in-person engagements.

A meeting then took place between the GSS, the Ministry of Health, and Ghana Health Service. Frequent partnerships between the GSS and Ghana Health Service, including the existing CDR project, had fostered a strong working relationship between these institutions. This facilitated quick liaison between the institutions, following the submission of initial letters of invitation. During this engagement, GSS described the available CDR dataset and its potential use cases. This prompted concerns over the representativeness of the available data, which comprises only CDRs from Vodafone Ghana, since other major telecommunications agencies are also prevalent across the country. The Ministry for Planning proposed that the Minister for Health could approach the Minister for Communication, who oversees the National Communications Authority, with regard to encouraging other mobile network operators to participate in the initiative in order to improve the representativeness of the available data.

The intention of the GSS in this engagement was to highlight the potential of using anonymized, aggregated CDRs to provide near real-time information about the emerging situation on the ground. However, the Ministry of Health enquired whether the data could be used to track individual phone numbers to assist with enhanced contact tracing. This was an exercise that the Government was conducting for the 20,000 passengers who had entered Ghana by international flight in the 2 weeks before the borders were closed. The GSS explained that the available data did not include phone numbers or any other personal identifiers, and also that the collaboration agreement explicitly disallowed the use of individual-level data, therefore meaning that contact tracing fell outside both the feasible and legal scope

of the project. Since the priority of the political leadership at the time was to engage mobile network operators to facilitate contact tracing, efforts to coordinate with the Ministry of Communication did not proceed, and the Ministry of Health demonstrated only limited interest in the idea of analysis based on the anonymized, aggregated CDR data.

The GSS and Flowminder decided to proceed with the analysis anyway, in order to be able to concretely demonstrate the type of information that it was possible to produce. They worked together to determine the metrics of interest to include in the analysis, and published two reports. The first was published on 3rd April—4 days after the introduction of lockdown restrictions in parts of Ghana, and the second on 15th May—3 weeks after lockdown restrictions had been lifted. The analyses are described in the following section.

2.2. Analysis

There are two fundamental indicators that form the basis of most mobility analyses using CDR data. They are the number of subscribers recorded to have been present at a specified location within a specified time period ("subscriber presence"), and the number of subscribers recorded to have travelled between two specified locations within a specified time period (commonly called "origin-destination matrices" or "transition matrices"). A detailed technical discussion about these indicators, and others that are relevant to the study of mobility and social-distancing interventions, can be found in Kishore et al. (2020). In the context of COVID-19, studying how these two indicators changed on a daily basis, as social-distancing and mobility restrictions were introduced and then lifted, provided insights into how restrictions were affecting the population's behavior. Insights could be obtained for the entire country, and with a time lag of just a few days. The initial report was published just 4 days after lockdown restrictions were introduced. This geographical coverage and timeliness would not have been possible from traditional data collection methods, that is manual surveys.

The first report (Flowminder, Ghana Statistical Service, & Vodafone Ghana, 2020a) included analyses on subscriber presence within regions (administrative Level 1) and districts (administrative Level 2) each day, and travel between regions and between districts each day, for the time period starting 4 weeks before initial social-distancing and mobility restrictions were introduced, until a few days after lockdown measures had been introduced. This showed how population movements changed in response to each set of restrictions. The report is available from GSS's website (https://statsghana.gov.gh/gsspublications. php?category=MTkwMDE4MjI2Ny4xMDg=/webstats/90r897632o; accessed 31 August 2020). A second report (Flowminder, Ghana Statistical Service, & Vodafone Ghana, 2020b) was released 3 weeks after lockdown measures had been lifted. This contained similar analyses to the first report, but extending to the post-lockdown period, and including several comments from the Government Statistician regarding policy implications. The report is available from GSS's website (https://statsghana.gov.gh/gsspublica tions.php?category=MTkwMDE4MjI2Ny4xMDg=/webstats/90r897632o). The reports also include technical details about the analysis and calculation of the studied indicators.

When interpreting the analyses, it is important to acknowledge that the data have known limitations. Firstly, it is known that the dataset will only include information about people that use a Vodafone Ghana SIM card. Therefore, the entire population is not represented in the dataset, and so it is not possible to make precise statistical inferences about the whole population using only that CDR dataset. Secondly, because CDRs are an actively collected telecommunications dataset—meaning that a record is generated only when a subscriber performs an action with their phone, for example makes a call—there are data available only when subscribers are active. If we observe that overall activity levels have remained constant across the period of study, then it is a reasonable assumption that any observed changes in mobility are due to actual changes in mobility. If, however, we observe that activity levels have changed during the period of study, then it is important to check whether those changes in activity are correlated with observed changes in mobility, as this would indicate that the observed changes in mobility may be artefacts of the changes in activity levels (e.g., more activity means that more movement will be observed). If so, this needs to be taken into account when drawing conclusions from the analysis.

In the current case, it was found that there were no significant correlations between activity levels and observed mobility levels.

In spite of the inherent limitations of the data, the insights provided by the data in this instance could not be produced at the same geographic scale, or in as timely a manner, by any other data source. Deeper studies, including the analysis of survey data, can indicate the degree to which the movements of the entire population are represented by the insights drawn from a CDR dataset. This will be specific to each dataset and context—see, for example, Wesolowski et al. (2013) and Arai et al. (2016).

2.3. Dissemination

Several channels were used in the dissemination of the two reports. A remote press launch resulted in 9 Ghanaian media outlets publishing 13 articles about the reports online and one major Ghanaian newspaper (Ghanaian Times) published an interview with the Government Statistician in print. The Global Partnership for Sustainable Development Data also published a case study on the analysis (Global Partnership for Sustainable Development Data (2020a)).

The reports were circulated widely through direct channels to the GSS' Sustainable Development Goals Secretariat's network of partners, which includes staff from the Ghana Health Service, the Ministry of Health, and other Government MDAs, as well as development partners, academics, and mobile network operators. The response was positive. In particular, staff from Ghana Health Service noted the retrospective value of the insights. For example, they compared travel behavior with case numbers, and observed a link between a spike in travel between the regions of Greater Accra and Oti preceding several local outbreaks in Oti—a region where there had previously been no cases. Questions arose about the possibility of applying a gendered lens to the analysis, but this was not possible within the scope of the current partnership agreement since no information on gender was available.

Finally, the GSS actively approached key Government partners. They delivered a printed copy of the report to the Minister of Health's office, and a presentation to the Office of the President. There were some difficulties arranging in-person dissemination meetings due to the number of Ministry of Health staff, including the Minister, having to withdraw into quarantine after testing positive for COVID-19.

3. Outcomes and impact

Following the launch of the first report, the Ministry of Health demonstrated an attitudinal change, becoming more interested in the analysis. They valued that the report gave an indication of whether the population were abiding by the Government-imposed restrictions. They also valued the insight into where people were still travelling to, remarking that this could be used to inform Ghana Health Service about the appropriate distribution of resources for contact tracing activities.

Following the publication of the second report, the GSS were invited to deliver a presentation at the Office of the President on both the CDR mobility analysis and the GSS COVID-19 Dashboard (https://statsghana.maps.arcgis.com/apps/opsdashboard/index.html#/a22ebfb6d9cb47ff9ce87619d53f68e5; accessed 31 August 2020), which was another initiative aimed at informing the response to the pandemic. The lead for technical coordination at the Ministry of Health attended this presentation and concluded that the mobility analysis was useful to inform future steps. An outcome of this engagement was that a senior GSS colleague was invited to join the National Technical Coordinating Committee on COVID-19, which is chaired by the Minister for Health. Appointment to this committee provides a direct platform for the GSS to disseminate evidence about the pandemic to decision makers.

Finally, the publications raised an awareness of how CDRs could be used in conjunction with other data sources to provide further insights. In early June, the Ministry of Health, as part of a consortium of organizations, approached the GSS with a request to integrate the mobility indicators produced in the CDR analysis with their existing epidemiological analysis. The Government Statistician and the Director of Social and Demographic Statistics were invited by the Director of Policy Planning at the Ministry of Health to discuss this request with the consortium. At this time, the Minister for Health tested positive for

COVID-19 and therefore had to self-quarantine. Meetings from that point were moved remotely, which disrupted progress and stalled the initiative. However, this was the first instance of a proactive request for CDR data being made by the Government of Ghana. This indicates the progress in sensitisation that had been made relative to before the pandemic, when the onus had been on the GSS to actively approach ministries and suggest use cases.

While the described engagement, analysis, and dissemination activities do not appear to have immediately shaped any policy decisions, there is evidence that they have elevated the profile of the use of CDRs as an alternate data source within the Government. The activities have also led to the national statistics office (GSS) being granted access to participate in additional policy spaces. This gives the GSS, and therefore official statistics, a more prominent role in the process of evidence-backed decision making in Ghana.

A process is also in place for parties outside of the immediate collaboration to request access to the anonymized, aggregated CDR data used to produce the reports. Parties that plan to use the data for a purpose that contributes to the national development of Ghana can submit a request to the Government Statistician of Ghana via email: gsadmin@statsghana.gov.gh.

4. Lessons Learned

4.1. Legal and technical challenges

This case study demonstrates how CDR data were used to produce a timely response to the COVID-19 pandemic in Ghana. This was possible because the necessary legal framework and technical infrastructure were already in place, prior to the pandemic. Some communications activities had also been conducted. Having these foundations in place enabled analyses to be conducted and disseminated in a timely manner.

In order to reach this point, the partners had to overcome several challenges during the initial stages of the project. They were:

- Constructing a project framework that is compliant with all the relevant data privacy regulations, including the GDPR, and that is also practical from an operational perspective. This was achieved through consultation with a number of legal and technical experts, and engaging with the Ghana Data Protection Commission.
- Designing technical solutions that satisfy all the necessary legal and privacy requirements, and that are feasible to operationalise in the given context. The implemented solution is one in which all personal data are stored and processed, physically, within Vodafone Ghana's premises. Only anonymized, aggregated data are released, and only to approved parties.
- Working across the cultural differences between private-sector and public-sector organizations. It is important to recognize that organizations from different sectors have very different structures, incentives, and requirements, and to accommodate those.
- Managing the perceptions of stakeholders and the public about how the data are being used. This was achieved through careful and transparent communication, to transmit the message that it is aggregated, anonymized data, and not data about any individuals, that are used to provide insights.

If the challenges listed above had had to be dealt with at the stage when the COVID-19 pandemic began, the response would have been much slower—taking months instead of days.

These types of challenges are common in public–private data partnerships, and are discussed as part of the larger public–private data-sharing ecosystem in Global Partnership for Sustainable Development Data (2020b).

4.2. Challenges with Government engagement

Several challenges arose during policy engagement: awareness of the data, bureaucratic processes, and the effects of the pandemic.

Firstly, a lack of awareness within the Government of the existence of indicators on population mobility, derived from aggregated and anonymized CDRs, meant that the GSS had to actively approach ministries to communicate the potential of the data. There was limited time available for engagements with policy makers, and some of that was taken up by communications to sensitise the audience, and outline the scope and limitations of the data. Furthermore, among the ministries who were already aware of the data, they primarily expressed an interest in contact tracing. However, contract tracing activities were not possible with the available data or within the existing project agreement, and correcting these false expectations resulted in a lack of overall interest in what was feasible. A wider sensitisation of ministries, ahead of time, may have prevented some of this confusion and disappointment.

A further institutional challenge was in the protocols of communication between ministries. It is common for new engagements between ministries to be initiated through physical signed letters, as the first step in building trust. This can result in delays due to waiting for letters to be signed, because letters have gone missing, and because of the need for in-person follow-ups to book appointments. The lengthy bureaucratic process meant that during the dissemination of the two analysis reports, the press releases that had originally been planned to follow the ministerial engagements came concurrently, or even before, a face-to-face ministerial dissemination was possible.

Finally, an unexpected challenge was the effect of public health policies on the availability of key stakeholders. The quarantining of two ministers and some staff from the Ministry of Health stilted communications between the GSS and the end-user institutions. This stilting was compounded during the partial lockdown because all remaining engagements were conducted remotely, and the transition to online meetings caused disruption and delays.

4.3. Suggestions to improve preparedness in the future

While the existing setup enabled a reasonably quick response to the need for dynamic sources of data to inform the COVID-19 response, there are three key approaches which would improve future preparedness.

Firstly, further efforts should be invested in sensitising Government MDAs to the potential applications of CDR data. Through a targeted dissemination of existing reports and the production of educational materials, such as videos for a Governmental audience, familiarity with the data can be developed in advance of any emergency situation. This could lead to proactive data requests, clearer understanding of the limitations of the data, and the wider adoption of aggregated, anonymized CDRs as a policy-relevant data source in the future.

Secondly, cautiousness about the public's reaction to the use of mobile phone data led to a lengthy approval process prior to any reports being published. Some initial public communications had been conducted prior to the pandemic, but only to a limited audience. Sign-off from the senior management level was therefore required from the collaboration partners, due to the need to manage reputational risks. To manage the public reaction and to build public trust for future releases, the partners agreed to produce resources, for a range of audiences, describing the privacy measures taken to protect personal data. These resources will be made widely available, and will accompany future media releases and sensitisation exercises.

Finally, the data used in the analysis originated from a single mobile network operator in Ghana. An inherent limitation of the data is therefore that it is unlikely to be representative of the full population, and that any insights about the mobility behavior of Vodafone Ghana subscribers are not necessarily representative of the behavior of the entire population. This was highlighted within the publications and during press engagements. To produce higher quality insights in the future, efforts should be made to onboard other mobile network operators onto the initiative, as well as to incorporate survey data about phone usage behavior in order to mitigate the biases in the current dataset.

5. Summary

In this article, we have described how a pre-existing partnership between Vodafone Ghana, the Ghana Statistical Service, and Flowminder enabled insights about population movements to be derived from

mobile phone data and delivered to the Ghanaian Government in a timely manner, in response to the COVID-19 pandemic. We have highlighted the importance of having a well-functioning partnership, a governance framework, and privacy-preserving technical infrastructure to process mobile phone data in place, prior to any emergency situation when mobile phone data are required. These are foundations that require a lot of time and effort to put in place. Without them, accessing and analyzing mobile phone data in an urgent situation is likely to be severely delayed.

However, having these foundations in place and being able to deliver timely insights does not necessarily lead to direct usage and impact. Firstly, a lack of understanding about the limitations of the data can lead to unrealistic expectations of what is possible, especially in a high-pressure situation. These expectations need to be managed through an extended process of sensitisation. Secondly, bureaucratic Government processes combined with social-distancing restrictions make ongoing communication with relevant stakeholders difficult, highlighting the importance of initiating and maintaining key relationships prior to emergency situations.

Although the data and insights do not appear to have been used directly by the Government to inform its COVID-19 response, there are indications that key stakeholders within the Government now have a better understanding of the data and how it can be used, and have shown an increased interest in the data. Ongoing engagement with those stakeholders will potentially lead to a scenario in which the data source is utilized in a routine and sustainable manner.

The views expressed in this work belong solely to the authors and should not be interpreted as the views of their institutions.

Abbreviations

CDR   call detail records
GSS   Ghana Statistical Service

References

Arai A, **Fan Z**, **Matekenya D and Shibasaki R** (2016) Comparative perspective of human behavior patterns to uncover ownership bias among mobile phone users. *ISPRS International Journal of Geo-Information* 5(6), 85.

Flowminder, Ghana Statistical Service and Vodafone Ghana (2020a). Mobility analysis to support the Government of Ghana in responding to the COVID-19 outbreak—Initial insights into the effect of mobility restrictions in Ghana using anonymised and aggregated mobile phone data. Available at https://statsghana.gov.gh/COVID-19%20press%20release%20report%20-%20analysis%20overview%20-%20final1.pdf (accessed 31 August 2020).

Flowminder, Ghana Statistical Service and Vodafone Ghana (2020b). Mobility analysis to support the Government of Ghana in responding to the COVID-19 outbreak—Insights into the effect of mobility restrictions in Ghana using anonymised and aggregated mobile phone data. Available at https://statsghana.gov.gh/COVID-19%20press%20release%20report%20-%20analysis%20overview%20-%20final.pdf (accessed 31 August 2020).

Global Partnership for Sustainable Development Data (2020a) How Anonymized Mobile Data are Helping Ghana Fight COVID-19, May 18. Available at https://www.data4sdgs.org/news/how-anonymized-mobile-data-are-helping-ghana-fight-covid-19 (accessed 31 August 2020).

Global Partnership for Sustainable Development Data (2020b) Unlocking Privately Held Data for Public Good. Real time collaborative learning for COVID-19 and beyond. Final report, July 2020. Available at https://www.data4sdgs.org/sites/default/files/services_files/Unlocking%20Privately%20Held%20Data%20for%20Public%20Good_Final%20Report.pdf (accessed 31 August 2020).

Kishore N, **Kiang MV**, **Engø-Monsen K**, **Vembar N**, **Schroeder A**, **Balsari S and Buckee CO** (2020) Measuring mobility to monitor travel and physical distancing interventions: A common framework for mobile phone data analysis. *The Lancet Digital Health 2*(11), e622–e628.

Wesolowski A, **Eagle N**, **Noor AM**, **Snow RW and Buckee CO** (2013) The impact of biases in mobile phone ownership on estimates of human mobility. *Journal of the Royal Society Interface*, *10*(81), 20120986.