Blogpost: Can we fight COVID-19 without resorting to mass surveillance?

Yves-Alexandre de Montjoye, Florimond Houssiau

Transcript of the blogpost at https://cpg.doc.ic.ac.uk/blog/fighting-covid-19/

Abstract – *Governments across the world are doing everything they can to fight the COVID-19 virus. Used correctly, data collected through mobile phones could help monitor the effectiveness of lockdown measures and track contacts of people who have been tested positive. We've had many people reaching out to ask if the data could be collected and used effectively without enabling mass surveillance. We thought we'd share our response.*

March 21, 2020

Governments across the world are doing everything they can to fight the COVID-19 virus. This includes increasing hospital capacity, procuring protective equipment and ventilators; but also using data collected through mobile phones. Used correctly, large-scale location data might help monitor the effectiveness of lockdown measures and quarantines, while mobile phone apps collecting close proximity data [1] could significantly increase our ability to track contacts of people who have been tested positive.

Multiple countries [2] such as Belgium [3], the UK [4] and the USA [5] are considering using mobility data from telecom providers or tech giants, while Israel [6] has already (controversially[7]) started. Smartphone apps collecting close-proximity data have been trialed in the last couple of years at the DTU in Denmark [8] and at MIT in Boston [9] and are being deployed by the Government of Singapore [10].

We've had many people reaching out to ask if the data shared to and by governments is anonymous, and whether it could be used effectively without enabling mass surveillance. We thought we'd share our response:

In short: this is a crisis situation and every day counts. However, proven solutions exist to help share data broadly without enabling mass surveillance.

Location data

Location data from mobile phones, collected by telcos and apps on your phone, is highly sensitive [11], and, despite claims of anonymity, is easy to re-identify. Our research [12] showed that knowing 4 places and times where someone was, is enough to uniquely identify them 95% of the time and that adding noise doesn't fundamentally help.

This data has however long been used to model [13] and fight the spread of diseases, such as malaria in Kenya [14] or dengue in Pakistan[15], and to help first responders after natural catastrophes such as for the 2015 Nepal earthquake [16].

In 2018, recognizing the need to find ways to safely use this data including in times of crisis, we put together a working group of privacy experts, telecommunication researchers, and practitioners. Together, we agreed on four broad models to use mobility data in privacy-conscientious ways [17]. Since then, these models have been used, in practice, by NGOs like Flowminder (the remote access model) [18] and the OPAL project (the query-based model) [19] for mobile phone data, while centers like CASD [20] have pioneered data access control mechanisms. These models exist and can be used today to enable mobile phone data to be used while preserving privacy.

Contact tracing

Contact tracing, recording close proximity between people using bluetooth, wifi, or gps data, could help efficiently notify people that they have earlier been in contact with someone now diagnosed with coronavirus and should self-isolate. The typical design for such an app would either be to have the phone detect nearby phones, e.g. through Bluetooth like in Copenhagen and Boston [21], or through GPS, with phones sending the data to a centralised server which would then warn users when a case is confirmed. While potentially very effective, this also enables the collection of an extremely large amount of sensitive data. In a thought experiment two years ago, we estimated that 1% of London installing a malicious app could allow a digital attacker to track half of the London population [22][1] potentially without their consent.

Here, we think proven cryptography techniques could help. Using technology such as oblivious transfer [23], the app could check whether a person they encountered in the last 14 days has been diagnosed as positive without the centralised server seeing personal data. These techniques have already been used at scale. For instance, the secure messaging service Signal [24] relies extensively on encryption, secure hash, and other advanced cryptographic techniques [25] to ensure that their server learns nothing of their user's messages or calls; while Google's private join-and-compute [26] tool allows two companies to check for shared customers without learning each other's full list of customers.

[1] This assumes a malicious app actively trying to collect data. It relies on many assumptions, especially regarding mobility but is meant to give an order of magnitude of what might be possible. More details are available in our preprint

Moving forward

Strong technical solutions can help but will not be sufficient to protect what we need the most in times like these: trust. We need strong technical solutions but also transparency and oversight. Why and how the data will be collected and used need to be clearly communicated and explained. Maybe we are comfortable sharing our data to help governments study the spread of COVID-19; but maybe less so if this data then surreptitiously used to crackdown on individuals not respecting quarantines or kept and used for unrelated purposes. Oversight by data protection authorities and experts will similarly be essential to ensure trust and proportionality.

Fighting the coronavirus is a major challenge for our societies that will require us to leverage every tool and technology at our disposal. But this doesn't have to mean mass surveillance. Good solutions exist, let's use them!

Bibliography

[1] L. Smith (2020), 'An app for tracking coronavirus in your community is almost here', *Fast Company*, March 19

[2] K. Grind, R. McMillan and A.W. Mathews (2020), 'To Track Virus, Governments Weigh Surveillance Tools That Push Privacy Limits', *The Wall Street Journal*, March 17

[3] A. Cloot (2020), 'Coronavirus: le cabinet De Block dit «oui» à l'utilisation des données télécoms', *Le Soir*, March 12

[4] M. Sweney and A. Hern (2020), 'Phone location data could be used to help UK coronavirus effort', *The Guardian*, March 19

[5] B. Fung (2020), 'Trump administration wants to use Americans' location data to track the coronavirus', *CNN Business*, March 18

[6] J. A. Gross and T. Staff (2020), 'Israel starts surveilling virus carriers, sends 400 who were nearby to isolation', *The Times of Israel*, March 18

[7] A. K. Sommer (2020), "Yuval Noah Harari Warns Against 'Coronavirus Dictatorship' in Israel – Netanyahu's Son Calls Him 'Stupid' ", *Haaretz*, March 19

[8] Sapiezynski, P., Stopczynski, A., Lassen, D.D. and Lehmann, S., 2019. 'Interaction data from the Copenhagen Networks Study.' *Scientific Data*, 6(1), pp.1-10.

[9] Aharony, N., Pan, W., Ip, C., Khayal, I. and Pentland, A., 2011, September. 'The social fMRI: measuring, understanding, and designing social mechanisms in

the real world.' In *Proceedings of the 13th international conference on Ubiquitous computing* (pp. 445-454).

[10] F. Ungku (2020) 'Singapore launches contact tracing mobile app to track coronavirus infections', *Reuters*, March 20

[11] J. Valentino-DeVries, N. Singer, M.H. Keller and A. Krolik (2018), 'Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret', *The New York Times*, December 10

[12] De Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: The privacy bounds of human mobility. Scientific reports. 2013 Mar 25;3:1376.

[13] Lima A, De Domenico M, Pejovic V, Musolesi M. Exploiting cellular data for disease containment and information campaigns strategies in country-wide epidemics. arXiv preprint arXiv:1306.4534. 2013 Jun 19.

[14] Wesolowski, A., Eagle, N., Tatem, A.J., Smith, D.L., Noor, A.M., Snow, R.W. and Buckee, C.O., 2012. Quantifying the impact of human mobility on malaria. Science, 338(6104), pp.267-270.

[15] Wesolowski, A., Qureshi, T., Boni, M.F., Sundsøy, P.R., Johansson, M.A., Rasheed, S.B., Engø-Monsen, K. and Buckee, C.O., 2015. Impact of human mobility on the emergence of dengue epidemics in Pakistan. Proceedings of the National Academy of Sciences, 112(38), pp.11887-11892.

[16] Wilson, R., zu Erbach-Schoenberg, E., Albert, M., Power, D., Tudge, S., Gonzalez, M., Guthrie, S., Chamberlain, H., Brooks, C., Hughes, C. and Pitonakova, L., 2016. Rapid and near real-time assessments of population displacement using mobile phone data following disasters: the 2015 Nepal Earthquake. PLoS currents, 8.

[17] de Montjoye, Y.A., Gambs, S., Blondel, V., Canright, G., De Cordes, N., Deletaille, S., Engø-Monsen, K., Garcia-Herranz, M., Kendall, J., Kerry, C. and Krings, G., 2018. On the privacy-conscientious use of mobile phone data. Scientific data, 5.

[18] https://web.flowminder.org/

[19] Oehmichen, A., Jain, S., Gadotti, A. and de Montjoye, Y.A., 2019, December. OPAL: High performance platform for large-scale privacy-preserving location data analytics. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 1332-1342). IEEE.

[20] https://www.casd.eu/

[21] Stopczynski, A., Sekara, V., Sapiezynski, P., Cuttone, A., Madsen, M.M., Larsen, J.E. and Lehmann, S., 2014. Measuring large-scale social networks with high resolution. PloS one, 9(4).

[22] Radaelli, L., Sapiezynski, P., Houssiau, F., Shmueli, E. and de Montjoye, Y.A., 2018. Quantifying surveillance in the networked age: Node-based intrusions and group privacy. arXiv preprint arXiv:1803.09007.

[23] Rabin, M.O., 2005. How To Exchange Secrets with Oblivious Transfer. IACR Cryptology ePrint Archive, 2005, p.187.

[24] https://signal.org

[25] moxie0 (2017), 'Technology preview: Private contact discovery for Signal', *Signal blog*, September 26

[26] A. Walker, S. Parel, M. Yung (2019), 'Helping organizations do more without collecting more data', *Google Security Blog*, June 19