2020 21st IEEE International Conference on Mobile Data Management (MDM)

COVID-19 Mobile Contact Tracing Apps (MCTA): A Digital Vaccine or a Privacy Demolition?

Demetrios Zeinalipour-Yazti
Department of Computer Science
University of Cyprus
1678 Nicosia, Cyprus
dzeina@cs.ucy.ac.cy

Christophe Claramunt
Naval Academy Research Institute
French Naval Academy, CC 600,
29240 Brest, France
christophe.claramunt@ecole-navale.fr

***Abstract*—The COVID-19 global pandemic emerged in the late 2019 causing so far a massive global health disruption with many fatalities and huge economy impact, enforcing most if not all governments to a global lockdown. Besides the battle on the medical front-line, governments and the industry also massively explored the deployment of information and communication technologies to track and curb the spread of the virus. On the front-line of these efforts, have been the so-called *Mobile Contact Tracing Applications (MCTA)*. These refer to mobile apps that exploit the rich ecosystem of mobile sensors (e.g., location, proximity) as well as social networks to facilitate the process of identification of persons who may have been previously into contact with a covid infected person and subsequent collection of further information about these contacts. Although MCTA can in theory help governments fight the rapid spread of diseases like COVID-19, there are important privacy considerations and many claim that these technologies will put in place a massive global surveillance infrastructure that will survive even when a vaccine for the COVID-19 disease has been found. This panel aims to discuss the major challenges and open topics surrounding MCTA. The panelists are expected to bring wealth of experience and vision from the academic, governmental and industrial sector to answer a set of challenging questions that are currently open to public debate as well as the global benefits one can expect when fighting the COVID-19 spread.**

Keywords*—*COVID-19, Contact Tracing, Privacy, Safety, Apps

I. Introduction

On December 31, 2019 a severe pneumonia of relatively unknown cause was reported from Wuhan, China to the World Health Organization (WHO). A highly infectious disease caused by a newly discovered coronavirus, the *SARS-CoV-2 (Severe Acute Respiratory Syndrome - CoronaVirus - 2)*, quickly spreads on a world-wide scale escalating very soon to a global pandemic, named *COVID-19 (Coronavirus Disease)* [1]. To this date, there have been around 5.31M confirmed cases, with 2.11M cases recovered and 342K deaths [2]. Bill Gates' greatest fear and early warning [3], inspired by a series of scientific studies [4], did realize to an unprecedented scale enforcing governments globally to shutdown borders and quarantine citizens, bringing the whole planet to a still-stand.

Besides the medical personnel, being on the front-line of combating the virus in intensive care units, epidemiologists and respective law-enforcement agencies attempted to as soon as possible identify COVID-19 clusters and manually tracking down persons who may have come into contact with an infected person, and subsequently collecting further information about these contacts, a process known as *Contact Tracing (CT)* [5]. The biggest problem with CT is that it is a relatively slow and an approximated process that cannot effectively cope with the speed a virus like SARS-CoV-2 spreads under loose distancing measures. Another problem is that CT is an arduous work that requires massive human resources, with experts in the US estimating that this process would require between 100,000 and 300,000 manual CTs.

Information and Communications Technology (ICT) experts on the other hand, attempted during this period to facilitate epidemiologists and governments with a variety of ICT tools to help curb the spread of the virus (with countries like South Korea leading this example and industrial fast developed applications setup by Apple/Google and telecommunication companies): Examples include, intelligent tracking and infection hotspots dashboards, AI-inspired Deep Learning self-questionnaires with Big Data, collaboration technologies to improve distancing and telework (e.g., e-conferences [6]), paper-less workflows in the cloud at a massive scale (e.g., social applications for COVID-19 benefits), rapid establishment of e-education pipelines (schools, universities, etc.), the uptake of faster and wider communication channels on the edge (5G and last mile fiber-optics).

On the front-line of these technological efforts, have been the so-called *Mobile Contact Tracing Applications (MCTA)* [7]. These refer to software artifacts exploiting the rich ecosystem of mobile sensors (e.g., location, proximity) to enable Contact Tracing of Custodians. The feature-rich interaction modalities of mobile device on the other hand, enabling opportunistic to participatory approaches, have brought forward very diverse propositions for MCTA from the research community: from handled-based approaches to cloud-based approaches, approaches utilizing BLE, Wi-Fi, Sound, etc., either for only outdoor settings or indoor settings as well (post-pandemic MCTA) with closed-source and open-source counterparts being promoted in virtually every country.

At the same time, MCTA have also sparkled global privacy concerns about basic human rights [8]–[11]. Many claim that these technologies will put in place a massive global surveillance infrastructure that will survive even when a vaccine for the COVID-19 disease has been found, similarly to stringent procedures that emerged and remained under in the wake of the 9/11 terrorist attacks [8]. Conspiracy theorists might even claim that this is all part of a global exploitation effort to establish big brother societies [8] with 5G [12]. Clearly, this also emphasizes socio-cultural differences and how different countries will apprehend and accept such technologies, not mentioning political differences in the way these technics might be developed and further used. Finally, the low and

DOI 10.1109/MDM48529.2020.00020

slow installation penetration rate of these apps 5-20% might constitute MCTA ineffective altogether, as opposed to GPS tracking devices used on confirmed COVID-19 victims to control their containment.

On the other hand, the pandemic caused global disruption and an economic devastation to many sectors that have been particularly significant to suffering economies. Consequently, a significant proportion of people eagerly have to return to work to avoid the bankruptcy of companies, organizations and respective states.

The question that is put to open debate is whether Mobile Contact Tracing Apps (MCTA) can be considered as an effective "digital covid stop" (easily and immediately supplement to curb the spread of the virus until a real biological vaccine is found) and/or whether these will only lead to the destruction and eventually to the demolition of user privacy?

This panel aims to discuss the major and open topics surrounding MCTA. The panelists are expected to bring wealth of experience from the academic, governmental and industrial sectors to answer a set of challenging technical, socio-economical and ethical questions that are currently open to public debate. The goal of this panel is to convey an advanced understanding of the unique characteristics, socio-technical challenges and opportunities in the sphere of contact tracing mobile apps to its audience. The discussion is expected to provide a solid understanding as of where the key technical challenges and opportunities lie in making MCTA a safety measure for public benefit without deconstructing or putting user privacy at risk.

II. Potential Questions

Through the short positioning statements of panelists and the active engagement of the audience, the goal will be to derive the current and future directions in the emerging domain of MCTA. The panelists will be asked to provide perspectives on the following indicative list of questions:

1) What are the most prominent Privacy and Ethical aspects you see with MCTA?

Location privacy refers to the ability of an individual to move in public spaces with the reasonable expectation that their location will not be systematically and secretly recorded for later use. A fundamental drawback of MCTA, is that these involve different types of location tracking technologies (e.g., beacons, bluetooth, Wi-Fi) that can potentially enable a service (and subsequently a company or government) to track custodians at a very fine resolution in real-time and in retrospective. On the one hand, MCTA can obviously also become attractive targets for hackers, aiming to steal location data and carry out illegal acts. Location tracking poses a serious imminent threat as it can occur at a very fine granularity in the scope of MCTA [13]. Complementary, mobility traces are inherently highly correlated and unique and, as such, the re-identification privacy threat is imminent. Most MCTA approaches have opted early on to rely on indirect tracking technologies (e.g., hashing of BLE, Wi-Fi MAC addresses through temporal cryptographic functions yielding Ephemeral IDs), again though the privacy threat now only becomes computationally more difficult rather than non-existent (i.e., Pseudonymization vs. Anonymization).

The panelists are expected to expand on whether those aspects present a legal barrier to the uptake of MCTA, under the prism of stricter global regulation on data privacy (e.g., GDPR in Europe).

2) Mobile Operating System vendors are developing their own MCTA. What are the pros and cons of this approach?

Google and Apple, as the predominant smartphone Operating System (OS) vendors with a 99% market share [14], have announced a joint effort to enable the use of Bluetooth technology to help governments and health agencies reduce the spread of the virus. According to these companies, their proposition has user privacy and security central to their design [15]. The particular idea uses a Temporary Exposure Key that is generated every 24 hours for privacy consideration and is used in local BLE broadcasts in the form of a Rolling Proximity Identifier that changes every 15 minutes.

The panelists are expected to expand on whether an open protocol that is possibly integrated deeply in the Smartphone OS has benefits or concerns over having individual organizations or governments rolling out their own apps. Should governments have administrative control over such designs? Should the final citizens have the right to opt-in these type of solutions or should the opt-out principle apply? Is blocking of 3rd party MCTA apps from the Google/Apple stores causing ethical issues or possibly invoking the antitrust law against these companies (currently only state-linked MCTA apps by national or regional health providers are admitted to the respective store markets)? So in summary, the question is whether smartphone OS support for MCTA will help privacy or hurt privacy in the MCTA mission?

3) What is the role of coordinated open-* (open protocols, open data and open source) in MCTA advancement?

There are a variety of protocols under development that aim to make MCTA more transparent and effective [16]. Examples include the following: the World-level Open-Source Anonymous Protocol, the Pan-European Privacy-Preserving Proximity Tracing (PEPP-PT), Whisper Tracing Protocol, the Decentralized Privacy-Preserving Proximity Tracing (DP-PPT), TCN Protocol, Contact Event Numbers (CEN), Privacy Sensitive Protocols And Mechanisms for Mobile Contact Tracing (PACT). Many of these protocols are merely propositions, aiming to consolidate human resources and effort at a global level, and have yet not evolved into concrete and finalized software artifacts. An alternative issue is that even though the protocol might be open, the software and/or the deployment might eventually be closed.

Complementary to open protocols one could also envision that systems promoting open data through crowdsourcing would promote awareness for common good. Eventually, these systems could integrate data from publicly-available APIs into information dashboards [17] but also release primary or mashed-up data in the form of semi-structured *Linked Open Data (LOD)*. However, this encompasses technical challenges in its own right, for example online entity disambiguation from streaming data as well as identification of false statements (possibly also rumor spreading, fake news detection).

Finally, the contribution of open-source MCTA applications during the COVID-19 pandemic has been significant (e.g., SafePaths by MIT [18]), as governments and organizations could utilize solutions off-the-shelf. This enabled organizations around the globe to quickly replicate and disseminate custom-made MCTA applications without having a deep technology understanding. At the same time, by running this type of MCTA applications regionally (e.g., within a country), also eased privacy concerns as regional authorities had the opportunity to engage the national privacy commissioners and provide a feeling of control to citizens. Yet, replicating open-source developments has also its risks, as the lack of deep domain understanding can potentially undermine important aspects of application integrity, security and correctness of collected data.

The question is whether open protocols, open data and open source are sufficient, either in isolation or in combination, to warrant the trustworthiness of MCTA.

4) What are the right architectures, functionalities and interfaces for MCTA?

The mass majority of proposed MCTA architectures have mainly focused mobile client / cloud architectures, where mobile handheld devices do the sensing of BLE, sound, Wi-Fi signals, or any combination of these and other local signals, and effectively retain this data on the handheld until the users decide to share it with a global repository in the cloud in an abstracted form (e.g., as a semantic trajectory with POIs visits or temporal keys as the Apple/Google approach). These approaches might suffer from typical privacy attacks as certain quasi-identifiers might be compromised in the future leading to privacy breaches (i.e., k-anonymity, l-diversity and t-closeness). On the other hand, there are approaches like Smarttrace [21], which will never reveal data to the cloud but will on the other hand participate in Query Answering for common good. Interactive databases on the other hand can also suffer from brute-force privacy attacks if appropriate measures are not effectively implemented (e.g., ϵ-differential privacy seeks to limit the knowledge that users obtain from query responses by adding Laplace noise so that brute-force queries don't lead to privacy revelations).

The question is whether specific communication architectures have benefits over other architectures to balance the Privacy-Safety dichotomy. What sort of advanced functionalities and real-time user-oriented interfaces should be provided to both citizens and epidemiologists? How to make a right balance between centralized and decentralized architectures as outlined in a recent manifesto [16]?

5) What are the technical challenges for MCTA?

The low installation penetration rate of MCTA apps (currently 5-20%) might constitute MCTA useless altogether, if these are not used by all users. A similar concept does also hold for real vaccines. For example, the Centers for Disease Control and Prevention [19] requires a vaccine coverage at 94 percent to provide a meaningful prevention for measles. Does a similar concept would also apply to a 'digital covid stop" like an MCTA? Preliminary studies suggest that a take-up threshold of 60% of the population can bring an outbreak under control [20], yet these numbers have not been verified and accepted by the scientific community. Several complementary obstacles have also been reported with preliminary MCTA apps due to technical challenges. For example Singapore's TraceTogether app [8], using its BlueTrace.io protocol, would deplete the battery of custodians. Additionally, a possible malfunctioning of sensors on the smartphone or device diversity issues, where different vendors report signal measurements on different scales, would also limit the reliable collection of data.

As such, the issue is how much practical and technical challenges are a limitation to the uptake of MCTA?

6) What are the correct localization/proximity technologies for MCTA given the wide spectrum of technologies?

The localization and proximity literature is very broad and diverse, as it exploits several technologies, including: GNSS, Infrared, Light, Bluetooth, visual or acoustic analysis, RFID, Inertial Measurement Units, Ultra-Wide-Band, Sensor Networks, Wireless LANs, Computer-Vision-based, etc.; including their combinations into hybrid systems. Most of these technologies deliver a high level of positioning/proximity accuracy, however some of these require the deployment and calibration of expensive equipment, such as custom transmitters, antennas or beacons, which are dedicated to positioning. This is time consuming and implies high installation costs. At the same time, it complicates data management queries and analytics, as solutions might be targeted to a specific set of technologies. Additionally, measurements from different sensor devices are not always consistent (i.e., device-diversity problem). Finally, shouldn't the primary underlying MCTA contact/proximity tracing technology (i.e., BLE advertising) be more thoroughly tested before being deployed at such large scale?

The question is whether MCTA should rely on a specific combination of localization and/or proximity technologies. Do we actually need this type of accuracy in MCTAs or are semantic trajectories (and semantic proximity) sufficient to protect people with the right level of accuracy?

7) What is the precise scope of MCTA? Should MCTA be used to speed up the Contact Tracing Process by Epidemiologists only or should these apps also protect us against a possible virus infection?

MCTA were initially proposed to allow Epidemiologists to track the virus spread faster and to curb in this way the spread of the virus. As such, the MCTA process has been considered an a posteriori process, rather than an a priori process. The question is whether MCTA should also alert users in an online manner (e.g., smartphone notifications when we approach a virus hotspot or an infected person). Moreover, global instantaneous statistics could be made available to epidemiologists as well as precise people's covid infection status. From the privacy perspective, such an additional functionality would introduce the notion of "online tracking" (i.e., custodians will share in an ongoing manner their contacts or their location to a central database) as opposed to "offline tracking", the former being considered even more severe from a privacy standpoint. One could of course also envision supplementary (domain specific) activity for MCTA (e.g., for restaurants, beaches).

As such, the question is if the scope of MCTA should be framed very precisely or whether it is better to allow companies, governments and users define their custom scope? How MCTA data could be made available across regions and countries for further epidemiology analysis and forecast? Finally, might MCTA replace fully CT in the future?

8) Is more legislative reform necessary before MCTA can be used with confidence?

We are witnessing a worldwide movement to more legislative reforms in the scope of MCTA. For instance, EUs eHealth voluntary Network has recently drafted a Common EU Toolbox for Member States that is expected to facilitate the development of MCTA [23] while the European Data Protection Board (EDPB) has also announced its own recommendations [24] to the EU member states and the same applies to Professional bodies like ACM Europe [26]. Similar initiatives are also under development across many countries such as the Privacy Amendment of Public Health Contact Information [25], recently established by law in Australia to provide stronger privacy protections for users in MCTA.

The question is whether these recommendation will be mandatory at some point and also who will be in charge of enforcing compliance? Are these centrally developed guidelines more effective than the open- approaches? Should cross-border MCTA approaches be encouraged (or targeted)?*

REFERENCES

[1] World Health Organization (WHO), Coronavirus disease (COVID-19) pandemic, https://www.who.int/

[2] Google News, Coronavirus (COVID-19), Accessed May 24, 2020, https://news.google.com/covid19/map

[3] "The next outbreak? We're not ready", Bill Gates, Apr 3, 2015, https://www.youtube.com/watch?v=6Af6b_wyiwI

[4] "Emerging infectious diseases and pandemic potential: status quo and reducing risk of global spread", B. McCloskey, O. Dar, A. Zumla, L. Heymann, The Lancet Infectious Diseases, 14(10), 1001-1010, 2014.

[5] "Contact Tracing, Coronavirus (COVID-19)", Accessed May 11, 2020, https://en.wikipedia.org/wiki/Contact_tracing

[6] "Holding a Conference Online and Live, Due to COVID-19", Angela Bonifati, Giovanna Guerrini, Carsten Lutz, Wim Martens, Lara Mazilu, Norman Paton, Marcos Antonio Vaz Salles, Marc H. Scholl, Yongluan Zhou, April 21, 2020 http://tiny.cc/9u3apz

[7] "COVID-19 apps", Accessed May 17, 2020, https://en.wikipedia.org/wiki/COVID-19_apps

[8] "Coronavirus contact tracing poses serious threats to our privacy", May 10, 2020, https://tiny.cc/6mk7oz

[9] "Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs", Hyunghoon Cho and Daphne Ippolito and Yun William Yu, arXiv:2003.11511, 2020. https://arxiv.org/abs/2003.11511

[10] "Coronavirus contact-tracing apps put users at risk, EU lawmaker says", May 10, 2020, https://tiny.cc/rdm8oz

[11] "The Privacy Factor in Ending the Lockdown", Stan Matwin, Project Syndicate, Apr 21, 2020, http://tiny.cc/g88gpz

[12] "What is the truth behind the 5G coronavirus conspiracy theory?", Euronews https://tiny.cc/w52apz

[13] "A State-of-the-Art Review on the Security of Mainstream IoT Wireless PAN Protocol Stacks", Kambourakis, G. et. al., Symmetry 12 (4), p.579, 2020,

[14] "Mobile operating systems' market share worldwide from January 2012 to December 2019", statista.com http://tiny.cc/apbhpz

[15] "Privacy-Preserving Contact Tracing", Apple.com https://www.apple.com/covid19/contacttracing

[16] "Give more data, awareness and control to individual citizens, and they will help COVID-19 containment", Nanni et al., 2020, Transactions on Data Privacy, 13(1), 61-66.

[17] COVID-19 Hong Kong Map https://covid19.vote4.hk/

[18] "Private Kit: Safe Paths; Privacy-by-Design Covid19 Solutions using GPS+Bluetooth for Citizens and Public Health Officials," https://safepaths.mit.edu/

[19] US Centers for Disease Control and Prevention https://www.cdc.gov/

[20] "Digital contact tracing: comparing the capabilities of centralised and decentralised data architectures to effectively suppress the COVID-19 epidemic whilst maximising freedom of movement and maintaining privacy", Fraser C. et. al., Preprint, May 7, 2020. http://tiny.cc/xd2mpz

[21] "Crowdsourced trace similarity with smartphones", Zeinalipour-Yazti et. al. IEEE Transactions on Knowledge and Data Engineering, Volume 25, Pages: 1240-1253, 2013. https://smarttrace.cs.ucy.ac.cy/

[22] "Coronavirus contact-tracing apps: can they slow the spread of COVID-19?", Mark Zastrow, Nature, Technology Feature, May 19, 2020 https://www.nature.com/articles/d41586-020-01514-2

[23] "Mobile applications to support contact tracing in the EUs fight against COVID-19", eHealth Network Common EU Toolbox for Member States, 15.04.2020, version 1, http://tiny.cc/ayeopz

[24] "Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak", EU Data Privacy Board, 21 April 2020, http://tiny.cc/azfopz

[25] Australian Privacy Amendment (Public Health Contact Information) Act 2020, No. 44, 2020 http://tiny.cc/d0vopz

[26] "Statement on Essential Priciples and Practices for COVID-19 Contact Tracing Applications", ACM Europe Technology Policy Committee, May 5, 2020, http://tiny.cc/bw5spz

SHORT BIOGRAPHIES OF PANEL CO-ORGANIZERS

Demetrios Zeinalipour-Yazti is an Associate Professor of Computer Science at the University of Cyprus, where he founded and directs the Data Management Systems Laboratory (DMSL). He holds a Ph.D. in Computer Science and Engineering from the University of California - Riverside, CA, USA. He has held short-term research visits at Akamai Technologies, Cambridge, MA, USA (2004), the University of Athens, Greece (2007) as a Marie-Curie Fellow and the University of Pittsburgh, PA, USA (2015). During 2016-2017, he was a Humboldt Fellow at the Max Planck Institute for Informatics, Saarbrücken, Germany. He is an ACM Distinguished Speaker (2017-2020), a Senior Member of ACM, a Senior Member of IEEE, and a Member of USENIX. He serves on the editorial board of Distributed and Parallel Databases (Elsevier), Big Data Research (Springer) and is an independent evaluator for the European Commission (Marie Skłodowska-Curie and COST actions) as well as the Hong-Kong RGC. For more information please visit: https://www.cs.ucy.ac.cy/˜dzeina/

Christophe Claramunt is a professor in computer science and research chair of the French Naval Academy. He was previously a senior lecturer in computing at the Nottingham Trent University in the United Kingdom and senior researcher at the Swiss Federal Institute of Technology in Lausanne, Switzerland. He holds a PhD in computer science from the University of Burgundy, France, and an "Habilitation à Diriger des Recherches" from the University of Rouen, France. His research is oriented towards theoretical and pluri-disciplinary aspects of geographical information science. His main research interests include, Spatio-temporal models and theories, Semantic and cognitive-based GIS; WEB, wireless and GIS systems; and Maritime, environmental and urban GISs. He is an associate editor of the International Journal of Geographical Information Science (IJGIS) and serves on the editorial boards of several area journals (CEUS, JOSIS, JLBS, IJAIS, GIS, IJAEIS) and conferences/workshops (COSIT, SeCoGIS, ACMGIS, SC, SDH, WEBIST, STAMI and ISA). For more information please visit: http://christophe.claramunt.free.fr/