On the Privacy of National Contact Tracing COVID-19 Applications: The *Coronavírus-SUS* Case

Jéferson C. Nobre[1], Laura R. Soares[1], Briggette O. R. Huaytalla[1], Elvandi da S. Júnior[1], Lisandro Z. Granville[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

Abstract. *The COVID-19 pandemic, caused by the quick dissemination of the SARS-CoV-2 vírus, deeply impacted the world. The use of digital technologies has been crucial in the effort to control it, and, among others, digital contact tracing (DCT) applications stood out. DCT was successfully employed to face other infectious diseases in the past. However, its use poses several privacy concerns due to the sensitiveness of the data it handles. These concerns are even more relevant when considering nationwide implementations, despite several countries having data protection regulations in place. This article analyzes the privacy features in national DCT COVID-19 applications and their overall adhesion. As a case study, we discuss in more depth Brazil's application, Coronavírus-SUS, since Brazil is one of the most impacted countries by the pandemic. Finally, as we believe DCT will remain relevant in health-related tasks, we present key research challenges.*

1. Introduction

Digital Contact Tracing (DCT) is a technological method of monitoring the progress of infectious diseases on a large population by detecting the contacts between infected and healthy individuals [4]. These contacts are traced through applications installed on mobile devices, usually smartphones. DCT apps use different techniques to trace proximity: Global Positioning System (GPS), triangulation of cellular operator antennas, electronic transaction data (*e.g.*, credit card data), and Bluetooth. Tracing data (*e.g.*, contact history) can be treated in a centralized manner (processed and stored on a central entity) or in a decentralized one (using users' devices for processing and storage). DCT apps were successfully employed in facing previous pandemics [1].

Countries worldwide implemented COVID-19 DCT apps in different stages of the pandemic and with varying penetration rates. To do that, several relied on the Google/Apple Exposure Notification (GAEN) system [5], which used a decentralized approach to foster the development of DCT apps by health authorities for both Android and iOS devices. In addition, data protection laws from several countries impose restrictions in data processing and detail the security measures required to store personal data. Some of the best practices in DCT applications are to design them to prevent the user's device from collecting personally identifiable information. Also, health agencies should be the only ones allowed to operate on collected data. Failing to consider these practices may result in user privacy violations, like improper access to personal data and even mass surveillance.

In this article, we analyze privacy features in national DCT apps designed to combat the COVID-19 pandemic. Since logging encounters and exposure notifications may

lead to undesirable privacy attacks, we discuss the privacy of national DCT COVID-19 apps considering variables such as user penetration, used technologies, and architecture. After almost two years in the pandemic, the attained adhesion of DCT apps is also discussed. Finally, given that Brazil is one of the most impacted countries, we evaluate in more detail the Brazilian application, *Coronavírus-SUS*.

This article is organized as follows. In the next section, we present a background on DCT, and then, in Section 3, we provide a general view of privacy in the context of COVID-19. Section 4 evaluates several DCT apps employed by national authorities around the world. We then present a case study regarding the *Coronavírus-SUS* application in Section 5. Finally, Section 6 ends this article by presenting key research challenges and final remarks.

2. Digital Contact Tracing in a Nutshell

National governments as well as the private sector are working towards developing computational tools to help on the effective management of the COVID-19 pandemic. DCT apps help break the chain of virus transmission because is enable the monitoring of interactions of infected and healthy users, thus detecting potential infections. There are precedents for such apps being used on other health crises as a part of the strategy for disease outbreak control [1]. The tasks performed by these apps can be aggregated into proximity tests, transmission, and exposure notification.

The proximity tests can be performed through several methods and technologies that are integrated into current mobile devices. Proximity tracing is a method usually performed using Bluetooth Low Energy (BLE) to transmit messages containing identifiers to nearby devices. Location tracking can be performed using data from the GPS or cell tower triangulation. Geotagging is a method where users scan the QR code with their mobile device to record their visits and their localization data.

Different architectures can be employed to collect users' data and contact events, that are depicted on Figure 1. Centralized architectures collect the raw contact history data of mobile phones. After that, this data is stored and processed in a central server, which generates reports and sends exposure notifications using the same network. However, this centralization brings concerns over dependability and performance. Decentralized architectures, employ local resources for data storage and processing, which is feasible since preserved only contact events of the last days. In both architectures, mobile devices are usually responsible for generating temporary (also called ephemeral) identifications.

Google and Apple formed a partnership to develop an interoperable interface for mobile devices, contact event detection based on BLE technology, called GAEN system, which presents an Application Programming Interface (API) that is implemented at the operating system level to avoid privilege problems. CTA from many countries use this API for their notification of exposure and the generation of temporary tokens (to preserve user privacy). However, GAEN API is not open source and its public documentation is limited, which brings concerns about the transparency of the API [6].

A DCT app collects and exchanges sensitive users' data on a regular basis. However, such ability comes at a cost: privacy concerns. Besides, characteristics of different architectures can impact the preservation of users' privacy, since the transmission, processing, and storage of users' data are performed in distinct ways.

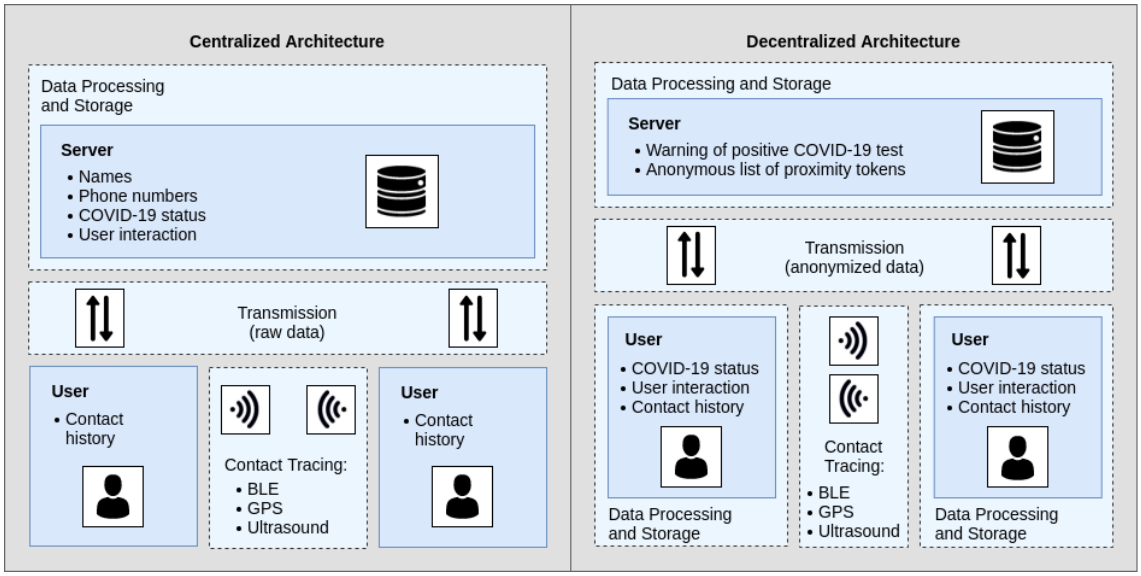

Figure 1. Centralized and decentralized architecture for contact-tracing applications.

3. Privacy in the Context of COVID-19

The COVID-19 pandemic spread quickly and challenged public health policies worldwide. Before the arrival of vaccines, social distancing was the most effective measure against COVID-19 [8]. At the beginning of 2020, as most of the global workforce started to work and socialize from home, threats related to personal data privacy became a major concern. Healthcare applications, such as DCT, are especially sensitive.

In addition to handling health data, DCT apps represent even higher risks to privacy since they must have extensive user coverage in a population to be effective. This is one of the major reasons why most of these apps were developed or funded by governmental offices[1]. However, this is no guarantee of success. In November 2020, a password leak inside the Ministry of Health of the Brazilian Government led to the exposure of data from at least 16 million COVID-19 patients[2]. Mass surveillance is also a concern. In January 2021, the Ministry of Home Affairs of Singapore confirmed that the police could access data from the TraceTogether COVID-19 DCT app[7].

Most countries have data protection regulations in place for handling the safety of personal information. The General Data Protection Regulation (GDPR)[3] from the European Union, for example, aims to protect a citizen's data in their territory but also within foreign enterprises handling it. Brazil's law on data protection, *Lei Geral de Proteção de Dados* (LGPD)[4], was created in 2018 using GDPR as a base but only took effect as of August 2021. In Singapore, despite the Personal Data Protection Act (PDPA), data from the DCT app TraceTogether is not exempt from the Criminal Procedure Code, which states that the police can access any data for criminal investigations.

Almost two years in the pandemic and the discussion on user privacy versus public

[1]https://tinyurl.com/mit-covid-tracing-traker
[2]https://tinyurl.com/estadao-ministerio-saude
[3]http://bit.ly/EUR-GDPR
[4]http://bit.ly/Planalto-LGPD

health in the context of DCT applications is far from over. Governments worldwide had time to refine these applications and strengthen, or not, the privacy of their citizens. In the next section, we analyze how the privacy of DCT apps evolved in the last year.

4. National DCT COVID-19 Apps and Privacy

Countries across the globe have employed DCT applications to face the COVID-19 pandemic. Governmental offices developed most of them due to the resources and logistics necessary to reach national coverage. After almost two years, several of these apps have faced problems, such as glitches, replacement of components, privacy concerns, and lack of adhesion.

Despite a privacy-friendly decentralized architecture employing the GAEN system, Japan's COCOA had to be suspended at least twice due to glitches and notification problems. Norway relaunched the *Smittestopp* app after switching from GPS coordinates and centralized architecture to the GAEN system and decentralization. Finland's and the United Kingdom also switched to the GAEN system after relaunch. Singapore's TraceTogether, despite privacy concerns and centralized architecture, is mandatory for entering public spaces such as shops and restaurants and reached 80% of the population. Israel's *HaMagen* and, later, *HaMagen 2.0* are estimated to have reached 2 million citizens, using GPS and Bluetooth data for precision. The last data on the adhesion of these applications, collected by the MIT Covid Tracing Tracker[5] as of January 2021, can be found in Table 1.

Location	Name	Penetration	Contact Detection	Distribution
Finland	Koronavilkku	45,31%	Bluetooth, GAEN	Decentralized
Israel	HaMagen 2.0	22,51%	Bluetooth, Location	Centralized
Japan	COCOA	6,09%	Bluetooth, GAEN	Decentralized
Norway	Smittestopp	2.94%	Bluetooth, GAEN	Decentralized
Singapore	TraceTogether	80%	Bluetooth, BlueTrace	Centralized
UK	NHS COVID-19 App	28,51%	Bluetooth, GAEN	Decentralized

Table 1. Exemples of DCT apps.

Several reasons might be behind the lack of adhesion of DCT apps. Elderly people and those without a smartphone, for instance, are groups of difficult reachability. Beyond that, the uncertainty regarding the effectiveness and the privacy risks of these applications might be keeping the general public from using them [3], e.g., fear of mass surveillance. Some of the best practices for handling sensitive data are to employ decentralized architectures, voluntary and consented use and sharing of tests results, to share only non-identifiable data with other application users, to collect only information strictly related to COVID-19, to make the source code publicly available, and to inform the user thoroughly of the DCT process [2].

The public concern on privacy issues might as well be one of the reasons why the adhesion of Brazil's application, *Coronavírus-SUS*, is not yet satisfactory. An overview of DCT in Brazil will be discussed in the next section.

[5]https://tinyurl.com/mit-covid-tracing-traker

5. Brazilian Digital Contact Tracing App: *Coronavírus-SUS*

The official application for COVID-19 DCT in Brazil, *Coronavírus-SUS*, was launched in February 2020 as a pilot version and upgraded with contact tracing functionalities only in September 2020. *Coronavírus-SUS* was promoted by the Ministry of Health and developed by DATASUS, the Department of Informatics of Brazil's public-funded healthcare system, the *Sistema Único de Saúde* (SUS).

Coronavírus-SUS uses GAEN API for anonymous Bluetooth sharing of tokens, making it a decentralized application. The users are informed about the application's privacy policy when first downloading it. The policy states that no personal data is collected, no GPS data is used, all communications are encrypted, and there is no way of finding out one's identity or contacts. Activating the contact tracing feature is optional. A positive COVID-19 test must be cross-validated with data from the *Rede Nacional de Dados em Saúde* (RNDS, National Network of Health Data) through a separate portal[6]. This portal has a separate privacy policy in compliance with Brazil's data protection regulation, LGPD. If compared with the other national apps in 4, Coronavírus-SUS is highly satisfactory regarding anonymity, decentralization, and storage of data. A point of improvement, however, would be making the source code available for public auditioning.

Despite meeting most privacy requirements, *Coronavírus-SUS* also faces a severe lack of adhesion. In November 2020, the application had around 10,5 million downloads, reaching roughly 5% of the population. Since the contact tracing feature is optional, the real percentage might be even lower. This scenario did not go through significant change in the last year. Several reasons might be behind the small adhesion, in addition to the ones stated in Section 4. The reachability of the cellular network, regional inequality, and socioeconomic challenges can make it difficult for a centralized effort to reach the population as a whole. However, the main issue is possibly the lack of official advertisement of *Coronavírus-SUS* from governmental offices and the major news outlets.

With most countries undergoing vaccination, the world takes the first steps to control the COVID-19 pandemic. However, DCT apps could still aid in this task. After two years of trials and errors, Coronavírus-SUS and most DCT applications should adapt to meet privacy and adhesion demands. The next challenges for DCT applications are discussed in the next session.

6. Key Research Challenges and Final Remarks

At the beginning of 2020, most countries could not carry out COVID-19 tests on the entire population at first notice due to cost and logistic challenges. DCT applications arrived to help control the pandemic. These apps are a special kind of healthcare app that integrates health data and communication technologies. Being so, governments and corporations responsible for these apps must treat the privacy of its users with utmost concern.

Almost two years have passed since the beginning of the COVID-19 pandemic, and the adhesion of contact tracing technologies everywhere is way less than originally expected. However, we believe these apps can still be useful. The pandemic is still going on in most of the world, and even countries with extensive vaccination coverage are retaking social distancing measures due to variants. DCT apps must adapt their governance

[6]http://validacovid.saude.gov.br

strategy and address both privacy and functionality matters in broad marketing campaigns in order to reach a high percentage of the population.

Coronavírus-SUS, Brazil's DCT application, has similar steps forward. Despite properly addressing the majority of the privacy concerns mentioned in this work, the application cannot fulfill its purpose without broad user coverage. Extensive advertising campaigns are necessary to inform the citizens that the application exists and that the personal data it handles is secure, in a language easily understandable for the majority of the population. To Brazil's advantage, the centralized public health system, the *Sistema Único de Saúde*, has the necessary structure already in place to reach the target population.

Developing and refining DCT applications is not wasted work, even after the COVID-19 pandemic is under control. The new paradigms and policies created for handling this sort of data can still be valuable for different healthcare-related tasks in the future, e.g., remote patient monitoring. As future work, it is necessary to evaluate the impact of different percentages of vaccination coverage on the performance of DCT apps, as well as to propose concrete guidelines for protecting users without occurring in mass surveillance or damaging the user's trust in the application.

References

[1] Use of a mobile application for Ebola contact tracing and monitoring in northern Sierra Leone: A proof-of-concept study. *BMC Infectious Diseases*, 19(1), sep 2019.

[2] Yoshua Bengio, Richard Janda, Yun William Yu, Daphne Ippolito, Max Jarvie, Dan Pilat, Brooke Struck, Sekoul Krastev, and Abhinav Sharma. The need for privacy with public digital contact tracing during the covid-19 pandemic. *The Lancet Digital Health*, 2(7):e342–e344, 2020.

[3] Alessandro Blasimme and Effy Vayena. What's next for covid-19 apps? governance and oversight. *Science*, 370(6518):760–762, 2020.

[4] European Commission. Annex iv: Inventory mobile solutions against covid-19. Available at `https://ec.europa.eu/health/sites/health/files/ehealth/docs/covid-19\_apps\_annex\_en.pdf`. Accessed on 2020/12/31.

[5] Google. Exposure notifications: Helping fight covid-19. Available at `https://www.google.com/covid19/exposurenotifications/`. Accessed on 2020/12/31.

[6] Douglas J Leith and Stephen Farrell. Gaen due diligence: Verifying the google/apple covid exposure notification api. *CoronaDef21, Proceedings of NDSS '21*, 2021, 2020.

[7] Smart Nation and Digital Government Office. Upcoming legislative provisions for usage of data from digital contact tracing solutions. Available at `https://www.sgpc.gov.sg/media_releases/sndgo/press_release/P-20210108-1`. Accessed on 2021/08/26.

[8] World Healh Organization. Coronavirus disease (covid-19) advice for the public. Available at `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public`. Accessed on 2021/08/24.