Slowing the Spread of Infectious Diseases Using Crowdsourced Data

By Tina White, James Petrie, Rhys Fenwick, Zsombor Szabo, Isaiah Becker-Mayer, Daniel Blank, Jesse Colligan, Mike Hittle, Mark Ingle, Oliver Nash, Victoria Nguyen, Jeff Schwaber, Akhil Veeraghanta, Mikhail Voloshin, Sydney Von Arx, Helen Xue

TABLE OF CONTENTS

This white paper was last updated with minor formatting alterations on April 8, 2020. The last alterations to its content were made on April 26, 2020.

INTRODUCTION

Covid Watch Mission Statement

We are a group of volunteers — researchers, software engineers, privacy and public health experts — who have developed a privacy-preserving mobile app intervention to reduce the spread of COVID-19. Our mobile app performs automatic decentralized contact tracing using Bluetooth proximity networks.

Our volunteers care strongly about preserving human life and human rights. All data we could collect is voluntary and fully anonymized. All code is transparent. It is open source and could be easily reviewed, reproduced and used anywhere on the planet.

The app could be installed by anyone with a Bluetooth-capable smartphone, alerting them to their risk of having been in contact with a confirmed case of COVID-19, and helping them to protect themselves and their friends, families, and other contacts altruistically.

We also believe scalable measures like an app are especially helpful in communities where contact tracing resources are too limited to match the scope of the pandemic. We're building this app to provide components and tools that public health agencies can use to supplement their pre-existing efforts to fight COVID-19, assisted by voluntary public action.

Privacy Focus

Existing mobile apps without a privacy focus have been an effective intervention to reduce the spread of COVID-19. However, invasive interventions carry significant human rights costs, including the temporary loss of personal freedom and fears around whether that freedom will be restored.

A mobile app with a strong privacy model may also have greater efficacy because people will be more likely to share accurate data if they know that data is safe. Ensuring privacy prevents COVID-19 patients from being ostracized or socially harmed on account of inadvertent potential data exposures. Also, mobile apps with poor privacy models may further undermine public confidence in responses and exacerbate existing mistrust.

In contrast, our mobile app research has focused on developing a strong privacy model while still providing effective intervention. Using the app, users can alert recent contacts without anyone being able to trace the information back to them. We believe this intervention has the potential to slow or stop the spread of COVID-19 and save lives.

Current Mobile Phone Interventions

South Korea and China have demonstrated two successful systems for containing COVID-19 that make extensive use of technology. The results they have seen match well with predictions from numerical models: with a sufficient diagnosis rate and contact tracing accuracy COVID-19 can be contained.

China was the first to create a mobile app intervention. Their app uses GPS history and other data to assign a risk score. This score is then used to control which individuals are allowed to move freely. China's intervention appears to have been successful,

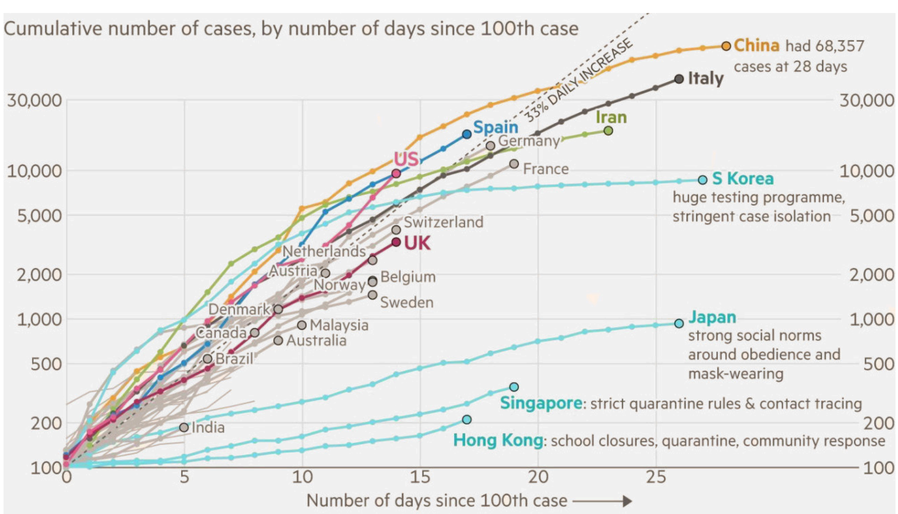


Figure 1: FT Graphic: @jburnmurdoch 19th March, John Hopkins data

COVID WATCH

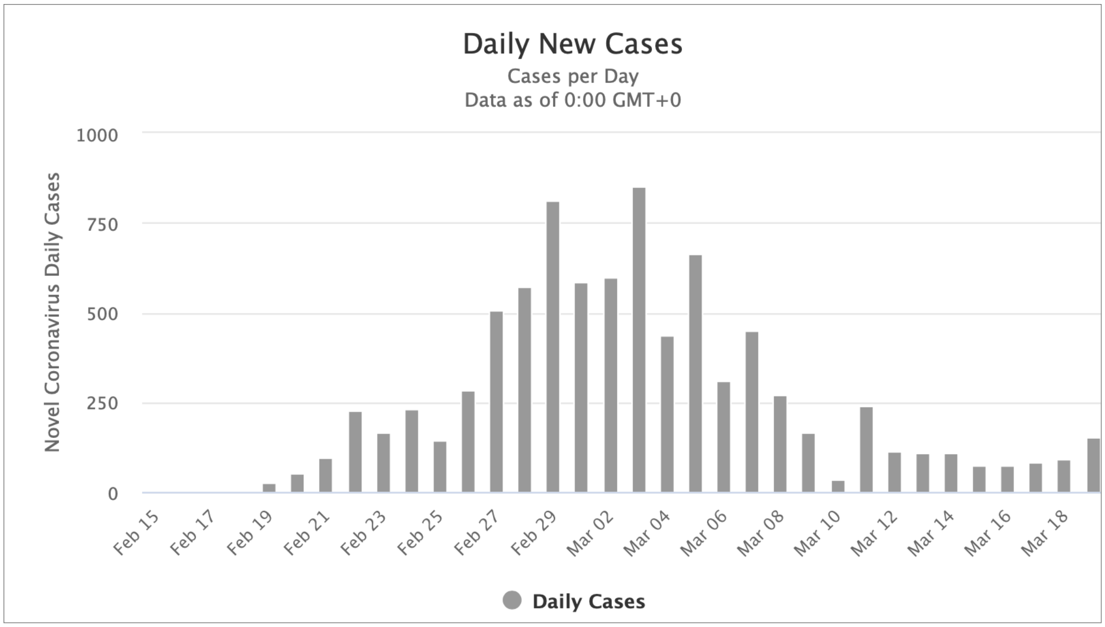


Figure 2: South Korean new cases per day

but required far-reaching state surveillance that, by the standards of most liberal democracies, would be considered highly invasive, likely unlawful, and politically unpalatable.

South Korea publicizes a large amount of information collected from the cellphones of infected patients so that others can determine if they had been in contact. South Korea's success has been attributed mostly to (1) widespread testing (2) contact tracing and (3) case isolation. However, their mobile alert solutions do not effectively anonymize patient data. They gather location data from interviews, mobile phone GPS history, surveillance cameras, and credit card records then send text alerts with the location history of patients. Much like the intervention in China, this appears to be effective, but takes a similarly high toll on personal privacy. We've built a privacy-preserving version of these successful interventions that we believe would have a regulatorily, publicly, and politically viable adoption process within the United States and other Western countries. The system as designed complies with existing regulations around medical information in the United States and does not reveal identifying patient information.

Making Interventions More Efficient

Non-pharmaceutical pandemic interventions fundamentally make a trade-off between two important social goods: (1) loss of life from the pandemic and (2) economic impact, which influences health and well-being outcomes indirectly. Mobile app interventions are a powerful public health tool because they can improve this trade-off.

In general, non-pharmaceutical approaches to infectious disease control have the following components:

- Filtering (picking a subset of the population)
- Intervention (modifying the behaviour of these people)

For example, quarantining patients with a positive diagnosis applies a filter based on testing and then applies the quarantine intervention. Other examples include travel restrictions for at-risk areas, cancelling public events in a specific city, or encouraging more handwashing in an entire country. Some of these interventions, especially self-isolation, are highly effective at preventing the transmission of infectious diseases like COVID-19. The dowside is that they can also be costly to use.

The quality of filtering plays a crucial role in determining the trade-off between loss of life and economic impact. If filtering is poor, a correspondingly larger economic impact will be needed to achieve the same loss of life reduction. Without good filtering, broad quarantines and social distancing are needed, incurring a huge cost in the form of negative impact on people's lives.

Unfortunately, traditional approaches to filtering, such as contact tracing, are labor intensive and don't scale well. So we expect filtering (and, correspondingly, the trade-off between loss of life and economic impact) to degrade in quality as a pandemic grows. But automated contact tracing solutions have the potential to be more scalable — and potentially even more accurate, with access to higher quality information than traditional contact tracing. This may allow for a better trade-off to be maintained in the midst of a pandemic.

PROPOSED SYSTEM: TWO PARTS

The system proposed here is intended to be used as part of a broader campaign to combat COVID-19. These methods focus on gathering and disseminating the information needed to perform targeted interventions.

There are two components to this system that work almost independently, but can be bundled into a single mobile app. Depending on privacy requirements and the needs of specific public health authorities a subset of these capabilities could be utilized.

1. Automated contact tracing at scale using anonymized Bluetooth proximity sensing
2. Recommendations from local health authorities and risk-aware suggestions about when to get tested

Part 1: Bluetooth Contact Tracing

CONTACT TRACING BACKGROUND

Non-pharmaceutical methods focused on social distancing reduce the spread of COVID-19. These methods are based on reducing contact between infected and susceptible people, even when it isn't known who is infected.

In the simplest form, this is being achieved by reducing all social events and increasing precautions like handwashing. This is an effective measure because the number of new infections is roughly proportional to the number of contact events with infectious people.

A more targeted approach employed by many health agencies is contact tracing. This system works by finding and monitoring contacts of patients that have been diagnosed. Individuals that are thought to be infected are then put into isolation to prevent further transmission, and individuals who have previously been in contact with infected individuals are quarantined.

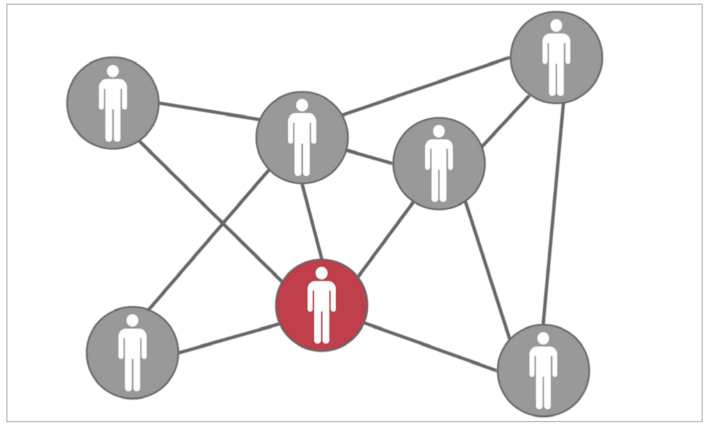


Figure 3: Contact Network

Hellewell et al analyzed the effectiveness of contact tracing as a method to contain COVID-19 at the beginning of an outbreak. **Their findings are promising: with 80% contact tracing accuracy and a mean detection time of 3.8 days after symptom onset, containment is likely.**

In the context of contact tracing there are three parameters that models show can strongly impact results:

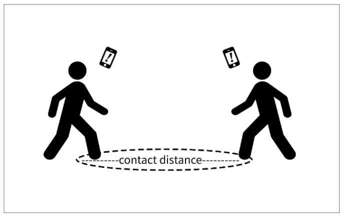
1. Reduction in overall transmission through social distancing
2. Testing rate and time to diagnosis
3. Contact tracing accuracy

MODEL DESCRIPTION

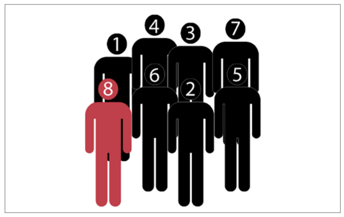
Mobile phones are carried by a majority of people in several countries, with an estimated 3.5 billion users worldwide. They are extremely common in Western society, with over 70% of the entire US population estimated to own one. Bluetooth is a radio protocol that can be used to wirelessly communicate between nearby mobile devices and the signal strength can be used to estimate distance.

Mobile devices can be made to proactively record contact events with other nearby devices by sending Bluetooth signals. By measuring the signal strength and discrete number of contact events, the duration and distance of contact between two phone users can be estimated.

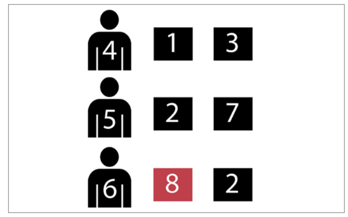
By recording all contact events, a high accuracy list of at-risk individuals can be generated automatically when a new person is diagnosed. These individuals can then be immediately notified to ensure they self-isolate before infecting more people.



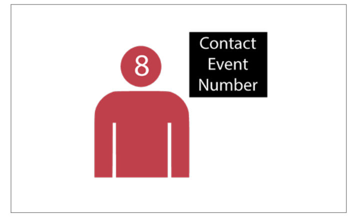
1. Every time two phones are close a contact event is detected.



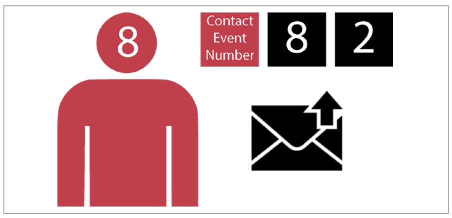
2. Each phone broadcasts a randomly generated contact event number. The number is updated every few minutes.



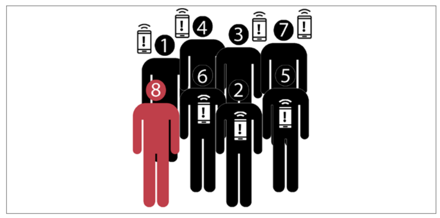
3. All phones keep a local log of every number they have transmitted or received.



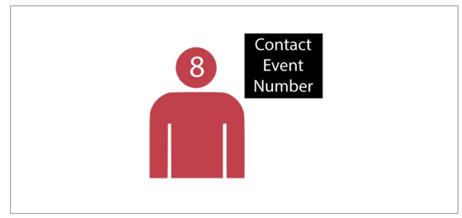
4. If one of the phone owners is diagnosed positive, they are given a permission number by health authorities.



5. This person sends a request to the public database with the permission number and their history of contact event numbers (both transmitted and received).



6. If the permission number is valid, the contact event numbers are stored in the database and transmitted to all other phones.



7. Each phone compares the publicly posted contact event numbers against their own history. If there are any matches this means they were close to an infected individual and are given instructions on what to do next.

Figure 4: Privacy model

Bluetooth proximity may be the most accurate crowdsourcing method for approximating close contact.

While GPS data is a more well-known general technology, there are significant advantages bluetooth has over GPS in terms of accuracy for contact tracing. With bluetooth, proximity can be approximated by signal strength that is reduced by obstructions like walls; therefore, it more accurately reflects functional proximity in high-risk environments for close contact: inside buildings, in vehicles and airplanes, and in underground transit.

Bluetooth communication also occurs directly between mobile devices. This means a decentralized system can be built more easily with and with stronger privacy protection than other crowdsourcing data types like GPS trajectories.

We are pursuing research in developing inexpensive external Bluetooth devices under the same automatic contact tracing alert system for use in countries with fewer smartphone users. These methods would face much steeper adoption challenges, but if mobile app users and external device users could be integrated under a single system, outcomes could be further improved over more global communities.

PRIVACY MODEL

The Bluetooth contact tracing system can be structured in a decentralized and anonymous way using randomly generated and locally stored 'Contact Event Numbers'. This allows the system to function fully without any private information being stored or transmitted.

By generating a new random number for each contact event, the system is able to operate without storing or transmitting any personal information. This method is designed so that only the phones involved in a contact event are able to identify messages on a public database.

The only authentication required is the permission number provided by a public health authority. This permission number is used so that malicious actors cannot send false alarms. After authentication the permission number is deleted from server memory.

The contact event numbers are random and only known by the message recipient and the message sender, so the database can be made public without risk of sensitive information being discovered.

While our current intervention is based on permission numbers, in regions where widespread testing is unavailable, a well-designed symptom sharing questionnaire may perform a similar function, with a higher number of false positives. Research in this direction is currently being done by the CoEpi team.

The most effective form of this intervention would occur in communities that implement widespread testing and where permission numbers are shared with the mobile app by public health departments.

Singapore just released an app that performs some contact tracing using Bluetooth proximity networks. However, our initial analysis suggests their privacy model is susceptible to attack because they did not take advantage of Bluetooth's potential for decentralization.

For a more thorough description of the TCN protocol implementation details, please see the TCN coalition github. We have moved our implementation details there, so that all of the other TCN-based apps can shared a central repo.

DATABASE

The specification for the database is very simple: it is shared across all installations of the app and stores anonymized Contact Event Numbers. If protections against hoaxes are required, permission numbers can be used, but are not an essential requirement of the system. If the database grows too large, it can also be fragmented based on general location. The code and a more in depth discussion of architecture are available on the open source github repo here.

IMPLEMENTATION

Bluetooth contact tracing is implemented via background processes on iOS and Android.

The approach currently being investigated utilizes BLE functionality for background advertisement and scanning. Due to different system requirements for Android and iOS, the protocol works differently depending on the operating systems of the devices involved. The key challenges are:

1. iOS devices acting as "peripherals" in the background can only be found by "centrals" that are scanning for their specific service UUID. These peripherals must establish a connection to transfer any data.
2. Android devices have several unfixed bugs where subsequent connections with many devices can cause the Bluetooth system to lock up.

The current solution is a hybrid model that is asymmetric for communication between iOS and Android. All devices will simultaneously act as peripherals and centrals, but only some devices will be able to detect others, and only some devices will need to establish a connection to exchange data. An extended description of the communication model and the code are available on github here for iOS and for Android.

This model has been successfully implemented as a proof-of-concept, as shown in the following video:

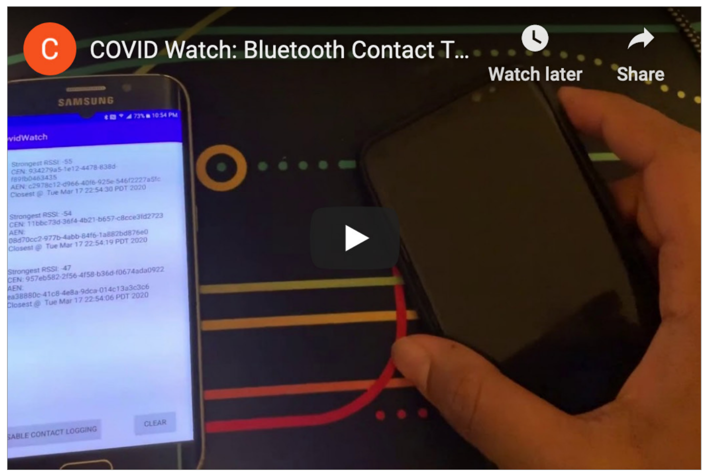


Figure 5: Bluetooth Contact Tracing Functionality Video

Part 2: User Recommendations

We have designed the app user interface (UI) and are building a beta app with the following features:

- CDC general COVID-19 advice, symptoms, and resources
- A notification system of potential COVID-19 contact risk via Bluetooth proximity networks
- Personalized advice: If close contact is detected, a popup instructs the user to call the local public health department, and looks up this number to call for them to inquire about next steps
- Future features may include: more personalized advice based on heat map location, a supplies map, self-reporting of symptoms, FAQs, news updates tailored by geographic region, travel counseling, or access to home-based testing

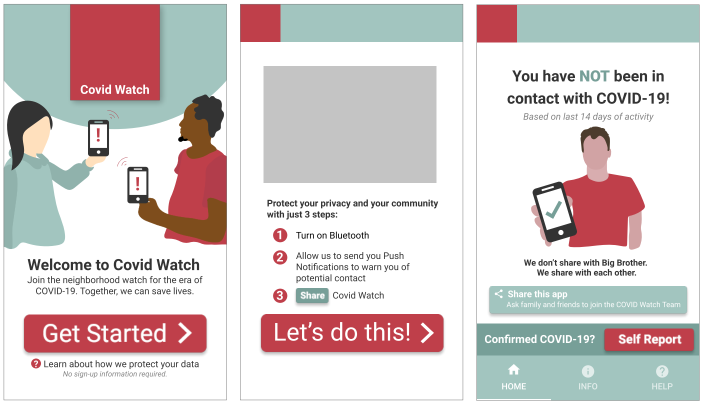


Figure 6: User Interface Design: New User OnBoarding Workflow. Downloading and using Covid Watch does NOT require any sign up of email, password, etc. of any kind. The only requirement is to enable Covid Watch to access bluetooth on your smartphone in order to detect other smartphones in close proximity to log a 'contact event'. If no other smartphones you have been in contact with are associated with a positive case COVID-19, Covid Watch informs you that you have not been in contact with COVID-19.

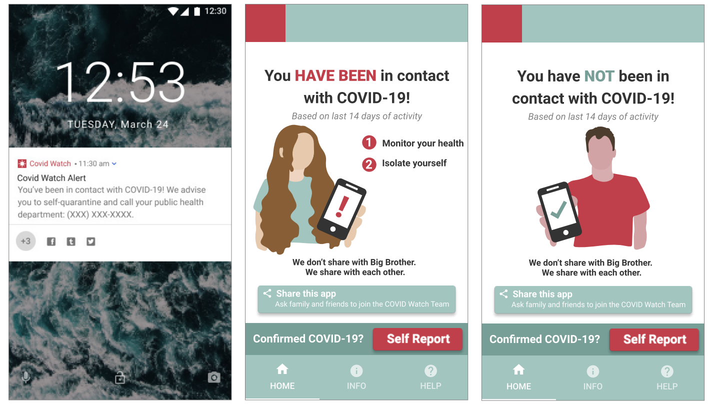


Figure 7: Figure 2: Contact Alert and Reporting. If Covid Watch detects another smartphone within bluetooth proximity that is associated with a positive COVID-19 case, you are alerted that you may have been exposed Covid Watch suggests steps of (1) monitoring your health and (2) isolate yourself. Additionally, you may update your own status as confirmed or not tested, along with the first date of symptoms.

WHY YOU SHOULD CARE

Incentives for Health Authorities

- High accuracy, instantaneous contact tracing
- Targeted interventions based on calculated risk and current health policy
- Easier communication of announcements and information

Incentives for Individuals

- Information about how to avoid contracting the disease
- Earlier warning to protect friends, family, and close contacts if you do get sick
- Friends, family, and close contacts who use the system will likely be warned before they can get you sick
- Broad health measures to contain the disease could be relaxed in favor of such targeted and private interventions, so regular life will not need to be disrupted as much

Quantitative Analysis of Impact

The impact of this technology will depend largely on the state of the system around it. Numerical models and ongoing campaigns suggest that with extensive testing, accurate contact tracing, and isolation of suspected cases, outbreaks can be contained.

For intermediate testing and contact tracing detection rates, a system like this would likely need to be used in combination with social distancing measures and manual contact tracing. However, the measures suggested by Ferguson et al. could potentially be relaxed if supplemented with sufficient targeted interventions.

Of the three target parameters (tracing accuracy, detection rate, base transmission), the potential influence of the app on tracing accuracy is the simplest to quantify. Any two users of the app who are at the same location at the same time will register a contact event. In theory, all transmission events except those by fomites would be detected. This includes all types of contact classified as being "high risk" by the ECDC.

Preventative measures encouraged by the app such as avoiding high-risk areas and increased precautions would reduce overall transmission rate, however it is difficult to quantify this impact. The current design of the app may also increase detection rate by informing users of symptoms to watch for and how to get tested. Ongoing research is being conducted on how to allocate testing resources to maximize the expected value of secondary cases detected.

To have a significant impact the technology will require a significant adoption rate. To detect contact events using Bluetooth both individuals must have the app installed at the time of transmission. Assuming 'P' percent of the population uses the app, 'A' percent of transmission events between app users are detected, and ‘T’ percent of infected are tested, then T*A*P^2 percent of total transmissions are detected (if app users are homogeneous in the population).

Figure 8 examines the percentage of detectable transmission events in countries with different testing rates as a function of app usage. The analysis assumes that these events are uncorrelated, which may significantly under-estimate detected transmissions due to missing chain-reaction testing. Ongoing modelling work is being done to investigate the dynamics of rapid tracing and testing along infection chains.

In order to investigate disease dynamics with clusters of app users, the model developed by Hellewel et al was modified to account for two sub-populations, one that uses the app and another that doesn’t. The modified implementation is available on Github and currently assumes 90% app tracing accuracy and 50% non-app tracing accuracy (assuming health systems are partially overwhelmed).

Figure 9 was used to estimate the average number of infections caused in 6 weeks due to a single imported case. Results show a significant reduction in risk for clusters of app users that are partially distanced from non app users. This is important because it provides an additional incentive for individuals to use the app even when the overall adoption rate is low. As the proportion of the population using the app increases the risk for the entire population is reduced and most outbreaks are contained.

SO FOR THE QUESTION: “CAN AN EFFECTIVE CONTACT TRACING PROGRAM REDUCE LOCAL TRANSMISSION SO THAT SUSTAINED LOCAL SPREAD DOES NOT OCCUR?”

The answer seems to be yes. With a comprehensive testing program, high contact tracing accuracy, and self-isolation of diagnosed individuals, our models predict that each new case could cause on the order of 10 other cases before the outbreak is extinguished. Even in parameter regimes where automated contact tracing alone is not enough, this technology can be used in combination with existing methodologies to provide greater protections with lower social cost.

Also, even in parameter regimes where automated contact tracing alone is not enough, this technology can be used in combination with existing methodologies to provide greater protection with lower social cost..

Timeline to Deployment

We estimate we could complete the remainder of the necessary technical requirements and deploy the first version of the mobile app for pilot within 2 weeks, and possibly sooner, based on the

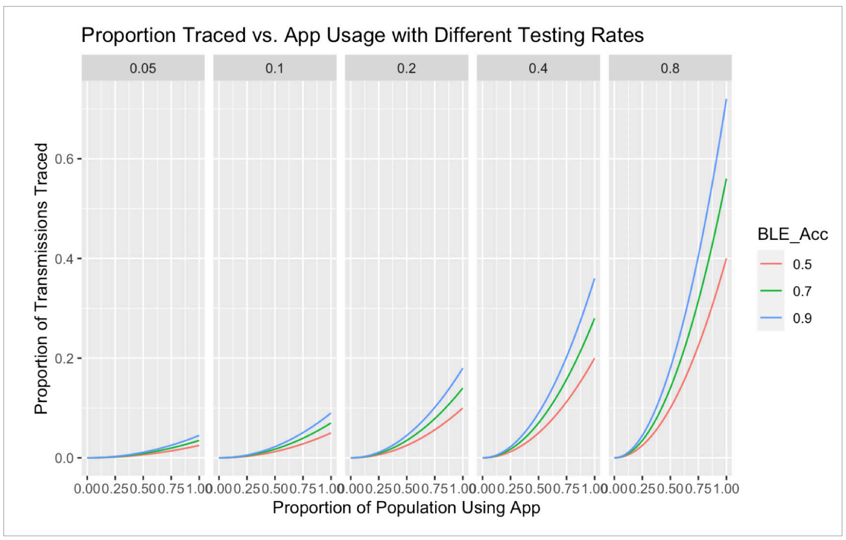


Figure 8: Transmission Detection vs. App Usage curves for testing rates [0.05, 0.1, 0.2, 0.4, 0.8]. BLE_Acc gives the detection rate of transmission events between app users. Work derived from this [Monte Carlo Model](#).

rapid rate of progress that has so far been achieved by the researchers and volunteers who have devoted their time to this technically challenging, but potentially high impact intervention.

The track to development requires (1) more volunteers to help us with skill sets listed on our collaborate page and (2) funding for, at a minimum, cloud services.

The app will be implemented and launched in two parts:

1. The first version will implement a Bluetooth proximity system to notify users of potential close contact exposure to SARS-CoV-2.

2. The second version will build upon the initial launch. Users may be able to self-report symptoms, receive personalized advice based on the local health department, and receive CDC advice and resources (Figure 6). When users come in contact with someone with COVID-19 or symptoms of it, their phones will send them a push notification of the alert. Users could then see their Transmission Phone Log within the application to determine where and when they may have contracted the virus. Additional features include incorporating a locally-stored questionnaire noting recent travel and symptoms to better determine the user's risk level (Figure 7).

We also want to emphasize that high user adoption as quickly as possible after release will facilitate the best possible outcomes for intervention. Historical data from Bluetooth contact events doesn't exist. Therefore, a user needs to download the app in order to start the clock and begin benefiting from the system as it anonymously logs their Bluetooth contact events.

CONCLUSIONS

Mobile technologies can provide instantaneous and high accuracy contact tracing, even between strangers at low social and economic cost. Instead of requiring thousands of healthcare workers to do this manually (as is the current approach in China) the process will be essentially cost-free. Also, because the system will be so accurate, a majority of people can continue to live their lives without the need for increased social distancing.

We are developing this technology as a high-quality filter to be used for the pandemic optimization problem. Combined with a comprehensive testing program, our filter may be powerful enough to stop further spread of COVID-19.

Who Can Help

We continue to seek more collaborators, donors, and support from health organizations and testing centers.

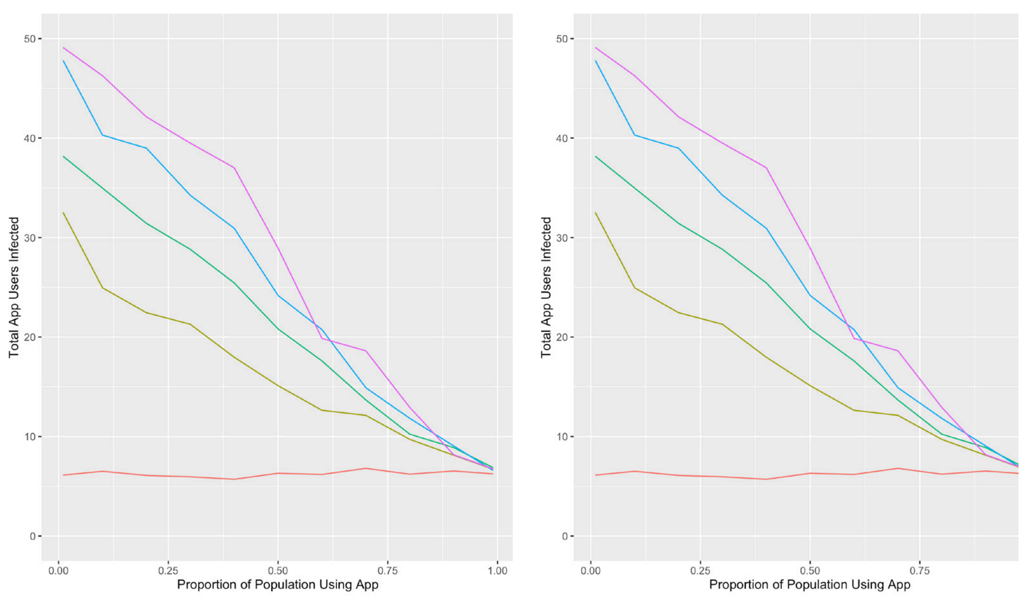


Figure 9: Expected infections for total population (1st) and app population (2nd) adjusted for relative population size.

Contributors, Advisors, and Acknowledgements

CONTRIBUTORS

NAME	ROLE	LOCATION
Tina White	*Executive Director*	Stanford University, USA
Rhys Fenwick	*Head of Communications*	Wollongong, Australia
Isaiah Becker-Mayer	*Software Engineer*	California, USA
James Petrie	*Head of Research*	University of Waterloo, Canada
Zsombor Szabo	*Bluetooth Team Lead*	Budapest, Romania
Daniel Blank	*Software Engineer - Heatmap*	Berkeley, USA
Jesse Colligan	*Software Engineer*	New York, USA
Mike Hittle	*Heat Map Team Epidemiologist*	Stanford University, USA
Mark Ingle	*Bluetooth Team Software Engineer*	South Carolina, USA
Oliver Nash	*Software Engineer*	London, United Kingdom
Victoria Nguyen	*UI Designer*	California, United States
Jeff Schwaber	*Software Engineer*	VP of Engineering at Bigtincan, USA
Akhil Veeraghanta	*Bluetooth Team Software Engineer*	University of British Columbia, Canada
Mikhail Voloshin	*Heat Map Team Lead*	North Carolina, United States
Sydney Von Arx	*Human Resources*	Stanford University, USA
Helen Xue	*Volunteer*	California, USA

PRIVACY AND SECURITY ADVISORS

NAME	ROLE	LOCATION
Tessa Alexanian	*Privacy and Security, iGEM, East Bay Biosecurity Group*	Berkeley, USA
Peter Eckersley	*Privacy and Security, stop-covid.tech and Distinguished Technology Fellow, EFF*	Australia
Jeffrey Ladish	*Privacy and Security, Security Consultant, Gordian Research*	Berkeley, USA

PUBLIC HEALTH AND EPIDEMIOLOGY ADVISORS

NAME	ROLE	LOCATION
Megan Coffee	*MD, PhD, Infectious Diseases, NYU and Public Health*	Columbia, USA
Jolene Elizabeth	*Public Health Research and Digital Health Innovation*	California, USA
Celine Gounder	*MD, ScM, FIDSA, Internist, Infectious Diseases Specialist, Epidemiologist*	New York, USA
Plinio Morita	*PhD, Public Health and Health Systems Specialist*	Waterloo, Canada
Julie Parsonnet	*MD, PhD, Infectious Diseases Specialist and Epidemiologist*	Stanford University, CA