**TraceDefense v2**

**It is possible to avoid many people dying and to have an economy.**

Running unchecked, the death toll of COVID-19 is estimated to be in the millions in the USA[[1]](#footnote-2). South Korea provides the best evidence that this is avoidable using *mobile contact tracing,* while maintaining a functioning economy; they dealt with an 8000-person outbreak generated by a super-spreading event in a church while avoiding severe economically-damaging lockdowns ([Propublica Link](https://www.propublica.org/article/how-south-korea-scaled-coronavirus-testing-while-the-us-fell-dangerously-behind) & [Science Mag Link](https://www.sciencemag.org/news/2020/03/coronavirus-cases-have-dropped-sharply-south-korea-whats-secret-its-success) ). Additional evidence is provided by China (outside of Hubei), Taiwan, and Singapore, all of which have effectively dealt with outbreaks without shutting down their economy. In addition to South Korea’s mobile contact tracing effort[[2]](#footnote-3), Singapore has recently released a mobile contact tracing app, TraceTogether ([Link](https://www.mobihealthnews.com/news/asia-pacific/singapore-government-launches-new-app-contact-tracing-combat-spread-covid-19)); Israel passed an emergency law permitting the use of mobile data for COVID-19 contact tracing ([Link](https://techcrunch.com/2020/03/18/israel-passes-emergency-law-to-use-mobile-data-for-covid-19-contact-tracing/)); the United Kingdom looks also to be developing a mobile contact tracing app ([Link](https://www.nytimes.com/2020/03/19/us/coronavirus-location-tracking.html), [Link](http://www.ox.ac.uk/news/2020-03-17-infectious-disease-experts-provide-evidence-coronavirus-mobile-app-instant-contact), and [Paper](https://github.com/BDI-pathogens/covid-19_instant_tracing/blob/master/Manuscript%20-%20Modelling%20instantaneous%20digital%20contact%20tracing.pdf)). To succeed, we want to *scalably* and *rapidly*: **Test** aggressively for the virus; **Trace** out the contacts with a >1% chance of infection; **Timeout** these traced contacts through evidence-based public health interventions based on level of exposure (e.g., immediate self-quarantine vs. serial symptom check) ([Link](https://www.newyorker.com/news/news-desk/keeping-the-coronavirus-from-infecting-health-care-workers?utm_source=twitter&utm_social-type=owned&utm_brand=tny&mbid=social_twitter&utm_medium=social)). The **Test/Trace/Timeout (TTT)** approach is only as strong as the weakest link. Our goal is to support the Tracing component of this strategy in an easy, highly scalable, and rapid fashion; we envision this as part of a combined strategy. **TTT** is the ‘backbone’ of the response strategy suggested by the WHO ([Link](https://www.weforum.org/agenda/2020/03/testing-tracing-backbone-who-coronavirus-wednesdays-briefing/)).   
  
**Summary:** Governments are currently dropping the ‘hammer’ of massive lockdowns, hoping to mitigate the spread and buy time for an effective solution. When the ‘hammer’ is removed, TraceDefense is designed to be in place, ready to prevent further deaths, while keeping the economy going. One could imagine this app being used by employers to verify that the employees who are coming in are in fact safe, using a barcode. The hope then is either: the virus can be directly strangled with **TTT** or a vaccine can be developed, while the local ‘fires’ are extinguished by **TTT**. The timeline is two weeks to have a tracing app, with the infrastructure to support 100M users. We are currently coordinating with the University of Washington to ensure access to validated positive tests; the University of Washington is conducting 90% of COVID19 tests in Washington State ([Link](http://depts.washington.edu/labmed/covid19/)). The next steps involve a large-scale coordination with all relevant parties (states, the CDC, government, testing providers, other tech companies, etc).

**Current status**: 2020-03-23

* Have an Android app prototype available which periodically logs GPS data.
* Coordinating with MIT safe paths project: <http://safepaths.mit.edu/> <https://github.com/tripleblindmarket/private-kit>
* Connections established with UW labs running COVID-19 tests and the Chief Information Officer at Boston Public health commission.

**Some concerns** we have encountered:   
1) Has the virus grown beyond the scale where this approach can work? The virus has a transmission time of ~1 week so a lockdown can reduce the number of cases as demonstrated in Hubei (see [Link](https://medium.com/@tomaspueyo/coronavirus-the-hammer-and-the-dance-be9337092b56) ) allowing the **TTT** approach to then take over and work.

2) Some epidemiologists (see [Link](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf) ) doubt that the suppression approach can work. However, when tracing is brought to this picture ([Link](https://www.sciencedirect.com/science/article/pii/S2214109X20300747)), modeling also demonstrates success. Furthermore, while some epidemiological models appear simplistic, the success from China and South Korea provide credible data, along with the WHO’s recommendation.

3) We do *not* need to achieve 100% success with this approach. All that’s required is succeeding a 1 – 1/R fraction of times where R is the reproduction number. Credible estimates place R at 2.5, implying 60% success is sufficient. We of course aim higher for a more robust solution.

## Solution concept

It is important that we create a solution that is easily accessible by all populations and people groups. We propose a smartphone application that performs the basic function of Contact Tracing.

**User Experience Flow:** It is currently captured as follows:

* **Initial signup experience:** Download the app, accept EULA terms of active data collection, and then provide two levels of privacy, users will be allowed two options to get started:
  + Microsoft Account – Sign in with an MSA for a secure experience that keeps your data completely private.
  + Phone number – Sign in with just a phone number, but display a warning message that indicates that using an MSA account gives more control over the data.
* **Standard interface:** Provide a simple risk status screen that displays a status to the user as to actions they should take. E.G. “No known transmission vectors, use personal social distancing” or “High risk of transmission, quarantine at home, a quarantine kit is on the way”.
* **Reporting positive test:** If a user gets a positive test results back for COVID, they can report it through the app by providing a unique test-id that is then verified with a central government or health resource (that we will have to partner with), e.g., 90% of COVID-19 tests done so far in Washington state are run by UW Virology. COVID-19 tests include information such as name, address, phone number date and time of sample collection. Approximately 24 hours after the sample is taken, data is in a SQL database maintained by UW IT. The vast majority of tests being conducted in the Seattle area are going to UW Virology and will be in the database and likely most tests from around the state as well. All data eventually sent to state DOH, but it first hits this UW database in Seattle. Similar entities might exist in other states.
* **Notification to user of possible exposure:** Whenever a positive test is recorded, we will run analytics on our secure servers to identify all other signals of people were close contacts of the COVID19 + individual. Close contacts will be defined as within 2 meters of individuals for 10 minutes or greater. Upon identification of a match, the people who were possibly exposed will receive a text message and app alert notifying them to take evidence-based public health recommendations (e.g., serial symptom check, immediate quarantine etc.) and initiation of a care package. (In the longer term, Machine Learning and other sensor modalities can improve the precision here.)
* **Retrospective data incorporation:** Using Project Bali, we can also allow the user to bring in their own historical geo-location data from other app sources (E.g. Google Maps, Facebook, Twitter, Instagram etc). This can be critical if for example the user just received a positive test result, but hadn’t been using the live tracking application up until this point. By using Bali to pull in these historical data sets, we can allow for user’s geo-location signals from other sources to be used to help fill in the data gap. Furthermore, we can use these partnerships to bring powerful partners on-board (E.G. Google / Apple), who can help socialize and bolster this effort.
* **Bluetooth / Ultrasonic token exchange:** Pursuant to the VirusWatch idea ([VirusWatch.docx](https://microsoft-my.sharepoint.com/:w:/p/sclundbe/EcUlhwF8I19DvlNyM1ELEUsBmR_WzkN3ePWasCBeseLjfQ?e=OsWc9x)), we can also build a token exchange based on proximity of smartphone devices to retroactively track possible virus contact points. Singapore has already implemented a form of this idea as documented here: [Trace Together Link](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.tracetogether.gov.sg%2F&data=02%7C01%7Cjolarso%40microsoft.com%7Cd4c537ed152647b4580208d7ce8351fb%7C72f988bf86f141af91ab2d7cd011db47%7C1%7C0%7C637204935737239946&sdata=pe83j4dXzJxhOrRCTX0GDDcVQNkagPmucelW7wZfBMk%3D&reserved=0). We are working to get open source access to the app. This is not necessary for the first MVP, but can be integrated in as soon as it is ready. It should also be noted that this is just but one tool in the toolbelt as this type of app does NOT trace geolocations and thus transmissions due to shared infected surfaces wouldn’t be captured.

**Data:** By default, all data is stored locally on the user’s phone with logs wiped on a rolling 14 day window. The following is the data the baseline data collected on a regular interval

* Location (Latitude / Longitude)
* Timestamp

Additional data from other sources (bluetooth, FB profiles, ultrasonics, etc...) will augment this log over time. The only data sent to the server is a location (with a random offset) and radius (perhaps 5 miles) stored under a salted phone number.

As new place-times of infection are discovered, users of the app can volunteer to provide their information or a trace log can be constructed in the traditional way. The infected log is uploaded to the server which checks the location/radius of every registered phone and provides the infected log for download to the nearby phones. The phone computes whether a contact event has occurred on downloaded traces and alerts the user if a contact is discovered. The phone app can then follow-up with quarantine suggestions, delivering a care kit, etc... as will help the process.

**Scaling:** It’s critical that the backend not flop as the system scales up. The core data might require ~24 bytes per minute per person, implying about ½ MB of storage on a phone over a 14 day time window. Allowing room for sloppy storage formats or additional information (such as bluetooth), perhaps 10MB is a reasonable upper limit.

On the server side, the minimum scale at which this is useful is probably about 100M people. Storing location/radius contact information for a phone requires ~100bytes, implying that 10GB of storage is required for contact information. Infected traces might be at most 10MB in size, implying that for ~100K traces, we need to about 1TB of data. Infected traces are wiped on a 14 day rolling basis.

For every positive case, we need to check whether it intersects with the location/radius of a phone on the server. This could be brute forced as an O(nm) computation where n is the number of participants and m is the number of positives, although it’s easy to achieve an O(m log n) solution using a tree data structure or locality sensitive hashing.

**Privacy:** Every trace is sequestered to a phone unless an infection is indicated. If an infection is detected, and the user allows the data to be uploaded, there is a fundamental limit to anonymity which applies even when healthcare professionals engage in contact tracing. In particular, contacts of an infected person could intersect their contact lists to discover to rapidly discover who is infected.

The design above implies that no additional information is provided to the user of an app.

A hacker who takes control of their phone would be able to observe the trace of infected people in their local area which may provide some additional de-anonymization. This level of de-anonymization appears to be comparable to what the South Korean health authorities simply released.

**Security** – All communication between a phone and the server uses a 128 bit cryptographic hash of a 128bit random number and the phone number to identify the phone to authenticate. The server is designed to support the following (minimal) set of API calls:

1. Register(location, hash(salt,phone)) // registers a new phone. May be called multiple times by the same phone if it moves around with the server keeping a trace of these calls.
2. Status(hash(salt,phone)) // returns any new nearby traces or nothing.
3. Upload(trace, authenticator) //uploads a trace for someone who is infected as well as evidence (via the authenticator) that this corresponds to an infected trace. Note that this interface allows the upload of authenticated traces which do \_not\_ necessary correspond to a phone log.

## Related Efforts

Our approach here is to provide the quickest possible response to the situation given the tools available. This project will aim to consolidate many of the smartphone telemetry projects across the COVID HackBox. Specifically, this project will incorporate VirusWatch, Bali, and several other projects that need to be merged and focused into a single point of effort. Exploring partnerships with other potential experienced app creators such as MileIQ is also worth investigation.

## Next steps and challenges

**Scaling, adoption, and other risks:** This is an effort that has the potential to save lives, get the USA economy going, and potentially help with the global pandemic beyond the USA. There are multiple challenges ahead and time is of the essence:

* **Health / Government Partnerships** – In order to prevent abuse, we need verifications of positive test results. Ultimately, we’d most like to see a central clearing house for all tests in a global database, but this is impractical at this stage given the many labs conducting tests. We need some partners or some method of validating a confirmed infection beyond simply the users’ word through the application. Ideally, we could partner with labs to securely double-check test ID’s and their status. State health departments, which are ultimately responsible for communicating with the CDC, are the place to go as labs within the state report to them efficiently following a positive test. This would enable rapid verification positivity within a state.
* **Adoption –** This exercise is only useful if it can achieve high levels of adoption. Many countries, including democracies, have laws to make contract tracing in public health emergencies. Adoption would involve advertising it on all available channels (news, social media, etc); gathering support and coordination with all relevant parties (the CDC, the state and federal governments, Google, Facebook, TraceTogether, etc); and helping with the individual incentive structure. For the latter, this may help those who are in quarantine to receive as much support as possible.
* **Incentives –** There needs to be a clear incentive, other than moral benefits, to the user for claiming they are infected. We need to work with partners like CDC/Govt to announce such benefits.
* **Privacy –** There needs to be a clear incentive, other than moral benefits, to the user for claiming they are infected. We need to work with partners like CDC/Govt to announce such benefits. We especially need to carefully approach the privacy for users who choose not to use an MSA account and to be mindful that a user may be traced back based on their actual movement patterns. Because users are voluntarily disclosing this information, we must assure them safety of their data.

## The Current Team:

**Microsoft**

John Langford\*, Sham Kakade\*, Jonathan Larson\*, Sreekanth Kannepalli\*, Chris White, Vikram Dendi, Scott Lundberg, Deepthi Uppala, Rafah Hosn, Rajan Chari, Jacob Alber, Alexey Taymanov, Jack Gerrits, and Cheng Tan (this list is open and will be updated as we are merging teams)

**University of Washington**

Shyam Gollakota\*, PhD, Jacob Sunshine\*, MD, Justin Chan, PhC, Sham Kakade, PhD

**Sound Life Sciences**

Shyam Gollakota\*, PhD, Jacob Sunshine\*, MD

\*are the point of contacts

## Appendix

## Survey of (selected) World Responses:

China, South Korea, Singapore and Taiwan all successfully halted the spread of Covid-19; these countries were prepared for the war with Covid-19 due to earlier outbreaks of SARS and MERS. Below is a brief summary of relevant world responses involving contact tracing.

* China: After severe lockdowns, teams of physical tracers were used (in Wuhan, 2K teams with > 5 people per team were used. [WHO report](https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf)).
* South Korea: The government avoided lockdowns and had a functioning economy ([Propublica Link](https://www.propublica.org/article/how-south-korea-scaled-coronavirus-testing-while-the-us-fell-dangerously-behind) & [Science Mag Link](https://www.sciencemag.org/news/2020/03/coronavirus-cases-have-dropped-sharply-south-korea-whats-secret-its-success)). After the MERS outbreak in 2015, the South Korea has the authorization to obtain cellphone records, credit card receipts, and other private data in the event of a public health emergency ([Propublica Link](https://www.propublica.org/article/how-south-korea-scaled-coronavirus-testing-while-the-us-fell-dangerously-behind)). Testing is widespread.
* Singapore: has widespread testing (>SouthKorea), teams of physical tracers. ([Link](https://www.straitstimes.com/singapore/health/how-contact-tracers-track-down-the-people-at-risk-of-infection) and [CDC Report](https://www.cdc.gov/mmwr/volumes/69/wr/mm6911e1.htm)), and surveillance. Recently, the mobile app TraceTogether was released (see below).
* Taiwan: [Link](https://www.nytimes.com/2020/03/17/world/asia/coronavirus-singapore-hong-kong-taiwan.html)
* Israel: passed an emergency law permitting the use of mobile data for contact tracing ([Link](https://techcrunch.com/2020/03/18/israel-passes-emergency-law-to-use-mobile-data-for-covid-19-contact-tracing/)).
* United Kingdom: The UK looks to be developing a mobile contact tracing app ([Link](https://www.nytimes.com/2020/03/19/us/coronavirus-location-tracking.html), [Link](http://www.ox.ac.uk/news/2020-03-17-infectious-disease-experts-provide-evidence-coronavirus-mobile-app-instant-contact), and [Paper](https://github.com/BDI-pathogens/covid-19_instant_tracing/blob/master/Manuscript%20-%20Modelling%20instantaneous%20digital%20contact%20tracing.pdf)).

## TraceTogether summary

Singapore recently released a mobile contact tracing app ([Link](https://www.mobihealthnews.com/news/asia-pacific/singapore-government-launches-new-app-contact-tracing-combat-spread-covid-19) and [Trace Together Link](https://nam06.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.tracetogether.gov.sg%2F&data=02%7C01%7Cjolarso%40microsoft.com%7Cd4c537ed152647b4580208d7ce8351fb%7C72f988bf86f141af91ab2d7cd011db47%7C1%7C0%7C637204935737239946&sdata=pe83j4dXzJxhOrRCTX0GDDcVQNkagPmucelW7wZfBMk%3D&reserved=0)):

* Find phones that are near (2 meters) for >10 minutes using bluetooth broadcasting.  If two phones meet this criterion, they store a record encrypted locally on the phone.
* It requires government intervention to be effective.  Upon detecting an infection, Singapore’s MOH will take the user’s data from the app and use that data to then contact others.
* When a person is contacted by Singapore MOH, he/she is required by law to assist in the activity mapping their movements and interactions or face possible prosecution under Singapore’s Infectious Disease Act.
* It would be useful to reach out to the Singapore team if they are willing to share code / infrastructure. This could help bootstrap the app that we are building.

## Sensor Modalities

Location: Using a location sensor (latitude/longitude/time) to create a trace is similar to the way that traditional contact tracing works. The information generated by it is quite strong for de-anonymization purposes since anyone with a memory of what happened along the trace can discover who it was from. This modality must be supported by any effective tracing app for backwards compatibility with laborious contact tracing techniques. For example, if someone is infected who is not running the app, it is desirable to allow a trace to be created by traditional techniques and then uploaded. Particularly for location data, it’s important that the information is simply stored on the phone rather than on a server in order to zero out the chance of inadvertant disclosure. In addition, storage and compute on the client side makes scaling easier on the cloud side.

Bluetooth: Bluetooth has an advantage over location services in providing more location precision and more anonymity. It has a disadvantage over location based approaches because fomite transmission modes cannot be tracked, not all phones do bluetooth, phones that periodically sleep may miss each other’s bluetooth beacons, and because it does not interoperate with traditional tracing. There are actually two kinds of bluetooth protocols. One of these simply does beacon recording from other nearby bluetooth devices. When someone is infected, the beacons the beacon of the corresponding phone is uploaded and broadcast, with other phones checking to see if they have seen the beacon. The second bluetooth creates greater anonymity by changing the beacon of the phone’s own bluetooth randomly every 15 minutes. Contact detection proceeds as before except that the beacons provide no inadvertent information about their source and cannot be tracked over time. In both cases, it’s important to add loose location information to focus the search to a finer degree.

Ultrasonic: Ultrasonic is a proximity sensor similar to Bluetooth, except that it may provide even finer proximity detection and always requires two cooperating devices.

1. If 50% of the USA is infected and with a 1% mortality (an optimistic estimate, assuming ICUs are not overloaded), this leads to >1.5 million deaths. That 50% will be infected is plausible because there is no evidence for natural immunity. [↑](#footnote-ref-2)
2. After learning from a MERS outbreak in 2015, the South Korean government legally has the authorization to obtain cellphone records, credit card receipts, and other private data in the event of a public health emergency ([Propublica Link](https://www.propublica.org/article/how-south-korea-scaled-coronavirus-testing-while-the-us-fell-dangerously-behind)). [↑](#footnote-ref-3)