REACT: Real-Time Contact Tracing and Risk Monitoring using Privacy-Enhanced Mobile Tracking

Li Xiong[1], Cyrus Shahabi[2], Yanan Da[1], Ritesh Ahuja[2], Vicki Hertzberg[3], Lance Waller[4], Xiaoqian Jiang[5], Amy Franklin[5]
[1]Department of Computer Science, Emory University, USA
[2]Department of Computer Science, University of Southern California, USA
[3]Nell Hodgson Woodruff School of Nursing, Emory University, USA
[4]Department of Biostatistics and Bioinformatics, Emory University, USA
[5] School of Biomedical Informatics, University of Texas Health Science Center, USA
{lxiong,yanan.da,vhertzb,lwaller}@emory.edu,
{shahabi,riteshah}@usc.edu,
{xiaoqian.jiang, amy.franklin}@uth.tmc.edu

Abstract

Contact tracing is an essential public health tool for controlling epidemic disease outbreaks such as the COVID-19 pandemic. Digital contact tracing using real-time locations or proximity of individuals can be used to significantly speed up and scale up contact tracing. In this article, we present our project, REACT, for REAal-time Contact Tracing and risk monitoring via privacy-enhanced tracking of users' locations and symptoms. With privacy enhancement that allows users to control and refine the precision with which their information will be collected and used, REACT will enable: 1) contact tracing of individuals who are exposed to infected cases and identification of hot-spot locations, 2) individual risk monitoring based on the locations they visit and their contact with others; and 3) community risk monitoring and detection of early signals of community spread. We will briefly describe our ongoing work and the approaches we are taking as well as some challenges we encountered in deploying the app.

1 Introduction

More than 6.5 million people in the U.S. have been infected with the coronavirus (COVID-19) and more than 200,000 have died as of September 2020[1]. While there has been a slowdown in new infections in recent weeks, tens of thousands of new cases are still reported daily nationwide.

Contact tracing [12] is an essential public health tool for controlling epidemic disease outbreaks such as the COVID-19 pandemic, involving identification and follow-up of all individuals who may have come into contact with an infected person. In traditional and current CDC-recommended practices[2], contact identification is conducted by asking about the person's activities. This process, however, does not scale. It is time-consuming and ultimately infeasible in public health crisis for large scale contact tracing, as is the case in COVID-19. Failure of traditional contact tracing necessitates alternatives with high degrees of community acceptance [11]. In addition, contact data collected in this way may be incomplete (limited to known contacts) or unreliable. Digital contact tracing using real-time locations

[1]`https://covid.cdc.gov/covid-data-tracker`

[2]`https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/contact-tracing.html`

or proximity of individuals can significantly speed up and scale up contact tracing, as demonstrated by many efforts in Asia and Europe [22, 29, 27, 21, 30, 41].

In March 2020, we submitted a RAPID proposal to NSF which was funded in April as a collaborative project among investigators at Emory University[3], University of Southern California[4], and UT Health Science Center[5]. The goal of our project is to develop techniques and a mobile application, REACT, for REAl-time Contact Tracing and risk monitoring via privacy-enhanced tracking of users' locations and symptoms.

In early April 2020, a draft of a seminal paper that was later published in Science [13] was released that made the case for digital contact tracing for COVID-19. The main observation from the paper was that since typically only symptomatic cases can be contact-traced, to bring R_0 below 1 (and hence stopping the spread), it was important to notify the symptomatic patient's contacts as soon as possible, which would only be possible through digital contact tracing. This was the first paper based on analysis of real-world COVID-19 data to make such a case, even though prior to that several studies (including our RAPID proposal to NSF in March) suggested the usefulness of digital contact tracing for COVID-19.

In mid April 2020, Apple and Google announced their proposed method of using mobile phone's bluetooth to exchange secure and anonymous tokens among nearby phones, which would then be used to notify device owners if they were in proximity of someone (actually someone's phone) who has been diagnosed with COVID-19 by health authorities. The approach was elegant but not very effective. In fact, one could argue[6] that the spatiotemporal data Apple and Google already collect from users, i.e., user mobility patterns, is much more useful for digital contact tracing.

Since then, discussions and efforts about creating contact tracing apps in the U.S. have become mired in battles over privacy concerns and inconsistent responses from different states and stakeholders. Our own project was politicized (ungroundedly) by Breitbart[7] in early May. There is also uncertainty about how much digital contact tracing would help the overall response to the pandemic compared to other measures including social distancing and mask wearing that are now largely adopted in the U.S.

Indeed, a critical issue in using real-time location traces of users for digital contact tracing is user privacy. A location trace can expose users to attacks such as unwanted spams/scams or physical danger, especially in the uncertain times at present. Location traces can be also linked to other information to disclose sensitive information about an individual, e.g., political views and religious inclinations.

Many contact tracing applications, including the ones from Apple and Google, use Bluetooth-based proximity only, not absolute locations, to protect privacy. Examples of this include official contact tracing apps from countries such as United Kingdom, Switzerland, Germany. Of those apps, some keep the contact data locally in the user's phone while others upload the contact data to a central location (e.g., Singapore, Australia). However, ignoring absolute locations sacrifices the ability to estimate the fine-grained transmission risk based on the type of the locations and identified hot spots, and the ability to trace indirect contacts. Another drawback is that bluetooth can "travel through walls" and wrongly identify someone in a neighboring room as a contact. A select few (e.g., Norway) collect both bluetooth contact data and GPS location data. This approach has led to privacy concerns and consequently a low adoption rate among citizens[8]. There have been apps that require mandatory location check-ins from citizens issued by governments like China [38]. While highly effective for containment interventions,

[3] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2027783
[4] https://www.nsf.gov/awardsearch/showAward?AWD_ID=202779
[5] https://www.nsf.gov/awardsearch/showAward?AWD_ID=2027790
[6] https://medium.com/@csatusc/why-we-need-more-than-bluetooth-data-to-fight-covid-19-64da29b3164e
[7] https://www.breitbart.com/asia/2020/05/04/usc-emory-creating-coronavirus-surveillance-system-similar-to-chinas-social-credit-scoring/
[8] https://www.bbc.com/news/technology-52355028

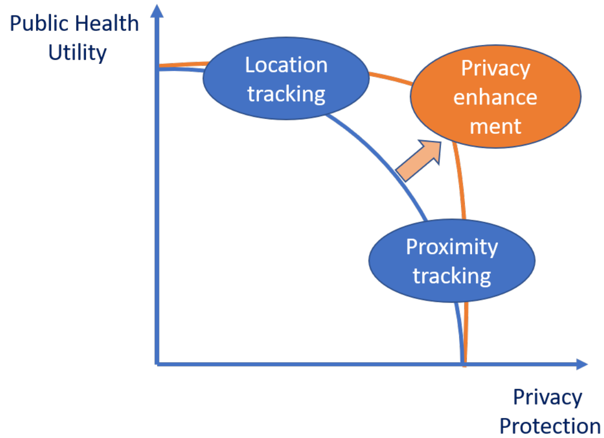

Figure 1: Public Health Utility and Privacy Tradeoffs

these apps have also heightened concerns about surveillance and data abuse.

Besides using cellphone apps, passive digital tracking is gaining popularity. The Chinese government now complements contact tracing efforts by using direct cellphone location information as well as train/plane travel data, while South Korea opts to send mass alerts describing the locations visited by infected individuals. Wi-Fi localization is being used to check for adherence to social distancing directives within industrial and university campuses. For example, using the beam-forming technology of routers to track precisely the flow of people through different regions/floors/buildings and to determine how well campus buildings are implementing the social distancing[9]. University campuses such as USC still opt for a fusion of manual contact tracing with digital contact advice, for example by requesting possible contacts of an infected individual to call the Public Health number for next steps. However, this approach does not eliminate the privacy or scaling problem discussed earlier.

We believe a pandemic like COVID-19 requires a careful design of privacy protection—with public health benefits and privacy enhancement approaches—that optimizes the tradeoffs (as shown in Figure 1). The acceptability of contact tracing technology and the ethical use of it mainly depend on privacy, voluntariness, and beneficence of the data [33]. As governments reopen activities and businesses, contract tracing and risk monitoring remain important components of the public health response along with other measures such as testing and support for quarantine.

The goal of our project is to develop techniques and a mobile application, REACT, via privacy-enhanced tracking of users' locations and symptoms. Figure 2 gives a schematic for our framework. Users can voluntarily submit their locations and symptoms to the server, in addition to the proximity information that is captured by Bluetooth. To enhance privacy, users can control and refine the precision other users with whom their information will be collected and used. We are developing a multi-stage privacy approach where users can upload perturbed locations and adjust the privacy level or precision of the location to be uploaded as their risk evolves. Given such privacy options and enhancements, we hope that REACT will enhance contact tracing of individuals who are exposed to infected cases and allow identification of hot-spot locations for decontamination or increased surveillance to control further spread. The key is to develop efficient and scalable spatiotemporal data structures and algorithms for contact tracing queries given the potentially large number of users and the multi-resolution or perturbed location traces. More importantly, our vision is to go beyond contact tracing and support *individual risk monitoring*. We hope to develop a learning-based approach to estimate the risk for the users based on the locations they visit and their contact with others, so they can receive a real-time exposure risk score and be informed and alerted, e.g., they can self-quarantine or get tested

[9] https://tippersweb.ics.uci.edu/covid19/d/IwAc1O9Wk/covid-19-effort-at-uc-irvine?orgId=1

when the risk is high. Finally, we also plan to use the data collected for community risk monitoring using a social network sensors approach by monitoring a random group and a friends group, to detect early signals of community spread to prepare for larger-scale infections.

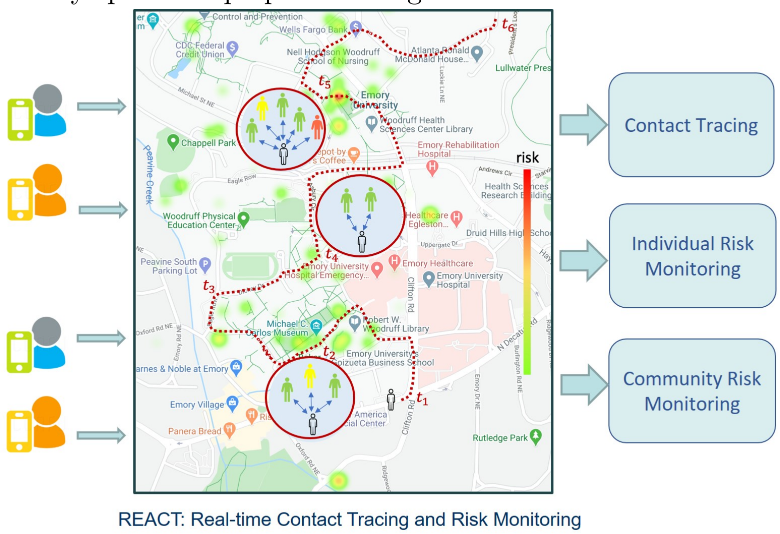

Figure 2: REACT Overview

In this article, we will briefly describe our ongoing work and the approaches we are taking including: 1) building efficient, scalable data structures and algorithms for contact tracing queries; 2) expanding a learning-based approach for modeling user's risks based on location risk factors, propagated risks from other users, and the user's self-reported risk factors such as symptoms, demographic data, existing conditions, and travel history; and 3) enabling a multi-stage privacy approach based on geo-indistinguishability and its variants [4, 42]. We also describe our progress to date in developing and deploying the app as we originally planned at our three collaborating institutions. The release of the app requires intensive security and human subject research reviews that are much more involved than what we had originally anticipated. In addition, we have encountered other nontechnical obstacles that we will describe in our case study at USC.

2 Approach

The main goal of our project is to enable contact tracing and risk monitoring via tracking of users' locations and symptoms. Upon user' consent, locations can be automatically uploaded at a user-selected frequency and granularity. Selected personal data will be collected on a voluntary basis including demographic information and existing conditions. In-app surveys will be used to collect user symptom data periodically or triggered by time/location following the Ecological Momentary Assessment methodology [34, 19]. We plan to ask users daily if they experience of any of the symptoms of upper respiratory illness: fever, chills, muscle aches, cough, congestion, runny nose, headaches, fatigue, and shortness of breath, as well as if they have a confirmed infection of COVID-19. Given the purpose of the app, we expect even if there are self-reported false positives, acting in an abundance of caution would benefit the community.

2.1 Contact Tracing

We envision that whenever a user self-reports onset of COVID-19 symptoms or a confirmed positive test, we can systematically identify all users who have been in contact with this infected case both directly and indirectly. Once identified, we can alert other users and update their risks. We discuss queries to identify contacts in this subsection and risk notification and modeling in next subsection.

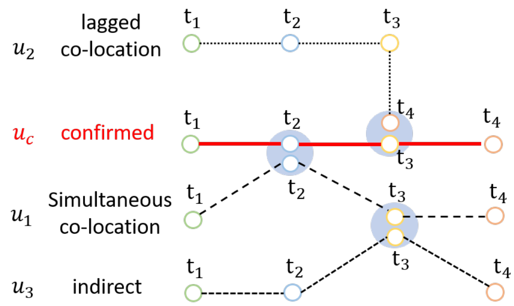

Figure 3: Transmission Scenarios

We consider three types of transmission of COVID-19: 1) direct person-to-person transmission, i.e. in close contact with someone infected (simultaneous co-location); 2) fomite transmission, i.e. in contact with a contaminated surface or object at a location visited by someone infected earlier (lagged co-location) [25]; and 3) indirect person-to-person transmission by contact with someone who is earlier in contact with someone infected. Figure 3 illustrates four user trajectories where u_c has a confirmed infection at time t_4 and the three transmission scenarios: 1) u_1 via simultaneous co-location at t_2, 2) u_2 via lagged co-location at time t_4, and 3) u_3 via direct transmission from u_1 at t_3 (indirect transmission from u_c).

We will first consider the first two scenarios which are the primary ways of transmission. Assume the location trace of each user is represented as a sequence of *visits* $< u, s, t >$ (user u at location s at time t). Given a user u_c with confirmed infection, we need to find all users who have simultaneous or lagged co-location with u_c in a back tracing window (this can be parameterized, e.g. as 5 days which is the currently known median incubation period for COVID-19 [35]). Defining a sequence of visits by a single user as a *trajectory*, this problem can be formulated as composition of *trajectory range queries* [43] which given a trajectory dataset, a spatial range and a time interval, returns all trajectories that intersect with both the spatial range and the time interval. For each location s visited by u_c at time t in the time window, we can compute trajectory range query using a spatial range of 6 feet centered at s and a time interval of up to 72 hours starting at t (the viable time of COVID-19 virus on surfaces [40]). These parameters can be changed with dynamic evidence such as the distance within closed/open spaces or modifications in aerosol range according to latest research and CDC recommendations. Consider the example in Figure 3, the query given u_c will return u_1 and u_2. There are many structures that index trajectories to efficiently answer such queries, but since we mainly work with visits, we can simply index visits using point-based indexes. Given the broad availability of R-trees [18] (e.g., with PostgreSQL) and the need to have our system up and running quickly, we will use an R-tree to index the visits on the three dimensions: latitude, longitude and time.

The indirect transmission is more challenging both from modeling and computational point of view. We need to find all users that are directly and indirectly in contact with the confirmed case such as u_3 in our example. This can be formulated as a *spatio-temporal reachability query* [37]. A straightforward approach is to first run the trajectory range queries using the confirmed case as source, then run the queries recursively by using the returned trajectories as sources. This will be computationally expensive, especially when the number of indirect transmission hops becomes more than a few. We plan to explore an alternative approach by leveraging our prior work [37] which proposed efficient grid and graph based indexes for answering "single source single destination" reachability queries. The main idea is to compute reachability on-the-fly by expanding the contact network starting from the query source and utilizing the spatio-temporal locality for enhanced performance.

2.2 Individual Risk Monitoring

While contact tracing is triggered when there is a confirmed infection, we envision REACT will also allow users to monitor their own risks in real time based on the locations they recently visited and the aggregated risks of other users they have come in contact with. Whenever a user visits a new location, we will use our risk model to update the risks for other users, so that they can be informed/alerted and take preventative measures when necessary. Specifically, we can define a risk score between 0 and 1 for each user u which represents the probability u will contract the virus and 1 means the user has a confirmed infection. We can gradually train a risk model (e.g. a logistic regression model) based on the risk factors as we collect more data including confirmed infections (which serves as ground truth).

The following factors can be considered in the risk model for each user u: 1) u's risk profile including demographic data and existing conditions, 2) aggregated risk of locations u has recently visited, and 3) aggregated risk of recent contacts with other users. Risk of a location can be dependent on type of location or Point of Interest (e.g. from Google Map API) and confirmed cases in the area [1]. Risk associated with each contact can be dependent on distance and duration. Due to infrequent or generalized location tracking for privacy enhancement, distance may not be accurate enough nor can duration of co-location be captured adequately. Hence, we plan to incorporate the strength of social relationships between two users as an additional factor. Duration of social contact is typically longer, suggesting a higher risk, for dyads with social ties rather than dyads of strangers [28, 6]. This social relationship can be partially explicitly collected from the users (see the subsection below) or implicitly inferred from their historical trajectories as we have demonstrated in our prior work [31]. The intuition is that if two people have frequent co-locations, especially at not-popular places, it is likely they are socially related.

In order to evaluate our risk monitoring algorithms, we need to have realistic data before we collect real data from the deployed app. We are developing an agent-based spread simulator based on a real mobility dataset[10] to generate test data. Most existing simulations are based on compartmental models which cannot account for real life mobility patterns. We are using real-world mobility patterns to inform disease spread and creating a simulator that can generate realistic spread data.

To complement our algorithm development for contact tracing queries and risk monitoring, we are also designing and developing a user facing dashboard (at USC) to enable public health practitioners to expedite contact tracing processing and provide recommendations based on risk parameters estimated from available co-occurrences and user data. The dashboard will enable decision makers to visualize and optimize interventions as powered by the Spread Simulator.

2.3 Community Risk Monitoring

To monitor the community, we will use the social network sensors approach [8]. Using a novel study design based on properties of social networks, the method yielded a 2 week advance signal of the H1N1 influenza outbreak in fall 2019 among Harvard undergraduates. In our project, we will invite a random sample of students at each of our three institutions to participate as a baseline (Random Group). We will ask the Random Group for contact information for 2-3 of their friends. Subsequently, we will invite the friends to participate (Friends Group). The value of the social network is that, in an epidemic, the more central members of a social network become infected earlier. Thus utility of the app as used with this study design is that not only will we have an estimation of risk, but we also expect a similar early signal of an outbreak by examining over time the differences between the Random Group and the Friends Group with its more central members [9, 28].

[10] the anonymized raw mobility dataset is provided to us by Veraset.

Specifically, we will track the numbers of users in each group who have a constellation of symptoms consistent with COVID-19 (fever, dry cough, shortness of breath). Without loss of generality we can similarly track symptoms consistent with influenza, providing generalized utility for the annual seasonal influenza outbreak beyond the current pandemic. We will use an adaptation of the cumulative sum (CUSUM) procedure to detect the first separation of the Friends Group from the Random Group with respect to prevalence of COVID-19 symptoms [20]. In particular, we will use the CUSUM count method to determine the separation time [14, 15]. Briefly, we will consider at the end of each day, t, the counts of the number of users in the Random and Friends groups reporting symptoms consistent with COVID-19, Y_t^R and Y_t^F respectively. Assuming that these counts are (conditionally) independent Poisson variables with mean μ_t^R and μ_t^F, the CUSUM method—based on the likelihood ratios of functions of sums of Y_t^R and Y_t^F as t increases—uses sequential hypothesis testing to determine the change point from $\mu_t^F = \mu_t^R$ to $\mu_t^F > \mu_t^R$ which suggests an imminent community outbreak.

2.4 Privacy Enhancements

Despite the utility of contact tracing, the technology also raised a lot of trust concerns [32], including cultural and behavioral issues [26], privacy and equity [5], legal issues [10], and individual autonomy, privacy, confidentiality, and social justice [36]. Medical professionals are facing ethical choices between the public good and individual's privacy [24]. Besides these intensely debated issues [23], data protection and user acceptability remain as major barriers [2].

To mitigate privacy risks while ensuring immediate public health impact, the key is to give users the options to control frequency and precision with which information will be collected. For instance, users will be able to choose and update frequency of tracking (or a manual check-in option) and the granularity of tracked locations (e.g., a generalized location range) as their risk evolves and to choose whether to report symptoms or not.

Given infrequent tracking or generalized locations, results of contact tracing queries may not be precise. We plan to have a multi-stage approach to address this challenge. In stage 1 (global computation), the server can perform "single source all destination" contact tracing queries to identify all "possible contacts" over the generalized or imprecise locations. This may lead to false positives and false negatives. We can adjust or relax both spatial range and temporal interval to account for generalized locations and infrequent tracking and ensure a low false negative rate. Possible contacts can then be alerted. In stage 2 (local refinement), alerted users can choose to upload his/her precise location trace stored locally in the recent window to confirm contact status with the confirmed case. The server then will perform a "single source single destination" query to get precise result. In our prior work [39], we have shown that such a multi-stage approach is promising for task assignment given uncertain/perturbed locations of workers and tasks for spatial crowdsourcing.

Currently, we have adopted and evaluated the geo-indistinguishability (GeoInd) privacy definition [4, 42, 39, 17] to enable users to protect their locations. Given app users u_1 (and u_2, respectively), the ε-GeoInd perturbation mechanism distorts their exact locations l_1 (l_2) to l_1' (l_2') by adding a spatial noise vector derived from a 2D Laplace distribution (with scale inversely proportional to ε). The challenge is then to accurately compute the range or reachability queries over the perturbed locations and to address the privacy risks associated with a location trace—a straightforward composition of geo-indistinguishability will render either the privacy or the utility not acceptable.

We extend probabilistic techniques from our previous work in [39] to calculate the range query over the pair of perturbed locations l_1' and l_2'. Recall that the range query captures whether or not two users actually made a contact (parameterized as a reachable distance R), which is indicative of the risk of a potential transmission. The objective is to then calculate their reachability probability $p(d \leq R|d')$, where d and d' are the Euclidean distances of their exact and perturbed location pairs, respectively.

Our preliminary studies using the Gowalla Geo-social Network checkin dataset [7] verifies that the probabilistic approach can outperform the baseline oblivious approach (which determines reachability using the perturbed locations directly), and it can achieve 80% precision and recall given a reasonable privacy level.

2.5 App Development

The current REACT[11] app is *forked* from an existing open source project named Covid Community Alert[12]. The REACT app collects proximal contacts (via Bluetooth) and locations (via GPS if permissed by user) for contact tracing. It maintains the anonymity of its users by recording ephemeral device IDs that persist for the duration the app is installed and can be reset by user by re-installing the app.

We extended the app with additional location privacy features. A UI page requests the user to input a desired privacy level (between low, medium and high) for sharing his/her locations. This privacy level is interpreted as the level of perturbation that is applied to user's location before it is transmitted to the receiving server. We implemented the GeoInd based location perturbation with predefined privacy levels. Another UI element provides the functionality for the users to self-report their COVID status (e.g., from symptomatic, tested positive/negative, recovered). The app works as follows. A user registers the device by sending a randomly generated device ID to the server when first time use the app (no personal information collected). The app keeps scanning surrounding Bluetooth signals and collects the IDs of the nearby devices. The interaction information including devices IDs, timestamp, interaction duration and GPS location are sent to the back-end server. When a confirmed case is reported, the back-end server finds the potential contacts and estimates their total risk score. If the risk score exceeds a preset level an alert is relayed to these users as a notification on their device.

3 Deployment: Challenges and Lessons

Our original plan was to develop and release a mobile app at our three collaborating institutions. At Emory University, at the time of writing, we are still under security review by the university IT office and IRB review for the human subject research, which will not proceed without satisfactory security review. At UT Health, the IRB review has been approved. At University of Southern California (USC), earlier versions of the app had gone through the security reviews and hence efforts were made to deploy it for real use. In this section, we discuss some of the practical and non-technical challenges we encountered in releasing the app for real use at USC.

As early as mid-May 2020, Shahabi's team at USC developed multiple contact-tracing app prototypes using a variation of techniques to detect and collect user location data. The very first app used the phone's location API to store mobility patterns. The later versions included a QR-code scanning capability to allow for scanning of location QR-codes as a way for the user to voluntarily check in and out of a location (e.g., a classroom). The final version of the app, developed on May 22nd 2020, also added bluetooth for proximity tracking. All these data collections had opt-out options. The goal was to modify the app later per our proposed ideas in the NSF RAPID project to collect location data at different spatial and temporal resolutions determined by the user and only store the detailed data on the user's device. Our first goal, however, was to influence one of our institutions, in this case USC (through Shahabi's participation in USC's contact-tracing subcommittee) to actually adopt and use

[11] `https://github.com/Emory-AIMS/react`

[12] https://coronavirus-outbreak-control.github.io/web/

the app for faculty, staff and students as they come back to campus. Unfortunately, these efforts were not successful for non-technical reasons. We review these reasons below.

Both Google and Apple restricted the release of any app on their app stores that was related to COVID-19 in general and for contact-tracing in particular. They required the backing of a health organization for COVID-19 apps. With this requirement, we could not release our app, even to be used by volunteers under IRB, unless we have the backing of our respective organizations. At USC, we tried to release different variations of our app, went through rigorous reviews of our IT offices for security and privacy, and at the end none of the variations (except for an early adaptation of a symptom collection app at USC, called SCORE) made it to the app store. Basically, the time cost of releasing an app was so high that it overshadowed the cost of developing the app itself. More importantly, even if we released the app, as discussed below, we did not have the backing of our institute to recruit students to use the app, which in turn rendered the releasing of the app useless.

The main hesitation of the organizations to support and release a contact-tracing app was to protect the users' location privacy. This was a surprise to us for three reasons. First, many apps already freely collect users' locations. Second, there is a decade of research on location privacy by our community that can be incorporated, starting from simple measures of allowing users to control the specificity of their reported locations (for instance, building level or shopping-center level), the frequency of reporting (for example, once or twice a day) and to remove sensitive locations. More sophisticated measures and technologies can also be incorporated, such as storing and searching all data in an encrypted form, similar to storing passwords or banking information. Both approaches have been studied extensively for location data in the past decade, e.g., [3] and [16]. Third, in case the integration of data about one's health status (in this case COVID-19 exposure) and location data was sensitive, we considered approaches to separate the location data from health data, each being stored and accessed by trusted parties within an organization. This is usually the first step in any privacy and security research, where threat models are clearly defined. However, unfortunately, we never proceeded sufficiently far to clearly define threat models. The main obstacle was the "perception" of privacy violation surrounding any contact-tracing app. We could not solve this concern of perception with technical solutions, and instead partnerships with colleagues from communication and journalism are needed to design and deploy broader messaging campaigns.

Finally, we tried to convince the health offices within our organization about the usefulness of digital contact tracing to get their support in convincing our institute to support the release of a contact-tracing app. Towards this end, working through the USC's contact-tracing subcommittee, we showcased numerous proof-of-concept tools, demonstrations and presentations, directly to USC's health practitioners and managers who were in charge of contact-tracing on campus to demonstrate what could have been done if location data were collected. The utilities included user-friendly dashboards to quickly find overlapping trajectories with the trajectory of an infected case, identifying hotspot locations (through density mapping of infected trajectories) and utilities to detect environmental and indirect contacts. However, at the end, the health organizations preferred traditional contact tracing approaches where individuals were interviewed thoroughly and then broad notifications were sent to anyone who could have potentially been collocated with the positive cases. Clearly the approach is not scalable but due to limited access of students and faculty to campuses (due to remote teaching), the issue of scalability has not been the main concern of health practitioners as they were dealing with other critical and time-sensitive issues.

Consequently, even though we designed and developed several useful solutions, we could not convince the decision makers at our institution to support the utilization of our tools for us to make a real impact in our community.

4 Conclusion

In this article, we presented our ongoing project for real-time contact tracing and risk monitoring via privacy-enhanced mobile tracking. We described the approaches we are taking for privacy enhanced contact tracing and risk monitoring, the preliminary results we have obtained which are encouraging, as well as some challenges we encountered in deploying the app. While there are continued privacy concerns, and other non-technical obstacles, we believe digital contact tracing remains an important component of the public health response and it requires careful designs and technical developments which we will continue to work on in order to ensure privacy protection and public health benefits.

References

[1] Coronavirus covid-19 global cases by the center for systems science and engineering (csse) at johns hopkins university (jhu). https://coronavirus.jhu.edu/map.html.

[2] J. Abeler, M. Bäcker, U. Buermeyer, and H. Zillessen. COVID-19 contact tracing and data protection can go together. *JMIR mHealth and uHealth*, 8(4):e19359, Apr. 2020.

[3] R. Ahuja, G. Ghinita, and C. Shahabi. A utility-preserving and scalable technique for protecting location data with geo-indistinguishability. In *Advances in Database Technology - 22nd International Conference on Extending Database Technology, EDBT 2019, Lisbon, Portugal, March 26-29, 2019*, pages 217–228. OpenProceedings.org, 2019.

[4] M. E. Andrés, N. E. Bordenabe, and K. Chatzikokolakis. Geo-indistinguishability: Differential privacy for location-based systems. *CoRR*, abs/1212.1984, 2012.

[5] I. Braithwaite, T. Callender, M. Bullock, and R. W. Aldridge. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. *The Lancet. Digital health*, Aug. 2020.

[6] S. Cauchemez, C. A. Donnelly, C. Reed, A. C. Ghani, C. Fraser, C. K. Kent, L. Finelli, and N. M. Ferguson. Household transmission of 2009 pandemic influenza a (h1n1) virus in the united states. *New England Journal of Medicine*, 361(27):2619–2627, 2009.

[7] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In C. Apte, J. Ghosh, and P. Smyth, editors, *International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pages 1082–1090. ACM, ACM, 2011.

[8] N. Christakis and J. Fowler. Social network sensors for early detection of contagious outbreaks. *PloS one*, 5(9), 2010.

[9] R. M. Christley, G. Pinchbeck, R. G. Bowers, D. Clancy, N. P. French, R. Bennett, and J. Turner. Infection in social networks: using network analysis to identify high-risk individuals. *American journal of epidemiology*, 162(10):1024–1031, 2005.

[10] A. Cioffi, C. Lugi, and C. Cecannecchia. Apps for COVID-19 contact-tracing: Too many questions and few answers. *Ethics, medicine, and public health*, 15:100575, Oct. 2020.

[11] E. Clark, E. Y. Chiao, and E. S. Amirian. Why contact tracing efforts have failed to curb coronavirus disease 2019 (covid-19) transmission in much of the united states. *Clinical Infectious Diseases*, 2020. https://doi.org/10.1093/cid/ciaa1155.

[12] P. C. Erwin and R. C. Brownson. *Principles of Public Health Practice*. Cengage Learning, 2016.

[13] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser. Quantifying sars-cov-2 transmission suggests epidemic control with digital contact tracing. *Science*, 368(6491), 2020.

[14] M. Frisén. Statistical surveillance. optimality and methods. *International Statistical Review*, 71(2):403–434, 2003.

[15] M. Frisén and P. Wessman. Evaluations of likelihood ratio methods for surveillance. *Communications in Statistics-Simulation and Computation*, 28(3):597–622, 1999.

[16] G. Ghinita, K. Nguyen, M. Maruseac, and C. Shahabi. A secure location-based alert system with tunable privacy-performance trade-off. *GeoInformatica*, 24(4):951–985, 2020.

[17] X. Gu, M. Li, Y. Cao, and L. Xiong. Supporting both range queries and frequency estimation with local differential privacy. In *2019 IEEE Conference on Communications and Network Security (CNS)*, pages 124–132, 2019.

[18] A. Guttman. R-trees: A dynamic index structure for spatial searching. In *Proceedings of the 1984 ACM SIGMOD international conference on Management of data*, pages 47–57, 1984.

[19] M. J, M. M, C. DL, and H. K. Geographic imputation of missing activity space data from ecological momentary assessment (ema) gps positions. 15, 2018.

[20] W. Jiang, L. Shu, H. Zhao, and K.-L. Tsui. Cusum procedures for health care surveillance. *Quality and Reliability Engineering International*, 29(6):883–897, 2013.

[21] M. Jonker, E. De Bekker-Grob, J. Veldwijk, L. Goossens, S. Bour, and M. Rutten-Van Mölken. COVID-19 contact-tracing apps: predicted uptake in the netherlands based on a discrete choice experiment. *JMIR mHealth and uHealth*, Aug. 2020.

[22] J. Jung, H. Jang, H. K. Kim, J. Kim, A. Kim, and K. P. Ko. The importance of mandatory COVID-19 diagnostic testing prior to release from quarantine. *Journal of Korean medical science*, 35(34):e314, Aug. 2020.

[23] K. Kaspar. Motivations for social distancing and app use as complementary measures to combat the COVID-19 pandemic: Quantitative survey study. *Journal of medical Internet research*, 22(8):e21613, Aug. 2020.

[24] J. Klaaren, K. Breckenridge, F. Cachalia, S. Fonn, and M. Veller. South africa's COVID-19 tracing database: Risks and rewards of which doctors should be aware. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde*, 110(7):617–620, June 2020.

[25] A. N. Kraay, M. A. Hayashi, N. Hernandez-Ceron, I. H. Spicknall, M. C. Eisenberg, R. Meza, and J. N. Eisenberg. Fomite-mediated transmission as a sufficient pathway: a comparative analysis across three viral pathogens. *BMC infectious diseases*, 18(1):540, 2018.

[26] F. Lucivero, N. Hallowell, S. Johnson, B. Prainsack, G. Samuel, and T. Sharon. COVID-19 and contact tracing apps: Ethical challenges for a social experiment on a global scale. *Journal of bioethical inquiry*, Aug. 2020.

[27] C. M. Manauis, M. Loh, J. Kwan, J. Chua Mingzhou, H. J. Teo, D. Teng Kuan Peng, S. Vasoo Sushilan, Y. S. Leo, and A. Hou. Bracing for impact: operational upshots from the national centre for infectious diseases screening centre (singapore) during the COVID-19 outbreak. *Journal of the American College of Emergency Physicians open*, June 2020.

[28] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*, 5(3), 2008.

[29] K. Nagarajan, M. Muniyandi, B. Palani, and S. Sellappan. Social network analysis methods for exploring SARS-CoV-2 contact tracing data. *BMC medical research methodology*, 20(1):233, 2020.

[30] J. Peto, J. Carpenter, G. D. Smith, S. Duffy, R. Houlston, D. J. Hunter, K. McPherson, N. Pearce, P. Romer, P. Sasieni, and C. Turnbull. Weekly COVID-19 testing with household quarantine and contact tracing is feasible and would probably end the epidemic. *Royal Society open science*, 7(6):200915, June 2020.

[31] H. Pham, C. Shahabi, and Y. Liu. EBM: an entropy-based model to infer social strength from spatiotemporal data. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 265–276. ACM, 2013.

[32] B. J. Ryan, D. Coppola, J. Williams, and R. Swienton. COVID-19 contact tracing solutions for mass gatherings. *Disaster medicine and public health preparedness*, pages 1–7, July 2020.

[33] A. S, K. OF, and A. KH. Covid-19 contact-tracing technology: Acceptability and ethical issues of use. *Patient Prefer Adherence*, 2020. https://doi.org/10.2147/PPA.S276183.

[34] S. S, S. AA, and H. MR. Ecological momentary assessment. 4, 2007.

[35] L. SA, G. KH, and e. a. Bi Q. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: Estimation and application. 2020.

[36] A. Shah and R. P. Aacharya. Combating COVID-19 pandemic in nepal: Ethical challenges in an outbreak. *JNMA; journal of the Nepal Medical Association*, 58(224):276–279, Apr. 2020.

[37] H. Shirani-Mehr, F. B. Kashani, and C. Shahabi. Efficient reachability query evaluation in large spatiotemporal contact datasets. *PVLDB*, 5(9):848–859, 2012.

[38] C. Thorbecke. China launches coronavirus tracking app as death toll surpasses 1,000. ABC News, Feb. 2020. https://abcnews.go.com/Business/china-launches-app-combat-coronavirus-spread/story?id=68907706.

[39] H. To, C. Shahabi, and L. Xiong. Privacy-preserving online task assignment in spatial crowdsourcing with untrusted server. In *IEEE Int. Conference on Data Engineering (ICDE)*, 2018.

[40] van Doremalen N, M. DH, and e. a. Holbrook MG. Aerosol and surface stability of sars-cov-2 as compared with sars-cov-1. March 19, 2020.

[41] M. Walrave, C. Waeterloos, and K. Ponnet. Adoption of a contact tracing app for containing COVID-19: A health belief model approach. *JMIR public health and surveillance*, 6(3):e20572, Sept. 2020.

[42] Y. Xiao and L. Xiong. Protecting locations with differential privacy under temporal correlations. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, October 12-6, 2015*, pages 1298–1309, 2015.

[43] M.-E. Yadamjav, F. Choudhury, Z. Bao, and H. Samet. *Efficient Multi-range Query Processing on Trajectories: 37th International Conference, ER 2018, Xi'an, China, October 22–25, 2018, Proceedings*, pages 269–285. 09 2018.