Smartphone Privacy in Citizen Science

Hannah M. Roth

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Masters of Science

in

Computer Science and Applications

Danfeng (Daphne) Yao, Chair

Kurt Luther

Gang Wang

May 8, 2017

Blacksburg, Virginia

Keywords: Citizen Science, Security, Android Mobile Software, Crowdsourcing, Crowdsensing, Groupsensing, Privacy Guarantee, Group Signature Scheme, User Study

Smartphone Privacy in Citizen Science

Hannah M. Roth

ABSTRACT

Group signature schemes enable anonymous-yet-accountable communications. Such a capability is extremely useful for modern applications such as smartphone-based crowdsensing and citizen science. A prototype named GROUPSENSE was developed to support anonymous-yet-accountable crowdsensing with SRBE in Android devices. From this prototype, an Android crowdsensing application was implemented to support privacy in citizen science. In this thesis, we will evaluate the usability of our privacy-preserving crowdsensing application for citizen science projects. An in person user study with 22 participants has been performed showing that participants understood the importance of privacy in citizen science and were willing to install privacy-enhancing applications, yet over half of the participants did not understand the privacy guarantee. Based on these results, modifications to the crowdsensing application have been made with the goal of improving the participants' understanding of the privacy guarantee.

Smartphone Privacy in Citizen Science

Hannah M. Roth

GENERAL AUDIENCE ABSTRACT

A group signature scheme is a security solution that allows any member of a group to create a digital signature without revealing his or her identity. This enables an application user to remain anonymous-yet-accountable during communication. Such a capability is extremely useful when collecting data for scientific research, referred to as citizen science, through a modern smartphone application. A prototype named GROUPSENSE was developed to support anonymous-yet-accountable data collection with SRBE, an advanced group signature scheme, in Android devices. From this prototype, an Android application was implemented to support privacy in citizen science. In this thesis, we will evaluate the usability of our privacy-preserving application developed for citizen science projects. An in person user study with 22 participants has been performed showing that participants understood the importance of privacy in citizen science and were willing to install privacy-enhancing applications, yet over half of the participants did not understand the specified privacy guarantee. Based on these results, modifications to the application have been made with the goal of improving the participants' understanding of the privacy guarantee.

Acknowledgments

Thank you Dr. Danfeng (Daphne) Yao for providing me with academic guidance and opportunities throughout my years of research and study at Virginia Tech. I would also like to acknowledge members of the Yao group for their continued support and answering any questions I had along the way. I am also grateful to Dr. Kurt Luther and Dr. Gang Wang for taking the time to serve on my evaluation committee.

In addition, I am grateful to the professors who provided the knowledge essential to the completion of my degree and the colleagues who collaborated with me to study for exams and complete group assignments. I would also like to acknowledge the Cybercorps Scholarship for Service program at Virginia Tech that has allowed me to continue my education and pursue a Master's Degree.

Finally, I would like to thank my parents for their unconditional support, whether providing advice or just always answering the phone. Also, thank you to my friends for the encouragement I received throughout my undergraduate and graduate degree programs. Without their constant support, obtaining a higher degree would not have been possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In order to have a clear understanding of this paper, a few key terms are defined as follows:

1. Crowdsourcing: the process of eliciting data collection from the general public.

2. Crowdsensing: the process of eliciting data collection from devices or sensors.

3. Citizen science: information for scientific research gathered by means of crowdsourcing or crowdsensing.

4. Group signature scheme: a type of cryptographic solution that allows any group member to create a digital signature (i.e. digitally signing a message) without revealing his or her identity.

5. Groupsensing: controlled crowdsensing scenario where only participants with proper authorization can contribute to the campaign.

6. Privacy guarantee: sensitive personal information (e.g. demographic or geographic descriptors) of a participant contributing to citizen science should be protected from (i.e. not exposed to) public or data servers.

Regardless of the project scale, privacy concerns are likely to moderate the acceptance and viability of citizen science. Privacy becomes important in citizen science because participants must believe they are protected against privacy threats and vulnerabilities. These threats could include a semi-honest data collector that attempts to track and identify participants under false pretenses; or perhaps a data breach to the data collector itself. Group signature schemes enable anonymous-yet-accountable communications. Such a capability is extremely useful for securing modern applications, such as smartphone-based crowdsensing and citizen science. However, revocation checking in most deterministic group signatures is linear to the size of the revocation list, significantly affecting their scalability. A cryptographic solution was developed that allows sensory data to be sent anonymously [61]. The solution is an innovative group signature scheme that supports sublinear revocation. It drastically improves the usability of the group signature primitive in real-world crowdsensing settings.

In addition to that rigorous cryptographic contribution, a groupsensing prototype named GROUPSENSE was developed. Accountability requires identity and privacy demands anonymity. To satisfy both requirements, it is desirable that only the participants with proper authorization can contribute in a crowdsensing campaign. This controlled crowdsensing scenario is referred to as groupsensing. An application that implements this GROUPSENSE prototype - GroupSensing - was developed. This prototype supports anonymous-yet-accountable crowdsensing in Android devices. It deploys cryptographic schemes to shield participants' information from exposure to data collectors. Multidisciplinary crowdsensing and citizen science projects require secure and privacy-preserving infrastructures. Secure crowdsensing encourages participation, which promotes an improved quality of data resources and increased discovery.

A user study was designed with the goal of evaluating:

- The usability of our privacy-preserving crowdsensing application for citizen science projects.

- Whether participants can reasonably understand the privacy guarantees of group signatures.
- Whether the group signatures would alleviate the privacy concern inherent among citizen science participants.

The user study was conducted and the results are detailed later on in this paper. The results from the study were a catalyst for much of the modifications to the application, in order to align the user experience with the above-mentioned goals.

1.1 Technical Challenges

1.1.1 Privacy Paradox: Quality of Service vs. Privacy

In general, people have a high regard for the sanctity of preserving their privacy, yet the mass population paradoxically actually favors quality of service and enhanced functionality over privacy [8]. Privacy preserving infrastructure affects the Quality of Service (QoS), which is a critical barrier to the effective deployment of privacy preserving infrastructure. For example, data obfuscation through adding noise or controlling granularity affects data integrity and reduces the service quality in return. Mixing traffics to destroy spatio-temporal correlations increases network latency. Performance enhancement techniques (e.g., caching), based on users' prior history, cannot be optimized while maintaining strict anonymity principles.

1.1.2 Why Privacy Is Important: Creating Mass Awareness

There is no simple answer to the direct question of "Why privacy is important?" [60]. Therefore, it is difficult for the mass population to attribute "proper" importance to the privacy in their daily life. As a result, people are likely to trade off their long-term privacy for short-term benefits [3], revealing the importance of mass awareness by educating the public on the importance of privacy.

1.2 Research Contributions

By presenting my work, this thesis makes the following contributions:

1. Insight into perceptions of privacy guarantees understood by participants of citizen science projects.
2. Results from a user study to support these participant perceptions.
3. An advanced prototype to be used for crowdsensing citizen science projects.

1.3 Thesis Layout

The remainder of this thesis is as follows. First, Chapter 2 provides background into various aspects of my research. Then, Chapter 3 details the user study conducted. Next, Chapter 4 explains the Android application and the improvements that were made based on the user study. Finally, the conclusion and opportunities for future work are presented.

Chapter 2

Background

2.1 Crowdsensing Platform

The new urban-scale crowdsensing vision promises useful applications, such as health monitoring [53], environment monitoring [40], and traffic prediction [58]. However, an open crowdsensing platform where anyone can submit data is undesirable as a portal for transferring sensitive data. An open platform that is unregulated exposes itself to malicious and erroneous participation. This constitutes a vulnerability which threatens privacy, data integrity, and reliability standards [56]. Accountability of participants for their data reports is a key requirement for crowdsensing platforms [68].

While accountability protects the data collector, the vast number of crowdsensing participants need to be protected against privacy threats and vulnerabilities [31], [42]. Examples of such threats to crowdsensing transactions include the semi-honest data-collection service provider, who attempts to track and identify participants, as well as data breaches on the data-collection servers. The adoption of sensing-time anonymity is an essential security requirement, especially for those participants who are involved in long-term crowdsensing campaigns.

The deployment of anonymous-yet-accountable crowdsensing controls appears technically challenging. Fortunately, cryptographic solutions for this paradoxical requirement type already exist. Although cryptographic solutions are not specifically designed for this application, custom-fitting for crowdsensing is a conceivable security control. These protocols can be broadly categorized into two groups: (1) pseudonym-based systems [10], [26] and (2) group signature-based systems [27], [15]. Both groups rely on having a trusted group manager to coordinate transactions between the signer (crowdsensing participant) and the verifier (semi-honest data-collection server).

In pseudonym-based systems (e.g., [63]), a participant would generate signatures using pseudonyms and refresh them periodically to preserve anonymity. The group manager, in this practice, must maintain a list of pseudonyms and public-key certificates to certify the public keys of participants (for accountability). However, frequent public-key certification and distribution are expensive to maintain under short-lived pseudonyms [51].

Group signature schemes (e.g., [27], [15], [9]), in comparison to pseudonym-based systems, allow signers to anonymously produce signatures. Maintaining this practice guarantees full anonymity without the requirement of frequent public-key certifications for the signers. This allows for a single public key to be used by all signers in the group. The complexity of revocation checking in modern group signature schemes is typically O(N), where N is the size of the revocation list. The list is maintained locally by the verifier. To check the revocation status of received signatures deterministically, a verifier must check whether any of the revocation tokens in the list can be mapped to the received signature. The use of efficient data structures to optimize the search is not feasible because signatures carry zero-knowledge identity information to ensure strong unlinkability. Consequently, this disconnect does not allow opportunity for straightforward comparison. As a result, sublinear revocation for group signature schemes has been a chronic open research problem [15]. This problem has been exacerbated by the negative consequences for group signature schemes based on blind signatures and anonymous credentials [19].

Whereas group signatures hold the promise to realize anonymous-yet-accountable crowdsensing with minimum management overhead, the methodology in which existing solutions scale up to real-time large crowdsensing campaigns (e.g., millions of participants as envisioned in [45]) is yet to be determined. If a group signature scheme is used to anonymously sign data to share with the data-collection server, deterministic revocation checking emerges as an unavoidable performance bottleneck. This negatively affects server side operations as revocation checking is executed with each signature check. More notably, the legitimate signatures face the worst-case complexity. Therefore, the allure of group signatures (e.g., [15], [13], [11], [18]) in crowdsensing is substantially dampened by the expensive revocation checking. For example, SPPEAR [38], a new comprehensive crowdsensing system, avoids using group signatures for sensory data submission. Instead, SPPEAR only uses group signatures for setting up pseudonyms, but resorts to the public-key certification approach for data. Such an approach always incorporates extra public-key certification management overhead (e.g., pseudonym certificate generation, acquisition, distribution, revocation). AnonySense [33] is a privacy-preserving crowdsensing framework that uses group signatures for data submission. However, AnonySense does not support membership revocation, resulting in a low accountability guarantee. There are several other active proposals that exist without accountability support [16], [35], [43].

Using SRBE group signatures, GROUPSENSE, a groupsensing prototype was developed. With GROUPSENSE, a participant anonymously submits signed data reports to the data-collection server - the signature does not reveal her identity but solely verifies membership.

The group manager can access blinded and discrete identifiable information to link compensation or revocation back to the participant.

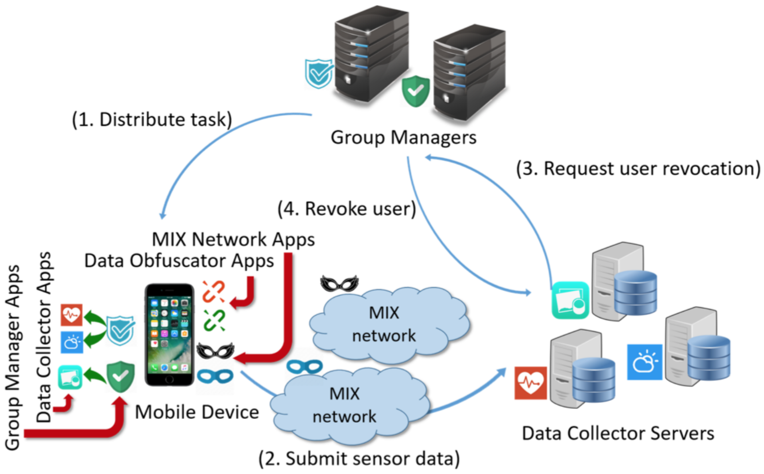


Figure 2.1: Proposed system model for the universal crowdsensing platform showing interactions among different entities during an active crowdsensing campaign.

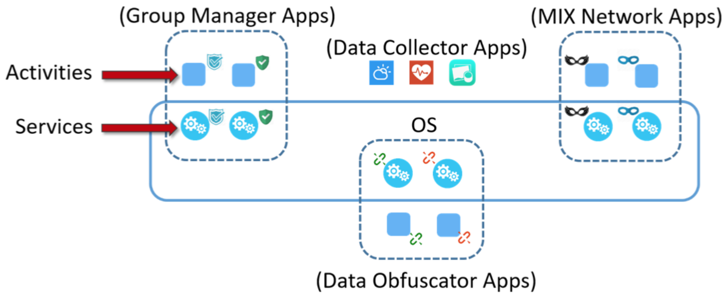


Figure 2.2: Necessary mobile applications from different entities.

2.1.1 Architecture Components

GROUPSENSE is composed of three types of entities: (1) participant's device (PD), (2) data collection server (DCS), and (3) trusted group manager (GM) (i.e., the group manager in SRBE). GROUPSENSE allows participants to anonymously sign and submit sensory data to a curious data collection server by employing a privacy preserving authentication protocol (e.g., pseudonym based signature scheme or group signature scheme), where the group manager takes on the role of the trusted entity. The data collector performs signature verification and revocation checking. However, it is unable to track participants even after the revocation, as data submitted by the same participant are backward unlinkable. The trusted group manager is responsible for checking credentials, revocation management, and possible reward distribution, but even the group manager cannot forge signatures for any participants due to the exculpable property of the SRBE scheme. In Figure 2.1 [62], the major entities of our system model are displayed. Figure 2.2 [62] shows different types of applications from different entities of the ecosystem to ensure anonymous-yet-accountable crowdsensing. The roles of these applications are discussed below. User facing components of an applications are referred to as *Activity* and background processor components as *Service*. The different application types are listed below:

- **Group Managers** have both activities and services. The major responsibilities involve joining the group (activity), obtaining signing parameters (service), signing the sensor data (service) and initiating tasks (activity), and receiving receipts of users' contributions from the data collection server (service). Existing certificate providers can assume the role of group manager.

- **Data Collectors** have different activities and services. They facilitate the initiation sensing tasks (activity) and receive receipts of usersâĂŹ contribution from the data collection server (service). It seems counterintuitive to initiate sensing tasks and receive the receipt of user contribution using the data collector application; however, this data collection model is the most realistic for many practical scenarios, where data collectors want fine-grained control over data quality.

- **Data Obfuscator** have both activities and services. The major responsibilities are to obfuscate sensor data by adding systematic variability [76], [77] or controlling the granularity of the sensed data [4], [24] (service), taking users' preferences to control the granularity (activity). It can also use k-anonymity based measures [42], [70] to obfuscate the sensing patterns of an individual. However, the service that manages communications between "k nearby users" or the trusted server should be independent from the one that provides real-time obfuscation service. The operating system or any third party vendor (other than an existing stakeholder) can play the role of the obfuscator.

- **MIX Networks** have services. The major responsibility of this application is to send data (service) using its network infrastructure. Existing anonymous network providers (e.g., TOR) can play the role of the MIX network.
- **Operating Systems** are referred to as OS (operating system) in Figure 3 to indicate that it has special privileges. This application (provided with OS) is responsible for coordinating the entire sensing ecosystem in the mobile phone. It communicates with different applications provided from different vendors through standard application program interfaces (APIs) to manage anonymous sensing.

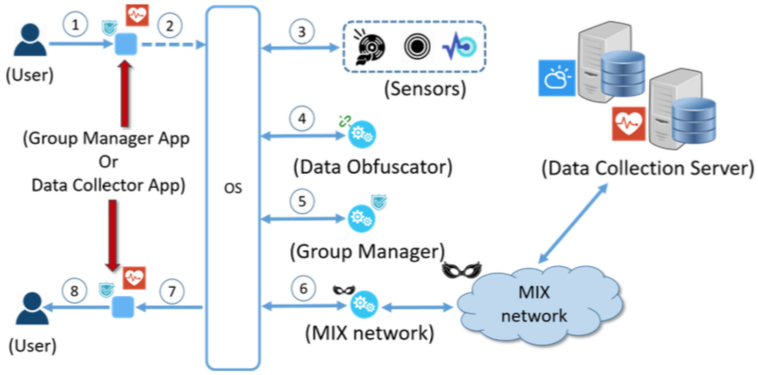


Figure 2.3: Component interactions during data submissions.

2.1.2 Operations

In this section, we define the key operations of our universal crowdsensing platform.

- Initialization and Recruitment - The data collection server (DCS) initiates a crowdsensing campaign by sending a group setup request to the group manager (GM) server who sets up the group. Depending on applications during this phase, the DCS may specify the desired sensing tasks including the target sensors, time period, and geographic area to sense. It may also specify the task budget and incentive scheme [65]. The GM is responsible for advertising the task and recruiting participants attributed to the particular crowdsensing campaign. Interested participants download the GM mobile application and join a data collection campaign. After successfully joining, the GM mobile application obtains signing keys and the information from the DCS. We assume that the communication between the user device and the GM during protocol is secured using end-to-end encryption.

- Task Assignment - In some of the applications, the GM may assign tasks to the users. In that scenario, the GM server sends task specifications to the GM application.

- Data Submission - In Figure 2.3, the data collection procedure is demonstrated. Users can initiate the data collection procedure with the help of two types of applications: (1) group manager application and (2) data collector application (Step 1). After initiation, the initiating application sends an asynchronous request to the OS applications (a new application provided with OS to coordinate the retrieval of sensor data) to collect sensor data (Step 2). The initiating application provides details of the target sensors, associated group manager application, and the location of the remote DCS. Then, the OS application collects sensor data (Step 3) and invokes an associated data obfuscator application (Step 4). Subsequently, the OS invokes the specified group manager application to obtain an anonymous signature (Step 5). After obtaining the signature, the OS encrypts the signature and data with the data collector's public key and then invokes the MIX network application to send the signed data to the DCS (Step 6). Next, the DCS verifies the signature. The DCS is responsible for storing and processing the collected data, including data aggregation and false data detection [29]. After each data submission, the DCS responds with a receipt (signed acknowledgement) targeting the group manager application in the user device. The MIX network forwards the receipt from the DCS to the MIX network application. Then, the MIX network application sends the receipt to the OS application who forwards it to the initiating application (Step 7). The initiating application then shows the notification to the user (Step 8).

- Revocation - After detection of a misbehaving user (e.g., data submitted by the user deviates from the normal pattern), the DCS sends the corresponding signature of that participant to the GM. After receiving the signature, the GM opens it to obtain the identity of the participant. Consequently, the GM sends the revocation token corresponding the user to the DCS.

- Reward Distribution - Metrics for distributing rewards may depend on applications. In general, the GM is responsible for the incentive distribution of the platform. If reward distribution demands the assessment of each participant's contribution, the user's device can send data submission receipts to the GM.

The GM and DCS may have multiple application components. The components capable of initiating sensing tasks and receiving messages from OS applications (steps 1 and 7 in Figure 2.3) do not have the permission to write or modify data in any external resource (file system or network). It is the responsibility of the MIX network to nullify any timingbased side channel attacks. For example, a data collecting adversary can track when it enters the sensing task initiator component and try to estimate the time when the data submission reaches its server to identify users. The MIX network should also obliterate any such patterns.

2.1.3 Related Work

Privacy concerns in participatory sensing were initially pointed out in [64], immediately followed by [41]. After the introduction of the privacy concerns in participatory sensing [65], [28], building anonymous-yet-accountable crowdsensing systems has been the prime focus in crowdsensing security research. Different solutions address different aspects of the problem. In [33], a generic framework for privacy preserving participatory sensing named AnonySense was proposed. AnonySense [33] offers strong privacy protection at the data collection server by decoupling data collection from the participant registration and task assignment modules. AnonySense was one of the earliest works that utilizes group signatures for crowdsensing. However, in AnonySense the accountability guarantee is low, because AnonySense does not support membership revocation. As explained in [37], AnonySense [33] employs group signatures by rendering it vulnerable to Sybil attacks [71]. Gisdakis et al. proposed another crowdsensing framework, named SPPEAR [38], that supports both anonymity and accountability. SPPEAR uses pseudonym based signatures to provide privacy preserving authentication. The authors for this framework argue that custom systems relying on group signatures are vulnerable to abuse. It is impossible to identify signatures from the same participant without having to open the signatures of all data reports. As a result, misbehavior detection becomes a lengthy and inefficient process; also requiring the identification, and therefore deanonymization, of benign reports. GROUPSENSE is able to overcome both of these challenges.

It is conceivable that the inherent openness of privacy preserving systems exposes itself to these vulnerabilities. Hence, the importance to hold malicious users accountable for their contributions is undeniable. In response, SPPEAR [38] and SPPEAR with enhanced incentive provisioning [37] focuses on both anonymity and accountability. In SPPEAR [38], BU-anonymity is achieved through pseudonym-based signature approach. As SPPEAR and GROUPSENSE aim to achieve similar goals, these alternative approaches to achieving both anonymity and accountability are compared:

- Because of using pseudonym based signature schemes to share data, SPPEAR incorporates extra public-key certificate management overhead (e.g., pseudonym certificate generation, acquisition, distribution, revocation), which may affect the scalability and performance of the system. An alternative approach that does not require public-key certificate management is provided.
- The signing delay of SPPEAR using ECDSA is shorter, compared with SRBE. However, with precomputation, the signing delay is comparable to SPPEAR.
- The size of the SRBE signature (3008 bits) is larger than a ECDSA (with SHA-1) signature (440 bits) of same security level. In addition, SPPEAR sends X.509 pseudonym certificates with each signature. Thus, the overall overhead becomes higher in SPPEAR.

2.2 Group Signatures

Recall that we define groupsensing to be a controlled crowdsensing scenario where data submission is restricted to members of an authorized group. Non-members without proper sensing group credentials cannot submit valid data reports. GROUPSENSE is a crowdsensing prototype that supports anonymity and accountability through the SRBE group signature scheme. We show how the SRBE group signature can be applied to realize anonymous-yet-accountable groupsensing. A new VLR-based group signature scheme, sublinear revocation with backward unlinkability and exculpability (SRBE), was presented by Rahaman et al. [61]. SRBE's security is guaranteed under the random oracle model [14]. The main feature of SRBE is that the revocation check complexity is O(log N), where N is the size of the revocation list.

SRBE introduces time bound pseudonyms for the signer to achieve this performance improvement. The use of pseudonyms allows the organization of revoked users in standard data structures such as binary search trees for fast revocation check. The main technical challenge to using these time-constrained and short-lived pseudonyms is to embed them in signatures without compromising security.

SRBE's anonymity is defined in terms of backward unlinkability, which intrinsically entails that even after the revocation of a signer, signatures produced by the signer before revocation remain anonymous. The unlinkability requirement demands pseudonyms to be pseudo-random, void of any correlations in-between. To revoke a signer, it is necessary to exhaustively revoke all the pseudonyms for the signer, hence the size of the revocation token increases in relation to the total number of pseudonyms per signer. To address the challenge of keeping the size of revocation tokens constant, pseudonyms are generated using a combination of forward and reverse cryptographic hash chains.

The exculpable property of SRBE protects signers from the group manager. Even the group manager cannot forge a signature of any honest signer (i.e., private key of the signer is not compromised), so that the signer cannot dispute. A limitation of SRBE is that signatures signed under the same pseudoID (within the same time period) can be linked by the verifier (i.e., the data-collection server). Similar limitations exist in other group signature schemes

[49].

An ongoing challenge is to design a group signature scheme with sublinear revocation and unlinkability support within a time interval and across intervals. The significance of the SRBE scheme is that it has the potential to make large-scale smartphone applications - whose privacy costs were previously formidable - become a reality. It provides a practical and fast alternative to existing group signatures, through reconciling the tradeoff between unlinkability and interval duration.

2.2.1 Related Work

Membership revocation has been a performance bottleneck for employing group signatures in systems with a large volume of dynamic users. After the introduction of group signature schemes [27], [9], different variants of group signature schemes were proposed [15], [63], [13], [11], [50]. In addition, Camenisch and Lysyanskaya [21] constructed a group signature scheme by combining blind signatures with an encryption scheme (e.g., [34], [22]). VLR-based group signature schemes [15], [11], [18], [32], [21] are known to be more practical than the other schemes. Some VLR-based group signatures [21], [18], [32] support backward unlinkability. The authors in [49] presented a group signature scheme with probabilistic revocation (GSPR) that significantly improves the performance of revocation checking. However, probabilistic revocation checking resulting in false positives (i.e., valid signatures mistaken as generated by revoked participants) may not be desirable in crowdsensing. Moreover, the experimental evaluation suggests that revocation check mechanisms of the SRBE scheme run faster than GSPR.

Anonymous credentials were initially envisioned by Chaum [25]. Anonymous credentials allow one to prove the ownership of any credential anonymously. Camenisch and Lysyanskaya proposed the first practical anonymous credential system [20]. Camenisch et al. realized the potential of the first prototype of anonymous credential system [23] for general purpose use. Most anonymous credential systems use blind signatures as the main building block (e.g., restrictive blind signatures for anonymous offline ecash [17], the first practical anonymous credential system [20], anonymous credentials based on bilinear maps [21], anonymous credentials light [7]). In [54], the authors proposed a practical VLR mechanism for anonymous credential systems supporting backward unlinkability. However, expanding this revocation mechanism for group signature schemes remains an ongoing problem. The main performance bottleneck of this VLR scheme [54] is the generation and distribution of revocation lists in each time period by the group manager. To generate a revocation list, it requires O(N) exponentiation operations of large numbers (i.e. not cost-effective), where N is the total number of revoked user. In the SRBE scheme, on the other hand, revocation lists are maintained by the verifiers. Upon receipt of a revocation token, verifiers update its revocation list with only $O(\log^2 N)$ comparisons.

One limitation across all the signature schemes providing privacy preserving accountability is the lack of accountability from the trusted entity. All these approaches lack appropriate mechanisms to prevent mass surveillance by the trusted entity. A case study of the problem is presented in [46].

2.3 Threat Model

We focus on three categories of threats to maintaining identity management integrity.

- Data forgery - Fake data reports (e.g. fake traffic congestion reports) may be submitted.
- Identity forgery - Unauthorized individuals and devices that are not part of the cre-

dentialed group may attempt to submit data reports. In addition, anyone - including the group manager - may attempt to forge the identity of a signer to submit malicious or fake data reports.

- Honest data collector - The data collection server complies with the protocol, but may nevertheless attempt to track a participant through her data reports. This type of adversary, also known as semi-honest, seeks unauthorized access to participant information for personal gain. For example, the data-collection server may examine (1) the signature to identify the user, (2) data reports to identify the context and location of the user, and (3) IP address history, as well as movement trajectory, to location the user.

The credential distribution between the group manager and the participants is assumed secure. In addition, the mobile application on the participant's device is assumed to be trustworthy. The security standard for this application includes being free of spyware, stealth tracking capability, and data-leak vulnerabilities. Advanced collusion and correlation attacks for de-anonymization, such as the semi-honest data-collection server colludes with a mobile service provider or correlates sensory data with known locations of a participant, are out of the research scope. We assume that external adversaries who may launch disruptive attacks such as DDoS and jamming can be identified with existing detection software. Traffic analysis threats from adversaries that are external to a groupsensing system (e.g., routers, access points, and other network intermediaries) are also out of the research scope.

2.4 Security and Privacy Goals

The crowdsensing prototype, GROUPSENSE, has three security and privacy goals:

- Accountability (Traceability) - The sensing group membership of a misbehaving participant can be identified and revoked efficiently.

- Identity Unforgeability - In groupsensing, this goal is two-fold: (1) The data-collection server can verify that received data reports are from valid group members so that any data submissions outside of group membership can be automatically discarded. (2) No one, including the group manager, can forge the identity of a valid signer.

- Sensing-time Anonymity - The data reports submitted by a participant do not provide any information that enables the data-collection server to link them with reports of the same participant in previous time periods, even after the signer is revoked.

Depending on specific application requirements, these security properties can be adjusted as tighter, more relaxed, or more fine-grained. For example, an application might require support for temporary revocation, unforgeability of group managers, partial anonymity (some of the data reports are linkable) or backward unlinkable anonymity (anonymity of data reports prior to the revocation), etc.

2.5 Android and Mobile Computing

The architecture of the Internet is constantly evolving with new features that encourage connectivity through mobile devices which seek to exchange information with one another. It is exceedingly apparent that the data consumed by the average smartphone user today is growing exponentially. Smartphones have been established as an essential part of users' daily lives. The increasing usage of mobile devices has introduced greater opportunities for adversaries interested in intercepting transmitted signals from technologies from baby monitors [5] to digital commerce transactions [36].

The perceived vulnerabilities associated with exchanging sensitive data (e.g., location, personal information, financial records, etc.) loom as a barrier to full-scale adoption of mobile devices for digital transactions. The far-reaching and dynamic way that stakeholders interface with systems, from traditional desktops to mobile devices, has made it increasingly difficult to institute universal security controls and measures [73].

The paradigm of citizen science and crowdsensing is based on individuals sharing observations or experiences in the physical world using their mobile sensing devices to generate collective intelligence. The cost-effectiveness of crowdsensing attracts diverse communities to build innovative applications using crowdsensing. Crowdsensing also opens new possibilities for interdisciplinary collaboration across various disciplines (e.g. physics, chemistry, ecology, life science, computer science, engineering, etc.) to address greater humanity and societal issues. Researchers have invented portable experimental devices that are powered by smartphones. For example, interdisciplinary works as shown in Figure 2.4 (e.g., smartphone-based PCR (polymerase chain reaction) [59], [75], microscopes [47], [48], [12]) present bioanalysis with smartphone-based portable devices, enable timely and effective management (previously unknown) of epidemics (e.g., Ebola, Zika), where early sample collection, diagnosis, and monitoring is extremely critical. These devices enable citizen science and motivate our study on privacy implications.

Unfortunately, modern mobile devices with elaborate sensing capabilities are also known to bring "Big Brother"[1] into effect. Crowdsensing has the potential to perpetuate this proclivity for surveillance by enabling new possibilities to "track" users for the "sake" of data collection [30], [43], [66] and sell their sensitive personal information without their consent. Anyone having access to this information is capable of doing malicious activities.

[1] Brother is a fictional character - subscribing to institutionalized surveillance as a tactical means for controlling society - in George Orwell's novel "Nineteen Eighty-Four".

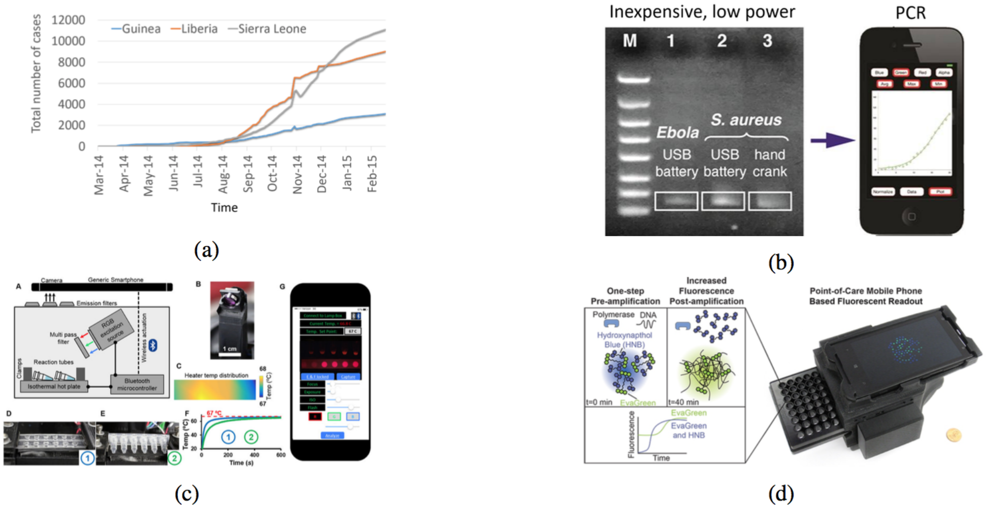


Figure 2.4: (a) Early spread of epidemic [1]; (b) Smartphone-base bioanalysis device [59]; (c) Smartphone-based bioanalysis for Zika screening [75]; (d) Smartphone-based microscope [47].

Attracting different specialized research communities from diverse backgrounds is essential for the sustainability and advancement of the mobile crowdsensing paradigm. However, without adhering to a systematic and disciplined approach to user security, the crowdsensing community could find itself in an undisciplined world. Barring strict security measures, application developers might struggle to align appropriate security solutions to their specific use cases, and security specialists might build systems assuming impractical system models. For example, it is known that most of the data collection servers collect data to enhance their quality of service, yet most of the frameworks for privacy preserving crowdsensing either do not assume the presence of data collectors' application in the users' mobile phone or the security model is inconsistent with the assumption [30].

Chapter 3

User Study

3.1 Study Design

User studies are essential catalysts to the evaluation process by eliciting and disseminating feedback from potential consumers. These studies can be conducted either through physical, remote, or even crowdsourcing mechanisms (such as Amazon's Mechanical Turk [44]). An Institutional Review Board (IRB) approved study was conducted in person to evaluate the usability of our privacy-preserving crowdsensing application. Participants recruited were both undergraduate and graduate students at Virginia Tech. A total of 22 students participated in the user study. Each participant received a detailed explanation of the study objective and was then asked to use the Android device that contained the crowdsensing application for fifteen minutes. Study participants were encouraged to use the device during this period for usual and customary activities, including but not limited to surfing the web, listening to music, watching videos, searching the map, and taking pictures. The participants were instructed not to enter any sensitive information into the device.

3.2 Participant Survey

Immediately following the user study, each participant was given an IRB approved survey to provide feedback on the existing application. The participant survey consisted of eight questions, taking each participant an initial estimate of ten minutes to complete. Table 3.1 shows responses to questions aimed towards gathering demographic information about the participants. The survey asked for participants to provide their exact age, which were consequently grouped into age ranges for quantitative readability. In addition, the survey asked for the participant's specific technical background, but the information has been simplified to a binary response for display purposes. Future studies may consider converting this element into an interval variable (e.g/ a scale of 1-10) for enhancing the relationship analysis between technical proficiency and participant responses. Based upon participant responses, the typical participant was a male between the ages of 21-24 with some type of technical background.

Age	
< 21	36.4%
21 - 24	54.5%
> 24	9.1%
Gender	
Male	68.2%
Female	31.8%
Prefer Not to Answer	0%
Technical Background	
Yes	86.4%
No	13.6%

Table 3.1: Demographic information of participants.

3.2.1 Survey Results

Table 3.2 shows the results of the quantitative questions asked on the participant survey. In addition to the quantitative questions, each question gave the participant space to provide a comment. Question 1 asked, "Would you be concerned of your privacy, if you were a contributor to a citizen science project using your smartphone?" The results show that almost 60% of participants indicated "No" for this question. In addition, participants who chose "Yes" were given room to provide an explanation. Out of the eight comments, some of the most interesting included:

- "I don't know what kind of information is vulnerable."
- "I should worry about my privacy because mobile phones have very sensitive personal info."
- "I wouldn't want sensitive data to be visible to anyone or vulnerable to attack."
- "How can I trust an app, what guarantees do I have?"

Question 2 asked, "Does a long-term citizen science project (i.e. 3 months) give you more privacy concerns than a shorter one (i.e. 1 week)?" The results show that almost 64% of participants indicated "Yes" for this question. In addition, participants were provided room to comment on their answer choice. Out of the 18 participants who provided comments, 5 responded "No" to the question and 13 responded "Yes". Out of those who responded "No", some of the comments included:

- "As long as the security procedures are sound, length of time shouldn't be an issue."
- "As long as my data is being protected."

Out of those who responded "Yes", some of the comments included:

- "More likely to hit all scopes of private info in a longer amount of time."
- "Longer time gives me more privacy concerns."

Question 3 asked, "If you decided to participate in a citizen science project, would you be willing to install privacy-enhancing apps on your smartphone to protect you?" For this question, the results show that almost 82% of participants responded "Yes". The participants who chose "No" were instructed to provide a comment. The responses included:

- "I am worrying about the performance of my phone, and battery consumption."
- "Inconvenient."

Question 4 asked, "If privacy-enhancing apps are not available, would you still be willing to participate in citizen science?" The results show that about 59% of participants chose "Yes" for this question. The higher percentage of participants who would still participate in the citizen science campaign if privacy-enhancing apps were not available may relate to the discussion in [67]. The authors explain how participants of their study were upset when informed of the information sharing practices of the applications on their smartphones, yet continued using their devices in the same manner. This correlation shows that if a participant is determined to contribute, the privacy concerns will not matter. An explanation was requested for those who chose "Yes", and some of the responses included:

- "I could still take my own security precautions."
- "I trust that most people would not misuse my data."
- "I have had minimal problems regarding security so far."

- "I'm ambivalent towards knowing how my privacy is being affected."

Question 5 asked, "The app in the study allows you to submit data anonymously. Do you feel like you understand its privacy guarantee?" The results show that almost 41% of participants chose "Somewhat" for this question. Participants were given the opportunity to provide a comment for this question. Of those who indicated "Somewhat", the most interesting responses included:

- "It's a little too technical for me."
- "Potential breaches of privacy were not fully explained."
- "Not sure how group signatures work."

Question 6 began by stating, "Suppose that you do not fully understand the privacy guarantee of the app. Please answer the following 2 questions:" Question 6a asked, "Would you still be willing to install the app on your smartphone?" The results showed that slightly over half of participants indicated "No" to this question. Participants were given the opportunity to provide a comment for this question. Some of the comments from participants who answered "Yes" for this question included:

- "I let Google on my phone and they take all my info."
- "If it's from a source I trust."
- "Just because something isn't fully understood doesn't mean it is bad."

Question 6b asked, "Would you still be willing to use your phone to participate in citizen science projects?" The results indicate that exactly half of participants responded "No" for this question. A few comments provided by participants who answered "Yes" included:

- "Privacy is not a huge concern to me."
- "I would take additional security precautions."
- "I would remove sensitive data from my phone."

Question 7 asked participants to rank the "Importance of privacy to you as a citizen science contributor." on a scale of 1 (not important) - 5 (important). The average result for this question was 4.14 which indicates that the group of participants viewed privacy as somewhat important. Finally, the participant was given the opportunity to provide any additional comments. This allowed us to gather information other than what was explicitly asked. Additional comments included:

- "It would be nice to know what data was gathered."
- "I need more explanation about what 'privacy' really entails and how it protects me."
- "It is very important to me that my data remain anonymous to other people."

3.3 Discussion

The three most noteworthy questions were questions 3, 5 and 7 that asked "The app in the study allows you to submit data anonymously. Do you feel like you understand its privacy guarantee?", "Importance of privacy to you as a citizen science contributor?", and "If you decided to participate in a citizen science project, would you be willing to install privacy-enhancing apps on your smartphone to protect you?", respectively. The responses from question 5 indicate that less than half of the user study participants understood the privacy guarantee. The average response to question 7 was 4.14 out of 5, indicating that

Question 1 - Privacy Concern: Citizen Science Participation using Smartphone	
Yes	40.1%
No	59.9%
Question 2 - Privacy Concern: Long-term vs. Short-term Project	
Yes	63.6%
No	46.4%
Question 3 - Willingness to Install Smartphone Application	
Yes	81.8%
No	18.2%
Question 4 - Willingness to Participate in Citizen Science	
Yes	59.1%
No	40.9%
Question 5 - Understanding of Privacy Guarantee	
Yes	54.6%
No	4.5 %
Somewhat	40.9%
Question 6a - Willingness to Install Smartphone Application Without Understanding of Privacy Guarantee	
Yes	45.5%
No	55.5%
Question 6b - Willingness to Participate in Citizen Science Without Understanding of Privacy Guarantee	
Yes	50%
No	50%
Question 7 - Importance of Privacy	
Average 1 - 5	4.14

Table 3.2: Results of participant survey.

the participants generally view privacy as important. The weighted importance and value placed upon privacy standards was underscored by the response rate to Question 3, with almost 82% of responses saying that they would still be willing to install the application on their smartphone.

It is important to note that this study was conducted using a sample of participants not necessarily representative of actual citizen science participants; thus, the behaviors and responses my differ. A demographic profile of real world citizen science participants is discussed in [74]. As shown in the user study results, the typical participant was a college student in their early twenties with at least some technical background. In reality, the typical citizen science participant will be composed of a broader age range with a background in scientific fields. Expanding the sample for future studies, as discussed later in this thesis, would be an enhanced predictor of behavioral variances among science-minded participants. In addition, because this study did not simulate a specific citizen science project (e.g. eBird [69]), the user was unaware of what data was being collected. It would be interesting to investigate how the type of data being collected would impact the participants' reactions. Another area to explore is whether the participants of this user study accurately self-report their behaviors. [72] has discovered for survey-based security research that behaviors involving awareness rather than action (e.g. automatic program update) are not self-reported by participants accurately. These responses that measure awareness lead the participant to skew her answer toward what the research administrator may want to hear. Thus, future studies should augment the survey instrumentation to include observation-based data collection.

The next chapter shows how the results of the user study were factored into application enhancements.

Chapter 4

Crowdsensing Android Application

4.1 What does the application do?

When the user clicks the "Start Sensing" button on the home page, the application starts collecting data on the sensors listed in Table 4.1.

Sensor Name	Type	Description	Use
Accelerometer	Hardware	Measures the acceleration force (including gravity) along the x, y, and z axes	Motion detection
Gravity	Hardware or Software	Measures the force of gravity along the x, y, and z axes	Motion detection
Gyroscope	Hardware	Measures the rate of rotation along the x, y, and z axes	Rotation detection
Linear Accelerometer	Hardware or Software	Measures the acceleration force (excluding gravity) along the x, y, and z axes	Monitors acceleration along a single axis
Magnetometer	Hardware	Measures the ambient geomagnetic field for the x, y, and z axes	Creates a compass
Rotation	Hardware or Software	Measures the orientation of a device by providing the x, y, and z axes of the rotation vector	Motion detection & Rotation detection

Table 4.1: Description of Android sensors the application uses for data collection [2].

On the current prototype, the data is collected but not actually stored to the device. Collecting data on the audio sensor is also an option, but it was not operational. The most important contribution that the application gives is the privacy guarantee. After 15 minutes, which is configurable to collection requirements, the application stops collecting data. Specifics regarding the user interface, for both the original and optimized applications are provided in sections below.

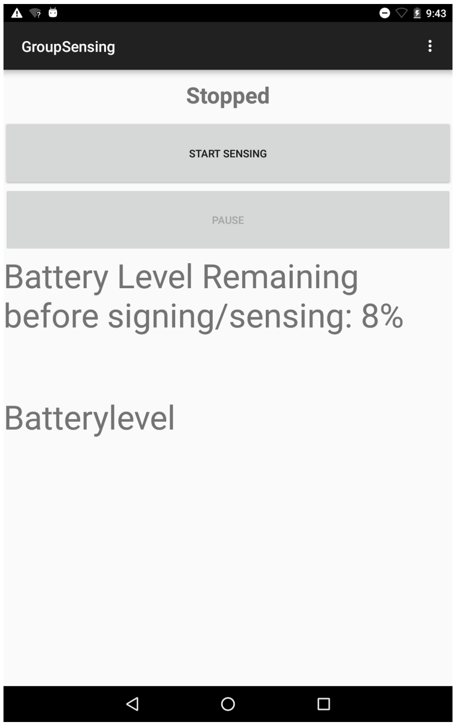
4.2 Original Application

4.2.1 User Interface

The user interface of the original crowdsensing application is only comprised of one main page. This page consists of three text fields and two large buttons. The text field at the top of the screen displays the current mode:

- Stopped - the application was not collecting data (Figure 4.1)
- Paused - the application was paused (Figure 4.3)
- Sensing - the application was collecting data (Figure 4.2)

When the "Start Sensing" button is clicked, the button text changes to "Stop Sensing" and the "Pause" button is enabled. When the "Pause" button is clicked, the button text changes to "Continue". Underneath the two buttons, the application displays information regarding how much battery life remains, expressed as a percentage, prior to signing or sensing. In addition, there is an unnecessary text field beneath the displayed battery percentage.



Stopped

Battery Level Remaining before signing/sensing: 8%

Batterylevel

Figure 4.1: Existing Interface Main Page - Stopped

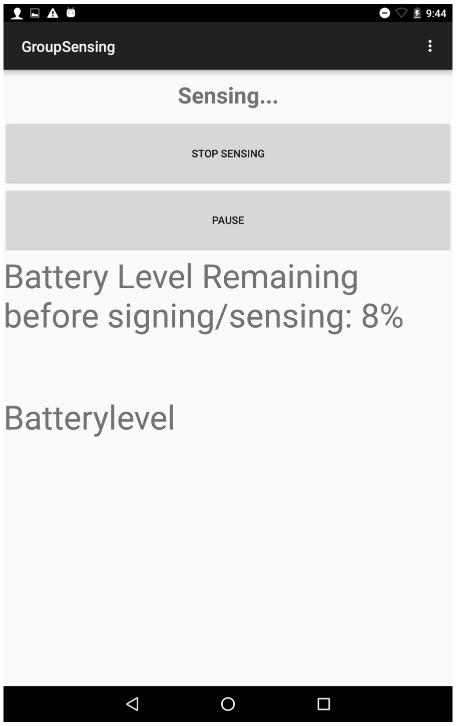
4.2.2 Notifications

The top text field notifies users of the current mode of the application (Sensing or Stopped). The user additionally receives a drop-down notification when the application is currently collecting data (Figure 4.4).

4.3 Optimized Application

4.3.1 Design Requirements

A requirement refers to a statement of what is needed to design a system to align with user expectations [39], and bridges the gap between analysis and design. This gap is mitigated by an initial elicitation of requirements, including the identification of user needs and design



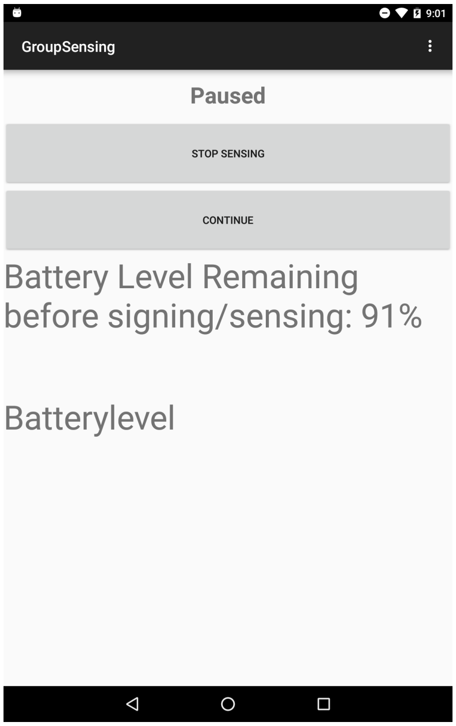
Sensing...

Battery Level Remaining before signing/sensing: 8%

Batterylevel

Figure 4.2: Existing Interface Main Page - Sensing

requirements from the analysis. This process is done by an iterative cycle of designing, prototyping, evaluating, and analyzing. Requirements can be identified from comments elicited from participant on a scope or need, or inferred from quantitative results. While the study identifies information, it does not directly correlate to the design needs. Thus, we identify interaction design requirements, where software and implementation specifics are excluded. This extraction of requirements produces documented business and functional specifications to be during the design process. Extracting requirements requires an iterative and deductive thinking method in order to construct each "requirement statement" [39]. To create the requirement statement, user needs, documented as business and functional requirements, are translated into measurable and testable design requirements, including specific user interface features. These interaction requirements turn into system requirements allowing for tie-back and traceability between design features and their input requirements.



Paused

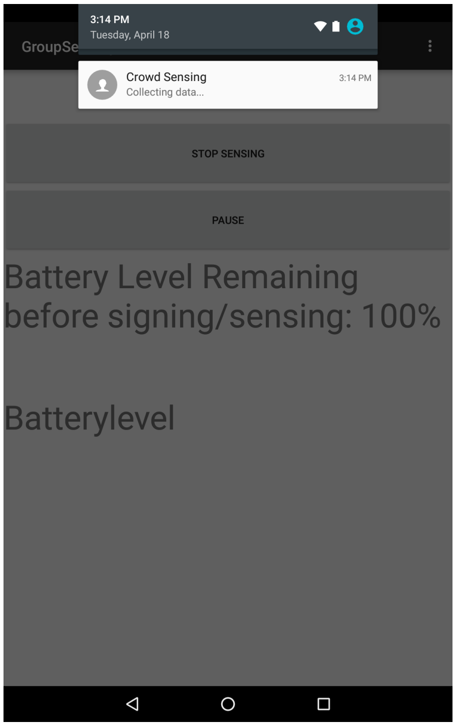
Battery Level Remaining before signing/sensing: 91%

Batterylevel

Figure 4.3: Existing Interface Main Page - Paused

The requirement statements formulated from the user study quantitative data and comments are shown below.

- **Understanding of Privacy Guarantee**
  - The user shall be able to view information about the privacy guarantee of the application.
  - The user shall be able to view information about the group signature scheme.
  - The user shall understand what information is vulnerable.
- **Understanding of Application Functionality**
  - The user shall know what data is being collected by the application.
  - The user shall know what stage of data collection (i.e. mode) the application is currently in.



Crowd Sensing
Collecting data...

Battery Level Remaining before signing/sensing: 100%

Batterylevel

Figure 4.4: Existing Interface Main Page - Sensing Notification

- The user shall know when submission is complete - data is no longer being collected.

- **Security Requirements**
  - The system shall confirm that the user would like to begin submitting data.
  - The system shall deploy the SRBE group signature scheme to protect sensitive participant data.

- **User Experience**
  - The user shall be aware of the battery life of the device prior to data submission.
  - The system shall adopt a cleaner appearance.

4.3.2 Additional Design Elements

From the identified design requirements, specific design elements to add were determined.

- **Understanding of Privacy Guarantee**
  - Add information page
    - Information about the privacy guarantee.
    - Information about the group signature scheme.
    - Information about what information is vulnerable.
- **Understanding of Application Functionality**
  - Add welcome page
    - Display what data is being collected
  - Display what mode the application is currently in.
  - Notification when data collection is complete.
- **Security Requirements**
  - Confirm that the user would like to begin submitting data.
  - Implement the SRBE group signature scheme.
- **User Experience**
  - Display the battery life of the device prior to submitting data.
  - Adopt an interface appearance that promotes overall enhanced application readability and usability for the participant.

4.3.3 User Interface

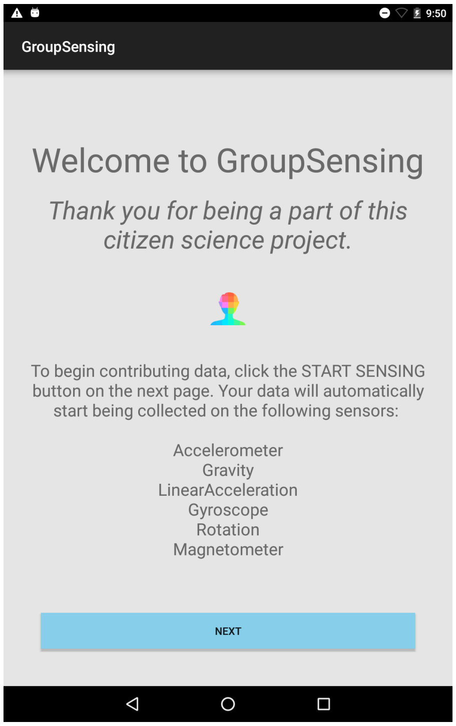
The new user interface is comprised of three pages. As pictured in Figure 4.5, the first page is the home page that contains information about the application. In addition to a brief welcome, there are basic instructions informing the user how to use the crowdsensing application. The sensors being used are listed below the instructions (Table 4.1), enabling the user to know what data is being collected. By providing this information, the user develops an expectation of which sensitive resources are being used. [52] discusses how the user's expectations could result in an increase in trust regarding the application. Finally, a button on the bottom on the page enables the user to navigate to the main page.

The main page can be seen in Figure 4.6. There are two large buttons that affect the current mode. On startup, the first button is enabled and the second button is disabled, reading "Start Sensing" and "Pause", respectively. The text field above the two buttons displays the current mode:

- Stopped - the application is not collecting data (Figure 4.6)
- Paused - the application is paused (Figure 4.7)
- Collecting Data - the application is collecting data (Figure 4.8)
- Signing Data - the application is signing the collected data (Figure 4.9)
- Sending Data - the application is sending the signed collected data to the data collector (Figure 4.10)

Below the buttons, the battery level remaining before signing is displayed.

The final page in the improved prototype is an information screen (Figure 4.11), which can be accessed by clicking the button at the bottom right-hand corner of the screen, noted with a standard information icon. The information page first provides the user with an explanation of GROUPSENSE and then mentions the purpose for data collection. The page then informs the user that the application is accessing her current location and displays this information. Beneath her current location is a statement explaining the importance of securing personal information and how this is done in GROUPSENSE. [6] mentions the importance of privacy nudges, such as this one, to ensure that the user is aware of the data being collected.



Welcome to GroupSensing

Thank you for being a part of this citizen science project.

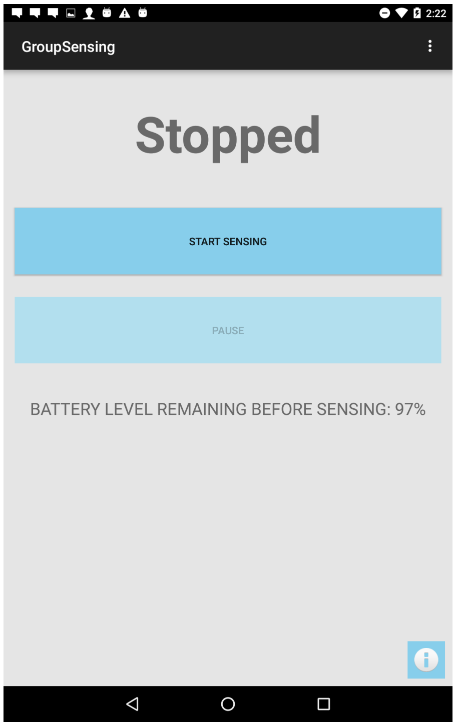
To begin contributing data, click the START SENSING button on the next page. Your data will automatically start being collected on the following sensors:

Accelerometer
Gravity
LinearAcceleration
Gyroscope
Rotation
Magnetometer

Figure 4.5: New Interface Welcome Page

4.3.4 Notifications

Iterative modifications to the interface were deployed to enhance the user experience. This generally consisted of augmenting the interface notification features in order to improve the user's understanding of what is occurring. On the main page, the current mode of the application was expanded to include: Stopped, Sensing, Signing, or Sending. In previous



GroupSensing

Stopped

BATTERY LEVEL REMAINING BEFORE SENSING: 97%

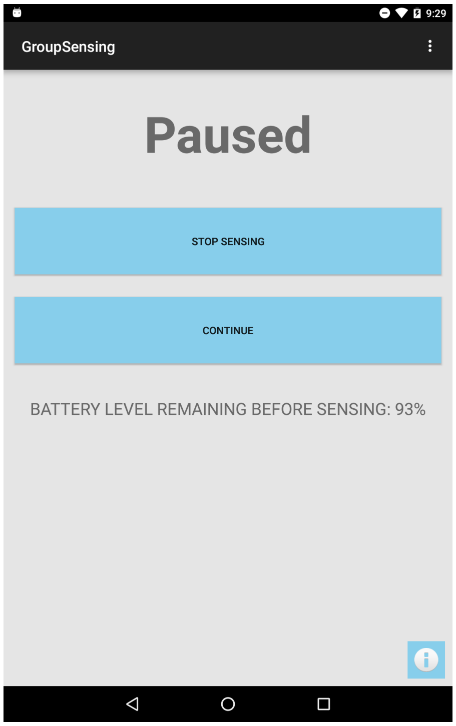
Figure 4.6: New Interface Main Page - Stopped

versions, only Stopped or Sensing were displayed for the user. In the current prototype, the data collected are not actually signed and operationally transacted to the data collector. The notifications, consequently, are simulated for demonstration purposes. In addition, the user now receives the pop-up notification, "Are you sure you want to begin sensing?" when clicking the "Start Sensing" button. Then, the user has the option to either cancel or continue the operation.

4.4 Summary

4.4.1 Initial Design Highlights

- Secure crowdsensing Android application



Paused

BATTERY LEVEL REMAINING BEFORE SENSING: 93%

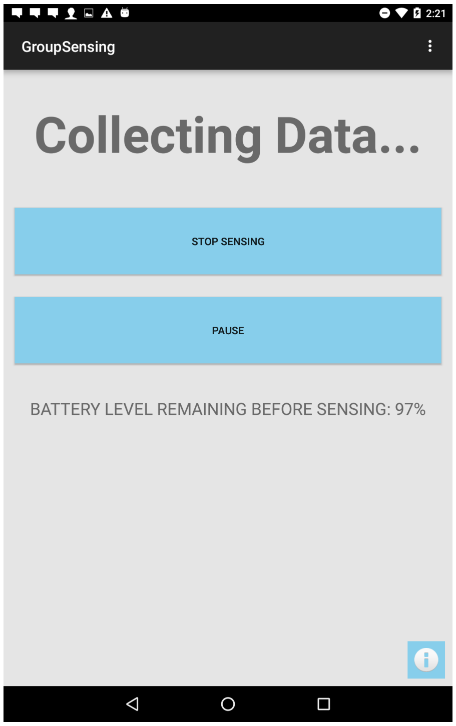
Figure 4.7: New Interface Main Page - Paused

- Novel group signature scheme

4.4.2 Further Design Additions

- Information displayed about application
- Enhanced notification system
- Usable application supported by user study

Without these modifications to the application, the user may not be aware of the privacy guarantees. As the primary portal to citizen science exchanges, it is paramount for the application to provide clear and precise notifications throughout the data collection transaction. In addition to the enhanced notification system, future releases of the application

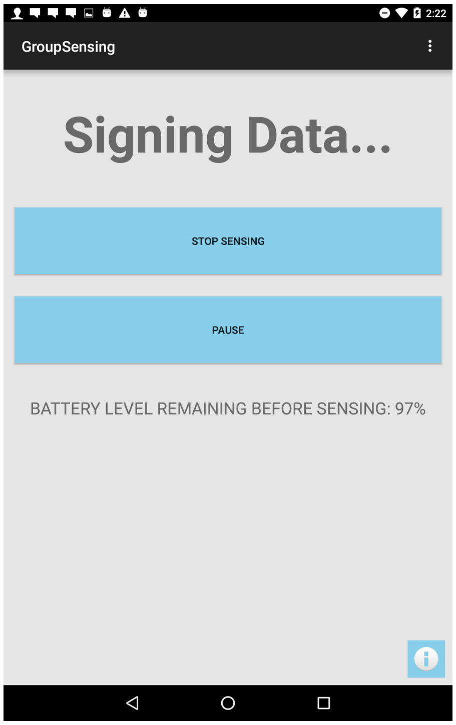


Collecting Data...

BATTERY LEVEL REMAINING BEFORE SENSING: 97%

Figure 4.8: New Interface Main Page - Collecting Data

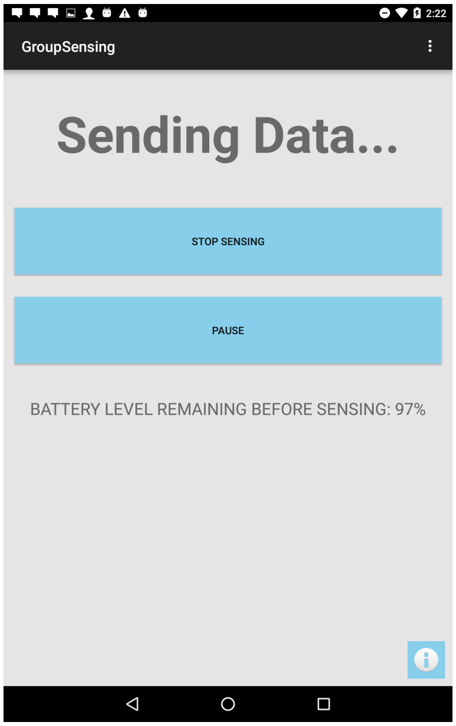
can provide a FAQ page, as well as contact information, to address issues or concerns that users may have. Alleviation of the privacy concerns, as supported by the responses in the user study, are critical to acceptance and usage by the crowdsourcing community.



Signing Data...

BATTERY LEVEL REMAINING BEFORE SENSING: 97%

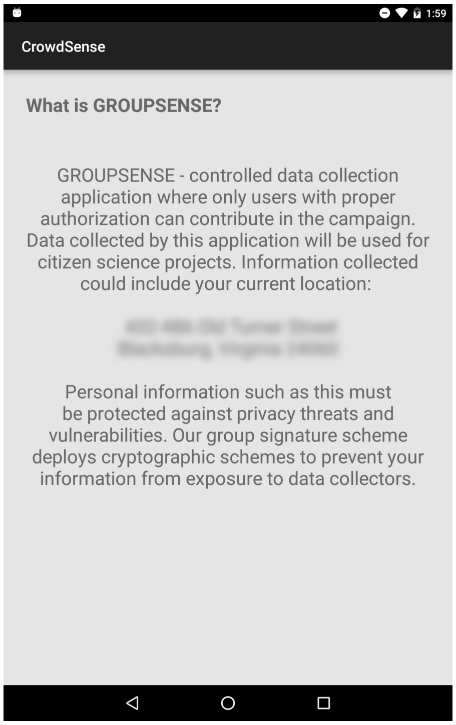
Figure 4.9: New Interface Main Page - Signing



Sending Data...

BATTERY LEVEL REMAINING BEFORE SENSING: 97%

Figure 4.10: New Interface Main Page - Sending Data



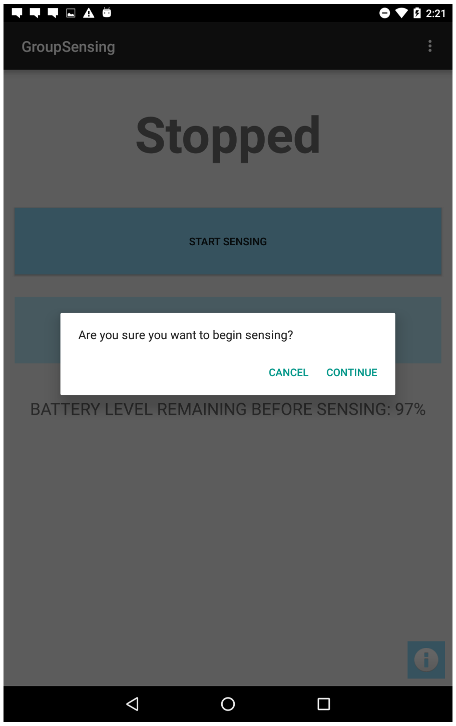
What is GROUPSENSE?

GROUPSENSE - controlled data collection application where only users with proper authorization can contribute in the campaign. Data collected by this application will be used for citizen science projects. Information collected could include your current location:

[redacted]

Personal information such as this must be protected against privacy threats and vulnerabilities. Our group signature scheme deploys cryptographic schemes to prevent your information from exposure to data collectors.

Figure 4.11: New Interface Information Page



Are you sure you want to begin sensing?

BATTERY LEVEL REMAINING BEFORE SENSING: 97%

Figure 4.12: New Interface Main Page - Confirmation

Chapter 5

Conclusion & Future Work

5.1 Future Work

5.1.1 Additional User Study

One area for future work is the scale of the user study. A follow-up user study should be conducted on the revised application to measure the effectiveness and user response to the interface modifications and feature enhancements. A similar sampling of users, using a consistent selection criteria, should be deployed in order to achieve reliable measurement outcomes (i.e. free of confounding bias). Further, participants from the first user study should be asked to participate in the follow-up user study. Hence, the study sample would be stratified into three user study groups: (1) those who only participated in the first user study; (2) those who only participated in the second user study; and (3) those who participated in both studies. The members of group three should demonstrate an improved understanding of the privacy guarantees in order to prove that the modifications to the applications were effective. Group two exists to remove any bias that the participant may experience having already seen and used the application.

In addition, the scope of the user study should be improved upon. First, the study participants should be selected in order to include a more diversified group of respondents. Also, the analysis should be extended to target potential predictors of user participation by analyzing how demographic information affects how a participant answers each question.

5.1.2 Application Modifications

As the application is currently in the prototype stage, many areas for enhancement remain prior to production release. First, the application has a lengthy initial load time because keys are generated for signing purposes at startup. This can be optimized by generating keys only once and reusing them by using caching techniques - storing them in the device's memory. In addition, the application should implement actual crowdsensing by storing the data collected onto the device to be signed and submitted to the data collector.

Relating to the interface, possible features that can be incorporated into future product enhancements include:

- Adding a dashboard for participants to view their history and outcomes.
- Integration of a financial exchange application (e.g. PayPal or Venmo) for participants to receive financial compensation.
- Leveraging identity management best practices, including multi-factor authentication, biometric authentication, and third-party identity verification systems.

In addition to these possible enhancements, crowdsourcing could be utilized to receive feedback on the interface design (such as the system created in [55].

5.1.3 Security and Performance Enhancements

The performance and the scalability of existing privacy preserving authentication protocols are inadequate for large-scale crowdsensing applications. Thus, more focused efforts in this direction are necessary. Side channel-based attacks play a critical role to evade any security shield [57]. Investigation and mitigation of any such side channel-based attacks are also important.

5.2 Conclusion

The motivation behind this work was the broad-based need for safe and secure large-scale anonymous smartphone applications, such as crowdsensing. SRBE, a provable secure group signature scheme realizes sublinear revocation checking, is the basis behind this technical contribution. Revocation checking is a frequently executed operation required for each signature verification. SRBE also provides typical group signature guarantees, such as backward unlinkability across time intervals in a periodic setup. The fast revocation checking is made possible through utilizing and integrating structural building blocks, including cryptographic, algorithmic, and data. We have highlighted risks, limitations, and constraints associated with the SRBE scheme in this paper. An additional noteworthy technical contribution is the SRBE-based crowdsensing prototype with Android support called GROUPSENSE. The design and privacy guarantees for this application can be expanded to meet the needs of any groupsensing application. The additional security guarantees and speed improvements make it a more viable prototype of groupsensing projects.

An Android application implementing GROUPSENSE was developed and evaluated through a user study. The user study was designed with the goal of evaluating the usability of the application, whether participants can reasonably understand the privacy guarantees of group signatures, and whether the group signatures would alleviate the privacy concern inherent among citizen science participants. The user study confirmed that the participants stressed their privacy with weighted importance, while only somewhat understanding the privacy guarantees of the application. Based on these findings, modifications were made to the application with the objective of improving participant understanding. Application modifications included additional screens that provide supplementary information as well as the deployment of a more sophisticated notification system.

The significance of this work is that it perpetuates the movement of provable secure group signatures closer to practical deployment in a large scale. Such effort on privacy is necessary with the ever-increasing number of user-centric applications. The migration of usage to mobile devices and the increasing connectivity to large volumes of social media sources perpetuates a more robust, yet data privacy vulnerable, portal of data collection. Privacy-preserving software that encrypts data in motion between mobile devices and the enterprise data storage platform and network will continue to gain traction. This movement stems from the adoption of mobile devices as a primary system interface. Studies such as this will further the cause toward increasing user dependency on mobile devices in alignment with their understanding of privacy risks and vulnerabilities. Falling short of this premise provides a channel for opportunistic 'Big Brother' and semi-honest users that exist within our digital society.

Bibliography

[1] Ebola outbreak in west africa - reported cases graphs. `https://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/cumulative-cases-graphs.html`, February 2016 accessed, 16-April-2017.

[2] Sensors overview. `https://developer.android.com/guide/topics/sensors/sensors_overview.html`, accessed 20-April-2017.

[3] Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE Security and Privacy*, 3(1):26–33, January 2005.

[4] Berker Agir, Thanasis G. Papaioannou, Rammohan Narendula, Karl Aberer, and Jean-Pierre Hubaux. User-side adaptive protection of location privacy in participatory sensing. *Geoinformatica*, 18(1):165–191, January 2014.

[5] K. Albrecht and L. Mcintyre. Privacy nightmare: When baby monitors go bad [opinion]. *IEEE Technology and Society Magazine*, 34(3):14–19, Sept 2015.

[6] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and Yuvraj Agarwal. Your location has been shared 5,398 times!: A field study on mobile app privacy nudging. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 787–796, New York, NY, USA, 2015. ACM.

[7] Foteini Baldimtsi and Anna Lysyanskaya. Anonymous credentials light. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer &#38; Communications Security*, CCS '13, pages 1087–1098, New York, NY, USA, 2013. ACM.

[8] Susan B. Barnes. A privacy paradox: Social networking in the united states. *First Monday*, 11(9), 2006.

[9] Mihir Bellare, Daniele Micciancio, and Bogdan Warinschi. Foundations of group signatures: Formal definitions, simplified requirements, and a construction based on general assumptions. In *Advances in Cryptology - EUROCRYPT 2003, International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw, Poland, May 4-8, 2003, Proceedings*, volume 2656 of *Lecture Notes in Computer Science*, pages 614–629. Springer, 2003.

[10] G. Bianchi, M. Bonola, V. Falletta, F. S. Proto, and S. Teofili. The sparta pseudonym and authorization system. *Sci. Comput. Program.*, 74(1-2):23–33, December 2008.

[11] Patrik Bichsel, Jan Camenisch, Gregory Neven, Nigel P. Smart, and Bogdan Warinschi. *Get Shorty via Group Signatures without Encryption*, pages 381–398. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[12] Isaac I. Bogoch, Hatice C. Koydemir, Derek Tseng, Richard K. D. Ephraim, Evans Duah, Joseph Tee, Jason R. Andrews, and Aydogan Ozcan. Evaluation of a mobile phoneâĂŞbased microscope for screening of schistosoma haematobium infection in rural ghana. *The American Journal of Tropical Medicine and Hygiene*, pages –, 2017.

[13] Dan Boneh and Xavier Boyen. Short signatures without random oracles and the sdh assumption in bilinear groups. *J. Cryptology*, 21(2):149–177, 2008.

[14] Dan Boneh, Ben Lynn, and Hovav Shacham. Short signatures from the weil pairing. In *Proceedings of the 7th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology*, ASIACRYPT '01, pages 514–532, London, UK, UK, 2001. Springer-Verlag.

[15] Dan Boneh and Hovav Shacham. Group signatures with verifier-local revocation. pages 168–177, 2004.

[16] Ioannis Boutsis and Vana Kalogeraki. Privacy preservation for participatory sensing data. In *2013 IEEE International Conference on Pervasive Computing and Communications, PerCom 2013, San Diego, CA, USA, March 18-22, 2013*, pages 103–113, 2013.

[17] Stefan Brands. Untraceable off-line cash in wallets with observers (extended abstract). In *Advances in Cryptology - CRYPTO '93, 13th Annual International Cryptology Conference, Santa Barbara, California, USA, August 22-26, 1993, Proceedings*, volume 773 of *Lecture Notes in Computer Science*, pages 302–318. Springer, 1993.

[18] Julien Bringer and Alain Patey. Backward unlinkability for a VLR group signature scheme with efficient revocation check. *IACR Cryptology ePrint Archive*, 2011:376, 2011.

[19] Jan Camenisch and Thomas Groß. Efficient attributes for anonymous credentials. *ACM Trans. Inf. Syst. Secur.*, 15(1):4:1–4:30, March 2012.

[20] Jan Camenisch and Anna Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques: Advances in Cryptology*, EUROCRYPT '01, pages 93–118, London, UK, UK, 2001. Springer-Verlag.

[21] Jan Camenisch and Anna Lysyanskaya. Signature schemes and anonymous credentials from bilinear maps. In *Crypto 2004, LNCS 3152*, pages 56 – 72. Springer Verlag, 2004.

[22] Jan Camenisch and Victor Shoup. Practical verifiable encryption and decryption of discrete logarithms. In *Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings*, volume 2729 of *Lecture Notes in Computer Science*, pages 126–144. Springer, 2003.

[23] Jan Camenisch and Els Van Herreweghen. Design and implementation of the idemix anonymous credential system. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, CCS '02, pages 21–30, New York, NY, USA, 2002. ACM.

[24] Supriyo Chakraborty, Kasturi Rangan Raghavan, Matthew P. Johnson, and Mani B. Srivastava. A framework for context-aware privacy of sensor data on mobile systems. In *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications*, HotMobile '13, pages 11:1–11:6, New York, NY, USA, 2013. ACM.

[25] David Chaum. Blind signatures for untraceable payments. In D. Chaum, R.L. Rivest, and A.T. Sherman, editors, *Advances in Cryptology Proceedings of Crypto 82*, pages 199–203, 1983.

[26] David Chaum. Security without identification: Transaction systems to make big brother obsolete. *Commun. ACM*, 28(10):1030–1044, October 1985.

[27] David Chaum and Eugène Van Heyst. Group signatures. pages 257–265, 1991.

[28] L. Cheng, L. Kong, C. Luo, J. Niu, Y. Gu, W. He, and S. Das. Deco: False data detection and correction framework for participatory sensing. In *2015 IEEE 23rd International Symposium on Quality of Service (IWQoS)*, pages 213–218, June 2015.

[29] Long Cheng, Linghe Kong, Chengwen Luo, Jianwei Niu, Yu Gu, Wenbo He, and Sajal K. Das. Deco: False data detection and correction framework for participatory sensing. In *23rd IEEE International Symposium on Quality of Service, IWQoS 2015, Portland, OR, USA, June 15-16, 2015*, pages 213–218, 2015.

[30] Delphine Christin. Privacy in mobile participatory sensing. *J. Syst. Softw.*, 116(C):57–68, June 2016.

[31] Delphine Christin, Andreas Reinhardt, Salil S. Kanhere, and Matthias Hollick. A survey on privacy in mobile participatory sensing applications. *J. Syst. Softw.*, 84(11):1928–1946, November 2011.

[32] Cheng-Kang Chu, Joseph K. Liu, Xinyi Huang, and Jianying Zhou. Verifier-local revocation group signatures with time-bound keys. In *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, ASIACCS '12, pages 26–27, New York, NY, USA, 2012. ACM.

[33] Cory Cornelius, Apu Kapadia, David Kotz, Dan Peebles, Minho Shin, and Nikos Triandopoulos. Anonysense: Privacy-aware people-centric sensing. In *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services*, MobiSys '08, pages 211–224, New York, NY, USA, 2008. ACM.

[34] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. *IACR Cryptology ePrint Archive*, 1998:6, 1998.

[35] Emiliano De Cristofaro and Claudio Soriente. Extended capabilities for a privacy-enhanced participatory sensing infrastructure (pepsi). *CoRR*, abs/1308.2921, 2013.

[36] S.M. Furnell and T. Karweni. Security implications of electronic commerce: a survey of consumers and businesses. *Internet Research*, 9(5):372–382, 1999.

[37] Stylianos Gisdakis, Thanassis Giannetsos, and Panagiotis Papadimitratos. Security, privacy, and incentive provision for mobile crowd sensing systems. *IEEE Internet of Things Journal*, 3(5):839–853, 2016.

[38] Stylianos Gisdakis, Thanassis Giannetsos, and Panos Papadimitratos. Sppear: Security &#38; privacy-preserving architecture for participatory-sensing applications. In *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless &#38; Mobile Networks*, WiSec '14, pages 39–50, New York, NY, USA, 2014. ACM.

[39] Rex Hartson and Partha S. Pyla. Chapter 5 - extracting interaction design requirements. In Rex Hartson, , and Partha S. Pyla, editors, *The {UX} Book*, pages 161 – 179. Morgan Kaufmann, Boston, 2012.

[40] Eiman Kanjo. Noisespy: a real-time mobile phone platform for urban noise monitoring and mapping. *Mobile Networks and Applications*, 15(4):562–574, 2010.

[41] Apu Kapadia, David Kotz, and Nikos Triandopoulos. Opportunistic sensing: Security challenges for the new paradigm. In *Proceedings of the First International Conference on COMmunication Systems And NETworks*, COMSNETS'09, pages 127–136, Piscataway, NJ, USA, 2009. IEEE Press.

[42] Leyla Kazemi and Cyrus Shahabi. A privacy-aware framework for participatory sensing. *SIGKDD Explor. Newsl.*, 13(1):43–51, August 2011.

[43] Leyla Kazemi and Cyrus Shahabi. Tapas: Trustworthy privacy-aware participatory sensing. *Knowl. Inf. Syst.*, 37(1):105–128, 2013.

[44] Aniket Kittur, Ed H. Chi, and Bongwon Suh. Crowdsourcing user studies with mechanical turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 453–456, New York, NY, USA, 2008. ACM.

[45] Mikkel Baun Kjærgaard, Jakob Langdal, Torben Godsk, and Thomas Toftkjær. Entracked: Energy-efficient robust position tracking for mobile devices. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services*, MobiSys '09, pages 221–234, New York, NY, USA, 2009. ACM.

[46] Markulf Kohlweiss and Ian Miers. Accountable metadata-hiding escrow: A group signature case study. *PoPETs*, 2015(2):206–221, 2015.

[47] Janay E. Kong, Qingshan Wei, Derek Tseng, Jingzi Zhang, Eric Pan, Michael Lewinski, Omai B. Garner, Aydogan Ozcan, and Dino Di Carlo. Highly stable and sensitive nucleic acid amplification and cell-phone-based readout. *ACS Nano*, 11(3):2934–2943, 2017. PMID: 28234452.

[48] Malte Kühnemund, Qingshan Wei, Evangelia Darai, Yingjie Wang, Ivan Hernandez-Neuta, Zhao Yang, Derek Tseng, Annika Ahlford, Lucy Mathot, Tobias Sjöblom, Aydogan Ozcan, and Mats Nilsson. Targeted dna sequencing and in situ mutation analysis using mobile phone microscopy. *Nature Communications*, 8, 2017.

[49] Vireshwar Kumar, He Li, Jung-Min (Jerry) Park, Kaigui Bian, and Yaling Yang. Group signatures with probabilistic revocation: A computationally-scalable approach for providing privacy-preserving authentication. In *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, CCS '15, pages 1334–1345, New York, NY, USA, 2015. ACM.

[50] Benoit Libert, Thomas Peters, and Moti Yung. Short group signatures via structure-preserving signatures: Standard model security from simple assumptions. Cryptology ePrint Archive, Report 2015/743, 2015. `http://eprint.iacr.org/2015/743`.

[51] Kiho Lim and D. Manivannan. An efficient protocol for authenticated and secure message delivery in vehicular ad hoc networks. *Veh. Commun.*, 4(C):30–37, April 2016.

[52] Jialiu Lin, Shahriyar Amini, Jason I. Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. Expectation and purpose: Understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 501–510, New York, NY, USA, 2012. ACM.

[53] Mu Lin, Nicholas D. Lane, Mashfiqui Mohammod, Xiaochao Yang, Hong Lu, Giuseppe Cardone, Shahid Ali, Afsaneh Doryab, Ethan Berke, Andrew T. Campbell, and Tanzeem Choudhury. Bewell+: Multi-dimensional wellbeing monitoring with community-guided user feedback and energy optimization. pages 10:1–10:8, 2012.

[54] Wouter Lueks, Gergely Alpár, Jaap-Henk Hoepman, and Pim Vullers. Fast revocation of attribute-based credentials for both users and verifiers. *IACR Cryptology ePrint Archive*, 2015:237, 2015.

[55] Kurt Luther, Amy Pavel, Wei Wu, Jari-lee Tolentino, Maneesh Agrawala, Björn Hartmann, and Steven P. Dow. Crowdcrit: Crowdsourcing and aggregating visual design critique. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work &#38; Social Computing*, CSCW Companion '14, pages 21–24, New York, NY, USA, 2014. ACM.

[56] Giovanni Merlino, Stamatis Arkoulis, Salvatore Distefano, Chrysa Papagianni, Antonio Puliafito, and Symeon Papavassiliou. Mobile crowdsensing as a service. *Future Gener. Comput. Syst.*, 56(C):623–639, March 2016.

[57] S. Narain, T. D. Vo-Huu, K. Block, and G. Noubir. Inferring user routes and locations using zero-permission mobile sensors. In *2016 IEEE Symposium on Security and Privacy (SP)*, pages 397–413, May 2016.

[58] Bei Pan, Yu Zheng, David Wilkie, and Cyrus Shahabi. Crowd sensing of traffic anomalies based on human mobility and social media. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL'13, pages 344–353, New York, NY, USA, 2013. ACM.

[59] Aashish Priye, Season Wong, Yuanpeng Bi, Miguel Carpio, Jamison Chang, Mauricio Coen, Danielle Cope, Jacob Harris, James Johnson, Alexandra Keller, Richard Lim, Stanley Lu, Alex Millard, Adriano Pangelinan, Neal Patel, Luke Smith, Kamfai Chan, and Victor M. Ugaz. Lab-on-a-drone: Toward pinpoint deployment of smartphone-enabled nucleic acid-based diagnostics for mobile health care. *Analytical Chemistry*, 88(9):4651–4660, 2016. PMID: 26898247.

[60] James Rachels. Why privacy is important. *Philosophy and Public Affairs*, 4(4):323–333, 1975.

[61] Sazzadur Rahaman, Long Cheng, Danfeng Yao, He Li, and Jung-Min. Park. Provably secure anonymous-yet-accountable crowdsensing with scalable sublinear revocation. In *The 17th Privacy Enhancing Technologies Symposium (PETS). Minneapolis, MN. Jul. 2017*, 2017. [To Appear].

[62] Sazzadur Rahaman, Hannah Roth, Danfeng Yao, and Victor. Ugaz. OrwellâĂŹs nightmare: Challenges toward privacy preserving crowdsensing. In *The 1st IEEE Symposium on Privacy-Aware Computing (IEEE PAC), Washington DC, Aug. 2017*, 2017. [Submitted].

[63] Maxim Raya and Jean-Pierre Hubaux. Securing vehicular ad hoc networks. *J. Comput. Secur.*, 15(1):39–68, January 2007.

[64] Sasank Reddy, Deborah Estrin, Mark Hansen, and Mani Srivastava. Examining micropayments for participatory sensing data collections. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, UbiComp '10, pages 33–36, New York, NY, USA, 2010. ACM.

[65] Francesco Restuccia, Sajal K. Das, and Jamie Payton. Incentive mechanisms for participatory sensing: Survey and research challenges. *ACM Trans. Sen. Netw.*, 12(2):13:1–13:40, April 2016.

[66] Minho Shin, Cory Cornelius, Apu Kapadia, Nikos Triandopoulos, and David Kotz. Location privacy for mobile crowd sensing through population mapping. *Sensors*, 15(7):15285–15310, June 2015.

[67] Irina Shklovski, Scott D. Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. Leakiness and creepiness in app space: Perceptions of privacy and mobile app use. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 2347–2356, New York, NY, USA, 2014. ACM.

[68] Reza Shokri, George Theodorakopoulos, Jean-Yves Le Boudec, and Jean-Pierre Hubaux. Quantifying location privacy. pages 247–262, 2011.

[69] Brian L. Sullivan, Christopher L. Wood, Marshall J. Iliff, Rick E. Bonney, Daniel Fink, and Steve Kelling. ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10):2282 – 2292, 2009.

[70] Khuong Vu, Rong Zheng, and Jie Gao. Efficient algorithms for k-anonymous location privacy in participatory sensing. In *2012 Proceedings IEEE INFOCOM*, pages 2399–2407, March 2012.

[71] Gang Wang, Bolun Wang, Tianyi Wang, Ana Nika, Haitao Zheng, and Ben Y. Zhao. Defending against sybil devices in crowdsourced mapping services. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '16, pages 179–191, New York, NY, USA, 2016. ACM.

[72] Rick Wash, Emilee Rader, and Chris Fennell. Can people self-report security accurately?: Agreement between self-report and behavioral measures. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 2228–2232, New York, NY, USA, 2017. ACM.

[73] Te-En Wei, A. B. Jeng, Hahn-Ming Lee, Chih-How Chen, and Chin-Wei Tien. Android privacy. In *2012 International Conference on Machine Learning and Cybernetics*, volume 5, pages 1830–1837, July 2012.

[74] Sarah Elizabeth West and Rachel Mary Pateman. Recruiting and retaining participants in citizen science: What can be learned from the volunteering literature? *Citizen Science: Theory and Practice*, December 2016. © 2016 The Author(s).

[75] Mohammad Zarei. Portable biosensing devices for point-of-care diagnostics: Recent developments and applications. *TrAC Trends in Analytical Chemistry*, 91:26 – 41, 2017.

[76] Fan Zhang, Li He, Wenbo He, and Xue Liu. Data perturbation with state-dependent noise for participatory sensing. In Albert G. Greenberg and Kazem Sohraby, editors, *INFOCOM*, pages 2246–2254. IEEE, 2012.

[77] Jindan Zhu, Kyu-Han Kim, Prasant Mohapatra, and Paul Congdon. An adaptive privacy-preserving scheme for location tracking of a mobile user. In *10th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, SECON 2013, New Orleans, LA, USA, 24-27 June, 2013*, pages 140–148, 2013.

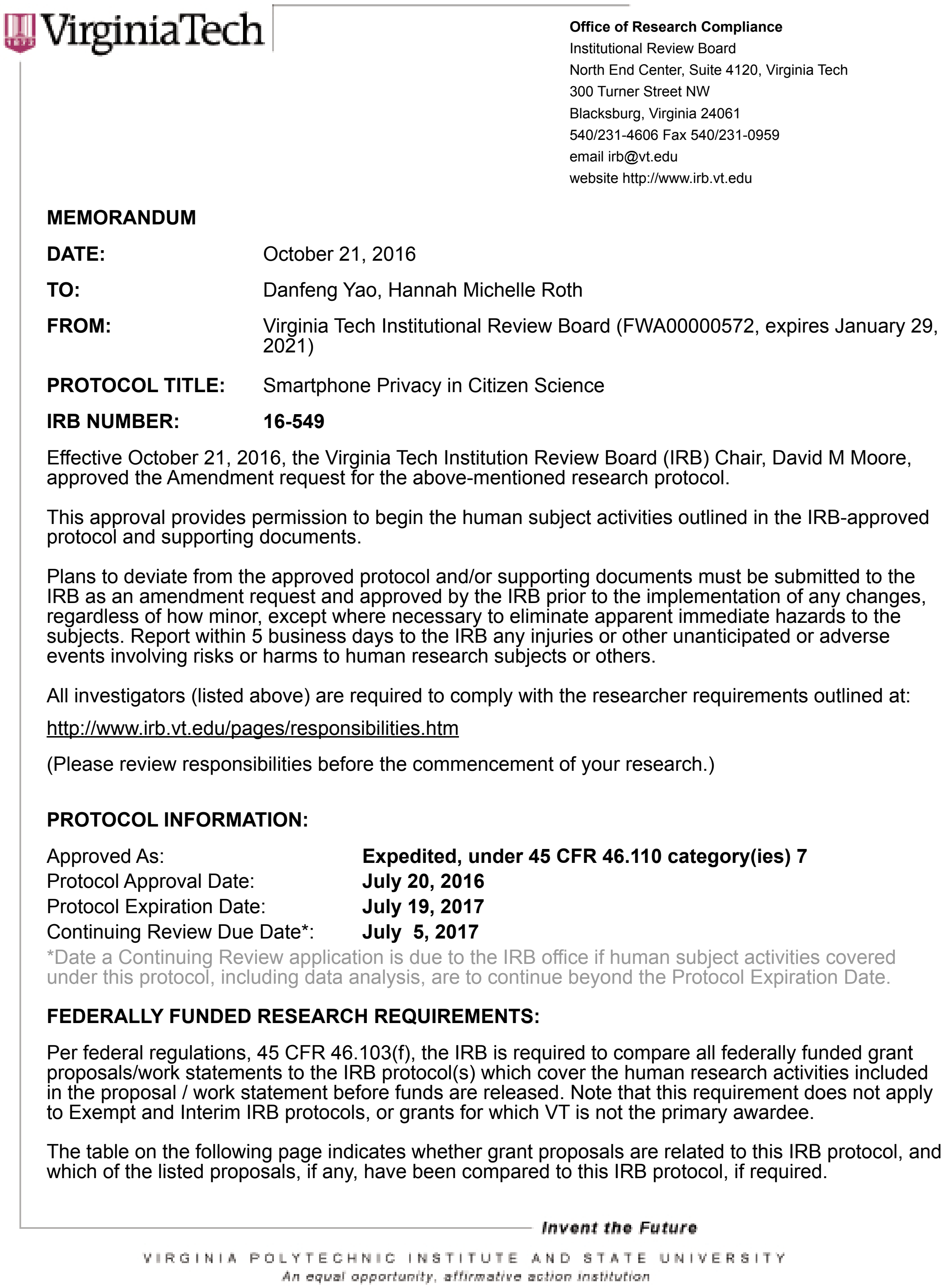
Appendix A User Study

A.1 Recruitment Announcement

I am conducting a user study as part of my Master's Thesis and am looking for participants for next week. The study will take 30 minutes per participant and you will be compensated $20 for your time. It will take place in KnowledgeWorks II in the CRC. The link to sign up is here: http://www.signupgenius.com/go/8050b4babae2da75-groupsensing. Thank you!

A.2 Participant Survey

A.3 IRB Approval Letter



VirginiaTech

Office of Research Compliance
Institutional Review Board
North End Center, Suite 4120, Virginia Tech
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-4606 Fax 540/231-0959
email irb@vt.edu
website http://www.irb.vt.edu

MEMORANDUM

DATE:	October 21, 2016
TO:	Danfeng Yao, Hannah Michelle Roth
FROM:	Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)
PROTOCOL TITLE:	Smartphone Privacy in Citizen Science
IRB NUMBER:	**16-549**

Effective October 21, 2016, the Virginia Tech Institution Review Board (IRB) Chair, David M Moore, approved the Amendment request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As:	**Expedited, under 45 CFR 46.110 category(ies) 7**
Protocol Approval Date:	**July 20, 2016**
Protocol Expiration Date:	**July 19, 2017**
Continuing Review Due Date*:	**July 5, 2017**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

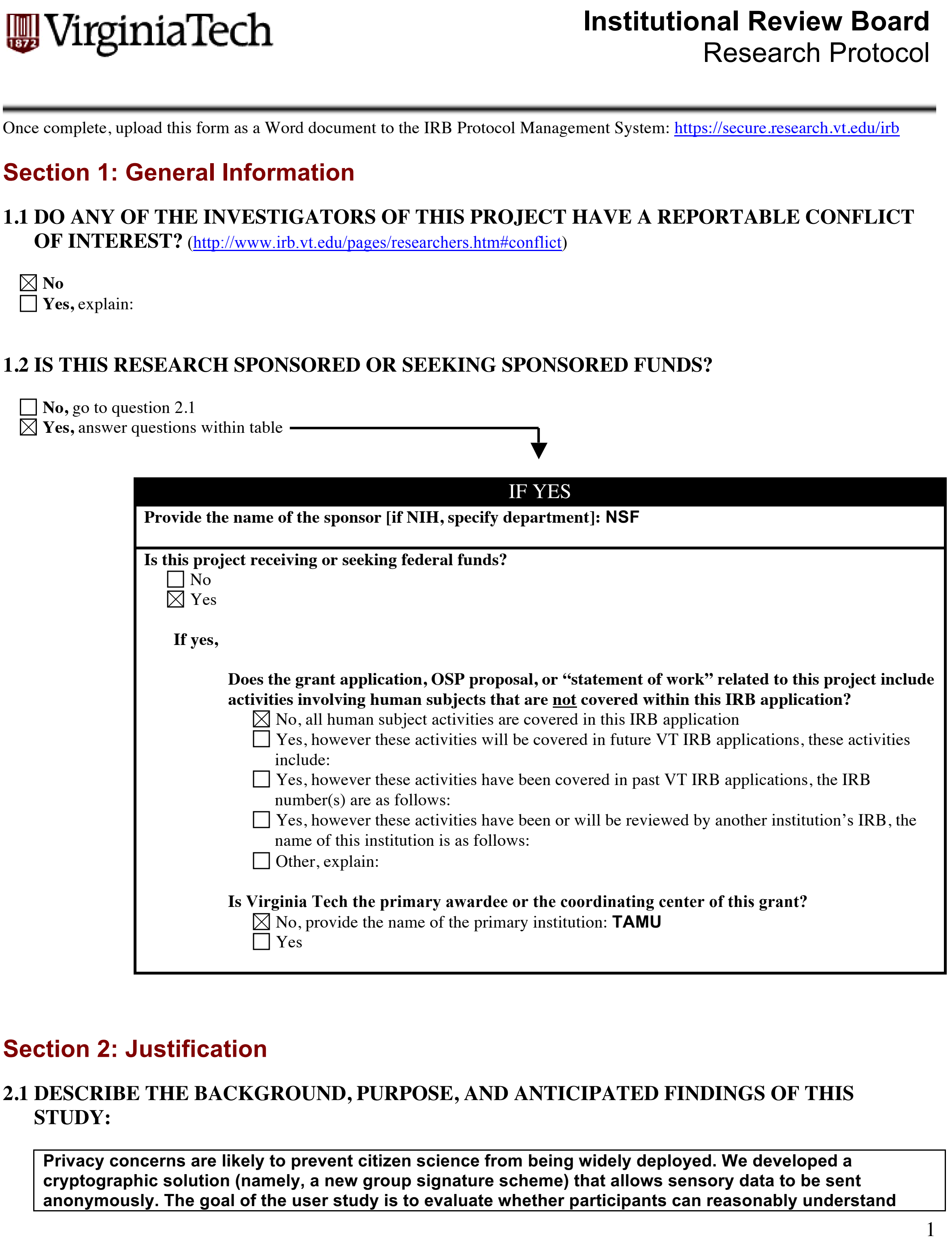
Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

A.4 IRB Research Protocol



VirginiaTech

Institutional Review Board
Research Protocol

Once complete, upload this form as a Word document to the IRB Protocol Management System: https://secure.research.vt.edu/irb

Section 1: General Information

1.1 DO ANY OF THE INVESTIGATORS OF THIS PROJECT HAVE A REPORTABLE CONFLICT OF INTEREST? (http://www.irb.vt.edu/pages/researchers.htm#conflict)

☒ **No**
□ **Yes,** explain:

1.2 IS THIS RESEARCH SPONSORED OR SEEKING SPONSORED FUNDS?

□ **No,** go to question 2.1
☒ **Yes,** answer questions within table

IF YES

Provide the name of the sponsor [if NIH, specify department]: **NSF**

Is this project receiving or seeking federal funds?
□ No
☒ Yes

If yes,

Does the grant application, OSP proposal, or "statement of work" related to this project include activities involving human subjects that are not covered within this IRB application?
☒ No, all human subject activities are covered in this IRB application
□ Yes, however these activities will be covered in future VT IRB applications, these activities include:
□ Yes, however these activities have been covered in past VT IRB applications, the IRB number(s) are as follows:
□ Yes, however these activities have been or will be reviewed by another institution's IRB, the name of this institution is as follows:
□ Other, explain:

Is Virginia Tech the primary awardee or the coordinating center of this grant?
☒ No, provide the name of the primary institution: **TAMU**
□ Yes

Section 2: Justification

2.1 DESCRIBE THE BACKGROUND, PURPOSE, AND ANTICIPATED FINDINGS OF THIS STUDY:

Privacy concerns are likely to prevent citizen science from being widely deployed. We developed a cryptographic solution (namely, a new group signature scheme) that allows sensory data to be sent anonymously. The goal of the user study is to evaluate whether participants can reasonably understand

group signature's privacy guarantees, and whether our group signatures would ease the privacy concern of citizen science participants.

Citizen science refers to "the collection and analysis of data relating to the natural world by members of the general public, typically as part of a collaborative project with professional scientists" [Wikipedia].

Group signature scheme refers to a type of cryptographic solutions that allow any group member to create a digital signature (i.e., digitally signing a message) without reveal his or her identity.

Privacy guarantee in our context means that sensitive personal information (such as location) of a participant contributing to citizen science should not be exposed to the public or citizen-science data servers.

Sensory data is the data collected from sensors (e.g., temperature) that are built in smartphones.

2.2 EXPLAIN WHAT THE RESEARCH TEAM PLANS TO DO WITH THE STUDY RESULTS:

For example - publish or use for dissertation

We intend to use the results to improve our design and publish the results.

Section 3: Recruitment

3.1 DESCRIBE THE SUBJECT POOL, INCLUDING INCLUSION AND EXCLUSION CRITERIA AND NUMBER OF SUBJECTS:

Examples of inclusion/exclusion criteria - gender, age, health status, ethnicity

Anyone who is 18 or older can participate in our study. We aim to have about 10 users.

3.2 WILL EXISTING RECORDS BE USED TO IDENTIFY AND CONTACT / RECRUIT SUBJECTS?

Examples of existing records - directories, class roster, university records, educational records

☒ **No,** go to question 3.3
☐ **Yes,** answer questions within table

IF YES

Are these records private or public?
☐ Public
☐ Private, describe the researcher's privilege to the records:

Will student, faculty, and/or staff records or contact information be requested from the University?
☐ No
☐ Yes, provide a description under Section 14 (Research Involving Existing Data) below.

3.3 DESCRIBE RECRUITMENT METHODS, INCLUDING HOW THE STUDY WILL BE ADVERTISED OR INTRODUCED TO SUBJECTS:

Put flyers on KnowledgeWorks II front door in CRC and also around the campus

3.4 PROVIDE AN EXPLANATION FOR CHOOSING THIS POPULATION:

Note: the IRB must ensure that the risks and benefits of participating in a study are distributed equitably among the general population and that a specific population is not targeted because of ease of recruitment.

Section 4: Consent Process

For more information about consent process and consent forms visit the following link: http://www.irb.vt.edu/pages/consent.htm

If feasible, researchers are advised and may be required to obtain signed consent from each participant unless obtaining signatures leads to an increase of risk (e.g., the only record linking the subject and the research would be the consent document and the principal risk would be potential harm resulting in a breach of confidentiality). Signed consent is typically not required for low risk questionnaires (consent is implied) unless audio/video recording or an in-person interview is involved. If researchers will not be obtaining signed consent, participants must, in most cases, be supplied with consent information in a different format (e.g., in recruitment document, at the beginning of survey instrument, read to participant over the phone, information sheet physically or verbally provided to participant).

4.1 CHECK ALL OF THE FOLLOWING THAT APPLY TO THIS STUDY'S CONSENT PROCESS:

☐ Verbal consent will be obtained from participants
☒ Signed consent will be obtained from participants
☐ Consent will be implied from the return of completed questionnaire. Note: The IRB recommends providing consent information in a recruitment document or at the beginning of the questionnaire (if the study only involves implied consent, skip to Section 5 below)
☐ Other, describe:

4.2 PROVIDE A GENERAL DESCRIPTION OF THE PROCESS THE RESEARCH TEAM WILL USE TO OBTAIN AND MAINTAIN INFORMED CONSENT:

We will store the signed consent forms

4.3 WHO, FROM THE RESEARCH TEAM, WILL BE OVERSEEING THE PROCESS AND OBTAINING CONSENT FROM SUBJECTS?

Danfeng Yao

4.4 WHERE WILL THE CONSENT PROCESS TAKE PLACE?

KnowledgeWork II at CRC, 2nd floor

4.5 DURING WHAT POINT IN THE STUDY PROCESS WILL CONSENTING OCCUR?

Note: unless waived by the IRB, participants must be consented before completing any study procedure, including screening questionnaires.

At the beginnig

4.6 IF APPLICABLE, DESCRIBE HOW THE RESEARCHERS WILL GIVE SUBJECTS AMPLE TIME TO REVIEW THE CONSENT DOCUMENT BEFORE SIGNING:

Note: typically applicable for complex studies, studies involving more than one session, or studies involving more of a risk to subjects.

☒ Not applicable

Section 5: Procedures

5.1 PROVIDE A STEP-BY-STEP THOROUGH EXPLANATION OF ALL STUDY PROCEDURES EXPECTED FROM STUDY PARTICIPANTS, INCLUDING TIME COMMITMENT & LOCATION:

1. The participant will read and sign the consent form.
2. The participant will be given an Android smartphone with the special app on it. The participant will deposit a photo ID during the study.
3. We will explain the objective of the user study, which is to evaluate the usability of our privacy-preserving crowdsensing app.
4. The participant will be asked to hang around KnowledgeWorks II (KWII) area (inside or outside) freely for about 15 minutes.
5. The participant is encouraged to use the phone during this period for their normal activities such as surfing the web, listening to music, searching the map, etc. We will ask participants not to enter any sensitive personal information to the phone.
6. The participant will be asked to return the phone to us. We will return the photo ID.
7. The participant will fill out a short survey on how he or she feels about the privacy and usability of the sensing app.

The total study will be around 30 minutes for each participant.

5.2 DESCRIBE HOW DATA WILL BE COLLECTED AND RECORDED:

The sensory data will be automatically collected by the Android app and sent to the server.
The paper survey will be manually entered by participants.

5.3 DOES THE PROJECT INVOLVE ONLINE RESEARCH ACTIVITES (INCLUDES ENROLLMENT, RECRUITMENT, SURVEYS)?

View the "Policy for Online Research Data Collection Activities Involving Human Subjects" at http://www.irb.vt.edu/documents/onlinepolicy.pdf

☒ **No,** go to question 6.1
□ **Yes,** answer questions within table

IF YES

Identify the service / program that will be used:
□ www.survey.vt.edu, go to question 6.1
□ SONA, go to question 6.1
□ Qualtrics, go to question 6.1
□ Center for Survey Research, go to question 6.1
□ Other

IF OTHER:
Name of service / program:
URL:
This service is…
□ Included on the list found at: http://www.irb.vt.edu/pages/validated.htm
□ Approved by VT IT Security
□ An external service with proper SSL or similar encryption (https://) on the login (if applicable) and all other data collection pages.
□ None of the above (note: only permissible if this is a collaborative project in which VT individuals are only responsible for data analysis, consulting, or recruitment)

Section 6: Risks and Benefits

6.1 WHAT ARE THE POTENTIAL RISKS (E.G., EMOTIONAL, PHYSICAL, SOCIAL, LEGAL, ECONOMIC, OR DIGNITY) TO STUDY PARTICIPANTS?

The risk of entering sensitive, personal, or identifiable information on the phone during the study is exposing the data to future phone users (including participants and researchers).

6.2 EXPLAIN THE STUDY'S EFFORTS TO REDUCE POTENTIAL RISKS TO SUBJECTS:

We will ask participants not to enter such information during the study. We will also clear the memory of the user-study app before another participant uses the phone.

6.3 WHAT ARE THE DIRECT OR INDIRECT ANTICIPATED BENEFITS TO STUDY PARTICIPANTS AND/OR SOCIETY?

To the participants, their understanding privacy tools is beneficial in the digital age.
To society, crowdsensing and citizen science will enable new scientific discovery, real-time environmental monitoring.

Section 7: Full Board Assessment

7.1 DOES THE RESEARCH INVOLVE MICROWAVES/X-RAYS, OR GENERAL ANESTHESIA OR SEDATION?

☒ **No**
☐ **Yes**

7.2 DO RESEARCH ACTIVITIES INVOLVE PRISONERS, PREGNANT WOMEN, FETUSES, HUMAN IN VITRO FERTILIZATION, OR INDIVIDUALS WITH MENTAL DISORDERS?

☒ **No,** go to question 7.3
☐ **Yes,** answer questions within table

IF YES

This research involves:
☐ Prisoners
☐ Pregnant women ☐ Fetuses ☐ Human in vitro fertilization
☐ Individuals with a mental disorder

7.3 DOES THIS STUDY INVOLVE MORE THAN MINIMAL RISK TO STUDY PARTICIPANTS?

Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily activities or during the performance of routine physical or psychological examinations or tests. Examples of research involving greater than minimal risk include collecting data about abuse or illegal activities. Note: if the project qualifies for Exempt review (http://www.irb.vt.edu/pages/categories.htm), it will not need to go to the Full Board.

☒ **No**
☐ **Yes**

IF YOU ANSWERED "YES" TO ***ANY ONE*** OF THE ABOVE QUESTIONS, 7.1, 7.2, OR 7.3, THE BOARD MAY REVIEW THE PROJECT'S APPLICATION MATERIALS AT ITS MONTHLY MEETING. VIEW THE FOLLOWING LINK FOR DEADLINES AND ADDITIONAL INFORMATION: http://www.irb.vt.edu/pages/deadlines.htm

Section 8: Confidentiality / Anonymity

For more information about confidentiality and anonymity visit the following link: http://www.irb.vt.edu/pages/confidentiality.htm

8.1 WILL PERSONALLY IDENTIFYING STUDY RESULTS OR DATA BE RELEASED TO ANYONE OUTSIDE OF THE RESEARCH TEAM?

For example – to the funding agency or outside data analyst, or participants identified in publications with individual consent

☒ **No**
□ **Yes,** to whom will identifying data be released?

8.2 WILL THE RESEARCH TEAM COLLECT AND/OR RECORD PARTICIPANT IDENTIFYING INFORMATION (E.G., NAME, CONTACT INFORMATION, VIDEO/AUDIO RECORDINGS)?

Note: if collecting signatures on a consent form, select "Yes."

□ **No,** go to question 8.3
☒ **Yes,** answer questions within table

IF YES
Describe if/how the study will utilize study codes: yes
If applicable, where will the key [i.e., linked code and identifying information document (for instance, John Doe = study ID 001)] **be stored and who will have access? The key will be stored in a locked cabinet in KnowledgeWorks II building. The data documents will be separately stored in a password-protected computer.** *Note: the key should be stored separately from subjects' completed data documents and accessibility should be limited.*
The IRB strongly suggests and may require that all data documents (e.g., questionnaire responses, interview responses, etc.) do not include or request identifying information (e.g., name, contact information, etc.) from participants. If you need to link subjects' identifying information to subjects' data documents, use a study ID/code on all data documents.

8.3 HOW WILL DATA BE STORED TO ENSURE SECURITY (E.G., PASSWORD PROTECTED COMPUTERS, ENCRYPTION) AND LIMITED ACCESS?

Examples of data - questionnaire, interview responses, downloaded online survey data, observation recordings, biological samples

We will take extreme precautions while managing the user-study data collected and analyzed in the experiments.

1.The sensory data and survey information collected during our experiments will be kept on the PI Yao's group server and students' workstations under strict access control policies.
2.The data collection will not record any sensitive personal data.
3.The server is physically located in the Virginia Tech computer science department technical staff office and is maintained by full-time dedicated technical staff.
4.The student workstations are in the building that has restricted after-work-hour access. The workstations are regularly patched by respective owners.
5.The accounts on the server and workstations are managed by PI Yao. Only active personnel related to this project will be given the access to the data.
6.The journal or conference publications will not contain any sensitive user data.

8.4 WHO WILL HAVE ACCESS TO STUDY DATA?

PI Yao and her students on this project.

8.5 DESCRIBE THE PLANS FOR RETAINING OR DESTROYING STUDY DATA:

The data will be kept on the group server. Currently, there is no plan for destroying the data.

8.6 DOES THIS STUDY REQUEST INFORMATION FROM PARTICIPANTS REGARDING ILLEGAL BEHAVIOR?

☒ **No,** go to question 9.1
☐ **Yes,** answer questions within table →

IF YES

Does the study plan to obtain a Certificate of Confidentiality?
☐ No
☐ Yes (Note: participants must be fully informed of the conditions of the Certificate of Confidentiality within the consent process and form)

For more information about Certificates of Confidentiality, visit the following link: *http://www.irb.vt.edu/pages/coc.htm*

Section 9: Compensation

For more information about compensating subjects, visit the following link: http://www.irb.vt.edu/pages/compensation.htm

9.1 WILL SUBJECTS BE COMPENSATED FOR THEIR PARTICIPATION?

☐ **No,** go to question 10.1
☒ **Yes,** answer questions within table →

IF YES

What is the amount of compensation? $20 for 30 minutes

Will compensation be prorated?
☐ Yes, please describe:
☒ No, explain why and clarify whether subjects will receive full compensation if they withdraw from the study? **Withdrawl participant will receive half of the compensation.**

Unless justified by the researcher, compensation should be prorated based on duration of study participation. Payment must not be contingent upon completion of study procedures. In other words, even if the subject decides to withdraw from the study, he/she should be compensated, at least partially, based on what study procedures he/she has completed.

Section 10: Audio / Video Recording

For more information about audio/video recording participants, visit the following link: http://www.irb.vt.edu/pages/recordings.htm

10.1 WILL YOUR STUDY INVOLVE VIDEO AND/OR AUDIO RECORDING?

☒ **No,** go to question 11.1
□ **Yes,** answer questions within table

IF YES
This project involves: □ Audio recordings only □ Video recordings only □ Both video and audio recordings
Provide compelling justification for the use of audio/video recording:
How will data within the recordings be retrieved / transcribed?
How and where will recordings (e.g., tapes, digital data, data backups) be stored to ensure security?
Who will have access to the recordings?
Who will transcribe the recordings?
When will the recordings be erased / destroyed?

Section 11: Research Involving Students

11.1 DOES THIS PROJECT INCLUDE STUDENTS AS PARTICIPANTS?

□ **No,** go to question 12.1
☒ **Yes,** answer questions within table

IF YES
Does this study involve conducting research with students of the researcher? ☒ No □ Yes, describe safeguards the study will implement to protect against coercion or undue influence for participation: *Note: if it is feasible to use students from a class of students not under the instruction of the researcher, the IRB recommends and may require doing so.*
Will the study need to access student records (e.g., SAT, GPA, or GRE scores)? ☒ No □ Yes

11.2 DOES THIS PROJECT INCLUDE ELEMENTARY, JUNIOR, OR HIGH SCHOOL STUDENTS?

☒ **No,** go to question 11.3
□ **Yes,** answer questions within table

IF YES
Will study procedures be completed during school hours?

☐ No
☐ Yes

If yes,

Students not included in the study may view other students' involvement with the research during school time as unfair. Address this issue and how the study will reduce this outcome:

Missing out on regular class time or seeing other students participate may influence a student's decision to participate. Address how the study will reduce this outcome:

Is the school's approval letter(s) attached to this submission?
☐ Yes
☐ No, project involves Montgomery County Public Schools (MCPS)
☐ No, explain why:

You will need to obtain school approval (if involving MCPS, click here: http://www.irb.vt.edu/pages/mcps.htm). Approval is typically granted by the superintendent, principal, and classroom teacher (in that order). Approval by an individual teacher is insufficient. School approval, in the form of a letter or a memorandum should accompany the approval request to the IRB.

11.3 DOES THIS PROJECT INCLUDE COLLEGE STUDENTS?

☐ **No,** go to question 12.1
☒ **Yes,** answer questions within table

IF YES

Some college students might be minors. Indicate whether these minors will be included in the research or actively excluded:
☐ Included
☒ Actively excluded, describe how the study will ensure that minors will not be included:

Will extra credit be offered to subjects?
☒ No
☐ Yes

If yes,

What will be offered to subjects as an equal alternative to receiving extra credit without participating in this study?

Include a description of the extra credit (e.g., amount) to be provided within question 9.1 ("IF YES" table)

Section 12: Research Involving Minors

12.1 DOES THIS PROJECT INVOLVE MINORS (UNDER THE AGE OF 18 IN VIRGINIA)?

Note: age constituting a minor may differ in other States.

☒ **No,** go to question 13.1
☐ **Yes,** answer questions within table

IF YES

Does the project reasonably pose a risk of reports of current threats of abuse and/or suicide?

□ No
□ Yes, thoroughly explain how the study will react to such reports:

Note: subjects and parents must be fully informed of the fact that researchers must report threats of suicide or suspected/reported abuse to the appropriate authorities within the Confidentiality section of the Consent, Assent, and/or Permission documents.

Are you requesting a waiver of parental permission (i.e., parent uninformed of child's involvement)?

□ No, **both** parents/guardians will provide their permission, if possible.
□ No, **only one** parent/guardian will provide permission.
□ Yes, describe below how your research meets **<u>all</u>** of the following criteria (A-D):

Criteria A - The research involves no more than minimal risk to the subjects:
Criteria B - The waiver will not adversely affect the rights and welfare of the subjects:
Criteria C - The research could not practicably be carried out without the waiver:
Criteria D - (Optional) Parents will be provided with additional pertinent information after participation:

Is it possible that minor research participants will reach the legal age of consent (18 in Virginia) while enrolled in this study?

□ No
□ Yes, will the investigators seek and obtain the legally effective informed consent (in place of the minors' previously provided assent and parents' permission) for the now-adult subjects for any ongoing interactions with the subjects, or analysis of subjects' data? If yes, explain how:

For more information about minors reaching legal age during enrollment, visit the following link: http://www.irb.vt.edu/pages/assent.htm

*The procedure for obtaining assent from minors and permission from the minor's guardian(s) must be described in **Section 4** (Consent Process) of this form.*

Section 13: Research Involving Deception

For more information about involving deception in research and for assistance with developing your debriefing form, visit our website at http://www.irb.vt.edu/pages/deception.htm

13.1 DOES THIS PROJECT INVOLVE DECEPTION?

☒ **No,** go to question 14.1
□ **Yes,** answer questions within table

IF YES

Describe the deception:

Why is the use of deception necessary for this project?

Describe the debriefing process:

Provide an explanation of how the study meets <u>all</u> the following criteria (A-D) for an alteration of consent:

Criteria A - The research involves no more than minimal risk to the subjects:
Criteria B - The alteration will not adversely affect the rights and welfare of the subjects:

Criteria C - The research could not practicably be carried out without the alteration: Criteria D - (Optional) Subjects will be provided with additional pertinent information after participation (i.e., debriefing for studies involving deception): *By nature, studies involving deception cannot provide subjects with a complete description of the study during the consent process; therefore, the IRB must allow (by granting an alteration of consent) a consent process which does not include, or which alters, some or all of the elements of informed consent.*
The IRB requests that the researcher use the title "Information Sheet" instead of "Consent Form" on the document used to obtain subjects' signatures to participate in the research. This will adequately reflect the fact that the subject cannot fully consent to the research without the researcher fully disclosing the true intent of the research.

Section 14: Research Involving Existing Data

14.1 WILL THIS PROJECT INVOLVE THE COLLECTION OR STUDY/ANALYSIS OF EXISTING DATA DOCUMENTS, RECORDS, PATHOLOGICAL SPECIMENS, OR DIAGNOSTIC SPECIMENS?

Please note: it is not considered existing data if a researcher transfers to Virginia Tech from another institution and will be conducting data analysis of an on-going study.

☒ **No,** you are finished with the application
□ **Yes,** answer questions within table

IF YES
From where does the existing data originate?
Provide a detailed description of the existing data that will be collected or studied/analyzed:
Is the source of the data public? □ No, continue with the next question □ Yes, you are finished with this application
Will any individual associated with this project (internal or external) have access to or be provided with existing data containing information which would enable the identification of subjects: ▪ **Directly** (e.g., by name, phone number, address, email address, social security number, student ID number), or ▪ **Indirectly through study codes** even if the researcher or research team does not have access to the master list linking study codes to identifiable information such as name, student ID number, etc or ▪ **Indirectly through the use of information that could reasonably be used in combination to identify an individual** (e.g., demographics) □ No, collected/analyzed data will be completely de-identified □ Yes, **If yes,** *Research will not qualify for exempt review; therefore, if feasible, written consent must be obtained from individuals whose data will be collected / analyzed, unless this requirement is waived by the IRB.* **Will written/signed or verbal consent be obtained from participants prior to the analysis of collected data?** -select one-

This research protocol represents a contract between all research personnel associated with the project, the University, and federal government; therefore, must be followed accordingly and kept current.

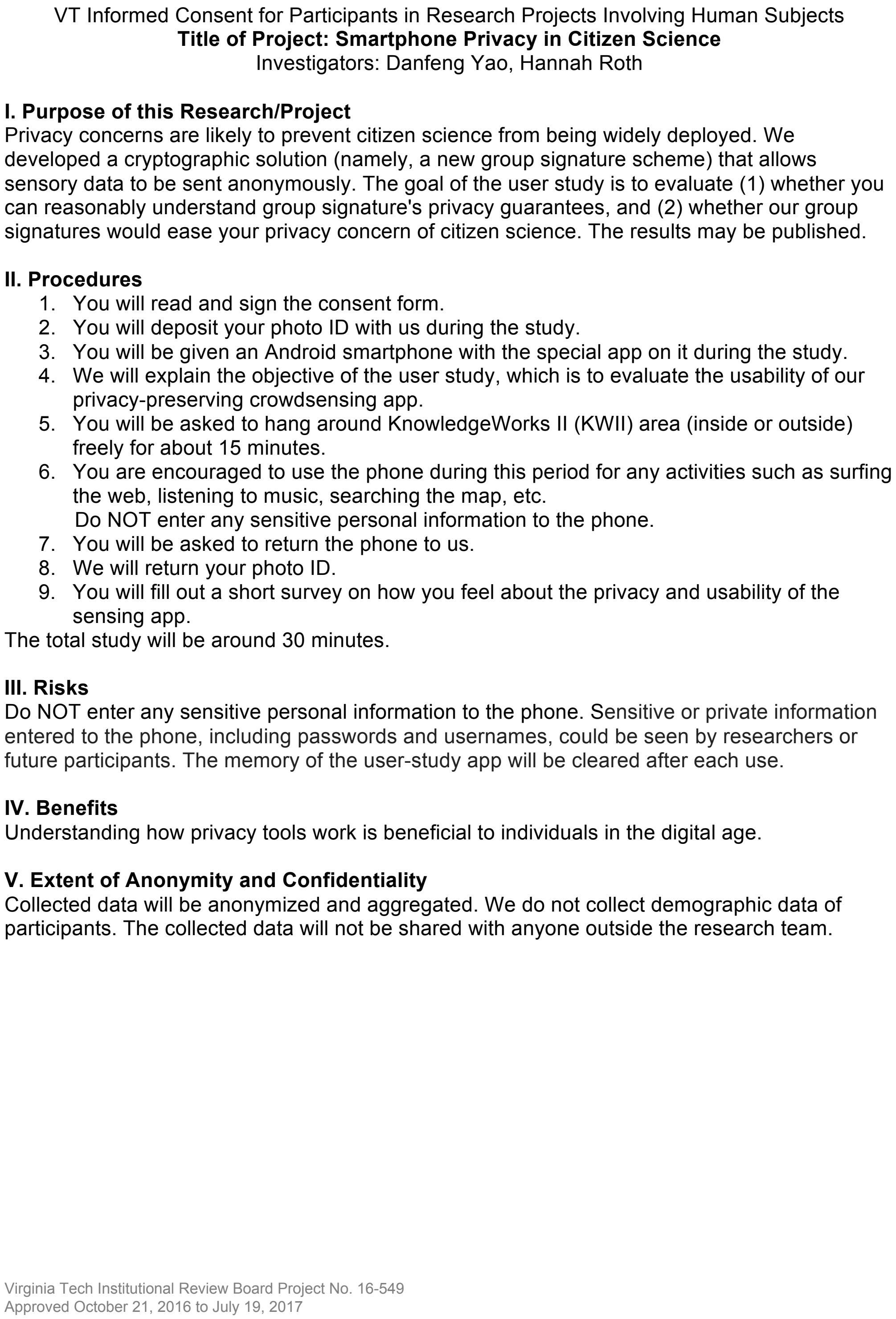
Proposed modifications must be approved by the IRB prior to implementation except where necessary to eliminate apparent immediate hazards to the human subjects.

Do not begin human subjects activities until you receive an IRB approval letter via email.

It is the Principal Investigator's responsibility to ensure all members of the research team who interact with research subjects, or collect or handle human subjects data have completed human subjects protection training prior to interacting with subjects, or handling or collecting the data.

----------END----------

A.5 Consent Form



VT Informed Consent for Participants in Research Projects Involving Human Subjects

Title of Project: Smartphone Privacy in Citizen Science

Investigators: Danfeng Yao, Hannah Roth

I. Purpose of this Research/Project

Privacy concerns are likely to prevent citizen science from being widely deployed. We developed a cryptographic solution (namely, a new group signature scheme) that allows sensory data to be sent anonymously. The goal of the user study is to evaluate (1) whether you can reasonably understand group signature's privacy guarantees, and (2) whether our group signatures would ease your privacy concern of citizen science. The results may be published.

II. Procedures

1. You will read and sign the consent form.
2. You will deposit your photo ID with us during the study.
3. You will be given an Android smartphone with the special app on it during the study.
4. We will explain the objective of the user study, which is to evaluate the usability of our privacy-preserving crowdsensing app.
5. You will be asked to hang around KnowledgeWorks II (KWII) area (inside or outside) freely for about 15 minutes.
6. You are encouraged to use the phone during this period for any activities such as surfing the web, listening to music, searching the map, etc.
   Do NOT enter any sensitive personal information to the phone.
7. You will be asked to return the phone to us.
8. We will return your photo ID.
9. You will fill out a short survey on how you feel about the privacy and usability of the sensing app.

The total study will be around 30 minutes.

III. Risks

Do NOT enter any sensitive personal information to the phone. Sensitive or private information entered to the phone, including passwords and usernames, could be seen by researchers or future participants. The memory of the user-study app will be cleared after each use.

IV. Benefits

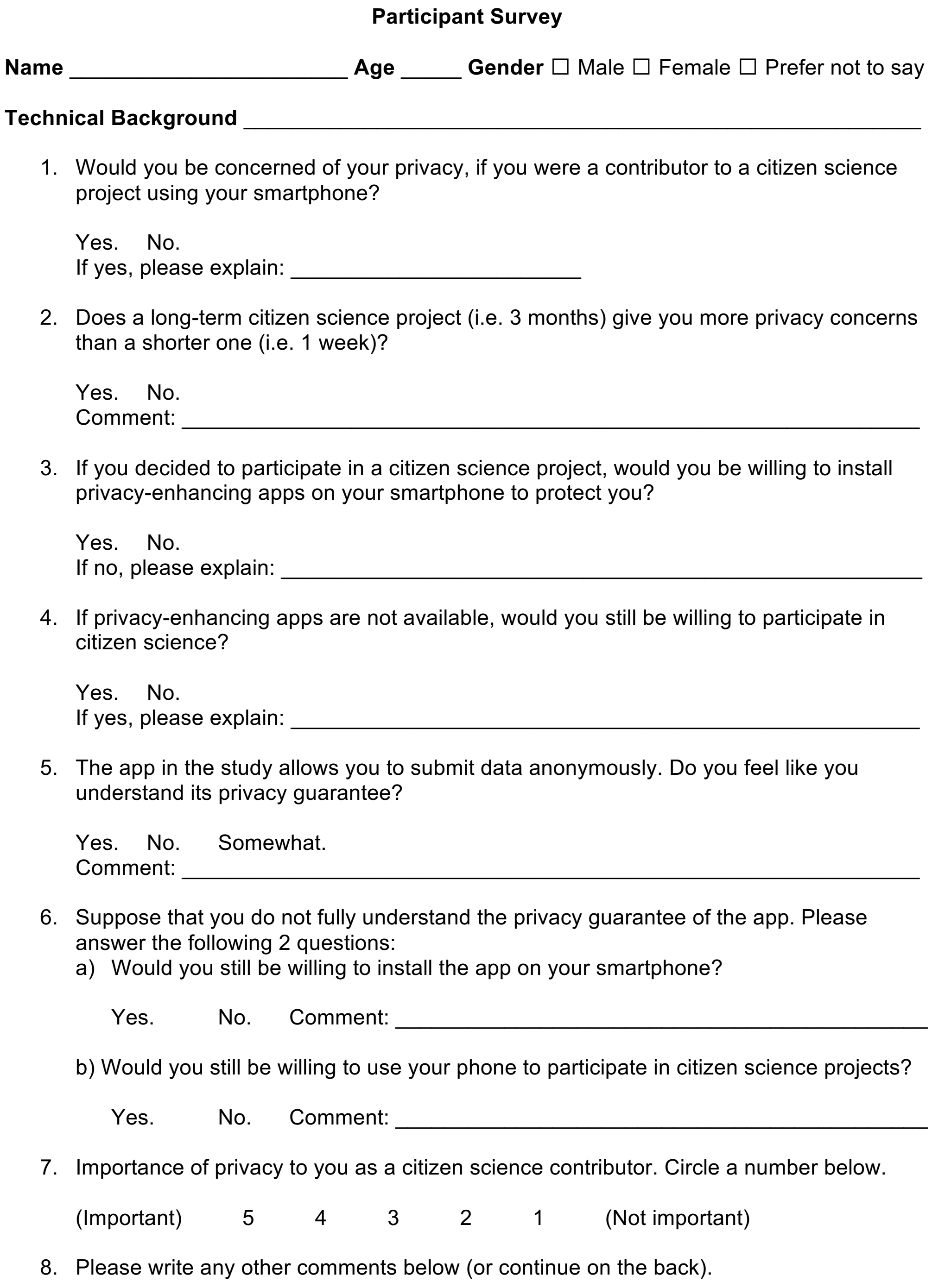
Understanding how privacy tools work is beneficial to individuals in the digital age.

V. Extent of Anonymity and Confidentiality

Collected data will be anonymized and aggregated. We do not collect demographic data of participants. The collected data will not be shared with anyone outside the research team.

Virginia Tech Institutional Review Board Project No. 16-549
Approved October 21, 2016 to July 19, 2017

A.6 Participant Survey Data



Participant Survey

Name ________________________ **Age** _____ **Gender** ☐ Male ☐ Female ☐ Prefer not to say

Technical Background ____

1. Would you be concerned of your privacy, if you were a contributor to a citizen science project using your smartphone?

   Yes. No.
   If yes, please explain: _________________________

2. Does a long-term citizen science project (i.e. 3 months) give you more privacy concerns than a shorter one (i.e. 1 week)?

   Yes. No.
   Comment: ____

3. If you decided to participate in a citizen science project, would you be willing to install privacy-enhancing apps on your smartphone to protect you?

   Yes. No.
   If no, please explain: ____

4. If privacy-enhancing apps are not available, would you still be willing to participate in citizen science?

   Yes. No.
   If yes, please explain: ____

5. The app in the study allows you to submit data anonymously. Do you feel like you understand its privacy guarantee?

   Yes. No. Somewhat.
   Comment: ____

6. Suppose that you do not fully understand the privacy guarantee of the app. Please answer the following 2 questions:
   a) Would you still be willing to install the app on your smartphone?

      Yes. No. Comment: ____

   b) Would you still be willing to use your phone to participate in citizen science projects?

      Yes. No. Comment: ____

7. Importance of privacy to you as a citizen science contributor. Circle a number below.

   (Important) 5 4 3 2 1 (Not important)

8. Please write any other comments below (or continue on the back).