A Privacy Assistant for the Internet of Things

Norman Sadeh
Martin Degeling
Anupam Das
Aerin Shikun Zhang
Alessandro Acquisti
Lujo Bauer
Lorrie Cranor
Anupam Datta
Daniel Smullen
Carnegie Mellon University
Pittsburgh, PA 15213, USA
sadeh@cs.cmu.edu
degeling@cs.cmu.edu
anupamd@cs.cmu.edu
acquisti@andrew.cmu.edu
lbauer@cmu.edu
lorrie@cs.cmu.edu
danupam@cmu.edu
dsmullen@cs.cmu.edu

Abstract
The rapid development of devices and sensors connected to the Internet of Things (IoT) is resulting in widespread collection of personally identifiable information. Since these devices and sensors are often small, embedded, and with no user interface, many people are unaware of the data collection, and are therefore unable to take control of their privacy in these contexts. Our work highlights the idea of an infrastructure which notifies users of data collection in heterogeneous IoT environments. We envision privacy assistants residing on individual smartphones, which selectively inform their owners about data collection in their vicinity. These assistants make invisible data collection visible, support users in configuring privacy settings, where available.

Author Keywords
Internet of Things; Intelligent assistants; Privacy assistant; Privacy profiles

ACM Classification Keywords
K.4.1 [Computers and Society]: Privacy

Introduction
With the emergence of the IoT and a data-centric economy, a growing number of products, services, and business processes rely on the collection and processing of user data. People are increasingly confronted with an unmanageable

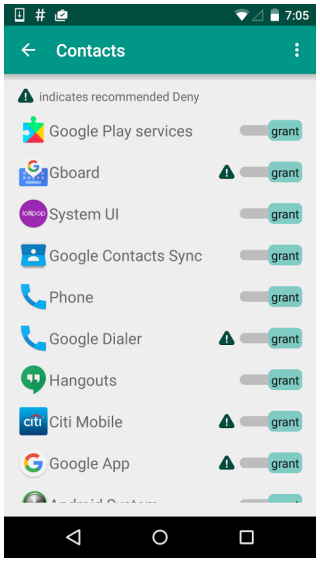

Figure 1: Screenshot of the Mobile Privacy Assistant application, making suggestions for Android application permissions

number of privacy decisions, and yet more situations where their privacy-related preferences go unheeded. While there is ample evidence that people care about their privacy [3], research shows that they are simply overwhelmed by the amount of information they would have to read and settings they are expected to configure [6] (e.g., smartphone settings, browser settings, smart thermostat settings, activity bracelet settings). What is needed is a new, more scalable paradigm that empowers users to regain control over their data. We have developed and piloted personalized privacy assistants, which are intelligent assistants capable of learning the privacy preferences of individual users over time. These assistants selectively inform users about data collection and use practices determined to be relevant to their preferences and concerns, helping them discover and configure available settings to enact these preferences.

Based on prior work on supporting users to manage permissions on their Android phones (see Fig. 1) we are now seeking to extend this functionality in support of Internet of Things scenarios. To do so, we have created an infrastructure along with protocols to enable the discovery of relevant IoT resources by privacy assistants. Following discovery, relevant elements of IoT resources' privacy policies and any available privacy settings are made available to users. We have also begun to develop models of individuals' privacy preferences and expectations, including notification preferences as they pertain to a growing number of emerging IoT scenarios.

Initial versions of the infrastructure have been deployed at both UC Irvine (UCI) [5], and Carnegie Mellon University (CMU) [1]. As we develop this infrastructure, we are placing emphasis on maximally promoting interoperability and usability. This means building interfaces that make this technology easy to deploy, as well as creating reusable templates and protocols. These templates enable device manufacturers to rapidly customize and publish device specifications with descriptions that can be semi-automatically interpreted to populate IoT resource registries. We are also exploring opportunities to go beyond just privacy, supporting the location-based discovery of IoT resources (e.g., apps to help navigate through buildings, to find printers or lounge spaces) and opportunities to support novel emerging standards (e.g., Manufacturing Usage Description).

Personalized Privacy Assistants for IoT

While people care about their privacy, the demands associated with reading privacy policies and configuring the diverse and growing collection of relevant privacy settings have become unrealistically high. Privacy assistants can help users to stay informed about privacy practices by selectively informing them about the specific practices they are likely to be concerned about and/or may not be expecting. They can also simplify the discovery of configurable privacy settings and help users to choose which settings to use (e.g., opting out or opting in to different data collection and use practices, enabling data aggregation or santization of features). Ideally, we would like privacy assistants to limit their interaction with users to only a few times during the day, and to otherwise be invisible. In IoT scenarios, where users may not even be aware of the presence of devices or services collecting and using their data, privacy assistants require the development of an infrastructure along with protocols that support dynamic discovery of resources and their policies.

The viability of privacy assistants has been proven in the context of mobile app permissions [2] and our privacy assistant for Android users has been released to the public. We have also developed a Privacy Assistant for the IoT that, so far, is capable of discovering and interacting with registries

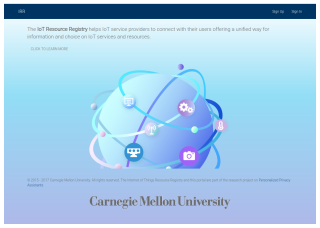

Figure 2: Frontpage of an IoT Resource Registry (IRR)

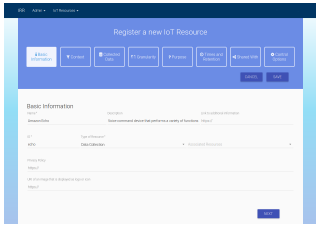

Figure 3: Registration form for adding IoT resources to the registry

that contain descriptions of IoT resources, their privacy policies, and their configurable privacy settings. As part of our ongoing research in this area, we are collecting and mining data about individuals' privacy preferences and expectations as they pertain to a growing collection of IoT scenarios, with the objective of building representative models to optimize the behavior of our privacy assistants.

IoT Infrastructure to Support the Discovery of IoT Resources, their Privacy Policies, and Settings

An important requirement for privacy assistants to work in IoT scenarios is the ability to dynamically discover IoT resources and the relevant elements of their privacy practices. We have developed an infrastructure that revolves around the deployment of IoT resource registries (see Fig. 2). Each registry is owned by a user or organization and is deployed with respect to a particular location (e.g., a registry in someone's home, an office building, in a city, or possibly across a larger area). We envision many tightly controlled registries in homes and corporate environments, but we intend to explore more open (and potentially less trustworthy) models of control as well. For example, a public registry which enables people in a city to advertise and/or identify relevant resources or services – with the risk of being exposed to spam and the need to deploy solutions that help mitigate this risk. The protocol we have developed enables users to dynamically discover resource registries in their vicinity and expose configurable privacy settings.

To support communication between the components of our infrastructure, we have developed a policy schema based on JSON that allows IoT resource owners to define a wide variety of sensors and associated data collection practices. It is based on concepts that are commonly included in privacy policies on the web [4], such as the data collection context (e.g., physical locality and responsibility), the purpose of the data collection effort (e.g., in support of a specific service), whether and with whom the collected data is shared, how long data is retained, and what choices the data subjects have with respect to management. In addition, the schema allows the specification of IoT-related information, such as descriptions of sensor types, and the granularity of captured sensory information.

We have implemented 2 different ways for resources owners to define the privacy practices of their IoT resources: (1) instantiating and customizing a reusable template associated with an existing off-the-shelf IoT device. (2) navigating a series of screens with drop-down menus and answering a series of basic questions (see Fig. 3). We plan to extend the functionality of our infrastructure in support of emerging IoT development platforms (e.g., in the context of Android Things) to make it as easy as possible for developers to publish templates for novel IoT devices and resources they create.

For the time being, we try to avoid assumptions about how detailed resource descriptions should be and do not assume that every resource owner would necessarily want to advertise the presence of their resources. Instead, our focus is on making it as easy as possible for resource owners to declare the presence of IoT resources and their privacy practices, if they so desire. In jurisdictions and contexts where notification and/or consent is required, we want to make it as easy as possible for resource owners to maintain compliance.

We also have come to realize that it is a good idea to open our infrastructure beyond privacy, supporting the discovery of resources and their functionality making our assistant an "IoT Assistant" (IoTA, see Fig. 4). The IoTA supports the discovery of IoT-connected apps available in different spaces to help users make the most of the IoT resources

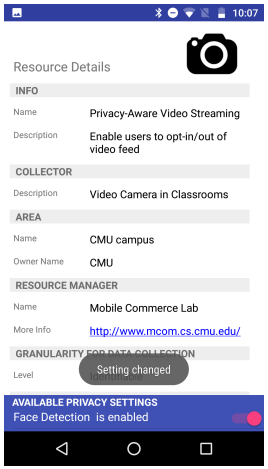

Figure 4: IoT Assistant

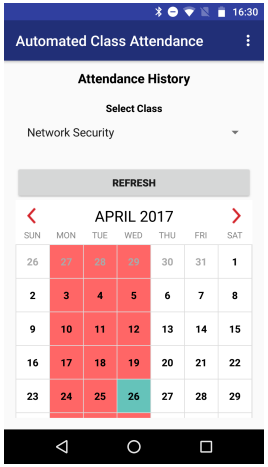

Figure 5: Class Attendance App

serving a given location (e.g., services for indoor navigation or finding local amenities).

Applications Using the Infrastructure

So far, we have developed three mobile applications that make use of IoT resources. Two are available on CMU's campus (friend finder and automated class attendance) and one on the UCI campus (indoor navigator) and offer a similar functionality. Both campuses are equipped with indoor location tracking services, using WiFi access points and Bluetooth beacons. WiFi access points offer a coarse grained location (e.g., imprecise location, distinguished by building, wing, or hallway). Fine grained location is based on Bluetooth beacons. Depending on the number and density of beacons that are deployed in a given area, Bluetooth beacons can be used for location detection precise enough to distinguish individual rooms. Pre-registered users of the location service can be located via WiFi access points using mobile phones. Bluetooth tracking requires a location service on a smartphone to scan for nearby Bluetooth beacons. In our deployment, the IoT Assistant notifies users about the availability of apps which use these location services. For example, the location sharing app enables users to share their location with friends, providing settings for location granularity. Additional apps may make use of the location tracking service and share that infrastructure. To simplify user interaction with the tracking system, allowing them to configure location granularity or arbitrarily disable location tracking at any given time, the IoTA exposes simple control options. When users configure these options, their settings are automatically sent to a policy enforcement server that was previously specified as part of the location service's resource registration in the IRR – the privacy policy associated with the resource on the IRR specifies what and how users may configure the resource.

A second application we have implemented uses facial recognition technology to automatically detect and record attendance for university lectures (see Fig. 5, described in detail in [1]). Participants register their face with the application using their phone. Once registered, as they walk past a camera when entering the lecture room, their attendance is recorded. Lecturers and students may use these records to keep track of who attended the class. Similar to applications that use the location tracking service, users can use the IoTA to change their privacy settings for the attendance tracking. This allows users to opt-in or out of the tracking, during the course of the semester. The application uses the same policy enforcement server as the location tracking service, which controls the facial detection processing service that the attendance tracking relies on. Each of these services may be part of shared infrastructure used to support other applications where facial recognition is required.

Conclusion

We envision that IoT Assistants will help users navigate through the complex process of making privacy decisions in IoT environments. This is particularly important as the emergence of the Internet of Things is making it increasingly complicated for individuals to keep track of what data is collected, when, by whom, and manage their preferences. We have previously shown that suggestions based on privacy profiles are effective in helping Android users to manage permissions on their phones. Now we have created an infrastructure that can provide similar notification, support and control for users over IoT-connected resources and applications. Once IoT instances of data collection are registered in an IoT Resource Registry, our IoTA is able to judiciously notify and inform users about the privacy practices, call attention to esoteric applications making use of these resources, expose relevant settings, and aid in personalized configuration.

Acknowledgements

This research has been supported in part by DARPA and the Air Force Research Laboratory under agreement number FA8750-15-2-0277 and by the National Science Foundation under grant SBE-1513957. The US Government is authorized to reproduce and distribute reprints for Governmental purposes not withstanding any copyright notation thereon. Additional support has also been provided by Google. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, the Air Force Research Laboratory, the NSF, Google, or the US Government.

REFERENCES

1. Anupam Das, Martin Degeling, Junjue Wang, Xiaoyou Wang, Mahadev Satyanarayanan, and Norman Sadeh. 2017. A Privacy-aware Infrastructure for using Facial Recognition. In *Workshop The Bright and Dark Sides of Computer Vision: Challenges and Opportunities for Privacy and Security (CV-COPS 2017)*. `http://www.cs.cmu.edu/~anupamd/paper/CV-COPS-2017.pdf`.
2. Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun (Aerin) Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. 2016. Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions. 27–41. `https://www.usenix.org/conference/soups2016/technical-sessions/presentation/liu`
3. Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Cranor, and Norman Sadeh. 2017. Privacy Expectations and Preferences in an IoT World. In *SOUPS 2017*.
4. Alessandro Oltramaria, Dhivya Piraviperumala, Florian Schauba, Shomir Wilsona, Norman Sadeha, and Joel Reidenbergb. 2016. PrivOnto: a Semantic Framework for the Analysis of Privacy Policies. *Semantic Web Journal* (2016).
5. Primal Pappachan, Martin Degeling, Roberto Yus, Anupam Das, Shikun Zhang, Sruti Bhagavatula, Pardis Emami Naeini, Sharad Mehrotra, Norman Sadeh, Alfred Kobsa, Lujo Bauer, and Nalini Venkatasubramanian. 2017. Towards Expressing Privacy Requirements in Smart Buildings. Atlanta, Georgia, USA. `https://www.cs.cmu.edu/~anupamd/paper/IoTCA2017.pdf`.
6. Ashwini Rao, Florian Schaub, Norman Sadeh, Alessandro Acquisti, and Ruogu Kang. 2016. Expecting the Unexpected: Understanding Mismatched Privacy Expectations Online. 77–96. `https://www.usenix.org/conference/soups2016/technical-sessions/presentation/rao`