HCI in Business: A Collaboration with Academia in IoT Privacy

Richard Chow[1](✉), Serge Egelman[2], Raghudeep Kannavara[1],
Hosub Lee[3], Suyash Misra[3], and Edward Wang[4]

[1] Software and Services Group, Intel Corporation, Santa Clara, USA
{richard.chow, raghudeep.kannavara}@intel.com
[2] Electrical Engineering and Computer Sciences,
University of California, Berkeley, USA
egelman@cs.berkeley.edu
[3] Donald Bren School of Information and Computer Sciences,
University of California, Irvine, USA
{hosubl, suyashm}@uci.edu
[4] Electrical Engineering Department, University of Washington,
Seattle, WA, USA
ejaywang@uw.edu

Abstract. The Internet of Things (IoT) integrates communication capabilities into physical objects to create a ubiquitous and multi-modal network of information and computing resources. The promise and pervasiveness of IoT ecosystems has lured many companies, including Intel, to devote resources and engineers to participate in the future of IoT. This paper describes a joint effort from Intel and two collaborators from academia to address the problem of IoT privacy.

Keywords: Internet of things · Privacy · Notifications · Trust · Information disclosure · Design · User experience

1 Introduction

The Internet of Things (IoT) is the next great technology challenge that integrates communication capabilities into physical objects to create a ubiquitous and multi-modal network of information and computing resources. While IoT stands at the cusp of new technological possibilities, coupled with it are notions of tracking, surveillance, and concerns about personal privacy. This paper describes three cooperative but loosely coupled research efforts in the area of IoT privacy undertaken by Intel, UC Berkeley, and UC Irvine.

According to Weiser [11], *"The problem, while often couched in terms of privacy, is really one of control. If the computational system is invisible as well as extensive, it becomes hard to know what is controlling what, what is connected to what, where*

Edward Wang: work done while at Intel.

F.F.-H. Nah and C.-H. Tan (Eds.): HCIB 2015, LNCS 9191, pp. 679–687, 2015.
DOI: 10.1007/978-3-319-20895-4_63

information is flowing, how it is being used…and what are the consequences of any given action." Weiser is discussing ubiquitous computing over a decade ago, but he may as well have been talking about IoT today. In the first part of this paper, we describe research performed by Intel that focuses on this invisible nature of IoT. On our computers, we have at least a semblance of control because we can in principle determine what applications are running and what data they are collecting. For IoT, traditional methods of control are largely absent. Hence, we describe below communication mechanisms between the user and the IoT ecosystem that can inform the user of data collection practices (and perhaps offer the option to block these practices).

Even assuming user control and transparency of purpose and function of IoT devices in the environment, there remain thorny user interface issues. For example, when nearby devices are collecting personal data, under what conditions is the user notified? These could be devices in the environment, like street cameras or WiFi access points that track MAC addresses, or mobile devices belonging to others, like smart phones or wearables recording audio. Due to the expected density of IoT devices, communicating all privacy practices may result in notification fatigue. Privacy preferences vary widely among individuals [8], and data collection that may be worthy of notification for one person may be ignored by another. Also, since user privacy preferences are complex, it is a challenge for users to define them [7]. In the second part of this paper we describe research from UC Irvine which investigates user privacy preferences in IoT. (Figs. 1, 2 and 3).

While UC Irvine researchers are concentrating on the problem of when a user wants to be notified, the problem of how to notify a user is equally complex. As most IoT devices are "interface-less" entities that are meant to function unobtrusively, this posits a unique challenge to the design of the human-computer interface. We describe in the third part of this paper how UC Berkeley researchers have demonstrated the difficulty in relying on traditional IoT notification mechanisms such as LED lights and also how to make a modern IoT notification system more meaningful through a technique based on crowd-sourcing.

2 Intel: System Infrastructure

The Intel authors have been looking at some of the infrastructure problems related to IoT devices. Specifically, how to determine what IoT devices are nearby, how to get information on what data they are collecting, and how to deliver privacy notifications to the end user. For example, Alice has installed surveillance cameras in her home. She would like her guests to know about the surveillance cameras in case they feel uncomfortable about being captured on video.

We propose extending a system designed by the Auto-ID Center for RFID devices. The Auto-ID Center system, called Object Name Service (ONS), acts much like a Domain Name Service, except using Electronic Product Codes attached to devices instead of hostnames. For ONS, the focus is more on actual physical devices, but more relevant for privacy are the services running on the devices. Hence, we propose modifying ONS to provide a lookup for IoT software services. This lookup provides a communication mechanism to the user for presence of IoT devices and associated privacy notifications. We call the system Private.iot.

Basic Requirements.

- One Device, Many Services: Services may draw data from and control multiple devices, while an individual device may provide data to multiple services.
- Query Privacy: Queries made about individual IoT devices could result in a leak of a user's location. Here, the adversary is the operator of the lookup service, whose query logs might be examined by insiders or hackers.
- Transient Services: Many services will be statically installed. However, there is a class of services that is more transient. One prime example is a smartphone using a camera app. When the camera turns on, the phone becomes a sensor like any other device in the environment. This architecture aims to support these transient services in the same way as static ones, as long as the device is discoverable.

Private.iot relies on existing device discovery mechanisms. By definition, IoTs are physical "things" that are connected with each other, through the Internet or through other networks. To achieve this, the IoTs have to be able to discover and be discovered. For example, a *gateway* provides a portal for devices to access the internet through a central hub either through connecting via WiFi, Bluetooth, Zigbee, etc. A *beacon* actively broadcasts the device's presence and provides a point of access to gain information about the device and how to interface with it. A commercial example is the iBeacon from Apple, which employs BLE to broadcast a device UUID. Finally, a *tag* passively provides a device's information when being scanned. The barcode and QR codes are optically scanned while NFC and RFID tags are read by RF scanners.

Private.iot builds on top of the work of the Auto-ID Center. The Electronic Product Code(EPC) was designed by Auto-ID in 2003 and have since been maintained by GS1. The code consists, at a basic level, of a URI and a namespace identifier, up to 96 bits. The full specification of EPC is updated in the official EPC Tag Data Standard documentation [4]. The Auto-ID center is focusing specifically on RFID tagged objects by scanning for and using their EPC to perform a lookup on a name service (ONS) [9]. This name service acts much like a DNS for domain name resolving to an IP address. The result of the look up is a product description in the format of a physical markup language(PML). The strength in this lookup architecture is the flexibility and amount of information that can be retrieved through a simple DNS like query. The Physical Markup Language (PML) is a standardized format that is designed to describe physically manufactured objects such as manufacturer, expiration date, physical dimensions, and any other relevant information [1]. Over the last decade, additional categories to the descriptors have been added to accommodate more features provided by IoT.

System Description

Private.iot uses the ONS-PML architecture to serve privacy notifications that may be of concern to a user. The system can be broken into four components: (1) Device Discovery, (2) ONS, (3) PML, (4) Privacy Browser.

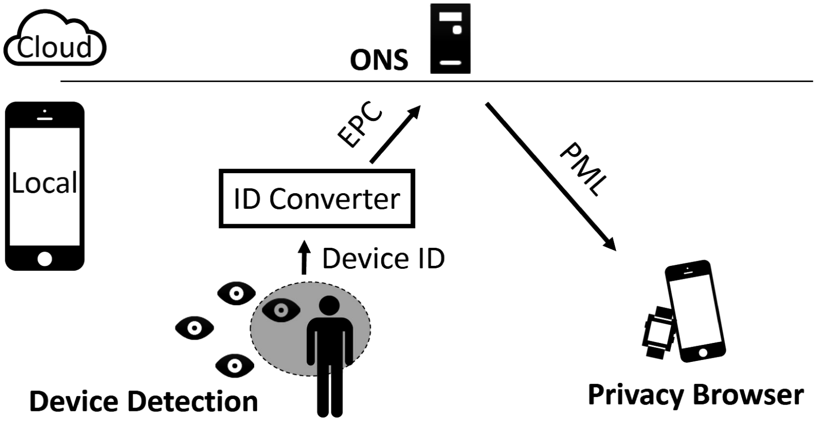


Fig. 1. ONS-PML architecture

(1) Device Discovery: Smartphones have become all-purpose communication protocol transceivers, including WiFi, Bluetooth, cellular, and NFC, all of which can be used to compose a list of device IDs that are in proximity. This is particularly true for beacons and tags. Note that devices connected through the gateway will not be discoverable directly through RF scanning. However, the gateway can control the discovery of devices connected to it. Because the proposed solution supports one-key-many-values, the gateway can register the services employed by the devices connected to it under its own device ID. In this way, the devices using gateways cannot be directly discovered but their services can. The final output from device discovery is a list of device IDs of devices around a person to be looked up in the ONS.

Different communication protocols present the issue that the lookup IDs are not all in the same format. In order to fit the current ONS infrastructure, the IDs can be reformatted to match the EPC format using a set of transformation functions done locally before the ONS query. This can be performed by constructing an EPC URI with the device ID type, device ID, and a namespace. This way, the ONS database schema will not have to be modified. To achieve this, a unique EPC is assigned to a device at the time of registration that encapsulates information about the device ID type, a device ID, and a corresponding namespace that matches the ID type broadcasted by the device (i.e. MAC, Bluetooth QPID). The device ID can then be extracted, assigned, and used by the device.

(2) Device-to-Service Lookup using ONS: A main goal of our system is to not modify the current ONS database. Given a device ID to EPC conversion, a user would be able to look up a device to retrieve a PML service description. One important criteria for the ONS lookup of this nature is the location privacy of the person querying. Unlike asking about a product, asking about a particular device could lead to exposing the location of the person. Fabian et al. describe various techniques to implement privacy preserving querying techniques for ONS [6]. The simplest is the adoption of a trusted server model where a server will purge the query data.

(3) Service Description using PML: The original ONS PML structure provided descriptions of objects belonging to physical categories. However, there have been additions to the original specifications over time as IoTs have become more capable. For example, objects that afford actuation have a “control” tag. In order to encapsulate services, we are proposing an additional tag category for services, which would include a privacy section. With this addition, the multiple service policies attached to a device can be captured in the following format:

```
<service>
<provider></provider>
<description></description>
<privacy>
<data raw="unprocessed sensor"
inference="processed sensor"/>
<notice>Human Readable Text</notice>
</privacy>
<optin></optin>
<optout></optout>
</service>
<service EPC="xxx.xxx...." />
```

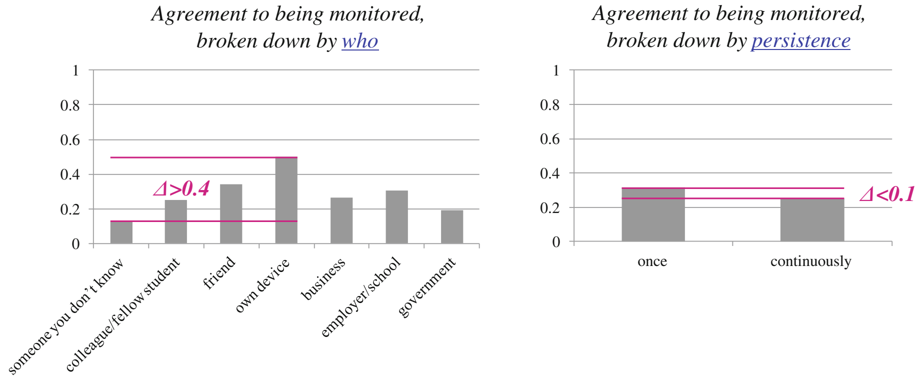
Services may draw data from and control multiple devices, while an individual device may provide data to multiple services. In this way, the concept of a virtual EPC should still be applied for referencing services that is deployed on various devices.

(4) Privacy Browser: After fetching the service information from the PML, an application can extract the privacy policy and data types that are being collected by the service and present the notice to the user. Through this browser, the user can again discover services relevant to him and when he sees that there is a service doing something against his privacy preferences, he can walk away or potentially opt-in or out given the retrieved service hooks.

3 UC Irvine: User Reactions to IoT Scenarios

The UC Irvine authors are gauging people’s attitudes, opinions and reactions to scenarios that involve IoT devices collecting potentially sensitive information. The specific goal of this project is to determine how the parameters of the scenarios influence participants’ need for notification and control. In a pre-study, 10 participants were interviewed about 9 scenarios. These scenarios differed from each other in terms of the device that is tracking the information (parameter *who*), what information is being tracked (parameter *what*), the purpose of the tracking (parameter *reason*), and the place and time at which the tracking occurs (parameter *where* and *persistence*). Participants were asked whether they felt comfortable with the scenario, whether they wanted to be informed about it, and whether they wanted to control it. The main reasons to feel discomfort with the scenarios were disagreement with the purpose, or the belief that the purpose did not justify the tracking. The main reasons to feel comfortable were trust in the entity who collects the information (e.g. government), and/or that the purpose justified the tracking.

In a follow-up study, the researchers at UC Irvine more comprehensively tested scenarios with all possible combinations of parameters. This study is designed to more determine the relative importance of each scenario parameter, and explore the opportunity to learn users' preferences regarding notification and control. To this end, researchers recruited 200 participants through Amazon Mechanical Turk, and asked them seven questions related to their privacy preferences for a randomly generated scenario. Researchers assigned 14 scenarios to each participant, therefore they gathered 98 responses from a single participant and 19,600 responses in total. Statistical analysis on the dataset showed that *who* is the most significant scenario parameter influencing people's privacy preferences. To be specific, we calculated average responses for agreement to being monitored according to all individual scenario parameters, and confirmed that there is the largest difference between responses regarding the *who* and *what* scenario parameters. Relatively, *purpose*, *where* and *persistence* have less impact, in order of significance. The data also showed that most people (more than 50 %) think monitoring activities in an IoT environment is not comfortable, safe, or appropriate.



Agreement to being monitored, broken down by who

Agreement to being monitored, broken down by persistence

Fig. 2. Average responses of allow (1: allow, 0: not allow) for *who* and *persistence* scenario parameters.

To verify whether the *persistence* parameter is truly the most insignificant, we conducted an in-depth analysis using a matrix slicing technique. We constructed partial tables to study the conditional association between participants' binary responses (i.e., allow monitoring or not) and possible combinations of scenario parameters, conditional on the *persistence* parameter. For instance, we generated scenarios using combinations of the *who* and *what* parameters, and calculated the percentage difference between responses of allow according to whether the *persistence* parameter is true or not. The results showed that the *persistence* parameter has a noticeable influence on most of the responses in subspaces of the scenarios.

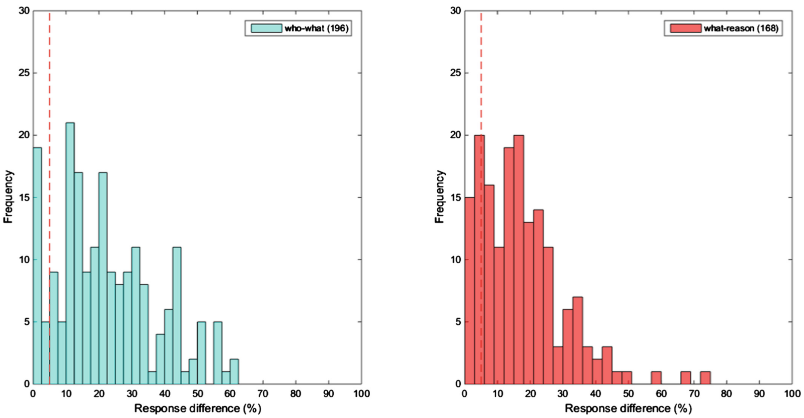


Fig. 3. Histogram of percentage differences between responses of allow for combinations of scenario parameters *who-what* and *what-reason*, conditional on the *persistence* parameter.

In summary, *who* and *what* scenario parameters affect people's privacy decisions globally, and the *persistence* parameter significantly interacts with subspaces of the scenarios, *who-what* and *what-reason* at least. In the future, we intend running additional live experiments to explore how people react to IoT notifications in realistic environments.

4 UC Berkeley: Communicating IoT Risks

IoT devices will capture a wide variety of data types, which will be accessible to numerous third-party applications. As a result, research is needed to understand the circumstances under which future users will want to be notified about an application's access to potentially-sensitive data. Providing too many notifications will lead to habituation, whereas providing too few notifications will lead to regret. Thus, in addition to studying *when* to notify, they are also examining *how* to notify.

As a precursor to designing privacy notifications for IoT devices, the UC Berkeley team studied an existing hardware privacy notification that has already been deployed for hundreds of millions of users: the webcam recording indicator. Popular media accounts suggest that many users fail to notice these indicators [2]. As a result, they performed a laboratory experiment to quantify this problem, as well as to uncover possible design improvements that could be applied to future privacy notifications [10]. They observed that fewer than half of their 98 participants noticed the indicators when performing computer-based tasks, and only 5 % noticed them when performing paper-based tasks within the computer's proximity. However, when redesigning the indicators to feature on-screen blinking glyphs, the rates at which participants noticed them increased to 93 % and 59 %, respectively.

As IoT devices become pervasive, they will need to communicate what data they are collecting (beyond raw video, as in the aforementioned experiments). We envision continuous sensing platforms employing trusted intermediaries to handle many of the sensing capabilities that third-party applications might require, similar to those described by D'Antoni et al. [3]. For instance, applications that use voice commands do not need to access recorded audio, which may contain identifiable information about the user or her surroundings. Instead, the trusted platform would process the audio so that only commands are shared with untrusted applications. One can imagine a plethora of system APIs that allow untrusted applications to benefit from audio and video input, while preventing them from collecting extraneous privacy-sensitive data. To facilitate this, users will need notifications, beyond simple recording indicators, to communicate the type of data an application is accessing.

The UC Berkeley researchers performed a series of experiments in collaborating with researchers at Intel to design these notifications for Intel's RealSense SDK [5], which performs some of the following functions:

- Age detection
- Emotion detection
- Gender detection
- Face detection (tracking)
- Face recognition
- Voice command
- Text to speech
- Language detection
- Gesture recognition
- Eye tracking
- Heart rate detection.

They conducted a series of experiments in order to create a set of intuitive icons that could be used to communicate to a user an application's use of these potentially sensitive functions. First, they described scenarios involving each of the above functions and asked participants to draw icons representing those functions, collecting a total of 240 pictograms. Through this process, they collected a wide variety of symbols representing each of the functions from participants of varying demographics and backgrounds. Next, multiple coders performed thematic analysis of the pictograms to determine the most prevalent themes. Based on the underlying themes, they iteratively created professional-looking icons and performed comprehension experiments. This iteratively improved set of icons is now going to be included in Intel's next RealSense SDK release.

5 Conclusion

The promise and pervasiveness of IoT ecosystems has lured many companies, including Intel, to devote resources and engineers to participate in the future of IoT. Privacy is part of the Internet of Things discussion because of the increased potential for sensitive data collection. This paper describes collaborative research undertaken in IoT privacy by Intel Corporation, UC Berkeley, and UC Irvine. We described some work by Intel on underlying communication protocols with an emphasis on enabling transparency with IoT devices. We saw how user studies by UC Irvine are determining which aspects of IoT are worrisome for end users. Finally, we described work by UC Berkeley researchers on the importance of the form of user notifications in IoT.

Acknowledgments. Work by UC Irvine supported by an Intel Software and Services Group Research Grant. Work by UC Berkeley supported by the Intel Science and Technology Center for Secure Computing (ISTC-SC).

References

1. Brock, D.L., Milne, T.P., Kang, Y.Y., Lewis, B.: The physical markup language core components: time and place. Auto-ID Center, Cambridge (2001)
2. Check Point Software Technologies Ltd: Are You Being Watched Through Your Webcam? http://www.zonealarm.com/blog/2013/10/are-you-being-watched-through-your-webcam/
3. D'Antoni, L., Dunn, A., Jana, S., Kohno, T., Livshits, B., Molnar, D., Moshchuk, A., Ofek, E., Roesner, F., Saponas, S., Veanes, M., Wang, H.J.: Operating system support for augmented reality applications. In: Proceedings of the 14th USENIX Conference on Hot Topics in Operating Systems, pp. 21–21 (2013)
4. EPC Tag Data Standard V1.9. GS1 (2014)
5. Egelman, S., Kannavara, R., Chow, R.: Is this thing on? Crowdsourcing privacy indicators for ubiquitous sensing platforms. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2015)
6. Fabian, B.: Secure name services for the internet of things. Dissertation, Humboldt-University (2008)
7. Kelley, P.G., Cranor, L.F., Sadeh, N.: Privacy as part of the app decision-making process. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3993–3402 (2013)
8. Knijnenburg, B.P., Kobsa, A., Jin, H.: Dimensionality of information disclosure behavior. Int. J. Hum. Comput. Stud. **71**, 1144–1162 (2013)
9. Mealling, M.: Auto-ID object name service (ONS) 1.0. Auto-ID Center Working Draft 12 (2003)
10. Portnoff, R.S., Lee, L.N., Egelman, S., Mishra, P., Leung, D., Wagner, D.: Somebody's watching me? Assessing the effectiveness of webcam indicator lights. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2015)
11. Weiser, M., Gold, R., Brown, J.S.: The origins of ubiquitous computing research at PARC in the late 1980s. IBM Syst. J. **38**(4), 693–696 (1999)