AFRL-RI-RS-TR-2020-083

RECON: REVEALING AND CONTROLLING PRIVACY LEAKS FROM NETWORK TRAFFIC

NORTHEASTERN UNIVERSITY

MAY 2020

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nations. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RI-RS-TR-2020-083 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE CHIEF ENGINEER:

/ S /
WALTER S. KARAS
Work Unit Manager

/ S /
JAMES S. PERRETTA
Deputy Chief, Information
Exploitation & Operations Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
MAY 2020	FINAL TECHNICAL REPORT	FEB 2017 – NOV 2019

4. TITLE AND SUBTITLE

RECON: REVEALING AND CONTROLLING PRIVACY LEAKS FROM NETWORK TRAFFIC

5a. CONTRACT NUMBER
FA8750-17-2-0145

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)

David Choffnes

5d. PROJECT NUMBER
DHSN

5e. TASK NUMBER
OR

5f. WORK UNIT NUMBER
TH

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Northeastern University
360 Huntington Ave
Boston MA 02115

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/RIGA
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSOR/MONITOR'S REPORT NUMBER
AFRL-RI-RS-TR-2020-083

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09

13. SUPPLEMENTARY NOTES

14. ABSTRACT

ReCon enables the auditing of PII leaks, addressing the key challenges of how to identify and control PII leaks when users' PII is not known a priori, nor is the set of apps or devices that leak this information. First, to enable auditing through improved transparency, we investigated how to use machine learning to reliably identify PII from network flows of mobile apps. Second, we extended our approach to IoT devices. Since IoT traffic is mostly encrypted, we developed tools and techniques for measuring the privacy exposure based on traffic patterns. We released L-ReCon, software for detecting PII leakage from consumer routers, and Mon(IoT)r, software for collecting and characterizing network traffic for IoT devices.

15. SUBJECT TERMS

Privacy, Security, Network traffic, Mobile apps, Internet of Things, Machine Learning, Fingerprinting

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			**WALTER KARAS**
U	U	U	UU	27	19b. TELEPHONE NUMBER (*Include area code*) **N/A**

TABLE OF CONTENTS

Section **Page**

1.0 SUMMARY

ReCon has the goal of enabling the auditing of Personally Identifiable Information (PII) leaks for mobile and IoT devices.

The first challenge we have addressed is the identification of PII leaked from mobile devices when there is no a priori knowledge of users' PII, the set of apps they are using, or the type of mobile devices we are using. The main technology we employed to address this challenge is machine learning, which is used to recognize traffic patterns that have shown the presence of PII during our past observations rather than recognizing the PII themselves. In this project, we have employed this approach to analyze popular and non-popular apps, and US apps as well as other countries apps. We have also extended the functionalities and the scalability of our approach by developing strategies for intercepting encrypted (TLS) traffic, automatic data labeling, automating apps experiments, as well as a decentralization of the approach. These have been implemented in a ReCon prototype, called L-ReCon (Local ReCon), which can be run on user premised on consumer home routers or Raspberry Pi devices without relying on VPN or a centralized cloud account, as it happens in state of the art solutions. Finally, we have publicly released a prototype ofL-ReCon.

The second challenge of this project is to identify information leaks also from Internet of Things (IoT) devices. To do this, we have created a *living* IoT lab (Mon(IoT)r Lab) at North-eastern University resembling a studio apartment with 91 IoT devices and set-up an IRB-approved user study where participants can use the devices for their intended purpose while we collect and study the data the devices exchange over the Internet. To help our analysis we have developed the Mon(IoT)r testbed software, which is able to collect and tag IoT traffic by device and experiment, intercept TLS connections, manage automated experiments by instrumenting IoT companion apps from Android devices, use a voice synthesizer to experiment with voice assistants, and show real-time statistics for all the devices. We have also extended our L-ReCon prototype to detect PII leaks from IoT devices as well. Finally, since we have seen that IoT traffic is mostly encrypted, we have developed tools and techniques for measuring the privacy exposure based on traffic patterns.

In conclusion, in this project we have developed and released three pieces of software: (*i*) L-ReCon, decentralized, deployable on user premises with no need to rely on VPN or cloud services, which helps users in detecting PII leaks from their mobile and IoT devices; (*ii*) Mon(IoT)r testbed software, which supports experiments and monitoring of IoT devices; (*iii*) traffic analysis scripts, which use the data collected with L-ReCon and Mon(IoT)r testbed software to analyze privacy trends and implications for mobile and IoT devices.

This project had significant impact to the academic community with the acceptance of four research papers in top conferences on security, privacy, and Internet measurements, and the recognition of a community contribution award for sharing our IoT data and analysis tools with the broad research community. Our software is currently in use in other prestigious labs: Imperial College London, Politecnico di Torino, America's Test Kitchen; and in the process of being deployed in Consumer Reports' premises. Articles highlighting our results have also appeared in prestigious US press such as the New York Times and the Financial Times, and translations of such articles have appeared in equally prestigious press in other countries.

2.0 INTRODUCTION

There has been a dramatic shift toward using mobile devices such as smartphones and tablets as the primary interface to access Internet services. Unlike their fixed-line counterparts, these devices also offer ubiquitous mobile connectivity via WiFi and cellular data access, and are equipped with a wide array of sensors (e.g., GPS, camera, and microphone). The combination of rich sensors and ubiquitous connectivity make these devices perfect vectors for invading the privacy for end users. In addition to mobile phones, other more specialized devices such as TVs, light bulbs, cameras, etc. are becoming Internet connected, which are known as the Internet of Things or simply IoT. As for mobile devices, these devices also have the potential to learn and expose extensive information about their users and their surrounding environment. As most of these devices lack any interfaces that indicate information exposure, there is an urgent need for research that provides transparency into such exposure at scale.

The main challenges of this project are the identification of privacy exposure from both mobile devices and the Internet of Things.

2.1 Challenges of Privacy Exposure for Mobile Devices

In the case of mobile devices, the main challenge is to identify and control privacy leaks in terms of Personally Identifying Information (PII), which are sent over the Internet by individual apps. This was possible since mobile devices allow us to install custom certificates and intercept encrypted traffic. Once we have access to plaintext traffic (both unencrypted, and intercepted from TLS connections), we can apply our analysis to extract PII.

Some problems we address in this project are how to advance the state of the art to detect such PII. Previous work addressed this using machine learning, manually engineered features, and using a centralized classifier. We advanced this by allowing automatic training of machine learning classifiers and making the system work in a decentralized and scalable way by developing a new tool called L-ReCon (Local ReCon). This way users do not need to send their traffic to the cloud for analysis and can simply install L-ReCon on their home router or other popular affordable devices such as a Raspberry Pi.

With the software we have produced we were able to do a longitudinal analysis of the app ecosystem spanning 8 years and 7,665 apps [19], release a tool for visualizing PII exposure [20], and to analyze non-US and non-popular apps [9].

2.2 Challenges of Privacy Exposure for the Internet of Things

Regarding IoT, the problem is similar, but the main challenge is different. From our preliminary experiments we have seen that most of the IoT network traffic is encrypted and IoT devices do not typically offer an easy way to install our TLS certificates. Specifically, IoT devices behave like black boxes where we can just analyze what enters and what exists, and what visible functionality has been triggered (e.g., a light bulb turning on/off).

For the few cases where IoT devices send traffic in plaintext, we were able to use the same methodology developed for mobile devices, including the support of our user-installable software L-ReCon. However, for the vast majority of cases for which the traffic is encrypted, we cannot extract plaintext PII, but we can still access traffic information such as destinations, ports, TLS

handshakes, and the timing of traffic. For this reason, in the case of IoT, we have built a new software based on L-ReCon, which is the Mon(IoT)r testbed software [15], deployed it in a lab replete of IoT devices, and used it to run controlled and uncontrolled experiments at scale (i.e., experiments run by the staff as well as others run as part of an IRB-approved user study with the collaboration of some voluntary user study participants). With this IoT testbed we were able to capture/tag IoT traffic, and eventually measure its information exposure in terms of: (*i*) suspicious destination, such as traffic directed to the device manufacturer, cloud/CDN services, or other third parties; (*ii*) adoption of encryption, by relying on protocol, port used, and information entropy; (*iii*) IoT behavior leakage, by correlating traffic patterns to the actual functionality of the IoT devices being triggered, so that we can detect what the device is doing in terms of observable functionality even if we cannot see what the device is transferring.

2.3 Impact

This project had significant impact both for the academic community in terms of research papers, technology transfer to companies and universities, and for the general public in terms of downloadable software and press coverage.

2.3.1 Academic Impact. Academic impact has been recognized by the acceptance of four research papers in top conferences: Network & Distributed System Security Symposium (NDSS '18) [19], Privacy Enhancing Technologies Symposium (PETS'18) [6], Internet Measurement Conference (IMC'19) [18], Network & Distributed System Security Symposium (NDSS '20) [22]. Moreover, the paper on IoT information exposure (IMC'19 [18]) was the only one awarded the prestigious *Community Contribution Award* since all the software and data has been released to the public to foster follow-up research [7, 15].

2.3.2 Technology Transfer. The Mon(IoT)r testbed software [15] we developed has been used at Imperial College London and Politecnico di Torino to create a copy of our IoT testbed, which resulted in a joint publication [18]. The same software has also been run to capture data from cooking IoT devices at America's Test Kitchen and is currently being deployed at Consumer Report's premises as well. In total, as part of this project, we collaborated with 12 organizations.

2.3.3 Public Impact. L-ReCon has been used for a public demo in front of 500 users at the New York Rooftop film festival in July 2017 and the data we have generated from our IoT experiments have been downloaded 59 times, facilitating follow-up work from other institutions.

Our work has been covered by the **New York Times** on an article about privacy concerns for smart speakers [3] and by the **Financial Times** regarding privacy concerns of popular smart TVs [5].

3.0 METHODS, ASSUMPTIONS, AND PROCEDURES

This section describes the list of research methods, assumptions, and procedures organized by tasks, following the order as they appeared in the initial project proposal.

3.1 Identifying Textual PII leaks in Network Traffic from Mobile Devices

This main activity was carried on over the first year of the project, and covered our first challenge of auditing mobile devices. The main methods we use are supervised machine learning and man-in-the-middle (MITM). Our main assumptions are the availability of ground truth regarding Personally Identifiable Information (PII) that we use to apply machine learning, and the fact that we can read plaintext traffic and use it to search for PII. The starting point for this activity is previous work where we have been able to detect PII from plaintext network traffic using a centralized classifier which was trained using manually labeled data. The following subsections explain the procedures we used to reach the goal of identifying textual PII leaks from mobile devices.

3.1.1 Extend ReCon analysis to include non-popular apps, and apps from other countries. In this task our goal was to analyze selected non-popular apps and apps from other countries with the help of our collaborators abroad. Regarding the analysis of non-popular apps (limited to the US market), we used semi-automated interactions to analyze about 1000 non-popular apps. For each app requiring a login, we created a new account using a previously unused email address. Regarding non-US apps, we applied the same interaction methodology, but selected apps from the Chinese and Indian popular app distribution stores for both iOS and Android devices: (*i*) For China, the official Android distribution store "Play store" is not available and the majority of the users use third-party app stores: "Tencent MyApps" and "360 Mobile Assistant." We focused on those two stores, since they have the most market shares, and selected popular apps whose encrypted traffic can be intercepted. (*ii*) The Indian market uses "Play store", thus we used the rankings in the official store. The details and the results of this study have been reported in project deliverable "Findings Regarding non-Popular and non-US apps" [9].

3.1.2 Automate use of information flow analysis (IFA) tools, allowing us to provide large volumes of labeled data without manual interaction. This research focused on designing a way for labeling data automatically. The approach we developed uses an iterative method in which, starting from an existing set of labels, we use such labels to train our first machine learning classifier. Then, the predictions of the trained classifier are used iteratively as labels to retrain the classifier itself. The methodology for producing training data has been delivered in project deliverable "Classifier data [8]." The actual data has been included in the L-ReCon distribution "Software to run ReCon on LANs [7, 14]."

3.1.3 Integrate recently developed automated app-interaction tools that go beyond MonkeyRunner, to further facilitate label generation. We have been able to improve MonkeyRunner and overcome its initial limitations to improve efficiency and to massively use it for testing different versions of the same app and also in situations when a login is required. This resulted in a paper publication in NDSS 2018 [19], where we studied privacy leaks from historical and current versions of 512 popular Android apps, covering 7,665 app releases over 8 years of app version history. Our app-interaction approach also works with companion apps that control Internet of Things devices, and therefore within this project this will be a fundamental component for exploring not only generic mobile apps, but Internet of Things functionalities as well. The code for this software has been released in [16].

3.1.4 Improve mapping between network traffic and app that generated it, to ensure correct app is blamed for PII leaks. This work started by recognizing traffic signatures we have previously encountered during our experiments. Such signatures are detected by analyzing the headers of HTTP(S) traffic flows generated by devices connected to ReCon. We have made this approach compatible with the consumer-router L-ReCon ("Software to run ReCon on LANs" [14]), which can be downloaded from [7]. We have further extended this approach to situations of apps that we have never encountered before and when man-in-the-middle is not available: our key observation is that mobile apps are composed of different modules that often communicate with a static set of destinations. We leverage this property to discover patterns in the network traffic corresponding to these different modules. On a high level, we group together (encrypted) network traffic flows based on their destination and find correlations in destinations frequently accessed together. We then combine these patterns into fingerprints, which may, among other use cases, be used for app recognition and unseen app detection. Details and results of this approach have been published to NDSS20 [22], with software available for download at [21].

3.1.5 Adapt ReCon to work on Raspberry Pi and/or home router. We developed L-ReCon ("Software to run ReCon on LANs" [7, 14]), which has the following improvements over the cloud-based ReCon software we used as our inspiration:

- L-ReCon works without requiring virtual private network (VPN) and cloud support. This required the creation of new distributed software components that replace VPN/cloud components.
- Added new man-in-the-middle interception software (mitmproxy4).
- Added automatic account creation when a device connects for the first time to L-ReCon. Individual devices are identified using their MAC-Address.
- No registration necessary as was needed in the original cloud-based ReCon.
- Added a simplified interface for installing user certificates when using smartphones and tablets with L-ReCon to easily enable man-in-the-middle.
- Added the possibility for users connected to L-ReCon to enable or disable TLS interception as they wish.
- We analyzed the performance of the latest version of L-ReCon: it can intercept and analyze 24 HTTPS flows per second using dd-wrt on a Netgear Nighthawk X10, which corresponds to decent performance for roughly 30 to 40 lowly active users. We consider this acceptable for such home router since it supports 32 devices per Wi-Fi band.

The L-ReCon documentation and software are available at [7].

3.1.6 Enable TLS interception using existing techniques such as mitmproxy, to enable ReCon analysis on encrypted flows. We developed a smart TLS interception system that allows us to intercept TLS connections of all devices connected to L-ReCon when possible, without disrupting the functionality of devices or apps that do not allow TLS interception (for example, because of

certificate pinning). Our approach can recognize when a TLS connection fails and can disable interception. This is currently done at IP address level, meaning that if man-in-them-middle (MITM) to a TLS connection of a device fails for a given IP, the next time such device connects to the same IP, no MITM is performed. This way a TLS connection only fails for the first time. This is not a problem since many apps (and IoT devices) usually try to connect more than once before giving up. The implementation is based on a modified version of mitmproxy that has been integrated in L-ReCon [7].

3.1.7 Allow training from decentralized crowdsourced labels, while ensuring privacy. We have designed a combined multi-classifier predictor that can use data trained by multiple classifiers. This allows L-ReCon to share its classifier to other users, which can be used to aid in its predictions. The automatically trained classifier can be shared by other users and then used by the multi-classifier predictor to improve the recognition capabilities of ReCon.

We have designed a software library with the new classifier component. Its description has been reported in deliverable "Software that allows decentralized crowd-sourced training [11]."

3.1.8 Develop framework for including analysis on data using other techniques. We developed several analysis tools to extract knowledge from the network traffic we collected that go beyond simple PII analysis. A list of the analysis methods we developed is the following.

- **Identify if traffic is sent in plaintext of encrypted**. We first developed a protocol-based approach, which we used and explained in our NDSS18 paper [19]. When such approach failed we used information entropy analysis by observing that encrypted traffic has entropy comparable to random traffic. This extension has been described and used in our IMC19 paper [18] and the code has been released in [16].

- **Party analysis**. We develop methods for distinguishing if traffic is sent to a *first party* (the developer of an app or related company) or a *third party* (everything that is not a first party such as advertisers and analytics providers). Our method employed a combination of domain name and domain name/IP ownership matching using public databases (such as WHOIS). This approach has been documented in [18, 19], with software available in [16].

- **App fingerprinting and similarity**. We developed an app fingerprinting approach that is unsupervised (i.e., it does not require ground truth) and that can be used to recognize automatically existing and new apps from network traffic [22]. One of the peculiarities of this approach is that it can distinguish if apps are browsers or not, and if apps are similar to other apps (for example, new versions of existing apps, or apps with libraries in common). The software for this analysis has been released in [21].

3.2 Extension of PII Analysis of IoT Devices

This second activity was carried on over the second year of the project, and covered our second challenge of auditing IoT devices. The main methods we use are the destination and encryption characterization methods we developed during the first year of this project, supplemented with new IoT-specific methods, such as IoT behavior fingerprinting based on interaction method and

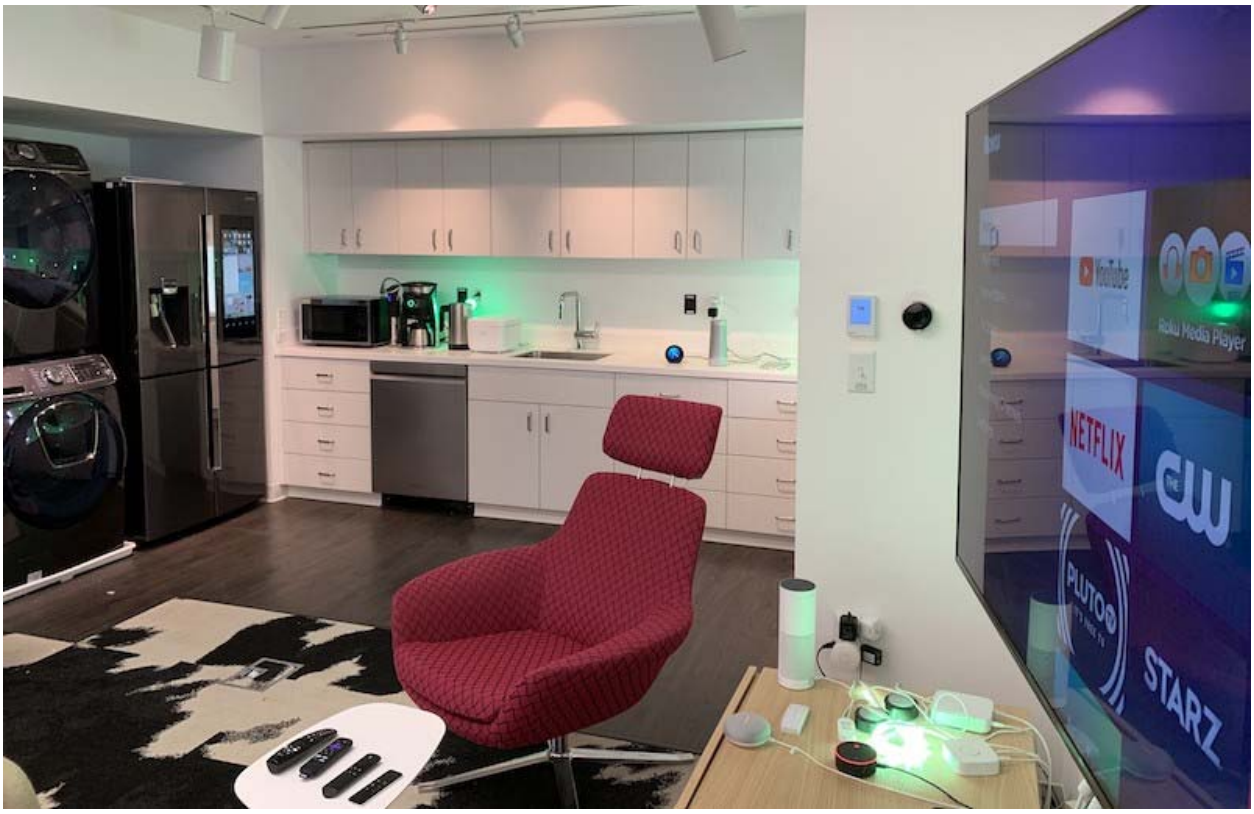

Figure 1. Mon(IoT)r IoT lab at Northeastern University, configured like a studio apartment that contains a large set of consumer IoT devices. The lab runs the Mon(IoT)r testbed software we developed as part of this project

functionality, and an IRB-approved user study to activate IoT devices for their intended purpose by user study participants. For this activity the assumption that we can use man-in-the-middle to intercept the leakage of PII is no longer true since we assume them as black boxes that do not let us install our self-signed certificates. For this reason, we mostly analyze what information can be inferred that does not require looking at IoT plaintext traffic (e.g., traffic destination, headers, and quantitative statistics based on sizes and timings).

The following subsections explain the procedures we used to reach the goal of identifying privacy exposure from IoT devices.

3.2.1 Identify the initial set of devices to acquire. Our research approach for analyzing IoT devices had to start with the selection of a list of initial IoT devices for our IoT Lab (named "Mon(IoT)r Lab", pronounced "Monitor Lab"). Since the Mon(IoT)r lab, as discussed in our original project proposal, is intended as a "living lab", where people will be using its devices for their personal enjoyment after being enrolled in a IRB-approved user study to collect their data, we prioritized devices that may typically be found in a smart home environment, that are easy to buy (i.e., they are available on Amazon), that are popular and cheap (according to Amazon sorting), and that do not require technical expertise or particular equipment to be used. Under these requirements we decided to buy devices in the following categories: smart refrigerator, smart

dishwasher, smart washer, smart dryer, smart TVs, smart speaker with voice assistant, smart bulbs, smart hubs, smart motion sensors, smart cameras, smart TV dongles, smart plug-in switches, smart label printer, smart bell, smart open/close sensor.

The list of devices we bought, including typology and vendor has been reported on our "Summary of testbed" deliverable available at [14].

3.2.2 Develop initial living IoT lab testbed. We obtained from Northeastern University a space configured as a studio apartment, where we deployed the IoT devices we acquired to create the Mon(IoT)r Lab. The final result is reported in Figure 1. For the Mon(IoT)r lab we developed a special version of L-ReCon that supports measurement of traffic from IoT devices. We named this version "Mon(IoT)r Testbed Software". This software is capable of autonomously analyzing the traversing traffic from devices that are connected to it (both mobile and IoT devices). For unencrypted traffic, no instrumentation is needed for either mobile devices and IoT devices. For encrypted traffic, mobile devices must add the TLS certificate used by L-ReCon using the configuration interface they offer, while IoT devices have to be physically rooted since they do not offer an interface for installing and trusting our TLS certificate. This Mon(IoT)r testbed software has been deployed on a dedicated server in our lab. The reason why we decided to use a dedicated server instead of a home router (which is the typical installation scenario for L-ReCon) is that a simple consumer router does not have the necessary computational power to run all our intensive experiments on all the 91 devices we have acquired. This particular setup is not intended for home use due to its additional complexity, however it is customized for our research needs and its detailed configuration has been reported in the "Summary of testbed" document [14].

Some additional features unique to the IoT testbed (in addition to the ones we developed for L-ReCon during the first year of the project) are the following:

- Possibility to capture the Wi-Fi-to-Wi-Fi traffic among the monitored devices (in addition to the traffic sent over the Internet and other intra-lan traffic). Normally, Wi-Fi adapters, especially USB ones and the ones integrated in Wi-Fi routers do not send the traffic between their clients to their host, thus making it impossible to capture the traffic. By trying different adapters, we have found out that Qualcomm Atheros 9k adapters allow such capture, thus allowing us to receive such traffic. We have then created a way to merge Wi-Fi-to-Wi-Fi traffic with the wired-to-wired traffic and Wi-Fi-to-wired traffic without adding the complexity of dealing with multiple simultaneous captures, by employing the use of virtual network interfaces and traffic mirroring. This way we have now all the traffic on a single interface that can be directly analyzed using tcpdump, or mirrored to an external port to feed a visualization engine that shows network statistics in real time. The major accomplishment is that we can now know if any devices on our testbed talk to each other regardless of the technology they use to be connected to the network. So far we have found out that many devices try to discover other devices on the network. In particular, we have seen this happening with Samsung Smartthings, which natively tries to access and control other devices (even from different vendors, such as the Philips Hue hub); moreover, we have seen that the Echo family of Amazon devices are constantly sending encrypted information to each other. We do not know yet what this information is, but we noticed that this is responsible for a certain degree of wireless connectivity degradation, which may not be desirable by its users.

- Ability to federate the IoT testbed with other testbeds running in different locations. We have created a permanent VPN link between Northeastern University and Imperial College London, which was running another instance of our Mon(IoT)r testbed software. Such a link can be used as default gateway for our IoT devices when we want to test whether US devices behave differently when used with a UK ip address and vice-versa. Our way of switching connectivity is entirely configurable and we are even able to have multiple networks at the same time, some with local connectivity and some with international connectivity. We expect this system to be extended with other third-party VPN services as well, to test IoT behavior with connectivities other than US and UK. So far we have seen that all the devices we have work correctly when used with a UK ip address. In certain cases, such as with the LG smart TV, we got an information message asking us to change the country in the device configuration and to accept the privacy policy of the new country.
- We have also developed tools to aid researchers in organizing controlled and uncontrolled experiments, which proved to be easy to learn for our undergraduate students.

A detailed description of all the Mon(IoT)r lab infrastructure, Mon(IoT)r testbed software, including the server, the network components, the experiment organization, and the data collection has been reported in the "Summary of testbed [14]", the Mon(IoT)r testbed software can be download from [15] and the tools for managing the experiments and analyzing the data collected are available at [16].

3.2.3 Identify additional devices to acquire and integrate into testbed. After doing some preliminary experiments with our initial set of IoT devices, we tried to cover more categories and to have additional samples of devices per category. To do this, we kept using the same methodology for adding new devices, i.e., we sorted Amazon results based on price and popularity for devices of multiple IoT categories.

The final list of devices we bought for this project, including typology and vendor has been also reported on the "Summary of testbed" document [14]. An up-to-date list with devices we are adding also after the end of this project has been maintained at [17].

3.2.4 Design Controlled Experiments for the Mon(IoT)r Lab. To create an experiment plan, we have defined a list of IoT devices interaction methods: local (physical interaction), voice interaction (using a smart speaker), companion app interaction from the same network (using a smart phone), companion app interaction from an external network, companion app interaction from different OSes and device types. Exploring different interaction methods allowed us to assess if the privacy risk is affected by how users interact with IoT devices. We have also defined a list of measurement metrics: domains contacted during the interaction, whether encryption is used or not, information being exposed in terms of network traffic patterns revealing the functionality and interaction method of the IoT devices. This is in addition to the PII recognition engine of L-ReCon.

The design of these experiments has been published in our IMC'19 paper [18].

3.2.5 Conduct Controlled Experiments at the Mon(IoT)r Lab, Analyze Data. To conduct controlled experiments, we have first defined what an IoT category is, and what actions such a category offers. This categorization allows us to repeat the same actions with several devices, thus

allowing for apples-to-apples comparisons. For each device, we have performed all actions of its category for each interaction method (e.g., local interaction, local network interaction, cloud interaction, voice interaction) and collected measurement data. Each experiment and analysis have been performed using our Mon(IoT)r testbed software. We have completed over 30000 controlled experiments for supervised behavior fingerprinting for the IoT devices we have in the Mon(IoT)r lab. Our behavior fingerprinting consists of analyzing encrypted traffic exchanged by each IoT device for specific patterns. The patterns found are then associated to the different type of interaction (and functionality) of the IoT device, including "power on" interactions, "local/physical" interactions, "remote interactions using mobile apps from the same LAN", "remote interactions using mobile apps" from different LANs, "remote interactions using voice assistants." This would enable the detection of IoT activity from generic encrypted traffic, and therefore reveal possible privacy exposure. While conducting our experiments, we have also captured the traffic generated by IoT devices to detect the presence of textual privacy leaks. To perform the IoT behavior fingerprinting, we developed a new IoT analysis tool, which is able, using the labeled encrypted traffic of an IoT device (where labels represent behavior, such as "turn off a light", "start a video stream", etc.), to learn to recognize the device itself, and the specific behavior when the traffic is unlabeled, with an accuracy of up to 88%. More in detail, this tool allows users to train a random forest model to either identify an IoT device or what state of interaction that smart device is in. To train a model, this tool receives a text file that contains a list of pcap (traffic) files that identify a specific device or interaction. The tool also expects to receive the number of packets used for grouping a set of packets (burst), and the label to train on. Finally, we have developed a way to perform automated experiments, thus allowing us to perform our large number of over 30,000 experiments in a short period of time with no human interaction. The idea behind this is that most IoT devices can be controlled using a companion app or voice interaction. Although we cannot easily automate physical interactions, we have been able to adapt existing tools to perform automated companion app interactions and automated voice interactions. To automate companion apps we used the possibility of the Android platform to send simulated user events to the devices. In this way, we are able to open/close apps, and to send click/swipe/type events that can effectively simulate user interactions and test different functionalities of the IoT device with minimal supervision. For what concerns voice-controlled devices, we used the Google voice synthesizer API, which proved to be recognized by such devices as a real human voice. By using such voice synthesizer we were able to issue voice commands that are also able to activate/deactivate IoT functionalities. It is possible that, in a small number of cases, the voice recognition fails and/or the automated interactions with an app fail (for example, when the app server is unavailable/slow, or when the app produces unexpected outputs, such as a request to rate the app). We are able to remove such cases from our dataset by finding experiments whose file size is very different from the one of other (successful) experiments. The technique we use to remove these outliers is the Interquartile range test method. Finally, we repeat any removed/failed experiments.

All these controlled experiments were also part of our IMC19 [18] publication, and we have also made all the analysis scripts and software available for public download [16].

3.2.6 Design and submit user study for conducting experiments using human subjects. We have written and submitted an IRB proposal for a user study related to ReCon and the Mon(IoT)r Lab. The proposal states that we can enroll any person that has access to the open space (room

ISEC 660 at Northeastern University) that hosts the Mon(IoT)r lab. The enrollment provides subjects with access to the lab and the possibility to use the smart devices inside for their intended purpose. As a return we are able to measure and use the data collected by our IoT for research purposes, including its analysis and dissemination in aggregate form.

The IRB of Northeastern University has approved our proposal. The initial duration was until the end of the project, but the study continues beyond this project and the IRB approval has been extended beyond that.

3.2.7 Conduct IRB-approved User Studies. We started the IRB user study in September 2018. To avoid violating the IRB protocol, we have added support to the Mon(IoT)r analysis infrastructure for the possibility to start and stop IRB experiments with an external button that can be operated by the research personnel of the project. This allows us to set the whole system in "off-the-record" mode while non-participants of our user study enter the lab. Since this IRB study is composed of uncontrolled experiments, there is no planning to do besides making the space available and having study participants using the lab

We enrolled 55 people to the user study (including the research personnel). The lab gets multiple visits every day, thus ensuring that some IoT devices are often actively and passively triggered. The data we captured has also been analyzed in our IMC'19 publication [18].

3.2.8 Adapt ReCon to Find Textual PII leaks from IoT Devices. The limitation of IoT devices such as vast predominance of encrypted traffic made it not possible to reuse the PII detection mechanism we developed for mobile devices. However, the Mon(IoT)r testbed software and in particular, the scripts to trigger IoT behavior, helped us to find a few case of PII leaks that are relatively easy to search from IoT network traffic. This included various forms of unique identifiers (MAC address , UUID, device ID), geolocation at the state/city level, and user specified/related device name (e.g., John Doe's Roku TV). A notable case that we found is the Samsung Fridge sending MAC addresses unencrypted to an EC2 domain. The implication is that it is now possible for an ISP to track this device. We also found that a Magichome LED Strip was sending its MAC address in plaintext to a domain hosted on Alibaba. The Insteon hub was sending its MAC address in plaintext to an EC2 domain. Each time the Xiaomi camera detected a motion, its MAC address, the hour and the date of the motion (in plaintext) was sent to an EC2 domain. We also noted that a video was included on the payload.

A more detailed description of the methodology we used to find these PII has been included in the IMC'19 paper [18] and the software has been released in [16].

3.2.9 Identify audio and video in network traffic from IoT devices. In carrying out this activity we used two approaches: first, looking for magic numbers in the traffic to detect audio and video content directly, assuming it unencrypted, and second using network traffic fingerprinting to recognize packet timings and sizes that can be correlated to a stream of audio or video. Regarding unencrypted audio and video leaks, we could not find a lot of cases since most of the communication is fortunately encrypted. We still have found some notable cases: we observed that the Ring doorbell performs a video recording action every time a user moves in front of it. However, this is unexpected behavior: the app used to set up the device does not warn the user that the doorbell performs such recording in real time, the doorbell offers no indication that recording is occurring,

and the only disclosure is in fine print as part of the privacy policy. Upon discovering this barely documented feature, we logged into our account to see the video, and learned that we must pay an additional monthly fee to access such recordings. We have not identified any way to turn off this feature. The Zmodo doorbell uploads camera snapshots when the device is first turned on, and also when anyone moves in front of the device. This feature is undocumented, and we were unable to prevent such snapshots from being taken, nor we were able to access them. A number of our user study participants complained that Alexa-enabled devices are frequently triggered during normal conversations, as if the keyword had been spoken. Upon investigation, we found that the default Alexa wake word "Alexa" is frequently triggered by many other unrelated words, much more so than the other voice assistants in our testbeds. A notable example is a sentence beginning with "I like [s-word]" such as "I like Star Trek." We understand that this may be a limitation of voice recognition technology, but it is still a potential privacy exposure since Amazon devices typically recognize false activations after sending the whole sentence to their servers (i.e., after such a sentence has been permanently stored). Also this work has been published as part of our IMC'19 [18] with software available in [16] and received press coverage on the Financial Times [5].

Another approach for identifying audio and video from IoT devices that integrate a voice assistant such as smart speakers, focused on exposing the IoT devices to some signals such as audio from popular Netflix shows, and check whether they start recording from their microphone or not without their wake word being pronounced. To do this we employed a combination of techniques such as network traffic pattern analysis (to recognize the presence of audio transmission from encrypted network traffic), camera feed analysis (to understand if the devices lights up, notifying a recording), and cloud data analysis (to understand if the audio from a device is being stored to the cloud). Preliminary results from this approach have been published in [2]. The preliminary results of this work also appeared on both the online and paper version of the New York Times [3].

3.2.10 Identify activity-triggered PII leaks from IoT devices. This work consisted in understanding the IoT devices behavior (with respect to privacy), in response to a given trigger. We have designed several controlled experiments where we perform an action, and we see what happens in the network traffic. We then analyzed such traffic for any possible information exposure by looking for PII, if the destinations are first or third parties by using DNS, WHOIS, and X509/SNI (Server Name Indication) information; if they were encrypted or not by using protocol analysis and information entropy analysis; and finally if the trigger can be inferred by performing IoT behavior fingerprinting through the use of machine learning. The advantage of behavior fingerprinting is that it allows us to detect if the device is exposing its behavior (i.e., its functionality) whether encryption is employed or not. To this end, during the controlled experiments we have produced a large set of labeled pcap files, each one including one action. These files are automatically organized by device, category, interaction method, and action by the Mon(IoT)r testbed software. We used these files and the associated ground truth to build a training/validation set for applying machine learning. We used Random Forest as ML approach, which proved to be good on previous work in literature. As features, we used statistical properties of the traffic (interarrival time between packets, number of packets, total traffic, traffic direction, destinations, protocols, grouped using averages and variances). Additional information of the methodology has been discussed in [18], with software downloadable at [16].

4.0 RESULTS AND DISCUSSION

The outcomes of this project can be summarized as follows:

- A list of technology transfers representing entities (companies and universities) that benefited from the outcomes of our project, from exploratory collaborations to running of software on their data or even on their premises (discussed in Section 4.1).
- A list of research papers that have been accepted in top conferences (listed in Section 4.2).
- A list of meetings and presentations where we presented and discussed our work with both the academic and the broader community (listed in Section 4.3).
- Finally, the list of deliverable and milestones, that summarize what has been produced from this project. Deliverables are a list of documents explaining the methodologies we developed, and the software we have released to the public (discussed in Section 4.4). For each deliverable we also provide a reference to its download links.

4.1 Technology Transfer

Telefonica. We shared our on-router L-ReCon deployment with Telefonica as part of our collaboration with the Data Transparency Lab. L-ReCon has been successfully demoed within Telefonica using a real consumer router. The plan for Telefonica is to experiment with the software and help us to improve the human-computer-interaction component of L-ReCon. We have already received some prototypes from Telefonica and we are waiting for usable components and/or additional feedback.

Northeastern University IT Services. We started a collaboration with IT services of Northeastern University to test L-ReCon on Enterprise equipment borrowed from them. Their help allowed as to be successful in our porting process and to be able to demo ReCon in front of an audience of 500 people, where each of them was allowed to connect to the system. With this approach we were able to scale L-ReCon performance by 20 times with respect to using consumer hardware, so allowing us to service even the whole crowd if needed.

CNIL. We started a collaboration with CNIL (Commission Nationale Informatique & Libertés), which is the FTC equivalent in France, to experiment with our software. We helped Franck Baudot install and use it and we are waiting for feedback.

FTC. We were told that the FTC (Federal Trade Commission) has been using L-ReCon, but as per their policy, they do not directly disclose what they are using, nor have they contacted us for support.

Mimic. We started a collaboration with Mimic about our IoT work. We got from Alec Rooney access to a large dataset of traffic generated by consumer IoT devices that we used to understand the behavior of IoT and to build analysis and behavior-recognition tools that can be used to aid our experiments in the IoT Lab.

Lookout. We shared ReCon with Lookout, after we met them at RSA 2018 and followed up in Boston. We agreed to use ReCon on their wide set of labeled pcap files to detect PII. We got access to data flows from their ElasticSearch traffic database.

America's Test Kitchen. We coordinated with America's Test Kitchen the deployment of a traffic-collecting component, based on our Mon(IoT)r lab infrastructure, which has been installed on their premises to produce some additional IoT data that we can use in our analysis.
Imperial College London. The research group of Hamed Haddadi at Imperial College London has built another IoT lab based on our setup. We are coordinating with them in order to have comparable experiments in the UK market. To this end we have given them the specifications and the code of our monitoring infrastructure. They have currently installed it and are in the process of collecting data. RBC. RBC has shown interest in IoT analysis for a bank environment.
Stanley Black & Decker. We met with Stanley Black & Decker to discuss a collaboration around their IoT products' security and privacy.
PricewaterhouseCoopers We are engaged in a one-year research agreement with PricewaterhouseCoopers on the topic of IoT security and privacy, and are planning to extend this to a multi-year agreement.
Politecnico di Torino. The research group of Marco Mellia at Politecnico di Milano has built a third IoT lab based on our Mon(IoT)r testbed code. They are using the software we developed as part of this project to perform IoT experiments for the Italian market.
Consumer Reports. We are drafting a data sharing agreement with Consumer Reports for using our tools to collect and analyze the data of the IoT devices they test. Our Mon(IoT)r software is soon expected to be deployed on their promises.

4.2 Research Papers

NDSS '18: Bug Fixes, Improvements, ... and Privacy Leaks - A Longitudinal Study of PII Leaks in Android Apps
Authors: Jingjing Ren, Martina Lindorfer, Daniel J. Dubois, Ashwin Rao, David Choffnes, Narseo Vallina-Rodriguez
Publication venue: NDSS 2018 (The Network and Distributed Systems Security Symposium)
Conference location and dates: San Diego, CA, USA - February 18-21, 2018
Paper reference: [19]
PETS '18: Panoptispy: Characterizing Audio and Video Exfiltration from Android Applications
Authors: Elleen Pan, Jingjing Ren, Martina Lindorfer, Christo Wilson, David Choffnes
Publication venue: Privacy Enhancing Technologies Symposium (PETS'18)
Conference location and dates: Barcelona, Spain - July 24-27, 2018
Paper reference: [6]
IMC '19: Information Exposure from Consumer IoT Devices (A Multidimensional Network-informed Measurement Approach)
Authors: Jingjing Ren, Daniel J. Dubois, Anna Maria Mandalari, Roman Kolkun, Hamed Haddadi
Publication venue: IMC 2019 (The Internet Measurement Conference)
Conference location and dates: Amsterdam, The Netherlands - October 21-23, 2019
Paper reference: [18]
NDSS '20: FLOWPRINT: Semi-Supervised Mobile-App Fingerprinting on Encrypted Network Traffic
Authors: Thijs van Ede, Riccardo Bortolameotti, Andrea Continella, Jingjing Ren, Daniel J. Dubois, Martina Lindorfer, David Choffnes, Maarten van Steen, and Andreas Peter

Publication venue: NDSS 2020 (The Network and Distributed Systems Security Symposium)
Conference location and dates: San Diego, CA, USA - February 23-26, 2020
Paper reference: [22]

4.3 Meetings and Presentations

ReCon kickoff
Meeting Purpose: Start of the project
Meeting Start and End Dates: May 2, 2017 (one day)
Meeting Location: ISEC building at Northeastern University, 805 Columbus Ave, Boston, MA
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel Dubois, Jingjing Ren

DHS S&T Cybersecurity Showcase
Meeting Purpose: Overview of project
Meeting Start and End Dates: July 13, 2017 (one day)
Meeting Location: Mayflower Hotel, Washington, D.C.
Meeting Attendees from this project: Prof. David Choffnes
Presentations Made: ReCon Overview

Rooftop films festival
Meeting Purpose: Show Harvest documentary [1] and involve the audience in a public demo of L-ReCon Meeting Start and End Dates: July 28, 2017 (one day)
Meeting Location: Industry City, Brooklyn, New York City, NY
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel Dubois, Jingjing Ren
Presentations Made: Harvest documentary, L-ReCon Public Demo

NDSS 2018 (The Network and Distributed Systems Security Symposium)
Meeting Purpose: Paper presentation.
Meeting Start and End Dates: February 18-21, 2018 (four days)
Meeting Location: Catamaran Resort Hotel & Spa, San Diego
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel Dubois, Jingjing Ren
Presentations Made: “Bug Fixes, Improvements, ... and Privacy Leaks - A Longitudinal Study of PII Leaks in Android Apps”

PrivacyCon 2018
Meeting Purpose: Paper presentation
Meeting Start and End Dates: February 28, 2018 (one day)
Meeting Location: Constitution Center, 400 7th St SW, Washington, DC
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel Dubois, Jingjing Ren
Presentations Made: “Bug Fixes, Improvements, ... and Privacy Leaks - A Longitudinal Study of PII Leaks in Android Apps”

PhD Student Open House 2018
Meeting Purpose: Poster and Demo presentation
Meeting Start and End Dates: March 16-17, 2018 (two days)

Meeting Location: ISEC building at Northeastern University, 308 Columbus Ave, Boston, MA
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel Dubois, Jingjing Ren
Presentations Made: ReCon and the Mon(IoT)r Lab

Clic Conference: Privacy Across the Disciplines
Meeting Purpose: Panel
Meeting Start and End Dates: April 5, 2018 (one day)
Meeting Location: Dockers Hall at Northeastern University, 65 Forsyth Street
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel Dubois, Jingjing Ren
Presentations Made: ReCon and the Mon(IoT)r Lab

RSA 2018
Meeting Purpose: Demo presentation
Meeting Start and End Dates: April 16-20, 2018 (four days)
Meeting Location: Moscone Center, San Francisco
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel J. Dubois, Jingjing Ren
Presentations Made: "L-ReCon Demo"

United Kingdom Ambassador to the US
Meeting Purpose: Outreach
Meeting Date: October 25, 2018
Meeting Location: Mon(IoT)r Lab, Northeastern University
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel J. Dubois, Jingjing Ren

Internet Measurement Conference
Meeting Purpose: Conference
Meeting Date: October 31, 2018
Meeting Location: Mon(IoT)r Lab, Northeastern University
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel J. Dubois, Jingjing Ren

Undergraduate Research Night / Cybersecurity and Privacy Institute
Meeting Purpose: Conference
Meeting Date: November 5, 2018
Meeting Location: Northeastern University
Meeting Attendees from this project: Jingjing Ren

Chair of the Board of Directors of Campus France (Prof. Bertrand Monthubert)
Meeting Purpose: Outreach
Meeting Date: December 6, 2018
Meeting Location: Mon(IoT)r Lab, Northeastern University
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel J. Dubois, Jingjing Ren

NBC Nightly Films Interview
Meeting Purpose: Interview, potential to be in the national Nightly News broadcast
Meeting Date: February 27, 2019

Meeting Location: Mon(IoT)r Lab, Northeastern University
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel J. Dubois

Northeastern Open House
Meeting Purpose: Present our research to local schools
Meeting Date: March 6-8, 2019
Meeting Location: Mon(IoT)r Lab, Northeastern University
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel J. Dubois

DHS Science & Technology Showcase
Meeting Purpose: Conference, L-ReCon presentation, poster, and demo
Meeting Date: March 18-20, 2019
Meeting Location: Marriott Wardman Park, Washington DC
Meeting Attendees from this project: Prof. David Choffnes, Dr. Daniel J. Dubois

4.4 Milestones and Deliverables

Analysis of additional benefit of non-popular and foreign-country apps. We used ReCon to find PII on non-popular apps and apps in China and in India and reported our findings in a deliverable report.

The deliverable (report) for this milestone is available at [9].

Train ReCon classifier with automated tools. As output for this milestone, we designed a method for automatically training the ReCon classifier in order to reduce reliance on manual training and help keep the classifier trained as new version of apps get released. The deliverable for this milestone is available at [8].

Release PII leaks found using automated tools. We used automated tools for finding PII leaks on several versions of several Android apps, which lead to the publication of a paper at NDSS'18 [19]. The deliverable (online application for visualizing PII leaks) for this milestone is available at [20].

Software to run ReCon in local area networks. This milestone consisted in the release of L-ReCon, which is software to run ReCon on a local device. A version of the software compatible with high-end Netgear routers (R9000) and Raspberry Pi II (or newer) has been released to the public and periodically updated with new features. The deliverable (L-ReCon software) for this milestone has been documented in [14] and available for download at [7].

Software that allows decentralized crowd-sourced training. For this milestone, we have designed a method for decentralized crowd-sourced training to integrate with L-ReCon.

The deliverable for this milestone is available at [11].

Final Report for the First Year. This deliverable consists in a report summarizing our accomplishments for the first year of the project. The deliverable for this milestone (report) is available at [12].

Demonstration of IoT testbed to DHS. This milestone consisted in demonstrating the latest version of L-ReCon to DHS in the RSA conference. For the demo we used an Android Phone, a Netgear home router R9000 with a USB stick with L-ReCon installed, a Ring Doorbell, and a multicolor LED strip with MagicHome app installed.

Construction of initial IoT testbed. In this milestone we have designed our IoT test and produced

software and documentation for it. The deliverable for this milestone (report) is available at the following link [14] and the software can be download from [15].

Initial IoT PII leak analysis. For this milestone we have analyzed the impact of interaction methods and functional categories on the privacy of consumer IoT devices. The deliverable (report document) for this milestone is available at the following link [10].

Human subjects study design. We have got our IRB study approved at Northeastern Univer- sity and were able to enroll 55 people. The deliverable (user study website) for this milestone is available at the following link [13].

Analysis of PII leaks from user study. For this milestone we performed an analysis of IoT privacy leaks detected during our IRB-approved user study. The results for this milestone have been delivered in the form of a website [16] and the research paper we published at IMC '19 [18].

Publish anonymized reports from controlled study.

Reports from controlled study have been delivered with the analysis of the PII leaks from user study (see previous milestone). They have been delivered together in the form of a website [16] and the research paper we published at IMC '19 [18].

Release IoT PII detection software. In this milestone we delivered the last iteration of L-ReCon, the Mon(IoT)r testbed software, plus the additional software for the identification of audio and video and other types of activities from network traffic from IoT devices.

The deliverable (software) for this milestone is publicly available: L-ReCon [7], Mon(IoT)r testbed software [15], additional IoT analysis tools and scripts [16].

5.0 CONCLUSIONS

This project has reached all the objectives that were contractually agreed for the first two years. Despite the early termination of the project, we were able to complete some of the tasks scheduled for the third year. Such work is still in progress, but we have already developed software for block- ing privacy exposure and for mitigating privacy risk in IoT [4]. A paper, currently under review, covers the privacy risks of one of the most popular category of IoT devices: smart speakers [2].

In conclusion, during its two years of funding, this project had a large impact on the broader community:

- L-ReCon has been used for a public demo in front of 500 people at the New York Rooftop film festival in July 2017. During this demo the documentary "Harvest" [1] was shown to the public, which is based on data obtained with ReCon.
- We published four top-conference papers in the security, privacy, and Internet measurement fields (NDSS'18 [19], PETS'18 [6], IMC'19 [18], NDSS'20 [22]).
- We deployed our software in two other universities (Imperial College London, Politecnico di Torino) to further foster research and obtain even greater impact.
- We deployed our software on the premises of an industrial partner (American's Test Kitchen) and we are in the process of installing it at Consumer Report's premises. Both parters work includes testing IoT devices and help public awareness on what they do, also from a privacy point of view.

- The data we have generated from our IoT experiments in pcap format [16] (several GB of pcap files that we made publicly available) have been downloaded 59 times, facilitating follow-up work from other institutions.
- Influential people expressed interest and visited our lab. For example the Ambassador of the United Kingdom and the First Lady of Germany.
- Our work has been covered by the press in several major and minor newspapers (online and paper-based) and TV shows in different countries. Two of the most recent influential articles citing this project's work have been published by the New York Times [3] and by the Financial Times [5].

Our future efforts will be spent on continuing to tackle the growing issues of data exposure from Internet-connected devices, including finding ways for controlling the leakage of IoT devices, building profiles of IoT devices that can be used to enforce policies, and continue running experiments with newer versions (and new firmware) of existing devices with the goal for searching for any possible behavior that may suggest risk of privacy exposure.

6.0 REFERENCES

[1] K. Byrnes. Harvest documentary, July 2017. http://www.indevu.com/harvest.

[2] D. J. Dubois, R. Kolcun, A. M. Mandalari, M. T. Paracha, D. Choffnes, and H. Haddadi. When Speakers are All Years: Understanding when smart speakers mistakenly record conversations, February 2020. https://moniotrlab.ccis.neu.edu/smart-speakers-study/.

[3] K. Hill. Activate This 'Bracelet of Silence,' and Alexa Can't Eavesdrop. *The New York Times*, February 2020. https://www.nytimes.com/2020/02/14/technology/alexa-jamming-bracelet-privacy-armor.html.

[4] A. M. Mandalari, R. Kolcun, H. Haddadi, D. J. Dubois, and D. Choffnes. Towards automatic identification and blocking of non-critical iot traffic destinations, February 2020. https://arxiv.org/abs/2003.07133.

[5] M. Murgia. Smart TVs sending private data to Netflix and Facebook. *The Finantial Times*, September 2019. https://www.ft.com/content/23ab2f68-d957-11e9-8f9b-77216ebe1f17.

[6] E. Pan, J. Ren, M. Lindorfer, C. Wilson, and D. R. Choffnes. Panoptispy: Characterizing Audio and Video Exfiltration from Android Applications. In *Privacy Enhancing Technologies Symposium (PETs '18)*, 2018.

[7] Recon Project. L-ReCon: Software for Revealing Privacy Leaks in Network Traffic Locally. https://moniotrlab.ccis.neu.edu/lrecon/, month=April, year=2019,.

[8] Recon Project. Classifier Data. *ReCon Project deliverable*, October 2017. https://moniotrlab.ccis.neu.edu/recon/d1.2.pdf.

[9] Recon Project. Findings regarding non-popular and non-US apps. *ReCon Project deliverable*, October 2017. https://moniotrlab.ccis.neu.edu/recon/d1.1.pdf.

[10] Recon Project. Analysis for Controlled Experiments. *ReCon Project deliverable*, July 2018. https://moniotrlab.ccis.neu.edu/recon/d2.3.pdf.

[11] Recon Project. Decentralized Crowdsourced Training. *ReCon Project deliverable*, April 2018. https://moniotrlab.ccis.neu.edu/recon/d1.5.pdf.

[12] Recon Project. Final Report (first year). *ReCon Project deliverable*, April 2018. https://moniotrlab.ccis.neu.edu/recon/d1.6.pdf.

[13] Recon Project. Mon(IoT)r / ReCon User Study, September 2018. https://moniotrlab.ccis.neu.edu/research-study/.

[14] Recon Project. Summary of Testbed. *ReCon Project deliverable*, July 2018. https://moniotrlab.ccis.neu.edu/recon/d2.2.pdf.

[15] Recon Project. Mon(IoT)r IoT Testbed Software, April 2019. https://moniotrlab.ccis.neu.edu/tools/.

[16] Recon Project. Tools for Measuring IoT Information Exposure (IMC '19 software), April 2019. https://moniotrlab.ccis.neu.edu/tools/.

[17] Recon Project. Mon(IoT)r IoT Testbed Device List, February 2020. https://docs.google.com/spreadsheets/d/1bOeVug_-9wBVRX7_r8IPPalMKNZoyzNq0b9yRmOjsGA/edit#gid=1449063789.

[18] J. Ren, D. J. Dubois, D. Choffnes, A. M. Mandalari, R. Kolcun, and H. Haddadi. Information Exposure for Consumer IoT Devices: A Multidimensional, Network-Informed Measurement Approach. In *Proc. of the Internet Measurement Conference (IMC '19)*, 2019.

[19] J. Ren, M. Lindorfer, D. J. Dubois, A. Rao, D. R. Choffnes, and N. Vallina-Rodriguez. Bug Fixes, Improvements, ... and Privacy Leaks – A Longitudinal Study of PII Leaks Across Android App Versions. In *Network & Distributed System Security Symposium (NDSS '18)*, 2018.

[20] J. Ren, M. Lindorfer, D. J. Dubois, A. Rao, D. R. Choffnes, and N. Vallina-Rodriguez. Should you update your app?, February 2018. https://recon.meddle.mobi/appversions/tool.html.

[21] T. Van Ede, R. Bortolameotti, A. Continella, J. Ren, D. J. Dubois, M. Lindorfer, D. R. Choffnes, M. van Steen, and A. Peter. Flowprint documentation and software, February 2020. https://flowprint.readthedocs.io/en/latest/.

[22] T. Van Ede, R. Bortolameotti, A. Continella, J. Ren, D. J. Dubois, M. Lindorfer, D. R. Choffnes, M. van Steen, and A. Peter. FLOWPRINT: Semi-Supervised Mobile-App Fingerprinting on Encrypted Network Traffic. In *Network & Distributed System Security Symposium (NDSS '20)*, 2020.

LIST OF ABBREVIATIONS AND ACRONYMS

API	application programming interface
app	application
CDN	content delivery network
CNIL	Commission Nationale Informatique & Liberte?s
DHS	Department of Homeland Security
DNS	domain name service
EC2	Amazon Elastic Compute Cloud
FTC	Federal Trade Commission
GPS	global positioning system
HTTPS	secure hypertext transfer protocol
ID	identifier
IFA	information flow analysis
IMC	Internet Measurement Conference
IP	Internet protocol
IRB	institutional review board
ISP	Internet service provider
IT	information technology
IoT	Internet of things
LAN	local area network
LED	light emitting diode
L-ReCon	local ReCon
lab	laboratory
MAC	media access control
MITM	man-in-the-middle
ML	machine learning
NDSS	Network and Distributed Systems Security Symposium
OS	operating system
PETS	Privacy Enhancing Technologies Symposium
PII	personally identifiable information
RSA	RSA Data Security conference
SNI	server name indication
TLS	transport layer security

UK	United Kingdom of Great Britain and Northern Ireland
US	United States of America
USB	universal serial bus
UUID	universally unique identifier
TV	television
VPN	virtual private network