DASH: Determine And Secure the House

Sahil Gandhi, Max Wang, Daniel Achee
University of California, Los Angeles

Abstract

Your abstract text goes here. Just a few facts. Whet our appetites. Not more than 200 words, if possible, and preferably closer to 150.

1 Introduction

2 Background and Motivations

In recent years there has been a proliferation of voice assistants in people's homes. These useful IoT devices provide a convenient voice interface for users, but also present some privacy concerns as they are an always-on listening device. Companies selling these products, such as Amazon, claim that the devices are constantly listening for a keyword, such as Alexa, and only after hearing this keyword, record and process what users say. Therefore, anything a user says before a keyword is issued and after their voice transaction has occurred is not recorded and private. Unfortunately, there is no easy way for a user to verify that the device is upholding public statements and the privacy agreement. Companies do admit that what a user says following this keyword is used for targeting advertisements and customizing content. This creates a conflict of interest as there is a financial incentive for companies to collect more voice data on their users and abuse their agreements.

These devices are relatively new and have emerged due to advances in natural language processing. These advances have allowed for fairly accurate keyword detection at the edge device in real time without the need to send the data to the cloud. With the increased accuracy of natural language processing and the ability to pull more semantic meaning from phrases, these devices are able to provide services that make them attractive to consumers.

There are many players in the market, with the most popular being Amazon Alexa, Google Home Assistant, Apple Siri, and Microsoft Cortana. There is no standardization amongst voice assistants, with each using different protocols and traffic patterns. These devices typically encrypt the traffic. This encryption is a double edged sword as it both protects a user from attackers sniffing their data, but also prevents the user from auditing the behavior of the device.

Given that voice assistants appear to be only expanding in popularity and are projected to be a 7840.82 million dollar market by 2023, it is imperative that users, even those that are not technically sophisticated, can easily audit and understand the behavior of their devices.

3 Related Work

Given that the data being transmitted across IoT devices is encrypted (end-to-end) it is quite difficult to discern what exactly is being transmitted, and even complicated man-in-the-middle style attacks can't be used since the data payload format and security cert/auth format are unknown to a random attacker. Despite these hindrances, there has been some work in the field by bloggers and researchers to capture the information being transmitted and infer some patterns.

Blog posts such as [1] and [2] go into great detail for setting up man-in-the-middle style proxies to capture all data being sent across the network by an IoT device (a Google Home and an Amazon Echo respectively for these blogs) and use an application such as Wireshark to sniff packets and determine whether they originated from an IoT device. We use and build off of these techniques to capture data from our own IoT devices during our experiments and for creating our deliverables mentioned below. Furthermore, these blog posts usually capture data for a short duration, minutes or an hour at most, but we wish to extend these investigations further into the tens of hours or several days timeframes. Every individual in the United States of America has on average several internet connected devices. These can range from laptops/cellphones to smart IoT connected devices, and determining what is an IoT device and what is not is imperative for our intended system to work. We build off of some of the research done in the following paper: [3] for classifying certain network streams as IoT devices and also determining the port numbers for IoT traffic so that they can be reused/saved across multiple sessions.

Another interesting paper in this space is [4] which aims to classify events in different IoT streams, in particular focusing on classifying when a Nest Camera is live streaming data vs motion detecting, when an Alexa is transmitting data back to the server, and a couple other applications. This is closely related to the classifications we intend to do, except at the current implementation, we wish to classify only smart assistants (and not all generic IoT devices) and we also wish to detect other behaviors within the IoT devices such as transmitting data during non-verbal sessions or other anomalies in the network traffic. The paper also only does some short polling of about 900 seconds for capturing data, and we want to increase that timeframe to something much longer to get a larger sense of what an IoT device is doing. This paper also looks into using DNS queries to map particular network traffic to an IoT device, and we intend to build off these findings to create our own mappings.

Lastly, as more modules ("skills" in the Alexa world) are created, our smart assistants become ever smarter and are able to do more, include such sensitive tasks like writing emails, calling other individuals, and more. The following paper, [5], determines the relative ease with which a hacker can mimic a user and send off malicious requests. Even someone without malicious intents can take advantage of the information and skills of a voice assistant and breach a user's privacy. Since 2014, when the previous paper was published, Google, Amazon and others have taken the initiative to alert users after their device was accessed (and send a transcript/recording of the voice data) and protect some sensitive information, the casual user is not aware of the repercussions of purchasing/using the device. We hope that our monitoring tool can serve as this "wake-up" call that the smart assistant is a very versatile device that should be treated with the same level of privacy and scrutiny that we treat our laptops/phones and monitoring what it is doing should be easy enough for the common non-technical person to set up on their own.

4 System Overview

4.1 Design

5 Evaluation

6 Discussion and Future Work

7 Conclusion

References