Alexa, How Secure Are You?

Reverse Engineering and Protocol Analysis of the Amazon Echo

Teresa T. Tseng[1], Aliza Isaacs[1], Jan Janak[2], and Prof. Henning G. Schulzrinne[2]

Barnard College[1], Internet Real-Time Laboratory, Columbia University[2]

Email: teresa.tseng@columbia.edu, janakj@cs.columbia.edu

Introduction

Internet of Things (IoT)

- IoT: a network of computers that interact with each other and with the physical world
- These devices often send data to the cloud for processing and storage
- Who has access to the data?

In this research project, we aim to evaluate authentication frameworks used by cloud-dependent IoT devices. We picked Amazon Echo as the first device to hack for its widespread use and potentially privacy-invasive nature.

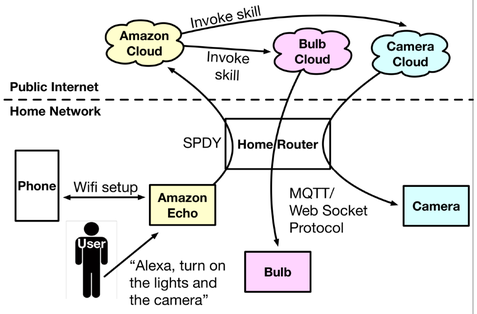

Fig. 1. *Amazon Echo and IoT device network ecosystem, illustrating the interactions between devices and cloud components. The Echo does not communicate directly with the other devices.*

Goals

The aim of this project is to gain an understanding of the mechanisms and protocols by which the Amazon Echo communicates with local devices as well as Amazon cloud services.

Approach

Overview:

1. Make the Amazon Echo susceptible to Man-In-The-Middle (MITM) attacks (Figs. 2,3)
2. Launch a MITM attack on the Echo (Fig. 4.)
3. Record, decrypt, and analyze the communication between the Echo, the cloud, and other IoT devices (Fig. 4.)

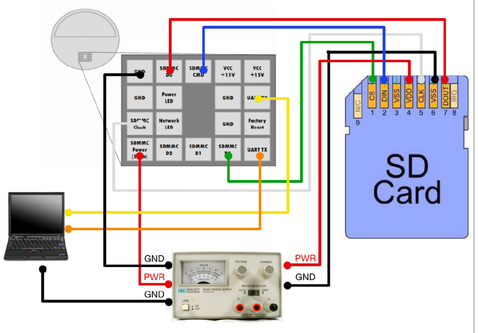

Fig. 2. *Hardware setup used to obtain write access to the Echo's file system. We used the technique developed by authors in [1,2]. Used to make the Echo trust a custom CA certificate generated by the MITM proxy. (adapted from [2])*

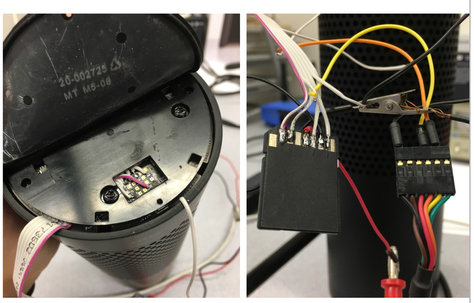

Fig. 3. *A rooted ~~HAL 9000~~ Amazon Echo. Debug pads visible in the left picture, SD card and wire connections shown on the right.*

System Architecture

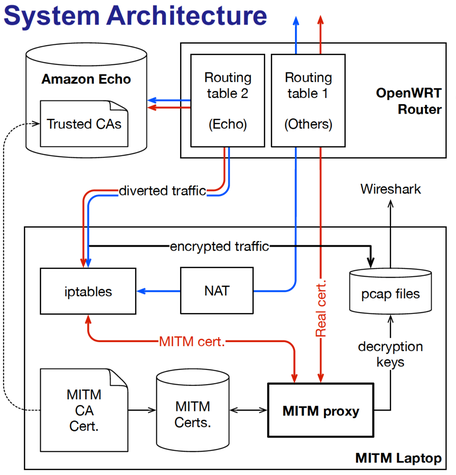

Fig. 4. *IoT device traffic capture and decryption architecture.*

Transparent Proxying and Traffic Capture Architecture

- Encrypted traffic from the Echo is diverted to the laptop running a MITM proxy
- TLS/SSL connections from the Echo are terminated by the proxy. Other traffic is forwarded unmodified
- Encrypted traffic is recorded together with decryption keys for later analysis

Techniques Involved

- Policy routing diverts traffic from the Amazon Echo to the MITM host
- Iptables rules transparently redirect selected traffic to the MITM proxy
- MITM proxy generates a custom certificate for each TLS/SSL connection

Preliminary Results

We analyzed three major features of the Echo: voice assistant, IoT device control, and intercom. We successfully decrypted all TLS/SSL traffic generated by the device.

Preliminary Findings:

- No certificate pinning (vulnerable against nation state actors)
- Only sends audio upon activation by the wake word
- All communication is encrypted
- Uses the SPDY and SIP protocols

Voice Assistant & IoT Device Control

- Maintains a persistent TLS/SSL connection to the cloud
- Uses the SPDY protocol (predecessor of HTTP/2)
- Echo only communicates with IoT devices via the cloud

Intercom & Voice Calling

- Sets up a call between two devices ("Alexa, drop in on Bob")
- Based on open protocols: SIP, RTP, STUN, ICE
- Voice data is end-to-end encrypted

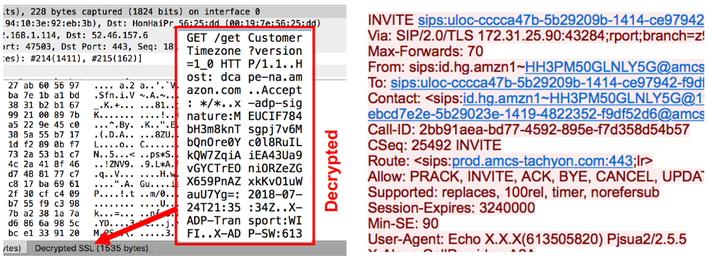

Fig. 5. *Decrypted SSL data (left) and decrypted intercom SIP signaling (right)*

Conclusions and Future Work

- Publish traffic recordings along with decryption keys to the Crawdad traffic repository [3]
- Further analyze the authentication mechanisms used by the Echo
- Rethink the procedure for setting up larger numbers of Echo devices
- Disseminate the results of this analysis in a technical report

Many thanks to the Henry Luce Foundation, Clare Booth Luce Scholars Program for funding this research. Credits go to Mark Barnes, Ike Clinton, Lance Cook, and Shankar Banik for pioneering work on hacking the Echo.

References

1. Barnes, Mark. "Alexa, Are You Listening?" *MWR Labs*, MRW Infosecurity, 1 Aug. 2017, labs.mwrinfosecurity.com/blog/alexa-are-you-listening.
2. Clinton, Ike, et al. *A Survey of Various Methods for Analyzing the Amazon Echo*. The Citadel, The Military College of South Carolina.
3. Crawdad https://crawdad.org/ Dartmouth College.