Who's Listening? Exploring the Security Vulnerabilities of the Amazon Echo Dot

Keerthi Naredla, Jayati Dev
Indiana University, Bloomington

Submitted in Requirement of I520/I544: Security of Networked Systems

1. Introduction

'Alexa, What's the weather today?'
'Alexa, Set me a 5 minute timer.'
'Alexa, Play me a Song.'
'Alexa, buy Only Time will Tell by Jeffrey Archer.'

Voice-controlled assistants are increasingly becoming a part of our homes and lives. Imagine going hands-free for most of your digital chores, with devices costing as low as $15. It comes as no surprise when statistics reveal that there were about 10.7 million IoT(Internet of Things) devices sold by Amazon running on the Alexa voice assistant. However, previous research has suggested a number of potential security vulnerabilities exist in IoT devices like Man-in-the-Middle attacks, IP Spoofing, Server SYN Flooding and Denial-of-Service attacks, many of which have been successfully performed and serve as current topics for exploration. Our goal in this study is to study the various types of attacks and verify whether they are feasible on the Second Generation Amazon Echo Dot. We find that Alexa's strong architecture partially or completely mitigates most of the attacks, except a few variants of sound attacks and possible Denial-of-Service attacks.

We further study successful attacks on the previous software versions of Alexa and the First Generation Echo Dot to compare with classroom knowledge. We describe our approach and methods, and also future scope of attacks with a theoretical perspective. Finally, we describe the Echo Dot hardware and software blocks and successful sound attacks at the application level of the software.

The rest of the paper is divided as follows - Section 2 describes the network architecture and the hardware description of the Echo Dot which we have found after disassembling the product. Section 3 reports previous attacks and mitigations. Section 4 is the core of our report which describes the different variations of sound attacks that we have performed on the Echo Dot. Section 5 describes conclusion.

2. Background

Confidentiality, Integrity and Availability form the three pillars of Computer Security. Any device that claims to be secure must follow these criteria, including voice assistants like Alexa, Siri, Cortana or Google. Voice assistants are increasingly being used nowadays for IoT devices like the Amazon Echo, Google Home, Kenmore Alfie, and similar devices. In our study, we explore the Amazon Echo Dot second generation, that runs on Amazon's Alexa voice assistant. We try to explore and replicate the different types of attacks that are possible on the protocol stack consisting of the application layer, transport layer, network layer and link layer. As a result, we find that the Amazon Echo Dot has a very robust architecture, with only application layer threats through sound attacks. We also study the network traffic using *WireShark*, which is a network scanning tool and analyze the different IP addresses for different servers to leave scope for future Denial of Service attacks. We also use *NetworkMiner 2.2*, which is a network forensics tool that we used for reverse IP lookup to get a better understanding of what each of these protocols mean.

2.1 Software Architecture

A conceptual model of architecture of the Amazon Echo Dot was proposed by Haack et al. that has been shown in Figure 1 below.

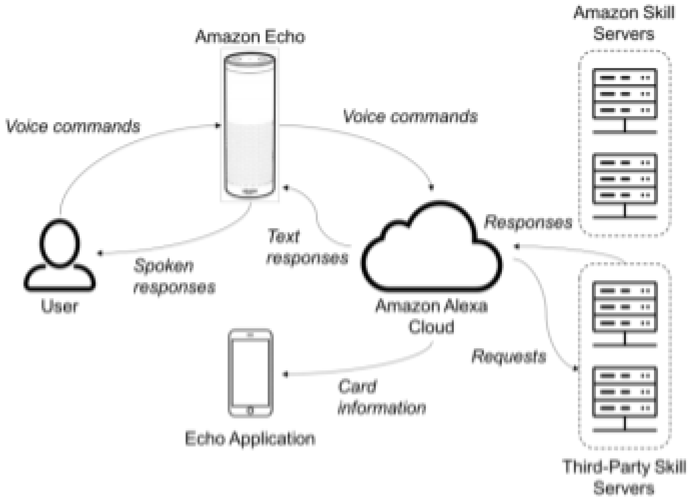


Figure 1: Architecture of the Echo Dot by Haack et al.[1]

When a user issues commands to Alexa enabled device, the speech is streamed to the Alexa service in the Amazon Web Service cloud. The Alexa Voice Service recognizes the speech, determines what the user wants, and then sends a structured request to a particular skill, which is hosted by AWS cloud in Alexa Skills Store. Thus, all speech recognition and conversion is handled by Alexa in the cloud and response is sent back to the device. Not only this, the user profile,purchase history, conversation history with Alexa enabled devices are stored and processed in AWS cloud, to provide personalized user experience. As a point to note, Amazon deploys more than 25000 skills for users of AVS on their IoT devices and each new skill should meet Amazon’s security requirements and policy guidelines. These security policies differ for each skill based on where the service for skill is hosted either AWS Lamada or own host, and type of user information that skill has access to.

Our work on replication of attacks, is more concerned with Alexa Voice Services, which is an intelligent voice control consisting of features like Natural Language understanding, Automatic Speech Recognition and Text to Speech conversion. The skills developed and deployed using Alexa skill kit comes to live with Alexa voice service. Hence, AVS serves as the core for Amazon echo dot or any other Alexa enabled devices, in terms of service as wells as user security.

2.2 Network Architecture

As can be seen in Figure 2, we have modified the network flow to include a Raspberry Pi which acts as the WiFi router. The purpose of this is to execute WireShark on the router traffic so that we can study the data transfer from and to the Echo Dot as well as the Alexa application, but do not pick up other IP addresses which might have been a part of the network. The IP addresses in blue represent the IP addresses of the respective devices and servers.

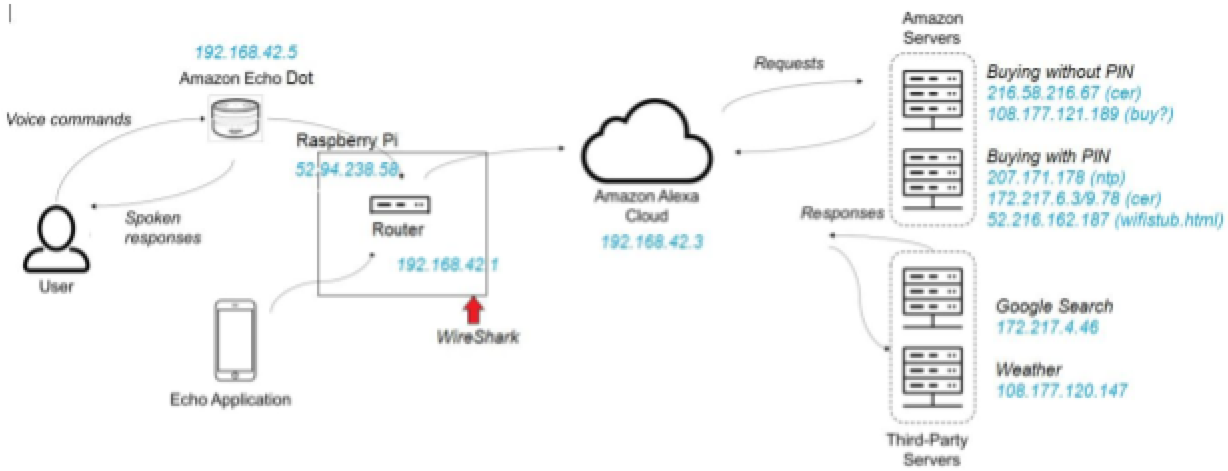


Figure 2: Modification in the model by Haack et al. [1]

2.3 Hardware Architecture

Surprisingly, after interaction with the device, we found significant hardware changes that Amazon has made to the Echo Dot which mitigated many of the previous attacks [2] done on the device. The Echo Dot's circuit boards are divided into two halves. The upper half consists of the microphone, the LEDS and the supporting circuits. It also contains analog-to-digital converter ICs for converting speech into electrical signals. A ribbon cable that connects the upper half to the lower half transfers these signals to the processor in the lower half.

The lower half of the circuit contains the major components - the processor, RAM, memory chip for storage, the power management module, among other components, many of which have not been described in the manual. The lower half also contains a mini USB port for power supply and a 3mm port for connection with external speakers. The changes in the hardware made in the lower half include -

a. **Removal of test points.** In the previous version of the Echo Dot as can be seen in Figure 3, a test port was left open for testing by developers. However, the second generation of the Dot does not have the same. In fact, all of the test ports have been removed except a few for continuity checking.
b. **Housing the processor, RAM and flash memory.** In the first version, these modules were directly embedded on the PCB without any separate housing. The current hardware upgrade includes housing the same using a metal container to prevent side-channel attacks or emf leakage.
c. **Change in the processor.** The processor has been changed to a faster and cheaper MEDIATEK processor in Dot v2.
d. **Change in flash memory.** The current version does not have the Samsung flash memory used earlier, but now includes a 4GB Micron chip, which is cheaper. Also, there are no data sheets available for the memory chip, unlike Samsung's, which makes it difficult for an attacker trying a hardware exploit.
e. **Change in the RAM.** The RAM is still 256MB, but has been upgraded from DDR to DDR3.

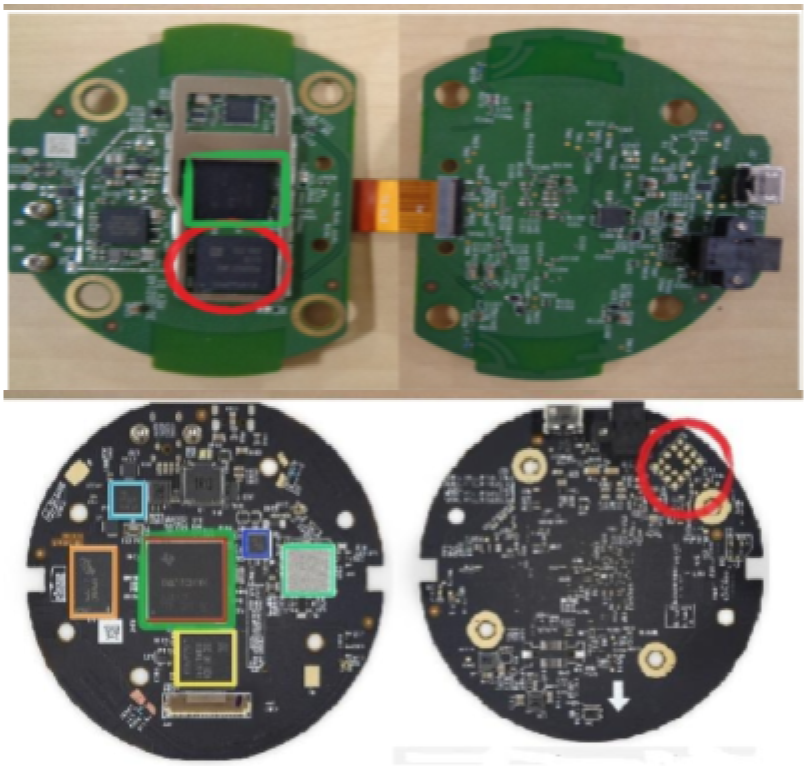


Figure 3: (a) Lower half of Echo Dot v2 showing the processor, memory chip (b) Other side of the lower half containing the ports

(c) Lower half of Echo Dot v1 showing the different components (d) Other side with the test port shown in red circle

3. Previous Work

As described above, the first generation of Echo Dot had a few security vulnerabilities, mostly due to its architecture. One of the first research works on the security vulnerabilities of the Alexa service explored the Amazon Echo v1 [1] and could perform sound attacks, which we will show to still exist in the current version of the Dot in the following sections. A different variant of the sound attack, which exploited network vulnerability was done by Sam Machin, who later also developed the web application for Amazon Alexa. It used a cell phone number linked to the Alexa web server to call into the Amazon Echo [7]. This vulnerability has since been converted into a *'Drop In'* feature in the Echo devices and subsequent software changes have removed it.

Another attack done on the Echo Dot specifically was a physical attack on the hardware that could successfully root the device and even install a malware. It was used to turn the Dot into a 'wiretapping' device and even access payment tokens [6]. Amazon has mitigated this problem by removing all the test ports which could be used to mount an external SD card.

4. Methods and Findings

Part I: Replication of Attacks

a. Sound Attacks

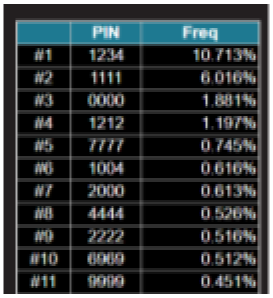
i. Default Configuration

Echo Dot usually comes with a default configuration which does not require a PIN for activation. The Alexa mobile application requires the user to enter Amazon account credentials for signing up, which means that the user's Amazon account gets linked, and any purchase can be made which would not require user authentication. This is a departure from natural authentication systems, where every purchase requires authentication. Also, the Echo Dot takes commands without verifying *'who'* or *'what'* initiated it. Thus, any human or mechanical voice can initiate a voice command any number of times, successfully.

ii. Voice PIN

In the second generation of Alexa devices, a Voice PIN was introduced for purchases, specifically any command that could add, delete or modify a purchase. When the Voice PIN is activated through the Alexa application, the user needs to speak a 4-digit numeric PIN to the Echo Dot to place/cancel an order. There are two attempts allowed every time an order is placed. However, the order can be placed any number of times. One-third of voice PINs can thus be broken using 61 combinations, which is just placing the order 30 times .

In our experiment, we required only two attempts to break the Voice PIN which the other person set, because the PIN was 1234. Consequently, most PINs can be narrowed down if we know some background of the user like memorable dates (birthdays, graduation dates, etc.) or their university or school ID numbers. Figure 4 shows a table of commonly used PINs and their frequencies.



	PIN	Freq
#1	1234	10.713%
#2	1111	6.016%
#3	0000	1.881%
#4	1212	1.197%
#5	7777	0.745%
#6	1004	0.616%
#7	2000	0.613%
#8	4444	0.526%
#9	2222	0.516%
#10	6969	0.512%
#11	9999	0.451%

Figure 4: List of commonly used PINs

iii. Speaker Recognition

The Speaker recognition is the newly built feature in Alexa Voice Service, called ‘Your Voice’. Amazon started working on it since 2015 and launched it in mid 2017. The main aim of Amazon in launching the feature was to give customers personalized user experience. This associate your voice with a specific voice ID. For each voice ID, a cloud storage space is allocated to keep track of the user history and preferences. Thus, only a specific user can place orders and Alexa can keep track of who is placing what. Although it provides speaker recognition and user authentication, its security effectiveness is yet to be explored. Since it is a machine learning model, if the user activity is less, then speaker recognition model might be weak, making it vulnerable to voice - impersonation attacks. Additionally, if the user’s own voice is recorded and replayed, it will still be a physical attack on the system.

b. Replay Attack

Replay Attack can be done in two ways - one is repeating the packets in the network and the other is by replaying the voice. Although Amazon doesn’t explicitly contradict the occurrence of duplicate packets, it doesn’t respond either which shows that Amazon has successfully mitigated this kind of replay attack.

The other type is replay voice which is repeating or imitating the owner voice which is also called voice impersonation attack. This is a threat to all the voice-enabled devices, voice biometric authentication and specifically IoT devices which more or less have voice as security interface.

A voice impersonating can be done in two ways, one machine-based voice which can be either by recording the owner voice, or by using modulators a owner voice can be generated from attacker voice command and the second one is live imitation of the owner voice. There are number of ways proposed to mitigate, detect such voice attacks.

Defense system for machine-based voice attack. For instance a machine generated voice can be detected with help of a magnetometer detecting the magnetic field that is generated by any conventional speakers. This alone might fail in case if the sound producing device is far away from the voice-enabled distance in such a way that magnetic field cannot be detected[3]. Using different use cases like sound source distance verification, sound field verification and,loudspeaker detection verification components they trained a binary classifier, Sound Field model using Support Vector Machine. This showed the ability to detect machine-generated voice on smartphone, with zero false rate. This can be further extended to voice-enabled devices like Amazon Echo Dot.

Defense system for user-based voice attack. An attacker imitating owner voice can be detected with help of state-of-the art Automatic Speaker Verification System like open-sourced Bob Spear verification toolbox developed by Khoury et al[5],which has been recognized for its performance in detecting against human-based impersonation attack. This implements machine learning algorithms to train the model to detect the speech with help of extracted features from the voice. Also deep learning methods to identify unique spectral and prosodic features of a user's voice is crucial.

Another study in this field, is done by collecting synchronous speech and ElectroGlottoGraphic, which provides a direct representation of the vocal fold vibration patterns for a given voice input, to analyze glottal and vocal tract measures including F0, speech rate, vowel formant frequencies, and timing characteristics of the vocal folds. This analysis confirmed that the impersonators modulated all four parameters in producing the voices, and provides a lower bound on the scale of variability that is available to impersonators. A no-reference objective metric which is formulated based on the vowel-dependent variance of the formants associated with each voice has the ability to detect impersonators from actual owners voice.[4]

Defense System for both machine-based and user-based. Continuous authentication VAuth is simple and powerful way to prevent both the above attacks. This enables the speaker/owner to continuously authenticate to the voice-enabled device devise, or in other words the voice-enabled device responds only if the speaker wear VAuth enabled device which can be eye glasses, watch , necklace etc.. This novel idea is based on mapping body surface vibrations with voice of the user. This proved to be effective regardless of VAuth's position, mobility and user's accent.With this a machine generated and voice imitation attack becomes very difficult as it requires the owner body as well as VAuth wearable device.[2]

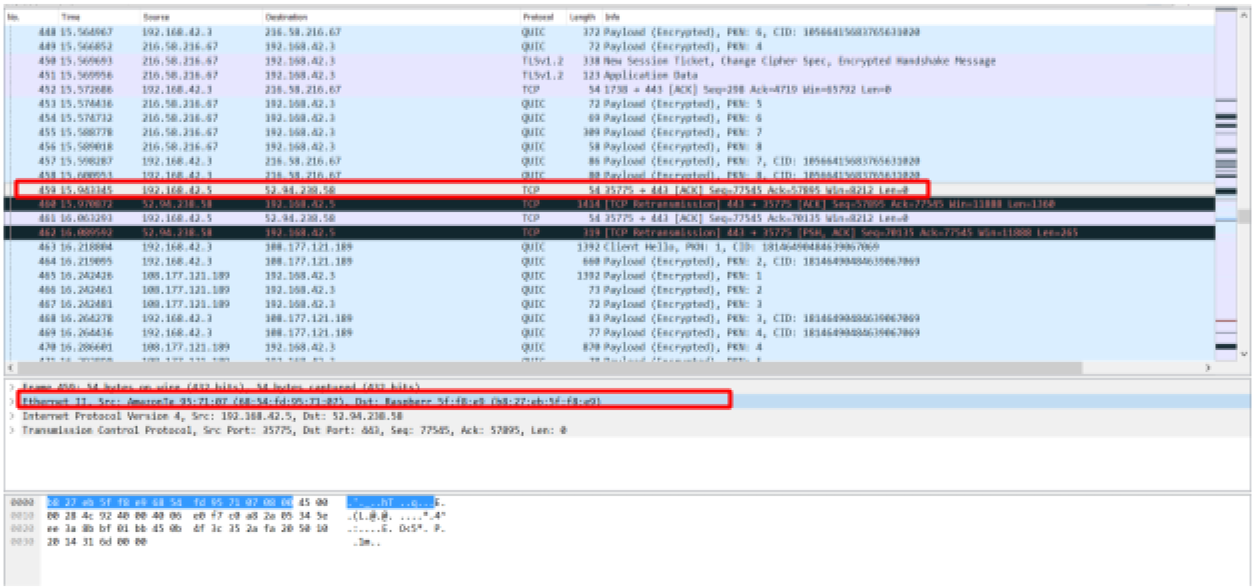
Apart from the necessity that these technologies must be implemented in Alexa voice Service, it is also important to increase the level of user authorization.

c. Dolphin Attack

In dolphin attack, inaudible ultrasound is used to control voice enabled devices. This is done by modulating the usual voice commands with high frequencies into an inaudible ultrasound of range. Since the devices are made receptive to inaudible ultrasounds for other purposes like pairing other devices etc, which is now a loophole for the attackers to launch this attack on any voice controlled devices like Siri.

We were successful in replicating this attack for the Amazon Echo Dot as well, with audio speed reduced to 0.67% and pitch increased to 800%, which was a bit within the borders of audible range, so it was still feebly audible. However, it is yet to be tested with oscillation output to generate actual ultrasonic waves since we have used *Audacity* to generate high frequency tones which might skew the actual tones.

Part II: Network Analysis and Potential Attacks



```
No.   Time        Source           Destination      Protocol  Length Info
  448 15.564967   192.168.42.3     216.58.216.67    QUIC       372 Payload (Encrypted), PKN: 6, CID: 10566415683765631020
  449 15.566852   216.58.216.67    192.168.42.3     QUIC        72 Payload (Encrypted), PKN: 4
  450 15.569693   216.58.216.67    192.168.42.3     TLSv1.2    338 New Session Ticket, Change Cipher Spec, Encrypted Handshake Message
  451 15.569956   216.58.216.67    192.168.42.3     TLSv1.2    123 Application Data
  452 15.572688   192.168.42.3     216.58.216.67    TCP         54 1738 → 443 [ACK] Seq=298 Ack=4719 Win=65792 Len=0
  453 15.574436   216.58.216.67    192.168.42.3     QUIC        72 Payload (Encrypted), PKN: 5
  454 15.574732   216.58.216.67    192.168.42.3     QUIC        69 Payload (Encrypted), PKN: 6
  455 15.588778   216.58.216.67    192.168.42.3     QUIC       309 Payload (Encrypted), PKN: 7
  456 15.589018   216.58.216.67    192.168.42.3     QUIC        58 Payload (Encrypted), PKN: 8
  457 15.598287   192.168.42.3     216.58.216.67    QUIC        86 Payload (Encrypted), PKN: 7, CID: 10566415683765631020
  458 15.600953   192.168.42.3     216.58.216.67    QUIC        88 Payload (Encrypted), PKN: 8, CID: 10566415683765631020
  459 15.943345   192.168.42.5     52.94.238.58     TCP         54 35775 → 443 [ACK] Seq=77545 Ack=57895 Win=8212 Len=0
  460 15.970072   52.94.238.58     192.168.42.5     TCP       1434 [TCP Retransmission] 443 → 35775 [ACK] Seq=57895 Ack=77545 Win=11888 Len=1360
  461 16.063293   192.168.42.5     52.94.238.58     TCP         54 35775 → 443 [ACK] Seq=77545 Ack=70135 Win=8212 Len=0
  462 16.089592   52.94.238.58     192.168.42.5     TCP        319 [TCP Retransmission] 443 → 35775 [PSH, ACK] Seq=70135 Ack=77545 Win=11888 Len=265
  463 16.218884   192.168.42.3     108.177.121.189  QUIC      1392 Client Hello, PKN: 1, CID: 18146490484639067069
  464 16.219095   192.168.42.3     108.177.121.189  QUIC       660 Payload (Encrypted), PKN: 2, CID: 18146490484639067069
  465 16.242426   108.177.121.189  192.168.42.3     QUIC      1392 Payload (Encrypted), PKN: 1
  466 16.242461   108.177.121.189  192.168.42.3     QUIC        73 Payload (Encrypted), PKN: 2
  467 16.242481   108.177.121.189  192.168.42.3     QUIC        72 Payload (Encrypted), PKN: 3
  468 16.264278   192.168.42.3     108.177.121.189  QUIC        83 Payload (Encrypted), PKN: 3, CID: 18146490484639067069
  469 16.264436   192.168.42.3     108.177.121.189  QUIC        77 Payload (Encrypted), PKN: 4, CID: 18146490484639067069
  470 16.286601   108.177.121.189  192.168.42.3     QUIC       870 Payload (Encrypted), PKN: 4

> Frame 459: 54 bytes on wire (432 bits), 54 bytes captured (432 bits)
> Ethernet II, Src: AmazonTe_95:71:07 (68:54:fd:95:71:07), Dst: Raspberr_5f:f8:e0 (b8:27:eb:5f:f8:e0)
> Internet Protocol Version 4, Src: 192.168.42.5, Dst: 52.94.238.58
> Transmission Control Protocol, Src Port: 35775, Dst Port: 443, Seq: 77545, Ack: 57895, Len: 0

0000  b8 27 eb 5f f8 e0 68 54  fd 95 71 07 08 00 45 00   .'._..hT ..q...E.
0010  00 28 4c 92 40 00 40 06  e0 f7 c0 a8 2a 05 34 5e   .(L.@.@. ....*.4^
0020  ee 3a 8b bf 01 bb 45 0b  4f 3c 35 2a fa 20 50 10   .:....E. O<5*. P.
0030  20 14 31 6d 00 00                                   .1m..
```

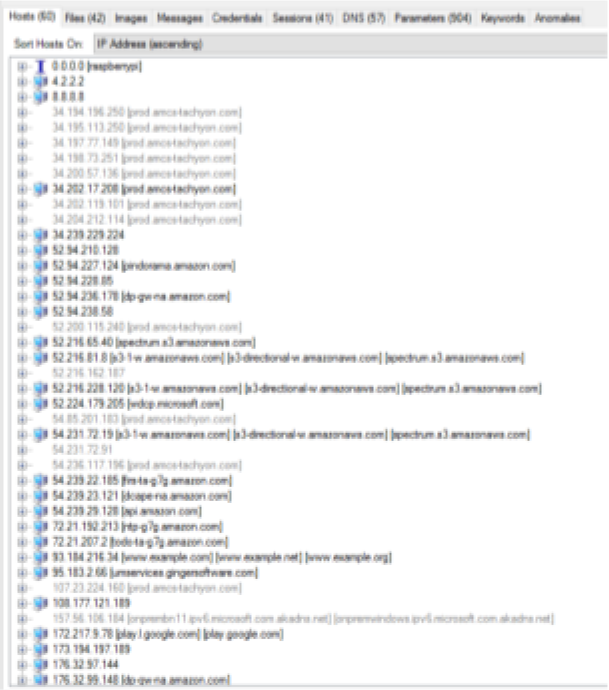
Figure 5: WireShark data for buying items on Amazon.com using Amazon Echo Dot

We first used a Raspberry Pi to connect it to a router to set it up as a WiFi access point and then run WireShark on it as described in Figure 2. This gave us traffic only pertaining to the Echo Dot. Figure 5 shows a snapshot of traffic through WireShark in the network. We collected these data for:
I. Placing an order (with and without password)
Ii. Cancelling an order (with and without password)
Iii. Checking the weather
Iv. Searching on Google

V. Re-ordering the shipment

We then used NetworkMiner 2.2 to run forensics on the obtained pcap data. Thus, we found that there were certain reserved IP addresses for Google or Amazon Web Services, but there were certain IP addresses that always went to the specific servers, For example, 108.177.120.147 was always the destination IP whenever we looked up weather information. Thus, we could conclude this particular Google server to be relaying weather data. Similarly, authentication information always passed through https://device-metrics-us.amazon.com which has its own login page that can be potentially exploited.

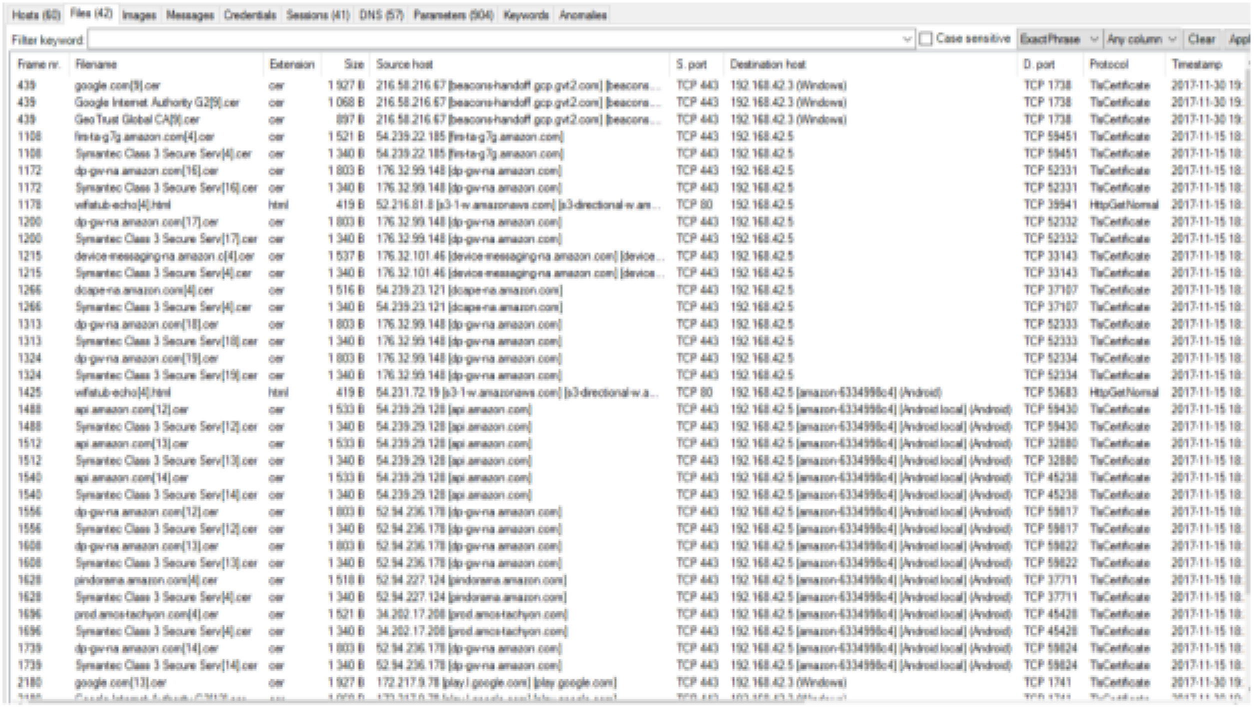
However, all the packets going from or to the Echo Dot was all encrypted. Figure 6 shows the host IPs in the NetworkMiner tool.



Sort Hosts On: IP Address (ascending)

```
0.0.0.0 [raspberrypi]
4.2.2.2
8.8.8.8
34.194.196.250 [prod.amcs-tachyon.com]
34.195.113.250 [prod.amcs-tachyon.com]
34.197.77.149 [prod.amcs-tachyon.com]
34.198.73.251 [prod.amcs-tachyon.com]
34.200.57.136 [prod.amcs-tachyon.com]
34.202.17.208 [prod.amcs-tachyon.com]
34.202.119.101 [prod.amcs-tachyon.com]
34.204.212.114 [prod.amcs-tachyon.com]
34.239.229.224
52.94.210.128
52.94.227.124 [pindorama.amazon.com]
52.94.228.85
52.94.236.178 [dp-gw-na.amazon.com]
52.94.238.58
52.200.115.240 [prod.amcs-tachyon.com]
52.216.65.40 [spectrum.s3.amazonaws.com]
52.216.81.8 [s3-1-w.amazonaws.com] [s3-directional-w.amazonaws.com] [spectrum.s3.amazonaws.com]
52.216.162.187
52.216.228.120 [s3-1-w.amazonaws.com] [s3-directional-w.amazonaws.com] [spectrum.s3.amazonaws.com]
52.224.179.205 [wdcp.microsoft.com]
54.85.201.183 [prod.amcs-tachyon.com]
54.231.72.19 [s3-1-w.amazonaws.com] [s3-directional-w.amazonaws.com] [spectrum.s3.amazonaws.com]
54.231.72.91
54.236.117.196 [prod.amcs-tachyon.com]
54.239.22.185 [fire-ta-g7g.amazon.com]
54.239.23.121 [dcape-na.amazon.com]
54.239.29.128 [api.amazon.com]
72.21.192.213 [ntp-g7g.amazon.com]
72.21.207.2 [todo-ta-g7g.amazon.com]
93.184.216.34 [www.example.com] [www.example.net] [www.example.org]
95.183.2.66 [umservices.gingersoftware.com]
107.23.224.160 [prod.amcs-tachyon.com]
108.177.121.189
157.56.106.184 [onpremipv11.ipv6.microsoft.com.akadns.net] [onpremwindows.ipv6.microsoft.com.akadns.net]
172.217.9.78 [play.l.google.com] [play.google.com]
173.194.197.189
176.32.97.144
176.32.99.148 [dp-gw-na.amazon.com]
```

Figure 7: NetworkMiner list of hosts



Frame nr.	Filename	Extension	Size	Source host	S. port	Destination host	D. port	Protocol	Timestamp
439	google.com[9].cer	cer	1 927 B	216.58.216.67 [beacons-handoff.gcp.gvt2.com] [beacons…	TCP 443	192.168.42.3 (Windows)	TCP 1738	TlsCertificate	2017-11-30 19:
439	Google Internet Authority G2[9].cer	cer	1 068 B	216.58.216.67 [beacons-handoff.gcp.gvt2.com] [beacons…	TCP 443	192.168.42.3 (Windows)	TCP 1738	TlsCertificate	2017-11-30 19:
439	GeoTrust Global CA[9].cer	cer	897 B	216.58.216.67 [beacons-handoff.gcp.gvt2.com] [beacons…	TCP 443	192.168.42.3 (Windows)	TCP 1738	TlsCertificate	2017-11-30 19:
1108	firs-ta-g7g.amazon.com[4].cer	cer	1 521 B	54.239.22.185 [firs-ta-g7g.amazon.com]	TCP 443	192.168.42.5	TCP 59451	TlsCertificate	2017-11-15 18:
1108	Symantec Class 3 Secure Serv[4].cer	cer	1 340 B	54.239.22.185 [firs-ta-g7g.amazon.com]	TCP 443	192.168.42.5	TCP 59451	TlsCertificate	2017-11-15 18:
1172	dp-gw-na.amazon.com[16].cer	cer	1 803 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52331	TlsCertificate	2017-11-15 18:
1172	Symantec Class 3 Secure Serv[16].cer	cer	1 340 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52331	TlsCertificate	2017-11-15 18:
1178	wifistub-echo[4].html	html	419 B	52.216.81.8 [s3-1-w.amazonaws.com] [s3-directional-w.am…	TCP 80	192.168.42.5	TCP 39941	HttpGetNormal	2017-11-15 18:
1200	dp-gw-na.amazon.com[17].cer	cer	1 803 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52332	TlsCertificate	2017-11-15 18:
1200	Symantec Class 3 Secure Serv[17].cer	cer	1 340 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52332	TlsCertificate	2017-11-15 18:
1215	device-messaging-na.amazon.c[4].cer	cer	1 537 B	176.32.101.46 [device-messaging-na.amazon.com] [device…	TCP 443	192.168.42.5	TCP 33143	TlsCertificate	2017-11-15 18:
1215	Symantec Class 3 Secure Serv[4].cer	cer	1 340 B	176.32.101.46 [device-messaging-na.amazon.com] [device…	TCP 443	192.168.42.5	TCP 33143	TlsCertificate	2017-11-15 18:
1266	dcape-na.amazon.com[4].cer	cer	1 516 B	54.239.23.121 [dcape-na.amazon.com]	TCP 443	192.168.42.5	TCP 37107	TlsCertificate	2017-11-15 18:
1266	Symantec Class 3 Secure Serv[4].cer	cer	1 340 B	54.239.23.121 [dcape-na.amazon.com]	TCP 443	192.168.42.5	TCP 37107	TlsCertificate	2017-11-15 18:
1313	dp-gw-na.amazon.com[18].cer	cer	1 803 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52333	TlsCertificate	2017-11-15 18:
1313	Symantec Class 3 Secure Serv[18].cer	cer	1 340 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52333	TlsCertificate	2017-11-15 18:
1324	dp-gw-na.amazon.com[19].cer	cer	1 803 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52334	TlsCertificate	2017-11-15 18:
1324	Symantec Class 3 Secure Serv[19].cer	cer	1 340 B	176.32.99.148 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5	TCP 52334	TlsCertificate	2017-11-15 18:
1425	wifistub-echo[4].html	html	419 B	54.231.72.19 [s3-1-w.amazonaws.com] [s3-directional-w.a…	TCP 80	192.168.42.5 [amazon-6334998c4] (Android)	TCP 53683	HttpGetNormal	2017-11-15 18:
1488	api.amazon.com[12].cer	cer	1 533 B	54.239.29.128 [api.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59430	TlsCertificate	2017-11-15 18:
1488	Symantec Class 3 Secure Serv[12].cer	cer	1 340 B	54.239.29.128 [api.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59430	TlsCertificate	2017-11-15 18:
1512	api.amazon.com[13].cer	cer	1 533 B	54.239.29.128 [api.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 32880	TlsCertificate	2017-11-15 18:
1512	Symantec Class 3 Secure Serv[13].cer	cer	1 340 B	54.239.29.128 [api.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 32880	TlsCertificate	2017-11-15 18:
1540	api.amazon.com[14].cer	cer	1 533 B	54.239.29.128 [api.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 45238	TlsCertificate	2017-11-15 18:
1540	Symantec Class 3 Secure Serv[14].cer	cer	1 340 B	54.239.29.128 [api.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 45238	TlsCertificate	2017-11-15 18:
1556	dp-gw-na.amazon.com[12].cer	cer	1 803 B	52.94.236.178 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59817	TlsCertificate	2017-11-15 18:
1556	Symantec Class 3 Secure Serv[12].cer	cer	1 340 B	52.94.236.178 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59817	TlsCertificate	2017-11-15 18:
1608	dp-gw-na.amazon.com[13].cer	cer	1 803 B	52.94.236.178 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59822	TlsCertificate	2017-11-15 18:
1608	Symantec Class 3 Secure Serv[13].cer	cer	1 340 B	52.94.236.178 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59822	TlsCertificate	2017-11-15 18:
1628	pindorama.amazon.com[4].cer	cer	1 518 B	52.94.227.124 [pindorama.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 37711	TlsCertificate	2017-11-15 18:
1628	Symantec Class 3 Secure Serv[4].cer	cer	1 340 B	52.94.227.124 [pindorama.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 37711	TlsCertificate	2017-11-15 18:
1696	prod.amcs-tachyon.com[4].cer	cer	1 521 B	34.202.17.208 [prod.amcs-tachyon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 45428	TlsCertificate	2017-11-15 18:
1696	Symantec Class 3 Secure Serv[4].cer	cer	1 340 B	34.202.17.208 [prod.amcs-tachyon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 45428	TlsCertificate	2017-11-15 18:
1739	dp-gw-na.amazon.com[14].cer	cer	1 803 B	52.94.236.178 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59824	TlsCertificate	2017-11-15 18:
1739	Symantec Class 3 Secure Serv[14].cer	cer	1 340 B	52.94.236.178 [dp-gw-na.amazon.com]	TCP 443	192.168.42.5 [amazon-6334998c4] [Android.local] (Android)	TCP 59824	TlsCertificate	2017-11-15 18:
2180	google.com[13].cer	cer	1 927 B	172.217.9.78 [play.l.google.com] [play.google.com]	TCP 443	192.168.42.3 (Windows)	TCP 1741	TlsCertificate	2017-11-30 19:

Figure 8: Files exchanged on the respective source and destination servers

As can be seen, https://beacons-handsoff.gcp.gvt.gvt2.com provides certificate information, which is an HTTPS protocol.

5. Conclusion

All of the attacks we have attempted were physical attacks that required the attacker to be present at the location or nearby. However, the network analysis data that we have obtained can be used in the future to perform distributed denial of service (DDoS) attacks on the servers. We find that the Echo Dot has a robust architecture and secure web servers which blocked traffic from non-Echo Dot devices, including web browsers. Communication to the secure servers also happened only after there was a certificate exchange protocol executed.

6. Acknowledgements

The Amazon Echo Dot and supporting hardware have been contributed by the IoT House research lab (612 N Park Avenue, Bloomington, IN 47405) at Indiana University under the supervision of Dr. L Jean Camp, Dr. Steve Myers, Dr. Ryan Henry, Joshua Streiff and their respective research teams. The Raspberry Pi 3 that has been used for wireless connection to study WireShark data is courtesy of Kaushik Srinivasan. We acknowledge their support in our project.

7. References

[1] Haack, W., Severance, M., Wallace, M., & Wohlwend, J. (2017). Security Analysis of the Amazon Echo.

[2]Huan Feng, Kassem Fawaz, and Kang G. Shin Continuous Authentication for Voice Assistants

[3]Si Chen, Kui Ren, Sixu Piao, Cong Wang, Qian Wang, Jian Weng, Lu Su, Aziz Mohaisen, You Can Hear But You Cannot Steal: Defending against Voice Impersonation Attacks on Smartphones

[4]Talal Bin Amin, Pina Marziliano,James Sneed German, Glottal and Vocal Tract Characteristics of Voice Impersonators

[5]Elie Khoury, Laurent El Shafey, Sebastien Marcel: SPEAR: AN OPEN SOURCE TOOLBOX FOR SPEAKER RECOGNITION BASED ON BOB

[6] https://labs.mwrinfosecurity.com/blog/alexa-are-you-listening

[7] http://sammachin.com/hacks-and-projects/call-alexa/

[8] http://www.datagenetics.com/blog/september32012/

[9] Zhang, G., Yan, C., Ji, X., Zhang, T., Zhang, T., & Xu, W. (2017). Dolphin Attack: Inaudible voice commands. arXiv preprint arXiv:1708.09537

[10]https://developer.amazon.com/docs/custom-skills/security-testing-for-an-alexa-skill.html

[11]https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/support/terms-and-agreements