Is the Amazon Echo Dot safe to Keep in a Domestic Setting?

Ian M. Henson

School for Science and Math at Vanderbilt, Nashville, TN, USA 37203

KEYWORDS. Security, Internet of Things, Amazon

BRIEF. How a popular Internet of Things device could allow criminals to listen in on nearby conversation.

ABSTRACT. Innovative home technology has received great merit and criticism by introducing new technologies into the domestic setting. One popular device that has been one of such trendsetters is the Amazon Echo Dot, an intelligent virtual assistant (IVA). The device is essentially an intelligent speaker that can receive voice commands and respond accordingly based on the Amazon Voice Services response system. Over the years similar devices entered the mainstream and are now well established in the Internet of Things (IoT) world. IoT is a general term for devices or "things" that are embedded with something to receive and process information. An example of this might be a smart thermostat or a floor cleaning robot. One of the potential problems with these devices is that their constant need to be connected to the internet can create serious security problems [1]. Furthermore, security concerns were raised after the use of reconnaissance techniques found that the Amazon Echo Dot's connection uses an outdated data transport protocol [1]. This study presents a theoretical attack that utilizes a flaw in that transport protocol to theoretically read all of the data being sent using it.

INTRODUCTION.

The Internet of Things (IoT) world is expanding faster than anyone could have expected, becoming integrated into everything we own. IoT is a general term for devices or "things" that are embedded with something to receive and process information. However, the security of these devices is moving at a much slower pace than the field's growth. The Mirai botnet, for example, was a mesh of connected IoT devices that had been infected with malware [2]. 400,000 of these devices were connected to the botnet at its peak. These types of devices included security cameras, smart TV's, and wireless presentation systems. These infected devices were used to form a network for Distributed Denial of Service attacks (DDoS) [2], which were used to knock multiple websites offline, including Twitter, Netflix, and GitHub [2]. Since then, the IoT has yet to receive a major security boost. Many companies have attempted to upgrade the security of their devices with limited success due to the hardware restrictions of embedded devices. The Intelligent Virtual Assistant (IVA) industry is a sector of IoT that has had security problems in the past [3]. The presence of 8.2 million of Amazon's IVA alone, could be a potential security risk to many [1].

The Amazon Echo Dot is a microphone and speaker controlled by the Amazon Voice Service (AVS) to function as an IVA. The Echo connects to AVS over Wi-Fi using Transport Layer Security (TLS) v1.2 protocol. That being said, the primary data pipeline for the device, one would hope that security is the number one concern for Amazon. Especially if the device is recording data without its owner explicitly using the "wake word" [1]. However, based on what was found in this research, it is not secure enough to transmit this valuable data. Amazon uses, in conjunction with TLS, an encryption algorithm called Elliptic Curve Encryption (ECC) [4], using its most lengthy key to "improve" security.

This specific information was found using a Man in the Middle attack (MitM) [5], commonly used to read actively transmitted data between two devices. This was initially done to find any non-encrypted data; however, after inspecting the captured data, it was found that everything was encrypted with some version of the TLS protocol. Many exploits were found after researching TLS vulnerabilities; however, many required hardware that wasn't accessible for this project. So, a new method was devised to test this connection based on a vulnerability that was theorized by the researcher. If an attacker were to gain access to an Amazon Echo Dot, they could listen in on all the sounds within the room that the device was in.

An experiment was created to test this attack's possible limitation: the packets generated could make decrypting and then re-encrypting this data challenging. So, if a room were to become louder, more packets would be sent to Amazon's Voice Service. An influx in packets would make it harder for an attacker to read each packet because, if they tried, a backlog could begin causing a latency issue for the device's user. The user could use a command, and it would take a few minutes to respond. This could possibly expose the attacker making this attack impractical and inefficient in practice.

MATERIALS AND METHODS.

Reconnaissance.

The device's traffic and running services had to be analyzed to find a means of attack or an exploitable service. Reconnaissance consisted of multiple scans on the device's open ports and analysis of different results. The tool used to do this is Nmap. The network traffic analysis was performed with the tool Wireshark and combined with a MitM attack to gain access to network traffic not ordinarily visible to a device sharing the network with the Echo Dot (Fig. 1). The original MitM attack was used to detect any network traffic that wasn't encrypted, perhaps offering some insight into the Echo and its running services.

Man in the Middle Attacks.

Performing a MitM attack with the ARP protocol can be done using Arpspoof, a tool most often used to impersonate another device's local IP address using the Address Resolution Protocol (ARP) [5]. In this situation, Arpspoof was used to position the "attacking" computer between the target (Echo Dot) and the router to intercept packets sent to and from AVS. The packets being sent were viewed through Wireshark, a tool designed to capture and view intercepted packets.

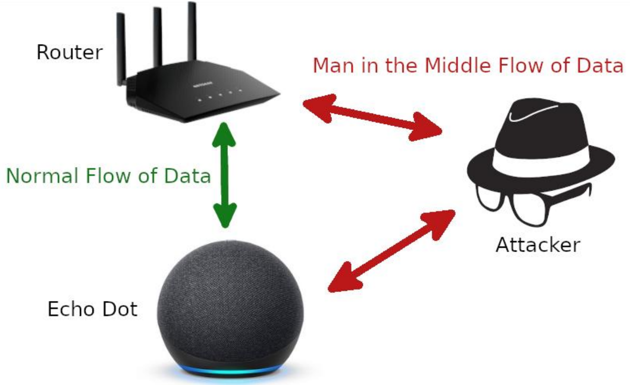

Figure 1. Diagram of a Man-in-the-Middle attack, showing how data moves comes and goes to the Echo Dot. During the attack data is sent to the attacking computer and then the router rather than directly to the router. This passive attack allows for data to be collected on the encryption method used to protect the data coming to and being sent from the Echo Dot. That knowledge could give the attacker information on how to further attack the Echo Dot.

Automation of Packet Analysis and Injection.

The attack script and some other scripts were developed to automatically gather data used to determine the efficiency of the attack script and how often target packets are sent. The script was done using Python and a library called Scapy to intercept packets, filter them, and inject a malicious payload.

Packet Injection Attack.

The packet injection tool was developed to automatically detect and modify handshake data sent to and from the Echo Dot to make its encrypted packets readable. This tool is in combination with Scapy and a MitM attack. This script has been developed; however, the execution of this attack requires approval from the Amazon corporation. The request has yet to be answered.

Data Collection.

A test was designed to be noninvasive and avoid sending any malicious packets to any privately-owned servers (AVS) that tested a possible limitation of the attack. The link between the server and the microphone on the Echo Dot is utilized to theoretically generate more packets of data. Noise in a room was theorized to affect the number of packets sent to the Amazon servers. A Python script was created; it used Scapy and Python to capture packets. Three noise levels were designed to emulate room noise, 45 dB silent noise condition: silent within the home the subject was kept in, 55 dB ambient noise condition: people were present and working within the room the device was in, and 70-75 dB loud noise condition: music was played to emulate a packed room or party. The decibel count was taken using the sound meter app on an android device. The script recorded for 6 hours and saved the time between each packet, the time recorded, and assigned a number to that packet. This data was analyzed using a one-way ANOVA to test if each sound level condition was significantly different from the other.

RESULTS.

After the 6-hour recording sessions for each noise level were completed, the data was analyzed, and the average number of packets sent at this time was 8,172. The silent noise condition session recorded 8,810 packets, the ambient noise condition session recorded 7,693 packets, and the loud noise condition session recorded 8,013 packets. The average time between all captured packets was 2.650 seconds. The encrypted data had an average of 2.685 seconds between packets and a time of 108.985 seconds between each key exchange.

The Silent, Ambient, and Loud noise condition time between packets was compared (Fig. 2) with a one-way ANOVA and found that there was no significant difference (F(2)=2.160, p=0.115) between all of the packets captured. For the key exchange packets, an ANOVA yielded no significant results with (F(2)=0.837, p=0.433) and the encrypted application data packets (F(2)=2.039, p=0.130).

DISCUSSION.

This study has found that the amount of noise in a room does not affect the amount of data sent and would not make the theorized attack harder. No significant difference was found between the three sound levels, which verifies this. The noise to time between packets is visualized by Figure 2, and most of these points stayed below the 50-second mark. This uniformity is assumed to be caused by the sensitivity of the microphone. It seems to pick up any amount of sound, even if it's considered background noise. That's why the frequency of packets didn't change based on the tested conditions. This conclusion disproves the hypothesized outcome of the experiment. Due to the simplicity of this vulnerability, even an unskilled hacker could run this attack.

The limitation of this study that most prominently affected progress was the lack of approval from Amazon to explore the possible outcomes of the attack, which is an area that will hopefully one day get the opportunity to be researched. This lack of consent prevented the actual attack from being tested and made the theory of a possible latency issue untestable. Besides this, the experiment and data collection were not limited at all by the environment or funding. At the same time, this type of attack is becoming increasingly obsolete because of the gradual upgrade to TLS 1.3. One thing that could bypass this improved security would be malformed packets affecting the receiving Amazon servers. This attack could cause something called a buffer overflow which has been a common exploit approach since the 1990s [6]. In the past, attacks have been developed that exploit the built-in capacity for this sent data, a buffer. If it exceeds this buffer, it could cause the system to do things the machine would not normally do [7]. Often, a system could get knocked offline or crash, which would generate a large service issue for Amazon. This attack is referred to as the ping of death. Another area of study that would be interesting to pursue in the future would be a more in-depth analysis of the device's services. Typical methods seem inappropriate as the device uses what appears to be some custom software.

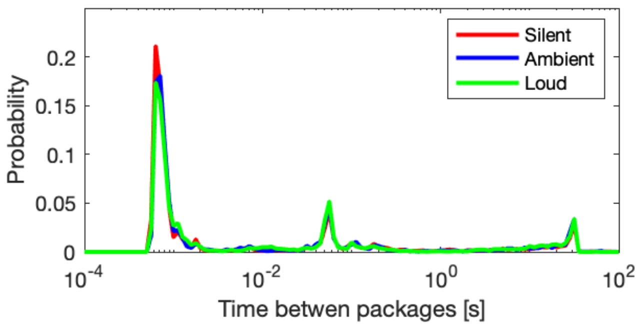

Figure 2. Histogram displaying the time between packets to the probability of that time. The tested sound variables were laid on top of each other to better visually compare the variables. The scaling used for the X axis was chosen because of the data's wide range of values. This histogram presents that no matter how much activity is within a room it will be just as easy for an attacker to listen in on its occupants.

CONCLUSION.

In conclusion, Amazon has not correctly protected their user's data, providing a criminal the opportunity to violate that user's privacy. If this data were to be compromised, an attacker could listen in on their victim throughout the day while the victim remains completely oblivious. What is much more frightening, however, is the Amazon device marketed for children. It remains unclear if there have been security improvements made to these devices. This raises the question about what is being done to protect the 8.2 million owners of these devices [1]. Amazon has made no effort to improve the wireless security of its devices. It should be noted that Amazon may not be the only company with these security concerns. It may be the case that other IVAs produced by Google and Apple could also be susceptible to this attack as well.

ACKNOWLEDGMENTS.

Thanks to the School for Science and Math at Vanderbilt for sponsoring this research. Thank you, Dr. Sarah Huneycutt, and Dr. Menton Deweese, for providing support throughout my research.

REFERENCES.

1. M. Ford, W. Palmer, Alexa, are you listening to me? An analysis of Alexa voice service network traffic. *Personal and Ubiquitous Computing*. (2019).
2. C. Kolias, G. Kambourakis, A. Stavrou, J. Voas, DDoS in the IoT: Mirai and Other Botnets. *Computer V.***50**, 80-84. (2017).
3. H. Chung, M. Iorga, J. Voas, and S. Lee, Alexa, Can I Trust You? *Computer*.**50**, 100-104. (2017).

4. Koblitz, N., Menezes, A. & Vanstone, S. The State of Elliptic Curve Cryptography. *Designs, Codes and Cryptography.***19**, 173–193. (2000).
5. H. Hwang, G. Jung, K. Sohn, and S. Park, A Study on MITM (Man in the Middle) Vulnerability in Wireless Network Using 802.1X and EAP. *International Conference on Information Science and Security*, Seoul, South Korea, 164-170. (2008).
6. AlephOne. Smashing stack for fun and profit. *Phrack*.**7**, 49. (1996)
7. J. N. Goel and B. M. Mehtre, Dynamic IPv6 activation-based defense for IPv6 router advertisement flooding (DoS) attack. *IEEE International Conference on Computational Intelligence and Computing Research*, 1-5. (2014).

Ian Henson is a student at Hillsboro High School in Nashville, TN; he participated in the School for Science and Math at Vanderbilt University.