A first look into Alexa's interaction security

Ismael Castell-Uroz
Universitat Politècnica de Catalunya

Xavier Marrugat-Plaza
Universitat Politècnica de Catalunya

Josep Solé-Pareta
Universitat Politècnica de Catalunya

Pere Barlet-Ros
Universitat Politècnica de Catalunya

ABSTRACT

With a rapidly increasing market of millions of devices, the intelligent virtual assistants (IVA) have become a new vector available to exploit security breaches. In this work we approach the third revision of the *Amazon Echo* ecosystem's device *Alexa* from a security perspective, focusing our efforts on the interaction between the user and the device. We found the client-server communications to be robust using encryption, but studying the voice message recognition system we discovered a method to execute voice commands remotely, a feature not available by default. This method could be used against the user if an attacker manages to perform a *session hijacking* attack on the web or mobile clients.

1 INTRODUCTION

With the advance in speech recognition techniques as well as the massive amounts of data that can be currently analyzed using artificial intelligence algorithms and cloud computing, a new device not present some years ago is increasingly getting relevance in computer interaction. Intelligent virtual assistants (IVA) like *Amazon Echo* or *Google Home* are gaining market rapidly due to the ability of performing simple tasks using only voice interaction. This can be considered the next step in human-machine interaction but at the same time it can be observed as a security and privacy threat.

Prior studies already demonstrated some vulnerabilities that most voice-recognition devices suffer. For instance, Carlini et al. in [2] and Zhang et al. in [11] studied how to produce voice commands using sounds innaudible by humans that are properly recognized by the devices.

In this work we dissect and analyze the communications between *Amazon*'s third revision of *Alexa Echo* devices, *Amazon* servers, and computer and mobile phone clients. On top of that, we also study the voice command interaction system. Based on our findings we show a way to exploit the device capabilities to be able to execute commands remotely sending reminders that will be self-executed, a feature that can be used against the user.

2 SECURITY IMPLICATIONS

As a cloud based service, Alexa makes use of big data and cloud computing systems to listen, interpret and execute a list of predefined voice commands available. All the ecosystem has several security implications; recognition security (voice commands), cloud security (Amazon servers), skills (Alexa's own applications), protocols (bluetooth, wifi) and client security (website and mobile phone applications). Some of them like skills security has already been investigated in the past [4] while some others like protocols security is dependent on very specific flaws [1] already covered. Our work focuses mainly on the interaction security, containing both, client security and voice recognition security.

2.1 Client security

In order to study the client communication security, first we analyzed the servers and services being accessed by Alexa on a daily basis. Figure 1 shows the device communications diagram when it is turned on. The starting one is *fireoscaptiveportal.com* and the next requests are performed clockwise. Whenever possible we analyzed the data being sent and received to try to infer the function of each one of the servers.

For non-secure communications sent by *http* it is not difficult to access the data. Apart from *fireoscaptiveportal.com*, portal from MarkMonitor [5], a company focused on computer security, the rest of the servers are accessed to check connectivity with some of their own Amazon services or to access the media file to be reproduced by the device.

To get information about the rest of the servers, we used the Burp Suite tool [8] to simulate a "*man in the middle*" attack. This way we can study the client-server encrypted communications. The structure of the solution is shown in Figure 2. We were able to see the metrics being sent to *device-metrics-us.amazon.com* as well as information about different endpoints and multimedia resources being set by *arcus-uswest.amazon.com*.

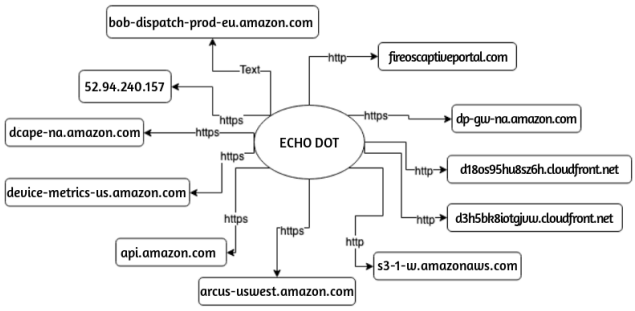

Figure 1: Device communications

Interestingly, we also found data about Alexa usage historic usually unaccessible by the user. This is sensible data that can be used to extrapolate useful patterns about the time frames when the users are far from the device. Table 1 shows a summary of the data that can be extracted from the communications between the server and the dashboard. Note that the last url permits the user to set a reminder on the device remotely, fact that will take special interest in the next section. We have published a small tool [6] that using a session cookie shows the usually unavailable historic data.

Information	**url** (https://alexa.amazon.es)
Name, email, address	/api/authentication
Additional emails	/api/eon/householdaccounts
Available devices	/api/devices-v2/device
Family members	/api/household
Buying activities history	/api/activities-with-range
Created lists	/api/namedLists
Lists elements	/api/namedLists/[listId]/items
Activity history	/api/activities-with-range
Reminder creation	/api/notifications/createReminder

Table 1: Information available inside resources

2.2 Voice interaction security

Looking that some IVAs reacted to TV commercials in the past [3] Amazon engineers implemented an on-the-cloud system to compare device wake voice commands with multiple other devices in real time [7]. This way they can detect broadcasting events like commercials.

Therefore, our research is focused on asynchronous events, like online videos, that could send voice commands to the device. Moreover, instead of looking directly for wake words, we search for speech recognition errors that could be used as a wake word to send the command. A similar approach has been taken in the past [4] to open a fake skill (Alexa's own application) whose name is similar in pronuntiation to a legitimate one.

To study speech recognition errors, an Spanish language audio database [9] with more than 1.700 different files was reproduced near the device while monitoring the reaction. We found some words that were commonly misunderstood (e.g."el examen", "economía") and made the device to wake up, listen and check with Amazon servers, but the given command was discarded. That means that Amazon also has a protection system to discard this type of false positives.

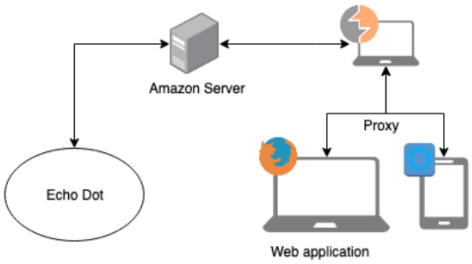

Figure 2: Burp Suite usage diagram

Surprisingly, during our experiments we observed that the device listens to himself while reproducing a message. Thus, we setup a reminder with a voice command inside it and let the device reproduce it. We discovered that if the volume setting is higher enough the device will execute the command just after reading it and will not be discarded.

This can be exploited in a number of different ways. For example, if an attacker gets the user session cookie by a *session hijacking* [10] attack (something plausible as we found that dashboard session cookies renewal period is longer than a decade) an attacking reminder can be set using the url mentioned in Section 2.1. The reminder will be automatically executed. This permits the attacker to do online purchases, unless the purchase verification code is turned on, or to operate other available intelligent devices like door locks.

3 CONCLUSION

In this work we presented a first look at Alexa's interaction security. We achieved some interesting findings as the remote execution of commands using a reminder, or finding the historic data hidden in the dashboard, that can potentially be used by an attacker against the user. We plan to continue exploring the available possibilities to detect new threats that could lead to new security breaches and privacy leakages.

4 ACKNOWLEDGMENTS

This work was supported by the Spanish MINECO under contract TEC2017-90034-C2-1-R (ALLIANCE).

REFERENCES

[1] Armis. 2017. *The Attack Vector "BlueBorne" Exposes Almost Every Connected Device.* Retrieved September 17, 2019 from https://www.armis.com/blueborne/

[2] Nicholas Carlini, Pratyush Mishra, Tavish Vaidya, Yuankai Zhang, Micah Sherr, Clay Shields, David Wagner, and Wenchao Zhou. 2016. Hidden Voice Commands. In *25th USENIX Security Symposium (USENIX Security 16).* USENIX Association, Austin, TX, 513–530.

[3] Nina Golgowski. 2017. *This Burger King Ad Is Trying To Control Your Google Home Device.* Retrieved September 05, 2019 from https://www.huffpost.com/entry/burger-king-launches-ad-on-google-home_n_58ee23dfe4b0c89f912307f9

[4] Deepak Kumar, Riccardo Paccagnella, Paul Murley, Eric Hennenfent, Joshua Mason, Adam Bates, and Michael Bailey. 2018. Skill Squatting Attacks on Amazon Alexa. In *Proceedings of the 27th USENIX Security Symposium (USENIX Security '18).* USENIX Association, Berkley, CA, 33–47.

[5] MarkMonitor. 2019. *MarkMonitor - Clarivate Analytics | Brand Protection, Domain Management, Anti Piracy, Anti Fraud.* Retrieved September 03, 2019 from https://www.markmonitor.com

[6] Xavier Marrugat-Plaza. 2019. *Alexa Cookied.* Retrieved September 17, 2019 from https://github.com/xampla/AlexaCookied/blob/master/alexaCookied.py

[7] Mike Rodehorst. 2019. *Why Alexa Won't Wake Up When She Hears Her Name in Amazon's Super Bowl Ad.* Retrieved September 05, 2019 from https://developer.amazon.com/es/blogs/alexa/post/37857f29-dd82-4cf4-9ebd-6ebe632f74d3/why-alexa-won-t-wake-up-when-she-hears-her-name-in-amazon-s-super-bowl-ad

[8] Portswigger Web Security. 2019. *Burp Suite.* Retrieved July 16, 2019 from https://portswigger.net/burp

[9] Eric Streit. 2016. *The Audio collections - Collection of spanish words (spa-wims-octavio).* Retrieved May 09, 2019 from https://fsi-languages.yojik.eu/audiocollections/audiocollections.html

[10] Wikipedia. 2019. *Session hijacking.* Retrieved July 21, 2019 from https://en.wikipedia.org/wiki/Session_hijacking

[11] Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. 2017. DolphinAttack: Inaudible Voice Commands. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17).* ACM, New York, NY, USA, 103–117. https://doi.org/10.1145/3133956.3134052