Improving Authentication in the Amazon Alexa Virtual Assistant by Using a Geofence

Jorge Fernández-García, Martiño Rivera-Dourado, Rubén Pérez-Jove, Cristian R. Munteanu, and Jose Vázquez-Naya

Grupo RNASA-IMEDIR, Facultade de Informática, Universidade da Coruña, Campus de Elviña, A Coruña, 15071, Spain
Centro de Investigación CITIC, Universidade da Coruña, 15071 A Coruña, Spain
IKERDATA S.L., ZITEK, University of Basque Country UPVEHU, Rectorate Building, Leioa, 48940, Spain
Correspondence: jorge.fgarcia@udc.es

DOI: *https://doi.org/10.17979/spudc.000024.50*

Abstract: Amazon Alexa processes voice commands as input to help users perform tasks. For protecting this commands, Amazon Alexa implements some security measures. These security measures, such as voice recognition and user's PIN, do not have the ability to mitigate replay attacks. In order to mitigate replay attacks, in this paper, we propose an authentication method based on Geofencing, consisting of (1) an Android application and (2) an Alexa Skill. By using the Android application, the user is able to configure a geofence near the Amazon Echo smart speaker. The developed Alexa Skill only accepts requests when the user is within the established geofence. This method mitigates replay attacks: an attacker could only try to use a replay attack when the legitimate user is close to the speaker, making it unfeasible.

1 Introduction

Amazon Alexa is a virtual assistant that processes voice commands quickly and efficiently. It provides basic functionalities and allows the extension of its capabilities through the creation of third-party functionalities, known as Alexa Skills (Amazon Alexa Skills, 2016). This is made possible by the development environment provided by Amazon (Amazon, 2023).

This virtual assistant has security measures such as voice recognition and user's PIN (Alattar et al., 2023), (Amazon, 2020). These security measures are used to process some sensitive requests, such as an online purchase. In particular, the user's PIN must be spoken aloud so that Amazon Alexa can authenticate the user and accept or reject the sensitive requests.

Despite these security measures, Amazon Alexa has a number of vulnerabilities published by INCIBE (Incibe, 2023). Some vulnerabilities to highlight are: the processing of voice commands at high frequency spectrum, the existence of backdoors and impersonation. Consequently, due to these vulnerabilities, differents attacks arise such as: dolphin attacks (Zhang et al., 2017), voice squatting (Zhang et al., 2018), voice masquerading (Zhang et al., 2018) and replay attacks (Malik et al., 2019).

As a result, these attacks can bypass the security measures, specifically replay attacks. Replay attacks consist of recording and replaying a victim's voice command to impersonate his identity. For example, attackers could record a user's spoken PIN and then replay it for Amazon Alexa to process the sensitive request.

In this paper, we developed an authentication method based on Geofencing to protect requests in Amazon Alexa, mitigating replay attacks that were previously possible. We propose

a Proof of Concept (PoC) where Geofencing technology is used to improve the authentication on requests. Geofencing uses GPS, Wi-Fi or Bluetooth to create virtual geographic barriers, named geofence, that monitor the location of a physical device or user, with the aim of detecting when the user enters or exits the geofence (Rahate and Shaikh, 2016). In our PoC, the geofence is configured at the location of the smart speaker using a developed Android application. When the user is inside the geofence, Amazon Alexa, through a developed Alexa Skill, processes the user's requests. Otherwise, it rejects them. With this approach, if an attacker tries to perform a replay attack, he needs the legitimate user to be inside the geofence, making this attack unworkable.

2 Material and methods

This paper seeks to create an authentication method that protect requests in Amazon Alexa using Geofencing technology. To achieve this, we employed a personal computer running the Windows 10 operating system with internet connection. Subsequently, we used the Android Studio IDE to develop the Android mobile application and tested it on a Samsung J7 2016 Android mobile device. Finally, we created accounts on Amazon Developer and Amazon Web Services (AWS, 2020) to develop the Alexa Skill and utilizing Amazon services.

We started this research by analyzing the security measures and vulnerabilities of Amazon Alexa. Next, we looked at what types of Skills exist in the market and how to develop them. Then, we searched for information to communicate Amazon Alexa and Android technology through Amazon services (AWS, 2020).

For software development, we adopted the agile Scrum methodology. This allowed us to work in Sprints and have different functional versions of the software. We began with a basic version of the application, which was refined during each Sprint until we reached the final product.

3 Development

To demonstrate how to improve authentication in Amazon Alexa, the development of a PoC was carried out. It is important to highlight three main blocks of this PoC: (1) the analysis and design of the authentication method based on Geofencing, (2) the development of an Android application for configuring the Geofencing technology and (3) the development of the Alexa Skill responsible for accepting or denying the requests. The following is a brief explanation of the three main blocks.

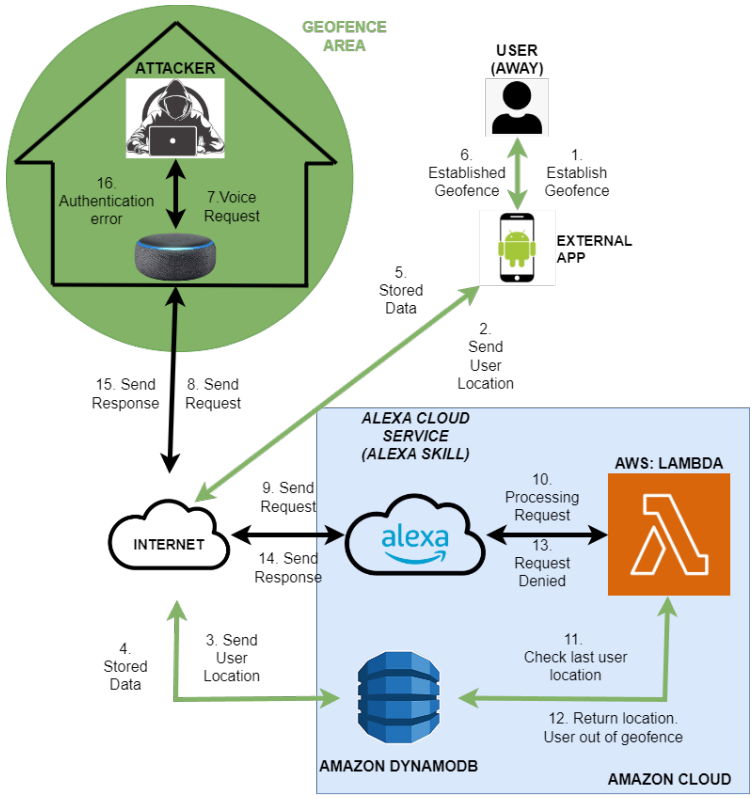

Figure 1: Request flow in PoC with the authentication method based on geofence. Black arrows represent the standard flow. In green, the additional flow added with our authentication method.

3.1 Analysis and design of the authentication method based on Geofencing

To design the authentication method, we first analyzed the flow of a request to any Alexa Skill. When a user sends a voice command to an Amazon Echo, it is forwarded to the Alexa cloud service, where the Skill is hosted. The request is then processed on the Skill's backend in AWS Lambda (Docs.Amazon, 2020) and a response is sent to the user via the Amazon Echo. This flow is represented by the black arrows in the figure 1. Without improved authentication, an attacker could impersonate the victim and use the Amazon Echo to send requests to the Alexa Skill, successfully performing a replay attack.

With this in mind, Geofencing technology was used to add two additional steps when processing a request. These additional steps are explained below. They can be seen as green arrows in figure 1.

1. The user starts the developed Android application and configures a geofence at the position of Amazon Echo. Additionally, this application continuously monitors the user's location with regard to the geofence and sends it to the DynamoDB service database (DynamoDB, 2020). Steps 1 to 6 in figure 1.
2. The developed Alexa Skill checks the user's last location in the DynamoDB service database before processing the request. If the user is within the geofence, the request is processed. Otherwise, the Alexa Skill returns an authentication error message and the request is not processed. Steps 7 to 16 in figure 1.

These two steps mitigate replay attacks. When the legitimate user is located outside the geofence, the attacker will receive an authentication error message. Therefore, if the attacker wants to perform a replay attack, he needs the legitimate user to be inside the geofence and therefore close to the Amazon Echo. This requirement makes the replay attack unfeasible.

3.2 Mobile application for geofence configuration

The development of the application, called "AlexaGeoApp", was carried out in the Android Studio IDE using the Java programming language, and Google SDKs, such as Geofencing and Google Maps (Geocoding, 2021),(GoogleMaps, 2021). The application renders a map and sets a geofence near the Amazon Echo smart speaker. It also continuously monitors the user's position with respect to the geofence and sends it to the AWS DynamoDB database. Finally, the application has an interface that informs the user and guides him through the whole setup process to establish the geofence.

3.3 Alexa Skill that uses our authentication method based of Geofencing

The developed Alexa Skill consists of two parts: the frontend and the backend. The frontend was developed in Alexa Developer Console (Console, 2020) and is responsible for detecting user requests and sending them to the back end. The backend, written in JavaScript and hosted in AWS Lambda, handle requests processing and provides responses to the frontend. Furthermore, a communication with AWS DynamoDB was implemented through the backend to verify the user's last location in relation to the geofence. If the user is within the geofence, the Alexa Skill accepts requests, otherwise it denies them.

4 Results

In this paper, we developed an authentication method based on Geofencing to protect requests in Amazon Alexa, mitigating replay attacks. Geofencing technology was used as an additional authentication factor when processing voice requests. To achieve this, a PoC was developed, consisting of the creation of an Android application and an Alexa Skill.

On the one hand, the user can configure a geofence at the position of the Amazon Echo smart speaker through the developed Android application. On the other hand, the developed Alexa Skill can use the configured geofence to accept or deny the user's requests. If the user is within the geofence the requests are accepted. Otherwise, they are denied. Therefore, if an attacker wants to perform a replay attack, he needs the legitimate user to be inside the geofence, making the replay attacks unfeasible.

Some important parts of the developed Android app and Alexa Skill are highlighted below.

4.1 AlexaGeoApp: an Android application

The developed Android application, called "AlexaGeoApp", features a user-friendly interface that allows the user to configure the geofence near Amazon Echo smart speaker. Once the geofence is configured, "AlexaGeoApp" renders a map representing this geofence as a red zone. The application monitors the user's location with respect to the geofence and sends it to the DynamoDB service. In addition, when the user enters or exits the geofence, the user is informed via a notification on his mobile device. This can be seen in the figure 2.

4.2 Secure Authentication: an Alexa Skill

The Alexa Skill, called "Secure Authentication", uses our authentication method based on Geofencing to process requests. "Secure Authentication" Skill connects to the DynamoDB database to make decisions based on the latest data stored by "AlexaGeoApp" to process user

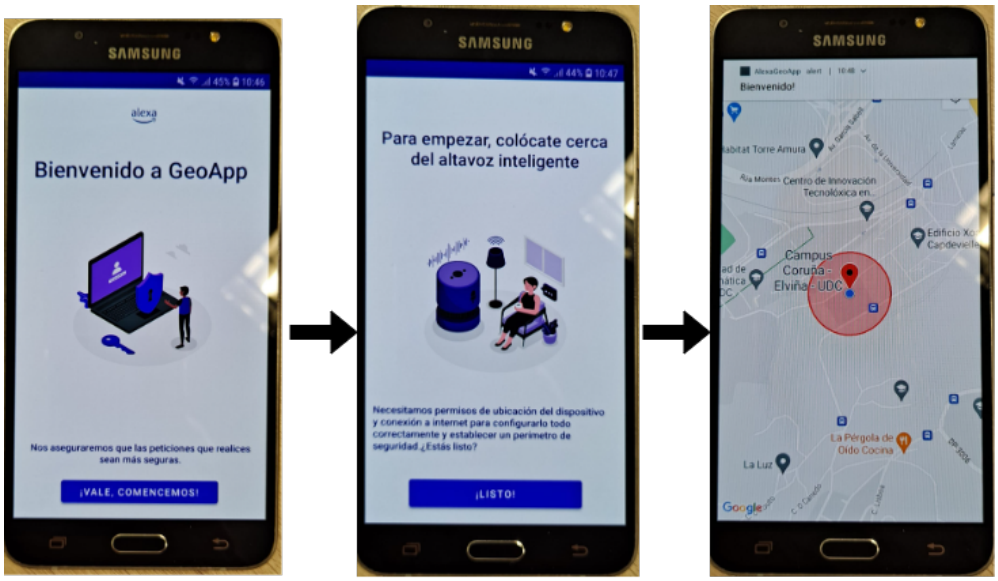

Figure 2: "AlexaGeoApp" flow. From left to right: (1) Welcome message to the user. (2) We inform the user to stand near the smart speaker. (3) We retrieve the user's location, set the geofence and inform the user with a notification.

requests. The user can verify his authentication status through the "Secure authentication" Skill. If the user is successfully authenticated by this method, he can make requests to the "Secure authentication" Skill. However, if authentication fails, the "Secure Authentication" Skill denies the request.

Finally, "Secure Authentication" has a validity window for the data stored in DynamoDB. This sets a maximum time for the validity of the data. If the data is very old and there have been no updates to the user's location, "Secure authentication" rejects all the requests, following the fail-safe principle (Leedeo, 2020). With this principle we avoid any denial of service issues on the mobile device that containing the developed "AlexaGeoApp", such as running out of battery or losing internet connectivity. "AlexaGeoApp" needs internet connectivity to detect user entries and exits in the geofence and update the DynamoDB database. If these state changes are not correctly detected, the authentication based on Geofencing would not work correctly. Thanks to the validity window we avoid this problem.

The resulting code of this work has been published as an open source project on the following Github repository: https://github.com/jorgefgarcia/AlexaGeoApp

5 Conclusions

The lack of security measures in the Amazon Alexa virtual assistant underscores the importance of addressing request security. Solutions that tackle the lack of authentication in this kind of settings, as the one proposed in this paper, may encourage manufacturers to develop more robust security measures for a safer user experience. Amazon environment could integrate our solution by default to protect sensitive requests, such as an online purchase through the Amazon Echo smart speaker, against replay attacks. In addition, it would be interesting to consider future vulnerabilities and adjust the design accordingly.

6 Future work

This paper has introduced an authentication method based on Geofencing that improves the authentication of requests directed at Amazon Alexa. However, it's important to consider future enhancements:

- **Utilization of other technologies:** Expanding the authentication enhancement to other virtual assistants such as Google Assistant, Bixby, Cortana or Siri. Additionally, developing a cross-platform application to configure security from multiple devices.
- **Responsive design:** Achieve a user interface design that adapts to different device screens.
- **Multi-user mode:** Add the capability of setting multiple geofences on a map with a group of users.
- **Wi-Fi usage:** Implementing a feature to process requests only when the user's mobile is connected to the same Wi-Fi network as the smart speaker.

Acknowledgements

This work was supported by the grant ED431C 2022/46 – Competitive Reference Groups GRC – funded by: EU and "Xunta de Galicia" (Spain). This work was also supported by CITIC, funded by "Xunta de Galicia" through the collaboration agreement between the "Consellería de Cultura, Educación, Formación Profesional e Universidades" and the Galician universities to strengthen the research centres of the "Sistema Universitario de Galicia" (CIGUS). Also, the work is founded by the "Formación de Profesorado Universitario" (FPU) grant from the Spanish Ministry of Universities to Martiño Rivera Dourado (Grant FPU21/04519).

Bibliography

Z. S. Alattar, T. Abbes, and F. Zerai. Smartphone-key: Hands-free two-factor authentication for voice-controlled devices using wi-fi location. *IEEE Transactions on Network and Service Management*, pages 1–1, 2023. Conference Name: IEEE Transactions on Network and Service Management.

Amazon. What is alexa voice ID? - amazon customer service. *https://www.amazon.com/gp/help/customer/display.html?nodeId=GYCXKY2AB2QWZT2X*, 2020. [Online; accessed 7-June-2023].

D. Amazon. Amazon developers. *https://developer.amazon.com/es-ES/home.html*, 2023. [Online; accessed 7-June-2023].

Amazon Alexa Skills. Alexa skills | amazon.com. *https://www.amazon.com/alexa-skills/b?ie=UTF8&node=13727921011*, 2016. [Online; accessed 6-September-2023].

AWS. ¿qué es AWS? *https://aws.amazon.com/es/what-is-aws/*, 2020. [Online; accessed 6-September-2023].

A. D. Console. Amazon alexa voice AI | alexa developer official site. *https://developer.amazon.com/en-US/alexa.html*, 2020. [Online; accessed 7-June-2023].

Docs.Amazon. ¿qué es AWS lambda? - AWS lambda. *https://docs.aws.amazon.com/es_es/lambda/latest/dg/welcome.html*, 2020. [Online; accessed 22-April-2023].

A. DynamoDB. ¿qué es amazon DynamoDB? - amazon DynamoDB. *https://docs.aws.amazon.com/es_es/amazondynamodb/latest/developerguide/Introduction.html*, 2020. [Online; accessed 7-June-2023].

G. D. A. Geocoding. Comenzar | geocoding API. *https://developers.google.com/maps/documentation/geocoding/start?hl=es-419*, 2021. [Online; accessed 5-September-2023].

G. D. A. GoogleMaps. Guía de inicio rápido del SDK de maps para android | maps SDK for android. *https://developers.google.com/maps/documentation/android-sdk/start?hl=es-419*, 2021. [Online; accessed 5-September-2023].

Incibe. El uso de altavoces inteligentes que riesgos de seguridad y privacidad nos exponen | ciudadanía | INCIBE. *https://www.incibe.es/ciudadania/formacion/infografias/el-uso-de-altavoces-inteligentes-que-riesgos-de-seguridad-y-privacidad-nos-exponen*, 2023. [Online; accessed 7-June-2023].

Leedeo. ¿qué significa que un sistema es fail safe? *https://www.leedeo.es/l/fail-safe/*, 2020. [Online; accessed 22-April-2023].

K. M. Malik, H. Malik, and R. Baumann. Towards vulnerability analysis of voice-driven interfaces and countermeasures for replay attacks. In *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, pages 523–528, 2019.

S. W. Rahate and D. M. Z. Shaikh. Geo-fencing infrastructure: Location based service. *IRJET*, 03(11), 2016.

G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu. Dolphinattack: Inaudible voice commands. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 103–117, New York, NY, USA, 2017. Association for Computing Machinery.

N. Zhang, X. Mi, X. Feng, X. Wang, Y. Tian, and F. Qian. Understanding and mitigating the security risks of voice-controlled third-party skills on amazon alexa and google home. *ArXiv*, abs/1805.01525, 2018.