Análisis de funcionamiento y de vulnerabilidades de Alexa

Trabajo de Fin de Grado, Informe inicial

Sergio Bachiller Rubia - 1426724

Marzo 2019, Universidad Autónoma de Barcelona

1 Introducción

El impulso del *Internet of Things* ha aumentado significativamente durante la última década, hasta llegar al punto en el que se calculan que a finales de 2017 había 8.400 dispositivos conectados a Internet [1].

Muchas compañías han aprovechado esta situación, en la que los usuarios buscan, principalmente, herramientas que hagan su vida diaria más fácil y cómoda. Es el caso de Alexa, el asistente virtual de la multinacional Amazon. Este asistente está disponible en hasta 5 modelos diferentes de Amazon Echo, sus altavoces inteligentes, además de dispositivos móviles y hasta de otros productos de terceros.

1.1 Arquitectura

Encontramos la parte funcional de Alexa en los servidores de Amazon, donde se reciben y contestan las peticiones de los usuarios, formando así una arquitectura que permite mejorar el servicio únicamente desde el servidor. Además, aprovechando la conexión a Internet, ofrece funciones como información metereológica, estado del tráfico a tiempo real o reproducción de audio y/o vídeo bajo demanda.

Además, Amazon ofrece la posibilidad a desarrollar *skills* o funciones, a terceros, lo que permite crear una comunidad que trabaje en la mejora continua del asistente virtual.

Figura 1: Visión general de la arquitectura de Amazon Alexa

A destacar, como podemos ver en la Figura 1.1, la comunicación entre el usuario y el dispositivo ejecutando Alexa siempre será mediante voz. Sin embargo, los comandos del usuario no se transcriben en el mismo dispositivo, sino que se envían al servidor de Alexa para ello. En cambio, el dispositivo sí que se encarga de transformar la respuesta en voz [1].

1.2 Potenciales vulnerabilidades

Viendo la arquitectura de Amazon Alexa pueden surgir diferentes ideas sobre las potenciales vulnerabilidades del sistema. Por ejemplo:

- Con tal de responder al usuario, los dispositivos estan en un estado de escucha constante, lo que hace pensar que el fabricante, en este caso Amazon, puede escuchar nuestras conversaciones privadas incluso sin estar haciendo uso de Alexa.
- Es posible la reproducción de un archivo de audio que imite una voz humana y pueda ejecutar instrucciones.
- De la misma forma que Alexa puede interactuar con diferentes dispositivos de domótica, es posible hacerlo desde un tercer dispositivo, por ejemplo, un ordenador.

2 Objetivos

El proyecto se enfocará, en una primera instancia, en conseguir los siguientes objetivos.

Con tal de conocer el estado del arte de Alexa, en lo que a funcionamiento y vulnerabilidades respecta, el primer objetivo es realizar una revisión de los estudios llevados a cabo hasta ahora.

Llegados a este punto, se empiezan los preparativos para realizar el análisis *per se*, realizando el diseño de un diagrama de red doméstica, con y sin Alexa, con tal de realizar una monitorización de estas redes.

Con tal de intensificar los resultados de la monitorización previa, también se fija como objetivo la instalación de un proxy entre el router de la red y Alexa, analizando el tráfico que pasa por éste.

Para comenzar el análisis de las vulnerabilidades, se realiza un estudio sobre los ataques que se pueden llevar a cabo, tanto a nivel de aplicación, *API* (*Application Programming Interface)*, como a nivel de red, con los resultados de haber monitorizado la red con anterioridad.

No hay que olvidar que Alexa se puede conectar con otros dispositivos de otros fabricantes, como bombillas o enchufes inteligentes, capaces de cambiar de estado a través de comando de voz. Por ello, se realizará un análisis del tráfico que intercambian Alexa y estos dispositivos, con tal de intentar vulnerar la seguridad de esta comunicación. Las brechas de seguridad que se plantean inicialmente van desde la suplantación de identidad de Alexa, pasando por intentar ejecutar comandos de voz con una voz emulada por ordenador hasta la grabación del entorno sin estar haciendo uso del asistente personal.

Para finalizar, se realiza una revisión de la documentación de los desarrolladores, tanto de Amazon como de terceros desarrolladores.

3 Metodología

Dado el tipo de proyecto a llevar a cabo, en el que es necesario la revisión continua de todas las tareas, se va a llevar a cabo mediante una metodología ágil, concretamente Kanban.

Kanban procede del japonés, que sería traducido literalmente como *tarjeta con signos* [2]. Este método casa muy bien con este proyecto, pues uno de sus principios es el implementar cambios incrementales y evolutivos, y en cualquier momento puede aparecer una nueva vulnerabilidad o cambio de funcionamento en el objeto de estudio.

De este modo, se trabaja con un tablero Kanban donde las tareas se reparten en cuatro categorías: por hacer, en proceso, revisión y finalizado. La categoría revisión es la más destacable, pues una vez se acabe una tarea se valorará si es posible revisarla e intentar mejorar el resultado conseguido hasta el momento, siempre respetando la planificación.

4 Planificación

En base a los objetivos propuestos anteriormente, se definen las siguientes tareas para el desarrollo del análisis, estando éstas sujetas a cambios o incluso a la aparición de nuevas:

- Búsqueda y documentación de vulnerabilidades de Alexa o Amazon Echo en el repositorio de CVE (*Common Vulnerabilities and Exposures*). (3h)
- Búsqueda y documentación de ataques basados en APIs. (10h)
- Búsqueda y documentación de ataques basados en sonido. (15h)
- Búsqueda y documentación de ataques de playback. (10h)
- Diseño del diagrama de red doméstica con y sin Alexa. (2h)
- Búsqueda de herramientas para monitorear la red. (5h)
- Monitorización de la red doméstica con las herramientas seleccionadas previamente. (10h)
- Búsqueda de información sobre el montaje de un proxy. (15h)
- Montaje de un proxy en la red. (20h)
- Búsqueda de herramientas para analizar el tráfico en el proxy. (5h)
- Análisis del tráfico capturado (5h)
- Análisis de ataques actualmente posibles, en base a la documentación previa. A nivel de API y de red. (20h)
- Conexión y captación de las claves de cifrado al conectar un dispositivo de terceros. (5h)

- Análisis del tráfico entre Alexa y dispositivos de terceros. (10h)
- Suplantación de identidad con las claves capturadas. (25h)
- Suplantación de voz humana con voz emulada. (25h)
- Grabación del entorno sin estar haciendo uso de Alexa. (15h)
- Otros posibles ataques que puedan ser descubiertos tras los análisis previos. (20h)
- Revisión de documentación de los desarrolladores de Amazon. (20h)
- Revisión de documentación terceros desarrolladores. (10h)
- Redacción y entrega I Informe de Seguimiento, II Informe de Seguimiento y Dossier final. (40h)

Encuentre un diagrama de Gantt de esta planificación en el Anexo I.

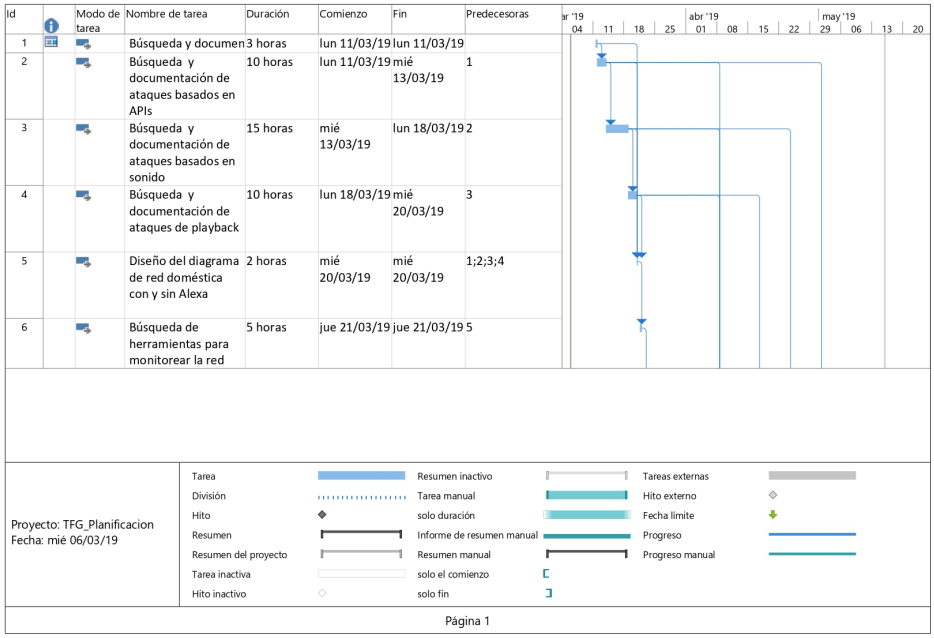
Bibliografía

[1] W. Haack, M. Severance, M. Wallace and J. Wohlwend, "Security Analysis of the Amazon Echo", MIT, 2017. [En línea]. Disponible en: https://courses.csail.mit.edu/6.857/2017/project/8.pdf. [Acedido: 04- Mar- 2019].

[2] "Qué es Kanban: Fundamentos - Kanbanize", Kanbanize.com, 2019. [En línea]. Disponible en: https://kanbanize.com/es/recursos-de-kanban/primeros-pasos/que-es-kanban/. [Accedido: 04- Mar- 2019].

5 Anexos

5.1 Anexo I

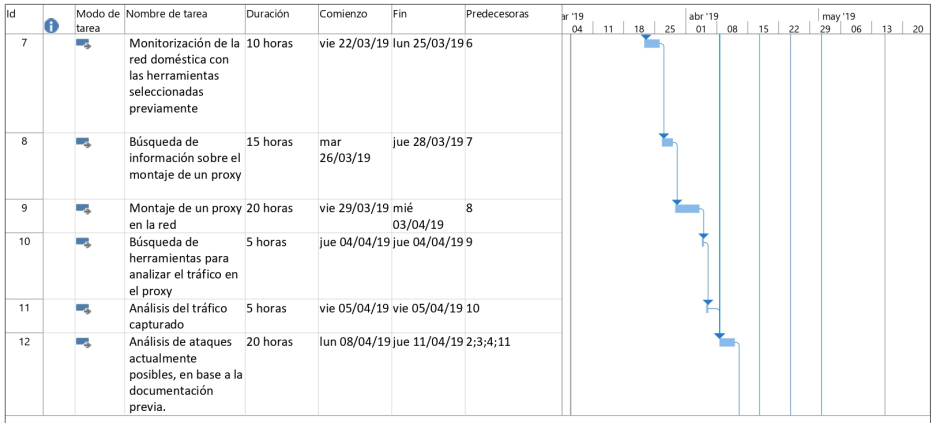


Id	Modo de tarea	Nombre de tarea	Duración	Comienzo	Fin	Predecesoras
1		Búsqueda y documen	3 horas	lun 11/03/19	lun 11/03/19	
2		Búsqueda y documentación de ataques basados en APIs	10 horas	lun 11/03/19	mié 13/03/19	1
3		Búsqueda y documentación de ataques basados en sonido	15 horas	mié 13/03/19	lun 18/03/19	2
4		Búsqueda y documentación de ataques de playback	10 horas	lun 18/03/19	mié 20/03/19	3
5		Diseño del diagrama de red doméstica con y sin Alexa	2 horas	mié 20/03/19	mié 20/03/19	1;2;3;4
6		Búsqueda de herramientas para monitorear la red	5 horas	jue 21/03/19	jue 21/03/19	5

Proyecto: TFG_Planificacion
Fecha: mié 06/03/19

Tarea, División, Hito, Resumen, Resumen del proyecto, Tarea inactiva, Hito inactivo, Resumen inactivo, Tarea manual, solo duración, Informe de resumen manual, Resumen manual, solo el comienzo, solo fin, Tareas externas, Hito externo, Fecha límite, Progreso, Progreso manual

Página 1

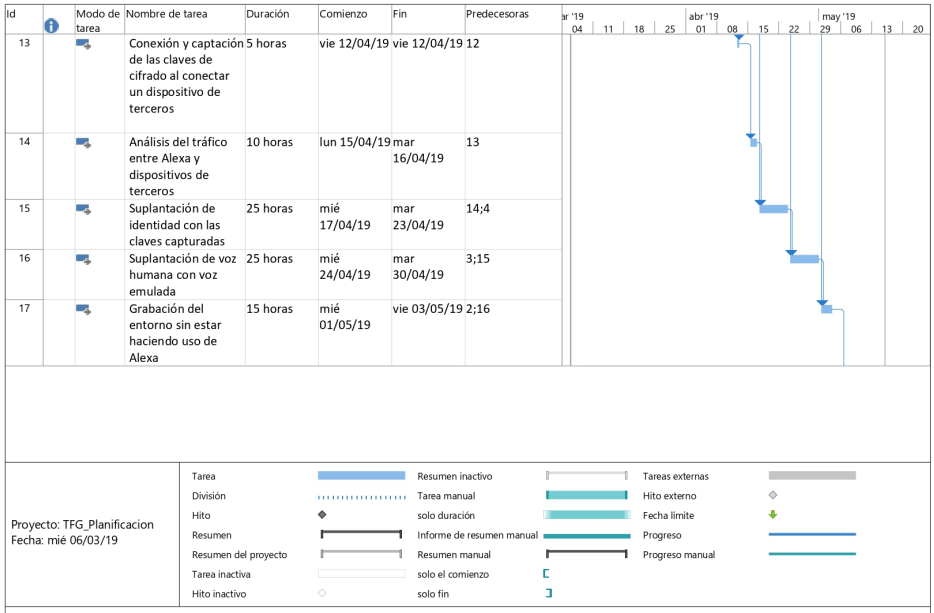


Id	Modo de tarea	Nombre de tarea	Duración	Comienzo	Fin	Predecesoras
7		Monitorización de la red doméstica con las herramientas seleccionadas previamente	10 horas	vie 22/03/19	lun 25/03/19	6
8		Búsqueda de información sobre el montaje de un proxy	15 horas	mar 26/03/19	jue 28/03/19	7
9		Montaje de un proxy en la red	20 horas	vie 29/03/19	mié 03/04/19	8
10		Búsqueda de herramientas para analizar el tráfico en el proxy	5 horas	jue 04/04/19	jue 04/04/19	9
11		Análisis del tráfico capturado	5 horas	vie 05/04/19	vie 05/04/19	10
12		Análisis de ataques actualmente posibles, en base a la documentación previa.	20 horas	lun 08/04/19	jue 11/04/19	2;3;4;11

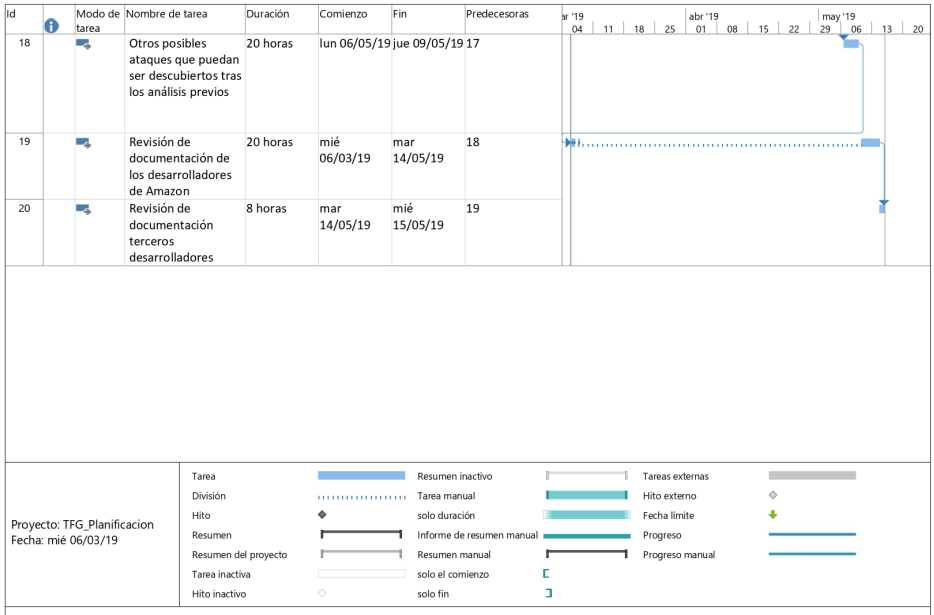
Proyecto: TFG_Planificacion
Fecha: mié 06/03/19

Tarea, División, Hito, Resumen, Resumen del proyecto, Tarea inactiva, Hito inactivo, Resumen inactivo, Tarea manual, solo duración, Informe de resumen manual, Resumen manual, solo el comienzo, solo fin, Tareas externas, Hito externo, Fecha límite, Progreso, Progreso manual

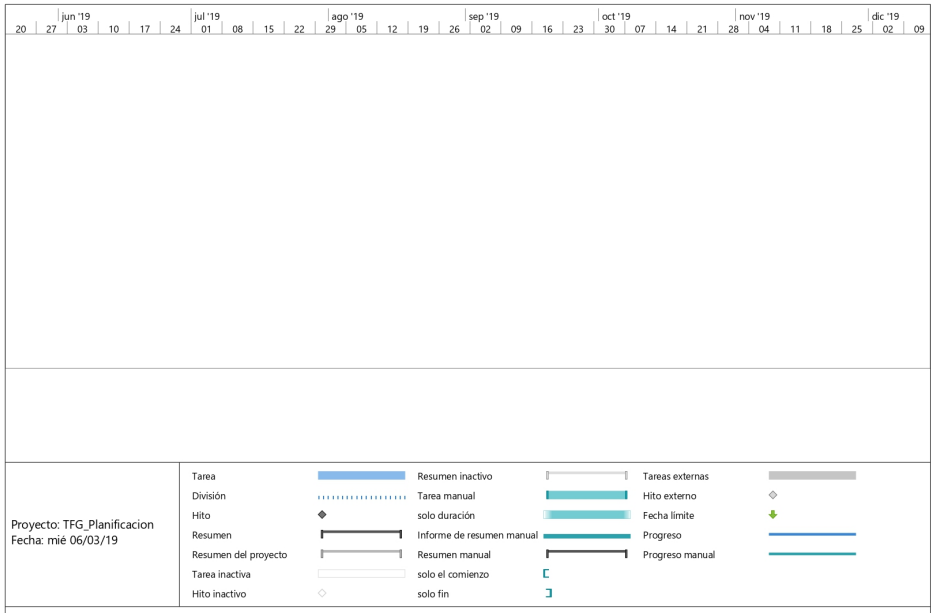
Página 2



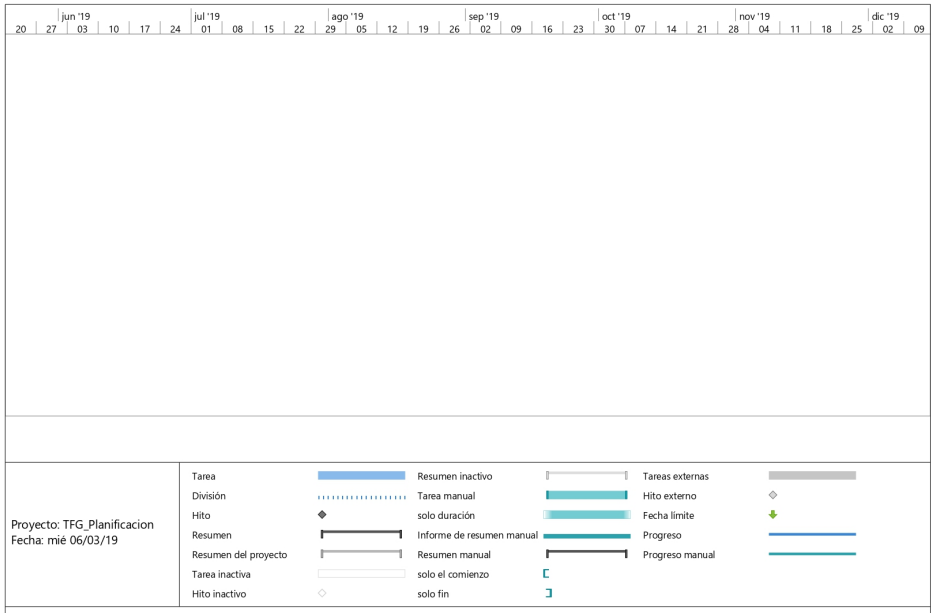
Página 3



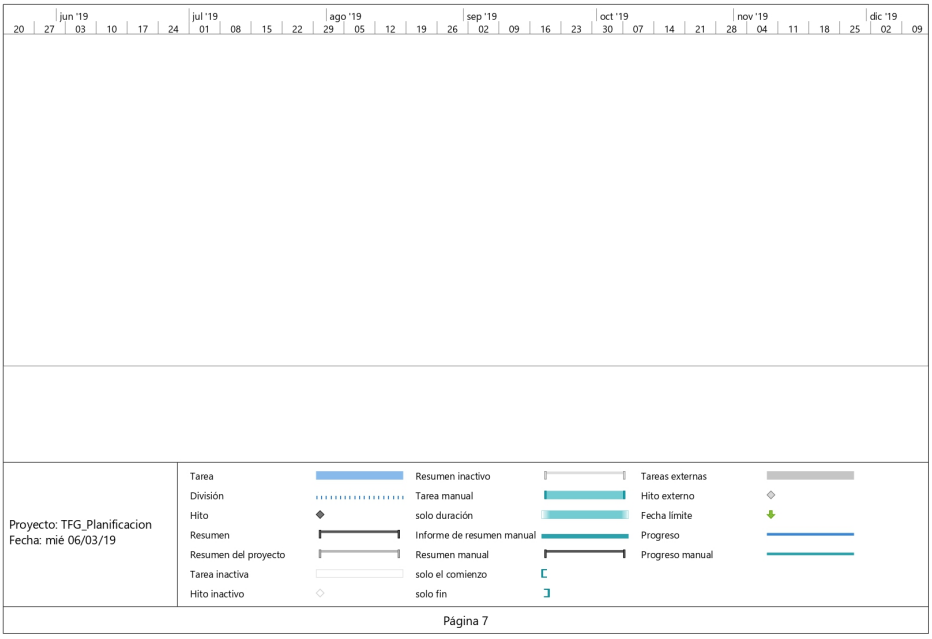
Página 4



Página 5



Página 6


jun '19
jul '19
ago '19
sep '19
oct '19
nov '19
dic '19
20 27 03 10 17 24 01 08 15 22 29 05 12 19 26 02 09 16 23 30 07 14 21 28 04 11 18 25 02 09
Proyecto: TFG_Planificacion
Fecha: mié 06/03/19
Tarea
División
Hito
Resumen
Resumen del proyecto
Tarea inactiva
Hito inactivo
Resumen inactivo
Tarea manual
solo duración
Informe de resumen manual
Resumen manual
solo el comienzo
solo fin
Tareas externas
Hito externo
Fecha límite
Progreso
Progreso manual

Página 7