International Journal of Innovation Engineering and Science Research
(ISSN: 2581-4591)
www.ijiesr.com

Forensic Analysis of Amazon Alexa and Google Assistant Built-In Smart Speakers

Ilkan Yildirim, ErkanBostanci, Mehmet SerdarGuzel
Department of Forensic Informatics, Institute of Forensic Sciences
Ankara University, Ankara, Turkey

ABSTRACT

People's communication with machines is evolving. The process that started with the buttons has evolved to the touchscreen and now people can command to machines just talking with them. The use of smart home assistants, which allows people to control their smart homes, access mail accounts and even order, is becoming increasingly popular. For this reason, it is possible that they will be found at crime scenes soon and carry the value of digital evidence. In this study, the best-selling products Alexa Echo and Google Home were examined in terms of forensic evidence and the data containing digital evidence were found. Then fake activities were created by changing device name, creating fake routine creation and custom skill development. As a result of the investigations, for the cyber security experts or academics working in this field the information was provided about which kind of digital evidence could be found in smart home assistant's activities. Also, difference between real activity and fake activity were elicited against anti-forensic.

Keywords—*IoT Forensic, Alexa, Google Assistant, Anti-Forensic, Fake Activity*

I. INTRODUCTION

People have discovered new ways to communicate with machines throughout history. People had started to use punch cards, and QWERT keyboards, and today continue with touch screens and speech commands with smart assistants like Google Assistant, Amazon Alexa, Siri and Cortana.

Speakers with integrated smart assistants are capable of control IoT devices like lights, TV, door and multimedia systems at home. Number of smart speaker sales increasing exponentially. Almost all home will have this kind of device and, they will be found in crime scenes soon.

Many of latest technology electronic devices have system logs such as cameras, drones and cars. These are valuable assets for the forensic investigations. Not only technology makes people life easier, but also criminals use it for anti-forensic activities[1]. For this reason, cyber security experts should take an action first to protect innocent people.

This study has two parts. In the first part Amazon Alexa with its smart speaker Echo Plus 2nd Generation and Google Assistant with its smart speaker Google Home Mini have been examined to find out whether there are digital evidences. In the second part, anti-forensic cases were created and compared with first part's digital evidences.

The rest of paper organized as follows: Section 2 includes review of background, Section 3 has related works, Section 4 has method, Section 5 includes evaluation, and last part Section 6 consist of conclusion and future work.

Smart speakers have integrated smart assistants that take input from users' speech, then convert it to text, apply NLP on this text, take an action and reply to user as an informative voice message. They can control IoT devices i.e. smart home, order something from anywhere and send money. People can do whatever they want with smart speakers like using their computers or mobile phones. The

difference between smart speakers and other computers, there is no GUI on smart speakers. They are using VUI as a user interface. Because of that, they are listening the environment all time until user uses the wake-up words to activate smart speakers.

To make user commands Alexa have skills and Google Assistant have actions. In this kind of devices more skills or actions means more capability. Skills and actions are almost same, both have development kits for the developers. At the end of 2018, Amazon announced that they have 70.000 skills, and they are increasing 192 per day worldwide [2]. Google Assistant actions reached 4.253 in January 2019; this number rose about 2.5 times last year compared [3].

At fig 1 number of unit sales of smart speakers are increasing exponentially. The study about global smart speaker unit sales shows that only the Q4 sales 8.5 times increased from 4.2 million to 38.5 million between 2016 and 2018 [4].According to research firm Arizton's report about smart speaker market, the size was $991 million in 2016 and will grow $4.8 billion in 2022 [5].

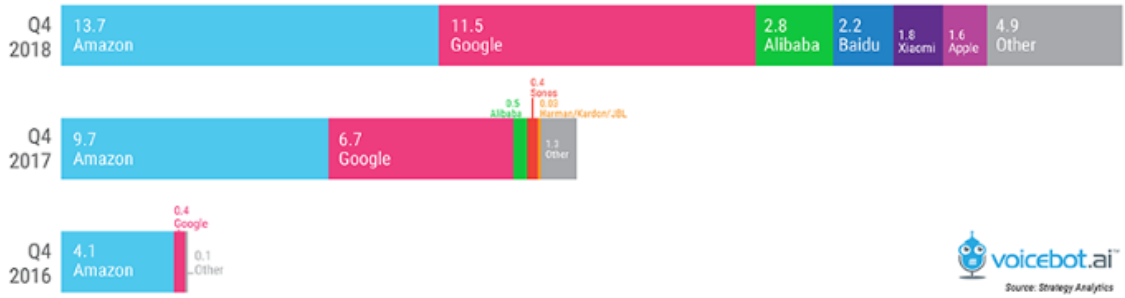

Fig.1. Number of Smart Speaker Unit Sales (Million)

There are studies about IoT forensics and smart speakers. Smart speakers are listening environment until hearing wake-up words. There is a study about smart speaker Alexa whether listening environment until wake-up words or recording everything to process data. They have used network forensic tools 21 days. The traffic was encrypted. They couldn't find any evidence about Alexa is listening all time [6]. Scope of another study consists of cloud-based asset analysis, voice command tests, application analysis and firmware analysis. The goal was emphasizing vulnerabilities of Alexa. As a result of that, they found that Alexa is recording environment sounds after hearing wake up words even if there are no meaning in this sentence[7].

Some of the studies are held on about forensic investigation processes for the smart home ecosystems. Researches were analyzed Echo Dot with Alexa, and they found evidences. First, they analyzed network traffic data to find API of Alexa. Then they send request to gather evidence. They found activity records, personal information etc. As a result, they created a tool theoretically for this purpose [8].

The study about Alexa vulnerability called "Skill squatting attacks" that focuses on accent of command. They used 11.460 American English words experimentally to see percentage of Alexa's misunderstanding. After realizing there is a systematic misunderstanding on Alexa, they attacked by using these words. They called this attack type called as "skill squatting attacks"[9]. Another vulnerability case study attack applied theoretically on Amazon Alexa and Google Home by ordering something fake and unlocking the door remotely.

The researchers mention about many of the forensic studies about smart speakers were holding on theoretically. They showed that there are some studies about smart speaker Amazon Echo and IoT

forensics Z-wave protocol practically and emphasized that researchers should give importance to this issue[10].

II. METHOD

In this study, experimental environment was arranged in order to produce data. LM-G710EM model number LG ThinQ Q7 mobile phone with Android 8.0.0 is used. For the smart home environment Google Home Mini H0A and Amazon Alexa Echo Plus 2nd Generation L9D29R used as smart speakers with their android mobile applications. 2.11.1.8 version of Google Home application and Amazon Alexa version 2.2.271281.0 version was installed. Amazon Alexa application does not work in Turkey. Therefore, the Google Play account settings have been changed to the United States. The SENGLED bulb R11-G13 was connected to Echo and GE C-Life Smart Bulb A19 was connected to Google Home Mini by using wireless.

Fig.2. Google Home Mini ve Amazon Echo Plus 2nd Generation

Alexa Skill was created by using Alexa Skills Kit in Amazon Web Services Lambda functions. In order to create fake activity, an intent code was written by using node.js. In this skill user commands for sum of two numbers, Alexa response as multiply of these two numbers. Same skill was built as a Google Action by using Google Dialog flow Fulfillment inline editor. In figure 3 shows that interaction between Alexa Skills Kit and Alexa Voice Services. Alexa allows developers to use both.

Fig.3. Alexa Web Services[11]

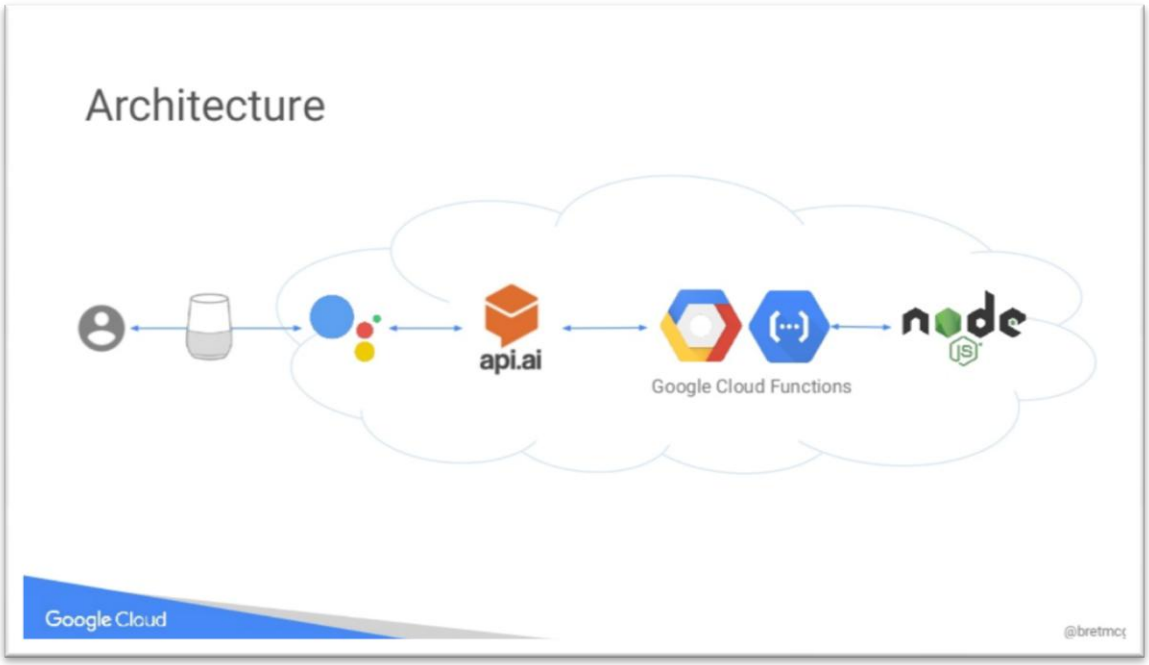

Fig.4. Google Assistant Architecture [12]

In figure 4 there is google cloud architecture, google device listens user that sends the audio to google assistant. Google Assistant uses the api.ai and Google Cloud Functions with nodejs.

TableI. CASE DETAILS

Cases	***Command***	***Command Type***	***First State***	***Expected Action***	***Expected Action***
1	Turn on light	Daily	Lights Off	Lights On	Lights On
2	Turn on horse	Edited Device Name	Lights Off	Lights On	Lights On
3	Turn on TV	Custom Routine	Lights Off	Lights On	Lights On
4	"Fake activity", "Add 5 and 6"	Skill/Action	-	Add	Multiply

In first case, as a regular user, "turn on light" command sent to both smart speakers. This is daily common usage of smart home devices. For this reason, this case was used as a reference point to compare other cases. The other cases are for the Anti-Forensic to mislead forensic investigator.

For the second case, name of device changed from "light" to "horse", then "turn on horse" command send to both smart speakers.

In third case, the routine was created. The invocation of the routine sentence was chosen as "Turn on TV", then routine action was set as turning on the light, and routine response set as "Ok." same as daily activity. After that "Turn on TV" command sent to both devices.

The last case was about creating custom skills by using developer kits of both devices. For the summation intent of two numbers, multiplication response was created. Firstly, invocation word, then "Add 5 and 4" commands sent to both devices.

Google Home application and Amazon Alexa application have activity histories. For the four cases, these activity histories were found in user privacy settings menu, after clicking this, applications redirect to the web pages.

III. EVALUATION

As a result of the first "turn on lights" daily case, the digital evidences were found in Alexa and Google Assistant's activity histories. Every user can see these histories on the mobile applications or the web interface of Amazon Alexa or Google Assistant's activity history. For the more advanced users, they can get these information by using the APIs. Table II shows the Google Assistant's activity details.

TABLE II. GOOGLE ASSISTANT ACTIVITY DETAILS

Google Assistant History	**"Turn on Lights" Command**
User Command's Text	Said turn on lights
User Command's Voice Recording	Play Button to Listen Recording
Time stamp	Today at 1:18 PM
Assistant's Response	Ok, turning the Light on.
Device Type	Smart Speaker
Device's Approximate Location	Google Map Image
Started By	Hotword
Action Type	com.google.homeautomation

Table III shows the Alexa's activity details.

TABLE III. AMAZON ALEXA ACTIVITY DETAILS

Alexa History	**"Turn on Lights" Command**
User Command's Text	"turn on light"
User Command's Voice Recording	Play Button to Listen Recording
Time stamp	Today at 02:07 PM
Assistant's Response	"ok"
Device Name	<user>'s Echo Plus

Both activity history has user command text, its voice recording, time stamps, and assistant's response. While Google is storing device type like Google Home or Android Application, Alexa is just storing the device name. In addition to Alexa, Google is storing approximate location of device, started by and action type. At the fig 5 shows that amazon Alexa history on its mobile application.

For the second case, the device names were changed from "light" to "horse" in both mobile applications to mislead forensic investigator. As a result, for the second case activity, there were no difference between daily case and this case in activity history records both Amazon Alexa and also Google Assistant. Thus, to find what horse is, forensic investigator should record device names by using mobile application if the user didn't change the name of devices.

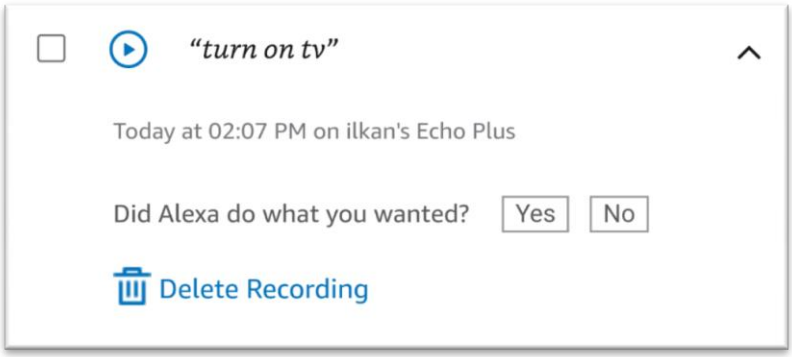

Fig.5. Amazon Alexa 3nd Case Activity History

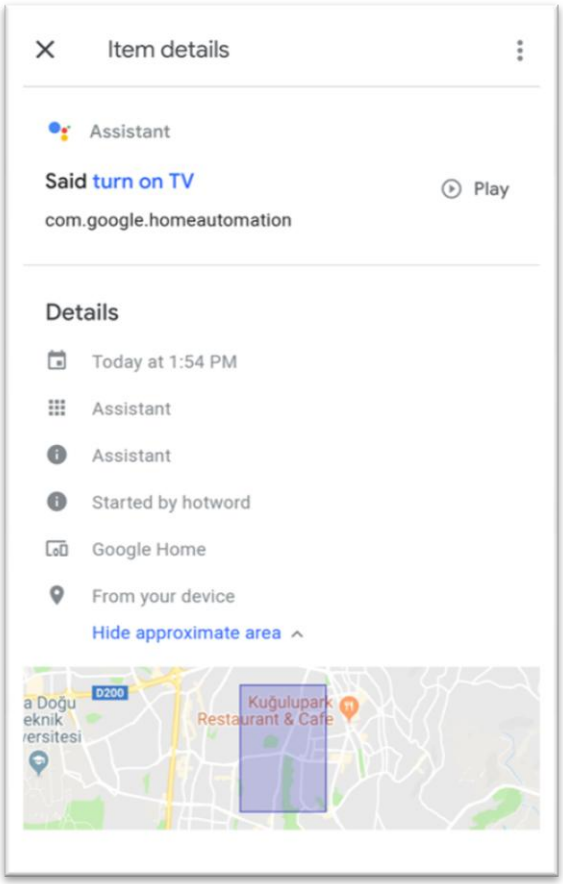

Fig.6. Google Assistant3nd Case Activity History

In third case, created routine command's activity history recordings were different according to daily activity history. In both activity history, the missing part was assistant's response although it was defined. This is the clue for the forensic investigators. At this point they should record all the routines carefully.

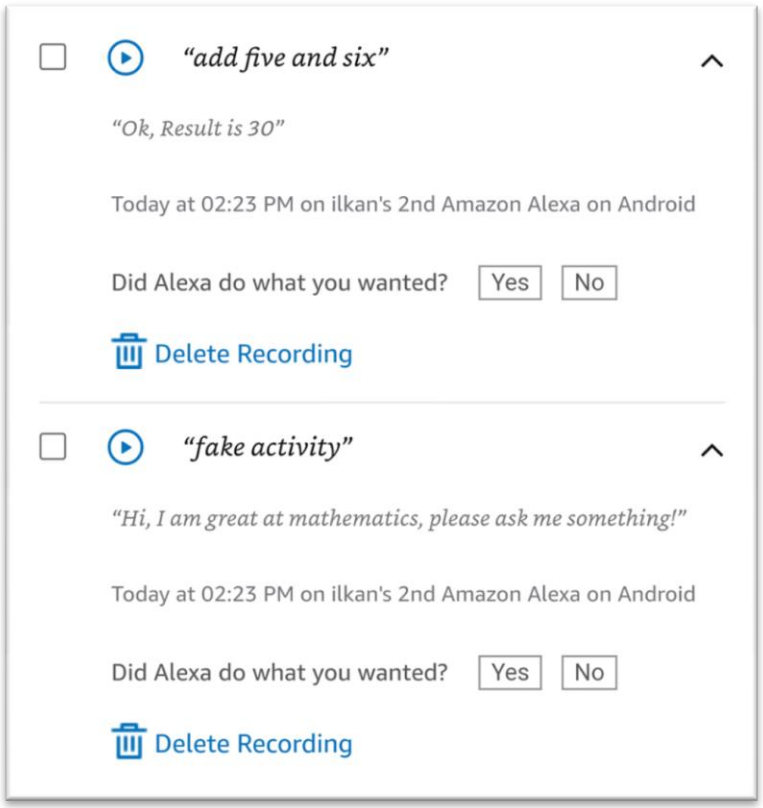

Fig.7. Amazon Alexa Custom Skill Activity History

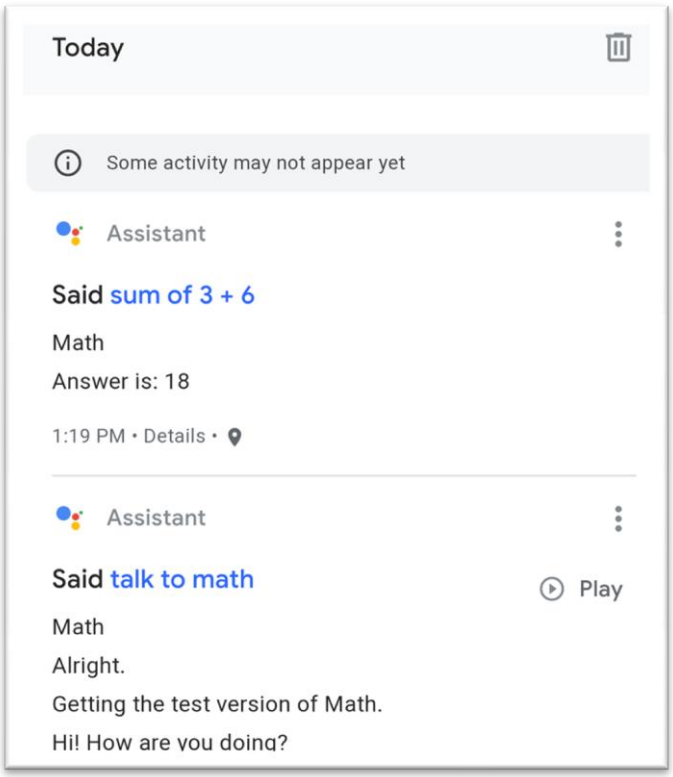

Fig.8. Google Assistant Custom Action Activity History

The last case creating custom actions' and skills' activity initiation history were same as daily case. At the fig 8 Google Assistant's intent activity histories were different than daily case which activity recording's missing parts were user command's voice recording and assistant's response. There was no missing part Alexa's in activity history which was same as daily case. At this point the only clue for the forensic investigators in Google Assistant's histories. At the fig 7 For the Alexa part, criminals can create anti-forensic fake activities easier by comparison to Google Assistant.

It was seen that both products store user data using cloud computing to improve their products and according to GDPR [13] they should present these data to users in readable format. That makes life easier for the forensic investigators.

IV. Conclusion

In this study, Amazon Alexa Echo and Google Home Mini smart assistants were selected by looking at the number of unit sales in the market and analyzed with forensic perspective. First, daily user command sent to both devices and its activity recording, then which compared with anti-forensic fake activities such as changing device name, creating routine and building custom skills in order to mislead forensic investigator. Clues and vulnerabilities were unearthed for these cases from activity histories.

Illogical requests with customized skills or actions allow users to perform different operations and create fake activity history recordings. Preventive studies can be done by Alexa and Google Assistant team.

Forensic analysis of smart assistants is a new field of study. The research has been done on Android devices, but it is obvious that there isn't any difference for IOS devices as both applications are directed to the web browser for activity histories. Criminal experts have no chance to take an image of device for analyzing, because all the data are uploading to cloud servers, but can request these data from Amazon and Google companies.

References

[1] D. Lillis, B. A. Becker, T. O'Sullivan and M. Scanlon, "Current Challenges and Future Research Areas for Digital Forensic Investigation," in *The 11th ADFSL Conference on Digital Forensics, Security and Law (CDFSL 2016)*, Daytona Beach, FL, USA, 2016.

[2] B. Kinsella, "There are Now More Than 70,000 Alexa Skills Worldwide, Amazon Announces 25 Top Skills of 2018," 14 12 2018. [Online]. Available: https://voicebot.ai/2018/12/14/there-are-now-more-than-70000-alexa-skills-worldwide-amazon-announces-25-top-skills-of-2018/. [Accessed 06 06 2019].

[3] "Google Assistant Actions Total 4,253 in January 2019, Up 2.5x in Past Year but 7.5% the Total Number Alexa Skills in U.S.," 15 02 2019. [Online]. Available: https://voicebot.ai/2019/02/15/google-assistant-actions-total-4253-in-january-2019-up-2-5x-in-past-year-but-7-5-the-total-number-alexa-skills-in-u-s/. [Accessed 06 06 2019].

[4] B. Kinsella, "Amazon Increases Global Smart Speaker Sales Share in Q4 2018, While Google Rise Narrows the Gap and Apple Declines," 20 02 2019. [Online]. Available: https://voicebot.ai/2019/02/20/amazon-increases-global-smart-speaker-sales-share-in-q4-2018-while-googles-rise-narrows-the-gap-and-apple-declines/. [Accessed 02 06 2019].

[5] Arizton, "Arizton Says Smart Speaker Market $4.8 Billion in 2022 - Voicebot," 4 1 2018. [Online]. Available: https://voicebot.ai/2018/01/04/arizton-says-smart-speaker-market-4-8-billion-2022. [Accessed 5 5 2019].

[6] M. Ford and W. Palmer, "Alexa, are you listening to me? An analysis of Alexa voice service network traffic," *Personal and Ubiquitous Computing 23(1),* pp. 1-13, July 2018.

[7] H. Chung, M. Iorga, J. Voas and S. Lee, "'Alexa, Can I Trust You?'," *Computer 50,* pp. 100-104, 2017.

[8] J. P. S. L. Hyunji Chung, "Digital forensic approaches for Amazon Alexa ecosystem," *Digital Investigation Volume 22, Supplement,* pp. Pages S15-S25, August 2017.

[9] D. Kumar, R. Paccagnella, P. Murley, E. Hennenfent, J. Mason, A. Bates and M. Bailey, "Skill squatting attacks on amazon alexa," *SEC'18 Proceedings of the 27th USENIX Conference on Security Symposium*, pp. 33-47, 2018.

[10] C. Shin, P. Chandok, R. Liu, S. J. Nielson and T. R. Leschke, "Potential Forensic Analysis of IoT Data: An Overview of the State-of-the-Art and Future Possibilities," in *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Exeter, 2017.

[11] Amazon Web Services, «Creating IoT Solutions with Serverless Architecture & Alexa,» 01 Şubat 2014. [Online]. Available: https://www.slideshare.net/AmazonWebServices/creating-iot-solutions-with-serverless-architecture-alexa. [Accessed: 01 Haziran 2019].

[12] B. McGowen, «Google Home and Google Assistant Workshop: Build your own serverless …,» 28 July 2017. [Online]. Available: https://www.slideshare.net/bretmc/google-home-and-google-assistant-workshop-build-your-own-serverless-action-on-google-app. [Accessed: 02 Haziran 2019].

[13] GDPR, "General Data Protection Regulation GDPR," 2018. [Online]. Available: https://gdpr-info.eu/art-12-gdpr/. [Accessed 01 06 2019].