



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής  
Πρόγραμμα Μεταπτυχιακών Σπουδών  
«Προηγμένα Συστήματα Πληροφορικής - Ανάπτυξη Λογισμικού και Τεχνητής  
Νοημοσύνης»

**Μεταπτυχιακή Διατριβή**

Τίτλος Διατριβής	<b>Επιθέσεις σε Φωνητικούς Βοηθούς με χρήση συντιθέμενης φωνής Attacks on Voice Assistants using synthesized voice</b>
Όνοματεπώνυμο Φοιτητή	<b>Δόμνα Μπίλικα</b>
Πατρώνυμο	<b>Νικόλαος</b>
Αριθμός Μητρώου	<b>ΜΠΣΠ/ 19033</b>
Επιβλέπων	<b>Κωνσταντίνος Πατσάκης, Αναπληρωτής Καθηγητής</b>

**Τριμελής Εξεταστική Επιτροπή**

(υπογραφή)

(υπογραφή)

(υπογραφή)

Κωνσταντίνος Πατσάκης  
Αναπληρωτής Καθηγητής

Ευθύμιος Αλέπης  
Αναπληρωτής Καθηγητής

Μαρία Βίρβου  
Καθηγήτρια

## **Acknowledgements**

First, I would like to thank my family who is by my side at every moment of my life and gives me strength and courage to continue. Also, my supervisor professor Mr. Patsakis Constantinos for the help and support he provided me, and the knowledge he passed on to me. Without the trust and patience of him and the other professors in the department, it would not have been easy to accomplish my goal. In addition, I would like to thank the Onassis Foundation for the scholarship it provided me during my studies. Finally, I would like to express my gratitude and respect towards my colleagues at work for accommodating to my university schedule on numerous occasions allowing me to work in my dissertation and studies when needed.

## Abstract

Nowadays, Digital Assistants are increasingly invading our daily lives by offering facilities. A typical example is the “Smart Home”, which promises to bring many positive changes in our lives. New technologies allow the creation of “Smart Homes”, incorporating devices that make the home a better place in terms of safety, comfort, and well-being. It consists of “smart” electronic devices connected to the internet, which can operate independently or in a network with other devices. These devices receive commands from the user to perform their operations. Users can train each device to understand their own voice and then respond to the commands users give. But are these devices only loyal to their “masters” or is there a risk of deception? This was the key question for starting this master's thesis.

The aim of the present study, therefore, is to investigate the security of DAs. Several attack attempts experiments have been performed using data from different media so that users are aware of the risks that they may be exposed to.

## Περίληψη

Στις μέρες μας οι Ψηφιακοί Βοηθοί εισβάλλουν όλο και περισσότερο στην καθημερινότητά μας προσφέροντας διευκολύνσεις. Χαρακτηριστικό παράδειγμα αποτελεί το «Έξυπνο Σπίτι» (“Smart Home”), το οποίο υπόσχεται να φέρει πολλές αλλαγές στην ζωή μας. Οι νέες τεχνολογίες επιτρέπουν τη δημιουργία «Έξυπνων Σπιτιών», ενσωματώνοντας συσκευές που καθιστούν το σπίτι ένα καλύτερο μέρος όσον αφορά την ασφάλεια, την άνεση και την ευημερία. Αποτελείται από «έξυπνες» ηλεκτρονικές συσκευές συνδεδεμένες στο διαδίκτυο, οι οποίες μπορεί να λειτουργούν αυτόνομα ή σε δίκτυο με άλλες συσκευές. Αυτές οι συσκευές λαμβάνουν εντολές από το χρήστη για την εκτέλεση των λειτουργιών τους. Οι χρήστες μπορούν να εκπαιδεύσουν κάθε συσκευή ώστε να καταλαβαίνει τη φωνή τους και έπειτα να ανταποκρίνονται στις εντολές που δίνουν οι χρήστες. Είναι όμως οι συσκευές αυτές πιστές μόνο στα «αφεντικά» τους ή υπάρχει κίνδυνος εξαπάτησής τους; Αυτό ήταν το βασικό ερώτημα για την έναρξη αυτής της μεταπτυχιακής διατριβής.

Στόχος, λοιπόν, της παρούσας έρευνας είναι ο έλεγχος της ασφάλειας των ψηφιακών βοηθών. Έχουν πραγματοποιηθεί αρκετές απόπειρες επιθέσεων χρησιμοποιώντας δεδομένα από διαφορετικά μέσα, έτσι ώστε να γνωρίζουν οι χρήστες τους κινδύνους στους οποίους ενδέχεται να εκτεθούν.

## Contents

1. Introduction .....	8
2. Related Work .....	10
2.1. The Evolution of Smart Home .....	10
2.2. Security and Privacy on Voice Assistants .....	12
3. Data Collection and Manipulation .....	13
3.1. Voice Synthesis .....	13
3.2. Face to Face .....	14
3.3. Via Call .....	14
3.4. Spy Applications .....	15
3.5. Youtube .....	16
4. Attacks .....	17
4.1. Attack via Android App .....	21
4.1.1. Project Structure and Permissions .....	21
4.1.2. Execution Example .....	22
4.1.3. Database .....	23
4.1.4. Code Presentation .....	24
4.2. Direct Recording Attack .....	25
5. Conclusion and Future Work .....	27
6. References .....	28
7. Appendix .....	29
7.1. Examples of Google Assistant Commands .....	29
7.1.1. Google Assistant Commands for Beginners .....	29
7.1.2. Controlling Music, Podcasts, Radio and Audiobooks with a Google Smart Speaker .....	29
7.1.3. Setting Timers and Alarms .....	30
7.1.4. Checking Calendars and Reminders .....	30
7.1.5. General Queries and Commands .....	30
7.1.6. Commands for Google Assistant Smart Home Control .....	31
7.1.7. Phone and Calls Commands .....	31
7.1.8. Commands for kids .....	31
7.1.9. Broadcasting Commands .....	32

## Abbreviations

The following table describes the significance of various abbreviations and acronyms used throughout the thesis. The page on which each one is defined or first used is also given. Nonstandard acronyms that are used in some places to abbreviate the names of certain white matter structures are not in this list.

Abbreviation	Meaning	Page
AI	Artificial Intelligence	8
BLE	Bluetooth Low Energy	24
DA	Digital Assistant	8
IDE	Integrated Development Environment	21
IoT	Internet of Things	9
IPA	Intelligent Personal Assistant	18
IVA	Intelligent Virtual Assistant	8
NLP	Natural Language Processing	8
SPA	Smart Home Personal Assistant	10
TTS	Text-To-Speech	14
VA	Voice Assistant	10
VAA	Voice & Audio Activity	9
VPA	Voice Personal Assistant	11

## Figures

Figure 1 Digital Assistant.....	8
Figure 2 Google Assistant: How it works.....	11
Figure 3 SV2TTS toolbox launch.....	13
Figure 4 Data collection via call.....	15
Figure 5 Spy Application.....	16
Figure 6 Permissions.....	16
Figure 7 Youtube upload.....	16
Figure 8 Comparison of notable assistants.....	17
Figure 9 Google Home Mini listening (white LEDs).....	18
Figure 10 Google Home Mini on a call (blue LEDs).....	18
Figure 11 Google Account connection on Duo application.....	19
Figure 12 Unlink Duo account.....	19
Figure 13 Duo set up.....	19
Figure 14 Google Home Mini training.....	19
Figure 15 Google Assistant settings.....	20
Figure 16 Enable Google Assistant when phone is locked.....	20
Figure 17 Project structure.....	21
Figure 18 Permissions.....	21
Figure 19 Location permission.....	22
Figure 20 Bluetooth permission.....	22
Figure 21 Turning Bluetooth on.....	22
Figure 22 Change Bluetooth connectivity.....	23
Figure 23 Detected devices.....	23
Figure 24 Storage Audio folder.....	23
Figure 25 Firebase storage.....	24
Figure 26 Database connection, Bluetooth and location permissions.....	24
Figure 27 Start devices scan.....	24
Figure 28 If Bluetooth device exists audio file is played.....	25
Figure 29 Voice Synthesis training.....	26
Figure 30 Voice attack.....	26
Figure 31 Google Assistant response.....	26

## 1. Introduction

Since 2017, Digital Assistants (DAs also known as Intelligent Virtual Assistant - IVA) are used more and more as their capabilities are rapidly increasing, with the manufacturing of new products. Apple and Google have big databases of users' information installed on smartphones. Microsoft, specifically, has a large installed database of Windows - based personal computers, smartphones, and smart speakers. Also, Amazon has a large installed database for smart speakers. DAs defined as devices - usually speakers - that use advanced artificial intelligence (AI) or, to be more precise, quite advanced algorithms. These systems can perform tasks or services for an individual based on commands or questions. DAs are extremely useful for people with mobility problems and the elderly citizens. They also use natural language processing (NLP), natural language understanding, and machine learning to learn as they go and provide a personalized, conversational experience. Combining historical information such as purchase preferences, home ownership, location, family size, and so on, algorithms can create data models that identify patterns of behavior and then refine those patterns as data is added. By learning a user's history, preferences, and other information, DAs can answer complex questions, provide recommendations, make predictions, and even initiate conversations.

DAs offer many benefits on an individual as well as on a business level. Most importantly, they facilitate the daily lives of users by performing functions to manage electrical appliances, even those related to home security. They also help save money for people as well as relieve them from tedious work. On the other hand, in a business environment they can help improve communication between both business members and customer service. They reduce the costs of the business spent on face-to-face contact with customers, relieving employees of time-consuming procedures they have had to carry out in the past. They often get accurate answers in a shorter time by using DAs as long as they do not have to wait for an employee to provide them the needed information. In addition, they improve the quality of the products by providing more immediate answers. DAs are always there for users to inform them of any important work to be done through reminders. A very important benefit is that they can simultaneously serve a large percentage of people. Finally, with prolonged use they can gather more useful information to improve the user experience.

DAs can be used in a variety of ways. For the needs of this master thesis, communication through the user's voice was utilized. To understand the commands they receive, they use NLP which translates the voice commands into executable, creating precise answers. Each voice command can consist of one or more sections that need to be addressed. The NLP can separate these sections and



Figure 1 Digital Assistant

edit them, providing comprehensible answers to users. Of course, not all DAs have the same capabilities. Based on the history and preferences of each user, with the help of advanced



algorithms, they can make predictions about their behavior. Each company needs its own DA activation command.

NLP is about interaction between people and computers. It processes and analyzes large amounts of natural language data. By using this a computer is able to understand the user's commands both in terms of text and in terms of perceiving the tone of his voice. Thus, it can be trained to obey only one user. Every DA requires user training before using it for the first time. Due to the increased volume of data they receive daily, they are becoming more and more efficient. However, training in a single voice is still quite difficult for them. It is worth noting that not all voice commands work in the same way, since there are commands that are executed by everybody, including the voice of the user who “trained” the DA, but there are also commands that require to have been uttered specifically by the user who trained this assistant.

As DAs become more popular, there are increasing legal risks involved. Enabling DAs to be used by voice commands has raised privacy concerns. This is because the user is required to grant permission so that the device always listens to the user's conversations without recording until the specific command is spoken. Each company processes the data in its own way based on the user's choices. For example, Google Assistant does not store data without the user's permission. Storage requires the user to go to Voice & Audio Activity (VAA) and activate the corresponding function so that the user's data is sent to the cloud and used by Google to improve the performance of the DA. Amazon's DA Alexa, on the other hand, always records conversations, once enabled, and stores them in the cloud. If the users wish, they can delete their recordings by going to the privacy settings, but they cannot avoid the initial save in the cloud storage. Apple DA uses audio copies to enhance it. The data sent for analysis is the most insignificant for the user, for example if the user asks the DA to read a message, which is significant to the user, they will not send this data for analysis. Of course, the users can prevent their data from being sent to the cloud through the corresponding settings.

AI is constantly evolving and along with it the DAs do as well. The future of AI promises smarter DAs with greater efficiency in even the most complex of questions. 5G is a technology that is expected to support this development. This will help DAs to respond faster and more accurately. Almost everyone will have at least one DA for personal use as well as in their the professional environment. This will save time and human resources. It will make people more creative by providing them with more free time and releasing them of unnecessary and lengthy tasks and procedures.

The term “Smart Home” refers to an installation where devices can be automatically controlled remotely from anywhere with an Internet connection using a mobile or other networked device. Many devices connected to the same network communicating with one another, increase automation much more through what is known as the Internet of Things (IoT). The user is able to control functions remotely such as access to the house, activation of a device, control of the alarm system etc. These devices are usually connected to a central “gateway”. Desktops, mobile applications, or web applications connected to the Internet are used for user control. This way the owners can control all the appliances, lighting, and so many more other functions through a personal device even if they are physically far away from it. At any time, they are aware of any operation of the house through relevant notifications. However, there are many risks, mainly in terms of security issues that affect both manufacturers and users. There is a possibility of violation of these devices with the most important being the entrance to these houses. Efforts are being made to address these issues by using strong passwords that are sometimes subject to encryption and connecting only trusted devices to the network.

In the next sections there is a refer to similar research on the safety of DAs. In addition, examples of voice synthesis are presented from data collected from various media. Attacks on DAs are analyzed and finally, conclusions are drawn and some thoughts for future research are suggested.

## **2. Related Work**

With progress, yet more additional needs are created to facilitate the daily life of people. Evidence of this, is that in 2020 around 4.2 billion Voice Assistant (VA) devices were purchased around the world. Forecasts suggest that by 2024, the number of VAs will reach 8.4 billion units – a number higher than the world’s population. DAs are a feature found in many consumer electronics devices. They can respond to commands, provide users with information, and assist in the control of other connected electronics. There are over 110 million DA users in the United States alone, and the software is especially common in smartphones and smart speakers. Gradually, more consumers demand interaction with their devices and companies have to offer new solutions to meet these needs. Tech consumers can now communicate with their connected homes and vehicles in much of the same way that they can do so with their smartphones. This raises concerns about the security that such infrastructures can offer to users.

### **2.1. The Evolution of Smart Home**

As early as 2010 the concept of “Smart Home” had begun to become popular. According to the article “Applications, Systems and Methods in Smart Home Technology: A Review” published in 2010 [1], the best definition of Smart Home technology is: the integration of technology and services through home networking for a better quality of living. People who are elderly or disabled benefit the most from a home automation system that employs AI. These systems offer to those who are less mobile, or in delicate health, the opportunity to be independent, rather than staying in an assisted living facility. Since then, security has been a major issue in Smart Home applications. Along with the development of “Smart Homes”, security issues were also be addressed. To identify these issues Robles, R. J. et al. [2] studied Smart Home and security and reviewed the tool related to Smart Home security.

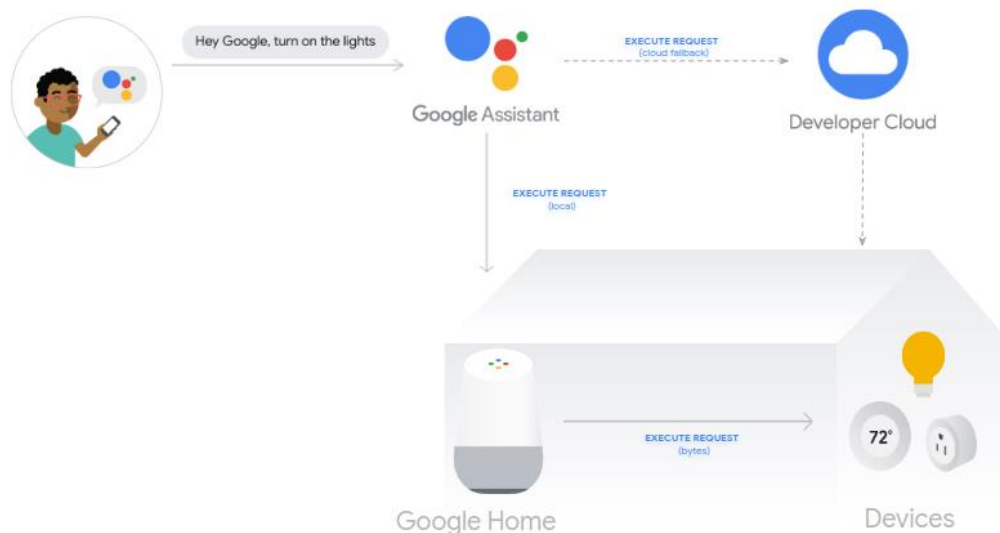
The article by Edu, J. S. et al. [3] presents an in-depth review of the security and privacy issues of Smart Home Personal Assistants (SPAs), categorizing the most important attack vectors and their countermeasures. SPAs are an emerging innovation that have been changing the means of interaction between users and technology. One of their key findings is that even though there has been a significant amount of recent research efforts in this area, it has so far focused on a small part of the attack surface, particularly on issues related to the interaction between the user and the SPA devices. As they claimed, this was the first article to conduct such a comprehensive review and characterization of the security and privacy issues and countermeasures of SPA. This article analyzes and classifies the security and privacy issues associated with the SPA and how a great number of malicious agents can exploit them to undermine the security and privacy of end users.

“Google Assistant Controlled Home Automation” [4] paper presents a proposal for home automation using voice via Google Assistant. As far as data shown in 2018, to make a Smart  
Επιθέσεις σε Φωνητικούς Βοηθούς με χρήση συντιθέμενης φωνής

Home a basic setup of 250\$ (USD) was needed. Google Home price was around 150\$ (USD) with an additional cost of the devices to be connected to, thus the total cost of the system reached over 250\$ (USD). Apple Home Kit proved to be 100\$ (USD) more than the Google Home just for a basic setup. Philips Hue, a smart light which is controlled by the Google Assistant, Amazon Echo and Siri, VA by Apple was priced around 145\$ (USD). Similarly, Belikin's Wemo light was priced around 44\$ (USD) per unit and this can be controlled both by Siri and Google Assistant. They accomplished to propose a cost-effective voice controlled (Google Assistant) home automation controlling general appliances found in one's home. This resulted in a highly reliable and efficient system for the elderly and people with different needs such as those on a wheelchair who are dependent on others.

The authors of the study "Design and Implementation of IoT-Based Smart Home Voice Commands for disabled people using Google Assistant" [5] aim to exploit the potentials of a Voice Personal Assistant (VPA) and to enable the user to control a device remotely and quickly with much more ease. This paper focused on Google Assistant using voice and the authors explore capabilities such as search on the internet, schedule events, set alarms, control appliances, etc. They built an application which can control the scrolling text message over a LED dot matrix display using this Google Assistant Platform over the phone. It was an IoT based application that utilizes another AI powered platform for controlling. So, Google Assistant controlled LED scrolling message display was successfully implemented in this work. This system provided user the liberty to use either a voice-based command or a text command. The developed system was highly responsive and performance wise.

Using a DA is very easy and offers many possibilities to the users. As shown in the image below after activating its function with a voice command such as 'Hey Google' or 'Ok Google', the user has access to a variety of features either locally or via the cloud. The ease of use of VAs was the reason for all the above research.



**Figure 2 Google Assistant: How it works**

Επιθέσεις σε Φωνητικούς Βοηθούς με χρήση συντιθέμενης φωνής

Despite the growing research on SPA security and privacy, little is known about users' security and privacy perceptions concerning SPA complex ecosystem, which involves several elements and stakeholders. To explore this, Abdi, N. et al. [6] considered the main four use case scenarios with distinctive architectural elements and stakeholders involved: using builtin skills, third-party skills, managing other smart devices, and shopping, through semi-structured interviews with SPA users. They concluded that users have incomplete mental models of SPAs, leading to different perceptions of where data is being stored, processed, and shared. Users' understanding of the SPA ecosystem is often limited to their household and the SPA vendor at most, even when using third-party skills or managing other smart home devices. This leads to incomplete threat models (few threat agents and types of attacks) and non-technical coping strategies they implement, so as to protect themselves. They also found that users are not making the most of the shopping capabilities of SPA due to security and privacy concerns and while users perceive SPA as intelligent and capable of learning, they would not like SPA learning everything about them. Users misunderstood SPA ecosystem, with most of them showing a very limited conception of SPA and inaccurate and incomplete mental models of the SPA ecosystem and related data activities (processing, storing, sharing, and learning).

## 2.2. Security and Privacy on Voice Assistants

Attempting to infringe on the Android operating system is not an emerging object of research. In every version of Android, an attempt is made to find vulnerabilities. The paper written by Diao, W. et al. [7] presents an approach (GVS-Attack) to launch permission bypassing attacks from a zero permission Android application (VoicEmployer) through the speaker. This application was installed in Android versions that existed in 2014 (5.0 – 5.1.1). The idea of GVSAttack utilizes an Android system built-in VA module – Google Voice Search. Through Android Intent mechanism, VoicEmployer triggers Google Voice Search to the foreground, and then plays prepared audio files (like “call number 1234 5678”) in the background. Google Voice Search can recognize this voice command and execute corresponding operations. GVS-Attack can forge SMS/Email, access privacy information, transmit sensitive data and achieve remote control without any permission. Also, they found a vulnerability of status checking in Google Search application, which can be utilized by GVS-Attack to dial arbitrary numbers even when the phone is securely locked with password. In theory, nearly all Android devices equipped with Google Services Framework can be affected by GVS-Attack.

The research of Alepis E. & Patsakis C. [8] examines the dangers lurking in mobiles with intelligent VAs. They deny that it is a fictitious threat but a real scenario that can greatly expose users. In this work, detailed real scenarios of attacks with voice commands were implemented. However due to the use of AI these systems are more difficult to break. An important point of their research is the fact that attacks on mobile devices were not limited. There are a variety of devices that have DAs such as smartwatches, personal computers, or even smartTVs. For this reason, in the present research, in addition to the Android mobile operating systems, attacks were also carried out on Smart Speaker Assistant.

In another research Zhang, R. et al. [9] proposed a stealthy attacking method targeting VAs on smartphones. They proposed an attacking method that could activate the VA and apply further attacks, such as leaking private information, sending forged SMS/Emails, and calling

arbitrary numbers. In order to hide the attack from users an optimal attacking time was chosen. Through their proof of-concept attack targeting Google Assistant on Android platform, they demonstrated the feasibility of the attack in real-world scenarios.

### 3. Data Collection and Manipulation

The aim of this study, as mentioned above, is to attack DAs. For this purpose, it is necessary to collect voice data from different sources. A user's voice was used to extract samples to demonstrate attacks. We are going through the information era where people share everything mainly through the internet. Due to the large amount of free data on the internet, it becomes easier to extract it for experiments. More specifically, four ways of data collection were selected which are analyzed in the following subsections.

#### 3.1. Voice Synthesis

For the voice synthesis, an application that has been developed in the context of a dissertation was used. This work was developed as a master thesis at the University of Liège [11]. For the development of this work, they relied on the 2018 publication of Jia Y. et al., concerning "Transfer learning from speaker verification to multispeaker text-to-speech synthesis" [12].

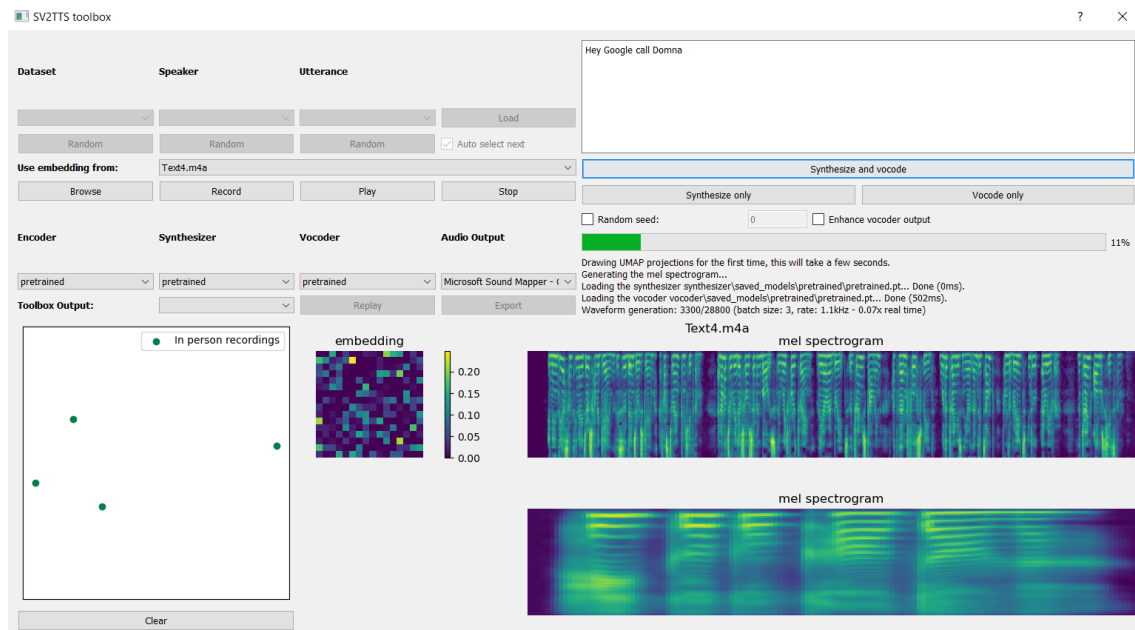


Figure 3 SV2TTS toolbox launch

With the use of "Real Time Voice Cloning" it is possible to compose voice in real time. The project is open source and therefore, anyone can access it. As input for training the system for its training needs audio samples of more than 5 seconds. The framework can capture in a digital format a meaningful representation of the voice spoken and then analyzed, resulting in the

utterance of a phrase or a text that is completed at that time. An important operation is to download the files so that they can be used later. The system neither requires the same voice to be re-edited nor restricts the user regarding the generated text. According to the authors, the application consists of three parts: a speaker encoder, a synthesizer and a vocoder. Because the application uses a neural network if sufficient data is given it can produce the voice of the "target" with the help of text-to-speech (TTS). This application has been developed in python and has as prerequisites the installation of Python 3.6 or 3.7, PyTorch (> = 1.0.1), ffmpeg and the download of pretrained models. "Real Time Voice Cloning" was the first public implementation of Jia Y. et al. research.

Figure 3 shows the toolbox for voice composition. The operator is asked to load the number of files he has and then pressing the "Synthesize and vocode" button starts the process of audio processing and voice production. The files must be at least 4 to start the process. At the top of the right side, there is a box available in order to write the phrase that user want to generate.

### 3.2. Face to Face

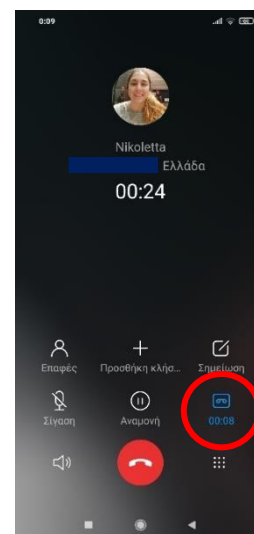
The first way of collecting data required personal contact with the "victim". The terms "victim" and "target" from now on represent the user who is going to be attacked. All that was needed was a recording device that was activated during the conversation with the "target". An important problem encountered in all cases is the maintenance of favorable conditions in terms of sound, as a lot of noise can distort the data. The attacker was called upon to obtain sufficient information from the user. This applies to both the duration of the conversation and the fluctuations of the tone of voice. In addition, care was taken to operate the recording device so that it would not be perceived by the "victim". It is the riskiest way of collecting data but at the same time the most direct. No data recording skills are required with this process nor any cost. Anyone can do it.

For the composition of the voice, 10 audio files were used that were recorded using a mobile phone. The created file was saved for later use. Initially, 4 files were imported, which was the minimum amount required. After the 10 recorded files were added, the system was re-trained. Adding more audio files led to better results.

### 3.3. Via Call

An alternative to the data collection used was call recording. In this case it was a less risky way as it did not bring the attacker into personal contact with the "victim" and therefore could not be perceived by the latter as being recorded. All that was needed was to make a call to the "target" under any pretext. They either need to know each other, or the attacker should try to offer a service to the "victim" that will attract his interest to keep him longer in the handset. In both cases there is a need-to-know information about him. In this process, the network quality and the location of the "victim" played an important role as these two factors can play a decisive role in the quality of the recording and therefore in the final result. The way that the data were extracted is presented in Figure 4.

To produce the synthesized voice file in this case, six recordings of calls with the same person were recorded. Each recording with the interlocutor lasted about one minute. During the call, in general, an effort was made to create ideal conditions. There was no noise in the background environment, such as the sound of cars from the street, and the final quality was quite satisfactory. Although it was initially considered one of the recording methods that would not lead to good results the final generated audio file was quite similar to the target's voice.



**Figure 4 Data collection via call**

### 3.4. Spy Applications

As is known from Google, there are over 2.5 billion active Android devices that have a DA pre-installed. In addition, more than 200 million smart speakers have been sold. Which leads to the conclusion that a very large percentage of people have at least one mobile device or a smart speaker. According to statistics, there are about 3.04 million applications for Android devices. Many surveys have been based on the major security vulnerabilities that appear in Android applications. Based on this conclusion, a master thesis was prepared aiming the understanding of the risks involved in the use of mobile applications as well as the vigilance of users who easily provide access to mobile capabilities, such as the use of microphone and camera [10]. This study presented the implementation of an Android application that made use of the device's microphone by recording the user without his knowledge. Emphasis was placed on the ease of creating this application, concluding that anyone with even a basic programming knowledge can steal data from other users. Therefore, another way of data collection that can be used, is applications such as the one presented in the research above. For the elaboration of the present work, the application of the forementioned master's thesis was utilized, which represents any application that has access to the microphone.

The files created using the application of this master's thesis are posted in a public database as it was created for research purposes. Having created a user in the application after registration and login, redirection to the main page is followed. There, pressing the button for the first time the recording starts and during the second press it stops. Recordings are uploaded to the database asynchronously. It is important to note that in the case of a longer recording, uploading the file will take some additional time. Because the application uses the Bluetooth name to separate users' files, it was enough to download the corresponding files from the database. 5 files were created this way. The quality of the files did not change during the upload and download. After

given as input to the voice synthesis application the final generated file was very close to the user's voice.

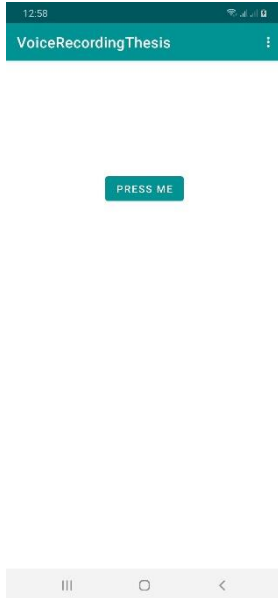


Figure 5 Spy application

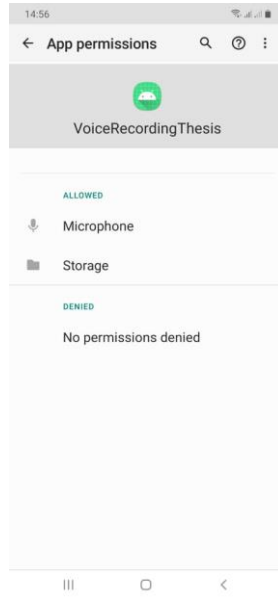


Figure 6 Permissions

### 3.5. Youtube

Technology is increasingly invading people's daily lives. Social media is gaining an irreplaceable role in their lives. A variety of images, videos and sounds are made public and the users freely provide access to them. For the needs of this dissertation, Youtube was chosen as it is one of the most popular means of sharing information. It has been in operation for 16 years and is available in over 100 countries to date. Additionally, 720,000 hours of video are uploaded every day to YouTube. Based on these data, the audio was extracted from videos that had been posted on Youtube and used to deceive the “victim”.

Channel content

Uploads Live

Filter





<input type="checkbox"/>	Video	Visibility	Restrictions	Date ↓	Views	Comments	Likes (vs. dislikes)
<input type="checkbox"/>	 20210327 102848 Add description 1:07	Private	Made for kids	Apr 12, 2021 Uploaded	0	0	-
<input type="checkbox"/>	 20210327 103240 Add description 1:00	Private	Made for kids	Apr 12, 2021 Uploaded	0	0	-
<input type="checkbox"/>	 20210327 103038 Add description 1:14	Private	Made for kids	Apr 12, 2021 Uploaded	0	0	-
<input type="checkbox"/>	 20210327 102510 Add description 0:57	Private	Made for kids	Apr 12, 2021 Uploaded	0	0	-

Figure 7 Youtube upload

Επιθέσεις σε Φωνητικούς Βοηθούς με χρήση συντιθέμενης φωνής



In order to properly test the recordings abstracted from the Youtube videos, posting and downloading the files preceded. After downloading each video with an external tool, the audio of the video was separated, and then given as input to produce the voice. In the four videos posted on YouTube, the "target" was asked to read some texts so that the video would be long. These four recordings that emerged were enough to make a voice composition. In Figure 7 the uploaded videos are shown.

## 4. Attacks

With the progress of technology enormous efforts are being made to meet all security needs. However, due to the increased speed of development, it becomes very difficult to control all the new operations in depth. A wide range of research focuses on finding security vulnerabilities that could be considered harmful to the user. One of them is the present research which focuses on the deception of Google Assistant on smartphones and Google Home Mini using the synthesized voice produced by the user's recordings. Appendix 7.1 lists the most used voice commands for Google Assistant and Google Home.

Intelligent personal assistant	Developer	Free software	Free and open-source hardware	HDMI out	External I/O	IOT	Chromecast integration	Smart phone app	Always on	Unit to unit voice channel	Skill language
Alexa (a.k.a. Echo)	Amazon.com	No	No	No	No	Yes	No	Yes	Yes	?	JavaScript
Alice	Yandex	No	N/A	N/A	N/A	Yes	No	Yes	Yes	N/A	?
AliGenie	Alibaba Group	No	No	N/A	N/A	Yes	No	Yes	Yes	N/A	?
Assistant	Speaktoit	No	N/A	N/A	N/A	No	No	Yes	No	N/A	?
Bixby	Samsung Electronics	No	N/A	N/A	N/A	No	No	Yes	N/A	N/A	?
BlackBerry Assistant	BlackBerry Limited	No	N/A	N/A	N/A	No	No	Yes	No	N/A	?
Braina	Brainasoft	No	N/A	N/A	N/A	No	No	Yes	No	N/A	?
Clova	Naver Corporation	No	N/A	N/A	N/A	Yes	No	Yes	Yes	N/A	?
Cortana	Microsoft	No	N/A	N/A	N/A	Yes	No	Yes	Yes	N/A	?
Duer	Baidu <sup>[44]</sup>										
Evi	Amazon.com True Knowledge	No	N/A	N/A	N/A	No	No	Yes	No	N/A	?
Google Assistant	Google	No	N/A	N/A	N/A	Yes	Yes	Yes	Yes	N/A	C++
Google Now	Google	No	N/A	N/A	N/A	Yes	Yes	Yes	Yes	N/A	?
M (discontinued) <sup>[45]</sup>	Facebook										
Mycroft <sup>[46]</sup>	Mycroft AI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Python
SILVIA	Cognitive Code	No	N/A	N/A	N/A	No	No	Yes	No	N/A	?
Siri	Apple Inc.	No	No	N/A	N/A	Yes	No	Yes	Yes	N/A	?
Viv	Samsung Electronics	No	N/A	N/A	N/A	Yes	No	Yes	No	N/A	?
Xiaowei	Tencent										?
Celia	Huawei	No	No	N/A	N/A	Yes	No	Yes	Yes	N/A	?

Figure 8 Comparison of notable assistants

Google Assistant is an artificial intelligence – powered virtual assistant developed by Google that is primarily available on mobile and “Smart Home” devices. Unlike the company’s previous virtual assistant, Google Now, the Google Assistant can engage in two-way Επιθέσεις σε Φωνητικούς Βοηθούς με χρήση συντιθέμενης φωνής

conversations. Users primarily interact with the Google Assistant through natural voice, though keyboard input is also supported. However, there are dozens of Intelligent Personal Assistants (IPA) in addition to Google Assistant whose features are shown in Figure 8.

Google Nest, formerly known as Google Home, is a series of smart speakers developed by Google under the Google Nest brand. Google Assistant is used for speaker-user interaction with voice commands. Google Nest devices also have integrated support for home automation. Users can control “Smart Home” devices with voice commands. Both in-house and third-party services are integrated, allowing users to listen to music, control playback of videos or photos, or receive news updates entirely by voice. Over time, new features are added through software updates to Google Nest and Google Assistant devices, such as the one that brought multi-user support.

The original Google Home speaker released in November 2016 featured a cylindrical shape with colored status LEDs on top. In October 2017, Google announced two additions to the product lineup, the miniature puck-shaped Google Home Mini and a larger Google Home Max. Google unveiled Google Home Mini (first generation) which is a variant of Google Home with the same overall functionality, but in a smaller shape, with a fabric top whose lights are white when it is on and is waiting for user's commands, while are blue when user is on a call (Figures 9 and 10). It has a mute switch rather than a mute button and uses a micro-USB connection for power.



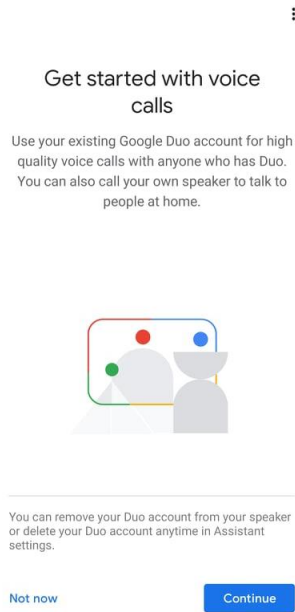
**Figure 9 Google Home Mini listening (white LEDs)**



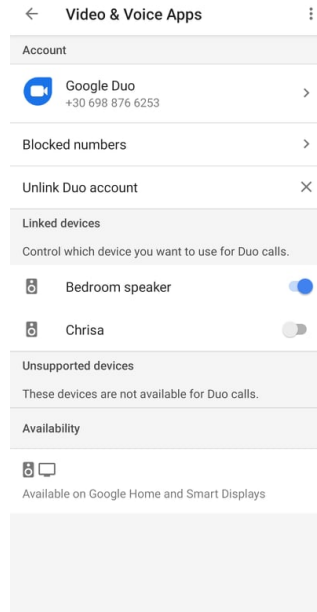
**Figure 10 Google Home Mini on a call (blue LEDs)**

Making video and audio calls through Google Home Mini requires the installation of the Duo application and its connection to the speaker. Anyone can have a list of contacts to connect through a Duo account without any cost. All that user is needed to do is to have Google Account connected to the Duo application on the smartphone (Figure 11). This action is necessary to find user's number. User can unlink Duo account from speaker or delete Duo account anytime in Assistant settings (Figure 12). During the first use, the user is asked to select the reason for using Google Home Mini (calls, music, movies) (Figure 13). In addition, is asked to train the device (Figure 14). Of course, this can be repeated later through the settings. In order to start a call,

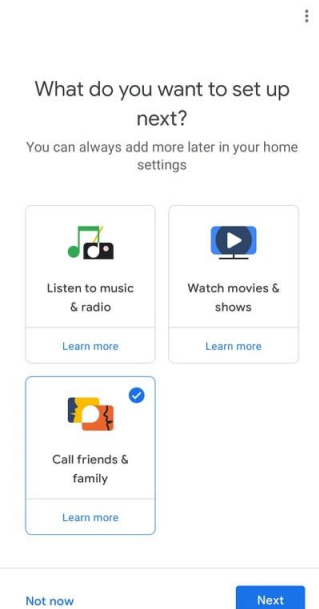
command "Hey Google call [contact]" or "Hey Google video call [contact]" is needed to be uttered from the user and command "Hey Google end the call" for end the call.



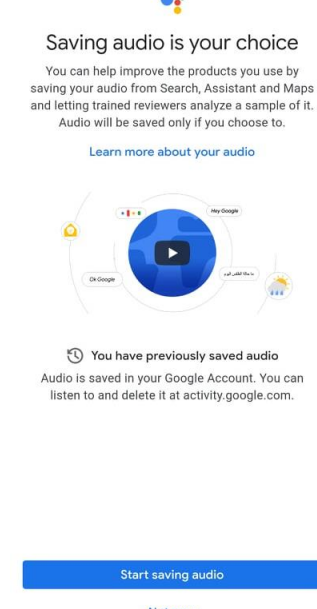
**Figure 11 Google Account connection on Duo application**



**Figure 12 Unlink Duo account**



**Figure 13 Duo set up**

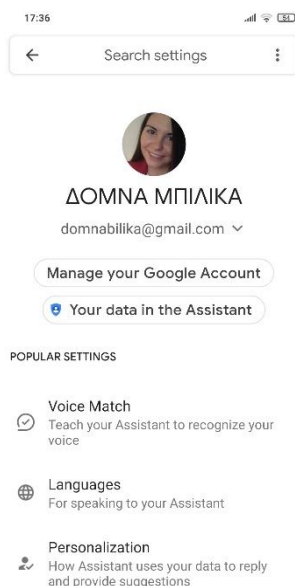


**Figure 14 Google Home Mini training**

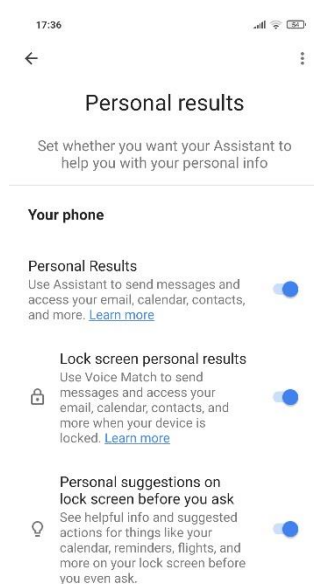
The attacks were carried out in two ways to better present the various ways in which malware can endanger users. An Android application was created for the first type of attack, which is presented below, while the second concerns the direct attack on the “victim’s” device.

In all cases it was taken for granted that the “victim’s” device was unlocked. This assumption was made as surveys show that a large percentage of users do not lock their devices. More specifically, based on research conducted by Kaspersky Lab [13], which is a global cybersecurity company operating in the market for over 20 years, 52% of people do not password-protect their mobile devices, and only 22% of consumers use anti-theft solutions on their phones. Humorously they report that pickpockets set to get lucky as many people leave data on their devices unprotected. Saving a lot of precious data stored on their mobile device does not necessarily make consumers more security conscious. They also noticed that less than half (48%) of those surveyed password-protect their mobile devices, and just 14% encrypt their files and folders to avoid unauthorized access.

Another reason why no further research was done on the above is the fact that through the settings of each Assistant the user can control what kinds of information the Assistant will say or show when the device is locked. Figures 15 and 16 show the relevant Google Assistant settings. Users can turn on lock screen personal results to allow the Google Assistant to send personal communications, call their contacts, and read or show personal results from their: email, including personal results from Gmail, like flight reservations and bills, Google Calendar, contacts, reminders, memory aids and shopping lists. If they turn off lock screen personal results, they will need to unlock their device to find and hear Assistant responses that include personal results. Thus, in order to make it work all that is needed is this feature to be enabled and the screen to be active.



**Figure 15 Google Assistant settings**



**Figure 16 Enable Google Assistant when phone is locked**

## 4.1. Attack via Android App

The Android application “DetectAndAttackThesis” was implemented, which detects nearby Android devices. As mentioned above, the “VoiceRecordingThesis” application created by Michopoulou N. was used for data collection [10]. This application records the users without them being aware and then stores the data in the Firebase under the name of its Bluetooth device. So, in this application the devices were detected based on the name of Bluetooth. Finally, if a device that has used the “VoiceRecordingThesis” application is detected nearby, the playback of the stolen recordings begins.

### 4.1.1. Project Structure and Permissions

In order to create “DetectAndAttackThesis” application, software development platform “Android Studio” was used. Android Studio is the official Integrated Development Environment (IDE) for Google's Android operating system, built on JetBrains' IntelliJ IDEA software and designed specifically for Android development. Also, Java is chosen as programming language. Java is a class-based, object-oriented programming language that is designed to have as few implementation dependencies as possible. It is a general-purpose programming language intended to let application developers write once, run anywhere (WORA), meaning that compiled Java code can run on all platforms that support Java without the need for recompilation.

The Android project that was created consists of folders of which three are the most important and worth presenting. The “manifest” folder contains the file “AndroidManifest.xml” which contains information about the activities and the necessary permissions for the operation of the application. The permissions needed are the device’s Bluetooth and location. As far as Bluetooth is concerned, no further explanation is needed as to why it is used. Instead, the need to use the location needs to be explained. By requiring location services to access Bluetooth, ensure that the user understands their location information may leak when they use Bluetooth. In versions of Android prior to Marshmallow, the user could use Bluetooth without location services enabled, but location information could still leak. The “java” folder contains the main class for running the application called “MainActivity”. While in the “res” folder is the layout folder that contains the code of the application’s screen.

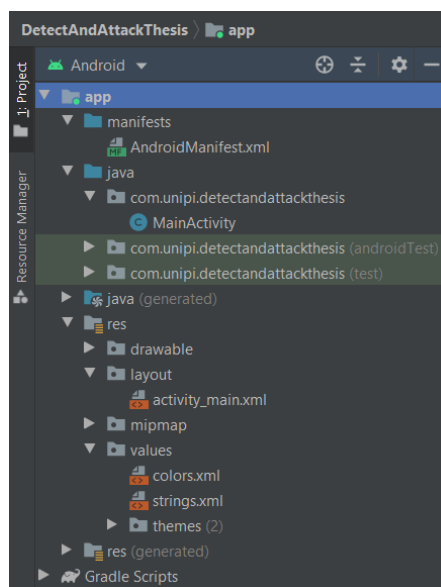


Figure 17 Project Structure

```
<uses-permission android:name="android.permission.BLUETOOTH_ADMIN" />
<uses-permission android:name="android.permission.BLUETOOTH" />
<uses-permission android:name="android.permission.ACCESS_COARSE_LOCATION" />
<uses-permission android:name="android.permission.ACCESS_FINE_LOCATION" />
```

Figure 18 Permissions

#### 4.1.2. Execution Example

This section presents an example of running the application. Initially the necessary permissions are requested from the user as shown in Figures 19 and 20. Figure 16 shows a screenshot of the application settings showing the use of a permission to change the status of Bluetooth. This is needed for refresh of list with devices when Bluetooth is turned on or to stop scanning when Bluetooth is turned off. After the user grants the right to access his location and allows the activation of Bluetooth he is redirected to the main screen of the application. Pressing the corresponding button detects nearby devices that have Bluetooth enabled. It is worth noting that if the user does not grant access to the location, it is not possible to detect the devices. Once it finds available devices it displays them in a list. Then searches in database if exists any recording for these devices. If it finds data in database starts playing a recording.

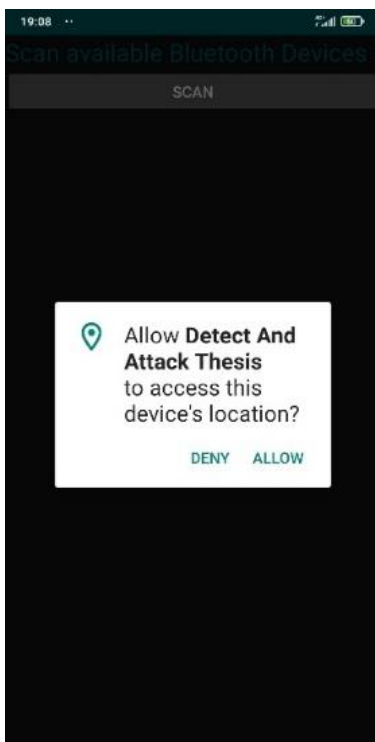


Figure 19 Location permission

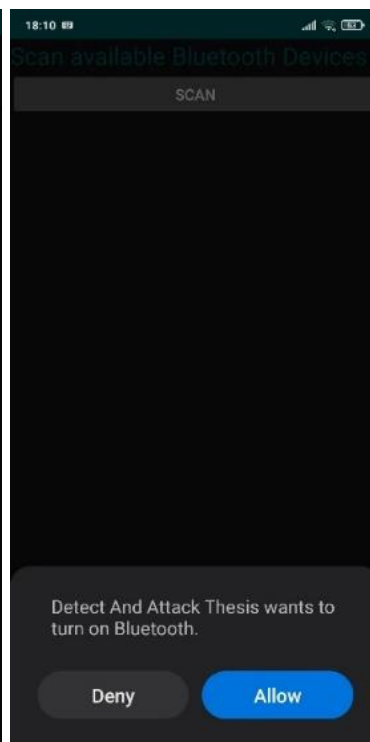


Figure 20 Bluetooth permission

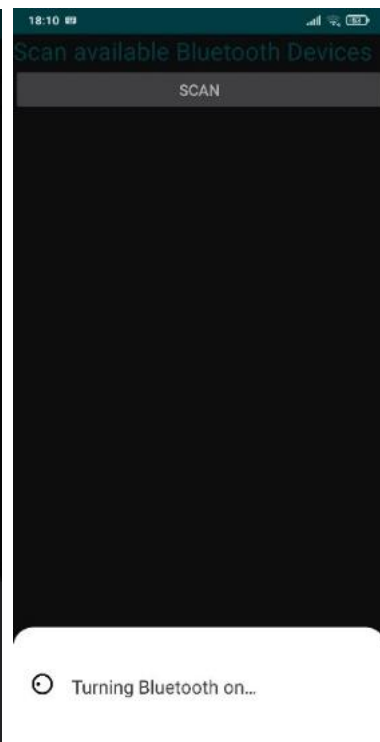


Figure 21 Turning Bluetooth on

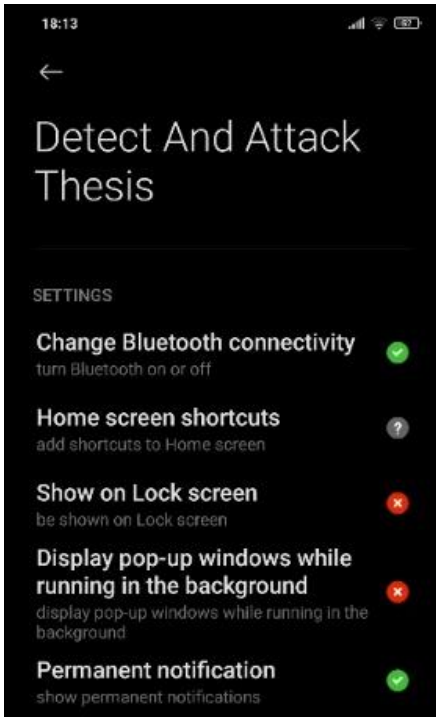


Figure 22 Change Bluetooth connectivity

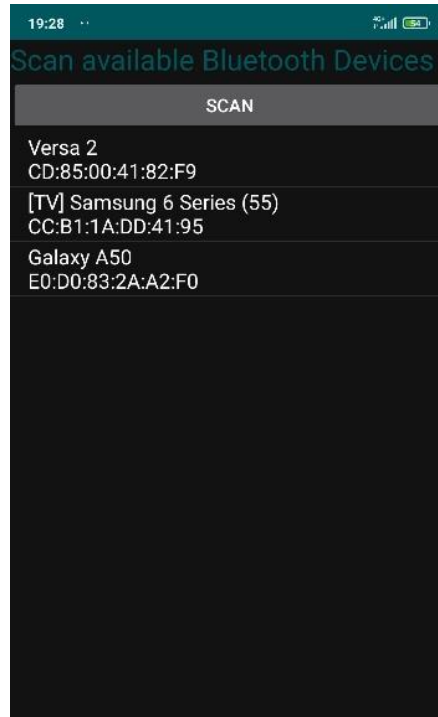


Figure 23 Detected devices

### 4.1.3. Database

Firestore was used as a database to implement the application. Firestore is a platform developed by Google for creating mobile and web applications. It was originally an independent company founded in 2011. In 2014, Google acquired the platform, and it is now their flagship offering for application development. This database is the same one used for the “VoiceRecordingThesis” application of the dissertation mentioned above [10]. The stolen data from the users, which stored in the “Audio” folder were used. Of course, anyone could take advantage of any data stored and even process it as they wish. Figures 24 and 25 are shown the storage.

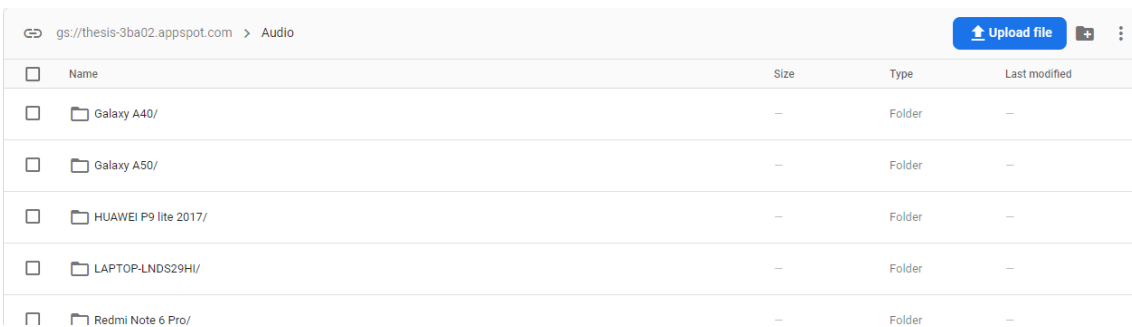


Figure 24 Storage Audio folder

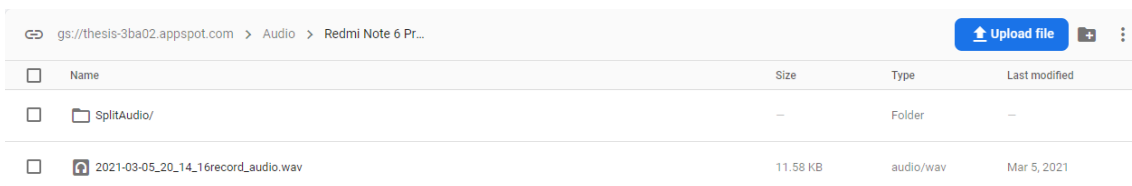


Figure 25 Firebase storage

#### 4.1.4. Code Presentation

The application consists of a class called "MainActivity". Initially a connection is made to the database to extract the necessary data. Next, Bluetooth connectivity is checked. If it is not enabled the user is asked to allow its activation otherwise it will not be possible to find the devices. Location permission is also required because a Bluetooth scan can be used to gather information about the user's location. This information may come from the user's own devices, as well as Bluetooth beacons in use at locations such as shops and transit facilities (Figure 26).

```
// Connection with database
mStorage = FirebaseStorage.getInstance().getReference();

// Turn on Bluetooth
if (myBluetoothAdapter==null)
    Toast.makeText( context: MainActivity.this, text: "Your device does not support Bluetooth", Toast.LENGTH_LONG).show();
else if (!myBluetoothAdapter.isEnabled()) {
    Intent btIntent = new Intent(BluetoothAdapter.ACTION_REQUEST_ENABLE);
    startActivityForResult(btIntent, requestCode: 0);
    Toast.makeText( context: MainActivity.this, text: "Turning on Bluetooth", Toast.LENGTH_LONG).show();
}

ActivityCompat.requestPermissions( activity: this, new String[]{RECORD_AUDIO, WRITE_EXTERNAL_STORAGE, READ_EXTERNAL_STORAGE, ACCESS_COARSE_LOCATION}, PackageManager.PERMISSION_GRANTED);
```

Figure 26 Database connection, Bluetooth and location permissions

Then the main procedure for finding the available devices using the Bluetooth adapter is executed. The BluetoothAdapter is used to perform fundamental Bluetooth tasks, such as initiate device discovery, query a list of bonded (paired) devices, instantiate a BluetoothDevice using a known MAC address, and create a BluetoothServerSocket to listen for connection requests from other devices, and start a scan for Bluetooth Low Energy (BLE) devices. With the registerReceiver command (FoundReceiver, new IntentFilter (BluetoothDevice.ACTION\_FOUND)); BroadcastReceiver starts running in the main activity thread.

```
// Scan for devices
scanb.setOnClickListener(new OnClickListener()
{
    public void onClick(View v)
    {
        btArrayAdapter.clear();
        myBluetoothAdapter.startDiscovery();
        Toast.makeText( context: MainActivity.this, text: "Scanning Devices", Toast.LENGTH_LONG).show();
    }
});
```

Figure 27 Start devices scan

The nearby devices are first searched by the Bluetooth name and then the database is searching if there are any recordings available. If it exists then the first audio file found in "Audio" folder is played, which has been recorded by the respective device.



```

private final BroadcastReceiver FoundReceiver = new BroadcastReceiver() {
    @Override
    public void onReceive(Context context, Intent intent) {
        String action = intent.getAction();
        // Get name of bluetooth device
        if (BluetoothDevice.ACTION_FOUND.equals(action)) {
            BluetoothDevice device = intent.getParcelableExtra(BluetoothDevice.EXTRA_DEVICE);
            if (device.getName() != null) {
                btArrayAdapter.add(device.getName().replace( oldChar: '[', newChar: ' ').replace( oldChar: ']', newChar: ' ') + "\n" + device.getAddress());
                btArrayAdapter.notifyDataSetChanged();

                // If device's name exists play audio
                mStorage.child("Audio").child(device.getName().replace( oldChar: '[', newChar: ' ').replace( oldChar: ']', newChar: ' ')).list( maxResults: 1)
                    .addOnSuccessListener(new OnSuccessListener<ListResult>() {
                        @Override
                        public void onSuccess(ListResult listResult) {
                            for (StorageReference item : listResult.getItems()) {
                                item.getDownloadUrl().addOnSuccessListener(new OnSuccessListener<Uri>() {
                                    @Override
                                    public void onSuccess(Uri uri) {
                                        MediaPlayer mediaPlayer = new MediaPlayer();
                                        try {
                                            mediaPlayer.setDataSource(getApplicationContext(), uri);
                                            mediaPlayer.prepare();
                                            mediaPlayer.start();
                                        } catch (IOException e) {
                                            e.printStackTrace();
                                        }
                                    }
                                }).addOnFailureListener(new OnFailureListener() {
                                    @Override
                                    public void onFailure(@NonNull Exception e) {
                                        Log.e( tag: "Error", Objects.requireNonNull(e.getMessage()));
                                    }
                                });
                            }
                        }
                    }).addOnFailureListener(new OnFailureListener() {
                        @Override
                        public void onFailure(@NonNull Exception e) {
                            Log.e( tag: "Error", Objects.requireNonNull(e.getMessage()));
                        }
                    });
            }
        }
    }
};

```

Figure 28 If Bluetooth device exists audio file is played

#### 4.1.5. Direct Recording Attack

The “victim’s” device was directly attacked using the data generated from the examples presented in the previous section. That means that data was collected via face-to-face communication, phone call, videos on Youtube and an application which records users. From the voice composition four final recordings were produced, one of each source, with the “target’s” voice which were then used for the attack. It is worth noting that the more recordings added to the training the greater the success rate. There must be no noise in the background of the recordings in order to lead to success for sure. Of course, many times a little bit of noise was not an obstacle. One last remark that emerged was about the volume during the pronunciation of the recording that was produced. When the volume was high the DA did not react to the commands, the same happened when it was too low. For greater success the volume had to be moderate and the “victim’s” device had to be at a short distance from the attacker.

After multiple tests, a high success rate was observed. There were 60 attacks of which 49 were successful, at 6 the DA reacted but did not understand the command that was addressed, while in the remaining 5 there was no response. Both basic and more personal voice commands were tested, such as “Hey Google, call [contact name]” and “Hey Google sent email”. Below are some sample snapshots from the Google Assistant attack.

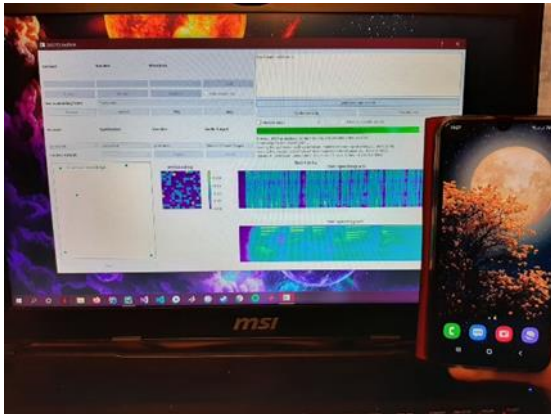


Figure 29 Voice Synthesis training

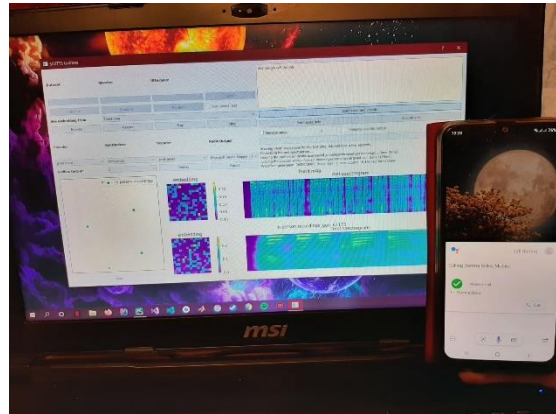


Figure 30 Voice attack

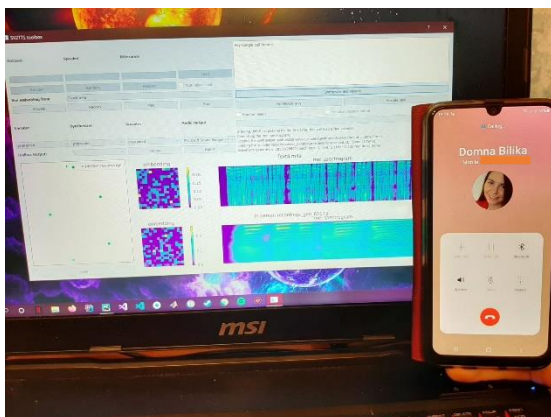


Figure 31 Google Assistant response

In all cases, for the most part, positive results emerged, which raised many questions. If a DA can be tricked so easily, that is, with applications that are either available for free or easy to implement, how ease could it be to trick a security system where all the devices are connected to the same network? For example, a Smart Home consists entirely of such devices. What if anyone could open a person's home at any time with just one voice command? Are we heading into a modern age faster than we should, leaving huge security gaps behind? Over the years more and more devices become part of Smart Home offering more convenience to users. As it has already been proven by research the number of VAs used to access Smart Home devices will reach 555 million by 2024, up from 105 million in 2019. By 2024, Juniper Research expects more than 90 per cent of VAs to be used to control smart home devices.

The latest estimates from Juniper Research circle the 2.5 billion magic number at the end of 2018, which is forecasted to reach as many as 8 billion VAs in use by 2023 [14]. The largest increase in the market based on the same research is observed in Smart TVs, Smart speakers, and Wearables. From the above it follows that over time the needs of users will increase leading companies to create new features to cover them in much less time, which will add even greater security problems.

## **5. Conclusion and Future Work**

In addition to unintentional actions that might be performed by manipulating the DAs or voice samples that could either accidentally be leaked or even be freely provided, such as in the case of Youtube, another security and privacy risk associated with IVAs is the creation of malicious voice commands. Since IVAs work off of voice commands, an attacker could impersonate the user and issue malicious voice commands: they could start a car, unlock a smart door to gain unauthorized entry to a home or garage or order items online without the user's knowledge. Although some DAs provide a voice-training feature to prevent such impersonation, it can be difficult for the system to distinguish between similar voices. Thus, someone with malicious intent, who is able to access a DA - enabled device might be able to fool the system into thinking that they are the real owner and carry out criminal or mischievous acts. As a future research someone could carry out the same attacks on different DAs and reveal their security vulnerabilities.

## 6. References

- [1] Robles, R. J., & Kim, T. H. (2010). Applications, systems and methods in smart home technology: A. *Int. Journal of Advanced Science And Technology*, 15, 37-48.
- [2] Robles, R. J., Kim, T. H., Cook, D., & Das, S. (2010). A review on security in smart home development. *International Journal of Advanced Science and Technology*, 15.
- [3] Edu, J. S., Such, J. M., & Suarez-Tangil, G. (2019). Smart home personal assistants: a security and privacy review. *arXiv preprint arXiv:1903.05593*.
- [4] Gupta, M. P. (2018). Google Assistant Controlled Home Automation. *International Research Journal of Engineering and Technology*, 5(5).
- [5] Isyanto, H., Arifin, A. S., & Suryanegara, M. (2020, February). Design and Implementation of IoT-Based Smart Home Voice Commands for disabled people using Google Assistant. In *2020 International Conference on Smart Technology and Applications (ICoSTA)* (pp. 1-6). IEEE.
- [6] Abdi, N., Ramokapane, K. M., & Such, J. M. (2019). More than smart speakers: security and privacy perceptions of smart home personal assistants. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS) 2019*.
- [7] Diao, W., Liu, X., Zhou, Z., & Zhang, K. (2014, November). Your voice assistant is mine: How to abuse speakers to steal information and control your phone. In *Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices* (pp. 63-74).
- [8] Alepis, E., & Patsakis, C. (2017). Monkey says, monkey does: security and privacy on voice assistants. *IEEE Access*, 5, 17841-17851.
- [9] Zhang, R., Chen, X., Lu, J., Wen, S., Nepal, S., & Xiang, Y. (2018). Using AI to hack IA: A new stealthy spyware against voice assistance functions in smart phones. *arXiv preprint arXiv:1805.06187*.
- [10] Michopoulou, N. (2021). Master thesis: Data interception from Android devices and voice synthesis.
- [11] Jemine, C. (2019). Master thesis: Real-Time Voice Cloning.
- [12] Jia, Y., Zhang, Y., Weiss, R. J., Wang, Q., Shen, J., Ren, F., ... & Wu, Y. (2018). Transfer learning from speaker verification to multispeaker text-to-speech synthesis. *arXiv preprint arXiv:1806.04558*.
- [13] Kaspersky Lab. (2018). Kaspersky Lab Finds Over Half of Consumers Don't Password-Protect their Mobile Devices.
- [14] Juniper Research. (2019). Digital Voice Assistants in Use to Triple to 8 Billion by 2023, Driven by Smart Home Devices.

## **7. Appendix**

### **7.1. Examples of Google Assistant Commands**

#### **7.1.1. Google Assistant Commands for Beginners**

Here are six of the most important Home commands to get started with, which you can choose to start with either "Hey Google" or "OK Google".

"Hey Google, stop."

"Hey Google, play [song title] by [artist] from Spotify."

"Hey Google, help."

"Hey Google, turn [up/down] the sound."

"Hey Google, what's the time?"

"Hey Google, add apples to my shopping list."

#### **7.1.2. Controlling Music, Podcasts, Radio and Audiobooks with a Google Smart Speaker**

"Hey Google, play some hip hop."

"Hey Google, play some Oasis."

"Hey Google, play Essex FM."

"Hey Google, play some music from Spotify."

"Hey Google, skip this track."

"Hey Google, play my Evening Jazz playlist."

"Hey Google, stop the music in 15 minutes."

"Hey Google, what's this song?"

"Hey Google, play some ambient noise"

"Hey Google, help me relax."

"Hey Google, pause the music."

"Hey Google, stop the music."

"Hey Google, skip to the next chapter."

"Hey Google, read Women and Power."

"Hey Google, how much time is left on this chapter?"

"Hey Google, move my music to the bedroom"

"Hey Google, play music on all the downstairs speakers."

### **7.1.3. Setting Timers and Alarms**

"Hey Google, set a timer for 10 minutes."  
"Hey Google, set an alarm for 7.15am tomorrow morning."  
"Hey Google, snooze the alarm."  
"Hey Google, cancel the alarm for 7.15."  
"Hey Google, do I have any alarms set?"  
"Hey Google, set a second timer for 30 minutes."

### **7.1.4. Checking Calendars and Reminders**

"Hey Google, what does my day look like?"  
"Hey Google, do I have anything scheduled for Friday?"  
"Hey Google, where is my first event tomorrow?"  
"Hey Google, remind me at 8am tomorrow to take my lunch in to work."  
"Hey Google, what are my reminders for this week?"  
"Hey Google, delete my reminder to pick up my dry cleaning."  
"Hey Google, remember that I left the spare keys in the kitchen cupboard."  
"Hey Google, where did I put the spare keys?"

### **7.1.5. General Queries and Commands**

"Hey Google, what's the weather like today?"  
"Hey Google, when is it going to rain?"  
"Hey Google, is it windy this morning?"  
"Hey Google, what's the traffic like on the way to work?"  
"Hey Google, how long will it take to get home?"  
"Hey Google, what's today's headlines?"  
"Hey Google, did the 49ers win yesterday?"  
"Hey Google, what's the French word for grapefruit?"  
"Hey Google, how do you make mushroom risotto?"  
"Hey Google, what's 100 dollars in British Pounds?"  
"Hey Google, spell manoeuvre."  
"Hey Google, when is Shake Shack open until?"  
"Hey Google, is Target open today?"  
"Hey Google, are there any Italian restaurants around here?"  
"Hey Google, how many calories are in a courgette?"  
"Hey Google, flip a coin."

Επιθέσεις σε Φωνητικούς Βοηθούς με χρήση συντιθέμενης φωνής

"Hey Google, how do you solve an algebraic equation?"

### **7.1.6. Commands for Google Assistant Smart Home Control**

"Hey Google, turn on the Bedroom Light."

"Hey Google, set the kitchen lights to 50%."

"Hey Google, dim the living room lights."

"Hey Google, brighten the office lamp."

"Hey Google, turn the living room lights to blue"

"Hey Google, play Stranger Things on Netflix."

"Hey Google, play The Crown on TV."

"Hey Google, set the thermostat to 20 degrees."

"Hey Google, raise the temperature three degrees."

"Hey Google, what's the temperature right now?"

"Hey Google, turn on all the switches."

"Hey Google, lock the front door."

"Hey Google, open the blinds."

"Hey Google, show me the backdoor camera."

"Hey Google, answer the door."

### **7.1.7. Phone and Calls Commands**

Google Assistant is able to integrate with Android messages, WhatsApp, Messenger and more, letting get read-outs of messages. However, someone can also handle calls with the commands below.

"Hey Google, call Elizabeth."

"Hey Google, call the nearest coffee shop."

"Hey Google, hang up."

"Hey Google, redial."

"Hey Google, Bluetooth pairing."

"Hey Google, is my phone connected over Bluetooth?"

"Hey Google, call my phone."

### **7.1.8. Commands for kids**

"Hey Google, what does an elephant sound like?"

"Hey Google, set a Teenage Mutant Ninja Turtle alarm for 8am."

"Hey Google, read a bedtime story."

"Hey Google, sing a lullaby."

"Hey Google, start the school day"

### **7.1.9. Broadcasting Commands**

"Hey Google, broadcast 'wake everybody up'."

"Hey Google, broadcast 'dinner is ready'."

"Hey Google, broadcast 'time to leave for school'."

"Hey Google, broadcast 'the football is about to start.'"