EXPOS

PLATFORM FOR INVESTIGATING PRIVACY CONCERNS IN INTELLIGENT VOICE ASSISTANTS

RAHUL LAO

Master Thesis

July 2018

Rheinische Friedrich-Wilhelms-Universität Bonn
Institute of Computer Science 4
Methods in Multi-Layer Usable Security Research

Platform for Investigating Privacy Concerns in Intelligent Voice Assistants
Master Thesis

Submitted by Rahul Lao
Date of submission: July 2018

First examiner / Erstgutachter: Dr. Emanuel von Zezschwitz
Second examiner / Zweitgutachter: Prof. Dr. Matthew Smith
Supervisor / Betreuer: Christian Tiefenau

Rheinische Friedrich-Wilhelms-Universität Bonn
Institute of Computer Science 4
Methods in Multi-Layer Usable Security Research

MOTIVATION

As Artificial Intelligence (AI) is growing there has been recent breakthroughs in Natural Language Understanding (NLU) systems like Echo by Amazon, Home by Google and Siri by Apple that are also known as Intelligent Voice Assistants (IVAs).
An intelligent voice assistant opens a new world, a world where you communicate with a machine as if it were a human and the machine will perform the work you requested [1]. For example, you may want to know the weather "Hey Alexa, what's the weather today?". These devices can not only respond to voice commands but also play music, search on internet, order items, interact with other devices connected to internet and much more [2].
For these devices to work they are usually powered by a brain (AI) that may reside on a virtual place e.g. a cloud [3]. Thus, all user queries have to be transferred over the network to the cloud where the actual processing and query resolution takes place. This raises many user privacy concerns like how the data is stored, where the data is stored, who all have access to user data? Simply watching television or listening to radio can also trigger interaction with smart speakers [4]. As these smart devices are always listening this could be big threat. User voice data has the potential to carry bio metric data and with the emergence of voice biometrics over the internet [5] a rogue IVA can be used to capture user voice commands and used for authentication.

RELATED WORK

(H Chung et al. 2017) published a paper in which they explain the ecosystem of intelligent voice assistants. Along with that they talk about identifying threat vectors of IVA and consequences of having a rogue IVA.
By utilizing the reviews posted online and responses to a survey for IVAs (Lydia Manikonda et al. 2017) provides a set of insights about the detected markers related to user privacy interests and privacy challenges.
In an ISTR special report by (Candid Wueest, 2017), risks IVAs possess to one's cyber security have been pointed out. It is even possible to control a voice assistant from outside the home by using ultrasound frequencies to avoid human detection.
Drawing from user reviews of the Echo posted to Amazon.com, (A Purington at al. 2017) explores the degree to which user reviews indicate personification of the device, sociability level of interaction, factors linked with personification, and influences on user satisfaction.
As intelligent voice assistants (IVA) operate based on the cloud computing architecture (Hyunji Chung and Sangjin Lee, 2018) shows and categorize types of IVA-related data that can be collected from popular IVA, Amazon Alexa. Forming an experimental dataset covering three months with Alexa service, they then analyze to characterize the properties of user's lifestyle and life patterns. The results presented provide important implications for and privacy threats to IVA vendors and users as well.

GOAL

Our goal is to develop a platform which when enabled doesn't allows smart voice assistants to listen and record any user commands, instead captures user voice commands and stores them on device for later analysis.

FIRST OUTLINE

Figure 1 shows three components of the entire platform. Raspberry Pi is connected with a device that runs Amazon Alexa (say Amazon Echo) to form the first component. We can call this component to be our private on device voice assistant (PODVA). As we know that Echo is always listening and transfers user voice data over the internet to the cloud upon hearing the wake word, we connect raspberry pi to stop Echo from listening and recording even when connected and on hearing wake word.
A web application hosted in Raspberry Pi will be the second component. From the web application the user can manage its interaction with PODVA and hence we can define this application as its management portal.
Finally, an app running on smart watch to give feedback upon intent classification of user query will form the third component. From Figure 1 we can see that user will interact will all three components.

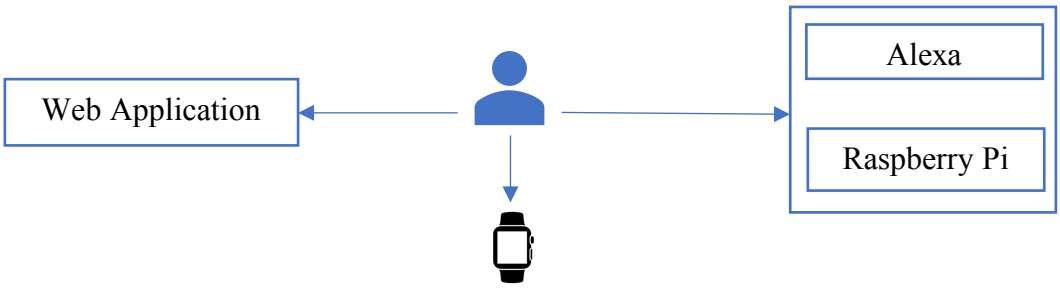

Figure 1

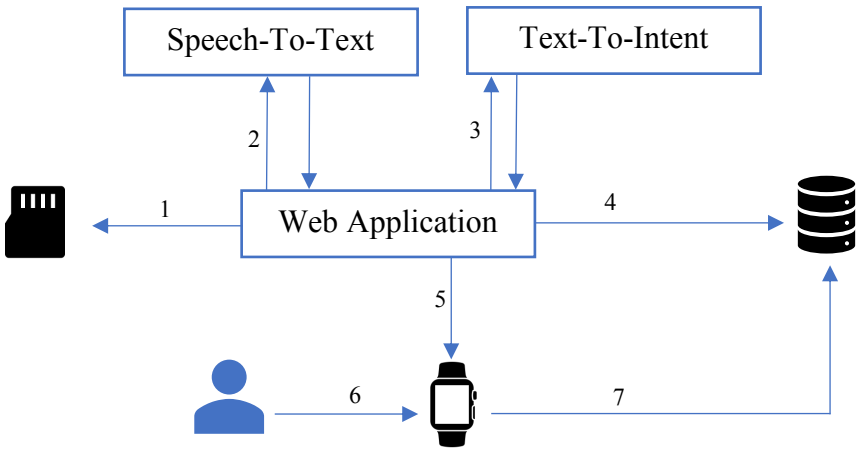

Figure 2

Once PODVA is setup and connected, user can access its management portal. From the management portal user can initiate a thread which will begin listening to user query. On completion of the query, user can stop the thread which then saves voice command on memory card attached with raspberry pi. Figure 2 shows the overall flow once a user has interacted with PODVA and voice data has been saved in memory card.

The user query saved as voice data file will be sent to an independent Speech-To-Text service running in raspberry pi to get the text from voice data. The obtained text query will then be sent to another independent Text-To-Intent service running in raspberry pi to obtain the intent of user query. Once the intent is obtained, it along with user query in text, wake word and timestamp will be stored in a database. A notification will be sent to an activated smart watch associated with user profile. User can then specify from an app running on smart watch if he or she was satisfied with the intent classification of the user query, which will then be saved in database.

OUTCOME

By having a private on device voice assistant (PODVA), no user data will be transferred over the network to the cloud and hence the final outcome of this thesis would be a platform to monitor privacy concerns related with intelligent voice assistants.

PLANNED TIMELINE / MILESTONES

Title	Start	End
Literature review	20.06.2018	29.06.2018
Proposal Writing	02.07.2018	08.07.2018
Design	09.07.2018	15.07.2018
Implementation 1.0	16.07.2018	31.07.2018
Feedback and changes	01.08.2018	12.08.2018
Implementation 2.0	13.08.2018	31.08.2018
Feedback and changes	01.09.2018	16.09.2018
Writing First Draft	24.09.2018	31.10.2018
Feedback and changes	01.11.2018	16.11.2018
Writing Final Draft	19.11.2018	15.12.2018
Presentation	16.12.2018	31.12.2018

SOURCES

[1], [3] H Chung, M Iorga, J Voas, S Lee: „Alexa, Can I Trust You? “. (2017)

[2] Virtual Assistant – Wikipedia, https://en.wikipedia.org/wiki/Virtual_assistant

[4] Candid Wueest: „ A guide to the security of voice-activated smart speakers “(2017)

[5] Laurent Besacier, Aladdin M. Ariyaeeinia, John S. Mason, Jean-Francois Bonastre, Pedro Mayorga, Corinne Fredouille, Sylvain Meignier, Johann Siau, Nicholas W. D. Evans, Roland Auckenthaler, Robert Stapert: „ Voice Biometrics over the Internet in the Framework of COST Action 275 “(2003)

[6] Lydia Manikonda, Aditya Deotale, Subbarao Kambhampati: „What's up with Privacy? User Preferences and Privacy Concerns in Intelligent Personal Assistants “(2017)

[7] H Chung, S Lee: „Intelligent Virtual Assistant knows Your Life “. (2018)

[8] A Purington, J G. Taft, S Sannon, N N. Bazarova, S H Taylor: „Alexa is my new BFF: Social Roles, User Satisfaction, and Personification of the Amazon Echo “. (2017)