Author Manuscript
Accepted for publication in a peer-reviewed journal

NIST National Institute of Standards and Technology • U.S. Department of Commerce

Published in final edited form as:
Computer (Long Beach Calif). 2017 September ; 50(9): 100–104. doi:10.1109/MC.2017.3571053.

Alexa, Can I Trust You?

Hyunji Chung,
National Institute of Standards and Technology

Michaela Iorga,
National Institute of Standards and Technology

Jeffrey Voas, and
National Institute of Standards and Technology

Sangjin Lee
Korea University

Abstract

Security diagnostics expose vulnerabilities and privacy threats that exist in commercial Intelligent Virtual Assistants (IVA) – diagnostics offer the possibility of securer IVA ecosystems.

Intelligent Virtual Assistants (IVAs) open a new world, a world where you can talk to a machine as if it were a human and the machine will perform the work you request. For example, when you wake up, “Hey, what’s on my schedule for today?” Before you leave the house, “Hey, what’s my commute?” For dinner, “Hey, order one large size pepperoni pizza.” When you go to sleep, “Hey, turn off the bed room lights.” Ideally, such conversations should be solely between you and the device assisting you. But are they? Do you know? Where is the trust?

IVAs may be new and mysterious to some consumers, but they are in the market place today. Gartner has said that the IVA market will reach $2.1 billion by 2020 [1]. Voice assistants such as Google Home, Apple’s Siri and Amazon’s Echo devices have always been susceptible to accidental hijack. A Google ad during the Super Bowl that used the phrase “OK, Google” reportedly set off people’s home devices that began reciting the definition of a Whopper, pulled from the website Wikipedia [2]. Since the website can be edited by users, the definition had been changed and “cyanide” was inserted as an ingredient in one version. Such kind of malicious information, if followed ad litter am, can cause harm.

In this column, we urge readers to think about the potential security and privacy concerns of this technology. For instance, (1) “Is my IVA secure?”, (2) “Is it listening to my conversations?”, (3) “Where is my voice data stored?”, etc. The fact that IVAs are installed in private homes makes this a public-facing challenge, and one that attracts instant media

attention when problems arise. To our knowledge, security and privacy threats of these IVAs have not received enough attention.

WHAT IS AN INTELIGENT VIRTUAL ASSISTANT?

Predecessors of Assisting IoT Devices

IoT devices for assistance are not new. IoT devices for assistance have evolved from half-century old chatbots programmed to pass the Turing test[1] (e.g., Eliza and Parry). A chatbot was a service that people interacted with in writing via a chat interface. They worked by examining a user's typed comments and identifying known keywords. If a keyword was found, a rule that transformed the user's comments was applied, and the resulting sentence was returned [3].

Today's newer versions of IVAs can not only respond to voice commands, but also can play music if asked, perform keyword searches, order items, turn on lights, open garage doors, and can even sustain conversations [4].

Defining IVA

There are various terms for this category of IoT devices. They include, but are not limited to, Smart Assistant, Intelligent Personal Assistant, Digital Assistant, Personal Virtual Assistant, Virtual Assistant Bot, etc. Among these terms we can recognize some common keywords: 'smart', 'assistant', 'intelligent', and 'virtual'. In this column, we employ the term *Intelligent Virtual Assistant (IVA)*, because, even though the communication is facilitated by devices in the proximity of the user, an assistant of this type is powered by *artificial intelligence,* and the "brain" of the assistant is in a *virtual* place, e.g., a cloud. These devices are communicating with the virtual assistant, sometimes by default, but more often only when configured to do so, and have no embedded intelligence. We will employ, in this column, the term IVA-enabled device when referring to such devices.

Well known IVAs

Table 1 summarizes known IVAs from major vendors such as Amazon, Apple, Google and Microsoft [5]. (in alphabetical order)

'IVAs' has an agent programs running on 'IVA-enabled devices' (endpoints) such as iPhone, iPad, Mac, Fire tablet, Echo, Google Home, etc. The main functionality, the "brain" of an IVA, is housed as a cloud service that processes voice data (converting voice-to-text, performing linguistic context analysis, and providing answers to questions.)

We divide IVAs into two types: (1) *built-in IVA* that use multi-purpose devices (endpoints) and (2) *stand-alone IVA* that use dedicated devices (endpoints). Examples of the built-in IVA include Siri (for Apple products) and Cortana (for Windows-based PCs). Examples of the stand-alone IVA include Alexa (that uses Echo, Echo Dot and Tab dedicated devices) and Google Assistant (that uses Google Home dedicated device.) The remainder of this article

[1]The **Turing test** is a test, developed by Alan Turing in 1950, of a machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human.

focuses on security and privacy *threat modeling* for *stand-alone IVAs* that are operating in peoples' homes.

IDENTIFYING THREAT VECTORS OF IVAs

To identify ways to secure IVAs, we begin by analyzing their security vulnerabilities. Then, since IVAs handle people's actual voice sounds which are Personally Identifiable Information (PII) [6], we extend the analysis to include user privacy. IVA vendors are already storing voice data, thus making it possible for unauthorized entities to use the data to identify individuals, to maliciously obtain access to systems that implement voice recognition, or simply to process data and construct voice artifacts that could be used to impersonate these individuals. These scenarios are problematic.

To identify the threat vectors, we have learned how IVAs operate along with their components through a variety of analysis methods such as voice command tests, firmware analysis, network traffic analysis, and application analysis. By doing so, we can unveil useful details about IVA ecosystems.

IVA ECOSYSTEMS

In general, IVAs consist of multiple components in heterogeneous environments. As shown in Figure 1, there are two user-side components: (1) companion applications, and (2) IVA-enabled devices. One of the IVAs we studied was Amazon's Alexa ecosystem. The main components of this ecosystem are grouped into two categories:

- **a.** The client side – that has 2 components:
  - **(a.1)** Endpoint1: an Alexa-enabled device (Echo);
  - **(a.2)** Endpoint2: a companion app that needs to be installed on user's device of choice;
- **b.** The cloud side - the 'intelligent assistant' Alexa that operates in the Amazon's cloud environment.

To test Alexa, we asked Echo questions and got answers. We learned that all the requests sent to Alexa (through Echo) were stored in a cloud in text format and in recorded voice. All the conversations and actual voice recordings were accessible through Alexa's companion app. Performing packet analysis we discovered what kind of data has been stored on the cloud side and how to get access to cloud-native data. In addition, analysis of the firmware and software of IVA-enabled devices helped us understand the overall ecosystem.

To utilize IVAs, the IVA-enabled devices need to run an agent program that communicates with the cloud services. Major vendors are providing this agent by integrating it into their operating systems: for example, the latest versions of iOS and OS X have the Siri agent installed by default. Microsoft Windows 10 has the Cortana agent as one of its default processes. IVA agents from Amazon and Google are similar in principle but use dedicated devices such as Echo and Google Home.

An interesting point here relates to IVA-enabled devices–these ‘Endpoints1’ are stand-alone products designed to only assist in the usage of IVA services. Because these home-embedded devices need to be connected to the Internet to communicate with the ‘intelligent assistant’, the vendors need to provide convenient interfaces for configuring them and managing activity history. Amazon and Google are providing companion applications (apps or web-sites) for completing these activities. It is also important to note that IVAs are expanding their features (often referred to as ‘skills’) by allowing third-party entities to add new compatible services. A few examples are: opening a garage door or unlocking a house door, ordering a pizza, or utilizing a social network service by voice.

ROGUE IVAs

Wiretapping the Internet

An IVA’s ecosystem network communication is divided into two parts: (1) IVA-enabled device-to-cloud or Endpoint1-to-cloud; and (2) companion application-to-cloud or Endpoint2-to-cloud. (Figure 2 - Case1)

On the left side of Figure 2 - Case1, the cloud services may use encrypted connections to protect customer’s personal data. In this environment, sniffing the network traffic between client’s companion application and the cloud may expose user’s security and privacy data. This is because identifying network communications helps attackers understand overall operations of an IVA ecosystem. For example, in laboratory environment, we used HTTPS interception tools, to analyze requests and responses, and then understand which APIs are used for sending and receiving data to and from the IVA running in the cloud.

On the right side of the Figure 2 - Case1, we illustrate IVA-enabled-device-to-cloud or Endpoint1-to-cloud communication. Our analysis reveals that, although most network traffic is encrypted, not everything may be sent over a secure protocol. There may be unencrypted connections, including but not limited to, checking the current network connectivity status, transmitting the firmware image upgrades, etc.

In the first case, it is possible to detect the presence of IVA devices inside of a home network. Also, if firmware data is transferred over unencrypted packets, a man-in-the-middle attack could take place. Even if the image is not altered, obtaining the firmware image is an important security concern because it provides a chance to understand the internal operations of a IVA-enabled device. Furthermore, a malicious attacker may be able to distribute modified firmware images [7]. The rest of the communication between the IVA-enabled device and the IVA running in the cloud is encrypted using HTTPS. So, what about an encrypted packet (HTTPS)? There are various existing studies on classifying network traffic through scientific approaches including machine-learning algorithms [8]. Even though the traffic is encrypted, various patterns including payload sizes and data rates could be utilized for identifying user’s behavior such as turning on the device, the idle status, talking to the assistant, listening to music, ordering products or services, and so on [9].

Compromised IVAs

There are well-known cases of compromised home-embedded devices that were connected to the Internet. Recently, DDoS attacks against Dyn LLC exploited vulnerabilities of 10s of millions of home-embedded devices such as webcams and DVRs, infecting them with Mirai botnet, and turning these devices into an army of bots used to attack Dyn's systems. Because gateway devices used for the IVA ecosystem are also embedded systems, there are similar possibilities for them to be compromised if they contain security vulnerabilities [10].

Figure 2-Case2 illustrates the vulnerability scenario of a 24/7 voice recording. In general, IVAs are not always recording, but always hearing. If an IVA-enabled device hears the 'wakeup word', the user's voice is recorded and transmitted to the IVA in the cloud. If the IVA-enabled device (Endpoint1) is compromised by a malicious attacker, it can play the role of a virtual spy. 'Always-on' voice recording in a user's private location can allow all sounds or voices to be recorded and sent to an attacker in real time. This is a privacy concern.

An IVA-enabled device is a remotely-controlled speaker, similar to a smart baby monitor. There was a recent case where a family living in Washington, spoke out about the horrors they experienced while using a baby monitor inside their 3-year-old son's bedroom. Parents discovered that a stranger had hacked into their baby monitor and was able to spy on their toddler and sometimes speaking disturbing messages into the device [11]. In a similar way, IVAs may be controlled by people pretending to be in the proximity of the IVA-enabled device, while, in fact, they are accessing a speaker positioned in the proximity, in the house.

Malicious voice commands

The third threat includes malicious voice commands as shown in Figure 2- Case3. In voice-activated services, users' voices may lead to dangerous outcomes. Some IVAs provide a voice training feature, but it is difficult to perfectly recognize user's voice, tones and accents. Therefore, an IVA could process requests and answer for someone else or for a malicious person. If a malicious person can come close enough to the targeted IVA-enabled device, he or she may be able to fool the system into thinking that the real owner is the person speaking.

Unintentionally recorded voice

The last scenario deals with data privacy. Voices can be recorded by accident and transmitted to a cloud (Figure 2-Case 4). Because speech recognition is not a perfect science, it is possible to eavesdrop on private conversations unintentionally. The potential for accidental recording means that users do not necessarily have complete control over what audio gets transmitted to the IVA in the cloud [12].

As more 'things' become connected to the Internet, there is a growing need for better understanding of security and privacy threats from IVAs. Our goal here is to provide an overview of IVA ecosystems and their potential threat vectors, and to explain four different cases involving IVAs that turned rogue.

References

1. Newsroom Gartner. Gartner Says Worldwide Spending on VPA-Enabled Wireless Speakers Will Top $2 Billion by 2020. https://www.gartner.com/newsroom/id/3464317
2. AP The Big Story. How Burger King revealed the hackability of voice assistants. http://bigstory.ap.org/2d8036d742504890b2f9edc3f98c77ef
3. Wikipedia. Chatbot. https://en.wikipedia.org/wiki/Chatbot
4. Wikipedia. Virtual assistant (artificial intelligence). https://en.wikipedia.org/wiki/Virtual_assistant_(artificial_intelligence)
5. Business Insider. Why Amazon's Echo is totally dominating — and what Google, Microsoft, and Apple have to do to catch up. http://www.businessinsider.com/amazon-echo-google-home-microsoft-cortana-apple-siri-2017-1
6. McCallister, E., Grance, T., Scarfone, K. NIST Special Publication 800-122. 2010. Guide to protecting the confidentiality of Personally Identifiable Information (PII).
7. Exploring the Amazon Echo Dot, Part 1: Intercepting firmware updates. https://medium.com/@micaksica/exploring-the-amazon-echo-dot-part-1-intercepting-firmware-updates-c7e0f9408b59-.dyktzwphz
8. Nguyen, T., Armitage, G. IEEE Communications Surveys and Tutorials. 2007. A survey of techniques for Internet traffic classification using machine learning.
9. Gu C, Zhang S, Sun Y. Real-time encrypted traffic identification using machine learning. Journal of software. 2011
10. ORACLE+Dyn. Dyn Statement on 10/21/2016 DDoS Attack. http://dyn.com/blog/dyn-statement-on-10212016-ddos-attack/
11. The San Francisco Globe. Stranger hacks family's baby monitor and talks to child at night. http://sfglobe.com/2016/01/06/stranger-hacks-familys-baby-monitor-and-talks-to-child-at-night/
12. The Christian Science Monitor. What do Alexa and Siri mean for privacy?. http://www.csmonitor.com/Technology/2017/0114/Devices-sprout-ears-What-do-Alexa-and-Siri-mean-for-privacy

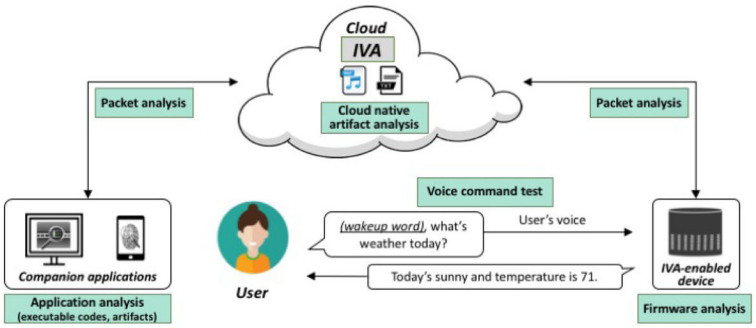

Figure 1.
IVA ecosystem

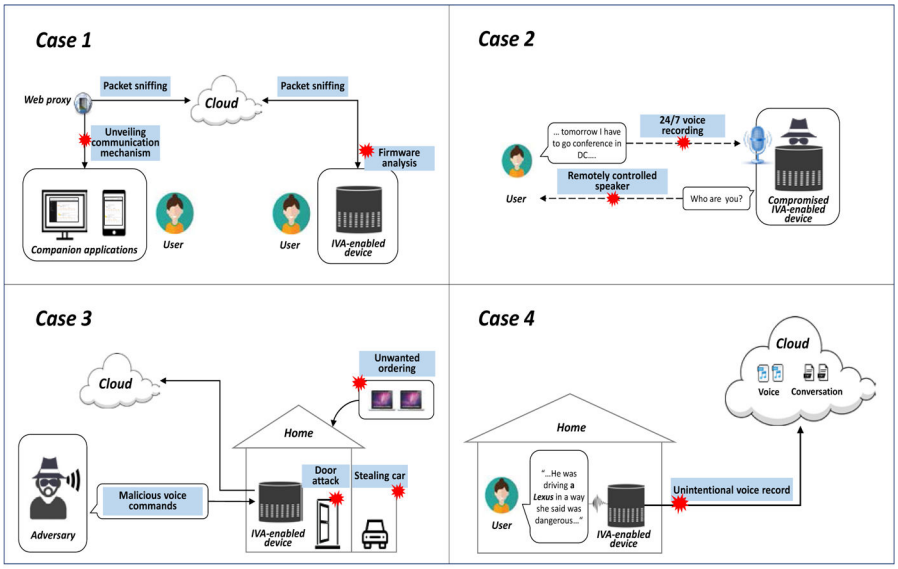

Figure 2.
Four cases when the IVA turns rogue

Table 1

Summary of best known IVAs

Vendor	IVA	IVA-enabled devices (Endpoints)
Amazon	Alexa	Echo, Dot, Tab, Fire Tablet
Apple	Siri	iPhone, iPad, Mac
Google	Google Now & Google	Any phone with Android, Google Home
Microsoft	Assistant Cortana	Any PC with Windows