Matt Sullivan
CSCI 5743
Survey of Security Attacks

Survey of Security Attacks

Matt Sullivan

Preface: For this assignment, you said we could either concentrate on a novel type of attack or an actual attack that took place. I chose to do one of each, with the cyber kill-chain concentrating on the novel rather than the historical attack.

1. Audio Attacks on IoT Based Voice Activation Systems.

1(A) Introduction:

In the quickly growing world of Internet of Things, the primary interface for many devices are voice commands. While there have been no widespread attacks recorded so far, it has been demonstrated that many of these interfaces are vulnerable to attacks simply by being in the same room as a television. Researchers are also discovering that it is possible to manipulate these interfaces at frequencies outside the range of human hearing, or in masking commands as unintelligible sounds. In light of this, it would be foolish to assume that these interfaces are safe and need not be secured, as it is likely that it is only a matter of time before there is an attack targeting these devices.

1(B) Description:

In September of 2017, many viewers of the show Southpark were distressed and annoyed to find that the episode of the television series issued a series of commands to their voice-operated devices, specifically Amazon Alexa and Google Home devices. The show added items to their shopping lists, made the devices repeat words, and set alarms for the next morning.[1] While in this instance the unauthorized access was harmless fun, it highlighted the fact that, while sophisticated, the devices were not able to distinguish live human speech from recorded human speech, and they were absolutely not secure in light of this. Accordingly, in the past few years, much research has been dedicated to understanding the vulnerability of these systems.

There are several methods of attack have been identified. The most basic is the simple “Replay” attack, where recorded phrases are used to manipulate the interface. This is the simplest method, and is easily detectible by a human user, but the devices may be vulnerable, particularly if there are no humans present. But this is just the tip of the iceberg. Researchers have been demonstrating that attacks are possible at the Operating System and Hardware level.[2] Operating system attacks involve the installation of malware on the device, allowing for gathering of recorded commands, monitoring if there are people that could possibly recognize an attack are present. Researchers are investigating multiple methods of masking commands using machine learning[3], injecting commands masked by other commands so as not to be detected[4], and even using frequencies outside the human hearing spectrum to initiate commands silently.[5] Additionally, it has been shown[6] that in some of these devices, the recording and execution are not mutually exclusive, creating the potential for a hacked device to project commands to itself. Each of these is a potential security vulnerability, but combined together they could create a significant attack vector.

1(C) Impact:

There have been hundreds of millions of Alexa and Google Home style interfaces sold since they've been introduced, and there are billions of IoT connected devices on the planet today. Many users (especially regular users) of the voice interfaces have them connected to their Google or Amazon accounts, with access to their contacts, financial data, internet history, and incredible amounts of private information. If these voice interfaces were used to exploit this data, it could be an incredibly large breach.

Additionally, many of these interfaces are linked to physical devices also connected to IoT. These interfaces could be used to physically open doors or control appliances. This presents a potential threat to physical security as well as information security.

1(D) Possible Detection or Prevention Mechanisms:

For the time being, the most obvious detection of these attacks is for the owner to monitor the voice operated interface for unusual or unprompted responses. All of the voice-operated interfaces have acknowledge responses to commands.

On the developer side, continuing to refine the speech interface to correctly identify commands and not be tricked by false input would be the most important deterrent. One of the studies (which itself used machine learning to create its attacks), suggested that machine learning was a possible avenue of approach to defeat its own attacks.[3]

Since all of the proposed attacks are using the audio sensor to direct without a human present, a solution would be to try to affirm that a human being was present and giving directions. However, most proposed solutions impinge on the utility of the voice-operated interface as a convenient interface[2]. Other safeguards would be password protecting/locking the interfaces when not in use, but again, as in all cybersecurity matters, security must be balanced against utility as an interface.

1(E) References:

[1] Consumerist, Consumer Reports 9/14/17 https://www.consumerreports.org/consumerist/south-park-screws-with-viewers-google-home-echo-devices/ (Retrieved 2/17/19)

[2]Y. Gong, C. Poellabauer, "An Overview of Vulnerabilities of Voice Controlled Systems" 1st International Workshop on Security and Privacy for the Internet-of-Things (IoTSec) **arXiv:1803.09156v1 [cs.CR]**

[3] M. Alzantot, B. Balaji, M. Srivastava, "Did you hear that? Adversarial Examples Against Automatic Speech Recognition" NIPS 2017 Machine Deception workshop, arXiv:1801.00554v1 **[cs.CL]**

[4] X. Liu , K. Wan , Y. Ding, "Adversarial Attack on Speech-to-Text Recognition Models" arXiv:1901.10300v1 **[eess.AS]**

[5]Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. 2017. DolphinAttack: Inaudible Voice Commands. In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17). ACM, New York, NY, USA, 103-117. DOI: https://doi.org/10.1145/3133956.3134052

[6]W. Diao, X. Liu, Z. Zhou, and K. Zhang, “Your voice assistant is mine: How to abuse speakers to steal information and control your phone,” in Proc. of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices. ACM, 2014, pp. 63–74.

2. NotPetya and the 2017 Cyberattack on the Ukraine

2(A) Introduction:

On June 27th of 2017, a devastating cyberattack tore its way across eastern Europe. An apparent ransomware attack appeared suddenly and apparently out of nowhere, disabling networks across entire companies and completely shutting down the Ukrainian government. As cybersecurity and technology experts scrambled to contain the damage, several factors emerged to show that this new threat, NotPetya, while superficially resembling it's namesake ransomware (Petya), was far more elaborate and destructive, and seemed to be developed as a weapon rather than a traditional ransomware. The resulting damage was devastating and is recorded as the largest (in terms of dollar amount) cyberattack in history, costing over 10 billion dollars to recover from.

2(B) Description:

NotPetya is a cyberweapon which was initially modeled after the Petya ransomware, sharing the same backdoor exploit EternalBlue (also used by wannacry) to gain access to target systems. However, this vulnerability had already been patched and was only effective on unpatched systems. NotPetya shared several key improvements over its ransomware sibling. Once a system was infected, the weapon triggered a shutdown, and upon restart started encrypting 65 types of files[2] with a uniquely generated 128bit AES key, which was then reencryped a public 2048 bit RSA key. While a unique ID was created to each victim, there seemed to be no correlation between the key and the ID.[1] It scanned for any unpatched systems to try to gain access using EternalBlue and EternalRomance (two exploits that had been patched). Simultaneously, it also used a modified version of the Mimikatz tool to steal the user's Windows credentials, and used those credentials along with PsExec to move laterally and spread throughout the network, meaning even computers patched against the two above exploits were vulnerable. It then encrypted the master file table and overwrote the boot record to display its message to the user.[2] As primary infection vector, the creators of NotPetya hacked and created a backdoor in MeDoc, a tax accounting software commonly used in the Ukraine. They used this backdoor to trigger an update to MeDoc to download and upload NotPetya into computers installed with MeDoc across Eastern Europe on June 27th.[1] The sophistication of the malware, the infection of primarily Ukrainian targets, the timing of the attack (attack took place right before a major holiday, potentially indicating maximum inconvenience rather than over a weekend as would be typical for a ransomware attack) and the fact that while it appeared to be ransomware at first, the actual methods of payment were dead ends and no computers were unencrypted led the CIA to say with high confidence that it was a state sponsored attack from Russia.[4]

2(C) Impact:

The attack spread quickly, overwhelming the networks of large corporations and the Ukrainian government alike in a matter of hours. The lateral movement techniques employed by the attack were extremely effective; disrupting entire companies from a single instance of MeDoc, affecting over 300 companies, and crippling the Ukrainian government and various aspects of their infrastructure. It was quickly apparent that, though the software appeared to be a ransomware, there was no way to retrieve/decrypt the files. All told, the lost revenue, compensation, and repair of the damage took the total cost of the attack to over 10 billion dollars, much of it being footed by a handful of large corporations who were almost completely crippled.

2(D) Possible Detection or Prevention Mechanisms:

Unfortunately, it was impossible to detect prior to the attack, though it was possible to stop the encryption to some degree by unplugging systems quickly enough. Quickly, cybersecurity companies worked to develop detections and prevention systems. From the perspective of prevention, during the attack, the initial footprint was inevitable due to the backdoor in the MeDoc software. Secondarily, however, the malware was able to traverse networks so quickly (in part) using exploits that had already been patched. In these cases it was the lack of patch application that left many systems vulnerable. The prompt application of the patch would have prevented the spread in these cases. Also, it was able to move freely on many networks due to the lateral structure of those networks, with administrators having power on areas of the network outside of their scope. Properly tiering and subdividing the network would have stopped the use of those administrator's credentials.

2(E) References:

[1]US Department of Homeland Security CISA Alert (TA17-181A)
https://www.us-cert.gov/ncas/alerts/TA17-181A (Retrieved 2/17/2019)

[2]Thomson, Ian, Everything you need to know about the Petya, er, NotPetya nasty trashing PCs worldwide, The Register https://www.theregister.co.uk/2017/06/28/petya_notpetya_ransomware/ (Retrieved 2/17/19)

[3] Greenberg, Andy, The Untold Story of NotPetya, the Most Devastating Cyberattack in History, Wired Magazine 8/22/18, https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/ (Retrieved 2/17/19)

[4] Nakashima, Ellen, Russian Military was behind 'NotPetya' Cyberattack in Ukraine, CIA Concludes, 1/12/18, https://www.washingtonpost.com/world/national-security/russian-military-was-behind-notpetya-cyberattack-in-ukraine-cia-concludes/2018/01/12/048d8506-f7ca-11e7-b34a-b85626af34ef_story.html?noredirect=on&utm_term=.7b8fc3447c8e (Retrieved 2/17/19)

3. Cyber Kill-Chain

3(A) Overview:

I will be examining a hypothetical intrusion on organization X using the audio attacks described in section 1 in light of the Cyber Kill-chain described by Lockheed Martin.

3(B) Reconnaissance

In the reconnaissance step, we gather information about the organization we want to infiltrate. Making our job difficult is the fact that voice operated interfaces are mostly tied to residences rather than organizations or companies. In fact, the most effective use of this attack to attack an organization (rather than an individual) may be to gain information for a different type of attack. Monitoring the target would allow you to know the patterns of when the target would be at home; as we've discussed, the absence of a physical person removes the single largest detection and prevention mechanism at this time. If you could obtain vocal samples from the target it could aid in the creation of the attack commands, depending on what type of replay attack we are attempting. Also, scanning the networks and seeing what/if any additional IoT devices are present and could be taken advantage of could be advantageous. The information obtained in this step could be used to determine the best type of audio attack.

3(C) Weaponization

The weaponization is the stage where we build commands to take advantage of the voice activated system. We would also consider the best method of avoiding detection and what delivery system we would like to use in line with our end goals. For example, if possible and we knew the schedule of the target, we could use another IoT compromised machine to pass the commands to the voice interface audibly. But if the target's schedule is unknown, we'd have to use one of the methods to obscure or silently pass the commands to the voice-operated interface. All of the reconnaissance and the goals of the infiltration determines the parameters of weaponization.

3(D) Delivery

The delivery can be in a number of ways, from a breached IoT speaker near the interface, to a portable speaker (if the interface is audibly accessible from outside). The most appropriate way is determined by the reconnaissance carried out prior to the delivery.

3(E) Exploitation

This is a somewhat unusual case, as the last four steps may happen somewhat simultaneously depending on the objectives. If the command is the straightforward unlocking of a door to gain access, there may be no need for installation of any further software. Whereas, if the objective is to set up malware on the user's interface to record additional reconnaissance for a different attack on the organization, the exploitation could be the beginning of a brand new reconnaissance phase by directing the voice-operated interface to give access to different information. If the objective is to further compromise the interface, this is the phase where the interface would be given appropriate commands to open itself up to further exploitation.

3(F) Installation

Again, this phase may or may not apply, depending on the objective. However, if it is appropriate, this is where malware would be installed on the machine to ease use in the future, and to allow the unit to gather and transmit further intelligence for your next attack.

3(G) Command and Control

Now that the unit is under our control, we can use it to obtain access to the personal information of the target it has available, as well as potentially the physical location (if it is connected to physical security systems, such as IoT door locks)

3(H) Actions on Intent

Again, this could have happened as early as the exploitation phase above, depending on the attack. We may be able to use the interface to gain stored information, infiltrate the physical locations it is tied to, or potentially set it up to eavesdrop on our target and gather additional information in the future.