Home Alone: The Insider Threat of Unattended Wearables and A Defense using Audio Proximity

Prakash Shrestha, Babins Shrestha, Nitesh Saxena
University of Alabama at Birmingham
Email: prakashs@uab.edu, babins@uab.edu, saxena@uab.edu

***Abstract*—In this paper, we highlight and study the threat arising from the unattended wearable devices pre-paired with a smartphone over a wireless communication medium. Most users may not lock their wearables due to their small form factor, and may strip themselves off of these devices often, leaving or forgetting them unattended while away from homes (or shared office spaces). An "insider" attacker (potentially a disgruntled friend, roommate, colleague, or even a spouse) can therefore get hold of the wearable, take it near the user's phone (i.e., within radio communication range) at another location (e.g., user's office), and surreptitiously use it across physical barriers for various nefarious purposes, including pulling and learning sensitive information from the phone (such as messages, photos or emails), and pushing sensitive commands to the phone (such as making phone calls, sending text messages and taking pictures). The attacker can then safely restore the wearable, wait for it to be left unattended again and may repeat the process for maximum impact, while the victim remains completely oblivious to the ongoing attack activity. This malicious behavior is in sharp contrast to the threat of stolen wearables where the victim would unpair the wearable as soon as the theft is detected. Considering the severity of this threat, we also respond by building a defense based on audio proximity, which limits the wearable to interface with the phone only when it can pick up on an active audio challenge produced by the phone.**

I. Introduction

Wearable computing is a rapidly emerging paradigm that is incorporated into items of clothing and accessories which can be comfortably worn by the users. Smartwatches, fitness trackers, Google Glass, and Emotiv headset, are some of the examples of already ubiquitous wearable devices. Wearable devices bring immense benefits to society and boast improved quality of life for wearers, ranging from interaction with virtual objects in an augmented reality world to healthier, "fitness-data inspired" lifestyles. However, being attached to the body of the wearer, in contrast to traditional devices, wearables raise unique security and privacy vulnerabilities, as shown by some recent studies [1]–[5].

One potentially critical threat pertains to the (typically) unguarded access to wearable devices due to their small form factor and interface-constrained nature. In particular, if a wearable device is stolen, anyone would have access to the information stored on the device, since wearable devices usually store data locally without encryption, PIN protection or user-authentication. Even if the device supports such authentication functionality, users are not willing to use it (as observed in our survey results presented in Section VIII-B) perhaps because they may perceive it as an inconvenient mechanism. This further emphasizes the seriousness of the threat of unattended wearables. Unlike wearable devices, smartphones offer various authentication schemes such as PIN/password, pattern, and fingerprints that are more widely adopted to secure the access to contents of the phone.

In this paper, we highlight and investigate a new attack vector — *Home Alone Wearables (HAW)* — arising from the *unattended* wearable devices pre-paired with a smartphone over a wireless communication medium (e.g., Bluetooth). As mentioned above, most users may not lock their wearables, and may strip themselves off of these devices often, leaving or forgetting them unattended while away from homes (or shared office spaces). An "insider" attacker, possibly a disgruntled friend, roommate, colleague or spouse, can therefore gain physical access to the wearable once left unattended, take it near the user's phone (i.e., within radio communication range) at another location (e.g., user's office), and surreptitiously use the wearable against the phone for various nefarious purposes across physical barriers and from some distance away. Specifically, *passive attacks* can be launched by pulling and learning sensitive information to the wearable from the phone, such as messages, photos or emails, and *active attacks* can be launched by pushing sensitive voice commands from the wearable to the phone, such as making phone calls, sending text messages and taking pictures. Once done, the attacker can safely restore the wearable exactly where it was left by the victim, wait for it to be left unattended again and may repeat the process for maximum damage. The attack can go on for as long as the attacker desires while the victim would remain completely unaware of the ongoing attack activity.

The threat posed by insiders is a known concern. In particular, the study of Muslukhov et al. [6] found that users are generally concerned about insiders accessing their data on smartphones, and presented evidence that the insider threat is a real problem impacting smartphone users. Our work takes the notion of such insider threats to the realm of wearable devices, pointing to a novel class of vulnerability not studied previously. We note that the need for the attacker to hide behind a physical barrier is an important attribute of our HAW vulnerability, since the attacker is an entity familiar to the victim and would not want to expose himself to the victim.

The malicious behavior exhibited in HAW lies in stark contrast to the traditional threat of stolen wearables. In case of stolen wearables, the victim would "unpair" the wearable as soon as the theft is detected, and therefore the stolen wearable would no longer be able to interface with the phone. In contrast, HAW will allow the wearable to "work with" the

978-1-5386-4586-4/18/$31.00 ©2018 IEEE

phone normally because the victim never unpairs the wearable as the wearable has not been stolen and the overall attack is invisible to the victim. Moreover, the threat of stolen wearables is due to "external" entities (e.g., burglars), likely not familiar to the victim. Whereas, the HAW attackers are anticipated to be "internal" entities, i.e., people familiar to the victim, potentially his/her closed ones, with the hidden intention of monitoring and stealing sensitive information without getting detected.

Considering the severity of the proposed threat, we take a step forward and go on to respond to the threat by building a defense based on *audio proximity*, which limits the wearable to interface with the phone only when it can pick up on an *active audio* challenge produced by the phone. Specifically, whenever the wearable queries to push/pull information to/from the phone, the phone generates a simple audio sound which will be recorded by both the phone and the wearable. Only if the phone deems that the audio sample recorded by itself and the audio sample recorded by the watch are highly similar to each other, the phone will respond to the wearable's queries. HAW attacker would not succeed since wearable would not be able to fully pick up on the sound challenge generated by the phone due to sound dissipation over distance and through physical barriers using which the attacker hides from the victim.

While the HAW threat/defense may be broadly applicable to different types of wearables that work in conjunction with a smartphone, in this paper, we focus our technical exposition on a *smartwatch* and a *smartglass*, given their popularity among the users, especially that of the smartwatch.

Our Contributions: We believe that our work makes the following contributions:

- ***Novel Offense-Defense of Home Alone Wearables***: We introduce a potentially serious vulnerability of unattended wearables and provide a viable fix based on a simple yet effective notion of active sound proximity.
- ***Evaluation of the Threat***: We design and test our attack in different settings involving multiple Android wearables, having varying communication ranges which allow the attacker to launch the attack across physical barriers (e.g., walls) while remaining invisible to the victim.
- ***Design, Implementation and Evaluation of the Defense***: We design, develop and evaluate our defense for the Android platform based on acoustic proximities using simple and short active notification sounds generated by the phone. Our defense may also be executed for only a randomly selected access attempt, which may reduce the distraction effect and still curb the chances for the success of the attacker. Our results show that the defense can significantly lower the impact of the threat without preventing the users from accessing their devices in most benign settings (also confirmed by our survey results) and without requiring any effort from the users. Also, the defense can completely block the attack executed across exterior walls or through building floors.
- ***Population Statistics Supporting Threat Assumptions***: We confirm the validity of our threat assumptions based on an online survey with Amazon Mechanical Turk participants that shows users are indeed prone to leaving or forgetting their wearables (smart or not) unattended frequently for long periods of time.

II. Threat Model

In the HAW threat model, we consider that a user has already paired his wearable device $\mathcal{W}$ with his smartphone which acts as a primary device $\mathcal{P}$. Through this pairing, $\mathcal{P}$ and $\mathcal{W}$ have established a security association and follow a standard cryptographic protocol for authenticating and encrypting all messages exchanged between them. We further assume that the user has not set any form of authentication in $\mathcal{W}$ due to reasons such as many wearable devices do not provide authentication mechanism by default or users may not want any authentication on $\mathcal{W}$ due to the inconvenience associated with the constrained user interface. As $\mathcal{P}$ device offers various authentication schemes (e.g., PIN/password, fingerprint), which are more widely adopted, we assume that the user has locked the $\mathcal{P}$ device using one of these schemes.

We further assume that the user has left his $\mathcal{W}$ device on desk (let's say for charging) or has forgotten it at home while the user is in possession of the $\mathcal{P}$ device at another location (let's say office). An *insider* attacker (possibly user's roommate, spouse, friend or colleague) gains the possession of the unattended/left $\mathcal{W}$ device. The attacker may either want to damage the victim's reputation or just want to sneak upon the victim out of curiosity. The attacker with $\mathcal{W}$ in possession knows the victim's whereabouts and attempts to interface $\mathcal{W}$ with $\mathcal{P}$. We also assume that interfacing $\mathcal{W}$ with $\mathcal{P}$ does not require any action from the user unless $\mathcal{W}$ is explicitly unpaired by the user. An approach of requiring an action from the user may be annoying to the user because he has to approve the pairing even when the wearable is few meters away. Since we assume that user has not enable authentication in $\mathcal{W}$, the attacker with $\mathcal{W}$ can just be in close proximity (within the Bluetooth range of $\mathcal{P}$) of the victim with $\mathcal{P}$ to pull/push some information from/to $\mathcal{P}$. The attacker can pursue this process over multiple rounds, where $\mathcal{W}$ device can be stowed back safely where it was, and the attacker can wait for $\mathcal{W}$ to be left unattended again by its user, and re-iterate.

We also assume that when a user attempts to use $\mathcal{W}$ to access $\mathcal{P}$, both two devices remains in the same room close to one another (within a few feet), whereas when the attacker wants to use $\mathcal{W}$ to access $\mathcal{P}$, the attacker has to be stealthy and cannot be in visible range. Hence, the attacker tries to hide behind a barrier such as a door, a wall or even on another floor. The attacker within a visible range would most likely create suspicion to the victim and raises the chances of getting caught. This fundamental difference between the benign setting and the attack setting (with barrier) serves as the main premise for the HAW defense proposed in this paper.

III. Attack Taxonomy and Experiments

A. Attack Taxonomy

We broadly divide HAW attacks into two categories - *Passive attacks*, and *Active attacks*, as described below. Some of our attack scenarios can be visualized in Figure 1.

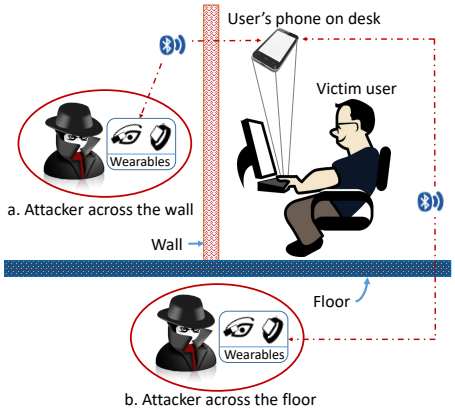

Fig. 1: HAW attack example scenarios. An attacker with an unattended wearable comes in the Bluetooth range of the phone while being hidden across physical barriers, e.g., (a) wall, and (b) floor.

Passive Attacks: In passive attacks, the attacker will execute only those commands from $\mathcal{W}$ possessed by the attacker that do not alert the victim user in possession of $\mathcal{P}$. This kind of attack is stealthy and the possession of the unattended device may go unnoticed. Reading call logs, emails, SMS, ongoing notifications, or viewing daily agenda or appointments, are a few of the examples. Besides executing commands to passively read sensitive information, the attacker can also quietly observe the notifications (of call or of SMS) that are pushed by $\mathcal{P}$ to $\mathcal{W}$. Hence, the victim's privacy would be compromised easily as a consequence of this attack.

Active Attacks: In active attacks, the attacker will execute commands from $\mathcal{W}$ to trigger some activity on $\mathcal{P}$. When an attacker possesses $\mathcal{W}$ powered with Android Wear [7], it may execute different voice commands with "Okay Google" [8]. Some $\mathcal{W}$ devices provide limited set of commands that can be executed. However, most of these devices allow to make calls/send SMS, take pictures, create events/alarms, etc. Some of the $\mathcal{W}$ devices such as Android powered wear or Google Glass even provide commands to update posts to Facebook.

It is clear that such active attacks could be quite devastating. One disadvantage is that these attacks may not be very stealthy, and the possession of the unattended device may be noticed by the victim as active alteration in the user interface of the device or creation of alerts may distract the victims. However, it is important to note that users often leave their phones unattended or inside pockets or bags [9], and therefore they may not frequently notice such alerts.

B. Attack Settings and Experiments

As mentioned above, to launch the HAW attacks, an attacker who has $\mathcal{W}$ needs to be within the Bluetooth range of $\mathcal{P}$ device. Bluetooth provides distance coverage between 1-100 m. The coverage range of Bluetooth depends on many factors such as communication configuration, surroundings, and radio performance [10], [11]. We set forth to do an analysis of Bluetooth range in some of the commonly available $\mathcal{P}$ and $\mathcal{W}$ devices. We use LG G3 (G3), Nexus 5 (N5) and Moto G4 (G4) as the $\mathcal{P}$ devices while we use LG G Watch R (W110) smartwatch (GR), Samsung Gear Live (GL) and Google Glass (GG) as the $\mathcal{W}$ devices. We pair N5 with GR, G4 with GL and G3 with GG. Among the $\mathcal{P}$ devices, N5 has Bluetooth v4.0, G3 has Bluetooth LE (Low Energy) v4.0 and G4 has Bluetooth LE v4.1. Among the $\mathcal{W}$ devices, GR has Bluetooth v4.0, GL has Bluetooth LE v4.0 and GG has Bluetooth v3.0.

TABLE I: Bluetooth Range: In between walls/barriers

Barrier	Air (No Wall)	Glass door	Two Glass doors	Wooden door	Interior Wall	Exterior Wall
N5 - GR	40m	15m	13m	5m	5m	4m
G4 - GL	40m	15m	13m	12m	8m	4m
G3 - GG	40m	15m	13m	7m	8m	6m

We tested the Bluetooth connection when two devices are separated by different kinds of barriers such as glass wall, wooden door, interior wall or exterior wall. We report the farthest distance from which we were able to execute the command with the barrier in between in the Table I. The results of our observation are as expected. Depending upon the type/material of the barrier the Bluetooth range differed. For example, in our experiment, devices were connected even up to 40m (1m = 3.28 feet) when there was no barrier while we could only send commands up to 8m (or even shorter distance) when the devices were separated by doors or walls. Moreover, we also observed that depending upon different kinds of Bluetooth (v4.0, v4.1, or BLE v4.0) embedded on the device, the Bluetooth signal varies significantly. N5 - GR pair with normal Bluetooth got disconnected after 5m when they were separated by wooden door or interior/exterior wall. However, G4 - GL (both with Bluetooth LE) were connected even up to 12m when separated by wooden door. In case of G3 - GG pair (one with Bluetooth LE and another with normal Bluetooth), they were disconnected at around 7m when they were separated by an wooden door. The Bluetooth connection range further decreased when there was an exterior wall in between the paired devices.

Note that as long as $\mathcal{P}$ and $\mathcal{W}$ are within the Bluetooth range (as shown in Tables I), the attacker with the possession of $\mathcal{W}$ can launch any of the active and passive attacks as mentioned earlier in Section III-A.

IV. Limitations of Traditonal Defenses

Location-based approaches, and distance-bounding protocols may be the potential approaches to address the HAW threats by estimating the physical proximity, or estimating the presence of barrier between the devices.

GPS-based Approaches: A location-based approach determines the physical proximity by comparing the location information extracted from the two devices at a given time. It requires both devices to be equipped with GPS. Since the measurement error of phone GPS lies above 5 meters [12], use of location-based approach would not be effective way to estimate the close physical proximity of few feet and cannot be used to defeat the HAW threat. Further, most of the wearables today do not come with on-board GPS sensors, e.g., LG G Watch R, the smartwatch that we use in our experiment and evaluation. Moreover, GPS sensors usually do not work well in indoor environment, which is a common use case of wearable and mobile devices.

Distance-bounding Protocols: A distance bounding protocol usually uses the received signal strength indicator (RSSI) and

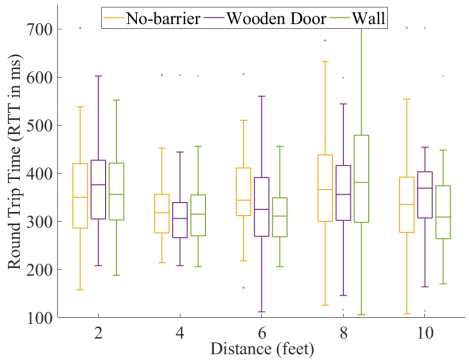

Fig. 2: Whisker diagram showing the distribution of Round-Trip Time (RTT) over various distances and across different barriers. The central rectangle spans the first quartile to the third quartile, and the line inside it presents the median. The *'whiskers'* above and below the box show the minimum and maximum value of RTT within 1.5 IQR.

the Time-of-flight (ToF) for distance estimation between two devices. RSSI is a measurement of strength of the radio signal received by the device. It changes over the distance from the source of signal. However, RSSI is not a reliable and secure method for distance estimation because signal strength can be altered by amplifying or attenuating the signal [13].

ToF system estimates the distance by measuring the time elapsed, in particular round-trip-time (RTT), during a message exchange. We implemented ToF system with a smartphone (Samsung Galaxy S6) and a smartwatch (LG G Watch R) utilizing Bluetooth for message exchange. We recorded RTT (in milliseconds) values 100 times by placing the watch and the phone at various distances from 2-10 feet with the increment of 2 feet on a smooth surface. We also recorded RTT values by positioning the phone and the watch across the wooden door and the interior wall as a barrier.

Figure 2 shows the distribution of RTT values for various distances between the phone and the watch across various barriers in the form of a *'whisker diagram'*. As can be seen from the figure, at each distance value, the RTT values are balanced (centered around the median) at some point for each type of barriers. All the RTT values seem to be balanced around 300-400ms for each of the distance value. This shows that there is no significant variation in the RTT values while placing a barrier between the devices (attack case) as compared to no-barrier in between the devices (benign case). It also shows no significant variation in RTTs across distance levels in presence of each type of barriers. There may be variation in RTT values across longer distances, and can be applied to make a distinction between close distance versus far distance. However, for estimating the close physical proximity and the presence of barrier, ToF is not a viable approach, and cannot be used to address the HAW threat.

V. Our Defense

The basic idea of our defense mechanism is limiting $\mathcal{W}$ to interface $\mathcal{P}$ based on their audio proximity between two devices, which is determined by the similarity score between the audio pairs recorded by these two devices. The concrete steps followed in HAW defense process are as follow. The user tries to access $\mathcal{W}$, for example, reading the texts/emails or sending various commands to $\mathcal{P}$ through $\mathcal{W}$. Before granting user access to the $\mathcal{W}$'s capabilities, $\mathcal{W}$ sends the "Trying to access" command to $\mathcal{P}$, and starts recording the audio. When $\mathcal{P}$ receives the "Trying to access" alert, it picks one of the sounds from the pool of notification sounds, and plays it back through its speaker and at the same time starts recording the audio. As soon as audio playback stops on the $\mathcal{P}$ device, it also stops recording and sends the "STOP" command to $\mathcal{W}$ to stop recording. The recordings from the $\mathcal{W}$ device is encrypted/authenticated and transmitted to $\mathcal{P}$ for proximity analysis. All messages between $\mathcal{W}$ and $\mathcal{P}$ are exchanged over the secure channel. The $\mathcal{P}$ device then computes the similarity score between the audio pairs and based on the similarity score, $\mathcal{P}$ decides whether $\mathcal{W}$ device access attempt should be granted.

In the benign case, when the user tries to push/pull information from $\mathcal{P}$ over to $\mathcal{W}$, $\mathcal{W}$ would be able to pick up on the active audio challenge generated by $\mathcal{P}$ as long as $\mathcal{W}$ is not too far from $\mathcal{P}$ and not shielded by barriers. Our wearable-phone usage pattern analysis (Section VIII) confirms that most users use their wearables while the phone is located near to them (e.g., on desk) with likely no physical barrier in between. This suggests that our active audio approach would work well in the benign scenario since $\mathcal{W}$ will be able to capture the sounds produced by $\mathcal{P}$ well, which may result in a high correlation between the corresponding recordings of $\mathcal{W}$ and $\mathcal{P}$. The active audio generated by $\mathcal{P}$ may be distracting to the users, so our approach uses simple and quick notification sounds, which people are already used to and may not be distracting as most of the users keep their phones in ringer mode at home and while asleep (Section VIII-B).

In contrast, in the HAW attack scenario, the attacker carrying $\mathcal{W}$ would usually be located at some distance from the victim ($\mathcal{P}$) and, perhaps more importantly, be separated by a physical barrier from the victim. In this condition, the audio sounds created by $\mathcal{P}$ would be shielded and dissipated significantly by the time they reach $\mathcal{W}$, which may result in a low correlation (or even no correlation depending upon the distance and the type of barrier) between the corresponding recording of $\mathcal{W}$ and $\mathcal{P}$. Therefore, the attack would be detected, preventing the attacker from pushing/pulling information from $\mathcal{P}$.

As an extra layer of security, the use of active audio sounds in our defense would also serve the purpose of alerting the user whenever $\mathcal{W}$ attempts to access $\mathcal{P}$. Our defense may also be executed not per access attempt but when a randomly selected access attempt is made. This may reduce the distraction effect and still curb the chances for the success of the attacker.

Use of Ultrasound: Nyquist principle states that recorders should record at a sampling rate greater than $40kHz$ to record and process the ultrasound ($> 20kHz$). However, since many of wearables, e.g., Sony Smartwatch 3, LG G Watch R, and Google Glass, have maximum sampling frequency of 22.05kHz, they cannot record and process ultrasound, thereby making the use of ultrasound in HAW system infeasible currently. In the near future, wearables may come with powerful microphone with the ability to process ultrasound that may

be used to process it transparently, thereby improving the usability of the system.

Fall-back to Ambient Audio: There may be scenarios where creating sounds may not be feasible, e.g., in a silent zone such as a library, hospital, or meeting. It is also possible that the phone is kept in silent mode or is connected to the earphone. In such scenarios, the HAW defense by nature will fall back to proximity detection based on ambient audio sounds. Here, instead of the audio created by $\mathcal{P}$, the ambient audio captured by both of the devices would be used for deciding whether to grant access. As ambient sounds across different locations within Bluetooth range of $\mathcal{P}$ are likely to be similar, which may lead the defense system to grant illegitimate access to the adversary. In other words, falling back to the ambient sound may not be as secure as its original implementation of using active-sound, but it only needs to be done on occasional basis when sounds can not be played back by the phone.

Comparison with Prior Audio-based Security Schemes: There are several works that utilize the audio signal to improve the usability and security level of existing security schemes, especially in the context of authentication. Sound-Proof [9] is one such security scheme, specifically a zero-effort two-factor authentication scheme, which utilizes ambient audio to determine the proximity between two devices – the phone and the login terminal (browser). Our approach of using audio correlation is similar to that of Sound-Proof, but we make use of actively generated audio sounds, rather than ambient sounds. The use of active sounds provides a high level of security in our scheme as even a nearby attacker may not be able to hear the sounds across barrier, in contrast to Sound-Proof which seems vulnerable to a proximity attacker. Moreover, our application domain is much different from that of Sound-Proof. Halevi et al. [14] employ ambient audio to detect the physical proximity between two devices in NFC transaction scenario. During an NFC transaction, they collect ambient audio from both the NFC phone and the NFC reader and validate if both devices are together by correlating the collected audio samples. Truong et al. [15] also use ambient audio (along with various Radio Frequency (RF) sensors such as Wi-Fi, Bluetooth and GPS) to detect the proximity of two devices. Again, in contrast to these studies, our work makes use of active sound proximity, not ambient sound proximity and focuses on a different application domain. SlickLogin [16], which has recently been acquired by Google, is another scheme that aims to minimize the user phone interaction during two-factor authentication. Our proposed HAW defense is in line with this approach but addresses the threat of "home alone wearables" not two-factor authentication. Also, in contrast to SlickLogin, we make use of simple notification sounds and audio correlation analysis.

VI. Defense Design and Implementation

Application Design: For our prototype design and implementation (and later testing) of HAW defense, we use LG Nexus 5 as the smartphone, and LG G watch R as the smartwatch. Both the smartphone and the smartwatch run Android version 6.0.1. We use Digital Sound Level Meter to measure the loudness of sound (dB) generated by the phone. The design of HAW defense consists of following two applications.

- *Phone Application*: HAW defense implements an Android phone app that has a simple button to control the audio recordings on the phone and on the watch. When the button is pressed to start recording, the phone sends the "start recording" trigger to the watch. The phone app picks a random notification sound (from its repository), and plays it back. It also starts recording the audio simultaneously while it is playing the notification sound. As soon as the phone finishes playing the audio sample, it automatically stops recording and sends the "stop recording" command to the watch, thereby stopping the recording on the watch. The phone stores the recording locally while the watch transmits the recording to the phone, for the purpose of our offline analysis as part of our evaluation.
- *Watch Application*: Our HAW defense implements another Android app as a watch/wear application. Android wear app remains idle in the background and is automatically activated when a "start recording" signal is received from the companion smartphone. Once activated, it starts recording audio, and stops recording as soon as it receives the "stop recording" signal from the phone.

Audio Correlation Analysis: We implemented correlation analysis between an audio pair (recorded by watch and phone) in a similar fashion as that proposed in [9] i.e., one-third octave band filtering with cross-correlation. One-third octave band is defined as a frequency band whose upper edge frequency is equal to its lower edge frequency times cube root of two. One-third octave band divides the audible frequency range (roughly 20Hz - 20kHz) into 32 unequal and non-overlapping bands. One-third octave bands of an audio recording provide its high resolution frequency information while retaining time-domain representation. Audio samples are divided into the bands ranging from 50Hz to 4kHz. As reported in Sound-Proof [9], these bands provided the best Equal Error Rate (EER). Hence, only the sixth band (50Hz) to the twenty sixth band (4kHz) are used, i.e. only twenty bands are considered. To split the audio sample into these bands, we use twentieth order Butterworth bandpass filter [17] in MATLAB. To correlate the audio pair, we use the same system that was implemented in [14]. To measure the similarity between two time-based signals X_i and X_j, we first normalize the signals according to their energy. Then, we calculate the correlation between each signal and use maximum correlation value. The correlation between two time-based signals X_i and X_j is computed as: $Corr(i,j) = max(CrossCorr(X_i, X_j))$

Similar design of the HAW defense can also be implemented with smartglasses. In the smartglass implementation of the HAW defense, all underlying techniques and steps will be same except that there will be a glass application instead of the watch application.

VII. Defense Experiments and Results

A. Data Collection and Settings

In our data collection experiment, we consider two factors that have a significant effect in the performance of our HAW defense: the *volume level* of the phone as it creates the active sounds, and the *distance* between the phone and the wearable. We also consider two different types of *physical barriers*

that can potentially be used by the attacker to hide itself while executing the attacks, i.e., pull/push information from/to the phone using the wearable. Our parameter settings are described below:

Volume Level: As users may have different preferences towards the volume level of the phone being used, we consider three different volume levels for our experiment. – *Full Volume, Average Volume, and Low Volume*. In *Full Volume*, the volume level of the phone is set to 100% (72dB) of the volume. In *Average Volume* and *Low Volume*, the volume level of the phone is set to 75% (66dB) and 50% (60dB) of the highest possible volume, respectively.

Distance: We consider two different upper limits of the distance based on the presence/absence of the physical barrier. In the absence of the physical barrier, which represents the benign setting with no attack, we consider 6 different distance settings starting from 0 feet (i.e., phone and watch being placed next to each other) to 10 feet, with the increment of 2 feet distance in succession. In the presence of the barrier, which represents the attack scenario, we collected samples at 3 different distance settings from the barrier starting from 2 to 6 feet, with increment of 2 feet distance in succession. We consider only up to 6 feet distance in the presence of barrier because if HAW defense can detect the illegitimate use of the wearable within this distance limit, it would also be able to detect the illegitimate use beyond this distance.

Physical Barriers: We chose two physical barriers.

(i) Wooden door: Wooden door is a potentially common physical barrier that can be utilized by an adversary to hide itself while executing the attack. To capture this setting, we selected one of the wooden doors (thickness of 1.5 inch) located in our lab, typically used in most of the buildings, to evaluate the performance of our defense.

(ii) Interior wall: Interior wall is another potential physical barrier that an adversary can use to hide itself while illegitimately perpetrating the attack. To evaluate the performance of our defense scheme in this setting, we chose a dry wall (thickness of approximately 5 inch) located in our lab, as an interior wall barrier.

We collected total of 1800 samples of audio recordings using our prototype implementation of the defense system. Out of these samples, 900 samples correspond to the setting – *without barrier*, while remaining 900 samples correspond to the setting – *with barrier*. Each sample consists of two recordings, one from the phone and the other from the watch. All the recordings were collected in a controlled environment, where there was no significant ambient noise. As the active sounds in our analysis, we chose 10 popular message notification sounds used by several popular applications and mobile devices. The notification sounds of Viber, Skype, Facebook messenger, Hangout, and default message notification sounds of iPhone, Samsung, and Sony devices are some of the examples we chose in our study. Each of these message notification sounds is less than 2 seconds long. For each notification sound, 5 samples of recordings were collected for each combination of distance and volume level settings, thereby making a total of 50 samples of recordings for each setting. In the real world implementation, besides the 10 notification sounds we selected, various other notification sounds can also be used.

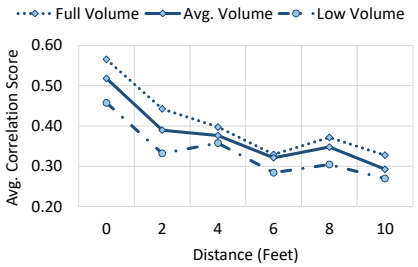

Fig. 3: Average correlation score between the phone recordings and the watch recordings for different settings – three different volume levels and 6 different distances between the watch and the phone.

B. Results

1) Preliminary Results: As a preliminary experiment, we tried to gauge the impact of volume level and distance on the correlation score between the audio pairs recorded by phone and watch. To this end, we used the recordings which are collected in the setting where phone is placed at several distance apart from the watch on a smooth surface without any physical barrier in between, and volume level set to three different levels. We computed the average correlation score between the phone and the watch recordings. The results (Figure 3) show that the correlation between the phone recording and the watch recording attenuate with the increase in the distance between the phone and the watch, and the decrease in the volume level.

Based on these initial results, we proceeded with the analysis of the collected samples to come up with the system's parameters, in particular, the correlation threshold for each volume level, that leads to the optimal results in terms of False Rejection Rate (FRR) and False Acceptance Rate (FAR). A false rejection occurs when the system denies a legitimate access to the device (benign setting), while a false acceptance occurs when the system grants a fraudulent access attempts (attack setting). When an attacker located within the Bluetooth range of the phone attempts to access the phone from the watch, phone creates an audio challenge, which is also recorded by the watch possessed by the attacker. The false access to the watch is granted if the audio recording from the watch possessed by the attacker and the one from the phone have similarity score greater than the threshold used to make the proximity decisions.

2) Analysis without Barrier – Benign Setting: In the benign case, the watch would be close to the phone with no physical barrier in between, and the watch would be able to pick the active audio created by the phone when the user tries to push/pull the information to/from the phone. We consider two distances – 4 feet and 6 feet – between the phone and the watch as legitimate distance range, and performed the analysis for each of them. We note that this analysis represents a benign setting, not the attack setting, since it is done without the presence of barriers. The results from this analysis, as we show in the next subsection, will be used to extend to the attack setting, i.e., in the presence of barriers.

Four feet Distance: We computed FAR and FRR to evaluate the performance of our defense mechanism using following

TABLE II: Threshold chosen for different volume level of the phone considering two different distances between the phone and the watch as a legitimate distance. Thresholds are chosen such that they offer very low FRR (highlighted cells) while providing reasonable FAR in the scenarios without barrier.

	Threshold Chosen	FRR	FAR
4 feet as a legitimate distance			
Full Volume	0.35	0.013	0.387
Average Volume	0.30	0.020	0.570
Low Volume	0.28	0.013	0.527
6 feet as a legitimate distance			
Full Volume	0.30	0.050	0.730
Average Volume	0.29	0.069	0.600
Low Volume	0.26	0.060	0.710

strategy. We used the recordings, which are collected in the settings where the phone is placed at $\leq$4 feet away from the watch on a smooth surface, without any physical barrier, to compute the FRR. To compute FAR, we used the recordings, which are collected in the settings where the phone is placed beyond 4 feet from the watch without any physical barrier. For the full volume setting, we achieved the Equal Error Rate (EER), defined as equilibrium point between FAR and FPR, of 0.22 when the similarity score is 0.39. Similarly, for average and low volume setting, we achieved EER (at correlation score) of 0.24 (0.34) and 0.17 (0.32), respectively.

Six feet Distance: The same strategy, that we used while evaluating 4 feet distance as the legitimate distance limit, is used while considering 6 feet as the legitimate distance limit to compute FAR and FRR. In this case, we achieved the EERs of 0.36, 0.33, and 0.31 when the similarity score were 0.37, 0.33, and 0.31 for full, average and low volume level, respectively.

From the analysis for 4 feet and 6 feet distances, we chose a similarity score threshold for each of the volume level settings that provides very low FRR while providing reasonable FAR. The need to establish a very low FRR is important considering the usability of the system as we do not want to block a legitimate access attempt by a benign user. The similarity score chosen as a threshold for each of the volume levels considering each of the legitimate distance range is as shown in Table II. As we can see, the corresponding FRRs when 4 feet is considered as a legitimate distance are $\leq$0.02 while they are $<$0.07 when 6 feet is considered as a legitimate distance.

3) Analysis with Barrier – Attack Setting: With the chosen thresholdization for each of the volume levels in the previous subsection (benign setting without barriers), we evaluate the performance of our defense in the presence of physical barriers, i.e., a dry wall and a wooden door, which represents the attack setting. We assume that the attacker would use some sort of a barrier to hide itself while performing illegitimate access attempts. We use FAR as a measure to evaluate the performance of our defense against the attack setting. To this end, we considered all the recordings which were collected in the setting where there was a barrier between the phone and the watch as the attacker samples. In the real-world attack scenario, the phone possessed by the victim would be at some distance (may be more than 2 feet away) from the barrier, in contrast to the watch possessed by the attacker, which could be very close to the barrier. So, we consider the distance of the phone from the barrier starting from 2 feet to 6 feet in both legitimate distance consideration.

TABLE III: False Acceptance Rate (FAR) of HAW while keeping the phone at different volume level at a varying distance from the barrier (wall and wooden door) considering up to 4 feet and 6 feet as a benign distance limit. Highlighted cells show the average FARs of corresponding volume level.

Distance (feet)	FAR			
	Wall-4	Door-4	Wall-6	Door-6
Full Volume				
2	0.020	0.033	0.040	0.333
4	0.000	0.000	0.000	0.080
6	0.000	0.000	0.120	0.180
Average	0.007	0.011	0.053	0.198
Average Volume				
2	0.020	0.040	0.039	0.220
4	0.000	0.000	0.020	0.060
6	0.000	0.000	0.040	0.039
Average	0.007	0.013	0.033	0.106
Low Volume				
2	0.020	0.060	0.200	0.300
4	0.000	0.040	0.220	0.340
6	0.020	0.060	0.220	0.280
Average	0.013	0.053	0.213	0.307

Table III shows the FAR for our defense in presence of the barrier (dry wall and wooden door) at different volume level and distance of the phone from the barrier while considering 4 feet and 6 feet as legitimate distance limit. Note that here we consider the recordings only from the attack settings, so no FRRs are shown for this setting. The FAR values are lower in case of the wall than in case of the door in all volume level settings and in both legitimate distance considerations (4 feet and 6 feet). The reasoning behind this may be that the audio created by the phone attenuates more through the wall than through the wooden door.

Considering 4 feet of distance between the watch and the phone as a legitimate access limit, we obtained the average FAR of 0.007 for full volume setting for wall, and average FAR of 0.011 for door case. Similarly, for average volume setting, we got FAR of 0.007 and 0.013 in wall and wooden door barrier, respectively. For low volume setting, the FAR is 0.013 for the wall and 0.053 for the door as physical barrier. When considering 6 feet of distance as the legitimate distance, the FAR values increases, as can be seen in Table III. This may be because the recordings, which is collected in the setting where the phone is kept at a distance of 6 feet has similar audio features as the ones collected in the setting where barrier is placed in between the phone and the watch.

With this analysis, we found that our defense works better with keeping legitimate distance limit to 4 feet than with 6 feet. We believe that the consideration of up to 4 feet of distance between the phone and the watch is a realistic assumption. In fact, our survey study (Section VIII-B) shows that when users are accessing their wearable devices through smartphone, both devices are usually nearby, either in the room or next to each other. In such a case, 4 ft of distance as the legitimate access limit would work well for benign users, while highly resisting the HAW attackers.

4) Analysis in Other Scenarios: We also performed audio correlation analysis when the phone is kept in a pocket (benign

case) as well as with exterior wall scenario (attack case). As indicated by our survey results, users may often keep their phones in a pocket while accessing the watch/glass. To evaluate our defense with respect to this benign scenario, we performed a simple audio analysis with the samples collected in such setting. We collected 20 samples of recordings by placing the phone inside a pocket of the user wearing the watch. During the experiment, the user was sitting on a chair and his hand was resting on a desk. We collected the samples at full and average volume level and computed the FRR considering 4 feet as legitimate distance limit. We found that our defense works well in such "phone-in-pocket" settings as well. The FRR of our approach in this setting is 0.00 for full volume case and 0.05 for average volume case.

In the exterior wall attack setting, keeping both the phone and the watch very close to the barrier but in two different sides of the wall, we collected 20 samples of recording by setting the volume level of the phone to its fullest. To evaluate the performance of our defense against such attack setting, we again consider 4 feet as legitimate distance range and computed FAR. We found that the FAR of our approach in such a setting is 0.00. This means that our approach can successfully detect the illegitimate access through such setting, since exterior walls may serve to shield the active audio sounds very well from the watch. We also repeated a similar experiment across two floors of a building (as in one of our tested attack in Section III). As expected, the attack was completely prevented yielding an FAR of 0.00.

VIII. Wearable Usage Statistics Supporting Our Attack and Defense

Since wearable-smartphone systems seem vulnerable in case participants intentionally or accidentally leave/forget their $\mathcal{W}$ devices at home or at desk, we conducted a survey to study the user habits with such wearables and smartphone by recruiting Amazon Mechanical Turk workers. The study was approved by our University's IRB. The participants in the study were strictly voluntary and they could opt out of the study at any time. The survey took about 15 minutes for each participant, for which they were compensated $0.5. In this section, we discuss the design and results from the survey.

A. Study Design

To better inform the design and execution of our HAW attacks in the real-world (as well as our defense mechanism), we asked the participants to answer several questions about their smartphone and wearable devices usage including their habits of using normal watch and normal glasses. We believe that the usage patterns of traditional watches and glasses may be generally aligned with that of smartwatches and smartglasses. Below we summarize the set of questions we posed during the survey.

Demographical Information: We asked the participants about their gender, age, education, industry or field they belong to, country of residence, and their general computer knowledge.

Smartphone Habits: We asked participants – how often they leave/forget their phone at home when they go to work or at desk for charging; in which mode (silent, vibrate, ring) they keep their phone while at work, at home and while asleep; how often they hear notification sounds due to calls/messages and how often they are distracted by such noises; if they use some form of authentication to lock/unlock their phone.

TABLE IV: Phone/Wearables placement when using companion device. Since user can keep one device at multiple places over different times while using corresponding companion device, the summation of percentage of the phone-wearable placements does not yield to 100%. "Wearable-worn": smartwatch is worn on the wrist and smartglass is worn on the head.

		Wearable-worn/ Phone-in-pocket	Desk	Purse	Next to them
Wearables placement when using phone	Smartwatch	75.0%	18.8%	9.4%	12.5%
	Smartglass	23.8%	19.0%	9.5%	23.8%
Phone placement when using wearable	Smartwatch	59.4%	40.6%	18.8%	12.5%
	Smartglass	38.1%	28.6%	23.8%	9.5%

(Smart)watch Habits: Since many participants may not have smartwatch, we first asked participants if they own a smartwatch and if they do not own a smartwatch, we queried them for traditional watch habits. We asked them – how often they take off their watches, forget their watch at home when they go to work/school; where they normally keep their phone when they use their watch; and, where they keep their watch when they access computer and phone. For those who have smartwatch, we also asked – how often their smartwatches run out of battery; how often they leave their watch at work/home for charging or other reasons; from how far they access their smartwatch from phone or phone from their watch; if they use some form of authentication on their smartwatch; if they read their emails/SMS from smartwatch and send commands to make calls/send SMS from their smartwatch to phone.

(Smart)glass Habits: Alike (smart)watch habits, we asked participants similar set of questions in the context of (smart)glass. They were asked about their habits towards (smart)glass rather than towards (smart)watch.

B. Study Results

Our survey polled a total 110 participants from within the pool of Amazon Mechanical Turk workers. In the survey, only participants who use smartphone and wear watch/glasses (either smart or normal) were considered eligible. Among the participants, 29.6% own a smartwatch and 19.4% own a smartglass. For others, we queried about the traditional watch/glass habits. We present the general placements of the phone/wearables when the users access corresponding companion device in Table IV and summarize the different habits of participants related to smartphone, (smart)watch, and (smart)glass in Table V.

Demographical Information: The survey participants consisted of 56.9% males and 43.1% females. Most of the participants fell in the age-groups of 25-34 (46.8%) and 35-44 (33.0%). The participants were from different industrial background with most from the Information Technology field. All the participants ranked their general computer skills to be either excellent (51.4%), good (37.6%), or average (11.0%). The demographic information shows that the survey covers a representative sample of diverse real-world users of devices central to the focus of our work.

TABLE V: Summary of how often the survey participants carry/wear, leave, forget their phones, (smart)watch, or (smart)glasses. The first column in each row shows the questions asked and the column headers show options provided. "NA" means the option was not provided for that question.

		Always	Most of the time	When leaving for school/work	Once in a while	Never
Carry phone		70.4%	27.8%	1.9%	0.0%	0.0%
Leave phone unattended		2.8%	8.3%	NA	66.7%	22.2%
Forget phone		1.9%	0.9%	NA	31.5%	65.7%
Connected to phone	**Smartwatch**	28.1%	56.3%	NA	12.5%	3.1%
	Smartglass	23.8%	23.8%	NA	42.9%	9.5%
Leave smartwatch unattended	**Home**	9.4%	18.8%	NA	65.6%	6.3%
	Work	6.3%	6.3%	NA	59.4%	28.1%
	Asleep	25.0%	31.3%	NA	31.3%	12.5%
Forget watch	**Smart**	0.0%	9.4%	NA	65.6%	25.0%
	Normal	4.8%	14.3%	NA	43.8%	37.1%
Leave smartglass unattended	**Home**	0.0%	52.4%	NA	38.1%	9.5%
	Work	14.3%	19.0%	NA	47.6%	19.0%
	Asleep	33.3%	38.1%	NA	23.8%	4.8%
Forget glasses	**Smart**	19.0%	19.0%	NA	47.6%	14.3%
	Normal	1.9%	6.5%	NA	16.7%	75.0%

Summary of Results: From the survey, we notice that most of the participants keep their phone in ringer mode while they are at home or while asleep. More precisely, the participants keep their phone in vibrate (45.4%), silent (22.2%), or ringer (32.4%) mode at work while they keep their phone in vibrate (16.7%), silent (4.6%), or ringer (78.7%) mode at home and vibrate (24.1%), silent (23.1%), or ringer (52.8%) mode while asleep. From the survey results, we also observe that many people often leave their wearable devices such as smartwatch or smartglass at home or on desk for charging or other reasons, while they generally keep their smartphone with them (i.e., "attended") protected. The smartphones are kept secured by using various lock screen mechanisms such as PIN (27.8%), password(16.7%), lock pattern (13%), fingerprint (17.6%), and others (16.7%). We also see that many users do not lock their wearable devices which may be due to the constrained user interface of the devices. We observe 43.8% of smartwatch users and 62.0% of smartglass users do not use any form of authentication to protect their smart devices. This supports our HAW attack assumptions regarding the wearable devices being left unattended and unlocked often. The survey results also show that when the users are accessing their wearable devices through smartphone, or smartphone through wearable devices, both devices are usually nearby, either in same room or next to each other. This serves to support our HAW defense that users are generally accessing their wearable-smartphone paired devices without any physical barrier in between them. The results also show that 63.9% of the participants either hear the notification sounds most of the time or even more frequently; however, only a few participants get distracted by such noises either always (3.7%) or most of the time (7.4%). This indicates that most people do not seem to be distracted with notification sounds created by their phones, which affirms the use of notification sound as active sounds in our defense mechanism.

IX. Conclusion

In this paper, we formalized and studied the threat of "home alone wearables", which allows an attacker to exploit an unattended wearable for secretly accessing information from, or submitting commands to, the companion smartphone in possession of the victim user. The attack simply works over the radio communication media and may remain nearly oblivious to the victim executed from several meters away through thick physical barriers such as doors, walls or even building floors. As our response to the threat, we proposed a new defense mechanism that makes use of audio proximity detection based on actively generated, yet simple and short sounds produced by the phone prior to granting access attempts from the wearable. Our evaluation of the defense shows it to be an effective means to significantly lower the impact of the vulnerability without imposing extra effort onto the device users.

References

[1] L. Lee, S. Egelman, J. H. Lee, and D. Wagner, "Risk perceptions for wearable devices," *arXiv preprint arXiv:1504.05694*, 2015.

[2] X. Liu, Z. Zhou, W. Diao, Z. Li, and K. Zhang, "When good becomes evil: Keystroke inference with smartwatch," in *Conference on Computer and Communications Security*. ACM, 2015, pp. 1273–1285.

[3] A. Maiti, M. Jadliwala, J. He, and I. Bilogrevic, "(smart) watch your taps: side-channel keystroke inference attacks using smartwatches," in *International Symposium on Wearable Computers*. ACM, 2015.

[4] I. Martinovic, D. Davies, M. Frank, D. Perito, T. Ros, and D. Song, "On the feasibility of side-channel attacks with brain-computer interfaces." in *USENIX Security Symposium*, 2012, pp. 143–158.

[5] A. Migicovsky, Z. Durumeric, J. Ringenberg, and J. A. Halderman, "Outsmarting proctors with smartwatches: A case study on wearable computing security," in *Financial Cryptography and Data Security*. Springer, 2014, pp. 89–96.

[6] I. Muslukhov, Y. Boshmaf, C. Kuo, J. Lester, and K. Beznosov, "Know your enemy: the risk of unauthorized access in smartphones by insiders," in *International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 2013, pp. 271–280.

[7] Google Inc., "Android wear," 2017. [Online]. Available: https://www.android.com/wear/

[8] ——, "Ok google," 2017. [Online]. Available: http://ok-google.io/

[9] N. Karapanos, C. Marforio, C. Soriente, and S. Capkun, "Sound-proof: usable two-factor authentication based on ambient sound," in *USENIX Security Symposium*, 2015, pp. 483–498.

[10] P. P. Parikh, M. G. Kanabar, and T. S. Sidhu, "Opportunities and challenges of wireless communication technologies for smart grid applications," in *Power and Energy Society General Meeting*. IEEE, 2010, pp. 1–7.

[11] N. Semiconductor, "Bluetooth range," 2017. [Online]. Available: http://blog.nordicsemi.com/getconnected/things-you-should-know-about-bluetooth-range

[12] P. A. Zandbergen and S. J. Barbeau, "Positional accuracy of assisted gps data from high-sensitivity gps-enabled mobile phones," *Journal of Navigation*, vol. 64, no. 3, pp. 381–399, 2011.

[13] A. Abu-Mahfouz and G. P. Hancke, "Distance bounding: A practical security solution for real-time location systems," *IEEE transactions on industrial informatics*, vol. 9, no. 1, pp. 16–27, 2013.

[14] T. Halevi, D. Ma, N. Saxena, and T. Xiang, "Secure proximity detection for nfc devices based on ambient sensor data," in *European Symposium on Research in Computer Security*. Springer, 2012, pp. 379–396.

[15] H. T. T. Truong, X. Gao, B. Shrestha, N. Saxena, N. Asokan, and P. Nurmi, "Comparing and fusing different sensor modalities for relay attack resistance in zero-interaction authentication," in *Pervasive Computing and Communications*. IEEE, 2014, pp. 163–171.

[16] G. Kumparak, "Google acquires slicklogin, the sound-based password alternative," 2014. [Online]. Available: http://techcrunch.com/2014/02/16/google-acquires-slicklogin-the-sound-based-password-alternative/

[17] MathWorks, "Butterworth filter design," 2017. [Online]. Available: http://www.mathworks.com/help/signal/ref/butter.html