Vibreaker: Securing Vibrational Pairing with Deliberate Acoustic Noise

S Abhishek Anand
University of Alabama at Birmingham
anandab@uab.edu

Nitesh Saxena
University of Alabama at Birmingham
saxena@uab.edu

ABSTRACT

Pairing between wireless devices may be secured by the use of an auxiliary channel such as audio, visuals or vibrations. A simple approach to pairing involves one of the devices initiating the transmission of a key, or keying material like a short password, over the auxiliary channel to the other device. A successful pairing is achieved when the receiving device is able to decode the key without any errors while the attacker is unable to eavesdrop the key.

In this paper, we focus on the security of the vibration channel when used for the key transmission. As shown in some recent work, sending the keying material over a clear vibrational channel poses a significant risk of an acoustic side channel attack. Specifically, an adversary can listen onto the acoustic sounds generated by the vibration motor of the sending device and infer the keying material with a high accuracy. To counteract this threat, we propose a novel pairing scheme, called *Vibreaker* (a "Vibrating speaker"), that involves active injection of acoustic noise in order to mask the key signal. In this scheme, the sending device artificially injects noise in the otherwise clear audio channel while transmitting the keying material via vibrations. We experiment with several choices for the noise signal and demonstrate that the security of the audio channel is significantly enhanced with Vibreaker when appropriate noise is used. The scheme requires no additional effort by the user, and imposes minimum hardware requirement and hence can be applied to many different contexts, such as pairing of IoT and implanted devices, wearables and other commodity gadgets.

1. INTRODUCTION

The wireless communication (Bluetooth, WiFi or RFID) is easy to eavesdrop and manipulate, and therefore a fundamental security objective is to secure this communication channel. "Pairing" is a term commonly used to refer to the operation of bootstrapping secure communication between two wireless devices, resistant against eavesdropping and man-in-the-middle attacks. Pairing is generally a hard problem due to the lack of a global infrastructure enabling devices to share an on- or off-line trusted third party, a certification authority, a PKI or any pre-configured secrets.

A well-researched approach to pairing is to leverage an auxiliary channel, also called an out-of-band (OOB) channel, which is governed by the users operating the devices. Examples of OOB channels include audio, visual, and vibrational channels. Unlike the radio communication channels, OOB channels are "human-perceptible", i.e., the underlying transmission/reception can be perceived by one or more of human senses. In other words, a user can validate the intended source of an OOB message and an adversary can not manipulate the OOB messages in transit (although he can eavesdrop). Prior research refers to such an authenticated OOB communication as A-OOB [1]. Using these protocols, a multitude of pairing methods based on a large variety of A-OOB channels have been proposed, as surveyed in [3].

Pairing protocols are challenging to implement on constrained devices that lack a good quality output interfaces (e.g, a speaker, display), input interfaces (e.g., keypads), or receivers (e.g., microphone, camera), and may not be physically accessible. Limited computational resources are also a limiting factor for establishing A-OOB channel on such devices. An alternative pairing approach involves the use *secret as well as authenticated* OOB channels (termed AS-OOB [1]). In this approach, the adversary is not only assumed to be incapable of manipulating OOB communication but also can not eavesdrop upon it.

Several prior proposals, including [2, 6], have taken the AS-OOB approach to pairing. The IMD Pairing scheme of [2] uses a low-frequency audio channel to pair an RFID tag – attached to an IMD (Implanted Medical Device) – with an authorized RFID reader. Basically, the tag generates a random key and broadcasts it to the reader which listens to it from a close distance (e.g., a microphone is placed in close proximity to the patient's chest in case of a cardiac implant). The PIN-Vibra method for pairing [6] uses an automated vibrational channel to pair a personal RFID tag with a mobile phone. The phone generates a PIN and transmits it to (an accelerometer-equipped) tag through its vibrations, while the user presses the phone against the tag. (The same channel is later used by the phone to authenticate to (or activate) the tag).

However, both of these pairing schemes have been subject to acoustic eavesdropping attacks and shown to be vulnerable [1]. The work of [1] demonstrated highly accurate attacks on IMD pairing (which uses direct acoustic signals), and PIN-Vibra (in which the acoustic signals are a by-product of the vibration – a side channel). These attacks serve to call the security of the whole AS-OOB model to pairing into question.

In this paper, we set out to enhance the security of the simple AS-OOB approach to pairing by cloaking the acoustic leakage underlying such schemes. In particular, we focus on a representative instance of vibrational pairing, the PIN-Vibra scheme, and aim to obfuscate the sounds created by the vibration motor of the sending device (the phone). The idea is to selectively jam the acoustic chan-

WiSec'16 , July 18-22, 2016, Darmstadt, Germany

DOI: http://dx.doi.org/10.1145/2939918.2939934

nel leakage by having the sending device produce deliberate sounds that would mask the sounds produced by the vibration motor.

Our Contributions: We propose a practical defense mechanism, called *Vibreaker*, against acoustic side channel attacks in vibrational pairing (PIN-Vibra), and evaluate its security against several attack vectors. The main contributions of this paper are summarized below.

1. *Design and Implementation of the Defense*: We build Vibreaker, a viable defense system that utilizes masking signals to mitigate vibration pairing side channel attacks. The defense system is designed to be a part of the device that is the source of acoustic leakage, which would be the phone in our case study. The intuition behind our defense model is to actively cloak the acoustic leakage emanating from the phone's vibration motor with other sounds that would be played back by the phone in the background. In this model, the insertion of acoustic noise only impacts the acoustic eavesdropper but does not at all affect the capability of the device which is receiving and decoding the vibrations (e.g., an RFID tag or possibly another phone).
2. *Evaluation of Security*: We evaluate the security of the Vibreaker system by testing its ability to reduce the accuracy of the acoustic side channel attack that we recreated in the initial step of our research by preventing the adversary from gaining usable information about the transmitted PIN (or any short content). In particular, we test two masking signals: white noise and sounds generated by the vibrations against the acoustic side channel attack. We also examine their performance against noise reduction techniques aimed at filtering out the masking signals. Our results show both types of masking signals to be effective against curbing the existing eavesdropping attack.

2. BACKGROUND AND RELATED WORK

2.1 A-OOB Pairing of Constrained Devices

A-OOB pairing of constrained wireless devices is challenging due to a number of reasons. Several prior pairing methods are based on bidirectional automated device-to-device (dtd) A-OOB channels (e.g., [3]). Such dtd channels require both devices to have transmitters and corresponding receivers (e.g., IR transceivers), which may not exist on constrained devices. In settings, where dtd channel(s) do not exist (i.e., when at least one device does not have a receiver), pairing methods can be based upon device-to-human (dth) and human-to-device (htd) channel(s) instead (e.g., based on transfer of numbers [7]). However, establishing such channels on constrained devices may also not be feasible.

One solution to the above problem is to use only unidirectional communication (from device A to B), but have the user transfer the result of pairing shown on B over to A, as shown in [5]. This, however, may lead to a critical security failure – a user may accept the pairing on A even though B indicates otherwise, as shown via the usability studies in [3]. (This is referred to as a *fatal* human error [3] which translates into a man-in-the-middle attack).

Another possible approach is based on manual comparison of audiovisual OOB strings over synchronized device-to-human (dth) channels, as shown in [3]. This would only require the two devices to be equipped with low-cost transmitters, such as LED(s) (and two buttons). However, the security of these approaches rely upon the decision made by the user and is prone to fatal human errors, as demonstrated in [3]. Even worse, a user who is in a rush to connect her devices may simply "accept" the pairing, without having to correctly take part in the decision process [3].

2.2 PIN-Vibra: AS-OOB Vibrational Pairing

Personal (passive) RFID tags (found, e.g., in access cards, e-passports, licenses) are increasingly becoming ubiquitous. Similar to other personal devices, personal RFID tags often store valuable information privy to their users, and are likely to get lost or stolen. However, unlike other personal wireless devices, such information can be easily be subjected to eavesdropping, relay attacks and unauthorized "reading", and can lead to owner tracking.

User authentication to an RFID device would allow a user to control when and where her RFID tag can be accessed and thus help solve some of the aforementioned problems. A road-block in developing an RFID user authentication mechanism is the lack of any input or output interfaces on RFID tags (RFID devices were not meant to interact with their users and vice-versa) and a somewhat atypical usage model (users often place RFID tags in their wallets and might not be in direct contact with them).

In [6], authors present PIN-Vibra, a novel approach for user authentication to RFID tags. PIN-Vibra leverages a pervasive device such as a personal mobile phone, motivated by its ubiquity. It uses the mobile phone as an authentication token, forming a unidirectional AS-OOB tactile communication channel between the user and her (accelerometer-equipped) RFID tags. Pairing of (and later authenticating to) an RFID tag requires the user to simply touch her vibrating phone with the tag or object carrying the tag (e.g., a wallet); the phone encodes a short PIN into vibrations which are read by the tag's accelerometer and decoded.

The security of PIN-Vibra relies on secrecy of the underlying vibrational channel, i.e., an adversary who is not in close physical contact with the phone should not be able to learn the transmitted PIN. In [1], the authors investigated the feasibility of eavesdropping the PIN-Vibra vibrational channel. In particular, they demonstrated how the acoustic emanations associated with a vibrating mobile phone can be eavesdropped upon from a short distance with off-the-shelf microphones. In Section 3, we will recreate and re-validate this attack, which will serve as a pre-requisite for the evaluation of our proposed defense Vibreaker to defeat this attack.

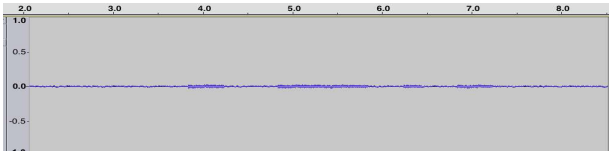
3. PIN-VIBRA ATTACK RECREATION

In this section, we provide the details of the vibration pairing scheme PIN-Vibra and show its vulnerability against an eavesdropping adversary over the audio channel.

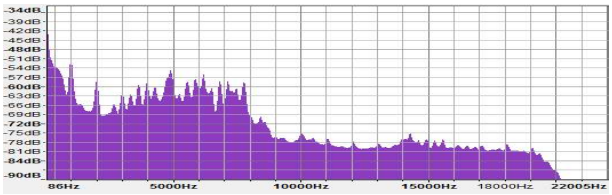
3.1 PIN-Vibra Details

In the vibration pairing scheme, the transmitter (a phone or a smart device) encodes the keying material into a series of vibrations through its vibration motor. The receiver, also a smart device equipped with an accelerometer reads the vibrations and decodes the transmitted information with the help of the accelerometer. For successful decoding of the vibrations by the receiver, both devices need to be in contact with each other.

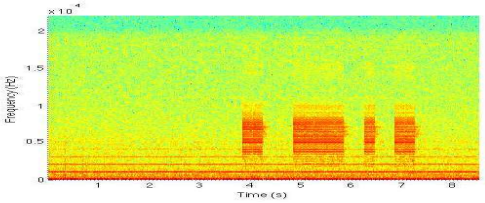
For our work, we implement the scheme proposed in [6]. We use four digit PINs as the data to be transmitted using a simple time interval based ON-OFF encoding mechanism. The PIN is treated as a decimal number and converted to its 14 bit binary equivalent. A preamble "110" is added at the beginning of the binary representation of the PIN to denote the start of the transmission bringing the total bit length of the sequence to 17 bits. Each '1' bit in the bit sequence is encoded into a vibration lasting for 200ms and each '0' bit is encoded as a silence period of 200ms. The total time to transmit a 17 bit sequence is therefore estimated to be $17 \times 200ms = 3.4s$.



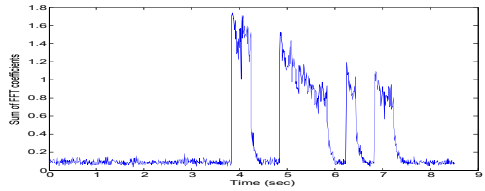
(a) Raw Audio Signal



(b) Amplitude vs Frequency



(c) Spectrogram of the Acoustic Leakage



(d) Sum of the FFT coefficients for Acoustic Leakage

Figure 1: Acoustic Characterstics of the Vibration for PIN "4562"

3.2 Threat Model

We follow a similar threat model as that of [1] where the adversary is able to eavesdrop on the devices in communication from a distance (10cm or more). This implies that while the adversary does not have access to the device's microphone, it can use a covert listening device for eavesdropping. Since the adversary is at a distance from the pairing devices, the listening device can be hidden in the surrounding environment. After recording, the adversary can process the recording offline for decoding the pairing key from the eavesdropped signal.

The environment is supposed to be quiet, devoid of any interfering background noise that lets the adversary eavesdrop on the vibration sounds with the best possible quality. The listening device could be any off the shelf recording devices available in the market that provide a decent recording quality. Thus, we have a realistic eavesdropping scenario in a clean environment with a moderate capability adversary.

3.3 Eavesdropping Attack Model

In any vibration based scheme, the bits are encoded into a series of vibrations. While the vibrations seem to be barely audible, an examination of the audio spectrum of the vibrations from a close distance reveals significant acoustic leakage in a particular frequency band.

For our experiment, we used Motorola Droid X2 android based phones as both the transmitter and the receiver. We also utilized a PC microphone to eavesdrop on the audio produced during bit transmission due to vibrations. Audio processing was done offline using Matlab software.

Figure 1a represents the raw audio signal captured during the vibration of the phone. The figure confirms the assumptions that the amplitude of the acoustic leakage from the vibrations is very low. However, upon examining the frequency spectrum in Figure 1c, the acoustic leakage from the vibrations seems to stretch from 3.5 kHz to 8.3 kHz. The power of the frequency spectrum appears to be consistently high in the frequency range 6.8 kHz to 7.8 kHz in both a normal scenario Figure 1d and on a dampened surface Figure 2. We therefore restrict the attack to this frequency band and use it in our evaluation as it seems to be the optimal one from the point of view of the attacker.

The first step to recover the transmitted data from the acoustic leakage involves the detection of the beginning of the transmission. Hence, we search for the preamble "110" in the eavesdropped audio signal. To detect a valid bit in the audio signal, we transform

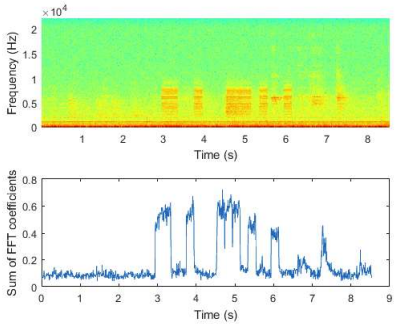


Figure 2: Spectrum Analysis on a Dampening Surface

the recorded audio signal from the time domain to the frequency domain by performing a Fast Fourier Transformation (FFT) of the signal using the *spectrogram* function of Matlab. We used a hamming window of size 441 samples with an overlap of half of the window size. Then, we sum up the FFT coefficients of the signal at each time instance in the frequency band 6.8 kHz - 7.8 kHz. The results obtained are similar to Figure 1d.

Once, we have calculated the sum of the FFT coefficients, we determine a threshold value above which we consider a vibration to have begun. This threshold value can be set as the maximum value of the sum of the FFT coefficients and can be reduced by 10% each time until the audio signal is correctly decoded.

As mentioned in the description of the vibration pairing scheme, each vibration lasts for a period of 200ms. We therefore divide the signal into audio bins of 200ms and find the average power in each bin. If the average power in the bin is more than the pre-defined threshold, we decode the corresponding bit as '1' otherwise '0'. Since the first three bits should be "110" as the preamble for a valid transmission, we continue checking the bits until we get a valid preamble. Once a valid preamble is found, we begin decoding the bits until all remaining 14 bits have been decoded.

We test our eavesdropping attack on ten random PINs using the above setup with the recordings done at a distance of 15cm. The attack was successful with 100% accuracy thereby demonstrating that communication using vibrations is highly susceptible to an acoustic eavesdropping adversary. This result confirms the results of the attack scheme proposed by Halevi et al. [1].

In the above discussed scenario, the two devices are assumed to be in contact during the transmission. We also studied other

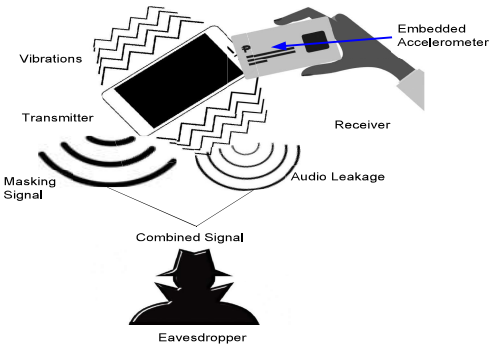


Figure 3: Defense Model Setup

scenarios where the transmitting device may be resting on a wallet (communicating with an RFID tag inside the wallet) or on an audio dampening surface (a thick layering of cloth). The results from the frequency spectrum analysis of the eavesdropped signal (Figure 2) show that even though the surface muffles the vibration sounds, it is still possible to decode the bits by lowering the threshold according to the obtained sum of FFT coefficients.

4. DEFENSE

In the previous section, we described an eavesdropping attack on pairing through vibrations. Now, we will briefly explore some potential defense mechanisms that could be used to mitigate this class of attack and follow it with the design of Vibreaker, the defense scheme proposed in this paper.

4.1 Defense Background

The acoustic side channel attacks on vibration pairing exploit the leaked audio that is generated due to the vibrations of the device. In order to prevent an adversary from gaining any useful information from the audio leakage, the vibration sound should be either canceled or masked with the help of another sound. In case of masking, the adversary has to filter out the masking signal in order to recover the audio signal and extract the keying material from it.

Signal cancellation usually involves introducing another signal having similar features (frequency characteristics and amplitude level) as the signal to be canceled but having opposite phase. Signal masking requires introducing a signal that contains all the frequencies of the signal to be masked and an amplitude level equal to or more than that of the signal to be masked. With the masking signal, we try to distort the original signal to an extent that the resultant signal has entirely different features from the original signal.

4.1.1 Audio Leakage Cancellation

Roy et al. [4] examined the possibility of canceling the sound of vibration (termed SoV) by creating an “anti-noisc” signal on the transmitter’s end. The challenges of creating an anti-noise signal involve an estimation of the surface in contact with the phone, the phase of the SoV signal and the limited real-time audio processing capabilities offered by the Android platform.

For estimating the effect of the surface on which the device has been placed, they transmitted a short preamble and recorded the resulting SoV. The FFT of the SoV was examined to look for the strongest overtones that are combined to create the "anti-noise" signal. For phase alignment, Ripple transmitter increases the sampling frequency of the "anti-noise" signal keeping track of the phase difference of the "anti-noise" and SoV and switching it back to its original value when the phase difference is minimum.

There exist multiple issues with this approach if we try to implement the scheme in our communication model. Our model is focused on short message exchange between two devices like pairing or authentication. Hence, a simple ON/OFF scheme is sufficient for all of our purposes. Since we implement our model on Android based smartphones, computationally exhaustive signal processing tasks such as calculating FFT for creating an "anti-noise" signal take more time than the entire duration of communication that is as mentioned in Section 3 is only 3.4s.

After combining the "anti-noise" signal with the SoV, the frequency domain of the reduced sound still remains unchanged though the amplitude is reduced significantly. However, a powerful adversary could scrutinize the frequency spectrum of the reduced sound looking for frequency footprints of the SoV. This can potentially reveal some information about the transmitted bits to the adversary.

If we do not consider the duration of the communication as a limiting factor by artificially increasing it via an addition of a preamble to the actual PIN, cancellation of audio signal may yet prove to be capable of mitigating the acoustic side channel attack. However, in this work, we restrict ourselves to the examination of easy to generate and computationally light signal masking technique.

4.1.2 Audio Leakage Masking

Signal masking mechanism borrows its motivation from a very common problem in signal communications where the presence of noise in the environment corrupts the signal. If the signal to noise ratio (SNR) is low, it becomes hard to differentiate the signal from the encompassing noise. We utilize this idea to intentionally introduce noise (referred as masking signal) during the vibration of the device so that it corrupts the audio leakage from the vibration to an extent that it becomes indistinguishable from the masking signal.

As the strength of the defense mechanism depends on the difficulty of the adversary’s task in filtering out the masking signal from the eavesdropped signal, we test out some types of masking signal that can be deployed to defend against eavesdropping attack. We also evaluate the effectiveness of these signals in masking the audio leakage and the difficulty of filtering these signals from the eavesdropped signal to recover partial or full audio signal (SoV).

4.2 Vibreaker Design and Setup

Vibreaker is designed to generate a masking signal that obscures the audio leakage in a fashion that makes it hard for the adversary to extract any information about the transmitted data. In order to accomplish this task, we test different types of sounds that could potentially be the masking signal and evaluate their security against an adversary as defined in Section 3.

In our setup for testing Vibreaker, the transmitter and the receiver are positioned similar to the attack recreation model where the devices are in contact. The transmitter transmits data by vibrating in a certain pattern that is decoded by the receiver. In addition, the transmitter is also equipped with speakers that emit the masking signal while the transmission is in progress. This ensures that any eavesdropping adversary will receive the combined signal (a mix of the audio leakage from vibrations and the masking signal) from which it would be difficult to recover the transmitted information. Figure 3 details the overview of the Vibreaker model.

5. EVALUATION OF THE DEFENSE

In this section, we will evaluate the efficiency of some types of masking signals against the attack setup described in Section 3. We will also investigate the prospect of filtering out the masking signals by the adversary and test if the resultant signal contains any relevant information about the transmitted data.

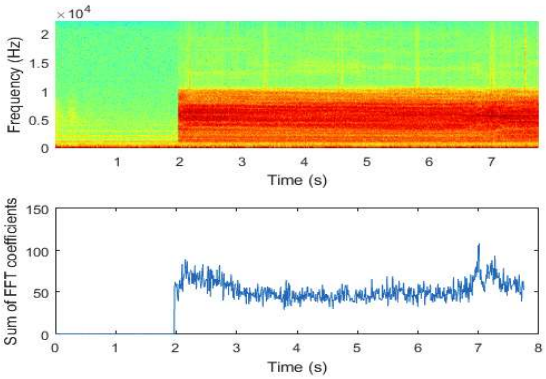


Figure 4: White Noise Masking Spectrum

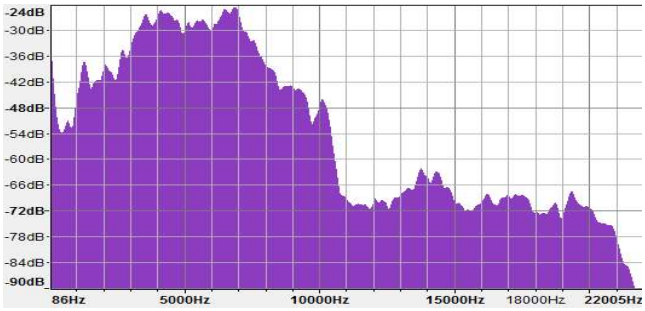


Figure 5: Amplitude vs Frequency for the White Noise

5.1 White Noise as Masking Signal

White noise is defined as a random signal having a constant power spectral density. White noise is constantly present in the environment for example the humming sound emanating from air conditioning units. It has also been used for sound masking in offices by suppressing other distracting sounds. Here, we use the white noise as the base level candidate signal for masking. It is not a sophisticated signal and can easily be generated. Filtering the white noise signal is fairly simple, but the process of filtering also affects the quality of the recovered signal. Since white noise has an equal distribution over all of the frequency spectrum, trying to filter it out also filters out the frequencies where the white noise overlaps with the frequency spectrum of the original signal (audio leakage from the vibrations).

Experimental Setup: We use the *wgn* function of Matlab to generate a 10 second sample of white Gaussian noise at a sampling frequency of 44.1 kHz. White Gaussian noise is a good approximation of real world white noise and hence sufficient for our intentions. We then apply a frequency filter to the noise sample to make sure that the white noise remains in the same frequency band as the audio leakage.

Once we have generated the white noise sample, we play it in the background while the phone vibrates. To make sure that the white noise suppresses all the audio leakage, we introduce a delay in the phone vibration at the beginning such that the phone starts vibrating only after the white noise has begun playing in the background.

Observations: To study the effectiveness of the white noise as a masking signal, we consider an adversary snooping at a distance of 15cm using the attack recreated in section 3.

Our observations for the recording done at a distance of 15cm (Figure 4) show that white noise completely masks the audio leakage from the vibrations. In addition, nothing can be learned about the vibrations in the frequency domain even after obtaining FFT of the eavesdropped signal. Apart from covering the spectrum in which the audio leakage from the vibrations lie, the sound level of the white noise as shown in Figure 5 is more than twice than that of vibrations alone (Figure 1b) thereby easily suppressing the leakage.

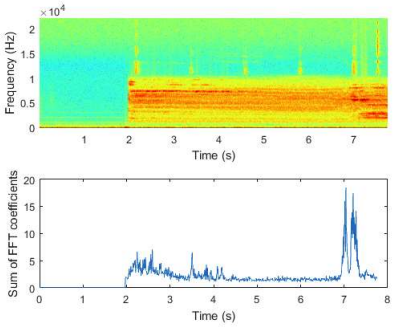


Figure 6: Signal Features for PIN "4562" after filtering the White Noise

Filtering the Masking Signal: For filtering the white noise, we use the *noise reduction* effect from Audacity software that allows us to select a small sample of noise as the noise profile and apply it to the whole signal for noise reduction. We used a noise reduction level of 15dB and a sensitivity value of 6 to get the best results for our scenario. Figure 6 shows the frequency spectrum of the captured audio signal after filtering out the noise.

The frequency spectrum in Figure 6 does not reveal any information about the audio leakage from the vibration. This effect is due to complete cloaking of the audio leakage by the white noise. Since vibration sounds are not loud, the attenuation of the audio leakage at the attacker's eavesdropping device makes it infeasible to extract any information from the recorded signal. In Figure 6, the plot of the sum of FFT coefficient vs time does not contain any vibration peaks. Hence, our attack would be unable to identify the vibration periods leading to a 0% success rate.

5.2 Vibration Noise as Masking Signal

Our next choice of masking signal is a close representation of the audio leakage itself. We pre-record a clip of the sound generated during the vibration and try to confuse the attacker by masking the audio leakage from the vibrations with the pre-recorded vibration noise (henceforth referred as fake vibrations).

Experimental Setup: We generate a random sequence of numbers and encode them as vibrations using the same protocol as PIN-Vibra [6]. However, in order to make sure that the fake vibrations completely overlap with the actual vibrations, we reduce the duration of silence from 200ms to 100ms between the vibrations. The resultant vibration sequence is recorded offline and stored for use as the masking signal.

When the user initiates the protocol for sending the PIN via vibrations, the device in addition to vibrating also begins playing the stored masking signal in the background. We adjust the timings of the masking signal such that it always begins playing at the approximately the same time as the vibrations. The adversary is presumed to be eavesdropping at a distance of 15cm.

Observations: The results show that fake vibrations are able to mask the audio leakage from the device's vibration. It is nearly impossible to distinguish between the fake vibration signals and the audio leakage by looking at the frequency spectrum. The FFT measure also shows only the response from the fake vibrations indicating that audio leakage has completely been masked.

Filtering the Masking Signal: We apply the same filtering process that was used for filtering out the white noise. Since sounds of fake vibration differ from actual vibration sound due to imper-

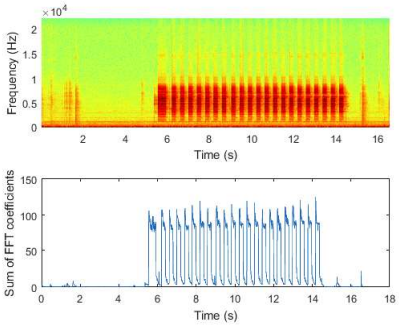


Figure 7: Signal Features for PIN "4562" in presence of Fake Vibrations

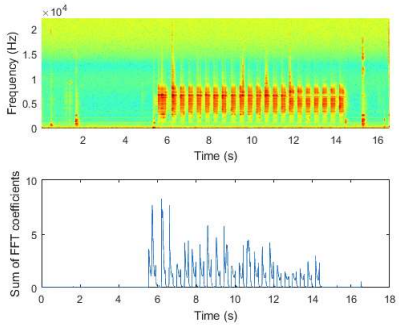


Figure 8: Signal Features for PIN "4562" after filtering Fake Vibrations

fect reproduction by the speakers, we (as an attacker) listen to the eavesdropped audio signal and select the part that is believed to be the fake vibrations. The selected part of the audio is used as the noise profile and applied to the full length of the eavesdropped audio signal for filtering the fake vibrations.

The results (Figure 8) after the filtering process reveal no additional information about the audio leakage from the vibration of the device. In the plot of the sum of FFT coefficient vs time (Figure 8), there does not exist a threshold that could differentiate between real and fake vibrations. Hence, our attack fails to decode the correct PIN leading to 0% success rate. Thus fake vibrations also serve a an efficient masking signal for obfuscating the vibration sounds.

Effect on the Receiver: We believe that the proposed defense mechanism does not affect the communication between the pairing devices because the the data transfer is through vibrations. Vibreaker relies on audio signal for masking the audio leakage. Receiver is using accelerometer to learn the real vibrations, fake vibrations is just sound and, as mentioned above, would have no effect on accelerometer readings.

6. CONCLUSION AND FUTURE WORK

In this paper, we explored possible ways to mitigate an acoustic side channel attack on vibration based pairing interactions. In particular, we focused our investigation on PIN-Vibra instance of vibrational pairing [6] and recreated the side channel attack on this pairing model, as earlier reported in [1].

We introduced Vibreaker, a novel defense mechanism to mitigate acoustic side channel attacks against PIN-Vibra by active injection of masking signal in the environment. The purpose of the masking signal was to obfuscate the acoustic leakage generated by the vibrations of the transmitting device. We studied some audio signals as candidates for masking the acoustic leakage and evaluated the security offered by them against the recreated attack.

Our results showed that both white noise and fake vibration sounds offer viable security against an adversary eavesdropping on the acoustic side channel leakage. Both types of masking signals were able to hide the acoustic leakage from an eavesdropping adversary making it difficult to distinguish between the masking signal and the acoustic leakage. We also studied the effect of noise filtering mechanism against Vibreaker and found out that even if the adversary tries to filter out the masking signal using noise reduction technique, it may not help the adversary in recovering any useful information from the eavesdropped signal.

A possible avenue for future work would be to evaluate the performance of Vibreaker against more powerful attackers who may use triangulation techniques based on multiple, strategically placed, audio recording devices or more powerful and direction oriented microphones. We believe that Vibreaker may be able to thwart such attacks due to the sources of audio leakage (vibration motor) and audio noise (speakerphone) being embedded to the same device (transmitter), essentially very close to each other. Due to the same reason, it would be viable to generate the masking signals at the receiving device (rather than the transmitting device). This may be useful when the transmitting device does not have a speaker while the receiving device is equipped with one.

However, while using the FFT features of the audio leakage was enough to create a successful attack in the absence masking signals, Vibreaker would also need to be tested against other sophisticated attacks that may utilize machine learning and feature classification (using FFT or MFCC) to detect vibration sounds in the eavesdropped signal. Another consideration may be to develop a stricter attack model where the distance of attacker is 0cm from the transmitter, or the attacker has a very powerful microphone, e.g., a parabolic microphone.

While Vibreaker transmission process is very short (less than 4 seconds), insertion of masking sounds may have an impact on the usability of the process. We believe that the use of fake vibration sounds may be less distracting to the users compared to white noise, since fake vibration sounds aim to match the sounds of the vibration itself. A formal usability study may need to be conducted to evaluate and compare the distraction effects of various masking signals in the Vibreaker system.

7. REFERENCES

[1] T. Halevi and N. Saxena. On pairing constrained wireless devices based on secrecy of auxiliary channels: the case of acoustic eavesdropping. In *ACM CCS*, 2010.

[2] D. Halperin, T. S. Heydt-Benjamin, B. Ransford, S. S. Clark, B. Defend, W. Morgan, K. Fu, T. Kohno, and W. H. Maisel. Pacemakers and implantable cardiac defibrillators: Software radio attacks and zero-power defenses. In *IEEE Symposium on S&P*, 2008.

[3] R. Kainda, I. Flechais, and A. W. Roscoe. Usability and security of out-of-band channels in secure device pairing protocols. In *SOUPS*, 2009.

[4] N. Roy, M. Gowda, and R. R. Choudhury. Ripple: Communicating through physical vibration. In *NSDI*, 2015.

[5] N. Saxena, J.-E. Ekberg, K. Kostiainen, and N. Asokan. Secure device pairing based on a visual channel. In *IEEE S&P*, 2006.

[6] N. Saxena, M. B. Uddin, J. Voris, and N. Asokan. Vibrate-to-unlock: Mobile phone assisted user authentication to multiple personal RFID tags. In *Percom*, 2011.

[7] E. Uzun, K. Karvonen, and N. Asokan. Usability analysis of secure pairing methods. In *USEC*, 2007.