Wearable Microphone Jamming

Yuxin Chen* **Huiying Li*** **Shan-Yuan Teng*** **Steven Nagels** **Zhijing Li**
Pedro Lopes **Ben Y. Zhao** **Haitao Zheng**
University of Chicago
{yxchen, huiyingli, tengshanyuan, stevennagels, zhijing,
pedrolopes, ravenben, htzheng}@cs.uchicago.edu
* *equal contribution*

ABSTRACT

We engineered a wearable microphone jammer that is capable of disabling microphones in its user's surroundings, including hidden microphones. Our device is based on a recent exploit that leverages the fact that when exposed to ultrasonic noise, commodity microphones will leak the noise into the audible range.

Unfortunately, ultrasonic jammers are built from multiple transducers and therefore exhibit blind spots, i.e., locations in which transducers destructively interfere and where a microphone cannot be jammed. To solve this, our device exploits a synergy between ultrasonic jamming and the naturally occurring movements that users induce on their wearable devices (e.g., bracelets) as they gesture or walk. We demonstrate that these movements can blur jamming blind spots and increase jamming coverage. Moreover, current jammers are also directional, requiring users to point the jammer to a microphone; instead, our wearable bracelet is built in a ring-layout that allows it to jam in multiple directions. This is beneficial in that it allows our jammer to protect against microphones hidden out of sight.

We evaluated our jammer in a series of experiments and found that: (1) it jams in all directions, e.g., our device jams over 87% of the words uttered around it in any direction, while existing devices jam only 30% when not pointed directly at the microphone; (2) it exhibits significantly less blind spots; and, (3) our device induced a feeling of privacy to participants of our user study. We believe our wearable provides stronger privacy in a world in which most devices are constantly eavesdropping on our conversations.

CCS Concepts

•**Human-centered computing** → **Interaction devices;** *Sound-based input / output;* Ubiquitous computing;

Author Keywords

Wearable; microphone; jamming; privacy; ultrasound

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376304

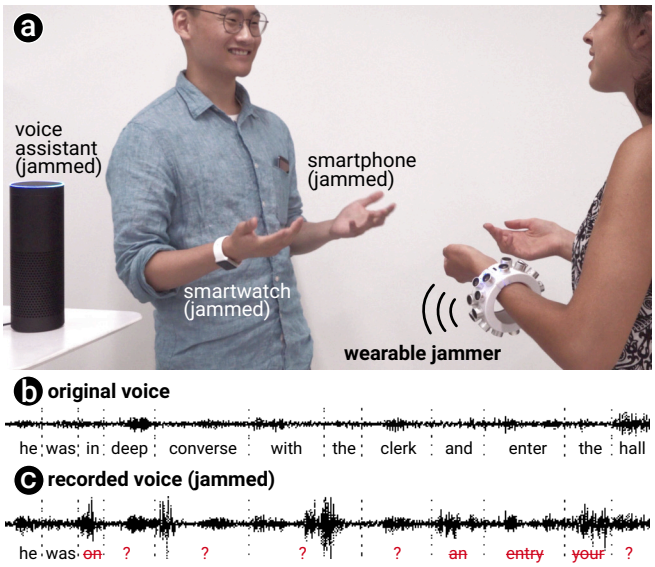

Figure 1. (a) We engineered a wearable ultrasound jammer that can prevent surrounding microphones from eavesdropping on a conversation. (b) This is the actual speech that the conversation partner hears, since our jammer does not disrupt human hearing. However, (c) is the transcript of what a state-of-the-art speech recognizer makes out of the jammed conversation.

INTRODUCTION

Despite the initial excitement around voice-based smart devices, consumers are becoming increasingly nervous with the fact that these interactive devices are, by default, *always* listening, recording, and possibly saving sensitive personal information [32, 55, 38, 26]. Take digital voice assistants, which are featured in most smartphones, smartwatches, and smart speakers, as an example. From the outside, these interactive assistants appear to only respond to designated wake-up words (e.g., "Alexa" and "Hey Google"). However, their implementation requires them to listen continuously to detect these wake-up words. It has been shown that these devices can monitor and record sounds and conversations in real time, either maliciously [56], by misconfiguration [26], or after compromise by attackers [51]. Leaked audio data can be processed to extract confidential information [56, 16, 15], track user activity [8], count human speakers [57], or even extract handwriting content [58]. These negative implications on users' security and privacy are significant and unacceptable. To make matters worse, many other acoustic attacks

(e.g., turning speakers into microphones [25], inferring the content of a printed page by recording its printing [9], inferring a 3D object's geometry by recording its printing [22], inferring typed text by listening to key presses [6, 61]) as well as many forms of espionage (e.g., industrial espionage [17, 28]) rely on eavesdropping via hidden microphones.

Therefore, it is critical to build tools that protect users against the potential compromise or misuse of microphones in the age of voice-based smart devices. Recently, researchers have shown that ultrasonic transducers can prevent commodity microphones from recording human speech [46]. While these ultrasonic signals are imperceptible to human ears, they leak into the audible spectrum after being captured by the microphones, producing a jamming signal *inside* the microphone circuit that jams (disrupts) voice recordings. The leakage is caused by an inherent, nonlinear property of microphone's hardware. Not only have researchers built prototypes using ultrasonic speakers [46], but also these jammers are currently commercially available to the public. However, all these devices exhibit two key limitations: (1) They are heavily directional, thus requiring users to point the jammer precisely at the location where the microphones are. This is not only impractical, as it interferes with the users' primary task, but is also often impossible when microphones are hidden. (2) They rely on multiple transducers that enlarge their jamming coverage but introduce *blind spots*—locations were the signals from two or more transducers cancel each other out. Such blind spots occur especially in close proximity to the jammer; in fact, 17% of all locations within 1.2m of a typical multi-transducer jammer are blind spots. If a microphone is placed in any of these locations it will not be jammed, rendering the whole jammer obsolete.

To tackle these shortcomings, we engineered a wearable jammer that is worn as a bracelet, which is depicted in Figure 1. By turning an ultrasonic jammer into a bracelet, our device leverages natural hand gestures that occur while speaking, gesturing or moving around to blur out the aforementioned blind spots. Furthermore, by arranging the transducers in a ring layout, our wearable jams in multiple directions and protects the privacy of its user's voice, anywhere and anytime, without requiring its user to manually point the jammer to the eavesdropping microphones.

We confirmed that an ultrasonic microphone jammer is superior to state-of-the-art and commercial stationary jammers by conducting a series of technical evaluations and a user study. These demonstrated that: (1) our wearable jammer outperformed static jammers in jamming coverage; (2) its jamming is effective even if the microphones are hidden and covered by various materials, such as cloths or paper sheets; and, (3) in a life-like situation our study participants felt that our wearable protected the privacy of their voice.

BACKGROUND AND RELATED WORK

Our work builds on top of ultrasonic emitters and wearables. Also, we discuss the implications of data leaks in interactive devices, especially those with microphones and cameras. Lastly, we introduce the underlying ultrasonic jamming principle that our device is based on.

Privacy issues with interactive devices

As various interactive devices being deployed into daily life, privacy issues arise since these devices often rely on constant capturing of multimedia, such as photos, videos, or sound, in order to provide services that assist users' activities [11, 42].

Researchers have proposed privacy-aware methods to collect user's data by, for instance, designing improved notifications [36] or exploring configurations that are privacy-conscientious [2]. These approaches are, however, developer-centric and thus require that the user trusts the interactive system. The result is that these approaches are beneficial but not a fail-proof solution, as devices are still exposed to attackers. Lastly, these solutions do not seek to empower the users to actively protect their privacy.

For example, as the privacy implications of cameras grew in importance, webcams started to use lights that indicate their recording state [43]. However, these indicators can be disabled by attackers [13], which led to many users opting for physically covering up the webcams [31].

More recently, digital assistant devices, such as Amazon Echo, have become very popular due to their interactive (conversational) ability. These interactive devices are built with a microphone and a speaker. To interact with the user when needed, these voice assistants are designed to respond to designated wake-up words (e.g., "Alexa" and "Hey Google"). However, continuously listening is required to detect these wake-up words—this has resulted in many worldwide security breaches, where it was found that these devices leaked or saved sensitive personal information from their users [32, 55, 38, 26]. It was shown that these devices can monitor and record all voices, sounds and conversations in real time, either maliciously [56], by misconfiguration [26], or after compromise by attackers [51]. The leaked audio can be further processed to extract confidential information [56, 16, 15], track user activity [8], count human speakers [57], and so forth. One sane option is certainly to turning these devices off one by one. Unfortunately, that still leaves eavesdropping devices that the user cannot control or that the user is simply not aware of. Instead of turning off all the devices manually, microphone jammers aim at empowering users with a tool to disrupt (jam) voice recordings whenever and wherever they want, providing a physical layer of privacy on demand.

Principles of ultrasonic microphone jamming

Recent work has demonstrated the feasibility of using ultrasonic transducers to disable nearby microphones [46]. The advantage of jamming by means of ultrasound is that it is "silent" to users, as ultrasound is inaudible to humans. We illustrate this type of jamming in Figure 2. Ultrasonic jamming is possible because these higher-frequency signals, after being captured by the microphone's non-linear diaphragm and power-amplifier, will create a lower-frequency "shadow" that happens to be in the microphone's filtering range—the audible range [46]. This technique works against billions of commodity microphones (found in phones, laptops, voice assistants, etc.), without any microphone modification. The fundamental exploit is due to the fact that acoustic amplifiers are only linear around the audible frequency range, while outside

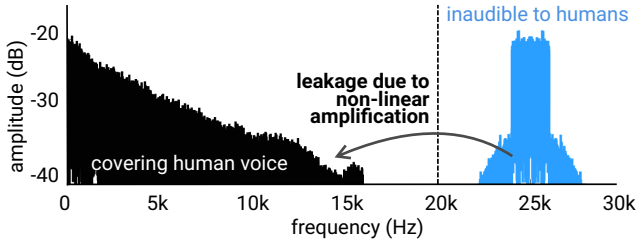

Figure 2. Working principle behind ultrasonic jamming (similar to [46] but here at the example of a 25kHz signal). Here, we depict how an ultrasonic jamming signal (shown in blue), which is inaudible to humans, still leaks into the recorded speech due to non-linear amplification of the microphone's circuit. The result is that the leaked signal covers up precisely the spectrum in which a user's voice is recorded (shown in black).

of the range (e.g., ultrasound), the amplifier's response exhibits non-linearities [46, 1]. This leakage from ultrasound to audible range adds so much audible noise on the microphone circuitry that it effectively renders voice recordings unusable.

Leveraging microphone non-linearity

This non-linearity in microphone circuitry was originally discovered by musicians and leveraged for sound synthesis [29]. Only more recently, have researchers leveraged these non-linearities as a potential tool for setting up hidden communication channels, disabling microphones, or as an adversarial avenue for injecting hidden voice commands. A series of projects leveraged this property to attack digital voice assistants [60, 51, 47]. Here, an adversary can play (arbitrary) voice commands modulated in the ultrasonic range to digital assistants and force these devices to decode them as normal voice commands. Since the original ultrasonic command is inaudible, the attacker can successfully issue commands without being detected (i.e., heard) by nearby users.

Similarly, *backdoor* [46] leverages non-linearity to build an inaudible communication channel among devices and to jam microphones. The *backdoor* jams based on either amplitude modulation (AM) or frequency modulation (FM). *backdoor* was tested in a limited set of experiments with the jammer pointing to a single microphone. Nowadays, there are commercial ultrasonic jammers, such as the *i4*. Unfortunately, although all of them are large, bulky (0.38kg–5kg), and pricey ($799–$6900) [21, 20, 39, 27]. These jammers have also a limited angular coverage and require the users to point directly at the microphone. This is disadvantageous in that these jammers require user's attention to operate and cannot be used against hidden microphones. Inspired by these devices, we propose a novel approach that, instead, leverages the advantages of a wearable design to enhance jamming effectiveness.

Wearable devices based on ultrasound

Researchers have used signals in ultrasonic bands [7, 35] and near-ultrasonic bands (e.g., 18.8kHz) [14, 24] to enable interaction with/among devices. As an example, Gupta et al., utilize Doppler shifts in emitted ultrasound to enable a laptop to perform gesture tracking [24]. A variety of smartphone applications use ultrasonic signals as beacons to perform device localization and tracking [5, 53, 23], again based on the aforementioned leakage to the audible band.

A WEARABLE JAMMER BRACELET

We engineered a microphone jammer in a wearable form factor, which effectively jams in more directions around the user than existing approaches. To assist the reader in replicating our device, we describe the implementation details and the key design elements that enabled our wearable jammer to outperform existing jammers.

We designed our wearable jammer as a bracelet so that it can be easily activated [59, 54, 19, 41, 3] whenever the user decides to engage in a private conversation. Having the device at users' reach at all times provides them with "always available input" [49], ensuring the user is the one in control.

Implementation

To help readers replicate our design, we now provide the necessary technical details. Furthermore, to accelerate replication, we provide all the source code, firmware, and schematics of our implementation [1].

Our prototype, which is depicted in Figure 3, is a self-contained wearable device comprised of the following components: a 3D-printed 9cm ring (outer diameter) with a slit that acts as a hinge, allowing the wearer to open up the bracelet and fit it around their arm; 23 ultrasound transducers (NU25C16T-1, 25kHz), featuring 12 on the lower ring and 11 on the top ring (one transducer was removed to make space for the aforementioned hinge); a low-power signal generator (AD9833, up to 12.5MHz with 0.004Hz programmable steps); an ATMEGA32U4 microprocessor; an LED status indicator; a tactile switch (not shown); a LiPo battery (3.7V, 500mAh); a 3W audio amplifier (PAM8403), and, a 3.7V to 5V step-up regulator. Our microprocessor controls the signal generator via Serial Peripheral Interface.

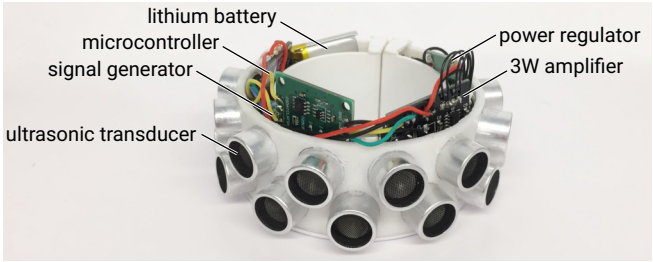

Figure 3. Our prototype is a self-contained wearable comprised of ultrasonic transducers, a signal generator, a microcontroller, a battery, a voltage regulator and a 3W amplifier.

Signal generation

We generate our ultrasound jamming signal via the AD9833 sine wave generator. This integrated circuit (IC) produces a pure sine wave at a desired frequency up to 12.5MHz. To select our sine wave's frequency, we control the AD9833 using our microcontroller via SPI[2]. In order to jam effectively, we produce not only one frequency but a range of frequencies. According to the principles of ultrasonic jamming, each of these will produce a shadow at an audible frequency; therefore, using multiple frequencies enhances jamming. We implement our signal by sweeping the frequency

[1] http://sandlab.cs.uchicago.edu/jammer
[2] https://github.com/Billwilliams1952/AD9833-Library-Arduino

of the sine wave randomly between 24kHz to 26kHz (i.e., 25kHz±1kHz) in steps of 1Hz. Our sine wave frequency changes every 0.45 ms. In our earlier designs, we employed a 92kHz wave player IC that played back the white noise (25kHz±1kHz). However, we found that via empirical testing that our randomly-sweeping sine wave yielded the same jamming power with significantly less power consumption than the overly complex wave-player IC. Lastly, we amplify our signal using a 3W amplifier (PAM8403). Note that we set the amplifier to operate below maximum amplification. This reduces our power consumption and preserves signal quality (low distortion). When measured directly at any of the transducers, the loudness of our device is around 92.3dBA.

Wearable characteristics: power and weight

We measured the energy consumption of our prototype bracelet. It consumes approximately 0.47W (3.7V $\times$ 127mA) when jamming, which is ten times less energy than that used by the commercially available *i4* jammer. Thus it can continuously jam for around four hours on our 500mA battery. Furthermore, our device and battery weigh 135 grams.

Key design elements

Our device was designed with three key elements that allow it to outperform state-of-the-art microphone jammers.

1. Multi-directional jamming using a ring layout. Existing microphone jammers, such as *backdoor* and *i4*, embed their ultrasonic transducers in a flat (1D or 2D) layout. As a result, these jammers are effective *only* when the user points them to the target microphone. This is disadvantageous as: (1) it requires the user to steer the device, making the jamming action a primary task; and, (2) it is practically impossible to use against hidden microphones. Instead, our prototype features all its ultrasonic transducers in a ring layout, effectively enabling jamming in *multiple directions* on a plane. We will later demonstrate that our design is superior by means of both simulations and experimental evaluations.

2. Reducing blind spots by leveraging naturally-occurring movements. A significant benefit of proposing a microphone jammer as a wearable device is that we can mitigate the traditional blind spot problem, which affects all transducer arrays, by leveraging naturally-occurring movements. While a user is wearing our jammer, the device is, most of the times, being moved as the user walks, gestures, points, types, etc. It is precisely these movements that we leverage to reduce blind spots, because as the device moves in space the signal emission map moves accordingly and creates new areas of increased signal strength that blur out the blind spot areas.

3. Collocation with the user's voice. The last design element that makes a wearable design superior is its ubiquitousness. A wearable jammer is collocated with the user that it protects, whereas stationary jammers need to be installed or moved around in every space the user inhabits. Furthermore, the short distance between the jammer and the speaker's mouth prevents the use of beamforming microphone arrays to separate the signals of the human speaker and the jammer [4], making the wearable jammer a stronger defense.

OVERVIEW OF EXPERIMENTS AND STUDY

In order to validate that wearable microphone jammers outperform existing approaches, we conducted **simulations** and three **experimental evaluations**. Lastly, to understand how participants perceive the effectiveness of our wearable jammer, we conducted a **user study**. To aid the reader in understanding the different validations we performed, we present an overview of our simulations, experiments and study:

1. Simulating jammer layouts. Prior to designing our jammer, we confirmed by means of simulation that a wearable bracelet with a ring-layout reduces blind spots when compared to stationary jammers with planar-layouts. To do so, we simulated the power of an ultrasonic signal in space after it leaves the transducer. With our simulations we found that (1) jammers with transducers in a planar layout jam mostly in one direction; (2) on the contrary, positioning the transducers in a ring layout increases jamming in multiple directions; and, (3) adding small (simulated) movement, which occurs naturally in a wearable device, results in a blind spot reduction, similar to what can be achieved using more complex control techniques with multi-frequency signals. This finding is critical because: (1) it allows us to keep the device's design and circuit simple (i.e., using a single signal source), which reduces power consumption, making it compatible with a wearable form factor; (2) our approach does not sacrifice jamming quality when compared to a more complex and hardware heavy approach (i.e., using multiple signal sources). These findings informed how we created our prototype, which we used in all subsequent experiments and user study.

2. Experiment#1: angular power distribution. We measured the angular power distribution of our wearable jammer and both existing devices (a planar jammer with 9 transducers and the commercially available *i4*). We found that our device provides a wide-spread angular coverage ($M = -3.3dBA$, $SD = 1.6dBA$), while the existing jammers are highly directional (planar jammer: $M = -19.2dBA$, $SD = 8.5dBA$; *i4*: $M = -17.0dBA$, $SD = 6.8dBA$).

3. Experiment#2: jamming speech recognizers. We measured how effectively our wearable device jams speech recognizers at different angles, when compared to a planar jammer and *i4*. We found that our wearable device jams more effectively in multiple directions with an increased word error rate (WER) when compared to the other jammers (our wearable: $M = 96.59\%$ WER, $SD = 3.97\%$; planar jammer: $M = 38.89\%$ WER, $SD = 21.72\%$; *i4*: $M = 57.55\%$ WER, $SD = 35.04\%$).

4. Experiment#3: jamming microphones covered by everyday materials. We evaluated how our wearable jams microphones that are covered with everyday materials (i.e., hidden microphones inside boxes, behind clothes, etc.); this stems from a unique feature of our device as it does not require pointing to the target microphone. We found that our device jams microphones hidden under a variety of objects, such as ordinary cloths, foam-based microphone windshields or paper sheets, with a word error rate above 97%.

5. User study. Lastly, we evaluated whether wearing our jamming bracelet impacted participants' perception of privacy. In our study, we asked groups of participants to engage in life-like conversations while they wore the bracelet one at a time. We found that participants felt our wearable protects their privacy (M=5.4 out of 7, SD=1.1).

SIMULATING JAMMER LAYOUTS

Prior to designing our jamming bracelet, we explored, via simulations, whether a wearable would be beneficial. We were interested in answering three questions: (1) how directional are jammers based on planar transducer layouts (e.g., *i4*)?; (2) how do blind spots affect a jammer with its transducers in a ring layout?; and, (3) how do the blind spots behave with respect to small movements of the jammer? All the following simulations were conducted using Matlab. For researchers interesting in replicating our simulations we provide their source code[3].

Simulation parameters

Generally speaking, our simulation computed the propagation of ultrasound from our sources to all points around the device. To model the directivity of our transducers, we utilized the piston model [37, 34] as a good approximation to the pattern supplied by the manufacturer's datasheet[4]. Our transducers are designed to operate at a central frequency of 25kHz, and our control technique sweeps the frequency of a sine wave randomly between 24kHz to 26kHz in steps of 1Hz, every 0.45ms. To simulate multiple signal sources, different random seeds are used in the generation of random frequency sweeping of each source. To simulate a planar jammer, we took the 3×3 array design by [46, 47], which features 9 transducers in a 3×3 planar grid. For the ring-layout, the transducers were placed in a diameter of 11cm. Our simulation runs on a 96kHz, i.e., larger than the Nyquist rate for 25kHz. Lastly, our simulation does not account for reflections.

Simulation algorithm

Our simulation algorithm is based of Morales et al. [37] and Marzo et al. [34]. Let S be the transducers in a jammer, with each transducer $s \in S$. Transducers are modeled as a piston source of radius $r = 8.2mm$. Let T be the time sampled in the simulation, with each time step $t \in T$. $\mathbf{P}_{ref}$ represents the transducers reference pressure; k is the wavenumber ($k = \omega/c_0$); $\mathbf{d}(p, p_s)$ is the distance between the transducer and the point; θ is the angle between the transducer's normal and the point; J_1 represents a Bessel function of the first kind, and $f_s(t)$ represents the signal transmitted by s at time t.

Given our transducer's model, the complex acoustic pressure $\mathbf{P}_{s,t}(p)$ contributed by each transducer s at a given position p and time t is computed as:

$$\mathbf{P}_{s,t}(p) = \frac{\mathbf{P}_{ref}}{\mathbf{d}(p,p_s)} \cdot \frac{2 \cdot J_1(k \cdot r \cdot sin\theta)}{k \cdot r \cdot sin\theta} \cdot \mathbf{f}_s(t - \frac{\mathbf{d}(p,p_s)}{c_0}) \quad (1)$$

The total far field generated by all the transducers at time t can be computed as the summation of the contribution of each individual transducer $\mathbf{P}_t(p) = \sum_{s \in S} \mathbf{P}_{t,s}(p)$. And the average far field generated over time can be computed as the root mean square of the contribution of each time step $\overline{\mathbf{P}(p)} = \sqrt{\frac{1}{|T|} \cdot \sum_{t \in T} \mathbf{P}_t(p)^2}$.

We simulated a total of 0.4s (roughly the average duration of a human spoken word [10]) with 13.573ms time gaps in between each sample, up to 1-meter radius around the jammer. To simulate a moving jammer, we update the position and orientation of each transducer at each sampled time step and simply repeat the aforementioned process. To simulate a small movement, we rotated all transducers by 15 degrees in 400ms – this depicts a relatively small microgesture of the wrist turning right.

Results

We performed four 3D simulations that suggested that a wearable jammer might outperform existing, planar or stationary, jammers. These are all depicted in Figure 4. For the sake of visual clarity, we plot only a 90° range of a 2D cross-section of the power distribution centered around the jammers.

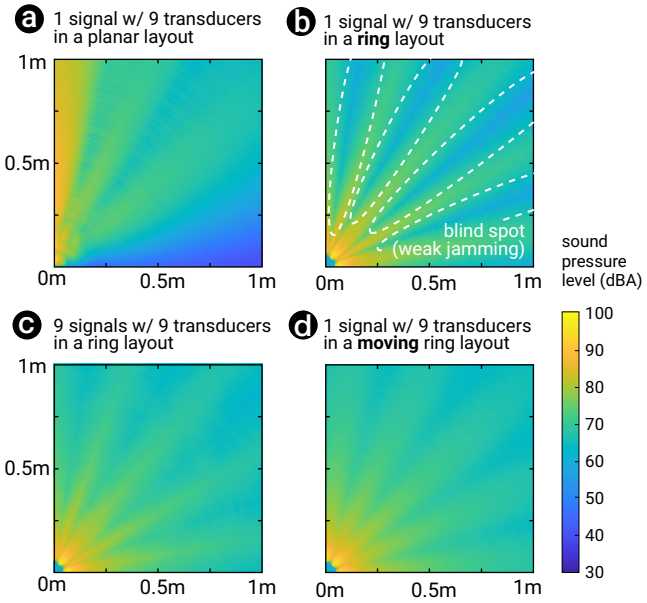

Figure 4. Our simulations depict how different transducer layouts radiate around the simulated device. We found that, when moving in space, a wearable jammer outperforms stationary jammers.

Simulating planar jammers

Figure 4(a) shows a simulation of a planar jammer; this is the design used in all known microphone jammers. We observed a rather limited angle coverage around the jammer, suggesting that planar jammers are mostly directional. From this insight, we decided to explore non-planar layouts.

Simulating ring-layouts

We simulated a ring-layout with 9 transducers. The result is depicted in Figure 4(b). We observed that, when compared to planar layouts, it radiates in all directions, with stronger components in the horizontal plane aligned with the transducers. However, we also observed the appearance of the blind spots

[3] http://sandlab.cs.uchicago.edu/jammer

[4] Ultrasonic transducer (NU25C16T-1), Jinci Technology. http://www.jinci.cn/showgoods/736.html

between transducer pairs (zones where their signals cancel each other out). These well-known blind spots are a key disadvantage of any multi-transducer jammer [33]; a microphone placed within a blind spot is unlikely to be jammed since the jamming signal intensity is weak.

One way to mitigate blind spots is to utilize a large number of out-of-phase sources. For instance, if one scales up to 9 independent signal generators it would limit phase collisions and thus reduce blind spots. We simulated this configuration and depicted it in Figure 4(c). We observed a smooth radiation pattern around the center—ideal for jamming. However, this approach drastically increases the number of components required to manufacture this design, e.g., 9 signal generators and 9 amplifiers (one per transducer). This approach is thus, highly impractical for a wearable implementation, both in its hardware footprint and its power consumption.

Blurring blind spots via movement

Thus, the ideal wearable implementation would find a way to mitigate the blind spots using only one signal source and one amplifier. Figure 4(d) demonstrates the power of making a jammer *into a wearable*. A wearable will move in space alongside the user's body. To simulate movement, we turned the jammer by 15 degrees during the 400ms of the simulation—as would occur when the user's wrist would turn to the right slightly. The result, depicted in Figure 4(d), is a smooth radiation map, containing almost no blind spots. We took this as the blueprint for our wearable jammer implementation. In the following laboratory experiments, we will empirically confirm these simulation results.

EXPERIMENT#1: ANGULAR POWER DISTRIBUTION

In this experiment, we measured the angular power distribution (i.e., the power emitted at different angles) of our wearable device in comparison to our aforementioned planar 3×3 jammer and the commercially available *i4*.

Experimental setup

We utilized three jammers in this study: (1) The *i4* (from *Amazon.com*, $799) consists of two perpendicular rows of ultrasonic transducers, five transducers on the side and two on the top. From our spectral analysis, the *i4* operates at the low end of ultrasonic frequency (20-24kHz), which allows its signals to travel further with slightly less power drop but unfortunately produces some disturbing audible sounds, likely due to signal leakage in its transducers resulting from the 20kHz signals. This device weighs 380 grams and consumes 4.2W of power. When measured directly at the transducers (with a sound pressure meter), its loudness is around 92.4dBA. (2) The planar jammer is an array of nine ultrasonic transducers in a 3 × 3 configuration. We built this jammer following [46, 47]; this device uses precisely the same transducers and amplifier as ours. The planar jammer used in this study operates at 25kHz±1kHz (the same signal as our device) and is completely inaudible. Similarly to [46, 47], this is not a stand-alone device and its power supply and circuitry are not integrated. When measured directly at the transducers, its loudness is around 92.6dBA. (3) Our wearable jammer was animated by a simple mechanical contraption. To move our bracelet, we used a servo motor. We programmed the servo to move 15° in 400ms, which is similar to slight wrist twist if the device was worn by a user. When measured directly at the transducers, our device's loudness is around 92.3dBA.

To measure the angular power distribution of all three devices, we placed the jammers on a table, one at a time. We measured all angles from 0° to 180° around the jammers at a distance of one meter, in steps of 5°. To obtain an accurate power measurement, we utilized the HT-80A sound level meter, which includes a well-calibrated microphone. When measuring our moving wearable, we took the average of the minimum and maximum power measured at each angle.

Results

The angular power distribution measured for our wearable jammer, planar jammer and *i4* are shown in Figure 5. We found that our device provides a wide-spread angular coverage ($M = -3.3dBA$, $SD = 1.6dBA$), while the existing jammers are highly directional (planar jammer: $M = -19.2dBA$, $SD = 8.5dBA$; *i4*: $M = -17.0dBA$, $SD = 6.8dBA$).

Furthermore, in the case of a planar jammer or the *i4*, even within the angular sector of [0°,40°], a subtle angle change of 2° leads to a 5-10dBA drop in their jamming power. This uneven distribution is due to the aforementioned blind spot problem [33]. Instead, the power of our wearable jammer has no dramatic drops across all angles, as the movement helps to blur out the blind spots.

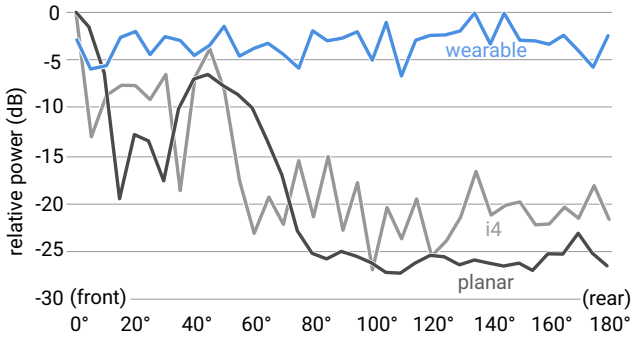

Figure 5. Real-world measurements of the jammer's angular coverage, in terms of the signal power as the jammer-to-microphone angle α increases from 0° to 180°, normalized by the maximum power of each jammer. The distance between the jammer and the microphone is kept at 1m. Angular coverage of the wearable jammer under movement. Jammer is 1m away from microphone.

EXPERIMENT#2: JAMMING SPEECH RECOGNIZERS

For an end-to-end evaluation of jamming effectiveness, we measured the ability of state-of-the-art speech recognizers to extract text from recordings of microphones jammed with our wearable or the baseline devices.

Experimental setup

We tested the jamming effectiveness of three jammers: our wearable device (animated by the same apparatus as in the previous experiment) and two baseline devices (the planar jammer and *i4*). We tested the jamming at multiple angles from 0° to 180°, in steps of 10° and always 1 meter away

from our jammer. This experiment used the built-in microphones of a Nexus 6 and a Xiaomi Mi 6. For the sake of visual clarity, we depict only the most conservative result, i.e., the device that best evaded our jamming–the Nexus 6.

To create a comparable experiment across multiple devices and angles, we cannot rely on a human speaker. Even a trained public speaker that would not make any pronunciation mistakes, would still introduce confounding variables in our measurements as their voice would not be perfectly replicable across multiple trials, i.e., its loudness (dBAs), its direction, its timbre, and so forth. Therefore, we utilized pre-recorded speech and played it back using a speaker (JBL GO, frequency response from 180Hz-20kHz). Our speaker was calibrated so as to play the pre-recorded human speech at a standard sound level of human conversation (55-66dBA measured at 1m away according to [40]). Lastly, the recorded speeches used in our experiment were ten 1-minute long sentences taken, at random, from the LibriSpeech dataset [18], which is commonly used by speech recognition researchers.

For each trial, we played back the pre-recorded speech via the speaker and recorded it with the smartphone's microphone. Then, we fed these recordings into the IBM Speech to Text [52]—a popular speech recognizer.

To compute the effectiveness of a jammer, we take the output of the recognizer and compare it to the transcript of each sentence in the dataset (ground truth). This results in the percentage of the words that were incorrectly transcribed by the text-to-speech; this is denoted as Word Error Rate (WER) and is a common metric in speech processing.

Results

Our results are depicted in Figure 6. We found that our wearable device jams more effectively in all directions ($M = 96.59\%$ WER, $SD = 3.97\%$) than the existing devices (planar jammer: $M = 38.89\%$ WER, $SD = 21.72\%$; *i4*: $M = 57.55\%$ WER, $SD = 35.04\%$). Since we did not measure much difference between the measurements obtained from the two different smartphones, our results depict an average of both. Furthermore, note that even without jamming, no text-to-speech system is perfect. In our experiment, we measured that in the absence of jamming the IBM Speech to Text had a WER around 30% for the smartphone.

Moreover, we observed a similar pattern to the angular power distribution found in the previous experiment. As depicted in Figure 6, both the planar jammer and *i4* exhibit WER drops at 30°and 60°, around their blind spots. On the contrary, our wearable jammer maintained a high WER throughout the measured angles. Furthermore, we observed a severe drop in WER, for planar jammer and *i4*, when the microphone was placed more than 90°away from the jammer (planar jammer: $M = 26.30\%$, $SD = 2.16\%$; *i4*: $M = 26.14\%$, $SD = 2.07\%$; our wearable: $M = 97.92\%$, $SD = 3.40\%$). First, this confirms that existing jamming approaches are highly directional. Secondly, it confirms that our approach is effective even when not pointing directly at the target device.

To exemplify the effectiveness of jamming with our wearable, we depict in Figure 7 three short sentences from our dataset.

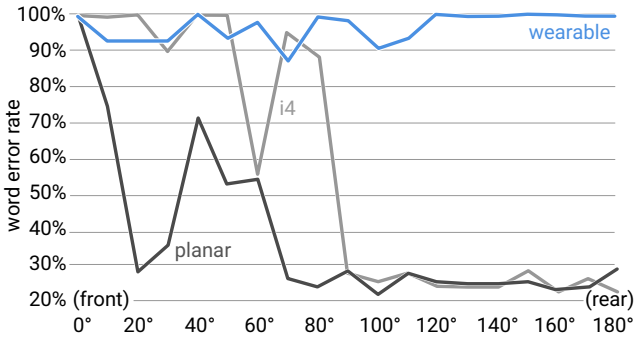

Figure 6. Word error rate (WER) of speech recognition for jamming with our wearable, planar jammer and *i4*. We found that the WER for planar and *i4* dropped drastically after 90°, while our wearable maintained a constant jamming effect >87%.

By contrasting the output of the text-to-speech when fed the jammed recording vs. when fed the clean recording, we observed that most words became unrecognizable. Yet, some words slipped through and were recognized, such as "space".

jammer off	jammer on
"now to bed boy"	"it"
"it is late and I go myself within a short space"	"space"
"most of all robin thought of his father what would he council"	

Figure 7. Examples of recognized sentences in clean speech case with perfect recognition and jamming case with our wearable jammer (WER 98.6%). Blank indicates nothing was recognized.

EXPERIMENT#3: JAMMING MICROPHONES COVERED BY EVERYDAY MATERIALS

As we observed in our last experiment, our wearable jammer has a wide angular coverage. Thus, it affords jamming even without the user needing to point to the target microphone. This feature allows it to also jam hidden microphones that the user might not be aware of. In this experiment, we evaluated whether this type of ultrasonic jamming is effective when the microphone is covered with a variety of materials, as it would be typical of a hidden microphone (e.g., in industrial espionage [17, 28]).

Experimental setup

We repeated our previous experiment (same apparatus), except we this time covered the microphones with different materials. In particular: a plastic bag (0.2mm thick Polyethylene), a plastic box (1mm thick Polypropylene), a paper sheet (from a 20lb set), a paper tissue (3-ply tissue), a cardboard box (3mm thick), a cloth (i.e., a cotton T-shirt), and two windshields (one fur and one foam) typically used in professional audio recordings. Additionally, we also recorded a baseline with no blockage applied.

Results

The results of the average WER are depicted in Figure 8. We found that the paper tissue, paper sheet, foam windshield and cloth had little impact on jamming performance, resulting in

an WER of 99%; in other words, our jammer was able to jam microphones hidden by these materials and only 1% of the words were correctly transcribed by the text-to-speech recognizer. Conversely, in the absence of our jammer, the text-to-speech recognizer recovered more than 60% of the words.

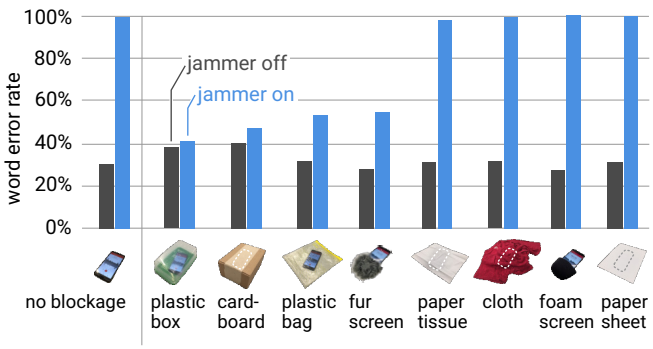

Figure 8. Speech recognition results when the microphone is covered up with various objects.

On the other hand, if the microphone was covered by a plastic box or cardboard box, we observed that the jamming performance dropped considerably to WER 41.01% and 46.76%, respectively; in other words, our device did not jam through the plastic box nor the cardboard box. In these two conditions, we also observed an increase of the WER even in the absence of jamming up to 38.13% and 40.29% respectively. To sum it up, materials such as paper, cloth and foam have little impact on jamming performance, while thicker or more complex blockage materials (e.g., plastic box, fur windshield, etc) decrease the jamming performance. This result is not a limitation of our wearable design but a limitation of acoustic jamming in general, since ultrasonic waves are reflected/absorbed differently from those at the audible spectrum for a given material. Therefore, practitioners and consumers should be aware of such limitations, and a more in-depth investigation of materials accordingly to their resonant properties and acoustic impedance is required.

USER STUDY

In our earlier experiments, we focused on controlled laboratory experiments that validated the jamming effectiveness of our wearable jammer. In our final study we, instead, aim to understand whether wearing our jamming bracelet impacts one's feeling of privacy. This study was reviewed and approved by our ethics committee (IRB19-0927).

Study design

Participants engaged in group conversations that lasted four minutes. Neither the topic nor the volume of the conversations was controlled. We asked participants to speak one at a time (otherwise the speech recognizer cannot make sense of it) and to not disclose any personal or sensitive information. During the group conversation, participants wore our privacy bracelet one at a time. They were asked to exchange the bracelet every minute, so that all could try it for an equal period. We recorded the conversation using four different commodity smartphones handed to each of the participants at the start of the study. We used the audio from all smartphones' recording for speech recognition. After the conversation was conducted, participants were presented with a transcript of the speech recognition (for the smartphone that they had during the study). After reading the transcript, they were asked to rate how much they felt that the bracelet had protected their privacy on a Likert scale (1-7). Lastly, note that the baseline of this study is implicit, as participants have a recollection of what they discussed in the group conversation and can judge how much the effect of the jammer influenced their perception of privacy. This study design does not allow us to repeat a non-staged conversation without the jammer nor were we interested in measuring actual word error rate (as we did that already in our previous controlled experiments).

Participants

To ensure that the English language level of each participant did not negatively reduce the fidelity of the speech recognizer, the candidates for this study were asked to read aloud sample sentences. Candidates who got over 70% accuracy were invited to participate in the study. As a result, we selected 12 participants (aged 18-26 years old; four self-identified as females and eight as males) from our local institution for this study. Ten of the participants had used some measure of privacy protection before, such as a laptop webcam cover, browser anti-tracking extensions, incognito mode, or VPN service. None of these participants had previously used a microphone jammer.

Apparatus

We used our jamming bracelet. We utilized four smartphones to record the conversation (models: Samsung S9+, Samsung S7, plus the aforementioned Nexus 6, and Xiaomi Mi 6). Lastly, we again used IBM's speech recognizer.

Results

Participants rated the feeling of privacy induced by the bracelet as $M = 5.4$ ($SD = 1.1$). This result, coupled with their positive comments, which we discuss below, suggested that the bracelet provided a sense of protection for the recorded conversation. While the wearable jammer did not jam the microphones completely in all recordings, the overwhelming majority of the transcripts of the four-minute conversations had only a dozen of mostly erroneous words.

When asked about their experience with the wearable jammer, most participants stated that they felt the bracelet was "definitely blocking out most words". Participants also noted that in certain cases specific words still made it through, such as the word *facebook* (P3). Three participants commented that the bracelet is bulky but not uncomfortable (P7, P8, P12). Two also added that while at the start, the bracelet was noticeable, once they focused on conversation, they "forgot about it" (P4) or "stopped feeling odd about wearing it" (P2).

Two participants (P10 and P8) added that they felt more protected either when wearing the bracelet or by simply seeing others wear it. To this, P8 added that at the current size the device would not be discreet enough to jam without others being unaware that you are doing so. Some participants (P5,

P2, P1) noted that they would have liked to better understand the range of the bracelet's efficiency.

Lastly, all twelve of the participants stated that they will use the bracelet again in the future. When asked specifically about the kinds of situations they would use it for, they noted, for instance: discussing private matters with their doctors (P1), discussing banking information (P6, P7, P10), talking to their employers (P9), or to strangers that joined a private conversation (P4).

DISCUSSION

Our experiments and user study provided insights into the advantages of a wearable microphone jammer. We found in our experiments that a wearable jammer in a ring layout is likely to outperform stationary jammers or jammers with planar layouts. Furthermore, we found that our jammer actually provided participants from our user study with a sense of increased privacy against eavesdropping microphones. Yet, there are a range of questions and limitations that we believe are relevant to address to move the field forward.

Limitations of our experiments

While we designed our studies to be as insightful and exhaustive as possible, it is simply not possible to test out the jammer against an infinite amount of existing microphone-based devices. Therefore, one must take into account that while our jammer was extremely effective against the microphones we used, these word rate errors cannot be easily generalized to other devices. Furthermore, our transducers are placed around the user's arm in a circular layout, which decreases its vertical coverage. In a preliminary experiment (using the apparatus of our experiment#2) we found that our device provides a vertical jamming of over 97% (WER) up to 75°; however, the jamming drops at 90° (precisely on top of the bracelet) to 75.54% (WER).

Non-linearities of microphone hardware

One speculative question is whether the non-linearity of today's microphone hardware is just a transient artifact of today's devices. We believe non-linearity is likely permanent for the foreseeable future, because the MEMS microphones used for smartphones and voice-based smart devices are designed for low-cost and small form-factors [30, 50, 12].

Counter-attacks to our wearable jamming

It is possible that attackers might craft exploits to circumvent our wearable jammer. That being said, the most likely attack would be noise canceling techniques intended to cancel out the jamming signals. To provide some validation against this attack, we de-noised the microphone recordings of our jammed signal over our speech library (same as in our Experiment#2) using two methods: (1) the deep neural network (DNN) denoising method from Rethage et al. [45], and (2) the widely used Wiener filter [44]. We observed no improvement in the denoised speech (WER 99.64% for the DNN-based method, and 100% for the Wiener filter), when compared to the original jammed speech audio (WER 99.64%). We believe that these current de-noising techniques will be of limited effect because of two key factors of our design: (1) we use randomly changing signals, which are hard to predict and cancel out; and, (2) the motion of the user's gestures and movements is also hard to predict, making it also extremely hard to cancel out these moving signal sources. Furthermore, to make it even harder to perform noise canceling of the jamming signals, one could even design signals that exhibit cadence patterns similar to human voice.

Safety

Our proposed system uses ultrasonic frequencies in the 25kHz range, while the upper limit frequency that the human ear can hear is around 15k-20kHz. The U.S. Occupational Safety and Health Administration (OSHA) warns that audible subharmonics can be harmful at intense sound pressures of 105 decibels or above [48]. Therefore, we ensured our jammer did not surpass this threshold. We measured the sound pressure of our bracelet directly at the transducer and found that its maximum sound pressure is below 92dB, well within the aforementioned safety limits.

Unintentional and selective jamming

As with any of the current ultrasonic jamming techniques (not only wearable jamming), it is possible that a jammer could accidentally jam legitimate microphones if these happen to be well inside the jamming range, including one's own smartphone, hearing aids or emergency response devices. More work is necessary to understand the impact of ultrasonic signals on these devices and to design workarounds.

Similarly, a user cannot selectively jam devices using ultrasound jamming: e.g., a user cannot choose to avoid jamming their own smartphone while still jamming another device. On this limitation, our approach does provide more control than existing stationary jammers. Stationary jammers, once activated will jam their entire range, requiring the user to walk all the way to the jammer to disable it. In our case, users can control the jammer's behavior by simply touching the bracelet. Moreover, moving forward, one would expect that adding intensity control to the wearable jammer might allow users to tune the jamming range.

Future form factors

While we found that our device outperformed existing jammer approaches, it is still larger than a typical bracelet. We believe our prototype offers a great blueprint towards a low-cost and ubiquitous microphone jammer. We expect this to inspire other wearable jammer designs, such as necklaces, earrings or even clothing.

CONCLUSIONS

We proposed, engineered and validated a wearable microphone jammer that is capable of disabling any microphones in the user's surroundings, including hidden microphones. Our wearable jammer takes the shape of a bracelet worn on the user's wrist and jams ubiquitously.

Our device is based on a recent exploit that leverages the fact that when exposed to ultrasonic noise, commodity microphones will leak the noise into the audible range. However,

previous ultrasound jammers that also exploited this principle, required users to point the jammer to the target microphone. This was necessary as these devices were built based on planar transducer layouts and were, therefore, highly directional. Unfortunately, this is impractical because it requires users to constantly worry and operate the jammer by pointing it to the surrounding microphones. Furthermore, this is sometimes impossible as users might desire to protect themselves from hidden eavesdropping microphones.

Instead, we found that our device outperforms these state-of-the-art jammers: (1) our wearable jams in multiple directions since its transducers are arranged in a ring layout; and, (2) our wearable jammer leverages natural hand gestures that occur while speaking to blur out blind spots, which are the main disadvantage of any jammer based on multiple transducers. We validated these advantages by means of simulation and three laboratory experiments.

Lastly, we conducted a user study with 12 participants that revealed that in a life-like situation participants felt that our wearable protected their voice privacy. We believe our wearable provides privacy in a world in which more and more devices are constantly eavesdropping on our conversations.

ACKNOWLEDGEMENTS

Our research was partially funded by the National Science Foundation, grants: CNS-1923778 and CNS-1705042. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Muhammad Taher Abuelma'atti. 2003. Analysis of the effect of radio frequency interference on the DC performance of bipolar operational amplifiers. *IEEE Transactions on Electromagnetic compatibility* 45, 2 (2003), 453–458.

[2] Rawan Alharbi, Tammy Stump, Nilofar Vafaie, Angela Pfammatter, Bonnie Spring, and Nabil Alshurafa. 2018. I Can'T Be Myself: Effects of Wearable Cameras on the Capture of Authentic Behavior in the Wild. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 90 (Sept. 2018), 40 pages. DOI: http://dx.doi.org/10.1145/3264900

[3] Leonardo Angelini, Maurizio Caon, Stefano Carrino, Luc Bergeron, Nathalie Nyffeler, Mélanie Jean-Mairet, and Elena Mugellini. 2013. Designing a Desirable Smart Bracelet for Older Adults. In *Proceedings of ACM Conference on Pervasive and Ubiquitous Computing Adjunct*. 425–434.

[4] Xavier Anguera, Chuck Wooters, and Javier Hernando. 2007. Acoustic beamforming for speaker diarization of meetings. *IEEE Transactions on Audio, Speech, and Language Processing* 15, 7 (2007), 2011–2022.

[5] Daniel Arp, Erwin Quiring, Christian Wressnegger, and Konrad Rieck. 2017. Privacy threats through ultrasonic side channels on mobile devices. In *Proc. of EuroS&P*.

[6] Dmitri Asonov and Rakesh Agrawal. 2004. Keyboard acoustic emanations. In *IEEE Symposium on Security and Privacy, 2004. Proceedings. 2004*. IEEE, 3–11.

[7] Md Tanvir Islam Aumi, Sidhant Gupta, Mayank Goel, Eric Larson, and Shwetak Patel. 2013. DopLink: Using the Doppler Effect for Multi-device Interaction. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*.

[8] Kazuki Awaki, Chun-Hao Liao, Makoto Suzuki, and Hiroyuki Morikawa. 2016. Speaker-less Sound-based 3D Localization with Centimeter-level Accuracy. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp)*.

[9] Michael Backes, Markus Dürmuth, Sebastian Gerling, Manfred Pinkal, and Caroline Sporleder. 2010. Acoustic Side-Channel Attacks on Printers.. In *USENIX Security symposium*. 307–322.

[10] Dom Barnard. 2018. Average Speaking Rate and Words per Minute. VIRTUALSPEECH. (January 2018). https://virtualspeech.com/blog/average-speaking-rate-words-per-minute.

[11] Victoria Bellotti and Abigail Sellen. 1993. Design for Privacy in Ubiquitous Computing Environments. In *Proceedings of the Third Conference on European Conference on Computer-Supported Cooperative Work (ECSCW'93)*. Kluwer Academic Publishers, Norwell, MA, USA, 77–92. http://dl.acm.org/citation.cfm?id=1241934.1241940

[12] Joseph A. Boales, Farrukh Mateen, and Pritiraj Mohanty. 2017. Micromechanical microphone using sideband modulation of nonlinear resonators. *Applied Physics Letters* 111, 9 (2017), 093504.

[13] Matthew Brocker and Stephen Checkoway. 2014. iSeeYou: Disabling the MacBook Webcam Indicator LED. In *Proceedings of the 23rd USENIX Conference on Security Symposium (SEC'14)*. USENIX Association, Berkeley, CA, USA, 337–352. http://dl.acm.org/citation.cfm?id=2671225.2671247

[14] Ke-Yu Chen, Daniel Ashbrook, Mayank Goel, Sung-Hyuck Lee, and Shwetak Patel. 2014. AirLink: Sharing Files Between Multiple Devices Using In-air Gestures. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*.

[15] H. Chung, M. Iorga, J. Voas, and S. Lee. 2017. Alexa, Can I Trust You? *Computer* 50, 9 (2017), 100–104.

[16] Hyunji Chung and Sangjin Lee. 2018. Intelligent Virtual Assistant knows Your Life. *CoRR* abs/1803.00466 (2018).

[17] Gregg D Colton. 1997. High-Tech Approaches to Breeching Examination Security. Espionage 101. (1997).

[18] LibriSpeech Dataset. 2017. http://www.openslr.org/12. (2017).

[19] Luigi De Russis, Dario Bonino, and Fulvio Corno. 2013. The Smart Home Controller on Your Wrist. In *Proceedings of ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp)*. 8.

[20] Silent Ultrasonic Microphone Defeater. 2019. https://www.uspystore.com/silent-ultrasonic-microphone-defeater. (2019).

[21] Hidden Microphone dictaphone Bug Recording supressor ultrasonic + Noise Generator by i4 Technology. 2019. https://www.amazon.com/Microphone-dictaphone-Recording-supressor-ultrasonic/dp/B01MG4WACJ/. (2019).

[22] Al Faruque, Mohammad Abdullah, Sujit Rokka Chhetri, Arquimedes Canedo, and Jiang Wan. 2016. Acoustic side-channel attacks on additive manufacturing systems. In *Proceedings of the 7th International Conference on Cyber-Physical Systems*. IEEE Press, 19.

[23] Carl Fischer, Kavitha Muthukrishnan, Mike Hazas, and Hans Gellersen. 2008. Ultrasound-aided Pedestrian Dead Reckoning for Indoor Navigation. In *Proceedings of the First ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments (MELT '08)*.

[24] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. SoundWave: Using the Doppler Effect to Sense Gestures. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems (CHI)*.

[25] Mordechai Guri, Yosef Solewicz, Andrey Daidakulov, and Yuval Elovici. 2017. SPEAKE(a)R: Turn Speakers to Microphones for Fun and Profit. In *11th USENIX Workshop on Offensive Technologies (WOOT 17)*. USENIX Association, Vancouver, BC. https://www.usenix.org/conference/woot17/workshop-program/presentation/guri

[26] Google is permanently nerfing all Home Minis because mine spied on everything I said 24/7 [Update x2]. 2017. https://www.androidpolice.com/2017/10/10/google-nerfing-home-minis-mine-spied-everything-said-247/. (2017).

[27] Speech jammer TOWER-A for blocking proffessional microphones / counter surveillance. 2019. https://www.detective-store.com/speech-jammer-tower-a-for-blocking-proffessional-microphones-counter-surveillance-1516.html. (2019).

[28] Ilias Kaperonis. 1984. Industrial espionage. *Computers & Security* 3, 2 (1984), 117–121.

[29] Gary S. Kendall, Christopher Haworth, and Rodrigo F. Cádiz. 2014. Sound Synthesis with Auditory Distortion Products. *Computer Music Journal* 38 (2014), 5–23. Issue 4.

[30] Junhong Li, Chenghao Wang, Wei Ren, and Jun Ma. 2017. ZnO thin film piezoelectric micromachined microphone with symmetric composite vibrating diaphragm. *Smart Materials and Structures* 26, 5 (2017), 055033.

[31] Dominique Machuletz, Stefan Laube, and Rainer Böhme. 2018. Webcam Covering As Planned Behavior. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 180, 13 pages. DOI: http://dx.doi.org/10.1145/3173574.3173754

[32] Sapna Maheshwari. 2018. Hey, Alexa, What Can You Hear? And What Will You Do With It? New York Times. (March 2018). https://mobile.nytimes.com/2018/03/31/business/media/amazon-google-privacy-digital-assistants.html.

[33] Robert J Mailloux. 1982. Phased array theory and technology. *Proc. IEEE* 70, 3 (1982), 246–291.

[34] Asier Marzo, Sue Ann Seah, Bruce W. Drinkwater, Deepak Ranjan Sahoo, Benjamin Long, and Sriram Subramanian. 2015. Holographic acoustic elements for manipulation of levitated objects. *Nature Communications* 6, 1 (2015), 8661.

[35] R. Mayrhofer and H. Gellersen. 2007. On the Security of Ultrasound as Out-of-band Channel. In *Proceedings of the 2007 IEEE International Parallel and Distributed Processing Symposium*.

[36] Saeed Mirzamohammadi and Ardalan Amiri Sani. 2016. Viola: Trustworthy Sensor Notifications for Enhanced Privacy on Mobile Systems. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '16)*. ACM, New York, NY, USA, 263–276. DOI: http://dx.doi.org/10.1145/2906388.2906391

[37] Rafael Morales, Asier Marzo, Sriram Subramanian, and Diego Martínez. 2019. LeviProps: Animating Levitated Optimized Fabric Structures using Holographic Acoustic Tweezers. In *Proc. of UIST*.

[38] Tim Moynihan. 2016. Alexa and Google Home Record What You Say. But What Happens to That Data? Wired. (December 2016). https://www.wired.com/2016/12/alexa-and-google-record-your-voice/.

[39] New Generation of High Grade Smartphone Scrambler. 2019. https://www.globaltscmgroup-usa.com/. (2019).

[40] Wayne O Olsen. 1998. Average speech levels and spectra in various speaking/listening conditions: A summary of the Pearson, Bennett, & Fidell (1977) report. *American Journal of Audiology* 7, 2 (1998).

[41] Minna Pakanen, Ashley Colley, Jonna Häkkilä, Johan Kildal, and Vuokko Lantz. 2014. Squeezy Bracelet: Designing a Wearable Communication Device for Tactile Interaction. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14)*. 305–314.

[42] Leysia Palen and Paul Dourish. 2003. Unpacking "Privacy" for a Networked World. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 129–136. DOI:http://dx.doi.org/10.1145/642611.642635

[43] Rebecca S. Portnoff, Linda N. Lee, Serge Egelman, Pratyush Mishra, Derek Leung, and David Wagner. 2015. Somebody's Watching Me?: Assessing the Effectiveness of Webcam Indicator Lights. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1649–1658. DOI: http://dx.doi.org/10.1145/2702123.2702164

[44] William K Pratt. 1972. Generalized Wiener filtering computation techniques. *IEEE Trans. Comput.* 100, 7 (1972), 636–641.

[45] Dario Rethage, Jordi Pons, and Xavier Serra. 2018. A wavenet for speech denoising. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5069–5073.

[46] Nirupam Roy, Haitham Hassanieh, and Romit Roy Choudhury. 2017. Backdoor: Making microphones hear inaudible sounds. In *Proceedings of ACM MobiSys*.

[47] Nirupam Roy, Sheng Shen, Haitham Hassanieh, and Romit Roy Choudhury. 2018. Inaudible Voice Commands: The Long-Range Attack and Defense. In *Proceedings of the 15th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*.

[48] U.S. Occupational Safety and Health Administration (OSHA). 2013. Occupational Safety and Health Administration Technical Manual. https://www.osha.gov/dts/osta/otm/new_noise/#appendixc. (August 2013).

[49] T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A. Landay. 2009. Enabling Always-available Input with Muscle-computer Interfaces. In *Proceedings of ACM Symposium on User Interface Software and Technology (UIST)*. 167–176.

[50] Woon Seob Lee and Seung S. Lee. 2008. Piezoelectric microphone built on circular diaphragm. 144 (06 2008), 367–373.

[51] Liwei Song and Prateek Mittal. 2017. Inaudible Voice Commands. *CoRR* abs/1708.07238 (2017).

[52] IBM Speech to Text. 2018. https://www.ibm.com/watson/services/speech-to-text/. (Jul. 2018).

[53] Your Phone Is Listening Literally Listening to Your TV. 2015. https://www.theatlantic.com/technology/archive/2015/11/your-phone-is-literally-listening-to-your-tv/416712/. (2015).

[54] Edward J. Wang, Tien-Jui Lee, Alex Mariakakis, Mayank Goel, Sidhant Gupta, and Shwetak N. Patel. 2015. MagnifiSense: Inferring Device Interaction Using Wrist-worn Passive Magneto-inductive Sensors. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*.

[55] Charlie Wood. 2017. Devices sprout ears: What do Alexa and Siri mean for privacy? Christian Science Monitor. (January 2017). https://www.csmonitor.com/Technology/2017/0114/Devices-sprout-ears-What-do-Alexa-and-Siri-mean-for-privacy.

[56] Candid Wueest. 2017. Everything You Need to Know About the Security of Voice-Activated Smart Speakers. Symantec. (Nov. 2017). https://www.symantec.com/blogs/threat-intelligence/security-voice-activated-smart-speakers.

[57] Chenren Xu, Sugang Li, Gang Liu, Yanyong Zhang, Emiliano Miluzzo, Yih-Farn Chen, Jun Li, and Bernhard Firner. 2013. Crowd++: Unsupervised Speaker Count with Smartphones. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*.

[58] Tuo Yu, Haiming Jin, and Klara Nahrstedt. 2016. WritingHacker: Audio Based Eavesdropping of Handwriting via Mobile Devices. In *Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*.

[59] Clint Zeagler. 2017. Where to Wear It: Functional, Technical, and Social Considerations in On-body Location for Wearable Technology 20 Years of Designing for Wearability. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*.

[60] Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. 2017. DolphinAttack: Inaudible voice commands. In *Proceedings of ACM Conference on Computer and Communications Security (CCS)*.

[61] Li Zhuang, Feng Zhou, and J Doug Tygar. 2009. Keyboard acoustic emanations revisited. *ACM Transactions on Information and System Security (TISSEC)* 13, 1 (2009), 3.