SURF: Eavesdropping on Underwater Communications from the Air

Poorya Mollahosseini[1,†], Sayed Saad Afzal[2,†], Fadel Adib[2], Yasaman Ghasempour[1]
poorya@princeton.edu,afzals@mit.edu,fadel@mit.edu,ghasempour@princeton.edu
[1] Princeton University, [2] Massachusetts Institute of Technology

ABSTRACT

This paper investigates how an airborne node can eavesdrop on the underwater acoustic communication between submerged nodes. Conventionally, such eavesdropping has been assumed impossible as acoustic signals do not cross the water-air boundary. Here, we demonstrate that underwater acoustic communications signals can be picked up and (under certain conditions) decoded using an airborne mmWave radar due to the minute vibrations induced by the communication signals on the water surface. We implemented and evaluated a proof-of-concept prototype of our method and tested it in controlled (pool) and uncontrolled environments (lake). Our results demonstrate that an airborne device can identify the modulation and bitrate of acoustic transmissions from an *uncooperative* underwater transmitter (victim), and even decode the transmitted symbols. Unlike conventional over-the-air communications, our results indicate that the secrecy of underwater links varies depending on the modulation type and provide insights into the underlying reasons behind these differences. We also highlight the theoretical limitations of such a threat model, and how these results may have a significant impact on the stealthiness of underwater communications, with particular concern to submarine warfare, underwater operations (e.g., oil & gas, search & rescue, mining), and conservation of endangered species. Finally, our investigation uncovers countermeasures that can be used to improve or restore the stealthiness of underwater acoustic communications against such threats.

CCS Concepts

• **Networks → Cyber-physical networks**; *Sensor networks*; *Mobile networks*; • **Security and privacy → Mobile and wireless security**.

[†] These authors contributed equally to this work.

ACM MobiCom '24, November 18–22, 2024, Washington D.C., DC, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0489-5/24/11.
https://doi.org/10.1145/3636534.3690663

Keywords

Wireless, Cross-Medium Communications, Security, Subsea Internet of Things

ACM Reference Format:
Poorya Mollahosseini, Sayed Saad Afzal, Fadel Adib, Yasaman Ghasempour. 2024. SURF: Eavesdropping on Underwater Communications from the Air. In *The 30th Annual International Conference on Mobile Computing and Networking (ACM MobiCom '24), November 18–22, 2024, Washington D.C., DC, USA*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3636534.3690663

1 INTRODUCTION

Over the years, acoustic wireless communication networks have gained widespread acceptance and adoption by the naval and oceanic communities [8, 16]. Generally, compared to optical and RF, acoustic frequencies are preferred for long-range communications due to their superior propagation characteristics under the water [17, 46]. These acoustic waves act as pressure waves that diverge as they propagate underwater. When these waves hit the water's surface, they cause minuscule vibrations, which can be sensed and decoded using an airborne radar receiver. This phenomenon has been recently exploited to establish a direct underwater-to-air communications link between two trusted cooperative parties [46]. Yet, the security implications of such cross-medium sensing are left unexplored. In other words, what if an air-borne non-cooperative eavesdropper attempts to intercept underwater communication links by picking up and extracting patterns in such vibration signatures?

This paper investigates the vulnerabilities of underwater acoustic communication against airborne adversaries. Specifically, we design, implement, and evaluate Snooping Underwater communications using Radio Frequency (SURF), the first eavesdropping framework that allows for intercepting acoustic underwater communication links from the air. Our system overview is captured in Fig. 1. An out-of-medium eavesdropper can be particularly concerning as its presence remains concealed from the victim, potentially even when equipped with the most advanced SONAR techniques. Indeed, because of the acoustic impedance mismatch between the two mediums of water and air, emitted sound by the underwater node reflects from the water-air interface [46]. Hence, SONAR techniques cannot be used to detect airborne nodes. There

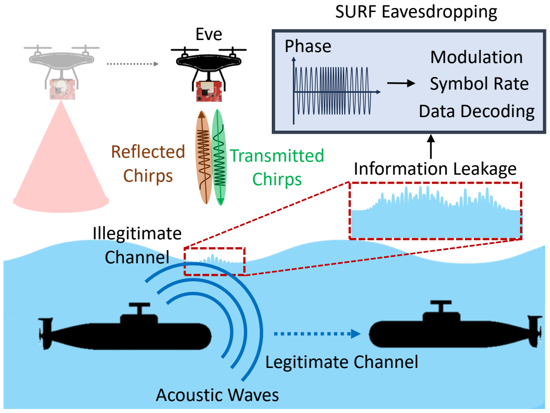


Figure 1: SURF's system overview. SURF uses a mmWave FMCW radar to eavesdrop on acoustic communication happening in underwater environments. It senses the vibration of the water surface and extracts the modulation and symbol rate of the data to decode the message.

are two key threats: First, the water vibrations caused by the legitimate acoustic transmitter may be picked up by adversary radars and successfully decoded (under certain conditions). Additionally, by picking up these vibrations, an adversary radar may, in principle, identify the rough location of the underwater transmitter. This is particularly important in the context of naval assets since submarine locations are closely regarded as trade secrets and even in conservation efforts where scientific expeditions want to preserve the privacy of the location of endangered species.

To understand these vulnerabilities, we first provide a comprehensive model of surface vibration as a function of the speaker's geometric and physical layer (PHY) parameters, as well as by taking into account how the displacement at one point propagates outward in the form of gravity-capillary waves.[1] Our model indicates that the spatial footprint of an acoustic signal is non-negligible on the water surface and indeed the area of vibration expands as the underwater speaker is submerged deeper into the water. This suggests that an airborne receiver may pick up vibrations from an underwater acoustic source over a wider area than was previously considered possible.

While our model demonstrates that the region of detection may be wide, an eavesdropper (Eve) still faces several challenges in decoding messages sent from an uncooperative underwater acoustic transmitter (Alice). First, because the adversary is outside of the medium and since Alice is uncooperative, extracting the underlying PHY layer parameters (including carrier frequency, modulation type, and symbol rate) is not straightforward for the adversary. Additionally, it is difficult to synchronize Eve with Alice's underwater speaker due to a lack of knowledge of the preamble and training sequence. All of this makes it challenging for Eve to identify and decode the underwater transmissions.

To address these challenges, we exploit the fundamental properties of the cross-medium channel. Specifically, we exploit that the underwater communication channel, the in-air RF propagation, and the water-air (translational) boundary are all linear. This means that the end-to-end channel acts as a linear system, preserving the spectral properties of the transmitted waveforms. Using this observation, we apply techniques from communications theory that aim to classify and determine the PHY properties of communication signals from uncooperative sources using spectral and temporal features, all achieved without any training data or prior knowledge. Adapting such past approaches to this cross-medium communication link requires SURF to take into account unique properties and constraints. For example, we show how the (significantly) higher sampling rate of chirp signals transmitted from the airborne radar enables higher-resolution measurements that can be used to infer the PHY parameters of the underlying acoustic communication. Our approach includes identifying the modulation and bitrate of the underwater acoustic transmissions, estimating empirical decision boundaries, as well as decoding the transmitter packets. Additionally, we quantify the secrecy metrics of different underwater acoustic modulation schemes.

A second challenge in eavesdropping on the underwater acoustic transmissions is that Eve does not know the location of the transmitting source. This is problematic because to obtain a high signal-to-noise ratio (SNR) of the underwater acoustic communication signals, Eve needs to hone in on the area of the surface where the induced vibrations are strongest. To address this challenge, SURF exploits the beamforming capabilities of the millimeter-wave radars and combines them with our derived model of the underwater-to-air channel. In particular, to maximize the SNR, SURF performs a beam search (using antenna array beamforming) to identify the location on the surface that has the highest vibration caused by the incident acoustic signals. Note that doing so requires first mitigating the impact of the naturally occurring surface waves (which are often 4-5 orders of magnitude higher than the induced vibrations); it also requires performing a range-vibration search over this surface, which SURF achieves using 2D FFTs.

We designed and built a prototype of SURF and evaluated it in different settings, including, controlled lab settings, in a swimming pool, and in a natural lake. Our prototype was built using a commercial off-the-shelf (COTS) millimeter-wave 77

[1] The same phenomenon causes rings of ripples when throwing a rock in a calm pond.

GHz Frequency Modulated Continuous Wave (FMCW) radar[2] and an underwater acoustic speaker. Our choice of COTS radar aimed to demonstrate how such eavesdropping may be feasible even without custom-designed high-end equipment. We have also demonstrated the feasibility of SURF when the radar is mounted on a flying drone. We measure and compare the bit error rate (BER) at the legitimate underwater hydrophone receiver (Bob) against the airborne radar. Our experimental evaluation demonstrates the following:

- SURF can accurately identify the underlying PHY layer properties of the victim. Specifically, SURF classifies the victim's modulation type with an accuracy of 97.58% and estimates the symbol rate with a root mean square error of 9 bps, without any prior knowledge.
- Different modulation types exhibit different resilience against a cross-medium eavesdropper. This has important new implications for enhancing the security of underwater links.
- SURF hints at the location of the underwater speaker based on the area of vibration caused on the water's surface.

Our results and observations provide interesting insights for securing the underwater communication channels against out-of-medium eavesdroppers: *(i)* The likelihood of eavesdropping is not the same for all modulations. In fact, frequency modulation is the most adopted modulation for underwater links and is also the least secure. Amplitude modulation exhibits the best secrecy property as natural environmental factors (e.g., surface waves and wind) mask the minute fluctuation in magnitude of vibration. *(ii)* the underwater link is more secure under certain data rates and transmit power levels, and *(iii)* the link secrecy improves if Alice transmits her data in several short intervals with idle times in between that are longer than the channel coherence time. This way, the eavesdropper does not have sufficient samples to estimate the PHY parameters and decision boundaries. Albeit, the enhanced secrecy is achieved at the cost of hindering communication goodput.

Contributions: Contrary to the established belief that underwater acoustic communication is secure from airborne eavesdroppers, this paper demonstrates a first-of-its-kind eavesdropping attack on such communication systems. It also contributes to a prototype implementation and experimental evaluation of such an attack in a lake, demonstrating the ability to sense, identify, and decode transmissions from an underwater acoustic source. The paper highlights the implications of such a threat, particularly on the stealthiness of underwater links, and identifies countermeasures that may restore - or at least improve - the security of these communication systems against airborne eavesdroppers.

2 SURF's DESIGN

2.1 Threat Model

Underwater communication mainly relies on acoustic signals that can travel long distances underwater [5, 22]. Recently, NATO's Centre for Maritime Research and Experimentation introduced JANUS, a standardized protocol for transmitting digital information underwater using sound [33]. Because of the limited bandwidth under the water, JANUS packets can only be 64 bits in size. This means that traditional encryption methods like Advanced Encryption Standard (AES) cannot be used effectively.

We consider an underwater acoustic communication link between an underwater transmitter (Alice) and a legitimate underwater receiver (Bob). We assume that a portion of the acoustic power aimed at Bob reaches the water's surface. This is reasonable as many underwater nodes deploy omnidirectional or quasi-omnidirectional antennas for SWaP-C constraints [31, 37, 42]. Even with deploying directional beams, the low attenuation of acoustic frequencies [34] and their large beam divergence [28] make it possible for at least a portion of the acoustic wave energy to reach the water's surface. However, we note that highly directional sound waves may necessitate a different model of vibration for their detection, interpretation, and positioning.

Upon reaching the surface, the energy in pressure waves causes minute displacements of the water. The magnitude and frequency of such displacements contain information about the underlying communication link between Alice and Bob. The attacker, Eve, is an airborne node located outside the water with radar-sensing capabilities and aims at mapping the vibration patterns to the stream of transmitted symbols by Alice. We highlight that a naive eavesdropper can in principle carry a tethered hydrophone from a drone and float it in the water (or very close to the water surface) for direct eavesdropping. However, such a setup is more challenging to implement (particularly in mobile settings) and easier to detect, e.g., using an underwater SONAR that detects the hydrophone.

We assume Eve has no prior knowledge about the PHY-layer parameters of the underwater nodes, including carrier frequency, modulation type, and symbol rate. Additionally, the eavesdropper is not cooperating nor synchronized with the legitimate underwater nodes. Furthermore, since the eavesdropper is not inside the medium (water), she cannot directly measure the channel conditions, i.e., dense vs. sparse multipath, delay spread, channel coherence time, etc. This imposes interesting new challenges and tradeoffs that often do not exist

[2] The choice of mmWave band for the radar is explained in more detail in [46]. A larger wavelength, such as those used in WiFi or cellular, leads to smaller phase variations, reducing robustness against noise. Conversely, a very small wavelength, like those in THz or optical frequencies, causes quick phase wrapping, which hinders the tracking of surface vibrations [14, 24].

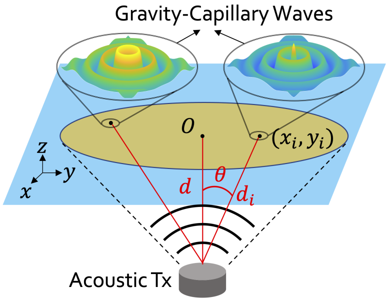


Figure 2: Modeling of water-air interface. The pressure waves emitted from the acoustic Tx hit the water's surface at the point of impact (x_i, y_i). The resulting displacement travels outward as gravity-capillary waves.

in typical threat models in which all parties belong to the same wireless medium.

2.2 Vibration Model

To infer the underlying modulated bit sequence through the vibration pattern on the water's surface, the first step is to systematically model the water vibration as it interacts with pressure waves. The amount of pressure is determined by the characteristics of the transmitter, including the transmitted power P_{tx} and directivity pattern $D(\theta)$. The wave attenuation due to expansion and absorption losses depends on the propagation distance to the water surface, as well as the medium properties: water density ρ and speed of sound c_w in the water. As shown in Fig. 2, we can model the amount of acoustic pressure at the impact point i on the water's surface with coordinates (x_i, y_i) as [36]:

$$p_i = \frac{D(\theta)}{d_i}\sqrt{\frac{P_{tx}\rho c_w}{4\pi}}e^{-\alpha d_i}, \tag{1}$$

where d_i is the distance between the speaker and point (x_i, y_i), α is the attenuation coefficient. The amplitude of the water displacement δ_i caused by the incident pressure p_i can be derived as:

$$\delta_i = \frac{p_i \cos\theta_i}{\omega\rho c_w}, \tag{2}$$

where $\cos\theta_i$ represents the normal component of the pressure wave hitting the surface, and ω is the angular frequency of the acoustic signal. Note that the tangential component does not contribute to the vibrations on the water's surface.

The displacement created at the impact point propagates outward in the form of exponentially decaying gravity-capillary waves which propagate outward in a circular pattern. The wave number of these waves, denoted by k_{gc}, can be obtained by solving the dispersion relation [36]. We define $r_i(x, y)$ as the distance between point i and an arbitrary point (x, y) on the water surface. Hence, the displacement at (x, y) caused by the vibration originated from point i is:

$$\eta_i(x, y, t) = \delta_i e^{-\beta r_i(x,y)} e^{j(k_{gc} r_i(x,y) + k d_i - \omega t)}, \tag{3}$$

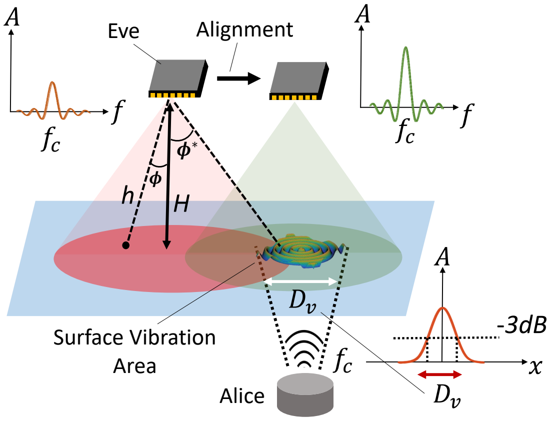


Figure 3: Localizing vibration area. SURF localizes the vibration area on the water's surface by exploiting the beamforming abilities of the radar.

where β represents the surface wave attenuation, and k is the wave number of the acoustic waves underwater.

Finally, Eq. (3) only considers the capillary waves generated by a single point, namely, point i. However, all the points that are displaced due to underwater pressure will be a source of capillary waves. Therefore, the total displacement is the superposition of all these complex forces:

$$R(x, y, t) = \sum_{j=1}^{\infty} \eta_j(x, y, t). \tag{4}$$

Since the displacement caused by capillary waves exponentially decays in space, in practice, only a fraction of spatial points j in close proximity of (x, y) will play a role in the overall displacement profile.

2.3 Vibration Detection

When the pressure waves emitted from the speaker reach the water's surface, it impacts some areas of the surface, depending on the speaker's location and its radiation pattern. It is advantageous for Eve to place herself directly over the vibration area to boost her SNR and increase her chances of eavesdropping. We assume that the vibration area is within the field of view (FOV) of the radar. If that is not the case, trajectory planning is needed, which can be addressed with mission-oriented path planning (well-studied in UAV and robotics). As shown in Fig. 3, when the radar is positioned directly above the speaker, it senses vibration frequency (f_c) with a considerably higher amplitude compared to when it is not directly above.

Eve can employ conventional beamforming techniques to find a rough estimate of the area of vibration and position herself strategically. Specifically, Eve uses a multi-antenna radar system with N virtual antennas[3] (number of transmit

[3] Low-cost COTS mmWave radar provides 8 to 12 virtual antennas [18, 19].

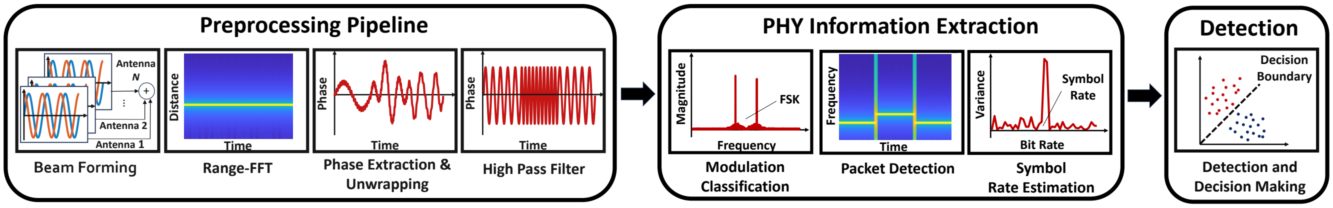


Figure 4: SURF's processing pipeline. First, we exploit FMCW signal processing for vibration detection. The middle blocks show the extraction of unknown PHY information. Finally, we implement detection and decision making.

antennas × number of receive antennas) and uses the spatial diversity of these channels to localize the vibration area. Note that merely comparing the reflected power captured by the radar under different beam conditions would not result in vibration localization. Instead, SURF exploits the fact that the vibration caused by acoustic signals falls within a certain frequency band B. Thus, it performs a 3D beam search – over angle ϕ, distance h, and frequency f_c – and localizes the vibration area. This can be formulated as 3D FFT (range FFT via FMCW signal processing, angle FFT, and Doppler FFT for vibration frequency) and can be performed using standard techniques [45, 46]. Mathematically, the angular location of vibration relative to the radar (denoted as ϕ^*) can be described as:

$$\phi^* = \arg\max_{\phi}[\max_{h,f_c\in B} \Psi(\phi, f_c, h)]. \quad (5)$$

where $\Psi(\phi, f, h)$ denotes the 3D FFT over vibration frequency, distance, and angle.

One challenge is that at angles other than normal, the reflected signal off the surface will not be pointed back to Eve's receiver as the water acts as a smooth mirror that obeys Snell's law. Hence, this implies that the amount of reflected power decreases at higher angles ϕ^*. While this trend is generally true, in practice, the presence of natural waves causes uneven surface levels and diversifies the reflection angle of chirps, leading to some power being reflected to the radar at incident angles other than normal. Additionally, Eve can average over several transmissions to improve the localization performance as she is primarily interested in tracking the phase of multiple chirps and not decoding any data at this stage.

As shown in Fig. 3, we denote the diameter of vibration as D_v by considering the distance at which the vibration amplitude drops by 3 dB from its peak. Since D_v is related to the receiver depth, Eve can potentially infer the depth of the speaker by measuring D_v.

2.4 PHY Inference and Demodulation Pipeline

The overall system architecture of SURF is shown in Fig. 4. There are three main components: Preprocessing, extracting PHY layer information, and non-coherent detection and decision-making.

Preprocessing. First, Eve combines the signal received by N virtual antennas to enhance her SNR. The consequent steps are standard FMCW radar processing steps that include deriving the range-FFT matrix, calculating the phase at the distance bin of maximum power, and unwrapping the phase, as shown in the first block of Fig. 4. The resulting phase pattern contains strong low-frequency components that stem from natural surface waves (typically in the range of 0-10 Hz [23]) that can mask the phase profile. Hence, Eve removes such unwanted low-frequency components using a high pass filter. The resulting phase profile contains information about the transmitted bit sequence by Alice. However, interpreting such information is not straightforward as Eve is unaware of the underlying PHY-level parameters including the modulation type, symbol boundaries, and rate.

2.4.1 PHY Information Extraction

Modulation Classification: A fundamental challenge for Eve is the unknown modulation used between Alice and Bob. To tackle this, Eve leverages two characteristics of the underwater channel: linearity and large channel coherence time[4]. Given that the underwater channel is linear [36, 46], Alice's carrier frequency is directly seen as a peak in the Doppler-FFT heat map. The presence of multiple dominant peaks suggests a multi-carrier modulation, e.g., FSK or OFDM. Furthermore, any temporal variation in the amplitude and phase at each carrier frequency is a sign of amplitude and phase modulation, respectively. Observing these features is possible because the long coherence time guarantees that such fast-changing variations in phase/amplitude are solely a function of Alice's transmission and not the channel.

The general strategy in identifying the modulation type involves examining the amplitude and phase variations of each carrier frequency in the short-time Fourier transform (STFT) of the filtered unwrapped phase. Our pseudo-code is shown in Algorithm 1. Specifically, by comparing the noise level with the peak vibration, we infer whether there is any ongoing acoustic transmission in that interval or not. Noise can be calculated from the power received at other ranges of frequency bins in the Doppler-FFT plot.

[4] Coherent time of a few hundred milliseconds reported in prior work [3]

Algorithm 1 Modulation Classification Algorithm

```
Require: The filtered unwrapped phase, $x[n]$.
Require: Peak detection algorithm $g$
Require: Noise level $a_n$, Hyper parameters $SNR_{min}$.
Ensure: Predicted class of $x[n]$
$X[f] \leftarrow$ FFT of $x[n]$
$\tilde{X}[f,t] \leftarrow$ STFT of $x[n]$
$(a_1,\dots,a_K,f_1,\dots,f_K) \leftarrow g(X[f])$ {Sorted peaks in the descending order, $a_1 > a_2 > \dots$}
if $\frac{a_1}{a_n} < SNR_{min}$ then
    return No Data
else if $K > 1$ then
    $A_k[t] \leftarrow$ Amplitude of $(\tilde{X}[f_k,t])$
    if $\exists(i,j,t)\ (i \neq j) : \frac{A_i[t]}{a_n} > SNR_{min}$ and $\frac{A_j[t]}{a_n} > SNR_{min}$ then
        return OFDM
    else
        return FSK
    end if
else
    $P_1[t] \leftarrow$ Phase of $(\tilde{X}[f_1,t])$
    if $P_1[t]$ is linear function of $t$ then
        return ASK
    else
        return PSK
    end if
end if
```

When the presence of surface vibration due to data-modulated symbols is confirmed, SURF proceeds to detect the modulation type. We denote $x[n]$ as the filtered unwrapped phase, and $X[f]$ as its FFT. Firstly, Eve runs a peak detection algorithm on $X[f]$ to identify the peaks in the frequency domain. The amplitude of these peaks is denoted by a_k and their frequency by f_k. The number of peaks at a given instant, denoted by K, determines whether the underlying modulation is single-carrier or multi-carrier (e.g., OFDM or FSK). Distinguishing between OFDM and FSK is straightforward as in FSK only one dominant peak exists in STFT intervals while OFDM has several sub-carriers.

For a given carrier frequency f_k, Eve assesses the temporal phase and amplitude variations (note that for single-carrier modulation, $K = 1$). This is achieved by computing the STFT of $x[n]$ and extracting the unwrapped phase at the frequency of interest. Therefore, Eve determines the modulation type by observing the variation in the unwrapped phase at the detected vibration frequency over consecutive chirps.[5]

Packet Detection: Another key challenge for an asynchronous eavesdropper is finding the start time of a symbol sent by the underwater Alice. Conventionally, the coarse and fine synchronization between wireless nodes is achieved via known preambles [47]. Unfortunately, Eve may not know the period and content of such preambles between Alice and Bob and hence cannot rely on conventional techniques. Instead, Eve exploits the fundamental discrepancy between the hardware capabilities of acoustic transmitters and high bandwidth mmWave radars. Specifically, the chirp transmission rate at a mmWave radar is often drastically higher than the symbol time. Hence, Eve can track abrupt changes in the vibration profile and find the beginning of a symbol. In practice, we define a threshold for the minimum amount of vibration fluctuations and we use this threshold to determine the beginning of a packet. We highlight that the accuracy of this coarse and non-cooperative method is a function of radar's hardware, carrier (or vibration) frequency, and symbol rate.

Symbol Rate Estimation: Finally, Eve needs to estimate Alice's symbol rate to be able to correctly segment the temporal vibration into symbol intervals. A common approach is auto-correlating the baseband signal and assessing its local peaks and troughs [6]. This approach, however, proves less effective for Eve since these typically work well in low noise (high SNR) conditions. Distortion due to surface waves adds significant noise to the extracted phase data and accurate inference of symbol rate is difficult.

To address this challenge, SURF employs a new strategy that involves finding the optimal symbol period (T_s^*) such that the segmented symbols with that period have the maximum variance in terms of their information-carrying parameter (phase, amplitude, or spectral content depending on the modulation). Our rationale is that under the accurate value of symbol period (T_s^*), the time-domain segmented signals of $s(t, t + T_s)$ would only contain the information of one symbol. Hence, the variance of the measured phase/amp/spectral features across all symbols would be maximized.

More specifically, SURF estimates the symbol rate (i.e. $\frac{1}{T_s^*}$) through the following process: First, SURF finds a range of possible candidates for T_s. Next, it divides the baseband waveform into individual symbols s_i and extracts the parameter of interest (phase/frequency/amplitude) according to the estimated modulation type. SURF searches for the symbol period that maximizes the $variance$ of the modulation-specific parameter V as follows:

$$T_s^* = \underset{T_s}{\operatorname{argmax}} \ var\{V(s_i, T_s)\}, \tag{6}$$

where var denotes the variance operation and V represents the phases of s_i for PSK, the energy of s_i for ASK, and for FSK it encompasses the frequency of each symbol.

2.4.2 Non-Coherent Detection and Decision Making

Utilizing the previously estimated PHY layer information, SURF aims to empirically estimate the decision boundaries for detection in the absence of channel knowledge. Conventionally, such boundaries are estimated at the receiver through channel sounding to remove all sorts of non-idealities like multipath. Without channel information, SURF has no choice but to

[5] The inference and measurement of QAM are reserved for future work.

infer such boundaries by observing the temporal variations of the water vibration profile in a given observation window. Assume the decision boundaries are estimated after obtaining L symbols at Eve. There exists one fundamental trade-off: When L is very large, it is likely to have an equal number of symbols from each constellation point (i.e., L/M for M constellation points). This enables Eve to calculate M uniformly distributed boundaries. However, the caveat is that the channel might not remain coherent during this time hindering the accuracy of the estimated boundaries. In contrast, with a short observation time, only a few symbols will be obtained. While the channel is likely stable during short intervals, all symbols might not appear equally likely and it is more difficult to eliminate the noise due to a limited number of symbols. We will evaluate the impact of the observation window in §4.3.

For M-ASK modulation, let us denote the average value of the baseband signal in the i-th symbol period by Q_i. Eve finds the decision boundaries by computing the $100k/M - th$ percentile of the ordered Q_i values, when $k = \{1, ..., M-1\}$. A similar approach is used for M-PSK modulation, however, in this case, Q_i is defined as the phase of carrier frequency at the i-th symbol. It is important to note that FSK does not require boundary selection in a conventional sense. Instead, Eve only needs to compare the averaged baseband power at all carrier frequencies relative to each other. While it is evident from Eq. (2) that the vibration magnitude is inversely related to frequency, this relationship is deterministic and known a priori. Hence, Eve can equalize the power accordingly before detection and decision-making.

Since FSK does not require absolute decision boundary estimation, we expect it to be more prone to eavesdropping. Interestingly, due to the inherent robustness of FSK against channel fluctuations and its superior performance at low SNR regimes, it is often the modulation of choice in long-range underwater communications [33]. We will compare the security vulnerabilities under different modulations in §4.

2.5 Eavesdropping Regimes and Bounds

Intuitively, the likelihood of successful eavesdropping would depend on Alice's PHY layer parameters (carrier frequency, symbol rate, etc) as well as Eve's hardware specifications. Here, we explain this dependency with the goal of determining the range of PHY layer parameters which may lead to secure underwater communication. Fig. 5 summarizes the relevant FMCW parameters including the fast sampling frequency of radar's ADC (F_s), the chirp slope (s), chirp duration (T_c), idle time (T_i), and the number of ADC samples per chirp (N_s).

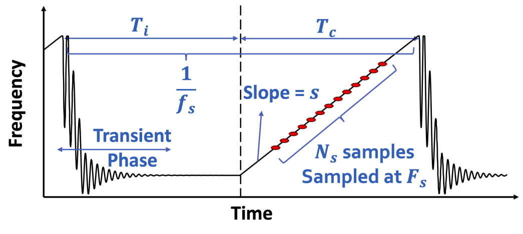


Figure 5: Important chirp parameters. Chirp configuration impacts the eavesdropping bounds.

First, vibration estimation is achieved by tracking the phase of multiple consecutive chirps, and the phase sampling rate is $f_s = \frac{1}{T_i+T_c}$. To maximize f_s, we need to either reduce T_i or T_c. Yet, T_i is not very flexible as it should be greater than the transient phase time as shown in Fig. 5. Further, we cannot make T_c arbitrarily small. This is because, if T_c is set too low, it will yield insufficient samples after de-chirping leading to poor range estimation. This is because, at a fixed ADC sampling rate of F_s, the number of samples (N_s) is determined by the chirp time, i.e., $N_s = T_c F_s$.

Second, for a center frequency of f_c and a bandwidth of W, according to Nyquist's criterion, $\frac{f_s}{2} \geq f_c + W/2$ must hold. Thus, the maximum detectable vibrating frequency is bounded by f_s. Furthermore, to avoid ambiguity and aliasing in vibration estimation, the phase evolution over several chirps should observe at least one complete period of the vibration and ideally many more. Due to hardware constraints, there is a limit on the number of continuous chirps that can be recorded, which we call N_p. Hence, for a vibration frequency f_c, we have: $\frac{1}{f_c} \ll \frac{N_p}{f_s}$. Lastly, generalizing this inequality to the data bandwidth of W and combining it with Nyquist's criterion, we can find the range of detectable vibrations at an airborne radar as:

$$\frac{f_s}{N_p} + \frac{W}{2} \ll f_c \leq \frac{f_s}{2} - \frac{W}{2}. \tag{7}$$

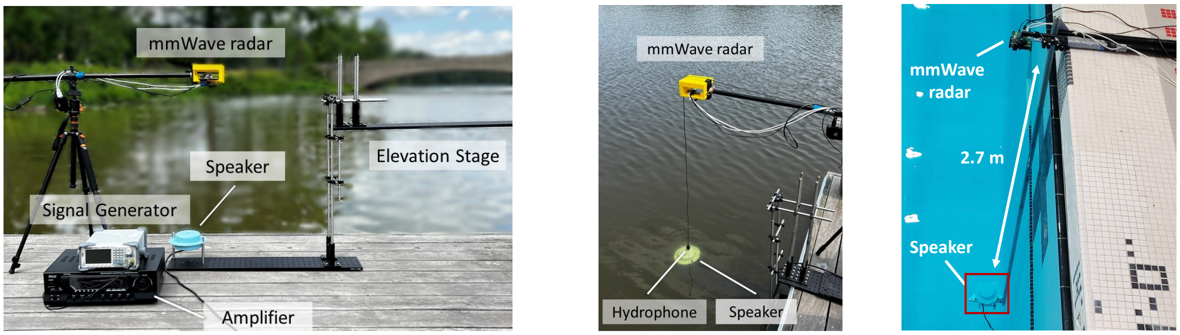
Eq. (7) suggests that if Alice chooses an extremely high or extremely low center frequency f_c, or employs a very high data rate/bandwidth W, she can prevent her information from being detected on the water surface. However, increasing the center frequency would cause high path loss at Bob and may hinder his successful reception. On the other hand, utilizing very low frequencies limits the achievable data rate at Bob.

2.6 Security Metrics

To quantify the ability of SURF to eavesdrop, we exploit well-known security metrics: *(i) Eve's BER:* We use BER at Eve as a measure of her ability to extract the correct bit sequence from the vibration signatures on the water. *(ii) Secrecy Capacity:* We look at Eve's eavesdropping *relative to* Bob's reception. In other words, high BER at Eve might, at first glance, suggest a secure channel, however, this security matters only if Bob can successfully decode his message. Therefore, for comparison purposes, we define secrecy capacity (SC) as:

$$SC = \log(1+SNR_{Bob}) - \log(1+SNR_{Eve}) [bits/sec/Hz], \tag{8}$$

where SNR_{Bob} and SNR_{Eve} are SNR at Bob and Eve in linear scale, respectively. Secrecy capacity is a well-known metric in



(a) (b) (c)

Figure 6: Experimental platform. (a) Different components of our setup in the lake experiments; (b) showing the underwater speaker (Alice - elevated for better visibility), hydrophone (Bob), and mmWave radar (Eve); (c) our pool setup.

information theory and refers to the maximum rate at which confidential or private information can be transmitted over a communication channel [26]. A higher secrecy capacity means the channel is more secure and vice versa.

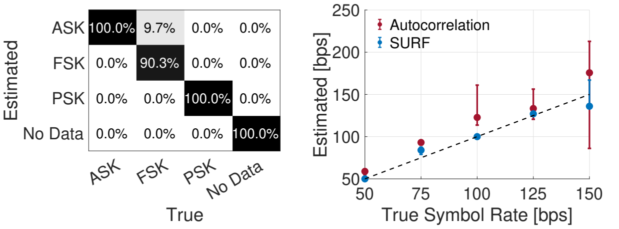
3 EXPERIMENTAL PLATFORM

We evaluate the vulnerabilities of underwater communications through extensive measurements in controlled environments and in the wild: *(i)* we perform controlled experiments in the lab using a 120 × 45 × 60 cm fish tank and motorized translational stages to precisely adjust and vary the position of the eavesdropper (radar); *(ii)* we run experiments in a natural lake to evaluate the performance of SURF in real-world settings. The experiments were conducted across different days, times of the day, and weather conditions; and *(iii)* we performed additional experiments in a swimming pool during the presence of other swimmers to test greater depths. The key components of our setup are shown in Fig. 6a.

Acoustic Transmitter (Alice). We use an Electro-Voice UW30 underwater speaker [12], which is connected to a Pyle PT270AIU amplifier [35]. To feed signals to the amplifier and the speaker, we use a combination of SIGLENT SDG 1032X waveform generator [43] for single-tone signals and the 3.5 mm jack on the Windows laptop for modulated data.

Acoustic Receiver (Bob). We use an Ambient ASF-2 hydrophone [1] connected to a MOTU M2 audio interface [29] to mimic the functionality of an acoustic receiver. All the recordings are done with a sampling rate of 48 KHz.

mmWave Radar (Eve). We use a low-cost commercial radar (TI IWR1642BOOST [18] and a DCA1000EVM [20]) for sensing the vibration signatures from the water surface. The recording of radar files is done through TI mmWave Studio software [21]. We set the radar parameters as follows: chirp time = 50μs, total number of chirps = 2.56×10^5, and total frame time = 12.8 s. These values were empirically obtained to satisfy the bounds derived in §2.5.



(a) (b)

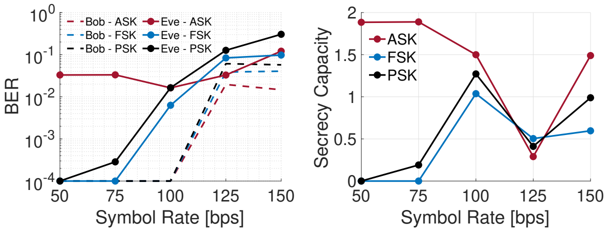
Figure 7: Inferring PHY parameters. SURF's performance in (a) detecting the modulation type and (b) estimating the symbol rate.

The experiments in the wild were conducted in the presence of standard conditions of wind and surface waves which represents a practical scenario where SURF would be used. Additionally, the radar was attached to a tripod that stood on top of a floating dock.

4 EVALUATION

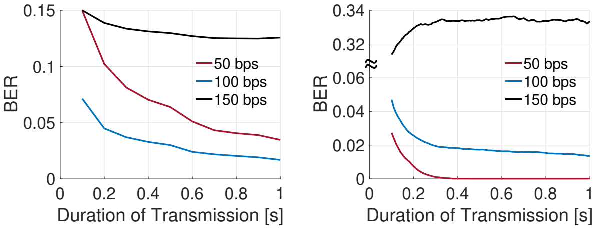
4.1 Inferring PHY Parameters of the Underwater Link

First, we explore SURF's capability to identify PHY parameters of underwater communication systems, notably without the cooperation of the underwater node. We evaluate the classification of modulation schemes and the estimation of symbol rates. We conduct lake-based experiments spanning various transmission modulations and symbol rates, ranging from 50 to 150 symbols per second. The speaker was submerged at a shallow depth (20 cm) beneath the water's surface. Experiments in §4.2 and §4.3 use the same setup. Fig. 7a presents the result of modulation detection via a confusion matrix encompassing four classes: ASK, FSK, PSK, and No Data. Eve achieves a classification accuracy of 97.58%. Misclassification errors, mostly mistaking FSK for ASK, are attributed to noise overshadowing one of the FSK tones in low SNR environments, thus resembling an ASK signal.



(a) (b)

Figure 8: SURF's eavesdropping performance. (a) BER at Bob and Eve under different modulation schemes and data rates; (b) corresponding secrecy capacity.



(a) (b)

Figure 9: Effect of duration of transmission. Demonstrating BER at Eve for (a) ASK and (b) PSK modulation as a function of the duration of continuous transmission at Alice.

Next, we evaluate the performance of SURF in estimating the symbol rate. Fig. 7b illustrates the estimated versus ground truth data rates across modulation types, with error bars representing the 25th and 75th percentiles. Eve obtains a root mean squared error (RMSE) of 9 bps across the different data rates. In comparison, the auto-correlation method has an RMSE of 72 bps.[6] This analysis demonstrates the superior accuracy of our symbol rate detection method compared to traditional auto-correlation schemes.

4.2 SURF's Eavesdropping Performance

In this section, we evaluate SURF's efficacy in intercepting the underwater communication between Alice and Bob. Our experiments, conducted in a lake, explored a range of modulations and data rates, transmitting 10^4 bits per configuration and recording the BER using both a hydrophone and our radar. The results are illustrated in Fig. 8a. As expected, higher data rates correlate with increased BER for Bob and Eve, attributed to reduced samples per symbol and subsequent SNR decrease.

Note that the rate of increase for BER varies between the radar and hydrophone across different bit rates, forming three separate data rate regimes. At lower data rates (50 and 75 bps), both Bob and Eve demonstrate low BERs with FSK and PSK modulations indicating that the channel is not secure. Increasing the data rate to 100 bps significantly increases Eve's BER without affecting Bob's, making this the optimal range for maximizing link secrecy. Further increases in data rate compromise Bob's reception quality, narrowing the performance gap with Eve and hence, degrading the link secrecy. Fig. 8b demonstrates that the secrecy capacity peaks at 100 bps for PSK and FSK, with a decline observed at both higher and lower rates, suggesting that Alice can optimize data rates to hinder Eve's eavesdropping attempts.

The eavesdropping performance for ASK is subpar, especially at lower bit rates, where Eve's detection accuracy suffers due to environmental factors (i.e. wind, surface waves, etc) affecting the vibration magnitude which is where the information is encoded. Despite employing high-pass filters to mitigate the effect of surface waves, the significant magnitude of environmental noise compared to acoustic vibrations degrades Eve's detection capability which means that ASK offers enhanced secrecy at lower data rates.

Interestingly, we observe an increase in secrecy capacity at 150 bps for all modulation schemes (which does not follow the general trend). This is because when Bob's reception is poor, the value of the secrecy rate is not an informative metric for evaluating the vulnerabilities of the system. In other words, even if the theoretical value of secrecy is high, operating in this regime is not desirable because of Bob's poor reception. Indeed, this is why we report the secrecy capability alongside the exact values of BER in Fig. 8a.

This comprehensive analysis reveals that by strategically adjusting data rates and modulation schemes, Alice can effectively minimize the risk of eavesdropping by Eve, with ASK modulation presenting a particularly intriguing option for enhancing communication secrecy at lower data rates.

4.3 Enhanced Secrecy with Bursty Transmission

We explained in §2.4.2 that Eve has to empirically estimate the decision boundaries by collecting several samples. Here, we evaluate the eavesdropping performance as a function of the total duration of Alice's transmission (i.e., the number of available samples). To this end, we vary the duration of transmission at Alice. With smaller transmission time (bursty traffic) fewer symbols are available at Eve for inference and decision making. Note that we consider long idle times in between consecutive transmissions. Hence, data cannot be aggregated across transmissions as the idle time is much longer than the channel coherence time, meaning the same decision boundaries would not apply.

Fig. 9 plots the BER for both PSK and ASK (we explained in §2.4 that FSK does not require absolute decision boundaries) as a function of the duration of the transmission. We observe that when the transmission duration is very short, Eve struggles to find the correct threshold due to random noise and the potential difference between the number of 0 and 1

[6] Increased error at 150 bps is mainly due to significant intersymbol interference (ISI).

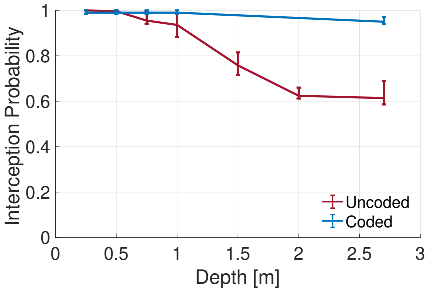


Figure 10: Impact of depth. The eavesdropping performance degrades as Alice is submerged deeper into the water. The error bars correspond to the 25th and 75th percentiles.

bits in that duration. Hence, Eve's BER improves when she can accumulate more symbols before forming the decision boundaries. Interestingly, the rate at which the BER changes depends on the data rate at Alice. In the low data rate regimes, even slightly increasing the transmission duration makes a big impact, whereas, for high data rates, there are already ample samples to make a reasonable estimate of the decision boundary. *A key security implication of this result is that Alice can enhance the security of its transmission by transmitting "bursty" traffic.* Specifically, Alice can send a small chunk of data, so that Eve cannot estimate the decision boundary correctly, and then wait for the channel to change before sending another chunk. The tradeoff is that Alice sacrifices the achievable goodput for communication.

4.4 Effect of Speaker's Depth

To understand the impact of depth on SURF's performance, we evaluated the system in a swimming pool setting and varied the depth of the underwater speaker from 25 cm to 2.7 m. We powered the speaker with the Crown XLi 2500 amplifier [2] and configured it to transmit BFSK modulated data (both coded[7] and uncoded) at a rate of 50 bps.

Fig. 10 shows our results. This figure plots the interception probability (defined as the ratio of the number of bits correctly decoded by Eve to the total number of data bits) as a function of depth. This metric represents Eve's likelihood of successfully intercepting messages from Alice/Bob. The error bars in the figure represent the 25th and the 75th percentiles. As expected, Eve's ability to eavesdrop worsens as the speaker's depth increases because of the reduced vibration amplitude on the surface. However, we can observe that SURF can reliably intercept with a probability of 93.6% when the speaker is at a depth of 1 m (for the uncoded case) and can intercept with a probability of 95% even when the depth of the speaker is increased to 2.7 m (coded case).[8] It is worth noting that even a small information leakage (or even detecting that there is a communication link) can have significant consequences for the victim in many sensitive underwater communication settings.

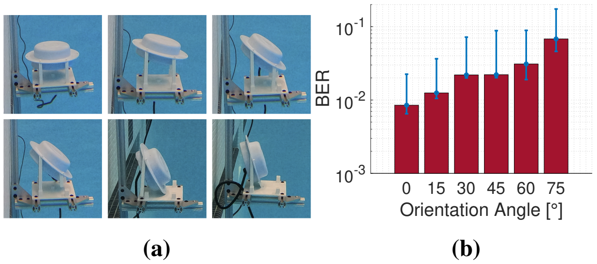


Figure 11: Impact of orientation. (a) The setup; (b) the BER at Eve as a function of different orientation angles.

Given the limitations of our setup, we cannot extend the depth beyond 2.7 m in the pool. To understand how this eavesdropping approach would perform in a more practical real-world scenario, we simulated the performance at a larger depth and a higher transmit power.

We first use the Sound Pressure Level (SPL) as a metric to characterize the power of our underwater transducer. SPL measures the pressure generated by an acoustic transducer at a distance of 1 m, relative to a reference pressure, such as 1 μPa, for a given input voltage. Specifically, our underwater speaker has an SPL of 157.1 dB re 1 μPa when driven with an input RMS voltage of 36.11 V [12]. In comparison, marine SONARs transmit with an SPL ranging from 210 to 240 dB re 1 μPa [50]. Since marine SONARs are significantly more powerful, they can be submerged much deeper while producing the same pressure at the water's surface as our underwater speaker, which operates at a shallower depth.

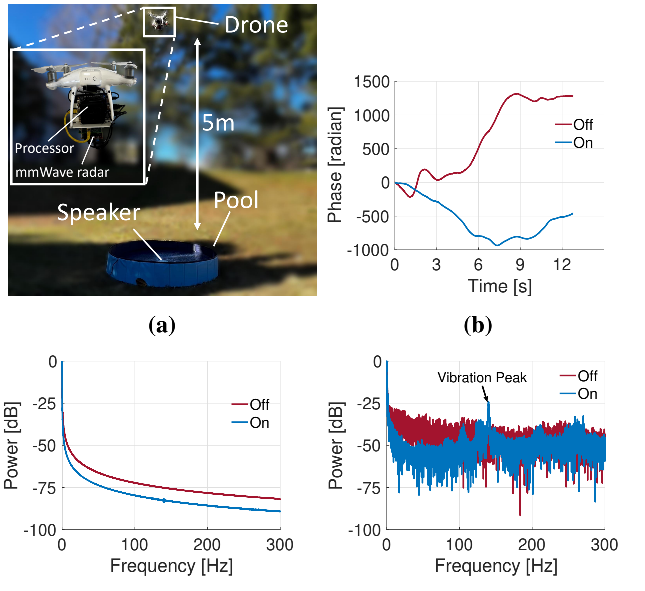
To be more precise, two speakers with SPL values of S_1 and S_2, submerged at depths d_1 and d_2, would generate the same pressure at the water's surface if $S_1 - S_2 = 20\log_{10}(d_1/d_2)$.[9]

For example, according to Fig. 10, SURF can intercept underwater communication with a probability of 93.6% for uncoded data at a depth of 1 m. Assuming a submarine transmits with an SPL of 220 dB re 1 μPa, we can calculate using the previous equation that this submarine's communication link can be intercepted with the same probability of 93.6% when it is submerged 1.4 km deep in the water. Furthermore, SURF can intercept underwater communication with a probability of 95% for coded data at a depth of 2.7 m. Using a similar calculation, we can show that SURF can achieve a 95% interception probability when the submarine transmits coded data at a depth of 3.8 km.

[7] The coded data was transmitted with a code rate of 0.1.

[8] In standard underwater communication protocols, coding is applied to packet headers, which is known [33].

[9] We use $20\log_{10}$(depth) because acoustic power decays as a function of $\frac{1}{d^2}$ with depth [46].



(a) (b) (c) (d)

Figure 12: Impact of drone vibration. (a) The setup; (b) large-scale drone movements over time; (c) vibration spectrum before processing; (d) vibration spectrum after processing.

4.5 Effect of Speaker Orientation

In previous experiments, the speaker faced the surface of the water. Here, we evaluate the impact of the speaker's orientation. In principle, if the speaker's transmissions are truly omnidirectional, the orientation of the speaker has little to no impact on the vibration patterns on the surface. However, in practice, tilting the speaker may yield a reduction of the vibration magnitude detected by Eve's radar. For this experiment, we positioned the speaker at a depth of 1 m and transmitted 2×10^3 BFSK modulated random bits at a rate of 50 bps. The measurements were repeated under the speaker's tilt angle of 0, 15, 30, 45, 60, and 75 degrees, as illustrated in Fig. 11a.

Fig. 11b shows the resulting BER at Eve as a function of the tilt angle. The BER gradually increases as the speaker tilts away from the water's surface due to its imperfect omnidirectional pattern, which reduces the amplitude of surface vibrations and hinders data extraction. This data was collected with the radar fixed directly above the speaker, meaning Eve decodes vibrations from above the speaker, not from the point of maximum vibrations, which shifts as the speaker rotates. Eve could improve interception by locating the maximum vibrations using the procedure explained in §2.3 and position herself to maximize SNR. However, a narrow beamwidth at Alice could still limit eavesdropping, as discussed in §6.

4.6 Effect of Drone Vibrations

So far we have only considered cases where the radar is perfectly stable. However, in practical scenarios, the radar is mounted on an airborne vehicle, e.g., a drone. The random movements of a hovering drone in the air (due to the wind) and the airflow generated by propellers create new challenges for SURF. Namely, since motion is relative, the drone displacement can distort the phase readings of the radar, as if the water is vibrating instead. Additionally, as the propellers of the drone push the air down to allow the drone to stay afloat, they might disturb the water's surface, further destroying the speaker's vibration signature. The amount of such disturbances is a function of the propellers' speed.

To remove such drone-related non-idealities, we exploit two key insights: First, the large-scale movement of the drone consists of much lower frequency components than the speaker's carrier frequency and hence can be mitigated using signal processing techniques. Second, the propeller effect on the water surface is negligible when the drone's altitude relative to the water is sufficiently high (i.e., above a certain minimum altitude threshold). Note that such minimum altitude threshold is a function of the propellers' speed and can be known by Eve in advance.

We placed a speaker in an inflatable outdoor pool filled with water at a depth of 15 cm. The speaker transmitted a single-tone signal with a frequency of 140 Hz. The airborne node (Eve) consists of a cheap COTS drone, DJI Phantom 4 Pro [11], carrying a LattePanda delta 3 processor [25] that communicates with the same radar as before. The setup is depicted in Fig. 12a. The drone is manually controlled to hover around 5 to 6 m to minimize the effect of propellers on the water's surface. To prove that Eve can detect vibrations while flying, we also repeated the same experiment while the speaker was turned off.

The results of this experiment are plotted in Fig. 12. Firstly, we can see the effect of the drone's movement on the unwrapped phase in Fig. 12b. The slow time-varying signal is tracking how the drone is moving as its distance to the water's surface changes. It also has the speaker vibrations added on top of it but it is not immediately visible. Note that the visible difference between the two cases of "On" and "Off" is mainly due to the different trajectories of the drone and not the speaker's vibration. The normalized frequency spectrum of these signals is plotted in Fig. 12c. As illustrated, the two cases share similar spectral contents with most of the power concentrated in low frequencies. We use signal-processing techniques to separate the speaker vibrations from the unwanted movements of the drone. Specifically, by exploiting a moving average filter, we first extract the movement of the drone and remove it from the time domain signal. Hence, we are left with the speaker's vibrations. The spectrum of the signal after post-processing is depicted in Fig. 12d. Clearly, the speaker vibrations at 140 Hz show a pronounced peak at the correct frequency and hence can be used by Eve for further demodulation and decoding.

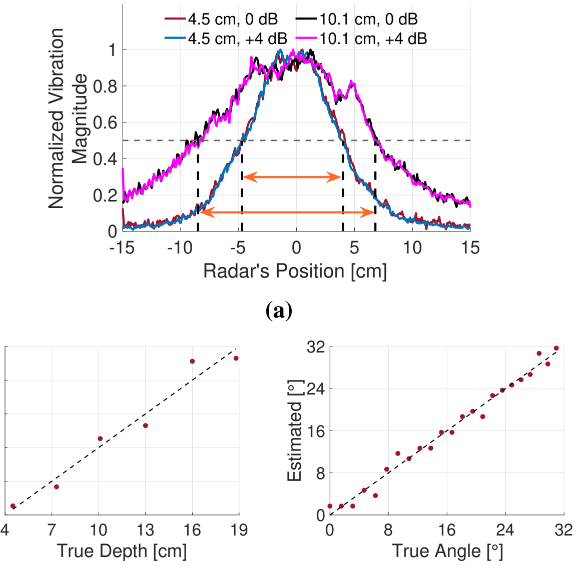
It is interesting to note that, although the speaker was placed at a shallow depth of 15 cm, we transmitted at low power using a signal with an RMS voltage of 3.8 V, which corresponds to an SPL of 137.6 dB re 1 μPa [12]. As discussed at the end of §4.4, a real submarine transmits with much higher acoustic power. Consequently, its vibrations can be detected at much greater depths than the 15 cm we tested with our system. A similar analysis to the one presented at the end of §4.4 suggests that the depth at which we could achieve comparable detection of the center frequency from marine SONARs is around 2 km. This implies that it is possible to detect the vibration profile from a drone hovering at a height of 5 m in real-world underwater scenarios.

This proof-of-concept experiment demonstrates that an airborne node with an FMCW mmWave radar can decode messages by detecting vibration frequencies using post-processing techniques. Once the frequency is identified, the temporal variation of the carrier can be monitored to demodulate bits, as shown in §4.2. Note that this experiment uses low-cost COTS equipment; without specialized hardware, such as more powerful radars or drones with higher payload capacities, decoding messages would be significantly more challenging. Advanced hardware could also extend the eavesdropper's range beyond 5 to 6 m, further reducing its visibility from the water. We leave this topic for future research.

4.7 Potential for Node Localization

Finally, we show the feasibility of pinpointing the area of vibration from the air through a combination of radar beamforming and mobility. We positioned the radar at a height of 1 m above the water's surface where the speaker was submerged 10 cm below the surface in the fish tank setup described in §3. The speaker transmits a single tone and we move the radar by 15 cm in both the positive and negative directions along the x axis with 1 mm increments using a motorized translational stage. At each point, the radar measures the amplitude of speaker vibrations on the water surface. We repeat the same experiment at different speaker depths and power levels.

First, Fig. 13a plots the *normalized* magnitude of vibration as a function of radar's position for two speaker depths and two transmit power levels. We observe that the diameter of vibration is solely a function of speaker depth and not the transmitted power. As expected, the diameter of vibration (shown with green arrows) increases with speaker depth because the acoustic beam impacts a larger area on the surface. This implies that one could create a one-to-one relationship between Alice's depth and the diameter of the vibration as seen by Eve. We emphasize that although the vibration magnitude depends on both the depth of the speaker and its transmitting power, normalized vibration magnitude is solely a function of depth. Such a model can be leveraged by the adversary for node localization. Fig. 13b shows that the estimated depth (extracted from the vibration diameter) matches closely to the ground truth speaker's depth in a controlled lab setting.



(a)

(b)

(c)

Figure 13: Localization performance. (a) Normalized vibration magnitude at different speaker depths and power levels; (b) estimated depth vs. true depth; (c) estimated angle vs. true angle. The dotted line in (b) and (c) represents the ideal case, i.e., no error.

Similarly, Eve applies beamforming to determine the AoA of the vibration center by implementing the beam search algorithm described in §2.3. Fig. 13c shows that the estimated angles align with the ground truth. Although these initial results are promising, further modeling and measurements are required to fully assess the technique's performance and limitations for localization, which we leave for future work.

5 RELATED WORK

Cross-Medium Communication. Various approaches establish communication links between air and water mediums, such as visible light methods [4, 7]. Closest to our work are TARF-based systems [36, 38, 46], utilizing mmWave radars to capture underwater speaker vibrations. Unlike prior works assuming transmitter-receiver cooperation, we adopt an adversarial perspective, unveiling eavesdropping tactics for non-cooperative entities. We evaluate the security performance under various PHY parameters and suggest appropriate countermeasures. Finally, we test cross-medium communication in a lake for the first time.

Vibration-Based Eavesdropping. Many researchers have taken advantage of the fact that acoustic signals vibrate their surrounding objects and have used sensors to pick up these vibrations and eavesdrop on potential victims [48, 49, 51]. However, in all existing efforts, the communication parties and the adversary belong to the same medium (air) and do not account for the challenges of out-of-medium eavesdropping.

Cross-Medium Eavesdropping. Conventional systems for cross-medium eavesdropping required deploying passive sonobuoys on the water surface to listen to messages underwater using hydrophones and then relay them to an aircraft using an RF transceiver [13]. Later approaches relied on (large) meta-material patches which are also deployed on the water surface and aim to amplify underwater vibrations to be picked up by a remote airborne receiver [27, 41, 52]. The challenge with these systems is that the eavesdropper itself becomes detectable (much more detectable than a drone that is above the surface) from an underwater transceiver. Moreover, these approaches have a limited coverage area of eavesdropping in comparison to drones that can, in principle, fly over and scan wider areas of the surface.

6 LIMITATIONS AND DISCUSSIONS

Effect of Radar's Height. Our experiments were performed at radar heights ranging from 1 m to 5 m. As the research evolves, it will be important to evaluate SURF at different heights and RF power levels. COTS radars, which have a transmit power of 12.5 dBm, could accurately measure vibrations with a negligible error (about 3.174 μm) at 5 m height when measuring a vibration amplitude of 100 μm [15]. Navy radars are much more powerful (having more than 10 kW of transmit power [9, 10]), and thus would be expected to detect these vibrations from significantly greater distances above the water.

Potential Counter Measures. The secrecy of underwater communications can be improved through several strategies: First, Alice and Bob could form a null toward the water surface to block pressure waves. However, this requires real-time orientation estimation and complex arrays, increasing complexity and power use. Null forming might not fully stop distant waves due to beam divergence. Another rather counter-intuitive approach is to place Alice close to the water's surface to minimize the area of impact (area of vibration on the water) hoping that Eve's radar cannot pick it up. We will investigate effective countermeasure strategies in the future.

Impact of Encryption. Although encryption could potentially complicate the eavesdropping process of SURF , it is often challenging to implement in power-constrained devices [39, 40], particularly in underwater environments. Encryption schemes are generally divided into two categories: asymmetric (no shared key) and symmetric (shared key). Asymmetric encryption requires significant computational power, leading to high energy consumption, which is unsuitable for the resource-limited devices commonly used in underwater environments [30]. Additionally, underwater sensor nodes often rely on batteries that are difficult to recharge or replace in harsh conditions [32], making asymmetric encryption impractical for many underwater applications. Symmetric encryption, such as AES, uses a pre-shared key among nodes for the encryption and decryption of messages. However, due to the ad-hoc nature of underwater communications, a significant challenge is presented when a new node without a pre-shared key attempts to join the network. Conventional solutions from terrestrial networks either require a pre-shared secret or impose communication overhead [44] or require a third party that is trusted by both nodes which might not always be available. As a result, symmetric encryption is also often not used in underwater communications. Additionally, even with encryption, traffic pattern analysis enabled by SURF can reveal sensitive information, such as packet length, timestamps, message frequency, and the identities of communicating parties. This makes encryption less effective in protecting communication between underwater nodes, especially in contexts involving naval assets.

Scaling to Real Naval Settings. In this paper, we have demonstrated the feasibility of a novel eavesdropping methodology targeting acoustic underwater nodes. However, deploying this technique in real-world naval scenarios presents a multitude of challenges that remain unresolved. Among these are the mobility of underwater nodes (that can cause Doppler shifts to the transmitted data, further complicating the demodulation process for the eavesdropper), large natural waves in the ocean, the challenge of operating at greater depths, and the need to accommodate increased data rates (e.g. Eve could use a radar with a higher sampling rate to eavesdrop on higher data rates - as discussed in §2.5), frequencies, and higher-order and more complex modulation schemes. Moreover, as the research evolves, more sophisticated techniques and hardware can be developed that allow for joint drone-based localization and detection during flight to enable efficient eavesdropping of sub-sea transmitters.

7 CONCLUSION

This paper presents SURF the first system that demonstrates eavesdropping of underwater links with an out-of-medium non-cooperative adversary. We demonstrate the feasibility of such attack using low-cost COTS mmWave radar. By sensing subtle surface disruptions, SURF intercepts underwater communication signals, without prior knowledge of channel and PHY parameters. Our rigorous evaluation validates its efficacy in diverse settings, including natural lakes. This paper holds important insights for securing underwater communication links, from maritime defense and submarine warfare to oil and gas exploration, search and rescue, mining, and aquatic species conservation.

8 ACKNOWLEDGEMENTS

We thank our shepherd and the anonymous reviewers for their constructive feedback. This work was supported by NSF, ONR, and Princeton's Project X Innovation Fund.

REFERENCES

[1] Ambient. 2023. ASF-2 MKII - Miniature Hydrophone. Available at https://ambient.de/en/brands/ambient/709/asf-2-mkii-miniature-hydrophone-w/48v-phantom-power-supply.

[2] Crown Audio. 2015. XLi 2500. Available at https://www.crownaudio.com/en-US/products/xli-2500.

[3] Brian Borowski. 2009. Characterization of a Very Shallow Water Acoustic Communication Channel. In *OCEANS 2009*. 1–10. https://doi.org/10.23919/OCEANS.2009.5422360

[4] Charles J. Carver, Zhao Tian, Hongyong Zhang, Kofi M. Odame, Alberto Quattrini Li, and Xia Zhou. 2020. AmphiLight: Direct Air-Water Communication with Laser Light. In *17th USENIX Symposium on Networked Systems Design and Implementation (NSDI 20)*. USENIX Association, Santa Clara, CA, 373–388. https://www.usenix.org/conference/nsdi20/presentation/carver

[5] Josko A Catipovic. 1990. Performance Limitations in Underwater Acoustic Telemetry. *IEEE Journal of Oceanic Engineering* 15, 3 (1990), 205–216.

[6] Y.T. Chan, B.H. Lee, R. Inkol, and F. Chan. 2009. Estimation of symbol rate from the autocorrelation function. In *2009 Canadian Conference on Electrical and Computer Engineering*. 547–550. https://doi.org/10.1109/CCECE.2009.5090190

[7] Yifei Chen, Meiwei Kong, Tariq Ali, Jiongliang Wang, Rohail Sarwar, Jun Han, Chaoyang Guo, Bing Sun, Ning Deng, and Jing Xu. 2017. 26 m/5.5 Gbps Air-water Optical Wireless Communication Based on an OFDM-Modulated 520-nm Laser Diode. *Optics express* 25, 13 (2017), 14760–14765.

[8] Mandar Chitre, Shiraz Shahabudeen, Lee Freitag, and Milica Stojanovic. 2008. Recent Advances in Underwater Acoustic Communications & Networking. *OCEANS 2008* (2008), 1–10.

[9] Chae K. Chong and William L. Menninger. 2010. Latest Advancements in High-Power Millimeter-Wave Helix TWTs. *IEEE Transactions on Plasma Science* 38, 6 (2010), 1227–1238. https://doi.org/10.1109/TPS.2010.2041940

[10] B.G. Danly, J. Cheung, V. Gregers-Hansen, G. Linde, and M. Ngo. 2002. WARLOC: A High-power Millimeter-wave Radar. In *Twenty Seventh International Conference on Infrared and Millimeter Waves*. 233–234. https://doi.org/10.1109/ICIMW.2002.1076170

[11] DJI. 2023. Phantom 4 Pro Drone. Available at https://www.dji.com/phantom-4-pro-v2.

[12] Electro-Voice. 2015. UW30 - Underwater Loudspeaker. Available at https://www.lubell.com/products/uw30pa/.

[13] Christine Erbe and Jeanette A Thomas. 2022. *Exploring Animal Behavior Through Sound: Volume 1: Methods*. Springer Nature.

[14] Yasaman Ghasempour, Claudio RCM Da Silva, Carlos Cordeiro, and Edward W Knightly. 2017. IEEE 802.11 ay: Next-Generation 60 GHz Communication for 100 Gb/s Wi-Fi. *IEEE Communications Magazine* 55, 12 (2017), 186–192.

[15] Junchen Guo, Yuan He, Chengkun Jiang, Meng Jin, Shuai Li, Jia Zhang, Rui Xi, and Yunhao Liu. 2023. Measuring Micrometer-Level Vibrations With mmWave Radar. *IEEE Transactions on Mobile Computing* 22, 4 (2023), 2248–2261. https://doi.org/10.1109/TMC.2021.3118349

[16] Robert Headrick and Lee Freitag. 2009. Growth of Underwater Communication Technology in the US Navy. *IEEE Communications Magazine* 47, 1 (2009), 80–82.

[17] John Heidemann, Wei Ye, Jack Wills, Affan Syed, and Yuan Li. 2006. Research Challenges and Applications for Underwater Sensor Networking. In *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006.*, Vol. 1. IEEE, 228–235.

[18] Texas Instruments. 2017. IWR1642 - Single-chip 76-GHz to 81-GHz mmWave Sensor Integrating DSP and MCU. Available at https://www.ti.com/product/IWR1642.

[19] Texas Instruments. 2023. AWR1843 - Single-chip 76-GHz to 81-GHz Automotive Radar Sensor Integrating DSP, MCU and Radar Accelerator. Available at https://www.ti.com/product/AWR1843.

[20] Texas Instruments. 2023. DCA1000EVM - Real-time Data-capture Adapter for Radar Sensing Evaluation Module. Available at https://www.ti.com/tool/DCA1000EVM.

[21] Texas Instruments. 2023. MmWave Studio. Available at https://www.ti.com/tool/MMWAVE-STUDIO.

[22] Daniel B Kilfoyle and Arthur B Baggeroer. 2000. The State of the Art in Underwater Acoustic Telemetry. *IEEE Journal of oceanic engineering* 25, 1 (2000), 4–27.

[23] Blair Kinsman. 1984. *Wind Waves: Their Generation and Propagation on the Ocean Surface*. Courier Corporation.

[24] Atsutse Kludze, Rabi Shrestha, Chowdhury Miftah, Edward Knightly, Daniel Mittleman, and Yasaman Ghasempour. 2022. Quasi-Optical 3D Localization using Asymmetric Signatures above 100 GHz. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 120–132.

[25] LattePanda. 2023. LattePanda 3 Delta. Available at https://www.lattepanda.com/lattepanda-3-delta.

[26] S. Leung-Yan-Cheong and M. Hellman. 1978. The Gaussian Wire-tap Channel. *IEEE Transactions on Information Theory* 24, 4 (1978), 451–456. https://doi.org/10.1109/TIT.1978.1055917

[27] Jingjing Liu, Zhengwei Li, Bin Liang, Jian-Chun Cheng, and Andrea Alù. 2023. Remote Water-to-Air Eavesdropping with a Phase-Engineered Impedance Matching Metasurface. *Advanced Materials* 35, 29 (2023), 2301799. https://doi.org/10.1002/adma.202301799 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.202301799

[28] M. S. Martins, C. Barardo, T. Matos, L. M. Gonçalves, J. Cabral, A. Silva, and S. M. Jesus. 2017. High Frequency Wide Beam PVDF Ultrasonic Projector for Underwater Communications. In *OCEANS 2017 - Aberdeen*. 1–5. https://doi.org/10.1109/OCEANSE.2017.8084677

[29] MOTU. 2023. MOTU M2 Audio Interface. Available at https://motu.com/de/products/m-series/m2/.

[30] Tingyuan Nie, Lijian Zhou, and Zhe-Ming Lu. 2014. Power Evaluation Methods for Data Encryption Algorithms. *IET software* 8, 1 (2014), 12–18.

[31] United States Department of Defense Navy. 2022. Directional Acoustic Communications Transmitters. Available at https://www.sbir.gov/node/2101761.

[32] Chunyan Peng, Xiujuan Du, Keqin Li, and Meiju Li. 2016. An Ultra-Lightweight Encryption Scheme in Underwater Acoustic Networks. *Journal of Sensors* 2016, 1 (2016), 8763528.

[33] John Potter, Joao Alves, Dale Green, Giovanni Zappa, Ivor Nissen, and Kim McCoy. 2014. The JANUS Underwater Communications Standard. In *2014 underwater communications and networking (UComms)*. IEEE, 1–4.

[34] James Preisig. 2007. Acoustic Propagation Considerations for Underwater Acoustic Communications Network Development. *SIGMOBILE Mob. Comput. Commun. Rev.* 11, 4 (oct 2007), 2–10. https://doi.org/10.1145/1347364.1347370

[35] PyleUSA. 2023. PT270AIU - 300 Watt Stereo Receiver. Available at http://manuals.pyleusa.com/PDF/PT270AIU_fd115492-a689-4c26-a5ee-8695a66c68d8.pdf.

[36] Jingyu Qian, Fengzhong Qu, Jiayi Su, Yan Wei, Mingyuan Cheng, Honghui Guo, Jiang Zhu, and Jie Wang. 2023. Theoretical Model and Experiments of Focused Phased Array for Cross-Medium Communication in Misaligned Transmitter/Receiver Scenarios. *IEEE Journal of Oceanic Engineering* (2023), 1–14. https://doi.org/10.1109/JOE.2023.3263202

[37] Gang Qiao, Songzuo Liu, Zongxin Sun, and Feng Zhou. 2013. Full-duplex, Multi-user and Parameter Reconfigurable Underwater Acoustic Communication Modem. In *2013 OCEANS - San Diego*. 1–8. https:

//doi.org/10.23919/OCEANS.2013.6741096

[38] Fengzhong Qu, Jingyu Qian, Jie Wang, Xuesong Lu, Minhao Zhang, Xuerui Bai, Zhouhua Ran, Xingbin Tu, Zubin Liu, and Yan Wei. 2021. Cross-medium Communication Combining Acoustic Wave and Millimeter wave: Theoretical Channel Model and Experiments. *IEEE Journal of Oceanic Engineering* 47, 2 (2021), 483–492.

[39] Hamed Rahmani and Aydin Babakhani. 2020. An Integrated Battery-Less Wirelessly Powered RFID Tag with Clock Recovery and Data Transmitter for UWB Localization. In *2020 IEEE/MTT-S International Microwave Symposium (IMS)*. IEEE, 460–463.

[40] Hamed Rahmani, Darshan Shetty, Mahmoud Wagih, Yasaman Ghasempour, Valentina Palazzi, Nuno B Carvalho, Ricardo Correia, Alessandra Costanzo, Dieff Vital, Federico Alimenti, et al. 2023. Next-Generation IoT Devices: Sustainable Eco-Friendly Manufacturing, Energy Harvesting, and Wireless Connectivity. *IEEE Journal of Microwaves* 3, 1 (2023), 237–255.

[41] Zhang Shaocong, Zhu Jiahui, Li Chenyang, Weng Jiaxuan, Wang Yanfeng, and Wang Yuesheng. 2023. Investigation on Modulation of Acoustic Waves Through Water-air Interface by Combined Metasurfaces. *Chinese Journal of Theoretical and Applied Mechanics* 56, 1 (2023), 1–14.

[42] Charles H Sherman and John L Butler. 2007. *Transducers and Arrays for Underwater Sound*. Vol. 4. Springer.

[43] SIGLENT. 2023. SDG1032X Waveform Generator. Available at https://siglentna.com/product/sdg1032x/.

[44] George Sklivanitis, Konstantinos Pelekanakis, Seçkin Anıl Yıldırım, Roberto Petroccia, Joao Alves, and Dimitris A Pados. 2021. Physical Layer Security Against an Informed Eavesdropper in Underwater Acoustic Channels: Reconciliation and Privacy Amplification. In *2021 Fifth Underwater Communications and Networking Conference (UComms)*. IEEE, 1–5.

[45] Bo Tan, Elena Simona Lohan, Bo Sun, Wenbo Wang, Taylan Yesilyurt, Christophe Morlaas, Carlos David Morales Pena, Kanaan Abdo, Fathia Ben Slama, Alexandre Simonin, et al. 2022. Improved Sensing and Positioning via 5G and mmWave radar for Airport Surveillance. *arXiv preprint arXiv:2202.13650* (2022).

[46] Francesco Tonolini and Fadel Adib. 2018. Networking across Boundaries: Enabling Wireless Communication through the Water-Air Interface. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication* (Budapest, Hungary) *(SIGCOMM '18)*. Association for Computing Machinery, New York, NY, USA, 117–131. https://doi.org/10.1145/3230543.3230580

[47] F. Tufvesson, O. Edfors, and M. Faulkner. 1999. Time and Frequency Synchronization for OFDM Using PN-sequence Preambles. In *Gateway to 21st Century Communications Village. VTC 1999-Fall. IEEE VTS 50th Vehicular Technology Conference (Cat. No.99CH36324)*, Vol. 4. 2203–2207 vol.4. https://doi.org/10.1109/VETECF.1999.797329

[48] Chao Wang, Feng Lin, Tiantian Liu, Ziwei Liu, Yijie Shen, Zhongjie Ba, Li Lu, Wenyao Xu, and Kui Ren. 2022. mmPhone: Acoustic Eavesdropping on Loudspeakers via mmWave-characterized Piezoelectric Effect. In *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*. 820–829. https://doi.org/10.1109/INFOCOM48880.2022.9796806

[49] Chao Wang, Feng Lin, Tiantian Liu, Kaidi Zheng, Zhibo Wang, Zhengxiong Li, Ming-Chun Huang, Wenyao Xu, and Kui Ren. 2022. MmEve: Eavesdropping on Smartphone's Earpiece via COTS MmWave Device. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking* (Sydney, NSW, Australia) *(MobiCom '22)*. Association for Computing Machinery, New York, NY, USA, 338–351. https://doi.org/10.1145/3495243.3560543

[50] Yuming Zeng, Siyi Shen, and Zhiwei Xu. 2023. Water Surface Acoustic Wave Detection by a Millimeter Wave Radar. *Remote Sensing* 15, 16 (2023), 4022.

[51] Jia Zhang, Yinian Zhou, Rui Xi, Shuai Li, Junchen Guo, and Yuan He. 2022. AmbiEar: MmWave Based Voice Recognition in NLoS Scenarios. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 151 (sep 2022), 25 pages. https://doi.org/10.1145/3550320

[52] Shao-Cong Zhang, Hong-Tao Zhou, Xiao-Tong Gong, Yan-Feng Wang, and Yue-Sheng Wang. 2024. Discrete Metasurface for Extreme Sound Transmission Through Water-air Interface. *Journal of Sound and Vibration* 575 (2024), 118269.