KITE: Exploring the Practical Threat from Acoustic Transduction Attacks on Inertial Sensors

Ming Gao[1,2], Lingfeng Zhang[1], Leming Shen[3], Xiang Zou[1,4], Jinsong Han[1,2*], Feng Lin[1,2], Kui Ren[1,5]
[1]Zhejiang University, Hangzhou, China
[2]ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou, China
[3]The Hong Kong Polytechnic University, Hong Kong, China
[4]Xi'an Jiaotong University, Xi'an, China
[5]Zhejiang Provincial Key Laboratory of Blockchain and Cyberspace Governance, Hangzhou, China
{gaomingppm,lingfengzhang,hanjinsong,flin,kuiren}@zju.edu.cn,leming.shen@connect.polyu.hk,Xiang_Zou@stu.xjtu.edu.cn

ABSTRACT

In cyber-physical systems, inertial sensors are the basis for identifying motion states and making actuation decisions. However, extensive studies have proved the vulnerability of those sensors under acoustic transduction attacks, which leverage malicious acoustics to trigger sensor measurement errors. Unfortunately, the threat from such attacks is not assessed properly because of the incomplete investigation on the attack's potential, especially towards multiple-degree-of-freedom systems, e.g., drones. To thoroughly explore the threat of acoustic transduction attacks, we revisit the attack model and design a new yet practical acoustic modulation-based attack, named KITE. Such an attack enables stable and controllable injections, even under frequency offset based distortions that limit the effect of prior attacking approaches. KITE exploits the potential threat of transduction attacks without the need of strengthening attackers' abilities. Furthermore, we extend the attack surface to multiple-degree-of-freedom systems, which are more widely deployed but ignored by prior work. Our study also covers the scenario of attacking moving targets. By revealing the practical threat from acoustic transduction attacks, we appeal for both the attention to their harm and necessary countermeasures.

CCS CONCEPTS

• **Security and privacy → Security in hardware**; **Mobile and wireless security**.

KEYWORDS

Cyber-physical system, inertial sensors, acoustic transduction attacks, spoofing attacks, IoT security

ACM Reference Format:
Ming Gao[1,2], Lingfeng Zhang[1], Leming Shen[3], Xiang Zou[1,4], Jinsong Han[1,2*], Feng Lin[1,2], Kui Ren[1,5]. 2022. KITE: Exploring the Practical Threat from Acoustic Transduction Attacks on Inertial Sensors. In *The 20th ACM Conference on Embedded Networked Sensor Systems (SenSys '22), November 6-9, 2022, Boston, MA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3560905.3568532

*Jinsong Han is the corresponding author.

SenSys '22, November 6-9, 2022, Boston, MA, USA

ACM ISBN 978-1-4503-9886-2/22/11…$15.00
https://doi.org/10.1145/3560905.3568532

1 INTRODUCTION

Cyber-physical systems (CPSs) are widely deployed in various areas, including consumer electronics, health care, industry, and military deployment [35, 37]. These systems, such as mobile devices (e.g., smartphones) and actuation systems (e.g., drones), rely on inertial sensors (i.e., accelerometers and gyroscopes) to identify motion states and make actuation decisions. As the increasing popularity of motion-driven applications, inertial sensors play integral roles [1].

Unfortunately, these inertial sensors have been reported to be vulnerable to acoustic interference with a specific frequency, namely, natural frequency [18, 19, 34, 51, 63, 75, 76]. With this property, attackers can disturb the operation of target systems. For example, a drone would crash under ultrasonic interference [63]. Even worse, existing countermeasures to those attacks, e.g., acoustic isolation [17, 64], seem ineffective in an embedded environment [71].

With the above effect, attackers will naturally develop more strategical attacks to maliciously control CPSs. State-of-the-art (SOTA) research has proposed to deliberately modulate acoustic signals [70, 71], instead of denial of service (DoS) attacks via disordered noise [63, 75]. However, the potential of such sophisticated attacks is not well studied due to the limited attacking scopes and scenarios targeted by existing approaches. First, existing approaches merely focus on **single-axis** inertial sensors. These targets' trajectories are restricted in the simplest motion mode, i.e., moving along ONE direction for an accelerometer or around ONE axis in a plane for a gyroscope. Second, existing research only involves **stationary** targets and ignores the influence of motion. In real-world scenarios, however, the systems' motion mode would be more complex. For example, a drone could fly with six degrees of freedom, consisting of three-dimensional linear motion and rotation. Therefore, besides the lack of effective defense, the threat level of such attacks is still unclear and not fully investigated. It boils down to a key problem: *to what degree acoustic transduction attacks can affect CPSs in practice.*

Answering this question is difficult because it remains challenging to realize the strategical acoustic transduction attack in real CPS systems, e.g., those with multiple-degree-of-freedom (MDOF). When extended from the single-axis to the multi-axis, i.e., injecting desired components of false signals into multiple axes respectively, the attack seems only to be able to disturb the target, instead of

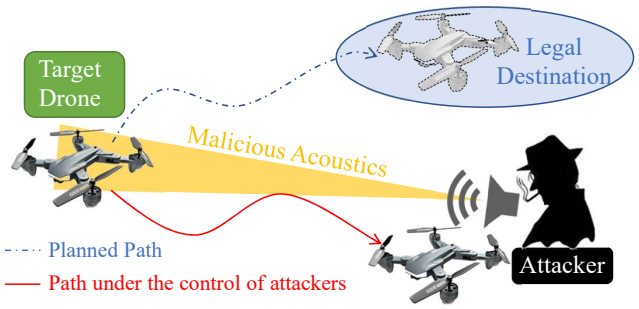


Figure 1: Acoustic transduction attacks aim at control over CPSs by spoofing inertial sensors.

freely controlling its movement according to the attacker's desire. This is because the injection on one axis would influence the components on other axes [25]. Thus, existing attacking approaches [70, 71] cannot guarantee to yield desired output on each axis of inertial sensors. As a result, the attackers cannot accurately control the target's orientation. Recalling the example of a drone, attackers want to tamper with the drone's yaw angle to modify its trajectory, but it may crash due to unexpectedly injected rolling or pitching. To fully understand the ability of attackers to real CPSs, we investigate the distribution of false signals in multi-axis sensors and leverage their spatial features for enabling stable adversarial controls. We extend the scope of acoustic transduction attacks to the multi-axis inertial sensors so as to cover commonly-seen MDOF systems, as illustrated in Fig. 1. In this way, we realize a sophisticated attack, namely KITE, which effectively controls drone-like systems.

To make the attack more realistic, we further consider a very common case, in which the target is **moving**. In this case, a slight distance variation between the malicious acoustic source and the moving target leads to nontrivial *phase fluctuation* which distorts false signals significantly. Moreover, motion signals may *couple* with false signals, producing abundant noise. The impacts of the attacks seem to be constrained against moving systems. The threat level of such attacks might be badly underestimated while their potential has not been fully dug. We realize remote attacks on moving targets with single-axis sensors. Furthermore, we explore the possibility of spoofing multi-axis sensors under the motion influence, based on our observation that transduction attacks are effective using acoustics that travels through solid. In many cases, there exists a possibility that attackers can perform a one-shot physical contact with target systems. For example, attackers can stealthily place a malicious unit under a mask of legal accessories (e.g., protective shells [23]). For such attacking scenarios, we design a malicious unit and enable adversarial control over moving systems.

Combining the above efforts together, we thoroughly display the practical threat of acoustic transduction attack. We redefine the threat model to make attackers more realistic and propose a novel method of acoustic modulation. Note that we adopt the identical attackers' abilities to existing work [30, 63, 70, 71] without any enhancement. Our method involves accurate frequency and phase estimation. It supports a stable and controllable injection (with identical effect to the attacks in [70]). In particular, we realize the automatic offset compensation, without which false signals would distort [25, 71] and the attack's effect would be constrained. In comparison, existing approaches [70, 71] merely work in ideal or well-controlled conditions (i.e., without sampling rate drifts), which are rare in reality. As a result, previous acoustic transduction attacks can hardly be performed on real IoT devices. In KITE, we propose a novel acoustic modulation method, which allows stable false injections, free from the tight constraints of no sampling rate drift. With this method, KITE allows the attackers to control the speed and orientation of a drone-like target. To our best knowledge, we are the first to accomplish adversarial control over moving targets using acoustic transduction attacks. Extensive evaluations demonstrate the effectiveness of KITE when attacking commercial devices, including a drone with the most popular autopilot (i.e., Pixhawk 4).

Our contribution can be summarized as follows:

- We perform a comprehensive analysis on practical threats to CPSs from acoustic transduction attacks. We extend the attack surface to MDOF systems, e.g., drones.
- We propose a new acoustic modulation method to manipulate the injected false signals as the attackers expect. By fully exploiting acoustic attacks, KITE is able to hijack the target CPSs using stable signal injections. This reflects the real risk of malicious control, which is underestimated in prior research.
- We model the response of moving systems under acoustics, which has not been studied in the literature. Accordingly, we launch KITE for the adversarial control in a more common scenario involving moving systems.

2 INERTIAL SENSORS

Inertial sensors comprise accelerometers for observing linear acceleration and gyroscopes for detecting angular velocity. They share a similar damping structure [5, 41]. The structure is composed of a movable seismic mass connecting with springs and capacitor electrodes. In an accelerometer, the linear acceleration causes the displacement according to Hooke's law. Then ,the displacement is converted into an electrical signal due to the proportional capacitance change. In a gyroscope, the angular velocity induces the Coriolis acceleration [31]. Similar to the process in an accelerometer, the Coriolis acceleration is transduced into an electrical signal. After amplification, filtering, and sampling, these motion-related electrical signals are transformed into digital signals. They jointly provide control systems with real-time inertial information.

Unfortunately, inertial sensors are sensitive to acoustic injections due to their damping structure and resonant features [18, 34, 76]. The resonance effect would occur when external signals' frequency matches or approaches the sensor's natural frequency. These natural frequencies usually fall into the acoustic band, about 0~10 kHz for accelerometers and 18~30 kHz for gyroscopes. Such a band is covered by speakers or transducers that are available to attackers. Accordingly, researchers pursue not only DoS attacks that disturb inertial sensors' operation and induce breakdowns or crashes [23, 63, 75], but also adversarial control on CPSs [70, 71]. These attacks succeed in manipulating stationary targets, e.g., self-balancing human transporters, self-balancing robots, and smartphones. They can also be applied to interfere in camera or computer vision based object detection systems by spoofing inertial sensors of image stabilizers [30]. However, SOTA attacks either select a target with a single-axis sensor or care only one axis of a multi-axis sensor.

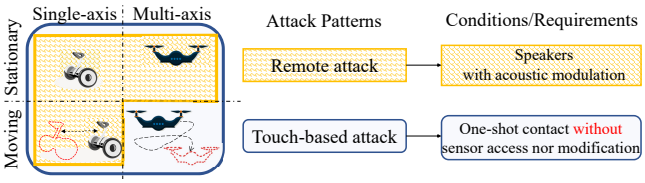


Figure 2: Threats of the proposed attacks.

3 THREAT ANALYSIS

We detail possible attack scenarios to investigate latent threats from acoustic transduction attacks. To make attackers more realistic, we refine their capability and attack means.

3.1 Attack Scenarios

We divide possible attack scenarios into 2×2 types, corresponding to the target systems' degree of freedom and motion state.

3.1.1 Single- vs. multi-axis sensors. Single-axis sensors only support single degree of freedom along or around one axis (represented by the obliquity sensor in a self-balancing robot [71]). These systems can only travel forth/back, rotate around one axis in a plane, or move in the pattern of combining the former two. Such systems are merely embedded with single-axis accelerometers, gyroscopes, or both (except redundant axes for anomaly detection, e.g., collisions).

MDOF systems, the more common systems, can move freely in space. A drone, a representative of those complex systems, is embedded with a three-axis accelerometer and a three-axis gyroscope. Although Tu et al. [71] test on systems based on multi-axis sensors (e.g., smartphones and stabilizers), they only care about outputs on one axis. Because the injection on one axis would influence components on the other axes in a multi-axis sensor [25], SOTA attacks cannot directly organize the desired false signals onto an assigned axis. They fail in the orientation control on MDOF systems.

3.1.2 Stationary vs.Moving. SOTA attacks conduct control over targets that are stationary or in a well-balanced status [70, 71], where inertial readings are originally zero. They respond merely to the acoustics and just output false signals.

In most cases, target CPSs are not still. The motion of targets is likely to cause the distance variation between the malicious acoustic source and the moving target. Such a distance variation will lead to phase fluctuation and therefore distort false signals. On the other hand, acoustic injections would never be the only input of inertial sensors in a moving target. These motion signals may couple with false signals and introduce additional noise.

3.2 Attackers' Capability and Patterns

We make the common assumptions [30, 30, 70, 71, 75] to describe attackers' capability: (1) they can synthesize any shape of acoustic signals using appropriate speakers or transducers and use auxiliary tools (e.g., optical/infrared camera and radar) to recognize the state (e.g., speed and orientation) of the targets [71] or their remote controllers [70]; (2) they have adequate knowledge about target systems, e.g., natural frequencies, and analyze the behavior of a device with the identical model in advance; (3) they cannot hack into target systems invasively because most CPSs prohibit such access rigorously without users' permission [4, 33].

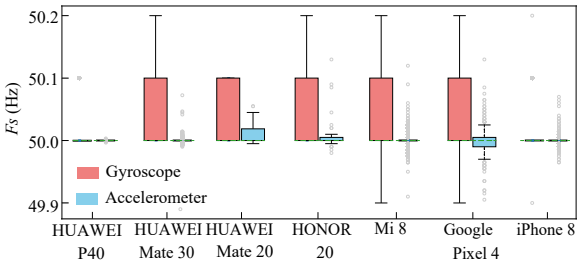


Figure 3: Sampling rate drifts are common in inertial sensors among COTS devices.

In reality, attackers would take various means to conduct attacks. We divide attackers' scope into two levels to cover most of possible non-invasive attack patterns as follows.

- *Remote Attack.* Attackers emit acoustic signals using nearby malicious sources to conduct remote transduction attacks. This can be done by playing sounds from speakers approaching the targets [71], or via means of tricking the users into visiting an email or a web page that auto-plays malicious audios [70].
- *Touch-based Attack.* Attackers afford the one-shot physical contact but they cannot physically alter the hardware. Neither can they directly access nor modify the inertial sensors. They can only attach a paster-like malicious acoustic transducer to the shell of target systems. For example, attackers can buy off a maintenance employee to place a malicious transducer under a mask of legal accessories, or they can attach the transducer by manipulating a miniature robot that approaches targets only once.

3.3 Attack Stage

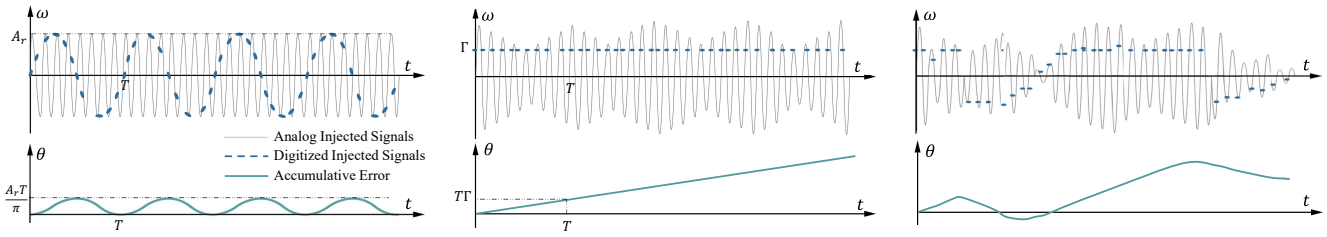
Combining the above analysis, we exploit the full potential of acoustic transduction attacks, as illustrated in Fig. 2. We first design an acoustic modulation for the stable injection (see Sec.4) with a controllable orientation (see Sec.5). We apply the proposed method to remote attacks for controlling stationary systems with both single-axis and multi-axis sensors. By investigating the motion influence (see Sec. 6.1), we extend remote attacks into moving systems with single-axis sensors while remote attacks merely pose DoS on the multi-axis under the impact of motion (see Sec.6.2). Against the most challenging targets, MDOF ones, we adopt the touch-based attacks for adversarial control (See Sec. 6.3).

4 ACOUSTIC MODULATION

We model the resonant characteristics of stationary inertial sensors under acoustic injections. Accordingly, we address the signal distortion caused by frequency offset and propose an acoustic modulation method to stably inject false signals.

4.1 Resonant Characteristic Modeling

Inertial sensors suffer from acoustic interference, due to the inner damping structure. We quantitatively model the resonant characteristics for the fine-grained modulation of malicious acoustic signals. We assume a malicious acoustic signal that resonates with an inertial sensor with the natural frequency $\omega_n = 2\pi f_n$. The signal exerts an oscillating pressure force $\boldsymbol{F} = F_0 sin\,(\omega_r t)$, where F_0 is the initial amplitude and ω_r is the initial frequency. The resonant response in



(a) Responses under unmodulated acoustics. (b) Envelope modulation for a stable output. (c) Amplitude adjustment for further control.

Figure 4: Basic idea of the proposed acoustic modulation.

one axis of the sensor [25, 68] is described as follows,

$$R(t) = A_r \cos(\omega_r t + \varphi_r), \tag{1}$$

where $A_r = -a_p a_r F_0$ is the overall amplitude, a_p and a_r are the constant gain coefficient during the analog process and resonance. Resonance introduces a phase lag φ_r,

$$\varphi_r = \arctan \frac{2\xi\omega_n\omega_r}{{\omega_n}^2 - {\omega_r}^2}, \tag{2}$$

where ξ is the constant damping ratio. For a given sensor, a_r and φ_r depend solely upon the injected frequency [68].

As the natural frequency typically exceeds the sampling rate in the analog-to-digital converter (ADC), aliasing migrates the high-frequency analogy signals into low-frequency digital ones according to the Nyquist sampling theory. In an ideal ADC, the sampling rate Fs keeps invariant. The injected signals are digitized as follows,

$$R[i] = A_r \cos(\omega_d \frac{i}{Fs} + \varphi_r), \ (i \in \mathbb{N}) \tag{3}$$

where ω_d is the frequency of digital injected signals in the target sensor, subject to the Fs as follows,

$$\omega_d = \omega_r - 2\pi n Fs, \ (|\omega_d| < \pi Fs, \ n \in \mathbb{N}). \tag{4}$$

Unfortunately, the sampling interval fails to keep constant. Instead, it drifts randomly within a range [8, 71]. Such drifts lead to unpredictable frequency offset, where $\omega_d^* = \omega_r - 2\pi n(Fs + \Delta Fs)$ replaces ω_d in Eq. 3. Therefore, false signals are significantly distorted and the attack is hard to perform.

We experimentally corroborate the randomness and universality of sampling rate drifts. We recruit seven volunteers[1]. Volunteers carry their smartphones as usual. These smartphones carry various modes of inertial sensors, including ICM-20690, BMI160, LSM6DSO, and the like. A third party application records the sampling rates of internal inertial sensors in these smartphones continuously for two weeks with the initialized sampling rate of 50 Hz. Results show that drift is common among inertial sensors in commercial off-the-shelf (COTS) devices with a range of 0.3 Hz. Among them, the Google Pixel 4 performs worst. Its sampling rate in the accelerometer ranges from 49.9 Hz to 50.1 Hz, and that in the gyroscope drifts up to 50.2 Hz. Even in the HUAWEI P40, the sampling rate changes intermittently. Because of the amplification effect [71], a slight drift might cause serious signal distortion.

[1]All experiments in this paper have obtained IRB approval. We have informed volunteers of the experiment purposes. Here, these data are merely used for the statistic on sampling rates, without any threat to privacy.

4.2 Stable and Controllable Injections

In pursuit of adversarial control, we modulate the acoustic signal by modifying its initial amplitude and phase. We leverage the unalterable characteristics to solve the problem of distortion caused by frequency offset and enable stable injections.

Goal. Attackers aim at a stable injection (i.e., constant outputs [70]) and then adjust it to desired waveforms.

Challenge. Frequency offset caused by the sampling rate drift [25, 71] would distort injections and degrade the attack effect into DoS. It is a challenge to *compensate the unpredictable and random offset.*

SOTA approaches. Existing approaches, e.g., WALNUT [70] and Poltergeist [30] set $A_r = \Gamma(t)$ and $\omega_r = 2\pi n Fs$ in Eq. 3. Therefore, they obtain a stable direct-current (DC) bias where $\omega_d = 0$. However, such treatments would be significantly distorted by frequency offset [71]. Or they may raise acoustic intensity to saturate the inner amplifier, yet produce non-adjustable outputs under audible injections with deafening volume.

Tu et al. [71] pace the acoustic phase (to be either always positive or always negative) to avoid the adverse impact of frequency offset. Although taking the initiative in spoofing gyroscopes in real systems, they merely obtain an accumulative error of the angular measurement, and thus, fail to produce stable false angular velocity.

Our solution. It has been proved that each amplitude of digital false signals can be modified independently by modulating acoustic amplitudes [71]. We observe that the final phases are also independently adjustable. Accordingly, we reshape the envelope of acoustic by carrying the reciprocal of the digital signal as follows,

$$F(t) = \frac{\Gamma(t)}{\cos(\omega_d t + \varphi_r)} \sin(\omega_r t). \tag{5}$$

Here, the additional phase φ_r is used to compensate for the phase lag introduced by resonance and two cosine items will be equal after sampling as Eq. 3, with the item $\Gamma(t)$ remained. Therefore, attackers are qualified to manipulate target sensors' readings into any designated waveform. Our basic idea is illustrated in Fig. 4. Under unmodulated acoustics, the digitized injected signals vary sinusoidally, with a tiny accumulative signal as shown in Fig. 4(a). Using our modulation method, we can obtain a constant digitized injected signal as presented in Fig. 4(b). By adjusting the acoustic intensity as illustrated in Fig. 4(c), we can generate false signals with arbitrary waveforms following the attackers' expectations. To achieve this, a fundamental issue is to estimate ω_d and φ_r.

4.2.1 Frequency Determination and Offset Compensation. It is difficult to calculate ω_d due to the lack of knowledge about targets'

sampling rate drifts. We exploit an unalterable frequency relationship to calculate ω_d and eliminate the influence of frequency offsets.

Frequency Difference. We observe that the frequency difference between acoustic signals also migrates into the low-frequency band after being digitized. To be specific, suppose that two signals of ω_{ri} $(i = 1, 2, \omega_{ri} = 2\pi nFs + \omega_{di})$ can resonate with the target sensor. We have the following unalterable frequency relationship,

$$\omega_{r1} - \omega_{r2} = \omega_{d1} - \omega_{d2}. \tag{6}$$

Offset compensation. This difference-based technique still works even if the sampling rate is drifting. Attackers can obtain an appropriate ω_{r1} with $\omega_{d1} = 0$ by analyzing the responses of a device of the identical model under ultrasonic resonance in advance. During a real attack, the actual digitized frequency is $\omega^*_{d1} = -n_p \Delta Fs$ due to the sampling drift. The drift brings about identical offsets in terms of both ω_{r1} and ω_{r2}. Under the guidance of Eq. 6, we can compensate the frequency offset by adjusting the frequency as,

$$\omega_{r2} = \omega_{r1} + \omega^*_{d1}. \tag{7}$$

Therefore, we have $\omega^*_{d2} = 0$ and the distortions caused by offsets are eliminated. Here the acoustic signal of ω_{r1} serves as a reference for the offset compensation. In practical attacks, the offset ω^*_{d1} can be measured by a remote camera or an attached malicious sensor.

4.2.2 Phase Estimation. Little existing literature notices that the phase under resonance lags significantly behind the original one, and the quantitative analysis on such a lag is also scarcely seen. Due to unknown parameters (i.e., ξ and ω_n in Eq. 2), we cannot obtain φ_r directly. Instead, we exploit the resonant phase-frequency characteristics to estimate the exact phase.

With the derivative of φ_r in Eq. 2, we obtain

$$\varphi'_r = \frac{2\xi(1 + (\frac{\omega_r}{\omega_n})^2)}{(1 - (\frac{\omega_r}{\omega_n})^2)^2 + (2\xi\frac{\omega_r}{\omega_n})^2} \approx 1/\xi, \tag{8}$$

where ω_r approaches ω_n and $|\omega_r - \omega_n| \ll \omega_n$ under resonance [25], $\frac{\omega_r}{\omega_n} \approx 1$, and φ'_r can be approximately recognized as a constant. It reveals that the phase lag φ_r has a positive linear correlation with acoustic frequency ω_r. φ'_r can be measured on sensors of the identical mode in advance. Considering that the reference signal supplies the feedback about both ω_{d1} and φ_{r1}, we can reckon malicious acoustic signals' phase lag φ_{r2} as follows,

$$\varphi_{r2} = \varphi_{r1} + \varphi'_r(\omega_{d2} - \omega_{d1}). \tag{9}$$

In practice, source speakers cannot support an excessive $A_r[i]$. Otherwise, the acoustic signals will distort. To mitigate the amplitude fluctuation, attackers should guarantee

$$|\cos(\omega_d i/Fs + \varphi_r + \varphi_0[i])| > \epsilon, \ (0 < \epsilon < 1), \tag{10}$$

where ϵ is a constant, satisfying that $\frac{\gamma}{\epsilon}$ is restricted within the output range of speakers. To meet this condition, we repetitively pace the acoustic initial phase as follows,

$$\varphi_0(t) = \begin{cases} -\varphi_r & |t - \frac{k\pi}{2\pi\omega_d}| < \frac{arccos\epsilon}{2\pi\omega_d}, \\ \pi - 2\arccos\epsilon - \varphi_r & \text{Others.} \end{cases} \tag{11}$$

In short, we modulate the malicious acoustic signals as

$$F(t) = \frac{\Gamma(t)}{\cos(\omega_d t + \varphi_a + \varphi_0(t))} \sin(2\pi\omega_r t + \varphi_0(t)). \tag{12}$$

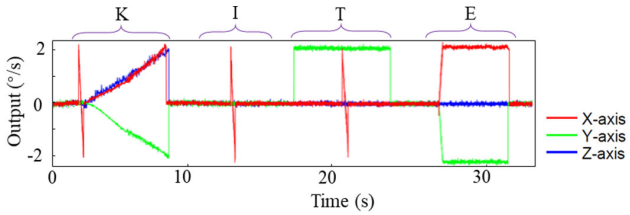


Figure 5: Spelling a 'KITE' trajectory by manipulating a BMI160 IMU using our proposed method.

Thus, we realize the stable and controllable injection $\Gamma(t)$. Figure 5 illustrates the threat from attacks adopting our proposed acoustic modulation method in manipulating a sensor's readings. Here we take the identical assumptions in SOTA attacks [30, 70, 71] without modifying attackers' capability. Moreover, such an injection could be achieved in both remote and touch-based attacks.

5 ORIENTATION CONTROL

Besides the single-axis systems, MDOF systems are widely used in real-world scenarios. The representatives include smartphones and drones, on which SOTA attacks barely investigate the potential of transduction attacks. With a full investigation into the distribution of false signals among axes under acoustic interference, we expand this attack surface into multi-axis sensors.

Goal. To completely control target MDOF systems, attackers should carefully arrange and inject appropriate false signals into each axis of the inner multi-axis inertial sensors. Therefore, target MDOF systems would face and go along an assigned orientation without any crash according to attackers' expectation.

Challenge. Injections on one axis would disturb those on other axes, because resonance would occur simultaneously on multiple axes in a sensor [25]. However, SOTA attacks [70, 71] ignore this issue, which still remains an open problem: *how to coordinate components of false signals among multiple axes accurately*?

Distribution among axes. Acoustic pressure force (vector) determines the false signals' amplitude and orientation [24]. We observe that in general the energy distribution of components in different axes is in line with the ray from an acoustic source to the target. One of our preliminary studies validates the directionality of such acoustic transduction attacks against inertial sensors. A speaker (JBL 750T, 30 W) is put 2 m away from a target sensor (a BMI055 chip) along each axis respectively. The mainly affected axes of sources from different orientations are illustrated in Fig. 6. That is, an acoustic source would influence the axis in an accelerometer that is parallel to the direction $\boldsymbol{e}_F$ from the acoustic source to the target, and the axis vertical to $\boldsymbol{e}_F$ in a gyroscope. The reason lies in the damping structure in inertial sensors [5, 41]. Imagine that an acoustic source is placed along the X-axis of an inertial sensor as shown in Fig. 6(a). It just interferes in the x-axial acceleration and the yaw angular velocity. We conclude the relationships as follows,

$$\boldsymbol{R}_{acc} \parallel \boldsymbol{e}_F, \ \boldsymbol{R}_{gyro} \perp \boldsymbol{e}_F, \tag{13}$$

where $\boldsymbol{R}_{acc}$ is the vector whose elements are the false signals on respective axes in an accelerometer and $\boldsymbol{R}_{gyro}$ is that in a gyroscope. In more common cases, $\boldsymbol{e}_F$ is not parallel to any axis. The influence of such a source can be decomposed into that of multiple orthogonal sources along each axis, due to the vector property of acoustics [78].

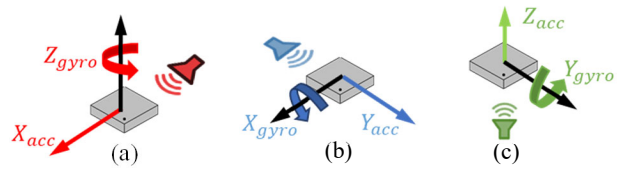


Figure 6: An illustration of affected axes of acoustic sources along (a) X-axis, (b) Y-axis, and (c) Z-axis.

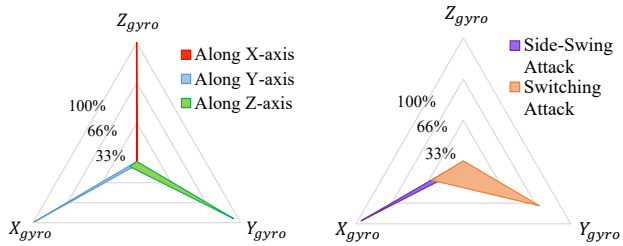


Figure 7: Energy distribution among axes. A smaller area means a better orientation control.

Solution. We utilize multiple acoustic sources to compensate for the orientation deviation. It is recommended to utilize three sources that constitute a set of three-dimensional (orthogonal, if possible) bases. By adjusting each source's acoustic intensity independently, false signals follow a given spatial vector with an assigned direction.

We represent energy distributions in the gyroscope of an iPhone 7 in Fig. 7, where attackers aim at generating appropriate false signals along each axis respectively. Compared with two demonstrations (Side-Swing and Switching Attacks) conducted on the identical device in [71], we successfully inject false signals into the target axis as expected, with little leakage into others. It maintains up to 99.13% of resonant energy in one desired orientation. In practice, the location of the inertial sensor in the target system can be inferred on a device with the same model as the target or by the aid of the datasheet beforehand. In addition, the multiple speakers should be aligned in a non-parallel manner, not necessarily orthogonally. The angle error keeps below 15° experimentally when the sources are non-orthogonal. By coordinating false signals using three acoustic sources, attackers are competent to drive target systems maliciously into any given orientation.

6 ATTACKS ON MOVING SYSTEMS

It is a common but complex scenario in which target systems are not stationary. We analyze the impact of targets' motion using a mathematical model, to describe the phase fluctuation and coupling effect quantitatively. Meanwhile, we explore possible threats after suppressing the influence of motion.

6.1 Motion Influence

Adversarial control over moving systems is an unsettled issue for acoustic transduction attacks. Motion interference distorts false signals under acoustic resonance. In this case, the effect of the attack would be currently constrained to uncontrollable disturbance.

6.1.1 Phase Fluctuation. The movement of a target alters the distance L between it and the sound source. The distance change provokes a phase fluctuation when acoustics travel in the air. Such a fluctuation results in the distortion of acoustic signals and attendant resonant responses. We denote the distance variation as ΔL. It will introduce an additional phase to Eq.3 as follows,

$$\Delta\varphi = \frac{\omega_r t \Delta L}{v}, \tag{14}$$

where v represents the acoustic speed and can be regarded as a constant. Because of this unexpected phase, the result of Eq. 12 on a moving target will be distorted, rather than the desired $\Gamma(t)$. Note that the motion also distorts false signals in all previous attacks [30, 70, 71] and limits their effect.

6.1.2 Coupling Effect. In inertial sensors, motion data will overlap, or even worse, couple with false signals. The coupling effect produces a force that introduces additional noise. We carry out the force analysis on a gyroscope using dynamic equations as follows,

$$\begin{aligned} m\ddot{y} + c\dot{y} + ky &= A_d \sin(\omega_n t) - 2m\Omega\dot{x} + F_y \sin(\omega_r t), \\ m\ddot{x} + c\dot{x} + kx &= 2m\Omega\dot{y} + F_x \sin(\omega_r t), \end{aligned} \tag{15}$$

where y and x are the driving and sensing displacements in the damping structure inside the gyroscope, k and m are constants, A_d is the amplitude of driving force at the frequency $\omega_n = \sqrt{k/m}$, F_x and F_y are components of the acoustic pressure F_0 on the driving and sensing directions, and Ω is the angular velocity around Z-axis to be measured. The angular velocity Ω will introduce the Coriolis forces $-2m\Omega\dot{x}$ and $2m\Omega\dot{y}$ into the sensing and driving directions respectively. Hence, we obtain the sensor's readings as follows,

$$\begin{aligned} x(t) = &\frac{2mA_d\Omega}{\omega_n c^2}\cos(\omega_n t) - a_r F_x \cos(\omega_r t + \varphi_r) \\ &+ 2ma_r^2 F_y \Omega\omega_r \cos(\omega_n t + \varphi_r). \end{aligned} \tag{16}$$

Here, the first term is the displacement that is proportional to the true angular velocity Ω; the second term is the false displacement triggered by the direct action of ultrasound on the sensing direction. Moreover, the acoustic action on the driving direction, coupled with the Coriolis force, is projected to the sensing direction as the third term. It would act as the noise and distort the false signals, which is jointly influenced by the system's motion and the component of acoustic pressure (that is related to the relative position from the target system to the acoustic source).

In addition, the movement of targets would result in the Doppler frequency offset. Nevertheless, this problem can be solved using the offset compensation method proposed in Sec. 4.2.1.

6.2 Remote Attacks

We propose a remote attack for the motion influence suppression and explore this method's limitations against moving MDOF systems. Advanced methods using auxiliary tools (e.g., optical/infrared camera and radar) enable accurate and real-time distance measurement. Therefore, attackers could measure ΔL to compensate for the phase fluctuation. In KITE, we adopt MVSCRF [83] due to its low measurement error (of below 1 mm in the original paper).

Then, we discuss the solution in terms of systems embedded with single- and multi-axis sensors respectively. The movement patterns of targets that carry single-axis sensors are usually simple, and thus attackers can easily predict the motion signals. By arranging malicious sources at appropriate places and aligning acoustics beams

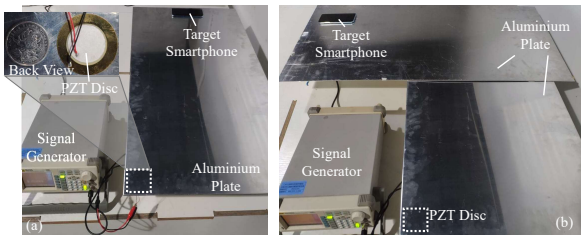


Figure 8: Experimental setup for the feasibility study of acoustic transduction attacks via (a) single-layer and (b) multi-layer (overlapping) solid media.

Table 1: Maximum Attack Distance on Different Materials

Material	Size	Distance†
Aluminium metal	1 m× 0.5 m× 2 mm	1.12 m +
	1 m× 0.5 m× 5 mm	1.12 m +
	1 m× 0.5 m× 7 mm	1.12 m +
Copper metal	1 m × 0.1 m × 0.2 mm	1.01 m +
Plastic	0.75 m × 0.7 m × 2 mm	1.02 m +
Glass	0.9m × 0.45 m × 10 mm	1.01 m +
Fiberboard	0.75 m × 0.7 m × 10 mm	1.03 m +
Log table	1.2 m × 0.75 m × 15 mm	0.32 m
Aluminium (2 mm)+Aluminium (5 mm)		1.12 m +
Aluminium (2 mm) overlaps Aluminium (5 mm)		1.50 m +
Aluminium (7 mm)+Copper		1.12 m +
Aluminium (7 mm) overlaps Plastic		1.25 m +
Aluminium (7 mm) overlaps Fiberboard		1.25 m +
Aluminium (7 mm)+Plastic+Fiberboard		0.65 m

†: '+' means that acoustics can affect target smartphones and such attacks remain at least as effective over a potentially longer distance via solid media.

along the target's trajectory, attackers can easily eliminate the coupling effect, i.e., F_y in Eq. 16. Hence, they continue manipulating those targets remotely, with evaluations in Sec. 7.3.

However, attackers cannot predict the complex movement of an MDOF system, so they fail to align the acoustics with the target's trajectory. Therefore, when attacking moving targets that are embedded with multi-axis sensors, current remote attacks merely act as DoS because existing methods fail in remote and real-time motion description on the centimeter scale. Our experimental results in Sec. 7.3.1 also demonstrate the limitation of remote attack on MDOF systems. In addition, camera-based methods can be influenced when there exists occlusion or the lightning condition is poor. Therefore, the application scenarios of remote attack are limited. In short, remote attacks cannot apply to manipulation of such systems that move freely in space.

6.3 Touch-based Attacks

In many cases, it is probable for attackers to have the one-shot physical contact with target systems. Therefore, they can perform a touch-based attack by attaching a paster-like malicious unit on targets, especially MDOF ones, so that they can continue malicious control on moving target. In the following, we first verify the feasibility of adopting acoustic propagation that travels in solid media to enable touch-based attacks. We then present our design of a malicious unit and its ability of attacking realistic systems.

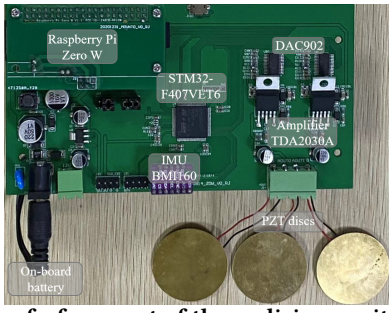


Figure 9: Proof-of-concept of the malicious unit (PCB board prototype) for touch-based attacks.

6.3.1 Acoustic attacks travelling in solid. Acoustic guided waves can propagate in solid media [78, 85]. Inspired by this, we divert acoustic interference into solid media (e.g., target systems' shells) by a piezoelectric (PZT) transducer instead of via air by speakers. A pilot study is launched to investigate its feasibility.

As shown in Fig. 8, we stick a miniature PZT disc (with 35 mm diameter and 0.3 mm thickness) to the underside of an aluminium metal plate (with 1 m × 0.5 m × 2 mm). A signal generator supplies sinusoidal signals that will be converted to acoustic guided waves by the PZT disc. The frequency response of the PZT disc ranges from 20 Hz to 40 kHz. The power consumption is about 55 μW. A smartphone (Samsung Galaxy S8) is placed at an arbitrary position on the plate. Its accelerometer generates false readings under the acoustic interference of 6.5 kHz. Similarly, the internal gyroscope resonates with the 19.5 kHz ultrasound.

We repeat the above experiments on other materials, including copper metal, plastic (polythene), wood (fiberboard and log table), and glass, which have covered most of common materials used in COTS CPSs [1]. The target devices still suffer from such acoustic injection via these solid media. Moreover, acoustic transduction attacks can cross multilayered media if they wholly or partially overlap as Fig. 8(b) shows. The maximum attack ranges via these media of various thicknesses are over 1 m, as listed in Tab. 1. Particularly, such attacks are powerful enough to affect devices within 32 cm through a wooden table board (15 mm thickness).

Compared with the ultrasound speakers used in the prior literature [63, 71, 75], the PZT transducers are cheap and much smaller in size. They can be covertly adhered to the target's shell for acoustic transduction attacks. Malicious acoustic signals are primarily localized in solid media, with little leakage into the air. Thus, attacks are conducted without victims' attention, with the evaluation on human inaudibility in Sec. 7.6.

6.3.2 Malicious unit design. With the purpose of suppressing motion interference, we design a malicious unit that adheres to the target stealthily. It facilitates touch-based attacks that propagate malicious acoustics via solid media.

The malicious unit carries a control center, a malicious inertial sensor, and PZT transducers. The control center supplies acoustic signals to the PZT transducer that emits sound waves through solid media (i.e., the shell and connections). In this case, the relative orientation and distance are unchangeable, and thus, phase fluctuations in Eq.14 are suppressed. The malicious inertial sensor measures the motion state of the target system (i.e. Ω in Eq. 16), and thus the control center could reduce the noise caused by the coupling effect.

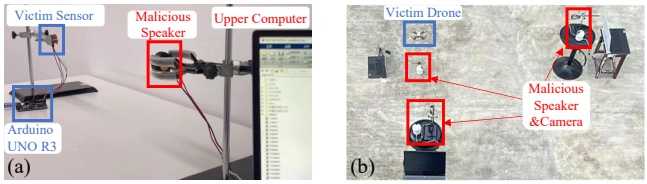


Figure 10: Experimental setups.

We integrate the touch-based attack into a printed circuit board (PCB) prototype as the malicious unit, as shown in Fig. 9. It carries an STM32-F407VET6 chip [65] and a Raspberry Pi Zero W [54] as the control center, a digital to analog converter (DAC902 [10]), a BMI160 inertial sensor [7] (here the on-board sensor is exchangeable and we choose one with resonant frequencies different from the targets to avoid being affected by the attacker itself), and an on-board battery (12 V, 1500 mA). It drives PZT discs that are attached closely onto targets to emit malicious acoustic signals. The size of the whole board is 13 cm in length and 7 cm in width. Note that this prototype is a proof-of-concept (without elaborated integration). In real implementation, the size of such a unit will be miniaturized into an extremely small (paster-like) size (e.g., within 4×4 cm^2) after customized manufacture. It costs merely about $26, without the need of expensive signal generators and loudspeakers, which are necessary in SOTA attacks [30, 70, 71].

In conclusion, we consider all possible attack scenarios and assess the practical threat level from transduction attacks. Remote attacks can threaten stationary targets and moving single-axis sensor embedded systems, while touch-based attacks cover all scenarios.

7 EVALUATION

We conduct remote and touch-based attacks on COTS devices and evaluate their effectiveness. Two end-to-end attacks demonstrate its attack effects by manipulating the route of a drone embedded in the most popular autopilot (Pixhawk 4) and imitating gaits to spoof the pedometer APP ('Pacer') on smartphones.

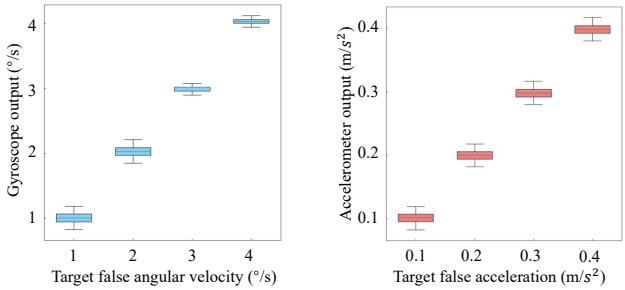
7.1 Experiment Setup

Target systems. We first carry out experiments on a BMI055 chip that is widely deployed in COTS CPSs (e.g., Oculus Rift and Pixhawk 4) for directly gathering the raw inertial data for quantitative analysis. A BMI055 chip contains a three-axis accelerometer and a three-axis gyroscope. We connect an Arduino board (UNO R3) to the sensor chip and samples its outputs at 50 Hz. Then we conduct KITE on COTS devices including self-balancing robots, smartphones, and drones, summarized in Sec. 7.5. In particular, we attack a quad-rotor drone (ATG-850 RTK) that carries Pixhawk 4, the most popular autopilot. It runs the open source PX4 controller [50] and carries two inertial measurement units, BMI055 and MPU-6000, which are both vulnerable. Here, we mainly evaluate the attacks on its BMI055 and the MPU-6000 performs similarly. The outputs of the BMI055 sensor are recorded locally and read by the upper computer after each experiment. Sampling rates are 50 Hz by default.

Acoustic source. In remote attacks, we use JBL 750T speakers [28] as the remote malicious acoustic source. Supplied by a 30 W power amplifier, it can emit acoustics from 20 Hz to 48 kHz with a peak intensity of 76 dB. A signal generator NI VituralBench 8012 [47], connected to an upper computer, modulates the signals and drives the speaker. In touch-based attacks, we exploit the PCB prototype in Fig. 9 as the malicious devices for attacks.

Table 2: Real Motion vs. False Signals (°/s)

Input		Median	Mean	Standard deviation	Range
Idle	0°/s	-0.025	-0.006	0.071	±0.155
Real	2°/s	2.006	1.985	0.066	±0.150
False	2°/s	2.013	2.031	0.077	±0.160
	-2°/s	-1.999	-1.976	0.085	±0.210



(a) On gyroscope (b) On accelerometer

Figure 11: Amplitude control.

Placement. In remote attacks, the JBL 750T speaker is placed 2 m away from target systems. We attack stationary targets (including inertial sensor chips and smartphones) in a quiet room with 46.6 dB ambient noise and moving targets (including a MITU robot and drones) in an open space with 55.9 dB ambient noise, as shown in Fig. 10. For the orientation control, we place three speakers centered around target sensors. In touch-based attacks, the PCB prototype of the malicious unit is attached to target devices' shells.

Metric. We adopt statistical characteristics including median, mean, standard deviation, and range to describe the performance on injecting assigned false signal in terms of amplitude. The orientation control is evaluated by the angle error denoted as $\Delta\vartheta$. It can be calculated by $\Delta\vartheta = \arccos(\boldsymbol{e}_t \cdot \boldsymbol{e}_o)$, where $\boldsymbol{e}_t$ is the unit direction vector of the target false signal and $\boldsymbol{e}_o$ is that of the achieved one.

7.2 Overall Performance

We evaluate KITE in injecting desired false signals with a controllable orientation on stationary targets.

7.2.1 Amplitude. We manage assigned injections with arbitrary amplitude at will. In most cases, a CPS rotates at a speed within 30°/s and accelerates within 0.5 m/s^2, and the speed of human activities is typically in this range. As representatives, we inject false signals of 1, 2, 3, and 4°/s into the yaw-axis of the gyroscope in target BMI055 chip. The setup is shown in Fig. 10(a). We first list the statistical characteristics of real motion and false signals under remote attacks in Tab. 2. Compared with the real motion where the target rotates at 2°/s, the false signals present insignificantly different results, only with a slight rise in terms of standard deviation. Fig. 11 further demonstrates the precision of our proposed acoustic modulation on diverse values. It obtains a low error with the standard deviation of about 0.08°/s on average. Such deviation would not increase with the amplitude of false signals. It peaks at 0.091°/s under the injection

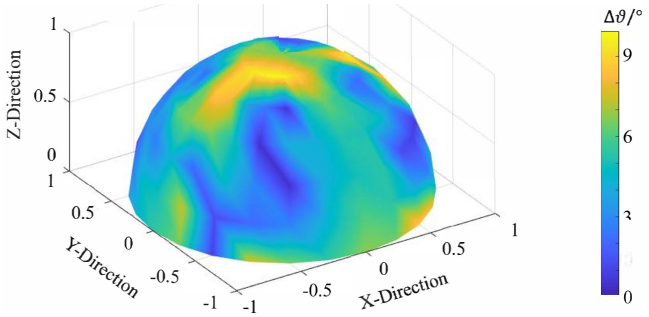
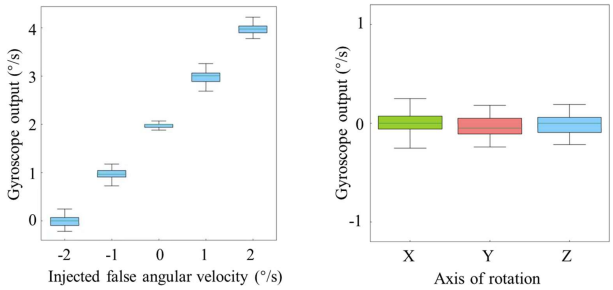


Figure 12: Orientation control over a drone. A smaller $\Delta\vartheta$ represents a better performance.



(a) Different injections (b) Different axes

Figure 13: Robustness against motion.

of 2°/s, while its minimum value just maintains 0.069°/s with tiny difference from the real motion (about 0.07°/s on average). Similarly, we repeat experiments on the X-axis of the accelerometer and find the average standard deviation is below 0.01 m/s^2. We manage to inject false signals ranging within ±50°/s into gyroscopes and ones ranging within ±0.8 m/s^2 into accelerometers. Accordingly, our attacks can induce any waveform and deceive the target system into following our preset trajectory. We adjust the sensor sampling rate as 5 Hz, 16.7 Hz, 100 Hz, and 200 Hz, which are typical values [4] and maintain the standard deviation below 0.09°/s and 0.012 m/s^2.

We further validate the effectiveness on a self-balancing robot, MITU robot [43], and aim at its single-axis gyroscope. The robot's embedded gyroscope is employed to detect and measure tilts (forward or backward) and accordingly the robot is actuated to move (backward or forward respectively according to the negative feedback mechanism). Using modulated acoustic signals, the robot would go along the direction following the false angle.

7.2.2 Orientation. We take orientation control over a stationary drone (ATG-850 RTK) remotely, as shown in Fig. 10(b). The standard deviation of false signal is below 0.1°/s. The thermodynamic diagram in Fig. 12 shows angle errors, where the other half shares a similar distribution. We find $\Delta\vartheta$ is below 9° globally, and it does not exceed 5° in over 70% of orientations. In some special cases, the location of acoustic sources may fail to be orthogonal. $\Delta\vartheta$ still keeps below 15° experimentally when the sources are non-orthogonal. Note that these sources should avoid being parallel, otherwise they cannot support the orientation control. Moreover, the attacks are still able to manipulate the yaw angle and the X-axial acceleration using only one acoustic source as illustrated in Fig. 6(a). It can command targets like unmanned cars (e.g., Baidu Apollo D-KIT) to alter orientation and speed up forward or backward.

The touch-based attack achieves similar performances. It maintains a low standard deviation of 0.071°/s in gyroscope and 0.009 m/s^2 in an accelerometer on average and small $\Delta\vartheta$ within 7.8°.

7.3 Robustness against Movement

We follow the proposed solutions in Sec. 6.1 to evaluate acoustic transduction attacks against a moving target.

7.3.1 Robustness of remote attack. We fix a 30 W powered JBL 750T speaker and a camera (Logitech C930e). They are connected to an upper computer that runs the MVSCRF algorithm [83] on a server with Intel(R) Xeon(R) Silver 4210R CPU@2.40GHz and two Nvidia GeForce RTX 3090 to measure the distance to targets and accordingly modify acoustic signals. MVSCRF realizes a low measurement error of below 2 mm. We place a BMI055 chip on a rotating table and keep it 2 m away from the speaker. The table rotates centered around the Z-axis of the BMI055 chip at a speed of 2°/s by default.

We inject false gyroscope signals of different amplitudes. The yaw angle velocities are shown in Fig. 13(a). They maintain the low deviation of 0.2°/s. In particular, we inject a false signal of -2°/s to neutralize real motion. Consequently, the gyroscope outputs zero and the target system would mistakenly regard itself in a stationary state. It would not respond to the real motion and lose the ability of perceiving the physical world. We adjust the rotating speed of the table from - 4°/s to 4°/s at a step of 1°/s. We inject the corresponding false signals to neutralize real motion. The output readings keep 0.07°/s on average with the deviation of below 0.26°/s. We repeat the experiments when the target rotates around other axes and obtain the similar performance with a deviation of 0.2°/s, as shown in Fig. 13(b). We further test attacks on a moving MITU robot. It moves at ±0.1 m/s and ±0.2 m/s or rotates at ±2°/s and ±5°/s respectively at most 3 m away from the speaker. KITE injects a false signal of 2°/s successfully, with a deviation of 0.18°/s.

However, when attacking moving targets with multi-axis sensors, remote attacks merely act as DoS. We conduct experiments on moving drones including a QQL RC UAV and a DJI Spark UAV. We cannot avoid the coupling effect and thus drones crash. The standard deviation of inertial readings in the DJI Spark UAV is 1.31°/s. By comparison, its standard deviation is approximately 1.45 under unmodulated acoustic injections using the same settings. It validates that remote attacks cannot apply to manipulating MDOF systems and frustrates SOTA attacks [70, 71] in practice.

7.3.2 Robustness of touch-based attack. We repeat the above experiments on moving drones using the PCB prototype. The standard deviations of inertial readings drop down to 0.08°/s. With touch-based attacks, we can adjust attitude of target drones without crash, but also inject false upward or downward accelerations to alter the target drone's flying altitude or order it to land or take off.

7.4 Effective Distance of Remote Attack

In remote attacks, the distance is positively correlated with the power supply of acoustic sources and varies among different targets due to their diverse sensitivity. We successfully manipulate readings of a Huawei P40's gyroscope 10.3 m away and that of the accelerometer 7.6 m away using a 30 W powered speaker with little

Table 3: Attack Experiments on COTS Devices‡

Device	IMU Model*	f_n (kHz)	
		Gyro.	Acc.
ATG-850 RTK drone	BS BMI055	24.4	1.45
	IS MPU6000	27.0	1.81
Huawei P40	Unknown	19.9	4.6
Huawei P20 Pro	IS ICM-20690	20.1	6.7
Samsung S20	Unknown	19.2	19.2
Samsung S8	STM LSM6DSL	19.4	6.5
Google Pixel 4	BS BMI160	23.1	-
Motorola Edge 5	Unknown	27.6	0.1
iPhone 11 Pro Max	BS BMI282	24.2	-
OPPO A32	Unknown	28.9	4.7
OPPO Find X2	Unknown	19.7	0.1
Reno 3 Pro	STM L2G2IS	39.1	0.1
Redmi K30 Pro	BS BMI270	38.9	6.5
MITU robot	IS ICM-20690	20.1	6.7
Baidu Apollo D-KIT	IS MPU6050	27.5	5.2
EAIBOT N1 UGV	M R6093U	27.2	6.5
QQL RC UAV	IS IMU3000	27.1	23.0
DJI Spark UAV	UnKonwn	23.8	5.5

‡The full list involving 28 COTS devices can be found in [3].

∗BS: Bosch, IS: TDK InvenSense, STM: STMicroelectronics, M: Microinfinity.

increase of the standard deviation (below 0.2°/s or 0.024 m/s^2) and $\Delta\vartheta$ (below 15°). This distance can be extended to over 13 m using a speaker powered by 50 W, also a common setting in COTS devices. Furthermore, better acoustic devices, e.g., professional speakers with power amplification techniques, could improve the attack distance to above 37 m [63].

7.5 Diversity of Target Devices

We evaluate our proposed attacks on more real devices equipped with inertial sensors. All tested devices are susceptible to adversarial control. We present partial results in Tab. 3, with the full list involving 28 COTS devices and attack demos in [3]. In particular, we can attack the accelerometer and gyroscope in a system simultaneously to spoof its controllers. Note that we can test the devices to measure the natural frequencies without knowledge of the IMU models. We observe the responses of robots/drones or inertial readings of smartphones (with zero-permission access [4]) under ultrasound whose frequency sweeps from 100 Hz to 30 kHz at an interval of 100 Hz first. When a rough range of the resonant frequency is found, we adjust the interval to 10 Hz and 1 Hz to determine the exact frequency with the maximum resonance, i.e., the natural frequencies. The measurement process is within several minutes. Moreover, multiple sensors can be measured simultaneously. Considering the prior work [30, 63, 70, 71, 75] and our results, we conclude that KITE could affect most CPSs.

7.6 Inaudibility

Acoustic transduction attacks should avoid being heard by surrounding people in case of being detected and defended against.

7.6.1 Remote attack. Gyroscopes and accelerometers are both vulnerable to acoustic interference, but sensitive to different frequency bands. Gyroscopes' natural frequencies typically exceed 19 kHz. This implies that malicious acoustics aimed at gyroscopes are beyond the human hearing [63, 76]. We recruit 22 volunteers aged from 18 to 45 when remotely attacking gyroscopes of the devices in Tab. 3. They report being unable to distinguish the existence of modulated ultrasound except when attacking OPPO Reno 3 Pro and Redmi K30 Pro. During the attacks on the two devices, the speakers would induce audible noise of about 18 kHz due to its poor performance at the high-frequency bands of over 35 kHz. We believe that using professional acoustic devices can overcome the fault for additional noise. Conversely, most accelerometers respond to sounds of below 10 kHz according to Tab. 3. Therefore, malicious sounds emitted from remote sources can be heard by humans. SOTA attacks aimed at controlling accelerometers [30, 70] alert surrounding people, unless on some exceptions, e.g., a Samsung S20, which is also selected as the only target in [30] due to the embedded accelerometer's high natural frequency of 19.2 kHz.

Table 4: Human Audibility Tests on A Drone

Motion status	Acoustic intensity		Human prediction
	10 cm	5 m	
Hanging	109.4 dB	68.7 dB	-
Rotating w/o attacks	109.9 dB	71.2 dB	3.29
Rotating w/ attack	109.5 dB	69.9 dB	3.41

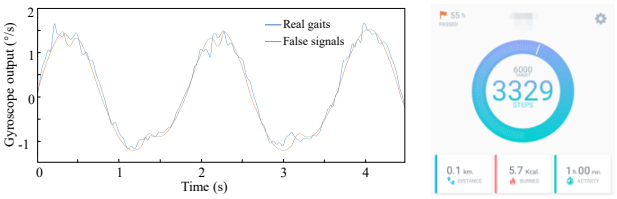
7.6.2 Touch-based attack. Touch-based attacks leverage malicious acoustics that are primarily localized in solid media, with little leakage into the air. Thus, attacks are covertly conducted without victims' attention. We place two microphones 10 cm away from the PZT disc placed under the 5 mm aluminium metal plate, following the setting in Fig. 8. The frequency of the attacking signal is 6.5 kHz for the accelerometer and 19 kHz for the gyroscope in the Samsung Galaxy S8, respectively. We use an NI USB-4431 sound measuring instrument and GRAS 46AM 1/2" CCP free-field standard microphones for measuring the unweighted sound pressure levels. The used GRAS 46AM microphone has a wide frequency range of 3.15 Hz to 31.5 kHz [26]. One microphone directly contacts the plate, and it measures that sound in solid reaches up to 73.7 dB. The other hanging in the air measures that sound remains 48.8 dB in a quiet room (46.6 dB). Such acoustic leakage is subtle and negligible, especially under mechanical noise from target systems. PZT transducers can also issue ultrasounds beyond the range of human hearing to attack gyroscopes. Surrounding people barely perceive such stealthy attacks travelling in the solid.

In short, gyroscopes are more at risk than accelerometers, and touch-based attacks are stealthier in terms of inaudibility.

7.7 End-to-End Attack Cases Study

We now evaluate the proposed attacks with end-to-end cases on COTS devices. We conduct the starting attacks on smartphones to spoof step counts and manipulate a drone.

7.7.1 On smartphones. In smartphones, inertial readings are utilized for navigation services, pedometer applications and the like. Using the remote attack, we accumulate a false yawing angle of up to approximately 6.23 rads or -6.19 rads in 1 minute in a Huawei P40,



(a) Comparison of false signals with imitates gaits with real ones.

(b) Screenshot of 3300 false steps on 'Pacer'.

Figure 14: Attacks on smartphones.

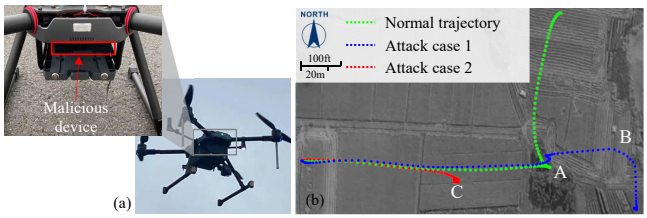


Figure 15: Attacks on a drone. (a) The malicious unit is attached onto the target covertly. (b) Touch-based attacks manage to manipulate the target's trajectory.

with few false signals in the pitch nor roll angle. Thus, attackers can deceive navigation services such that they would misguide users into wrong routes. Furthermore, we modulate malicious acoustic signals to produce false gaits. We can adjust these 'gaits' at speed of 5 to 55 steps in a minute without any real walking. The comparison with the inertial readings of gaits from a real user is shown in Fig. 14(a). With no difference from real ones, these false gaits could trick motion-driven pedometer applications. We register around 3300 steps in 1 hour on a pedometer APP 'Pacer', which is one of the most popular step-counting APPs in Google's Play. The screenshot is shown in Fig. 14(b). Here we have not claimed in-app rewards.

7.7.2 On a drone. Unmanned vehicles, e.g., drones, depend on inertial readings for attitude estimation, autonomous navigation, and actuation decision. The PCB prototype is attached to an ATG-850 RTK drone, with a camouflage shell whose color is similar to the target, as shown in Fig. 15(a). The drone flies about 200 m above the ground in an open space. In the drone, a complementary filter [22] is employed for attitude estimation using the data collected by the BMI055 inertial sensor. GPS is forbidden to reveal the threat on inertial sensors here. In addition, GPS signals in some cases may lose (e.g., due to the electromagnetic interference), and the attacks can continue with GPS spoofing [32, 48, 59, 69].

If no attack occurs, the drone follows a preset path, i.e., a normal trajectory as the baseline. It goes east at the speed of around 4 m/s and then turns north at the angular velocity of about 5°/s, as the green line in Fig. 15(b). We conduct two attacks that manipulate trajectories as the blue and red lines in Fig. 15(b).

In the attack case 1, we successfully deflect the target drone to drift away and move under adversarial control as the blue line in Fig. 15(b). When arriving at the 'A' point, the target drone is supposed to alter its orientation and turn north following the preset path. The PCB prototype detects this rotation and launches an attack, where the target gyroscope produces false readings of 5°/s and tells the controller that it has faced north (actually still faces east, with an angle error of about 11°). Hence, the drone goes straight rather than turns left. After the 'A' point, the drone is intended to move straightly without veering. The PCB prototype keeps idle until the drone arrives at 'B' point. It injects false anticlockwise gyroscope readings of 5°/s here. This unreal rotation reported by the attacker-controlled inertial sensor misleads the actuation system to the belief that it is pushed by a real external force. Due to the negative feedback mechanism for balance, the drone sheers off clockwise, and thus, faces south. In this way, attackers manipulate the target drone into following the malicious trajectory. Ultimately, at the assigned location under adversarial control, the drone will have an 'illusion' of arriving at the legal destination and stop its flying (this can be done by using the following attack case 2).

In the attack case 2, we stop the target drone, as the red line in Fig. 15(b). A forward false signal of 0.4 m/s^2 is injected into the accelerometer. Due to the negative feedback mechanism, the target drone actuates backward under the misperception of the existence of an unreal forward acceleration. As a result, the drone slows down and then (consuming about 10 s) stops at the 'C' point.

In comparison with existing approaches [70, 71], our proposed attacks realize stable injection into all inertial sensors, free from the disturbance from frequency offsets. We extend attacks to moving targets, and in particular, the touch-based attacks cover the most complex scenarios where the MDOF targets are moving.

8 DEFENSE AND DISCUSSION

In this section, we discuss the limitations of our proposed attacks and countermeasures for protecting inertial sensors.

8.1 Countermeasure

Considering the wide deployment of inertial sensors, it is urgent to develop effective countermeasures. We have informed relevant manufacturers of the attack and the following defending methods

8.1.1 Existing Approaches. We summarize the limits of current methods that are potentially against acoustic transduction attacks.

Dampening and Isolation. An intuitive idea is to weaken or eliminate the acoustic injection before it acts on sensors. Using acoustic dampening materials, such as acoustic foams, can attenuate over-the-air acoustic waves before they penetrate sensors [17]. Advanced dampening materials reach 90% acoustic reduction [64]. However, this method undoubtedly introduces significant cost. Besides, its resilience is unclear against attacks via solid propagation.

Filtering. Using low pass filters is another option to weaken acoustic effect [70]. However, the attacks still work even if the cut-off frequency is limited within 10 Hz due to hardware defects [34]. Sun et al. [67] propose a filter based on orthogonal demodulation, but the I/O dual channel is rare in existing inertial sensors.

Redundancy. Redundancy techniques that leverage multiple sensors for double checking are believed to enhance the resilience. Nevertheless, the vulnerability of Pixhawk 4 implies that acoustic transduction attack can jointly influence multiple inertial sensors simultaneously. Although other types of signals can be fused [14, 52], those signals are not always reliable. For example, GPS signals may lose in some cases (e.g., under the electromagnetic interference). Even worse, spoofing attacks [21, 49, 72] threat various sensors, including GPS [32, 48, 59, 69], LiDARs [12, 60, 66], camera sensors

[16], microphones [36, 56, 85–87]. An advanced power-switching method [61, 80, 89] is effective against electromagnetic interference. However, it is inapplicable to detecting transduction attacks.

Sampling. Normally, attackers modulate acoustics based on target sensors' sampling rate. Conversely, it is feasible to modify sampling intervals. Trippel et al. [70] propose two defense mechanisms. One is the randomized sampling that adds a random delay to each sampling period. It prevents false DC signals but with the penalty of accumulating growing measurement errors. The other requires out-of-phase sampling with two samples at a 180° phase delay. Its essence lies in doubling the sampling rate. However, it performs ineffectively when $\omega_d = 4\pi nFs$ in Eq. 4. Tu et al. [71] recommend a dynamic Fs on the basis of the above randomized sampling mechanism. Nevertheless, it has not yet alleviated the problem of degraded accuracy in inertial measurements.

8.1.2 Our Suggestion. Though the standard deviation of false signals is slightly higher than real ones, this difference is too tiny to separate false signals. Instead, we alter sampling rate and reduce its side effect. We minimize the accuracy loss by regulating jitters into the sampling period, with a theoretical analysis on its effectiveness.

Sampling jitters t_a would limit the outputs' signal to noise ratio (SNR) according to the frequency ω as follows [9],

$$SNR = -20log_{10}(\omega \times rms(t_a)), \tag{17}$$

where $rms(t_a)$ is the root mean square jitters. For injected signals, $\omega = \omega_r$ is far greater than that of real motion. Therefore, sampling jitters significantly disturb spoofing attacks.

Instead of a fixed sampling interval $\frac{1}{Fs}$, we design alternate intervals $\frac{1}{Fs} + t_a[i]$ with the cyclic jitters $t_a[i]$ as follows,

$$t_a[i] = \alpha_m,\ (m = i\ mod\ C,\ i \in \mathbb{N}) \tag{18}$$

where α_ms are small constants and C is an arbitrary constant. Here we set $C = 2$ and $\alpha_0 = -\alpha_1$. In the comparison of random [70] or dynamic [71] ones, the periodically alternating jitters have a smaller root mean square with the adjustable α_ms. Hence, our countermeasure significantly mitigates the effect of spoofing attacks, at the cost of exerting little adverse influence on inertial measurements.

8.2 Discussion

Here we discuss the potential influence of the small-sized malicious device and limitations of our proposed attack.

8.2.1 Impact of small-sized integration. Indeed, the small-sized integration could be double-edged. After the integration, the malicious devices will become sufficiently small to perform more covert attacks. Considering that the voltage supplied by the integrated power component is related to the intensity of the malicious acoustics, we should maintain a voltage supply of 12 V. Otherwise, the range of false signals' amplitudes would reduce. However, in this case, the battery volume might be reduced, resulting in a small endurance. Fortunately, it does not need to constantly emit malicious acoustics and the small-sized malicious device can still support practical attacks.

8.2.2 Limitations. Our remote attacks on moving targets with single-axis sensors are assisted by a camera running the MVSCRF algorithm. However, MVSCRF requires non-trivial computing resources. In our experiments, the MVSCRF algorithm presents an execution latency of 2 s. Such latency can be compensated when the targets move at a low speed (e.g., the MITU robot in Sec. 7.3.1) or approximately uniform speed by multiple speed measurements. However, if a target keeps changing the moving speed, MVSCRF would produce extensive errors. These errors limit the ability of attackers to generate a stable false signal and degrade the remote acoustic transduction attacks to be DoS.

9 RELATED WORK

Privacy Leakage through Inertial Sensors. Different from the pursuit of inertial data tampering, several attacks utilize IMUs for privacy exfiltration, including speech [2, 4, 27, 44], keystroke [11, 40, 42, 45, 73, 82], physical activity [29, 74, 77], localization [15, 38, 46, 53], and device identification [20, 62, 88]. Moreover, inertial sensors can also leak users' behavioral biometrics [13, 39, 79, 81].

Sensor Spoofing Attacks. Such spoofing attacks are increasingly risking the security of CPSs [84]. A slew of sensors are suffering from electromagnetic interference (EMI) [21, 36, 49, 57, 58, 72]. LiDARs systems [12, 60, 66], GPS [32, 48, 59, 69] and camera sensors [16] are also vulnerable. Ultrasound can inject inaudible commander into VAs as well, which benefits from the acoustic nonlinearity [56, 85–87]. As a countermeasure, researchers usually utilize power-switching for defence [61, 80, 89]. Unfortunately, the power-switching method defends mainly against EMI, but it could not apply to resisting acoustic transduction attacks.

Acoustic Sensitivity of Inertial Sensors. Inertial sensors are vulnerable to acoustic injection [18, 34, 76]. Not content with DoS attacks [63, 75], researchers [70, 71] pursue adversarial control but are unable to achieve controllable waveform and orientation due to the frequency offset, multiaxial resonance, and the target's motion. In contrast, KITE achieves this goal, not to mention that KITE also has other advantages, such as orientation control, motion robustness, and low cost. In addition, sensitive inertial sensors can be utilised to establish covert channels [6, 25, 55].

10 CONCLUSION

We conduct a thorough threat analysis of acoustic transduction attacks against CPSs. We model acoustic effect on inertial sensors and organize our study covering most of the possible attack scenarios. A new acoustic modulation-based attacking method is proposed to exploit the practical potential threat of a realistic attacker under all these scenarios. Combining the performed investigations together, we expand the attack surface into MDOF systems and suppress the motion influence. In particular, we accomplish control over COTS in an automatic manner using the designed PCB prototype. End-to-end attack cases appeal for people to take necessary countermeasures to resist such threats.

ACKNOWLEDGES

This paper is partially supported by the National Key R&D Program of China (2021QY0703), National Natural Science Foundation of China under grant U21A20462, 61872285, 62032021, 61772236, 62172359, and 61972348, Research Institute of Cyberspace Governance in Zhejiang University, Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Grant No. 2018R01005), and Ant Group Funding No.Z51202000234.

REFERENCES

[1] Analog Devices, Inc. 2017. The five motion senses: Using MEMS inertial sensing to transform applications. https://www.analog.com.

[2] S. A. Anand and N. Saxena. 2018. Speechless: Analyzing the Threat to Speech Privacy from Smartphone Motion Sensors. In *IEEE Symposium on Security and Privacy*.

[3] Anonymous User. 2022. KITE. https://github.com/KITE-anonymous-user/KITE.git.

[4] Z. Ba, T. Zheng, X. Zhang, Z. Qin, B. Li, X. Liu, and K. Ren. 2020. Learning-based Practical Smartphone Eavesdropping with Built-in Accelerometer. In *Network and Distributed System Security Symposium*.

[5] J. J. Bernstein, S. Cho, A. T. King, A. Kourepenis, and M. Weinberg. 1993. A micromachined comb-drive tuning fork rate gyroscope. In *IEEE Micro Electro Mechanical Systems*.

[6] K. Block, S. Narain, and G. Noubir. 2017. An Autonomic and Permissionless Android Covert Channel. In *ACM Conference on Security and Privacy in Wireless and Mobile Networks*.

[7] Bosch, Inc. 2018. BMI160 Datasheet. https://www.bosch-sensortec.com/products/motion-sensors/imus/bmi160.html.

[8] Bosch, Inc. 2020. BMI055 Datasheet. https://www.mouser.cn/datasheet/2/783/BST_BMI055_DS000-1509583.pdf.

[9] B. Brannon and A. Barlow. 2006. Aperture Uncertainty and ADC System Performance. https://www.analog.com/media/en/technical-documentation application-notes/an-501.pdf.

[10] Burr-Brown Products from Texas Instruments. 2002. DAC902 Datasheet. https://www.ti.com/lit/ds/symlink/dac902.pdf.

[11] L. Cai and H. Chen. 2011. TouchLogger: Inferring Keystrokes on Touch Screen from Smartphone Motion. In *USENIX Workshop on Hot Topics in Security*.

[12] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao. 2019. Adversarial Sensor Attack on LiDAR-based Perception in Autonomous Driving. In *ACM Conference on Computer and Communications Security*.

[13] W. Chen, L. Chen, Y. Huang, X. Zhang, l. Wang, R. Ruby, and K. Wu. 2019. Taprint: Secure Text Input for Commodity Smart Wristbands. In *Annual International Conference on Mobile Computing and Networking*.

[14] H. Choi, W. Lee, Y. Aafer, F. Fei, Z. Tu, X. Zhang, D. Xu, and X. Xinyan. 2018. Detecting Attacks Against Robotic Vehicles: A Control Invariant Approach. In *ACM Conference on Computer and Communications Security*.

[15] A. Das, N. Borisov, and M. Caesar. 2016. Tracking Mobile Web Users Through Motion Sensors: Attacks and Defenses. In *Network and Distributed System Security Symposium*.

[16] D. Davidson, H. Wu, R. Jellinek, V. Singh, and T. Ristenpart. 2016. Controlling UAVs with Sensor Input Spoofing Attacks. In *USENIX Workshop on Offensive Technologies*.

[17] R. Dean, N. Burch, M. Black, A. Beal, and G. Flowers. 2011. Microfibrous metallic cloth for acoustic isolation of a MEMS gyroscope. *International Society for Optical Engineering* 7979 (2011), 797909.

[18] R. N. Dean, S. T. Castro, G. T. Flowers, G. Roth, A. Ahmed, A. S. Hodel, B. E. Grantham, D. A. Bittle, and J. P. Brunsch. 2011. A Characterization of the Performance of a MEMS Gyroscope in Acoustically Harsh Environments. *IEEE Transaction on Industrial Electronics* 58, 7 (2011), 2591–2596.

[19] R. N. Dean, G. T. Flowers, A. S. Hodel, G. Roth, S. T. Castro, R. Zhou, A. Moreira, A. Ahmed, R. Rifki, B. E. Grantham, D. Bittle, and J. Brunsch. 2007. On the Degradation of MEMS Gyroscope Performance in the Presence of High Power Acoustic Noise. In *IEEE International Symposium on Industrial Electronics*.

[20] S. Dey, N. Roy, W. Xu, R. R. Choudhury, and S. Nelakuditi. 2014. AccelPrint: Imperfections of Accelerometers Make Smartphones Trackable. In *Network and Distributed System Security Symposium*.

[21] J. L. Esteves, E. Cottais, and C. Kasmi. 2018. Unlocking the access to the effects induced by IEMI on a civilian UAV. In *International Symposium on Electromagnetic Compatibility*.

[22] M. Euston, P. Coote, R. Mahony, J. Kim, and T. Hamel. 2008. A complementary filter for attitude estimation of a fixed-wing UAV. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

[23] B. Farshteindiker, N. Hasidim, A. Grosz, and Y. Oren. 2016. How to Phone Home with Someone Else's Phone: Information Exfiltration Using Intentional Sound Noise on Gyroscopic Sensors. In *USENIX Workshop on Offensive Technologies*.

[24] T. F.Hueter and R.H.Bolt. 1955. *Sonics*. John Wiley & Sons.

[25] M. Gao, F. Lin, W. Xu, M. Nuermaimaiti, J. Han, W. Xu, and K. Ren. 2020. Deaf-Aid: Mobile IoT Communication Exploiting Stealthy Speaker-to-Gyroscope Channel. In *Annual International Conference on Mobile Computing and Networking*.

[26] GRAS Acoustics, Inc. 2000. GRAS 46AM 1/2" CCP Free-field Standard Microphone Set, Wide Frequency. https://www.grasacoustics.com/products/measurement-microphone-sets/constant-current-power-ccp/product/551-46am.

[27] J. Han, A. J. Chung, and P. Tague. 2017. PitchIn: Eavesdropping via Intelligible Speech Reconstruction Using Non-acoustic Sensor Fusion. In *ACM/IEEE International Conference on Information Processing in Sensor Networks*.

[28] Harman Inc. 2019. JBL STADIUM GTO750T. https://www.onlinecarstereo.com/CarAudio/p_51143_JBL_STADIUMGTO750T.aspx.

[29] J. Hou, X. Li, P. Zhu, Z. Wang, Y. Wang, J. Qian, and P. Yang. 2019. SignSpeaker: A Real-time, High-Precision SmartWatch-based Sign Language Translator. In *Annual International Conference on Mobile Computing and Networking*.

[30] X. Ji, Y. Cheng, Y. Zhang, K. Wang, C. Yan, W. Xu, and K. Fu. 2021. Poltergeist: Acoustic Adversarial Machine Learning against Cameras and Computer Vision. In *IEEE Symposium on Security and Privacy*.

[31] V. Kaajakari. 2009. *Practical MEMS: Design of Microsystems, Accelerometers, Gyroscopes, RF MEMS, Optical MEMS, and Microfluidic Systems*. Small Gear Publishing.

[32] A. J. Kerns, D. P. Shepard, J. A. Bhatti, and T. E. Humphreys. 2014. Unmanned Aircraft Capture and Control Via GPS Spoofing. *Journal of Field Robotics* 31, 4 (2014), 617–636.

[33] S. K. Khaitan and J. D. McCalley. 2015. Design Techniques and Applications of Cyberphysical Systems: A Survey. *IEEE Systems Journal* 9, 2 (2015), 350–365.

[34] S. Khazaaleh, G. Korres, M. A. Eid, M. Rasras, and M. F. Daqaq. 2019. Vulnerability of MEMS Gyroscopes to Targeted Acoustic Attacks. *IEEE Access* 7 (2019), 89534–89543.

[35] M. Kraft and N. M. White. 2013. MEMS for automotive and aerospace applications. *MEMS for automotive tire pressure monitoring systems* (2013), 54–77.

[36] D. F. Kune, J. Backes, S. S. Clark, D. Kramer, M. Reynolds, K. Fu, Y. Kim, and W. Xu. 2013. Ghost talk: Mitigating EMI signal injection attacks against analog sensors. In *IEEE Symposium on Security and Privacy*.

[37] A. G. Kuznetsov, Z. S. Abutidze, B. I. Portnov, V. I. Galkin, and A. A. Kalik. 2011. Development of MEMS sensors for aircraft control systems. *Gyroscopy and Navigation* 2, 1 (2011), 59–62.

[38] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao. 2012. A Reliable and Accurate Indoor Localization Method Using Phone Inertial Sensors. In *ACM Conference on Ubiquitous Computing*.

[39] J. Liu, C. Wang, Y. Chen, and N. Saxena. 2017. VibWrite: Towards Finger-input Authentication on Ubiquitous Surfaces via Physical Vibration. In *ACM Conference on Computer and Communications Security*.

[40] X. Liu, Z. Zhou, W. Diao, Z. Li, and K. Zhang. 2015. When Good Becomes Evil: Keystroke Inference with Smartwatch. In *ACM Conference on Computer and Communications Security*.

[41] N. C. Macdonald, K. A. Shaw, and S. G. Adams. 2001. Microelectromechanical accelerometer for automotive applications. *Smart Materials Bulletin* 2001, 5 (2001), 16.

[42] P. Marquardt, A. Verma, H. Carter, and P. Traynor. 2011. (sp)iPhone: decoding vibrations from nearby keyboards using mobile phone accelerometers. In *ACM Conference on Computer and Communications Security*.

[43] MI, Inc. 2016. Mi Robot Builder. https://www.mi.com/global/mi-robot-builder.

[44] Y. Michalevsky, D. Boneh, and G. Nakibly. 2014. Gyrophone: Recognizing Speech from Gyroscope Signals. In *USENIX Security Symposium*.

[45] E. Miluzzo, A. Varshavsky, S. Balakrishnan, and R. R. Choudhury. 2012. Tapprints: your finger taps have fingerprints. In *International Conference on Mobile Systems, Applications, and Services*.

[46] S. Narain, T. D. Vo-Huu, K. Block, and G. Noubir. 2016. Inferring User Routes and Locations Using Zero-Permission Mobile Sensors. In *IEEE Symposium on Security and Privacy*.

[47] National Instruments, Inc. 2017. NI VituralBench 8012. https://www.ni.com/pdf/manuals/371527e.pdf.

[48] J. Noh, Y. Kwon, Y. Son, H. Shin, D. Kim, J. Choi, and Y. Kim. 2019. Tractor Beam: Safe-hijacking of Consumer Drones with Adaptive GPS Spoofing. *ACM Transactions on Privacy and Security* 22, 2 (2019), 12:1–12:26.

[49] Y. Park, Y. Son, H. Shin, D. Kim, and Y. Kim. 2016. This Ain't Your Dose: Sensor Spoofing Attack on Medical Infusion Pump. In *USENIX Workshop on Offensive Technologies*.

[50] P. PX4. 2021. Open Source Autopilot for Drones - PX4 Autopilot. https://px4.io.

[51] O. Pöllny, A. Held, and F. Kargl. 2021. The Effect Of Sound On The Gyroscopes In Your Car. In *IEEE Vehicular Technology Conference*.

[52] R. Quinonez, J. Giraldo, L. E. Salazar, E. Bauman, A. A. Cárdenas, and Z. Lin. 2020. SAVIOR: Securing Autonomous Vehicles with Robust Physical Invariants. In *USENIX Security Symposium*.

[53] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen. 2012. Zee: Zero-Effort Crowdsourcing for Indoor Localization. In *Annual International Conference on Mobile Computing and Networking*.

[54] Raspberry Pi. 2021. Raspberry Pi Zero W. https://www.raspberrypi.org/pi-zero-w/.

[55] N. Roy, M. Gowda, and R. R. Choudhury. 2015. Ripple: Communicating through Physical Vibration. In *USENIX Symposium on Networked Systems Design and Implementation*.

[56] N. Roy, H. Hassanieh, and R. Roy Choudhury. 2017. BackDoor: Making Microphones Hear Inaudible Sounds. In *ACM SIGMOBILE International Conference on Mobile Systems, Applications, and Services*.

[57] F. Sabath. 2011. What can be learned from documented Intentional Electromagnetic Interference (IEMI) attacks?. In *URSI General Assembly and Scientific Symposium*.

[58] J. Selvaraj, G. Y. Dayanıklı, N. P. Gaunkar, D. Ware, R. M. Gerdes, and M. Mina. 2018. Electromagnetic Induction Attacks Against Embedded Systems. In *ACM on Asia Conference on Computer and Communications Security*.

[59] J. Shen, J. Y. Won, Z. Chen, and Q. A. Chen. 2020. Drift with Devil: Security of Multi-Sensor Fusion based Localization in High-Level Autonomous Driving under GPS Spoofing. In *USENIX Security Symposium*.

[60] H. Shin, D. Kim, Y. Kwon, and Y. Kim. 2017. Illusion and dazzle: Adversarial optical channel exploits against lidars for automotive applications. In *International Conference on Cryptographic Hardware and Embedded Systems*. 445–467.

[61] Y. Shoukry, P. Martin, Y. Yona, S. Diggavi, and M. Srivastava. 2015. PyCRA: Physical Challenge-Response Authentication For Active Sensors Under Spoofing Attacks. In *ACM Conference on Computer and Communications Security*.

[62] Y. Son, J. Noh, J. Choi, and Y. Kim. 2018. GyrosFinger: Fingerprinting Drones for Location Tracking Based on the Outputs of MEMS Gyroscopes. *ACM Transactions on Privacy and Security* 21, 2 (2018), 10:1–10:25.

[63] Y. Son, H. Shin, D. Kim, Y. Park, J. Noh, K. Choi, J. Choi, and Y. Kim. 2015. Rocking Drones with Intentional Sound Noise on Gyroscopic Sensors. In *USENIX Security Symposium*.

[64] P. Soobramaney, G. Flowers, and R. Dean. 2015. Mitigation of the Effects of High Levels of High-Frequency Noise on MEMS Gyroscopes Using Microfibrous Cloth. In *Asme International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*.

[65] STMicroelectronics. 2018. STM32-F407VET6 Datasheet. https://www.mouser.cn/datasheet/2/389/cd00191185-1796739.pdf.

[66] J. Sun, Y. Cao, Q. A. Chen, and Z. M. Mao. 2020. Towards Robust LiDAR-based Perception in Autonomous Driving: General Black-box Adversarial Sensor Attack and Countermeasures. In *USENIX Security Symposium*.

[67] Y. Sun, P. Guo, L. Feng, C. Xing, and J. Wu. 2020. A Filtering Algorithm of MEMS Gyroscope to Resist Acoustic Interference. *Sensors* 20, 24 (2020), 7352.

[68] W. T. Thomson. 1981. *Theory of vibration with applications*. Prentice Hall.

[69] N. O. Tippenhauer, C. Pöpper, K. B. Rasmussen, and S. Capkun. 2011. On the requirements for successful GPS spoofing attacks. In *ACM Conference on Computer and Communications Security*.

[70] T. Trippel, O. Weisse, W. Xu, P. Honeyman, and K. Fu. 2017. WALNUT: Waging Doubt on the Integrity of MEMS Accelerometers with Acoustic Injection Attacks. In *IEEE European Symposium on Security and Privacy*.

[71] Y. Tu, Z. Lin, I. Lee, and X. Hei. 2018. Injected and Delivered: Fabricating Implicit Control over Actuation Systems by Spoofing Inertial Sensors. In *USENIX Security Symposium*.

[72] Y. Tu, S. Rampazzi, B. Hao, A. Rodriguez, K. Fu, and X. Hei. 2019. Trick or Heat? Manipulating Critical Temperature-Based Control Systems Using Rectification Attacks. In *ACM Conference on Computer and Communications Security*.

[73] C. Wang, X. Guo, Y. Wang, Y. Chen, and B. Liu. 2016. Friend or Foe?: Your Wearable Devices Reveal Your Personal PIN. In *ACM on Asia Conference on Computer and Communications Security*.

[74] H. Wang, T. T. Lai, and R. R. Choudhury. 2015. MoLe: Motion Leaks through Smartwatch Sensors. In *Annual International Conference on Mobile Computing and Networking*.

[75] Z. Wang, K. Wang, B. Yang, S. Li, and A. Pan. 2017. Sonic Gun to Smart Devices: Your Devices Lose Control Under Ultrasound/Sound. In *Blackhat USA*.

[76] Z. Wang, W. Zhu, J. Miao, H. Zhu, C. Chao, and O. K. Tan. 2005. Micromachined thick film piezoelectric ultrasonic transducer array. *Sensors & Actuators A Physical* 130-131 (2005), 485–490.

[77] H. Wen, J. Ramos Rojas, and A. K. Dey. 2016. Serendipity: Finger Gesture Recognition Using an Off-the-Shelf Smartwatch. In *Conference on Human Factors in Computing Systems*.

[78] Wikipedia. 2021. Acoustic wave. https://en.wikipedia.org/wiki/Acoustic-wave.

[79] C. Wu, K. He, J. Chen, Z. Zhao, and R. Du. 2020. Liveness is Not Enough: Enhancing Fingerprint Authentication with Behavioral Biometrics to Defeat Puppet Attacks. In *USENIX Security Symposium*.

[80] W. Xu, C. Yan, W. Jia, X. Ji, and J. Liu. 2018. Analyzing and Enhancing the Security of Ultrasonic Sensors for Autonomous Vehicles. *IEEE Internet Things Journal* 5, 6 (2018), 5015–5029.

[81] X. Xu, J. Yu, Y. chen, Q. Hua, Y. Zhu, Y.-C. Chen, and M. Li. 2020. TouchPass: Towards Behavior-Irrelevant on-Touch User Authentication on Smartphones Leveraging Vibrations. In *Annual International Conference on Mobile Computing and Networking*.

[82] Z. Xu, K. Bai, and S. Zhu. 2012. TapLogger: inferring user inputs on smartphone touchscreens using on-board motion sensors. In *ACM Conference on Security and Privacy in Wireless and Mobile Networks*.

[83] Y. Xue, J. Chen, W. Wan, Y. Huang, C. Yu, T. Li, and J. Bao. 2019. MVSCRF: Learning Multi-View Stereo With Conditional Random Fields. In *IEEE/CVF International Conference on Computer Vision*.

[84] C. Yan, H. Shin, C. Bolton, W. Xu, Y. Kim, and K. Fu. 2020. SoK: A Minimalist Approach to Formalizing Analog Sensor Security. In *IEEE Symposium on Security and Privacy*.

[85] Q. Yan, K. Liu, Q. Zhou, H. Guo, and N. Zhang. 2020. SurfingAttack: Interactive Hidden Attack on Voice Assistants Using Ultrasonic Guided Waves. In *Network and Distributed System Security Symposium*.

[86] X. Yuan, Y. Chen, Y. Zhao, Y. Long, X. Liu, K. Chen, S. Zhang, H. Huang, X. Wang, and C. A. Gunter. 2018. CommanderSong: A Systematic Approach for Practical Adversarial Voice Recognition. In *USENIX Security Symposium*.

[87] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu. 2017. Dolphinattack: Inaudible voice commands. In *ACM conference on computer and communications security*.

[88] J. Zhang, A. R. Beresford, and I. Sheret. 2019. SensorID: Sensor Calibration Fingerprinting for Smartphones. In *IEEE Symposium on Security and Privacy*.

[89] Y. Zhang and K. Rasmussen. 2020. Detection of electromagnetic interference attacks on sensor systems. In *IEEE Symposium on Security and Privacy*.