# Exploring Practical Acoustic Transduction Attacks on Inertial Sensors in MDOF Systems

Ming Gao, Lingfeng Zhang, Leming Shen, Xiang Zou, Jinsong Han, *Senior Member, IEEE,*
Feng Lin, *Senior Member, IEEE,* and Kui Ren, *Fellow, IEEE*

**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 (MDOF) 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.

**Index Terms**—Cyber-physical system, inertial sensors, acoustic transduction attacks, spoofing attacks, IoT security

## 1 Introduction

Cyber-physical systems (CPSs) are widely deployed in various areas, including consumer electronics, health care, industry, and military deployment. 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 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 [2], [3], [4]. With this property, attackers can disturb the operation of target systems. For example, a drone would crash under ultrasonic interference [3]. The influence raised by acoustic transduction attacks is significantly devastating. Even worse, existing countermeasures to those attacks, e.g., acoustic isolation [5], [6], seem ineffective in an embedded environment [7].

With the above effect, attackers will naturally develop more strategical attacks to maliciously control CPSs [8]. State-of-the-art (SOTA) research has proposed to deliberately modulate acoustic signals [7], [9], instead of denial of service (DoS) attacks via disordered noise [3], [4]. However, the potential of such sophisticated attacks is not well studied due to the limited attacking scopes and scenarios targeted by existing approaches. First, in SOTA attacks, a unsettled issue is that desired false signals are distorted by unpredictable frequency offset due to sampling rate drifts [7], [10]. The impacts of SOTA attacks seem to be constrained because the desired stable false signals are distorted by the frequency offset. Second, 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. Third, existing research only involves **stationary** targets and ignores the influence of motion. In real-world scenarios, however, the systems' motion mode is 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.*

Answering this question is difficult because it is confronted with two challenges to realize the strategical acoustic attack in real CPS systems. (1) *How to manipulate multi-degrees-of-freedom (MDOF) devices that can move freely in space?* 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 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 [10]. Thus, existing attacking approaches [7], [9] 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

- *Ming Gao and Lingfeng Zhang, are with the School of Cyber Science and Technology, Zhejiang University, Hangzhou 310027, China, and also with the ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou 311200, China.*
*Email: gaomingppm@zju.edu.cn, lingfengzhang@zju.edu.cn.*
- *Leming Shen is with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong 999077, China.*
*Email: leming.shen@connect.polyu.hk.*
- *Xiang Zou is with the School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China.*
*Email: Xiang_Zou@stu.xjtu.edu.cn.*
- *Jinsong Han (corresponding author), Feng Lin, and Kui Ren are with the School of Cyber Science and Technology, Zhejiang University, Hangzhou 310027, China, and Zhejiang Provincial Key Laboratory of Blockchain and Cyberspace Governance, Hangzhou, 310000, China.*
*E-mail: hanjinsong@zju.edu.cn, flin@zju.edu.cn, kuiren@zju.edu.cn.*

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

(2) *How to suppress the influence of the movement of the target on false signals?* 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 influence of acoustic transduction attacks on moving targets is badly underestimated because their potential has not been fully dug. With the aid of camera-based distance measurement, remote attacks are competent to cover moving systems with a single-axis sensor, while such remote attacks cannot inject stable false signals into moving systems with multi-axis sensors. 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 [11]). 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 and realize a sophisticated attack, namely KITE, which effectively controls drone-like systems, as illustrated in Fig. 1. 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 [3], [7], [9], [12] without any enhancement. Our method involves accurate frequency and phase estimation. It supports a stable and controllable injection (with the identical effect to the attacks in [9]). In particular, we realize the automatic offset compensation, without which false signals would distort [7], [10] and the attack's effect would be constrained. In comparison, existing approaches [7], [9] 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). In addition, we experimentally observe that the natural frequency of an accelerometer is typically below 10 kHz while that of a gyroscope exceeds 18 kHz. Thus, due to the inaudibility of the malicious ultrasound to humans, attacks on gyroscopes become stealthier and more covert.

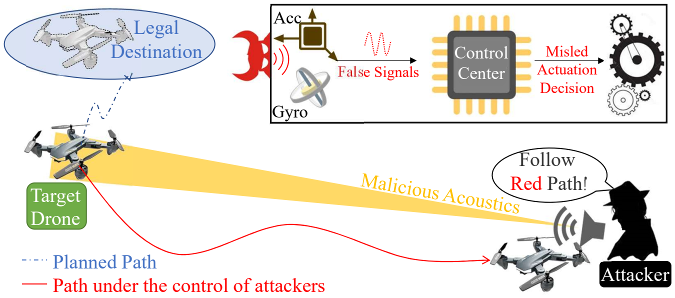


Fig. 1. KITE aims at controlling CPSs by spoofing inertial sensors.

Our contribution can be summarized as follows:

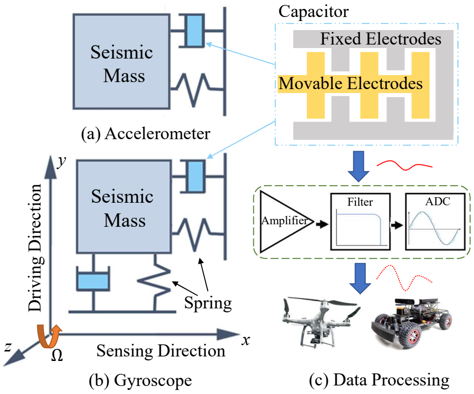
- We perform a comprehensive analysis on practical threats to CPSs from acoustic transduction attacks. We extend the attack surface to MDOF systems and demonstrate that such attacks pose a serious threat to inertia-based systems.
- We propose a new acoustic modulation method to manipulate the injected false signals as the attackers expect.
- 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.
- We release our source code [13] to facilitate successive research on CPSs' security and corresponding defenses.

## 2 Inertial Sensors

In this section, we provide the background about inertial sensors and their vulnerability to acoustics interference.

Inertial sensors comprise accelerometers for observing linear acceleration and gyroscopes for detecting angular velocity. They share a similar damping structure [14], as illustrated in Fig. 2. 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 [15]. 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 [2]. 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 [3], [4], [11], but also adversarial control on CPSs [7], [9]. These attacks succeed in manipulating stationary targets, e.g., self-balancing human transporters, self-balancing



(a) Accelerometer (b) Gyroscope (c) Data Processing

Fig. 2. Typical structures and data processing in inertial sensors.

robots, and smartphones. They can also be applied to interfere in computer vision based object detection systems by spoofing inertial sensors of image stabilizers [12]. However, SOTA attacks either select a target with a single-axis sensor or care only one axis of a multi-axis sensor.

# 3 Threat Analysis

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

## 3.1 Attack Scenarios

We divide possible attack scenarios into 2×2 types, according to the target's 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 [7]). 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., collision detection).

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. [7] 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 [10], SOTA attacks cannot directly organize the desired false signals onto an assigned axis. Taking a drone as an example, attackers merely want to modify its yaw angle readings, but it may crash rather than follow the misguided path due to unexpected injected rolling or pitching under SOTA attacks. In short, they fail in the orientation control on MDOF systems.

### *3.1.2 Stationary vs.Moving*

SOTA attacks conduct control on targets that are stationary or in a well-balanced status [7], [9], where inertial readings are originally zero. They respond merely to the acoustics and just output false signals.

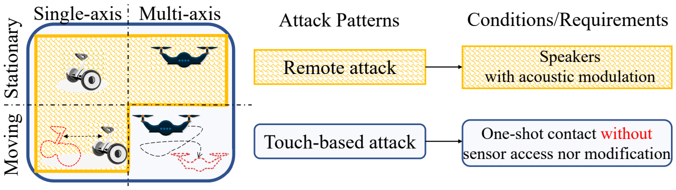


Fig. 3. Threats of the proposed attacks.

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 [4], [7], [9], [12] 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 [7] or their remote controllers [9]; (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 targets invasively because most CPSs prohibit such access rigorously without users' permission [16].

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

- *Remote Attack*. Attackers emit acoustic signals using nearby malicious sources, for example, placing a Bluetooth speaker or a smartphone adjacent to the target devices (including smartphones, unmanned vehicles, and other CPSs). In particular, the attackers can conduct a drive-by auto-play attack, in which malicious audios are distributed on the Internet via browsers and emails in a stealthy manner. It unintentionally auto-plays malicious audios onto the victim's computer or mobile device. During the attacks, the victim does not have to click on anything, press download, or open a malicious email attachment to become infected [17].
- *Touch-based Attack*. Attackers afford the one-shot temporary 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 (that can effectively emit acoustic signals into water or sound, represented by a piezoelectric (PZT) transducer) to the shell of target systems and emits malicious acoustics following the attackers' expectations. For example, attackers can buy off an 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. With the assumption that attackers have physical access to the target, our proposed touch-based attack has the minimum requirement, compared with other touch-based attacks. We merely require a one-shot temporary

physical contact with the target. The requirement can be easily met by either a passer-by or a robot. In comparison, other touch-based attacks require additional efforts or permission to the inner structure of the target, such as downloading and running malicious codes for hacking the system, accessing and modifying the sensors' data. Such operations are relatively time-consuming and prone to alerting the user. Moreover, if the attackers are able to conduct other attacks, our touch-based attack can be combined together for more serious threats or serve as an alternative in case the other attacks are interrupted by the inner defense mechanism in the target systems.

### 3.3 Attack Stage

Combining the above analysis, we exploit the full potential of acoustic transduction attacks, as illustrated in Fig. 3. 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 (which would result in crashes, and also deserves being concerned and taking countermeasures) on the multi-axis under the impact of motion (see Sec.6.2). That is, the remote attack can control both stationary and moving single-axis targets and stationary multi-axis targets as the attackers expect, which indeed threats the security of CPSs in practice. 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. It represents the real damage to the security of CPSs.

### 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 acoustic modulation. 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 $F_0$ and $\omega_r$ are the initial amplitude and frequency, respectively. The resonant response in one axis of the sensor [10] 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 $\varphi_r$,

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

where $\xi$ is the constant damping ratio. For a given sensor, $a_r$ and $\varphi_r$ depend solely upon the injected frequency [7].

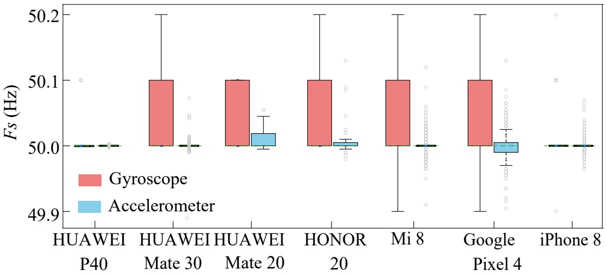


Fig. 4. Sampling rate drifts are common in inertial sensors.

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 $\omega_d$ 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 [7], leading to unpredictable frequency offset, where $\omega_d^* = \omega_r - 2\pi n(Fs + \Delta Fs)$ replaces $\omega_d$ 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 in Fig. 4 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 [7], a slight drift might cause serious signal distortion.

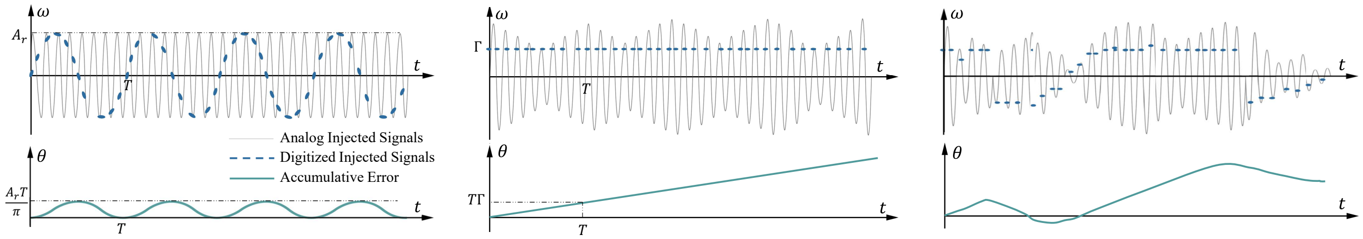
### 4.2 Stable and Controllable Injections

In pursuit of adversarial control, we modulate acoustic signals by modifying the 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 [9]) and then adjust it to desired waveforms.

**Challenge**. Frequency offset caused by the sampling rate drift [7], [10] would distort injections and degrade the attack effect into DoS. It is a challenge to *compensate the unpredictable and random offset*.

[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.



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

Fig. 5. Basic idea of the proposed acoustic modulation.

**SOTA approaches**. Existing approaches, e.g., WALNUT [9] and Poltergeist [12] set $A_r = \Gamma(t)$ and $\omega_r = 2\pi nFs$ 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 [7]. 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. [7] 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 [7]. 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,

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

Here, the additional phase $\varphi_r$ 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. 5. Under unmodulated acoustics, the digitized injected signals vary sinusoidally, with a tiny accumulative signal as shown in Fig. 5(a). Using our modulation method, we can obtain a constant digitized injected signal as presented in Fig. 5(b). By adjusting the acoustic intensity as illustrated in Fig. 5(c), we can generate false signals with arbitrary waveforms following the attackers' expectations. To achieve this, a fundamental issue is to estimate $\omega_d$ and $\varphi_r$.

### 4.2.1 Frequency Determination and Offset Compensation

It is difficult to calculate $\omega_d$ due to the lack of knowledge about targets' sampling rate drifts. We exploit an unalterable frequency relationship to calculate $\omega_d$ and eliminate the influence of frequency offsets.

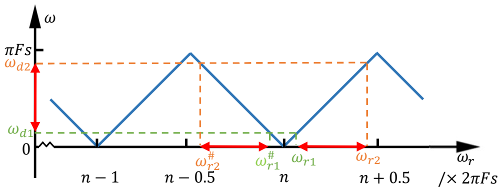


Fig. 6. Attackers can infer and determine $\omega_r$ according to the unalterable frequency difference, where $\omega_{r2} - \omega_{r1} = \omega^{\#}_{r1} - \omega^{\#}_{r2} = \omega_{d2} - \omega_{d1}$ (indicated by red arrows).

**Frequency Difference**. We observe that the frequency difference between acoustic signals also migrates into the low-frequency band after being digitized, as illustrated in Fig. 6. To be specific, suppose that two signals of $\omega_{ri}$ $(i = 1, 2, \omega_{ri} = 2\pi nFs + \omega_{di})$ can resonate with the target sensor. We have the following unalterable frequency relationship,

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

**Offset compensation**. This difference-based technique still works even if the sampling rate is drifting. Attackers can obtain an appropriate $\omega_{r1}$ 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 $\omega_{r1}$ and $\omega_{r2}$. 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 $\omega_{r1}$ serves as a reference for the offset compensation. In practical attacks, the offset $\omega^*_{d1}$ 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., $\xi$ and $\omega_n$ in Eq. 2), we cannot obtain $\varphi_r$ directly. Instead, we exploit the resonant phase-frequency characteristics to estimate the exact phase.

With the derivative of $\varphi_r$ 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 $\omega_r$ approaches $\omega_n$ and $|\omega_r - \omega_n| \ll \omega_n$ under resonance [10], $\frac{\omega_r}{\omega_n} \approx 1$, and $\varphi'_r$ can be approximately recognized as a constant. It reveals that the phase lag $\varphi_r$ has a positive linear correlation with acoustic frequency $\omega_r$. $\varphi'_r$ can be measured on sensors of the identical mode in advance. Considering that the reference signal supplies the feedback about both $\omega_{d1}$ and $\varphi_{r1}$, we can reckon malicious acoustic signals' phase lag $\varphi_{r2}$ as follows,

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

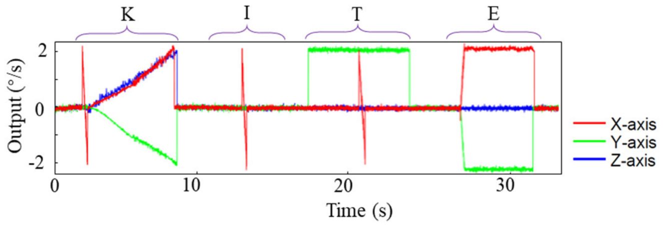


Fig. 7. Spelling a 'KITE' trajectory by manipulating a BMI160 IMU using our proposed method.

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 $\epsilon$ 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

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

Thus, we realize the stable and controllable injection $\Gamma(t)$. Figure 7 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 [7], [9], [12] 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' expectations.

**Challenge**. Injections on one axis would disturb those on other axes, because resonance would occur simultaneously on multiple axes [10]. However, SOTA attacks [7], [9] ignore this issue, which 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. 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. 8. 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 [14]. Imagine that an acoustic source is placed along the X-axis of an inertial sensor as shown in Fig. 8(a). It just interferes in the x-axial acceleration and the yaw ($Z_{gyro}$) angular velocity. In particular, the acoustic source along the X-axis does not affect the pitch ($Y_{gyro}$) angular velocity in the gyroscope. Recalling Fig. 2(b), the malicious acoustics along the X-axis would resonate with the damping structure on the sensing direction (i.e., the X-axis in this case) and thus produce false signals on the yaw axis due to the Coriolis force [14]. If targeted at the pitch ($Y_{gyro}$) angular velocity in the gyroscope, the malicious source should be placed along the Z-axis, which is the sensing direction for pitch ($Y_{gyro}$) angular velocity, as shown in Fig. 8(c). The sensing directions for the roll ($X_{gyro}$), pitch ($Y_{gyro}$), and yaw ($Z_{gyro}$) are typically along the Y-, Z-, and X-axis respectively according to the conventional regulations in the relevant manufacturers [1], [18]. We conclude the relationships as follows,

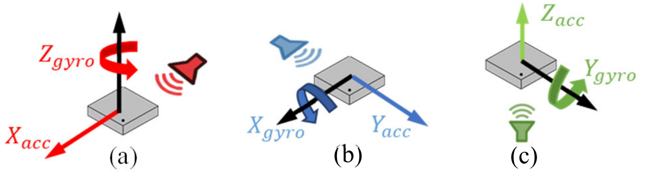
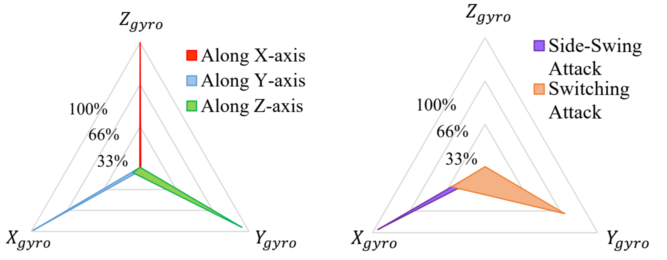


Fig. 8. An illustration of affected axes of acoustic sources along (a) X-axis, (b) Y-axis, and (c) Z-axis.



(a) Our solution (b) Two demonstrations in [7]

Fig. 9. Energy distribution among axes. A smaller area means a better orientation control.

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

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 [19].

**Solution**. We utilize multiple acoustic sources to compensate for the orientation deviation. By adjusting each source's acoustic intensity independently, the attackers do not need to move in the physical world and the false signals would follow a given spatial vector with an assigned direction. In comparison, in the attack using one source at arbitrary azimuth, the attackers have to adjust the location of the source to modify the orientation of false signals, which is cumbersome. It is recommended to utilize three sources that constitute a set of three-dimensional (orthogonal, if possible) bases. Note that a set of non-orthogonal bases are also effective after a coordinate system conversion [20].

We represent energy distributions in the gyroscope of an iPhone 7 in Fig. 9, 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 [7], 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 model the impact of targets' motion 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, resulting in the distortion of acoustic signals and attendant resonant responses. We denote the distance variation as $\Delta L$. 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 [7], [9], [12] 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 $\Omega$ is the angular velocity around Z-axis to be measured. The angular velocity $\Omega$ 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 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 $\Omega$; 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 would result in the Doppler frequency offset. Nevertheless, this problem can be solved using the offset compensation method 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 $\Delta L$ to compensate for the phase fluctuation. In KITE, we adopt MVSCRF [21] 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 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 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.

### 6.3.1 *Acoustic Attacks Travelling in Solid*

Acoustic guided waves can propagate in solid media [19]. 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. 10, 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 $\mu$W. 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 experiments on other materials, including copper metal, plastic (polythene), wood (fiberboard and log table), and glass, which have covered most of the 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. 10(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). In addition, we consider acoustic damping material, which is made of fibre materials and is able to reduce the acoustic intensity. KITE can be implemented successfully on aluminium metal, copper metal, plastic, glass, fiberboard, and log table, but it fails on the acoustic damping material. The results show the material would affect the possibility of implementing the attack. Fortunately, such an acoustic damping material is not commonly used in COST CPSs. Therefore, KITE can cover most CPSs in practice. Furthermore, the material mainly affects the attacking distance with little influence on the accuracy of the manipulation once the attack is conducted.

The attack distance of ultrasonic attacks on gyroscopes may differ from that of audible attacks on accelerometers. There are two reasons. On the one hand, the propagation of ultrasound is commonly shorter than audible sound. On the other hand, the sensibility of a gyroscope to ultrasound is different from that of an accelerometer to sound.

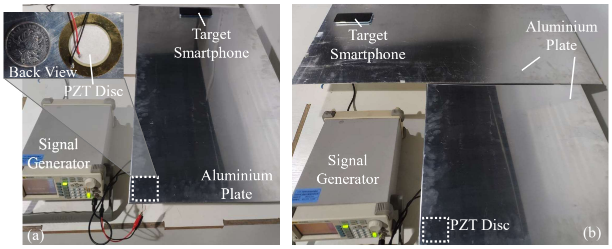


Fig. 10. Experimental setup for the feasibility study of acoustic attacks via (a) single-layer and (b) multi-layer (overlapping) solid media.

TABLE 1
Maximum Attack Distance on Different Materials

| Material | Size | Attack Distance[1] | |
|---|---|---|---|
| | | On Acc. | On Gyro. |
| Aluminium metal | 1 m× 0.5 m× 2 mm | 1.12 m+ | 1.12 m+ |
| | 1 m× 0.5 m× 5 mm | 1.12 m+ | 1.12 m+ |
| | 1 m× 0.5 m× 7 mm | 1.12 m+ | 1.12 m+ |
| Copper metal | 1 m×0.1 m×0.2 mm | 1.01 m+ | 1.01 m+ |
| Plastic | 0.75 m×0.7 m×2 mm | 1.02 m+ | 1.02 m+ |
| Glass | 0.9m×0.45 m×10 mm | 1.01 m+ | 1.01 m+ |
| Fiberboard | 0.75 m×0.7 m×10 mm | 1.03 m+ | 1.03 m+ |
| Log table | 1.2 m×0.75 m×15 mm | 0.33 m | 0.32 m |
| Aluminium (2 mm)⊕[2] Aluminium (5 mm) | | 1.12 m+ | 1.12 m+ |
| Aluminium (2 mm) ⊗[3] Aluminium (5 mm) | | 1.50 m+ | 1.50 m+ |
| Aluminium (7 mm)⊕ Copper | | 1.12 m+ | 1.12 m+ |
| Aluminium (7 mm) ⊗ Plastic | | 1.25 m+ | 1.25 m+ |
| Aluminium (7 mm) ⊗ Fiberboard | | 1.25 m+ | 1.25 m+ |
| Aluminium (7 mm) ⊕ Plastic ⊕ Fiberboard | | 0.71 m | 0.65 m |

[1]: ‘+’ means that acoustics can affect both accelerometers and gyroscopes in the target smartphones and such attacks remain at least as effective over a potentially longer distance via solid media.
[2]: ⊕ means that the multilayered media wholly overlap.
[3]: ⊗ means that the multilayered media partially overlap.

Compared with the ultrasound speakers used in the prior literature [3], [4], [7], the PZT transducers are cheaper and smaller. They can covertly adhere to the target's shell for 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. 7.8.

### 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. $\Omega$ in Eq. 16), and thus the control center could reduce the noise caused by the coupling effect.

We integrate the touch-based attack into a printed circuit board (PCB) prototype, as shown in Fig. 11. It carries an STM32-F407VET6 chip and a Raspberry Pi Zero W as the control center, a digital to analog converter (DAC902), a BMI160 inertial sensor (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 onto targets to emit malicious acoustics. The size of the whole board is 13 cm in length and 7 cm in width. Note that this prototype is a proof-

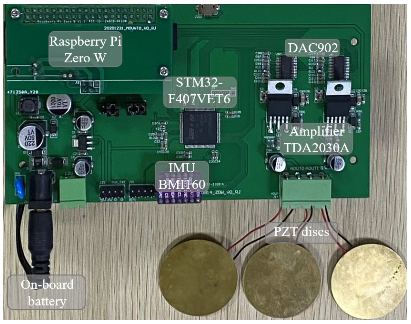


Fig. 11. Proof-of-concept of the malicious unit (PCB board prototype) for touch-based attacks.

of-concept (without elaborated integration). In a 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 \$26, without the need of expensive signal generators and loudspeakers, which are necessary in SOTA attacks [7], [9], [12].

### 6.4 Automatic Attack

A further requirement of practical attacks is the automatic conduction. The key challenge is how to *adaptively generate desired false signals according to targets' timely motion states and cancel the frequency offset*.

The on-board inertial sensor of the malicious unit induces two-fold advantages. First, given a trajectory that can be loaded in advance or sent using wireless signals, the control center of the target compares the current state with the one corresponding to the trajectory, and then prepares acoustic injections for desired false signals automatically. Second, it detects the frequency offset and leverages reference signals in Sec. 4.2.1 for automatic compensation. Thus, attackers can manipulate a CPS to follow the maliciously assigned trajectory without manual adjustment.

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.

### 7.1 Experiment Setup

**Target systems.** We first carry out experiments on a BMI055 chip [22] that is widely deployed in COTS CPSs (e.g., Oculus Rift and Pixhawk 4) for directly gathering the raw inertial data for quantitative analysis. 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. 7.7. 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 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.

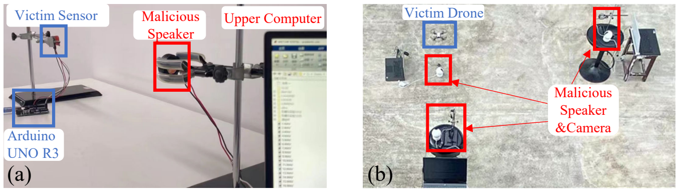


Fig. 12. Experimental setups.

**Acoustic source.** In remote attacks, we use JBL 750T speakers 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, connected to an upper computer, modulates the signals and drives the speaker. In touch-based attacks, we exploit the PCB prototype in Fig. 11 as the malicious devices for attacks.

**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. 12. 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

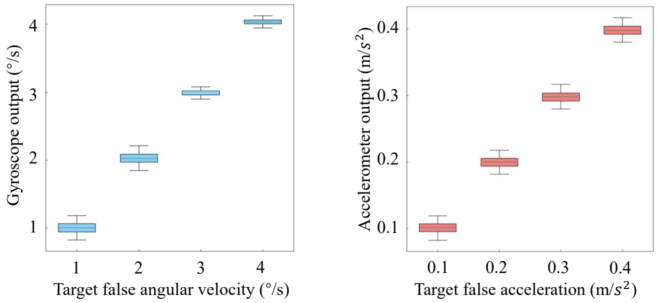
We evaluate KITE in injecting desired false signals with a controllable orientation on stationary targets.

#### *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. 12(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

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

Fig. 13. Amplitude control.

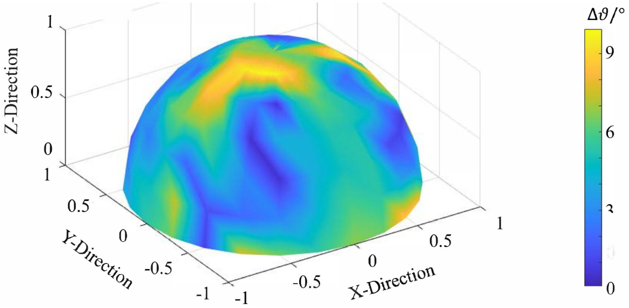


Fig. 14. Orientation control over a drone. A smaller $\Delta\vartheta$ represents a better performance.

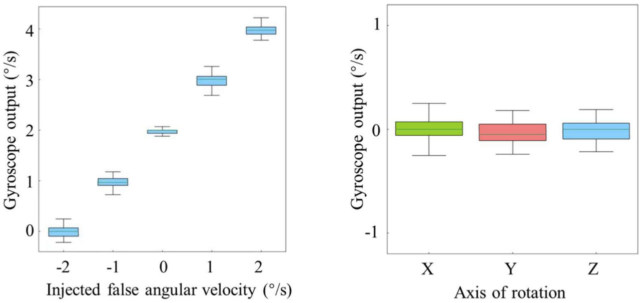
results, only with a slight rise in terms of standard deviation. Fig. 13 further demonstrates the precision of our proposed acoustic modulation on diverse values. It obtains a low error with the standard deviation of about 0.08°/s on average. Such deviation would not increase with the amplitude of false signals. It peaks at 0.091°/s under the injection 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 [16] 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, 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. 12(b). The standard deviation of false signal is below 0.1°/s. The thermodynamic diagram in Fig. 14 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. 8(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°.



(a) Different injections (b) Different axes

Fig. 15. Robustness against motion.

## 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 [21] 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. 15(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 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. 15(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.

When attacking moving targets with multi-axis sensors, remote attacks merely act as DoS. We conduct experiments

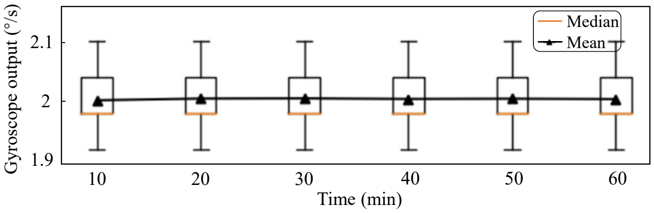


Fig. 16. Stability of long-time automatic attacks.

TABLE 3
Robustness against Environmental Noise (°/s)

| Noise Level | Median | Mode | Mean | Standard deviation | Range |
|---|---|---|---|---|---|
| No | 2.013 | 2.013 | 2.031 | 0.077 | ±0.160 |
| Low | 2.001 | 2.001 | 2.009 | 0.071 | ±0.157 |
| Medium | 2.015 | 2.015 | 2.026 | 0.065 | ±0.157 |
| High | 1.999 | 1.999 | 2.002 | 0.088 | ±0.213 |

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 °/s under unmodulated acoustic injections using the same settings. It validates that remote attacks cannot apply to manipulating MDOF systems and frustrates SOTA attacks [7], [9] 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 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 [3]. In practical attacks, once the target enters the coverage of our proposed remote attack, the attacker could manipulate the target' trajectory. Therefore, the attacker can avoid the target moving out of the attack's effective range. Moreover, even if a moving target moves away due to the effect of inertia, the attacker is able to track it using a drone or a robot to continue the attack.

### 7.5 Automatic Touch-based Attack

To assess the stability of the automatic touch-based attack proposed in Sec. 6.4, we inject a prolonged false signal into an ATG-850 RTK drone that is hanging over the air. A reference signal with a small amplitude (below 0.3°/s) is emitted intermittently (0.1 s in every 2 s) in case offset occurs during attacks. The drone rotates under the attack that lasts for an hour. The planned value is 2°/s and actual values range from 1.93°/s to 2.09°/s. As shown in Fig. 16, the standard deviation maintains 0.106°/s. The deviation fluctuates irregularly, which reveals its independence of attack duration. It greatly extends the valid time of automatic attacks while SOTA attacks [7] merely last several minutes without manual adjustment. The total power consumption of our malicious unit measures 324 mA. The on-board 1500 mA battery supports continuous attacks of 4 hours experimentally. *Such stability and endurance enable a prolonged and automatic attack, exacerbating the threat of acoustic injections.*

### 7.6 Impact of Environmental Noise

We generate random noise using additional speakers as ambient noise. The noise intensity around the target measures 45 dB (no noise), 55 dB (low-noise), 70 dB (medium-noise), and 80 dB (high-noise) respectively. The value of target false signals is 2°/s. Table. 3 shows that the impact of environmental noise on KITE is nearly negligible. Moreover, common ultrasonic applications, e.g., medical examination, prefer frequency bands of over 40 kHz [7], [9]. Most gyroscopes resonate in the frequency band between 18 kHz and 30 kHz, where few devices work. Although accelerometers resonate with audible sound, the intensity of the environmental noise that has the same frequency as the accelerometer's natural frequency is tiny. Otherwise, the accelerometer cannot normally work due to interference from the environmental noise. We test KITE in six typical scenes in urban cities, i.e., an office, a café, a mall, a street side, a bus station, and a metro station. The attacking effect does not degrade to DoS with a low standard deviation below 0.1°/s when we inject a false signal of 2°/s.

Especially, motors and rotors would make tremendous noise during operation. We test the robustness of both remote and touch-based attacks against environmental noise by repeating false signals of 2°/s when an ATG-850 RTK drone moves at ±0.5m/s and ±1m/s and rotates at ±5°/s and ±10°/s respectively. The standard deviation of the false signals keeps within 0.19°/s. The above results confirm that KITE is robust to environmental noise.

### 7.7 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. 4, with the full list involving 28 COTS devices in [13]. In particular, there are multiple proportion integration differentiation (PID) controllers in the drone-like MDOF systems, for example, the ATG-850 RTK drone in Tab. 4, which has two PID controllers based on two inertial measurement units. Experimentally, 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 [16])

TABLE 4
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 |
| DJI Spark UAV | UnKonwn | 23.8 | 5.5 |
| QQL RC UAV | IS IMU3000 | 27.1 | 23 |
| Huawei P40 | Unknown | 19.9 | 4.6 |
| Huawei P20 Pro | IS ICM-20690 | 20.1 | 6.7 |
| HONOR V30 | IS ICM-20690 | 27.3 | - |
| Samsung Note 10 Plus | Unknown | 20.9 | 0.2 |
| 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 6 | Unknown | 26.9 | - |
| iPhone 6s Plus | IS MP67B | 27.2 | - |
| iPhone 7 | IS 773C | 27.2 | - |
| iPhone XS | BS BMI282 | 26.0 | - |
| 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 |
| iPad Air 3 | Unknown | 25.8 | - |
| iPad Pro 2020 | Unknown | 26.4 | - |
| MITU robot | IS ICM-20690 | 20.1 | 6.7 |
| Baidu Apollo D-KIT | Unknown | 27.5 | 5.2 |
| EAIBOT N1 UGV | M R6093U | 27.2 | 6.5 |
| Apple Watch Series 6 | BS BMI282 | 25.9 | - |
| AMAZFIT Mi | BS BMI160 | 19.8 | - |

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

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 [3], [4], [7], [9], [12] and our results, we conclude that KITE could affect most CPSs.

## 7.8 Inaudibility

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

### 7.8.1 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 [3]. We recruit 22 volunteers aged from 18 to 45 when remotely attacking gyroscopes of the devices in Tab. 4. They report being unable to distinguish the 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 their 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. 4. Therefore, malicious sounds emitted from remote sources can be heard by humans. SOTA attacks aimed at controlling accelerometers [9], [12] alert surrounding people, unless on some exceptions, e.g., a Samsung S20, which is also selected as the only target in [12] due to the embedded accelerometer's high natural frequency of 19.2 kHz.

TABLE 5
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 |

### 7.8.2 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 the 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. 10. 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. 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.

We further conduct a proof-of-concept attack on a drone (DJI Spark UAV) to validate the inaudibility of touch-based attacks. This drone hangs about 2.5 m over the ground. 20 volunteers (aged between 18 and 35) stand 2 to 5 m away from the drone. After 10 demonstrations of legal operation, volunteers are asked to report the possibility of being attacked using scores ranging from 1 to 5, where '5' represents that the target is under adversarial control in all likelihood. The drone moves along 15 tracks for 14 times, half of which are legal and the others are under the acoustic transduction attack. Tab. 5 lists the audibility results. The volunteers distinguish attacks with a score of 3.41 on average, and that of real motion is 3.29 by comparison. The results reflect that volunteers cannot judge the existence of attacks at all. They claim to never tell the difference between real and false motion. Besides, two microphones are placed 10 cm and 5 m away at the same height as the drone to record the acoustic intensity, where the intensity of ambient noise is approximately 50 dB. As a baseline, the intensity when the drone hangs without attacks is also measured. Results in Tab. 5 demonstrate that the malicious acoustic signals are covered totally by the mechanical noise during the operation of the target. In short, touch-based attacks on both gyroscopes and accelerometers remain inaudible.

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

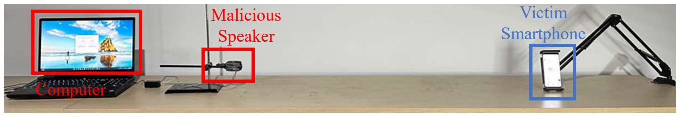


Fig. 17. Experimental setup for cases study on smartphones.

## 8 End-to-End Attack Cases Study

We now evaluate the proposed attacks with end-to-end cases on COTS devices. Four end-to-end attacks demonstrate the attack effects by manipulating the route of a drone embedded in the most popular autopilot (Pixhawk 4) and spoofing the motion-driven applications on smartphones.

### 8.1 On Smartphones

In smartphones, inertial readings are utilized for navigation services, pedometer applications and the like. We select three typical motion-driven applications on smartphones (i.e., navigation services, pedometers, and gait-based authentications) as the targets in our end-to-end attack cases. The experimental setup is shown in Fig. 17. The malicious JBL 750T speaker emits modulated acoustic signals to spoof victim smartphones that are 1 m away.

#### 8.1.1 Navigation services

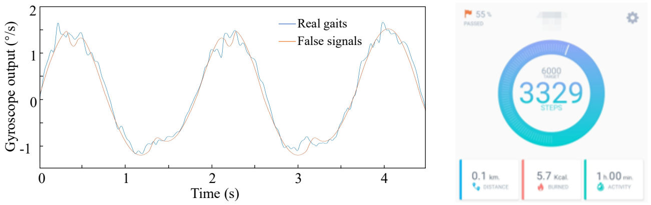
Using the remote attack, we accumulate a false yawing angle of up to approximately 6.23 rads or -6.19 rads in 1 minute in a Huawei P40, with few false signals in the pitch nor roll. Thus, attackers can deceive navigation services such that they would misguide users into the wrong routes. We conduct KITE on 'Baidu Map', a popular navigation application. We can remotely alter the 'orientation' reported by the navigation service, during which the victim smartphone has never moved or turned physically. The victim smartphones are placed towards the north, while the 'Baidu Map' are spoofed by KITE. It reports that the smartphones are rotating and finally face west or east under the control of remote attackers, with the video demo in [13]. Moreover, KITE attacks successfully on other popular navigation services, including 'Gaode Map' and 'Google Maps'.

#### 8.1.2 Pedometer

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. 18(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. 18(b). Here we have not claimed in-app rewards.

#### 8.1.3 Gait-based Authentication

We further consider a attack on gait-based authentication. We implement an authentication system based on a smartphone accelerometer for gait recognition [23] and achieve an 18% false positive rate (FPR) (approaching the FPR of 10% in [23]) when the smartphone is fixed at the hip of a user. By collecting 30 minutes of gait data from the user, KITE generates false gait signals to imitate the user and achieves a 41% chance of bypassing the authentication system [23] (much greater than that without our attack, i.e., the FPR of 18%). The result shows that KITE is capable of misleading gait-based authentication. This also reveals the vulnerability of other motion-driven applications to KITE, e.g., inertial sensor based identification [24], [25], [26], [27].



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

(b) Screenshot of 3300 false steps.

Fig. 18. Attacks on pedometer applications on smartphones.

### 8.2 On A Drone

Unmanned vehicles, e.g., drones, depend on inertial readings for attitude estimation and autonomous navigation.

#### 8.2.1 DoS under Remote Attacks

We conduct a DoS attack on drones using the proposed remote attacks. Note that the remote attack can only control single-axis targets and stationary multi-axis targets (e.g., smartphones in Sec. 8.1) and merely conduct DoS attacks on moving multi-axis targets (e.g., drones) as analyzed in Sec. 6.2 and evaluated in Sec. 7.3.1. A DJI Spark UAV rotates at about 1°/s at a height of 4 meters in the air. We set a malicious JBL 750T speaker to emit modulated acoustic signals that are 1 meter away horizontally at a height of 1 meter, with a camera running the MVSCRF algorithm for calculating the distance accordingly to Sec. 6.2. Under the effect of the remote attack, the standard deviation of inertial readings in the DJI Spark UAV can reach up to 8.89°/s. In this case, the inertial sensor cannot measure the real motion state of the drone. The drone is shaky and unsteady, and finally crashes, with the attack dome presented in [13].

#### 8.2.2 Trajectory Manipulation under Touch-based Attacks

In the case of the touch-based attack, our designed 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. 19(a). The drone flies about 200 m above the ground in an open space. In the drone, a complementary filter [28] is employed for attitude estimation using the data collected by the BMI055 inertial sensor. GPS is forbidden to reveal the threat to inertial sensors here. In addition, GPS signals may lose in some cases (e.g., due to electromagnetic interference), and the attacks can continue with GPS spoofing [29], [30].

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 following Line 'OA' and then turns north at the angular velocity of about 5°/s to follow Line '$AD_{A0}$', as the green line in Fig. 19(b). In the practical trajectory, the maximum deviation in Line 'OA' is 3.6 m and that in Line '$AD_{A0}$' is 14.2 m, while the practical destination point '$D_A$' is 4.9 m away from the preset destination point '$D_{A0}$'. We

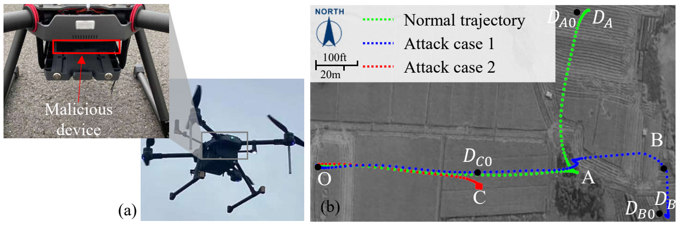


Fig. 19. Attacks on a drone. (a) The malicious unit is attached onto the target covertly. (b) Touch-based attacks manage to manipulate the target's trajectory.

conduct two attacks that manipulate trajectories as the blue and red lines in Fig. 19(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. 19(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 the '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. The maximum deviation of the practical trajectory to the desired trajectory in Line 'AB' is 15.1 m. Compared to that in Line 'A$D_{A0}$' of 14.2 m, such a deviation is acceptable for the adversarial control. The practical destination point '$D_B$' is 0.8 m away from the desired destination point '$D_{B0}$', which also demonstrates that the attacker can manipulate the drone to precisely follow the attacker's desired 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. 19(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. The maximum deviation of the practical trajectory to the desired trajectory in Line 'O$D_{C0}$' is 9.8 m and the practical destination point 'C' is 9.8 m away from the desired destination point '$D_{C0}$. Such deviations are tolerable considering the braking time of about 10 s (which is to blame for the deviation, while the maximum deviation in Line 'O$D_{C0}$' except for the last 10 m is merely 5.1 m).

Therefore, the attacker is able to manipulate the drone to precisely follow the attacker's desired trajectory. We can also achieve the similar effect on other drones including a QQL RC UAV and a DJI Spark UAV.

# 9 Discussion and Defense

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

## 9.1 Comparison with SOTA Attacks

We implement two typical SOTA attacks [7], [9] on inertial sensors and compare them with KITE in Tab. 6.

WALNUT [9] performs acoustic transduction attacks on the accelerometer in a smartphone. It plays malicious ultrasound using the on-board speaker of the same smartphone (remote controller), which remotely controls a car (the target) to go forth or back. This attack is easily disturbed by frequency offsets due to the sampling rate drift. Or they may raise acoustic intensity to saturate the inner amplifier, yet produce non-adjustable outputs under audible injections with deafening volume. In detail, WALNUT produces stable false signals via the amplitude modulation which 'modulates amplitude of clipping at the amplifier [9]. For a given amplifier, the amplitude of clipping is a fixed value, which is decided by its hardware. For example, the accelerometer amplitudes of clipping in a BMI055 [22] are ±2 g, ±4 g, ±8 g, ±16 g (respectively accordingly to different visions), and the gyroscope amplitudes of clipping are ±125°/s, ±250°/s, ±500°/s, ±1000°/s, and ±2000°/s. To achieve the amplitudes of clipping of ±2 g in a BMI055 accelerometer, it requires a malicious sound source of 109 dB experimentally, which is placed 1 m away from the sensor. As for the other (higher) amplitudes of clipping, the required acoustic intensity is higher proportionally. The intensity of over 109 dB would indeed damage the hearing of surrounding people (in comparison with the car whistle noise of about 80∼110 dB). To achieve the amplitudes of clipping of ±125°/s in a BMI055 gyroscope, it requires an ultrasound source of 121 dB experimentally. Although the ultrasound is beyond the human hearing, the human-exposure ultrasonic limit suggested by the International Non-Ionizing Radiation Committee is less than 110 dB. Otherwise, a high-volume ultrasound would harm human health, let alone consumes a lot of power. Poltergeist [12] follows the same principles and suffers from the identical shortcomings.

Tu et al. [7] bypasses the problem of offsets by generating controllable accumulated errors while the injections are still unstable, oscillating in a fixed pattern that can serve as a detector for defense. Though they have developed automatic attacks on Google Maps and can already adaptively control the direction of the navigation app. The injected false orientation changes in an unnatural and periodic pattern, while our injection follows a stable orientation change, which seems like an actual user's real actions. Moreover, the audio-based attack [9] works only without the sampling rate drift that occurs frequently [7], and thus it requires manual adjusting once the sampling rate drifts. The existing automatic attacks [7] merely last several minutes without manual adjustment, while the valid time of automatic attacks in KITE can be extended to over 1 hour.

In short, prior works hardly conduct controllable and stable injections. In contrast, KITE achieves this goal of stable injection, not to mention that KITE also has other advantages, such as orientation control, motion robustness,

TABLE 6
Comparison with SOTA Attacks

| Attacks | Attack Effect | Stability against Offset | Attack Target | | | |
|---|---|---|---|---|---|---|
| | | | Single-axis sensor | | Multi-axis sensor | |
| | | | Stationary | Moving | Stationary | Moving |
| WALNUT [9] | Controllable but unstable (under offset) or fixed injections | × | ✓ | × | × | × |
| Tu et al. [7] | Controllable accumulated errors (angle) but unstable injections (angular velocity) | ✓ | ✓ | ✓ | × | × |
| KITE | Controllable and stable injections | ✓ | ✓ | ✓ | ✓ | ✓ |

and low cost. We list the innovations compared with SOTA attacks that are aimed at adversarial control over CPSs [7], [9] in Tab. 6. In comparison with existing approaches [7], [9], 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.

## 9.2 Discussion

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

### 9.2.1 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.

### 9.2.2 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.

## 9.3 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

### 9.3.1 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 [5]. Advanced dampening materials reach 90% acoustic reduction [6]. However, this method undoubtedly introduces significant costs. Besides, its resilience is unclear against attacks via solid propagation.

**Filtering.** Using low pass filters is another option to weaken acoustic effect [9]. However, the attacks still work even if the cut-off frequency is limited within 10 Hz due to hardware defects [2]. Sun et al. [31] propose a filter based on orthogonal demodulation, but the I/O dual channel is rare in existing inertial sensors.

**Common-mode difference.** Analog Devices, Inc. assembles a dual-core structure into industrial MEMS gyroscopes [32]. This structure outputs differential inertial data to suppress common-mode noise. They mainly focus on resisting vibrations. However, due to the tiny spatial distance, acoustics expose forces on these two cores with a significant phase difference, rather than common-mode ones. In this case, the differential structure is unprofitable or even unfavourable for inertial sensors against acoustic inference, where false signals might be amplified rather than eliminated.

**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 [33], [34], 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 threat various sensors, including GPS [29], [30], LiDARs [35], [36], camera sensors [37], microphones [19], [38], [39], [40], [41]. An advanced power-switching method [42], [43] is effective against electromagnetic interference. However, it is inapplicable to detecting transduction attacks.

**Sampling.** Normally, attackers modulate acoustics based on the target sensors' sampling rate. Conversely, it is feasible to modify sampling intervals. Trippel et al. [9] propose two defense mechanisms. One is the randomized sampling which 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 $\omega_d = 4\pi n Fs$ in Eq. 4. Tu et al. [7] recommend a dynamic $Fs$ based on the randomized sampling mechanism. Nevertheless, it has not yet alleviated the problem of degraded accuracy in inertial measurements.

### 9.3.2 *Our Solutions*

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 the 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 signal to noise ratio (SNR) [44] according to the frequency $\omega$ as follows,

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

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 the adversarial control and degrade such attacks to DoS.

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 $\alpha_m$s are small constants and $C$ is an arbitrary constant. Here we set $C = 2$ and $\alpha_0 = -\alpha_1$. In the comparison of random [9] or dynamic [7] ones, the periodically alternating jitters have a smaller root mean square with the adjustable $\alpha_m$s. Hence, our countermeasure significantly mitigates the effect of spoofing attacks, at the cost of exerting little adverse influence on inertial measurements.

To further eliminate the threat from acoustic transduction attacks, we leverage the resonant characteristics among axes in an inertial sensor to detect and separate the injected false signals. Recalling Sec. 5, there still remains approximately 1% of the resonate energy in the non-target axes. Although such tiny remnants pose no impact on the malicious orientation control, they can be exploited for the detection of acoustic transduction attacks. The false signals on different axes share the identical frequency and keep a fixed phase difference [10]. In comparison, the real motion in gyroscopes barely presents identical frequency in all axes at the same time. This phenomenon matches the intuitive expectations, in which the normal CPSs would not rotate around its three axes at the same rate simultaneously. We recruit seven volunteers and record the inertial data on their smartphones for two weeks, during which these volunteers work, walk, run, bike, drive, and so on as usual. We also collect a total of 720-hour operating data from inertial sensors on four drones and one unmanned ground vehicle and one robot. Over 99.9% of the above gyroscope data show diverse frequencies on different axes at the same time. In below 0.1% gyroscope data, two axes in a gyroscope share the same frequency but the other axis rotates at a different frequency. The results demonstrate that the frequency difference can serve as a standard to detect acoustic transduction attacks. Correspondingly, we analyze the frequencies of all axes using the short-time Fourier transform (STFT). The components with the frequency that occurs at all axes simultaneously and with the amplitude of over 0.06°/s empirically are removed. A pilot experiment on a BMI160 gyroscope verifies the effectiveness of our proposed defending method. In our future work, we will improve the real-time defense by optimizing the computational complexity and expand it onto the accelerometer protection, for example, by leveraging both the frequency and phase difference among axes in the resonant characteristics.

## 10 Related Work

**Privacy Leakage through Inertial Sensors.** Different from the pursuit of inertial data tampering, several attacks utilize IMUs for privacy theft, including speech [16], [45], keystroke [46], [47], physical activity [48], [49], and localization [50], [51], [52]. Through stealing readings from inertial sensors, to which most systems do not restrict access, attackers can know what you say, what you do and where you are. Moreover, inertial sensors can also leak users' behavioral biometrics [24], [25], [26], [27]. Additionally, there exists different fixed offsets in the outputs of different accelerometers and gyroscopes due to the manufacturing errors, which can be considered as a fingerprint for device identification [53].

**Sensor Spoofing Attacks.** Such spoofing attacks are increasingly risking the security of CPSs. A slew of sensors are suffering from electromagnetic interference (EMI). With EM injection, the attacker can make the sensor produce erroneous outputs, or even further seize control of the target system. Related researches show that LiDARs systems [35], [36], GPS [29], [30] and camera sensors [37] are also vulnerable to spoofing attacks, which is a great threat to the corresponding application systems. Additionally, ultrasound is often used as another medium for spoofing attacks on acoustic system. Inaudible commands can be injected into VAs using ultrasound, which benefits from the acoustic non-linearity [19], [39], [40], [41]. As a countermeasure, researchers usually utilize power-switching for defence [42], [43]. 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 [2]. Inertial sensors will produce abnormal outputs under the interference of a specific frequency of ultrasound. Not content with DoS attacks [3], [4], researchers [7], [9] pursue adversarial control by trying to control the outputs of the inertial sensors quantitatively. However, they are unable to achieve controllable waveform and orientation due to the frequency offset, multiaxial resonance, and the target's motion. These unavoidable factors make their attack method not practical. 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 [10], [54]. Through commercial off-the-shelf (COTS) loudspeakers and devices containing inertial sensors, using special coding methods, covert communication between the two can be established and the identification of the device can be carried out automatically.

## 11 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.

## ACKNOWLEDGMENTS

This paper is partially supported by the National Key R&D Program of China (2021QY0703), National Natural Science Foundation of China under grant U21A20462, 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), Basic Research Project of Shenzhen Science and Technology Innovation Commission (Project Number: JCYJ20190812155213250), and Shenzhen Longhua District Science and Technology innovation special fund project (Project Number: JCYJ201903).

## REFERENCES

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

[2] S. Khazaaleh, G. Korres, M. A. Eid, M. Rasras, and M. F. Daqaq, "Vulnerability of MEMS gyroscopes to targeted acoustic attacks," *IEEE Access*, vol. 7, pp. 89534–89543, 2019.

[3] Y. Son, H. Shin, D. Kim, Y. Park, J. Noh, K. Choi, J. Choi, and Y. Kim, "Rocking drones with intentional sound noise on gyroscopic sensors," in *USENIX Security Symposium*, pp. 881–896, 2015.

[4] Z. Wang, K. Wang, B. Yang, S. Li, and A. Pan, "Sonic gun to smart devices: Your devices lose control under ultrasound/sound," in *Blackhat USA*, pp. 1–50, 2017.

[5] R. Dean, N. Burch, M. Black, A. Beal, and G. Flowers, "Microfibrous metallic cloth for acoustic isolation of a mems gyroscope," *International Society for Optical Engineering*, vol. 79, pp. 1–9, 2011.

[6] P. Soobramaney, G. Flowers, and R. Dean, "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*, pp. 9–14, 2015.

[7] Y. Tu, Z. Lin, I. Lee, and X. Hei, "Injected and delivered: Fabricating implicit control over actuation systems by spoofing inertial sensors," in *USENIX Security Symposium*, pp. 1545–1562, 2018.

[8] M. Gao, L. Zhang, L. Shen, X. Zou, F. Lin, J. Han, and K. Ren, "Kite: Exploring the practical threat from acoustic transduction attacks on inertial sensors," in *ACM Conference on Embedded Networked Sensor Systems*, pp. 1–14, 2022.

[9] T. Trippel, O. Weisse, W. Xu, P. Honeyman, and K. Fu, "WALNUT: waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks," in *IEEE European Symposium on Security and Privacy*, pp. 3–18, 2017.

[10] M. Gao, F. Lin, W. Xu, M. Nuermaimaiti, J. Han, W. Xu, and K. Ren, "Deaf-aid: Mobile iot communication exploiting stealthy speaker-to-gyroscope channel," in *Annual International Conference on Mobile Computing and Networking*, pp. 1–13, 2020.

[11] B. Farshteindiker, N. Hasidim, A. Grosz, and Y. Oren, "How to phone home with someone else's phone: Information exfiltration using intentional sound noise on gyroscopic sensors," in *USENIX Workshop on Offensive Technologies*, 2016.

[12] X. Ji, Y. Cheng, Y. Zhang, K. Wang, C. Yan, W. Xu, and K. Fu, "Poltergeist: Acoustic adversarial machine learning against cameras and computer vision," in *IEEE Symposium on Security and Privacy*, pp. 160–175, 2021.

[13] KITE, "Kite." https://github.com/KITE-anonymous-user/KITE.git, 2022.

[14] J. J. Bernstein, S. Cho, A. T. King, A. Kourepenis, and M. Weinberg, "A micromachined comb-drive tuning fork rate gyroscope," in *IEEE Micro Electro Mechanical Systems*, pp. 143–148, 1993.

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

[16] Z. Ba, T. Zheng, X. Zhang, Z. Qin, B. Li, X. Liu, and K. Ren, "Learning-based practical smartphone eavesdropping with built-in accelerometer," in *Network and Distributed System Security Symposium*, 2020.

[17] A. K. Sood and S. Zeadally, "Drive-by download attacks: A comparative study," *IT Professional*, vol. 18, no. 5, pp. 18–25, 2016.

[18] Bosch, Inc., "Bmi160 datasheet." https://www.bosch-sensortec.com/products/motion-sensors/imus/bmi160.html, 2018.

[19] Q. Yan, K. Liu, Q. Zhou, H. Guo, and N. Zhang, "Surfingattack: Interactive hidden attack on voice assistants using ultrasonic guided waves," in *Network and Distributed System Security Symposium*, 2020.

[20] W. Qiu, *Analytical Geometry (Third Edition)*. Peking University Press, 2015.

[21] Y. Xue, J. Chen, W. Wan, Y. Huang, C. Yu, T. Li, and J. Bao, "MVSCRF: learning multi-view stereo with conditional random fields," in *IEEE/CVF International Conference on Computer Vision*, pp. 4312–4321, 2019.

[22] Bosch, "Bmi055 datasheet.." https://www.bosch-sensortec.com/products/motion-sensors/imus/bmi055/documents, 2020.

[23] Y. Ren, Y. Chen, M. C. Chuah, and J. Yang, "User verification leveraging gait recognition for smartphone enabled mobile healthcare systems," *IEEE Transactions on Mobile Computing*, vol. 14, no. 9, pp. 1961–1974, 2015.

[24] C. Wu, K. He, J. Chen, Z. Zhao, and R. Du, "Liveness is not enough: Enhancing fingerprint authentication with behavioral biometrics to defeat puppet attacks," in *USENIX Security Symposium*, pp. 2219–2236, 2020.

[25] X. Xu, J. Yu, Y. chen, Q. Hua, Y. Zhu, Y.-C. Chen, and M. Li, "Touchpass: Towards behavior-irrelevant on-touch user authentication on smartphones leveraging vibrations," in *Annual International Conference on Mobile Computing and Networking*, pp. 1–13, 2020.

[26] W. Chen, L. Chen, Y. Huang, X. Zhang, l. Wang, R. Ruby, and K. Wu, "Taprint: Secure text input for commodity smart wristbands," in *Annual International Conference on Mobile Computing and Networking*, pp. 1–16, 2019.

[27] J. Liu, C. Wang, Y. Chen, and N. Saxena, "Vibwrite: Towards finger-input authentication on ubiquitous surfaces via physical vibration," in *ACM Conference on Computer and Communications Security*, pp. 73–87, 2017.

[28] M. Euston, P. Coote, R. Mahony, J. Kim, and T. Hamel, "A complementary filter for attitude estimation of a fixed-wing uav," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 340–345, 2008.

[29] N. O. Tippenhauer, C. Pöpper, K. B. Rasmussen, and S. Capkun, "On the requirements for successful GPS spoofing attacks," in *ACM Conference on Computer and Communications Security*, pp. 75–86, 2011.

[30] J. Shen, J. Y. Won, Z. Chen, and Q. A. Chen, "Drift with devil: Security of multi-sensor fusion based localization in high-level autonomous driving under GPS spoofing," in *USENIX Security Symposium*, pp. 931–948, 2020.

[31] Y. Sun, P. Guo, L. Feng, C. Xing, and J. Wu, "A filtering algorithm of MEMS gyroscope to resist acoustic interference," *Sensors*, vol. 20, no. 24, p. 7352, 2020.

[32] Analog Devices, Inc., "Adxl362 datasheet." https://www.analog.com/media/en/technical-documentation/data-sheets/ADXL362.pdf, 2017.

[33] R. Quinonez, J. Giraldo, L. E. Salazar, E. Bauman, A. A. Cárdenas, and Z. Lin, "SAVIOR: securing autonomous vehicles with robust physical invariants," in *USENIX Security Symposium*, pp. 895–912, 2020.

[34] H. Choi, W. Lee, Y. Aafer, F. Fei, Z. Tu, X. Zhang, D. Xu, and X. Xinyan, "Detecting attacks against robotic vehicles: A control invariant approach," in *ACM Conference on Computer and Communications Security*, pp. 801–816, 2018.

[35] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao, "Adversarial sensor attack on lidar-based perception in autonomous driving," in *ACM Conference on Computer and Communications Security*, pp. 2267–2281, 2019.

[36] J. Sun, Y. Cao, Q. A. Chen, and Z. M. Mao, "Towards robust lidar-based perception in autonomous driving: General black-

box adversarial sensor attack and countermeasures," in *USENIX Security Symposium*, pp. 877–894, 2020.

[37] D. Davidson, H. Wu, R. Jellinek, V. Singh, and T. Ristenpart, "Controlling uavs with sensor input spoofing attacks," in *USENIX Workshop on Offensive Technologies*, 2016.

[38] D. F. Kune, J. Backes, S. S. Clark, D. Kramer, M. Reynolds, K. Fu, Y. Kim, and W. Xu, "Ghost talk: Mitigating emi signal injection attacks against analog sensors," in *IEEE Symposium on Security and Privacy*, pp. 145–159, 2013.

[39] X. Yuan, Y. Chen, Y. Zhao, Y. Long, X. Liu, K. Chen, S. Zhang, H. Huang, X. Wang, and C. A. Gunter, "Commandersong: A systematic approach for practical adversarial voice recognition," in *USENIX Security Symposium*, pp. 49–64, 2018.

[40] N. Roy, H. Hassanieh, and R. Roy Choudhury, "Backdoor: Making microphones hear inaudible sounds," in *ACM SIGMOBILE International Conference on Mobile Systems, Applications, and Services*, pp. 2–14, 2017.

[41] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, "Dolphinattack: Inaudible voice commands," in *ACM conference on computer and communications security*, pp. 103–117, 2017.

[42] Y. Shoukry, P. Martin, Y. Yona, S. Diggavi, and M. Srivastava, "Pycra: Physical challenge-response authentication for active sensors under spoofing attacks," in *ACM Conference on Computer and Communications Security*, pp. 1004–1015, 2015.

[43] Y. Zhang and K. Rasmussen, "Detection of electromagnetic interference attacks on sensor systems," in *IEEE Symposium on Security and Privacy*, pp. 203–216, 2020.

[44] B. Brannon and A. Barlow, "Aperture uncertainty and adc system performance." https://www.analog.com/media/en/technical-documentation application-notes/an-501.pdf, 2006.

[45] Y. Michalevsky, D. Boneh, and G. Nakibly, "Gyrophone: Recognizing speech from gyroscope signals," in *USENIX Security Symposium*, pp. 1053–1067, 2014.

[46] E. Miluzzo, A. Varshavsky, S. Balakrishnan, and R. R. Choudhury, "Tapprints: your finger taps have fingerprints," in *International Conference on Mobile Systems, Applications, and Services*, pp. 323–336, 2012.

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

[48] J. Hou, X. Li, P. Zhu, Z. Wang, Y. Wang, J. Qian, and P. Yang, "Signspeaker: A real-time, high-precision smartwatch-based sign language translator," in *Annual International Conference on Mobile Computing and Networking*, pp. 1–15, 2019.

[49] H. Wang, T. T. Lai, and R. R. Choudhury, "Mole: Motion leaks through smartwatch sensors," in *Annual International Conference on Mobile Computing and Networking*, pp. 155–166, 2015.

[50] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Annual International Conference on Mobile Computing and Networking*, pp. 293–304, 2012.

[51] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *ACM Conference on Ubiquitous Computing*, pp. 421–430, 2012.

[52] S. Narain, T. D. Vo-Huu, K. Block, and G. Noubir, "Inferring user routes and locations using zero-permission mobile sensors," in *IEEE Symposium on Security and Privacy*, pp. 397–413, 2016.

[53] J. Zhang, A. R. Beresford, and I. Sheret, "Sensorid: Sensor calibration fingerprinting for smartphones," in *IEEE Symposium on Security and Privacy*, pp. 638–655, 2019.

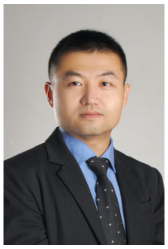
[54] N. Roy, M. Gowda, and R. R. Choudhury, "Ripple: Communicating through physical vibration," in *USENIX Symposium on Networked Systems Design and Implementation*, pp. 265–278, 2015.

**Ming Gao** is a Ph.D. candidate at the school of cyber science and technology, Zhejiang University. He received the Master and Bachelor degree from Xi'an Jiaotong University. His research interests include cyber-physical security, mobile computing, and privacy protection. He is a recipient of the Best Paper Award Nomination from SenSys'21.

**Lingfeng Zhang** is pursuing M.Sc. at the School of Cyber Science and Technology, Zhejiang University, under the supervision of Prof. Jinsong Han. His research interests include cyber-physical security and smart sensing.

**Leming Shen** received BS degree from Zhejiang University in 2022. He is working toward the Ph. D. degree at Department of Computing, The Hong Kong Polytechnic University. His research interests include mobile computing, wireless sensing, IoT security and applications.

**Xiang Zou** received his B.S. degree from XI'AN University of Posts Telecommunications, Xi'an, China, in 2014, the M.S. degree from the Chang'an University in 2018. He is pursuing the Ph.D. degree at Xi'an Jiaotong University, Xi'an, China. His research interests include RFID, smart sensing, and mobile computing.



**Jinsong Han** is now a professor at the School of Cyber Science and Technology, Zhejiang University. He is a senior member of the ACM and IEEE. His research interests focus on IoT security, smart sensing, wireless and mobile computing.

**Feng Lin** received the Ph.D. degree from the Department of Electrical and Computer Engineering, Tennessee Technological University, USA, in 2015. He is currently a Professor with the School of Cyber Science and Technology, College of Computer Science and Technology, Zhejiang University, China. He was an Assistant Professor with the University of Colorado Denver, USA, a Research Scientist with the State University of New York (SUNY) at Buffalo, USA, and an Engineer with Alcatel-Lucent (currently, Nokia). His current research interests include mobile sensing, wireless sensing, Internet of Things security, biometrics, and AI security. Dr. Lin was a recipient of the ACM SIGSAC China Rising Star Award, the Best Paper Awards from ACM MobiSys'20, IEEE Globecom'19, IEEE BHI'17, the Best Demo Award from ACM HotMobile'18, and the Best Paper Award Nomination from SenSys'21 and INFOCOM'21.

**Kui Ren** (Fellow, IEEE and ACM) received the PhD degree from Worcester Polytechnic Institute. He was the SUNY Empire Innovation professor with the State University of New York at Buffalo. He is currently a professor and the associate dean of College of Computer Science and Technology, Zhejiang University, where he also directs the Institute of Cyber Science and Technology. His research interests include cloud and IoT security.