SOUNDBOOST: Effective RCA and Attack Detection for UAV via Acoustic Side-Channel

Haoran Wang*, Zheng Yang*, Sangdon Park†, Yibin Yang*, Seulbae Kim†,
Willian Lunardi‡, Martin Andreoni‡, Taesoo Kim*, Wenke Lee*
*Georgia Institute of Technology, †Pohang University of Science and Technology, ‡Technology Innovation Institute
{haoran.wang, ianyang, yyang811, taesoo, wenke}@gatech.edu,
{sangdon, seulbae}@postech.ac.kr, {WillianTessaro.Lunardi, Martin.Andreoni}@tii.ae

***Abstract*—Unmanned Aerial Vehicles (UAVs), or drones, are emblematic examples of cyber-physical systems where computational components and physical processes integrate to enable autonomous navigation. UAVs rely heavily on sensors such as *Inertial Measurement Units* (IMU) and *Global Positioning System* (GPS) for accurate environmental awareness and control. However, the trust placed in these sensors makes UAVs vulnerable to adversarial attacks that compromise the UAV's operational integrity. While prior work focuses on detecting attacks against specific sensors, there remains a critical gap in performing *Root Cause Analysis* (RCA) to determine which component failed and why – especially under ambiguous or conflicting sensor reports. To address this gap, we propose SOUNDBOOST, a novel RCA framework that leverages the UAV's *acoustic side-channel* (i.e., sound) to diagnose navigation failures and attribute them to specific sensor compromises. While SOUNDBOOST detects attacks by validating GPS and IMU sensor data, it focuses on post-incident diagnosis. SOUNDBOOST conducts post-incident RCA by extracting robust acoustic signatures and using machine learning to cross-validate reported kinematics against physical behavior. We deploy SOUNDBOOST on a UAV and evaluate it under real-world GPS spoofing attacks and synthesized IMU biasing attacks. SOUNDBOOST achieves 100% true positive rate for IMU attacks and over 80% for GPS spoofing, outperforming the state-of-the-art by 21% – demonstrating its effectiveness as a practical forensic tool for sensor attack RCA.**

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs), commonly known as drones, are transforming domains such as surveillance [46], delivery [8], mapping [36], entertainment [24], and beyond [3], [25]. Their ability to perform autonomous flights has made them indispensable for both routine operations and high-stakes missions. However, the growing reliance on autonomy also expands the attack surface – particularly through the real-time sensors that underpin navigation and stability control. Among them, the *GPS* and the *IMU* are especially critical. Autonomous flight depends on them not only for basic positioning, but also for real-time responsiveness to environmental changes and overall flight integrity. For example, sensor attacks such as *GPS spoofing* [18], [38], [45], [52] and *IMU biasing attacks* [17], [48], [53], [55] severely compromise the UAV's navigation integrity and safety, posing risks of misnavigation, crashes or even hijacking. A notable incidence occurred in 2011 when Iran used GPS spoofing to land and capture a U.S. Lockheed Martin RQ-170 Sentinel drone [58].

Although existing defense approaches [10], [31], [40] perform well in distinguishing targeted GPS attacks from benign flight scenarios, they become ineffective when the trusted sensors (i.e., the IMU) utilized for attack detection are also under attack. For example, SAVIOR [40] utilizes the integrity of the IMU to detect GPS spoofing attacks. While this approach is effective under certain conditions, the *Rocking Drone* attack [48] highlights that the IMU itself can also be subject to attacks. This raises important considerations that when GPS and/or IMU are compromised, it is challenging to understand which sensor is the root cause. This highlights the pressing need for a solution that conducts Root Cause Analysis (RCA) for UAVs– explaining post-failure behaviors and providing forensics for determining sensor compromise.

Our System. Motivated by the above issue in the existing attack detection approaches, we propose SOUNDBOOST, a novel RCA framework designed to diagnose UAV navigation failures by leveraging the UAV's *acoustic side-channel* to validate the data from GPS and IMU sensors. While SOUNDBOOST detects attacks, it does not focus on real-time attack detection. It performs post-incident RCA to determine if a failure was attack-induced and, if so, identifies the compromised sensors using the UAV's acoustic emissions. Unlike the data from GPS or IMU sensors, which are susceptible to spoofing, the acoustic signatures generated by UAV's physical components during operation reliably reflect its movements. Thus, these acoustic signals are inherently resistant to spoofing attacks (§ IV-D), providing a trustworthy source for attack detection and RCA. By integrating the acoustic side-channel data with onboard sensor data, SOUNDBOOST provides a robust mechanism for identifying compromised sensors in the UAV navigation system. To the best of our knowledge, this is the *first work* capable of detecting IMU and GPS attacks altogether, conducting comprehensive RCA for UAV navigation systems.

Key Idea. SOUNDBOOST is built on a real-world observation: during flights, UAVs emit *acoustic signals (i.e., sounds)* that provide *supplementary* and *reliable* information of actuation. These signals are not only hard to spoof but also reflect physical outcomes of the UAV's behavior, making them ideal for post hoc validation of sensor data. By employing a microphone array to capture these sounds and analyzing them, SOUNDBOOST builds a detailed acceleration profile of UAVs that serves as a signature for performing RCA and detecting sensor attacks targeting GPS and IMU sensors. SOUNDBOOST is particularly valuable as it capitalizes on additional *reliable* information collected during *real flights* – previously unexploited – offering a new course to improve UAV security.

Workflow. SOUNDBOOST collects the acoustic signals from the UAV during real flight missions using a microphone array. These signals are filtered to isolate the sound frequencies primarily emitted by the UAV motors. Applying Fast-Fourier-Transform (FFT) analysis, SOUNDBOOST obtains the acoustic signatures for the flights (§ III-A). Next, in the sensory mapping phase, SOUNDBOOST trains an acoustic Deep Learning model to learn the correlation between the acceleration vector $\vec{a_x}, \vec{a_y}, \vec{a_z}$ of the UAV based on the ground truth data provided by an intact IMU. To minimize the impact of environmental factors (e.g., wind) and IMU measurement noise, augmented training data is employed (§ III-B). Once trained, SOUNDBOOST uses the model to predict acceleration vectors $\vec{a_x}', \vec{a_y}', \vec{a_z}'$ during flight missions to determine whether the IMU or GPS sensors are under attack (§ III-C).

More specifically, SOUNDBOOST conducts RCA for UAV navigation systems in two stages:

1) SOUNDBOOST determines whether the IMU is under attack by examining the distribution of distances between the predicted $\vec{a_x}', \vec{a_y}', \vec{a_z}'$ and the actual readings $\vec{a_x}, \vec{a_y}, \vec{a_z}$ from the IMU over a 0.5-second window (a hyper-parameter found in § IV-B). If the IMU is under attack, the observed distribution deviates significantly from the normal distribution observed under benign conditions, as shown in Fig. 6.
2) To detect GPS spoofing attacks, SOUNDBOOST uses the predicted $\vec{a_x}', \vec{a_y}', \vec{a_z}'$ to derive the prediction of the velocity vectors $\vec{v_x}', \vec{v_y}', \vec{v_z}'$ and compares them with the velocity readings $\vec{v_x}, \vec{v_y}, \vec{v_z}$ from the GPS. If it finds that the IMU is not compromised in the previous step, it fuses the sane IMU readings $\vec{a_x}, \vec{a_y}, \vec{a_z}$ with the predicted $\vec{a_x}', \vec{a_y}', \vec{a_z}'$ using a customized Kalman Filter to offer even more robust detection of GPS spoofing attacks.

Evaluation. We evaluate SOUNDBOOST against two classic sensor spoofing attacks: IMU biasing attacks [55] and GPS spoofing attacks [45] (§ IV-C). Our experiments show that SOUNDBOOST accurately identifies attacks on specific sensors and performs RCA with high precision. We highlight:

1) SOUNDBOOST's acoustic model for predicting can produce robust $\vec{a_x}', \vec{a_y}', \vec{a_z}'$, even in the presence of malicious interference. Concretely, SOUNDBOOST achieves to attribute the navigation system anomaly to IMU and identify *all* IMU biasing attack with a false positive rate as low as 10% in 20 flights.
2) Our experiment (Tab. II) shows that SOUNDBOOST can identify GPS anomaly and detect GPS spoofing attacks with 79% true positive rate when considering the IMU is not trustworthy, outperforming the state-of-the-art by 21% and the baselines [10], [15] by 53% and 11%, respectively.
3) SOUNDBOOST's detection rate is even higher at 89% when the IMU is determined intact.

II. BACKGROUND

This section provides the technical foundation for understanding UAV navigation systems and their vulnerabilities and motivates the need for robust a RCA technique.

A. UAV Control Systems

A typical UAV control system includes several key components: sensors (such as GPS and IMU), Proportional-Integral-Derivative (PID) controllers, actuators, and control algorithms. We review how these components interact to control a drone.

1) **Sensors** continuously provide data on position, speed, and acceleration. UAVs rely on a combination of sensors.
2) **Control Algorithms** estimate the UAV's states and drive it along a path. It continuously refines the estimations based on real-time sensor data to adapt to changing conditions.
3) **PID Controllers** interpret sensor data and compute control signals. The proportional component responds proportionally to the current deviation; the integral component addresses accumulated errors; and the derivative component reacts to error rate changes. They adjust motor speeds based on the difference between the desired setpoint and the current state.
4) **Actuators** are motor-driven propellers that respond to the control signals from the PID controllers to adjust the drone's thrust, roll, pitch, and yaw.
5) **Control Loop** integrates sensor inputs and control outputs. Sensor inputs (e.g., IMU and GPS data) are used for state estimations, which are processed by PID controllers, which then adjust the actuators. This control loop continuously updates, allowing the UAV to respond dynamically to the environment and flight commands.

B. UAV Navigation System

The GPS and IMU are the core technologies enabling real-time and precise UAV navigation, forming the foundation of most autonomous systems. GPS provides global positioning by leveraging satellite data to pinpoint locations within meters. However, its low update frequency limits effectiveness for real-time localization and rapid decision-making. IMU provides high-frequency local movement and orientation data using accelerometers and gyroscopes, enabling real-time stabilization and control. However, IMU is inaccurate compared to GPS and thus is prone to cumulative drift errors over time if used alone. To address these limitations, UAVs fuse GPS and IMU data using a Kalman Filter (KF), which reduces sensor noise and precisely estimates physical states by combining the precision of GPS with the responsiveness of IMU.

C. Attack Surfaces

Nevertheless, this reliance on GPS and IMU technologies exposes the UAV navigation system to larger attack surfaces. Among the prevalent risks, GPS spoofing and IMU biasing attacks stand out as substantial threats to UAV security [12], [35], [37], [42], [47], [48]. GPS spoofing injects false GPS signals that the GPS receiver mistakes as legitimate satellite signals. With spoofed GPS signals integrated, the control loop of UAV makes wrong perceptions of its position, resulting in deviations from the planned path. On the other hand, the IMU biasing attack involves injecting systematic errors into the IMU sensor through physical manipulations. This causes incorrect readings in accelerations or angular velocities, causing the flight controller to make incorrect adjustments of the UAV, resulting in erratic flight behaviors. These attacks severely undermine the operational integrity of UAVs, as evidenced by existing works, highlighting their practical implications and the urgent need for defense measures [10], [34], [40], [45].

Challenges. Despite existing research efforts focusing on monitoring, detection, and mitigation of specific UAV at-

tacks [10], [31], [40] — primarily through sensor fusion techniques using KF and identifying predictable patterns — these measures often fall short against novel or unidentified threats. The assumption that a specific sensor is compromised by a known type of attack limits the effectiveness of such defensive strategies. The fact that either IMU or GPS or both of the sensors can be compromised challenges us with a solution capable of conducting accurate post hoc RCA to pinpoint compromised sensors amidst unknown attacks. Unlike the existing works, our goal is not immediate intervention but post-failure analysis to detect attack, identify compromised sensor, and make decisions on whether to trust the specific sensor for future state estimations. We provide the design of SOUNDBOOST in § III-C.

D. Acoustic Side-Channel on UAVs

To conduct reliable RCA and unknown sensor attack detection, we need a reliable information source other than the existing sensors for UAVs. The acoustic side-channel signals that are physically emitted during flights provide a trustworthy reflection of the flight status which we deem reliable.

Thrust. UAV's actuation results in propeller motion that generates acoustic signals as a physical byproduct that *cannot* be eliminated. A rotating propeller creates a pressure difference between the front and back sides of its blades. As a result, the air accelerates in the backward direction, creating a reaction force in the opposite direction, known as thrust, which generates three types of noises.

1) **Blade passing noise**: A *low-frequency* sound that depends on the number of blades and propellers' rotational speed.
2) **Mechanical noise**: A *mid-frequency* sound from the propeller system's mechanical components and electrical noise from the Electronic Speed Controller (ESC). Motor speed affects electromagnetic forces, influencing vibrations and acoustics in this range.
3) **Aerodynamic noise**: A *higher-frequency* sound is caused by blade-air interactions. Blade design, shape, and angles of attack influence airflow patterns, turbulence, and vortices, shaping this noise.

Dynamics Captured through Acoustic Side-Channels. The utilization of acoustic signals as a diagnostic tool for capturing UAV dynamics stems from the analogous methods employed in cyber-physical additive manufacturing systems [7]. Specifically, UAV flight dynamics, including acceleration, deceleration, and turnings, manifest in distinctive acoustic patterns. For instance, during ascent, all four motors spin faster simultaneously, generating a sound pattern with higher amplitude and higher pitch. When the UAV turns, the motors at diagonally opposed positions exhibit varying speeds, contributing to a nuanced acoustic profile. Furthermore, each propeller emits a distinct sound, louder and higher-pitched during acceleration and quieter with a lower pitch during deceleration.

By placing a microphone array with four microphones at an *off-center* location onboard, each propeller can be located by employing the Time-Difference-of-Arrival (TDoA) technique. This method calculates the differences in the time it takes for the sound waves from each propeller to reach the microphones, allowing for triangulation of the position of each sound source [51]. The strategic off-center placement of the microphone array further enhances this differentiation, with microphones closer to a propeller receiving louder sounds, while those farther ones receive quieter sounds. This setup not only enables the distinct identification of each propeller based on its acoustic characteristics but also facilitates the inference of the UAV's acceleration dynamics. By mapping these acoustic patterns to known maneuvers, it enables inference of the UAV's acceleration in *three dimensions*. Thus, the real-time changes in the acoustic signals emitted during flight provide a reliable means for RCA after mission failures, detecting attacks against the UAV navigation system and identifying compromised sensors.

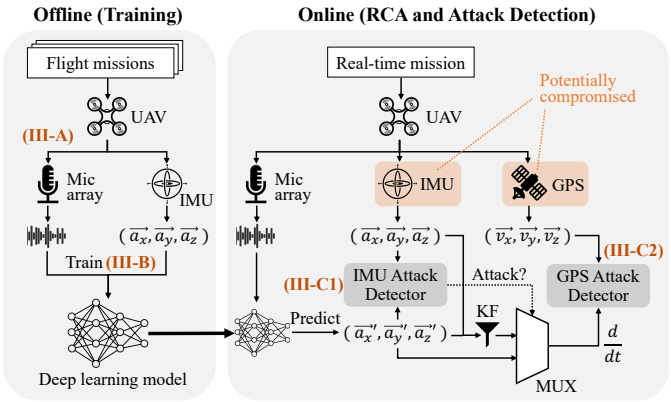


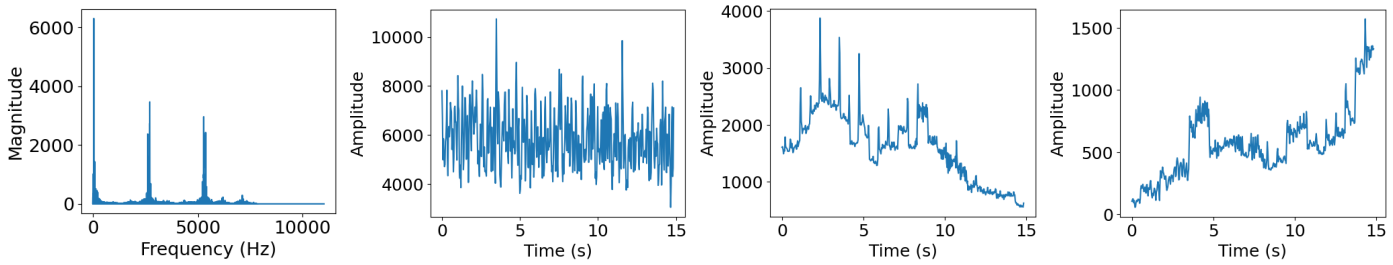
Fig. 1: Overview and workflow of SOUNDBOOST.

Challenges. Currently, there is no comprehensive sound dataset available for UAV flights that can be used to correlate the IMU measurements, nor is there a dataset for UAV sound under sensor attacks. Although there are existing works using the acoustic signals for UAV indoor location estimation [30], [51], finding the accurate correlation between the acoustic signals and IMU data for RCA and sensor attack detection in unseen real-world scenarios is still a challenge. It is also necessary to consider environmental variability, such as winds, which challenges us to augment the dataset for a more robust model across various environmental conditions. We provide our solution to this challenge in § III-B.

E. Threat Model & Assumptions

Assumptions. We assume the UAV is benign, and its hardware and firmware are free of attacks. Additionally, we assume the UAV only relies on GPS and IMU for navigation, as LiDAR or vision sensors are optional to most of the UAV models. Moreover, by default, during the flight, the UAV operates in low-sensitivity failsafe mode, aiming for an emergency landing in case of sensor failure.

Threat Model. We assume that the attacker possesses full knowledge of the current flight status of the UAV (e.g., flight path, location, etc). This attacker can launch sensor spoofing attacks to inject false signals into either the GPS [45] or the IMU sensors [55] in the navigation system from the ground for which the distance between the attacker and the UAV is limited by the flight altitude. The attacker can perform the attacks at any given time after the UAV finishes taking off and before it starts landing. The primary objective of the attacker



(a) Audio signal frequency domain (b) Hovering (stable amp) (c) Decelerating (shrinking amp) (d) Accelerating (rising amp)

Fig. 2: (a) shows the frequency distribution of the audio signal captured by a microphone. The signals are mainly concentrated around three frequencies; 200 Hz group (blade passing noise), 2500 Hz group (mechanical noise), and 5500 Hz group (aerodynamic noise). (b)-(d) shows the correlation between the amplitude (i.e., amp) of acoustic signals and measured acceleration - hovering (constant acceleration), decelerating (negative acceleration), and accelerating (positive acceleration).

is to control the UAV's positioning and influence its flights in a stealthy way rather than crashing the UAV. As for the detection sensor, we assume that the attacker can perform a reasonable sound spoofing attack using an off-the-shelf portable speaker for record-and-replay attacks up until around 100 dB, the max volume of a portable speaker [43]. Considering *practicality* and *stealthiness*, we do not consider cases in which the attacker is capable of using a directional speaker or a loudspeaker to spoof our microphone. Such methods are not cost-effective and would require the attacker to create precise phase-shifted audio signals to cancel out the audio signals from the moving drone and inject the loud signals into the detector onboard, which is practically intractable.

III. Design

Fig. 1 illustrates the overview of SoundBoost, including four components: (1) an acoustic signature generation stage where SoundBoost captures and extracts the acoustic signatures of the UAV (§III-A); (2) an offline learning stage where SoundBoost trains a DL model to map the acoustic signatures with the IMU measurements (§III-B); (3) an RCA IMU attack detection stage where acoustic signatures are fed into the model for sensory prediction, which is then compared with the IMU measurements for IMU attack detection (§III-C1); (4) an RCA GPS attack detection stage where SoundBoost learns the parameters of the UAV's physical invariants for physical state estimations and uses the sensory predictions from the DL model for GPS attack detection (§III-C2).

A. Acoustic Signature Generation

The acoustic signatures for the flight are generated online within a specific time window. SoundBoost performs a *Fast Fourier Transform* (FFT) analysis of the audio signal, converting a time-domain audio signal into the frequency domain. This enables SoundBoost to identify the presence and strength of different frequency components within audio signals. SoundBoost identifies specific frequency components related to noise sources. As shown in Fig. 2a, the signals concentrate mainly on three frequency domains. The signals around 200 Hz identify the blade passing noise, which we call blade passing frequency group. The signals around 2500 Hz identify the mechanical noise, which we refer as the mechanical frequency group. The signals in the higher frequency domain around 5500 Hz identify as aerodynamic noise, for which we call the aerodynamic frequency group.

SoundBoost filters out audio frequencies larger than the aerodynamic frequency groups (i.e., 6 kHz), such that it includes these characteristic frequency bands and is not impacted by IMU spoofing attacks. Since these attacks typically utilize frequencies above the audible range (above 20 kHz) [48], our deliberate focus on the audible spectrum ensures SoundBoost is innately immune to such interference. This inherent filtering capability enhances the security of the microphone array sensor and protects it against ultrasonic manipulations aimed at disrupting IMU sensors.

Acoustic Signatures v.s. Acceleration. Fig. 2 shows the amplitude pattern of the aerodynamic frequency groups during different actuation processes, demonstrating the correlation between acoustic signatures and UAV acceleration. When the UAV is hovering (Fig. 2b), the amplitude of the acoustic signals remains almost unchanged. Fig. 2c shows that when the UAV is decelerating, the amplitude of the acoustic signals decreases over time. This phenomenon is intuitive in that decelerating requires less motor power, which makes less noise. Fig. 2d demonstrates the opposite effect. Leveraging such a correlation, SoundBoost can map the acoustic signals with the acceleration vectors in x, y, z directions

B. Acoustic Signature & Sensory Mapping

To map the generated acoustic signature to the real-world IMU measurements, SoundBoost utilizes a DL model to find their correlation. To achieve this goal, SoundBoost needs to determine the best time window of collected data so that the correlation is not over-fitting or under-fitting. Moreover, the DL model has to be robust to the environmental impact, specifically the winds. To achieve this goal, SoundBoost leverages data augmentation. The modeling phase is conducted *offline* using the acoustic signatures as the features to learn the corresponding raw IMU measurements over a certain time window. The model captures patterns and *correlations* between acoustic signatures and IMU labels, recognizing relationships in frequency components of aerodynamic and mechanical noise. This mapping enables the

model to generalize and predict UAV postures and behaviors, providing insights into flight missions by analyzing audio data.

DL Model Selection. We employed three DL models popularly used for processing audio data: *ResNet101*, *MobileNetV2*, and a *Neural Ordinary Differential Equations* (ODE) model. We use a sliding window to align the acoustic signals and onboard IMU readings and use the IMU readings as the ground truth. Then, we train each model with *mean squared error loss (MSE)* as the loss function, ensuring that the predicted sensory from the audio signature would be closely aligned with IMU measurements. This training process also ensures that the measurement noise of the onboard IMU is minimized through the DL model.

Time Window. The choice of the time window for acoustic signature is critical for capturing the acoustic signals in relation to the IMU measurements. Since the IMU measurements are sampled at a fixed rate, we have to select a time window at least as large as the IMU sampling window. The time window should be large enough to capture the entire acceleration process and short enough to capture the rapid changes in different maneuvers. We choose the time window empirically through experiments for our UAV model, which enables us to have a precise correlation between the audio and motion data of the UAV and for further RCA and post hoc attack detection.

Data Augmentation. To build a robust model, we need to consider the impact of winds. Collecting data under different wind conditions is challenging. Therefore, we choose data augmentation to enrich the training dataset. Varying wind speeds cause the UAV to frequently change its actuation to maintain the acceleration or velocity setpoint, resulting in fluctuations in the acoustic patterns (§ III-A). When encountering winds, due to the nature of the PID controller, UAVs may drastically change the actuation initially to quickly counteract the wind's effect and vary their actuation time to maintain the acceleration setpoint and smooth the flight. As illustrated in Fig. 3, with tailwinds, the UAV's propellers spin slower, generating softer sounds, and it requires less time to reach a desired setpoint (e.g., a desired velocity). With headwinds, the UAV's motors have to rotate faster, generating louder sounds, to fight against a continuous adverse wind, for which the UAV has to actuate for a longer period.

With this theory, we use "Time Shift" (Fig. 3) to augment the dataset, exposing the model to a broader variety of temporal scenarios within the data. Specifically, we augment the data by varying the time window: a larger window simulates headwinds, and a smaller window simulates tailwinds. For example, the entire sequence of data in Fig. 2d represents the effort of fighting against various speeds of headwinds, a larger window with fixed number of data points of the sound amplitude can capture the entire process. Similarly, the effect of various speeds of tailwinds can be captured using a smaller window. Since the shortened actuation process under tailwinds is covered by the focused window, we expect limited performance improvements by augmenting with small time windows. We focus on augmenting the dataset with various expanded time windows to better expose the whole actuation process under any headwind scenarios to the model.

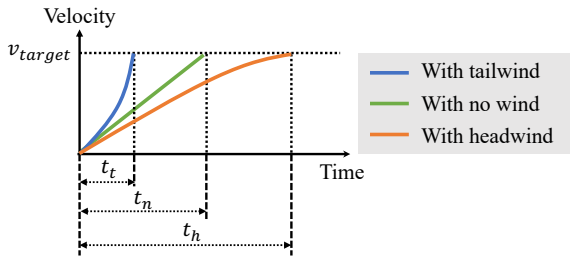


Fig. 3: Time shift augmentation. Wind conditions affect the time required to reach a target velocity (v_{target}). With no wind, this time is t_n, while tailwinds reduce it (t_t) and headwinds extend it (t_h). To account for these variations, we augmented the dataset by using sampling windows of different lengths.

C. Root Cause Analysis Framework

The primary objective of root cause analysis is to identify the origin of the problem or the fault within the navigation system after mission failure. By leveraging the acoustic side-channel, SOUNDBOOST generates detailed audio signatures that enable post hoc RCA and detection of attacks on the drone. SOUNDBOOST conducts two layers of RCA to pinpoint the compromised sensor during a sensor attack. To pinpoint the compromised sensor, SOUNDBOOST first identifies the discrepancies in the IMU measurements and corresponding sound patterns, then the GPS measurements with the motion estimation output (described in § III-C2).

1) IMU Attack Detection: SOUNDBOOST leverages audio acceleration predictions from its ML models to detect IMU biasing attacks. During attacks, where the IMU readings are manipulated by the attacker, the unaffected acoustic signatures act as a reliable benchmark, revealing any disparities in the IMU measurements. Under normal conditions, IMU measurements should closely match audio acceleration predictions. In contrast, when a spoofing attack targets the UAV's IMU, the attacker provides false measurements. While the acoustic signatures, being a physical byproduct of the UAV's operation, are not subject to the same spoofing, the audio acceleration prediction, therefore, remains accurate. Consequently, discrepancies between the IMU readings and the audio acceleration predictions will be observed, indicating a potential attack.

Threshold. Attack detection is achieved through out-of-distribution detection. SOUNDBOOST begins with analyzing the distribution of the differences between the acceleration measurements and the audio acceleration predictions in benign settings. We observe that the distribution of the residuals closely approximates a normal distribution in benign settings, as demonstrated in the evaluation (§ IV-A). Leveraging this insight, SOUNDBOOST analyzes whether an attack is occurring within the time window. This post hoc analysis is performed as part of RCA, for each time window, SOUNDBOOST calculates the residuals between the predicted audio accelerations and the actual IMU measurements. To determine the presence of an attack, we subject these residuals to a Kolmogorov-Smirnov Test against the normal distribution observed under benign conditions. If the residuals within the detection window deviate significantly from this normal distribution, SOUNDBOOST flags the occurrence of a potential IMU biasing attack.

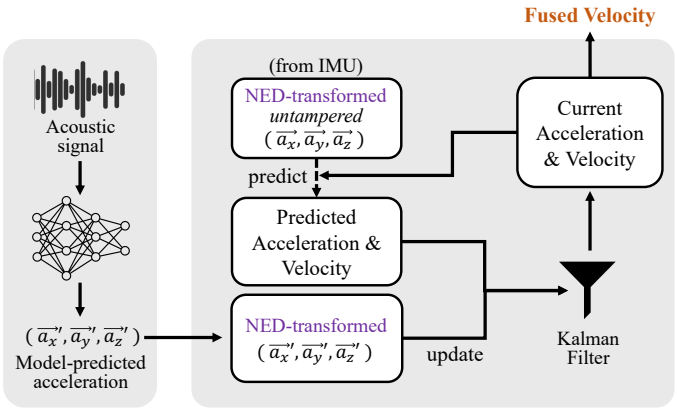


Fig. 4: Our Customized Kalman Filter. The KF framework incorporates the acceleration predictions from the audio and the IMU measurements for control analysis.

2) GPS Attack Detection: For GPS attack detection, SOUNDBOOST first conducts control analysis for motion estimations of the UAV using the audio acceleration prediction as well as the IMU measurements if determined to be trustworthy. To estimate the UAV motions, SOUNDBOOST adopts two versions of Kalman Filters (KF) for accurate sensor fusion for control analysis in the pipeline. Then, SOUNDBOOST uses the state estimation output from the control analysis step for GPS attack detection.

Version 1: Audio Only KF (with compromised IMU). In the case when the IMU is deemed not trustworthy, SOUNDBOOST focuses exclusively on the audio acceleration predictions. The KF receives audio-derived acceleration as the primary input. This acceleration prediction, obtained through audio signal analysis, serves as the only basis for velocity estimation.

In the KF prediction step, SOUNDBOOST utilize the *North-East-Down* transformed audio acceleration predictions to forecast the velocity of the UAV. By employing the first kinematic formula in the state prediction step:

$$v_1 = v_0 + a \cdot t$$

For prediction:

$$\hat{v}_{k|k-1} = F_k \hat{v}_{k-1|k-1} + B_k a_k$$
$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

where the Predict State and Predict State Covariance are calculated. $\hat{v}_{k|k-1}$ predicts velocity and acceleration state estimate at time k given information until time $k-1$; $\hat{v}_{k-1|k-1}$ is the updated velocity and acceleration state estimate at time $k-1$ given measurement at time $k-1$; F_k is the state transition model for velocity; B_k is the control input model for acceleration; and a_k is the control vector for acceleration; $P_{k|k-1}$ predicts state covariance matrix for velocity and acceleration at time k given information until time $k-1$; Q_k is the process noise covariance. SOUNDBOOST predicts the velocity at the next time step based on the updated current velocity. This prediction provides an initial estimate of the UAV's velocity state and the state covariance matrix for the subsequent fusion process. In the KF update step, SOUNDBOOST leverage the North-East-Down transformed *audio acceleration prediction* to refine the previous velocity estimate.

For the update, SOUNDBOOST calculates the Kalman Gain:

$$K_k = P_{k|k-1} H_k^T \left(H_k P_{k|k-1} H_k^T + R_k \right)^{-1}$$

and updates state estimate and state covariance with audio acceleration predictions:

$$\hat{v}_{k|k} = \hat{v}_{k|k-1} + K_k \left(z_k - H_k \hat{v}_{k|k-1} \right)$$
$$P_{k|k} = \left(I - K_k H_k \right) P_{k|k-1}$$

where $\hat{v}_{k|k}$ updates velocity and acceleration state estimate at time k given measurement at time k; H_k is the observation model for audio prediction; R_k is the observation noise covariance; K_k is the Kalman Gain; z_k is the audio acceleration prediction at time k; $P_{k|k}$ updates the state covariance matrix for velocity at time k; and I is the identity matrix.

Version 2: Audio + IMU KF (with benign IMU). For the case when the IMU is benign and can be used for analysis, SOUNDBOOST utilizes a customized KF design for a more reliable velocity estimation using both the IMU measurements and the audio predictions. The overview of this framework is shown in Fig. 4. For our customized design, SOUNDBOOST uses the IMU measurements in the prediction step and the audio prediction in the update step for weighted measurements. We customized the KF prediction step to utilize the *North-East-Down* transformed IMU acceleration to forecast the velocity of the UAV.

Predict State and Predict State Covariance are given by:

$$\hat{v}_{k|k-1} = F_k \dot{z}_{k-1} + B_k a_k$$
$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

where $\hat{v}_{k|k-1}$ predicts velocity and acceleration state estimate at time k given information until time $k-1$; $\dot{z}_{k-1}$ is the IMU measured velocity and acceleration state estimate at time $k-1$; F_k is the state transition model for velocity; B_k is the control input model for acceleration; and a_k is the control vector for acceleration; $P_{k|k-1}$ predicts state covariance matrix for velocity and acceleration at time k given information until time $k-1$; Q_k is the process noise covariance.

SOUNDBOOST predicts the velocity at the next time step based on the current velocity and measured IMU acceleration. This prediction provides an initial estimate of the UAV's velocity state and the state covariance matrix for the subsequent fusion process. For the update step, SOUNDBOOST uses the same algorithm as in the audio-only KF, with z_k being the audio acceleration prediction at time k for the update. By incorporating the velocity calculated from the audio acceleration prediction, SOUNDBOOST updates the previous estimation with a weighted combination of the two velocity sources. The weights assigned to the IMU and audio acceleration predictions reflect their respective reliabilities and are updated dynamically, enabling the KF to fully utilize the information from both sources.

The fusion of the IMU and audio acceleration predictions in the KF allows SOUNDBOOST to capitalize on the strengths of each sensor while compensating for their individual limitations for the post hoc RCA process. The IMU provides high-frequency measurements that capture rapid changes in the UAV's velocity, while the audio acceleration prediction offers additional insights into the UAV's movement based on

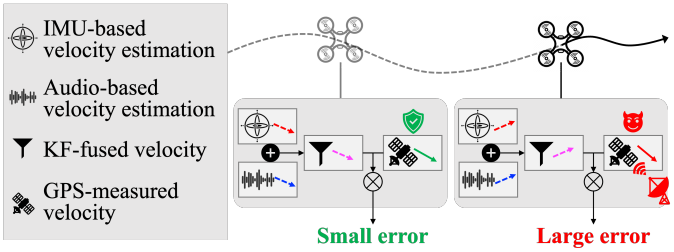


Fig. 5: GPS Spoofing Detection. Through prediction and state estimation, our framework assigns weights to different data sources and estimates the velocity for attack detection.

acoustic cues. By combining these sources through the KF, SoundBoost obtains a more robust and accurate estimation of the UAV's velocity. This fused velocity information serves as a crucial input for the overall control analysis, enabling more precise velocity estimations for post hoc attack detections.

Detection. In the RCA attack detection step, SoundBoost uses the fusion result from its KF framework to identify potential attacks against the UAV. Fig. 5 illustrates the overview of the attack detection procedure. SoundBoost first measures the error between each observation of the fused location and the GPS measurements in benign cases without GPS spoofing. This error measurement serves as a reference to gauge the expected behavior of the system. Then, SoundBoost repeats the error measurement process on datasets where attacks are known to be present. SoundBoost accumulates the difference between the GPS velocity and the velocity reference and monitors the running mean of the error. When analyzing the testing data, SoundBoost deems a potential attack to be present if the running mean of the error exceeds a predefined threshold, which is the maximum running mean error of the benign cases after removing outliers. This comparison allows SoundBoost to identify continuous deviations from the expected behavior, which signals the potential occurrence of an attack. By continuously monitoring the velocity deviations and the cumulative error, SoundBoost can detect abnormal patterns deviating from expected behavior. These patterns indicate external disturbances or attacks on the vehicle, allowing for further timely intervention and appropriate countermeasures.

D. Design Choices and Onboard Implementation

Our design explicitly targets the constraints of typical UAV platforms by prioritizing low computational overhead and ease of integration. We use the PX4 autopilot [33] to showcase the practical efficacy of our framework, benefiting from PX4's open-source nature and extensive user base. We use a Raspberry Pi, functioning as a companion computer, to supplement the PX4 autopilot, providing the computational capacity for SoundBoost. We use MAVSDK [32], a companion software development kit, to directly access, control, and communicate with the UAV during flight using the MAVLink protocol. For sound collection, we use the Seed Studio ReSpeaker USB Mic Array [50] that connects to the Raspberry Pi. This microphone array features an array of four digital microphones strategically positioned *off-center* on the UAV's frame, a location that permits sound localization from distinct propellers. Data from the UAV and the microphone array are gathered and synchronized on the Raspberry Pi. All the RCA steps of SoundBoost can be easily completed by the Raspberry Pi. Unlike many onboard systems that rely on powerful edge computing modules such as the NVIDIA Jetson or cloud processing due to real-time demands, our design emphasizes lightweight, fully self-contained post hoc RCA capabilities. The choice reflects a practical trade-off between inference latency, power consumption, and platform compatibility, ensuring that SoundBoost can operate fully onboard without compromising flight stability or significantly affecting mission duration. Furthermore, by co-locating acoustic sensing and telemetry processing, our design achieves tight synchronization and avoids bandwidth bottlenecks that arise sharing data elsewhere for processing. This design ensures reproducibility and shows that post hoc RCA is feasible without specialized hardware, enabling deployment on resource-constrained UAV systems.

IV. Evaluation

In this section, we evaluate the overall effectiveness of SoundBoost as a post hoc RCA framework. While SoundBoost performs anomaly detection as part of RCA pipeline, SoundBoost only triggers *after* a mission failure observed. We demonstrate the effectiveness and robustness of SoundBoost by evaluating its ability to detect IMU and GPS spoofing attacks as part of RCA. Specifically, we show:

- The acoustic signals obtained from the UAV's motors are effective in predicting its acceleration and are *not* affected by sound spoofing attacks (§ IV-A).
- SoundBoost can accurately detect IMU biasing attacks based on trained acoustic signatures of UAV (§ IV-B).
- SoundBoost can accurately pinpoint GPS spoofing attacks even when UAV is under IMU biasing attack (§ IV-C).
- SoundBoost is robust to sound spoofing attacks (§ IV-D).

A. Efficacy of Acoustic Side-Channel

Setup and Data Collection. We used a Holybro X500 drone with a Raspberry Pi as the companion computer for the experiments. We utilized a microphone array, strategically positioned *off-axis* on top of the body of the UAV to capture the acoustic signals generated from the four motors while conducting different flight maneuvers. To train our model, we collected 36 benign flights. These flights covered a wide range of maneuvers, including hovering, ascent, descent, forward flight, and turns across 6 carefully designed extended navigation scenarios leading to diverse acoustic patterns. For testing, we conducted 21 additional flights (10 with IMU attack, 11 with GPS spoofing attack) that are unseen in training flights to assess the generalization capability of SoundBoost. Specifically, for GPS spoofing evaluations, testing flights involved different flight lengths, speeds, and mission trajectories compared to the training set, ensuring robustness against trajectory variations. To ensure that the data represents real-world scenarios, all flights were conducted outdoors in *diverse* environments, including calm and windy conditions with various ambient noise levels. This dataset is designed to ensure that SoundBoost is evaluated on a wide range of flight scenarios, demonstrating its applicability to unseen trajectories and real-world conditions.

TABLE I: Data Augmentation Choice. By experimenting with different window size augmentations, augmenting the chosen window (0.5 seconds) five times performs the best.

Augmentation Window Size	Train MSE	Validation MSE	Test MSE
w/ 0.5x	0.2322	0.3924	0.3371
No Aug.	0.2215	0.3667	0.3325
w/ 1x	0.1478	0.4694	0.5468
w/ 2x	0.0571	0.4463	0.3963
w/ 3x	0.1343	0.4924	0.5244
w/ 5x	**0.1059**	**0.3450**	**0.3366**

Acoustic Signature Window and Data Augmentation. Window size is an important parameter that impacts the mapping of the acoustic signatures to the IMU measurements. It defines the span of audio data used to predict IMU readings and significantly impacts the model's accuracy. To figure out the ideal window size for the MobileNetV2 model we used, we experimented with window sizes spanning from 0.1 to 2 seconds. We found out that smaller window sizes enable the model to perceive and learn intricate patterns in the acoustic signatures, thus improving prediction accuracy for IMU measurements. However, as the window size increases beyond 0.5, this granularity is lost, causing the Mean Squared Error (MSE) to rise, leading to decreased performance. The optimal window size was 0.5 seconds, balancing detailed pattern recognition with sufficient context for robust mapping.

To further improve the model's resilience against environmental factors like wind, time-shift data augmentation is used (§ III-B). This approach aimed to capture subtle variations of wind directions by augmenting the training dataset. We experimented with augmentation window sizes ranging from 0.5x to 5x, the optimal window. As shown in Tab. I, when trained with 5x augmentation of the selected window size (0.5 seconds), SOUNDBOOST achieves the best performance. Notably, the model is intentionally trained with slight overfitting due to the physical constraints governing UAV behavior in real-world conditions. A highly generalized model could potentially violate these constraints, which is undesirable in practical applications. Despite the lower training MSE, the testing MSE remains lower than the validation MSE, indicating that the model is not degrading on truly unseen data.

Model Performance. With the chosen window size (0.5 seconds) and augmented data (5x) to mitigate environmental impact such as wind, the blue histogram in Fig. 6 illustrates the strong correlation between the predicted $\vec{a_z}'$ and the reading $\vec{a_z}$ from the IMU sensor. Specifically, the means of the differences between the predicted acceleration values and the one read from the IMU sensor in the three axes are close to 0, and the standard deviation is relatively low as well, showing that the model correlates with the IMU measurements well, with a very low bias and small modeling error.

Frequency Importance. To understand the contribution of different frequency components to our model's predictions, we conducted a counterfactual feature importance analysis [57]. Specifically, we analyzed the impact of removing key frequency groups—including aerodynamic frequencies, blade passing frequencies, mechanical frequencies, and other noise frequencies—and measured how this affected the model's MSE. Removing aerodynamic frequencies increases the MSE to 1.2698 (3.77x), indicating that aerodynamic noise carries the strongest signal relevant to the model's ability to predict UAV accelerations. The increases after removing blade passing and mechanical frequency groups are relatively small (less than 0.12x), indicating that, while these frequencies contribute to the model, they are not the dominant features for predictions. Removing all other frequencies, which represent general environmental and ambient noise, results in only a minor change in MSE (less than 0.05x), suggesting that background noise has a negligible impact on the model performance.

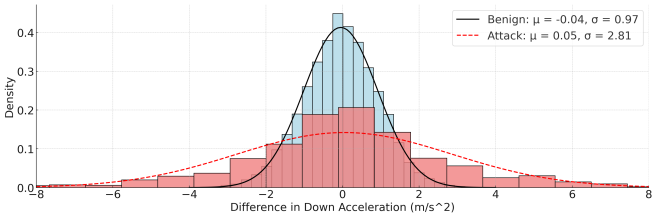


Fig. 6: IMU Attack Detection. The acceleration residuals between predicted acceleration and real readings from IMU along with the fitted curves for benign (blue and normal curve) and attack (red and dash curve) cases. Attack shows a different distribution with a larger standard deviation than benign.

B. IMU Attack Detection

IMU Biasing Attack Setup. For an IMU biasing attack, given the constraints of our hardware-attacking device, we opted to synthesize the attacks by attacking the UAV's firmware. Through this approach, we simulated two primary types of attacks: the accelerometer Denial-of-Service (DoS) attack and the gyroscope Side-Swing attack [55].

The premise of Side-Swing attacks relies on amplifying the sensor output in a designated target direction. We simulate a side swing attack where spoofing of the gyroscope is proved to be controllable to some extent [55]. To recreate this in our synthesized setup, we injected a sequence of positive-biased signals into the sensor output. This attack attempts to skew the UAV's interpretation of its orientation and trajectory. Our synthesized attack is conducted empirically, incrementally adding small biases for a short time period enough to observe obvious abnormal flight behaviors but not causing a crash. The accelerometer DoS attack is simulated by injecting a sequence of random noise, as control of the accelerometer cannot be achieved [55]. However, the oscillatory nature of the injected signal implies that it contributes almost equivalently to both directions. Although we still observe obvious abnormal behaviors, the disruption of this attack by nature is not capable of causing mission failure, which aligns with our threat model. We conducted 10 hovering flights under IMU biasing attacks, with 5 flights each for side-swing and DoS attacks. Each spoofing event lasted 10 seconds while the drone hovered. Overall, we conducted 20 different flights, 10 normal flights, and 10 subjected to the synthesized attacks.

Result. In this section, we show RCA detection results of IMU biasing attacks to demonstrate the effectiveness of SOUNDBOOST in pinpointing the anomalies in the IMU. We

utilize the audio acceleration predictions obtained from the previous step of sensory mapping and compare them directly with the acceleration measurements from the IMU. Our aim is to detect significant discrepancies between the two sets of data that may indicate the presence of a spoofing attack. Among the three models we tested, MobileNetV2 has the best performance in general to fulfill the requirements.

In benign cases, we typically observe that the acceleration predictions closely align with the IMU measurements, exhibiting only minor discrepancies that roughly follow a normal distribution. This underscores the accuracy of our model in normal operating conditions, instilling confidence in its reliability for subsequent control analysis. In IMU biasing attacks from §III-C1, instability in the UAV's flight is induced by tampering with the z-axis (downward). We found that the model predictions and the sensor measurements deviated significantly during the attack periods, enabling SOUNDBOOST to easily reject the normal hypothesis. Fig. 6 shows the distribution of differences between the IMU measured downward accelerations and the acoustic predicted downward accelerations for an attack dataset along with the normal curve from a benign case. We also fit a normal curve for the attack distribution of residuals. As can be seen, the distribution of the residuals from the attack dataset is significantly different from the benign distribution, with a large standard deviation of 2.81, showing the unstable flight behavior from the attack.

SOUNDBOOST is able to identify *all* the attacks during poc hoc analysis with an average delay of 2.3 seconds from the attack onset as observed in the flight log, with *one* benign flight showing a false positive. We attribute it to the unstable flight caused by critically low battery levels, highlighting the robustness of our detection approach. This result highlights the and reliability of identifying IMU biasing attacks, contributing to the overall effectiveness of SOUNDBOOST.

C. GPS Spoofing Attack Detection

GPS Spoofing Attack Setup. For a GPS spoofing attack, we use an open-sourced tool called GPS-SDR-SIM [54] and a software-defined radio device (SDR) called HackRF One [16]. We use GPS-SDR-SIM to generate GPS baseband signals by providing it with the ephemeris data, i.e., the longitude and latitude of the spoofing location, with the spoofing length. Then, we use HackRF One to convert the generated baseband signal into radio frequencies that mimic the GPS satellites.

In the experiments of the GPS spoofing attack, we programmed the UAV to perform two missions. For the first mission, we had the UAV hover at a certain location. For the GPS spoofing attacks, we set the spoofing location of the attack to be static throughout the attack period. For the first case, we spoof at a location 10 meters from the UAV; and for the second case, we spoof the UAV at a location within the mission path. Overall, we evaluate the accuracy of SOUNDBOOST's GPS attack detection component using 30 benign and 19 attack flight periods. Across the 19 attack periods, each GPS spoofing period lasts between 60 and 90 seconds.

Failsafe IMU Only Baseline. To establish comprehensive baseline comparisons, we conduct an ablation study of SOUNDBOOST considering using only the IMU measurements in KF, the same algorithm inherently adopted by ArduPilot for failsafe motion estimations [2]. In this baseline model, we solely rely on the IMU measurements for physics prediction in KF, similar to the *audio only* KF of SOUNDBOOST. By leveraging the UAV's onboard IMU, which provides information about its orientation, acceleration, and angular velocity, we construct a state-space model for the KF. The IMU acceleration is used in the prediction step to estimate the UAV's velocity based on the first kinematic equation. The fusion algorithm then updates this velocity estimation with only the IMU measurements and further performs RCA attack detection using the same algorithm as the *Audio Only KF* (§ III-C1).

TABLE II: GPS spoofing attack detection result. The table summarizes the system inputs, the number of benign and attack flights tested, the number of alerts reported, and the TPR and FPR for each detector.

	SOUNDBOOST		Baselines				
System Inputs	**audio only**	**audio & IMU**	**Failsafe [33] IMU only**	**LTI [10] yaw**	**LTI [10] vx**	**LTI [10] vy**	**DNN [15] (LSTM)**
# Benign Flights	30	30	30	30	30	30	30
# Alerted	7	2	5	3	0	1	22
# Attack Flights	19	19	19	19	19	19	19
# Alerted	15	17	11	5	1	1	13
TPR	**0.79**	**0.89**	0.58	0.26	0.05	0.05	0.68
FPR	0.23	**0.10**	0.17	0.10	0	0.03	0.73

Control Invariant (LTI) Baseline. We include the state-of-the-art *Control Invariant* method [10] as our baseline. This method employs System Identification (SI) to construct a linear time-invariant (LTI) model of the UAV's kinematic system from observed data. The LTI system is represented as:

$$x_{k+1} = Ax_k + Bu_k \text{ and } y_k = Cx_k.$$

A, B, and C are matrices that characterize the specific UAV process. The established LTI model serves as an invariant monitor for attack detection based on a set threshold. The UAV's gyroscope and velocity data are used in our implementation.

Deep Neural Network (DNN) Baseline. We also include a data-driven baseline that uses the LSTM model to approximate the UAV's control dynamics [15]. This approach learns the UAV's normal control behavior as a time series from benign flight data and estimates the next control outputs. Any deviation between the predicted and actual control outputs beyond a learned threshold is flagged as an attack.

Result. Tab. II demonstrates the effectiveness of SOUNDBOOST in detecting spoofing attacks, showing the comparisons of various input configurations in SOUNDBOOST against the baseline systems. For each, it lists the number of benign and attacked flight periods and the number of alerts raised during the flight period and calculates the true positive rate (TPR) and false positive rate (FPR).

When the IMU sensor is flagged as anomalous, SOUNDBOOST relies solely on audio acceleration predictions for GPS spoofing detection. It successfully identified 15 out of the 19 attacks, achieving a true positive rate of 79%. For the

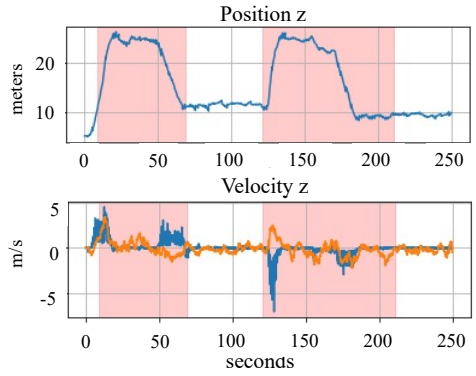


Fig. 7: Z-axis Position and Velocity Estimation. SOUNDBOOST position estimation is shown on top. Velocity estimations from the GPS (blue) and from SOUNDBOOST (orange) are shown at the bottom. The GPS spoofing period is highlighted in pink.

30 benign flight periods, SOUNDBOOST *audio only* detector reports 7 false alarms, resulting in a false positive rate of 29%. While this false positive may sound high, with IMU being flagged introduces irregular in flight behaviors, the method still maintains a good true positive rate in identifying GPS spoofing attacks. This highlights the value of using acoustic signatures for GPS spoofing attack RCA and detection while suggesting that the *audio only* approach is prone to false positives, which emphasizes the need to incorporate both acoustic signature predictions and IMU measurements to improve post hoc detection accuracy if IMU data is available for use.

When the IMU sensor is confirmed benign, SOUNDBOOST combines acceleration predictions with IMU measurements, using a weighted approach of sensor fusion based on the covariance matrix to improve the precision of velocity estimates. *audio + IMU* outperforms other methods in detection rates, identifying attacks with higher accuracy and maintaining a reasonably low false alarm rate. Specifically, of the 19 attacks, with *audio + IMU*, SOUNDBOOST successfully detects 17 of them within the attack period, achieving a true positive rate of 89%. Fig. 7 illustrates the estimated position and velocity from SOUNDBOOST during a GPS spoofing attack, where the drone is instructed to execute a hovering mission. Large discrepancies observed in velocity estimations indicate the effectiveness of this method. While *audio + IMU* reports 2 false alarms out of the 30 benign periods tested, resulting in a false positive rate of 8%, we attribute these false positives to the UAV's recovery attempts from the previous GPS spoofing attack period. When we label the dataset, we cut the flight into periods of attack and benign cases, but in reality a benign case usually follows an attack case. Thus, these false positives are possible to be minimized with deliberate thresholding.

The runtime overhead of SOUNDBOOST signature generation is on average of 2.4%. While offline training and parameter tuning are computationally intensive, they only need to be performed once for each UAV model. SOUNDBOOST achieves efficient post hoc detection, identifying GPS spoofing attacks in an average of 18.1 seconds after attack starts – timely alerts with high accuracy and low false-positive rates.

Comparisons with Baseline. To underscore the effectiveness of our GPS spoofing detection methodology, we evaluate our framework using only acoustic signatures (*audio*) with and without IMU measurements against the ablation study baseline using only IMU data (*IMU*). We also benchmarked against LTI control invariant methods that rely on x-axis velocity, y-axis velocity, and yaw measurements and a DNN-based method that leverages time series for control signal estimations.

As shown in Tab. II, the baseline approach using *IMU only* for detection indicates lower detection rates for attacks compared to both approaches in SOUNDBOOST. This approach shows fewer false positives compared to the *audio only* approach, with an FPR 0.17 along with a TPR of 0.58, showing that the detector using only IMU is less sensitive compared to both of the SOUNDBOOST's detectors. The control invariant approaches (*yaw*, *vx*, *vy*) exhibit varying levels of success. *Yaw-based* control invariants outperform vx and vy in detection, with both vx- and vy-based detectors showing low alert rates. This indicates that the control invariant approaches, especially the vx- and vy-based detectors, are not very applicable to real-world attack detection cases. The DNN-based approach shows a TPR of 0.68, which is higher than the *IMU only* and control invariant methods, indicating better sensitivity to attacks. However, it also has a significantly higher FPR of 0.73 on benign flights, suggesting a lack of specificity that is not suitable for practical RCA deployment.

This analysis validates the performance of SOUNDBOOST's methods using *audio only* and *audio+IMU*, highlighting the effectiveness of the acoustic side-channel in real-world scenarios. The evaluation emphasizes the role of SOUNDBOOST as a post-incident RCA and attack detection framework, providing an accurate diagnosis to attribute UAV navigation failures.

D. Adversarial Experiments Against Sound

To evaluate SOUNDBOOST's robustness in adversarial settings, we leverage two key insights: 1) SOUNDBOOST filters out acoustic signals beyond aerodynamic frequency group, making inaudible signals ineffective. 2) Acoustic signal intensity decreases with distance [4], requiring interference to occur in close proximity. To this end, we conducted two adversarial experiments: 1) a real-world interference attack using a second UAV and an off-the-shelf speaker, and 2) a simulated attack where an adversary manipulates sound phase and amplitude in an idealized setting.

Real-world Interference. To execute a real-world sound spoofing attack, the attacker has to interfere with the sound signals received by a target UAV. We conducted two experiments to achieve this. First, we use a second spoofing UAV of the same model to fly at a radius of 2, 1.5, 1, and 0.5 meters around the hovering UAV, attempting to interfere with SOUNDBOOST using its own sound. Secondly, we perform record-and-replay attacks using a portable speaker mounted on the attacking UAV playing pre-recorded UAV sounds at maximum volume while maintaining a minimal distance (i.e., 0.5 meters) avoiding collisions. The results indicate that neither attack have any measurable effect on SOUNDBOOST 's acceleration predictions. FFT analysis revealed that aerodynamic frequency alterations were negligible. We attribute the failure to the spoofing sound not phase-synchronized with the UAV's sound signal. To better understand the attacks, we further analyzed the FFT of the audio recorded by a fixed microphone array

TABLE III: SOUNDBOOST *audio + IMU* under adversarial phase-synchronized sound attack at the aerodynamic frequencies. The baseline TPR and FPR are 0.89 and 0.1 (Tab. I).

		Channels							
		1		2		3		4	
Attacks	**Amplitude**	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR
Canceling	0%	0.74	0.41	0.74	0.47	0.76	0.56	0.70	0.57
	25%	0.79	0.40	0.76	0.46	0.76	0.48	0.76	0.57
	50%	0.84	0.38	0.81	0.40	0.78	0.47	0.79	0.43
	75%	0.84	0.28	0.84	0.38	0.84	0.38	0.79	0.40
Amplifying	125%	0.79	0.13	0.64	0.06	0.59	0.06	0.53	0.07
	150%	0.62	0.08	0.52	0.04	0.45	0.06	0.42	0.07
	175%	0.59	0.06	0.47	0.04	0.39	0.07	0.37	0.07
	200%	0.55	0.04	0.42	0.04	0.38	0.07	0.37	0.07

positioned 0.5 meters from a hovering UAV and compared it to recordings from the microphone array mounted on the UAV. The result shows that the maximum magnitude of the aerodynamic frequency group is approximately 12000 near the UAV, but drops to 5500 at the microphone array 0.5 meters away, representing only 46% of the original intensity.

Simulated Interference. Since real-world tools cannot affect SOUNDBOOST, we simulate interference assuming attackers have powerful wave emitters targeting individual microphones in the worst case. This tests the robustness of sound-based detection. We start by manipulating the amplitudes of the most important frequencies § IV-A, assuming an adversary with precise control over phase and amplitude using low-latency equipment. Specifically, we amplify (125%-200%) or cancel (75%-0%) aerodynamic frequencies on one to four microphone channels. Tab. III shows SOUNDBOOST *audio + IMU* detection results (averaged by channel combinations). Amplifying by a factor of 2 reduces the baseline TPR from 0.89 (Tab. I) to 0.37 on four channels and 0.55 on one. Fully cancellation only reduces the TPR to 0.70 on four channels and 0.74 on one. This is likely because cancellation resembles weaker signals of benign variations, limiting how much the model can be misled, whereas amplification introduces abnormal signals that may violate UAV's physical constraints. Amplification reduces both TPR and FPR by skewing error distribution, making downstream detection more susceptible. Cancellation increases FPR while keeping TPR high, making detection overly sensitive but not easily bypassed.

Despite these simulated effects, real-world execution of this attack is highly impractical. First, precise phase synchronization across moving microphones requires advanced audio processing and precise control over the propagation of sound waves. Any phase drift weakens the attack. Second, reflections, air turbulence, and sound absorption by surrounding objects further disrupt alignments at all microphones. Third, even if an attacker succeeds in producing phase-synchronized sounds to attack all microphones whether using a second drone or a powerful, portable speaker, our first adversarial experiment shows that without crashing the target, signal strength at the target only reaches 46% due to noise diffusion. Thus, effective amplitudes are limited between 54% (cancellation) and 146% (amplification). At 54% amplitude across all channels, SOUNDBOOST *audio + IMU* achieves a TPR of 0.79 and an FPR of 0.43, comparable to SOUNDBOOST *audio only* benign performance. At 146%, SOUNDBOOST *audio + IMU* performs slightly worse than Failsafe *IMU only* benign performance with a TPR of 0.47, but a better FPR of 0.07.

Our results show that real-world sound spoofing from a drone fails to impact SOUNDBOOST due to noise diffusion and lack of phase alignment. Simulated phase-synchronized attacks can degrade performance but are nearly impossible in practice. Even in worst-case simulations, SOUNDBOOST remains robust through increased false positives allowing minimal bypasses. These results confirm SOUNDBOOST 's reliability against both practical and theoretical sound spoofing threats.

V. DISCUSSION

A. Robustness to Concurrent GPS and IMU Spoofing Attacks

First, we highlight that there is no reliable way to launch simultaneous GPS and IMU spoofing attacks without crashing the UAV. Although existing research primarily focuses on single-sensor attacks, concurrent attacks on both sensors would severely disrupt the UAV's stability, leading to unpredictable UAV behavior or crashes. This is supported by the "Rocking Drone" attack [48], which demonstrated a significant impact of real-world IMU biasing attack on UAV stability. Thus, launching simultaneous spoofing attacks on both GPS and IMU without crashing the UAV is highly challenging, if not unfeasible. Second, even if concurrent attacks occur and UAV remains airborne, SOUNDBOOST remains functional. Its IMU RCA attack detection approach remains unchanged, and GPS spoofing can still be identified using SOUNDBOOST *audio-only* (outlined in §III-C2). This demonstrates that SOUNDBOOST is capable of performing post-incident RCA and detecting attacks even under complex, concurrent sensor attack scenarios.

B. Generality of SOUNDBOOST

Unseen Maps or Trajectories. SOUNDBOOST's model is trained on a diverse set of flight missions incorporating various speeds and maneuvers to ensure its generality to unseen trajectories or maps. By correlating acoustic signatures with acceleration vectors, SOUNDBOOST RCA capability remains independent of specific mission plans. This ensures the performance of SOUNDBOOST in completely novel environments, providing reliable RCA across diverse fields.

Ambient Environment. The robustness of SOUNDBOOST extends to varying ambient environmental conditions. The model is trained and tested under different scenarios, incorporating frequency filters for out-of-range noise and data augmentations for various wind conditions. SOUNDBOOST is tested to be robust against sound spoofing attacks from the same UAV model. This ensures that SOUNDBOOST is not only adaptable but also resilient to environmental factors.

UAV Models. SOUNDBOOST's methodology is designed to be adaptable to different UAV configurations, including variations in UAV frames, motors, propellers, and microphone setups. While the framework is designed to be generalized across different hardware, it is necessary to retrain the machine learning model to accommodate the specific acoustic and vibrational characteristics of each new hardware setup. This retraining process is crucial for maintaining the accuracy of the spoofing detection mechanism, ensuring that SOUNDBOOST remains effective across different hardware setups.

Actuator and Multiple IMUs Attacks. Beyond GPS spoofing and IMU biasing attacks, SOUNDBOOST generalizes to other threats as well by learning sound patterns during actuation. It can detect actuator denial-of-service attacks exploiting block waveforms on PWM-controlled actuators [13], as the model predicts near-zero acceleration when actuators stop. However, SOUNDBOOST fails if an attack fully controls all actuators – though this is unlikely in a quadcopter, where opposing motor pairs prevent uniform waveform attacks. Additionally, SOUNDBOOST can detect attack on UAV with multiple IMUs as well, but detection thresholds must be learned separately for different IMU models.

C. Ethical Considerations

All experiments were conducted in a carefully chosen nonurban location to ensure safety and compliance. The GPS spoofing device (HackRF One) operated at a minimal transmission power of 5 dBm, limiting its range to a small, controlled radius of 29.2 m. Locations were selected to be isolated, located over 15 times the transmission radius away from critical human activity areas. Additionally, all flights were FAA-authorized and conducted within controlled airspace at or below 400 ft, adhering to legal, safety, and ethical guidelines.

VI. RELATED WORK

The widespread use of UAVs has raised significant security concerns spanning cyber and physical dimensions. Cyber attacks can compromise communication, control systems, and data integrity, resulting in unauthorized access or information leaks [20], [22], [23], [27], [35]. Physical attacks disrupt operations through interference, sensor manipulation, or hardware tampering [9], [13], [28]. Studies highlight the feasibility and consequences of such attacks [10], [34], [40], [45], emphasizing the need for robust defenses.

UAV Sensor Spoofing Attacks. Sensor spoofing, especially GPS spoofing and IMU biasing attacks, poses critical risks to the UAVs [12], [37], [48]. GPS spoofing exploits vulnerabilities in unauthenticated signals[1], enabling UAV seamless takeovers [38] and real-time control in a *Human-in-the-Loop* manner [45]. IMU biasing attacks use out-of-band signals like high-frequency sound to interfere with or manipulate gyroscopes and accelerometers, injecting false data [48], [53], [55]. Other attacks exploit physical feedback [56] or electromagnetic channels [17], further highlighting sensor vulnerabilities.

Sensor Spoofing Detection. Various mechanisms have been developed aim to detect sensor spoofing. Control-invariant frameworks build physical models to monitor deviations [10], and have been extended with non-linear invariants for greater robustness [40]. Data-driven models use time-series patterns [11], [15] or safety-critical variables [39] to flag control anomalies. Hardware-based defenses include spoofing-resistant GPS receivers that employ advanced signal processing [42] or *data fusion* [44], [47] and an Angle-of-Arrival method using satellite constellations [29] to validate GPS signal legitimacy. For IMU biasing attacks, a recovery technique based on denoising has been proposed, establishing the first testbed for such attacks and providing a method to mitigate signal injection impacts [19]. Despite these advances, most defenses rely on specific attack assumptions. They often lack resilience to *unknown* attacks or the ability to identify compromised sensors. Additionally, GPS spoofing defenses often lack validation in real world. And few methods conduct post-incident RCA after UAV has experienced a mission failure. By addressing these limitations, our framework demonstrates improved effectiveness in both scenarios.

Cyber-Physical Defenses via Acoustic Side-Channel. The use of acoustic side-channels for defense has been explored in prior research. The acoustic side-channel is used for UAV model identification [26], authentication [14], [41], [59] and fault detection and isolation [6], [21]. These studies demonstrate the acoustic side-channel's effectiveness in monitoring UAV operational status. In broader CPS domains, researchers have reverse-engineered 3D printer outputs using microphone and IMU recordings [1], [5], [49], showing the side-channel's diagnostic value. Building on these foundational studies, our work uses the acoustic side-channel for post-incidence RCA and sensor attack detections to secure UAV operations. SOUNDBOOST 's novelty lies in leveraging acoustic side-channel to enable post hoc diagnosis, offering forensic insight into failures even after real-time defenses may have failed.

VII. CONCLUSION

We present SOUNDBOOST, a novel framework that leverages the acoustic side-channel of UAVs to conduct root cause analysis and detect sensor spoofing attacks targeting UAV navigation systems. We demonstrate that the acoustic side-channel has a strong correlation with the acceleration vectors of the UAV and is robust for post hoc sensor spoofing attack RCA. Unlike previous works, we evaluate SOUNDBOOST with real-world flights. We collect the first audio dataset of the UAV under real-world GPS spoofing. These extensive real-world experiments demonstrate that SOUNDBOOST achieved a 100% attack detection rate for IMU biasing attacks, which is the first work to detect IMU sensor attacks, and 79% GPS spoofing attack detection rate using acoustic signatures only and 89% with additional trusted IMU measurement, outperforming the state-of-the-art baselines by over 21%.

VIII. ACKNOWLEDGMENTS

We thank the anonymous reviewers for their constructive feedback. We thank Professor Saman Zonouz for his insightful comments on earlier drafts of this paper. This material is supported in part by the National Science Foundation (NSF) under grant No. CNS-2229876, Office of Naval Research (ONR) under grant No. N00014-17-1-2179, Defense Advanced Research Projects Agency (DARPA) under contract N66001-21-C-4024; and by the Korea government (MSIT) Institute of Information & communications Technology Planning & Evaluation (IITP) grant: No.RS-2019-II191906 (Artificial Intelligence Graduate School Program at POSTECH), RS-2024-00457882 (National AI Research Lab Project), and RS-2024-00509258 (Global AI Frontier Lab). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF, ONR, DARPA, or MSIT.

[1]It is well-known that GPS signal is *unauthenticated*.

REFERENCES

[1] Mohammad Abdullah Al Faruque, Sujit Rokka Chhetri, Arquimedes Canedo, and Jiang Wan. Acoustic side-channel attacks on additive manufacturing systems. In *2016 ACM/IEEE 7th international conference on Cyber-Physical Systems (ICCPS)*, pages 1–10. IEEE, 2016.

[2] ArduPilot Development Team. Ekf failsafe - copter documentation.

[3] Muhammet Fatih Aslan, Akif Durdu, Kadir Sabanci, Ewa Ropelewska, and Seyfettin Sinan Gültekin. A comprehensive survey of the recent studies with uav for precision agriculture in open fields and greenhouses. *Applied Sciences*, 12(3):1047, 2022.

[4] Henry E Bass, Louis C Sutherland, Allen J Zuckerwar, David T Blackstock, and DM Hester. Atmospheric absorption of sound: Further developments. 1995.

[5] Christian Bayens, Tuan Le, Luis Garcia, Raheem Beyah, Mehdi Javanmard, and Saman Zonouz. See no evil, hear no evil, feel no evil, print no evil? malicious fill patterns detection in additive manufacturing. In *26th USENIX Security Symposium (USENIX Security 17)*, pages 1181–1198, 2017.

[6] Adam Bondyra, Marek Kołodziejczak, Radosław Kulikowski, and Wojciech Giernacki. An acoustic fault detection and isolation system for multirotor uav. *Energies*, 15(11), 2022.

[7] Sujit Rokka Chhetri, Arquimedes Canedo, and Mohammad Abdullah Al Faruque. Kcad: Kinetic cyber-attack detection method for cyber-physical additive manufacturing systems. In *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–8, 2016.

[8] Wen-Chyuan Chiang, Yuyu Li, Jennifer Shang, and Timothy L Urban. Impact of drone delivery on sustainability and cost: Realizing the uav potential through vehicle routing optimization. *Applied energy*, 242:1164–1175, 2019.

[9] Hongjun Choi, Sayali Kate, Yousra Aafer, Xiangyu Zhang, and Dongyan Xu. Cyber-physical inconsistency vulnerability identification for safety checks in robotic vehicles. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, CCS '20, page 263–278, New York, NY, USA, 2020. Association for Computing Machinery.

[10] Hongjun Choi, Wen-Chuan Lee, Yousra Aafer, Fan Fei, Zhan Tu, Xiangyu Zhang, Dongyan Xu, and Xinyan Deng. Detecting attacks against robotic vehicles: A control invariant approach. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, CCS '18, page 801–816, New York, NY, USA, 2018. Association for Computing Machinery.

[11] Pritam Dash, Guanpeng Li, Zitao Chen, Mehdi Karimibiuki, and Karthik Pattabiraman. Pid-piper: Recovering robotic vehicles from physical attacks. In *2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 26–38. IEEE, 2021.

[12] Drew Davidson, Hao Wu, Rob Jellinek, Vikas Singh, and Thomas Ristenpart. Controlling UAVs with sensor input spoofing attacks. In *10th USENIX Workshop on Offensive Technologies (WOOT 16)*, Austin, TX, August 2016. USENIX Association.

[13] Gökçen Yılmaz Dayanıklı, Sourav Sinha, Devaprakash Muniraj, Ryan M Gerdes, Mazen Farhood, and Mani Mina. {Physical-Layer} attacks against pulse width {Modulation-Controlled} actuators. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 953–970, 2022.

[14] Yufeng Diao, Yichi Zhang, Guodong Zhao, and Mohamed Khamis. Drone authentication via acoustic fingerprint. In *Proceedings of the 38th Annual Computer Security Applications Conference*, pages 658–668, 2022.

[15] Aolin Ding, Praveen Murthy, Luis Garcia, Pengfei Sun, Matthew Chan, and Saman Zonouz. Mini-me, you complete me! data-driven drone security via dnn-based approximate computing. In *Proceedings of the 24th International Symposium on Research in Attacks, Intrusions and Defenses*, pages 428–441, 2021.

[16] Great Scott Gadgets. Hackrf one, 2023. Hardware device.

[17] Joon-Ha Jang, Mangi Cho, Jaehoon Kim, Dongkwan Kim, and Yongdae Kim. Paralyzing drones via emi signal injection on sensory communication channels. In *NDSS*, 2023.

[18] Kai Jansen, Matthias Schäfer, Daniel Moser, Vincent Lenders, Christina Pöpper, and Jens Schmitt. Crowd-gps-sec: Leveraging crowdsourcing to detect and localize gps spoofing attacks. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 1018–1031. IEEE, 2018.

[19] Jinseob Jeong, Dongkwan Kim, Joon-Ha Jang, Juhwan Noh, Changhun Song, and Yongdae Kim. Un-rocking drones: Foundations of acoustic injection attacks and recovery thereof. In *NDSS*, 2023.

[20] Chijung Jung, Ali Ahad, Yuseok Jeon, and Yonghwi Kwon. Swarmflawfinder: Discovering and exploiting logic flaws of swarm algorithms. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 1808–1825, 2022.

[21] Sai Srinadhu Katta, Kide Vuojärvi, Sivaprasad Nandyala, Ulla-Maria Kovalainen, and Lauren Baddeley. Real-world on-board uav audio data set for propeller anomalies. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 146–150, 2022.

[22] Hyungsub Kim, Muslum Ozgur Ozmen, Antonio Bianchi, Z Berkay Celik, and Dongyan Xu. Pgfuzz: Policy-guided fuzzing for robotic vehicles. In *NDSS*, 2021.

[23] Hyungsub Kim, Muslum Ozgur Ozmen, Z. Berkay Celik, Antonio Bianchi, and Dongyan Xu. Pgpatch: Policy-guided logic bug patching for robotic vehicles. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 1826–1844, 2022.

[24] Si Jung Kim, Yunhwan Jeong, Sujin Park, Kihyun Ryu, and Gyuhwan Oh. A survey of drone use for entertainment and avr (augmented and virtual reality). *Augmented Reality and Virtual Reality: Empowering Human, Place and Business*, pages 339–352, 2018.

[25] Taegyu Kim, Aolin Ding, Sriharsha Etigowni, Pengfei Sun, Jizhou Chen, Luis Garcia, Saman Zonouz, Dongyan Xu, and Dave (Jing) Tian. Reverse engineering and retrofitting robotic aerial vehicle control firmware using dispatch. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, MobiSys '22, page 69–83, New York, NY, USA, 2022. Association for Computing Machinery.

[26] Harini Kolamunna, Thilini Dahanayaka, Junye Li, Suranga Seneviratne, Kanchana Thilakaratne, Albert Y Zomaya, and Aruna Seneviratne. Droneprint: Acoustic signatures for open-set drone detection and identification with online data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(1):1–31, 2021.

[27] Peng-Yong Kong. A survey of cyberattack countermeasures for unmanned aerial vehicles. *IEEE Access*, 9:148244–148263, 2021.

[28] Depeng Li. Prevent malicious controller in a physical, human and cyber triad. In *23rd conference on USENIX Security Symposium (USENIX Security'14)*, pages 1–2, 2014.

[29] Shinan Liu, Xiang Cheng, Hanchao Yang, Yuanchao Shu, Xiaoran Weng, Ping Guo, Kexiong Curtis Zeng, Gang Wang, and Yaling Yang. Stars can tell: a robust method to defend against {GPS} spoofing attacks using off-the-shelf chipset. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 3935–3952, 2021.

[30] Wenguang Mao, Zaiwei Zhang, Lili Qiu, Jian He, Yuchen Cui, and Sangki Yun. Indoor follow me drone. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '17, page 345–358, New York, NY, USA, 2017. Association for Computing Machinery.

[31] Ruben Mascaro, Lucas Teixeira, Timo Hinzmann, Roland Siegwart, and Margarita Chli. Gomsf: Graph-optimization based multi-sensor fusion for robust uav pose estimation. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1421–1428, 2018.

[32] MAVSDK Development Team. Mavsdk, 2023. Software Development Kit.

[33] Lorenz Meier, Daniel Agar, and Beat etc. Küng. Px4-autopilot, 2022.

[34] Sashank Narain, Aanjhan Ranganathan, and Guevara Noubir. Security of gps/ins based on-road location tracking systems. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 587–601, 2019.

[35] B. Nassi, R. Bitton, R. Masuoka, A. Shabtai, and Y. Elovici. Sok: Security and privacy in the age of commercial drones. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 1434–1451, 2021.

[36] Francesco Nex and Fabio Remondino. Uav for 3d mapping applications: a review. *Applied geomatics*, 6:1–15, 2014.

[37] Tyler Nighswander et al. Gps software attacks. In *Proceedings of the 2012 ACM Conference on Computer Communication Security (CCS'12)*, pages 450–461, New York, NY, USA, 2012. ACM.

[38] Juhwan Noh, Yujin Kwon, Yunmok Son, Hocheol Shin, Dohyun Kim, Jaeyeong Choi, and Yongdae Kim. Tractor beam: Safe-hijacking of consumer drones with adaptive gps spoofing. *ACM Trans. Priv. Secur.*, 22(2), apr 2019.

[39] Sangbin Park, Youngjoon Kim, and Dong Hoon Lee. Scvmon: Data-oriented attack recovery for rvs based on safety-critical variable monitoring. In *Proceedings of the 26th International Symposium on Research in Attacks, Intrusions and Defenses*, pages 547–563, 2023.

[40] Raul Quinonez, Jairo Giraldo, Luis Salazar, Erick Bauman, Alvaro Cardenas, and Zhiqiang Lin. SAVIOR: Securing autonomous vehicles with robust physical invariants. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 895–912. USENIX Association, August 2020.

[41] Soundarya Ramesh, Thomas Pathier, and Jun Han. Sounduav: Towards delivery drone authentication via acoustic noise fingerprinting. In *Proceedings of the 5th Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, pages 27–32, 2019.

[42] Aanjhan Ranganathan, Hildur Ólafsdóttir, and Srdjan Capkun. Spree: A spoofing resistant gps receiver. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, MobiCom '16, page 348–360, New York, NY, USA, 2016. Association for Computing Machinery.

[43] RTINGS.com. The 6 loudest bluetooth speakers - summer 2024: Reviews, 2024. Accessed: 2024-07.

[44] Harshad Sathaye, Gerald LaMountain, Pau Closas, and Aanjhan Ranganathan. Semperfi: Anti-spoofing gps receiver for uavs. In *Network and Distributed Systems Security (NDSS) Symposium 2022*, 2022.

[45] Harshad Sathaye, Martin Strohmeier, Vincent Lenders, and Aanjhan Ranganathan. An experimental study of GPS spoofing and takeover attacks on UAVs. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 3503–3520, Boston, MA, August 2022. USENIX Association.

[46] Eduard Semsch, Michal Jakob, Dušan Pavlicek, and Michal Pechoucek. Autonomous uav surveillance in complex urban environments. In *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, volume 2, pages 82–85. IEEE, 2009.

[47] Junjie Shen, Jun Yeon Won, Zeyuan Chen, and Qi Alfred Chen. Drift with devil: Security of multi-sensor fusion based localization in high-level autonomous driving under gps spoofing (extended version), 2020.

[48] Inc. Sonalysts, Yongdae Kim, Hojoon Lee, Jiho Kim, Junyoung Kim, Yonghwi Lee, Juyong Jang, Jihoon Yun, and Yongdae Kim. Rocking drones with intentional sound noise on gyroscopic sensors. In *24th USENIX Security Symposium (USENIX Security 15)*, pages 881–896, 2015.

[49] Chen Song, Feng Lin, Zhongjie Ba, Kui Ren, Chi Zhou, and Wenyao Xu. My smartphone knows what you print: Exploring smartphone-based side-channel attacks against 3d printers. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 895–907, 2016.

[50] Seeed Studio. Respeaker usb mic array, 2023. Hardware device.

[51] Yimiao Sun, Weiguo Wang, Luca Mottola, Ruijin Wang, and Yuan He. Aim: Acoustic inertial measurement for indoor drone localization and tracking. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, SenSys '22, page 476–488, New York, NY, USA, 2023. Association for Computing Machinery.

[52] Nils Ole Tippenhauer, Christina Pöpper, Kasper Bonne Rasmussen, and Srdjan Capkun. On the requirements for successful gps spoofing attacks. In *Proceedings of the 18th ACM conference on Computer and communications security*, pages 75–86, 2011.

[53] Timothy Trippel, Ofir Weisse, Wenyuan Xu, Peter Honeyman, and Kevin Fu. Walnut: Waging doubt on the integrity of mems accelerometers with acoustic injection attacks. In *2017 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 3–18, 2017.

[54] Osamu Tsuboi. GPS-SDR-SIM: GPS Satellite Signal Simulator, 2023. GitHub repository.

[55] Yazhou Tu, Zhiqiang Lin, Insup Lee, and Xiali Hei. Injected and delivered: Fabricating implicit control over actuation systems by spoofing inertial sensors. In *27th USENIX Security Symposium (USENIX Security 18)*, pages 1545–1562, Baltimore, MD, August 2018. USENIX Association.

[56] Yazhou Tu, Sara Rampazzi, and Xiali Hei. Towards adversarial process control on inertial sensor systems with physical feedback side channels. In *Proceedings of the 5th Workshop on CPS&IoT Security and Privacy*, pages 39–51, 2023.

[57] Sandra Wachter, Brent Mittelstadt, and Chris Russell. Counterfactual explanations without opening the black box: Automated decisions and the gdpr. *Harv. JL & Tech.*, 31:841, 2017.

[58] Wikipedia. Iran–u.s. rq-170 incident - wikipedia. https://en.wikipedia.org/wiki/Iran%E2%80%93U.S._RQ-170_incident, December 2011. (Accessed on 04/22/2024).

[59] Chuxiong Wu and Qiang Zeng. Turning noises to fingerprint-free "credentials": Secure and usable drone authentication. *IEEE Transactions on Mobile Computing*, page 1–14, 2024.