ESP Spoofing: Covert Acoustic Attack on MEMS Gyroscopes in Vehicles

Zhen Hong, *Member, IEEE*, Xiong Li, Zhenyu Wen, *Member, IEEE*, Leiqiang Zhou, Huan Chen, Jie Su

***Abstract*—Electronic Stability Program (ESP) is widely used in modern vehicles. Its safety and stability largely depend on the strength and reliability of the MEMS gyroscope. However, the tight coupling between this sensor and the environment brings significant safety hazards to the vehicle. In this study, we describe the physical vulnerability of gyroscopes to high-frequency acoustics and introduce methods for finding resonant frequencies. We devised two methods to inject the attack signal into audio files to make the acoustic attack more stealthy. The realized attack is non-intrusive and does not require tampering with the ESP hardware device, making attack detection more difficult. We also consider a neural network-based defense strategy and verify its effectiveness. The construction of the vehicle simulation system and the above experiments are completed in the co-simulation environment of Carsim and Simulink.**

***Index Terms*—MEMS gyroscope, Resonant frequencies, Acoustic attack, Non-intrusive, Neural Network, Carsim, Simulink**

I. Introduction

Allied Market Research [1] reported that the global autonomous vehicle market is growing significantly with 39.4 percent annual growth from 2019 to 2026, and will reach 556.67 billion by 2026. The safety of these autonomous or semi-autonomous cars dramatically relies on the deployed sensors to collect environmental information and make reactions based on the collected data. For example, if the wheel speed sensors report that the wheel rotating is significantly slower than the vehicle's speed, the Anti-Lock Braking Systems (ABS) will reduce the force on the wheel to turn them faster to avoid wheel lock. As a result, if the wheel speed sensors are attacked and manipulated by hackers, it may cause serious problems.

Many works studied sensor-based physical attacks on vehicle systems. Roosta [2] divided them into two types: invasive and non-invasive attacks. In invasive attacks, the components of the system are physically tampered such as changing the circuitry and wiring. On the contrary, non-invasive attacks leverage the vulnerabilities of the sensors in a vehicle and make the sensors fail to infer the physical environment. Compared to invasive attacks, non-invasive attacks are more challenging to be detected because monitored physical environments are tough to be verified [3]. Shoukry et al. exploited a non-invasive vulnerability in [4] to attack an ABS and demonstrated that the proposed attack can lead to severe security issues.

In invasive attacks, it is already known that malicious acoustic interference can affect the output of software-trusted sensors in various real systems[5]. Yunmok Son et al.[6] studied the resonant frequency of MEMS gyroscopes and used high-frequency noise to incapacitate drones equipped with MEMS gyroscopes. After that, Timothy Trippel et al.[7] further investigated how high-frequency noise could be used to achieve complete adversarial control of sensor output for MEMS accelerometers. For this, they verified it in the toy remote control car. However, it can be found that the attack object systems in the above work are not complicated, and the threat to human beings from the attacks is limited. In addition, in consumer-grade speakers, the audible component of high-frequency noise poses a challenge to the concealment of attacks. For some complex systems in which humans intervene, the realization of attacks is not easy. On the basis of their work, we investigate non-intrusive vulnerabilities in onboard electronic stability program (ESP) with MEMS gyroscope as a key sensor and propose a new non-intrusive attack.

Specifically, we inject the high-frequency noise into an ordinary sound wave to attack the MEMS gyroscope to paralyze ESP. Our simulation experiments based on actual sensors show that the attack can cause serious consequences, such as vehicle drift and rollover. In addition, we play the role of defender and discuss how to defend against such attacks effectively.

Our contributions can be summarised as follows:

- We propose and design a non-intrusive sound wave attack to ESP system and use audio overlay technology to improve the diversity of our attack.
- We design and build a closed-loop control vehicle simulation system based on fuzzy Proportion Integration Differentiation (PID) controller, which combines the hardware and simulation tools, to verify the effectiveness of the attack.
- We formulate an active defense strategy against the above attacks, and design experiments to evaluate its robustness.

II. Preliminaries

In this section, we briefly introduce the ESP and its critical sensors and draw out the possible risks in the system. The main parameters used in this paper are listed in Table I.

*This work was supported by National Natural Science Foundation of China under Grant 62072408, Zhejiang Provincial Natural Science Foundation of China under Grant LY20F020030, and New Century 151 Talent Project of Zhejiang Province. (*Corresponding author: Jie Su.*)

Z. Hong, X. Li, Z. Wen and L. Zhou are with the Institute of Cyberspace Security, and College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China (e-mail: zhong1983@zjut.edu.cn; lx.3958@gmail.com; zhenyuwen@zjut.edu.cn; leiqiangzhou@gmail.com).

H. Chen was with the Faculty of Mechanical Engineering & Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China (e-mail: ch950228@163.com).

J. Su is with the Open Lab, School of Computing, Newcastle University, Newcastle Upon Tyne, NE4 5AX, UK (e-mail: jieamsu@gmail.com).

TABLE I: Main parameters

Parameter	Sign	Unit
Automobile Stability Factor	K	/
Distance from Centroid to Front Axle	a	m
Distance from Centroid to Rear Axle	b	m
Front Wheel Cornering Stiffness	k_1	N/rad
Real Wheel Cornering Stiffness	k_2	N/rad
Vehicle Quality	m	kg
Ground Adhesion Coefficient	μ	/
Left Front Wheel Braking Torque	T_{bfl}	$N \cdot m$
Left Rear Wheel Braking Torque	T_{brl}	$N \cdot m$
Right Front Wheel Braking Torque	T_{bfr}	$N \cdot m$
Right Rear Wheel Braking Torque	T_{brr}	$N \cdot m$
Left Front Wheel Vertical Load	F_{zfl}	N/mm^2
Left Rear Wheel Vertical Load	F_{zrl}	N/mm^2
Right Front Wheel Vertical Load	F_{zfr}	N/mm^2
Right Rear Wheel Vertical Load	F_{zrr}	N/mm^2
Compensation for Yawing Moment	ΔM	$N \cdot m$
Front Wheel Steering Angle	δ_f	deg
Yaw Rate	ω	deg/s
Ideal Value of Yaw Rate	ω_d	deg/s
Centroid Slip Angle	β	deg
Ideal Value of Centroid Slip Angle	β_d	deg
Centroid Longitudinal Velocity	v_x	m/s
Centroid Lateral Velocity	v_y	m/s

A. Electronic Stability Program

ESP is a computerized module that utilizes high sensitive sensors to detect the loss of traction of a vehicle system. If a loss of traction is detected (e.g., driving on a slippery road), it automatically helps the driver to steer the car in the right direction. Fig. 1 illustrates the workflow of the ESP. First, the *driver's intent information* can be predicted by the *steering angle*. Then, the *basic vehicle status information* is monitored by the *horizontal/vertical acceleration detection* module and *yaw velocity detection* module. The *sub-stabilizer control* module analyzes the driver's intent information and the actual vehicle status information to decide whether the current status can achieve the driver's requirements. If not, the *traction control* module will request to increase or decrease the output of engine torque.

There are two common scenarios, *understeer* and *oversteer* which may cause severe results without the support of ESP. Fig. 2(a) shows the case of *understeer* that a car steers less than the driver requested. To overcome this, ESP triggers an additional amount of horizontal pendulum counterclockwise torque to pull the vehicle back to the expected direction. Similarly, when a car is steered too much, an additional yawing moment clockwise is requested by ESP to correct the path back to normal. The torque compensation strategy for one-sided wheels is given in Table II, where δ_f denotes the front wheel steering angle of the car, and ΔM denotes the compensation torque.

B. MEMS gyroscope

The MEMS gyroscope [8] is one of the most critical components of ESP. It measures the angular velocity of rigid

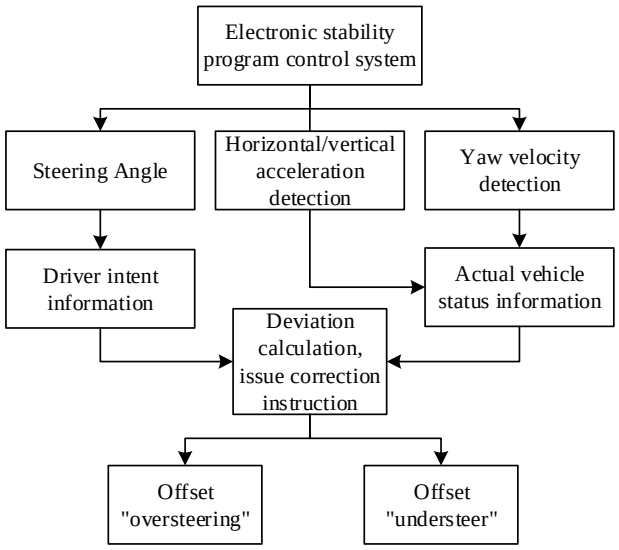
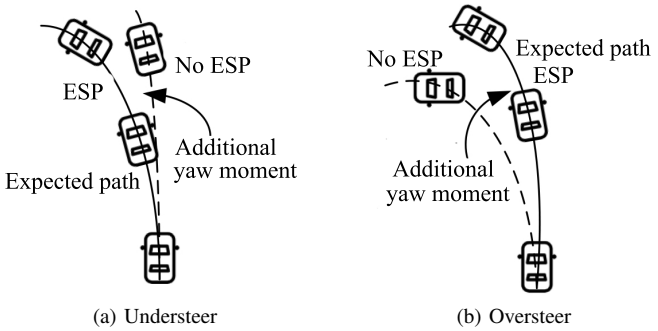


Fig. 1: The workflow of the ESP



(a) Understeer (b) Oversteer

Fig. 2: ESP braking force application strategy

body rotation. In other words, it measures the rotating speed of the Z-axis while the car moves, as shown in Fig. 3.

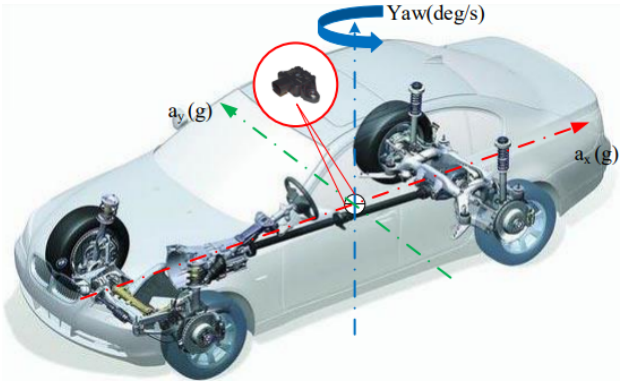


Fig. 3: Yaw Angle sensor in an automobile

The MEMS gyroscope follows the law of physics known as the Coriolis effect[9], which describes the deflection of a moving object in a rotating reference frame. The yaw rate w of the car can be computed by

$$w = -\frac{a_y}{2v_x} \tag{1}$$

where a_y denotes the acceleration in the Y-axis direction generated by the Coriolis effect. v_x denotes the velocity in the X-axis direction, which is measured by the mass continuously

TABLE II: Strategy of compensation for yawing moment

Front wheel angle	Steering feature	Yaw moment	Brake side wheel
$\delta_f > 0$	Oversteer	$\Delta M > 0$	Right wheel
	Understeer	$\Delta M < 0$	Left wheel
$\delta_f < 0$	Understeer	$\Delta M > 0$	Right wheel
	Oversteer	$\Delta M < 0$	Left wheel
$\delta_f = 0$	Excessive left turn	$\Delta M > 0$	Right wheel
	Insufficient left turn	$\Delta M < 0$	Left wheel
Any value	Stablize	$\Delta M = 0$	None

vibrating at a specific frequency concerning the X-axis (see Fig. 4).

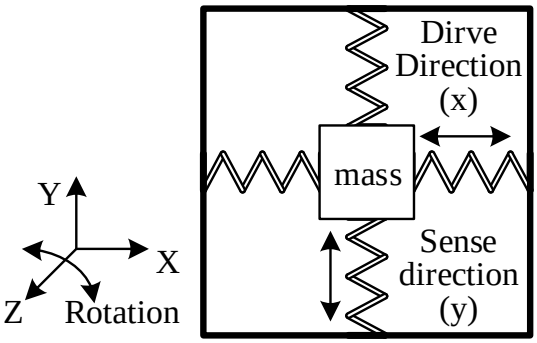


Fig. 4: Gyroscope structure

C. The impact of acoustic noise

Many works [6, 8, 10] have reported that harsh acoustic noise can degrade the accuracy of MEMS gyroscopes. [6] found that some MEMS gyroscopes generate ghost outputs when the attacker injects sound noise to cause frequency resonance. Moreover, the authors in [11] theoretically modeled the effect of acoustic noise for MEMS gyroscopes, and the model shows the false angular velocity reading has a positive correlation with displacement emanating from the ultrasonic excitation.

III. Attack Design

In this section, we discuss and determine the best resonance frequency of MEMS gyroscopes, and then use this feature to discover the vulnerabilities in ESP. Based on building the attack model, we further design the attack music and propose our simulation framework.

A. Determination of MEMS resonance frequency

To determine the resonance frequency, in this paper, we choose commonly used 5 gyroscope chips for testing, including MPU9250, MPU6050, MPU6500, L3G4200D, and L3GD20. Fig. 5 shows the entire experimental design framework for determining the resonance frequency, including a function signal generator, a wide-band power amplifier, a full-range speaker, and a personal computer (PC).

The malicious high-frequency signal is generated by the function signal generator, and the amplifier amplifies the signal to drive the speaker. Then, the sound wave is applied to the gyroscope chip. We connect the STM32[12] chip and the gyroscope chip through the integrated circuit bus IIC[13]. The STM32 chip can convert the abnormal hexadecimal number generated by the gyroscope into a decimal number. Finally, anomalous data is passed into the PC via the USB cable to attack the vehicle ESP system.

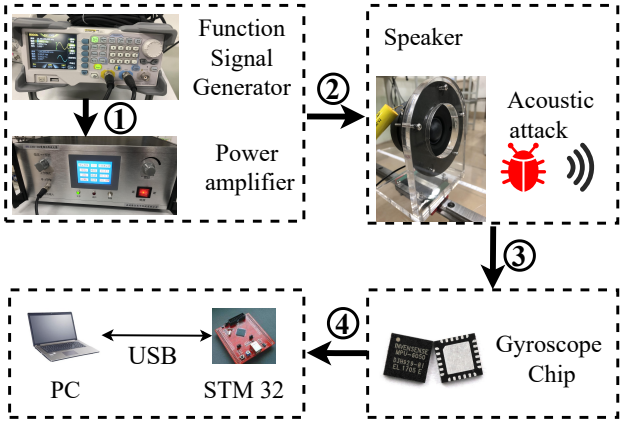


Fig. 5: Experimental framework for determining the resonance frequency. ① Sound wave signal generation. ② The sound wave signal is transmitted to the speaker. ③ Acoustic attack. ④ STM32 reads yaw rate data

Since the sound wave is a pressure wave and it exists in the medium (air or water), the gyroscope is set 10cm in front of the speaker. We control the frequency of the speaker from 100Hz to 34400Hz and collect 5000 samples at each frequency from the target gyroscope. Scanning the sound frequency range can be probed to determine the resonant frequency. Table III summarizes the resonant frequency of each gyroscope chip determined in the experiment.

TABLE III: The resonant frequency of the gyroscope chip in the experiment

Sensor model	Resonant frequency	
	Theoretical value	Test value
MPU9250	27±2 KHz	26.48~26.51KHz
MPU6050	27±3 KHz	26.90~27.30KHz
MPU6500	27±2 KHz	26.50~27.90KHz
L3G4200D	Null	28~8.13KHz
L3GD20	Null	19.70~19.92KHz

Fig. 6 shows the frequency sweep response of MPU9250. The X-axis represents the frequency range of scanning noise and the Y-axis represents the abnormal output amplitude of the gyroscope. It can be found from Table III that the resonant frequency range of the MPU9250 chip is 26.48KHz 26.51KHz. By calculating the average amplitude of the sample, it is found that the noise frequency that makes the maximum abnormal amplitude of the gyroscope output is 26.495KHz. At this frequency, the maximum abnormal amplitude generated by the gyroscope is 2 $degree$. When the distance between the speaker and the MEMS gyroscope chip increases from 10cm to 40cm, the attenuation rate of the maximum abnormal value is only 7.2%, as shown in Fig. 7.

B. Attack model

Our goal is to inject adversarial noise into the gyro chip and change the vehicle trajectory. To achieve that, the following assumptions are required.

Target system access. The attacker can approach the target vehicle, but he cannot directly access the system, cannot change the target system settings, or install malware on the target system controller. Moreover, the attacker cannot directly

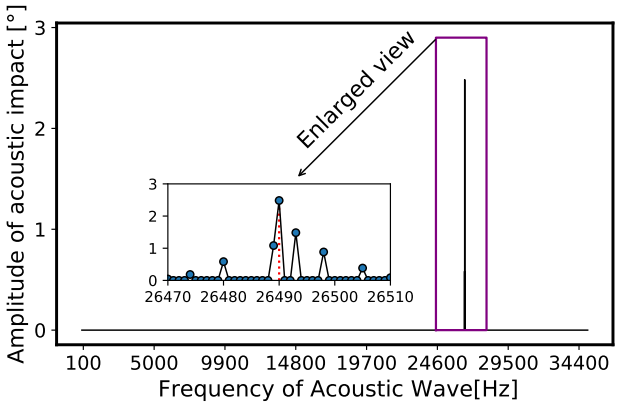


Fig. 6: The best frequency for resonance phenomenon

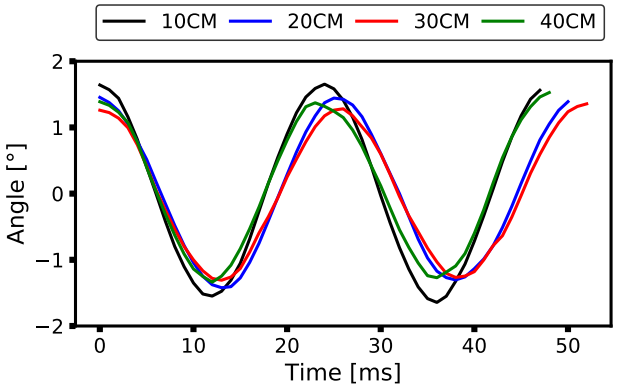


Fig. 7: The maximum abnormal value produced by MEMS gyroscope and the influence of distance attenuation

damage the sensor physically. However, this paper assumes that the attacker can learn about the control algorithm used in the target system by consulting manuals, etc.

Sensor evaluation. The attacker understands the basic principles of the sensor system. By investigating the second-hand car markets or car dealers, they can also obtain the sensor design parameters in advance, such as package, model, installation location, etc., to further explore the vulnerabilities of the sensor. The attacker may be proficient in hardware design and can use off-the-shelf hardware to complete the assessment and implement the attack. Based on the above assumptions, two possible attack models are discussed below.

External attack. On urban roads, the attacker can follow the car and use high-power ultrasonic equipment such as remote acoustic equipment and acoustic call equipment (AHDS) to follow the target vehicle within an effective distance. The attack distance may be several meters. In other words, the attacker has sufficient resources to make the attack farther, as shown in Fig. 8.

However, this scenario only applies when the victim vehicle is driving on a road segment with no other obstacles between it and the attacker's vehicle. In addition, the attacker can use a drone that equips a high-frequency sound wave transmitter, sending the attack sound wave to the target vehicle.

Insider attack. Attackers can use modified music to attack the ESP system deployed on the target vehicle as shown in Fig. 9. The attacker, for example, can inject the malicious

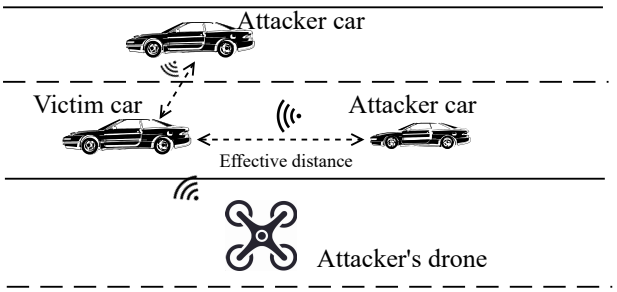


Fig. 8: External attack scenario

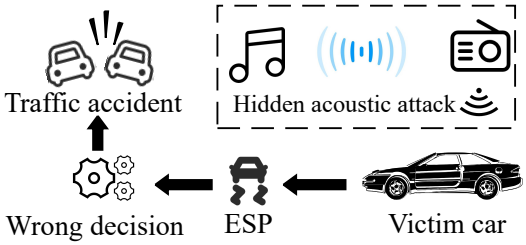


Fig. 9: Insider attack scenario

sound signal into a music file.

When people play the audio in the car, the hidden malicious sound wave attack can continuously and covertly affect the performance of the sensor, which may cause the sensor system to malfunction. In addition, attackers can use low-cost hardware devices that support software-defined radio (SDR) to broadcast a radio embedded with malicious sound waves at a specific frequency, thereby mimicking a radio station.

C. Inject the attacking signal to the music

To achieve the two attack scenarios mentioned in the previous section, we aim to superimpose the attack signal with the normal audio signal. The combined attack signal should meet the following two conditions: i) The frequency of the attack signal should be able to cause the MEMS sensor to produce a resonance effect. ii) The generated attack audio should be able to be played in the car's audio playback system.

Hardware-based injection method. The hardware-based solution is able to apply to the external attack. The required hardware is deployed on the attacker's vehicle to launch attacks while tracking the victim's vehicle. As shown in Fig. 10, we use a multi-channel adder to superimpose the resonant signal and the ordinary audio signal by adjusting the appropriate gain value and the amplitude of the attack signal. Then, the power amplifier will amplify the weak electrical signal from the signal source and drive the speaker to emit sound.

Software-based injection method. To perform the insider attack, we develop a method that reads the music signal from the original audio file and then injects the simulated digital attack signal into the music by Eq. 2.

$$S_{attack} = S_{music} + Asin(2\pi f_c t) \quad (2)$$

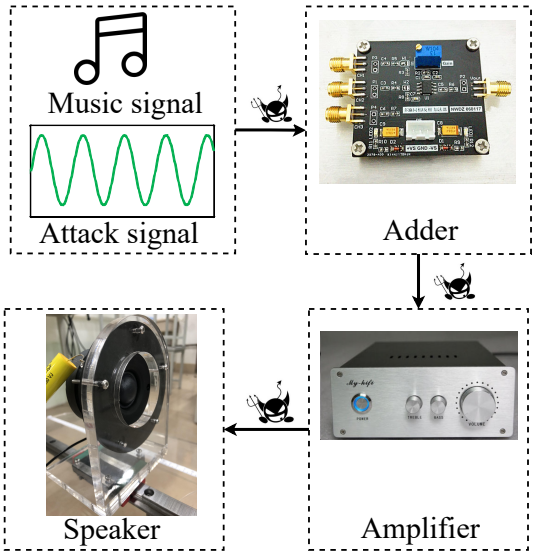


Fig. 10: Multiplexer makes malicious audio

where S_{attack} is the synthetic attack signal, S_{music} is the normal music signal, A and f_c are the attack signal's gain and frequency, respectively.

To save a digital signal into a playable audio file, we need to determine the playing time of the audio file by

$$duration = \frac{\lambda}{samplerate \cdot depth \cdot channel} \tag{3}$$

Here λ is binary digits, which is computed as $\lambda = filesize \cdot 8 \cdot 1024^2$. The *filesize*, *depth* and *channel* are the size of the audio file, bit depth and the number of channels, respectively, and the *samplerate* is a changeable parameter. For example, the frequency range of normal music is between 20Hz and 20KHz, so the sampling rate should be two times greater than the maximal frequency according to the Nyquist theory [14]. It usually is from 40KHz to 50KHz and its default value is 44.1KHz. The resonant frequency of the gyroscope is generally higher than 18KHz. For instance, the resonant frequency of the MPU9250 gyroscope is 26.5KHz. As a result, the frequency of the attack signal must be greater than 26.5KHz. If we want to inject the attack signal into the music, we have to increase the sampling rate to 53KHz.

In order to insert the attack signal into an audio file, the sample rate needs to be two times greater than or equal to the resonant frequency of the gyroscope. However, if we directly modify the sample rate to save an audio file, the duration of the original music will be severely distorted. We, therefore, develop a simple music signal rewriting method that duplicates the original digital single to allow the attack signal to be injected, as shown in Eq. (4), where n (positive integer) and f_c are the augment parameter and the attack signal frequency, respectively. SR_{music} represents the sample rate of the given audio file. The expanded music data is superimposed with the attack signal of equal length, and a new audio file is generated at n times the original sampling rate. For example, if the sample rate of the original audio is 44.1KHz and the frequency of the attack signal is 26.5KHz, we have to repeat the music 2 or more times.

$$n \geq \frac{2 \cdot f_c}{SR_{music}} \tag{4}$$

A set of resulting plots is shown in Fig. 11, where n is equal to 3. Figs. 11a, 11b, and 11c show the local information in the time domain of the original audio, the rewritten audio, and the mixed audio, respectively, while Figs. 11d, 11e, and 11f depict the corresponding complete audio from the frequency domain, i.e., the spectrogram of the audio. It can be seen from Figs. 11a and 11b that the waveforms of the original audio and the rewritten audio in the time domain are very close, so the human's ear usually cannot distinguish them (see Appendix A). The subtle changes in the frequency domain embodied in Fig. 11d and Fig. 11e can be completely accepted by the original playback equipment. Through the spectrograms shown in Fig. 11e, and Fig. 11f, it is not difficult to find that the constructed attack signal is perfectly superimposed into the rewritten music signal.

IV. Defense Strategy

In this section, we discuss the possible defense strategies for our proposed attack.

Passive defense method. Passive defense refers to hardening measures that are prepared in advance against a specific attack. The energy of the ultrasonic wave can be reduced by physical occlusion. Thus, we can wrap the sensors with a protective film such as a metal shell to reduce the possibility of resonance. However, this protection may fail for the following reasons: 1) the energy of the ultrasonic wave is strong enough to penetrate the protection. 2) Some covered sensors may affect their heat dissipation. Adding a low-pass filter (LPF) is another way to effectively mitigate high-frequency noise. However, in practical applications, LPF cannot completely eliminate high-frequency noise[15] (see Appendix B).

Active defense method. Active defense requires the ability to quickly respond to changes in threats. Appropriate detection mechanisms can also be used to detect and defend against such attacks. However, it is a challenge to accurately predict the sensor reading. "Long Short-Term Memory (LSTM)" outperforms other statistical and machine learning methods for nonlinear and complex time series data [16–18]. Inspired by these works, in this paper, we design an anomaly detection component based on LSTM-CUSUM, which is configured in front of the original ESP to filter outliers as shown in Fig. 12.

The vehicle model generates multi-sensor data in real-time, which is fed into the LSTM model in the form of a sliding time window with length m, and the model predicts the yaw rate at the next moment through a point-by-point prediction method [19]. Then, we compare the predicted outputs with the actual value of the sensor. If the difference is greater than the threshold, we will feed the predicted value to the ESP system to prevent the attacks.

Details of the LSTM model. The construction and training of the network model are based on the neural network toolbox provided by Matlab. The input of the network includes three dimensions of yaw rate theoretical value, lateral acceleration, and steering wheel angle, and the output is the predicted value

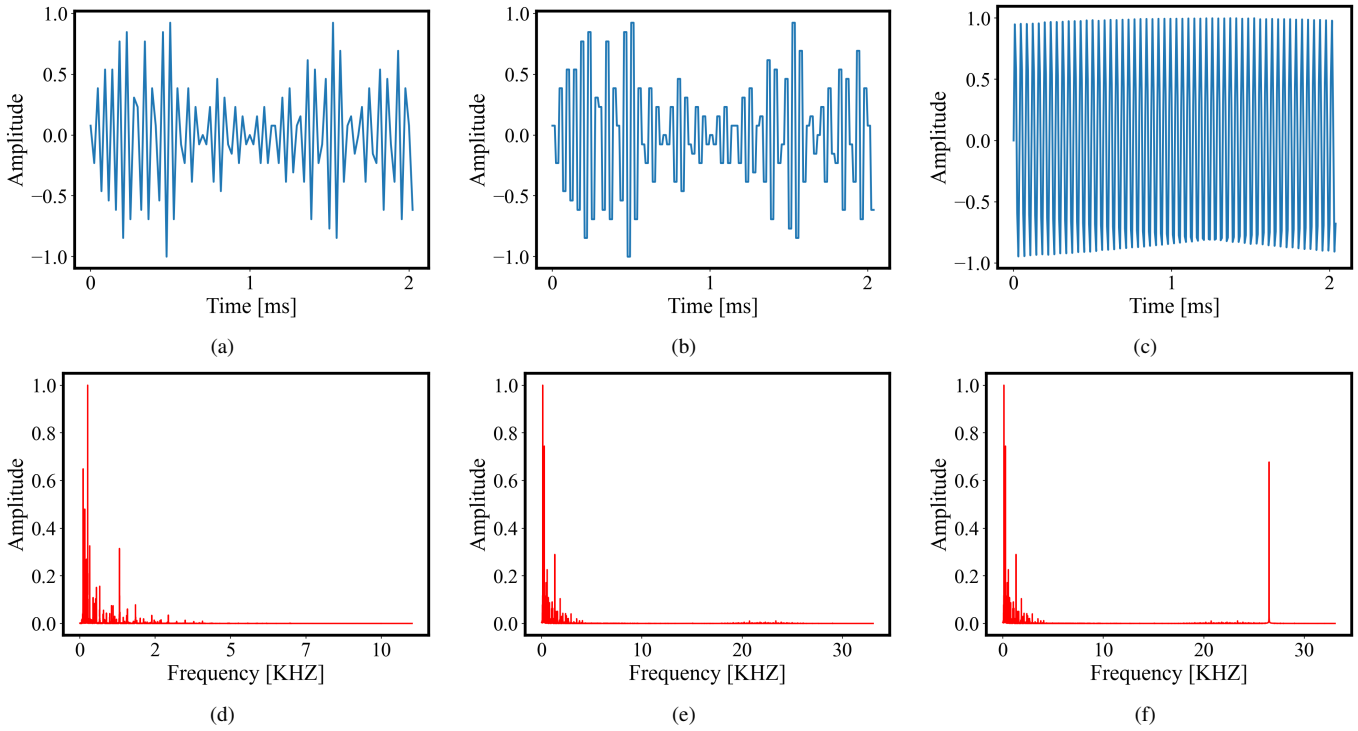


Fig. 11: (a) Original audio file, (b) An amplified audio file, (c) Superimposed audio file, (d)(e)(f) Corresponding spectrogram

of yaw rate. It consists of a 4-layer network with 30 neurons as input (sliding window size of 10, number of time series 3) and output of 1 neuron. The number of hidden layers is 2, the first layer contains 90 neurons, the second layer contains 180 neurons, and the loss function is Cross-Entropy. Before model training, the collected time series data needs to be pre-processed. For example, a set of data whose length is $3 * N$, can be divided into $3 * (N - m)$ sets of short sequence data whose length is $m + 1$. After that, $N - m$ training samples can be constructed based on them, the length of a single sample is $3 * m$, and the corresponding label is a single data. After preprocessing, all samples will be mixed and shuffled and put into training, the purpose is to make the prediction model also robust under changing operating conditions. Specifically, the collected data is divided into a training set and test set, wherein the specific gravity of the training set and test set is set to 4:1. During the training process, the neural network is only used as a simple predictor, and the loss is calculated by the difference between the predicted value and the real reading of the sensor and the gradient is updated in the reverse direction. To the end, the trained network can realize real-time tracking and prediction of test data (see Appendix C).

Determining the threshold. To define the threshold, we need to identify the impact of the environment noise and the real attack on the MEMS gyroscope. Thus, we obtain the thresholds T in the CUSUM algorithm via observing the experimental results. We simulate ten different road conditions(that is, the arrangement and combination of different driving conditions and road environments), and set ten different road noises (stones, puddles, etc.) to collect data. Then we calculate the cumulative error between the network prediction and the real sensor reading in a fixed time window. The threshold T is the average of the cumulative error after performing the experiment one hundred times.

V. Implementation

Our experimental setup consists of physical components and a simulator, as shown in Fig. 13. It mainly consists of three parts: 1) Malicious music generation unit, 2) Sensor data acquisition and transmission unit, and 3) Simulink and CarSim co-simulation unit. We use mobile phone music as a normal audio signal, superimpose with the attack signal generated by the signal generator, and then attack the gyroscope after power amplification. The sensor value fluctuations caused by the attack will be fed into the simulator in real-time with the help of STM32.

A. Simulating car system

To evaluate our attacking model, we refer to the method of [20–22] to develop a pipelined simulator for simulating automotive systems operating in various environments. Fig. 14 shows the pipelined simulator. In this paper, Carsim[20] is used to provide a holistic vehicle environment under various conditions, including vehicle body parameters, aerodynamic model, transmission model, suspension model, and road surface model. The ESP closed-loop control model is built with Simulink. Simulink uses the vehicle's system information (e.g., the yaw rate (w), the longitudinal velocity (A_x), and the front wheel angle ($Steer_{L1}$)) generated from Carsim and applies the ESP algorithm to generate control commands (i.e., the braking torque (T_b)) that are fed to Carsim. Our attack

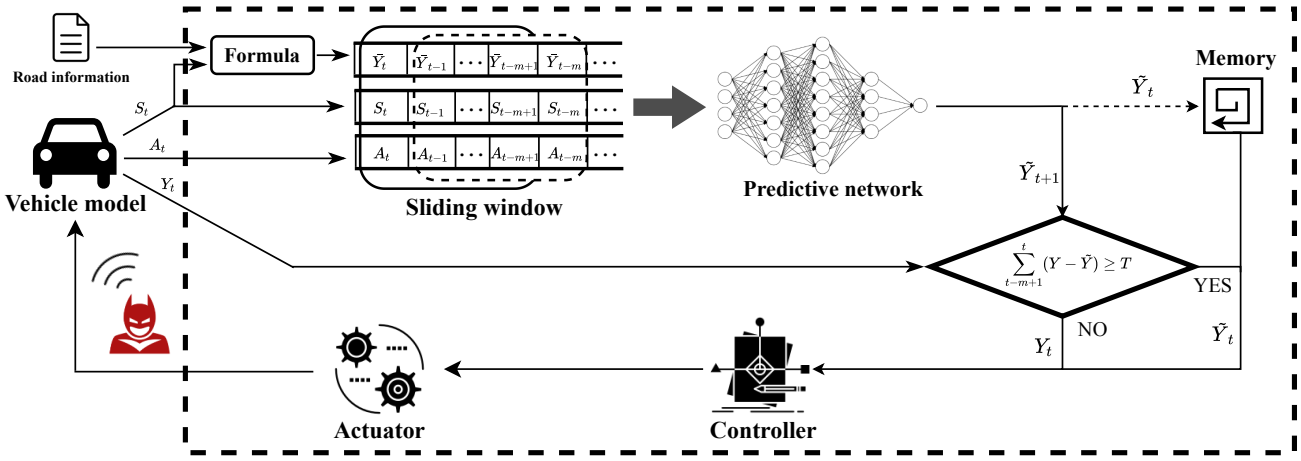


Fig. 12: The framework of the detection method. The thick dashed line inside is the proposed reinforced ESP structure. A, S, Y, $\bar{Y}$, $\tilde{Y}$ represent lateral acceleration, steering wheel angle, yaw rate, the theoretical value of yaw rate, and predicted value of yaw rate, respectively. The solid line represents the current moment, and the dashed line represents the previous moment.

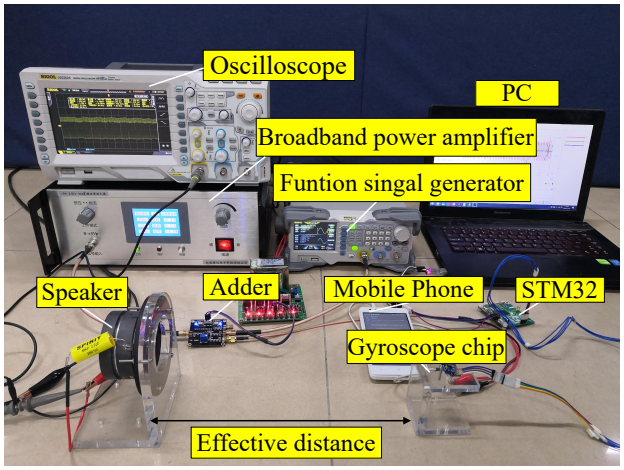


Fig. 13: Overall experiment setup

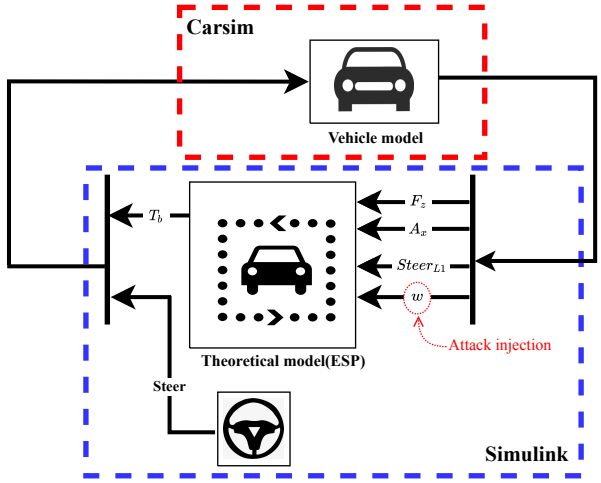


Fig. 14: Piplelined simulator

raises an ESP exception by changing the value of the yaw rate.

Build a vehicle model. We use a 2-DOF (degrees-of-freedom) dynamic model to describe the motion state of the moving vehicle (see Fig. 15). The mathematical expression is shown in Eq. (5), and the main parameters involved are given in Table I. $\dot{v_y}$ and $\dot{\omega_d}$ represent the derivatives of v_y and ω_d, respectively. A detailed derivation of this system of equations can be found in [22].

$$\begin{cases} (k_1+k_2)\beta_d + \frac{(ak_1-bk_2)\omega_d}{\mu} - k_1\delta_f = m(\dot{v_y}+\mu\dot{\omega_d}) \\ (ak_1-bk_2)\beta_d + \frac{(a^2k_1-b^2k_2)\omega_d}{\mu} - k_1a\delta_f = I_z\dot{\omega_d} \end{cases} \tag{5}$$

[22] indicates that the ideal yaw rate of the vehicle (ω_d) and the vehicle's stability coefficient (K) are given by

$$\omega_d = \frac{\mu/(a+b)}{1+K\mu^2}\delta_f \tag{6}$$

$$K = \frac{m}{(a+b)^2}\left(\frac{a}{k_2} - \frac{b}{k_1}\right) \tag{7}$$

ESP Based on Fuzzy PID Control. Fuzzy PID is a control algorithm based on intelligent reasoning, which is more suitable for nonlinear scenarios than ordinary PID. It can adaptively adjust PID coefficients to achieve a faster response.

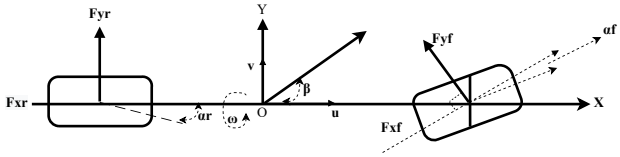


Fig. 15: The 2-DOF reference model of the vehicle

After setting the vehicle model, the error $e(\omega)$ and $e(\beta)$ will be used as the input of the controller, the controller output is the yaw moment compensation $\triangle M$ of the vehicle, and the controlled object is the vehicle model. Then, the vehicle model will give feedback to update $e(\omega)$ and $e(\beta)$. Therefore, when the input is a continuous signal, a continuous closed-loop control system can be formed (see the closed-loop structure in Fig. 16). The $e(\omega)$, $e(\beta)$ and $\triangle M$ are respectively given as

$$\begin{cases} e(\omega) = \omega - \omega_d \\ e(\beta) = \beta - \beta_d \end{cases} \tag{8}$$

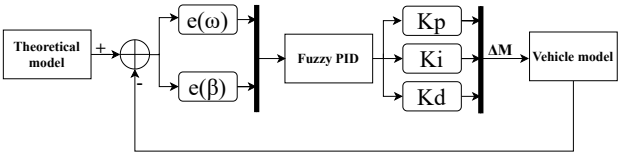


Fig. 16: Vehicle closed-loop control block diagram

$$\Delta M(t) = K_p(t) + K_i(t) \cdot \int_0^t \left(e_\omega(t) + e_\beta(t)\right) dt + K_d(t) \frac{d(e_{\omega(t)} + e_{\beta(t)})}{d_t} \tag{9}$$

where K_p, K_i, K_d represent the proportional coefficient, integral coefficient, and differential coefficient of the PID controller, respectively. After manually assigning an initial value of the PID parameters, the controller will optimize the parameters in real-time according to certain fuzzy rules[23].

B. Attack strategy

In order to improve the destructiveness and flexibility of the attack, we obtain the following attack strategy through analysis. By simply adjusting the positive, negative, and magnitude of the attack signal, the precise control of the steering state of the victim's vehicle can be achieved.

Table II shows that the torque distribution of the wheels depends on the positive and negative of ΔM, so we consider that the attack signal can be used to control ΔM, and thereby control the steering of the victim's vehicle. Eq. (9) gives the relationship between ΔM and e_ω and e_β. It can be seen from [24] that $e_\beta << e_\omega$, so the positive and negative of ΔM is completely determined by e_ω, which can be simplified as

$$\Delta M(t) = C_e[\omega(t) - \omega_d(t)] \tag{10}$$

In Eq. (10), C_e is always a positive number. Therefore, under the condition that the signal transmitting power is large enough, if the attack signal is positive, then $\omega > \omega_d$ is established, and the ESP will take braking measures to the right wheel. Conversely, if the attack signal is negative, the brake is applied to the left wheel.

VI. Evaluation

In this section, we will test the attack effect of malicious audio through the hardware-in-the-loop simulation platform (section V), and verify the effectiveness of the proposed active defense method.

A. Attack evaluation

We simulate the following two common high-speed driving scenarios to verify the effect of the acoustic attack. Specifically, the simulation duration is set to 3s and the sampling frequency is 50Hz. The power of the speaker is 15W.

Scenario 1. The vehicle runs in a straight line at a speed of 100km/h, and the road surface is a cement road with an adhesion coefficient of 0.7. Set the steering wheel input (unit is deg) to always 0. Let the attack signal be a positive pulse, and the frequency is set to 26.495KHz. The numerical fluctuation generated by the gyroscope is connected to the closed-loop control system 1s after the simulation starts. The trajectory of the victim's vehicle is shown in Fig. 17. It can be seen that the vehicle is in an unstable state after being attacked, and the trajectory appears to obviously deviated to the right.

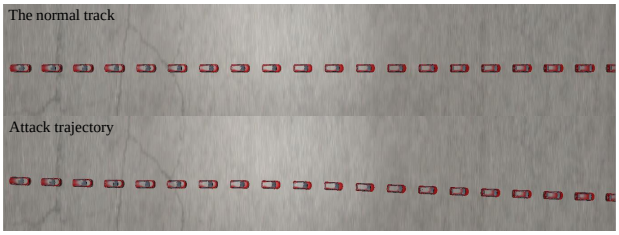


Fig. 17: Snapshot of straight-line driving trajectory

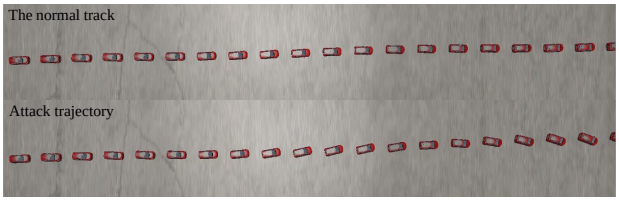


Fig. 18: Trajectory snapshot of vehicle lane change

Scenario 2. The vehicle changes lanes at a speed of 100km/h. The road setting is the same as in Scenario 1. Set the steering wheel input to a sine wave with a period of 3s and an amplitude of 30. The parameter settings of the attack signal and the time point of attack injection remain unchanged. The trajectory of the victim's vehicle is shown in Fig. 18. It can be seen that the normal lane change of the vehicle is damaged, and there is a large tail drift phenomenon, which has a great risk of rollover.

The yaw rate change of the vehicle in the above test environment is shown in Fig. 19. We can see that after the attack is injected, the real data of the sensor increases sharply, and there is a large deviation from the ideal value. Therefore, the ESP mistakenly believes that the vehicle is in an abnormal steering state, and the controller issues an incorrect torque compensation command, causing the vehicle to quickly lose control and deviate from the track.

In addition, we increase the signal transmitting power to 25W and further evaluate the proposed attack strategy on the basis of Scenario 2. We extend the simulation time to 5 seconds, and reduce the vehicle's speed to 50km/h, while the steering wheel input remains unchanged for the first 3 seconds, and 0 for the next 2 seconds. The vehicle's driving trajectory is shown in Fig. 20. It can be seen that the positive attack makes the target vehicle deviate to the right, while the negative attack makes the vehicle deviate to the left. It proves the effectiveness of the attack strategy.

B. Defense evaluation

For the two aforementioned attack scenarios, we embed the proposed LSTM-CUSUM framework into a closed-loop control system to evaluate the defense effect.

The change in the yaw rate of the vehicle is shown in Fig. 21. We can see that under the same attack, the real data, ideal

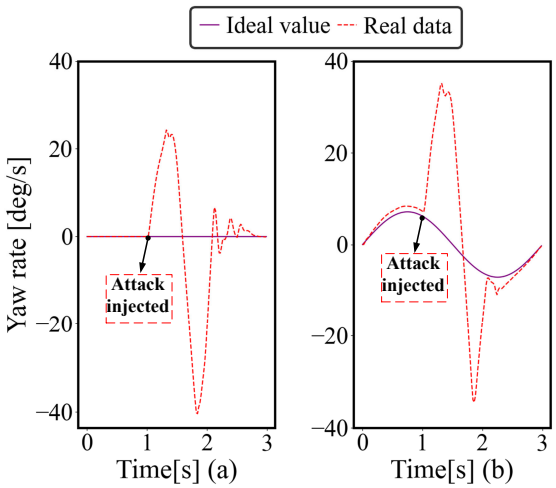


Fig. 19: (a) and (b) correspond to the yaw angular velocity values when the vehicles are attacked in the straight ahead and lane changing conditions, respectively.

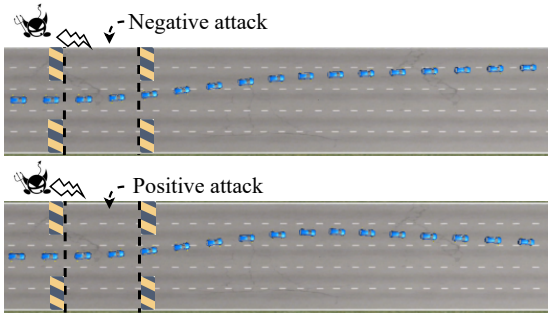


Fig. 20: Control vehicle's steering state through acoustic attack

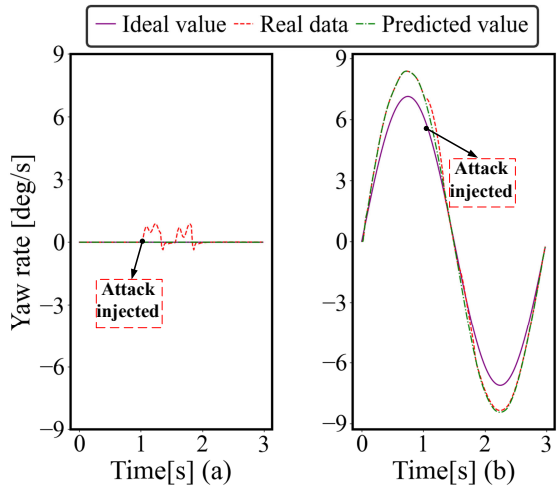


Fig. 21: (a) and (b) correspond to the yaw angular velocity values of the vehicle under the conditions of going straight and changing lanes after the defense is turned on, respectively.

Conversation language(60dB) The urban traffic(90dB)

Fig. 22: Noise decibel measurement

value, and predicted value of the sensor are relatively close. This indicates that the attack signal is not successfully expressed, because the neural network detects abnormal changes in the sensor value, and replaces the real data under attack with the predicted value into the ESP controller. Since the predicted value can well express the current steering state of the vehicle, the deviation from the ideal value is more realistic, so that the controller maintains a relatively stable working state.

C. Impact Quantification

The impact of background noise. In a real environment, the background noise is very common such as conversational speech (60dB) and urban traffic (90dB), which may affect our attack. In this section, we study the impact of background noise by performing our attack in a noisy environment.

Fig. 22 shows that we use a mini speak to create a background noise with 60dB and 90dB, and then perform the attack on the MEMS gyroscope. In Fig. 23, we can see that our attacks perform in an environment with background noise can achieve a similar performance as compared to the attack in a quiet environment.

The impact of plastic shells. The short wavelength determines that the diffraction ability of ultrasonic waves is poor, so it has strong penetration to obstacles[25]. In a real car, the sensor is not exposed to the real environment directly, only part of the energy penetrates into the chip, we need to determine the conditions or range of attack benefits. For example, you need to find out the power of the possible sound source.

Consequently, we also set the same environment and test at the same distance, and then we install the common plastic protective shell of equipment in the sensor. Fig. 24 shows the comparison of the attack effect between the plastic shell installed and not-installed, at the rated power of 15W. Obviously, the curve of a not-installed plastic shell fluctuates significantly more than that with plastic protection.

In terms of attack amplitude, under the condition of shell, the attack effect is reduced by 49.19%. The influence effect and attenuation percentage of the speaker output power on the sensor are shown in Fig. 25. We can see that the transmitted power can be proportional to the attack amplitude, and the shell attenuation of each transmission power is about 42%~57%. We calibrate the current attack range of 20cm and the speaker power of 15W as the effective attack amplitude. Therefore, if the attacker wants to realize the attack in the experimental environment in the real automobile, the transmitting power of the loudspeaker must be increased and achieve the best attack effect.

The impact of speed. Our attack is more powerful when the victim's vehicle is traveling at high speed. It is clear that a vehicle traveling at high speed will have a larger offset in a shorter time while being attacked. Fig. 26 shows that at a speed

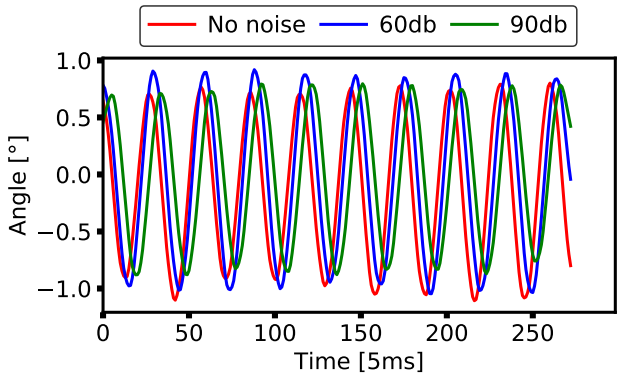


Fig. 23: Attack effects at different decibels

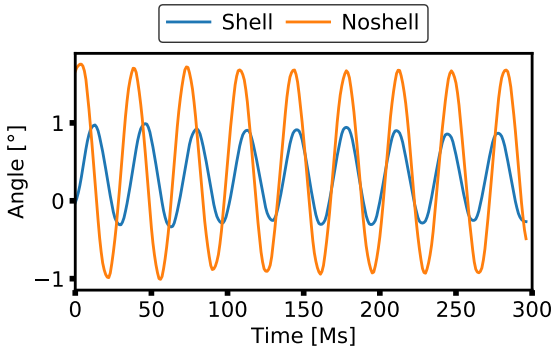


Fig. 24: Compare the effect of mounting shell and non-mounting shell

of 120 km/h, the victim's vehicle has a large tail drift within 2s. To have a similar offset it takes 5s when the speed is 30 km/h. The lower the vehicle speed, the more sufficient reaction time is left for the driver to manually adjust the direction of the car. However, the safety speed threshold depends on both the driver and the car, which can not be accurately measured. On the one hand, the driver's operating experience and safety awareness are also important reference factors in actual situations. On the other hand, each type of car has a different brake response delay, that is, the time elapsing from the moment when the braking force is applied to the moment when the braking system reaches the value of deceleration expected by the driver. Thus, the safety speed threshold can not be accurately measured. We may further discuss it in detail in future work.

VII. Discussion

In this section, we discuss the correlation between our attack and the type of carrier music. During the experiment, we select a total of twenty pieces of music as carrier signals and test them under the same attack setting, which covered various genres including pop music, pure music, rock, and classical music. At the same time, we also evaluate the detector. Since the attack signals contained in these attack music are the same, their attack effects are almost identical and Our detector can play an approximate adversarial effect on them.

VIII. Conclusion

In this paper, we act as both attacker and defender to illustrate that the vehicle's ESP can be spoofed by the acoustic

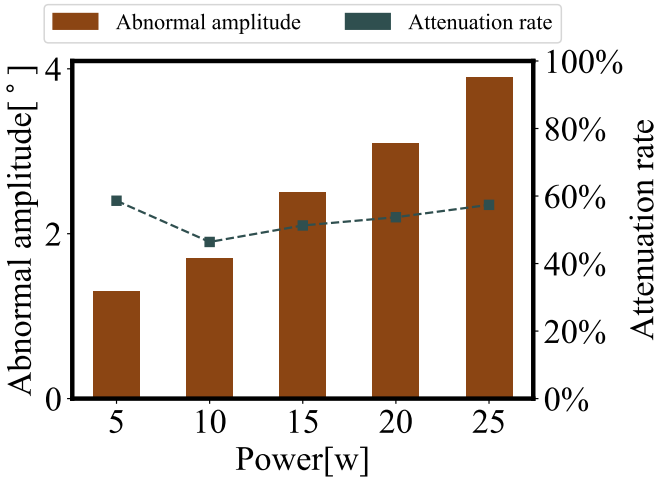


Fig. 25: Attack amplitude and attenuation percentage after adding shell under different power

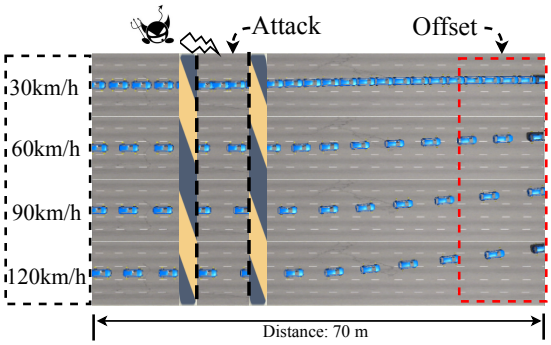


Fig. 26: Trajectory snapshots under attack at different driving speeds.

attack. For this purpose, firstly, we use "frequency sweep" to find the resonant frequency of MEMS gyro, and then construct a non-intrusive acoustic attack using the audio superposition method. Moreover, we build a semi-physical and semi-simulation platform to simulate the real environment to evaluate the attack. On the other hand, we propose effective defense strategies from two perspectives and systematically discuss the impact of other possible factors on attack effectiveness. Finally, we fully confirm that even a small part of a vehicle's key sensors are attacked, which can have very serious consequences on vehicle safety. In the future, we will consider conducting our offensive and defensive tests on real vehicles.

IX. Related Work

A. Cyber Attacks

For years, the auto industry has been investing heavily in driverless cars and connected cars. This tight coupling between network components [26] and the physical world in driverless cars often leads to more complex systems. Although this design has contributed to the functional and efficient development of modern cars, it has also introduced a range of potential cyber-attack problems. Koscher[27] were the first to demonstrate that it was possible to hack into vehicles,

many researchers[28] have discovered vulnerabilities in vehicle networks and control units, demonstrating the dangers of remote hacking to real vehicles[29][30]. In recent years, the research hotspots of cyber attacks mainly focus on Global Positioning System (GPS) and communication protocols [31]. Attacking GPS consists of two main approaches: GPS Jamming and Spoofing. GPS Jamming aims to block the vehicle to receive the GPS signals [32, 33]. Moreover, GPS Spoofing attack creates and transmits a fake GPS signal to the vehicle system, thereby deviating the system to a wrong destination [34, 35]. In response to GPS spoofing attacks, Zhang et al. [36] developed a game-theoretic security mechanism to defend against such attacks by portraying Bayesian equilibrium (PBE). Researchers can inject packets to the in-vehicle network to compromise electronic control units (ECUs) via remote vehicle network (e.g., Bluetooth, Cellular) [37] and this compromising has remotely stopped a Jeep Cherokee running on a highway [30]. Shin et al. [38] proposed a clock-based intrusion detection system. It collects periodic interval vehicle information to perform fingerprint identification on the electronic control unit. Then they used the recursive least squares (RLS) algorithm to build the baseline of the ECU clock behavior. The intrusion detection system can identify any abnormal changes that deviate from this baseline to achieve the purpose of rapid intrusion detection. Radio frequency identification (RFID) technology has been widely used for remote keyless entry (RKE) of modern vehicles. Still, many studies [39–41] have broken the security of the majority of RFID immobilizers. The vulnerabilities V2V (vehicle-to-vehicle) systems that utilize vehicular ad-hoc networks (VANETs) have also been studied in [42, 43].

B. Physical Attacks

Compared with the security in the automobile network, the physical security of vehicle sensors is also crucial for self-driving cars, but little research has been conducted. Petit et al. [44] successfully induced the generation of multiple obstacles based on the automatic Lidar. These obstacle points are not from real objects, but signals generated by injection. This is the first work to reveal that autonomous vehicle sensors can be easily affected by external stimuli. Another notable work of Yan et al. is to conduct a comprehensive safety analysis of the environmental awareness sensor installed on the Tesla Model S of an actual vehicle [45]. They point to a number of sensor vulnerabilities, such as their success in jamming ultrasonic sensors and injecting false signals, and in interfering with millimeter wave radar, and they also demonstrate, as Petit et al., that cameras are highly susceptible to strong light sources. In addition, Shoukry et al.[46] eliminated the legitimate magnetic field of the sensor by launching the reverse wave of the wheel magnetic encoder, and the ABS system would not be able to brake correctly. Xu et al.[47] took advantage of the vulnerability of ultrasonic sensors to design and cheat the obstacle detection system of autonomous vehicles, so as to make the vehicles crash. We believe that different types of sensors use different underlying physics, so the vehicle sensor safety challenges are diverse, Researchers have verified attacks against other sensors, such as cameras, fingerprint sensors, medical infusion pumps, analog sensors, and MEMS sensors [48][44][49][50][45]. However, There has been no prior work Attacking cars with MEMS gyroscope sensors and this paper is a work in that direction. In particular, we also proposed a defense measure against such sensor attacks and verified its feasibility.

C. Resonance on MEMS Gyroscopic Sensor

The sensor resonance is a type of mechanical resonance. When a mechanical system's oscillations are at the same frequency as its natural vibrational frequency (also known as its resonance frequency or resonant frequency), mechanical resonance occurs. This phenomenon causes mechanical systems to respond at greater amplitude on resonance frequencies than at other frequencies. Resonant frequency has been identified as a problem that causes the performance degradation of MEMS gyroscopes [6].

The typical architecture of a MEMS gyroscope consists of a resonating microstructure [8]. An electrostatic comb-driven actuator is used in this microstructure to create oscillations along one sensor's in-plane axis (i.e., the actuation axis). Another orthogonal in-plane axis is called the sense axis, while the orthogonal axis normal to the plane of the device is called the rotation axis. When the sensor is rotated about the rotation axis, the Coriolis force produces sinusoidal microstructure motion along the sense axis, the amplitude of which is proportional to the applied angular rate [51]. Since the microstructure, with a high mechanical quality factor, is intended to oscillate at its resonant frequency along the actuation axis, the sensor may be susceptible to external vibrations near that frequency in the working environment[52, 53].

Recently, many works[53–55] studied the susceptibility of MEMS gyroscopes to mechanical shock and high-frequency vibration. Geen [56], Weinberg *et al.* [55], and Weber *et al.* [57] presented that acoustic stimuli could adversely impact the performance of MEMS gyroscopes, but they did not present any experimental data to corroborate this. Later, Robert *et al.* [8] demonstrated that the MEMS gyroscopes are susceptible to high-power high-frequency acoustic noise when acoustic energy frequency components are close to the resonating frequency of the gyroscope's proof mass. Yunmok *et al.* [6] further investigated the effect of the resonant output of MEMS gyroscopes on the flight control of drones via software analysis. Moreover, this study designed a novel approach to attacking drones equipped with vulnerable MEMS gyroscopes using intentional sound noise.

D. Mitigation of High-Frequency Noise

When the frequency of the noise is high enough to be consistent with the natural frequency of the gyroscope, the resonance effect will destroy the output of the gyroscope. This poses a potential threat to some gyroscope-based applications, so researchers have explored ways to mitigate the effects of high-frequency noise.

A simple and feasible way is physical shielding, that is, using a shell to wrap the gyroscope. The absorption capacity

of the shell material to sound waves directly determines the attack mitigation effect. In this paper[58, 59], the acoustic characteristics of different materials are discussed, and a special sound insulation cover is designed using nickel microfiber material in wet papermaking process. This physical shielding method has a significant effect on the reduction of high-frequency noise. However, Redesigning hardware to tolerate acoustic interference is not an option for gyroscopics already deployed in the field. Another typical solution is to use multiple sensors to make decisions together. For example, triple module redundancy (TMR) uses three sensors to measure the same physical properties and produces a single output by majority voting or weighted average. In article[58], a differential measurement system consisting of two gyroscopes is designed and its robustness is verified in a high-frequency noise environment. Such solutions will not only add additional costs but will fail when multiple sensors are affected at the same time. Therefore, some studies explore defense mechanisms that can be implemented in software and deployed to actual systems as firmware updates. There are a series of studies based on the wavelet threshold denoising method[60–62]. Specifically, wavelet transform can be used to obtain high-frequency coefficients representing noise and low-frequency coefficients representing useful signals from noisy signals, and then denoising can be realized based on appropriate thresholds. This kind of method only has better performance for random noise. Since the denoised signal retains the characteristics of noise, it is not suitable for filtering our attack signal.

The basic idea of a recent study[63] is similar to ours. They also try to predict the output value of the sensor online and use this prediction value instead of the actual value to access the closed-loop control loop when attacked. The difference is that they achieve prediction by building a state space model, and we use neural networks to achieve this function. With the rapid iterative development of artificial intelligence technology, our method may gain more attention and be more expandable in the future.

REFERENCES

[1] Allied Market Research. *Autonomous Vehicle Market Outlook - 2026*, 2020 (accessed may 30, 2020). https://www.alliedmarketresearch.com/autonomous-vehicle-market.

[2] Tanya Roosta, Shiuhpyng Shieh, and Shankar Sastry. Taxonomy of security attacks in sensor networks and countermeasures. In *The first IEEE international conference on system integration and reliability improvements*, volume 25, page 94, 2006.

[3] Yasser Shoukry, Paul Martin, Paulo Tabuada, and Mani Srivastava. Non-invasive spoofing attacks for anti-lock braking systems. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pages 55–72. Springer, 2013.

[4] Yasser Shoukry, Paul D. Martin, Paulo Tabuada, and Mani B. Srivastava. Non-invasive spoofing attacks for anti-lock braking systems. In Guido Bertoni and Jean-Sébastien Coron, editors, *Cryptographic Hardware and Embedded Systems - CHES 2013 - 15th International Workshop, Santa Barbara, CA, USA, August 20-23, 2013. Proceedings*, volume 8086 of *Lecture Notes in Computer Science*, pages 55–72. Springer, 2013.

[5] Kevin Fu and Wenyuan Xu. Risks of trusting the physics of sensors. *Communications of the ACM*, 61(2):20–23, 2018.

[6] Yunmok Son, Hocheol Shin, Dongkwan Kim, Youngseok Park, Juhwan Noh, Kibum Choi, Jungwoo Choi, and Yongdae Kim. Rocking drones with intentional sound noise on gyroscopic sensors. In *24th {USENIX} Security Symposium ({USENIX} Security 15)*, pages 881–896, 2015.

[7] 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. IEEE, 2017.

[8] Robert Neal Dean, Simon Thomas Castro, George T Flowers, Grant Roth, Anwar Ahmed, Alan Scottedward Hodel, Brian Eugene Grantham, David Allen Bittle, and James P Brunsch. A characterization of the performance of a mems gyroscope in acoustically harsh environments. *IEEE Transactions on Industrial Electronics*, 58(7):2591–2596, 2010.

[9] Shyh-Chin Huang and Werner Soedel. Effects of coriolis acceleration on the free and forced in-plane vibrations of rotating rings on elastic foundation. *Journal of sound and vibration*, 115(2):253–274, 1987.

[10] Simon Castro, Robert Dean, Grant Roth, George T Flowers, and Brian Grantham. Influence of acoustic noise on the dynamic performance of mems gyroscopes. In *ASME 2007 International Mechanical Engineering Congress and Exposition*, pages 1825–1831. American Society of Mechanical Engineers Digital Collection, 2007.

[11] Shadi Khazaaleh, Georgios Korres, Mohammed Eid, Mahmoud Rasras, and Mohammed F Daqaq. Vulnerability of mems gyroscopes to targeted acoustic attacks. *IEEE Access*, 7:89534–89543, 2019.

[12] Geoffrey Brown. Discovering the stm32 microcontroller. *Cortex*, 3(34):64, 2012.

[13] Shutao Wei, Lin Chen, Lili Cui, and Zhenxing Wang. Interface design of eeprom and single chip micro computer by iic bus. In *2009 International Conference on Information Management and Engineering*, pages 554–557. IEEE, 2009.

[14] RA DeCarlo, J Murray, and R Saeks. Multivariable nyquist theory. *International Journal of Control*, 25(5):657–675, 1977.

[15] Umut Engin Ayten, Revna Acar Vural, and Tulay Yildirim. Low-pass filter approximation with evolutionary techniques. In *2011 7th International Conference on Electrical and Electronics Engineering (ELECO)*, pages II–125. IEEE, 2011.

[16] Anita Yadav, CK Jha, and Aditi Sharan. Optimizing lstm for time series prediction in indian stock market. *Procedia Computer Science*, 167:2091–2100, 2020.

[17] Alaa Sagheer and Mostafa Kotb. Time series forecast-

ing of petroleum production using deep lstm recurrent networks. *Neurocomputing*, 323:203–213, 2019.

[18] Vinay Kumar Reddy Chimmula and Lei Zhang. Time series forecasting of covid-19 transmission in canada using lstm networks. *Chaos, Solitons & Fractals*, 135:109864, 2020.

[19] Kai Chen, Yi Zhou, and Fangyan Dai. A lstm-based method for stock returns prediction: A case study of china stock market. In *2015 IEEE international conference on big data (big data)*, pages 2823–2824. IEEE, 2015.

[20] Yong Li, Huifan Deng, Xing Xu, and Wujie Wang. Modelling and testing of in-wheel motor drive intelligent electric vehicles based on co-simulation with carsim/simulink. *IET Intelligent Transport Systems*, 13(1):115–123, 2019.

[21] ShengHui Pan and HuangQian Zhou. An adaptive fuzzy pid control strategy for vehicle yaw stability. In *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pages 642–646. IEEE, 2017.

[22] Liangmo Wang, Li Tan, Li-hua An, Zhi-lin Wu, and Li Li. Study on the esp system based on fuzzy logic pid control and multibody dynamics. *Journal of Electrical Systems*, 8(1):57–75, 2012.

[23] Kit-Sang Tang, Kim Fung Man, Guanrong Chen, and Sam Kwong. An optimal fuzzy pid controller. *IEEE transactions on industrial electronics*, 48(4):757–765, 2001.

[24] Andrei Aksjonov, Klaus Augsburg, and Valery Vodovozov. Design and simulation of the robust abs and esp fuzzy logic controller on the complex braking maneuvers. *Applied Sciences*, 6(12):382, 2016.

[25] Vincent Chan and Anahi Perlas. Basics of ultrasound: Pitfalls and limitations. In *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, pages 11–15. Springer, 2018.

[26] Marko Wolf, André Weimerskirch, and Thomas J. Wollinger. State of the art: Embedding security in vehicles. *EURASIP J. Embed. Syst.*, 2007, 2007.

[27] Karl Koscher, Alexei Czeskis, Franziska Roesner, Shwetak N. Patel, Tadayoshi Kohno, Stephen Checkoway, Damon McCoy, Brian Kantor, Danny Anderson, Hovav Shacham, and Stefan Savage. Experimental security analysis of a modern automobile. In *31st IEEE Symposium on Security and Privacy, S&P 2010, 16-19 May 2010, Berleley/Oakland, California, USA*, pages 447–462. IEEE Computer Society, 2010.

[28] Charlie Miller and Chris Valasek. Adventures in automotive networks and control units. *Def Con*, 21:260–264, 2013.

[29] Stephen Checkoway, Damon McCoy, Brian Kantor, Danny Anderson, Hovav Shacham, Stefan Savage, Karl Koscher, Alexei Czeskis, Franziska Roesner, and Tadayoshi Kohno. Comprehensive experimental analyses of automotive attack surfaces. In *20th USENIX Security Symposium, San Francisco, CA, USA, August 8-12, 2011, Proceedings*. USENIX Association, 2011.

[30] Charlie Miller and Chris Valasek. Remote exploitation of an unaltered passenger vehicle. *Black Hat USA*, 2015:91, 2015.

[31] CG Leela Krishna and Robin R Murphy. A review on cybersecurity vulnerabilities for unmanned aerial vehicles. In *2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*, pages 194–199. IEEE, 2017.

[32] Esat Elezi, Göksel Çankaya, Ali Boyaci, and Serhan Yarkan. A detection and identification method based on signal power for different types of electronic jamming attacks on GPS signals. In *30th IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2019, Istanbul, Turkey, September 8-11, 2019*, pages 1–5. IEEE, 2019.

[33] Ahmad Y Javaid, Weiqing Sun, and Mansoor Alam. Single and multiple uav cyber-attack simulation and performance evaluation. *EAI Endorsed Trans. Scalable Information Systems*, 2(4):e4, 2015.

[34] Ahmad Y Javaid, Farha Jahan, and Weiqing Sun. Analysis of global positioning system-based attacks and a novel global positioning system spoofing detection/mitigation algorithm for unmanned aerial vehicle simulation. *Simulation*, 93(5):427–441, 2017.

[35] Soon Heng Mavric Tan and Chai Kiat Yeo. GPS location spoofing and FM broadcast intrusion using software-defined radio. *Int. J. Interdiscip. Telecommun. Netw.*, 12(4):104–117, 2020.

[36] Tao Zhang and Quanyan Zhu. Strategic defense against deceptive civilian gps spoofing of unmanned aerial vehicles. In *International Conference on Decision and Game Theory for Security*, pages 213–233. Springer, 2017.

[37] Karl Koscher, Alexei Czeskis, Franziska Roesner, Shwetak Patel, Tadayoshi Kohno, Stephen Checkoway, Damon McCoy, Brian Kantor, Danny Anderson, Hovav Shacham, et al. Experimental security analysis of a modern automobile. In *2010 IEEE Symposium on Security and Privacy*, pages 447–462. IEEE, 2010.

[38] Kyong-Tak Cho and Kang G Shin. Fingerprinting electronic control units for vehicle intrusion detection. pages 911–927, 2016.

[39] Ygal Bendavid, Nasour Bagheri, Masoumeh Safkhani, and Samad Rostampour. Iot device security: Challenging "a lightweight RFID mutual authentication protocol based on physical unclonable function". *Sensors*, 18(12):4444, 2018.

[40] Thomas Hänel, Alexander Bothe, René Helmke, Christoph Gericke, and Nils Aschenbruck. Adjustable security for rfid-equipped iot devices. In *IEEE International Conference on RFID Technology & Application, RFID-TA 2017, Warsaw, Poland, September 20-22, 2017*, pages 208–213. IEEE, 2017.

[41] Roel Verdult, Flavio D Garcia, and Baris Ege. Dismantling megamos crypto: Wirelessly lockpicking a vehicle immobilizer. In *Supplement to the Proceedings of 22nd {USENIX} Security Symposium (Supplement to {USENIX} Security 15)*, pages 703–718, 2015.

[42] Maria Azees, Pandi Vijayakumar, and Lazarus Jegatha

Deboarh. Eaap: Efficient anonymous authentication with conditional privacy-preserving scheme for vehicular ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, 18(9):2467–2476, 2017.

[43] Dimitrios Karagiannis and Antonios Argyriou. Jamming attack detection in a pair of RF communicating vehicles using unsupervised machine learning. *Veh. Commun.*, 13:56–63, 2018.

[44] Jonathan Petit, Bas Stottelaar, Michael Feiri, and Frank Kargl. Remote attacks on automated vehicles sensors: Experiments on camera and lidar. *Black Hat Europe*, 11:2015, 2015.

[45] Chen Yan, Wenyuan Xu, and Jianhao Liu. Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle. *DEF CON*, 24(8):109, 2016.

[46] Yasser Shoukry, Paul Martin, Yair Yona, Suhas Diggavi, and Mani Srivastava. Pycra: Physical challenge-response authentication for active sensors under spoofing attacks. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pages 1004–1015, 2015.

[47] Wenyuan Xu, Chen Yan, Weibin Jia, Xiaoyu Ji, and Jianhao Liu. Analyzing and enhancing the security of ultrasonic sensors for autonomous vehicles. *IEEE Internet of Things Journal*, 5(6):5015–5029, 2018.

[48] Radoslav Ivanov, Miroslav Pajic, and Insup Lee. Attack-resilient sensor fusion for safety-critical cyber-physical systems. *ACM Transactions on Embedded Computing Systems (TECS)*, 15(1):1–24, 2016.

[49] Paul Y Montgomery, Todd E Humphreys, and Brent M Ledvina. Experimental results of a multi-antenna receiver defense against a portable civil gps spoofer. In *Proceedings of the 2009 international technical meeting of the institute of navigation, Anaheim*, pages 26–28, 2009.

[50] Hocheol Shin, Yunmok Son, Youngseok Park, Yujin Kwon, and Yongdae Kim. Sampling race: Bypassing timing-based analog active sensor spoofing detection on analog-digital systems. In *10th {USENIX} Workshop on Offensive Technologies ({WOOT} 16)*, 2016.

[51] Min-Hang Bao. *Micro mechanical transducers: pressure sensors, accelerometers and gyroscopes*. Elsevier, 2000.

[52] Robert Dean, George Flowers, Nicole Sanders, Roland Horvath, Michael Kranz, and Michael Whitley. Micromachined vibration isolation filters to enhance packaging for mechanically harsh environments. *Journal of microelectronics and electronic packaging*, 2(4):223–231, 2005.

[53] T Gordon Brown. Harsh military environments and microelectromechanical (mems) devices. In *SENSORS, 2003 IEEE*, volume 2, pages 753–760. IEEE, 2003.

[54] Robert Dean, George Flowers, Scotte Hodel, Ken MacAllister, Roland Horvath, and Alex Matras. Vibration isolation of mems sensors for aerospace applications. In *SPIE proceedings series*, pages 166–170, 2002.

[55] Marc S Weinberg and Anthony Kourepenis. Error sources in in-plane silicon tuning-fork mems gyroscopes. *Journal of Microelectromechanical systems*, 15(3):479–491, 2006.

[56] John A Geen. Very low cost gyroscopes. In *SENSORS, 2005 IEEE*, pages 4–pp. IEEE, 2005.

[57] Mark Weber, Mark Bellrichard, and Chris Kennedy. High angular rate and high g effects in the mems gyro. *Sensors Magazine, November*, 2004.

[58] Pregassen Soobramaney. *Mitigation of the effects of high levels of high-frequency noise on mems gyroscopes*. PhD thesis, 2013.

[59] Pregassen Soobramaney, George Flowers, and Robert Dean. Mitigation of the effects of high levels of high-frequency noise on mems gyroscopes using microfibrous cloth. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 57113, page V004T09A014. American Society of Mechanical Engineers, 2015.

[60] Zhang Ruoyu, Gao Shuang, and Cai Xiaowen. Modeling of mems gyro drift based on wavelet threshold denoising and improved elman neural network. In *2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, pages 1754–1761. IEEE, 2019.

[61] Zan-ping Li, Qiong-jian Fan, Li-min Chang, and Xiaohong Yang. Improved wavelet threshold denoising method for mems gyroscope. In *11th IEEE International Conference on Control & Automation (ICCA)*, pages 530–534. IEEE, 2014.

[62] Jianguo Yuan, Yantao Yuan, Feilong Liu, Yu Pang, and Jinzhao Lin. An improved noise reduction algorithm based on wavelet transformation for mems gyroscope. *Frontiers of Optoelectronics*, 8(4):413–418, 2015.

[63] Hongjun Choi, Sayali Kate, Yousra Aafer, Xiangyu Zhang, and Dongyan Xu. Software-based realtime recovery from sensor attacks on robotic vehicles. In *23rd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2020)*, pages 349–364, 2020.

[64] Raymond E Wright. Logistic regression. 1995.

[65] Anthony J Myles, Robert N Feudale, Yang Liu, Nathaniel A Woody, and Steven D Brown. An introduction to decision tree modeling. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 18(6):275–285, 2004.

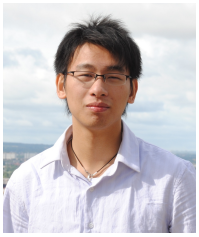
[66] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[67] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4, 2015.

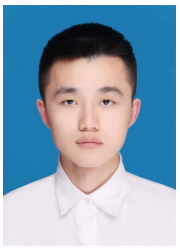
[68] Hirotugu Akaike. Autoregressive model fitting for control. In *Selected Papers of Hirotugu Akaike*, pages 153–170. Springer, 1998.

Zhen Hong received the B.S. degree from Zhejiang University of Technology (China) and University of Tasmania (Australia) in 2006, respectively, and the Ph.D. degree from the Zhejiang University of Technology (ZJUT) in Jan. 2012. He is a full professor with the Institute of Cyberspace Security, and College of Information Engineering, Zhejiang University of Technology, China. Before joining to ZJUT, he was an associate professor with the Faculty of Mechanical Engineering & Automation, Zhejiang Sci-Tech University, China. He has visited at the Sensorweb Lab, Department of Computer Science, Georgia State University in 2011. He also has been at CAP Research Group, School of Electrical & Computer Engineering, Georgia Institute of Technology as a research scholar in 2016 to 2018. His research interests include Internet of things, wireless sensor networks, cyberspace security, cybersecurity, and data analytics. He received the first Zhejiang Provincial Young Scientists Title in 2013 and the Zhejiang Provincial New Century 151 Talent Project (The Third-Level) in 2014. He is a member of IEEE, CCF and senior member of CAA, and serves on the Youth Committee of Chinese Association of Automation and Blockchain Committee and CCF YOCSEF, respectively.

Xiong Li received the B.S. degree from Jiangxi University of Science and Technology, Ganzhou, China, in 2020. He is currently pursuing the Ph.D. degree in control theory and control engineering with the College of Information Engineering, Zhejiang University of Technology, Hangzhou, China. His current research interests include unmanned system security and machine learning.



Zhenyu Wen (Member, IEEE) is currently a Tenure-Tracked Professor with the Institute of Cyberspace Security, and College of Information Engineering, Zhejiang University of Technology. His current research interests include IoT, crowd sources, AI system, and cloud computing. For his contributions to the area of scalable data management for the Internet of Things, he was awarded the the IEEE TCSC Award for Excellence in Scalable Computing (Early Career Researchers) in 2020.



Leiqiang Zhou received the B.S. degree from Hunan Institute of Engineering, Xiangtan, China,in 2020. He is currently working toward the master's degree in electronic information at the College of Information Engineering, Zhejiang University of Technology, Hangzhou, China. His current research interests include unmanned system security and machine learning.

Huan Chen received a master's degree in engineering from Zhejiang Sci-Tech University in 2021, and he is now working at DBAPPSecurity Co., Ltd Research Institute. His main research interests cover IoT security and machine learning.

Jie Su received B.S. degree in computer science and technology from China Jiliang University, China, in 2017 and M.S. degree in Data Analytics from University of Southampton (Distinction), UK, in 2018. He is currently pursuing the Ph.D. degree in Computer Science in OpenLab, Newcastle University, UK. His current research interests focus on the application of deep learning, and IoT security.

APPENDIX A
ATTACK CONCEALMENT ASSESSMENT

Regarding the 'concealment' assessment of the generated attack audios, we interviewed 20 volunteers to evaluate the 'reality' of generated attack audios, that is, whether the attack audios can be distinguished from normal audios. Specifically, we designed a questionnaire (i.e., Table IV) which include the evaluation of 15 audios with 5 pairs of normal/attack, 5 pairs of normal/normal and 5 pairs of attack/attack audios. The volunteers need to judge if the given audios are identical or different. The collect response from the interview and the results are shown in Fig. 27.

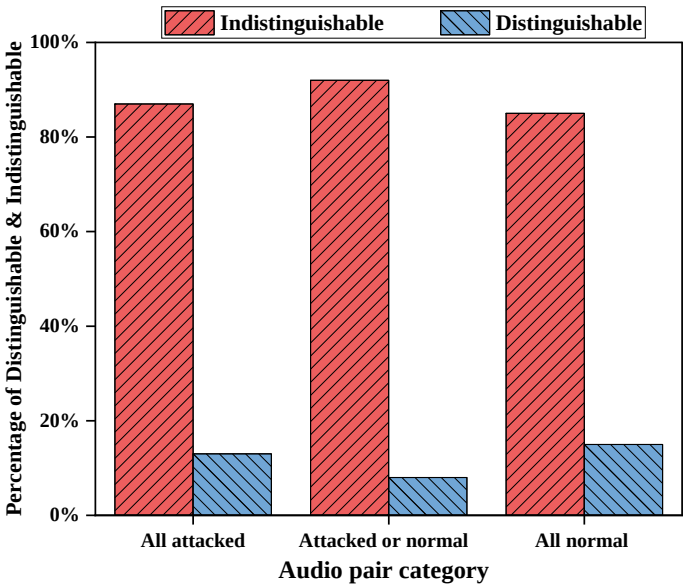
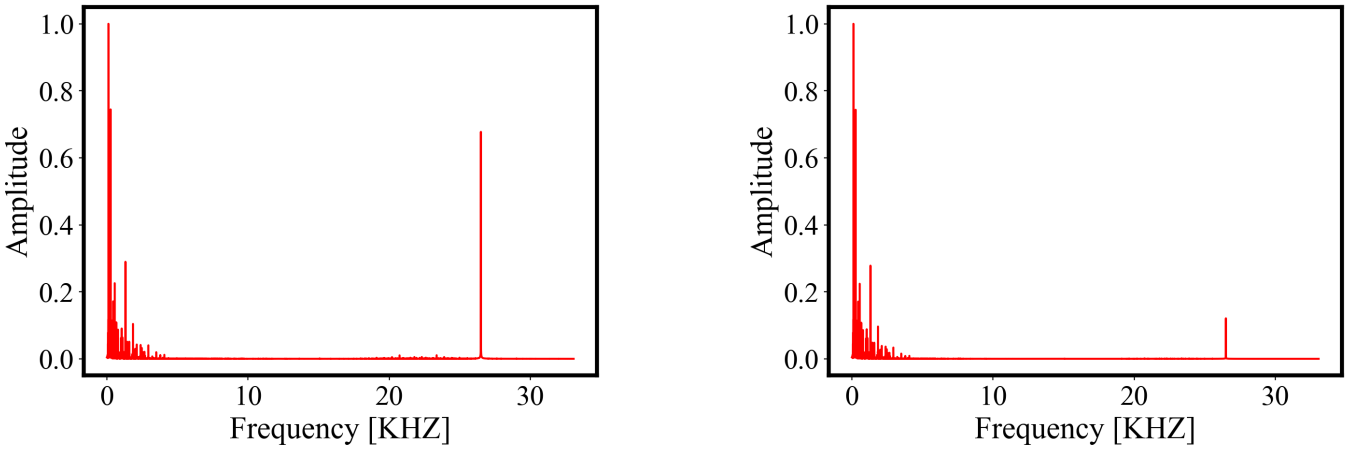


Fig. 27: The results of the questionnaire are about whether the respondents can accurately identify "Attaked audios" and "Normal audios", "Audio pair category" represents three groups (five pieces of audio each) of different situations of audio pairs.

TABLE IV: The questionnaire

Music types	Music lists	Question	Answer(YES/NO)
All attacked	As It Was	can you hear the difference?	
	Running Up That Hill		
	Afraid To Feel		
	Green Green Grass		
	Glimpse Of Us		
Attacked or normal	Break My Soul		
	Layla		
	About Damn Time		
	Beautiful Girl		
	Follow		
All normal	Kleiner Prinz		
	Dicht Im Flieger		
	Powerade		
	The Motto		
	We Made It		



(a) (b)

Fig. 28: (a)Spectrogram of original audio, (b)Spectrogram of filtered audio

TABLE V: Performance comparison

	Logistic regression[64]	Decision tree[65]	Random forests[66]	XGboost[67]	Autoregressive model[68]	Ours
SSE	0.9477	0.1123	**0.0081**	0.0331	0.0912	0.0113
MAE	0.2136	0.0212	0.0112	0.0027	0.0467	**0.0053**
MSE	0.5243	0.0105	0.0053	0.0086	0.0316	**0.0030**
RMSE	0.7241	0.1027	0.0728	0.0927	0.1778	**0.0551**

APPENDIX B
LOW-PASS FILTERING EXPERIMENT

We tested how well a low-pass filter could filter our attacks. For prepared malicious audio, we compared the original spectrogram with the spectrogram after loss-pass filtering. The results are shown in Fig. 28.

APPENDIX C
PERFORMANCE COMPARISON

We compared our method (LSTM) with 6 other machine learning methods such as logistic regression etc. Furthermore, we utilize (MSE (Mean Squared Error), RMSE (Root Mean Squared Error), etc.) to evaluate the predictive performance of different methods. In TableV, each experiment was performed independently 20 times and averaged. We can see that the method (LSTM) significantly outperforms most traditional machine learning methods on the yaw rate time series under each sequence of operations.