AttackDeceiver: Anti-spoofing Automotive Radar System using an Interleaving Chirp Waveform

Kaiyi Huang, Shengding Liu, Yanlong Qiu, Yanjiao Chen,
Senior Member, IEEE, and Jin Zhang, *Member, IEEE,*

***Abstract*—Millimeter-wave (mmWave) radars are indispensable components of the safety-critical Advanced Driver Assistance Systems (ADAS), facilitating robust and weather-resilient sensing of the surrounding environment for autonomous vehicles. Despite the advanced capabilities, mmWave radars remain vulnerable to adversarial attacks. Malicious users can spoof victim radars, leading to distorted environmental observations and potentially hazardous driving behaviors. Prior anti-spoofing techniques have been proposed to counter specific attacks. However, they may not be effective against adaptive adversaries. To ensure driving safety, we introduce AttackDeceiver, a novel anti-spoofing system that enables resilient environmental sensing under diverse spoofing attacks through an interleaved chirp waveform. AttackDeceiver exploits the comparison of estimates derived from two independent channels to detect and mitigate false targets. Additionally, it implements a deception strategy to induce attackers to generate false targets with unreasonable velocities. The prototype of AttackDeceiver is realized using a compact setup equipped with commercial-off-the-shelf (COTS) radars. Experimental results demonstrate the efficacy of our system, achieving an impressive false target recall rate exceeding 95% and a substantial enhancement in SINR exceeding 10 dB.**

***Index Terms*—Autonomous Vehicles, Frequency-modulated Continuous Wave (FMCW) Radar, Millimeter-wave (mmWave) Radar, Radar Countermeasures, Spoofing Attack, Vehicle Security.**

I. Introduction

Autonomous vehicle (AV) [1] related research has become an extremely prominent topic in the last two decades, due to the vast improvement in sensor technologies such as ultrasonic sensors [2], cameras [3], LiDAR [4], and millimeter-wave (mmWave) radars [5]. These sensors collaborate to form the advanced driver assistance systems (ADAS), which sense the surrounding physical environment and make safety-critical decisions, such as emergency braking [6] and line change assist [7]. In particular, mmWave frequency modulated continuous wave (FMCW) radars serve as a core component of ADAS, as they provide precise and robust object detection even under inclement weather such as fog and low light, where other light-based sensors such as cameras and Lidars would fail [8]–[10]. Besides, they are also employed for pedestrian and blind spot detection [11], adaptive cruise control [12], multi-lane traffic monitoring [13], and intersection management [14].

Given the broad and crucial application scenarios, researchers have extensively explored various attacks on automotive radars. In these attacks, adversarial sensors attempt to manipulate the received signals of the victim radars to mount attacks. Generally, these attacks can be organized into two main categories: spoofing and jamming [15]. In jamming attacks, the attackers intentionally emit high-energy signals within the bandwidth of the victim radars, lowering their signal-to-noise ratio (SNR), thereby making surrounding objects undetectable [16], [17]. Fortunately, jamming attacks can be easily detected, as they cause the radars to malfunction. In contrast, spoofing attacks are difficult to detect and mitigate. In this case, attackers inject false targets that mimic the physical patterns of real objects, forcing the victim radars to detect objects not existing [18]–[25]. Pseudo-obstacles can cause the ADAS to perform dangerous driving behaviors such as emergency braking, which can lead to accidents endangering the lives of passengers.

In recent years, research in spoofing attacks has made tremendous progress. Most prior arts have achieved false target injection through time delays [18]–[23] or frequency shifts [24], [25]. By estimating the waveform of the victim radars, the attackers can realize customized spoofing, where the false targets have a reasonable attacker-selected range and velocity [22], [23], [25]. However, the existing spoofing approaches have certain limitations. They assume that all chirps in a frame have the same shape and apply the same time delay or frequency shift to the entire chirp sequence, which provides an opportunity to resist this coarse-grained attack. Nashimoto et al. [24] propose an alternating slope modulation scheme to identify false targets by detecting varying distance measurements. However, this approach cannot assist when the attacker applies fine-grained frequency shifts to each chirp. Based on the random start frequency technique discussed in [21]–[23], attackers can easily extend existing coarse-grained spoofing attacks to fine-grained attacks that adaptively adjust configuration for each chirp. Existing countermeasures [21], [23], [24] are ineffective against such attacks, as they are aimed at counteracting specific corresponding attacks, leav-

Kaiyi Huang is with the Department of Computer Science and Engineering, Southern University of Science and Technology, China (email: huangky2019@mail.sustech.edu.cn).

Shengding Liu is with the Department of Computer Science and Engineering, Southern University of Science and Technology, China (email: 12110813@mail.sustech.edu.cn).

Yanlong Qiu is with the Department of Computer Science and Engineering, Southern University of Science and Technology, China, and also with the Department of Computer and Information Science, Temple University, U.S. (email: qiuyl@mail.sustech.edu.cn).

Yanjiao Chen is with the College of Electrical Engineering, Zhejiang University, China (email: chenyanjiao@zju.edu.cn).

Jin Zhang is with Shenzhen Key Laboratory of Safety and Security for Next Generation of Industrial Internet, Research Institute of Trustworthy Autonomous Systems, Department of Computer Science and Engineering, Southern University of Science and Technology, China (email: zhangj4@sustech.edu.cn).

ing the sophisticated adaptive spoofing attacks beyond their consideration.

The core challenge of counteracting fine-grained adaptive attacks is classifying the detected target. The signals reflected by environmental objects are time-delayed versions of the transmitted signals with a certain Doppler shift. Thus, with knowledge of the transmitted signal waveform, malicious users can transmit carefully designed spoofing signals, to introduce false targets that follow similar patterns to real objects. As shown in Fig. 1, the commercial-off-the-shelf (COTS) radars simultaneously receive the reflected signal (blue arrow) and spoofing signal (red arrow) to sense the surrounding environment. Without any waveform design, they can not distinguish the real targets from the false ones, since all detected targets appear as bright spots at a certain distance and speed in the range-Doppler (RD) profiles.

Can we design an anti-spoofing system to defend the fine-grained adaptive attacks? Our key insight is that we can distinguish between false targets introduced by the attackers and real targets from the environment, leveraging cross-validating the ambient sensing results of two independent channels. Building upon this insight, we present AttackDeceiver, a robust and versatile anti-spoofing automotive radar system using an interleaving chirp waveform. We apply distinct phase offsets to the chirp sequences of these two channels, so that the real targets remain in their proper position after the phase calibration process, while the false targets are differently shifted in the Doppler domain. The phase modulation process can be seen as a form of signal encryption. The automotive radars are well aware of the phase offsets of each chirp, thus they can compensate to bring the real targets back to where they should be. In contrast, the false targets are forced to undergo distinct defender-selected Doppler shifts, as the attackers have no knowledge of the key. The inconsistency of the results between the two channels enables AttackDeceiver to detect and mitigate the injected false targets, thus making the anti-spoofing system robust to state-of-the-art (SOTA) spoofing attacks.

Moreover, the phase modulation provides additional benefits. Firstly, it can mislead the attackers to perform inaccurate relative velocity estimation. Through dynamic phase offset modulation, we can introduce unreasonable velocity fluctuations for false targets. By analyzing these patterns, we can effectively identify and mitigate these false targets, thereby expanding the spectrum of countermeasures. Secondly, it makes the proposed countermeasures difficult to detect and counteract. The phase offset of each chirp is much more minute compared to other waveform configurations, such as frequency slope, making it notoriously difficult to be accurately estimated. Even if malicious users could achieve precise estimation with expensive circuitry, it is almost mission impossible for them to know the proportionality between the motion-induced phase and the artificial-introduced phase. By employing frequency slope alternation within an interleaving chirp waveform, we propose a novel anti-spoofing system that effectively discriminates between false and real targets. We summarize our contributions as follows:

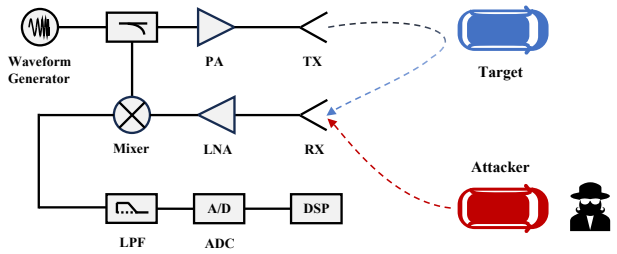


Fig. 1. Block diagram of COTS mmWave radars.

- We propose an anti-spoofing automotive radar system called AttackDecive to counter SOTA spoofing attacks. By employing an interleaving chirp waveform, COTS radars would be able to detect and mitigate false targets injected by the attackers by comparing results from multiple channels.
- We design a versatile interleaving chirp waveform to differentiate between false and real targets in diverse adversarial scenarios. Moreover, we present a novel mitigation algorithm to neutralize the effects of spoofing attacks and restore the expected sensing results.
- We implement a prototype of AttackDeciver using COTS mmWave radar. We conduct extensive experiments to show the effectiveness and robustness of our system. The experimental results show that AttackDeciver can achieve impressive false target recall over 95%, along with an SINR enhancement exceeding 10 dB under various attack scenarios.

Roadmap. The rest of this paper is organized as follows. Background and related work are presented in Section II. Section III demonstrates the threat model. Section IV offers a brief overview of our system, while Sections V and VI delve into the intricate design specifications of the transmitter and receiver components, respectively. The practical implementation of our system, along with a thorough evaluation of its effectiveness and robustness, is presented in Sections VII and VIII. Section IX addresses the inherent limitations of our approach and explores potential countermeasures. Finally, Section X concludes the paper.

II. Background and Related Work

In this section, we first review the workflow of FMCW radar and introduce the principles of range and velocity estimation. We then provide a comprehensive survey of the spoofing attacks and countermeasures.

A. Principle of FMCW Radar

Millimeter-wave radars leverage FMCW signals to sense environments and capture information about surrounding objects. With a fine wavelength, they can achieve high-resolution range and velocity estimation. The transmitted signal of each frame is a sequence composed of M chirps. Each chirp sweeps a bandwidth B over a period T_c in the radio-frequency (RF)

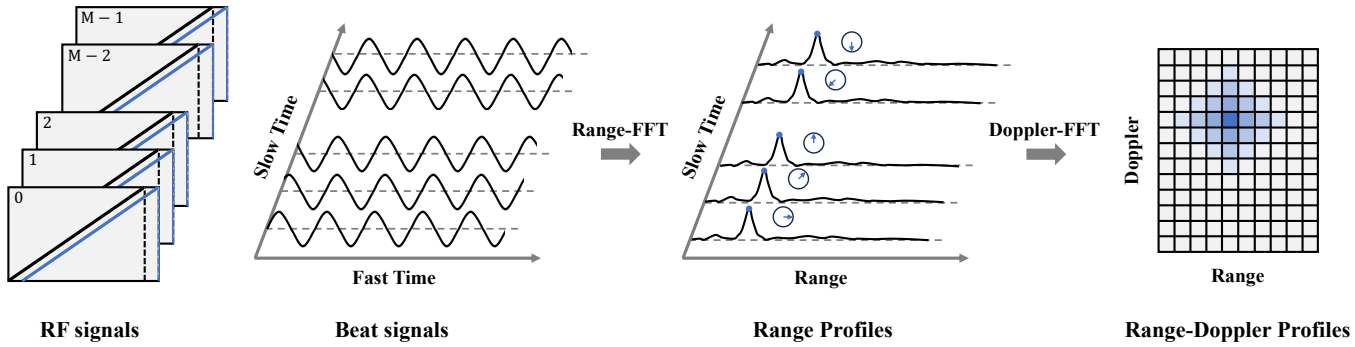


Fig. 2. Signal processing workflow of mmWave FMCW radars.

domain. The m^{th} chirp of the transmitted signal $x(t)$ in a frame can be represented as:

$$x(t) = e^{-j(2\pi f_c t_s + \pi\kappa t_s^2 + \phi_0)}, \tag{1}$$

where $t_s = t - mT_c$ is the within-chirp time, f_c is the start frequency, κ is the frequency slope, and ϕ_0 is the initial phase.

The transmitted signal propagates through the environment and is then reflected by ambient objects, generating echoes that can be viewed as scaled and delayed versions. These echoes are superimposed and received by the receiving antennas. Without loss of generality, we assume there are K objects in the environment. Then, the m^{th} segment of the received signal $y(t)$ in a frame can be expressed as:

$$y(t) = \sum_{k=1}^{K} A_k x(t_s - \tau_k), \tag{2}$$

where A_k and τ_k represent the scaling factor and time delay of the k^{th} target.

Afterward, the mixer within the radar circuit performs a dechirp operation, producing the conjugate product of the transmitted and received signals. The output is then downconverted by a low-pass filter (LPF), to produce the beat signal or the intermediate frequency (IF) signal. The m^{th} segment of the beat signal $h(t)$ in a frame can be denoted as:

$$h(t) = x(t) \cdot \overline{y(t)} = \sum_{k=1}^{K} A_k e^{-j(2\pi f_k t_s + \phi_k)}, \tag{3}$$

where f_k and ϕ_k are the beat frequency and beat phase of the k^{th} target. These beat frequencies and phases embed the range and velocity information of the targets. Utilizing digital signal processing methods, such as fast Fourier transform (FFT), we can obtain the RD profiles and extract the range and velocity information, thus realizing contactless environment sensing.

Range Estimation: FMCW radars determine the range of detected targets by analyzing the received signal. Assume only the k^{th} target is in the scene. As shown in Fig. 2, there is a time delay τ_k between the received signal (blue line) and the transmitted signal (black line), resulting in a certain frequency difference Δf_k across each chirp. With the time delay τ_k as an intermediate variable, we can establish the relationship between the range d_k and the frequency difference Δf_k:

$$\tau_k = \frac{2d_k}{c} = \frac{\Delta f_k}{\kappa}. \tag{4}$$

Given that the frequency difference Δf_k precisely corresponds to the beat signal frequency f_k, we can derive the range of the k^{th} target by performing the range-FFT on the beat signal and identifying the corresponding spectral peak.

$$d_k = \frac{c}{2\kappa} f_k. \tag{5}$$

Velocity Estimation: FMCW radars estimate the velocity of targets by monitoring the phase variation between consecutive chirps within a frame. Given the short duration of each chirp, the displacement of targets during a chirp is typically negligible compared to the range resolution. Consequently, the spectral peaks corresponding to chirps within a frame exhibit the same range bin in the range profile, as displayed in Fig. 2. However, these peaks have different phases due to the motion. Assuming uniform chirp spacing and constant target velocity within the frame, the phase difference between adjacent peaks remains constant. Therefore, we can perform another FFT operation, called Doppler FFT, to compute the target velocity:

$$v_k = \frac{\Delta d_k}{\Delta T} = \frac{\lambda \omega_k}{4\pi T_c}, \tag{6}$$

where λ is the wavelength, and $\omega_k = \phi_{k,m+1} - \phi_{k,m}$ is the phase difference.

In summary, by employing 2D-FFT on the beat signals, we can construct the RD profiles to reveal the range and velocity information of the targets. The concrete estimation values of range and velocity can be determined by locating the bins with significant intensity in the RD profiles.

B. Radar Spoofing Attacks

Millimeter-wave spoofing attacks have been intensively conducted. Recent works have attempted to launch spoofing attacks by actively generating spoofing signals or passively receiving and modifying radar transmission signals. A comparison among representatives of the generation-based and reflection-based spoofing attack is listed in Table I.

TABLE I
COMPARISON OF SOTA MMWAVE SPOOFING ATTACK WORKS

Work	Type	Techniques	Spoofing Capability	Attacking Device
S. Nashimoto et al. [21]	Generation-based	Time delay	Range	24GHz FMCW radar
R. Komissarov et al. [22]	Generation-based	Time delay	Range, velocity	Software-defined radio
Z. Sun et al. [23]	Generation-based	Time delay	Range, velocity, angle	Software-defined radio
P. Nallabolu et al. [24]	Reflection-based	Frequency shift	Range	SSB RF mixer
R. Reddy Vennam et al. [25]	Reflection-based	Frequency shift	Range, velocity	Phased array based USRP
X. Chen et al. [26]	Reflection-based	Meta-surface	Range, velocity, angle	Meta-material tags

Generation-based Spoofing: In these attacks, the attackers would emit self-generated spoofing signals at the proper time, causing the victim to detect the false targets [19], [21]–[23], [27]. To realize this spoofing attack, the attackers need to satisfy two fundamental conditions. Firstly, they need to know the transmitted signal waveform of the victims, to ensure that their injected false targets are similar to the real ones. Secondly, they need to synchronize with the victims, to ensure that the spoofing signals are transmitted at the expected time. Any minor error in synchronization will expose false targets due to the invalid range or velocity, and cause spoofing failures. Nashimoto et al. [21] achieve precise time synchronization through a wired connection and evaluate the usability of their range spoofing approach in an indoor environment. Komissarov et al. [22] and Sun et al. [23] take a step further. They use expensive specialized circuitry, the software-defined radio (SDR), to achieve wireless nanosecond-level synchronization, unleashing the unrealistic wired setup. Komissarov et al [22] propose a velocity spoofing method employing phase compensation, which significantly enhances the difficulty of false target detection. Sun et al [23] demonstrate a joint attack strategy, achieving multi-dimensional spoofing of range, velocity, and angle by coordinating multiple malicious radars.

Reflection-based Spoofing: In these attacks, the attackers intercept and manipulate the transmitted signals of the victim radars to achieve spoofing. Given that the intercepted signal embeds the waveform and clock information of the victim radar, the attacker can bypass the time synchronization process and flexibly adjust the attacking strategy to spoof different types of radars. Due to these advantages, many studies on reflective-based spoofing are flourishing. Several existing studies induce frequency shifts in intercepted signals to launch spoofing attacks [24], [25]. Nallabolu et al. [24] present a range spoofing approach employing a single sideband (SSB) RF mixer. However, this approach neglects the velocity information, making false targets easy to detect due to the random velocities. Vennam et al. [25] achieve independent control over the range and velocity of the false targets with phased array-based universal software radio peripheral (USRP), and demonstrate the spoofing capability in a realistic on-road environment. Some other approaches use meta-surface to spoof victim radars [26], [28], [29]. However, this spoofing attack would be easily detected. Because, the spoofing signal is filled with strong harmonic frequencies, leading to the appearance of suspected equidistant targets [25].

Spoofing Countermeasures: To safeguard the reliability of autonomous vehicles, researchers actively investigate countermeasures against the potential threats of spoofing attacks. Dutta et al. [30] present an on-off modulation scheme to mitigate generation-based spoofing attacks. This approach involves intermittently stopping transmission to detect attacks based on the presence of reflected signals. However, this countermeasure is ineffective against reflection-based attacks, as the adversary can immediately cease spoofing in response to transmission interruptions. Thus, we need to propose a robust mechanism for false target identification while continuously sensing the environment. In addition, Sun et al. [23] present a randomized phase modulation technique to disrupt the attacker from injecting meaningful false targets, by dispersing the target energy across the Doppler domain. However, this method would inadvertently interfere with other legitimate radars, leading to malfunction in their velocity estimation. Therefore, we should adopt a gentle approach to achieve spoofing resistance while avoiding mutual interference. Nashimoto et al. [24] propose a slope alternation scheme to detect reflection-based spoofing attacks based on mismatch range estimates. However, this method would fail when the attacker applies fine-grained frequency shifts to each chirp. Thus, we attempt to deceive the attackers and classify the detected targets by judging the plausibility of their velocity estimates. Furthermore, we would introduce waveform designs that are challenging for adversaries to detect, thereby obstructing them from fabricating false targets that mirror the characteristics of real ones.

III. THREAT MODEL

Attack Scenario: In this paper, we focus on a scenario where AVs are equipped with mmWave radars as the main information source for object detection. Specifically, we assume that the victim AVs are running on the road and malicious users are in the front or on the roadside conducting generation-based or reflection-based spoofing attacks. The objective of the attack is to continuously generate false targets with specific range and velocity to force the ADAS to make dangerous driving maneuvers such as hard braking and emergency lane changing.

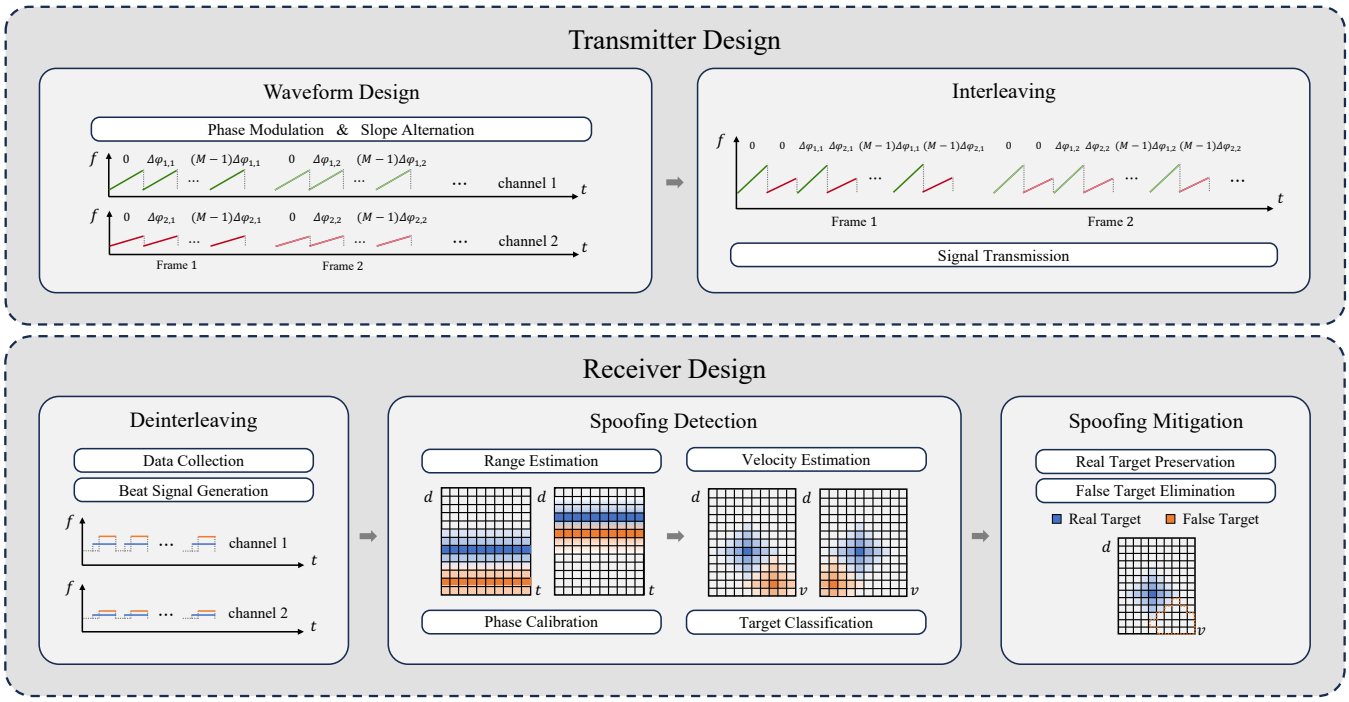


Fig. 3. System architecture of AttackDeceiver.

Attacker Knowledge: We assume that attackers possess knowledge of the hardware parameters of the victim mmWave radars, including the available RF band and maximum sampling rate. Furthermore, they have comprehensive information about the waveform configuration, such as frequency slope, start frequency, and chirp cycle time. The assumption is reasonable as mmWave sensors must follow standards specified by spectrum licenses, and the information can be accessed through open-source documentation, patent disclosure, and reverse engineering. In addition, methods for estimating radar waveform have been well researched [25], [31], which offer ways of realizing accurate waveform estimation.

Attacker Capability: In our setup, the attackers can not directly modify the collected data on the victim AVs. However, we assume they could achieve or bypass high-precision time synchronization to execute generation-based or reflection-based spoofing attacks, thereby inducing victim radars to generate erroneous estimates of the environment.

IV. AttackDeceiver: Overview

Existing countermeasures are inadequate in addressing the threat of adaptive spoofing attacks, as they primarily focus on counteracting static uniform adversary strategies. To bridge this critical vulnerability and ensure the safety of AVs, we propose AttackDeceiver, an anti-spoofing system designed to proactively deceive the attacker and induce them to generate false targets with ranges and velocities divergence from the expected values. The architecture of AttackDeceiver is shown in Fig. 3. In detail, we employ two independent channels with distinct frequency slopes and varying phase offsets to identify false targets effectively. Subsequently, we standardize the RD profiles obtained from both channels and apply a novel mitigation algorithm to preserve all real targets while eliminating false ones. By offering RD profiles immune to spoofing attacks, AttackDeceiver enables robust and precise environmental sensing.

V. Transmitter Design

A. Waveform Design

Conventional automotive radars are configured with a fixed frequency slope and zero phase offset for each chirp. With such a waveform design, malicious users can easily cause the victim radars to mistake false targets for real surrounding objects by transmitting spoofing signals at the right time. To make false targets not share similar patterns with the real ones, we employ two independent channels to sense the surroundings. Without loss of generality, we denote the frequency slopes of the two channels as κ_1, κ_2, where κ_1 is the steeper slope while κ_2 is the flatter one. Assuming each channel contains M chirps for each frame, the instantaneous frequency of the m^{th} chirp for channels 1 and 2 can be expressed as:

$$\begin{aligned} f_1(t) &= \kappa_1 t_s + f_c, \\ f_2(t) &= \kappa_2 t_s + f_c, \end{aligned} \tag{7}$$

where $m = 0, 1, \ldots, M-1$, $t_s = t - mT_c$ is the within-chirp time, and f_c is the start frequency.

In addition to the design of the frequency slope, we elaborate on the phase offsets of the channels. Distinct phase offset units $\Delta\phi_1$, $\Delta\phi_2$ are assigned to each channel, while the phase offset of each chirp is meticulously determined based on its corresponding index m. With the combination of slope alternation and phase modulation, we could represent the m^{th} chirp signal for channels 1 and 2 as:

$$
\begin{aligned}
x_1(t) &= e^{-j(2\pi\Phi_1(t)+m\Delta\phi_1)},\\
x_2(t) &= e^{-j(2\pi\Phi_2(t)+m\Delta\phi_2)},\\
\Phi_1(t) &= \textstyle\int f_1(t)\,dt = f_c t_s + \frac{1}{2}\kappa_1 t_s^2,\\
\Phi_2(t) &= \textstyle\int f_2(t)\,dt = f_c t_s + \frac{1}{2}\kappa_2 t_s^2,
\end{aligned}
\tag{8}
$$

where $\Phi_1(t), \Phi_2(t)$ are the phase portions due to instantaneous frequency accumulation, and $m\Delta\phi_1, m\Delta\phi_2$ are the artificially added phase offsets. Furthermore, we alternate the phase offsets of two channels over frames, to induce the attacker to generate false targets with fluctuating velocity.

$$
\Delta\phi_1, \Delta\phi_2 = \begin{cases} \Delta\phi_{1,1},\ \Delta\phi_{2,1}, & \text{when frame is odd,}\\ \Delta\phi_{1,2},\ \Delta\phi_{2,2}, & \text{when frame is even.}\end{cases}
\tag{9}
$$

B. Interleaving

Generally, radar systems employ multiple transmitting and receiving antennas to achieve multi-channel sensing. However, this approach is incompatible with certain low-cost commercial radars. To circumvent the associated limitations, we utilize time division multiplexing (TDM) to interleave the chirp sequences from two channels, thereby reducing the antenna requirement to a single pair of antennas. We formulate the transmission waveform $x(t)$ containing a total of $2M$ chirps per frame as:

$$
x(t) = \begin{cases} e^{-j(2\pi\Phi_1(t)+\frac{m}{2}\Delta\phi_1)}, & m=0,2,\ldots,2M-2,\\ e^{-j(2\pi\Phi_2(t)+\frac{m-1}{2}\Delta\phi_2)}, & m=1,3,\ldots,2M-1.\end{cases}
\tag{10}
$$

VI. Receiver Design

A. Deinterleaving

Automotive radars sense the environment by analyzing the received signal. Without loss of generality, we can decompose the received signal $y(t)$ into two components: the reflected signal $y_r(t)$ and the spoofing signal $y_f(t)$, where $y_r(t)$ and $y_f(t)$ corresponds to the K real targets and S false targets, respectively. By denoting the phase offset units of the attacker as $\Delta\phi_1'$, $\Delta\phi_2'$, we can express the received signal and its components as follows:

$$
\begin{aligned}
&y(t) = y_r(t) + y_f(t),\\
&y_r(t) = \begin{cases} \sum_{k=1}^{K} A_k e^{-j(2\pi\Phi_1(t-\tau_k)+\frac{m}{2}\Delta\phi_1)}, & m=0,2,\ldots,2M-2,\\ \sum_{k=1}^{K} A_k e^{-j(2\pi\Phi_2(t-\tau_k)+\frac{m-1}{2}\Delta\phi_2)}, & m=1,3,\ldots,2M-1,\end{cases}\\
&y_f(t) = \begin{cases} \sum_{s=1}^{S} A_s e^{-j(2\pi\Phi_1(t-\tau_s^1)+\frac{m}{2}\Delta\phi_1')}, & m=0,2,\ldots,2M-2,\\ \sum_{s=1}^{S} A_s e^{-j(2\pi\Phi_2(t-\tau_s^2)+\frac{m-1}{2}\Delta\phi_2')}, & m=1,3,\ldots,2M-1,\end{cases}
\end{aligned}
\tag{11}
$$

where A_k, A_s are the scaling factors of the real and false targets, τ_k is the round trip time delay and τ_s^c, $c=1,2$ is the attacker-introduced delay of the c^{th} channel.

After collecting the received signal, the mixer would then be applied to obtain the conjugate element-wise product of the received signal $y(t)$ and the reference signal $s(t)$, thereby generating the beat signal $h(t)$:

$$
\begin{aligned}
h(t) &= s(t)\cdot\overline{y(t)},\\
s(t) &= \begin{cases} e^{-j2\pi\Phi_1(t)}, & m=0,2,\ldots,2M-2,\\ e^{-j2\pi\Phi_2(t)}, & m=1,3,\ldots,2M-1,\end{cases}
\end{aligned}
\tag{12}
$$

where $s(t)$ is the reference signal devoid of artificially introduced phase offset. Subsequently, we would perform deinterleaving to split the beat signal $h(t)$ into two channels. The m^{th} segments of beat signal from the channels can be denoted as:

$$
\begin{aligned}
h_1(t) &= \underbrace{\sum_{k=1}^{K} A_k e^{-j(2\pi f_k^1 t_s+\phi_{k,m}^1-m\Delta\phi_1)}}_{real\ targets} + \underbrace{\sum_{s=1}^{S} A_s e^{-j(2\pi f_s^1 t_s+\phi_{s,m}^1-m\Delta\phi_1')}}_{false\ targets},\\
h_2(t) &= \underbrace{\sum_{k=1}^{K} A_k e^{-j(2\pi f_k^2 t_s+\phi_{k,m}^2-m\Delta\phi_2)}}_{real\ targets} + \underbrace{\sum_{s=1}^{S} A_s e^{-j(2\pi f_s^2 t_s+\phi_{s,m}^2-m\Delta\phi_2')}}_{false\ targets},
\end{aligned}
\tag{13}
$$

where $f_k^c = \kappa_c\tau_k$, $f_s^c = \kappa_c\tau_s^c$ are the beat frequencies, and $\phi_{k,m}^c$, $\phi_{s,m}^c$ are the inherent motion-induced phases.

B. Spoofing Detection

We can detect spoofing attacks by identifying false targets with inconsistent or unreasonable values based on the range and velocity estimates derived from the beat signals.

1) Range Estimation: We first perform the range-FFT on the beat signals from channels 1 and 2, to obtain the beat frequencies of the targets. By analogy to Eq. 5, we can derive the range estimates for both real and false targets:

$$
\begin{aligned}
\hat{d}_k^1 &= \tfrac{c}{2\kappa_1} f_k^1 = c\tau_k,\\
\hat{d}_k^2 &= \tfrac{c}{2\kappa_2} f_k^1 = c\tau_k,\\
\hat{d}_s^1 &= \tfrac{c}{2\kappa_1} f_s^1 = \tfrac{c}{2\kappa_1}(f_a^1 + \Delta f_s^1),\\
\hat{d}_s^2 &= \tfrac{c}{2\kappa_2} f_s^2 = \tfrac{c}{2\kappa_2}(f_a^2 + \Delta f_s^2).
\end{aligned}
\tag{14}
$$

Since the beat frequencies of real targets are proportional to the frequency slopes, the range estimates in both channels are equal to $c\tau_k$. In contrast, the beat frequencies of false targets include the component due to the victim-attacker distance f_a^c and the offset introduced by the attacker Δf_s^c. By adjusting the frequency offsets, the attacker can shift the false targets from their actual range to any desired spoofing range.

2) Phase Calibration: After completing the range estimation, the next step is to compute the phase differences between adjacent chirps and deduce the velocity. By subtracting the phases of neighboring chirps, we can obtain the phase differences for real and false targets in channels 1 and 2:

$$\begin{aligned}\omega_k^1 &= (\phi_{k,m+1}^1 - \phi_{k,m}^1) - \Delta\phi_1,\\ \omega_k^2 &= (\phi_{k,m+1}^2 - \phi_{k,m}^2) - \Delta\phi_2,\\ \omega_s^1 &= (\phi_{s,m+1}^1 - \phi_{s,m}^1) - \Delta\phi_1',\\ \omega_s^2 &= (\phi_{s,m+1}^2 - \phi_{s,m}^2) - \Delta\phi_2'.\end{aligned} \tag{15}$$

Since we introduce additional phase offsets to the transmission waveform, we would need to calibrate the phase of each chirp, according to the phase offset units of the channels $\Delta\phi_1$, $\Delta\phi_2$. The calibrated phase differences can be then expressed as:

$$\begin{aligned}\omega_k^{1\prime} &= \omega_k^1 + \Delta\phi_1 = \tfrac{4\pi T_c v_k}{\lambda},\\ \omega_k^{2\prime} &= \omega_k^2 + \Delta\phi_2 = \tfrac{4\pi T_c v_k}{\lambda},\\ \omega_s^{1\prime} &= \omega_s^1 + \Delta\phi_1 = \tfrac{4\pi T_c v_{a,s}}{\lambda} + (\Delta\phi_1 - \Delta\phi_1'),\\ \omega_s^{2\prime} &= \omega_s^2 + \Delta\phi_2 = \tfrac{4\pi T_c v_{a,s}}{\lambda} + (\Delta\phi_2 - \Delta\phi_2').\end{aligned} \tag{16}$$

where v_k is the relative velocity between the victim radar and the k^{th} target, and v_a is the attacker-victim relative velocity.

3) Velocity Estimation: As illustrated in Eq. 6, the velocity of targets depends on the inter-chirp phase difference. With the calibrated phase differences, we can provide accurate and consistent velocity estimates for the real targets:

$$\begin{aligned}\hat{v}_k^1 &= \tfrac{\lambda}{4\pi T_c}\omega_k^{1\prime} = v_k,\\ \hat{v}_k^2 &= \tfrac{\lambda}{4\pi T_c}\omega_k^{2\prime} = v_k.\end{aligned} \tag{17}$$

In contrast, besides the component embedding the spoofing velocity, the velocity estimates for false targets are compelled to include the component resulting from phase calibration:

$$\begin{aligned}\hat{v}_s^c &= \tfrac{\lambda}{4\pi T_c}\omega_s^{c\prime} = v_s^c + \tfrac{\lambda\Delta\phi_c}{4\pi T_c},\\ v_s^c &= \tfrac{\lambda}{4\pi T_c}\omega_s^c \;= v_a + \Delta v_s^c = v_a - \tfrac{\lambda\Delta\phi_c'}{4\pi T_c}.\end{aligned} \tag{18}$$

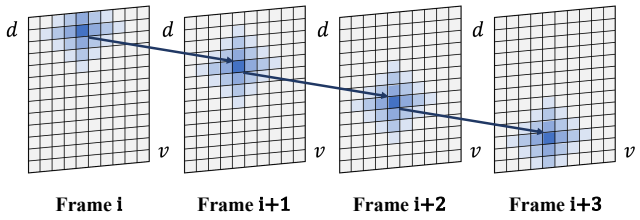
By adjusting the phase offset unit $\Delta\phi_c'$, the attacker could control the velocity offset Δv_s^c, thereby manipulating the spoofing velocity v_s^c and the estimate on the victim side $\hat{v}_s^c$.

4) Target Classification: We classify detected targets by comparing their range and velocity estimates from two channels. Specifically, we identify all false targets through three metrics: distinct range estimates, distinct velocity estimates, and similar yet unreasonable velocity estimates.

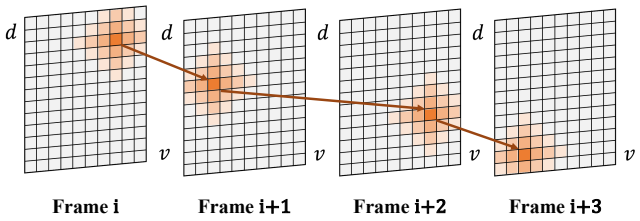
Case 1 False Target: We first detect false targets induced by reflection-based spoofing attacks through range inconsistencies. Typically, the attacker employs a uniform frequency shift Δf to launch reflection-based attacks, resulting in the frequency offsets Δf_1, Δf_2 are both equal to the same value. As a result, the range estimates for these false targets $\hat{d}_s^1$, $\hat{d}_s^2$ would have distinct values:

$$\begin{aligned}\hat{d}_s^1 &= \tfrac{c}{2\kappa_1}(f_a^1 + \Delta f_1) = d_a + \tfrac{c}{2\kappa_1}\Delta f,\\ \hat{d}_s^2 &= \tfrac{c}{2\kappa_2}(f_a^2 + \Delta f_2) = d_a + \tfrac{c}{2\kappa_2}\Delta f,\end{aligned} \tag{19}$$

where d_a is the distance between the victim and the attacker. However, when the attacker attempts to launch generation-based spoofing attacks, the frequency offsets Δf_1, Δf_2 vary proportionally with the frequency slopes, leading to the difference in range estimates disappears.



(a) The variation pattern of the real target.



(b) The variation pattern of the false target.

Fig. 4. Inter-frame phase offset alternation.

Case 2 False Target: We utilize the incoherence in velocity to identify false targets caused by generation-based attacks. From Eq. 18 we can derive the following expression:

$$\hat{v}_s^1 - \hat{v}_s^2 = \tfrac{\lambda}{4\pi T_c}[(\Delta\phi_1 - \Delta\phi_2) - (\Delta\phi_1' - \Delta\phi_2')], \tag{20}$$

which indicates that the velocity estimates of the two channels are consistent only when the difference in phase offsets imposed by the attacker match those of the victim. Given that the attacker has minimal knowledge of the victim's phase shifts, equalizing the velocity estimates is challenging.

Furthermore, if the attacker adheres to this condition, the choice of the spoofing velocity would be restricted. We denote the desired spoofing velocity in both channels as $\bar{v}_s$. Then, the velocity offset Δv_s^c can be determined as:

$$\Delta v_s^c = \bar{v}_s - \hat{v}_a^c. \tag{21}$$

where $\hat{v}_a^c$ is the relative velocity estimate on the attacker side. Since we apply additional phase offsets, the attacker would be deceived and obtain biased relative velocity estimate $\hat{v}_a^c$:

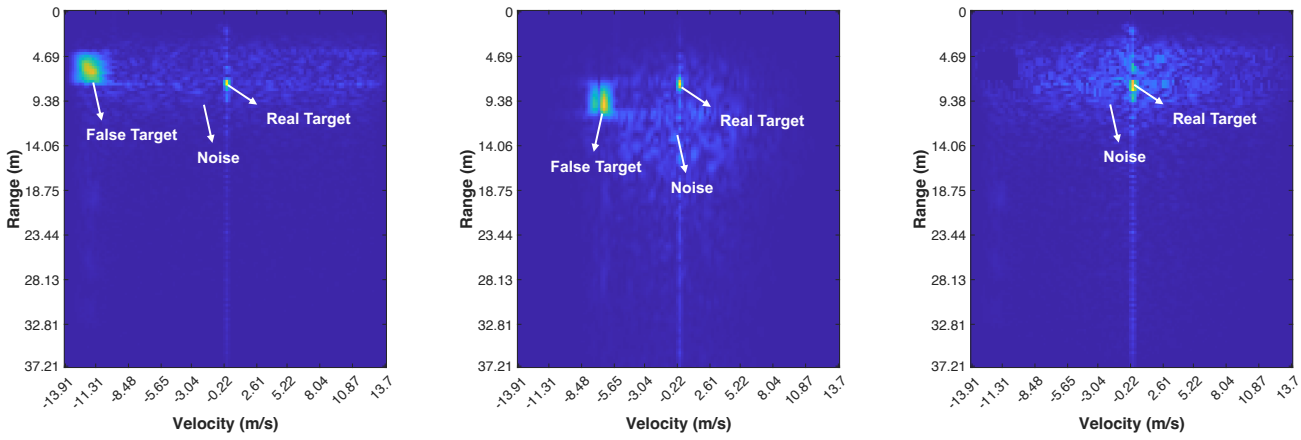
$$\hat{v}_a^c = \tfrac{\lambda}{4\pi T_c}\omega_a^c = v_a - \tfrac{\lambda\Delta\phi_c}{4\pi T_c}, \tag{22}$$

where ω_a^c is the phase difference on the attacker side. As a result, the velocity estimates on the victim side deviate differently from the expected spoofing value $\bar{v}_s$:

$$\hat{v}_s^c = v_a + \Delta v_s^c + \tfrac{\lambda\Delta\phi_c}{4\pi T_c} = \bar{v}_s + \tfrac{\lambda\Delta\phi_c}{2\pi T_c}. \tag{23}$$

Therefore, we can generally detect these false targets based on diverse velocity estimates.

Case 3 False Target: We identify false targets by detecting unreasonable fluctuations in velocity estimates when the attacker executes sophisticated adaptive attacks with subtly varying spoofing velocities. According to Eq. 22, we can introduce controlled deviations to the attacker's velocity estimates $\hat{v}_a^c$, by alternating the phase offset unit $\Delta\phi_c$. In response, the attacker would dynamically adjust the velocity offset Δv_s^c to maintain the expected spoofing velocity. However, the adaptive



(a) Channel 1. (b) Channel 2. (c) Mitigated result.

Fig. 5. An illustration of the input and output of the spoofing mitigation algorithm.

behavior results in the victim radar observing fluctuating target velocities, as depicted in Fig. 4(b). In contrast, the real targets remain unaffected, their range changes smoothly at a constant relative speed, as illustrated in Fig. 4(a). By leveraging the continuity of motion, we can effectively distinguish false targets from real ones.

C. Spoofing Mitigation

Based on the range and velocity estimates of the detected targets, we could localize their regions in the RD profiles. By preserving the regions corresponding to real targets and interpolating the regions associated with false targets, we could mitigate spoofing attacks and derive unaffected RD profiles. In practical implementation, the beat signals of the channels $h_1(t), h_2(t)$ are sampled and arranged into $N \times M$ matrices $\mathbf{H_1}$, $\mathbf{H_2}$ for the computation of RD profiles. We denote the beat signal matrix of the c^{th} channel as:

$$\begin{aligned} &\mathbf{H_c}[n,m] = h_c[n+mN],\ c{=}1,2, \\ &n{=}0,1,\ldots,N{-}1,\ m{=}0,1,\ldots,M{-}1, \end{aligned} \tag{24}$$

where N is the number of samples in the chirp, M is the number of chirps in the frame, $h_c[l]$ is the discrete beat signal, and $l{=}0,\ 1,\ \ldots,\ NM{-}1$ is the index of the sampled sequence.

We employ Range FFT along columns of the matrices i.e., the fast time dimension, to obtain the range profiles:

$$\mathbf{R_c}[n,m] = FFT(\mathbf{H_c}[n,m]). \tag{25}$$

Subsequently, we apply phase calibration and adopt Doppler FFT along rows of the matrices i.e., the slow time dimension, to derive the RD profiles:

$$\begin{aligned} \mathbf{P_c}[n,m] &= FFT(\mathbf{R_c}[n,m] * e^{-jm\Delta\phi_c}), \\ \mathbf{A_1}[n,m] &= ABS(\mathbf{P_1}[n,m]), \\ \mathbf{A_2}[n,m] &= INTERP(ABS(\mathbf{P_1}[n,m])). \end{aligned} \tag{26}$$

Since different frequency slopes are adopted, the RD profiles from the two channels have different unambiguity ranges and resolutions. Therefore, we standardize the RD profiles from channel 2 through interpolation. Afterward, we obtain the amplitude spectra of RD profiles through modulus calculation. An illustration of the amplitude spectra for channels 1 and 2 is shown in Fig. 5(a) and 5(b). With all the preprocessing complete, we can then mitigate spoofing attacks by applying Algorithm 1 to the amplitude spectra of the channels.

In Algorithm 1, we sequentially localize the false targets

Algorithm 1: Spoofing Mitigation.

Input: Amplitude spectra of RD profiles $\mathbf{A_1}$, $\mathbf{A_2}$;
Estimates of false targets with unreasonable values in channel 1: $\mathbf{E_u^1} = [(\hat{d}_1^1, \hat{v}_1^1), \ldots, (\hat{d}_U^1, \hat{v}_U^1)]$;
Small quantities of range and velocity ϵ_d, ϵ_v.

Output: Mitigation result $\mathbf{A_m}$.

```
// Case 1, 2 false targets localization
D = ReLU(A_1 − A_2);
M_d = CFAR(D);
// Case 3 false targets localization
M_u = 0;
if U > 0 then
    M_t = CFAR(A_1);
    [M_1, M_2, ..., M_{K+S}] = DBSCAN(M_t);
    for i = 1 to K+S do
        d̂_i^1, v̂_i^1 = EST(M_i);
        for j = 1 to U do
            if ||d̂_i^1 − d̂_j^1|| ≤ ε_d & ||v̂_i^1 − v̂_j^1|| ≤ ε_v then
                M_u = M_u + M_i;
M = M_d + M_u;
A_r = A_1 * (1 − M);        // Real preservation
A_f = INTERP(A_1, M);       // False elimination
A_m = A_r + A_f
return A_m;
```

with distinct estimates and those with unreasonable estimates. Initially, we compute the difference matrix $\mathbf{D}$ by applying the rectified linear unit (ReLU) function to the element-wise difference of the amplitude spectra $\mathbf{A_1}$, $\mathbf{A_2}$ (line 1). Since the mitigation result $\mathbf{A_m}$ is derived by eliminating the false targets in $\mathbf{A_1}$, we focus solely on the corresponding regions in channel 1. The regions associated with false targets in channel 2 are irrelevant. Therefore, we employ the ReLU function, a piecewise function that returns the maximum of zero and its input, to set the values in those regions to zero. Thereby, we could obtain the binary mask matrix $\mathbf{M_d}$ that represents the regions of false targets with distinct estimates, using the constant false alarm rate (CFAR) algorithm (line 2).

Considering the inherent difficulty of introducing false targets with implausible estimates, we initialize the binary mask matrix $\mathbf{M_u}$ with an all-zero matrix of identical dimensions (line 3). In scenarios where false targets with unreasonable estimates exist, we employ the CFAR algorithm to determine the regions corresponding to these false targets in $\mathbf{A_1}$ (lines 5). Subsequently, we apply the DBSCAN algorithm to partition $\mathbf{M_u}$ into K+S matrices, each representing an individual target (lines 6). For each target, we estimate its range $\hat{d}_i^1$ and velocity $\hat{v}_i^1$ based on the region center and compare them with the values of entries in the array $\mathbf{E_u^1}$. If the discrepancies in range and velocity are sufficiently small, we classify this target as a false target with unreasonable estimates and incorporate its matrix into the matrix $\mathbf{M_u}$ (lines 7-11). By integrating $\mathbf{M_d}$ and $\mathbf{M_u}$, we obtain the mask matrix $\mathbf{M}$, which denotes the regions of all false targets in channel 1 (line 12). Afterward, we preserve the original results from $\mathbf{A_1}$ in regions corresponding to the real targets (line 13) and interpolate the vacancies within the regions corresponding to false targets based on the edge values (line 14). By merging $\mathbf{A_r}$ and $\mathbf{A_f}$, we obtain the final mitigated result $\mathbf{A_m}$. As illustrated in Fig. 5(c), the mitigated result $\mathbf{A_m}$ successfully eliminates false targets from channel 1, retaining only real targets and noise.

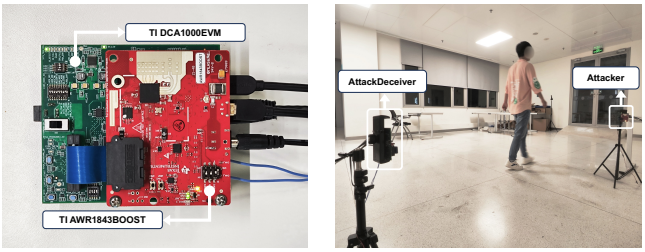
VII. Implementation

Hardware Platform: We implement AttackDeceiver and the attacker using the COTS mmWave radar kits provided by Texas Instrument (TI). As illustrated in Fig. 6(a), the radar kit includes a TI AWR1843BOOST sensor board operating at 76–81 *GHz* [32] and a TI DCA1000EVM data collection board [33]. The AWR1843BOOST features three transmitting antennas and four receiving antennas. Through the interleaving and de-interleaving, we enable robust environment sensing with only a pair of antennas. The chirp waveform configurations are detailed in Table II. Specifically, we set the sampling rate to $10\,MHz$, the chirp cycle time to $35\,\mu s$, and the frequency slope of channel 1 to $20\,MHz/\mu s$, to satisfy the requirements of $0.25\,m$ range resolution and $37.5\,m$ maximum range in blind spot detection. Furthermore, we configure the number of chirps per frame in each channel to 128, to accommodate the speed limit of $100\,km/h$.

Software: We use TI mmWave studio version 02.01.01 for radar configuration and MATLAB R2022(a) for radar signal processing. Our software runs on Windows personal computers with an AMD Ryzen 7 3700X CPU equipped with an 8-core processor and 16GB of RAM.

TABLE II
Chirp waveform configurations

Parameter	Value	Parameter	Value
Chirp Idle Time	$6\,\mu s$	Chirp Cycle Time	$35\,\mu s$
Chirps Per Frame	256	Frame Periodicity	$10\,ms$
Samples Per Chirp	256	Sampling Rate	$10\,MHz$
Phase Offset Unit 1	180°	Frequency Slope 1	$20\,MHz/\mu s$
Phase Offset Unit 2	90°	Frequency Slope 2	$10\,MHz/\mu s$



(a) AttackDeceiver setup. (b) Evaluation scenario.

Fig. 6. AttackDeceiver implementation and evaluation scenario.

VIII. Evaluation

A. Evaluation Setup

We set up our evaluation scenario in an open space ($6.5\,m \times 12.5\,m$) in the research building. Fig. 6(b) depicts the experimental environment and equipment. We employ one mmWave radar kit as AttackDeceiver and the other as the attacker. Meanwhile, we have individuals walking around within the detection field of AttackDeceiver as the real targets. We establish precise time synchronization among hardware devices via a wired connection to facilitate consistent and effective spoofing attacks by adversaries. In the static scenarios, we conduct extensive experiments with various frequency slopes and phase offsets to assess the impacts of different parameter configurations. Additionally, we evaluate robustness by experimenting with different distances and directions. Furthermore, we investigate the effect of motion in three dynamic cases.

B. Evaluation Metrics

We evaluate the spoofing detection performance of AttackDeceiver by analyzing precision and recall for false targets. The precision and recall are defined as:

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN}. \tag{27}$$

In our context, TP represents the false targets correctly identified as false. Whereas, FP, FN represents the real targets incorrectly identified as false, and the false targets incorrectly identified as real. Therefore, precision reflects the probability that a target identified as false is indeed false, while recall reveals the probability that a target with a false truth value is correctly identified. By utilizing the probabilities of these two

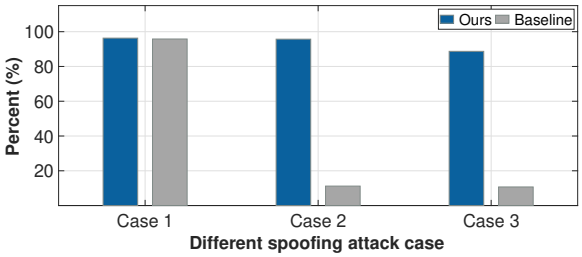


Fig. 7. Spoofing detection rate under three attack scenarios. Case 1: reflection-based spoofing attack. Case 2: generation-based spoofing attack. Case 3: adaptive spoofing attack.

types of misclassifications, we can comprehensively assess the detection performance of our system.

We evaluate the spoofing mitigation performance by comparing the signal to the interference plus noise ratio (SINR) before and after mitigation, which is defined as:

$$SINR = 10\lg(\frac{P_s}{P_i + P_n}), \tag{28}$$

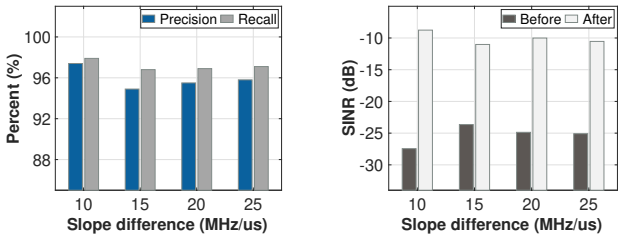
where P_s, P_i, P_n denote the power of the real targets, false targets, and noise, respectively. Since SINR integrates the information from targets and noise, it reflects the channel condition comprehensively. By comparing SINR before and after mitigation, we can effectively evaluate the mitigation performance through the SINR enhancement.

C. Overall Performance

We first evaluate the spoofing detection rate of AttackDeceiver under three distinct attack scenarios. To demonstrate the superiority of our anti-spoofing method, we compare AttackDeceiver with a baseline approach proposed by Nashimoto et al. [24]. As illustrated in Fig. 7, the baseline method proves ineffective against case 2 and 3 spoofing attacks. It can only sporadically detect attacks exploiting distinct range estimates when significant cumulative cycle drift in the synchronization occurs. In contrast, our method does not rely on errors in the spoofing attack implementation. It effectively recognizes case 2 and 3 attacks through distinct velocity patterns and unreasonable velocity fluctuations. While case 2 attacks can be detected based on single-frame analysis, case 3 attacks necessitate results from multiple frames. Consequently, the detection rate for case 3 attack is lower. The experimental results demonstrate the effectiveness of our system against diverse SOTA spoofing attacks.

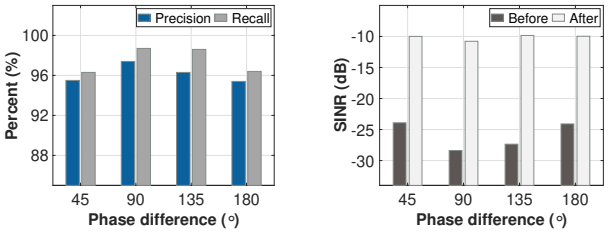
We then conduct extensive experiments to investigate the influence of parameter configurations on the overall performance, as AttackDeceiver offers the flexibility to adjust the frequency slopes and phase offsets of the two channels

Impact of Frequency Slopes: The frequency slopes of the channels directly determine the maximum unambiguous range and range resolution of the RD profiles, as we adopt a fixed chirp cycle time. Excessive slope differences may complicate the spoofing mitigation. Therefore, we evaluate the impact of frequency slopes by increasing the slope difference between two channels from $10\,MHz/\mu s$ to $25\,MHz/\mu s$ in increments of $5\,MHz/\mu s$. Fig. 8(a) demonstrates the detection



(a) Detection performance. (b) Mitigation performance.

Fig. 8. Impact of frequency slopes.



(a) Detection performance. (b) Mitigation performance.

Fig. 9. Impact of phase offsets.

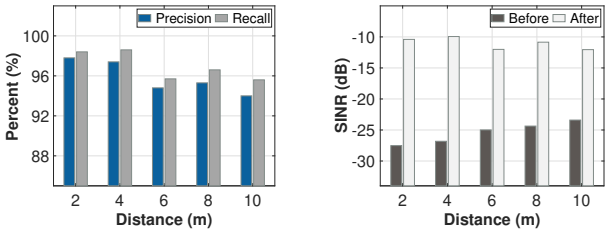
performance of our system. Although the slope difference increase yields a slight decrease in the precision and recall, the overall precision remains above 95% and the overall recall remains above 97%. Besides, as shown in Fig. 8(b), the SINR before and after mitigation is approximately $-25.27\,dB$ and $-10.08\,dB$. The average enhancement of $15.18\,dB$ indicates that our system can effectively mitigate false targets.

Impact of Phase Offsets: Besides, we evaluate the impact of the phase offsets by varying the phase offset difference between two channels from $45°$ to $180°$ in increments of $45°$. As illustrated in Fig. 9, the system performance exhibits a symmetric trend with increasing phase difference. The optimal performance is achieved when the phase difference is $90°$. At this point, the precision reaches 97.4%, the recall reaches 98.7%, and the SINR has an enhancement of $17.60\,dB$. Furthermore, the performance remains robust in the worst-case scenario, where the precision attains 95.3%, the recall reaches 96.4%, and the SINR enhances $13.93\,dB$. The performance result implies that our system enables effective spoofing detection and mitigation.

D. Robustness

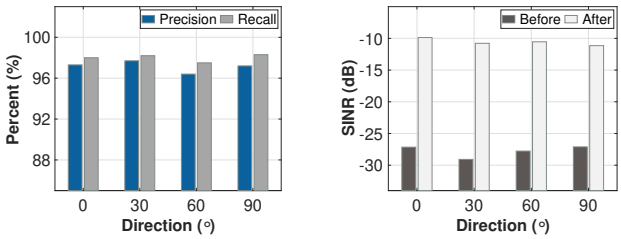
Robustness is a crucial issue for the anti-spoofing system. Thus, we conduct extensive experiments under varying distances and directions to demonstrate the robustness of AttackDeceiver.

Impact of Distance: We commence our evaluation on the impact of distance, by adjusting the separation between the attacker and AttackDeceiver from $2\,m$ to $10\,m$ with a $2\,m$ step. To exclude the influence caused by different directions, we place the attacker in the same incident direction of $0°$. As shown in Fig. 10(a), the precision and recall slightly decline when the distance exceeds $4\,m$. Nevertheless, the average precision and recall beyond $4\,m$ remain above 94% and 95%,



(a) Detection performance. (b) Mitigation performance.

Fig. 10. Impact of distance.

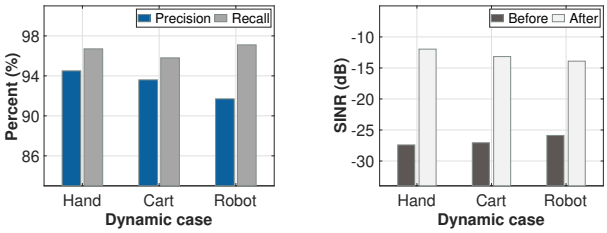


(a) Detection performance. (b) Mitigation performance.

Fig. 11. Impact of direction.

(a) Handheld. (b) Cart. (c) ROS robot.

Fig. 12. Different dynamic cases.



(a) Detection performance. (b) Mitigation performance.

Fig. 13. Performance of dynamic case.

indicating that our system sustains high detection performance. As displayed in Fig. 10(b), the variation in distance results in a minor reduction in the SINR enhancement, averaging $0.72\,dB$ per meter. Nonetheless, the SINR still enhances by over $11.35\,dB$ after mitigation. These performance outcomes suggest that our system is distance-resistant.

Impact of Direction: We then evaluate the influence of incidence angle on our system. We maintain the attacker at a constant distance and control the incidence direction within the COTS radar field-of-view (FOV), varying from $0°$ to $90°$ in $30°$ increments. As presented in Fig. 11(a), the precision and recall are minimally affected by the incidence direction, with average values of 96.9% and 98.1%. Additionally, Fig. 11(b) demonstrates a steady enhancement in the SINR across different directions. The lowest SINR enhancement still reaches $15.96\,dB$ at $90°$. The remarkable performance results reveal the directional robustness of our system.

E. Dynamic Case Study

In the AV scenario, the relative positions of the attacker and the victim vary frequently. Therefore, in this section, we evaluate the performance of our system in certain dynamic scenarios. Since the motion between the attacker and Attack-Deceiver is relative, we only test the situations where the attacker is moving. As shown in Fig. 12, we move the attacker using three approaches: handheld, mounted on a cart, and mounted on a ROS-based robot. In these dynamic cases, the attacker moves back and forth facing the AttackDeceiver at a speed of $3\,km/h$, $4\,km/h$, and $8\,km/h$. Meanwhile, we request individuals to walk around within the FOV as the detection objectives.

The performance results for the dynamic cases are shown in Fig 13. We can observe that the recall exceeds 96% across various speeds. Whereas, the precision decreases as the speed of the attacker increases. This result indicates that our system tends to misclassify real objects as false ones while the attacker moves fast. The misclassification can result in the erroneous exclusion of real targets, leading to a precision rate of 91.7% in the ROS robot scenario. Furthermore, such misclassification adversely affects mitigation performance. As depicted in Fig. 13(b), the SINR enhancement decreases with increasing speed. The worst SINR enhancement attains $11.97dB$ when the attacker is mounted on the ROS robot. Considering the human ability to sense surroundings also diminishes at higher speeds, we deem these results acceptable.

IX. Discussion

Weakly Reflective Object: In scenarios where pedestrians unexpectedly emerge from concealed areas, presenting a limited reflective surface area, the strength of the reflected signals can be substantially attenuated. Therefore, these targets may be erroneously classified as ambient noise before false target cancellation. To prevent collisions, we employ the spoofing mitigation algorithm to restore the RD profiles to their pre-attack state and employ existing weak target detection algorithms to warn these weakly reflective objects.

Dynamic Power Attack:: Attackers can strategically adjust their transmission power to ensure that false targets exhibit comparable signal strength to real targets, based on the distance between the victim and the attacker. Our system consistently resists spoofing attacks, even when false targets cannot be identified through excessive signal strength. Rigorous testing demonstrates our system's efficacy, achieving a recall exceeding 94%, and a precision surpassing 96%. Additionally, the SINR following spoofing mitigation stabilizes at approximately $-10\,dB$ in most scenarios, indicating the effectiveness of our mitigation algorithm. For a detailed analysis, please refer to Section VII.

Multi-radar Hybrid Attack: Even when multiple radars simultaneously employ various spoofing attacks, our system

can mitigate these threats without mutual interference. This is achieved by modifying only the areas where attacks are detected, thereby preserving unaffected regions. According to Algorithm 1, three types of spoofing attacks are addressed sequentially. Moreover, the induced false targets would be filtered out due to the excessively high beat frequency, when the attacker and the victim operate in different frequency bands (e.g., 24 GHz and 77 GHz).

X. Conclusion

In this work, we present AttackDeceiver, a novel anti-spoofing automotive radar system that detects and mitigates fine-grained adaptive attacks using the COTS mmWave radar kit. It employs an interleaving chirp waveform to detect false targets with different or unreasonable estimates and utilizes multi-channel fusion to reveal the expected unaffected environment leveraging RD profiles. Extensive experiments demonstrate that AttackDeceiver is capable of achieving accurate sensing in a variety of realistic environments. We believe it can be deployed to secure AVs in various scenarios with flexible spoofing attacks.

References

[1] J. Van Brummelen, M. O'brien, D. Gruyer, and H. Najjaran, "Autonomous vehicle perception: The technology of today and tomorrow," *Transportation research part C: emerging technologies*, vol. 89, pp. 384–406, 2018.

[2] J. Borenstein and Y. Koren, "Obstacle avoidance with ultrasonic sensors," *IEEE Journal on Robotics and Automation*, vol. 4, no. 2, pp. 213–218, 1988.

[3] J. Kim, S. Hong, J. Baek, E. Kim, and H. Lee, "Autonomous vehicle detection system using visible and infrared camera," in *2012 12th International Conference on Control, Automation and Systems*. IEEE, 2012, pp. 630–634.

[4] S. Royo and M. Ballesta-Garcia, "An overview of lidar imaging systems for autonomous vehicles," *Applied sciences*, vol. 9, no. 19, p. 4093, 2019.

[5] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 ghz frequency band," *IEEE transactions on microwave theory and techniques*, vol. 60, no. 3, pp. 845–860, 2012.

[6] S. Rajendar and V. K. Kaliappan, "Recent advancements in autonomous emergency braking: A survey," in *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2021, pp. 1027–1032.

[7] V. K. Kukkala, J. Tunnell, S. Pasricha, and T. Bradley, "Advanced driver-assistance systems: A path toward autonomous vehicles," *IEEE Consumer Electronics Magazine*, vol. 7, no. 5, pp. 18–25, 2018.

[8] M. Steinhauer, H.-O. Ruoß, H. Irion, and W. Menzel, "Millimeter wave-radar sensor based on a transceiver array for automotive applications," *IEEE transactions on microwave theory and techniques*, vol. 56, no. 2, pp. 261–269, 2008.

[9] K. Bansal, K. Rungta, S. Zhu, and D. Bharadia, "Pointillism: Accurate 3d bounding box estimation with multi-radars," in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, 2020, pp. 340–353.

[10] M. Ulrich, S. Braun, D. Köhler, D. Niederlöhner, F. Faion, C. Gläser, and H. Blume, "Improved orientation estimation and detection with hybrid object detection networks for automotive radar," *arXiv preprint arXiv:2205.02111*, 2022.

[11] D. Solomitckii, C. B. Barneto, M. Turunen, M. Allén, G. P. Zhabko, S. V. Zavjalov, S. V. Volvenko, and M. Valkama, "Millimeter-wave radar scheme with passive reflector for uncontrolled blind urban intersection," *IEEE transactions on vehicular technology*, vol. 70, no. 8, pp. 7335–7346, 2021.

[12] M. E. Russell, A. Crain, A. Curran, R. A. Campbell, C. A. Drubin, and W. F. Miccioli, "Millimeter-wave radar sensor for automotive intelligent cruise control (icc)," *IEEE Transactions on microwave theory and techniques*, vol. 45, no. 12, pp. 2444–2453, 1997.

[13] S. Ingle and M. Phute, "Tesla autopilot: semi autonomous driving, an uptick for future autonomy," *International Research Journal of Engineering and Technology*, vol. 3, no. 9, pp. 369–372, 2016.

[14] P. Nimac, A. Krpič, B. Batagelj, and A. Gams, "Pedestrian traffic light control with crosswalk fmcw radar and group tracking algorithm," *Sensors*, vol. 22, no. 5, p. 1754, 2022.

[15] C. Yan, W. Xu, and J. Liu, "Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle," in *Def Con*, vol. 24, no. 8, 2016, p. 109.

[16] H.-R. Chen and P. Pace, *FMCW radar jamming techniques and analysis*. Monterey, California: Naval Postgraduate School, 2013.

[17] R. Poisel, *Modern communications jamming principles and techniques*. Artech house, 2011.

[18] S. Roome, "Digital radio frequency memory," *Electronics & communication engineering journal*, vol. 2, no. 4, pp. 147–153, 1990.

[19] R. Chauhan, "A platform for false data injection in frequency modulated continuous wave radar," Master's thesis, Utah State University, 2014.

[20] R. Chauhan, R. M. Gerdes, and K. Heaslip, "Demonstration of a false-data injection attack against an fmcw radar," *Embedded Security in Cars (ESCAR)*, 2014.

[21] S. Nashimoto, D. Suzuki, N. Miura, T. Machida, K. Matsuda, and M. Nagata, "Low-cost distance-spoofing attack on fmcw radar and its feasibility study on countermeasure," *Journal of Cryptographic Engineering*, vol. 11, 09 2021.

[22] R. Komissarov and A. Wool, "Spoofing attacks against vehicular fmcw radar," in *Proceedings of the 5th Workshop on Attacks and Solutions in Hardware Security*, 2021, pp. 91–97.

[23] Z. Sun, S. Balakrishnan, L. Su, A. Bhuyan, P. Wang, and C. Qiao, "Who is in control? practical physical layer attack and defense for mmwave-based sensing in autonomous vehicles," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3199–3214, 2021.

[24] P. Nallabolu and C. Li, "A frequency-domain spoofing attack on fmcw radars and its mitigation technique based on a hybrid-chirp waveform," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 11, pp. 5086–5098, 2021.

[25] R. Reddy Vennam, I. K. Jain, K. Bansal, J. Orozco, P. Shukla, A. Ranganathan, and D. Bharadia, "mmspoof: Resilient spoofing of automotive millimeter-wave radars using reflect array," in *2023 IEEE Symposium on Security and Privacy (SP)*, 2023, pp. 1807–1821.

[26] X. Chen, Z. Li, B. Chen, Y. Zhu, C. X. Lu, Z. Peng, F. Lin, W. Xu, K. Ren, and C. Qiao, "Metawave: Attacking mmwave sensing with meta-material-enhanced tags," *Proceedings 2023 Network and Distributed System Security Symposium*, 2023.

[27] M. Ordean and F. D. Garcia, "Millimeter-wave automotive radar spoofing," 2022.

[28] A. Lazaro, A. Porcel, M. Lazaro, R. Villarino, and D. Girbau, "Spoofing attacks on fmcw radars with low-cost backscatter tags," *Sensors*, vol. 22, no. 6, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/6/2145

[29] Y. Zhu, C. Miao, H. Xue, Y. Yu, L. Su, and C. Qiao, "Malicious attacks against multi-sensor fusion in autonomous driving," in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, ser. ACM MobiCom '24. Association for Computing Machinery, 2024, p. 436–451.

[30] R. G. Dutta, X. Guo, T. Zhang, K. Kwiat, C. Kamhoua, L. Njilla, and Y. Jin, "Estimation of safe sensor measurements of autonomous system under attack," in *2017 54th ACM/EDAC/IEEE Design Automation Conference (DAC)*, 2017, pp. 1–6.

[31] Y. Qiu, J. Zhang, T. Sun, Y. Chen, J. Zhang, and B. Ji, "Waston: Inferring critical information to enable spoofing attacks using cots mmwave radar," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–14, 2024.

[32] Texas Instruments Incorporated, "AWR1843, Single-chip 76-GHz to 81-GHz automotive radar sensor integrating DSP, MCU and radar accelerator," 2020, https://www.ti.com/product/AWR1843.

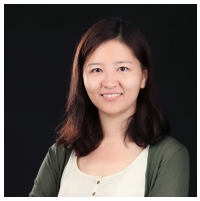
[33] Texas Instruments Incorporated, "DCA1000EVM, Real-time data-capture adapter for radar sensing evaluation module," 2020, https://www.ti.com/tool/DCA1000EVM.

Kaiyi Huang received his B.E. degree in the Department of Computer Science and Engineering in 2023. He is currently a master student in Computer Science and Engineering, at Southern University of Science and Engineering. His research interests include wireless sensing and security.

Shengding Liu will receive his B.E. degree in the Department of Computer Science and Engineering from Southern University of Science and Technology in 2025. His research interests include mmwave sensing and security and UWB localization.

Yanlong Qiu received his B.E. degree in the Department of Electrical and Electronic Engineering (EEE) from the Southern University of Science and Technology in 2017. He received his Ph.D. in Computer and Information Science at Temple University, USA, and in Computer Science and Engineering at Southern University of Science and Technology, China. His research interests include wireless sensing and security.

Yanjiao Chen received her B.E. degree in Electronic Engineering from Tsinghua University in 2010 and Ph.D. degree in Computer Science and Engineering from Hong Kong University of Science and Technology in 2015. She is currently a Bairen researcher at the College of Electrical Engineering, Zhejiang University, China. Her research interests include computer networks, network security, and the Internet of Things.



Jin Zhang is currently an associate professor with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen. She received her B.E. and M.E. degrees in electronic engineering from Tsinghua University, Beijing, in 2004 and 2006 respectively, and received her Ph.D. degree in computer science from Hong Kong University of Science and Technology, Hong Kong, in 2009. She was then employed in HKUST as a research assistant professor. Her research interests are mainly in mobile healthcare and wearable computing, wireless communication and networks, network economics, cognitive radio networks, and dynamic spectrum management. She has published more than 70 papers in top-level journals and conferences. She is the principal investigator of several research projects funded by the National Natural Science Foundation of China, the Hong Kong Research Grants Council, and the Hong Kong Innovation and Technology Commission.