

Adversarial Attack on Radar-based Environment Perception Systems

Amira Guesmi^{1,2} and Ihsen Alouani^{1,3}

¹IEMN CNRS-UMR 8520, Université Polytechnique Hauts-de-France

²eBrain Laboratory, New York University Abu Dhabi, UAE

³CSIT, Queen's University Belfast, UK

Abstract

Due to their robustness to degraded capturing conditions, radars are widely used for environment perception, which is a critical task in applications like autonomous vehicles. More specifically, Ultra-Wide Band (UWB) radars are particularly efficient for short range settings as they carry rich information on the environment. Recent UWB-based systems rely on Machine Learning (ML) to exploit the rich signature of these sensors. However, ML classifiers are susceptible to adversarial examples, which are created from raw data to fool the classifier such that it assigns the input to the wrong class. These attacks represent a serious threat to systems integrity, especially for safety-critical applications. Several adversarial attacks have been developed during the recent years targeting different application domains such as computer vision, speech recognition, healthcare, etc. While these works highlighted the vulnerability of ML systems to adversarial noise, few of the underlying attack scenarios are practical in real-life.

In this work, we present a new adversarial attack on UWB radars in which an adversary injects adversarial radio noise in the wireless channel to cause an obstacle recognition failure. First, based on signals collected in real-life environment, we show that conventional attacks fail to generate robust noise under realistic conditions. We propose a-RNA, i.e., Adversarial Radio Noise Attack to overcome these issues. Specifically, a-RNA generates an adversarial noise that is efficient without synchronization between the input signal and the noise. Moreover, a-RNA generated noise is, by-design, **robust** against pre-processing countermeasures such as filtering-based defenses. Moreover, in addition to the undetectability objective by limiting the noise magnitude budget, a-RNA is also efficient in the presence of sophisticated defenses in the spectral domain by introducing a frequency budget.

We believe this work should alert about potentially critical implementations of adversarial attacks on radar systems that should be taken seriously.

1 Introduction

Intelligent Transportation Systems (ITS) and robotics drove tremendous effort in both industry and research community that led to promising ML-driven environment perception and scene understanding solutions. One of the major challenges in building high-performance environment perception for ITS is designing robust obstacle detection/recognition systems.

Why Radars? – While cameras are the by-default sensors for this task, they are generally limited under poor/degraded capturing conditions such as fog, rain, etc [19]. Therefore, other complementary modalities such as radar technologies have been proposed in enhanced environment perception [15, 19]. In fact, since radars use electromagnetic waves, they are not impacted by lighting or weather conditions and challenges.

Why UWB Radars? – Most radar systems are used to detect the existence, location, and trajectory of objects by analyzing the electromagnetic waves reflected by the environment. However, Ultra-Wide Band (UWB) radars have even higher utility and are particularly efficient for short range settings as they carry richer information on the environment. In fact, UWB radars deliver high-resolution signals that can be exploited to not only detect, but also recognize obstacles. This radar technology transmits very short electromagnetic pulses with low energy in the order of nanoseconds. These initial pulses are received as reflected echo signals with distortions that are directly impacted by the obstacle physical properties, and thereby represent the object signature. In fact, this signature contains information that go beyond the distance and the velocity; it is shaped by the object material, geometry and size [20]. Due to these interesting properties, recent deep learning techniques were applied to exploiting UWB data and achieved promising environment perception results [11, 18, 19].

In spite of Deep Neural Networks (DNNs) outstanding performance, they are shown to be vulnerable to adversarial noise [6–8, 16]. Computer vision is the mainstream application that caught the attention of the community from adversarial machine learning perspective [3, 6, 16]. However,

the vulnerability of ML in other application domains has also been explored. For example, several papers have explored adversarial attacks on automatic speech/speaker recognition [1, 5, 25], Lidar [32] or electrocardiograms [9]. While some of the state-of-the-art work is practical in real-world conditions [1, 14, 25, 30, 31], most of the works require very specific settings to be applicable.

Electromagnetic waves propagate in a broadcasted manner within a wireless channel that may contain a variety of radio signals. This allows a malicious actor to inject noise in the receiver side by propagating a specific adversarial noise using a rogue emitter. This makes the ML-based radio applications, a potential target for adversarial attacks that are practical under real-world conditions [4]. Radar systems are widely used in security-sensitive and safety critical applications, and are also vulnerable to adversarial attacks. More specifically, short range devices such as UWB radar technologies represent a highly practical attack setting because of the possibility of line-of-sight transmission conditions. Few papers in the literature target radar systems [10, 23]. Authors in [10] consider X-band spotlight mode radar which is utilized for hundreds of kilometers range and not practical for injecting adversarial noise due to the channel complexity. [23] targets short-range Frequency-modulated continuous-wave (FMCW) radars and is the closest paper to our work. However, FMCW radars give only velocity and range information, while UWB delivers a complete signature of the obstacle. An more detailed overview on the related work could be found in Section 11.

To our knowledge, this is the first work that proposes adversarial attacks on UWB radar systems. We propose a systematic pipeline to generate robust physical adversarial examples against real-world UWB-based object detectors. Robustness is achieved in three ways:

(i) Shift Robustness. This refers to the robustness against de-synchronization. In fact, while adversarial patches location is not highly influential in some computer vision cases, we surprisingly found that even a minor shift between the initial signal and the adversarial noise crafted with state-of-the-art methods results in a low to no efficiency of the attack. Therefore, we included an aggregation of random noise locations within the noise generation to craft shift-resistant adversarial patches.

(ii) Spectral Domain Robustness. We noticed that the generated shift-resistant adversarial noise has a wide spectral signature; the generated noise contains frequency components that are beyond the expected range of an UWB radar echo. This results in a direct vulnerability against signal pre-processing defenses. To bypass these countermeasures, we clip the generated adversarial noise iteratively along with the noise magnitude budget, to keep the noise in a defined frequency range, and thereby generating shift and **filtering-resistant** noise.

(iii) Undetectability. In addition to the magnitude budget

and the frequency clipping, we also consider a limit on the noise application time, i.e., the size in time-domain of the generated noise. The motivation are behind exploring this property is that a shorter noise in time has lower risks to be observable, and hence detectable. Moreover, even with being in the same frequency domain of the victim device, adversarial noise could be detected by spectrum sensing. Shorter adversarial noise injection periods lead to lower magnitude in the spectral domain, and hence higher chances to be undetected.

The paper will explore each of these three perspectives to illustrate their strengths and limitations before presenting our final solution.

Contributions. In summary, the contributions of this paper are as follows:

- We present a-RNA, the first attack leveraging unique characteristics of Ultra-Wide Band radar signals to inject real-world adversarial noise in a radar-based environment perception system.
- While in simulation, the state-of-the-art attacks show high efficiency, we show that they fail under realistic conditions. First, baseline attacks are neutralized with a slight timing shift, i.e., de-synchronization between the noise and the signal. Moreover, these attacks generate noise outside the spectral range expected from the reflected echos, and can be easily defended against by pre-processing countermeasures.
- Our approach generates adversarial noise that is: **(i) input-agnostic** to be practical in real-time, **(ii) robust to incidence delay**, i.e., to de-synchronized settings, and **(iii) filtering-robust** by tailoring adversarial noise that is in the same spectral signature of the raw signal. Therefore, a-RNA represents a practical threat to UWB-based environment perception systems.
- To further anticipate adaptive defender that uses spectrum sensing to detect a potential adversarial noise, we include a additional constraint on the noise generation mechanism on the time domain patch sizes and show that these adversarial patches are hard to detect because of their low magnitude spectral components.
- We also analyze the robustness of our attack in the presence of adversarially trained networks. While adversarial training reduces the attack efficiency, we show that further measures need to be taken to preserve UWB radar systems integrity.
- We open-source our codes and collected data for the community to encourage further investigations of this direction.

¹omitted for blind review

2 Background

Ultra-wideband (UWB) is a short-range radio communication technology that allows fast and stable data transmission. UWB is generally the technology of choice for localization of moving assets in complex and space-sensitive locations due to its precision, reliability and rich information.

The UWB transmitter emits a narrow pulse at a target's direction, and the reflected signal is detected by the UWB receiver. When a UWB pulse encounters a boundary between two types of medium with different dielectric properties during propagation, a portion of the incident electromagnetic energy is reflected back to the original medium with a reflection angle θ_r (zero reflection angle if the incident wave path is parallel to the normal line), while the other portion propagates through the next medium. The extremely short pulses (usually in order of few nanoseconds) provide a very wide bandwidth, which has numerous advantages, including high throughput, covertness, jamming resistance, lower power, and coexistence with existing radio services [28, 36]. UWB not only has the potential to transmit a rich data over a short distance while using very low power, but it can also pass through physical objects that tend to reflect signals with narrow bandwidth.

UWB radar technology is also employed in the automotive field [19, 34] owing to the richness of information it provides, and its robustness to degraded capturing conditions. The preeminent characteristic of such technology consists in the deformation of the emitted pulse. This distortion depends on the obstacles characteristics, thereby it is labeled as the object signature. This signature is affected by the shape, material and size of the object. For example, the signature of a metallic object has a higher amplitude than that of a pedestrian. Consequently, the use of such technology remains promising for detecting objects at short range. The data acquired from the UWB radar can be represented in two forms: a one-dimensional (1D) signal, which is the reflected echo, and a two-dimensional (2D) data that can be a 2D feature map or a converted image.

Figure 1 shows samples of UWB radar signals with corresponding images for illustration purposes.

3 Threat Model

3.1 Attack scenario

An adversary wants to remotely compromise an UWB-based environment perception system such as an autonomous vehicle by causing an obstacle detection failure. To do so, the adversary corrupts the reflected radar signal by injecting carefully crafted adversarial radio noise in the channel.

Physical setting. We assume the adversary can be in the surrounding environment of the victim device. We assume

that an adversary cannot physically touch the victim's devices, alter the device settings, or install malware apps.

Attacker knowledge. We assume the adversary has access to the model, i.e., a white-box setting. Therefore, the attacker is aware of the victim classifier's parameters and architecture. This information is used by the attacker to construct adversarial examples. However, at inference time, the adversaries have no access to the victim device functional parameters. They have no prior knowledge on when the system starts to send/receive the UWB signals.

Attack equipment. We assume that the adversaries possess wireless equipment such as a USRP and a directional antenna that allows them to broadcast random signal in the channel and specifically in the direction of the victim device.

Figure 2 gives an overview on our threat model in the context of environment perception setting for Intelligent Transportation Systems. An embedded UWB radar is used for obstacle detection/recognition in an autonomous vehicle. The radar signals are fed to a CNN which classifies the obstacles based on received signatures. A malicious actor injects adversarial noise in the channel to compromise the receiver side using a rogue emitter.

3.2 Problem definition

An adversary, using information learnt about the structure of the classifier, tries to craft perturbations added to the input to cause incorrect classification. For illustration, given an original input x and a target classification model $C(\cdot)$, the problem of generating an adversarial example x^* can be formulated as a constrained optimization [35]:

$$\begin{aligned} x^* = \arg \min_{x^*} & \mathcal{D}(x, x^*), \\ \text{s.t. } & C(x^*) = l^*, \quad l \neq l^* \end{aligned} \quad (1)$$

Where \mathcal{D} is a distance metric used to quantify similarity between two inputs (images/signals) and the goal of the optimization is to minimize the added noise, typically to avoid detection of the adversarial perturbations. l and l^* are the two labels of x and x^* , respectively: x^* is considered as an adversarial example if and only if the label of the two inputs are different ($C(x) \neq C(x^*)$) and the added noise is bounded ($\mathcal{D}(x, x^*) < \epsilon$ where $\epsilon \geq 0$).

4 Experimental Setup

4.1 Data Collection

We collected UWB radar signals in real-world conditions, i.e., an outdoor environment within a University campus in which we have a variety of classes and scenarios. The radar considered in the dataset is an UWB radar developed by the UMAIN Inc company [33] named HST-D3 with an efficient range of 6 – 10 meters and a frequency range in 3 – 4 GHz. The waveform of the generated pulse by the UWB radar is the



Figure 1: Illustration of different radar signatures of different classes (with the corresponding scenes).

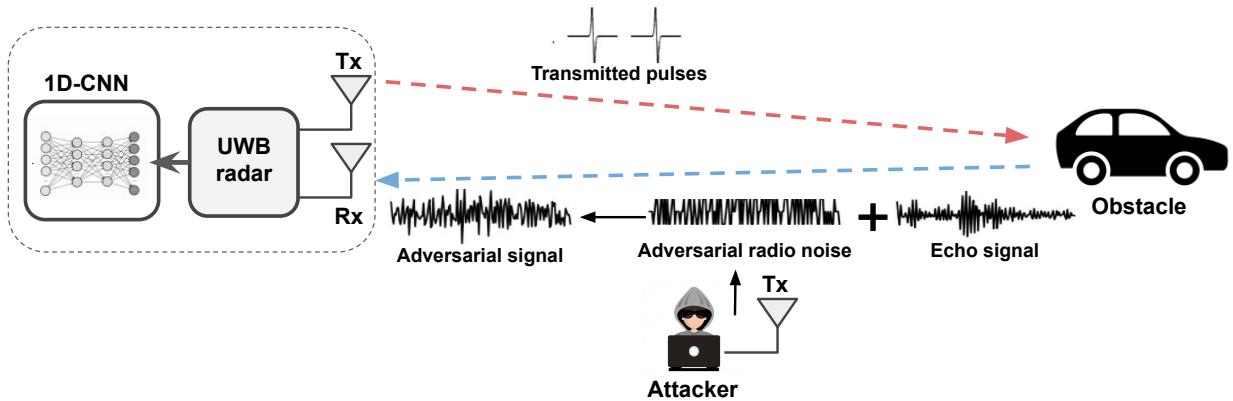


Figure 2: Illustration of the threat model.

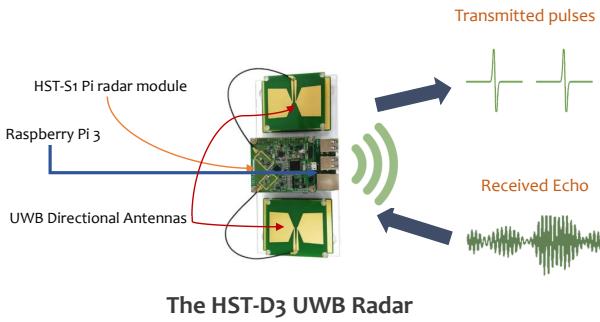


Figure 3: An overview on the UWB radar system.

1st order differential Gaussian pulse. We exploit the available directional antenna since it guarantees better target echo-to-clutter and noise ratio. The HST-D3 radar is a combination of HST-S1 Pi module radar and a Raspberry Pi 3 for the acquisition. The hardware connection is presented in Figure 3 and Table 1 shows the radar specifications.

The dataset was captured during 3 months under different weather conditions and corresponds to 4 different classes chosen for their representativity of an urban transportation environment, namely: Car, Pedestrian, Cyclist and Tramway.

Table 1: Umain radar specifications

Parameter	Value and comments
Frequency range	3 ~ 4 Ghz
Output Power	Typ. -25dBm
Antenna Specification	UWB Directional Antenna : Gain = Avg.7 dBi Antenna angle (@-3dB) = 56°(X-Z plane) 77.5°(Y-Z plane) Size = 76mm x 58.5mm x 17mm
Sampling	660 samples per frame
Sampling frequency	7.69Ghz

4.2 UWB-based System for Obstacle Detection

Deep learning methods have shown great potential in solving complex problems due to their ability to automatically learn features through multiple levels of abstraction, which frees the designer from the dependence on hand-engineered features. Particularly, Convolution Neural Networks (CNNs) are very efficient in projecting raw information to learnt features' space, which represents the input space of a classifier. More specifically, 1D-CNNs can be used to extract local 1D patches (subsequences) from uni-dimensional data samples. For obstacle recognition based on UWB signals, we consider a 1D-CNN based on 3 convolutional layers with filters of size

3x1 and 2 fully connected layers. LeakyReLU, a variant of the linear rectifier function (ReLU) is used as activation function. The detailed architecture hyper-parameters are presented in Table 2.

Table 2: Architecture of the 1D-CNN.

Layer type	Unit	Output shape	# of Parameters
Conv1D (LReLU)	(16,3)	(16, 658)	64
Maxpool	(2)	(16, 329)	0
Conv1D (LReLU)	(32,3)	(32, 327)	1,568
Maxpool	(2)	(32, 163)	0
Conv1D (LReLU)	(64,3)	(64, 161)	6,208
Maxpool	(2)	(64, 80)	0
Flatten	-	-	5,120
Linear (LReLU)	16	-	81,936
Linear	5	-	85
Total parameters	-	-	89,861

Training and validation samples are split in 7:3 way which gives 800 training examples and 294 test examples. We use a learning rate of $1e^{-4}$ and a batch size of 4 for the training. The Adam optimizer ($\beta_1 = 0.9$ and $\beta_2 = 0.999$) was used with the negative log likelihood loss function. In Table 3, we present the results of classification accuracy of the model after training for 100 epochs.

Table 3: 1D-CNN classification accuracy.

Classification Accuracy	99.31%

Once we have a system with high utility, in the next sections we evaluate the vulnerability of such system to adversarial noise under realistic conditions.

5 Baseline Adversarial Attack on UWB systems

In this section we proceed to a preliminary analysis of UWB-based environment perception systems vulnerability to state-of-the-art adversarial attacks. For this reason, we consider two situations:

(i) Input-specific adversarial attacks, where the adversarial noise is generated to target a specific sample.

(ii) Universal noise that tries to alter the model output regardless of the specific sample.

We first present the attack generation methods, then show their corresponding results, and discuss their limits.

5.1 Input-Specific attacks

In this section, we build a baseline input-specific attack, where we use two state-of-the-art white-box methods, i.e., Fast Gradient Sign Method and Projected gradient descent attacks to

generate adversarial examples against the UWB-based system.

Fast Gradient Sign Method (FGSM): FGSM [6] is a single-step, gradient-based, attack. An adversarial example is generated by performing a one step gradient update along the direction of the sign of gradient at each element of the signal as follows:

$$x^* = x + \epsilon \text{sign}(\nabla_x J_\theta(x, y)) \quad (2)$$

Where $\nabla J()$ computes the gradient of the loss function J and θ is the set of model parameters. The $\text{sign}()$ denotes the sign function and ϵ is the perturbation magnitude.

Projected gradient descent (PGD): PGD [16] is a stronger iterative method where the adversarial example is generated as follows:

$$x^{t+1} = \mathcal{P}_{S_x}(x^t + \alpha \cdot \text{sign}(\nabla_x \mathcal{L}_\theta(x^t, y))) \quad (3)$$

Where $\mathcal{P}_{S_x}()$ is a projection operator projecting the input into the feasible region S_x and α is the added noise at each iteration. PGD tries to find the perturbation that maximizes the loss of a model on a particular input while keeping the size of the perturbation smaller than a specified amount.

5.2 Input-Agnostic attacks

We also consider the universal adversarial perturbations (UAP) [21]. UAP has been initially proposed against computer vision systems; it generates an input-agnostic adversarial patch after optimizing over a given dataset. Let $x \in \mathbb{R}^d$ be an input of dimension d that follows a distribution μ ($x \sim \mu$). The main objective of a UAP is to fool a target model $C(\cdot)$ on almost all inputs sampled from μ . This problem can be formulated as finding a vector δ such that:

$$C(x + \delta) \neq C(x), \text{ for "most" } x \sim \mu \quad (4)$$

Where δ represents the adversarial patch and must satisfy the following two constraints:

- $\|\delta\|_p \leq \xi$
- $\mathbb{P}_{x \sim \mu}(C(x + \delta) \neq C(x)) \geq 1 - \rho$

The parameter ξ controls the magnitude of the perturbation vector δ , and ρ quantifies the desired fooling rate for all images sampled from the distribution μ .

5.3 Results

We use l_∞ -norm as a distance metric of the noise generation. The attack success rate (defined as 1 - Classification Accuracy) represents the proportion of total perturbed signals for which the adversarial noise forces the model to output a

wrong label. Table 4 and Table 5 show the success rate of the input-specific and the UAP, respectively. As expected, these methods adapted to the UWB signals are able to generate effective adversarial examples. In the next section we discuss to which extent these results hold under realistic conditions.

Table 4: Attack success rate of baseline attacks.

Attack	Epsilon				
	0.001	0.002	0.005	0.007	0.01
PGD	50%	87%	96%	98%	98%
FGSM	64%	93%	98%	98%	98%

Table 5: Attack success rate of UAP.

Attack	Epsilon				
	0.01	0.02	0.03	0.04	0.05
UAP PGD-based	57%	60%	70%	82%	92%
UAP FGSM-based	69%	89%	90%	90%	90%

5.4 Limits

While the previous results show the vulnerability of UWB-based ML systems to adversarial attacks, the attack assumptions do not take into account the specificity of the application, nor the real-life conditions. First, the input-specific attacks are not practical unless they are generated on-the-fly. Moreover, the adversarial noise has been applied under a perfect synchronization with the input samples, which is not practical in real-world scenarios.

In this section, we investigate the impact of de-synchronization on adversarial attacks success rate. We apply random *time shifts* to the noise incidence, which correspond to the delays in the adversarial noise incidence at receiver side. The results are depicted in Figure 4; we notice a huge drop in attacks effectiveness. In fact, the baseline attacks with small noise magnitudes are totally neutralized if not synchronized with the input. For instance, we noticed a drop from 87% to 1% for a noise magnitude $epsilon = 0.002$ and for higher noise budget constraints, we notice a drop of at least 61% in success rate.

The same trend has been shown by UAP as shown in Figure 5; we notice considerably lower attack success rates. With a noise budget equal to 0.05 the attack is 36% less efficient. These attacks require perfect synchronization in order to be effective which is not practical since usually it's extremely hard for the attacker to predict the exact time the system starts to receive the signal.

Terminology. Henceforth, we consider universal adversarial noise only, and we use adversarial *patch* and adversarial *noise* interchangeably.

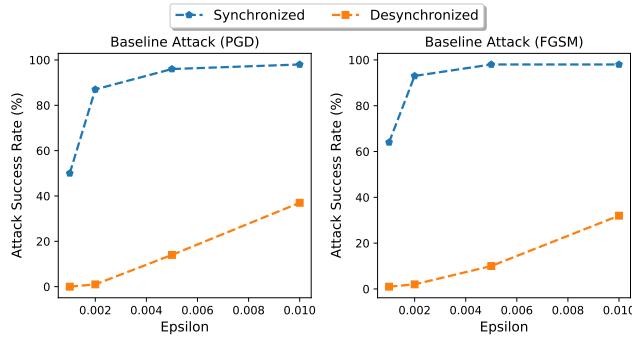


Figure 4: Attack success rate of baseline attacks under synchronized and desynchronized setting.

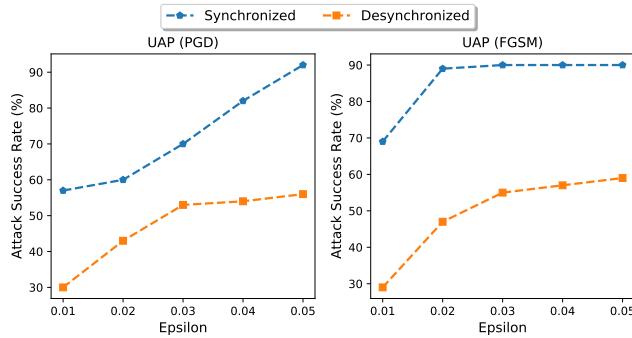


Figure 5: Attack success rate of UAP under synchronized and desynchronized adversarial settings.

6 Shift Resistant Patch (SRP)

In this section, we propose Shift Resistant Patch (SRP), an adversarial adversarial patch generation methods that overcomes desynchronization with the target sample.

6.1 Approach

As previously shown, a slight timing shift that desynchronizes the adversarial noise with the input neutralizes the attack. To overcome this limitation, we propose to generate adversarial noise that can efficiently fool the classifier regardless of the noise incidence delay. For this reason, we include the noise incidence time within the optimization problem. Specifically, we iteratively update the adversarial noise by aggregating its impact over different incidence times.

Our main objective is to fool a victim model $C(\cdot)$ on almost all the input signals from a distribution μ in \mathbb{R}^d whenever the adversarial noise incidence time within the victim signal. This problem can be formulated as finding a vector δ such that:

$$C(x + \text{shift}(\delta, k)) \neq C(x), \text{ for most } x \sim \mu \quad (5) \\ \text{s.t. } \|\delta\|_p \leq \varepsilon$$

Where the parameter ε controls the magnitude of the noise vector δ and $\text{shift}(\cdot)$ is a function that quantifies the adversarial patch signal δ , relatively with regard to a target signal x given a time incidence $k \in [0, d]$. Given an adversarial patch $\delta = \{\delta_j\}_{j \in [0, d]}$ that is *repeatedly broadcasted in a continuous loop* within the channel, the function $\text{shift}(\cdot)$ could be expressed as:

$$\text{shift}(\delta, k) = \begin{cases} \delta(d - k + j) & \text{if } j \in [0, k] \\ \delta(j - k) & \text{else} \end{cases} \quad (6)$$

In Algorithm 1, we iterate across the data in the batch X gradually updating the adversarial patch. For each input example we generate the corresponding adversarial noise using *attack()* function for a randomly chosen locations k . The function *attack()* is detailed in Algorithm 2; it is a gradient-based attack that performs m steps along the direction of the sign of gradient at each element of the signal in a way that maximizes the loss of a model. Therefore, to update the patch every iteration, we use $\nabla_{x+\text{shift}(\delta, k)}(\cdot)$ instead of using $\nabla_{x+\delta}(\cdot)$. If the generated adversarial example x_i^{adv} fools the model, we undo the shifting and rearrange the adversarial noise back to

Algorithm 1 Shift Resistant Patch (SRP).

```

1: Input: Data batch:  $X$ , classifier:  $C$ , noise magnitude:  $\epsilon$ , step size:  $\alpha$ , number of iterations:  $N$ , raw signal size:  $d$ .
2: Output:  $\delta$  adversarial patch.
3: Initialize  $\delta \leftarrow 0$ 
4: while  $iter < N$  do
5:   for each data point  $x_i \in X$  do
6:     if  $C(x_i + shift(\delta, k)) == C(x_i)$  then
7:       Select a random location  $k$ 
8:        $x_i^{adv} \leftarrow attack(C, x_i + shift(\delta, k), y_i, \epsilon, \alpha)$ 
9:       if  $C(x_i) != C(x_i^{adv})$  then
10:         $\Delta\delta_i \leftarrow concat((x_i^{adv} - x_i)[k : d], (x_i^{adv} - x_i)[0 : k])$ 
11:         $\delta \leftarrow \mathcal{P}_{l_p, \epsilon}(\delta + \Delta\delta_i)$ 
12:      end if
13:    end if
14:   end for
15:    $iter += 1$ 
16: end while
17: function  $shift(\delta, k)$ 
18:   Initialize  $shifted\_delta \leftarrow 0$ 
19:    $shifted\_delta \leftarrow concat(\delta[d - k : d], \delta[0 : d - k])$ 
20: end function

```

Algorithm 2 $attack$ function.

```

1: Input: a classifier:  $C$  with loss  $J$ , noise budget:  $\epsilon$ , step size:  $\alpha$ , input signal:  $x$ , label:  $y$ , number of iterations:  $m$ .
2: Output:  $x^{adv}$ 
3: Initialize  $x^{adv} \leftarrow 0$ 
4: for  $i = 0 \dots m-1$  do
5:    $x_{i+1}^{adv} = Clip\{x_i + \alpha sign(\nabla_{x_i}^{adv} J_\theta(C(x_i^{adv}), y))\}$ 
6: end for

```

the original form (before performing the *shift*). This noise is aggregated to the current instance of the patch and than projected on the L_p norm ball of size ϵ .

6.2 Results

In this section, we evaluate SRP under random noise incidence delays. We generate adversarial noise using Algorithm 1, and evaluate the attack success rate for different noise magnitudes. The evaluation is performed comparatively with to UAP under random incidence time of the adversarial noise.

As shown in Figure 6, SRP is able to recover the adversarial attack effectiveness. For instance SRP has 30% higher success rate than UAP for a noise budget equal to 0.05.

6.3 Limits

While the proposed attack addresses the synchronization limit, the threat model assumes a baseline victim device without

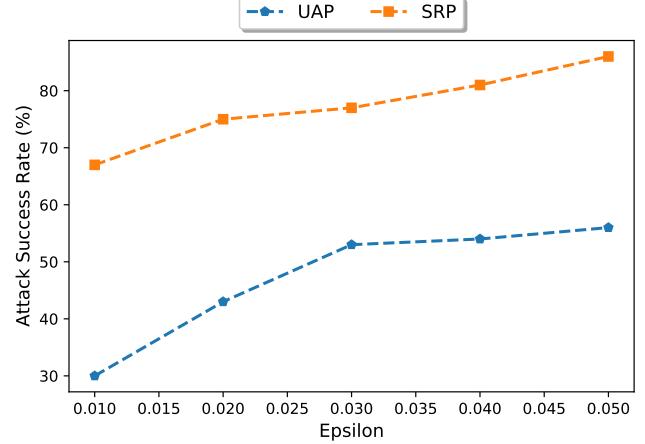


Figure 6: Attack success rate of proposed SRP compared to UAP.

any countermeasures. However, one of the straightforward protection of such systems is to pre-process the input to filter out undesired signals from non relevant frequency ranges.

6.3.1 Impact of a pre-processing defense

In this section, we consider a system that is comprehensively pre-processing the input signal. The defender defines its target frequency range based on the expected echos from the environment. We evaluate the efficiency of SRP under this setting for different noise magnitudes.

The frequency spectrum is obtained using the Fast Fourier Transform (FFT), which is defined by

$$F(k) = \sum_{n=0}^{N-1} f(n) e^{-i \frac{2\pi}{N} nk} \quad (7)$$

Where N is the length of the spectral signature and f is the original time domain signal.

For each window, an FFT generates a frequency domain representation of the signal referred to as magnitude spectrum. The magnitude spectrum details each frequency and the corresponding intensity that make up a signal.

The Inverse Fast Fourier Transform (IFFT) converts frequency domain signal to time domain signal and is expressed as follows

$$f(n) = \frac{1}{N} \times \sum_{k=0}^{N-1} F(k) e^{i \frac{2\pi}{N} nk} \quad (8)$$

Where $F(k)$ is the frequency domain magnitudes and $f(n)$ is the recovered time domain samples.

We first explore the power spectral density of the raw data to identify the frequency range that represents the region of interest. Specifically, we identify the average minimum and maximum frequencies ($[f_{min}, f_{max}]$) that contain 95% of the

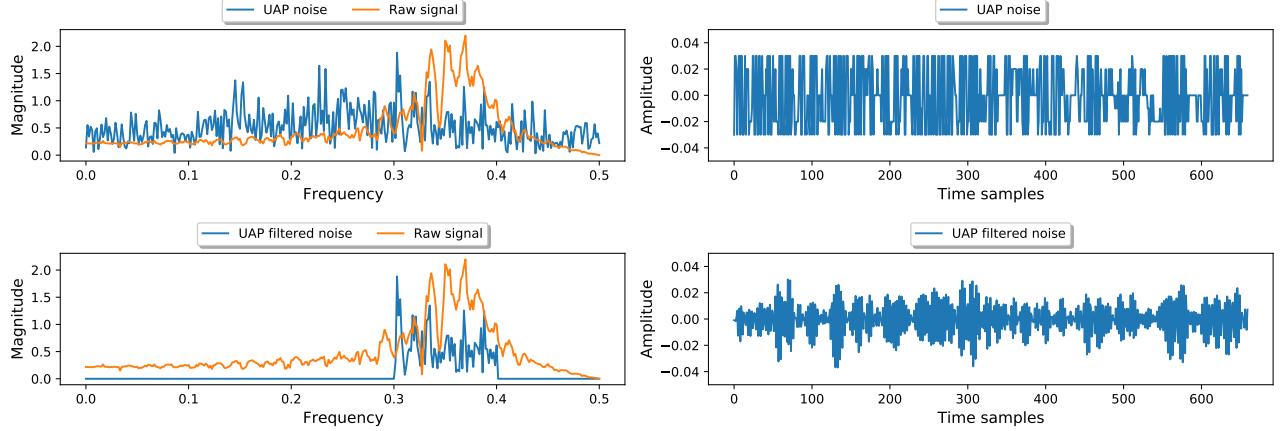


Figure 7: UAP adversarial noise before and after filtering in the spectral domain (left) and the time domain (right).

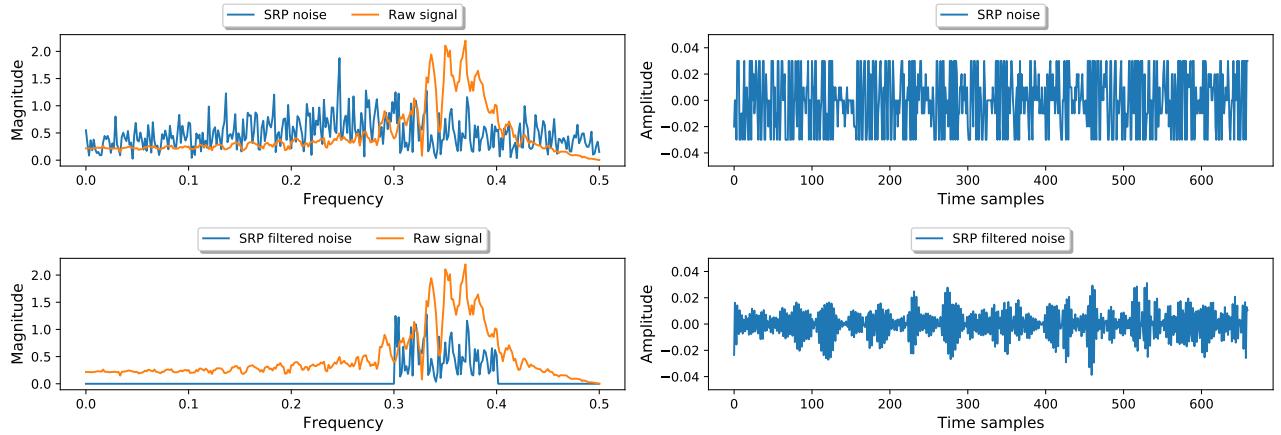


Figure 8: SRP adversarial noise before and after filtering in the spectral domain (left) and the time domain (right).

total spectral power. Accordingly, we propose to use a pass-band filter corresponding to this frequency range, and make sure that there is no baseline accuracy drop is noticed under this pre-processing.

Algorithm 3 gives an overview on the pre-processing procedure. We first transfer the signal to the spectral domain using the FFT. Then we apply a low-pass filter using f_{min} as a cut-off frequency, followed by a high-pass using f_{max} as cut-off. To finish, we recover the time domain signal using the IFFT. The resulting procedure represents a clipping operation in the spectral domain that we will use later in Section 7 for an adaptive attack.

Figures 7 and 8 illustrate the frequency spectrum of the UAP and SRP-generated adversarial noise compared to the raw signal. Since the generation process of SRP has no restriction on the frequency, we notice a wide spectral signature, i.e., the generated noise contains frequency components that are beyond the expected range of an original raw signal which makes it vulnerable to filtering-based defenses.

Figures 7 and 8 also show the impact of the filter on the

time domain signal for UAP and SRP-generated noise. While the filtered noise signals are significantly changed, we confirm this by studying their adversarial impact post-filtering.

Algorithm 3 filter: Pass-band Filter.

- 1: **Input:** time domain signal: x , Min Frequency: f_{min} , Max Frequency: f_{max}
 - 2: **Output:** $filtered_x$
 - 3: Initialize $filtered_x \leftarrow 0$
 - 4: $X \leftarrow FFT(x)$
 - 5: $X \leftarrow LowPass(X, f_{min})$
 - 6: $X \leftarrow HighPass(X, f_{max})$
 - 7: $filtered_x \leftarrow IFFT(X)$
-

Figure 9 presents the attack success rate of UAP and SRP. As a matter of fact, the filter significantly degrades the performance of both attack methods. For UAP, we notice a drop of 34% for a noise budget of 0.05, while SRP-generated noise (Algorithm 1) is 20% less effective when limiting the fre-

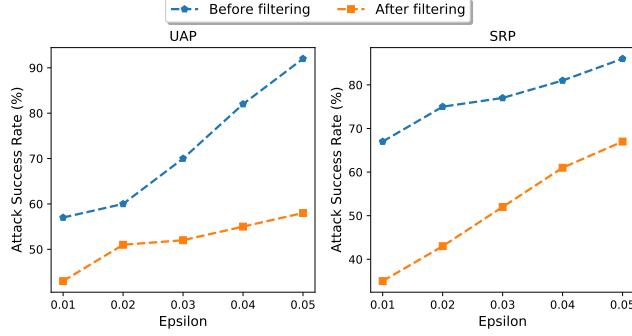


Figure 9: Attack success rate of UAP and SRP under passband filter.

quency range.

The next section proposes an adaptive attack to this defense mechanism.

7 Shift & Filtering Resistant Patch (SFR)

In this section we propose SFR (Shift & Filtering Resistant Patch) to generate adversarial noise that is robust against incidence delay and additionally adaptive to filtering defense.

7.1 Approach

The objective of SFR is to generate adversarial noise in a specific frequency range which corresponds to the expected range of the radar echoes. Consequently, the defender will not be able to cut-off the impact of the injected noise unless with a loss of utility. For this reason, SFR not only clips the noise magnitude in the time domain to fit a noise budget, but also adds a new constraint on the noise generation in the spectral domain. In fact, SFR also projects the noise in a subset of the frequency domain to restrain its spectral components. The problem can therefore be formulated as follows:

$$\begin{aligned} C(x + \text{shift}(\delta, k)) &\neq C(x), \text{ for most } x \sim \mu \text{ s.t.} \\ \|\delta\|_p &\leq \epsilon \\ \text{FFT}(\delta) &\in [f_{min}, f_{max}] \end{aligned} \quad (9)$$

Where ϵ controls the magnitude of the noise vector δ and $\text{shift}(\cdot)$ is a function expressed by Equation 6 that quantifies the adversarial patch signal δ , relatively with regard to a target signal x given a time incidence $k \in [0, d]$. The new constraint on δ limits the noise frequency components to an acceptable range defined by f_{min} and f_{max} .

Algorithm 4 details the noise generation mechanism. In SFR, we include the filtering step within the adversarial noise generation procedure. For each iteration, after updating the noise using $\text{attack}()$ function, we project the noise back to the target spectral domain using the pass-band filter defined previously. Therefore, we only retain noise samples with the

Algorithm 4 Shift & Filtering Resistant Patch (SFR).

```

1: Input: Data batch:  $X$ , classifier:  $C$ , noise magnitude:  $\epsilon$ , number of iterations:  $N$ , min frequency:  $f_{min}$ , max frequency:  $f_{max}$ .
2: Output:  $\delta$  trained patch.
3: Initialize  $\delta \leftarrow 0$ 
4: while  $iter < N$  do
5:   for each data point  $x_i \in X$  do
6:     if  $C(x_i + \text{shift}(\delta, k)) = C(x_i)$  then
7:       Select a random location  $k$ 
8:        $x_i^{adv} \leftarrow \text{attack}(C, x_i + \text{shift}(\delta, k), y_i, \epsilon, \alpha)$ 
9:       if  $C(x_i) \neq C(x_i^{adv})$  then
10:         $\Delta\delta_i \leftarrow \text{concat}((x_i^{adv} - x_i)[k : d], (x_i^{adv} - x_i)[0 : k])$ 
11:         $\delta \leftarrow \delta + \Delta\delta_i$ 
12:         $\delta \leftarrow \text{filter}(\delta, f_{min}, f_{max})$ 
13:         $\delta \leftarrow \mathcal{P}_{l_p, \epsilon}(\delta)$ 
14:      end if
15:    end if
16:   end for
17:    $iter += 1$ 
18: end while

```

same frequency range as the raw signal, these samples are then aggregated to the current patch instance.

7.2 Results

Figure 10 gives an illustration of the filtering impact on a SFR-generated noise both in the frequency and time domains. In contrast with the previous results, the illustration shows that SFR-generated noise is not substantially impacted by the filter.

Figure 11 shows the attack success rate of SFR for different noise budgets comparatively with the baseline UAP. The results show that SFR can efficiently bypass a defender that uses a filtering stage. For instance, SFR success rate reaches 84% for 0.05 noise budget while being input-agnostic and incidence delay-resistant, while UAP shows limited efficiency.

7.3 Limits

While SFR jointly addresses the robustness to incidence time and to filtering defense, its construction opens a new weakness that can be used to detect the adversarial noise. In fact, while generating adversarial noise in the same frequency range of the victim signal bypasses any filter-based pre-processing, it makes the adversarial noise broadcasting detectable by spectrum sensing techniques. In fact, cognitive radars can use different techniques such as energy detection to check for the channel availability. While a baseline noise has a wide spectrum and hence a more sparse spectral power distribution, SFR concentrates the spectral components in a relatively

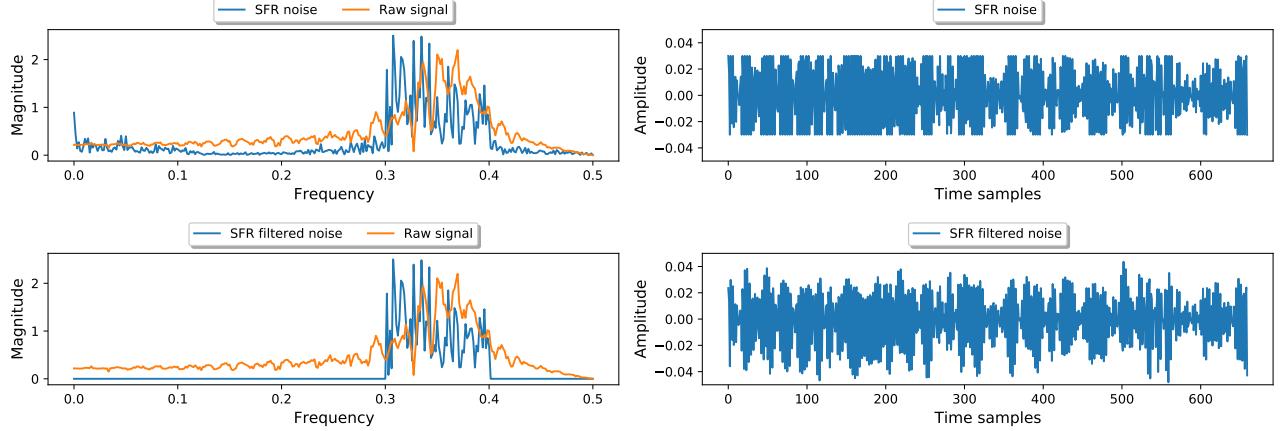


Figure 10: Frequency spectrum of SFR adversarial noise.

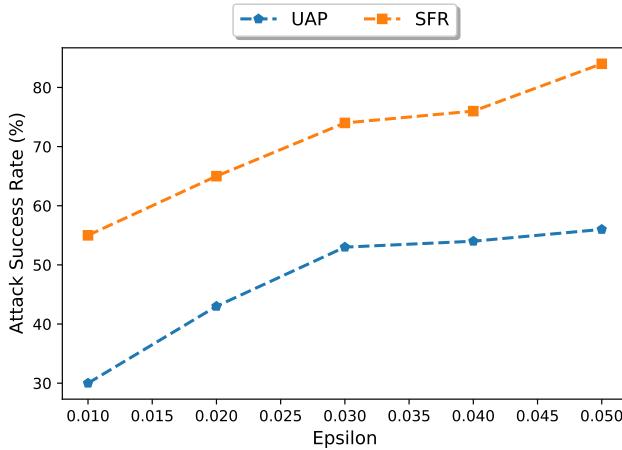


Figure 11: Attack success rate of proposed SFR compared to UAP.

narrow band, which makes it detectable by spectrum sensing.

The next section presents the a-RNA, which keeps all the previous characteristics of SFR while attempting to evade spectrum sensing.

8 Adversarial Radio Noise Attack (a-RNA)

In this section, we propose a-RNA which is an attack that satisfies the robustness to the previously detailed challenges, i.e., noise incidence delay and filtering, while overcoming SFR limits.

8.1 Approach

The weakness of SFR is being potentially observable in the spectral domain through idle spectrum sensing. The direct reason of this observability consists of the constrained optimization approach that forces the spectral components in

a specific range, i.e., the victim operating range. However, the magnitude of the adversarial noise spectral components depend also on the time window for which the noise is broadcasted. For this reason, we propose to add a new constraint on the noise generation in order to limit the magnitude of the frequency components. Specifically, we introduce a time domain window budget that limits the size of the adversarial noise. Therefore, a-RNA attempts to generate an adversarial patch with smaller size rather than using a full size patch to evade detection. The problem can be formulated as follows:

$$\begin{aligned}
 C(x + shift(\delta, k)) &\neq C(x), \text{ for most } x \sim \mu \text{ s.t.} \\
 \|\delta\|_p &\leq \epsilon \\
 FFT(\delta) &\in [f_{min}, f_{max}] \\
 size(\delta) &\leq \sigma
 \end{aligned} \tag{10}$$

Where ϵ controls the magnitude of the noise vector δ and $shift(\cdot)$ is a function expressed by Equation 6 that quantifies the adversarial patch signal δ , relatively with regard to a target signal x given a time incidence $k \in [0, d]$. The noise δ has frequency components constrained in a range defined by f_{min} and f_{max} . The new constraint on δ is σ which represents the budget in terms of noise size in time domain.

Algorithm 6 details the noise generation procedure to solve the optimization problem in Equation 10. In this algorithm we introduce a temporal mask that extracts a signal that includes δ with a position k relative to the input signal x_i . The signal is padded with trailing zeros to equal the size of the original raw signal. Therefore, the noise is active in a time window of size $s \leq \sigma$ and null elsewhere.

The function $mask(\cdot)$ could be expressed as:

$$mask(\delta, k) = \begin{cases} \delta(j - k) & \text{if } j \in [k, k + s] \\ 0 & \text{else} \end{cases}$$

Therefore, $mask()$ function jointly implements the time shift and the size constraint on the noise. This signal is added

Algorithm 5 Adversarial Radio Noise Attack (a-RNA).

```

1: Input: Data points:  $X$ , classifier:  $C$ , noise magnitude:  $\epsilon$ , number of iterations:  $N$ , patch size:  $s$ , raw signal size:  $d$ .
2: Output:  $\delta$  trained patch.
3:
4: Initialize  $\delta \leftarrow 0$ 
5: while  $iter < N$  do
6:   for each data point  $x_i \in X$  do
7:     if  $C(x_i + mask(\delta, k)) = C(x_i)$  then
8:       Select a random location  $k$ 
9:        $x_i^{adv} \leftarrow attack(C, x_i + mask(\delta, k), y_i, \epsilon, \alpha)$ 
10:      if  $C(x_i) \neq C(x_i^{adv})$  then
11:         $\Delta\delta_i \leftarrow (x_i^{adv} - x_i)[k : k + s]$ 
12:         $\delta \leftarrow \delta + \Delta\delta_i$ 
13:         $\delta \leftarrow filter(\delta, f_{min}, f_{max})$ 
14:         $\delta \leftarrow P_{l_p, \epsilon}(\delta)$ 
15:      end if
16:    end if
17:   end for
18:    $iter += 1$ 
19: end while
20: function  $mask(\delta, k)$ 
21:   if  $k \leq d - s$  then
22:      $Mask = Padding(\delta, (k, d - (k + s)), (0, 0))$ 
23:   else
24:      $Mask = Append(zeros(k), \delta[0 : d - k])$ 
25:   end if
26: end function

```

to the input signal and fed to the $attack()$ function. We modify $attack()$ to only use the gradient sign of the $[k, k + s]$ signal portion to update the adversarial noise.

Algorithm 6 $attack$ function.

```

1: Input: a classifier  $C$  with loss  $J$ , noise budget  $\epsilon$ , step size  $\alpha$ ,  $x$  input image,  $y$  label,  $m$  number of iterations,  $k$  random position,  $s$  noise size.
2: Output:  $x^{adv}$ 
3: Initialize  $x^{adv} \leftarrow 0$ 
4: for  $i = 0 \dots m-1$  do
5:    $x_{i+1}^{adv} = Clip\{x_i + \alpha sign(\nabla_{x_i}^{adv} J_\theta(C(x_i^{adv}), y)) \times mask(ones(s), k)\}$ 
6: end for

```

8.2 Results

In this section, we evaluate a-RNA from different perspectives:

Impact of patch size on magnitude spectrum. We first investigate the impact of the patch size budget on the magnitude spectrum, which directly implies its detectability using spectrum sensing. The magnitude indicates the strength of the

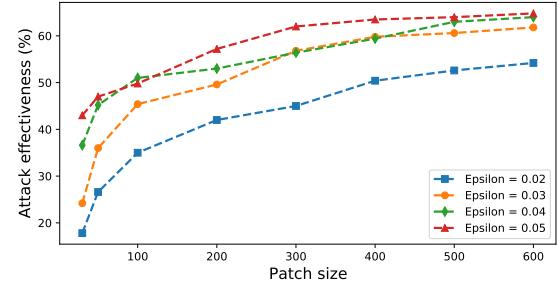


Figure 12: Impact of patch size on patch effectiveness under different noise magnitude constraints.

frequency components relative to other components. In Table 6, we report the average magnitude of the frequency spectrum of each generated patch. We use a-RNA to generate patches with a noise constraint equal to 0.03. Notice that for a size budget of 600 (full size), the attack corresponds to SFR case.

Adversarial noises with shorter injection periods have lower magnitudes in the spectral domain, and hence lower chances to be detected. A patch with a size $s = 400$ has a $2 \times$ higher magnitude than a patch with a size $s = 100$. A short patch with a size of 50 samples has around $4 \times$ lower average magnitude and $7 \times$ lower maximum magnitude than a SFR patch.

Impact of Patch size on attack efficiency. We vary the patch size s from 30 to 600 and investigate its impact on model accuracy. We use algorithm 5 to train the patches. For each test sample, only one patch is diffused at a random time with different noise magnitude ($epsilon$). We notice that the bigger the size of the patch, the more powerful the patch (see Figure 12).

Impact of random noise In this section we compare our proposed technique a-RNA to injecting random white noise in the wireless channel to cause miss-classification of obstacles. We use patches of Gaussian white noise with the same magnitude and with different sizes varying from 50 to 600. We use two noise magnitude constraints 0.02 and 0.05 for both a-RNA and random white noise. As for the white noise, we use a normal distribution and we vary the mean and the standard deviation to reach the desired noise magnitude. As shown in Figure 13, our technique is more efficient than injecting random noise. For instance, for a patch size equal to 600 and for a noise budget equal to 0.02, a-RNA insures more than 54% attack success rate, however, under the same constraints, the random noise is 17% of the times is successful.

Table 6: Magnitude of signals in the spectral domain for different noise sizes (ϵ = 0.03).

Patch size	30	50	100	200	300	400	500	600
Average Magnitude	0.0555	0.0665	0.1186	0.1368	0.1854	0.2085	0.2343	0.2615
Max Magnitude	0.4352	0.4596	0.8011	1.1412	1.6528	2.3231	2.8343	3.6812

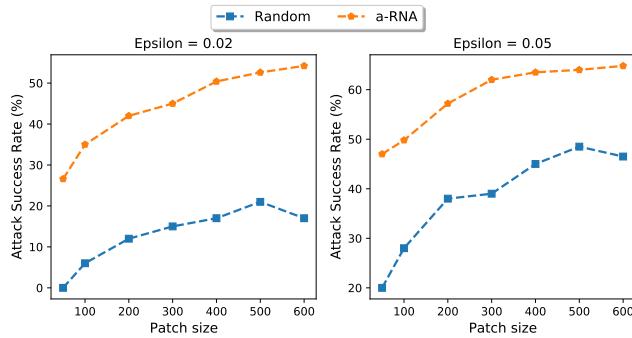


Figure 13: Attack success rate of a-RNA compared to random white noise.

9 Adversarial Training

In this section, we evaluate a-RNA in the case of an adversarially trained network. Adversarial training (AT) [17] is a state-of-the-art defense strategy against adversarial attacks. It can be formulated as follows [17]:

$$\min_{\theta} \mathbb{E}_{(x,y) \sim \mathcal{D}} \left[\max_{\delta \in B(x,\epsilon)} \mathcal{L}_{ce}(\theta, x + \delta, y) \right] \quad (11)$$

Where θ indicates the parameters of the classifier, \mathcal{L}_{ce} is the cross-entropy loss, $(x,y) \sim \mathcal{D}$ represents the training data sampled from a distribution \mathcal{D} and $B(x,\epsilon)$ is the allowed perturbation set. The interpretation of this is that the inner maximization problem is finding the worst-case samples for the given model, and the outer minimization problem is to train a model robust to adversarial examples [17].

We used the PGD algorithm [16] to solve the inner maximization problem. We set the noise magnitude (ϵ) equal to 0.002 with a step size of 0.0005 with a number of iterations equal to 20. We train the model for 500 epochs to reach 99% classification accuracy on clean input samples.

In Figure 14, we report the difference in impact of patches generated with a-RNA on adversarially-trained models and undefended models in the case of one random localized patch. As expected, AT decreases the effectiveness of the attack for low adversarial noise magnitude, and is practically bypassed for larger epsilons. Notice that adversarial training for high noise magnitudes may come at a baseline accuracy cost and lead to model utility drop.

To further explore the AT limits, we repeatedly (continuously) broadcast the adversarial patches; results in Figure 15 show that AT can be evaded with such setting.

In the next section we provide a discussion of our work and point out possible perspectives.

10 Discussion

This paper presents an adversarial attack against UWB-based ML systems for environment perceptions. This is the first work that shows a practical attack against such systems, which are used in several safety-critical applications. We investigate the attack from different perspectives and assumptions. We comprehensively consider real-world constraints such as random incidence time, and different defender models going from basic pre-processing to sophisticated cognitive radio defenders and an adversarially trained model. The proposed approach generates adversarial patches that are: (i) applicable under real-time constraints since they are input-agnostic, (ii) robust to time incidence delay, (iii) robust to filtering techniques, and (iv) can be undetectable even under a defender that deploys both filtering and spectrum sensing. Notice that there is a straightforward solution to avoid spectrum sensing: it consists of a cognitive adversarial noise emission by which the attacker broadcasts the adversarial noise exclusively when the victim device is active. This is known in other wireless communication contexts as reactive jamming [22, 24, 29]. In this work, we do not consider this method and focus our effort on the noise generation itself.

The assumption of this work is a line-of-sight (LoS) communication channel between the adversary and the victim device. While we tested under different distances (with LoS) and the attack remains efficient as far as the noise magnitude is increased to compensate the path loss, the attack is not successful in no-line-of-sight (No-LoS) settings. In future work, we will study this specific case by integrating the channel model in the noise generation method.

We believe this work should alert about the potential vulnerabilities of such systems to adversarial attacks in real-world settings, especially for critical applications. We hope these results would encourage the community to further investigate this direction.

11 Related Work

Computer vision is the mainstream application that caught the attention of the community from adversarial machine learning perspective [3, 6, 16]. However, the vulnerability of ML in other application domains has also been explored. For

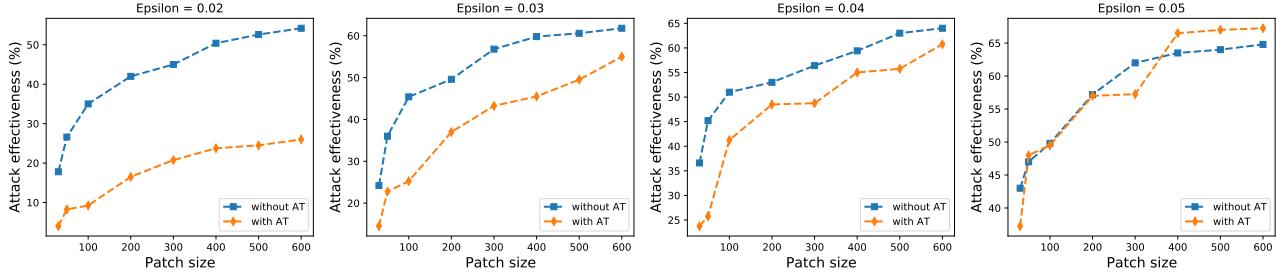


Figure 14: Impact of adversarial training (AT) on patch effectiveness.

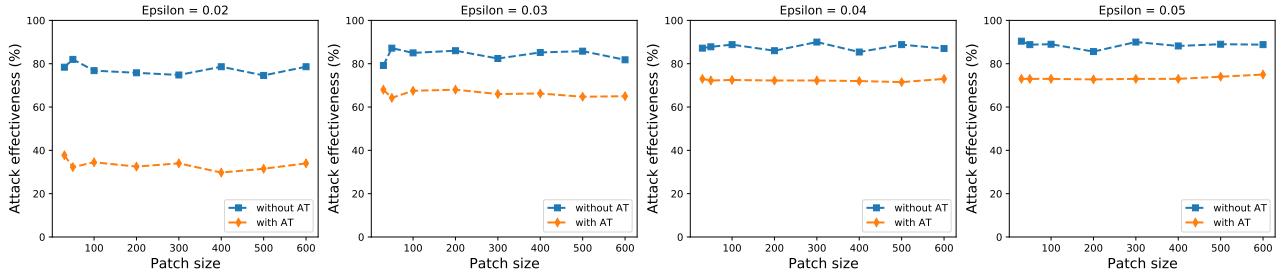


Figure 15: Impact of adversarial training (AT) on continuously broadcast patch effectiveness.

example, several papers have explored adversarial attacks on Lidar sensors [32] or electrocardiograms [9]. Some of the literature on automatic speech/speaker recognition have been designed under practical real-world assumptions [1, 14, 25, 30, 31]. Yet, most of the works require very specific settings to be applicable.

Several application of wireless communication use ML. These radio applications represent a potential target for adversarial attacks that are practical under real-world conditions. Recent work proposed crafting adversarial attacks against ML-based wireless systems. However, most of the existing works [2, 12, 13, 26, 27] use a input-specific perturbation to attack the target wireless model and do not provide the key properties for a practical attack, i.e., robustness against realistic conditions and active defenses. A recent paper present an attack against DNN-based wireless communication system [4]. This paper is the closest to our work and considers three applications: autoencoder communication, modulation recognition and channel estimation. While the authors presented an input-agnostic noise and discuss its robustness to potential defense techniques, the application case prevents the attack from being implementable in real life, and this is mainly due to the complex propagation channel of the adversarial noise itself.

Among wireless applications, radars are used for environment perception, and similar to other ML-based applications, they are vulnerable to adversarial examples. More specifically, short range systems such as some radar technologies represent a highly practical attack setting because of the possibility of line-of-sight transmission conditions. Few papers in the liter-

ature target these systems [10, 23]. Authors in [10] consider X-band spotlight mode radar which is utilized for hundreds of kilometers range and not practical for injecting adversarial noise due to the channel complexity. [23] targets short-range Frequency-modulated continuous-wave (FMCW) radars and is the closest paper to our work. However, FMCW radars give only velocity and range information, while UWB delivers a complete signature of the obstacle.

In this work, we propose an input-agnostic, undetectable, and robust adversarial attack against ML-based UWB radar systems. We design tailored universal patches to perform the attack and discuss their efficiency from practical perspectives. The short range aspect of the UWB radars represent a practical case due to the simple channel model.

12 Conclusion

We present a new adversarial radio noise attack (a-RNA) on UWB radars to generate a noise robust by design against realistic conditions and adaptive against defensive countermeasures. To our knowledge, this is the first approach to generate such practical attacks against DNN-based UWB systems. We believe a-RNA should alert the community about the feasibility of real-world attacks against radar systems.

References

- [1] Hadi Abdullah, Washington Garcia, Christian Peeters, Patrick Traynor, Kevin Butler, and Joseph Wilson. Practical Hidden Voice Attacks against Speech and Speaker

- Recognition Systems. In *Network and Distributed System Security Symposium (NDSS)*, 2019.
- [2] Abdullatif Albaseer, Bekir Sait Ciftler, and Mohamed M. Abdallah. Performance evaluation of physical attacks against e2e autoencoder over rayleigh fading channel. In *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, pages 177–182, 2020.
- [3] Anish Athalye, Nicholas Carlini, and David Wagner. Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 274–283. PMLR, 10–15 Jul 2018.
- [4] Alireza Bahramali, Milad Nasr, Amir Houmansadr, Dennis Goeckel, and Don Towsley. Robust adversarial attacks against dnn-based wireless communication systems. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, CCS ’21, page 126–140, New York, NY, USA, 2021. Association for Computing Machinery.
- [5] Guangke Chen, Sen Chenb, Lingling Fan, Xiaoning Du, Zhe Zhao, Fu Song, and Yang Liu. Who is real bob? adversarial attacks on speaker recognition systems. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 694–711, 2021.
- [6] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples, 2014.
- [7] Amira Guesmi, Ihsen Alouani, Mouna Baklouti, Tarek Frikha, and Mohamed Abid. Sit: Stochastic input transformation to defend against adversarial attacks on deep neural networks. *IEEE Design & Test*, 2021.
- [8] Amira Guesmi, Ihsen Alouani, Khaled N. Khasawneh, Mouna Baklouti, Tarek Frikha, Mohamed Abid, and Nael Abu-Ghazaleh. *Defensive Approximation: Securing CNNs Using Approximate Computing*, page 990–1003. Association for Computing Machinery, New York, NY, USA, 2021.
- [9] Xintian Han, Yuxuan Hu, Luca Foschini, Larry Chinitz, Lior Jankelson, and Rajesh Ranganath. Deep learning models for electrocardiograms are susceptible to adversarial attack. *Nature Medicine*, 26(3):360–363, 2020.
- [10] Teng Huang, Yongfeng Chen, Bingjian Yao, Bifen Yang, Xianmin Wang, and Ya Li. Adversarial attacks on deep-learning-based radar range profile target recognition. *Information Sciences*, 531:159–176, 2020.
- [11] Changhui Jiang, Jichun Shen, Shuai Chen, Yuwei Chen, Di Liu, and Yuming Bo. Uwb nlos/los classification using deep learning method. *IEEE Communications Letters*, 24(10):2226–2230, 2020.
- [12] Brian Kim, Yalin E. Sagduyu, Kemal Davaslioglu, Tugba Erpek, and Sennur Ulukus. Over-the-air adversarial attacks on deep learning based modulation classifier over wireless channels. In *2020 54th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6, 2020.
- [13] Brian Kim, Yalin E. Sagduyu, Kemal Davaslioglu, Tugba Erpek, and Sennur Ulukus. Channel-aware adversarial attacks against deep learning-based wireless signal classifiers, 2021.
- [14] Stepan Komkov and Aleksandr Petiushko. Advhat: Real-world adversarial attack on arface face id system. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 819–826, 2021.
- [15] Yongqiang Lu, Hongjie Ma, Edward Smart, and Hui Yu. Real-time performance-focused localization techniques for autonomous vehicle: A review. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–19, 2021.
- [16] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks, 2017.
- [17] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks, 2019.
- [18] Julien Maitre, Kévin Bouchard, Camille Bertuglia, and Sébastien Gaboury. Recognizing activities of daily living from uwb radars and deep learning. *Expert Systems with Applications*, 164:113994, 2021.
- [19] Amira Mimouna, Ihsen Alouani, Anouar Ben Khalifa, Yassin El Hillali, Abdelmalik Taleb-Ahmed, Atika Menhaj, Abdeldjalil Ouahabi, and Najoua Essoukri Ben Amara. Olimp: A heterogeneous multimodal dataset for advanced environment perception. *Electronics*, 9(4), 2020.
- [20] Amira Mimouna, Anouar Ben Khalifa, Ihsen Alouani, Najoua Essoukri Ben Amara, Atika Rivenq, and Abdelmalik Taleb-Ahmed. Entropy-based ultra-wide band radar signals segmentation for multi obstacle detection. *IEEE Sensors Journal*, 21(6):8142–8149, 2021.
- [21] S. Moosavi-Dezfooli, A. Fawzi, O. Fawzi, and P. Frossard. Universal adversarial perturbations. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 86–94, 2017.

- [22] Danh Nguyen, Cem Sahin, Boris Shishkin, Nagarajan Kandasamy, and Kapil R. Dandekar. A real-time and protocol-aware reactive jamming framework built on software-defined radios. In *Proceedings of the 2014 ACM Workshop on Software Radio Implementation Forum*, SRIF '14, page 15–22, New York, NY, USA, 2014. Association for Computing Machinery.
- [23] Utku Ozbulak, Baptist Vandersmissen, Azarakhsh Jalalvand, Ivo Couckuyt, Arnout Van Messem, and Wesley De Neve. Investigating the significance of adversarial attacks and their relation to interpretability for radar-based human activity recognition systems. *Computer Vision and Image Understanding*, 202:103111, 2021.
- [24] Konstantinos Pelechrinis, Marios Iliofotou, and Srikanth V. Krishnamurthy. Denial of service attacks in wireless networks: The case of jammers. *IEEE Communications Surveys Tutorials*, 13(2):245–257, 2011.
- [25] Yao Qin, Nicholas Carlini, Garrison Cottrell, Ian Goodfellow, and Colin Raffel. Imperceptible, robust, and targeted adversarial examples for automatic speech recognition. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 5231–5240. PMLR, 09–15 Jun 2019.
- [26] Meysam Sadeghi and Erik G. Larsson. Adversarial attacks on deep-learning based radio signal classification. *IEEE Wireless Communications Letters*, 8(1):213–216, 2019.
- [27] Meysam Sadeghi and Erik G. Larsson. Physical adversarial attacks against end-to-end autoencoder communication systems. *IEEE Communications Letters*, 23(5):847–850, 2019.
- [28] Akihiko Saito, Hiroshi Harada, and Atsuhiro Nishikata. Development of band pass filter for ultra wideband (uwb) communication systems. *IEEE Conference on Ultra Wideband Systems and Technologies*, 2003, pages 76–80, 2003.
- [29] Mario Strasser, Boris Danev, and Srdjan Čapkun. Detection of reactive jamming in sensor networks. *ACM Trans. Sen. Netw.*, 7(2), sep 2010.
- [30] Bilel Tarchoun, Ihsen Alouani, Anouar Ben Khalifa, and Mohamed Ali Mahjoub. Adversarial attacks in a multi-view setting: An empirical study of the adversarial patches inter-view transferability. In *2021 International Conference on Cyberworlds (CW)*, pages 299–302, 2021.
- [31] Simen Thys, Wiebe Van Ranst, and Toon Goedeme. Fooling automated surveillance cameras: Adversarial patches to attack person detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019.
- [32] James Tu, Mengye Ren, Sivabalan Manivasagam, Ming Liang, Bin Yang, Richard Du, Frank Cheng, and Raquel Urtasun. Physically realizable adversarial examples for lidar object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [33] UMAIN Inc. HST-D3 evaluation kit. <https://umain.en.ec21.com/>.
- [34] Changqiang Wang, Aigong Xu, Xin Sui, Yushi Hao, Zhengxu Shi, and Zhijian Chen. A seamless navigation system and applications for autonomous vehicles using a tightly coupled gnss/uwb/ins/map integration scheme. *Remote Sensing*, 14(1):27, 2022.
- [35] Xiaoyong Yuan, Pan He, Qile Zhu, Rajendra Rana Bhat, and Xiaolin Li. Adversarial examples: Attacks and defenses for deep learning. *CoRR*, abs/1712.07107, 2017.
- [36] Rudolf Zetik, Jurgen Sachs, and Reiner S. Thoma. Uwb short-range radar sensing - the architecture of a baseband, pseudo-noise uwb radar sensor. *IEEE Instrumentation Measurement Magazine*, 10(2):39–45, 2007.