Poster: Practical Removal Attacks on LiDAR-based Object Detection in Autonomous Driving

Takami Sato*†, Yuki Hayakawa*‡, Ryo Suzuki*‡, Yohsuke Shiiki*‡, Kentaro Yoshioka‡, and Qi Alfred Chen†
†University of California, Irvine; ‡Keio University

***Abstract*— LiDAR (Light Detection And Ranging) is an indispensable sensor for precise long- and wide-range 3D sensing, which directly benefited the recent rapid deployment of autonomous driving (AD). Recent research demonstrates that one can manipulate the LiDAR point cloud and fool object detection. However, these efforts evaluate only a specific LiDAR (VLP-16) and do not consider the recent LiDARs, so-called next-generation LiDARs. In this poster, we report our recent progress in the security analysis of the next-generation LiDARs. We identify a new type of LiDAR spoofing attack applicable to a much more general and recent set of LiDARs. We find that our attack can remove >72% of points in a 10×10 m^2 area and can remove real vehicles in the physical world. We also discuss our future plans.**

I. Introduction

LiDAR (Light Detection And Ranging) is one of the most innovative sensors in the past decade. While highly beneficial to our everyday life and society, the major security concern of LiDARs is the fundamental vulnerability against *LiDAR spoofing attacks* [1]. However, we find that prior efforts evaluate only a specific LiDAR (VLP-16) and do not consider the state-of-the-art defense mechanism in the recent LiDARs. Although prior attacks are valid on VLP-16, there is no guarantee that these results are still valid in more recent LiDARs, known as next-generation (or *next-gen*) LiDARs [2], as opposed to the first-generation (or *first-gen*) ones such as VLP-16. The next-gen LiDARs have more advanced spoofing-related features, such as laser timing randomization and pulse fingerprinting.

In this poster, we report our recent progress in designing powerful and practical *asynchronized* (§II-A) spoofing attacks to rigorously measure the vulnerability status of next-gen LiDARs since synchronized ones are directly foiled by their laser timing randomization. We identify a new asynchronized attack design called *High-Frequency Removal (HFR)* attack, which is much more powerful and practical than prior ones (e.g., can remove points in a 10×10 m^2 area, while the latest prior one can only remove points in a 41×42 cm^2 area [3]). We finally discuss the future plans of this research.

II. Background and Related Works

A. LiDAR Spoofing Attacks

Table I shows a taxonomy of LiDAR spoofing attacks based on (1) the requirement of *synchronization* with the LiDAR scanning pattern (row); and (2) the *attack effect*: object injection or removal (column). The spoofing mechanisms are illustrated in Fig. 1. *Synchronization* means to synchronize the malicious laser firing timing with the victim LiDAR scanning (i.e., laser firing) timing.

Synchronization is no longer possible on Next-Gen LiDAR. Among the new features of next-gen LiDARs, the laser timing randomization makes the synchronized attacks virtually impossible because it can directly foil the fundamental attacker's assumption that needs to predict the scan pattern of LiDAR. However, if the timing of laser firing and receiving is randomized, the attacker can no longer synchronize it or even know when the laser will fire.

TABLE I: Taxonomy of existing LiDAR spoofing attacks and ours. Rows correspond to whether the attack requires the synchronization with the LiDAR scanning pattern. Columns correspond to attack effects: object injection or removal.

	Object Injection Attack	Object Removal Attack
Async. (Black-box)	Relay [4], Saturating [3]	Saturating [3], **HFR∗ (ours)**
Sync. (White-box)	Adv-LiDAR∗ [5], Occlusion∗ [6], Frustum∗ [7]	PRA∗ [1], ORA [4]

∗ Attack effectiveness against AD has been considered.

III. Attack Experiment Setup and Methodology

A. Our Improvements on Spoofer Design

We generally follow the common setup adopted in the prior works [1], [5]–[7], but we improve the electronics and optical setup of the spoofer and achieve a significant improvement enabling the injection and removal of more than 6,000 points, which is 30 times more than the prior injection works [6], [7] and 1.5 times more than the prior removal attack [1].

B. High-Frequency Removal (HFR) Attack

As mentioned in §I, to measure the vulnerability status of next-gen LiDARs, we need powerful and practical *asynchronized* attacks since synchronized ones are directly foiled by the timing randomization. In this poster, we report our recent progress in designing a new type of asynchronized removal attack called high-frequency removal (HFR) attack, which is illustrated in Fig. 1. The key idea of it is to fire a large number of attack laser pulses to the victim LiDAR at a very high *frequency*, which, more specifically, is higher than the laser-firing frequency of the victim LiDAR. This allows the attack laser to hit every laser-firing event of the victim LiDAR in the scanning range hit by the spoofer, which can thus achieve the spoofing effect for every point in that range without any knowledge or synchronization with the victim scan pattern.

IV. Evaluation

Fig. 2 shows the attack demonstrations of the PRA [1] and our HFR attacks in the indoor setup. We place the spoofer 2 m away from the target LiDAR. As shown, the person and the majority of the room wall are removed by the attacks. For our HFR attack, there are points like a salt-and-pepper noise in the removed area. This is because as the key design feature, the HFR attack is asynchronized and thus achieves removal by moving points to a random location or undetectable area. Table II lists the results of the PRA attack and our

*co-first authors

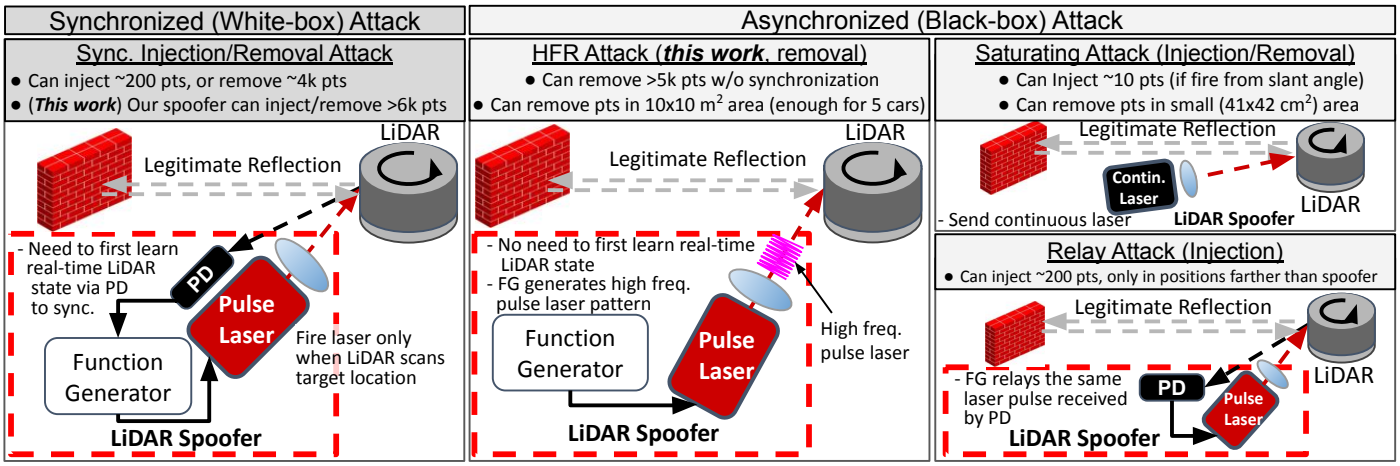


Fig. 1: Illustration of 4 LiDAR spoofing attack types. Asynchronized attacks do not need synchronization, and thus are generalizable to next-gen LiDARs. Our HFR attack is the first asynchronized removal attack on par with synchronized attacks.

TABLE II: Evaluation of PRA [1] and our HFR attack. $\mathcal{N}$ is the maximum # of injected points. θ is the attacked azimuthal range. $\mathcal{R}$ is the spoofing success rate in the azimuthal range.

		VLP-16	VLP-32c	NG-LiDAR①
PRA [1] w/ our spoofer	$\mathcal{N}$	6,621	9,711	N/A
	$\mathcal{R}$	96.9%	82.9%	N/A
	θ	85.4°	73.2°	N/A
HFR (ours)	$\mathcal{N}$	5,358	8,778	19,182
	$\mathcal{R}$	78.1%	72.2%	79.9%
	θ	85.8°	76.0°	81.7°

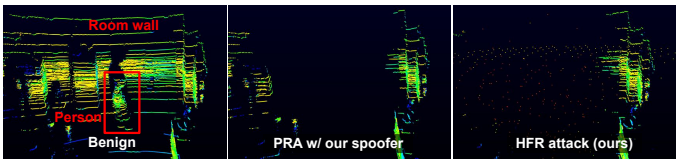


Fig. 2: Attack demos of removal attacks. A person and the room wall are totally removed by PRA [1] and our HFR attack.

newly-identified HFR attack on the first-gen LiDARs and an anonymized next-gen LiDAR (*NG-LiDAR*①) with the timing randomization. Note that we anonymize the next-gen LiDAR for security reasons in this poster.

As shown, due to the reliance on synchronization, PRA is only applicable to the first-gen LiDARs (VLP-16 and VLP-32c); for the next-gen LiDAR (NG-LiDAR①), the synchronization is directly foiled by the time randomization. On the other hand, our HFR attack can still be effective on next-gen LiDARs with time randomization, since it does not depend on the synchronization with the fixed scanning pattern.

A. Real Vehicle Removal with HFR attack

We further test the effectiveness of the HFR attack in the physical world. Fig. 3 shows the point clouds in the benign and attack scenarios. We target VLP-16 LiDAR with the dual return mode. We detect objects with the PointPillars model in Baidu Apollo 6.0. As shown, our HFR attack is found to successfully remove 5 front vehicles at ~5 meters away, out of which all can be correctly detected in the benign scenario. Such an attack effect is found consistent across all the 100 continuous frames we collected, leading to a 100% attack success rate over 10 s.

V. Concluding Remarks

In this poster, we report our recent progress in the HFR attack, which is the first removal attack that can attack a more general and recent set of LiDARs, which shows high effectiveness in physical-world experiments. In the future, we plan to conduct a large-scale measurement study on LiDAR spoofing attack capabilities on object detectors with multiple next-gen LiDARs. We will also evaluate the defense side.

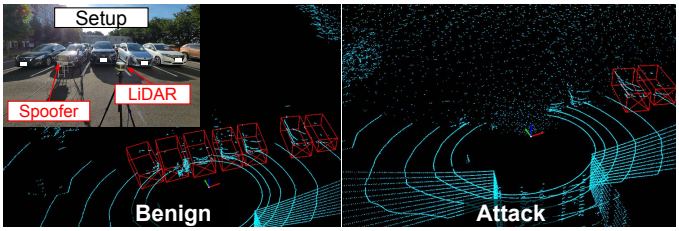


Fig. 3: Front-vehicle removal attack effect against real vehicles using our HFR attack. The 5 front vehicles become undetected.

Acknowledgements

This research was supported in part by the NSF CNS-1932464, CNS-1929771, CNS-2145493, USDOT UTC Grant 69A3552047138, JST SPRING JPMJSP2123, JST PRESTO JPMJPR22PA, and JSPS KAKENHI 21K20413.

References

[1] Y. Cao, S. H. Bhupathiraju, P. Naghavi, T. Sugawara, Z. M. Mao, and S. Rampazzi, "You Can't See Me: Physical Removal Attacks on LiDAR-based Autonomous Vehicles Driving Frameworks," in *USENIX Security*, 2023.

[2] K. Yoshioka, "A Tutorial and Review of Automobile Direct ToF LiDAR SoCs: Evolution of Next-Generation LiDARs," *IEICE Transactions on Electronics*, vol. E105.C, no. 10, pp. 534–543, 2022.

[3] H. Shin, D. Kim, Y. Kwon, and Y. Kim, "Illusion and Dazzle: Adversarial Optical Channel Exploits Against Lidars for Automotive Applications," in *ICCHES*, 2017.

[4] J. Petit, B. Stottelaar, M. Feiri, and F. Kargl, "Remote Attacks on Automated Vehicles Sensors: Experiments on Camera and Lidar," *Black Hat Europe*, vol. 11, p. 2015, 2015.

[5] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao, "Adversarial Sensor Attack on Lidar-Based Perception in Autonomous Driving," in *ACM CCS*, 2019.

[6] J. Sun, Y. Cao, Q. A. Chen, and Z. M. Mao, "Towards Robust LiDAR-based Perception in Autonomous Driving: General Black-box Adversarial Sensor Attack and Countermeasures," in *USENIX Security*, 2020.

[7] R. S. Hallyburton, Y. Liu, Y. Cao, Z. M. Mao, and M. Pajic, "Security Analysis of Camera-LiDAR Fusion Against Black-Box Attacks on Autonomous Vehicles," in *USENIX Security*, 2022.

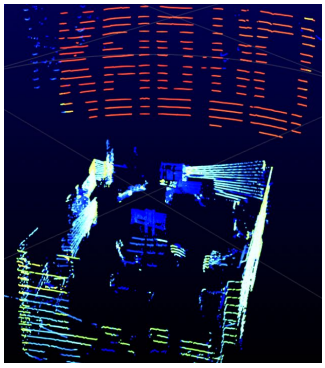
Poster: Practical Removal Attacks on LiDAR-based Object Detection in Autonomous Driving

Takami Sato[1]*, Yuki Hayakawa[2]*, Ryo Suzuki[2]*, Yohsuke Shiiki[2]*, Kentaro Yoshioka[2], Qi Alfred Chen[1]
[1]University of California, Irvine [2]Keio University
(* denotes co-first authors)

Limitations of Current LiDAR Spoofing Attack Research

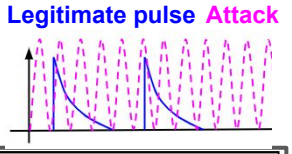
Evaluating only on a specific LiDAR
- Almost all attacks are only evaluated on **VLP-16**, which is known as a ***1st-generation*** **LiDAR**

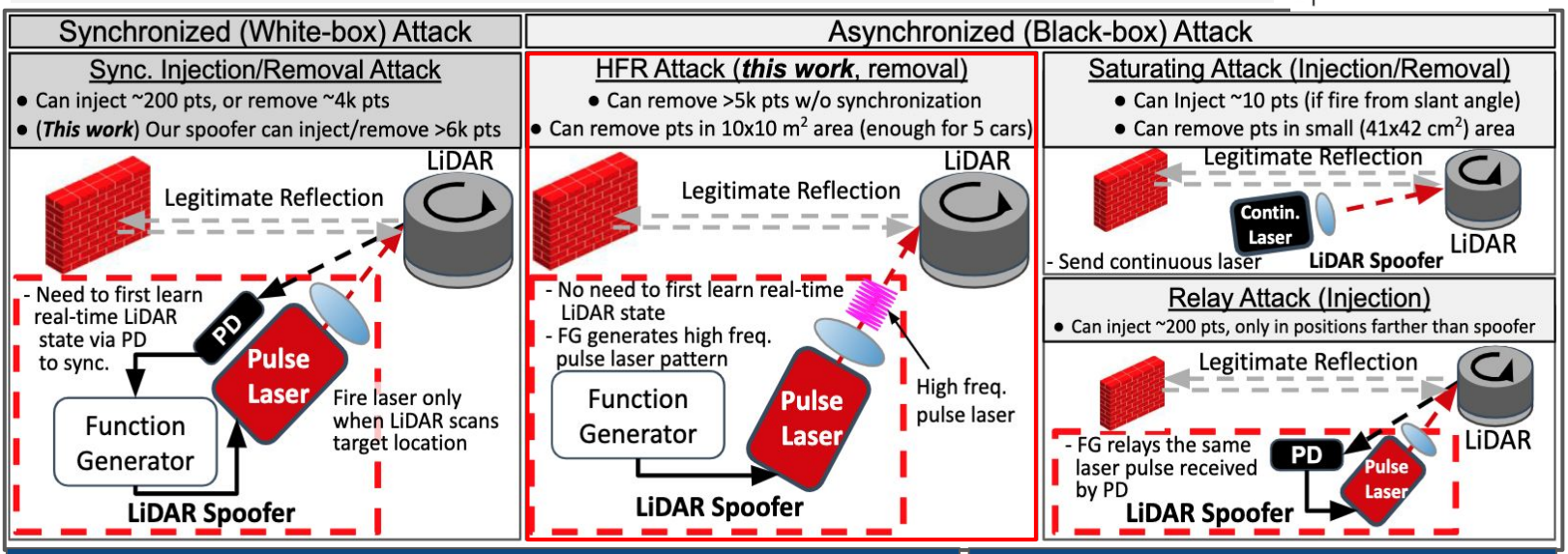
Did not consider spoofing-related features in recent next-gen. LiDARs [1]
- Recent next-gen LiDARs have spoofing-related features, which can potentially foil the existing LiDAR spoofing attacks
- Particularly, **timing randomization** can directly foil the attacker's fundamental assumption



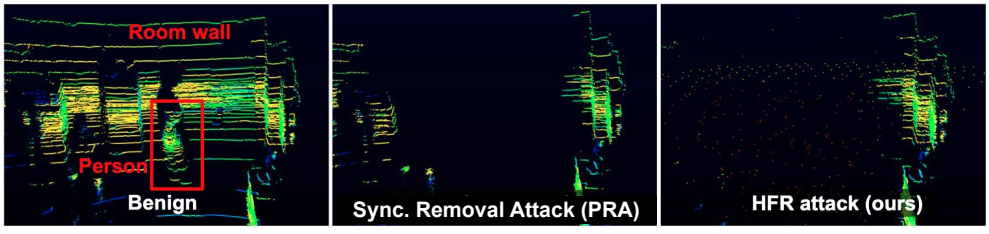
High Frequency Removal (HFR) Attack: Practical Removal Attack against Next-Gen LiDARs

- **Key idea:** Fire **a large number of attack laser pulses** to the victim LiDAR **at a very high frequency**
- Attack laser moves each point to a **random position** or **undetectable area** of the victim LiDAR
- Can theoretically **hit any LiDAR scan pattern even with randomization** if frequency is high enough

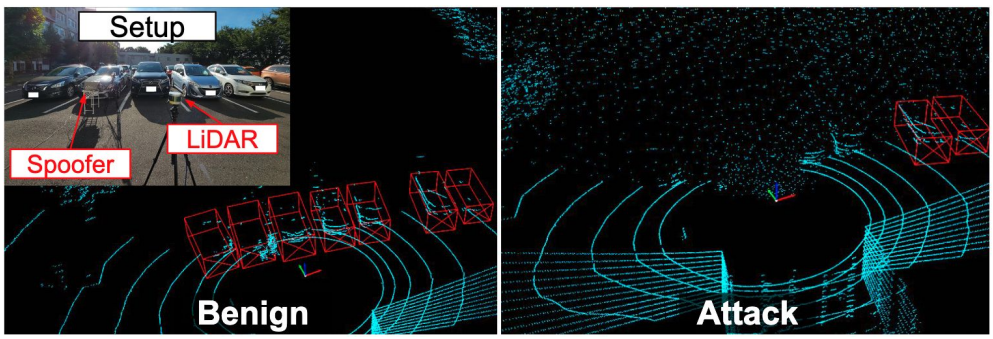




Attack Demo (Indoor)



Attack Demo (Outdoor)



Quantitative Analysis (Indoor)

		VLP-16	VLP-32c	NG-LiDAR①
Sync. Removal Attack (PRA)	$\mathcal{N}$	6,621	9,711	N/A
	$\mathcal{R}$	96.9%	82.9%	N/A
	θ	85.4°	73.2°	N/A
HFR (ours)	$\mathcal{N}$	5,358	8,778	19,182
	$\mathcal{R}$	78.1%	72.2%	79.9%
	θ	85.8°	76.0°	81.7°

NG-LiDAR①: An anonymized next-gen LiDAR
N: Maximum number of points injected by spoofing.
θ: Attacked azimuthal range.
R: the spoofing success rate in the azimuthal range.
N/A: Attack is not applicable on the LiDAR

Conclusion & Future Plan

- Identify **HFR attack**, the **first removal attack** that can attack a **more general and recent set of LiDAR**.
- Plan to conduct **first large-scale measurement study** on next-gen LiDARs
- Plan to identify and evaluate the defense capability of **other spoofing-related features**
- Plan to evaluate **end-to-end safety implications** of LiDAR spoofing attack on **autonomous driving**

[1] YOSHIOKA, Kentaro. "A Tutorial and Review of Automobile Direct ToF LiDAR SoCs: Evolution of Next-Generation LiDARs." IEICE Transactions on Electronics (2022)