LiDAR Spoofing Meets the New-Gen:
Capability Improvements, Broken Assumptions, and New Attack Strategies

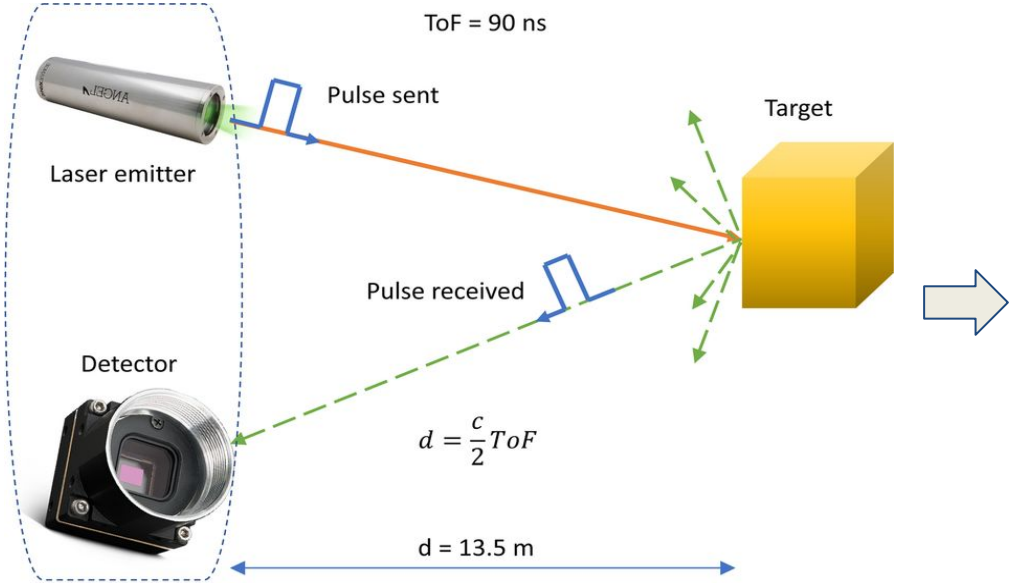
Takami Sato, Yuki Hayakawa, Ryo Suzuki, Yohsuke Shiiki,
Kentaro Yoshioka , Qi Alfred Chen

NDSS'24

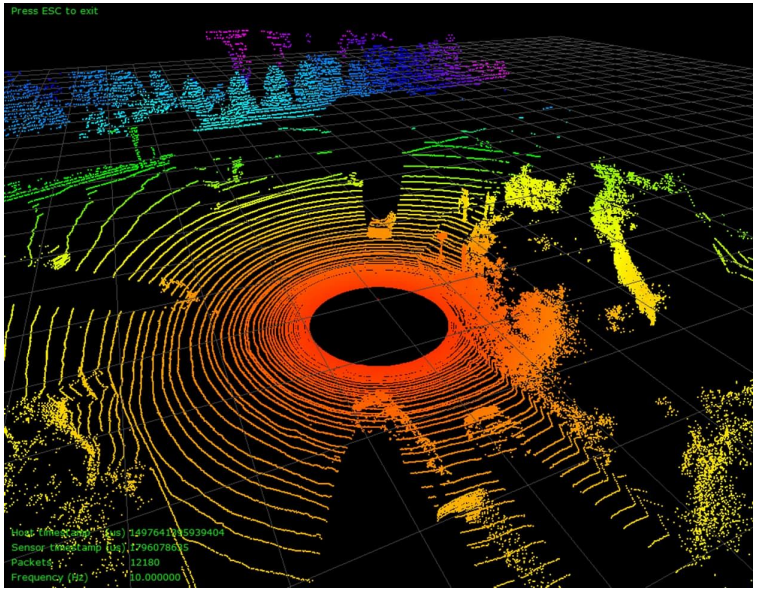
Presentor : Jiwoo Suh

Introduction

- What is LiDAR?



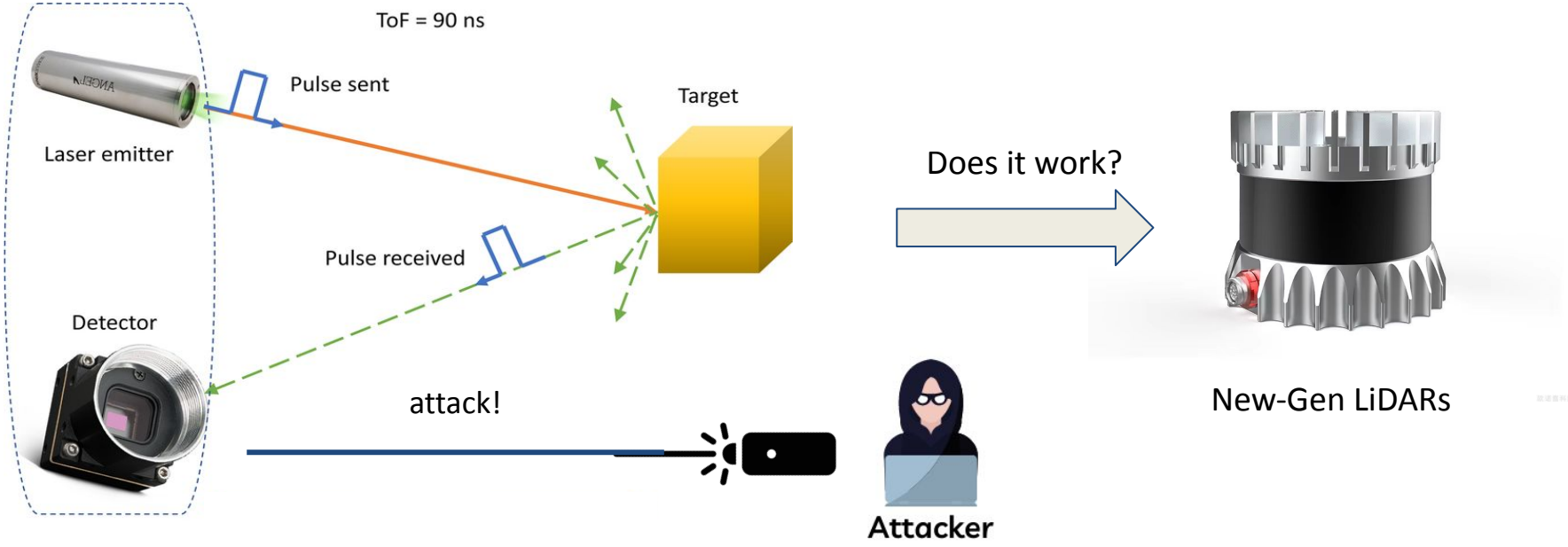
measuring time-of-flight



point cloud

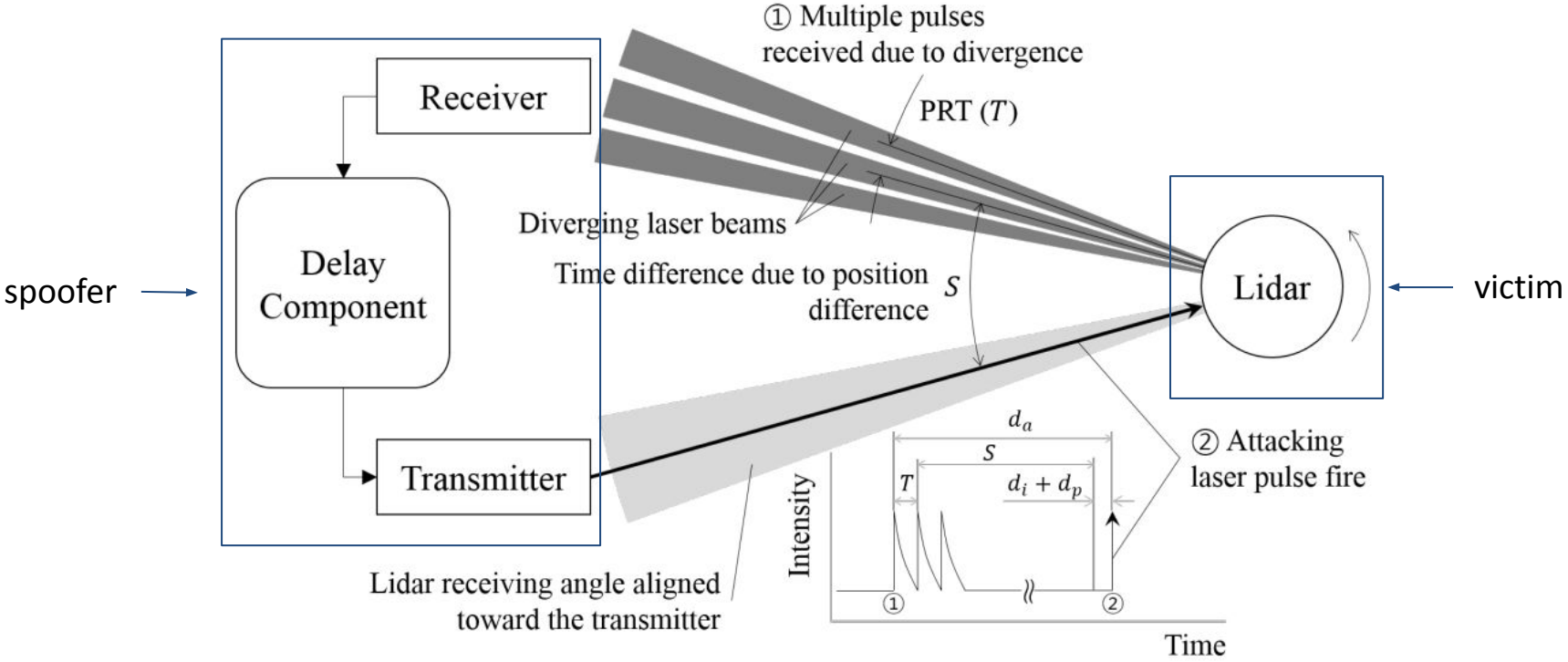
Introduction

- What if an attacker shoots a laser at the LiDAR detector?



Background

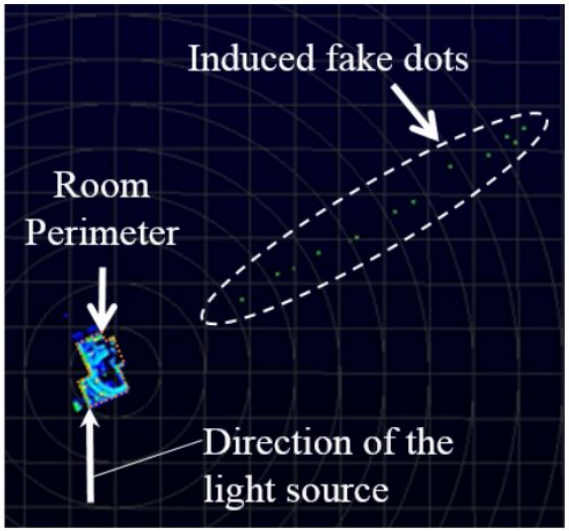
- LiDAR spoofing first tested by Shin, CHES'17[1)]



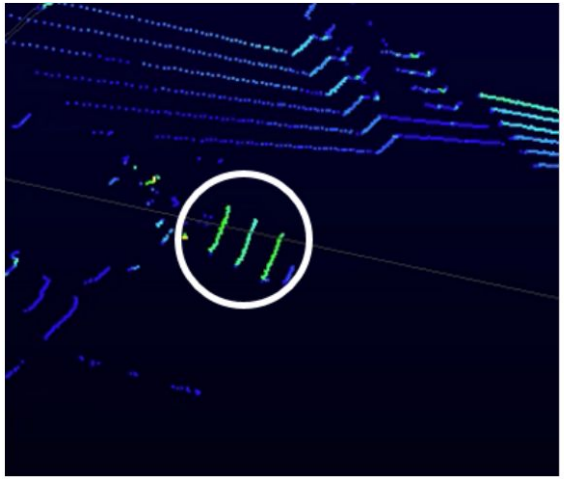
1) Hocheol Shin, Dohyun Kim, Yujin Kwon, and Yongdae Kim "Illusion and Dazzle: Adversarial Optical Channel Exploits Against Lidars for Automotive Applications." CHES'17

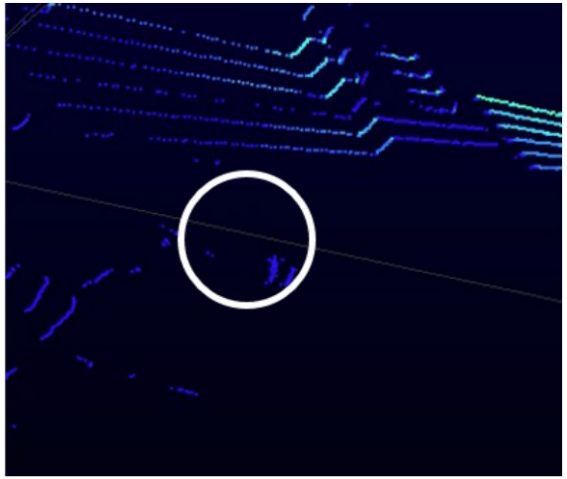
Background

- Limitation?



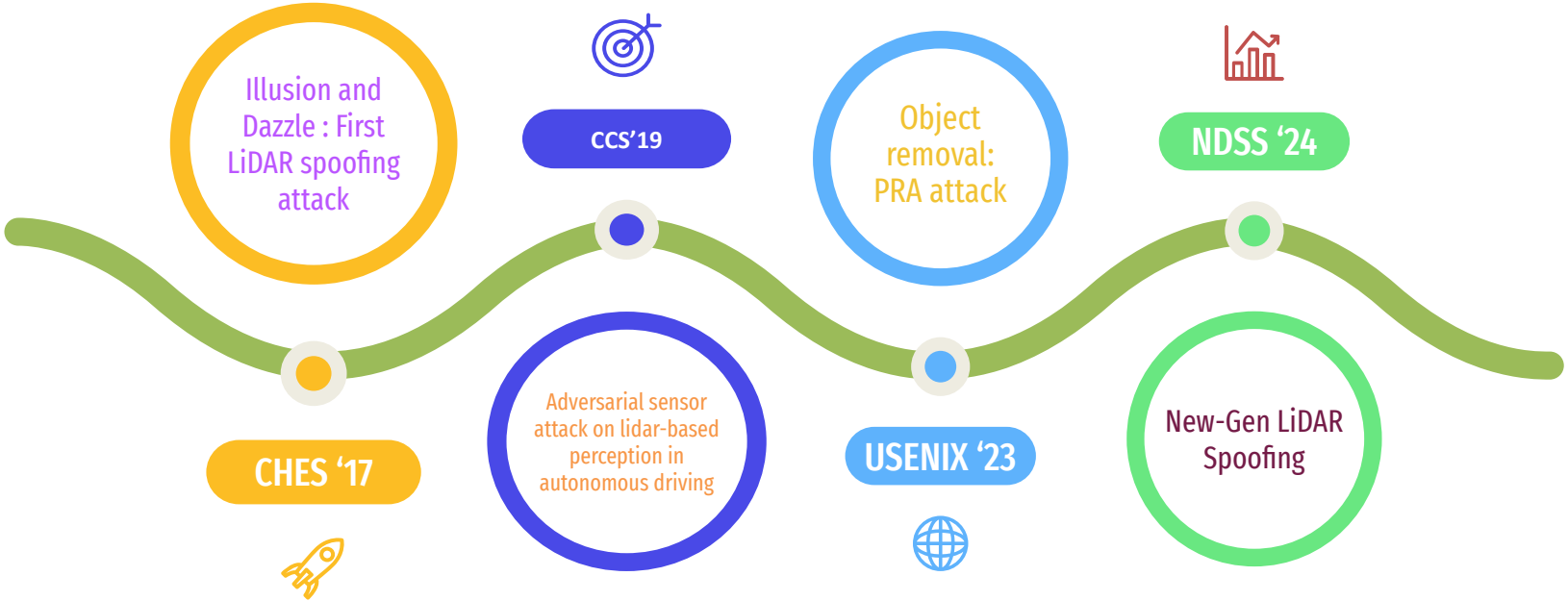
spoofed point cloud





Saturation attack : Object removal attack!

Previous work



New-gen LiDAR

- Previous work mainly focuses on the Velodyne VLP-16.
  - Older attacks are not guaranteed to succeed on new-gen LiDARs!

VLP-16

VLS-128

1st-Gen LiDARs

OS1-32

Realsense L515

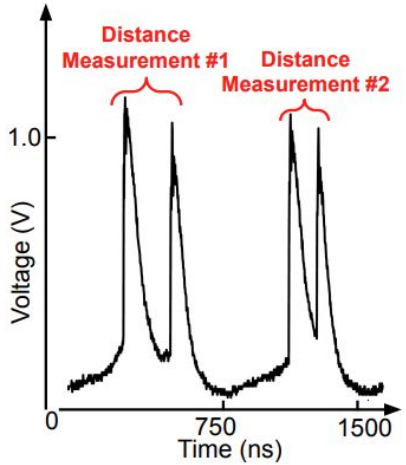
New-Gen LiDARs

New-gen LiDAR

- New-gen LiDARs have new features that counter spoofing attacks
  - Timing randomization
  - Pulse fingerprinting

New-Gen LiDAR

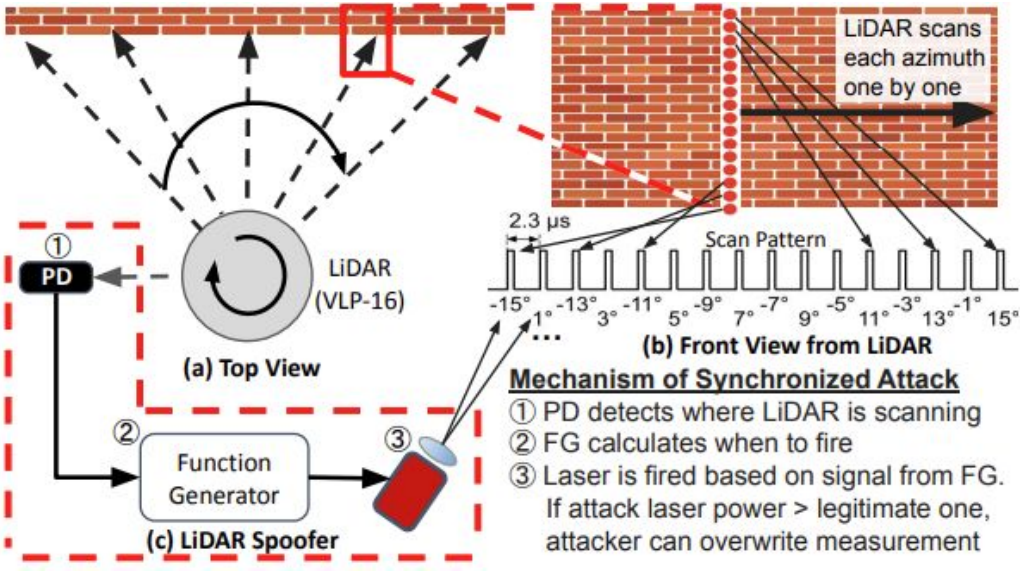
1st-Gen LiDAR

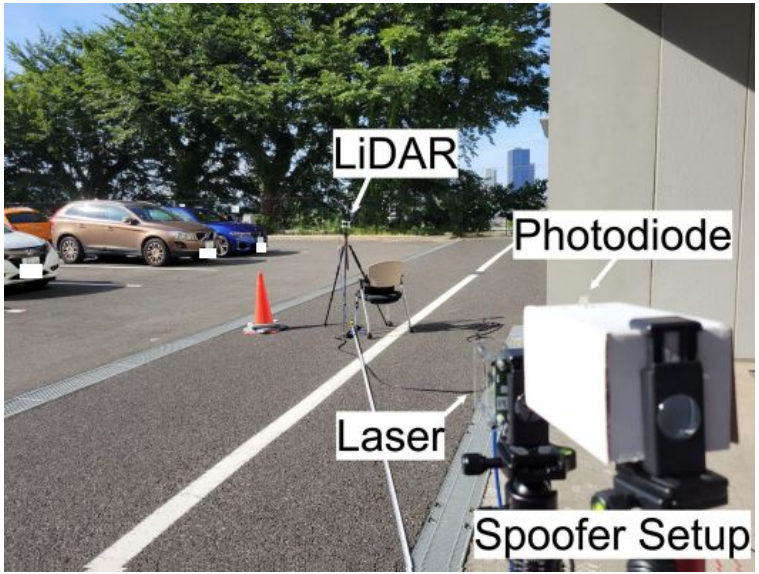


pulse fingerprinting

Threat model

- The attacker synchronizes the malicious laser firing timing with the victim LiDAR
- The attacker aims to inject/remove points from point cloud



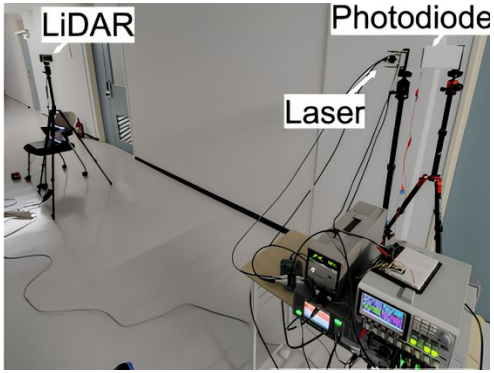


Research Question

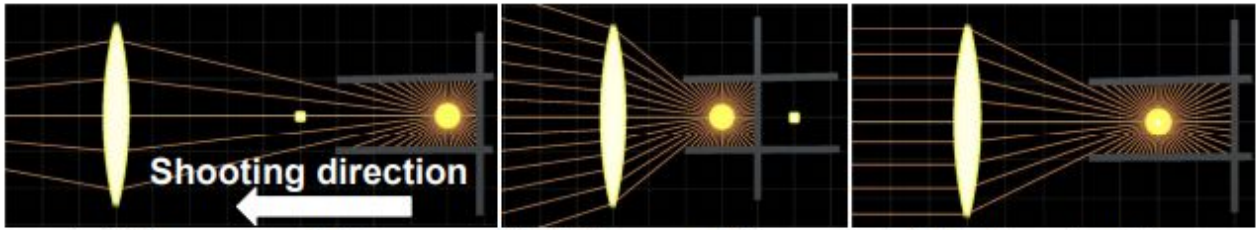
- RQ1 : Is Chosen Pattern Injection actually feasible?
- RQ2 : Do new-gen LiDAR features defend well against spoofing attacks?
- RQ3 : Do new-gen LiDAR systems exhibit different vulnerability characteristics?

RQ1 : Spoofer Improvements

- Fixed the inadequate design of previous spoofers
  - Fixed optical design



spoofer setup

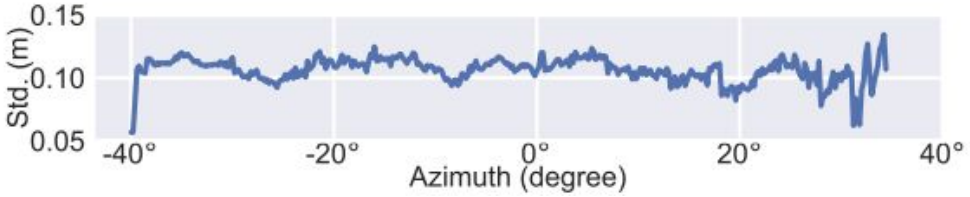


CPI attack capability can be achievable in well-calibrated spoofer!

RQ1 : CPI attack on VLP-16

d	Indoor $\mathcal{N}$	Indoor $\mathcal{R}$	Indoor θ	Outdoor (Daytime: 70 lux) $\mathcal{N}$	Outdoor (Daytime: 70 lux) $\mathcal{R}$	Outdoor (Daytime: 70 lux) θ
2 m	6,523 (-)	98.5%	82.7°	7,705 (-)	94.9%	100.5°
(2.5 m)	(<4k)			(-)		
4 m	6,386 (-)	96.9%	82.5°	7,950 (<1.8k)	96.9%	101.5°
6 m	6,575 (-)	98.6%	83.4°	7,357 (<1.5k)	87.2%	99.6°
8 m	6,213 (-)	93.8%	82.8°	6,702 (<1k)	97.7%	83.4°
10 m	6,131 (-)	93.2%	82.1°	6,514 (<1k)	93.3%	84.2°

N : Number of injected points by spoofing R: Point injection success rate within θ



Standard deviations of inner-frame error on VLP-16

CPI attack capability can be achieved with a well-calibrated spoofer!

RQ2 : CPI attack on new-gen LiDAR

CPI attack is not feasible on new-gen LiDAR!

	First-Gen		New-Gen					
			w/ Timing Randomization				w/ Fingerprint	
	VLP-16	VLP-32c	OS1-32	Helios	Horizon	L515	XT32	
$\mathcal{N}$	6,523	9,711	28	3,203	19,182	321	113	
$\mathcal{R}$	98.50%	82.90%	43.80%	19.4%	79.90%	0.1%	2.10%	← low attack success rate on new-gen LiDAR!
θ	82.7°	73.2°	0.72°	34.2°	103.4°	81.7°	70°	

N : Number of injected points by spoofing R: Point injection success rate within θ

RQ2 : Spoofing attack capability modeling

- How does the paper mathematically model the point injection capability?

$$\mathcal{P}_I(x_{ij}) = x_{ij} + (\delta_{ij}^{\text{rand}} + \delta_{ij}^{\text{inner}} + \delta^{\text{inter}}) \cdot g(x_{ij}),\ x_{ij} \in \mathcal{C}_n \subset \mathcal{C}$$

point injected by the attack at i-th altitude and j-th azimuth

error part

Attacker's chosen pattern (e.g., point cloud of a vehicle)

Parameters	Explanation
δ_{ij}^{rand}	error caused by timing randomization
δ_{ij}^{inner}	inner frame error
δ^{inter}	inter frame error

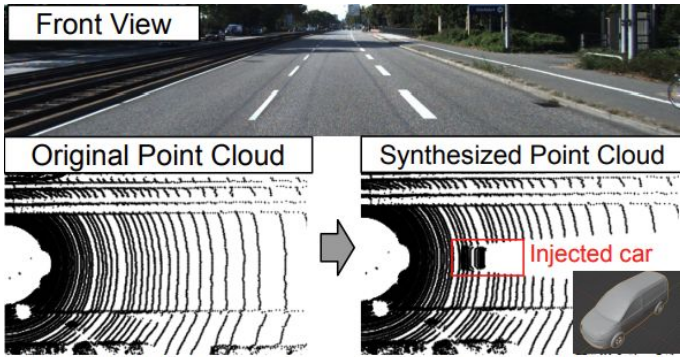
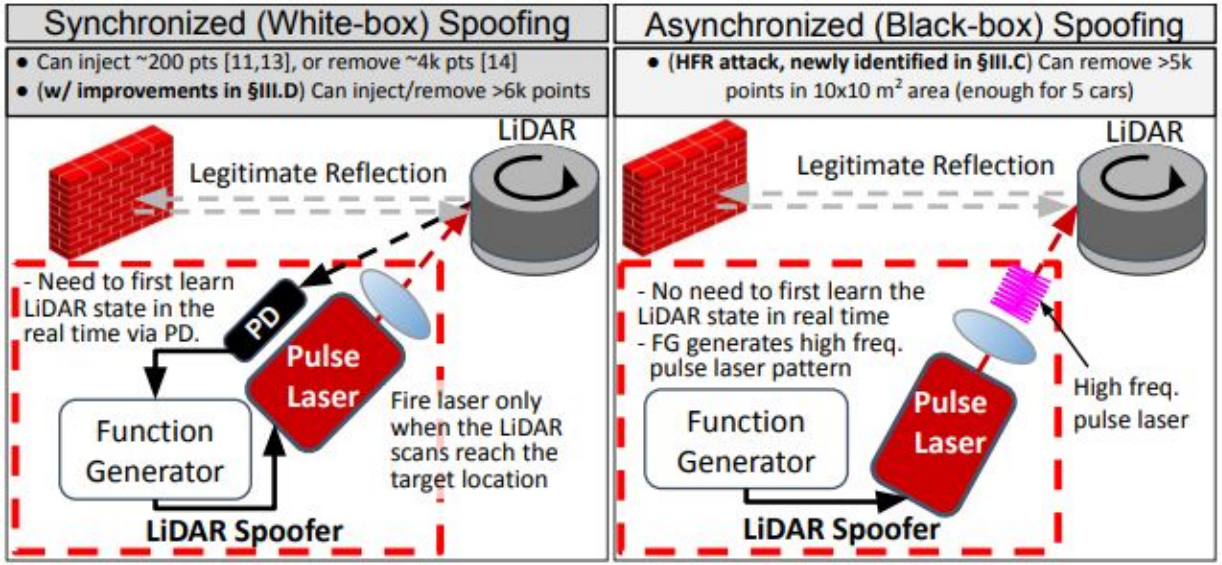


Figure 17: Targeted experiment scenario from KITTI [32].

High frequency removal attack

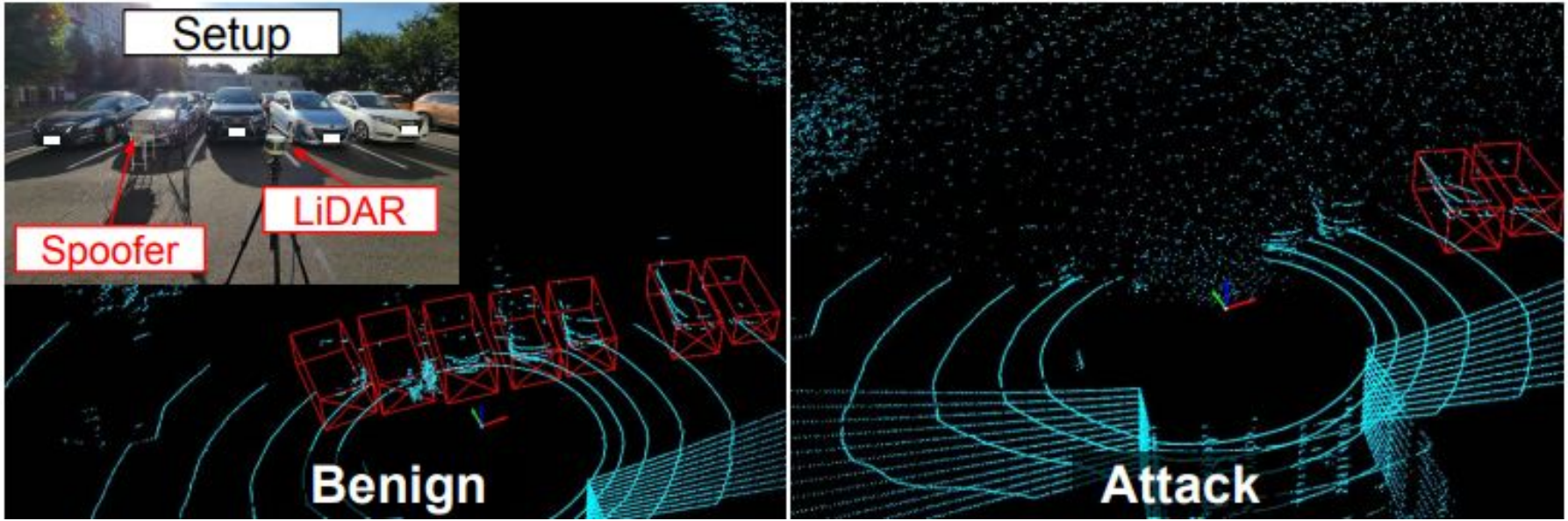
- Object removal attack (PRA) requires synchronization
  - New-gen LiDARs have timing randomization!



The author suggests HFR (high frequency removal) attack!

High frequency removal attack

High frequency removal attack works on new-gen LiDARs!



High frequency removal attack

High frequency removal attack works on new-gen LiDARs!

		First-Gen		New-Gen				
				w/ Timing Randomization				w/ Fingerprint
		VLP-16	VLP-32c	OS1-32	Helios	Horizon	L515	XT32
PRA [13]	$\mathcal{N}$	6,621	9,711	N/A	N/A	N/A	N/A	N/A
	$\mathcal{R}$	96.9%	82.9%	N/A	N/A	N/A	N/A	N/A
	θ	85.4°	73.2°	N/A	N/A	N/A	N/A	N/A
HFR (§III-C)	$\mathcal{N}$	5,358	8,778	28	4,108	19.2k	206k	113
	$\mathcal{R}$	78.1%	72.2%	43.8%	24.8%	79.9%	91.3%	2.1%
	θ	85.8°	76.0°	0.72°	103.4°	81.7°	70.0°	34.2°

* N/A: Attack is not applicable to the LiDAR

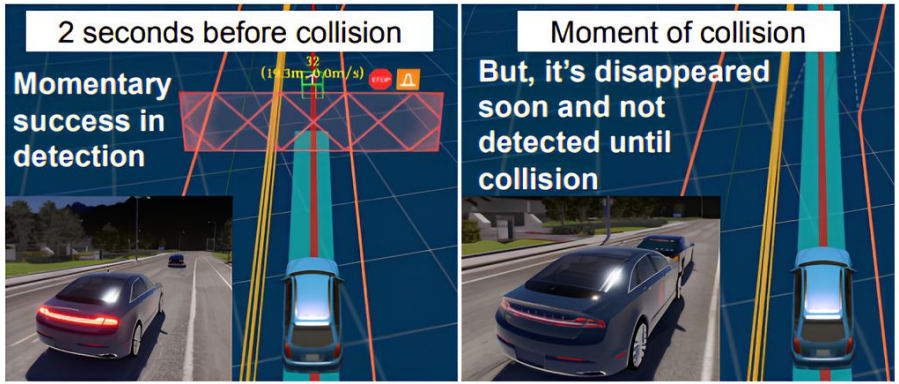
object removal attack success rate

Evaluation

- Used Baidu apollo and LGSVL

AD system

simulator



HFR attack in a simulator

Evaluation

- Tested PRA and HFR on different LiDARs

		Benign	10m	15m	16m	17m	18m	19m	20m
PRA	VLP-16	0/10	0/10	5/10	8/10	9/10	**10/10**	**10/10**	**10/10**
HFR	VLP-16	0/10	0/10	6/10	7/10	8/10	**10/10**	**10/10**	**10/10**
	VLP-32c	0/10	1/10	9/10	8/10	**10/10**	**10/10**	**10/10**	**10/10**
	XT32	0/10	0/10	0/10	0/10	0/10	0/10	0/10	0/10
	Helios	0/10	0/10	6/10	5/10	**10/10**	**10/10**	**10/10**	**10/10**

← attack distance

Vehicle collision rate over 10 trials using PRA and HFR

Demo video

Defense

1) Sensor-level defense: More complex fingerprinting is required
2) Software-level defense: Detect the unique characteristics of an HFR attack

Features	Effectiveness		Limitations		
	Injection	Removal	Eye safety	Latency	Range↓*
Timing Random.	**High**	**High**	**No risk**	**Low impact**	**None**
Pulse Fingerprint	Mid	**High**	High risk	Mid impact	High
Simul. Firing	Low	None	**Low risk**	**Low impact**	**Low**

* Range↓: Degradation of the effective sensing range of LiDAR

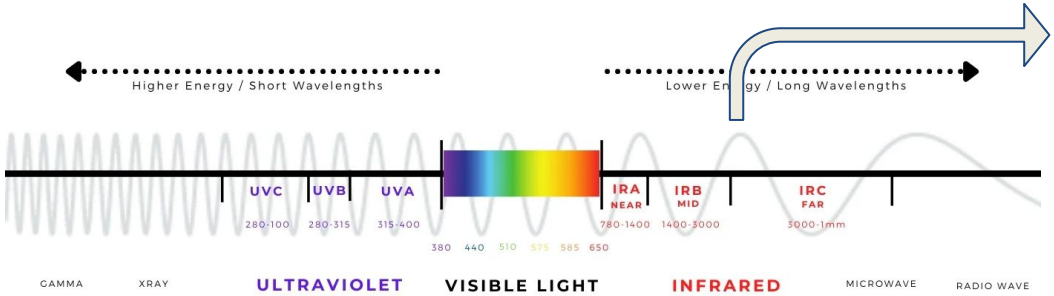
Defense effectiveness of new-gen LiDARs

Limitation

- Aiming at Driving AD vehicle : Is HFR attack deployable in real-world?
- LiDAR model coverage

Future work

- Pulse fingerprinting coding design
  - complexity ⟺ eye safety and detection range



use long range wavelength?

- Other possible spoofing attack on new-gen LiDARs?

Conclusion

- Contribution
  - First large scale measurement study on LiDAR spoofing attack
    - Tested with 9 popular LiDARs
  - Spoofer improvements
  - Identify new LiDAR attack : High frequency removal attack
  - Mathematical modeling for LiDAR attacks

- Personal opinion
  - Advantages of HFR attack over saturation attack?
  - Mathematical modeling?

Good questions

- Reliable experimental method for LiDAR spoofing like a real environment(fast moving cars)?
- Can we make autonomous vehicle more secure from spoofing attacks by combining multiple sensors like camera, radar, or LiDAR?
- Can we add amplitude modulation to pulse fingerprinting?
- Can we also attack analog sensors using techniques like 'Ghost Talk' to spoof LiDAR systems?
- Can absorbing or reflecting a laser sent from the LiDAR sensor induce object removal effect as well?
- Sharing GPS coordinates and a 3D mapping of their surroundings with other cars?

Best questions

Yuanxin Pang : For the HFR attack, I wonder whether the LiDAR can detect unexpected frequency distribution through the signal's frequency spread, interference can be identified?

Wonyoung Kim : If someday a robot with sensors and object recognition capabilities similar to humans were to drive instead, wouldn't it be safer than autonomous driving? They might be able to do some actions such as turning their heads

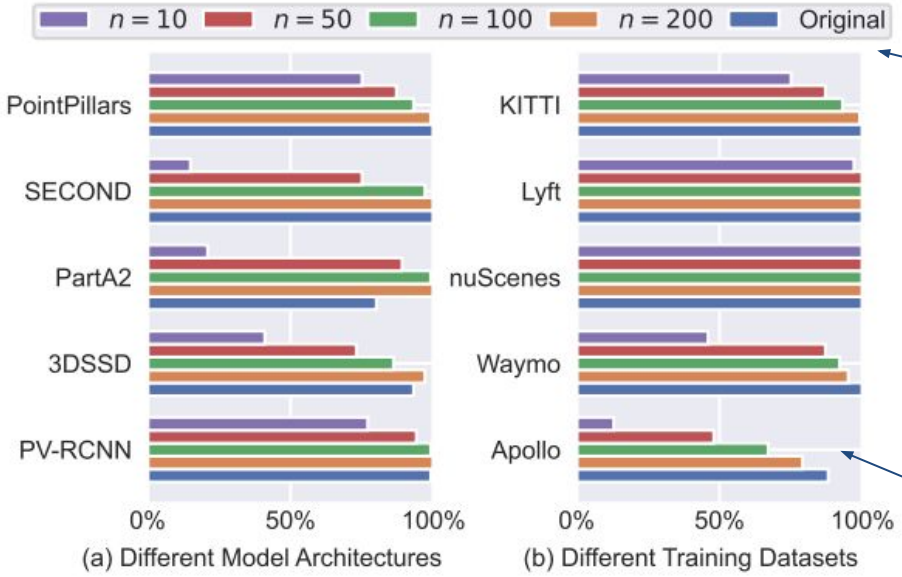
Boris Antoine Testud : LiDARs seem to be the most accurate sensors we have today to measure distances and create 3D mappings of environments. What do you think could be the reason why Tesla is choosing to move away from using LiDARs in their cars and replacing them with cameras and computer vision? (compare Lidar and camera, best question)

LiDAR vs Camera

	LiDAR	Camera
Cost	Expensive	Cheap
Depth sensing	Accurate	Requires stereo vision setup
Object recognition	Limited	Good(can see traffic signs)
Environment	Works without light, robust to fog and dust	Vulnerable to weather conditions
Range	Limited	Long

RQ3 : Impact of Pulse Fingerprinting

- Downsample the point cloud as a modeling of the fingerprinting effect



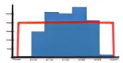
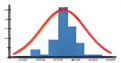
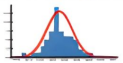
Object injection attack success rate

lower n implies higher fingerprinting complexity

With sufficient complexity, pulse fingerprinting demonstrates a high defense capability!

RQ3 : Impact of timing randomization

- Impact of timing randomization?
  - δ_{ij}^{rnd} follows uniform or gaussian distribution!

	OS1-32 [22]	Horizon [42]	L515 [41]	Pixell [40]	Helios [23]
					
Dist. [μs]	$\mathcal{U}_{1.4,1.8}$	$\mathcal{U}_{4.0,4.3}$	$\mathcal{N}_{51,0.025}$	$\mathcal{U}_{4.5,5.8}$	$\mathcal{N}_{1.6,0.005}$
Std. σ	33.3 m	26.0 m	7.5 m	110.4 m	1.5 m
Max. Δ	57.7 m	45.0 m	20.1 m	191.3 m	5.3 m

$\mathcal{U}_{\min,\max}$ - Uniform distribution, $\mathcal{N}_{\text{mean,std}}$ - Gaussian distribution

Distribution of laser firing intervals

LiDAR	Rand. model [m]	PointPillars	SECOND	PartA^2	3DSSD	PV-RCNN
VLP-16	$\emptyset$	100%	100%	80%	93%	97%
Helios	$\mathcal{N}(0, 1.5)$	2%	54%	41%	7%	24%
L515	$\mathcal{N}(0, 7.5)$	**0%**	24%	14%	7%	**0%**
Horizon	$\mathcal{U}(-45, 45)$	39%	35%	21%	30%	17%
OS1-32	$\mathcal{U}(-58, 58)$	47%	38%	21%	28%	23%
Pixell	$\mathcal{U}(-191, 191)$	60%	21%	20%	8%	43%
	Avg.	30%	34%	23%	16%	21%
With fingerprinting effect $n = 100$:						
	Avg.	38%	20%	16%	18%	41%

object injection attack success rates under different randomization levels

Timing randomization can have significant defense capability against object injection attack!