(2020 USENIX)Drift with Devil: Security of Multi-Sensor Fusion based Localization in High-Level Autonomous Driving under GPS Spoofing

1 Summary

In this paper, the author perform the first study about GPS sproofing attack upon Multi-Sensor-Fusion localization algorithms used in the state-of-art autonomous vehicles. To systematically understand the security the author first evaluated the upper-bound attack effectiveness and then used the take-over effect found to design a novel GPS sproofing attack method, FushionRipper. The author evaluated FusionRipper on real sensor traces and found that it can get a high success rates. With the observation that the profiling of parameters play an important role in the final effectiveness, the author also designed an offline method that can identify attack parameters with over 80% success rates. The time cost for this method is at most half a day which is affordable for most scenarios.

2 Challenge

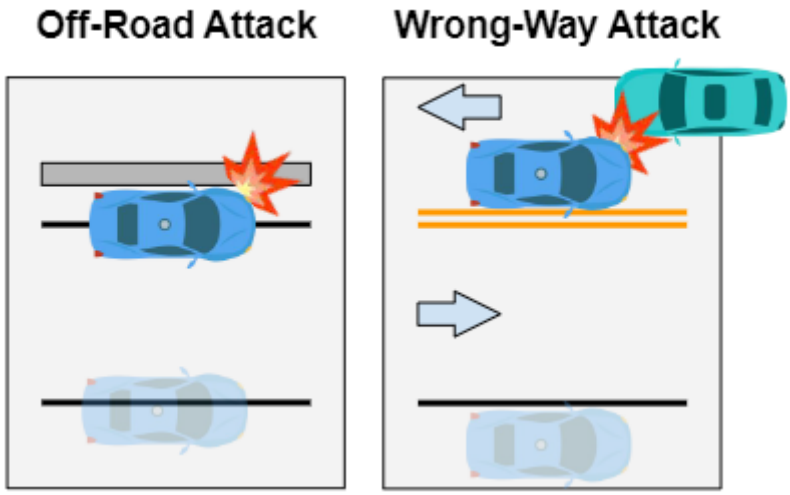
The MSF-based localization method used to be seemed as security to GPS sproofing attack, because the sensor fusion process will use the data from LIDAR and IMU to calibrate the result. So the attackers must manage to compromise the sensor fusion in localization.

3 Main Idea

3.1 Attack goal

In this paper, the author proposed two attack goals:

- Off-road attack
- Wrong-way attack

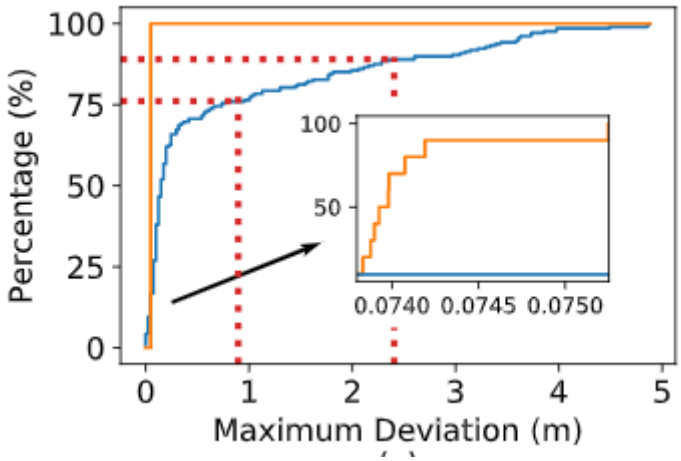


The main ideas of both attacks are to generate a later deviation in the localization output, which will make the car deviate from the traffic lane. The difference between two attacks is the distance needed to achieve the goal:

Attack Goal	Required Deviation (m)	
	Local	Highway
Off-Road Attack	0.895	1.945
Wrong-Way Attack	2.405	2.855

3.2 Upper-bound attack effectiveness and cause analyze

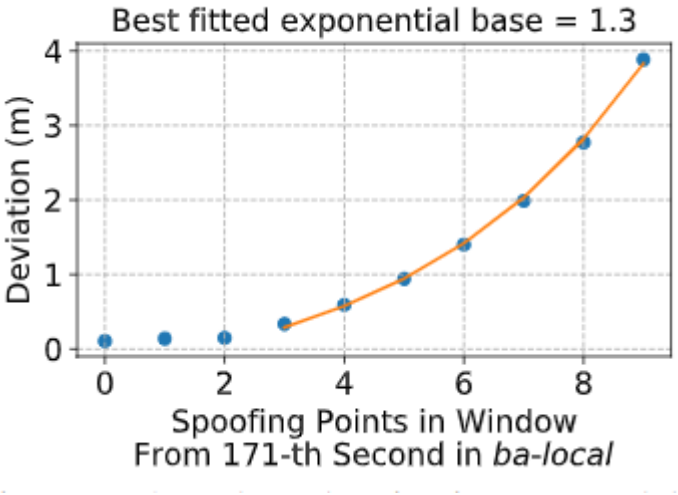
To find out the upper-bound attack effectiveness, the author evaluate a exhaustive search for each attack window in both an actual sensor trace and synthesized sensor trace. Here are the results:



form this CDF figure, we can find out that all of the windows in the synthesized window fail to achieve either of the attack goals. The majority can't either. However there are also some windows which get a maximum deviation above 2 meters. This result proves that there is the feasibility two perform successful GPS sproofing attack.

The author identified four potential factors which contribute to this phenomenon:

- Initial MSF state uncertainty(P_0)
- LiDAR measurement uncertainty(R_1^{lidar})
- Difference between LiDAR position and the original MSF output without attack(Δ_{lidar})
- IMU measurement(imu_1)
  The author also plot the deviation in those windows that meet the limitation in both attacks.



A take over is found in which the deviation is increased in a exponential pattern compared two common pattern.

3.3 FusionRipper

The analysis above tell us that the take over effect happen when the algorithm has a large uncertainty on the LiDAR and IMU measurement. In this situation, the result from the sproofed GPS input may have bigger weight in one window. Base on this observation the author proposed a two-phase two achieve the attack goal.

1. Vulnerability profiling. The attacker continues emit sproofed signal with a constant deviation. When the offset of the vehicle is larger than a threshold, then it's thought the sensor is in a vulnerable state.
2. Aggressive spoofing. After identify the vulnerable state, the main attack process is performed. In this stage, a exponential deviation is used to generate sproofed GPS input.

3.4 Offline attack parameter profiling

The results of the experiments prove that there always exists a combination of attack parameters that can achieve high success rates. However, these experiments also show that the selection of parameters has a significant impact on the attack result. To achieve effective attack, a parameter profiling method should be proposed to integrate the process.
In this paper, the author proposed a exhaustive method. For each combination, the attack performs several attack attempt. If the successful rate is higher than a threshold, the algorithm stops and reports current combination as profiling result. If there is non combination meets the requirement, the algorithm will return the combination with the highest success rates.

4 Strength

- Propose a novel attack GPS sproofing method that defeats sensor fusion.
- Prove the method can work relatively well in common scenarios.

5 Weakness

- The paper doesn't perform real-world attack. All the works are down by simulation.
- The offline method is quite primitive.
- The method is only evaluated in one particular production-level MSF implement.

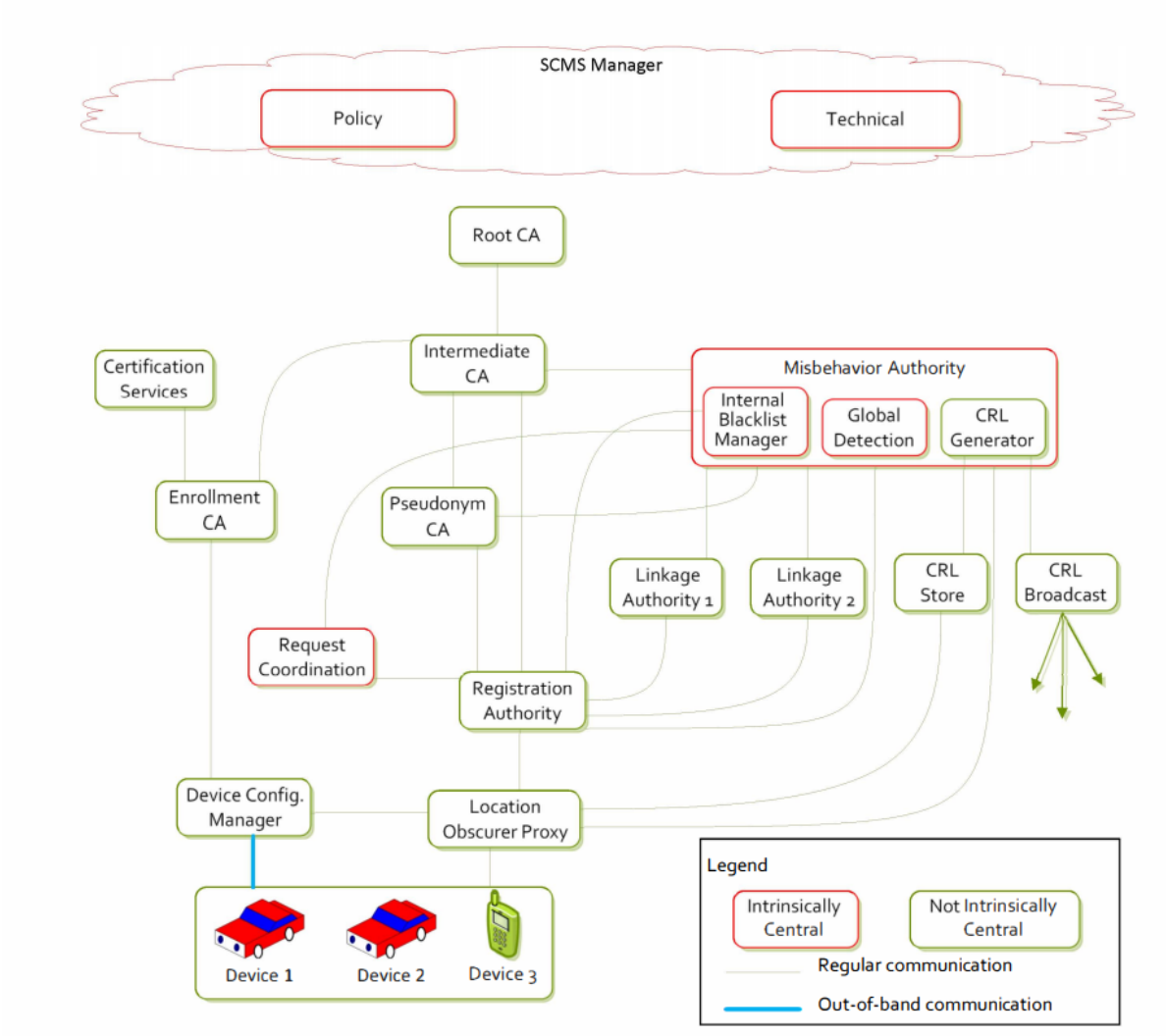
(2013 VNC)A Security Credential Management System for V2V Communications

1 Solved problem

In this paper, the author proposed a novel secure credential management system for vehicle-to-vehicle communication. The design of this system tries to get a balance between enough security and reasonable cost and flexibility.

2 Main idea

The system's overall architecture can be seen follow:



The design goal for this system is to prevent the following attack:

- Attacks on end-users' privacy from SCMS outsiders.
- Attacks on end-users' privacy from SCMS insiders.
- Authenticated messages leading to false warnings.
  A novel pseudonym certificate provisioning model and a butterfly cryptographic algorithm is used to achieve these goals.

3 Highlights worth learning

- The design philosophy that no single component knows information enough to track a device.
- The butterfly key expansion used to encrypt the certificate.

Security of Deep Learning based Lane Keeping System under Physical-World Adversarial Attack

1 Solved problem

In this paper, the author proposed a new attack method which is specific to the autonomous vehicles. The method uses dirty road patch on road line to compromise the Lane-Keeping Assistance System(LAKS).
The author formulate the attack as an optimization problem and use a gradient-based method to generate perturbations drawn on the road patch.
The experiments shows that the method proposed can compromise LAKS within 1.3 seconds, which is less than the average response time of drivers.

2 Main idea

The author regard this attack problem as an optimization problem. But the problem has a challenge that the attackers should affect a consecutive frames to achieve the attack goal.
So the author proposed a method that uses a perspective transformation to generate consecutive frames. Because the sizes of the patch in consecutive frames are different a aggregation gradient is used to align result from different frames.

3 Highlights worth learning

- The prospective transformation used to generate continuous frames of camera.
- The align method used to generate perturbation patches.