Attack-Resilient Sensor Fusion for Cooperative Adaptive Cruise Control

Pengyuan Lu∗1, Limin Zhang∗2, B. Brian Park3, Lu Feng1

Abstract—Cooperative adaptive cruise control (CACC) has the potential to enable vehicle platooning and achieve benefits including improved highway throughput and reduced energy consumption. However, malicious attacks such as sensor jamming or data injection can lead to security vulnerabilities of vehicle platooning and cause catastrophic crashes. We present a novel attack-resilience sensor fusion method for vehicle platooning with CACC, which exploits spatial information provided by multiple vehicles and combines sensor readings to achieve more precise estimation. We demonstrate the feasibility of our method in a set of simulated vehicle platooning experiments with different CACC controllers and malicious attacks.

I. INTRODUCTION

Cooperative adaptive cruise control (CACC) is a promising technology that can enable vehicle platooning where a group of vehicles’ movements are coordinated, with each vehicle autonomously follows the one in front of it. Several field demonstrations including California PATH program [1] and European Truck Platooning Challenge [2] have showed advantages of vehicle platooning via CACC, including significant mobility improvement (e.g., vehicle throughput could be tripled if CACC can maintain 0.6 seconds between vehicles compared to typical vehicles having 1.8 seconds between vehicles) and fuel efficiency (e.g., platooning trucks can achieve up to 17% fuel savings [3]). Meanwhile, safety and security vulnerabilities of vehicle platooning are drawing increasing attention. There are a variety of potential cyberattacks on automated vehicles [4], ranging from GPS spoofing, to sensor jamming, to CAN messages injecting, etc. One study [5] shows that a single, maliciously controlled vehicle can destabilize the entire vehicle platoon, causing catastrophic effect. A more recent study [6] demonstrates the effect of attacking CACC systems through jamming and data injection.

Modern automotive vehicles have many sensors (e.g., GPS, Radar, and Lidar) that can be used to measure and estimate the same physical variable (e.g., velocity, acceleration, distance to front vehicle). Sensor fusion is a technique that can combine the readings of diverse sensors with different precisions to achieve a more accurate estimation of the physical variable. An attack-resilient sensor fusion method

was proposed in [7], which can obtain precise estimations by combining the data received from all sensors, even when a subset of these sensors are compromised under malicious attacks. This method was validated in a case study involving a single automated ground vehicle.

In this paper, we present a new attack-resilient sensor fusion method for multiple vehicle platooning with CACC. Our key insight is that sensor fusion accounting for spatial information provided by multiple vehicles in the platooning can achieve better performance than the previous method that only takes into account sensor data from a single vehicle. We implement a highway vehicle platooning scenario inspired by the Grand Cooperative Driving Challenge [8] in the PreScan simulation platform [9]. We compare the performance of our sensor fusion algorithm with the methods proposed in [7] in this simulated scenario, with two different CACC control approaches (i.e., linear controller and model predictive controller) and under a wide variety of attacks (e.g., jamming, data injection, sensor manipulation).

The rest of the paper is organized as follows. We introduce the motivating vehicle platooning scenario and CACC control approaches in Section II. We describe the attack methods in Section III and present sensor fusion algorithms in Section IV. We discuss the experimental results in Section V and draw conclusions in Section VI.

II. MOTIVATING CACC SCENARIO

We consider a motivating scenario of highway vehicle platooning inspired by the Grand Cooperative Driving Challenge [8]. Figure 1 shows the scenario with a leading vehicle (i.e., red vehicle), which introduces acceleration disturbances by braking and accelerating, and two competitive vehicle platoons. In our simulation experiments, we apply the attacks and sensor fusion to one vehicle platoon, while keeping the other vehicle platoon as the control group for comparison.

Vehicles in the platoon uses CACC control system to automatically accelerates and decelerates so as to keep a desired distance to the preceding vehicle. Each vehicle constantly intakes physical measurements of environmental variables and outputs corresponding throttle or brake force for vehicle. In addition, CACC architecture allows vehicle-to-vehicle (V2V) communication among vehicles in the platoon.
Hence, the vehicles are able to transmit their current states, such as position, velocity and acceleration, and utilize the information received for better coordination. The V2V channel is digital, with a reliable transfer protocol. We assume that each vehicle in the motivating scenario has four radars, one GPS, and one V2V component for digital communication. Radars and GPS sensors are noisy and vulnerable to attacks. The V2V communication is assumed noiseless, but it is a potential target for malicious attacks.

Figure 2 illustrate the CACC concept with sensor measurements and V2V communication in our scenario. Vehicle i in the platoon uses its onboard Radar sensors to measure the distance $d_i$ to the preceding vehicle $i-1$ and transmits its GPS position $p_i$, velocity $v_i$ and acceleration $a_i$ to the following vehicle $i+1$. The CACC controller then decides the throttle or brake percentage based on the estimated headway distance and relative velocity to the front car. We consider two different types of CACC controllers proposed in [8].

1) Linear Controller: The linear controller utilizes the errors in the state variables and computes its transition as a linear combination of the errors. Vehicle i’s position error $ep_i$ and velocity error $ev_i$ are respectively defined as:

$$ep_i = (p_{i-1} - p_i) - d$$  
$$ev_i = v_i - v_{i-1}$$

where $d$ is an ideal distance between two consecutive cars. Consequently, the linear controller gives the acceleration of a vehicle by:

$$a_i = -(K_p \cdot ep_i + K_v \cdot ev_i)$$

where the gains $K_p$ and $K_v$ are empirical results which produce a stable platoon when free of attacks or defenses, i.e. the errors are reasonably bounded and eventually converge to 0.

2) Model predictive controller: The model predictive controller predicts the next state after optimizing an objective function and satisfying a set of constraints. Here, the constraints are based on each vehicle’s safety, performance and passenger convenience. In other words, the position error, the velocity error and the jerk are minimized while fulfilling the vehicle dynamics and actuator limits.

Both the linear controller and the model predictive controller aim to achieve the safety and string stability of the vehicle platoon. In the motivating scenario, safety property requires that a safe minimum distance should be maintained from the preceding vehicle to reduce the risk of collision, while string stability of a platoon is defined with respect to the spacing error, i.e., the spacing errors between vehicles are not amplified when propagate toward the tail of the platoon.

### III. ATTACKS

In this paper, we consider abstract sensor model [10] which interprets the measurement of a sensor at a time as a random interval: The larger the uncertainty, the larger its size. We assume that a healthy sensor guarantees its ground truth to lie within the interval. In reality, all sensors transmitting and receiving analog signals have noises. To restore the reality, we assume fixed relative errors for radar measurements and fixed absolute errors for GPS measurements. On the other hand, the V2V communication channel transmits digital signals with an assumed reliable protocol, which means negligible packet loss or corruption. Nonetheless, all the sensors are vulnerable to malicious attacks.

Different from noises, malicious attacks are assumed to be conducted by an adversary, who has full knowledge of the system and intends for the most severe traffic accidents. Via various ways can the attacker undermine a car platoon, from physical damage to highly skilled hijacking [4]. While all the attacks cause certain levels of damages, some are worth no research due to little expertise required, whereas some demand so much cost and skills that they merely occur. In this paper, we specifically focus on the following three possible attacks: jamming, data injection and sensor manipulation. They worth the most discussion due to the facts that these attacks: (1) are capable of causing vehicle performance issues, potentially leading to accidents; (2) require medium cost to be executed; (3) require medium level of expertise to be executed and defended.

**Jamming.** Radar jamming is usually caused by interference of a malicious signal, causing an additional and usually significant noise at the exploited sensors[11]. In the motivating scenario, we implement jamming by adding a band-limited white noise at the already noisy output of sensor j on vehicle i without corrupting the interval size:

$$s'_{ij}(t) = [s_{ij}^{min}(t) + w(t), s_{ij}^{max}(t) + w(t)]$$

where $W[0,T] = \int_0^T w(t)dt$ is a band-limited white noise, during the entire time of simulation $T = 63$ seconds.

**Data Injection.** Data injection hijacks one or some of the sensors and purposely gives predefined false information. We consider a typical data injection, namely ghost vehicle, which deceives the sensor that the obstacle is at a different distance than in reality. Usually, the adversary cheats the sensor with a larger range, causing the vehicles to collide:

$$s'_{ij}(t) = [s_{ij}^{min}(t) + d', s_{ij}^{max}(t) + d']$$

where $d'$ is the data injected to deceive the sensor, shifting the position of the obstacle. Such attack method is also known as a ghost vehicle attack.

**Sensor Manipulation.** Literally, all attacks on the signals at the sensors are sensor manipulations. Here we define the
attack specifically for undermining a sensor’s precision by enlarging its output interval:

\[ s'_{ij}(t) = [s_{ij}^{\text{min}}(t) + \gamma_l, s_{ij}^{\text{max}}(t) + \gamma_u] \]

where \( \gamma_u - \gamma_l > 0 \). In implementation, we have fixed the minimal value while enlarging the maximal value, cheating the car with a larger headway distance to induce collision.

Figure 3(a) shows three internals representing the measurement of three healthy sensors \( s_{i1}, s_{i2}, \) and \( s_{i3} \). The red dashed line represents the ground truth value \( \theta_i \). The larger the internal, the less precision the sensor. Figure 3(b), (c) and (d) illustrate the abstract sensor models when \( s_{i2} \) is under different attacks of jamming, data injection and sensor manipulation, respectively. The grey dashed lines represent the measurement of \( s_{i2} \) under attacks.

### IV. SENSOR FUSION

Sensor fusion is a strategy to recover the ground truth of a physical measurement, using multiple sensors with noises and fewer than half vulnerable to attacks. In the following, we first describe two existing sensor fusion algorithms in the literature, and then present a new sensor fusion method for vehicle platooning.

#### A. Naive and Pairwise Sensor Fusion

The simplest naive sensor fusion algorithm [7] takes intersection of all intervals returned. For a vehicle \( i \) with \( n \) sensors, the sensor fusion yields

\[ s_i = \bigcap_{j=1}^{n} s_{ij} \]  

(7)

Each abstract sensor model internal shall contain the ground truth value. When attack-free, the intersection recovers the smallest interval that guarantees the real measurement. If the intervals are vulnerable, the order of intersection matters. Figure 4(a) illustrates the naive sensor fusion with abstract sensor models \( s_{i1}, s_{i2}', \) and \( s_{i3} \), of which \( s_{i2}' \) is comprised by malicious attacks. The fusion algorithm takes the intersection of all three intervals and the output is an interval \( s_i \), which contains the ground truth value \( \theta_i \).

A second fusion algorithm uses additional temporal information was proposed in [7] to enhance the sensor fusion. The pairwise intersection uses historical intervals from some previous time, maps them to present with the vehicle’s dynamics and intersects corresponding interval pairs. Then, all resulting intervals involve in a final intersection to recover the ground truth. For measurement \( x \) with mapping function \( m(x) \) from time \( t - \delta t \) to \( t \):

\[ s_i(t) = \bigcap_{j=1}^{n} (s_{ij}(t) \cap m(s_{ij}(t - \delta t))) \]

(8)

This algorithm is effective to exclude sensors with inconsistent information at two times. Faulty results occur if both measurements are incorrect. We can further improve this algorithm by shutting down a sensor forever as soon as its pairwise intersection is empty. Nevertheless, it requires extra buffer for the historical data and the mapping function \( m(x) \) might be complicated in reality, with intermediate variables also noisy and vulnerable. Figure 4(b) shows the pairwise sensor fusion using historical data. Assume that sensor \( s_{i1} \) is under attack. Note that since the intersection of \( s_{i1}'(t) \) and the mapping of its historical measurement \( m(s_{i1}(t - 1)) \) is empty, this pair of internals are excluded from the fusion.

#### B. A New Sensor Fusion Method for Vehicle Platoons

We now present a new sensor fusion method by leveraging the spatial information of multiple vehicles in a platoon. As shown in Figure 5, suppose vehicle \( i \) not only receives the range information \( p_{i-1} - p_i \), but also the additional range information passed by vehicle \( i - 1 \) (i.e. \( p_{i-2} - p_{i-1} \)), and the position of vehicle \( i - 2 \) (i.e. \( p_{i-2} \)). Ideally, the distance between two cars can be computed as:

\[ p_{i-1} - p_i = (p_{i-2} - p_i) - (p_{i-2} - p_{i-1}) \]

(9)

We can measure \( p_{i-2} - p_{i-1} \) by the sensors of vehicle \( i - 1 \). Consequently, using the distance from vehicle \( i \) and \( i - 2 \) to subtract all the intervals from the last car, \( s_{i-1,j} \), we receive an extra set of intervals for sensor fusion. We name this new algorithm triangular pairwise intersection, since it uses the vector difference of two edges in a triangle to enhance the
measurement of the third edge:

\[ s_i = \bigcap_{j=1}^{n} (s_{ij} \cap (p_{i-2} - p_i - s_{i-1})) \]  

(10)

where \( p_{i-2} - p_i - s_{i-1} \) is the interval

\[ [p_{i-2} - p_i - s_{i-1}^{max}, p_{i-2} - p_i - s_{i-1}^{min}] \].

**Algorithm 1 Triangular pairwise intersection**

1: procedure TRIANGULAR\((S_i, S_{i-1}, p_{i-2})\)
2: \( p_i \leftarrow \text{gps\_get\_position()} \)
3: \( s_i \leftarrow U \)
4: while \( S_i \neq \emptyset \land S_{i-1} \neq \emptyset \) do
5: \( s_{ij}^{\min} \leftarrow p_{i-2} - p_i - s_{i-1}^{max} \)
6: \( s_{ij}^{\max} \leftarrow p_{i-2} - p_i - s_{i-1}^{min} \)
7: \( s_{ij} \leftarrow [s_{ij}^{\min}, s_{ij}^{\max}] \)
8: \( s_i \leftarrow s_i \cap (s_{ij} \setminus s_{ij}) \)
9: \( S_i.\text{pop}(s_{ij}) \)
10: \( S_{i-1}.\text{pop}(s_{i-1}) \)
11: return \( s_i \)

Figure 5 illustrates the idea of triangular sensor fusion. Algorithm 1 shows the algorithmic procedure of computing an fusion output of triangular pairwise intersection. Triangular pairwise intersection can be generalized in higher dimensions; unlike pairwise intersection using historical data, it does not require a well-designed mapping function. Comparing to the topology where all members calibrate with the lead car, this neighboring transmission helps reduce error propagation and lower the risk of lead car being attacked. In addition, triangular pairwise intersection can also produce fusion output with higher confidence than pairwise intersection with temporal information, which may intersect false intervals at both time \( t - \delta t \) and \( t \). The experimental results in Section V will demonstrate that the new sensor fusion method is an effective defense for vehicle \( i \), assuming vehicle \( i - 1 \) is not under attack.

V. RESULTS AND DISCUSSION

In this section, we show the results of running experiments of the motivating scenarios in the PreScan simulation platform [9]. PreScan is physics-based simulation platform that is used in the automotive industry for designing and evaluating autonomous driving applications with realistic sensor technologies such as radar, laser/lidar, camera, and GPS, as well as vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication applications. In the experiment, we use GPS and radar. Tracking the GPS data along time gives the position and velocity. Both simulated sensors are real and we hard-coded their noises. We compare the performance of our new sensor fusion method with existing algorithms when the vehicle platoon is under different attacks and controlled by different CACC control methods.

A. Results of the Attacks

Figures 7-9 show the PreScan simulation screen shots and the vehicle position errors (i.e., a metric showing the string stability of the platoon) when the platoon is equipped with linear CACC controller and under the attacks of jamming, data injection and sensor manipulation, respectively. The bottom vehicle platoon is attacked, while the top platoon is attack-free and serves as a control group for comparison. We use the empirical coefficients \( K_p = 0.8 \) and \( K_v = 5 \) for the linear controller, which ensures the baseline performance, i.e. the platoon performs stably and efficiently with absence of attacks or defenses. In this case, the controller simply takes the average reading of all sensor intervals. All the attacks are performed on the most precise radar and the GPS.

The graphs plot the positional errors of all vehicles against time, with vehicle 3 (i.e. the third vehicle from left to right) attacked defenselessly. Figure 7 shows that the position error \( ep_3 \) significantly exceeds that of all the others due to the jamming attack, which oscillates the measurements in a random pattern. Consequently, the linear controller keeps falsely consider the front vehicle vibrating from very far to near. The perceived near points matter: they prevent the controller to give sufficient throttle force and hence hinder the car. As a result, vehicle 3 has under-performed.

More severe results are shown under data injection and sensor manipulation, where the vehicles collide. Figures 8 and 9 both show that \( ep_3 \) has gone below -15 meters at some point. Due to the default distance set is 15 meters, the
results present negative distances between vehicle 2 and 3; in other words, traffic accidents occur and the data afterwards is meaningless. Such collisions are due to both attacks deceiving the controller that the front vehicle is further than the reality. In data injection attack, the ghost vehicle of vehicle 2 is placed 20 meters ahead of the actual position, and in sensor manipulation, the attacked interval is fixed at the minimum while increased at the maximum. Hence, in both cases the average of the perceived headway distance is larger than the fact, leading the controller to falsely produce more throttle force than sufficient. Were the perceived headway distance to be smaller, the results will be under-performance as well.

B. Results of the Sensor Fusion Algorithms

All the three sensor fusion algorithms excellently recover the positional information. The comprehensive quantitative results are visualized in the table shown in Figure 11 which presents the range of $e_p \triangleq$ each experiment trial. The results show that: (1) Regardless of the controller, the type of attack and the attack power, all the three algorithms recover the vehicle performance, allowing the vehicle to be safe and efficient. (2) Generally, based on the range that $e_p$ is controlled inside, the three sensor fusion algorithms behave increasingly favorable in the order of: naive, pairwise intersection using historical data and triangular pairwise intersection.

Figures 11-13 show that the positional errors of triangular pairwise intersection is generally smaller than the other two sensor fusion methods under different attacks, assuming that the front vehicle is not attacked. Hence, the additional intervals from V2V give accurate spatial information. On the other hand, pairwise intersection using temporal data uses previous corrupt information to support itself, and the pair usually gets rejected based on the algorithm, leaving the healthy yet inaccurate sensors for the controller. An alternative way to analyze the results is to check the interval sizes after fusion. All the three graphs show that the interval sizes generally shrink from the naive to the temporal then to the spatial algorithm, which means the algorithms increasingly succeed in reducing uncertainties in that order.

VI. CONCLUSION

We present a novel attack-resilient sensor fusion method for vehicle platooning, using spatial information exchanged in cooperative adaptive cruise control. Experimental results via PreScan simulation show that our new algorithm outper-
forms traditional sensor fusion methods, in both maintaining the positional error in a smaller range and suppressing the uncertainty of the fusion result. For the future work, we will explore the proposed method in vehicle platoons with a larger number of vehicles and heterogeneous types of vehicles. Also, comparing different topological structures can be a promising research topic.

REFERENCES