Distributed Cyber Attacks Detection and Recovery Mechanism for Vehicle Platooning

Eman Mousavinejad, *Student Member, IEEE*, Fuwen Yang, *Senior Member, IEEE*, Qing-Long Han, *Fellow, IEEE*, Xiaohua Ge, *Member, IEEE*, and Ljubo Vlacic, *Senior Member, IEEE*

***Abstract*—This paper is concerned with the distributed attack detection and recovery in a vehicle platooning control system, wherein inter-vehicle information is propagated via a wireless communication network. An active adversary may launch malicious cyber attacks to compromise both sensor measurements and control command data due to the openness of the wireless communication. First, a distributed attack detection algorithm is developed to identify any of those attacks. The core of the algorithm lies in that each designed filter can provide two ellipsoidal sets: a state prediction set and a state estimation set. Whether a filter can detect the occurrence of such an attack is determined by the existence of intersection between these two sets. Second, two recovery mechanisms are put forward, through which the adversarial effects of cyber attacks can be mitigated in a timely manner. The recovery mechanisms depend on reliable modifications of the attacked signals required for the computation of the two ellipsoidal sets. Finally, simulation is provided to validate the effectiveness of the proposed method in both detection and recovery phases.**

***Index Terms*—Vehicle platooning, cyber attacks, attack detection, recovery mechanism, set-membership filtering.**

I. Introduction

With the growing demand for mobility and development of urbanization, the number of vehicles has been significantly increased in recent years. As a result, there has been mounting concern about modern transportation systems due to traffic congestion, traffic accidents, energy waste, and pollution. To deal with these issues, developing intelligent transportation systems (ITS) technologies for the driving pattern change from individual driving to platoon-based driving is of utmost importance. As an advanced automated technology, vehicle platooning is aimed at steering a team of vehicles into a platoon on a road where all vehicles' speed can be automatically adjusted such that the inter-vehicle distance is reduced while not compromising safety.

The cooperative adaptive cruise control (CACC), which extends the conventional adaptive cruise control (ACC) technology [1], is deemed as one of the most developed ITS technologies employed in vehicle platooning systems. The objective of CACC is to guarantee that all vehicles in the platoon achieve a prescribed leader vehicle's velocity/acceleration and meanwhile preserve a desirable inter-vehicle distance. To fulfil this, vehicles are usually equipped with suitable on-board sensors, such as radar, camera, and lidar, enabling each of them to monitor its predecessor's behavior. Moreover, vehicles exchange their inter-vehicle data used by CACC among their neighbors via short-range to medium-range wireless communication channels so as to coordinate their behaviors. From this perspective, a platoon-based vehicular control system can be regarded as a multi-agent system, where vehicles represent mobile agents connected together via a vehicular ad-hoc network (VANET).

Recently in ITS industry, considerable research interest has been paid to the impact of network-induced phenomena in vehicle platooning and the improvement of string stability performance, see [2]–[15]. Besides, with the ever-increasing utilization of communication networks in ITS industry, a multitude of studies have investigated the impact of cyber attacks on VANET security because of the openness of the communication networks. Readers are referred to the survey papers [16]–[20] and many references therein for some latest results. Generally, there exist mainly two types of cyber attacks which can deteriorate the performance of VANET, namely *denial-of-service (DoS) attacks* and *deception attacks* [21], [22]. DoS attacks are aimed at jamming the communication channels in order to prevent information exchange between vehicles. From a technological perspective, an adversary can perpetrate DoS attacks through disrupting the radio frequencies on wireless communication channels, which leads to congestion of those channels. Platoon behavior under DoS attacks is investigated in [23], where a drone flying above the platoon, namely the adversary, disrupts the platoon by using its limited power. It is shown that the best possible location to carry out an attack with a huge adverse impact is above the second vehicle and the impact decreases when the adversary moves down in the string. A real-time detection scheme for DoS attacks is presented in [24]. The effects of the DoS attacks are modeled by time-delayed data transmission via a communication network and the attack is diagnosed through tracking the delay in information processing by a set of observers. Deception attacks, on the other hand, represent a kind of attempt to violate the integrity of sensor and/or control data, thus manipulating the vehicle toward the desired behavior of the adversary. To launch a successful deception attack, the adversary usually has sufficient knowledge of the information transmitted through VANET in real time. For example, an adversary may be a trusted insider, i.e., an authenticated member of the platoon, so that the adversary can discard the pseudonymous certificates and digital signatures employed to

This work was supported by the Australian Research Council Discovery Project under Grant DP160103567. (*Corresponding author: Q.-L. Han*).

E. Mousavinejad, F. Yang, and L. Vlacic are with the School of Engineering and Built Environment, Griffith University, Gold Coast, QLD 4222, Australia (Emails: eman.mousavinejad@griffithuni.edu.au (E. Mousavinejad); fuwen.yang@griffith.edu.au (F. Yang); l.vlacic@griffith.edu.au (L. Vlacic)).

Q.-L. Han and X. Ge are with the School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, VIC 3122, Australia (Emails: qhan@swin.edu.au (Q.-L. Han); xge@swin.edu.au (X. Ge)).

protect the content of beacon messages. Two typical types of deception attacks on VANET are *message falsification attacks* [25]–[27] and *replay attacks* [16], [17]. In message falsification attacks, an adversary starts collecting the information transmitted through the wireless medium and simultaneously modifies the content meaningfully such that the adversary can severely damage the string stability by rebroadcasting bogus information. The impacts of cyber attacks and sensor tampering are described in [25] through launching a message falsification attack by an external vehicle into the platoon. It is shown that degrading the CACC strategy to ACC can be a potential countermeasure to this attack. In [26], a two-component controller consisting of a radar-based proportional-derivative (PD) feedback component and a communication-based feedforward component is proposed to design an attack detection strategy. The proposed attack detection algorithm is based on the estimate of the expected behavior of the front vehicle and the switching policy to ACC if any abnormal behavior exists. String stability of a vehicular platoon under message falsification attacks is investigated in [27], where an attacker is able to force the platoon to oscillate at a resonant frequency by modifying the control signal through which he may cause a serious accident. Replay attacks, however, do not require prior knowledge of the system components. It can be easily implemented by an adversary in two phases. In the first phase, which is known as disclosure phase, the adversary overhears and stores the packets transmitted over communication channels without injecting any malicious input to the system. In the second phase, the adversary manipulates the system by replaying the recorded data as if they are new packets received from those tampered channels. Even though the content of the vehicle data packets is not falsified by the replay attacks, any outdated information may mislead the platoon members, having a great potential to destroy the platoon stability. For example, the attacker may modify a following vehicle's control command, which contains the desired acceleration received from its predecessor, by some outdated information with the aim of speeding up the following vehicle and causing a collision.

The lack of attention to detection of cyber attacks and recovery of the vehicular platoon system has led the vehicle industry to raise concern about the risks of cyber attacks in this system, which constitutes one of the most challenging issues when a vehicle platoon control system is commercialized in highways [28]. Therefore, the ongoing research on vehicle platooning includes significant efforts to develop effective attacks detection and protection strategies. This motivates us to develop an attack detection strategy and its associated recovery mechanism so as to not only detect attacks that maliciously disrupt the performance of VANET but also recuperate the system from the detected attacks in a timely fashion.

Note that most of attack detection approaches [29]–[40], particularly the χ^2-detector based on the celebrated Kalman filters, require system noises in a stochastic framework. Whereas, a stochastic noise model requires particular statistical properties of the noise such as known mean and covariance, which may be unrealistic in some practical situations in vehicular systems. For example, acceleration of a vehicle is always deterministically bounded because of the engine dynamics and construction of the vehicle, but becomes uncertain when the vehicle is running. In this case, noise on control data and sensor measurements of a vehicle can be assumed to be unknown-but-bounded (UBB). Indeed, an assumption of UBB noise eliminates prior knowledge of the accurate statistical characteristics of noise because only the knowledge of a bound on the realization is needed. Therefore, an alternative filtering method called set-membership filtering is developed in the literature [41]. This filtering method aims to calculate a bounding ellipsoidal state estimation set in state space, which always encloses the true state of the system [42], [43]. During the past decade, such a filtering method has been intensively studied for different problem formulations of various system models (see [44]–[48]).

Motivated by the discussion above, the main objective of this paper is twofold. 1) A distributed attack detection issue for vehicle platooning will be addressed. In our preliminary work [49], a centralized cyber attack detection method based on ellipsoidal filtering has been developed for general linear time-varying systems. However, this method is not straightforwardly applicable for a large-scale vehicular platoon because it requires full knowledge of the entire platoon information. Furthermore, the computational overhead for this detection method is quite high, thereby rendering the detection system inadvisable. To overcome these limitations, a distributed attack detection method will be developed in this paper in the sense that each vehicle is equipped with its own detection system. More specifically, each vehicle has its own filter to estimate the state of the vehicle without having the full knowledge from other agents states, which helps apply such a system to a large-scale platoon and decrease the computational overhead. 2) It is our intention to introduce two recovery mechanisms to mitigate the adversarial impacts of the attacks on the performance of the vehicle platooning system, which represents a distinct difference between this paper and our preliminary work [49]. With these two recovery mechanisms, the system can be brought back to the normal condition after detection of the attacks. In other words, the recovery mechanisms will make the state estimation secure against the attacks and then the controller will use the secure estimation to generate the desired control command.

The major contribution is threefold. 1) *A set-membership filtering technique in a distributed framework* is developed such that each vehicle can determine a group of confidence ellipsoids and each vehicle's true state always resides in bounding ellipsoidal sets regardless of UBB process noise and measurement noise, providing the vehicle is free of any attack. In particular, each vehicle is equipped with a set-membership filter constructing two ellipsoidal sets: a prediction set and an estimation set. 2) *A refined attack detection method* is proposed to discern when the occurrence of an attack can be detected and alarmed. The alarm is triggered once there exists no intersection between the two ellipsoidal sets. When the system is free of attacks, the two sets must both enclose the true state, thus guaranteeing the existence of the intersection. When the system is subject to attacks, however, the center of at least one of the two sets is biased by the attacks, which excludes the intersection between the ellipsoidal sets. 3) *Two recovery mechanisms* are introduced into the proposed algorithm so that

each ellipsoidal set will adopt the compensated state prediction and/or state estimation to increase the resilience of estimation system. In other words, the recovery mechanisms promise reliable modifications of the attacked signals required for the computation of the prediction and estimation ellipsoidal sets. Based on the proposed recovery mechanisms, the controllers in CACC system are able to alleviate the adversarial effect of an attack through accessing the secure state estimation.

The rest of this paper is organized as follows. Section II formulates the longitudinal vehicle dynamics and the control strategy, which together form the CACC-equipped vehicle that will be employed in this paper. Besides, the models of some potential cyber attacks that will be studied in this paper are constructed in this section. The attack detection criteria under a two-step set-membership filtering method are also provided in this section. Section III presents the design of the prediction and updated estimation ellipsoidal sets, and then the associated attack detection algorithm and recovery mechanisms in a distributed manner are proposed at the end of this section. The simulation results are demonstrated in Section IV to illustrate the effectiveness of the proposed method. Finally, Section V concludes the paper.

Notation: For a symmetric matrix P, the notation $P > 0$ means that P is positive definite. $P \geq 0$ means that P is positive semi-definite. The notation $\mathrm{Tr}(P)$ denotes the trace of P. An ellipsoidal set is denoted as $\mathcal{X} \triangleq \{\zeta \in \mathbb{R}^n : (\zeta - c)^T P^{-1}(\zeta - c) \leq 1\}$, where $c \in \mathbb{R}^n$ is the *center* and $P = P^T > 0$ is the *shape matrix* of the ellipsoid. Let $E \in \mathbb{R}^{n\times m}$ with $\mathrm{rank}(E) = m \leq n$ be a lower triangular matrix in which each diagonal element is positive. By a Cholesky factorization, one has that $P = EE^T > 0$. Hence, an alternative representation of the ellipsoidal set is $\mathcal{X} \triangleq \{\zeta : \zeta = c + Ez, \|z\| \leq 1\}$. The *size* of the ellipsoid is a function of the squared shape matrix P which can be measured by means of $\mathrm{Tr}(P)$, i.e., the sum of squared semiaxes lengths. Other notations in this paper are quite standard.

II. Problem Formulation

A. CACC Law

A platoon-based vehicular control system with some possible attack points is shown in Fig. 1. To model the dynamics of each vehicle in the string, the linear longitudinal dynamic model presented in [5] is adopted for each vehicle in the string as

$$\dot{q}_i = v_i, \quad \dot{v}_i = a_i, \quad \dot{a}_i = -\eta_i^{-1} a_i + \eta_i^{-1} u_i, \quad i \in \mathbb{N}_{[1,n]} \tag{1}$$

where q_i, v_i, and a_i denote the absolute position, velocity, and acceleration of vehicle i, respectively; η_i is the internal actuator dynamics parameter; and u_i represents the commanded acceleration for vehicle i. Let us consider vehicle i's state variables as $x_i = [d_i, v_i, a_i, v_{i-1} - v_i, a_{i-1} - a_i]^T \in \mathbb{R}^{n_x}$, where d_i, $v_{i-1} - v_i$, and $a_{i-1} - a_i$ represent the inter-vehicle distance between two consecutive vehicles, their relative velocity, and acceleration, respectively.

Considering process and measurement noises in a UBB sequence for each vehicle in the string, the discrete-time dynamics model of the i-th CACC-equipped vehicle at sampling instants $t_k = kh, k \in \mathbb{Z}^+$ can be expressed as

$$x_{k+1}^i = A_i x_k^i + B_{s,i} u_k^i + B_{c,i} u_k^{i-1} + F_i w_k^i \tag{2}$$

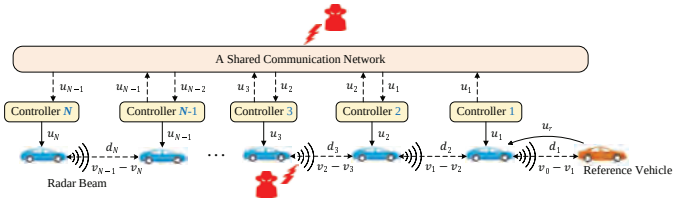


Fig. 1. A platoon-based vehicular control system with some possible attack points.

where A_i is the state matrix of vehicle i; $B_{s,i}$ denotes the input matrix of vehicle i; and $B_{c,i}$ represents the input matrix from vehicle $i-1$ through the wireless communication network. These matrices can be obtained by

$$A_i = e^{\bar{A}_i h}, \; B_{s,i} = \int_0^h e^{\bar{A}_i s} ds \bar{B}_{s,i}, \; B_{c,i} = \int_0^h e^{\bar{A}_i s} ds \bar{B}_{c,i} \tag{3}$$

with

$$\bar{A}_i = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\eta_i^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -\eta_i^{-1} \end{bmatrix}, \bar{B}_{s,i} = \begin{bmatrix} 0 \\ 0 \\ \eta_i^{-1} \\ 0 \\ -\eta_i^{-1} \end{bmatrix}, \bar{B}_{c,i} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \eta_i^{-1} \end{bmatrix}.$$

The measurement output obtained from each individual distributed sensor on each vehicle is given by

$$y_k^i = C_i x_k^i + D_i v_k^i \tag{4}$$

where $y_k^i \in \mathbb{R}^{n_y}$ is the measurement on sensor i which is based on the inter-vehicle distance, the vehicle absolute velocity, and the relative velocity; F_i, C_i and D_i are known matrices with appropriate dimensions; $w_k^i \in \mathbb{R}^{n_w}$ and $v_k^i \in \mathbb{R}^{n_v}$ are the process noise and measurement noise of vehicle i, respectively.

Assumption 1: The noises w_k^i and v_k^i are UBB and confined to the following specified ellipsoids

$$\mathcal{W}_k^i \triangleq \{w_k^i : {w_k^i}^T {Q_k^i}^{-1} w_k^i \leq 1\} \tag{5}$$

$$\mathcal{V}_k^i \triangleq \{v_k^i : {v_k^i}^T {R_k^i}^{-1} v_k^i \leq 1\} \tag{6}$$

where $Q_k^i = {Q_k^i}^T > 0$ and $R_k^i = {R_k^i}^T > 0$ are known matrices with compatible dimensions.

Remark 1: A reference vehicle (denoted by index $i = 0$ and with state x_k^0) must be introduced as a trajectory generator in the lead vehicle in case there are no preceding vehicles. The discrete-time dynamics of the reference vehicle can be described by

$$x_{k+1}^0 = A_0 x_k^0 + B_{s,0} u_k^r \tag{7}$$

where $x_k^0 = [v_k^0 \quad a_k^0]^T$, $u_k^r = u_r(t_k)$ is the sampled reference acceleration profile at the sampling instants $t_k = kh$, and A_0 and $B_{s,0}$ are calculated in the same method as A_i and $B_{s,i}$ in (3) when $i = 0$, with

$$\bar{A}_0 = \begin{bmatrix} 0 & 1 \\ 0 & -\eta_0^{-1} \end{bmatrix}, \quad \bar{B}_{s,0} = \begin{bmatrix} 0 \\ \eta_0^{-1} \end{bmatrix}.$$

In practice, the reference vehicle normally serves as a command generator, thus not being affected by the followers. For this purpose, it is assumed that the reference vehicle's dynamics are not subject to UBB process noise but deterministic to the leader. In this sense, the leader can be informed by the reference vehicle. It should be also pointed out that the lead vehicle (denoted by index $i = 1$ and with state x_k^1) in the string requires special consideration.

Assumption 2: The reference acceleration profile u_k^r is locally available to the lead vehicle without any network-induced imperfection or any cyber attack impact.

It is noteworthy that Assumption 2 is mild since u_k^r is generated locally by the lead vehicle. Therefore, the dynamics of the lead vehicle can be represented by (2) with $i = 1$ and $u_k^{i-1} = u_k^0 = u_k^r$.

Assumption 3: The platoon is composed of homogenous vehicles whose longitudinal dynamic properties are identical.

In actual traffic with different types of vehicles, Assumption 3 can be achieved through adequately designed pre-compensators, i.e., low-level acceleration controllers. Hence, for a homogenous platoon, the index $i \in \mathbb{N}_{[1,n]}$ can be omitted from the system's matrices A_i, $B_{s,i}$, $B_{c,i}$, C_i, F_i, D_i and the vehicle parameter η_i in the rest of the paper.

In this study, the control strategy is similar to the one that is presented in [5], which is compatible with the structure shown in Fig. 1. This control strategy requires a combination of a communication-based forward component $u_{ff,k}$ and a measurement-based feedback component $u_{fb,k}$. Thus, the total control signal can be written as

$$u_k^i = u_{ff,k}^i + u_{fb,k}^i \tag{8}$$

with the discretized feedforward control signal in the form of

$$u_{ff,k+1}^i = (1 - hh_d^{-1})u_{ff,k}^i + hh_d^{-1}u_k^{i-1} \tag{9}$$

where h_d is the headway-time for vehicle i to arrive at the same position as its predecessor under a constant time headway (CTH) spacing policy [5].

Different from [5], which utilizes the state measured by sensors directly into calculation of the control signal's feedback component, we are interested in employing the state estimate. This is because in real-world applications, the vehicle system's state may not be fully available or not be implicitly trusted due to the restricted sensor perceiving and computing capabilities or faulty sensors. However, full availability of state information is vital for detection of an attack and its integrity is crucial in recovery of the violated system in a timely manner. Motivated by this discussion, the feedback signal can be expressed in an observer-based control protocol as

$$u_{fb,k}^i = K\Omega\hat{x}_k^i \tag{10}$$

where $K = [k_p \quad k_d]$ with $k_d = \omega_c$ and $k_p = \omega_c^2$, ω_c is the bandwidth of the controller and is chosen such that $\omega_c \ll 1/\eta$, $\Omega = \begin{bmatrix} 1 & -h_d & 0 & 0 & 0 \\ 0 & 0 & -h_d & 1 & 0 \end{bmatrix}$, and $\hat{x}_k^i$ is the estimate of state x_k^i.

Note that the control signal is sampled at the sampling instants $t_k = kh$ and then transmitted over the wireless network. Hence, the sampled and transmitted data may typically experience some network-induced constraints, such as data communication delays, data packet dropouts/disorders, channel fading, and quantization effects [50]. In this paper, however, we pose the following assumption on the wireless communication channels, which allows us to focus on the cyber attack detection and protection issue.

Assumption 4: The wireless communication network is free of network-induced constraints but suffers from cyber threats.

Remark 2: The control objective is to regulate the spacing error $e_i = d_i - d_{r,i}$ such that e_i goes to zero asymptotically, where the desired spacing between two consecutive vehicles under CTH policy is represented as

$$d_{r,i} = h_d v_i.$$

The spacing error is produced by the leader's acceleration and deceleration. *String stability* requires spacing error attenuation as vehicles move upstream in the string. This phenomenon is well investigated in [5] and tested experimentally for the adopted control strategy (8). The time domain definition of string stability [23] is expressed as

$$\max_{t_k}\|e_n(t_k)\| < \max_{t_k}\|e_{n-1}(t_k)\| < \ldots < \max_{t_k}\|e_1(t_k)\| \tag{11}$$

Since string stability may be degraded in the presence of cyber attacks, a detection method will be introduced such that the attack can be detected in a timely manner and then, two recovery mechanisms will be considered so as to provide the control unit with the secure state prediction and estimation required to calculate the control signal. In this manner, the string maintains its normal performance and the string stability is satisfied in the existence of cyber attacks, which will be further investigated in Section IV.

B. Cyber Attack Model

In this paper, we investigate the scenario that there exists an adversary whose aim is to manipulate the data transmitted through the communication network, namely Υ_k^i, which represents either the inter-vehicle signal or the sensor measurement output of vehicle i. The under-attack signal $\tilde{\Upsilon}_k^i \in \mathbb{R}^q$ on vehicle i can be expressed as

$$\tilde{\Upsilon}_k^i = \Upsilon_k^i + \Lambda_k^i\sigma_k^i \tag{12}$$

where $\sigma_k^i \in \mathbb{R}^q$ is an attack signal carefully designed by the adversary and $\Lambda_k^i \in \mathbb{R}^{q\times q}$ represents the physical constraints imposed on the attack signals and is assumed to be of the following diagonal structure

$$\begin{cases} \Lambda_k^i = \text{diag}\{\lambda_{1,k}^i, \lambda_{2,k}^i, \ldots, \lambda_{q,k}^i\} \\ \underline{\lambda}_{j,k}^i \le \lambda_{j,k}^i \le \overline{\lambda}_{j,k}^i,\ j = 1, 2, \ldots, q \end{cases} \tag{13}$$

where $\underline{\lambda}_{j,k}^i$, $\overline{\lambda}_{j,k}^i \in [0, 1]$ are the scalars representing lower- and upper-bounds on $\lambda_{j,k}^i$.

Remark 3: In engineering practice, a realistic adversary may suffer from some physical constraints, such as device saturations, finite battery, and limited interference capacities. This means that the adversary may not be able to launch the attacks at all times. For this reason, a diagonal matrix Λ_k^i is exposed on the attack signal, as shown in (12). Note that the scalar parameters $\lambda_{j,k}^i$ can also be interpreted as the healthy status of the j-th component of the transmitted data under attacks. Specifically, if $\lambda_{j,k}^i = 0$, there is no attack on the j-th component of the transmitted data at time k and therefore, this component is securely transmitted. When $0 < \lambda_{j,k}^i < 1$, the j-th component is partially contaminated by the attack signal. The case $\lambda_{j,k}^i = 1$ represents the worst case scenario that the j-th component is fully corrupted by the attack signal.

Remark 4: The under-attack signal model in (12) includes different attack strategies as its special cases. For example, if the attack signal $\sigma_k^i = -\Upsilon_k^i$, there exist DoS attacks compromising the signal at time k. For replay attacks, it is assumed that the attacker can record sequences of data from a signal Υ_k^i from k_o till k_r with the window size $\tau = k_r - k_o$ in the first phase. In the second phase, the attacker replays the recorded data to the system from $k = k_r + d$ till the end of the attack at $k = k_f$, where d is the delay between the recording time and the replaying time. Hence, the replay attacks can be implemented if $\sigma_k^i = \Upsilon_{k-\tau}^i - \Upsilon_k^i$. The case $\sigma_k^i = \delta_k^i$, where

δ_k^i is any arbitrary signal chosen by the attacker, represents message falsification attacks.

C. Set-Membership State Estimation

Before proceeding further, let us provide some definitions about set-membership state estimation which is the foundation of the cyber attack detection algorithm. In what follows, a state estimation problem is formulated for each vehicle such that its state prediction and estimation ellipsoids can be calculated to guarantee the enclosing of the true vehicle state in an attack-free scenario.

The state estimation problem is formulated at two steps described as follows.

Prediction Step: We are interested in constructing a set-membership filter which runs the following prediction

$$\hat{x}_{k+1|k}^i = G_k^i \hat{x}_k^i + B_s u_k^i + B_c u_k^{i-1} \tag{14}$$

where G_k^i is the filter gain matrix sequence to be determined. The system described by (2) and (4) is said to achieve *set-membership state estimation* at the prediction step if there exists G_k^i such that each vehicle's state x_{k+1}^i resides in a prediction ellipsoid $\mathcal{X}_{k+1|k}^i$, which always contains the true state of vehicle i, where

$$\mathcal{X}_{k+1|k}^i \triangleq \{x_{k+1}^i : (x_{k+1}^i - \hat{x}_{k+1|k}^i)^T \times P_{k+1|k}^{i^{-1}} (x_{k+1}^i - \hat{x}_{k+1|k}^i) \leq 1\} \tag{15}$$

for any values of the process noise $w_k^i \in \mathcal{W}_k^i$ and the measurement noise $v_k^i \in \mathcal{V}_k^i$.

Measurement Update Step: The set-membership filter updates its state based on the current measurement of vehicle i. More specifically, the state update of the filter is given in the form of

$$\hat{x}_{k+1}^i = \hat{x}_{k+1|k}^i + L_k^i (y_{k+1}^i - \hat{y}_{k+1|k}^i) \tag{16}$$

where L_k^i is the filter gain matrix sequence to be determined and $\hat{y}_{k+1|k}^i = C\hat{x}_{k+1|k}^i$. The system described by (2) and (4) is said to achieve *set-membership state estimation* at the measurement update step if there exists L_k^i such that each vehicle's state x_{k+1}^i resides in an estimation ellipsoid $\mathcal{X}_{k+1}^i$, which always contains the true state of vehicle i, where

$$\mathcal{X}_{k+1}^i \triangleq \{x_{k+1}^i : (x_{k+1}^i - \hat{x}_{k+1}^i)^T P_{k+1}^{i^{-1}} (x_{k+1}^i - \hat{x}_{k+1}^i) \leq 1\} \tag{17}$$

whenever the equality

$$y_{k+1}^i = C\hat{x}_{k+1|k}^i + CE_{k+1|k}^i z + Dv_{k+1}^i \tag{18}$$

holds for some $\|z\| \leq 1$ and any values of the process noise $w_k^i \in \mathcal{W}_k^i$ and the measurement noise $v_k^i \in \mathcal{V}_k^i$.

The initial state x_0^i satisfies the following assumption.

Assumption 5: The initial state of vehicle i belongs to the given ellipsoid

$$\mathcal{X}_0^i \triangleq \{x_0^i : (x_0^i - \hat{x}_0^i)^T P_0^{i^{-1}} (x_0^i - \hat{x}_0^i) \leq 1\}. \tag{19}$$

D. Cyber Attack Detection Criteria

To achieve successful attack detection, the following detection criteria are developed to identify attacks violating the data measured by each vehicle's sensors in the platoon and the inter-vehicle data transmitted through the communication network.

1) If the prediction ellipsoidal set $\mathcal{X}_{k+1|k}^i$ has no intersection with the previous updated ellipsoidal set $\mathcal{X}_k^i$, i.e., $\mathcal{X}_k^i \bigcap \mathcal{X}_{k+1|k}^i = \varnothing$, it indicates that the communication network channel transmitting inter-vehicle data is affected by the attack.
2) If the updated ellipsoidal set $\mathcal{X}_{k+1}^i$ has no intersection with the prediction ellipsoidal set $\mathcal{X}_{k+1|k}^i$, i.e., $\mathcal{X}_{k+1}^i \bigcap \mathcal{X}_{k+1|k}^i = \varnothing$, it indicates that the sensor measurement of vehicle i is affected by the attack.

Remark 5: In the case that an attack is present, there exists one among the two sets not enclosing the true state because the center of that set is biased by the attack. From (14), if the attack violates the network channel that transmits the inter-vehicle signal u_k^{i-1} sent from vehicle $i-1$ to vehicle i, at time k, the center of the i-th prediction ellipsoidal set $\mathcal{X}_{k+1|k}^i$ is affected by the attack. Hence, the condition (15) may not be satisfied, which leads the prediction ellipsoidal set not to include the true state of the system. However, the previously updated estimation ellipsoidal set $\mathcal{X}_k^i$ is based on its prediction ellipsoidal set at time $k-1$ and thus not affected by the attack. Thus, one can conclude that if there is no intersection between $\mathcal{X}_{k+1|k}^i$ and $\mathcal{X}_k^i$, the transmitted signal u_k^{i-1} and its corresponding network channel are compromised by an adversary.

Remark 6: If the sensor of vehicle i is manipulated by an adversary at time $k+1$, its current measurement output y_{k+1}^i is violated. Hence, from (16), the center of the i-th estimation ellipsoidal set $\mathcal{X}_{k+1}^i$, which is updated with the current measurement, is affected. This renders the condition (17) unsatisfied and further leads the estimation ellipsoidal set not to confine the true state of the system. However, the i-th prediction ellipsoidal set $\mathcal{X}_{k+1|k}^i$, which is based on the measurement at time k, is not affected by the attack. Thus, it can be indicated that if there is no intersection between $\mathcal{X}_{k+1|k}^i$ and $\mathcal{X}_{k+1}^i$, the sensor measurement of vehicle i is compromised by an attack.

Remark 7: Note that attacks can be carefully designed to be undetectable and stealthy. For example, if the attacks have slow dynamics, such attacks are difficult to be distinguished from model uncertainty and noise. Therefore, most attack detection approaches presented in the literature involve a threshold and if the abrupt changes are below of this threshold, they cannot be detected, and thus remain undetected [51]. Hence, if malicious falsification delivered by an attacker is fairly small, the proposed detection strategy may not be able to recognize the attack in a timely fashion, i.e., there might be an intersection between the two ellipsoidal sets. Thus, the size of the ellipsoidal sets must be minimized, which necessitates the usage of an optimization approach aimed at minimizing the threshold. This concept will be outlined in Section III-B. In other words, the optimization approach in our proposed attack detection algorithm is aimed at softening the adversarial impact of attacks which may be restored by a well-designed resilient control strategy. The main trust of this paper is in its attack detection strategy while the design of a resilient control method constitutes one of our future research work.

III. Distributed Attack Detection and Recovery Mechanism Design

A. Design Criteria

We first present the following theorems, which establish sufficient conditions on the existence of the prediction and

estimation ellipsoidal sets that guarantee to contain the true state of each vehicle.

Theorem 1: For vehicle i described by (2) and (4) subject to UBB noises $w_k^i \in \mathcal{W}_k^i$ and $v_k^i \in \mathcal{V}_k^i$, suppose that at time k the vehicle's state x_k^i belongs to its state estimation ellipsoid $(x_k^i - \hat{x}_k^i)^T P_k^{i^{-1}}(x_k^i - \hat{x}_k^i) \leq 1$. Then the one-step ahead state x_{k+1}^i resides in its state prediction ellipsoid $(x_{k+1}^i - \hat{x}_{k+1|k}^i)^T P_{k+1|k}^{i^{-1}}(x_{k+1}^i - \hat{x}_{k+1|k}^i) \leq 1$, if there exist matrix sequences $P_{k+1|k}^i > 0$, G_k^i, and scalar sequences $\tau_{m,k}^i > 0$, $m = 1, 2$ such that

$$\begin{bmatrix} -P_{k+1|k}^i & \Theta_{i,k}^1 & AE_k^i & F \\ * & \Theta_{i,k}^2 & 0 & 0 \\ * & * & -\tau_{2,k}^i I & 0 \\ * & * & * & \Theta_{i,k}^3 \end{bmatrix} \leq 0 \tag{20}$$

where

$$\Theta_{i,k}^1 = (A - G_k^i)\hat{x}_k^i,\ \Theta_{i,k}^2 = -1 + \tau_{1,k}^i + \tau_{2,k}^i$$
$$\Theta_{i,k}^3 = -\tau_{1,k}^i Q_k^{i^{-1}}.$$

Proof: See the Appendix. ∎

Theorem 2: For vehicle i described by (2) and (4) subject to UBB noises $w_k^i \in \mathcal{W}_k^i$ and $v_k^i \in \mathcal{V}_k^i$, suppose that the one-step ahead state x_{k+1}^i belongs to its state prediction ellipsoid $(x_{k+1}^i - \hat{x}_{k+1|k}^i)^T P_{k+1|k}^{i^{-1}}(x_{k+1}^i - \hat{x}_{k+1|k}^i) \leq 1$. Then such a state resides in its state estimation ellipsoid $(x_{k+1}^i - \hat{x}_{k+1}^i)^T P_{k+1}^{i^{-1}}(x_{k+1}^i - \hat{x}_{k+1}^i) \leq 1$, which is updated with the measurement output y_{k+1}^i, if there exist matrix sequences $P_{k+1}^i > 0$, L_k^i, N_k^i, and scalar sequences $\tau_{m,k}^i > 0$, $m = 3, 4$ such that

$$\begin{bmatrix} -P_{k+1}^i & 0 & \bar{\Theta}_{i,k}^1 & -L_k^i D \\ * & \bar{\Theta}_{i,k}^2 & \bar{\Theta}_{i,k}^3 & \bar{\Theta}_{i,k}^4 \\ * & * & \bar{\Theta}_{i,k}^5 & \bar{\Theta}_{i,k}^6 \\ * & * & * & \bar{\Theta}_{i,k}^7 \end{bmatrix} \leq 0 \tag{21}$$

where

$$\bar{\Theta}_{i,k}^1 = (I - L_k^i C)E_{k+1|k}^i,$$
$$\bar{\Theta}_{i,k}^2 = -1 + \tau_{3,k}^i + \tau_{4,k}^i + M_k^i[1,1],$$
$$\bar{\Theta}_{i,k}^3 = M_k^i[\alpha_1, \beta_1]_{\alpha_1=1, \beta_2 \in \mathbb{N}_{[2,7]}},\ \bar{\Theta}_{i,k}^4 = M_k^i[1,8],$$
$$\bar{\Theta}_{i,k}^5 = -\tau_{4,k}^i I + M_k^i[\alpha_2, \beta_2]_{\alpha_2 \in \mathbb{N}_{[2,7]}, \beta_2 \in \mathbb{N}_{[2,7]}},$$
$$\bar{\Theta}_{i,k}^6 = M_k^i[\alpha_3, \beta_3]_{\alpha_3 \in \mathbb{N}_{[2,7]},\ \beta_3 = 8}, \bar{\Theta}_{i,k}^7 = -\tau_{3,k}^i R_k^{i^{-1}},$$
$$M_k^i = N_k^{i^T} \Pi_k^i + \Pi_k^{i^T} N_k^i,$$
$$\Pi_k^i = \begin{bmatrix} C\hat{x}_{k+1|k}^i - y_{k+1}^i & CE_{k+1|k}^i & D \end{bmatrix}.$$

Proof: See the Appendix. ∎

Remark 8: It is clearly shown in Theorems 1 and 2 that the proposed ellipsoidal set-membership filtering problem can be cast into the feasibility problem of a set of recursive linear matrix inequalities (RLMIs) (20) and (21). Thus, Theorems 1 and 2 provide criteria for designing two ellipsoidal sets which always have intersection since they both guarantee to contain the true state x_{k+1}^i in an attack-free system.

B. Recursive Convex Optimization and Cyber Attack Detection Algorithm with Recovery Mechanism

In light of Theorems 1 and 2, for the i-th attack-free vehicle, the one-step ahead state x_{k+1}^i always resides in both the state prediction ellipsoidal set $\mathcal{X}_{k+1|k}^i$ and its updated state estimation ellipsoidal set $\mathcal{X}_{k+1}^i$ if (20) and (21) hold. Therefore, there exists a vector z satisfying $\|z\| \leq 1$ such that $x_{k+1}^i = \hat{x}_{k+1|k}^i + E_{k+1|k}^i z$ and $x_{k+1}^i = \hat{x}_{k+1}^i + E_{k+1}^i z$, respectively. Furthermore, the center of the state prediction ellipsoid $\hat{x}_{k+1|k}^i$ and the center of the state estimation ellipsoid $\hat{x}_{k+1}^i$ are determined by (14) and (16), respectively. Thus, Theorems 1 and 2 outline the principles of determining the state prediction ellipsoidal set and the state estimation ellipsoidal set calculated through updating the prediction set via the current measurement data. However, Theorems 1 and 2 do not provide optimal state prediction and state estimation ellipsoids. Therefore, a convex optimization approach is applied in (22) and (23), in which the traces of $P_{k+1|k}^i$ and P_{k+1}^i are optimized at each time step in an effort to find the prediction ellipsoidal set and the estimation ellipsoidal set with minimal size.

$$\underset{P_{k+1|k}^i > 0, G_k^i, \tau_{1,k}^i > 0, \tau_{2,k}^i > 0}{\text{minimize}} \quad \text{Tr}(P_{k+1|k}^i) \tag{22}$$
subject to (20).

$$\underset{P_{k+1}^i > 0, L_k^i, N_k^i, \tau_{3,k}^i > 0, \tau_{4,k}^i > 0}{\text{minimize}} \quad \text{Tr}(P_{k+1}^i) \tag{23}$$
subject to (21).

Remark 9: In Theorems 1 and 2, one can observe that RLMIs (20) and (21) are linear to $P_{k+1|k}^i$, P_{k+1}^i, G_k^i, L_k^i, N_k^i, and $\tau_{m,k}^i, m = 1, 2, \cdots, 4$. As a result, the optimization problems (22) and (23) can be solved by some existing semidefinite programming via an interior-point algorithm at each time step. In practice, interior-point methods for semidefinite programs are competitive with other methods for small programs and substantially faster for medium and large-scale problems [52].

Based on the optimization problems (22) and (23), we next present an algorithm that recursively computes the state ellipsoidal sets for each vehicle in the platoon such that cyber attacks can be detected. Besides, two recovery mechanisms are provided to secure each vehicle in the platoon against cyber attacks violating its data measured by sensors and the inter-vehicle data transmitted through the communication network.

Recursive Convex Optimization Algorithm / Attack Detection & Recovery

---

Step 1. Initialization

Given initial conditions x_0^i, $\hat{x}_0^i$, and x_0^0. Choose suitable Q_k^i, R_k^i, and P_0^i such that (5), (6), and (19) hold. Calculate E_0^i according to $P_0^i = E_0^i E_0^{i^T}$. Set $k = 0$, $x_k^i = x_0^i$, $\hat{x}_k^i = \hat{x}_0^i$, $x_k^0 = x_0^0$, and $E_k^i = E_0^i$. Define the reference acceleration profile u_k^r. Let the simulation run recursively for N steps.

Step 2. Prediction

1) Solve the optimization problem (22) to calculate $P_{k+1|k}^i$ and G_k^i. Obtain $E_{k+1|k}^i$ such that $P_{k+1|k}^i = E_{k+1|k}^i E_{k+1|k}^{i^T}$.
2) Compute the center of the state prediction ellipsoid $\hat{x}_{k+1|k}^i$ by (14).

Step 3. Diagnosis of Cyber Attack on Inter-Vehicle Signal

1) If $\mathcal{X}_k^i \bigcap \mathcal{X}_{k+1|k}^i \neq \varnothing$, report "No Attack" and go to Step 5.
2) If $\mathcal{X}_k^i \bigcap \mathcal{X}_{k+1|k}^i = \varnothing$, report "Attack" and go to Step 4.

Step 4. Recovery Mechanism

1) Set $u_k^{i-1} \leftarrow u_{k-1}^{i-1}$.
2) Recompute $u_{ff,k}^i$ and u_k^i by (9) and (8), respectively.
3) Recalculate the center of the state prediction ellipsoid by (14) with the new u_k^{i-1} and u_k^i.
4) Set $E_{k+1|k}^i \leftarrow E_{k|k-1}^i$.

Step 5. Measurement Update

1) Solve the optimization problem (23) to calculate P_{k+1}^i and L_k^i. Obtain the new E_{k+1}^i such that $P_{k+1}^i = E_{k+1}^i E_{k+1}^{i^T}$.

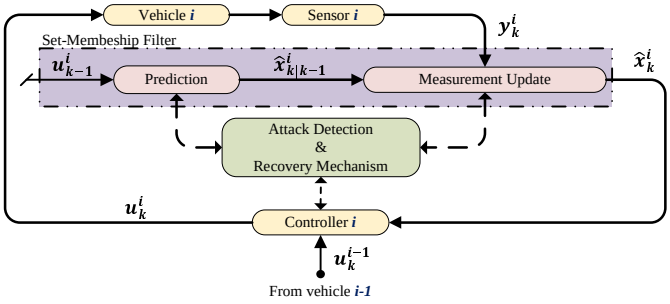


Fig. 2. Schematic of distributed attack detection with recovery mechanisms for vehicle i under set-membership filtering technique.

2) Compute the center of the updated state estimation ellipsoid $\hat{x}_{k+1}^i$ by (16).

Step 6. Diagnosis of Cyber Attack on Sensor Signal

1) If $\mathcal{X}_{k+1}^i \bigcap \mathcal{X}_{k+1|k}^i \neq \varnothing$, report "No Attack" and go to Step 8.
2) If $\mathcal{X}_{k+1}^i \bigcap \mathcal{X}_{k+1|k}^i = \varnothing$, report "Attack" and go to Step 7.

Step 7. Recovery Mechanism

Set $x_{k+1}^i \leftarrow x_{k+1|k}^i$ and $E_{k+1}^i \leftarrow E_{k+1|k}^i$

Step 8. Loop

If $k = N$ then Exit, Else $k \leftarrow k+1$ and go to Step 2.

Remark 10: The impacts of an attack on the centers of the prediction and estimation ellipsoidal sets result in infeasibility of the conditions (15) and (17), i.e., they may not contain the true state of the system. Therefore, it is of utmost importance to recover these sets once the attack is detected, which not only ensures the feasibility of those conditions but also preserves the system's string stability while mitigating the adversarial effects of the attack. To address this, two recovery mechanisms are proposed in the recursive algorithm. 1) In *step 4*, if $\mathcal{X}_k^i \bigcap \mathcal{X}_{k+1|k}^i = \varnothing$, the prediction ellipsoidal set must be reobtained through modifying the transmitted signal u_k^{i-1} with its one-step-behind value. This modification requires a command signal received by the i-th controller from which the controller recalculates the feedforward control signal $u_{ff,k}^i$ and sends a new control signal u_k^i to vehicle i. 2) In *step 7*, if $\mathcal{X}_{k+1}^i \bigcap \mathcal{X}_{k+1|k}^i = \varnothing$, the estimation ellipsoidal set must be replaced with its prediction ellipsoidal set, which is still based on the previous measurement output. Therefore, a command signal is sent to the measurement update step so that the replacement occurs in this step. From Fig. 2, one can conclude that the i-th controller always uses the free-of-attack state prediction and state estimation so as to generate the control signal for vehicle i through which the string stability can be satisfied in the entire platoon.

Remark 11: The information flow topology (IFT) defines how information is exchanged between connected vehicles. So far, in this paper, the proposed attack detection algorithm and recovery mechanisms have been constructed over a typical IFT, which is the predecessor following (PF) topology shown in Fig. 1. In addition to the PF topology, representative approaches for IFT in the literature [53] can be arguably categorized as the bidirectional (BD), predecessor-following leader (PFL), bidirectional leader (BDL), two predecessor-following (TPF), and two predecessor-following leader (TPFL) topologies. Considering different topologies is currently beyond the scope of this paper. However, the proposed attack detection algorithm and recovery mechanisms have a great potential to be extended for the specific communication topologies aforementioned and therefore, it constitutes one of our future research work.

Remark 12: Note that faults in the system's components, i.e., sensors and actuators, commonly exist in many practical engineering systems. For example, consider $f_{1,k}$ and $f_{2,k}$ as a UBB actuator's fault and a UBB sensor's fault added into (2) and (4), respectively. The proposed two-step filter in Section II-C can now be refined to provide a state prediction ellipsoidal set and a state estimation ellipsoidal set which always contain the true state of the system in the presence of the faults. If the system is affected by an attack with a different pattern of the fault signal, this attack may lead to empty set intersection and thus can be successfully identified through the proposed detection method and its impact on string stability can be eliminated by the proposed recovery mechanisms. However, if the attack hides in the fault (e.g., follows a similar UBB pattern of the fault signal), the intersection of the two sets may not always be empty, the proposed detection method fails to distinguish the fault-like attack from the fault. However, the refined filter is able to tolerate this type of attack in the same way as it does with respect to the fault. From this perspective, the proposed detection method is only sensitive to those attacks whose patterns are different from those of fault signals. As a matter of fact, one key merit of this method is that attack must occur when it is detected by the designed filter.

IV. Simulation Results and Case Studies

In this section, the developed attack detection algorithm and the two recovery mechanisms will be applied to a platooning system of five vehicles to validate their effectiveness under various cases of attacks.

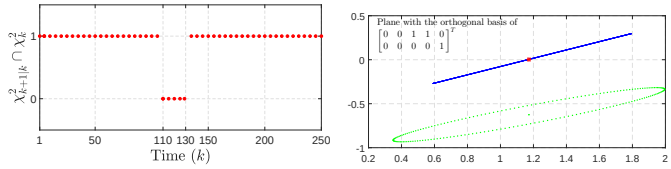
A. Simulation Setup

The platoon consists of one leader ($i = 1$) and four followers ($i = 2, 3, 4, 5$) whose desired acceleration informed by the reference vehicle in the platoon ($i = 0$) is in the form of

$$u_k^r = \begin{cases} 1 \text{ m/s}^2, & \text{if } 50 < k < 150 \\ 0, & \text{otherwise,} \end{cases}$$

and its initial velocity and acceleration are 15 $\mathrm{m/s}$ and 0 $\mathrm{m/s^2}$, respectively. The parameters for each vehicle in the platoon are assumed to be identical, i.e., $\eta = 0.1$, $h_d = 0.7$ s, which indicates a homogeneous vehicle platoon. As discussed in [54], the bandwidth of the ACC controller must be taken as $\omega_c = 0.05\eta^{-1}$ to satisfy the internal stability of the vehicle dynamics. Hence, the PD controller gains are chosen as $k_p = 0.25$ and $k_d = 0.5$. The continuous-time model is discretized with a sampling period $h = 0.1$ s.

The initial velocity of each vehicle is taken as 15 $\mathrm{m/s}$ with the initial inter-vehicle distance 10.5 m and zero initial acceleration, which means zero initial spacing error from the definition of the spacing error in Remark 2. Therefore, the initial condition in the platoon is $x_0^i = [10.5, 15, 0, 0, 0]^T, i \in \mathbb{N}_{[1,5]}$. For each vehicle, suppose the UBB noises are $w_k^i = 0.1\sin(2k)$ and $v_k^i = 0.2\cos(5k)$ and set $Q_k^i = 2$ and $R_k^i = 4$. Then it can be easily checked that w_k^i and v_k^i belong to the ellipsoidal sets defined in (5) and (6), respectively. Furthermore, let $F = [0.2, 0.2, 0.1, 0.2, 0.1]^T$



(a) (b)

Fig. 3. (a) Intersection between $\chi^2_{k+1|k}$ and χ^2_k ("1" indicates there exists intersection, "0" indicates there is no intersection); (b) $\chi^2_{111|110}$ (green, dotted line), χ^2_{110} (blue, solid line), true state (red, star).

and $D = [1, 1, 1]^T$. The initial conditions of estimators are chosen as $\hat{x}^1_0 = [10.42, 14.98, 0, 0.02, 0]^T$, $\hat{x}^2_0 = [10.431, 14.972, 0, 0.018, 0]^T$, $\hat{x}^3_0 = [10.458, 14.96, 0, 0.012, 0]^T$, $\hat{x}^4_0 = [10.474, 14.942, 0, 0.018, 0]^T$, $\hat{x}^5_0 = [10.498, 14.911, 0, 0.031, 0]^T$, which define the centers of initial ellipsoidal sets (19) for each vehicle with the shape matrix taken as $P^i_0 = 50 I_{5\times 5}$. The simulation results are obtained during 250 sampling instants.

B. DoS Attack on Communication Network

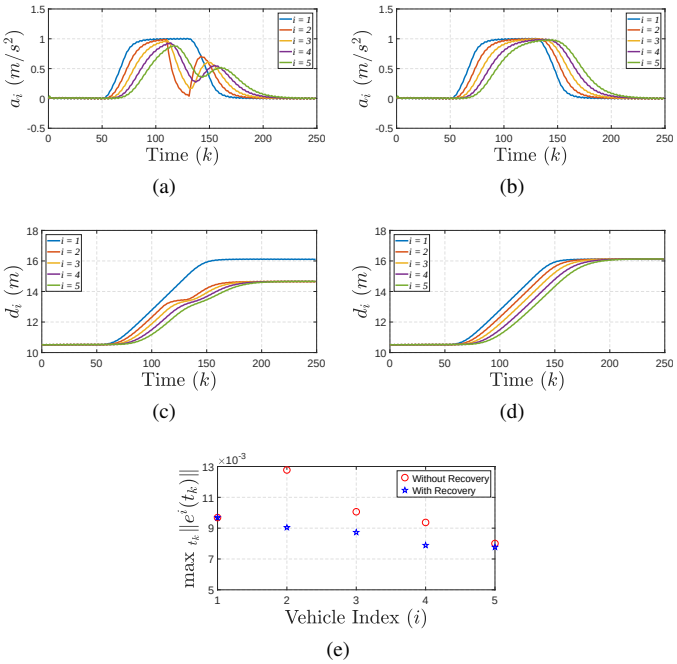
In this case, the adversary launches a DoS attack on the control signal u^1_k, which is sent by vehicle $i = 1$ and received by vehicle $i = 2$, through jamming the communication network channel between these two vehicles from $k = 110$ to $k = 130$. In the simulation, the lower- and upper-bounds on Λ^1_k in (13) are chosen as $0.8 \leq \Lambda^1_k \leq 1$. Hence, the actual control signal received by vehicle $i = 2$ is

$$\begin{cases} \tilde{u}^1_k = u^1_k - \Lambda^1_k u^1_k, & \text{if } 110 \leq k \leq 130 \\ \tilde{u}^1_k = u^1_k, & \text{otherwise,} \end{cases}$$

where, $\Lambda^1_k = 1$ indicates that the signal u^1_k is completely lost during its transmission; $0.8 \leq \Lambda^1_k < 1$ represents that vehicle $i = 2$ receives up to 20% of the signal u^1_k.

Fig. 3a shows the sequences of intersection between the prediction ellipsoidal set and the estimation ellipsoidal set updated at the previous time for vehicle $i = 2$ during the simulation time. Since the DoS attack starts at $k = 110$, the prediction ellipsoidal set for vehicle $i = 2$, i.e., $\mathcal{X}^2_{111|110}$, is affected by the attack as its center is obtained from the signal data received from vehicle $i = 1$ at this time. Therefore, from Remark 5, it can be concluded that the prediction set $\mathcal{X}^2_{111|110}$ does not enclose the true state as shown in Fig. 3b. However, the attack does not have any influence on the previous updated estimation ellipsoidal set, i.e., $\mathcal{X}^2_{110}$, since it is based on its prediction set at time $k = 109$. Thus, from *step 3* in the proposed algorithm, it is expected that there must be no intersection between these two sets in the existence of the attack at time $k = 110$, i.e., $\mathcal{X}^2_{110} \bigcap \mathcal{X}^2_{111|110} = \varnothing$.

Once the communication channel between vehicles $i = 1$ and $i = 2$ is jammed, vehicle $i = 2$ in the string does not completely receive its desired acceleration from the leader ($i = 1$) and therefore, all its following vehicles' acceleration in the string are affected by this attack as shown in Fig. 4a. In consequence, as depicted in Fig. 4c, the inter-vehicle distance between all followers cannot reach the desired distance defined under CTH policy, which opposes the objective of increasing the traffic throughput. Therefore, from Fig. 4e, the string stability (11) cannot be achieved as $\max_{t_k} \|e^3(t_k)\|$ and $\max_{t_k} \|e^2(t_k)\|$ are greater than $\max_{t_k} \|e^1(t_k)\|$, which



(a) (b) (c) (d) (e)

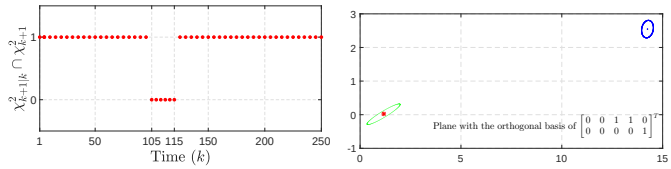
Fig. 4. (a) and (b) Vehicle's acceleration without and with recovery mechanisms; (c) and (d) Inter-vehicle distance without and with recovery mechanisms; (e) String stability analysis.

indicates that the spacing error does not decrease upstream the string.

The next step after the detection of the attack is to recover the predicted state $\hat{x}^2_{111|110}$ that is used to calculate the estimated state $\hat{x}^2_{111}$. From *step 4*, the control signal u^1_{110} must be replaced with its one-step-behind value from which the feedforward control signal $u^2_{ff,110}$ is recalculated. Then, the predicted state is reobtained from these modified signals and it will be sent to *step 5*, the measurement update step, in order to generate the secure estimated state. These modifications are required for each instant until the attack finishes at $k = 130$. With the recovery mechanism, the platoon obtains this capability to maintain its string stability in the existence of the DoS attack since the spacing error is attenuated toward the tail of the platoon as shown in Fig. 4e. Thus, all vehicles in the string follow the leader's desired acceleration and reach the safety inter-vehicle distance in accordance with CTH policy, as shown in Fig. 4b and Fig. 4d, respectively.

C. Replay Attack on Sensor Data

To manipulate the measurement output with a replay attack, it is considered that the attacker uses his ability to perform a disclosure attack so that he can store the data measured through on-board sensors. In the simulation, it is assumed that he obtains access to vehicle $i = 2$'s sensor through which its velocity is measured and records the signal data from $k = 65$ till $k = 75$. In the second phase of the attack, the attacker modifies the current data of the signal with his recorded data from $k = 105$ to $k = 115$. The lower- and upper-bounds on Λ^2_k are chosen as $\text{diag}\{0, 0.8, 0\} \leq \Lambda^2_k \leq \text{diag}\{0,1,0\}$, which indicates that there is only a replay attack on the second component of the measurement output y^2_k, i.e., the measured velocity of vehicle $i = 2$. Therefore, the actual measurement



(a) (b)

Fig. 5. (a) Intersection between $\chi^2_{k+1|k}$ and χ^2_{k+1}; (b) $\chi^2_{105|104}$ (green, dotted line), χ^2_{105} (blue, solid line, magnified with the ratio of 10^2), true state (red, star)

.

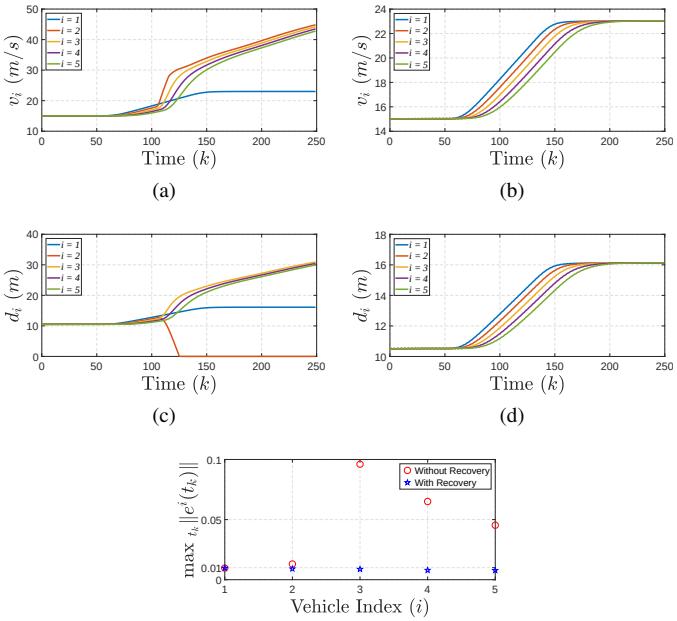
output is

$$\begin{cases} \tilde{y}^2_k = y^2_k + \Lambda^2_k(y^2_{k-40} - y^2_k), & \text{if } 105 \le k \le 115 \\ \tilde{y}^2_k = y^2_k, & \text{otherwise.} \end{cases}$$

Fig. 5a demonstrates the sequences of intersection between the prediction ellipsoidal set and the estimation ellipsoidal set updated at the current time for vehicle $i = 2$ during the simulation time. In particular, as the replay attack starts at $k = 105$, the prediction ellipsoidal set $\mathcal{X}^2_{105|104}$ is calculated from the measurement data obtained at $k = 104$ when there is no attack. However, the estimation ellipsoidal set $\mathcal{X}^2_{105}$ is updated with the current sensor measurement output at $k = 105$, which is compromised by the attack. Therefore, one can conclude from Remark 6 that the estimation set $\mathcal{X}^2_{105}$ does not include the true state as depicted in Fig. 5b. Hence, it is expected from *Step 6* in the proposed algorithm that there must be no intersection between the prediction ellipsoidal set and the estimation ellipsoidal set updated at the current time when the attack starts at $k = 105$, i.e., $\mathcal{X}^2_{105} \bigcap \mathcal{X}^2_{105|104} = \varnothing$.

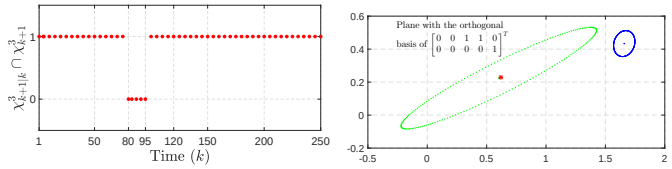
Replacing the second vehicle's velocity from $k = 105$ till $k = 115$ with the recorded value from $k = 65$ till $k = 75$ causes the controller to consider that the driver reduces the velocity. Thus, the controller suddenly increases the acceleration of the vehicle which results in an abrupt increase in the real-time velocities of the vehicle and its followers that may exceed the limitations of vehicle dynamics causing the vehicles to lose their internal stability, as demonstrated in Fig. 6a. It can be easily seen from Fig. 6c that the controller fails to mitigate the effect of the attack which causes vehicle $i = 2$ not to be able to properly adjust its inter-vehicle distance based on CTH policy and hence, it crashes into its predecessor. Furthermore, the string stability (11) is violated by the replay attack since $\max_{t_k} \|e^i(t_k)\|$, $i = 2, 3, \ldots 5$ is greater than $\max_{t_k} \|e^1(t_k)\|$ as shown in Fig. 6e, which contradicts the definition of string stability.

Once the attack is detected, the estimated state $\hat{x}^2_{105}$, which is utilized to generate the control command, must be recovered from the attack. To achieve this, the center and the shape matrix of the estimation ellipsoidal set are replaced with the same properties of its prediction ellipsoidal set as concluded from *Step 7*. In doing so, the controller uses the state prediction instead of its estimation since the predicted state is secure from the attack. This replacement must be made till the end of the attack at $k = 115$. With the recovery mechanism, the controller receives the secure predicted state through which it can compensate the malicious effect of the attack. As a consequence of the proposed recovery mechanism, the spacing error does not propagate upstream the string as shown in Fig. 6e, which satisfies the definition of string stability in (11)



(a) (b) (c) (d) (e)

Fig. 6. (a) and (b) Vehicle's velocity without and with recovery mechanisms; (c) and (d) Inter-vehicle distance without and with recovery mechanisms; (e) String stability analysis.



(a) (b)

Fig. 7. (a) Intersection between $\chi^3_{k+1|k}$ and χ^3_{k+1}; (b) $\chi^3_{80|79}$ (green, dotted line), χ^2_{80} (blue, solid line, magnified with the ratio of 10), true state (red, star)

.

and hence, all vehicles in the string follow the velocity pattern of the leader and reach the desired inter-vehicle distance, as depicted in Fig. 6b and Fig. 6d, respectively.

D. Message Falsification Attack on Sensor Data

To perform a message falsification attack, it is assumed that the attacker is able to collect the information from vehicle $i = 3$'s sensor through which its distance from vehicle $i = 2$ is measured. Simultaneously, the attacker manipulates the platoon toward his desired performance through adding the signal $A^3_k = 0.2 + 0.05\sin(k)$ into the content of the measurement output from $k = 80$ to $k = 95$. The lower- and upper-bounds on Λ^3_k in (13) are taken as $\text{diag}\{0.6, 0, 0\} \le \Lambda^3_k \le \text{diag}\{1, 0, 0\}$, which represents that there only exists a message falsification attack on the first component of the measurement output y^3_k, which is the measured distance between vehicles $i = 2$ and $i = 3$. Thus, the actual measurement output is

$$\begin{cases} \tilde{y}^3_k = y^3_k + \Lambda^3_k A^3_k, & \text{if } 80 \le k \le 95 \\ \tilde{y}^3_k = y^3_k, & \text{otherwise.} \end{cases}$$

Fig. 7a illustrates the sequences intersection between the prediction ellipsoidal set and the estimation ellipsoidal set updated at the current time for vehicle $i = 3$ over the simulation time. When the attack starts at $k = 80$, the prediction ellipsoidal set $\mathcal{X}^3_{80|79}$ is calculated from the measurement data

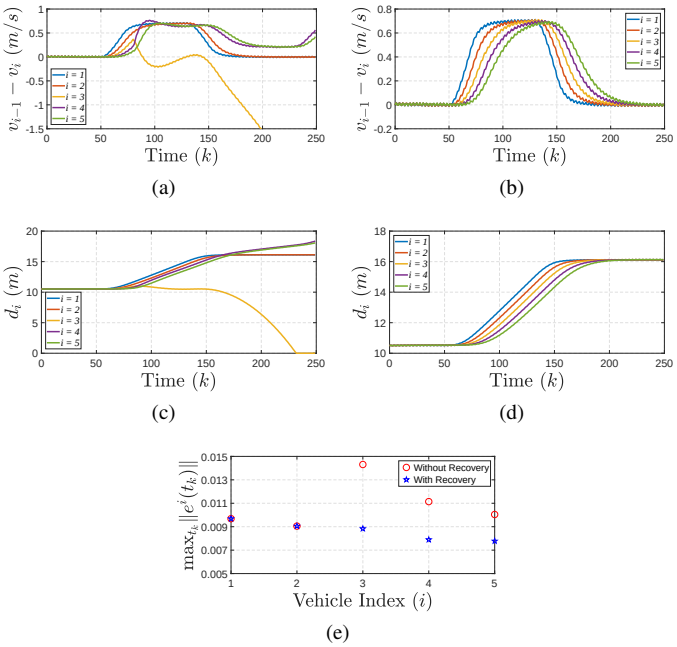


Fig. 8. (a) and (b) Relative velocity without and with recovery mechanisms; (c) and (d) Inter-vehicle distance without and with recovery mechanisms; (e) String stability analysis.

received at $k = 79$ when there is no attack. However, the estimation ellipsoidal set $\mathcal{X}^3_{80}$ is updated with the current measurement output at $k = 80$, which is affected by the attack. Therefore, as discussed in Remark 6, the estimation set $\mathcal{X}^3_{80}$ does not include the true state as shown in Fig. 7b. Thus, from *step 6* in the proposed algorithm, one can conclude that there must be no intersection between the prediction ellipsoidal set and the estimation ellipsoidal set updated at the current time when the attack starts at $k = 80$, i.e., $\mathcal{X}^3_{80} \bigcap \mathcal{X}^3_{80|79} = \varnothing$.

Tampering the content of the measured inter-vehicle distance between vehicles $i = 2$ and $i = 3$ with the bogus signal A^3_k deceives the controller into considering that there exists an increase in the measured distance. Therefore, the controller generates its command so as to reduce the relative velocity between vehicles $i = 2$ and $i = 3$ as shown in Fig. 8a, which results in a sudden decrease in the inter-vehicle distance between these vehicles. Hence, vehicle $i = 3$ fails to follow its required inter-vehicle distance and crashes into vehicle $i = 2$ as shown in Fig. 8c. These abrupt modifications degrade the string stability of the platoon since $\max_{t_k} \|e^i(t_k)\|$, $i = 3, 4, 5$ is greater than $\max_{t_k} \|e^2(t_k)\|$ as shown in Fig. 8e.

To tackle this, the estimated state $\hat{x}^3_{80}$ must be recovered from the attack so as to generate the control command by using the secure state estimation. Hence, from *step 7* in the proposed algorithm, the center and the shape matrix of the estimation ellipsoidal set $\mathcal{X}^3_{80}$ are replaced with the same properties of its prediction ellipsoidal set $\mathcal{X}^3_{80|79}$ so that the controller employs the predicted state which is secure from the attack. This recovery strategy must be implemented for each instant till the end of the attack at $k = 95$. As a result of the recovery mechanism, the string stability (11) is achieved since the spacing error decreases toward the end of the platoon as shown in Fig. 8e. Moreover, all vehicles in the platoon follow the relative velocity pattern introduced by the leader and reach the desired inter-vehicle distance, as demonstrated in Fig. 8b and Fig. 8d, respectively.

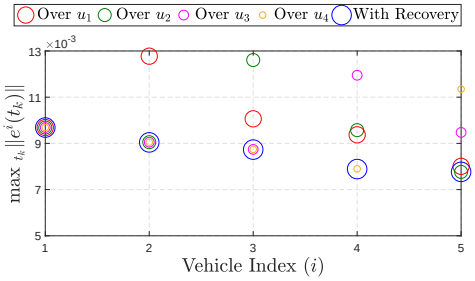


Fig. 9. Stability analysis under DoS attack on different communication channels.

E. String Stability Analysis under Different Attack Locations

In this section, we investigate the performance of the proposed algorithm on string stability while the attacker launches an attack on different vehicles in the string. For the brevity of presentation, only the DoS attack with the same properties as discussed in Section IV-B is considered over the communication network through which the desired control signal of each vehicle is received by its predecessor in the string. The discussion about the results for a replay attack and a message falsification attack on sensor data of different vehicles in the string is similar to that of the DoS attack and therefore, it is omitted here.

As depicted in Fig. 9, once the attacker moves toward the tail of the platoon, although the spacing error propagates in the string from vehicle i to vehicle $i+1$, the maximum magnitude of $\max_{t_k} \|e^i(t_k)\|$ decreases. This is due to the fact that the more spacing error is attenuated by CACC controllers as the attacker goes far away from the lead vehicle. Thus, the further attacker moves away from the lead vehicle, the less he is able to destabilize the platoon. However, as shown in Fig 9, no matter where the attack is launched, the platoon can achieve its string stability with the proposed recovery mechanisms once the attack is detected at its time of occurrence by the proposed distributed detection method.

V. Conclusion

A novel distributed attack detection method has been developed to address the problem of the detection of cyber attacks compromising the shared communication network and on-board sensors employed in a vehicle platooning system. To build a foundation of the attack detection method, a recursive state estimation algorithm characterized by ellipsoidal set-membership filters in a distributed framework is developed. The distributed set-membership filtering algorithm allows each vehicle in the platoon to estimate their states with no need of having full knowledge of the entire platoon, which reduces the computational overhead for this model-based detection method and makes the method suitable for a large-scale platoon. To mitigate the malicious effects of cyber attacks, two recovery mechanisms are introduced into the proposed algorithm. The idea of these two recovery mechanisms is based on reliable modifications of the attacked signals required for the computation of the prediction ellipsoidal set and estimation ellipsoidal set so that each ellipsoidal set will adopt the compensated state prediction and/or state estimation to increase security of the estimation system. The simulation results demonstrated that with the proposed recovery methods, the controllers in

CACC system are able to compensate the adversarial effect of an attack through accessing the secure state estimation so that the string stability of the platoon is satisfied while the attack is present. In the coexistence of the faults in the system, the proposed set-membership state estimation method can be refined as a fault tolerant filter by considering the UBB fault signals in the system dynamics. In doing so, the refined filter can tolerate the impact of the fault-like attack in the same way as it does with respect to the fault.

APPENDIX

Proof of Theorem 1:

At time k, if $(x_k^i - \hat{x}_k^i)^T P_k^{i^{-1}} (x_k^i - \hat{x}_k^i) \leq 1$, by Schur complement [55], the inequality is equivalent to $(x_k^i - \hat{x}_k^i)(x_k^i - \hat{x}_k^i)^T \leq P_k^i$. From Cholesky factorization, one has $P_k^i = E_k^i E_k^{i^T}$. Let $z = E_k^{i^{-1}}(x_k^i - \hat{x}_k^i)$, then we have

$$z^T z = (x_k^i - \hat{x}_k^i)^T P_k^{i^{-1}} (x_k^i - \hat{x}_k^i) \leq 1 \tag{24}$$

which means that there exists a vector z satisfying $\|z\| \leq 1$ such that $x_k^i = \hat{x}_k^i + E_k^i z$. The prediction tracking error can be calculated as

$$x_{k+1}^i - \hat{x}_{k+1|k}^i = \Psi_k^i \psi_k^i \tag{25}$$

where $\Psi_k^i = [\Theta_{i,k}^1, AE_k^i, F]$ and $\psi_k^i = [1, z^T, w_k^{i^T}]^T$. Thus, the condition in (15) can be rewritten as

$$\psi_k^{i^T} (-\text{diag}\{1,0,0\} + \Psi_k^{i^T} P_{k+1|k}^{i^{-1}} \Psi_k^i) \psi_k^i \leq 0. \tag{26}$$

From (5) and (24), the unknown variables w_k^i and z satisfy

$$\begin{cases} w_k^{i^T} Q_k^{i^{-1}} w_k^i \leq 1 \\ \|z\| \leq 1 \end{cases}$$

which can be expressed in ψ_k^i as

$$\begin{cases} \psi_k^{i^T} \Phi_k^1 \psi_k^i \geq 0 \\ \psi_k^{i^T} \Phi_k^2 \psi_k^i \geq 0 \end{cases} \tag{27}$$

where $\Phi_k^1 = \text{diag}\{1, 0, -Q_k^{i^{-1}}\}$ and $\Phi_k^2 = \text{diag}\{1, -I, 0\}$.

By virtue of $\mathcal{S}$-procedure [55], one obtains from (26) and (27) that the inequality (26) is satisfied if there exist scalar sequences $\tau_{m,k}^i > 0$, $m = 1, 2$ such that

$$-\text{diag}\{1,0,0\} + \Psi_k^{i^T} P_{k+1|k}^{i^{-1}} \Psi_k^i + \tau_{1,k}^i \Phi_k^1 + \tau_{2,k}^i \Phi_k^2 \leq 0. \tag{28}$$

Using Schur complement to (28) straightforwardly gets (20). This completes the proof. ■

Proof of Theorem 2:

From Theorem 1, if x_{k+1}^i belongs to the ellipsoid $(x_{k+1}^i - \hat{x}_{k+1|k}^i)^T P_{k+1|k}^{i^{-1}} (x_{k+1}^i - \hat{x}_{k+1|k}^i) \leq 1$, then there exists a vector z satisfying

$$z^T z = (x_{k+1}^i - \hat{x}_{k+1|k}^i)^T P_{k+1|k}^{i^{-1}} (x_{k+1}^i - \hat{x}_{k+1|k}^i) \leq 1 \tag{29}$$

such that $x_{k+1}^i = \hat{x}_{k+1|k}^i + E_{k+1|k}^i z$, where $E_{k+1|k}^i$ is a factorization of $P_{k+1|k}^i = E_{k+1|k}^i E_{k+1|k}^{i^T}$. The one-step ahead state estimation error can be obtained as

$$x_{k+1}^i - \hat{x}_{k+1}^i = \tilde{\Psi}_{k+1}^i \tilde{\psi}_{k+1}^i \tag{30}$$

where $\tilde{\Psi}_{k+1}^i = [0, \bar{\Theta}_{i,k}^1, -L_k^i D]$ and $\tilde{\psi}_{k+1}^i = [1, z^T, v_{k+1}^{i^T}]^T$. Hence, the condition in (17) can be rewritten as

$$\tilde{\psi}_{k+1}^{i^T} (-\text{diag}\{1,0,0\} + \tilde{\Psi}_{k+1}^{i^T} P_{k+1}^{i^{-1}} \tilde{\Psi}_{k+1}^i) \tilde{\psi}_{k+1}^i \leq 0. \tag{31}$$

From (6) and (29), it is clear that the unknown variables v_{k+1}^i and z satisfy

$$\begin{cases} v_{k+1}^{i^T} R_{k+1}^{i^{-1}} v_{k+1}^i \leq 1 \\ \|z\| \leq 1 \end{cases}$$

which can be rewritten as

$$\begin{cases} \tilde{\psi}_{k+1}^{i^T} \tilde{\Phi}_{k+1}^1 \tilde{\psi}_{k+1}^i \geq 0 \\ \tilde{\psi}_{k+1}^{i^T} \tilde{\Phi}_{k+1}^2 \tilde{\psi}_{k+1}^i \geq 0 \end{cases} \tag{32}$$

where $\tilde{\Phi}_{k+1}^1 = \text{diag}\{1, 0, -R_{k+1}^{i^{-1}}\}$ and $\tilde{\Phi}_{k+1}^2 = \Phi_k^2$.

Applying $\mathcal{S}$-procedure, (31) is satisfied if there exist scalar sequences $\tau_{m,k}^i > 0$, $m = 3, 4$ such that

$$-\text{diag}\{1,0,0\} + \tilde{\Psi}_{k+1}^{i^T} P_{k+1}^{i^{-1}} \tilde{\Psi}_{k+1}^i + \tau_{3,k}^i \tilde{\Phi}_{k+1}^1 + \tau_{4,k}^i \tilde{\Phi}_{k+1}^2 \leq 0. \tag{33}$$

The output constraint (18) can be expressed in terms of $\tilde{\psi}_{k+1}^i$ as

$$\Pi_k^i \tilde{\psi}_{k+1}^i = 0 \tag{34}$$

with Π_k^i defined in (21). Applying Finsler's lemma [55], (33) under constraint (34) (i.e. (18)) holds if there exists an N_k^i such that

$$\tilde{\Psi}_{k+1}^{i^T} P_{k+1}^{i^{-1}} \tilde{\Psi}_{k+1}^i + \tau_{3,k}^i \tilde{\Phi}_{k+1}^1 + \tau_{4,k}^i \tilde{\Phi}_{k+1}^2 + N_k^{i^T} \Pi_k^i + \Pi_k^{i^T} N_k^i \leq 0. \tag{35}$$

Using Schur complement to (35) yields (21). This completes the proof. ■

REFERENCES

[1] D. Jia, K. Lu, J. Wang, X. Zhang, and X. Shen, "A survey on platoon-based vehicular cyber-physical systems," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 263–284, Firstquarter 2016.

[2] K. C. Dey, L. Yan, X. Wang, Y. Wang, H. Shen, M. Chowdhury, L. Yu, C. Qiu, and V. Soundararaj, "A review of communication, driver characteristics, and controls aspects of cooperative adaptive cruise control (cacc)," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 491–509, Feb. 2016.

[3] J. K. Hedrick, Y. Chen, and S. Mahal, "Optimized vehicle control/communication interaction in an automated highway system," California Partners for Advances Transit and Highways, BERKELEY, Tech. Rep. UCB-ITS-PRR-2001-29, Oct. 2001.

[4] X. Liu, A. Goldsmith, S. S. Mahal, and J. K. Hedrick, "Effects of communication delay on string stability in vehicle platoons," in *Proc. IEEE Intell. Transp. Syst.*, CA, USA, Aug. 2001, pp. 625–630.

[5] S. Öncü, J. Ploeg, N. van de Wouw, and H. Nijmeijer, "Cooperative adaptive cruise control: Network-aware analysis of string stability," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1527–1537, Aug. 2014.

[6] P. Seiler and R. Sengupta, "Analysis of communication losses in vehicle control problems," in *Proc. Amer. Control Conf.*, VA, USA, Jun. 2001, pp. 1491–1496.

[7] R. Teo, D. M. Stipanovic, and C. J. Tomlin, "Decentralized spacing control of a string of multiple vehicles over lossy datalinks," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 2, pp. 469–473, Mar. 2010.

[8] C. Lei, E. M. van Eenennaam, W. K. Wolterink, G. Karagiannis, G. Heijenk, and J. Ploeg, "Impact of packet loss on cacc string stability performance," in *Proc. 11th Inter. Conf. ITS Telecommun.*, Aug. 2011, pp. 381–386.

[9] S. Wen, G. Guo, B. Chen, and X. Gao, "Event-triggered cooperative control of vehicle platoons in vehicular ad hoc networks," *Inform. Sci.*, vol. 459, pp. 341–353, Aug. 2018.

[10] G. Guo and S. Wen, "Communication scheduling and control of a platoon of vehicles in VANETs," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1551–1563, Jun. 2016.

[11] S. Santini, A. Salvi, A. S. Valente, A. Pescapè, M. Segata, and R. L. Cigno, "A consensus-based approach for platooning with inter-vehicular communications," in *Proc. IEEE Conf. Computer Commun.*, Kowloon, Hong Kong, Apr. 2015, pp. 1158–1166.

[12] G. Guo and W. Yue, "Sampled-data cooperative adaptive cruise control of vehicles with sensor failures," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 6, pp. 2404–2418, May 2014.

[13] Y. Liu, C. Pan, H. Gao, and G. Guo, "Cooperative spacing control for interconnected vehicle systems with input delays," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10692–10704, Dec. 2017.

[14] Y.-L. Wang, Q.-L. Han, M.-R. Fei, and C. Peng, "Network-based T-S fuzzy dynamic positioning controller design for unmanned marine vehicles," *IEEE Trans. Cybern.*, vol. 48, no. 9, pp. 2168–2267, Sep. 2018.

[15] Z. Peng, J. Wang, and D. Wang, "Distributed maneuvering of autonomous surface vehicles based on neurodynamic optimization and fuzzy approximation," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 3, pp. 1083–1090, May 2018.

[16] R. G. Engoulou, M. Bellaïche, S. Pierre, and A. Quintero, "VANET security surveys," *Comput. Commun.*, vol. 44, pp. 1–13, May 2014.

[17] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, "VANET security challenges and solutions: a survey," *Veh. Commun.*, vol. 7, pp. 7–20, Jan. 2017.

[18] R. W. van der Heijden, T. Lukaseder, and F. Kargl, "Analyzing attacks on cooperative adaptive cruise control (CACC)," in *Proc. IEEE Veh. Netw. Conf.*, Italy, Nov. 2017, pp. 45–52.

[19] S. Parkinson, P. Ward, K. Wilson, and J. Miller, "Cyber threats facing autonomous and connected vehicles: future challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 2898–2915, Nov. 2017.

[20] L. Liu, O. De Vel, Q.-L. Han, J. Zhang, and X. Yang, "Detecting and preventing cyber insider threats: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1397–1417, Secondquarter 2018.

[21] S. Hu, D. Yue, Q.-L. Han, X. Xie, X. Chen, and C. Dou, "Observer-based event-triggered control for networked linear systems subject to denial-of-service attacks," *IEEE Trans. Cybern.*, to be published, DOI: 10.1109/TCYB.2019.2903817

[22] D. Ding, Z. Wang, Q.-L. Han, and G. Wei, "Security control for discrete-time stochastic nonlinear systems subject to deception attacks," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 48, no. 5, pp. 779–789, May 2018.

[23] A. Alipour-Fanid, M. Dabaghchian, H. Zhang, and K. Zeng, "String stability analysis of cooperative adaptive cruise control under jamming attacks," in *Proc. IEEE 18th Inter. Symp. High Assur. Syst. Eng.*, Jan. 2017, pp. 157–162.

[24] Z. A. Biron, S. Dey, and P. Pisu, "Real-time detection and estimation of denial of service attack in connected vehicle systems," *IEEE Trans. Intell. Transp. Syst.*, 2018, DOI: 10.1109/TITS.2018.2791484.

[25] M. Amoozadeh, A. Raghuramu, C. n. Chuah, D. Ghosal, H. M. Zhang, J. Rowe, and K. Levitt, "Security vulnerabilities of connected vehicle streams and their impact on cooperative driving," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 126–132, Jun. 2015.

[26] B. DeBruhl, S. Weerakkody, B. Sinopoli, and P. Tague, "Is your commute driving you crazy?: A study of misbehavior in vehicular platoons," in *Proc. 8th ACM Conf. Se. & Pri. Wireless Mob. Netw.*, NY, USA, Jun. 2015, pp. 22:1–22:11.

[27] S. Dadras, R. M. Gerdes, and R. Sharma, "Vehicular platooning in an adversarial environment," *in Proc. 10th ACM Symp. Inf., Comput. Commun. Secur.*, Singapore, Apr. 2015, pp. 167-178.

[28] J. Petit and S. E. Shladover, "Potential cyberattacks on automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 546–556, Apr. 2015.

[29] R. W. van der Heijden, S. Dietzel, T. Leinmüller, and F. Kargl, "Survey on misbehavior detection in cooperative intelligent transportation systems," *IEEE Commun. Surveys Tuts.*, 2018, DOI: 10.1109/COMST.2018.2873088.

[30] F. Sakiz and S. Sen, "A survey of attacks and detection mechanisms on intelligent transportation systems: VANETs and IoV," *Ad Hoc Netw.*, vol. 61, pp. 35–50, 2017.

[31] A. Rawat, P. Anand, H. Chen, and P. Varshney, "Collaborative spectrum sensing in the presence of byzantine attacks in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 774–786, Feb. 2011.

[32] B. Kailkhura, Y. S. Han, S. Brahma, and P. K. Varshney, "Distributed bayesian detection in the presence of byzantine data," *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5250–5263, Oct. 2015.

[33] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[34] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 612–621, Mar. 2014.

[35] G. Chaojun, P. Jirutitijaroen, and M. Motani, "Detecting false data injection attacks in AC state estimation," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2476–2483, Sep. 2015.

[36] B. M. Sanandaji, E. Bitar, K. Poolla, and T. L. Vincent, "An abrupt change detection heuristic with applications to cyber data attacks on power systems," in *Proc. Amer. Control Conf.*, Portland, OR, 2014, pp. 5056–5061.

[37] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on SCADA systems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1396–1407, Jul. 2014.

[38] K. Manandhar, X. Cao, F. Hu, and Y. Liu, "Detection of faults and attacks including false data injection attack in smart grid using kalman filter," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 4, pp. 370–379, Dec. 2014.

[39] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 93–109, Feb. 2015.

[40] D. B. Rawat and C. Bajracharya, "Detection of false data injection attacks in smart grid communication systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1652–1656, Oct. 2015.

[41] A. B. Kurzhanski and I. Valyi, *Ellipsoidal calculus for estimation and control*, 1st ed. Boston, MA: Birkhäuser Basel, 1997.

[42] F. C. Schweppe, "Recursive state estimation: Unknown but bounded errors and system inputs," *IEEE Trans. Autom. Control*, vol. 13, no. 1, pp. 22–28, Feb. 1968.

[43] D. Bertsekas and I. Rhodes, "Recursive state estimation for a set-membership description of uncertainty," *IEEE Trans. Autom. Control*, vol. 16, no. 2, pp. 117–128, Apr. 1971.

[44] F. Yang and Y. Li, "Set-membership filtering for systems with sensor saturation," *Automatica*, vol. 45, no. 8, pp. 1896–1902, Aug. 2009.

[45] Z. Wu, F. Yang, and Q.-L. Han, "A novel islanding fault detection for distributed generation systems," *Int. J. Robust Nonlin. Control*, vol. 24, no. 8-9, pp. 1431–1445, 2014.

[46] X. Ge, Q.-L. Han, and F. Yang, "Event-based set-membership leader-following consensus of networked multi-agent systems subject to limited communication resources and unknown-but-bounded noise," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5045–5054, Jun. 2017.

[47] F. Yang and Y. Li, "Set-membership filtering with state constraints," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 4, pp. 1619–1629, Oct. 2009.

[48] X. Ge, Q.-L. Han, and Z. Wang, "A dynamic event-triggered transmission scheme for distributed set-membership estimation over wireless sensor networks," *IEEE Trans. Cybern.*, vol. 49, no. 1, pp. 171–183, Jan. 2019.

[49] E. Mousavinejad, F. Yang, Q.-L. Han, and L. Vlacic, "A novel cyber attack detection method in networked control systems," *IEEE Trans. Cybern.*, vol. 48, no. 11, pp. 3254–3264, Nov. 2018.

[50] X. Ge, F. Yang, and Q.-L. Han, "Distributed networked control systems: A brief overview," *Inf. Sci.*, vol. 380, pp. 117–131, Feb. 2017.

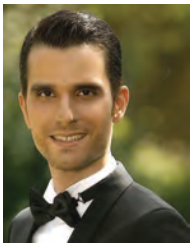
[51] A. Teixeira, I. Shames, H. Sandberg and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[52] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[53] S. E. Li, Y. Zheng, K. Li, Y. Wu, J. K. Hedrick, F. Gao, and H. Zhang, "Dynamical modeling and distributed control of connected and automated vehicles: challenges and opportunities," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 3, pp. 46–58, 2017.

[54] J. Ploeg, B. T. M. Scheepers, E. van Nunen, N. van de Wouw, and H. Nijmeijer, "Design and experimental evaluation of cooperative adaptive cruise control," in *Proc. 14th Inter. IEEE Conf. Intell. Transp. Syst.*, WA, USA, Oct. 2011, pp. 260–265.

[55] R. E. Skelton, T. Iwasaki, and K. Grigoriadis, *A Unified Algebraic Approach to Linear Control Design*. Bristol, PA: Taylor & Francis, 1998.



Eman Mousavinejad (S'14) received the B.Sc. degree in mechanical engineering from Azad University, Tehran, Iran, in 2006, the M.Sc. degree in mechatronics engineering (control) from Sharif University of Technology, Tehran, Iran, in 2009, and the M.Phil. degree in mechatronics engineering (control) from Griffith University, Gold Coast, QLD, Australia, in 2014, where he is currently pursuing the Doctor of Philosophy degree in mechatronics engineering (control) with the Griffith School of Engineering.

His current research interests include networked control and filtering, multiagent systems, cyber security, vehicle dynamics and control, autonomous vehicles, drive assistance systems, by-wire systems, and linear and nonlinear systems.

Fuwen Yang (M'99-SM'04) received the Ph.D. degree in control engineering from Huazhong University of Science and Technology, China, in 1990.

He is currently an Associate Professor with the Griffith School of Engineering and Built Environment, Griffith University, Gold Coast, QLD, Australia. He was a Research Fellow with Brunel University London, Uxbridge, U.K., and King's College London, London, U.K., a Professor with Fuzhou University, Fuzhou, China, and the East China University of Science and Technology, Shanghai, China, and an Associate Professor with Central Queensland University, Rockhampton, QLD, Australia. He had also held visiting professor appointments with the University of Manchester, Manchester, U.K., and the University of Hong Kong, Hong Kong. He has published over 200 journal and conference papers. His current research interests include networked control systems, distributed filtering and sensing, reliable fault detection and diagnosis, distributed control and filtering for smart girds, and solar PV power generation systems.

Dr. Yang is an Associate Editor of the IEEE CSS Conference Editorial Board.

Qing-Long Han (M'09-SM'13-F'19) received the B.Sc. degree in Mathematics from Shandong Normal University, Jinan, China, in 1983, and the M.Sc. and Ph.D. degrees in Control Engineering and Electrical Engineering from East China University of Science and Technology, Shanghai, China, in 1992 and 1997, respectively.

From September 1997 to December 1998, he was a Post-doctoral Researcher Fellow with the Laboratoire d'Automatique et d'Informatique Industielle (currently, Laboratoire d'Informatique et d'Automatique pour les Systémes), École Supérieure d'Ing'enieurs de Poitiers (currently, École Nationale Supérieure d'Ingé nieurs de Poitiers), Université de Poitiers, France. From January 1999 to August 2001, he was a Research Assistant Professor with the Department of Mechanical and Industrial Engineering at Southern Illinois University at Edwardsville, USA. From September 2001 to December 2014, he was Laureate Professor, an Associate Dean (Research and Innovation) with the Higher Education Division, and the Founding Director of the Centre for Intelligent and Networked Systems at Central Queensland University, Australia. From December 2014 to May 2016, he was Deputy Dean (Research), with the Griffith Sciences, and a Professor with the Griffith School of Engineering, Griffith University, Australia. In May 2016, he joined Swinburne University of Technology, Australia, where he is currently Pro Vice-Chancellor (Research Quality) and a Distinguished Professor. In March 2010, he was appointed Chang Jiang (Yangtze River) Scholar Chair Professor by Ministry of Education, China. His research interests include networked control systems, neural networks, time-delay systems, multi-agent systems and complex dynamical systems.

Professor Han is one of The World's Most Influential Scientific Minds: 2014-2016, and 2018. He is a Highly Cited Researcher according to Clarivate Analytics (formerly Thomson Reuters). He is a Fellow of The Institution of Engineers Australia. He is an Associate Editor of a number of international journals, including IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE INDUSTRIAL ELECTRONICS MAGAZINE, IEEE TRANSACTIONS ON CYBERNETICS, CONTROL ENGINEERING PRACTICE, INFORMATION SCIENCES, and IEEE/CAA JOURNAL OF AUTOMATICA SINICA.

Xiaohua Ge (M'18) received the B.Eng. degree in electronics and information engineering from Nanchang Hangkong University, Nanchang, China, in 2008, the M.Eng. degree in control theory and control engineering from Hangzhou Dianzi University, Hangzhou, China, in 2011, and the Ph.D. degree in computer engineering from Central Queensland University, Rockhampton, QLD, Australia, in 2014.

From 2011 to 2013, he was a Research Assistant with the Centre for Intelligent and Networked Systems, Central Queensland University, Rockhampton, QLD, Australia, where he was a Research Fellow in 2014. From 2015 to 2017, he was a Research Fellow with the Griffith School of Engineering, Griffith University, Gold Coast, QLD, Australia. He is currently a Lecturer with the School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, VIC, Australia.

His research interests include distributed estimation over sensor networks, distributed coordination in multi-agent systems, security and privacy preserving in cyber-physical systems.

Ljubo Vlacic (M'92-SM'96) received the Graduate Diploma degree in electrical engineering and the M.Phil. and Ph.D. degrees in control systems engineering from the University of Sarajevo, Sarajevo, Bosnia and Herzegovina, in 1973, 1976, and 1986, respectively.

He is with the Institute for Integrated and Intelligent Systems, Griffith University, Gold Coast, QLD, Australia. He is a Control Systems Scientist and a Practitioner, renowned for his contributions to cooperative driverless vehicles and intelligent control systems research and development. He has held a number of leading roles in both industry and academia.

Professor Emeritus Vlacic was a recipient of 19 awards, including the IEE Achievement Medal (world-wide), the Sir Lionel Hooke Award, the Queensland Professional Engineer of the Year Award, and the Gold Coast Business Events Ambassador Award. His research achievements made news headlines and were broadcast through media outlets throughout the world. He is currently the Editor-in-Chief of the IEEE Intelligent Transportation Systems Magazine, the Chair of the Board, Engineers Australia ITEE College and the IET Network Queensland, and the General Chair of the 2019 IEEE Intelligent Transportation Systems Conference. He hosted ten national and international scientific conferences.