Attack detection and estimation in cooperative vehicles
A sliding mode observer approach

Jahanshahi, Niloofar; Ferrari, Riccardo

DOI
10.1016/j.ifacol.2018.12.037

Publication date
2018

Document Version
Final published version

Published in
IFAC-PapersOnLine

Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 51-23 (2018) 212–217

Attack Detection and Estimation in Cooperative Vehicles Platoons: A Sliding Mode Observer Approach ⋆

Niloofar Jahanshahi, Riccardo M.G. Ferrari

Delft Center for Systems and Controls,
Delft University of Technology, Delft, The Netherlands
{n.jahanshahi,r.ferrari}@tudelft.nl

Abstract: Platoons of autonomous vehicles are currently being investigated by academic and industrial researchers as a way to increase road capacity and fuel efficiency. In order to fully reach such goals, a platoon must be endowed with cooperative capabilities, such as Cooperative Adaptive Cruise Control (CACC). This technique is based on the vehicles' sensors and on wireless communication between them, in order to control their longitudinal dynamics. However, the use of wireless communication exposes individual vehicles to cyber-attacks that aim at disrupting the platoon. Detecting and estimating a class of such attacks is the challenge considered in this paper, where an adaptive sliding mode observer is designed for this purpose. Theoretical results on the observer stability and robustness and simulation results are provided.

Keywords: Cooperative adaptive cruise control, Adaptive sliding mode control, Cyber attack, Attack reconstruction

1. INTRODUCTION

Rising traffic congestion is an alarming condition in populous urban areas across the world, having a negative effect on air pollution and energy consumption. Driver assistant systems, such as *Adaptive Cruise Control* (ACC), are currently being featured in individual vehicles in order to increase safety and provide a smoother ride. An extension of ACC, called *Cooperative Adaptive Cruise Control* (CACC), is currently being investigated and will enable several vehicles to coordinate themselves, thus increasing fuel efficiency and road capacity (see Naus et al. (2010); Ploeg et al. (2011)). In ACC, a feedback loop uses measurements from a vehicle local sensors to track a desired velocity and keep a safe distance from preceding vehicles. CACC, in addition, allows neighbouring vehicles to communicate and form a *platoon*, which is a string of vehicles travelling together while following a common velocity profile. Communication is the key to reaching a desirable property called *string stability*, which results in dampening of velocity oscillations along the platoon and which allows tighter inter-vehicle spacings than ACC.

As cooperative vehicles platoons employing CACC do posses sensing, actuation, computation and communication capabilities, they can be classified as a kind of cyber-physical systems (CPS), and as such are exposed to the same types of cyber-attacks that can threaten other known CPS, such as networked control systems and critical infrastructures (see Cárdenas et al. (2008, 2009); Teixeira et al. (2015a)). Cyber attacks that were considered in the literature include *Denial of Service* (DoS), routing, replay and stealthy data injection attacks, for instance.

The vulnerability of automotive communication and automation networks to cyber attacks has been extensively investigated and practical attacks have been demonstrated, as reported by Studnia et al. (2013), Miller and Valasek (2014), Amoozadeh et al. (2015) and Ploeg (2017). While CACC can provide limited robustness to network induced effects such as random packet losses (see Lei et al. (2011); Ploeg et al. (2013)), a malicious attacker targeting the vehicle-to-vehicle (V2V) network used by CACC-enabled vehicles can disrupt a platoon and possibly endanger its members. For this reason, passive robustness should be complemented by dedicated detection methods. The problem of designing such methods for generic CPS has been the subject of active research in the last years. Centralised and decentralised monitors for detecting and identifying attacks in linear time invariant descriptor systems were described in Pasqualetti et al. (2013), while Teixeira et al. (2012) and Ferrari and Teixeira (2017), amongst others, proposed techniques to detect attacks on sensor outputs. For the specific case of autonomous vehicles formations Meskin and Khorasani (2009) and Quan et al. (2018) proposed an observer-based approach for fault detection, while Biron et al. (2017) et al considered the problem of designing a model based observer for detecting DoS attacks, which they characterised as an equivalent time delay in the communication network. Finally, Bißmeyer et al. (2012) is an example of a contribution from the Computer Science community, where the trustworthiness of vehicles participating to a CACC-enabled platoon is evaluated using a particle filter.

In this paper we design an adaptive sliding mode observer for a CACC-equipped string of vehicles under a class of attacks affecting their V2V communication network. As shown in Figure 1, each vehicle can sense the position

⋆ This work has received funding from the European Union H2020 Programme under grant no. 707546 (SURE).

10.1016/j.ifacol.2018.12.037

and velocity of the preceding one using a frontal radar, and can receive via a wireless V2V network its intended acceleration. The local CACC controller of each vehicle is designed to use the measured and received data to control the vehicle acceleration and keep some desirable inter-vehicle distance. We propose to implement on each vehicle also an adaptive sliding mode observer designed in a way to estimate the longitudinal position, velocity and acceleration of the *preceding* vehicle. By using such observer both a detection *residual* and a dynamic robust *threshold* can be computed, allowing to detect a class of attacks that will alter the data communicated by the preceding vehicle. The novelty of our approach is that, differently from Quan et al. (2018), we propose to use the sliding mode observer control input as a detection residual. The advantage of this choice is that during an attack the same residual will estimate the attack magnitude as well, thus avoiding the need to design a separate identification algorithm.

The remainder of the paper is organized as follows. Section 2 formulates the CACC problem for a vehicle platoon where the V2V wireless communication channel is under attack, and provides a mathematical description of the effect of the attack on the platoon dynamics. Section 3 presents the adaptive sliding mode observer design procedure and provides theoretical results on the observer stability and its robustness to sensors measurement uncertainty. In section 4, the proposed method is applied to a simulated CACC–equipped string of three vehicles in order to exemplify its detection capabilities. Conclusion and future work are presented in the final section.

2. PROBLEM FORMULATION

In this section we will initially provide the main equations describing the dynamics of a CACC–equipped vehicles platoon, following the formulation introduced by Ploeg et al. (2011). Then we will introduce a generic attack term affecting data communicated over the V2V network and show how it affects the vehicles and CACC dynamics.

2.1 Platoon longitudinal dynamics and CACC equations

Consider a string of $m \in \mathbb{N}$ homogenous vehicles as shown in Figure 1. The dynamics of the i-th vehicle, $2 \leq i \leq m$, can be modelled as

$$\begin{bmatrix} \dot{p}_i(t) \\ \dot{v}_i(t) \\ \dot{a}_i(t) \end{bmatrix} = \begin{bmatrix} v_i(t) \\ a_i(t) \\ -\frac{1}{\tau} a_i(t) + \frac{1}{\tau} u_i(t) \end{bmatrix}, \tag{1}$$

where $p_i(t)$, $v_i(t)$, $a_i(t)$ and $u_i(t) \in \mathbb{R}$ are the position, velocity, acceleration and the input of the i-th vehicle, respectively. The time constant τ accounts for the engine's dynamics: for a discussion on the validity of the simple linear model in (1) the reader is referred to Ploeg et al. (2011). Each vehicle's objective is to match the velocity of the preceding one and keep a desired inter-vehicle distance $d_{r,i}$, which depends on a constant time headway policy:

$$d_{r,i}(t) = r_i + h v_i(t)\,.$$

The parameters r_i and h represent the desired distance at stand still and, respectively, the time headway between the i-th and the $i-1$-th vehicle. Gehring and Fritz (1997) have shown that this choice of spacing policy improves *string stability*, where a platoon is defined to be string stable if relative position, velocity or acceleration errors between pairs of adjacent vehicles are not amplified downstream the string (Ploeg et al., 2011, Definition 1).

The distance error between adjacent vehicles is defined as the error between their distance $d_i(t) \triangleq (p_{i-1}(t) - p_i(t) - L)$ and the desired distance as

$$\begin{aligned} e_i(t) =& d_i(t) - d_{r,i}(t) \\ =& (p_{i-1}(t) - p_i(t) - L) - (r + h v_i(t))\,, \end{aligned} \tag{2}$$

where L is the vehicle's length.

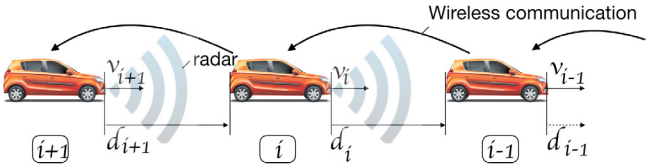

Fig. 1. CACC equipped string of vehicles. The V2V communication network is implemented via a wireless channel, and is assumed to be subjected to a class of cyber attacks.

The input $u_i(t)$ represents the desired acceleration that is commanded to the vehicle drivetrain by the local CACC controller, and is computed as the solution to the following dynamical equation

$$\begin{aligned} \dot{u}_i(t) =& \frac{1}{h} u_i(t) + \frac{1}{h} \big(k_p e_{1,i}(t) \\ & + k_d e_{2,i}(t) \big) + \frac{1}{h} u_{i-1}(t), \end{aligned} \tag{3}$$

where

$$\begin{bmatrix} e_{1,i}(t) \\ e_{2,i}(t) \\ e_{3,i}(t) \end{bmatrix} = \begin{bmatrix} e_i(t) \\ \dot{e}_i(t) \\ \ddot{e}_i(t) \end{bmatrix}$$

are the distance, velocity and acceleration errors between adjacent vehicles and k_p and k_d are design parameters. By substituting (3) and (4) in (1), the error dynamics in normal conditions can be written as

$$\begin{bmatrix} \dot{e}_{1,i}(t) \\ \dot{e}_{2,i}(t) \\ \dot{e}_{3,i}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{k_p}{\tau} & -\frac{k_d}{\tau} & -\frac{1}{\tau} \end{bmatrix} \begin{bmatrix} e_{1,i}(t) \\ e_{2,i}(t) \\ e_{3,i}(t) \end{bmatrix}, \tag{4}$$

Stability of (4) is obtained for any $h > 0$, $k_p > 0$ and $k_d > 0$. So, if the CACC local controller is implemented as described here, (4) becomes the dynamics of a stable autonomous linear system and the errors $e_{1,i}$, $e_{2,i}$ and $e_{3,i}$ converge to zero.

As it can be seen from (3) and (4), the local CACC controller for vehicle i needs to know at each time the position, velocity and the intended acceleration of the preceding vehicle in order to compute the local input u_i. As said, the first two variables are assumed to be measured by local sensors, but the feedforward term $u_{i-1}(t)$ is received through a wireless V2V communication network. A cyber attack to the V2V network will thus affect only the received value of the intended acceleration $u_{i-1}(t)$.

2.2 Effects of a cyber attack on the V2V network

In the present paper we assume that the effect of an attack, either implemented by forging or blocking packets transmitted on the V2V network or by installing malicious hardware or software on a vehicle, is to make the i–th vehicle receive the attacked value $\tilde{u}_{i-1}$ instead of the true one u_{i-1}. The signal

$$\Delta u_{i-1}(t) = u_{i-1}(t) - \tilde{u}_{i-1}(t)$$

is defined as the *attack signal* and is thus equal to the difference between the *physical* value u_{i-1} and the *received* value $\tilde{u}_{i-1}$ of the preceding vehicle intended acceleration.

Remark 1. The attack model assumed here is general enough to allow to model a wide class of attacks. For instance, a DoS attack can be obtained by choosing $\tilde{u}_{i-1}(t) = u_{i-1}(T_a)$ for every $t \geq T_a$, where T_a is the attack start time. A replay attack, instead, corresponds to $\tilde{u}_{i-1}(t) = u_{i-1}(t - T_a + T_r)$, where T_r denotes the attack recording start time and $T_a > T_r$ is the attack replay start time. Finally, a data injection attack can be described by an arbitrary attack signal $\Delta u_{i-1}(t)$, as described in Teixeira et al. (2015b).

Remark 2. We highlight once more that an attack will only change the value $\tilde{u}_{i-1}(t)$ received by the i–th vehicle, but the preceding vehicle will still locally command its drivetrain using the intended acceleration value $u_{i-1}(t)$. For this reason, we termed the last value as the *physical* value, to indicate it cannot be affected by a cyber attack. This fact will be enabling in letting us design in the next section an observer capable of detecting a cyber attack.

As in the attacked case it holds $\Delta u_{i-1}(t) \neq 0$, then the error dynamics for the i–th vehicle become

$$\begin{aligned}\begin{bmatrix}\dot{e}_{1,i}(t)\\ \dot{e}_{2,i}(t)\\ \dot{e}_{3,i}(t)\end{bmatrix} &= \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -\frac{k_p}{\tau} & -\frac{k_d}{\tau} & -\frac{1}{\tau}\end{bmatrix}\begin{bmatrix}e_{1,i}(t)\\ e_{2,i}(t)\\ e_{3,i}(t)\end{bmatrix}\\ &\quad + \begin{bmatrix}0\\ 0\\ 1\end{bmatrix}\Delta u_{i-1}(t)\,.\end{aligned} \tag{5}$$

Since the state matrix in (5) is not changed, the stability analysis holds even during an attack. Anyway, an attacker may design $\Delta u_{i-1}(t)$ to be big enough to cause $e_i(t) \leq -d_{r,i}(t)$, which would correspond to a collision between adjacent vehicles.

3. ADAPTIVE SLIDING MODE OBSERVER

In this section, an adaptive sliding mode observer is designed in order to detect additive attacks $\Delta u_{i-1}(t)$ on the data communicated to the generic i–th vehicle by the preceding one. First of all, we need to complete the state space equation (5) with the following output equation

$$y_i(t) = \begin{bmatrix}e_{1,i}(t)\\ e_{2,i}(t)\end{bmatrix},$$

which accounts for the fact that the i–th vehicle can compute the distance and velocity error with respect to the preceding one, thanks to its frontal radar. As the acceleration error is instead not measurable, we will redefine the state of the error dynamics as $z_i \triangleq [z_{1,i},\ z_{2,i}]^\top$, where $z_{1,i}(t) = [e_{1,i}^\top(t),\ e_{2,i}^\top(t)]^\top$ and $z_{2,i}(t) = e_{3,i}(t)$ represent the measurable and the unmeasurable states, respectively. Equation (5) can thus be rewritten as

$$\begin{aligned}\begin{bmatrix}\dot{z}_{1,i}(t)\\ \dot{z}_{2,i}(t)\end{bmatrix} &= \begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}z_{1,i}(t)\\ z_{2,i}(t)\end{bmatrix}\\ &\quad + \begin{bmatrix}0_{2\times 1}\\ b\end{bmatrix}\Delta u_{i-1}(t),\\ y_i(t) &= z_{1,i}(t),\end{aligned} \tag{6}$$

where

$$A_{11} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix},\qquad A_{12} = \begin{bmatrix}0\\ 1\end{bmatrix},$$
$$A_{21} = \begin{bmatrix}-\frac{k_p}{\tau} & -\frac{k_d}{\tau}\end{bmatrix}, A_{22} = -\frac{1}{\tau}, b = \frac{1}{\tau}\,.$$

A change of coordinates $\begin{bmatrix}I_{2\times 2} & 0_{2\times 1}\\ M & 1\end{bmatrix}$, is introduced, dependent on a matrix $M \in \mathbb{R}^{2\times 1}$ to be described later, resulting in the following state space:

$$\begin{aligned}\begin{bmatrix}\dot{z}'_{1,i}(t)\\ \dot{z}'_{2,i}(t)\end{bmatrix} &= \begin{bmatrix}A'_{11} & A'_{12}\\ A'_{21} & A'_{22}\end{bmatrix}\begin{bmatrix}z'_{1,i}(t)\\ z'_{2,i}(t)\end{bmatrix}\\ &\quad + \begin{bmatrix}0_{2\times 1}\\ b'\end{bmatrix}\Delta u_{i-1}(t),\\ y_i(t) &= z'_{1,i}(t),\end{aligned} \tag{7}$$

where $z'_{1,i}(t) = z_{1,i}(t), z'_{2,i}(t) = Mz_{1,i}(t) + z_{2,i}(t)$, and

$$\begin{aligned}A'_{11} &= A_{11} - A_{12}M,\ A'_{12} = A_{12},\\ A'_{21} &= MA_{11} + A_{21} - A'_{22}M,\\ A'_{22} &= MA_{12} + A_{22},\quad b' = b.\end{aligned}$$

We now need the following

Assumption 1. The attack signal $\Delta u_{i-1}(t)$ and its time derivative are bounded, that is $|\Delta u_{i-1}(t)| \leq \overline{\Delta u}_{i-1}$ and $|\Delta \dot{u}_{i-1}(t)| \leq \overline{\Delta \dot{u}}_{i-1}$, with $\overline{\Delta u}_{i-1} \in \mathbb{R}_+$ and $\overline{\Delta \dot{u}}_{i-1} \in \mathbb{R}_+$ being unknown but finite quantities.

We can then design the following adaptive sliding mode observer for system (7)

$$\begin{aligned}\dot{\hat{z}}'_{1,i}(t) &= A'_{11}\hat{z}'_{1,i}(t) + A'_{12}\hat{z}'_{2,i}(t) - \nu_i(t)\\ \dot{\hat{z}}'_{2,i}(t) &= A'_{21}y_i(t) + A'_{22}\hat{z}'_{2,i}(t)\\ \hat{y}_i(t) &= \hat{z}'_{1,i}(t)\end{aligned} \tag{8}$$

where the control signal $\nu_i(t) \in \mathbb{R}^2$ makes the states of the system to slide along the *sliding surface* $\epsilon_{y,i}(t) = \hat{y}_i(t) - y_i(t)$, and is defined as

$$\begin{aligned}\nu_i(t) &= (A'_{11} + P)\epsilon_{y,i}(t) + \Lambda_i(t)\mathrm{sgn}(\epsilon_{y,i}(t)),\\ \dot{\lambda}_{i_j}(t) &= \bar{\lambda}_{i_j}|\epsilon_{y,i_j}(t)|,\quad \text{for } j = 1, 2.\end{aligned} \tag{9}$$

with $P \in \mathbb{R}^{2\times 2}$ being a positive matrix, $\bar{\lambda}_{i_j} > 0$ and $\Lambda_i(t) = \mathrm{diag}\{\lambda_{i_1}, \lambda_{i_2}\}$.

By defining the observer errors as

$$\begin{aligned}\epsilon_{1,i}(t) &= \hat{z}'_{1,i}(t) - z'_{1,i}(t),\\ \epsilon_{2,i}(t) &= \hat{z}'_{2,i}(t) - z'_{2,i}(t),\\ \epsilon_{y,i}(t) &= \hat{z}'_{1,i}(t) - z'_{1,i}(t) = \epsilon_{1,i}(t),\end{aligned}$$

then the error dynamics can be expressed in the following form

$$\dot{\epsilon}_{1,i}(t) = A'_{11}\epsilon_{1,i}(t) + A'_{12}\epsilon_{2,i}(t) - \nu_i(t), \tag{10a}$$
$$\dot{\epsilon}_{2,i}(t) = A'_{22}\epsilon_{2,i}(t) - b'\Delta u_{i-1}(t), \tag{10b}$$

By suitably choosing the matrix M, the matrix A'_{22} can be designed to be negative definite matrix and nonsingular. This can be obtained, for instance, by solving a corresponding LMI. The solution to (10b) is

$$\epsilon_{2,i}(t) = e^{A'_{22}t}\epsilon_{2,i}(0) - \int_0^t e^{A'_{22}(t-k)}b'\Delta u_{i-1}(k)\mathrm{d}k,$$

from which, as A'_{22} is *Hurwitz*, it follows that

$$e^{A'_{22}t}\epsilon_{2,i}(0) \to 0 \quad \text{as} \quad t \to \infty,$$

and thus we get

$$\epsilon_{2,i}(t) = -\int_0^t e^{A'_{22}(t-k)}b'\Delta u_{i-1}(k)\mathrm{d}k,$$

By use of the Laplace transformation properties and the assumption that the attack is slow varying and can be approximated as constant, *after sliding has occurred* $\epsilon_{2,i}(t)$ will have the following form:

$$\epsilon_{2,i}(t) = A'^{-1}_{22}b'\Delta u_{i-1}(t) \tag{11}$$

Theorem 1. (Stability). The control signal $\nu_i(t)$ in (9) will force the states of the system (6) to slide along the defined sliding surface, resulting in the error $\epsilon_{y,i}(t)$ in (10a) converging to zero.

Proof. Consider the following Lyapunov function

$$\begin{aligned} V(\epsilon_{y,i}(t), \tilde{\lambda}_i(t)) =& \frac{1}{2}\epsilon^\top_{y,i}(t)\epsilon_{y,i}(t) \\ &+ \frac{1}{2}\tilde{\lambda}^\top_i(t)\bar{\Lambda}^{-1}_i\tilde{\lambda}_i(t), \end{aligned} \tag{12}$$

where $\tilde{\lambda}_i(t) = \lambda_i(t) - \beta$, $\beta \in \mathbb{R}^2$ is a positive constant vector and $\bar{\Lambda}_i(t) = diag\{\bar{\lambda}_1, \bar{\lambda}_2\}$. The derivative of the Lyapunov candidate (12) along the trajectories of $\epsilon_{y,i}(t)$ and $\tilde{\lambda}_i(t)$ is as follows

$$\begin{aligned} \dot{V}(\epsilon_{y,i}(t), \tilde{\lambda}_i(t)) =& \epsilon^\top_{y,i}(t)\Big[A'_{11}\epsilon_{y,i}(t) + A'_{12}\epsilon_{2,i}(t) \\ &- (A'_{11} + P)\epsilon_{y,i}(t) \\ &- \Lambda_i(t)\mathrm{sgn}(\epsilon_{y,i}(t))\Big] + \dot{\tilde{\lambda}}^\top_i(t)\bar{\Lambda}^{-1}_i\tilde{\lambda}_i(t) \\ =& -\epsilon^\top_{y,i}(t)P\epsilon_{y,i}(t) + \epsilon^\top_{y,i}(t)\big(\bar{A}_{12}\epsilon_{2,i}(t)\big) \\ &- |\epsilon^\top_{y,i}(t)|\lambda_i(t) + |\epsilon^\top_{y,i}(t)|\tilde{\lambda}_i(t), \end{aligned}$$

By considering the fact that $\epsilon_{2,i}(t)$ is bounded as A'_{22} can be designed to be negative by proper choice of the matrix M, it can be said that there exists a positive constant vector Ω such that

$$|A'_{12}\epsilon_{2,i}(t)| \leq \Omega,$$

resulting in the following inequality

$$\begin{aligned} \dot{V}(\epsilon_{y,i}(t), \tilde{\lambda}(t)) \leq& -\epsilon^\top_{y,i}(t)P\epsilon_{y,i}(t) + |\epsilon^\top_{y,i}(t)|\Omega \\ &- |\epsilon^\top_{y,i}(t)|\lambda_i(t) + |\epsilon^\top_{y,i}(t)|\tilde{\lambda}_i(t) \\ =& -\epsilon^\top_{y,i}(t)P\epsilon^\top_{y,i}(t) + |\epsilon^\top_{y,i}(t)|(\Omega - \beta), \end{aligned}$$

where the constant vector β can be chosen is such way that Ψ defined as $\Psi = \beta - \Omega$, results in a positive vector and hence

$$\begin{aligned} \dot{V}(\epsilon_{y,i}(t), \tilde{\lambda}(t)) \leq& -\epsilon^\top_{y,i}(t)P\epsilon_{y,i}(t) - |\epsilon^\top_{y,i}(t)|\Psi \\ \leq& |\epsilon^\top_{y,i}(t)|\Psi \leq 0, \end{aligned} \tag{13}$$

thus it is concluded that the derivative of the Lyapunov candidate (12) is less than or equal to zero, it follows that $\epsilon_{y,i}(t)$ and $\tilde{\lambda}(t)$ are bounded. By integrating both sides of the inequality (13), it follows

$$\begin{aligned} \lim_{t\to\infty}\int_0^t |\epsilon^\top_{y,i}(t)|\Psi\mathrm{d}\tau \leq& \lim_{t\to\infty}[V(\epsilon_{y,i}(0), \tilde{\lambda}_i(0)) \\ &- V(\epsilon_{y,i}(t), \tilde{\lambda}_i(t)] \leq \infty. \end{aligned}$$

Since $V(\epsilon_{y,i}(t), \tilde{\lambda}(t))$ is bounded, it can be concluded that $\lim_{t\to\infty}\int_0^t |\epsilon^\top_{y,i}(t)|\Psi\mathrm{d}\tau$ is also bounded and hence $\dot{\epsilon}_{y,i}(t)$ is bounded as $\epsilon_{y,i}(t), \epsilon_{2,i}(t)$ and $\nu_i(t)$ are bounded. Therefore, $\Psi^\top|\epsilon_{y,i}(t)|$ is uniformly continuous in t. By use of *Barbarat's Lemma* it follows that $\lim_{t\to\infty}\Psi^\top|\epsilon_{y,i}(t)| = 0$ and, as Ψ is a positive vector, that: $\lim_{t\to\infty}|\epsilon_{y,i}(t)| = 0$. ■

By using the theorem's result and the fact that the attack is assumed to be slow varying, the attack can be detected and reconstructed by substituting (11) in (10a), as follows:

$$0 = A'_{12}\Big(A'^{-1}_{22}b'\Delta u_{i-1}(t)\Big) - \nu_i(t)$$

thus we get

$$\Delta u_{i-1}(t) = b'^{-1}A'_{22}A'^{-1}_{12}\nu_i(t). \tag{14}$$

3.1 A robust dynamic detection threshold

In this section, we will introduce the attack detection logic and account for the inevitable measurement uncertainties affecting each vehicle local sensor. We will so design a robust detection threshold such that measurement noises do not cause a false alarm. First of all, in the presence of measurement uncertainties the output equation in (7) can be written as:

$$y_i(t) = z'_{1,i}(t) + \zeta_i(t),$$

where $\zeta_i(t) = [\zeta_{1,i}(t),\ \zeta_{2,i}(t)]^\top$ represents the measurement uncertainties on the i-th vehicle's sensors. We will assume the following on $\zeta_{j,i}(t)$, with $j \in \{1, 2\}$:

Assumption 2. The measurement uncertainty $\zeta_{j,i}(t)$ and its time derivative are bounded, that is $|\zeta_{j,i}(t)| \leq \overline{\zeta}_{j,i}(t)$ and $|\dot{\zeta}_{j,i}(t)| \leq \overline{\dot{\zeta}}_{j,i}(t)$, with $\overline{\zeta}_{j,i}(t) \in \mathbb{R}_+$ and $\overline{\dot{\zeta}}_{j,i}(t) \in \mathbb{R}_+$ being known and finite quantities.

Now, we introduce the detection logic used in this paper:

Definition 1. (Detection). A cyber attack affecting system (5) is said to be *detected* if there exist at least one time instant t and one component $j \in \{1, 2\}$ such that

$$|\nu_{j,i}(t)| > \overline{\nu_{j,i}}(t)\,,$$

where the observer control input $\nu_{j,i}(t)$ acts as a *dynamic detection residual* and the signal $\overline{\nu_{j,i}}(t)$ as a *dynamic detection threshold*.

Remark 3. As anticipated, a novel contribution of the present paper is to use the observer control input $\nu_{j,i}(t)$ as a detection residual, as opposed to existing works where the observer output estimation error ϵ_y would have been used for the same purpose.

Now, the problem consists in selecting a detection threshold $\overline{\nu_i}$ with suitable properties, which in our case corresponds to being robust to measurement uncertainties. We start by writing the observer estimation error dynamics in the case of measurement uncertainties:

$$\begin{aligned} \dot{\epsilon}_{1,i}(t) &= A'_{11}\epsilon_{1,i}(t) + A'_{12}\epsilon_{2,i}(t) - \nu_i(t), \\ \dot{\epsilon}_{2,i}(t) &= A'_{21}\zeta_i(t) + A'_{22}\epsilon_{2,i}(t), \\ \dot{\epsilon}_{y,i}(t) &= A'_{11}\epsilon_{1,i}(t) + A'_{12}\epsilon_{2,i}(t) - \nu_i(t) - \dot{\zeta}_i(t)\,. \end{aligned} \tag{15}$$

After sliding has occurred, we have

$$\epsilon_{y,i}(t) = \dot{\epsilon}_{y,i}(t) = 0 \Rightarrow \epsilon_{1,i}(t) = \zeta_i(t),$$

By solving for the uncertain error dynamics (15), remembering (9) and applying the Comparison Lemma, it holds that

$$|\nu_{j,i}(t)| \leq \overline{\nu_{j,i}}(t)$$

for all components $j \in \{1, 2\}$ and time instants t before the attack time T_a, when the threshold $\overline{\nu_i}(t) = [\overline{\nu_{1,i}}(t),\ \overline{\nu_{2,i}}(t)]$ is defined as

$$\overline{v, i}(t) \triangleq A'_{12}A'^{-1}_{22}(1 - e^{A'_{22}t})A'_{21}\bar{\zeta}_i + A'_{11}\bar{\zeta}_i - \dot{\bar{\zeta}}.$$

4. SIMULATION RESULT

In this section, the designed adaptive sliding mode observer is implemented on a simulated CACC-equipped string of three vehicles. The value of the CACC parameters are: $\tau = 0.1$, $k_p = 0.2$, $k_d = 0.7$, $h = 0.7$, $L = 2$ and $r = 1.5$. The design parameters are chosen as: $\bar{\lambda} = 12$, $P = 20 \cdot I_{2\times 2}$. It is assumed that the communication link between the leader vehicle and the second vehicle at time $t = 20s$ is subjected to an attack with the value 5, i.e.,

$$\tilde{u}_{i-1}(t) = 5 + u_{i-1}(t) \Rightarrow \Delta u_{i-1}(t) = 5\,\mathrm{m}\cdot\mathrm{s}^{-2}.$$

The measurement uncertainty ζ is chosen to be a uniform random variable with magnitude limited between $\pm$ 2% of sensor range, and with its derivative magnitude bounded by 0.2. This results in a threshold $\overline{\nu_i}$ with components equal to 0.24.

Figure 2 shows the distance and the desired distance between the vehicles in case of no attack, and as it can be seen, the vehicles start with an initial distance and reach the desired distance after some finite time. Figure 3 illustrates the distance and the desired distance for the case where the system has been subjected to the attack, in the absence of the adaptive sliding mode observer and it can be seen that the attack will cause significant changes in the distance between the vehicles by altering the inter vehicle distance and resulting in a collision, which is a disastrous consequence. Figure 4 shows the estimation of the attack by use of the proposed adaptive sliding mode observer. Therefore, estimation can be used in order to compensate the effect of the attack, resulting the system to behave normally as in Figure 2.

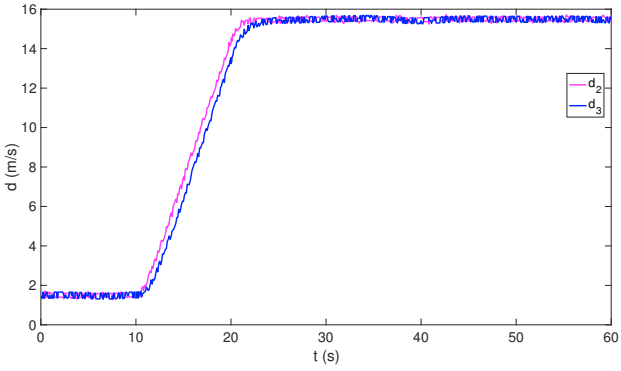

Fig. 2. The distance between the vehicles in case of no attack

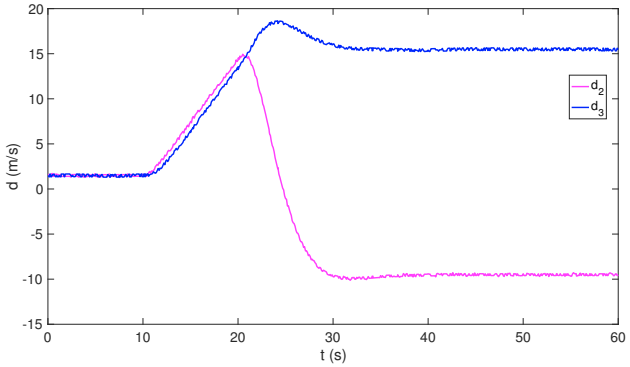

Fig. 3. The distance between the vehicles in case of attack and absence of the adaptive sliding mode observer

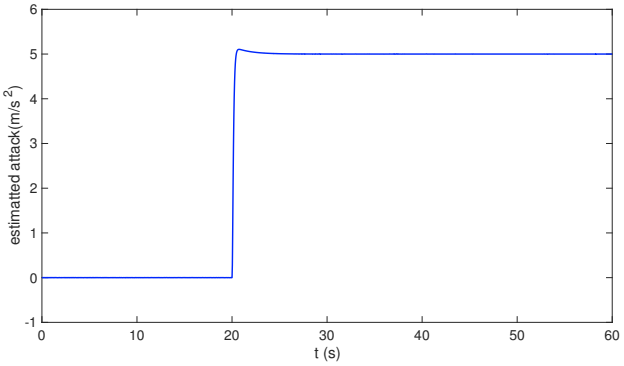

Fig. 4. The estimated attack by use of the adaptive sliding mode observer. As the threshold is equal to 0.24, we can see that detection is almost instantaneous.

5. CONCLUDING REMARKS

Platoons of cooperative autonomous vehicles, such as those equipped with CACC, need to communicate via a wireless V2V network. In particular, each vehicle needs to receive from the preceding one the value of its intended acceleration. By using this information and measurement of inter-vehicle relative distance and velocity, CACC–equipped platoons enjoy string-stability and a lower safe inter-vehicle distance than ACC–equipped ones. Anyway, the use of a wireless V2V network exposes such platoons to the effects of cyber-attacks that may alter, or block the communication between adjacent vehicles. While CACC algorithms feature a limited robustness to such effects, there is a need to develop attack detection methods as a first step toward full attack tolerance.

In this paper, we proposed to endow each vehicle with an adaptive sliding mode observer, whose aim is to estimate not its local dynamics, but those of the *preceding* vehicle. By using such estimates, the local measurements and the values received via the V2V network, it is possible to compute a detection residual and a threshold. Indeed, it could be said that by using an observer based on the *physics* of the preceding vehicle, it is possible to detect anomalies in its *cyber* part.

As a novel contribution, in this paper the sliding mode observer equivalent control input is used as a residual,

as opposed to other approaches where the observer estimation error is chosen for this. Current theoretical and simulation results are based on the assumption that the attack is slowly varying and can be approximated as being constant. As a future work, a relaxation of this assumption will be investigated, and a larger class of cyber attacks will be considered. Furthermore, more complex vehicle dynamical models, e.g. larger order and/or nonlinear models, will be considered.

REFERENCES

Amoozadeh, M., Raghuramu, A., n. Chuah, C., Ghosal, D., Zhang, H.M., Rowe, J., and Levitt, K. (2015). Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. *IEEE Commun. Mag.*, 53(6), 126–132.

Biron, Z.A., Dey, S., and Pisu, P. (2017). Resilient control strategy under denial of service in connected vehicles. In *2017 American Control Conference (ACC)*, 4971–4976.

Bißmeyer, N., Mauthofer, S., Bayarou, K.M., and Kargl, F. (2012). Assessment of node trustworthiness in VANETs using data plausibility checks with particle filters. In *2012 IEEE Vehicular Networking Conference (VNC)*, 78–85.

Cárdenas, A.A., Amin, S., and Sastry, S.S. (2008). Secure control: Towards survivable Cyber-Physical systems. In *First International Workshop on Cyber-Physical Systems*.

Cárdenas, A.A., Amin, S., Sinopoli, B., Giani, A., Perrig, A., and Sastry, S.S. (2009). Challenges for securing cyber physical systems. In *Workshop on Future Directions in Cyber-physical Systems Security*.

Ferrari, R.M.G. and Teixeira, A.M.H. (2017). Detection and isolation of replay attacks through sensor watermarking. *IFAC-PapersOnLine*, 50(1), 7363–7368.

Gehring, O. and Fritz, H. (1997). Practical results of a longitudinal control concept for truck platooning with vehicle to vehicle communication. In *Intelligent Transportation System, 1997. ITSC'97., IEEE Conference on*, 117–122. IEEE.

Lei, C., van Eenennaam, E.M., Wolterink, W.K., Karagiannis, G., Heijenk, G., and Ploeg, J. (2011). Impact of packet loss on CACC string stability performance. In *2011 11th International Conference on ITS Telecommunications*, 381–386.

Meskin, N. and Khorasani, K. (2009). Actuator fault detection and isolation for a network of unmanned vehicles. *IEEE Trans. Automat. Contr.*, 54(4), 835–840.

Miller, C. and Valasek, C. (2014). A survey of remote automotive attack surfaces. *black hat USA*, 2014.

Naus, G., Vugts, R., Ploeg, J., van de Molengraft, R., and Steinbuch, M. (2010). Cooperative adaptive cruise control, design and experiments. In *American Control Conference (ACC), 2010*, 6145–6150. IEEE.

Pasqualetti, F., Dorfler, F., and Bullo, F. (2013). Attack detection and identification in Cyber-Physical systems. *IEEE Trans. Automat. Contr.*, 58(11), 2715–2729.

Ploeg, J. (2017). Cooperative vehicle automation: Safety aspects and control software architecture. In *2017 IEEE International Conference on Software Architecture Workshops (ICSAW)*, 6–6.

Ploeg, J., Scheepers, B.T.M., van Nunen, E., de Wouw, N.v., and Nijmeijer, H. (2011). Design and experimental evaluation of cooperative adaptive cruise control. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 260–265.

Ploeg, J., Semsar-Kazerooni, E., Lijster, G., de Wouw, N.v., and Nijmeijer, H. (2013). Graceful degradation of CACC performance subject to unreliable wireless communication. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 1210–1216.

Quan, Y., Chen, W., Wu, Z., and Peng, L. (2018). Distributed fault detection and isolation for leader–follower multi-agent systems with disturbances using observer techniques. *Nonlinear Dyn.*

Studnia, I., Nicomette, V., Alata, E., Deswarte, Y., Kaâniche, M., and Laarouchi, Y. (2013). Survey on security threats and protection mechanisms in embedded automotive networks. In *2013 43rd Annual IEEE/IFIP Conference on Dependable Systems and Networks Workshop (DSN-W)*, 1–12. ieeexplore.ieee.org.

Teixeira, A., Shames, I., Sandberg, H., and Johansson, K.H. (2012). Revealing stealthy attacks in control systems. In *50th Annual Allerton Conference on Communication, Control, and Computing*.

Teixeira, A., Shames, I., Sandberg, H., and Johansson, K.H. (2015a). A secure control framework for Resource-Limited adversaries. *Automatica*, 51(1), 135–148.

Teixeira, A., Shames, I., Sandberg, H., and Johansson, K.H. (2015b). A secure control framework for resource-limited adversaries. *Automatica*, 51, 135–148.