Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.
Detection of Cyber-Attacks in Collaborative Intersection Control

Twan Keijzer, Fabian Jarmolowitz, and Riccardo M.G. Ferrari

Abstract—Road intersections are widely recognized as a lead cause for accidents and traffic delays. In a future scenario with a significant adoption of Cooperative Autonomous Vehicles, solutions based on fully automatic, signage-less Intersection Control would become viable. Such a solution, however, requires communication between vehicles and, possibly, the infrastructure over wireless networks. This increases the attack surface available to a malicious actor, which could lead to dangerous situations. In this paper, we address the safety of Intersection Control algorithms, and design a Sliding-Mode-Observer based solution capable of detecting and estimating false data injection attacks affecting vehicles’ communication. With respect to previous literature, a novel detection logic with improved detection performances is presented. Simulation results are provided to show the effectiveness of the proposed approach.

I. INTRODUCTION

Classical road intersections for human-driven vehicles are managed via fixed signage and traffic lights, which allows for a sub-optimal vehicle throughput while guaranteeing an adequate safety level. Still, it is well known that classical intersections are a major cause of accidents, due to human error, and traffic delays [1]. In the early 2000s, based on the projected introduction of Cooperative Autonomous Vehicles (CAVs), works such as [2] proposed replacing classical intersections with safer, automated solutions. Intersection Control (IC) would thus automatize the tasks of negotiating, planning and executing the trajectories of CAVs in order to increase the safety and vehicle throughput of the junction.

A key enabling technology of an IC solution are wireless networks, allowing real-time communication of measurements and control signals between CAVs, and to the roadside infrastructure. The security and robustness of such Vehicle-to-Everything (V2X) networks, is thus of paramount importance for the safety of the IC itself. As a first line of defence, current V2X protocols include encryption and authentication mechanisms to prevent intrusion. For instance, the Autosar standard with End-to-End Protection (E2E), following the ISO 26262 standard, is such a preventive protection measure and is analyzed in [3]. The case of communication disruption or false data injection by a malicious attacker [4], [5] that can circumvent these protections would instead require a different approach. Indeed, a major difference between an ordinary fault and a smart attacker is that the latter actively seeks to cause great harm to the IC, while minimizing the possibility of being detected [6], [7]. This lead to development of increasingly sophisticated anomaly detection approaches for V2X communication traffic, as a second line of defence that can provide a guarantee against either inside attackers, or sophisticated attackers that successfully infiltrated the system. These techniques include plausibility checks based on elementary models of the CAVs and IC [8], [9], [10], as well as more advanced, model-based approaches that have been proposed in the literature for detecting attacks in general Cyber-Physical Systems (CPS), such as [11], [12], [13], [14], [15]. Nevertheless, these techniques were not yet applied to CAVs on an automated intersection.

In this paper, we will address the problem of designing a second line of defence for IC subjected to false-data injection attacks. To this end, a novel cyber-attack detection method is presented based on a Sliding Mode Observer (SMO). In previous work by the authors [14] a detection logic based on the so-called Equivalent Output Injection (EOI) term of the SMO was presented. The novel approach proposed in the present paper no longer requires this EOI, leading to better detection performance. The proposed technique is verified in simulation of an IC scenario. Here a control approach called Virtual Platooning (VP, [16]) is used. However, also potentially better-performing optimization-based approaches could be applied [17], [18], [19], [20] without affecting the proposed detection method.

The structure of the paper is as follows: Section II introduces the problem addressed in this paper, while Sections III and IV, respectively, introduce the SMO design and the detection thresholds on which the proposed attack detection strategy is built. Simulation results are shown in Section V, while concluding remarks are finally drawn in Section VI.

II. PROBLEM FORMULATION

In this paper, detection of cyber-attacks on the inter-vehicle communication is considered in a collaborative intersection control scenario. Each car is modeled as

\[
\begin{bmatrix}
\dot{p}_i \\
\dot{v}_i \\
\end{bmatrix} = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix} \begin{bmatrix}
\dot{p}_i \\
\dot{v}_i \\
\end{bmatrix} + \begin{bmatrix}
\frac{1}{\tau} \\
0 \\
\end{bmatrix} u_i \\

y_i = \begin{bmatrix}
p_i - p_{i-1} - L_i \\
\gamma_i - v_{i-1} \\
\end{bmatrix} + \xi_i,
\]

where the subscripts \( i \) and \( i - 1 \) denote the variables are related to cars \( i \) and \( i - 1 \) respectively. \( p, v, a, u, \gamma, \xi, \tau, \) and \( L \) are, respectively, the distance from the rear of the vehicle to the intersection (negative when approaching the intersection), velocity, acceleration, input, measurements, sensor noise, engine time constant, and length of the cars.

The considered IC scenario is depicted in Figure 1. In this scenario, as can be seen from Equation (1), each car measures the relative distance to the intersection, and relative velocity from the preceding car, as well as its own velocity and acceleration. Mandatory for having these
measurements for all cars entering the intersection are either a central infrastructure with appropriate sensors and V2X communication and/or cars equipped with lateral sensors, e.g. [21]. Furthermore, each car receives, via wireless V2V communication, the input of the preceding car.

The interaction between two cars in an IC scenario can be modeled, from the perspective of car \( i \), as

\[
\begin{bmatrix}
\hat{x}_{i-1} \\
\hat{y}_i
\end{bmatrix} = \begin{bmatrix}
A_{i-1} & 0 \\
B_{i-1} & 0
\end{bmatrix} \begin{bmatrix}
\hat{x}_{i} \\
u_{i-1}
\end{bmatrix} + \begin{bmatrix}
0 \\
B_i
\end{bmatrix} \begin{bmatrix}
u_i
\end{bmatrix},
\]

(2)

where

\[
y_i = C_{\hat{x}_i} + \begin{bmatrix}
-L_i \\
0_{3 \times 1}
\end{bmatrix} e_i.
\]

Here, \( C \) can be derived from equation (1). Furthermore, one can see that the coupling between the vehicles appears in the measurement equation only. These measurements, as well as the communicated input \( u_{i-1} \), can be used by car \( i \) to calculate a control input \( u_i \) such that the IC objective is achieved. In this work, which primarily deals with the cyber-attack detection, any control law for IC can be chosen without affecting the detection method.

A. Model Uncertainty & Cyber-attack

In this section, System (2) will be rewritten to make the model uncertainty and cyber-attack explicit. To quantify the model uncertainty, the following assumption is made.

Assumption 1: Each car \( i \) is assumed know its own dynamics, represented by \( \tau_i \) and \( L_i \). It will, however, only have access to the nominal dynamics of the preceding car.

The nominal value of \( \tau_i \), available to car \( i \), is defined as

\[
\tau_i = r_i \tau_{i-1}.
\]

Furthermore, the cyber-attack is defined as a “man in the middle” attack, such that each car \( i \) will receive

\[
u_{i-1, r} = u_{i-1} + \Delta u_{i-1},
\]

where \( \Delta u_{i-1} \) is the cyber-attack. To make the model uncertainty and cyber-attack explicit a model of the form

\[
\begin{bmatrix}
\hat{x} \\
\hat{y}_i
\end{bmatrix} = \begin{bmatrix}
A & 0 \\
B & 0
\end{bmatrix} \begin{bmatrix}
\hat{x} \\
u_i
\end{bmatrix} + \begin{bmatrix}
0 \\
\eta_i
\end{bmatrix} + \begin{bmatrix}
u_i
\end{bmatrix},
\]

(3)

is proposed. Here \( A, B, \) and \( \eta \) are redefined using

\[
\tau_{i-1} \leftarrow \tau_i \text{ and } u_{i-1} \leftarrow u_{i-1, r},
\]

such that they are known to car \( i \), i.e. they can be used in its detection logic. Furthermore, \( E, F, \) and \( \eta \) are defined as

\[
E = \begin{bmatrix}
0_{2 \times 1} \\
0_{3 \times 1}
\end{bmatrix} ; F = \begin{bmatrix}
0_{2 \times 1} \\
0_{3 \times 1}
\end{bmatrix} ; \eta = (r_i - 1)(u_{i-1} - a_{i-1}),
\]

where \( E \) and \( F \) are known to car \( i \) and \( \eta \) is unknown uncertainty. Note that Systems (3) and (2) are only reformulated.

The following assumptions are made on system (3).

Assumption 2: The sensor noise \( \zeta_i \) is zero-mean and bounded by a known value \( \tilde{\zeta} \geq |\zeta_i| \).

Assumption 3: The uncertainty \( \eta \), and cyber-attack \( \Delta u_{i-1} \) are bounded by known values \( \tilde{\eta} \geq |\eta| \) and \( \tilde{\Delta} \geq |\Delta u_{i-1}| \). These bounds are defined for the IC scenario in section V.

B. Model Transformation

In System (3) the cyber-attack appears as an unknown input. This allows for the use of an SMO for estimation and detection of the cyber-attack [14]. In order to implement this SMO based approach, the system is transformed to

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix} \begin{bmatrix}
\hat{x}_1 \\
\hat{x}_2
\end{bmatrix} + \begin{bmatrix}
B_1 \\
B_2
\end{bmatrix} u + \begin{bmatrix}
E_1 \\
E_2
\end{bmatrix} \eta + \begin{bmatrix}
F_1 \\
F_2
\end{bmatrix} \Delta u_{i-1} + \begin{bmatrix}
0 \\
0
\end{bmatrix},
\]

(4)

using a transformation introduced by [23]. Here it is required that \( A_{11} \prec 0 \) to ensure that the observer error dynamics are stable. The following assumption ensures \( A_{11} \prec 0 \) [23].

Assumption 4: The invariant zeros of \((A,F,C)\) lie in \( \mathbb{C}_- \).

III. SLIDING MODE OBSERVER DESIGN

Based on [14], the following SMO is introduced

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix} \begin{bmatrix}
\hat{x}_1 \\
\hat{x}_2
\end{bmatrix} + \begin{bmatrix}
B_1 \\
B_2
\end{bmatrix} u - \begin{bmatrix}
A_{12} \\
A_{22}
\end{bmatrix} \eta + \begin{bmatrix}
F_1 \\
F_2
\end{bmatrix} \Delta u_{i-1} + \begin{bmatrix}
0 \\
0
\end{bmatrix},
\]

(5)

where \( A_{22} \Delta u = 22_2 - 22_x \), \( 22_x \prec 0 \) is the linear gain, the diagonal matrix \( M \prec 0 \) is the switching gain, and \( e_i \triangleq y_i - \hat{y}_i \).

Based on the system dynamics of equation (4) and the SMO of equation (5), the observer error dynamics become

\[
\begin{bmatrix}
\dot{e}_1 \\
\dot{e}_2
\end{bmatrix} = \begin{bmatrix}
A_{11} & 0 \\
A_{21} & A_{22}
\end{bmatrix} \begin{bmatrix}
e_1 \\
e_2
\end{bmatrix} + \begin{bmatrix}
A_{12} \\
A_{22}
\end{bmatrix} \eta - \begin{bmatrix}
F_1 \\
F_2
\end{bmatrix} \Delta u_{i-1} + \begin{bmatrix}
0 \\
0
\end{bmatrix}. \]

(6)

where \( e_1 \triangleq \hat{x}_1 - x_1 \) and \( e_2 \triangleq \hat{x}_2 - x_2 \). Furthermore, \( e_i \) can be expressed in terms of \( e_2 \) as \( e_i = e_2 - \hat{\zeta}_i \).

Lemma 1 presents bounds on \( e_1, e_2, \) and \( \hat{\zeta}_i \) in healthy and attacked conditions. These bounds are used in the detection logic design and detectability analysis in section IV.

Lemma 1: Define \( \epsilon \leq |e_1| \), max \( (|e_1|, |e_2|) = \epsilon_1, \epsilon_2 \) \leq \( \epsilon_2 \). Furthermore, \( \epsilon_1 \) denote bounds in healthy conditions, when \( \Delta u_{i-1} = 0 \), with a superscript 0.

If, \( \epsilon_1 \) elementwise, \( \text{diag}(M) > |A_{21}|\hat{\epsilon}_1 + |A_{22}|\hat{\zeta}_i + |E_2|\hat{\eta} + |F_2|\Delta \), then the following conditions hold

1) \( e_1 \leq \epsilon_1 \triangleq \epsilon_1 \triangleq \epsilon_1 \prec \epsilon_1 \)
2) \( e_1 \leq \epsilon_1 \triangleq \epsilon_1 \triangleq \epsilon_1 \prec \epsilon_1 \)
3) \( e_1 \leq \epsilon_1 \triangleq \epsilon_1 \triangleq \epsilon_1 \prec \epsilon_1 \)
4) \( e_1 \leq \epsilon_1 \triangleq \epsilon_1 \triangleq \epsilon_1 \prec \epsilon_1 \)
5) \( e_2 \leq \epsilon_2 \triangleq \epsilon_2 \triangleq \epsilon_2 \prec \epsilon_2 \)
6) \( e_2 \leq \epsilon_2 \triangleq \epsilon_2 \triangleq \epsilon_2 \prec \epsilon_2 \)
7) \( e_2 \leq \epsilon_2 \triangleq \epsilon_2 \triangleq \epsilon_2 \prec \epsilon_2 \)

Authorized licensed use limited to: TU Delft. Downloaded on January 06, 2022 at 12:23:21 UTC from IEEE Xplore. Restrictions apply.
8) \( \bar{e}_2 = e_2 + \text{sgn}(e_2) r(\Delta t_{(l-1)}) \)
9) \( \bar{e}_2 = e_2 + \text{sgn}(e_2) r(\Delta t_{(l-1)}) \)
10) \( \text{sgn}(e_2) = -\text{sgn}(e_2) \)

where \( r(\Delta t_{(l-1)}) = \int_0^t F_1 \Delta t_{(l-1)}(s) e^{\alpha_1(t-s)} ds \) and \( r(\Delta t_{(l-1)}) = -\frac{A_1}{2} r(\Delta t_{(l-1)}) + F_2 \Delta t_{(l-1)} \).

Proof: Statements 3)-7) and 10) are proven [14]. Statements 1), 2), 8) and 9) are proven in the appendix.

IV. DETECTION DESIGN

In this section, the novel cyber-attack detection method will be described. This method directly analyses the behaviour of observer error \( e_2 \), and uses this to detect cyber-attacks. For comparison, the EOI based detection method presented in previous work [14] is presented in subsection IV-D. The detection performance of the two methods in a collaborative IC scenario will be compared in section V.

The novel proposed detection logic uses thresholds on the observer error \( e_2 \) based on the bounds in Lemma 1, and the relation \( e_2 = e_2 - \bar{e}_2 \). The resulting thresholds, \( \bar{e}_2 \leq e_2 \leq \bar{e}_2 \), will be used for cyber-attack detection. Preferably one would directly monitor this condition, and detect a cyber-attack when it is violated. However, as \( e_2 \) is not known to the observer, this is not possible. Alternatively, the condition

\[
\bar{e}_2 > \bar{e}_2 \tag{7}
\]

can be monitored. Satisfying this condition implies violation of \( \bar{e}_2 \leq e_2 \leq \bar{e}_2 \), and thus serves as a detection condition.

A. Design of Error-bounds

At all time, \( |e_2| < \tilde{\zeta}_i \) (Lemma 1.5), and when a new measurement arrives to the observer \( e_2 - \tilde{\zeta}_i \leq e_2 \leq e_2 + \tilde{\zeta}_i \). Denote the sequence of measurement times as \( \{t_m\} \), which do not need to be equidistant. Then, for any time \( t_m \)

\[
\bar{e}_2(t_m) = \min(e_2(t_m) + \tilde{\zeta}_i, \tilde{\zeta}_i) \\
e_2(t_m) = \max(e_2(t_m) - \tilde{\zeta}_i, -\tilde{\zeta}_i) \tag{8}
\]

Furthermore, bounds on \( \bar{e}_2 \) are known from Lemma 1. With these, \( e_2 \) can be bound during each period \( [t_{m-1}, t_m] \) as

If \( e_2(t_{m-1}) < 0 \)

\[
\bar{e}_2(t) = \int_{t_{m-1}}^t \bar{e}_2(T) dT; e_2(t) = \int_{t_{m-1}}^t \bar{e}_2(T) dT. \tag{9}
\]

If \( e_2(t_{m-1}) > 0 \)

\[
\bar{e}_2(t) = -\int_{t_{m-1}}^t \bar{e}_2(T) dT; e_2(t) = -\int_{t_{m-1}}^t \bar{e}_2(T) dT. \tag{9}
\]

The above bounds require further inspection. At first sight they seem to depend only on the modeled system behaviour through \( \bar{e}_2 \) and \( \bar{e}_2 \). However, the bounds also depend on the real behaviour through \( e_2 \). The integration duration is dictated by the sign of \( e_2 \). The two bounds in equations (8) and (9) can be combined for \( m \geq 1 \) as

If \( e_2(t_{m-1}) < 0 \)

\[
\bar{e}_2(t_m) = \min(e_2(t_m) + \tilde{\zeta}_i, \tilde{\zeta}_i) \\
e_2(t_m) = \max(e_2(t_m) - \tilde{\zeta}_i, -\tilde{\zeta}_i)
\]

If \( e_2(t_{m-1}) > 0 \)

\[
\bar{e}_2(t_m) = \min(-\int_{t_{m-1}}^{t_m} e_2(T) dT, e_2(t_m) + \tilde{\zeta}_i, \tilde{\zeta}_i) \\
e_2(t_m) = \max(-\int_{t_{m-1}}^{t_m} e_2(T) dT, e_2(t_m) - \tilde{\zeta}_i, -\tilde{\zeta}_i)
\]

Equation (8) can be used to obtain \( \bar{e}_2(t_0) \) and \( e_2(t_0) \). Based on these bounds, the detection criterion (7) can be monitored at every measurement time \( t_m \).

B. Detectability Analysis

In this section, conditions are presented for which the proposed novel detection method can detect an attack. Furthermore, it is proven that in healthy conditions, the approach will never cause a detection.

First, introduce an assumption which is required in the presented proofs. This assumption is a relaxation of the matching condition commonly used in SMO literature[14].

Assumption 5: \( (F_2 - A_2 A_1^{-1} F_1) \) is full column rank.

First, it will be proven that no detection occurs in healthy conditions.

Theorem 1: Consider system (4), observer (5) and detection criterion (7). In healthy conditions, i.e. if \( \Delta t_{(l-1)} = 0 \forall t \), the detection criterion will never be satisfied.

Proof: Define the sequence \( \{t_{i}\} \) as the times where \( e_2 \) changes sign, \( e_2^j \) as the average \( |e_2| \) while \( e_2 > 0 \), and \( e_2^n \) as the average \( |e_2| \) while \( e_2 < 0 \). Furthermore, without loss of generality, assume \( e_2 \) becomes positive at every \( t_{2j} \), allowing to write \( t_i = t_{2j} - t_{2j-2} \).

Then, denote for the true dynamics of \( e_2 \) as

\[
e_2(t_{2j+2}) = e_2(t_{2j}) + e_1, \\
e_2(t_{2j+2n}) = e_2(t_{2j}) + \sum_{j=0}^{N} c_i \tag{10}
\]

where \( c_i = t_i \tilde{\zeta}_i - t_i \tilde{\zeta}_i \). Now \( c_i \) can be bounded, using the bounds on \( e_2 \) from lemma 1 and \( e_2 = e_2 - \zeta_i \), as

\[
-\bar{e}_2(t_{2j}) + \min(e_2(t_{2j}) - \tilde{\zeta}_i, -\tilde{\zeta}_i) \leq \sum_{j=0}^{N} c_i \leq -\bar{e}_2(t_{2j}) + \min(e_2(t_{2j}) - \tilde{\zeta}_i, \tilde{\zeta}_i) \forall N \in \mathbb{Z}.
\]

Furthermore, from equation (10), it can be derived that \( \frac{e_2}{t_i} = \bar{e}_2 + \frac{e_2}{t_i} \).

With this, \( \bar{e}_2 \) in equation (9) can be rewritten as

\[
\bar{e}_2(t_{2j+2}) = \bar{e}_2(t_{2j}) + t_i \frac{e_2}{t_i} (e_2(e_2 - \tilde{\zeta}_i) + \bar{e}_2(e_2 - \tilde{\zeta}_i) + \bar{e}_2(e_2 - \tilde{\zeta}_i)), \tag{11}
\]

which can be extended for \( e_2(t_{2j+2n}) \) as

\[
e_2(t_{2j+2n}) = e_2(t_{2j}) + \sum_{j=0}^{N} \frac{t_i}{t_i} \left( e_2(e_2 - \tilde{\zeta}_i) + \bar{e}_2(e_2 - \tilde{\zeta}_i) + \bar{e}_2(e_2 - \tilde{\zeta}_i) \right). \tag{12}
\]

It can be seen that in healthy conditions, when \( \frac{e_2}{t_i} \leq \bar{e}_2 \leq \frac{e_2}{t_i} \).
and \( \bar{e}_2^0 \leq \bar{e}_2^+ \leq \bar{e}_2^0 \),
\[
\bar{e}_2(t_{2s+2nN}) - \bar{e}_2(t_{2sN}) \geq \sum_{j=0}^{N-1} \frac{\bar{e}_2^0}{\bar{e}_2^0} c_{i+j} \geq \sum_{j=0}^{N-1} c_{i+j}
\]
(13)

By subtracting these inequalities it can be found that \( \bar{e}_2(t_{2s+2nN}) - \bar{e}_2(t_{2sN}) \geq \bar{e}_2(t_{2sN}) - \bar{e}_2(t_{2sN}) \), i.e., considering the behaviour in equation (9), the difference between \( \bar{e}_2 \) and \( \bar{e}_2 \) is non-decreasing. This only leaves to prove that no detection occurs if the bounds are affected by equation (8).

If both bounds are affected by equation (8), \( \bar{e}_2 - \bar{e}_2 = 2\bar{e}_2 - 2\bar{e}_2 \geq 0 \). If only the lower bound is affected, use equation (13) to derive
\[
\bar{e}_2(t_{2s+2nN}) - \bar{e}_2(t_{2sN}) \geq \bar{e}_2(t_{2sN}) - \bar{e}_2(t_{2sN}) + \max(e_j(t_{2sN}) - \bar{e}_2(t_{2sN})) \geq \bar{e}_2(t_{2s+2nN})
\]
This proves the theorem.

Then two lemmas are introduced to support the proof of theorem 2, where sufficient conditions for attack detection are presented.

Lemma 2: consider \( r(\Delta u_{t-1}) \) as defined in Lemma 1. Then the following statements can be proven

1) \( r(\Delta u_{t-1}) = 0 \) \( \forall t \) if \( \Delta u_{t-1} = 0 \) \( \forall t \), i.e., healthy conditions.
2) There always exists \( \gamma \geq 0 \) such that within finite time \( r(\Delta u_{t-1}) - (F_2 - A_{21}A_{11}^1 F_1) \Delta u_{t-1} \leq \gamma \).

Proof: By substituting \( \Delta u_{t-1} = 0 \) \( \forall t \) in the function for \( r(\Delta u_{t-1}) \), it can directly be seen that \( r(\Delta u_{t-1}) = 0 \) \( \forall t \). This proves statement a). For a constant \( \Delta u_{t-1} \),
\[
\lim_{t \to \infty} r(\Delta u_{t-1}) = (F_2 - A_{21}A_{11}^1 F_1) \Delta u_{t-1}.
\]
As \( r(\Delta u_{t-1}) \) is a smooth function, this means that within finite time \( r(\Delta u_{t-1}) - (F_2 - A_{21}A_{11}^1 F_1) \Delta u_{t-1} \leq \gamma \).

Lemma 3: Consider the behaviours of \( \bar{e}_2 \) and \( \epsilon_i \) from equation (11) and (12). Assume there exist \( \epsilon^+ > 0 \) and \( \epsilon^- > 0 \) such that

- \( \bar{e}_2^0 < \bar{e}_2^- \) and \( \bar{e}_2^2 > \bar{e}_2^+ \) for the period \( [t_{2s}, t_{2s+2nN}] \).
- OR \( \bar{e}_2^0 > \bar{e}_2^- + \epsilon^+ \) and \( \bar{e}_2^2 < \bar{e}_2^- + \epsilon^- \) for the period \( [t_{2s}, t_{2s+2nN}] \).

Then, there exists an \( \epsilon \) such that \( \bar{e}_2(t_{2s+2nN}) < \epsilon_i(t_{2s+2nN}) \), if \( N > \frac{\epsilon_k}{\epsilon^+} \) where \( \phi \leq \frac{t_{2s+2nN}}{\epsilon^+} \forall i,j \).

Proof: First, use \( \bar{e}_2^0 < \bar{e}_2^- - \epsilon^- \) and \( \bar{e}_2^2 > \bar{e}_2^+ + \epsilon^+ \) to derive
\[
\bar{e}_2^0 \bar{e}_2^+ - \bar{e}_2^0 \bar{e}_2^- < -(\bar{e}_2^- + \bar{e}_2^+ - \epsilon^- - \epsilon^+).
\]
Then substitute \( \bar{e}_2^0 \bar{e}_2^- - \bar{e}_2^0 \bar{e}_2^- < -\epsilon^- \) and \( \phi \leq \frac{t_{2s+2nN}}{\epsilon^+} \forall i,j \) in equation (11) giving
\[
\bar{e}_2(t_{2s+2nN}) - \bar{e}_2(t_{2sN}) < -N \epsilon \phi + \sum_{j=0}^{N-1} c_{i+j}.
\]
Using the bound on \( c_{i+j} \) gives
\[
\bar{e}_2(t_{2s+2nN}) < \bar{e}_2(t_{2sN}) - N \epsilon \phi - \epsilon_i(t_{2sN}) + \min(e_j(t_{2sN}) + \bar{e}_2(t_{2sN})) < -N \epsilon \phi + 2 \min(e_j(t_{2sN}) + \bar{e}_2(t_{2sN})) - \max(e_j(t_{2sN}) + \bar{e}_2(t_{2sN})) < \bar{e}_2(t_{2sN}) - \bar{e}_2(t_{2sN}) + \max(e_j(t_{2sN}) + \bar{e}_2(t_{2sN})) - \max(e_j(t_{2sN}) + \bar{e}_2(t_{2sN}))
\]
Meanwhile, always \( \epsilon_i(t_{2s+2nN}) > \max(e_j(t_{2sN}) + \bar{e}_2(t_{2sN})) \), which with some simplification leads to
\[
\bar{e}_2(t_{2s+2nN}) - \bar{e}_2(t_{2sN}) < -N \epsilon \phi + 4 \bar{e}_2(t_{2sN})
\]
So, \( \bar{e}_2(t_{2s+2nN}) < \epsilon_i(t_{2s+2nN}) \) if \( N > \frac{4 \bar{e}_2(t_{2sN})}{\epsilon_k} \).

The same result can be obtained by using \( \bar{e}_2^0 > \bar{e}_2^+ + \epsilon^+ \) and \( \bar{e}_2^2 < \bar{e}_2^- - \epsilon^- \) to obtain \( \bar{e}_2^0 \bar{e}_2^+ - \bar{e}_2^0 \bar{e}_2^- < \bar{e}_2^0 \bar{e}_2^- < \epsilon^- \). This is equivalent to the first condition in Lemma 3. The second condition holds if \( \bar{e}_2(t_{2s+2nN}) < \frac{\epsilon_k}{\epsilon^+} + \epsilon^- + \epsilon^- \) and can be proven similarly.

Furthermore, in Lemma 2 it is shown that there exists a \( \gamma > 0 \) such that \( r(\Delta u_{t-1}) - (F_2 - A_{21}A_{11}^1 F_1) \Delta u_{t-1} \leq \gamma \) within finite time. Therefore, if assumption 5 holds, there always exists a \( \Delta u_{t-1} \) such that \( r(\Delta u_{t-1}) \geq \delta \).

C. Equivalent Output Injection based detection

In this subsection the equivalent output injection (EOI) based detection, as previously introduced in [14], is presented for comparison with the novel detection method. First the EOI is defined as
\[
\bar{e}_i = K(\nu - \bar{v}_i),
\]
where \( \bar{v}_i \) is the EOI, and \( K > 0 \) is a diagonal gain matrix.

The EOI was originally introduced to estimate the cyber-attack. In [14] the following was proven.

Proposition 1: Consider noiseless system (4), where \( \bar{e}_2 \) is known. If \( \text{diag}(M) > |A_{21}| \bar{e}_2 + |A_{22}| \bar{e}_2 + |E_2| \bar{e}_2 + |F_2| \bar{e}_2 \), then \( \bar{e}_2 \to 0 \) and \( \bar{e}_2 \to 0 \). Furthermore, assuming a constant cyber-attack, the cyber-attack estimate
\[
\hat{\Delta} u_{t-1} = (F_2 - A_{21}A_{11}^1 F_1) \bar{v}_{\bar{v}_i} has an accuracy of
\[
|\Delta u_{t-1} - \hat{\Delta} u_{t-1}| \leq |(F_2 - A_{21}A_{11}^1 F_1)^\top (A_{21}A_{11}^1 |E_1| + |E_2| |\bar{e}_2| |\bar{e}_2|)
\]
\]
The EOI can also be used for cyber-attack detection. Based on the bounds presented in Lemma 1, a threshold for EOI-based cyber-attack detection is introduced in [14], globally bounding the healthy EOI behaviour. In the threshold, each element \( i \) of \( \bar{v}_{\bar{v}_i} \) is defined as
\[
\bar{v}_{\bar{v}_i} = e^{-k_{i}^{t_{\max}}} \hat{U}_{i} + (1 - e^{-k_{i}^{t_{\max}}}) m,
\]
where \( k = K_{i,\bar{v}_i}, |M_{i,\bar{v}_i}|, \quad \bar{v}_{\bar{v}_i} = \lim_{t \to \infty} \frac{\bar{e}_2(t_{2sN})}{\epsilon^+} \), and \( \hat{U} = \lim_{t \to \infty} |A_{12}| \bar{e}_2 + |A_{22}| \bar{e}_2 + |E_2| \bar{e}_2 + |A_{22}| \bar{e}_2 \). A lower threshold
\( \nu_{fil} = -\bar{\nu}_{fil} \) is derived similarly. A cyber-attack is detected if
the condition \( \nu_{fil} \leq \nu_{fil} \leq \bar{\nu}_{fil} \) is violated.

V. SIMULATION OF INTERSECTION CONTROL

A simulation is performed with 2 cars approaching an intersection. The car closest to the intersection will be referred to
as the leader car, for which the input sequence is pre-defined.

The cyber-attack is illustrated in figure 5. This estimate of the attack
slower than the detection with the novel detection method. However,
the lead vehicle drives at a constant
speed, and at \( t = 0.5 \) [s] a step attack is performed on the
communication. In figure 3 it is shown that this attack causes
the lead vehicle to crash. Figure 4 shows the detection results obtained with the novel detection
logic presented in Section IV. Detection first occurs for a
very short period at \( t = 0.64 \) [s]. This is not visible in the
figure. More consistent detection occurs at \( t = 0.82 \) [s]. This
consistent detection occurs well before the crash occurs at \( t = 4.8 \) [s].

In Figure 6, it is shown that the EOI based detection method from previous work also detects the attack.
However, the detection only occurs at \( t = 2.73 \) [s]. This is significantly slower than the detection with the novel detection method.

The capability of the EOI based method to also estimate
the attack is illustrated in figure 5. This estimate of the attack
can be very useful in designing an effective control strategy
to deal with the attack.

As both detection methods depend on the same observer,
it is feasible to implement both methods concurrently. In this way
the best properties of both methods can be combined.

VI. CONCLUSIONS

Intersections, in the way they are currently designed for
human drivers, are a major cause of accidents as well as traffic delays. Therefore, automated control of vehicles in intersections offers great potential for improvement. However,
intersection Control systems do rely on wireless
V2X networks for traffic coordination, security of such communication channel is paramount for safety. V2X networks
are a first line of preventive security measures are already
in place at the protocol level to make cyber-attacks more
difficult. Still the possibility of an inside attacker, or one
able of overcoming preventive security measures cannot
be ruled out. To protect against such a scenario, a second
line of defenses based on a cyber-attack detection method
is proposed in this paper. In particular, a novel detector
based on a Sliding Mode Observer and a corresponding
set of thresholds was designed. With respect to previous
results, the novel detection approach is shown to be faster
and more sensitive, as the filtering of the observer Equivalent
Output Injection is avoided. Theoretical results certifying the
robustness and detectability of the proposed approach were
provided, as well as a simulation study.

In the future, a two-stage approach based on a fault
detector and a fault identification scheme may be envisaged,
thus paving the way for fully autonomous accommodation
of faults and cyber-attacks in Intersection Control systems
based on cooperative autonomous vehicles. Furthermore,
adaptations to the detection method are envisioned for which
boundedness of the attack is no longer required.

**TABLE I**

| Parameter      | Value  | Parameter      | Value  |
|----------------|--------|----------------|--------|
| \( p_0(0) \)   | \(-40 \) [m] | \( v_0(0) \)   | \( 8 \) [m/s] |
| \( a_0(0) \)   | \( 10 \) [m/s] | \( p_1(0) \)   | \( -50 \) [m] |
| \( \tau_0 \)   | \( 0.11 \) [s] | \( \tau_1 \)   | \( 0.1 \) [s] |
| \( L_1 \)      | \( 4 \) [m] | \( r \)        | \( 0.9 \) |
| \( k_p \)      | \( 0.2 \) [s\(^{-1}\)] | \( k_d \)      | \( 0.7 \) [s\(^{-1}\)] |
| \( \Delta \)    | \( 10 \) [m/s] | \( \bar{\eta} \) | \( 1 \) [m/s] |
| \( \bar{\xi}_1 \) | \( 0.15 \) [s] | \( \lambda_1 \) | \( 0.03 \) |
| \( \lambda_2 \) | \( 0.15 \) [s] | \( A_{22} \)    | \( -0.1 \) [s\(^{-1}\)] |
| \( M \)        | \( \text{diag}((0.5, 11.5, 2.0, 2.0)) \) | \( A_{22} \) | \( 0 \) [m/s] |

Based on [25], IC is initiated when the cars are within
50 [m] from the intersection. Furthermore, the intersection is
approached at \( 8 \) [m/s] \( \approx 30 \) [km/h], which is a common
standard speed in urban areas. In figure 2, the input of
the leader car, and the cyber-attack considered are show.

It can be seen that the lead vehicle drives at a constant speed, and at \( t = 0.5 \) [s] a step attack is performed on the
communication. In figure 3 it is shown that this attack
causes the cars to gradually drive closer together, eventually leading
to a crash at the intersection. The crash occurs when
the follower vehicle enters the intersection, i.e. \( p_f = -4 \) [m].
At this point the lead vehicle has not yet left the intersection, i.e. \( -4 \) [m] < \( p_f \) < 0 [m], resulting in a crash. Figure 4
shows the detection results obtained with the novel detection
logic presented in Section IV. Detection first occurs for a
very short period at \( t = 0.64 \) [s]. This is not visible in the
figure. More consistent detection occurs at \( t = 0.82 \) [s]. This
consistent detection occurs well before the crash occurs at \( t = 4.8 \) [s].

Authorized licensed use limited to: TU Delft Library. Downloaded on January 06, 2022 at 12:23:21 UTC from IEEE Xplore. Restrictions apply.
Fig. 4. Error bounds and detection performance of novel detection logic.

Fig. 5. Attack estimation by EOI based estimation method.

APPENDIX

From the first row of equation (6), using lemma 1.1.1 in [26] we obtain $e_1(t) = e^{A_{11}(t)} e_1(0) - r_{\zeta}(\zeta(t)) - r_{\eta}(\eta(t)) - r_{A} (\Delta u_{t-1})$, where $r_{\zeta}(\zeta(t)) = \int_{0}^{t} A_{21} \zeta(s) e^{A_{11}(t-s)} ds$ and $r_{\eta}(\eta(t)) = \int_{0}^{t} E_{1} \eta(s) e^{A_{11}(t-s)} ds$.

From this it can be concluded that $e_1(t) = e^{A_{11}(t)} e_1(0) - r_{A} (\Delta u_{t-1})$, and therefore $\dot{e}_1(t) = e^{A_{11}(t)} e_1(t) - r_{A} (\Delta u_{t-1})$ and $e_1(t) = e^{A_{11}(t)} e_1(0) - r_{A} (\Delta u_{t-1})$, which proves statements 1) and 2) in Lemma 1.

Furthermore $\dot{e}_2$ from the second row of equation (6) is $\dot{e}_2 = A_{21} e_1 + A_{22} e_2 + A_{23} \bar{u}_t - E_{2} \eta - F_{2} \Delta u_{t-1} - M \text{sgn}(e_r)$. With this we can write,

$$\dot{e}_2 = e^{A_{22}} e_2 + A_{23} (e_1 - e_2) - F_{2} \Delta u_{t-1} = e^{A_{22}} e_2 + A_{23} r_{A} (\Delta u_{t-1}) - F_{2} \Delta u_{t-1}.$$  (14)

To prove statements 8) and 9) in Lemma 1 we finally observe that $\dot{e}_2 = e^{A_{22}} e_2 + \text{sgn}(e_r) r(\Delta u_{t-1})$ and $\dot{e}_2 = e^{A_{22}} e_2 + \text{sgn}(e_r) r(\Delta u_{t-1})$.

REFERENCES

[1] European Commission, “Traffic safety basic facts on junctions,” 2016.
[2] K. Dresner and P. Stone, “Multagent traffic management: An improved intersection control mechanism,” in AAMAS ’05. Proc. of 2005.
[3] T. Arts, M. Dorigatti, and S. Tonetta, “Making Implicit Safety Requirements Explicit: An AUTOSAR Safety Case,” in Computer Safety, Reliability, and Security, A. Bondavalli and F. D. Giudomenico, Eds. Cham: Springer International Publishing, 2014, vol. 8666, pp. 81–92.
[4] M. Amoosoade, A. Raghuramu, C.-N. Chuah, D. Ghosal, H. M. Zhang, J. Rowe, and K. Levitt, “Security vulnerabilities of connected vehicle systems and their impact on cooperative driving,” IEEE Communications Magazine, vol. 53, no. 6, pp. 126–132, 2015.
[5] C. Yan, W. Xu, and J. Liu, “Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle,” DEF CON, vol. 24, no. 8, p. 109, 2016.
[6] R. G. Engoulou, M. Bllaiche, S. Pierre, and A. Quintero, “Vehnet security surveys,” Computer Communications, vol. 44, pp. 1–13, 2014.
[7] R. van der Heijden, T. Lukaseder, and F. Kargl, “Analyzing attacks on cooperative adaptive cruise control (cacc),” in 2017 IEEE Vehicular Networking Conference (VNC). IEEE, 2017, pp. 48–52.
[8] N. Bilbeyner, S. Mauthofer, K. M. Bayarou, and F. Kargl, “Assessment of node trustworthiness in vanets using data plausibility checks with particle filters,” in VNC. Proc. of, 2012.
[9] S. Dietzel, J. Petit, G. Heijenk, and F. Kargl, “Graph-based metrics for insider attack detection in vanet multihop data dissemination protocols,” IEEE Trans. on Vehic. Tech., vol. 62, no. 4, pp. 1505–1518, 2012.
[10] G. Yan, S. Olariu, and M. C. Weigle, “Providing vanet security through active position detection,” Computer communications, vol. 31, no. 12, pp. 2883–2897, 2008.
[11] T. Arts, M. Dorigatti, and S. Tonetta, “Making Implicit Safety Requirements Explicit: An AUTOSAR Safety Case,” in Computer Safety, Reliability, and Security, A. Bondavalli and F. D. Giudomenico, Eds. Cham: Springer International Publishing, 2014, vol. 8666, pp. 81–92.
[12] M. Amoosoade, A. Raghuramu, C.-N. Chuah, D. Ghosal, H. M. Zhang, J. Rowe, and K. Levitt, “Security vulnerabilities of connected vehicle systems and their impact on cooperative driving,” IEEE Communications Magazine, vol. 53, no. 6, pp. 126–132, 2015.
[13] C. Yan, W. Xu, and J. Liu, “Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle,” DEF CON, vol. 24, no. 8, p. 109, 2016.
[14] R. G. Engoulou, M. Bllaiche, S. Pierre, and A. Quintero, “Vehnet security surveys,” Computer Communications, vol. 44, pp. 1–13, 2014.
[15] R. van der Heijden, T. Lukaseder, and F. Kargl, “Analyzing attacks on cooperative adaptive cruise control (cacc),” in 2017 IEEE Vehicular Networking Conference (VNC). IEEE, 2017, pp. 48–52.
[16] N. Bilbeyner, S. Mauthofer, K. M. Bayarou, and F. Kargl, “Assessment of node trustworthiness in vanets using data plausibility checks with particle filters,” in VNC. Proc. of, 2012.
[17] S. Dietzel, J. Petit, G. Heijenk, and F. Kargl, “Graph-based metrics for insider attack detection in vanet multihop data dissemination protocols,” IEEE Trans. on Vehic. Tech., vol. 62, no. 4, pp. 1505–1518, 2012.
[18] G. Yan, S. Olariu, and M. C. Weigle, “Providing vanet security through active position detection,” Computer communications, vol. 31, no. 12, pp. 2883–2897, 2008.
[19] T. Arts, M. Dorigatti, and S. Tonetta, “Making Implicit Safety Requirements Explicit: An AUTOSAR Safety Case,” in Computer Safety, Reliability, and Security, A. Bondavalli and F. D. Giudomenico, Eds. Cham: Springer International Publishing, 2014, vol. 8666, pp. 81–92.
[20] M. Amoosoade, A. Raghuramu, C.-N. Chuah, D. Ghosal, H. M. Zhang, J. Rowe, and K. Levitt, “Security vulnerabilities of connected vehicle systems and their impact on cooperative driving,” IEEE Communications Magazine, vol. 53, no. 6, pp. 126–132, 2015.
[21] C. Yan, W. Xu, and J. Liu, “Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle,” DEF CON, vol. 24, no. 8, p. 109, 2016.
[22] R. G. Engoulou, M. Bllaiche, S. Pierre, and A. Quintero, “Vehnet security surveys,” Computer Communications, vol. 44, pp. 1–13, 2014.