Contents lists available at ScienceDirect

European Journal of Control

journal homepage: www.elsevier.com/locate/ejcon

Resilient average consensus on general directed graphs in presence of cyber-attacks[☆]

Mahdieh S. Sadabadi[a,*], Azwirman Gusrialdi[b]

[a] School of Electronic Engineering and Computer Science, Queen Mary University of London, London, United Kingdom
[b] Faculty of Engineering and Natural Sciences, Tampere University, Tampere 33014, Finland

ARTICLE INFO

Article history:
Received 8 April 2022
Accepted 5 June 2022
Available online xxx

Recommended by Prof. T Parisini

Keywords:
Average consensus
Cooperative control
Resilient control
False data injection (FDI) cyber-attacks

ABSTRACT

This paper proposes a resilient distributed control scheme that ensures average consensus in multi-agent systems with continuous-time single-integrator kinematics in the presence of cyber-attacks. Potential cyber-attacks considered on such systems are in the form of unknown uniformly bounded false data injection (FDI) to control input channels (actuators) and also eavesdropping attacks. The purpose of such cyber-attacks is to disturb the average consensus in multi-agent systems and also to disclose the states of agents. The proposed resilient distributed average consensus protocol includes a set of virtual variables/states being exchanged via a communication network given by a general directed graph (digraph) and is designed to make the closed-loop system stable and preserve the average consensus regardless of the existence of cyber-attacks. Unlike the existing literature, the proposed distributed average consensus framework does not require any conditions on directed graphs to be strongly connected, balanced, or symmetric. A graph-theoretical approach, Lyapunov direct method, and LaSalle's invariance principle are used to guarantee the rigorous stability of multi-agent systems augmented with the proposed distributed algorithm. Simulation results validate the theoretical contributions of this paper.

1. Introduction

1.1. Motivation and literature review

Distributed multi-agent systems are composed of a set of agents that can exchange information via distributed communication represented by a directed graph (digraph) or an undirected graph. One of the important problems in multi-agent systems is distributed average consensus in which agents aim to reach an agreement on the average value of their initial states via communication and information exchanges amongst their neighbors [15].

The standard average consensus algorithm in [12] guarantees average convergence as long as the corresponding directed communication topology is bidirectional and connected or both strongly connected and balanced. For the last decade, there has been a number of work that aim to develop distributed average consensus algorithms on general (unbalanced) directed graphs. For example, the work in [1,7] proposes a distributed average consensus algorithm by introducing additional variables to each agent (called "surplus") for both continuous and discrete-time versions on an arbitrary strongly connected directed graph that is not necessarily balanced. A discrete-time distributed strategy is proposed in [2] that ensures finite-time average consensus based on interconnection topologies described by strongly connected digraphs. A continuous-time fixed-time average consensus over a strongly connected digraph is proposed in [11]. Finally, the authors in [9] developed a continuous-time and sampled-data-based average consensus method based on logarithmic quantizers where digraphs are assumed to be balanced and contain a spanning tree.

While the introduction of ICT (information and communication technology) facilitates the implementation of distributed algorithms for achieving global objectives, the use of open ICT such as wireless communication makes cooperative systems vulnerable to cyber intrusions/attacks. Furthermore, due to the tight coupling between the cyber (i.e., ICT) and the cooperative (physical) systems, cyber-attacks can destabilize and damage the cooperative system as witnessed from the coordinated attack on the Ukraine power grid in 2015 [13]. As cyber-attacks cannot be foreseen in advance, it is of importance to design distributed/cooperative control algorithms capable of maintaining the overall system's perfor-

[☆] The work by A. Gusrialdi is supported by Academy of Finland under academy project decision no. 330073.

[*] Corresponding author.:
E-mail address: m.sadabadi@qmul.ac.uk (M.S. Sadabadi).

https://doi.org/10.1016/j.ejcon.2022.100669
0947-3580/© 2022 The Authors. Published by Elsevier Ltd on behalf of European Control Association. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Please cite this article as: M.S. Sadabadi and A. Gusrialdi, Resilient average consensus on general directed graphs in presence of cyber-attacks, European Journal of Control, https://doi.org/10.1016/j.ejcon.2022.100669

mance under unknown attacks. This type of cooperative control is also known as resilient cooperative control.

There has been a number of existing approaches that address resilient cooperative control for cyber-attacks on communication and/or sensors and/or actuators for either leaderless consensus and leader-following consensus problems, see for example [4,6,8,10,16,17,21–24]. However, to the best of our knowledge, there are a few existing methods addressing the design of resilient average consensus algorithms on general directed graphs in the presence of cyber-attacks. Recent work on this topic is presented in [19] where the authors proposed a homomorphic encryption-based resilient average consensus algorithm on strongly connected digraphs in the presence of attacks on the communication network. However, the proposed method has restrictions on the number of cyber-attacks in the network and the communication network topology. In addition, the graph is assumed to be strongly connected and the network topology needs to satisfy the so-called r-robustness property that depends on the upper-bound of the number of attacks in each agent's neighbors. In [18], a privacy-preserving strategy has been proposed for the problem of average consensus in multi-agent systems where the initial states of agents are hidden in random values. However, the underlying communication graph in this approach is assumed to be strongly connected; moreover, the case of cyber-attacks on actuators has not been considered in either [19] or [18].

1.2. Statement of contributions

In this paper, we first address the problem of designing a resilient average consensus algorithm on general directed graphs in presence of cyber-attacks on actuators (control input channels) and eavesdropping attacks that are privacy threats disclosing the local information of agents. In contrast to [20], we do not impose any restrictions on the number of attacks in the network. Similar to [1,5,16], our proposed distributed resilient average consensus strategy is based on some auxiliary variables, called virtual states. Instead of exchanging the physical states (i.e., the states on which we are interested in achieving average consensus), the agents exchange the virtual states via a communication network. By means of a graph theoretical approach, Lyapunov direct method, and LaSalle's invariance principle, the stability of the proposed algorithm is analyzed. We show that by the proper design of the proposed distributed control algorithm and the initialization of some of the virtual variables, the convergence to a small neighborhood of the average consensus value is guaranteed provided that digraphs contain a rooted-out tree. In addition, we show that in the absence of cyber-attacks, all the agents' states under the proposed cooperative control algorithm converge to the exact average consensus value. In contrast to the existing literature on average consensus on general directed graphs, our proposed distributed protocol does not require the directed graphs to be strongly connected, balanced, or symmetric. Furthermore, as the proposed distributed average consensus in this paper does not require exchanging agents' true (physical states) amongst their neighbors, it facilitates the privacy-preserving of agents' states and avoids potential privacy threats such as eavesdropping attacks. In summary, the contributions of this paper are threefold:

1. The proposed distributed control approach in this paper does not impose any restrictions on the connectivity of underlying communication digraphs. The only condition it requires is that the digraphs contain a rooted-out tree.
2. We demonstrate that by a proper design of the control parameters, the proposed distributed control approach is resilient against false data injection attacks on actuators.
3. The proposed distributed control approach enhances privacy-preserving feature in the average consensus problem as it relies only on exchanging virtual (auxiliary) states amongst agents.

1.3. Organization and notation

Paper Organization. The paper is organized as follows. Section 2 formulates distributed average consensus problem in the presence of actuator attacks. Section 3 presents a distributed average consensus strategy that rigorously guarantees the asymptotic stability and average consensus in first-order multi-agent systems. The attack-resilient feature of the proposed distributed average consensus algorithm is analyzed in Section 4. Section 5 provides numerical results. The concluding remarks and future works are given in Section 6.

Notation. The notation used in this paper is standard. In particular, $\mathbf{1}_n$, $\mathbf{0}_n$, $\mathbf{I}_n$, and $\mathbf{0}_{n\times m}$ are an $n \times 1$ vector of ones, an $n \times 1$ zero vector, an $n \times n$ Identity matrix, and a zero matrix of dimension $n \times m$, respectively.

2. Problem statement

Consider a first-order multi-agent system composed of n agents where each agent is labeled by a natural number i, $i = 1, \dots, n$. The dynamics of each agent are presented as follows:

$$\begin{aligned} \dot{x}_i(t) &= u_i(t), \\ x_i(0) &= x_{i,0}, \end{aligned} \tag{1}$$

where $x_i(t) \in \mathbb{R}$ is the state, $u_i(t) \in \mathbb{R}$ is the control input, and $x_{i,0} \in \mathbb{R}$ is the initial state of agent i. As demonstrated in [14], in a heterogeneous nonlinear/linear multi-agent system if the individual dynamic systems are input passivity short, then their dynamic behavior and control design at the network level can equivalently be investigated in terms of the first-order dynamics in (1). Therefore, in the remainder of this paper, the simpler yet equivalent dynamics in (1) are considered.

The information flow amongst agents is modeled by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The node set $\mathcal{V} = \{1, \dots, n\}$ and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represent agents and information exchange links, respectively. An edge $(i, j) \in \mathcal{E}$ indicates that node j can receive information from node i. In this case, node i is an in-neighbor of node j. The set of the in-neighbors of node i is denoted by $\mathcal{N}_i^{\text{in}}$. The following assumption is made on the communication digraphs:

Assumption 1. The digraphs contain a rooted-out tree.

Assumption 1 means that there exists a (root) node from which all other nodes can be reached (direct/indirectly) by following a directed path.

The first objective in this paper is to design a distributed average consensus algorithm for first-order multi-agent systems whose dynamics are given in (1) under a directed graph satisfying Assumption 1 so that all agents reach an average consensus. This means that the state of all agents will converge to

$$\lim_{t\to\infty} x_i(t) = x_{ave}, \quad \forall i \in \mathcal{V}. \tag{2}$$

where $x_{ave} = \frac{1}{n}\sum_{j=1}^{n} x_{j,0}$.

Next, it is assumed that an adversary may launch false data injection (FDI) attack on the input channel (actuator) of the individual agent, i.e., [10]

$$\tilde{u}_i(t) = u_i(t) + \delta_{u_i}(t), \tag{3}$$

where $\tilde{u}_i(t)$ is the corrupted control input sent to agent i and $\delta_{u_i}(t)$ represents the false data injection that are unknown to agents. Furthermore, we make the following assumption on the injection

$\delta_{u_i}(t)$. Moreover, it is assumed that potential eavesdropping attacks might have access to the exchanged physical states of agents ($x_i(t)$).

Assumption 2. Injection $\delta_{u_i}(t)$ in (3) is uniformly bounded and does not depend on agents' states $x_i(t)$, $i = 1, \dots, n$.

The assumption on uniformly bounded injection has also been considered in related works, e.g., [4,6,17], and is reasonable in practice due to the following two reasons: (i) From the attackers perspective, an intelligent attacker would aim at destabilizing the system with a limited change to avoid any detection and (ii) from the defenders perspective, an injection with an unbounded magnitude can easily be rejected by a threshold check [6]. Other than Assumption 2, no other assumptions are imposed for the attacker.

Under the existence of these actuator-attacks, the control input $u_i(t)$ in (1) is then replaced by $\tilde{u}_i(t)$. The second objective in this paper is to design a control law $u_i(t)$ in (3) for achieving approximate average consensus given by

$$|\lim_{t\to\infty} x_i(t) - x_{ave}| \le \epsilon, \ \ \forall i \in \mathcal{V}. \tag{4}$$

where ϵ is a sufficiently small non-negative scalar.

3. Distributed average consensus algorithm in the absence of actuator-attacks

3.1. Proposed distributed average consensus strategy

First, consider the case when there are no actuator attacks, i.e., $\delta_{u_i}(t) = 0$ for all agents. To achieve the average consensus objective in (2), in addition to the physical state $x_i(t)$, each agent i maintains auxiliary (virtual) variables $v_i(t) \in \mathbb{R}$, $\theta_i(t) \in \mathbb{R}$, and $z_i(t) \in \mathbb{R}$. We propose the following control input of node i in (1):

$$u_i(t) = kx_i(t) + k_P(v_i(t) - \kappa x_i(t)) + k_I z_i(t) \tag{5}$$

where $k \neq 0$, $k_P \neq 0$, $k_I \neq 0$, and auxiliary states $(v_i(t), z_i(t))$ are updated according to:

$$\dot{v}_i(t) = -\alpha(v_i(t) - \kappa x_i(t)) - \sum_{j=1}^{n} \gamma_{j,i}\big(\theta_i(t) - \theta_j(t)\big) \tag{6a}$$

$$\dot{\theta}_i(t) = -\eta\theta_i(t) + \sum_{j=1}^{n} \gamma_{i,j}\big(v_i(t) - v_j(t)\big) \tag{6b}$$

$$\dot{z}_i(t) = \alpha(v_i(t) - \kappa x_i(t)) \tag{6c}$$

with α, κ, and η are positive scalars. In addition, the initial value of $z_i(t)$ in (6c) is set to be equal to

$$z_i(0) = -v_i(0) + x_{i,0}\left(\kappa - \frac{k}{k_I}\right) \tag{7}$$

while $(v_i(0), \theta_i(0))$ are set to arbitrary values.

According to (6c), the virtual state $v_i(t)$ asymptotically tracks $\kappa x_i(t)$; hence, instead of reaching consensus on $x_i(t)$, the consensus can be reached on $\kappa^{-1}v_i(t)$, $i \in \mathcal{V}$. That is the main reason why $v_i(t)$ is exchanged amongst neighboring agents in (6b). Furthermore, as it will be discussed in Theorem 1, the dynamics of $\dot{v}_i(t)$ in (6a) enhance resilience against false data injection $\delta_{u_i}(t)$ in (3) while achieving approximate average consensus in (4) provided that the value of k_I in (5) is chosen to be sufficiently large. Finally, we present a set of stabilizing controller in Proposition 1 based on the choice of k, k_P, and k_I.

Information exchange between the nodes in (6) is illustrated in Fig. 1. As one can observe from (6) and Fig. 1, instead of the physical states $x_i(t)$, each node exchanges the auxiliary variables $v_i(t)$ and $\theta_i(t)$ via the communication network. In contrast to the physical states $x_i(t)$, the auxiliary states $v_i(t), z_i(t)$, and $\theta_i(t)$ do not have any physical meaning and thus are also called as virtual states. Note that according to (6) and Fig. 1, virtual states are being sent on different directions based on two different directed graphs (communication network topologies), one associated with a graph Laplacian matrix $\mathbb{L}$ and the other one associated with $\mathbb{L}^T$. It is assumed that the digraph corresponding to each virtual states being exchanged contains a rooted-out tree. The physical states $x_j(t)$ are not exchanged on any directions as this will expose the physical state (private information) to the adversary. On the other hand, exchanging different virtual states on each direction (and updating them according to the proposed algorithm in (6)) not only achieve average consensus (refer to Theorem 1 in Section 4.1) but also protect the physical state from eavesdropping attacks (see the discussion in Section 4.2).

Let $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]^T$, $\mathbf{v}(t) = [v_1(t), \dots, v_n(t)]^T$, $\theta(t) = [\theta_1(t), \dots, \theta_n(t)]^T$, and $\mathbf{z}(t) = [z_1(t), \dots, z_n(t)]^T$, we can then write the update rule (6) for all nodes $i \in \mathcal{V}$ in a compact form as

$$\begin{aligned} \dot{\mathbf{v}}(t) &= -\alpha(\mathbf{v}(t) - \kappa\mathbf{x}(t)) - \mathbb{L}^T\theta(t) \\ \dot{\theta}(t) &= -\eta\theta(t) + \mathbb{L}\mathbf{v}(t) \\ \dot{\mathbf{z}}(t) &= \alpha(\mathbf{v}(t) - \kappa\mathbf{x}(t)) \end{aligned} \tag{8}$$

where $\mathbb{L}$ is the Laplacian matrix corresponding to the directed graph $\mathcal{G}$. The closed-loop system for the open-loop dynamics (1) under (5) and (8) can then be written as follows:

$$\dot{\mathbf{v}}(t) = -\alpha(\mathbf{v}(t) - \kappa\mathbf{x}(t)) - \mathbb{L}^T\theta(t) \tag{9a}$$

$$\dot{\theta}(t) = -\eta\theta(t) + \mathbb{L}\mathbf{v}(t) \tag{9b}$$

$$\dot{\mathbf{z}}(t) = \alpha(\mathbf{v}(t) - \kappa\mathbf{x}(t)) \tag{9c}$$

$$\dot{\mathbf{x}}(t) = k\mathbf{x}(t) + k_P(\mathbf{v}(t) - \kappa\mathbf{x}(t)) + k_I\mathbf{z}(t) \tag{9d}$$

Remark 1. Note that even though (9) looks similar to the proposed algorithm in [16], the virtual network in [16] is designed to achieve resilient leader-follower consensus. However, in this paper, we are interested in average consensus and the extension from [16] to the average consensus problem is not trivial. Moreover, unlike [16], the proposed consensus algorithm in this paper requires a special initialization of some of the auxiliary variables in (7) in order to achieve the average consensus objective given in (2).

3.2. Analysis of equilibria

The following lemma analyzes the existence and the uniqueness of the equilibria of the closed-loop dynamics in (9).

Lemma 1. *Consider the dynamical system (9) whose initial values $z_i(0)$ are chosen as in (7) and the directed graph associated with the Laplacian matrix $\mathbb{L}$ satisfies Assumption 1. There exists a unique equilibrium $(\bar{\mathbf{x}}, \bar{\mathbf{v}}, \bar{\theta}, \bar{\mathbf{z}})$ for the closed-loop system in (9) that is presented as follows:*

$$\begin{aligned} \bar{\mathbf{x}} &= \mathbf{1}_n x_{ave} \\ \bar{\mathbf{v}} &= \kappa\mathbf{1}_n x_{ave} \\ \bar{\theta} &= \mathbf{0}_n \\ \bar{\mathbf{z}} &= -\frac{k}{k_I}\mathbf{1}_n x_{ave}. \end{aligned} \tag{10}$$

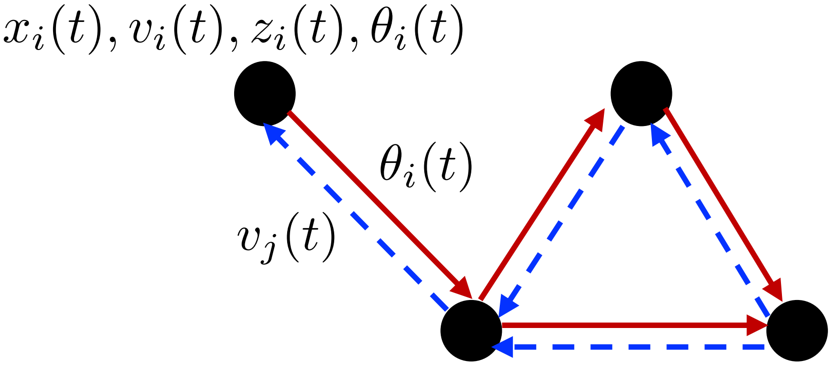

Fig. 1. Information exchange in the proposed average consensus algorithm.

Proof. Consider the distributed system in (9). The equilibria $(\bar{\mathbf{x}}, \bar{\mathbf{v}}, \bar{\theta}, \bar{\mathbf{z}})$ of (9) can be found by solving the following algebraic equations:

$$\mathbf{0}_n = -\alpha(\bar{\mathbf{v}} - \kappa\bar{\mathbf{x}}) - \mathbb{L}^T\bar{\theta} \tag{11a}$$

$$\mathbf{0}_n = -\eta\bar{\theta} + \mathbb{L}\bar{\mathbf{v}} \tag{11b}$$

$$\mathbf{0}_n = \alpha(\bar{\mathbf{v}} - \kappa\bar{\mathbf{x}}) \tag{11c}$$

$$\mathbf{0}_n = k\bar{\mathbf{x}} + k_P(\bar{\mathbf{v}} - \kappa\bar{\mathbf{x}}) + k_I\bar{\mathbf{z}}. \tag{11d}$$

From (11b) and (11c), one obtains that $\bar{\theta} = \frac{1}{\eta}\mathbb{L}\bar{\mathbf{v}}$ and $\bar{\mathbf{x}} = \kappa^{-1}\bar{\mathbf{v}}$. By replacing $\bar{\mathbf{x}} = \kappa^{-1}\bar{\mathbf{v}}$ and $\bar{\theta} = \frac{1}{\eta}\mathbb{L}\bar{\mathbf{v}}$ in (11a), one obtains that

$$-\tfrac{1}{\eta}\mathbb{L}^T\mathbb{L}\bar{\mathbf{v}} = \mathbf{0}_n. \tag{12}$$

Invoking the properties of the Laplacian matrix $\mathbb{L}$ as $\mathbb{L}\mathbf{1}_n = \mathbf{0}_n$ (see Assumption 1), from (12), it follows that $\bar{\mathbf{v}} = \mathbf{1}_n v^*$, where v^* is a scalar. Hence, $\kappa\bar{\mathbf{x}} = \bar{\mathbf{v}} = \mathbf{1}_n v^*$ and $\bar{\theta} = \mathbf{0}_n$. By replacing $\bar{\mathbf{x}}$ and $\bar{\mathbf{v}}$ in (11d), it follows that $\bar{\mathbf{z}} = -\frac{k}{k_I}\bar{\mathbf{x}} = -\frac{k}{k_I}\kappa^{-1}\mathbf{1}_n v^*$. Note that $k_I \neq 0$ and $\kappa \neq 0$.

In the next step, we will show that $v^* = \kappa x_{ave}$. To this end, we consider the closed-loop dynamics in (9). By replacing $\alpha(\mathbf{v}(t) - \kappa\mathbf{x}(t))$ with $\dot{\mathbf{z}}(t)$ in (9a), one obtains that:

$$\dot{\mathbf{v}}(t) \quad = -\dot{\mathbf{z}}(t) - \mathbb{L}^T\theta(t) \tag{13}$$

By left-multiplying of the above equation with $\mathbf{1}_n^T$, we have:

$$\mathbf{1}_n^T\dot{\mathbf{v}}(t) = -\mathbf{1}_n^T\dot{\mathbf{z}}(t) \tag{14}$$

Hence,

$$\mathbf{1}_n^T\mathbf{v}(t) = -\mathbf{1}_n^T\mathbf{z}(t) + \beta, \ \forall t \geq 0 \tag{15}$$

where β is a scalar. Since the above equation must hold for $t = 0$ (initial conditions) and $t \to \infty$ (steady-state conditions), the following equations are obtained:

$$\mathbf{1}_n^T\mathbf{v}(0) = -\mathbf{1}_n^T\mathbf{z}(0) + \beta \tag{16a}$$

$$\mathbf{1}_n^T\bar{\mathbf{v}} = -\mathbf{1}_n^T\bar{\mathbf{z}} + \beta \tag{16b}$$

By replacing $\bar{\mathbf{v}} = \mathbf{1}_n v^*$ and $\bar{\mathbf{z}} = -\frac{k}{k_I}\kappa^{-1}\mathbf{1}_n v^*$ in (16b), β is obtained as follows:

$$\beta = \mathbf{1}_n^T\bar{\mathbf{v}} + \mathbf{1}_n^T\bar{\mathbf{z}} = n(1 - \frac{k}{k_I}\kappa^{-1})v^* \tag{17}$$

Replacing β in (16a) yields that

$$\mathbf{1_n^T}\mathbf{v}(0) + \mathbf{1}_n^T\mathbf{z}(0) = n(1 - \tfrac{k}{k_I}\kappa^{-1})v^*. \tag{18}$$

As $\mathbf{z}(0) = -\mathbf{v}(0) + \mathbf{x}(0)(\kappa - \frac{k}{k_I})$ (see (7)), v^* is obtained as follows:

$$v^* = \kappa\,\tfrac{\mathbf{1}_n^T\mathbf{x}(0)}{n} = \kappa x_{ave}. \tag{19}$$

As a result, $\bar{\mathbf{x}} = \kappa^{-1}\mathbf{1}_n v^* = \mathbf{1}_n x_{ave}$. This completes the proof. □

3.3. Theoretical analysis of satability and convergence

Let $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) - \bar{\mathbf{x}}$, $\tilde{\mathbf{v}}(t) = \mathbf{v}(t) - \bar{\mathbf{v}}$, $\tilde{\theta}(t) = \theta(t) - \bar{\theta}$, and $\tilde{\mathbf{z}}(t) = \mathbf{z}(t) - \bar{\mathbf{z}}$. Then, the dynamics of the shifted closed-loop system are obtained as follows:

$$\begin{aligned}
\dot{\tilde{\mathbf{v}}}(t) &= -\alpha(\tilde{\mathbf{v}}(t) - \kappa\tilde{\mathbf{x}}(t)) - \mathbb{L}^T\tilde{\theta}(t) \\
\dot{\tilde{\theta}}(t) &= -\eta\tilde{\theta}(t) + \mathbb{L}\tilde{\mathbf{v}}(t) \\
\dot{\tilde{\mathbf{z}}}(t) &= \alpha(\tilde{\mathbf{v}}(t) - \kappa\tilde{\mathbf{x}}(t)) \\
\dot{\tilde{\mathbf{x}}}(t) &= k\tilde{\mathbf{x}}(t) + k_P(\tilde{\mathbf{v}}(t) - \kappa\tilde{\mathbf{x}}(t)) + k_I\tilde{\mathbf{z}}(t).
\end{aligned} \tag{20}$$

Remark 2. From (15), one can obtain that $\mathbf{1}_n^T\tilde{\mathbf{v}}(t) = -\mathbf{1}_n^T\tilde{\mathbf{z}}(t)$.

The stability and convergence of (9) is analyzed in the following proposition.

Proposition 1. *Consider the dynamical system* (9) *whose initial values $z_i(0)$ are chosen as in* (7) *and the directed graph associated with the Laplacian matrix $\mathbb{L}$ satisfies Assumption* 1. *If $\alpha > 0$, $\eta > 0$, $\kappa > 0$, $k < 0$, $k_P > 0$, and $0 < k_I < -kk_P$, the following statements hold:*

1. *The origin in* (20) *is globally asymptotically stable.*
2. *The average consensus in* (2) *is ensured.*

Proof. We choose $k < 0$, $k_P > 0$, and $0 < k_I < -kk_P$ and compute following positive scalars:

$$\rho \quad = \tfrac{k}{kk_P + k_I}, \quad \nu = -\tfrac{k_I}{k}. \tag{21}$$

We then choose the following Lyapunov function:

$$\mathcal{V} \quad = \tfrac{1}{2}\mathbf{x_{cl}}^T(t)\mathcal{P}\mathbf{x_{cl}}(t), \tag{22}$$

where $\mathbf{x_{cl}}(t) = \left[\tilde{\mathbf{v}}^T(t), \tilde{\theta}^T(t), \tilde{\mathbf{x}}^T(t), \tilde{\mathbf{z}}^T(t)\right]^T$ and

$$\mathcal{P} = \begin{bmatrix} \mathbf{I_n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{I}_n & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \kappa\rho\mathbf{I}_n & -\kappa\rho\nu\mathbf{I}_n \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & -\kappa\rho\nu\mathbf{I}_n & \kappa\nu(1+\rho\nu)\mathbf{I}_n \end{bmatrix} \tag{23}$$

It can be shown that $\mathcal{P}$ is a positive definite matrix. The time derivative of $\mathcal{V}$ in (22) along the closed-loop trajectories (20) is obtained as follows:

$$\dot{\mathcal{V}} = \tfrac{1}{2}\mathbf{x_{cl}}^T(t)(\mathbf{A_{cl}}^T\mathcal{P} + \mathcal{P}\mathbf{A_{cl}})\mathbf{x_{cl}}(t), \tag{24}$$

where

$$\mathbf{A_{cl}} = \begin{bmatrix} -\alpha\mathbf{I}_n & -\mathbb{L}^T & \alpha\kappa\mathbf{I}_n & \mathbf{0}_{n\times n} \\ \mathbb{L} & -\eta\mathbf{I}_n & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ k_p\mathbf{I}_n & \mathbf{0}_{n\times n} & (k-k_P\kappa)\mathbf{I}_n & k_I\mathbf{I}_n \\ \alpha\mathbf{I}_n & \mathbf{0}_{n\times n} & -\alpha\kappa\mathbf{I}_n & \mathbf{0}_{n\times n} \end{bmatrix}. \tag{25}$$

By direct calculations and taking into account (21), $\dot{\mathcal{V}}$ in (24) is simplified as follows:

$$\dot{\mathcal{V}} = \sum_{i=1}^{n} \kappa[\tilde{x}_i(t)\ \tilde{z}_i(t)]^T Q[\tilde{x}_i(t)\ \tilde{z}_i(t)] - \eta\tilde{\theta}^T(t)\tilde{\theta}(t) - \alpha(\tilde{\mathbf{v}}(t) - \kappa\tilde{\mathbf{x}}(t))^T(\tilde{\mathbf{v}}(t) - \kappa\tilde{\mathbf{x}}(t)) \tag{26}$$

where

$$Q = \rho\begin{bmatrix} k & -k\nu \\ -k\nu & \nu^2 k \end{bmatrix}. \tag{27}$$

Since $Q \in \mathbb{R}^{2\times 2}$, $Q = Q^T$, $trace(Q) = \rho k(1+\nu^2) < 0$, and $det(Q) = 0$, $Q \preceq 0$. Consequently, $\dot{\mathcal{V}} \leq 0$. To show the globally asymptotic stability of the origin in (20), we should illustrate that the only solution of (20) that satisfies $\dot{\mathcal{V}} = 0$ is the origin (LaSalles invariance principle). Note that $\dot{\mathcal{V}} = 0$ implies that $\tilde{\mathbf{v}} = \kappa\tilde{\mathbf{x}}$, $\tilde{\theta} = \mathbf{0}_n$, and $[\tilde{x}_i(t)\ \tilde{z}_i(t)]^T \in ker(Q)$, $i \in \mathcal{V}(\mathcal{G})$. The null-space of Q is characterized as $\tilde{x}_i(t) = \nu\tilde{z}_i(t)$. As $\tilde{\theta} = \mathbf{0}_n$, the closed-loop trajectories in (20) imply $\mathbb{L}\tilde{\mathbf{v}} = \mathbf{0}_n$. Therefore, $\tilde{\mathbf{v}} = \kappa\tilde{\mathbf{x}} = \mathbf{1}_n\hat{\nu}$ where $\hat{\nu}$ is a constant, and $\tilde{\mathbf{z}} = -\frac{k}{k_I}\kappa^{-1}\mathbf{1}_n\hat{\nu}$. From Remark 2, we have $\mathbf{1}_n^T\tilde{\mathbf{v}} = -\mathbf{1}_n^T\tilde{\mathbf{z}}$. Hence, $\hat{\nu} = \frac{k}{k_I}\kappa^{-1}\hat{\nu}$. This implies that $\hat{\nu} = 0$ (note that $\frac{k}{k_I}\kappa^{-1} < 0$); hence, $\tilde{\mathbf{v}} = \tilde{\mathbf{x}} = \tilde{\mathbf{z}} = \mathbf{0}_n$ Thus, the origin is the only solution that satisfies $\dot{\mathcal{V}} = 0$. Therefore, by invoking LaSalles invariance principle, we can conclude that the origin in (20) is globally asymptotically stable.

Due to the globally asymptotic stability of the origin in (20), the closed-loop states of (9) converge to their equilibria given in (10). As a result, $\mathbf{x}(t)$ asymptotically converges to $\bar{\mathbf{x}} = \mathbf{1}_n x_{ave}$. This completes the proof of Proposition 1. □

4. Resilient distributed average consensus while under cyber-attacks

In this section, the resilience of the proposed distributed average consensus system in (9) is analyzed with respect to potential FDI attacks on actuators, as modeled in (3) and eavesdropping attacks.

4.1. Resilience to FDI cyber-attacks on actuators

The resilience to FDI attacks on actuators is analyzed in the following theorem.

Theorem 1. *Consider the dynamical system (9) whose initial values $z_i(0)$ are chosen as in (7) and the directed graph associated with the Laplacian matrix $\mathbb{L}$ satisfies Assumption 1. Furthermore, the injections $\delta_{u_i}(t)$ in (3) satisfy Assumption 2. If $\alpha > 0$, $\eta > 0$, $\kappa > 0$, $k < 0$, $k_P > 0$, and $0 < k_I < -kk_P$, the states of the cooperative distributed system in (9) are bounded in the presence of actuator attacks. In addition, for a sufficiently large value of k_I, $x_i(t)$, $\forall i \in \mathcal{V}$, is forced to converge to an arbitrarily small neighborhood around x_{ave}.*

Proof. In the presence of actuator attacks in (3), the cooperative distributed system in (20) can be rewritten in the new coordinates as follows:

$$\dot{\mathbf{x}}_{\mathbf{cl}}(t) = \mathbf{A_{cl}}\mathbf{x_{cl}}(t) + \mathbf{B_{cl}}\delta_u(t), \tag{28}$$

where $\mathbf{x_{cl}}(t) = \left[\tilde{\mathbf{v}}^T(t), \tilde{\theta}^T(t), \tilde{\mathbf{x}}^T(t), \tilde{\mathbf{z}}^T(t)\right]^T$, $\delta_u(t) = \left[\delta_{u_1}(t), \ldots, \delta_{u_n}(t)\right]^T$, $\mathbf{A_{cl}}$ is given in (25), and $\mathbf{B_{cl}}$ is defined as follows:

$$\mathbf{B_{cl}} = \begin{bmatrix}\mathbf{0_{n\times n}} & \mathbf{0}_{n\times n} & \mathbf{I}_n & \mathbf{0}_{n\times n}\end{bmatrix}^{\mathbf{T}}. \tag{29}$$

Due to the globally asymptotic stability of the origin in (20), $\mathbf{A_{cl}}$ in (29) is a Hurwitz matrix. Thus, the cooperative distributed system in (28) is input-to-state stable. This implies that the states in (28) are bounded in the presence of bounded injections $\delta_u(t)$. The state vector $\mathbf{x_{cl}}(t)$ in (28) can be obtained as follows:

$$\mathbf{x_{cl}}(t) = e^{\mathbf{A_{cl}}t}\mathbf{x_{cl}}(0) + \int_0^t e^{\mathbf{A_{cl}}(t-\tau)}\mathbf{B_{cl}}\delta_u(\tau)d\tau. \tag{30}$$

Thus,

$$\lim_{t\to\infty}\|\mathbf{x_{cl}}(t)\| \leq \lim_{t\to\infty}\left\|e^{\mathbf{A_{cl}}t}\mathbf{x_{cl}}(0)\right\| + \lim_{t\to\infty}\left\|\int_0^t e^{\mathbf{A_{cl}}(t-\tau)}\mathbf{B_{cl}}\delta_u(\tau)d\tau\right\| \tag{31}$$

Since $\mathbf{A_{cl}}$ is Hurwitz, $\lim_{t\to\infty}\left\|e^{\mathbf{A_{cl}}t}\mathbf{x_{cl}}(0)\right\| = 0$. Moreover, as $\delta_u(t)$ is assumed to be uniformly bounded (see Assumption 2), there exists a constant vector $\bar{\delta}_u \in \mathbb{R}^n$ so that $\left\|\int_0^t e^{\mathbf{A_{cl}}(t-\tau)}\mathbf{B_{cl}}\delta_u(\tau)d\tau\right\| \leq \left\|\int_0^t e^{\mathbf{A_{cl}}(t-\tau)}\mathbf{B_{cl}}\bar{\delta}_u d\tau\right\|$ [16]. As a result, one can obtain that

$$\lim_{t\to\infty}\|\mathbf{x_{cl}}(t)\| \leq \lim_{t\to\infty}\left\|\int_0^t e^{\mathbf{A_{cl}}(t-\tau)}\mathbf{B_{cl}}\bar{\delta}_u d\tau\right\| = \left\|-\mathbf{A_{cl}}^{-1}\mathbf{B_{cl}}\bar{\delta}_u\right\|. \tag{32}$$

For a sufficiently large value of k_I, $\left\|-\mathbf{A_{cl}}^{-1}\mathbf{B_{cl}}\bar{\delta}_u\right\|$ converges to zero. As a result, $\lim_{t\to\infty}\|\mathbf{x_{cl}}(t)\| \approx 0$. This implies that $\lim_{t\to\infty}\mathbf{x}(t)$ converges to $\bar{\mathbf{x}}$. Since $\bar{\mathbf{x}} = \mathbf{1}_n x_{ave}$ (see (10)), the approximate average consensus is guaranteed provided that k_I is chosen to be sufficiently high. □

Remark 3. The value of k_I characterizes a trade-off between resilience to cyber-attacks in (3) and the transient response of $x_i(t)$. In other words, the higher value of k_I enhances the resilience to cyber-attacks while makes the state trajectories of agents more oscillatory.

4.2. Resilience to eavesdropping cyber-attacks

Eavesdropping attacks are privacy threats that disclose (sensitive) local information of agents ($x_i(t)$). Eavesdropping attacks read some or all of exchanged data and save them for later processing. As stated in [3], this type of cyber-attacks could be the first stage of a more disruptive cyber-attack such as replay attacks. Hence, to avoid potential privacy threats, it is important to preserve the privacy of local information. Resilience to eavesdropping attacks is essential in cases where the privacy of state information is of concern.

The common approach in distributed average consensus is based on exchanging agents' physical states with their neighboring agents, leading to the disclosure of state information. This approach does not provide resilience to eavesdropping cyber-attacks. As one can observe from (9), the proposed distributed average consensus in this paper does not require exchanging agents' physical states $\mathbf{x}(t)$ amongst their neighbors. Instead, it allows exchanging auxiliary variables (virtual states) $(\mathbf{v}(t), \theta(t))$ that do not have any physical meaning and thus are less interesting for attackers. This feature of the proposed algorithm has the potential to facilitate the privacy-preserving of agents' physical states and avoid potential privacy threats such as eavesdropping attacks.

5. Simulation results

In this section, the performance of the proposed resilient distributed average consensus in (9) is evaluated by the following numerical examples.

Example 1. We consider a multi-agent system with single-integrator kinematics in (1) that is composed of $n = 4$ agents. It is

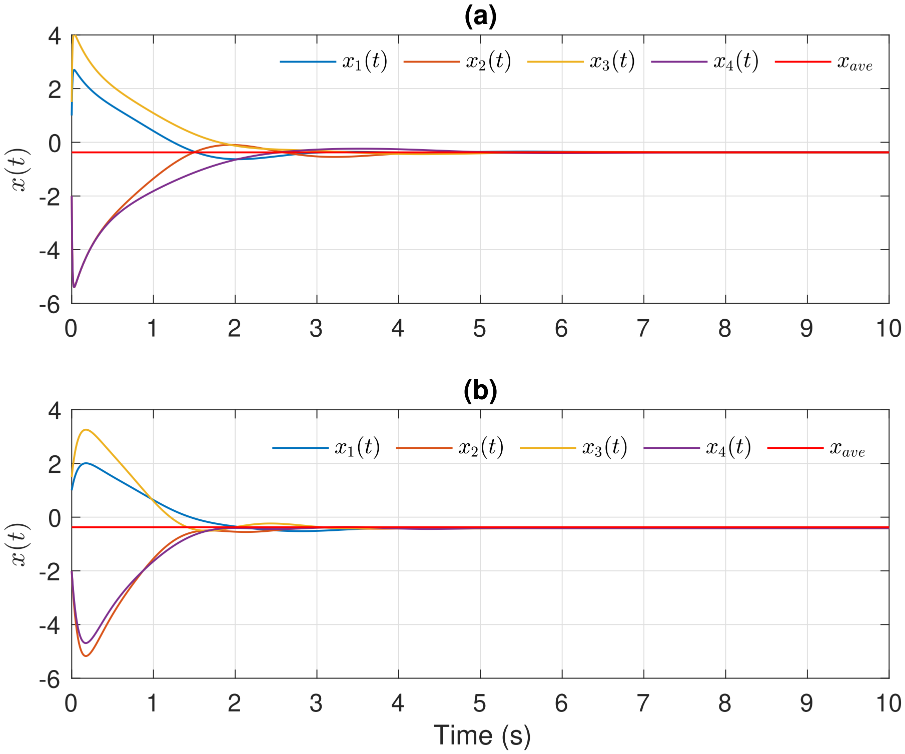

Fig. 2. Performance of the proposed resilient distributed average consensus in (9) in the absence of cyber-attacks with (a) the graph Laplacian $\mathbb{L}_1$ and (b) with the graph Laplacian $\mathbb{L}_2$.

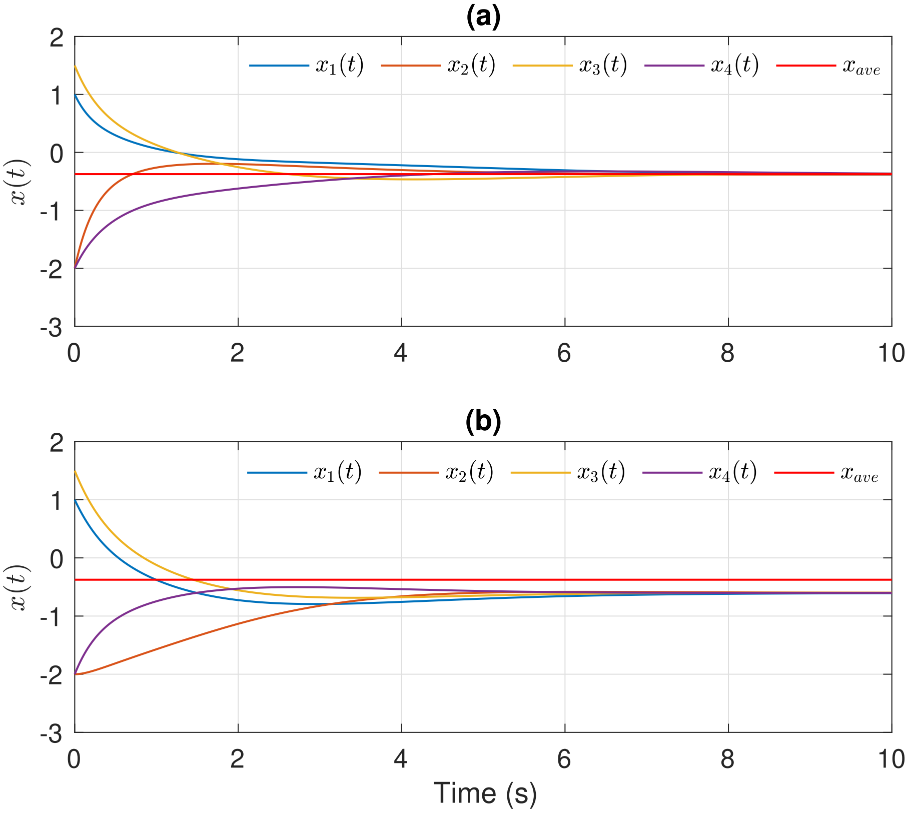

Fig. 3. Performance of the proposed surplus-based average consensus algorithm in [7] in the absence of cyber-attacks with (a) the graph Laplacian $\mathbb{L}_1$ and (b) with the graph Laplacian $\mathbb{L}_2$.

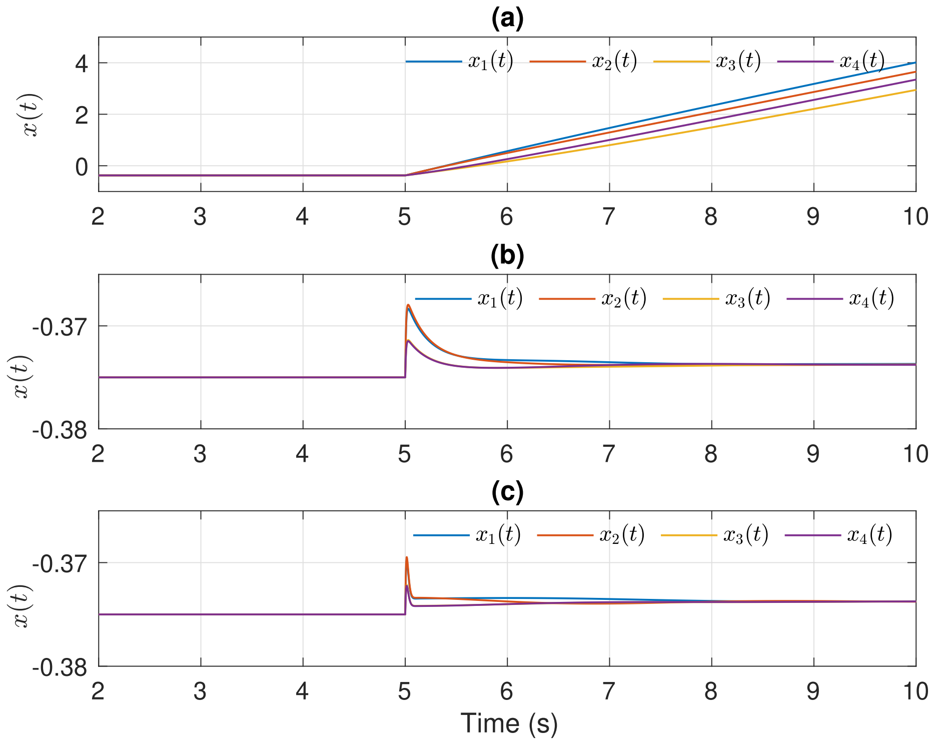

Fig. 4. State trajectories of agents in the presence of cyber-attacks launched at $t = 5$ s using (a) the proposed surplus-based averaging algorithm in [7] with the graph Laplacian $\mathbb{L}_1$, (b) the proposed distributed algorithm with the graph Laplacian $\mathbb{L}_1$, and (c) the proposed distributed algorithm with the graph Laplacian $\mathbb{L}_2$.

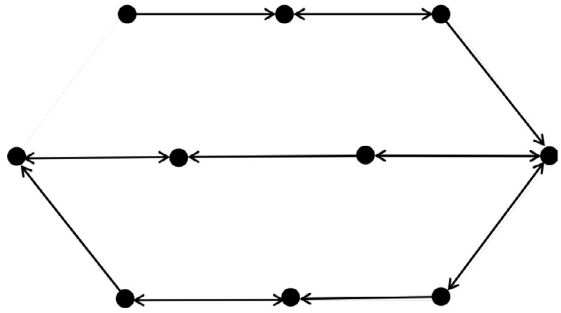

Fig. 5. Digraph associated with Laplacian matrix $\mathbb{L}$ in Example 2.

assumed that $\mathbf{x}(0) = [1 \ -2 \ 1.5 -2]^T$ and $x_{ave} = -0.375$. The initial values of $\mathbf{v}(0)$ and $\theta(0)$ are randomly chosen. We consider two different communication digraphs, one is strongly connected and unbalanced and the other one contains a rooted-out tree. The Laplacian matrices for both digraphs are given as follows:

$$\mathbb{L}_1 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \ \mathbb{L}_2 = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}. \tag{33}$$

Fig. 2 shows the average consensus using the proposed distributed algorithm for both digraph topology. The results in this figure illustrate that although the digraph with the Laplacian matrix $\mathbb{L}_2$ is not strongly connected and balanced, it ensures the average consensus for the multi-agent system.

To highlight the superiority of the proposed distributed average consensus in this paper over the existing literature, two comparative case studies are carried out. In the first case study, we consider a case where agents' actuators are not subject to cyber-attacks. We show that although the proposed surplus-based averaging algorithm in [7] ensures an exact average consensus over digraphs with a Laplacian matrix $\mathbb{L}_1$ (see Fig. 3(a)), it cannot guarantee average consensus over digraphs with a Laplacian matrix $\mathbb{L}_2$ (see Fig. 3(b)). The results of this case study are depicted in Fig. 3.

In the second comparative case study, we test the attack-resilient performance of the proposed resilient distributed average consensus scheme. It is assumed that all agents' actuators are subject to a mixture of constant and time-varying bounded FDI attacks launched at $t = 5$ s. The state trajectories of agents in the presence of cyber-attacks using the proposed distributed surplus-based averaging algorithm in [7] and the proposed distributed average consensus algorithm over digraphs whose Laplacian matrices are given in (33) are depicted in Fig. 4. The parameters of the proposed resilient distributed algorithm in (5) and (8) are designed as $\eta = 1$, $\alpha = 10$, $\kappa = 2$, $k = -10$, $k_P = 11$, and $k_I = 100$.

As one can observe from Fig. 4, in the absence of cyber-attacks ($t < 5$), the average consensus is achieved using both cooperative distributed consensus approaches. However, the existence of cyber-attacks, launched at $t = 5$ s, adversely affects the average consensus in the surplus-based average consensus algorithm in [7] while the proposed distributed algorithm in (5) and (8) is resilient to such attacks. Hence, the following approximate average consensus is ensured in the presence of actuator attacks:

$$\lim_{t\to\infty} x_i(t) - x_{ave} = 0.001, \ \ i = 1, \ldots, 4.$$

Example 2. In this example, we consider a digraph in Fig. 5 with $n = 10$ nodes and 15 edges. The digraph is not strongly connected but includes a rooted-out tree. It is assumed that the initial conditions of each node ($x_{i,0}$) are random. In this example, we have used the same parameters for the resilient controller as in Example 1.

Fig. 6 depicts the state trajectories of agents in the absence ($t < 5$) and presence ($t \geq 5$) of false data injection attacks in the form of (3) launched at $t = 5$ s for a random initial state $x_{i,0}$ with the average of $x_{ave} = 0.1302$. From this figure, one can observe that the

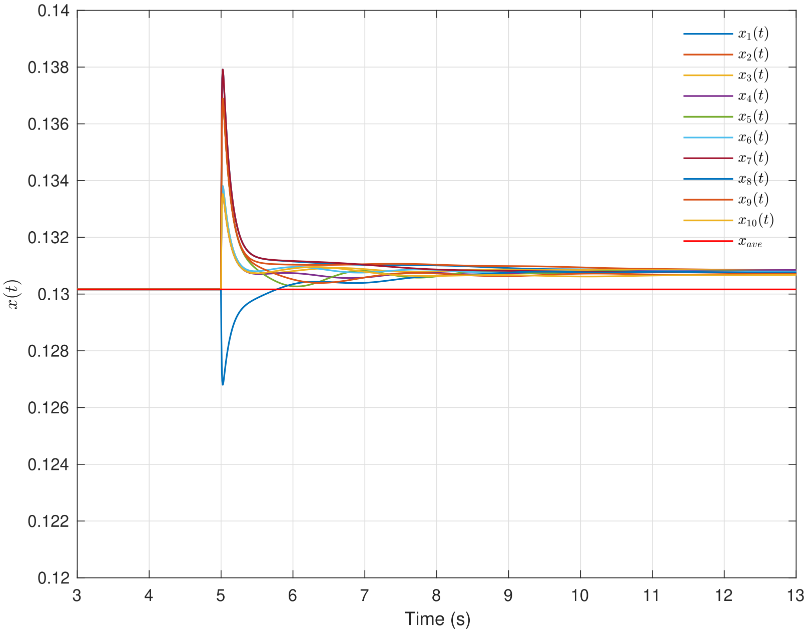

Fig. 6. State trajectories of agents in the presence of cyber-attacks launched at $t = 5$ s using the proposed distributed algorithm with the communication digraph in Fig. 5.

exact average consensus is achieved in the absence of FDI cyber-attacks and the approximate average consensus in (4) is resulted in the presence of cyber-attacks, i.e.,

$$\lim_{t\to\infty} x_i(t) - x_{ave} = 0.0006, \quad i = 1, \ldots, 10.$$

It is worth mentioning that the proposed resilient average consensus algorithms in [7] and [19] cannot be applied to this example as the underlying digraph in Fig. 5 is not strongly connected.

6. Conclusions and future works

This paper deals with the problem of average consensus of continuous-time multi-agent systems with single-integrator kinematics subject to bounded and unknown false data injection cyber-attacks and eavesdropping attacks. The proposed distributed approach is based on the introduction of virtual variables being exchanged via a communication network represented by a directed graph that contains a rooted-out tree. The stability of the overall distributed average consensus system is guaranteed by using a Lyapunov-based approach. The future scope of this work will focus on the design of optimal parameters of the proposed distributed average consensus algorithm and the extension of results to FDI attacks on sensors and communication links as well as higher-order multi-agent systems. Also, a theoretical analysis on the privacy-preserving feature of the proposed distributed algorithm will be considered as the future work of this paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] K. Cai, H. Ishii, Average consensus on general strongly connected digraphs, Automatica 48 (11) (2012) 2750–2761.

[2] T. Charalambous, Y. Yuan, T. Yang, W. Pan, C.N. Hadjicostis, M. Johansson, Distributed finite-time average consensus in digraphs in the presence of time delays, IEEE Trans. Control Netw. Syst. 2 (4) (2015) 370–381.

[3] M.S. Chong, H. Sandberg, A.M.H. Teixeira, A tutorial introduction to security and privacy for cyber-physical systems, in: Proceedings of the 18th European Control Conference (ECC), Naples, Italy, 2019, pp. 968–978. 10.23919/ECC.2019.8795652

[4] H. Dong, C. Li, Y. Zhang, Resilient consensus of multi-agent systems against malicious data injections, J. Frankl. Inst. 357 (4) (2020) 2217–2231.

[5] M. Franceschelli, A. Giua, C. Seatz, Consensus on the average on arbitrary strongly connected digraphsbased on broadcast gossip algorithms, in: Proceedings of the First IFAC Workshop on Estimation and Control of Networked Systems, Venice, Italy, 2009, pp. 66–71.

[6] A. Gusrialdi, Z. Qu, M.A. Simaan, Competitive interaction design of cooperative systems against attacks, IEEE Trans. Autom. Control 63 (9) (2018) 3159–3166.

[7] S. Kawamura, K. Cai, M. Ye, Z. Lin, Tight bound on parameter of surplus-based averaging algorithm over balanced digraphs, Int. J. Control 93 (8) (2020) 1859–1866.

[8] Z. Li, Z. Li, Y. Liu, Resilient control design of the third-order discrete–time connected vehicle systems against cyber-attacks, IEEE Access 8 (2020) 157470–157481.

[9] S. Liu, T. Li, L. Xie, M. Fu, J.-F. Zhang, Continuous-time and sampled-data-based average consensus with logarithmic quantizers, Automatica 49 (11) (2013) 3329–3336.

[10] A. Mustafa, H. Modares, Attack analysis and resilient control design for discrete-time distributed multi-agent systems, IEEE Robot. Autom. Lett. 5 (2) (2020) 369–376.

[11] J. Ni, L. Liu, C. Liu, X. Hu, S. Li, Further improvement of fixed-time protocol for average consensus of multi-agent systems, in: Proceedings of the 20th IFAC World Congress, volume 50, Toulouse, France, 2017, pp. 2523–2529.

[12] R. Olfati-Saber, R.M. Murray, Consensus problems in networks of agents with switching topology and time-delays, IEEE Trans. Autom. Control 49 (9) (2004) 1520–1533.

[13] T. Pultarova, Ukraine grid hack is wake-up call for network operators [news briefing], Eng. Technol. 11 (1) (2005) 12–13.

[14] Z. Qu, M.A. Simaan, Modularized design for cooperative control and plug-and–play operation of networked heterogeneous systems, Automatica 50 (9) (2014) 24052414.

[15] W. Ren, R.W. Beard, E.M. Atkins, Information consensus in multivehicle cooperative control, IEEE Control Syst. Mag. 27 (2) (2007) 71–82.

[16] M.S. Sadabadi, A. Gusrialdi, On resilient design of cooperative systems in presence of cyber-attacks, in: Proceedings of the European Control Conference (ECC), 2021, pp. 946–951. 10.23919/ECC54610.2021.9654990

[17] G.D.L. Torre, T. Yucelen, Adaptive architectures for resilient control of networked multiagent systems in the presence of misbehaving agents, Int. J. Control 91 (3) (2018) 495–507.

[18] A. Wang, W. Liu, T. Li, T. Huang, Privacy-preserving weighted average consensus and optimal attacking strategy for multi-agent networks, J. Frankl. Inst. 358 (6) (2021) 3033–3050, doi:10.1016/j.jfranklin.2021.01.039.

[19] D. Wang, N. Zheng, M. Xu, Y. Wu, Q. Hu, G. Wang, Resilient privacy-preserving average consensus for multi-agent systems under attacks, in: Proceedings of the 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), 2020a, pp. 1399–1405, doi:10.1109/ICARCV50220.2020.9305459. Shenzhen, China

[20] D. Wang, N. Zheng, M. Xu, Y. Wu, Q. Hu, G. Wang, Resilient privacy-preserving average consensus for multi-agent systems under attacks, in: Proceedings of the 16th International Conference on Control, Automation, Robotics and Vision, IEEE, 2020b, pp. 1399–1405.

[21] E. Yildirim, S.B. Sarsilmaz, A.T. Koru, T. Yucelen, On control of multiagent systems in the presence of a misbehaving agent, IEEE Control Syst. Lett. 4 (2) (2020) 456–461.

[22] W. Zeng, M. Chow, Resilient distributed control in the presence of misbehaving agents in networked control systems, IEEE Trans. Cybern. 44 (11) (2014) 2038–2049.

[23] J. Zhou, Y. Lv, G. Wen, X. Yu, Resilient consensus of multiagent systems under malicious attacks: appointed-time observer-based approach, IEEE Trans. Cybern. (2021) 1–13, doi:10.1109/TCYB.2021.3058094.

[24] S. Zuo, D. Yue, Resilient output formation containment of heterogeneous multigroup systems against unbounded attacks, IEEE Trans. Cybern. 52 (3) (2022) 1902–1910.