On Resilient Design of Cooperative Systems in Presence of Cyber-Attacks

Mahdieh S. Sadabadi and Azwirman Gusrialdi

Abstract— **This paper develops a resilient cooperative control system for leader-follower consensus problems subject to false data injection attacks. The attackers are assumed to inject unknown bounded exogenous signals to the actuators of followers and/or the communication networks of the leader and follower states. In order to attenuate the effects of such attacks on the consensus and stability of the system, we develop a cooperative control system augmented with a virtual network and interconnected with a leader and followers so that the leader-follower consensus is guaranteed under unknown attacks. A Lyapunov-based design framework is proposed to guarantee stability and leader-follower consensus against attacks. The effectiveness of the theoretical results is evaluated through a simulation example.**

I. Introduction

Cooperative control systems has recently received significant attention due to their applicability in various problems including smart grids and energy systems, DC microgrids, intelligent transportation system, robotics, and sensor networks [1]–[5]. Despite the potential benefits of the cooperative systems over centralized counterparts such as improved scalability, reliability, resilience to a single point of failure, and reduced cost, the use of communication network makes cooperative systems vulnerable to cyber-physical attacks. A real-world example of such cyber-attack is the coordinated attack on the Ukraine power grid in 2015 which caused several hours of blackout and affected hundred thousands of customers [6].

To address the challenges associated with cyber-physical attacks in cooperative control systems, a number of control strategies have been proposed in the literature. The existing methods can be categorized as attack-detection-based approaches and resilient control techniques.

The first category is mainly based on the identification of malicious nodes and their removal. Examples of attack detection approaches can be found in [7]–[10] and reference therein. The main drawback of these approaches is that there is usually restriction on the number of compromised nodes, the local number of adversarial nodes in the neighborhood of each intact node, or the connectivity of the communication graph. More importantly, the system stability may already have been compromised before the attack is detected. Since cyber-attacks cannot be foreseen in advance, it is thus desirable to design cooperative control algorithms so that the cooperative system becomes resilient against unknown attacks.

To address the issues regarding the first category, resilient cooperative control systems have been developed (e.g. [11]–[19]) that attenuate the adverse effects of malicious agents/nodes and maintain an acceptable performance level of the system against attacks, without detecting and removing misbehaving agents. A distributed adaptive control strategy for multi-agent systems in the presence of misbehaving agents is developed in [16] and is extended to the case of directed communication graphs in [17]. A resilient distributed adaptive $\mathscr{H}_\infty$ control for the leader-follower synchronization under attacks on sensors and actuators is presented in [15]. A cooperative control method, based on a virtual layer, for the leader-follower consensus has been proposed in [11] which provides resilience against attacks on communication networks. In [19], a resilient distributed control algorithm composed of four phases (detection, mitigation, identification, and update) is developed for leader-following consensus problems under some misbehaving followers. The control approaches in [11]–[19] require the connectivity of the communication graph, as well as the knowledge on the neighboring agents. In summary, the existing results on resilient cooperative control have limitations on the connectivity requirement of the communication network, number of compromised nodes, the local number of malicious nodes, and the centralized design of the control parameters. Yet, a systematic resilient cooperative control approach, which does not rely on the above-mentioned limitations and guarantees the stability and consensus while under unknown attacks, is highly desirable.

Motivated by aforementioned challenges, this paper provides a theoretical framework for leader-follower consensus problems with an emphasis on the resilience against false data injection cyber-attacks. The attackers aim to destabilize the consensus dynamics by intercepting the system's communication network and injecting false data to actuators (control input channels). The main objective of this paper is to develop a cooperative control strategy to ensure the leader-follower consensus and guarantee the stability of the cooperative system against potential unknown attacks. Inspired by [11], our proposed cooperative control approach consists of a virtual system with a hidden network where the communication between the virtual system and follower/leader nodes are decentralized. By virtue of the Lyapunov stability theory, we show that by an appropriate choice of control parameters,

M. S. Sadabadi is with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, United Kingdom (e-mail: m.sadabadi@sheffield.ac.uk). The work of M. S. Sadabadi was supported by 2020-2021 National Productivity Investment Fund (NPIF).

A. Gusrialdi is with Faculty of Engineering and Natural Sciences, Tampere University, Tampere 33014, Finland (e-mail: azwirman.gusrialdi@tuni.fi). The work of A. Gusrialdi was supported by Academy of Finland under academy project decision no. 330073.

the origin of the overall system, i.e., the interconnection of the follower/leader nodes and the virtual system is globally asymptotically stable.

The proposed resilient cooperative control mechanism offers the following main features: (i) The proposed control strategy does not require any information about the nature and/or location of cyber-attacks and does not have any restriction on the number of malicious nodes. (ii) The controller design for each follower is decentralized without requiring any knowledge on neighboring nodes. (iii) In contrast to the virtual system proposed in [11]–[14], the communication graph in the virtual network does not need to be necessarily connected. Instead, it is assumed that the communication digraph contains a rooted-out tree. These important features as well as the decentralized design of control variables facilitate creating a plug-and-play environment, where follower nodes can be easily plugged in/out as long as the updated digraph has a rooted-out tree. Furthermore, the physical states of cooperative systems are not being exchanged with other nodes in the virtual system. As a result, the risk of the virtual system being exposed to the adversary might be minimized. (iv) By means of the proposed cooperative control system, leader-follower consensus is guaranteed in the presence of the aforementioned attack types.

Throughout this paper, $\mathbf{1}_n$ is an $n \times 1$ vector of ones, $\mathbf{0}_n$ is an $n \times 1$ zero vector, $\mathbf{I}_n$ is an $n \times n$ Identity matrix, and $\mathbf{0}_{n\times m}$ is a zero matrix of dimension $n \times m$. Throughout the paper, $\mathrm{col}(x) = \begin{bmatrix} x_1^T & x_2^T & \dots & x_n^T \end{bmatrix}^T$ and $[a] = \mathrm{diag}(a_1, a_2, \dots, a_n)$. For a symmetric matrix X, the positive definite and positive semidefinite operators are respectively shown by $X \succ 0$ and $X \succeq 0$. We define $\mathbb{R}_+ := \{x \in \mathbb{R} \ |x > 0\}$ and $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \ |x \geq 0\}$.

II. Problem Statement

Consider a cooperative system consisting of $n+1$ nodes, where a leader node is labeled by 0 and the follower nodes are labeled by i, $i = 1, \dots, n$. The information flow among the nodes is modeled by a directed graph $\mathscr{G} = (\mathscr{V}(\mathscr{G}), \mathscr{E}(\mathscr{G}))$, where the node set $\mathscr{V}(\mathscr{G})$ and the edge set $\mathscr{E}(\mathscr{G})$ represent nodes and integrant information exchange links, respectively. Let $x_i(t) \in \mathbb{R}$ denotes the state of node i whose dynamics are given by

$$\dot{x}_i(t) = u_i(t), \tag{1}$$

for $i \in \mathscr{V}(\mathscr{G})$, where $u_i(t) \in \mathbb{R}$ is the control input of node i. The main objective is to design the control input $u_i(t)$ such that

- The cooperative system in (1) reaches a consensus, i.e.

$$\lim_{t\to\infty} (x_i(t) - x_j(t)) = 0, \ i, j \in \mathscr{V}(\mathscr{G}), \tag{2}$$

- The follower nodes track the leader node, i.e.

$$\lim_{t\to\infty} x_i(t) = x_0, \ i \in \mathscr{V}(\mathscr{G}), \tag{3}$$

where $x_0 \in \mathbb{R}$ is the state of the leader node.

A. Leader–Following Consensus

Consider the following distributed control protocol [20]:

$$u_i(t) = a_{i0}(x_0 - x_i(t)) + \sum_{j=1}^{n} a_{ij}(x_j(t) - x_i(t)), \tag{4}$$

for $i \in \mathscr{V}(\mathscr{G})$, $a_{ij} \in \{0, 1\}$, and $a_{ij} = 1$ if the follower node i receives information from node j including the leader node 0; otherwise, $a_{ij} = 0$. The cooperative system with the control protocol (4) can be written in a compact form as

$$\dot{\mathbf{x}}(t) = -(\mathscr{L} + \mathscr{A})\,\mathbf{x}(t) + (\mathscr{L} + \mathscr{A})\,\mathbf{1}_n x_0, \tag{5}$$

where $\mathbf{x}(t) = \mathrm{col}(x(t))$, $\mathscr{A} = \mathrm{diag}(a_{10}, \dots, a_{n0})$, and $\mathscr{L}$ is the Laplacian matrix associated with the digraph $\mathscr{G}$. Defining the error state $\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{1}_n x_0$, the error dynamics of the closed-loop system (5) can be written as

$$\dot{\mathbf{e}}(t) = -(\mathscr{L} + \mathscr{A})\,\mathbf{e}(t). \tag{6}$$

If there exists a node which has access to the state information of its neighbours and the leader node and the graph contains a directed spanning tree with the leader node as the root node, it can be shown that $-(\mathscr{L} + \mathscr{A})$ in (6) is Hurwitz [21]. As a result, we have $\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{1}_n x_0$, that is the consensus and the tracking objectives in (2) and (3) are achieved [20].

B. Cyber-Attack Modeling

In practice, malicious attackers might inject unknown exogenous signals to the control input channels of control nodes and/or the communication links of the physical states. Note that in this paper, we assume that sensors are secured and are not attacked. Without loss of generality, it is assumed that cyber-attacks are bounded. This is a reasonable assumption since from the attacker's perspective any intelligent attacker would aim at destabilizing the system with a bounded injection to avoid the attack detection [11]. On the other hand, from the defender's perspective, in the case of unbounded injection, simple filtering can be applied to each node in order to remove excessively large signals received from its neighbors [11]. Similarly, excessively large signals observed in actuators or can be also ignored. To this end, a filtering and bad-date rejection technique based on a thresholding mechanism has been proposed in [11].

Under the potential attack $\delta_{u_i}(t)$ on the control input channel i (actuator), the false data injection cyber-attack can be modeled as follows:

$$\hat{u}_i(t) = u_i(t) + \lambda_{u_i}\delta_{u_i}(t), \tag{7}$$

where $\hat{u}_i(t)$ is the corrupted control input and $\lambda_{u_i} \in \{0, 1\}$, where $\lambda_{u_i} = 1$ indicates the presence of an attack on the control input of the follower node i. Moreover, the measurement state $x_i(t)$ communicated from node i (including the leader node) to the follower node j in the presence of the attack is modeled as

$$\hat{x}_{[i,j]}(t) = x_i(t) + \lambda_{x_{[i,j]}}\delta_{x_{[i,j]}}(t), \tag{8}$$

for $i=0,1,\ldots,n$, where $x_i(t)$ is the unattacked measurement of state of node i, $\hat{x}_{[i,j]}$ is the disrupted state measurement sent to node j, $\lambda_{x_{[i,j]}}=1$ if there is an attack on the communication of $x_i(t)$ from node i to node j, and 0 otherwise.

Cooperative system in (5) is not resilient against attacks on actuators and/or communication links and does not guarantee the consensus and tracking objectives in (2) and (3) in the presence of false data injection attacks in (7) [11], as will be shown in Fig. 2 (b) in Section IV. The main objective of this paper is to develop an attack-resilient distributed control strategy such that the objectives given in (2) and (3) are guaranteed in the presence of the unknown potential attacks in (7) and (8).

III. Resilience Cooperative Control

This section is devoted to the development of a new resilient cooperative control system. The equilibria and stability analysis are then presented.

A. Proposed Attack-Resilient Distributed Control Strategy

In order to guarantee the consensus and the tracking in the presence of cyber-attacks, a control layer with a virtual (and possibly hidden) distributed network (also called as *virtual system*) is introduced in addition to the (peer-to-peer) communication network utilized to implement cooperative control (5). The number of nodes in the virtual network is equal to n. The dynamics of the virtual layer are given as follows:

$$\begin{aligned} T_{v_i}\dot{v}_i =& -\alpha_i\left(v_i-x_i\right)-K\sum_{j=1}^{n}\gamma_{j,i}\left(\theta_i-\theta_j\right)-\beta\gamma_{i0}\left(v_i-x_0\right),\\ T_{\theta_i}\dot{\theta}_i =& -\eta_i\theta_i+\sum_{j=1}^{n}\gamma_{i,j}\left(v_i-v_j\right),\\ T_{w_i}\dot{w}_i =& \alpha_i\left(v_i-x_i\right),\\ u_i =& k_{1,i}\alpha_i\left(v_i-x_i\right)+k_{2,i}x_i+k_{3,i}w_i, \end{aligned} \tag{9}$$

for $i=1,\ldots,n$. The parameters $T_{w_i}\in\mathbb{R}_+$, $T_{v_i}\in\mathbb{R}_+$, $T_{\theta_i}\in\mathbb{R}_+$, $K\in\mathbb{R}_+$, $\eta_i\in\mathbb{R}_+$, $\gamma_{ij}\in\mathbb{R}_{\geq 0}$, $\alpha_i\in\mathbb{R}_+$, $\beta\in\mathbb{R}_+$, and $(k_{1,i},k_{2,i},k_{3,i})$ are the design parameters of the distributed control protocol (virtual system) that can be used to guarantee the closed-loop stability in the presence of unknown attacks. Scalar $\gamma_{i0}\in\{0,1\}$, where $\gamma_{i0}=1$ if the virtual node i receives information from the leader; otherwise, we set $\gamma_{i0}=0$.

The schematic diagram of the closed-loop system and the interconnection between the control layer and the agents is depicted in Fig. 1. Interpretation and rationale of the virtual system proposed in (9) are described below. The virtual network can be realized by using the cloud computing and communication technologies while the information flow in the virtual network can be directed by taking advantage of the flexibility offered by software-defined networking [22]. Each follower node has a corresponding virtual node in cloud which can perform computation and send back the signal used for resilient control. Moreover, each follower node and the leader publish its state $x_i(t)$ and x_0 to the cloud (see Fig. 1). The resilient control $u_i(t)$ in (9) implemented by local controller of each node can also be written as

$$\begin{aligned} u_i(t)=&\mu\left(a_{i0}\left(x_0-x_i(t)\right)+\sum_{j=1}^{n}a_{ij}\left(x_j(t)-x_i(t)\right)\right)\\ &+k_{1,i}\alpha_i\left(v_i(t)-x_i(t)\right)+k_{2,i}x_i(t)+k_{3,i}w_i(t), \end{aligned}$$

with scalar $\mu=0$ which means that each follower node ignores the information that it receives via peer-to-peer communication network used to implement (4) from the leader node and its neighbors. In other words, instead of using the information that they receive over the peer-to-peer network, the virtual nodes first try to "randomize" the local states $x_i(t)$ using the dynamics of the auxiliary variables v_i(t), $w_i(t)$, and $\theta_i(t)$ in (9) before exchanging these values with other virtual nodes in the virtual network (which can be randomly chosen as will be shown later).

Remark 1: Note that the virtual variables v_i, w_i, and θ_i in (9) do not have any physical meaning and their initial values can be set to any values. Furthermore, the virtual network might be secured; however, this is not a must. Even though the virtual network is not secured, the adversary might find it hard to destabilize the system since it would be very difficult for the adversary to recognize the auxiliary variables among the huge amount of other variables. The adversary might corrupt the information exchange from the follower i and/or the leader to j ($x_{[i,j]}(t)$); however, according to (9) $x_{[i,j]}(t)$ is not exchanged in the cloud or utilized in (9). Hence, the existence of cyber-attacks on communication channels used for cooperative control (5) does not affect the proposed resilient control. This is one of the main advantages of virtual system proposed in this work compared to the ones presented in [11]–[14], as it does not require $x_{[i,j]}(t)$ to be exchanged in the virtual layer. As a result, the hidden layer might be secured with a high probability.

Dynamics of the virtual system including its interconnection with the original cooperative system are designed to satisfy the following properties: (i) its interconnection with the cooperative system does not impact the convergence of the states of cooperative systems to the leader's value x_0; (ii) the robustification strategy is automatically activated when attacks appear anywhere in the cooperative system; (iii) the virtual system maintains tracking objectives in (2) and (3) in the presence of unknown cyber-attacks (see Theorem 1 in Section III). In Proposition 1 given in Section III-C, conditions on the decentralized design of the control parameters are proposed.

The overall system, i.e., the interconnection of the follower nodes, the leader, and the cooperative control system with virtual network in (9) in the presence of unknown false data injection attacks can be described as follows:

$$\begin{aligned} \left[T_v\right]\dot{\mathbf{v}}(t) =& -\left[\alpha\right]\left(\mathbf{v}(t)-\mathbf{x}(t)\right)-K\mathscr{L}_h^T\theta(t)-\beta\mathscr{A}_h\left(\mathbf{v}(t)-\mathbf{1}_n x_0\right)\\ \left[T_\theta\right]\dot{\theta}(t) =& -\left[\eta\right]\theta(t)+\mathscr{L}_h\mathbf{v}(t),\\ \dot{\mathbf{x}}(t) =& \left[k_1\right]\left[\alpha\right]\left(\mathbf{v}(t)-\mathbf{x}(t)\right)+\left[k_2\right]\mathbf{x}(t)+\left[k_3\right]\mathbf{w}(t)+\delta(t),\\ \left[T_w\right]\dot{\mathbf{w}}(t) =& \left[\alpha\right]\left(\mathbf{v}(t)-\mathbf{x}(t)\right), \end{aligned} \tag{10}$$

where $\mathbf{w}(t)=\mathrm{col}(w(t))$, $\mathbf{v}(t)=\mathrm{col}(v(t))$, $\theta(t)=\mathrm{col}(\theta(t))$, and $\delta(t)=\mathrm{col}(\delta(t))$, where $\delta_i(t)=\lambda_{u_i}\delta_{u_i}(t)+\mu\left(a_{i0}\lambda_{x_{[0,i]}}\delta_{x_{[0,i]}}(t)-\sum_{j=1}^{n}a_{ij}\lambda_{x_{[j,i]}}\delta_{x_{[j,i]}}(t)\right)$.

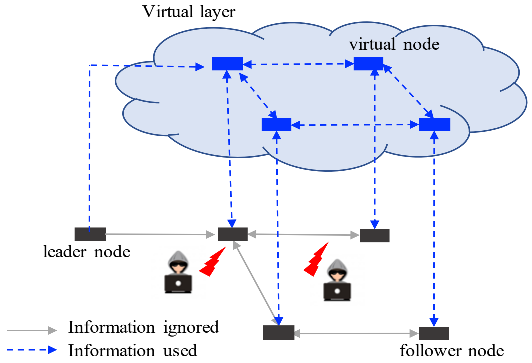

Fig. 1. Resilient design of cooperative systems with the proposed virtual layer in (9). The solid black lines represent the information flow from the leader to the follower nodes, while the dashed blue lines denote the communication links between the agents (physical layer) and the control layer as well as the communication amongst the virtual nodes in the control layer.

$\mathscr{A}_h = \mathrm{diag}(\gamma_{10}, \ldots, \gamma_{n0})$, and $\mathscr{L}_h \in \mathbb{R}^{n\times n}$ is the Laplacian matrix associated with the communication digraph in the virtual node which is not necessarily equal to $\mathscr{L}$. Recall again that in the above cooperative system, there are two types of communication networks: physical and virtual communication. The physical uses a peer-to-peer communication network, while the virtual layer uses a cloud-based communication.

Assumption 1: It is assumed that the communication digraph in the virtual layer contains a rooted-out tree. As a result, $rank(\mathscr{L}_h) = n-1$.

Next, let us define $\mathbf{x}_{\mathbf{cl}}(t) = \begin{bmatrix} \mathbf{e}_v^T & \mathbf{e}_\theta^T & \mathbf{e}_x^T & \mathbf{e}_w^T \end{bmatrix}^T$, where $\mathbf{e}_v(t) = \mathbf{v}(t) - \bar{\mathbf{v}}$, $\mathbf{e}_\theta(t) = \theta(t) - \bar{\theta}$, $\mathbf{e}_x(t) = \mathbf{x}(t) - \bar{\mathbf{x}}$, $\mathbf{e}_w(t) = \mathbf{w}(t) - \bar{\mathbf{w}}$, and $(\bar{\mathbf{v}}, \bar{\theta}, \bar{\mathbf{x}}, \bar{\mathbf{w}})$ are the equilibria of (10) in the absence of $\delta(t)$. Then, the cooperative system in (10) can be rewritten in the new coordinates as follows:

$$\dot{\mathbf{x}}_{\mathbf{cl}}(t) = \mathbf{A}_{\mathbf{cl}}\mathbf{x}_{\mathbf{cl}}(t) + \mathbf{B}_{\mathbf{cl}}\delta(t), \tag{11}$$

where $(\mathbf{A}_{\mathbf{cl}}, \mathbf{B}_{\mathbf{cl}})$ are defined as follows:

$$\begin{aligned}\mathbf{A}_{\mathbf{cl}} &= \begin{bmatrix} -[T_v]^{-1}([\alpha]+\beta\mathscr{A}_h) & -K[T_v]^{-1}\mathscr{L}_h^T & [T_v]^{-1}[\alpha] & \mathbf{0}_{n\times n} \\ [T_\theta]^{-1}\mathscr{L}_h & -[T_\theta]^{-1}[\eta] & \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \\ [k_1][\alpha] & \mathbf{0}_{n\times n} & [k_2]-[k_1][\alpha] & [k_3] \\ [T_w]^{-1}[\alpha] & \mathbf{0}_{n\times n} & -[T_w]^{-1}[\alpha] & \mathbf{0}_{n\times n} \end{bmatrix} \\ \mathbf{B}_{\mathbf{cl}} &= \begin{bmatrix} \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \mathbf{I}_n & \mathbf{0}_{n\times n} \end{bmatrix}^T.\end{aligned} \tag{12}$$

In the following, we discuss the existence of the equilibria $(\bar{\mathbf{v}}, \bar{\theta}, \bar{\mathbf{x}}, \bar{\mathbf{w}})$ and the stability analysis of the cooperative system in (11).

B. Existence and Uniqueness of Equilibria

First, the following lemma discusses the existence of the equilibrium points $(\bar{\mathbf{v}}, \bar{\theta}, \bar{\mathbf{x}}, \bar{\mathbf{w}})$ of the cooperative system (10) in the absence of the attack vector $\delta(t)$.

Lemma 1: Consider the cooperative system in (11) with the proposed control scheme in (9) in the absence of the attack $\delta(t)$. Let Assumption 1 hold. If $k_{3,i} \neq 0$ for $i \in \mathscr{V}(\mathscr{G})$, there exists a unique equilibrium $(\bar{\mathbf{v}}, \bar{\theta}, \bar{\mathbf{x}}, \bar{\mathbf{w}})$ satisfying

$$\bar{\mathbf{v}} = \mathbf{1}_n x_0, \;\; \bar{\theta} = \mathbf{0}_n, \;\; \bar{\mathbf{x}} = \mathbf{1}_n x_0, \;\; \bar{\mathbf{w}} = -[k_3]^{-1}[k_2]\bar{\mathbf{x}}. \tag{13}$$

Proof: See Appendix V-A. ■

C. Stability Analysis

The following results illustrate that for appropriately chosen parameters in (10) and in the absence of cyber-attacks, the interconnected system (11) is globally stable, that is the virtual system does not impact the convergence of the state x_i to the leader's value x_0.

Proposition 1: Let Assumption 1 hold. If $\mathscr{A}_h \succeq 0$ has at least one positive diagonal element and $K \in \mathbb{R}_+$, $\beta \in \mathbb{R}_+$, $[\alpha] \succ 0$, $[\eta] \succ 0$, $[T_v] \succ 0$, $[T_\theta] \succ 0$, $[T_w] \succ 0$, and $(k_{1,i}, k_{2,i}, k_{3,i})$ belongs to the set

$$\mathscr{Z}_{[i]} = \left\{ k_{1,i} > 0, \; k_{2,i} < 0, \; 0 < \frac{k_{3,i}}{T_{w_i}} < -k_{1,i}k_{2,i} \right\}, \; i \in \mathscr{V}(\mathscr{G}) \tag{14}$$

then, $\mathbf{A}_{\mathbf{cl}}$ given in (12) is a Hurwitz matrix.

Proof: See Appendix V-B. ■

Finally, the following theorem shows that using the proposed method, the objectives (2) and (3) are achieved in the presence of bounded attacks.

Theorem 1: Let Assumption 1 hold. Moreover, let us choose $\mathscr{A}_h \succeq 0$ to have at least one positive diagonal element, $K \in \mathbb{R}_+$, $\beta \in \mathbb{R}_+$, $[\alpha] \succ 0$, $[T_v] \succ 0$, $[T_\theta] \succ 0$, $[T_w] \succ 0$, and $(k_{1,i}, k_{2,i}, k_{3,i})$ belongs to the set (14). The states of the cooperative system in (10) are then bounded for any bounded adversary attack $\delta(t)$. Furthermore, for a sufficiently large value of $k_{3,i}$, $\forall i \in \mathscr{V}(\mathscr{G})$, $\lim_{t\to\infty} x_i(t) = x_0$, $i \in \mathscr{V}(\mathscr{G})$.

Proof: Since $\mathbf{A}_{\mathbf{cl}}$ in (12) is a Hurwitz matrix as shown in Proposition 1, the cooperative system in (10) is input-to-state stable. This implies that if $\delta(t)$ is bounded, the states of the cooperative system are bounded too.

From the closed-loop system in (11), the closed-loop state vector $\mathbf{x}_{\mathbf{cl}}(t)$ can be obtained as follows:

$$\mathbf{x}_{\mathbf{cl}}(t) = e^{\mathbf{A}_{\mathbf{cl}}t}\mathbf{x}_{\mathbf{cl}}(0) + \int_0^t e^{\mathbf{A}_{\mathbf{cl}}(t-\tau)}\mathbf{B}_{\mathbf{cl}}\delta(\tau)d\tau. \tag{15}$$

Since $\delta(t)$ is uniformly bounded, there exists a constant vector $\bar{\delta} \in \mathbb{R}^n$ such that $\left\| \int_0^t e^{\mathbf{A}_{\mathbf{cl}}(t-\tau)}\mathbf{B}_{\mathbf{cl}}\delta(\tau)d\tau \right\| \leq \left\| \int_0^t e^{\mathbf{A}_{\mathbf{cl}}(t-\tau)}\mathbf{B}_{\mathbf{cl}}\bar{\delta}d\tau \right\|$. Hence,

$$\lim_{t\to\infty} \|\mathbf{x}_{\mathbf{cl}}(t)\| \leq \lim_{t\to\infty} \left\| \int_0^t e^{\mathbf{A}_{\mathbf{cl}}(t-\tau)}\mathbf{B}_{\mathbf{cl}}\bar{\delta}d\tau \right\| = \left\| -{\mathbf{A}_{\mathbf{cl}}}^{-1}\mathbf{B}_{\mathbf{cl}}\bar{\delta} \right\|. \tag{16}$$

It can be shown that ${\mathbf{A}_{\mathbf{cl}}}^{-1}\mathbf{B}_{\mathbf{cl}}\bar{\delta}$ can be obtained as follows:

$${\mathbf{A}_{\mathbf{cl}}}^{-1}\mathbf{B}_{\mathbf{cl}}\bar{\delta} = \begin{bmatrix} \mathbf{0}_{n\times n} \\ [k_3]^{-1} \\ \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} \end{bmatrix} \bar{\delta}, \tag{17}$$

From the above equation, it follows that for a sufficiently large value of $k_{3,i}$, $\forall i \in \mathscr{V}(\mathscr{G})$, $\left\| -{\mathbf{A}_{\mathbf{cl}}}^{-1}\mathbf{B}_{\mathbf{cl}}\bar{\delta} \right\| \approx 0$. As a result, $\lim_{t\to\infty} \|\mathbf{x}_{\mathbf{cl}}(t)\| \approx 0$. This implies that $\lim_{t\to\infty} \mathbf{x}(t) \approx \bar{\mathbf{x}}$.

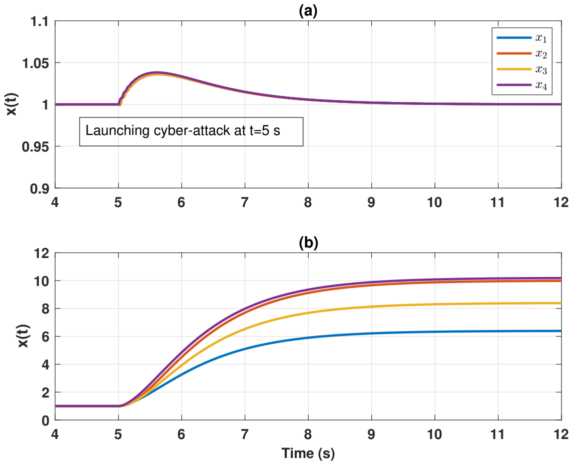

Fig. 2. Trajectories of the states of the followers in the presence of the cyber-attack $\delta(t)$: (a) with the proposed attack-resilient cooperative system in (10) and (b) without using the virtual layer.

As a result, the leader-follower consensus is guaranteed as illustrated by $\bar{\mathbf{x}}$ in (13). ■

As can be seen from Theorem 1, the Laplacian matrix $\mathscr{L}_h$ associated with the communication graph in the virtual layer is not necessarily connected. Moreover, the controller design for each follower is decentralized without requiring any knowledge on the neighboring nodes.

IV. Simulation Results

In this section, the performance of the proposed resilient distributed control approach is verified through the following example.

Example. Consider the following multi-agent system:

$$\dot{x}_i(t) = u_i(t),\ x_0 = 1, \tag{18}$$

for $i = 1, \ldots, 4$, where $x_i(t) \in \mathbb{R}$ and $u_i(t) \in \mathbb{R}$. The parameters of the cooperative system in (10) are given as follows:

$$\mathscr{L}_h = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & -1 & 2 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix},\ \mathscr{A}_h = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{19}$$

$[T_\theta] = 10^{-3} \times \mathbf{I}_4$, $[T_v] = 10^{-3} \times \mathbf{I}_4$, $[T_w] = 10^{-1}\mathbf{I}_4$, $[\eta] = 10^{-1}\mathbf{I}_4$, $K = 10$, $[\alpha] = 10 \times \mathbf{I}_4$, $\beta = 1$, $[k_1] = 11 \times \mathbf{I}_4$, $[k_2] = -120 \times \mathbf{I}_4$, and $[k_3] = 120 \times \mathbf{I}_4$.

The performance of the proposed control technique in (9) is assessed under the following the dynamic cyber-attacks, launched at $t = 5\ s$:

$$\dot{\delta}(t) = A_d\delta(t) + B_d\delta_0(t), \tag{20}$$

where

$$\delta_0(t) = 4 \times \mathbf{1}_4,\quad A_d = -\mathbf{I}_4,\quad B_d = \begin{bmatrix} 1 & 2 & 4 & 2 \\ -9 & 4 & 1 & 3 \\ -4 & 3 & 1 & 2 \\ 2 & 1 & 4 & 3 \end{bmatrix}. \tag{21}$$

Note that the above attack dynamics are not known to agents. Fig. 2 shows the states of the follower nodes and illustrates that the effects of the cyber-attack $\delta(t)$ on $\mathbf{x}(t)$ are compensated by means of the attack-resilient cooperative system in (10). As expected from Theorem 1, the proposed control framework achieves consensus; moreover, the followers track the leader node with a zero steady-state error. We also compare our results with the performance of the cooperative system (5) with $\mathscr{L} = \mathscr{L}_h$ and $\eta = 1$, as depicted in Fig. 2 (b). The results reveal that without the virtual layer the consensus and tracking objectives in (2) and (3) are no longer achieved.

V. Conclusion

In this paper, a resilient cooperation control strategy for the leader-follower consensus problem in the presence of cyber-attacks is proposed. The attackers are assumed to infiltrate actuators and affect the communication networks of the leader and follower states by injecting false data. An attack-resilient cooperative control framework consisting of a virtual layer is developed and investigated under unknown bounded attacks. In contrast to the existing literature, our proposed solution does not require the connectivity of the communication graph. By virtue of the Lyapunov stability method and network control theory, a concise stability certificate is derived and the leader-follower consensus is guaranteed against attacks. An illustrative example verifies the effectiveness of the proposed cooperative control strategy.

Appendices

A. *Proof of Lemma 1*

Consider the cooperative system in (10). The equilibrium points of (10) in the absence of attack, i.e. $\delta(t) = 0$, can be found by solving the following equations:

$$\mathbf{0}_n = -[\eta]\bar{\theta} + \mathscr{L}_h\bar{\mathbf{v}}, \tag{22a}$$

$$\mathbf{0}_n = [\alpha](\bar{\mathbf{v}} - \bar{\mathbf{x}}), \tag{22b}$$

$$\mathbf{0}_n = -[\alpha](\bar{\mathbf{v}} - \bar{\mathbf{x}}) - K\mathscr{L}_h^T\bar{\theta} - \beta\mathscr{A}_h(\bar{\mathbf{v}} - \mathbf{1}_n x_0), \tag{22c}$$

$$0 = [k_1][\alpha](\bar{\mathbf{v}} - \bar{\mathbf{x}}) + [k_2]\bar{\mathbf{x}} + [k_3]\bar{\mathbf{w}} + \bar{\delta}. \tag{22d}$$

From (22a), one obtains that $\bar{\theta} = [\eta]^{-1}\mathscr{L}_h\bar{\mathbf{v}}$. Since $[\alpha]$ is a non-singular matrix, from (22b) we have $\bar{\mathbf{x}} = \bar{\mathbf{v}}$. By replacing $\bar{\mathbf{x}} = \bar{\mathbf{v}}$ and $\bar{\theta} = [\eta]^{-1}\mathscr{L}_h\bar{\mathbf{v}}$ in (22c), one obtains that

$$-\left(K\mathscr{L}_h^T[\eta]^{-1}\mathscr{L}_h\bar{\mathbf{v}} + \beta\mathscr{A}_h(\bar{\mathbf{v}} - \mathbf{1}_n x_0)\right) = \mathbf{0}_n. \tag{23}$$

Invoking the properties of the Laplacian matrix $\mathscr{L}_h$ as $\mathscr{L}_h\mathbf{1}_n = \mathbf{0}_n$, from the above equation, it follows that

$$\underbrace{\left(K\mathscr{L}_h^T[\eta]^{-1}\mathscr{L}_h + \beta\mathscr{A}_h\right)}_{\mathscr{X}}(\bar{\mathbf{v}} - \mathbf{1}_n x_0) = \mathbf{0}_n. \tag{24}$$

Since $\mathscr{L}_h^T[\eta]^{-1}\mathscr{L}_h \succeq 0$, $\mathscr{A}_h \succeq 0$, and $rank(\mathscr{L}_h^T[\eta]^{-1}\mathscr{L}_h) = n-1$, it can be shown that $\mathscr{X} \succ 0$; hence, it is invertible. As a result, the above equality leads to $\bar{\mathbf{v}} = \mathbf{1}_n x_0$. As a result, $\bar{\mathbf{x}} = \bar{\mathbf{v}} = \mathbf{1}_n x_0$. By replacing $\bar{\mathbf{x}}$ and $\bar{\mathbf{v}}$ in (22d), it follows that $\bar{\mathbf{w}} = -[k_3]^{-1}[k_2]\bar{\mathbf{x}}$. Furthermore, from $\bar{\theta} = [\eta]^{-1}\mathscr{L}_h\bar{\mathbf{v}}$ and $\bar{\mathbf{v}} = \mathbf{1}_n x_0$, one obtains that $\bar{\theta} = \mathbf{0}_n$.

B. Proof of Proposition 1

Let $\mathbf{d}(t)=\mathbf{0}_n$ in (11). Then, it suffices to show that the origin in (11) is globally asymptotically stable. To this end, the following quadratic-type Lyapunov function is considered:

$$\begin{aligned}\mathscr{V}(\mathbf{x_{cl}}) &= \frac{1}{2}\mathbf{e}_v^T(t)\,[T_v]\,\mathbf{e}_v(t)+\frac{K}{2}\mathbf{e}_\theta^T(t)\,[T_\theta]\,\mathbf{e}_\theta(t)\\ &+\frac{1}{2}\sum_{i=1}^{n}\left[e_{x_i}(t)\ e_{w_i}(t)\right]P_i\left[e_{x_i}(t)\ e_{w_i}(t)\right]^T,\end{aligned} \tag{25}$$

where $P_i \in \mathbb{R}^{2\times 2}$ is defined as follows:

$$P_i=\left[\begin{array}{cc}\rho_i & -\frac{1}{T_{w_i}}\rho_i\nu_i\\ -\frac{1}{T_{w_i}}\rho_i\nu_i & \nu_i\left(1+\frac{1}{T_{w_i}^2}\rho_i\nu_i\right)\end{array}\right], \tag{26}$$

where $\rho_i>0$ and $\nu_i>0$ are determined based on any values of $(k_{1,i},k_{2,i},k_{3,i},T_{w_i})$ in $\mathscr{Z}_{[i]}$ given in (14) as follows:

$$\rho_i=\frac{k_{2,i}}{k_{2,i}k_{1,i}+\frac{1}{T_{w_i}}k_{3,i}},\ \nu_i=-T_{w_i}\frac{k_{3,i}}{k_{2,i}}. \tag{27}$$

Note that $trace(P_i)>0$ and $det(P_i)>0$, hence $P_i\succ 0$. The time derivative of $\mathscr{V}(\mathbf{x_{cl}})$ in (25) along the trajectories (11) is expressed as

$$\begin{aligned}&\dot{\mathscr{V}}(\mathbf{x_{cl}})=-\frac{1}{2}\left(\mathbf{e}_v^T\,[\alpha]\,(\mathbf{e}_v-\mathbf{e}_x)-(\mathbf{e}_v-\mathbf{e}_x)^T\,[\alpha]\,\mathbf{e}_v\right)\\ &-K\mathbf{e}_\theta^T\,[\eta]\,\mathbf{e}_\theta-\frac{K}{2}\left(\mathbf{e}_v^T\mathscr{L}_h^T\mathbf{e}_\theta+\mathbf{e}_\theta^T\mathscr{L}_h\mathbf{e}_v\right)-\frac{\beta K}{2}\mathbf{e}_v^T\left(\mathscr{A}_h+\mathscr{A}_h^T\right)\mathbf{e}_v\\ &+\frac{K}{2}\left(e_\theta\mathscr{L}_h\mathbf{e}_v+\mathbf{e}_v^T\mathscr{L}_h^T\mathbf{e}_\theta\right)+\frac{1}{2}\sum_{i=1}^{n}\left[e_{x_i}\ e_{w_i}\right]Q_i\left[e_{x_i}\ e_{w_i}\right]^T\\ &+\frac{1}{2}\sum_{i=1}^{n}\alpha_i\left(\left[e_{x_i}\ e_{w_i}\right]P_iH_i(e_{v_i}-e_{x_i})+(e_{v_i}-e_{x_i})^TH_i^TP_i\left[e_{x_i}\ e_{w_i}\right]^T\right),\end{aligned} \tag{28}$$

where

$$Q_i=P_i\left[\begin{array}{cc}k_{2,i} & k_{3,i}\\ 0 & 0\end{array}\right]+\left[\begin{array}{cc}k_{2,i} & k_{3,i}\\ 0 & 0\end{array}\right]^TP_i,\ H_i=\left[\begin{array}{c}k_{1,i}\\ \frac{1}{T_{w_i}}\end{array}\right]. \tag{29}$$

Taking into account (26)-(27), it follows that

$$\begin{aligned}\rho_i\left(k_{1,i}-\frac{1}{T_{w_i}^2}\nu_i\right)=1,\ Q_i=2\rho_i\left[\begin{array}{cc}k_{2,i} & -\frac{k_{2,i}}{T_{w_i}}\nu_i\\ -\frac{k_{2,i}}{T_{w_i}}\nu_i & \frac{\nu_i^2}{T_{w_i}^2}k_{2,i}\end{array}\right],\\ P_iH_i=\left[1\ \ 0\ \ \right]^T.\end{aligned} \tag{30}$$

Therefore, considering (30), $\dot{\mathscr{V}}(x)$ can be rewritten as

$$\begin{aligned}\dot{\mathscr{V}}(\mathbf{x_{cl}})=&-\beta\frac{K}{2}\mathbf{e}_v^T\left(\mathscr{A}_h+\mathscr{A}_h^T\right)\mathbf{e}_v+\frac{1}{2}\sum_{i=1}^{n}\left[e_{x_i}\ e_{w_i}\right]Q_i\left[e_{x_i}\ e_{w_i}\right]^T\\ &-K\mathbf{e}_\theta^T\,[\eta]\,\mathbf{e}_\theta-(\mathbf{e}_v-\mathbf{e}_x)^T\,[\alpha]\,(\mathbf{e}_v-\mathbf{e}_x).\end{aligned} \tag{31}$$

In can be shown that $trace(Q_i)=2\rho_ik_{2,i}(1+\frac{\nu_i^2}{T_{w_i}^2})<0$ and $det(Q_i)=0$; therefore, $Q_i\preceq 0$. Since $Q_i\preceq 0$, $[\alpha]\succ 0$, and $\mathscr{A}+\mathscr{A}^T\succeq 0$, $\dot{\mathscr{V}}(\mathbf{x_{cl}})\leq 0$. Now, let define $\mathscr{S}=\left\{\mathbf{x_{cl}}(t):\ \dot{\mathscr{V}}(\mathbf{x_{cl}})=\mathbf{0}_n\right\}$. If $\dot{\mathscr{V}}(x)=0$, then $\mathbf{e}_v=\mathbf{e}_x$, $\mathbf{e}_\theta=\mathbf{0}_n$, $\mathscr{A}\mathbf{e}_v=0$, and $\left[e_{x_i}\ e_{w_i}\right]^T\in ker(Q_i)$, $i\in\mathscr{V}(\mathscr{G})$. The null-space of Q_i is characterized as $e_{x_i}=T_{w_i}^{-1}\nu_ie_{w_i}$. Taking into account $\mathscr{S}$, the closed-loop trajectories in (11) imply $\mathscr{L}_h\mathbf{e}_v=\mathbf{0}_n$. Therefore, $\mathbf{e}_v=\mathbf{e}_x=\mathbf{0}_n$ and $\mathbf{e}_w=\mathbf{0}_n$. Thus, the only solution that can stay identically in $\mathscr{S}$ is $\mathbf{x_{cl}}(t)=\mathbf{0}_{4n}$. Therefore, the origin in (11) is the globally asymptotically stable. As a result, $\mathbf{A_{cl}}$ in (12) is Hurwitz.

References

[1] A. Maknouninejad and Z. Qu, "Realizing unified microgrid voltage profile and loss minimization: A cooperative distributed optimization and control approach," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1621–1630, 2014.

[2] A. Gusrialdi, Z. Qu, and M. A. Simaan, "Distributed scheduling and cooperative control for charging of electric vehicles at highway service stations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2713–2727, 2017.

[3] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[4] Wei Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proceedings of the 2005, American Control Conference, 2005.*, vol. 3, 2005, pp. 1859–1864.

[5] M. S. Sadabadi, "A distributed control strategy for parallel DC-DC converters," *IEEE Control Systems Letters*, vol. 5, no. 4, pp. 1231–1236, Oct. 2021.

[6] T. Pultarova, "Ukraine grid hack is wake-up call for network operators [news briefing]," *Eng. Technol.*, vol. 11, no. 1, pp. 12–13, 2005.

[7] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 90–104, 2012.

[8] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "Distributed fault detection and isolation resilient to network model uncertainties," *IEEE Transactions on Cybernetics*, vol. 44, no. 11, pp. 2024–2037, 2014.

[9] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, p. 1495–1508, Jul. 2011.

[10] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 4, pp. 766–781, 2013.

[11] A. Gusrialdi, Z. Qu, and M. A. Simaan, "Competitive interaction design of cooperative systems against attacks," *IEEE Trans. Automatic Control*, vol. 63, no. 9, pp. 3159–3166, Sept. 2018.

[12] H. Dong, C. Li, and Y. Zhang, "Resilient consensus of multi-agent systems against malicious data injections," *Journal of the Franklin Institute*, vol. 357, no. 4, pp. 2217–2231, 2020.

[13] S. Zuo and D. Yue, "Resilient output formation containment of heterogeneous multigroup systems against unbounded attacks," *IEEE Transactions on Cybernetics*, pp. 1–9, 2020.

[14] Z. Li, Z. Li, and Y. Liu, "Resilient control design of the third-order discrete-time connected vehicle systems against cyber-attacks," *IEEE Access*, vol. 8, pp. 157 470–157 481, 2020.

[15] H. Modares, B. Kiumarsi, F. L. Lewis, F. Ferrese, and A. Davoudi, "Resilient and robust synchronization of multiagent systems under attacks on sensors and actuators," *IEEE Transactions on Cybernetics*, vol. 50, no. 3, pp. 1240–1250, Mar. 2020.

[16] G. De La Torre, T. Yucelen, and J. D. Peterson, "Resilient networked multiagent systems: A distributed adaptive control approachy," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 5367–5372.

[17] G. D. L. Torre and T. Yucelen, "Adaptive architectures for resilient control of networked multiagent systems in the presence of misbehaving agents," *International Journal of Control*, vol. 91, no. 3, pp. 495–507, 2018.

[18] E. Yildirim, S. B. Sarsilmaz, A. T. Koru, and T. Yucelen, "On control of multiagent systems in the presence of a misbehaving agent," *IEEE Control Systems Letters*, vol. 4, no. 2, pp. 456–461, Apr. 2020.

[19] W. Zeng and M. Chow, "Resilient distributed control in the presence of misbehaving agents in networked control systems," *IEEE Transactions on Cybernetics*, vol. 44, no. 11, pp. 2038–2049, 2014.

[20] H. Zhang, F. L. Lewis, and A. Das, "Optimal design for synchronization of cooperative systems: State feedback, observer and output feedback," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1948–1952, Aug. 2011.

[21] Z. Qu, *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Springer-Verlag, 2009.

[22] A. Darabseh and N. M. Freris, "A software-defined architecture for control of IoT cyberphysical systems," *Cluster Computing*, vol. 22, no. 4, pp. 1107–1122, 2019.