Competitive Interaction Design of Cooperative Systems Against Attacks

Azwirman Gusrialdi, *Member, IEEE*, Zhihua Qu, *Fellow, IEEE*, and Marwan A. Simaan, *Life Fellow, IEEE*

***Abstract*—This technical note proposes a resilient cooperative control design for networked cooperative systems when subjected to external attacks. The systems considered in this paper can have any information topology described by a leader-follower digraph. A potential attack on such systems consists of unknown bounded signals generated from any linear or nonlinear finite-L_2-gain exogenous dynamical system and injected distributively into nodes of the system's network. The purpose of the attack is to destabilize the consensus dynamics by intercepting the system's communication network and corrupting its local state feedback. The proposed resilient control design consists of introducing a virtual system with hidden network such that the overall system consisting of the original consensus system, the virtual system, and the attack dynamics is stable without requiring any information about the locations or nature of the attack. This is accomplished by utilizing the concept of competitive interaction to provide explicit design criteria for the hidden network of the virtual system to interact with the original system. A graph theoretical approach and a Lyapunov direct method are used to analyze the overall system and show that the proposed design ensures stability of the overall system and preserves the consensus of the original system. An example, which includes several scenarios, is used to illustrate the results.**

***Index Terms*—Cooperative systems, leader-following consensus, resilient control design.**

I. Introduction

Advances in wireless communication networks allow for the implementation of cooperative control in cyber-physical systems so as to achieve better robustness, scalability, and efficiency while requiring only local information exchange. Cooperative control has been applied to various problems including formation control [1], smart grids [2], and analysis of interconnected systems [3]. A common goal for the outputs of a networked system is to reach a consensus [4]. Typical examples of consensuses are the clock speed for wireless sensor networks [5], and the velocity or displacement with respect to a formation for a team of autonomous vehicles [1].

While the use of communication networks in cyber-physical systems is necessary and has many advantages, it has also made the systems more vulnerable to data attacks. As an example, a computer worm called Stuxnet was discovered in 2010 [6], which was designed to attack industrial programmable logic controllers. An another example is data attacks on power transmission networks operated by supervisory control and data acquisition systems [7] that could be used to disrupt the system such as by taking control of electrical facilities.

The leaderless consensus problem in the presence of faulty/misbe having nodes and attacks are considered in [8]–[12]. Graph-theoretical methods are proposed in [8]–[10] to detect, identify, and isolate the faulty nodes. Specifically, it is shown in [8] and [9] that the misbehaving nodes can be correctly identified, as long as the connectivity of the communication graph is sufficiently high. However, the proposed methods are computationally expensive, requiring nonlocal information on the network. A set of local filtering algorithms to mitigate the influence of adversaries, which depends on the number of attacks, is proposed in [10], where each node in the network removes the extreme values with respect to its own value. De La Torre *et al.* [12] propose a distributed adaptive control strategy to ensure that the agents in undirected network reach consensus in the presence of misbehaving agents modeled by exogenous bounded disturbances and interagent uncertainties. The results are extended to the case of directed graph and switching topology in [13]. Moreover, strategies to detect attack and mitigate its affect are proposed in [14] for consensus of cooperative systems with double integrator dynamics and undirected network topology. Game theory provides a natural framework to model and analyze the interaction between the attacker and defender that have conflicting interests in security problems. A game-theory-based consensus learning algorithm under persistent adversaries, where the problem is modeled as a minimax optimization, is proposed in [15]. The work in [16] considers leader-following consensus problem where some of the follower agents are misbehaving. The authors propose resilient distributed control algorithm consisting of four phases (detection, mitigation, identification, and update) to guarantee the tracking of the leader's state. Zhu and Martinez [17] consider the problem of distributed formation control where a group of vehicles is remotely controlled by a network of operators. It is assumed that each vehicle–operator pair is attacked by an adversary who corrupts the commands sent from the operator to the vehicle. A resilient distributed algorithm is proposed that allows the operator to adjust their policy so that the vehicles converge to the desired formation. To summarize, the existing results on resilient control for cooperative system have limitations on communication network topology, type of attacks (attacker's strategy) or number of compromised nodes and cooperative control problem under consideration (that is either leader-follower or leaderless consensus problem).

This paper considers the leader-following consensus (cooperative tracking) problem in the presence of unknown attacks. The attacker may use injections generated from nonlinear/linear dynamics to corrupt the commands sent from the leader to the followers, intercept the communication signals among the followers, or corrupt the state estimates of the nodes by interconnecting with the system's communication network. It

Manuscript received May 25, 2017; revised December 25, 2017; accepted January 8, 2018. Date of publication January 12, 2018; date of current version August 28, 2018. This work was supported in part by the U.S. National Science Foundation under Grant ECCS-1308928 and Grant CCF-0956501, in part by the U.S. Department of Energy awards DE-EE0006340, DE-EE0007327, and DE-EE0007998, in part by the U.S. Department of Transportation award DTRT13-G-UTC51, in part by the L-3 Communication contract 11013I2034, in part by the Leidos Contract P010161530, and in part by Texas Instruments grants. Recommended by Associate Editor H. Lin. *(Corresponding author: Azwirman Gusrialdi.)*

The authors are with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando 32816, USA (e-mail: azwirman.gusrialdi@ucf.edu; qu@ucf.edu; simaan@eecs.ucf.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2018.2793164

is obvious that without protection, the adversary can easily destabilize the consensus dynamics. Note that the leader-following consensus has been applied to the control of a smart grid [2], a transportation system [18], [19], a mobile sensor network [20], and human–swarms interaction [21]. The objective of this paper is to develop a control design method to ensure that the cooperative system remains stable against potential attacks and the consensus is also maintained. To this end, a virtual system with hidden network (which acts as a controller) interconnected with the original consensus network is introduced. The virtual system with hidden network is designed to maintain the stability of the overall system by competitively interacting with the original consensus network. The strategy is based on the competitive interaction concept introduced in [22] in which leaderless networks of undirected graphs and attacks with linear dynamics are investigated and is extended in [23] to the case of strongly connected directed graphs. In this paper, networks of directed graphs and attacks generated from nonlinear dynamics are investigated, and Lyapunov analysis and design methods are presented to provide explicit guidelines on how to interconnect both the original consensus and the hidden networks of the virtual system to make the overall system robust against attacks. In contrast to alternative approaches proposed in the existing literature, the method proposed in this paper can be applied to directed graph and both leaderless and leader-following consensus problems. Moreover, the introduction of hidden network has the advantage of maintaining the robustness of the systems under attacks, before and after the attacks are appropriately identified. Any estimation-based robustification approaches can be applied only after successful identification. The existing approaches (such as high-gain control) could attempt to achieve robustness by overcompensating for potential attacks, whereas the proposed hidden network does not interfere the system during normal operations. Another important feature of the proposed strategy is that the overall system can still maintain its stability even if the hidden network is also subjected to a similar attack.

The remainder of this paper is organized as follows. The problem statement is formulated in Section II. A robustification method to protect cooperative system against unknown and destabilizing attacks based on a virtual system with hidden network is presented and analyzed in Section III. An illustrative example is included in Section IV, and conclusions are drawn in Section V.

II. Problem Statement

Consider a *cooperative* system Σ_s consisting of $n+1$ nodes, where a node labeled by 0 is assigned as a leader and the nodes indexed by $1,\dots,n$ are referred to as followers. The communication structure (information flow) among the nodes is modeled by a directed graph $\mathcal{G}_s$ consisting of a vertex set $\mathcal{V}=\{v_0,\cdots,v_n\}$ and an edge set $\mathcal{E}\subset\mathcal{V}\times\mathcal{V}$. A directed edge $(j,i)\in\mathcal{E}(\mathcal{G}_s)$ denotes that node i can obtain information from node j. The set of neighbors of node i is denoted by $\mathcal{N}_i=\{j|(j,i)\in\mathcal{E}\}$. A directed path from node i to node j is defined as a sequence of adjacent edges $(v_i,v_l),(v_l,v_p),\dots,(v_q,v_j)$. A graph has a *spanning tree* if there is a root node that has a path to every other node.

The cooperative system is a cyber-physical system in the sense that its local communication is represented by $\mathcal{G}_s$ while physical systems at the nodes are heterogeneous and have following dynamics:

$$\dot{y}_i=F_i(y_i,u_i),\quad x_i=M_i(y_i,u_i)$$

where y_i is the internal state of the ith system, u_i is the control, and $x_i\in\Re$ is the output, that is the state of physical variables of node i to be controlled toward a *consensus* (in the sense that the followers track the leader node, i.e., $x_i\to x_0$, $x_0\in\Re$ for $i=1,\cdots,n$). A systematic and modular design is presented in [24] for dealing with networked control of heterogeneous nonlinear/linear systems; in particular, if the individual dynamic systems are input passivity short, then their dynamic behaviors at the network level as well as their network control design can equivalently be studied in terms of the systems of $\dot{x}_i=u_i$. Furthermore, potential attacks could be modeled as injections through local communication networks, and broad implications can be drawn for general passivity-short systems by focusing upon their impacts on cooperative systems of simpler yet equivalent dynamics. Therefore, for the rest of this paper, we will investigate the corresponding equivalent problem in which each node updates its states as follows:

$$\dot{x}_0=0$$
$$\dot{x}_i=a_{i0}(\tilde{x}_{i,0}-\tilde{x}_{i,i})+\sum_{j=1}^{n}a_{ij}(\tilde{x}_{i,j}-\tilde{x}_{i,i}),\ i=1,\dots,n \tag{1}$$

where $\tilde{x}_{i,i}$ and $\tilde{x}_{i,j}$ denote the feedback of x_i and x_j at node i, respectively, and $a_{ij}=1$ if follower node i receives information from node j, including leader node 0 (i.e., $(j,i)\in\mathcal{E}$) and $a_{ij}=0$ otherwise. Note that, when attacks are not present, $\tilde{x}_{i,i}=x_i$ and $\tilde{x}_{i,j}=x_j$. It can be observed from (1) that the leader takes no action in response to the followers' states (though it may receive information about the followers). In this note, it is assumed that graph $\mathcal{G}_s$ contains a spanning tree with the leader node v_0 as the root node. As a result, it is shown that, if there is no attack, consensus is guaranteed [1], i.e., $x_i\to x_0$ for $i=1,\dots,n$.

Since it is not always possible to ensure that all communication networks/channels are secure, the system may be subject to attack. Specifically, the attacker distorts the communication channels (and/or the sensors) by adding exogenous signals to perturb the state feedback of nodes and/or to modify (some of) the neighbors' (including the leader's) information that a specific node receives. Hence, the potentially corrupted feedback at node i can be expressed as

$$\tilde{x}_{i,j}=x_j+\delta_{ij},\ j\in\{\mathcal{N}_i\cup i\} \tag{2}$$

where $\delta_{ij}(x,t)$ denotes the potential injection inserted by the attacker and may be a function of x and time. The attacker may have the full knowledge of the network, i.e., all a_{ij} in (1), as well as have access to both local communication of x_j and measurement of x_i. By inserting injection δ_{ij}, the attacker aims at a destabilizing system (1). In practice, the adversary would have a limited budget to launch an attack and any intelligent attacker would aim at destabilizing the system by inserting a "bounded" injection δ_{ij}. In the event that the attacker does insert an injection of infinite magnitude, the cooperative system appropriately designed can enable each of its nodes to simply reject such injection by removing excessively large values it receives. This defensive mechanism can be incorporated into (1) as follows:

$$a_{ij}=\begin{cases}1 & \text{if } \tilde{x}_{i,j}\in\Omega \text{ and } j\in\mathcal{N}_i\\ 0 & \text{if otherwise}\end{cases} \tag{3}$$

where Ω is a compact set describing all feasible values of state variables. The choice of compact set Ω depends on the operational range of physical variables to be controlled. By performing distributively and real time at each node rejection of bad data as prescribed by (3), cooperative system (1) under potential attacks (2) can be written in a compact form as

$$\dot{x}=Ax+Bx_0+d \tag{4}$$

where $x=[x_1,\dots,x_n]^T$, $d=[d_1,\dots,d_n]^T$ is the attack vector, and $\sum_{j\in\{\mathcal{N}_i\cup i\}}\delta_{ij}=d_i$. Moreover, vector $B=[a_{10},\dots,a_{n0}]^T$ and

matrix A is given by

$$A = \begin{bmatrix} -\sum_{j=0}^{n} a_{1j} & a_{12} & \cdots & a_{1n} \\ a_{21} & -\sum_{j=0}^{n} a_{2j} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \cdots & -\sum_{j=0}^{n} a_{nj} \end{bmatrix}.$$

Without loss of any generality, for the rest of this paper, we consider the following assumptions. In particular, Assumption 1 includes as a special case that injection d is uniformly bounded, i.e., $\|d(x,t)\| \leq \bar{d}$ for all $x \in \Re^n$. Assumption 1 also include typical choices of attack vector d that aim to make the cooperative system unstable, as will be elaborated further in Section IV.

Assumption 1: The exogenous injection δ_{ij} may be unbounded but the system operational set Ω is bounded. Hence, under the thresholding mechanism of (3), the effective injection $d(x,t)$ in (4) is always bounded for any bounded x; in particular, should it satisfy any differential equation of general form

$$\dot{d} = f(d,x) \tag{5}$$

system (5) would have a finite L_2 gain.

Assumption 2: Upon implementing bad-data rejection (3), the leader node remains to be a global reachable node.

Assumption 2 implies that matrix A in (4) is Hurwitz [1]. Should some of the nodes become isolated (from the leader node) because of all their a_{ij} being zero, their stability is maintained but they may not reach the consensus as the rest of connected nodes. Accordingly, these isolated nodes will be excluded from subsequent analysis, and hence Assumption 2 can be made without loss of any generality.

In summary, the objective of this paper is to develop a network-enabled defense mechanism so the cooperative system become robust against all possibly destabilizing attacks in the form of (5).

III. Resilient Network-Level Design

The consensus can be ensured at the network level by introducing a hidden layer Σ_h (shielded from the attacker), that is, a virtual system with hidden network whose number of nodes is equal to n, as shown in Fig. 1, whose state z is designed to maintain the stability of the overall system in the presence of (destabilizing) attacks. To this end, let us consider the following *overall* system, i.e., the interconnection of cooperative systems and a virtual system with hidden network:

$$\begin{aligned} \dot{x} &= Ax + \beta K z + Bx_0 + d \\ \dot{z} &= Hz - \beta G x + \beta D x_0 \end{aligned} \tag{6}$$

where $z \in \Re^n$ is the state of the virtual system and injection d satisfies Assumption 1. The parameter and matrices to be designed are scalar $\beta > 0$, vector D, matrix H, and interconnection matrices K and G. The choice of the same scalar β for matrices K, G, and D is without loss of any generality. In addition, it also facilitates the stability analysis as will be shown later and simplifies the hidden network design.

The hidden layer (which is a virtual layer) could be implemented as an internal signal component at every node of the networked system Σ_s whose security is guaranteed such as in SCADA (supervisory control and data acquisition) system, which makes it difficult (i.e., requires high cost) for the attacker to compromise. Moreover, additional information flow in the hidden layer can be achieved using different communication network/channel (for better security) such as internet technology or software-defined networking [25], which is a recent advancement in cloud computing and network management. It is worth noting that in contrast to the state x of the networked system, the hidden state z of the hidden layer has no physical meaning (e.g., can be transferred by the internet), and thus is less observable to the attacker (i.e., unlikely to be attacked) or it would be difficult for the attacker to associate the information flow in the internet with the measurements of physical variables used by the cooperative systems. In order to further reduce the risk of exposing the additional information flow used by the cooperative system, the defender could also build multiple hidden layers on top of each other. Furthermore, the addition of a hidden network comes at a price of an increased computational and communication expenses. The computational expense is small since simple multiplications and additions are performed distributively at each of the nodes. Communication expense could be small as well since the hidden network is not necessarily physical (as the primary network of cooperative systems), and hence could be done using one of the standard network technologies (such as wireless or internet).

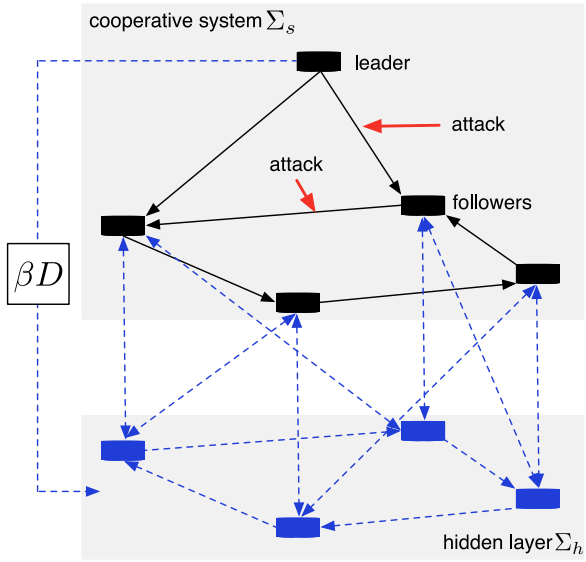

Fig. 1. Resilient design of cooperative system: Interconnection of the cooperative system and the virtual system with hidden network. The blue nodes denote hidden nodes of the virtual system (Σ_h). The solid black lines denote the information flow (communication topology) from the leader to the followers and among the followers. The dashed lines represent the interconnection between cooperative system (Σ_s) layer and virtual system (hidden layer). The red-arrowed links are examples where attacks may occur (unknowingly).

Sparsity and distributed implementation of the hidden layer are established in Section III-A, and these features keep the computational and communication expenses reasonable. The nodes in the hidden layer serve as *synthetic anchors* by which the overall system is shown to be robust against attacks and maintain resilient operation in Section III-A.

A. Nominal Design

First, observe that since matrix A is Hurwitz, there exists a symmetric matrix $P_s > 0$ such that $A^T P_s + P_s A < 0$. To be more precise, the closed form for computing matrix $P_s > 0$ is given by [26]

$$\begin{aligned} P_s &= \text{diag}(q_i/p_i) \\ p = [p_1, \ldots, p_n]^T &= (-A)^{-1}\mathbf{1} \\ q = [q_1, \ldots, q_n]^T &= (-A)^{-T}\mathbf{1}. \end{aligned} \tag{7}$$

As a first step of the design, matrix H in (6) is chosen by the designer to be Hurwitz and sparse. Therefore, there exists a matrix $P_h > 0$

satisfying $H^T P_h + P_h H < 0$. For simplicity, matrix H can be chosen to be equal to matrix A. In general, matrix H is chosen so that $(-H)$ is a sparse nonsingular M-matrix [26], which also allows matrix P_h to be computed similarly as in (7). Note that the design of the virtual system with hidden network Σ_h has to be done in such a way that, when connected with the cooperative system, its presence does not interfere with the (steady state) operation of the original cooperative system Σ_s (since Σ_s has physical meaning, whereas Σ_h is completely virtual). The following lemma presents topological conditions on vector D and interconnection matrices K and G so that, in the absence of attacks, the state x converges to the leader state x_0.

Lemma 1: Consider the overall system (6) with $d = 0$. Then, the consensus value of x of the system (6) is the same as that of system (1) (that is $x \to \mathbf{1}x_0$ as $t \to \infty$) provided that D, K, and G are chosen to satisfy

$$K^T P_s = P_h G$$
$$G\mathbf{1} = D. \tag{8}$$

Proof: Let us define the transformed state $\tilde{x} = x - \mathbf{1}x_0$. Using the transformed state $\tilde{x}$ and assuming $d = 0$, the dynamical system (6) can be written as

$$\dot{\tilde{x}} = A\tilde{x} + \beta K z + A\mathbf{1}x_0 + Bx_0$$
$$\dot{z} = Hz - \beta G\tilde{x} - \beta(G\mathbf{1} - D)x_0.$$

Substituting $G\mathbf{1} = D$ and noting that $A\mathbf{1}x_0 = -Bx_0$ yields

$$\dot{\tilde{x}} = A\tilde{x} + \beta K z$$
$$\dot{z} = Hz - \beta G\tilde{x}. \tag{9}$$

Next, consider the Lyapunov function candidate

$$V(\tilde{x}, z) = \tilde{x}^T P_s \tilde{x} + z^T P_h z.$$

Taking the time derivative of V along the trajectory of (9) yields

$$\begin{aligned}\dot{V} =& \ 2\left(A\tilde{x} + \beta K z\right)^T P_s \tilde{x} + 2\left(Hz - \beta G\tilde{x}\right)^T P_h z \\ =& \ -\tilde{x}^T (A^T P_s + P_s A)\tilde{x} - z^T (H^T P_h + P_h H) z \\ & + 2\beta z^T (K^T P_s - P_h G)\tilde{x}.\end{aligned}$$

It follows from (8) and the fact that matrices A and H are Hurwitz that

$$\dot{V} = -x^T Q_s x - z^T Q_h z < 0$$

where

$$Q_s = A^T P_s + P_s A, \quad Q_h = H^T P_h + P_h H.$$

Hence, it can be concluded that in the absence of attacks ($d = 0$), the equilibrium of (9) is stable and thus $x \to \mathbf{1}x_0$ as $t \to \infty$. ■

Note that Lemma 1 includes the nominal designs of leaderless consensus [23] as the special case of $D = 0$. Lemma 1 and its conditions in (8) provide the design of the hidden layer and its interconnections with the cooperative system. To summarize, the design process consists of the following simple steps:

Hidden layer itself: Matrix H can be chosen to be any sparse Hurwitz matrix.

Interconnection matrix K: Interconnection matrix K can also be chosen to be any invertible sparse matrix.

Interconnection matrix G: Once K is chosen, G is given by the second equation in (8) and can be computed according to

$$G = P_h^{-1} K^T P_s \tag{10}$$

where P_s and P_h are computed from (7).

Vector D: Vector D can be computed as $D = G\mathbf{1}$.

Remark 3.1: As discussed in the beginning of Section III-A, matrix H can be chosen to be sparse and Hurwitz so that P_h is a diagonal matrix. Therefore, by choosing interconnection matrix K to be sparse, it can be observed from (10) that matrix G will also be sparse whose structure is similar to that of K^T. Similarly, it can be observed from (10) that matrix K can also be chosen to be sparse and invertible such that vector D is also sparse.

Remark 3.2: As revealed in Remark 3.1, the matrices H, K, G, and D can be chosen/designed to be sparse. Hence, the hidden layer and its interconnections with the networked system can be realized through networking, and their implementation is all distributed.

Remark 3.3: The design of hidden network, such as matrices P_s, P_h, and G, might be performed in a centralized manner. Note that the matrices H and K can be designed distributively since there are no connectivity requirements on their network topology and their Hurwitz and nonsingularity conditions can be guaranteed distributively using Gershgorin theorem, see for example [3]. In addition, if the network topology of the follower nodes in both the cooperative system and hidden layer are given by strongly connected digraphs and each node knows its out-neighbors (i.e., a set of neighbors to whom that node sends information), then the hidden network can also be designed in a distributed fashion. Specifically, matrix P_s (similarly P_h) in (7) can be computed by solving a set of linear equations distributively. To this end, from (7), we can write

$$(-A)p = \mathbf{1}. \tag{11}$$

Hence, the vector p can be computed from linear equations (11) distributively, for example, using the method proposed in [27]. Similarly, for q in (7), we can also write

$$(-A)^T q = \mathbf{1}$$

which can be solved distributively similar to computation of p since each node knows its out-neighbors (i.e., node i knows the ith row of A^T). Moreover, since matrices P_s and P_h can be chosen to be diagonal and using the information of out-neighbors available to each node, matrix G in (10) can also be computed distributively. If matrix G is designed in a distributed manner, it then follows that vector D can also be computed distributively.

B. Resilient Control Design

Based on the results developed in Section III-A, consider the following overall system with injection vector d satisfying Assumption 1:

$$\dot{x} = Ax + \beta K z + Bx_0 + d$$
$$\dot{z} = Hz - \beta P_h^{-1} K^T P_s x + \beta(P_h^{-1} K^T P_s \mathbf{1})x_0$$
$$\dot{d} = f(d, x) \tag{12}$$

where matrices H, K, P_s, and P_h are chosen according to the procedure described in Section III-A. The following theorem shows the stability of overall systems (12) against attacks.

Theorem 1: For all sufficiently large values of $\beta > 0$, the cooperative system (12) is uniformly bounded for any injections d satisfying Assumption 1. Furthermore, by increasing β, x is forced to converge to an arbitrarily small neighborhood around $x_0\mathbf{1}$. Specifically

$$\lim_{t\to\infty} x(t) = x_0\mathbf{1} - (A + \beta^2 K H^{-1} P_h^{-1} K^T P_s)^{-1} d^e \tag{13}$$

where d^e is the steady state of the injection d.

Proof: As shown in the proof of Lemma 1 using the transformed state $\tilde{x} = x - x_0\mathbf{1}$, system (12) can be rewritten as

$$\begin{aligned}\dot{\tilde{x}} &= A\tilde{x} + \beta K z + d\\ \dot{z} &= Hz - \beta P_h^{-1} K^T P_s \tilde{x}.\end{aligned} \tag{14}$$

Equilibria of the above-mentioned system satisfy

$$\begin{aligned}0 &= A\tilde{x}^e + \beta K z^e + d^e\\ 0 &= Hz^e - \beta P_h^{-1} K^T P_s \tilde{x}^e.\end{aligned} \tag{15}$$

Defining error vectors $\overline{x} = x - x^e$ and $\overline{z} = z - z^e$, we have the error system as

$$\begin{aligned}\dot{\overline{x}} &= A\overline{x} + \beta K\overline{z} + (d - d^e)\\ \dot{\overline{z}} &= H\overline{z} - \beta P_h^{-1} K^T P_s \overline{x}\end{aligned} \tag{16}$$

where d^e is defined by the solution to equation $0 = f(d^e, x^e)$. Moreover, by the Lyapunov converse theorem [28], it is known that injection d satisfying Assumption 1 has the following property: Given equilibrium x^e from which equilibrium d^e of dynamics (5) satisfies $0 = f(d^e, x^e)$, there exists a Lyapunov function $V_d(d - d^e)$ such that

$$\begin{cases}\gamma_1 \|d - d^e\|^2 \le V_d(d - d^e) \le \gamma_2 \|d - d^e\|^2\\ \frac{\partial V_d}{\partial d} f(d, x^e) \le -\gamma_3 \|d - d^e\|^2\\ \left\|\frac{\partial V_d}{\partial d}\right\| \le \gamma_4 \|d - d^e\|\end{cases} \tag{17}$$

and that

$$\begin{cases}\|f(d, x) - f(d, x^e)\| \le \gamma_5 \|x - x^e\|\\ \|f(d, x^e) - f(d^e, x^e)\| \le \gamma_6 \|d - d^e\|\end{cases} \tag{18}$$

where $\|\cdot\|$ denotes the standard Euclidean norm, and γ_i for $i = 1, \ldots, 6$ are positive constants and $\gamma_1 \le \gamma_2$.

To show that system (16) together with $\dot{d} = f(d, x)$ is asymptotically stable for all large values of β, we choose a Lyapunov function candidate as

$$V' = \beta \overline{x}^T P_s \overline{x} + \beta \overline{z}^T P_h \overline{z} + V_d(d - d^e) + 2\overline{z}^T P_h K^{-1}(d - d^e) \tag{19}$$

where $V_d(\cdot)$ is the Lyapunov function given in (17). It should be noted that the resilient control designer does not need to know the expression of V_d but simply its existence. It is clear that V' is positive definite for all values of β satisfying

$$\beta > \frac{\lambda^2_{\max}(P_s)}{\gamma_1 \lambda_{\min}(P_s)}.$$

Computing the time derivation of Lyapunov function (19) along the trajectory of (16) yields

$$\begin{aligned}\dot{V}' = \; & -\beta \overline{x}^T Q_s \overline{x} - \beta \overline{z}^T Q_h \overline{z} + \dot{V}_d + 2\overline{z}^T P_h K^{-1}[f(d, x)\\ & - f(d^e, x^e)] + 2\overline{z}^T H^T P_h K^{-1}(d - d^e).\end{aligned} \tag{20}$$

We can rewrite (20) as

$$\begin{aligned}\dot{V}' = \; & -\beta \overline{x}^T Q_s \overline{x} - \beta \overline{z}^T Q_h \overline{z} + \dot{V}_d\\ & + 2\overline{z}^T P_h K^{-1}[f(d, x^e) - f(d^e, x^e)]\\ & + 2\overline{z}^T P_h K^{-1}[f(d, x) - f(d, x^e)]\\ & + 2\overline{z}^T H^T P_h K^{-1}(d - d^e).\end{aligned}$$

Invoking (17) and (18), we have

$$\dot{V}_d \le -\gamma_3 \|d\|^2 + \gamma_4 \gamma_5 \|d - d^e\| \|\overline{x}\|$$

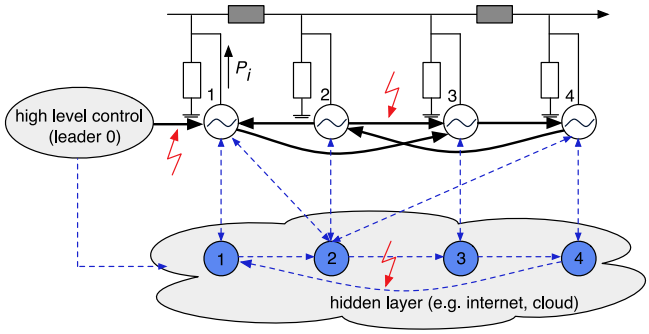

Fig. 2. Interconnection of cooperative systems Σ_s of four PVs and a virtual system with hidden network (hidden layer), which can be implemented using software-defined networking. The solid black lines represent communication topology of the PVs (followers). The blue dashed lines denote the interconnection between the cooperative system layer and the hidden layer. The red solid arrows represent the examples where attacks may occur.

and hence

$$\begin{aligned}\dot{V}' \le \; & -\beta \overline{x}^T Q_s \overline{x} - \beta \overline{z}^T Q_h \overline{z} - \gamma_3 \|d\|^2\\ & + \gamma_4 \gamma_5 \|d - d^e\| \|\overline{x}\| + 2\|\overline{z}\| \|P_h K^{-1}\| \gamma_6 \|d - d^e\|\\ & + 2\|\overline{z}\| \|P_h K^{-1}\| \gamma_5 \|\overline{x}\| + 2\|\overline{z}\| \|H^T P_h K^{-1}\| \|d - d^e\|.\end{aligned}$$

It can be observed that $\dot{V}'$ is negative definite for all large values of β, and thus system (16) together with $\dot{d} = f(d, x)$ is asymptotically stable. Given asymptotic convergence/stability, the ultimate bound in (13) can be established by solving for x^e from (15). ■

Remark 3.4: The analysis of Theorem 1 is based on the robust control framework. A different analysis could be done from a game theoretical point of view by adopting the approach originally presented in [29] for leaderless consensus where the problem is cast as a two-player nonzero-sum differential game since both the attacker and controller have conflicting but not necessarily exactly opposite objectives.

C. Robustness Against Attacks on the Hidden Network

Next, let us consider the case that the virtual system with hidden network, designed according to the procedure described in Section III-A, may also be subject to attacks. To this end, consider the following overall system with injection vectors d and d' satisfying Assumption 1

$$\begin{aligned}\dot{x} &= Ax + \beta K z + Bx_0 + d\\ \dot{z} &= Hz - \beta P_h^{-1} K^T P_s x + \beta(P_h^{-1} K^T P_s \mathbf{1})x_0 + d'\\ \dot{d} &= f(d, x)\\ \dot{d}' &= f'(d', z).\end{aligned} \tag{21}$$

We then have the following results.

Theorem 2: For all sufficiently large values of $\beta > 0$, the cooperative system (21) is uniformly bounded for any injections d and d' satisfying Assumption 1. Furthermore, by increasing β, x is forced to converge to an arbitrarily small neighborhood around $x_0\mathbf{1}$. Specifically

$$\begin{aligned}\lim_{t\to\infty} x(t) = x_0\mathbf{1} + \big[&(A + \beta^2 K H^{-1} P_h^{-1} K^T P_s)^{-1}\\ &(\beta^2 K H^{-1} d'^e - d^e)\big]\end{aligned} \tag{22}$$

where d^e and d'^e are the steady state of the injections d and d', respectively.

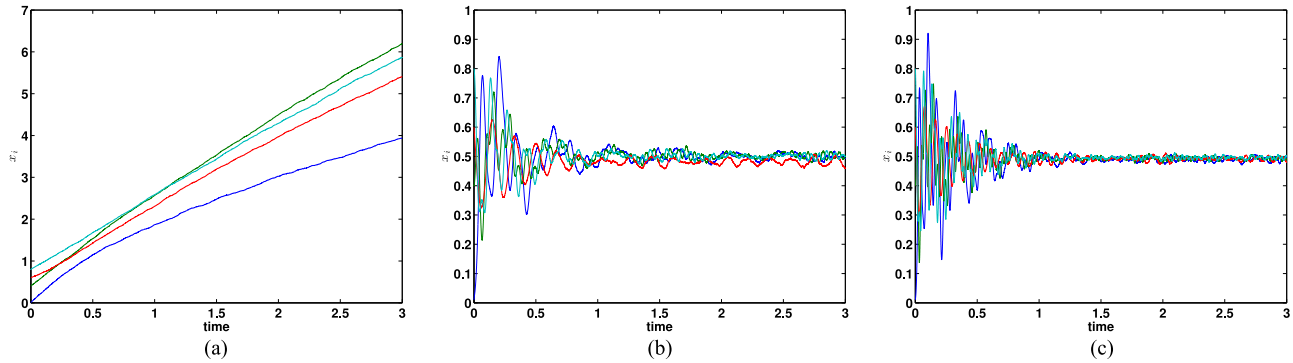

Fig. 3. Trajectories of the followers' state of Σ_s (with $x_0 = 0.5$) under attack with uniformly bounded injection $\|d(t)\| \leq 4$ for the cases: (a) without the hidden network Σ_h, (b) with the hidden network Σ_h and $\beta = 5$, and (c) with the hidden network Σ_h and $\beta = 10$.

Proof: Similar to the proof of Theorem 1, defining error vectors $\overline{x} = x - x^e$ and $\overline{z} = z - z^e$, we have the error system for (21) as

$$\begin{aligned} \dot{\overline{x}} &= A\overline{x} + \beta K\overline{z} + (d - d^e) \\ \dot{\overline{z}} &= H\overline{z} - \beta P_h^{-1} K^T P_s \overline{x} + (d' - d'^e). \end{aligned} \tag{23}$$

Computing the derivative of the following Lyapunov candidate function

$$\begin{aligned} V' = \beta \overline{x}^T P_s \overline{x} + \beta \overline{z}^T P_h \overline{z} + V_d(d - d^e) + V_{d'}(d' - d'^e) \\ + 2\overline{z}^T P_h K^{-1}(d - d^e) - 2\overline{x}^T (K^T)^{-1} P_h (d' - d'^e) \end{aligned}$$

(with $P_s, P_h > 0$ and satisfying $A^T P_s + P_s A < 0, H^T P_h + P_h H < 0$) along (23) and following the similar steps as in the proof of Theorem 1, it can then be shown that the system (23) together with $\dot{d} = f(d, x)$ and $\dot{d}' = f'(d', z)$ is asymptotically stable and (22) can also be similarly obtained. ■

IV. Numerical Example

In order to illustrate the results, we consider the problem of regulating distributively the power output of groups of photovoltaic (PV) generators in a distribution network [2]. The aggregated power output of all PVs in each group can be dispatched and controlled by coordinating the output level of each PV within each group. To this end, one simple strategy is to prescribe certain utilization profile for all PVs in a group, given by the ratio of power output versus available power. The objective is to make all PVs in a group converge to any given utilization profile while only requiring local information from neighboring generators and providing the desired power dispatched from the group. Mathematically, at the equilibrium point, we aim to have

$$\frac{P_1}{P_1^{\max}} = \cdots = \frac{P_n}{P_n^{\max}} = x_0 \tag{24}$$

where P_i and $P_i^{\max}$ denote the active power output and maximum capacity of generator i, respectively. The value x_0 is the desired utilization profile computed by a high-level control (i.e., the leader) such as distribution control center. Note that for the sake of simplicity, we only consider the control of active power since the reactive power can also be addressed in a similar manner. Furthermore, it is assumed that the high-level control has sufficient information, including the parameters and the states of the distribution network, and as a result the desired utilization profile can be computed directly. For the simulation, we consider a group of four PVs and set $x_0 = 0.5$. A leader-following consensus based distributed control law was proposed in [2] to control the power output of PVs so that (24) is achieved. It is shown that the closed-loop system can be expressed as in (1) with $x_i = \frac{P_i}{P_i^{\max}}$. The communication topology of the PVs (which communicate, e.g., via ZigBee) for the simulation is shown in Fig. 2, namely the matrix A and vector B are given by

$$A = \begin{bmatrix} -2 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 1 & -2 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \; B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

When there is no attack (i.e., $d = 0$) and under the communication structure given in Fig. 2, the states of the followers converge to $x \to 0.5\mathbf{1}$. Now assume that there is an attacker that aims to destabilize the system by injecting bounded input d into the system, as shown in (4). We consider two types of injections given by Assumption 1, namely uniformly bounded injection $\|d(t)\|_\infty \leq 4$ and attack dynamics

$$\dot{d} = F_a d + B_a x \tag{25}$$

with

$$F_a = -I, \;\; B_a = \begin{bmatrix} 1 & 2 & 4 & 2 \\ -9 & 4 & 1 & 3 \\ -4 & 3 & 1 & 2 \\ 2 & 1 & 4 & 3 \end{bmatrix} \tag{26}$$

which are unknown to the cooperative system Σ_s. As a result, it can be observed from Figs. 3(a) and 4(a) that without protection (i.e., without the hidden layer) both attacks destabilize the cooperative system. Next, we make the systems robust against the attacks by connecting the cooperative system to a virtual system with hidden network (e.g., using software-defined networking [30]), as illustrated in Fig. 2. For the hidden network, we set matrices H and K as

$$H = \begin{bmatrix} -4 & 0 & 0 & 1 \\ 2 & -5 & 0 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 2 & -4 \end{bmatrix}, \; K = \begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 0 & 1 \\ 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & -2 \end{bmatrix}.$$

Moreover, matrices P_s, P_h, G, and vector D are computed according to the design procedure described in Section III-A. Figs. 3(b) and (c) and 4(b) and (c) show the trajectories of the cooperative systems interconnected with the virtual system in (12) for the values of $\beta = 5$ and $\beta = 10$, respectively. It can be observed that by increasing β, all the followers' state x_i are forced to converge to the neighborhood around the desired utilization profile $x_0 = 0.5$ while the trajectories become more oscillatory during the transient. Therefore, the robustness of the cooperative system against the unknown attacks is ensured. Next, we

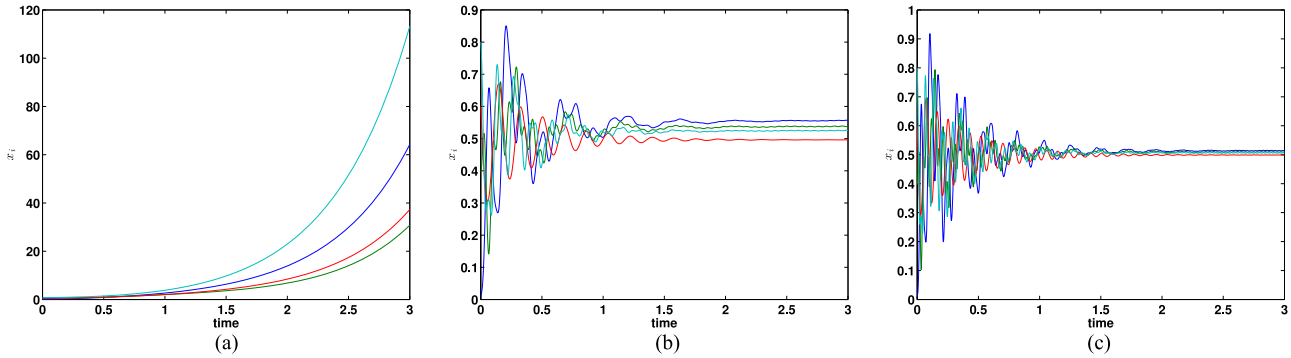

Fig. 4. Trajectories of the followers' state of Σ_s (with $x_0 = 0.5$) under attack in (25) for the cases: (a) without the hidden network Σ_h, (b) with the hidden network Σ_h and $\beta = 5$, and (c) with the hidden network Σ_h and $\beta = 10$.

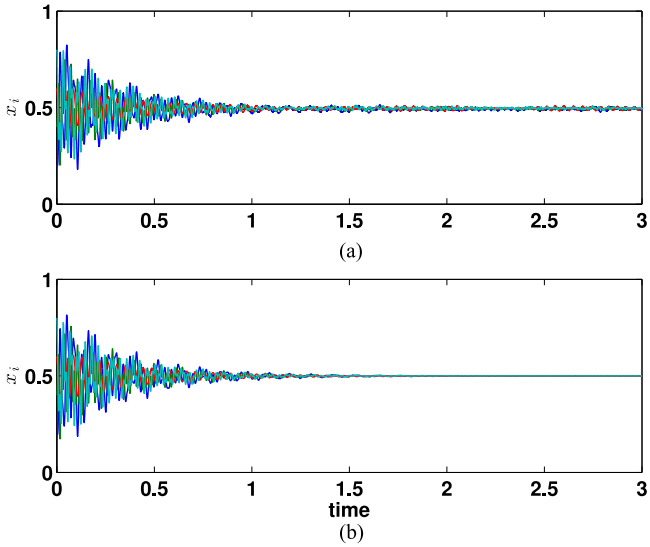

Fig. 5. Trajectories of the followers' state of Σ_s (with $x_0 = 0.5$) under attacks in (21) with $\beta = 20$ and for the cases: (a) uniformly bounded injection $\|d(t)\| \leq 4$, $\|d'(t)\| \leq 4$ and (b) attack dynamics (26), (27).

assume that the virtual system with hidden network is also attacked as described in (21) and the attacks at the hidden layer are also given by either uniformly bounded injection $\|d'(t)\| \leq 4$ or attack dynamics $\dot{d'} = F'_a d' + B'_a z$ with

$$F'_a = -2I, \quad B'_a = \begin{bmatrix} 1 & 3 & 4 & 5 \\ -4 & 4 & 0 & 3 \\ -4 & 9 & 1 & 2 \\ 2 & 6 & 4 & 3 \end{bmatrix}. \tag{27}$$

Using the same hidden network and by setting gain $\beta = 20$, it is shown in Fig. 5 that the stability of the cooperative system Σ_s is still guaranteed despite of the attacks on both networks.

Remark 4.1: As demonstrated in [2], the cooperative systems (1) are robust against loss of a few nodes, e.g., due to a cut power line. In this case, the node becomes decoupled from the cooperative systems, whereas the rest of nodes continue its operation. Similarly, when the adversary spoofs the internet traffic at the site, this means that one of the virtual nodes (together with the corresponding physical node) is out of commission, which also does not affect the operation of the rest of nodes and the hidden network can be adjusted for the rest of the nodes.

V. Conclusion and Future Work

In this technical note, leader-following consensus dynamics of a cooperative system under attack by an adversary are studied in the general setting, where the communication network is represented by a directed graph. The adversary may interact with the consensus network and employ either linear or nonlinear dynamics to alter the local distributed feedback control with the intent of destabilizing the overall system. A Lyapunov-based design is proposed to ensure the stability and consensus against all possible such attacks by the adversary. The approach consists of interconnecting a virtual system with hidden network to the original cooperative system, designed without requiring information about the adversary and whose purpose is to maintain stability of the overall system. Explicit conditions on the interconnection between the two networks are obtained for any consensus network and for any attack model. Future work include the design of optimal control gain β using optimal control framework and also analysis of the proposed approach under control input constraints.

References

[1] Z. Qu, *Cooperative Control of Dynamical Systems*. London, U.K.: Springer-Verlag, 2009.
[2] H. Xin, Z. Qu, J. Seuss, and A. Maknouninejad, "A self-organizing strategy for power flow control of photovoltaic generators in a distribution network," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1462–1473, Aug. 2011.
[3] A. Gusrialdi and Z. Qu, "Distributed estimation of all the eigenvalues and eigenvectors of matrices associated with strongly connected digraphs," *IEEE Control Syst. Lett.*, vol. 1, no. 2, pp. 328–333, Oct. 2017.
[4] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
[5] L. Schenato and F. Fiorentin, "Average TimeSynch: A consensus-based protocol for clock synchronization in wireless sensor networks," *Automatica*, vol. 47, no. 9, pp. 1878–1886, 2011.
[6] N. Falliere, L. Murchu, and E. Chien, "W32.stuxnet dossier," Feb. 2011.
[7] S. Gorman, "Electricity grid in U.S. penetrated by spies," *Wall Street J.*, pp. A1–A2, Apr. 8, 2009. [Online]. Available: https://www.wsj.com/articles/SB123914805204099085
[8] S. Sundaram and C. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Trans. Automat. Control*, vol. 56, no. 7, pp. 1495–1508, Jul. 2011.
[9] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Trans. Automat. Control*, vol. 57, no. 1, pp. 90–104, Jan. 2012.
[10] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 766–781, Apr. 2013.
[11] H. Park and S. Hutchinson, "A distributed robust convergence algorithm for multi-robot systems in the presence of faulty robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Hamburg, Germany, Sep. 28–Oct. 2, 2015, pp. 2980–2985.

[12] G. De La Torre, T. Yucelen, and J. Peterson, "Resilient networked multiagent systems: A distributed adaptive control approach," in *Proc. IEEE Conf. Decis. Control*, Los Angeles, CA, USA, Dec. 15–17, 2014, pp. 5367–5372.

[13] G. D. L. Torre and T. Yucelen, "Adaptive architectures for resilient control of networked multiagent systems in the presence of misbehaving agents," *Int. J. Control*, pp. 1–13, 2017. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/00207179.2017.1286040

[14] Y. Dong, N. Gupta, and N. Chopra, "Content modification attacks on consensus seeking multi-agent system with double-integrator dynamics," *Chaos*, vol. 26, no. 11, 2016, Art. no. 116305.

[15] K. G. Vamvoudakis, L. R. Garcia Carrillo, and J. P. Hespanha, "Learning consensus in adversarial environments," in *Proc. SPIE*, 2013, Paper 87410K.

[16] W. Zeng and M. Y. Chow, "Resilient distributed control in the presence of misbehaving agents in networked control systems," *IEEE Trans. Cybern.*, vol. 44, no. 11, pp. 2038–2049, Nov. 2014.

[17] M. Zhu and S. Martinez, "On attack-resilient distributed formation control in operator-vehicle networks," *SIAM J. Control Optim.*, vol. 52, no. 5, pp. 3176–3202, 2014.

[18] A. Gusrialdi, Z. Qu, and M. A. Simaan, "Distributed scheduling and cooperative control for charging of electric vehicles at highway service stations," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 10, pp. 2713–2727, Oct. 2017.

[19] A. Gusrialdi and Z. Qu, "Analysis of cooperative systems with time delay: Applications to transportation systems," in *Proc. IEEE Conf. Control Appl.*, Buenos Aires, Argentina, Sep. 19–22, 2016, pp. 392–397.

[20] A. Gusrialdi and C. Yu, "Exploiting the use of information to improve coverage performance of robotic sensor networks," *IET Control Theory Appl.*, vol. 8, no. 13, pp. 1270–1283, 2014.

[21] J. P. Croix and M. Egerstedt, "A control Lyapunov function approach to human–swarm interactions," in *Proc. Amer. Control Conf.*, Chicago, IL, USA, Jul. 1–3, 2015, pp. 4368–4373.

[22] B. Gharesifard and T. Basar, "Resilience in consensus dynamics via competitive interconnections," in *Proc. 3rd IFAC Workshop Estimation Control Netw. Syst.*, Santa Barbara, CA, USA, Sep. 14–15, 2012, pp. 234–239.

[23] A. Gusrialdi, Z. Qu, and M. Simaan, "Robust design of cooperative systems against attacks," in *Proc. Amer. Control Conf.*, Portland, OR, USA, Jun. 4–6, 2014, pp. 1456–1462.

[24] Z. Qu and M. A. Simaan, "Modularized design for cooperative control and plug-and-play operation of networked heterogeneous systems," *Automatica*, vol. 50, no. 9, pp. 2405–2414, Sep. 2014.

[25] D. Kreutz, F. M. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proc. IEEE*, vol. 103, no. 1, pp. 14–76, Jan. 2015.

[26] H. Zhang, Z. Li, Z. Qu, and F. L. Lewis, "On constructing Lyapunov functions for multi-agent systems," *Automatica*, vol. 58, pp. 39–42, 2015.

[27] S. Mou, J. Liu, and A. S. Morse, "A distributed algorithm for solving a linear algebraic equation," *IEEE Trans. Automat. Control*, vol. 60, no. 11, pp. 2863–2878, Nov. 2015.

[28] H. K. Khalil, *Nonlinear Systems*, 3rd ed. London, U.K.: Prentice-Hall, 2002.

[29] A. Gusrialdi, Z. Qu, and M. Simaan, "Game theoretical designs of resilient cooperative systems," in *Proc. Eur. Control Conf.*, Linz, Austria, Jul. 15–17, 2015, pp. 1699–1705.

[30] D. Jin *et al.*, "Toward a cyber resilient and secure microgrid using software-defined networking," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2494–2504, Sep. 2017.