Attack Resilient Interconnected Second Order Systems: A Game-Theoretic Approach

Mohammad Pirani, Joshua A. Taylor, and Bruno Sinopoli

***Abstract*— This paper studies the resilience of second-order networked dynamical systems to strategic attacks. We discuss two widely used control laws, which have applications in power networks and formation control of autonomous agents. In the first control law, each agent receives pure velocity feedback from its neighbor. In the second control law, each agent receives its velocity relative to its neighbors. The attacker selects a subset of nodes in which to inject a signal, and its objective is to maximize the $\mathcal{H}_2$ norm of the system from the attack signal to the output. The defender improves the resilience of the system by adding self-feedback loops to certain nodes of the network with the objective of minimizing the system's $\mathcal{H}_2$ norm. Their decisions comprise a strategic game. Graph-theoretic necessary and sufficient conditions for the existence of Nash equilibria are presented. In the case of no Nash equilibrium, a Stackelberg game is discussed, and the optimal solution when the defender acts as the leader is characterized. For the case of a single attacked node and a single defense node, it is shown that the optimal location of the defense node for each of the control laws is determined by a specific network centrality measure. The extension of the game to the case of multiple attacked and defense nodes is also addressed.**

I. Introduction

A. Motivation

The resilience of cyber-physical systems to strategic attacks is one of the primary concerns in the design level and real-time operation of interconnected systems. Examples of such systems include power networks, water and gas networks, and transportation systems. A subtle difference between faults and attacks is that in the latter, the attacker uses knowledge of vulnerabilities to maximize its effect and/or minimize its visibility or effort to attack. The defender thus has to adopt an intelligent strategy to counter the attacker. One approach to modeling interactions between intelligent attackers and defenders is via game theory.

B. Related Work

Security and resilience of cyber-physical systems from the game-theoretic standpoint has attracted attention in the past decade; see [1]–[6] and references therein. The notion of *games-in-games* in cyber-physical systems reflects two interconnected games, one in the cyber layer and the other in the physical layer, for which the payoff of each game affects the result of the other one [7]. Some approaches discussed appropriate game strategies, e.g., Nash or Stackelberg, based on the type of adversarial behavior (active or passive) [3], [8]. The evolution of networked systems are modeled as cooperative games [9] and the resilience of these games to adversarial actions and/or communication failures are investigated [10], [11]. There is a large literature on the security of first and second order systems [12]–[14]. To date, no approach uses game theory to model the actions of intelligent attackers and defenders in second order systems.

M. Pirani and J. A. Taylor are with the Department of Electrical and Computer Engineering, University of Toronto. E-mail: {pirani}@kth.se, {josh.taylor}@utoronto.ca. Bruno Sinopoli is with the Department of Electrical and Systems Engineering, Washington University in St. Louis E-mail: {bsinopoli}@wustl.edu.

C. Contributions

The contributions of this paper are as follows:

- We discuss an attacker-defender game on the resilience of two canonical forms of second order systems. The attacker targets a set of nodes in the network to maximize the system $\mathcal{H}_2$ norm from the attack signals to the output, while the defender chooses a set of nodes (to install feedback control) in order to minimize this system norm (or mitigate the effect of the attack).
- Necessary and Sufficient conditions for the existence of Nash equilibrium (NE) for the game for each of the two second-order dynamics is discussed (Propositions 2 and 3). For the cases where there is no NE, a Stackelberg game is discussed when the defender acts as the game leader (Theorems 1 and 3 and Corollary 1).
- For the case of a single attacked node and a single defense node, it is shown that the optimal location of the defense node in the network for each of the second order systems introduces a specific network centrality measure (Remark 3).
- The extension of the game to the case of multiple attacked and defense nodes is also addressed (Theorems 2, and 4).[1]

It worth noting that for resilient distributed control algorithms proposed in the literature, a large level of network connectivity is required to bypass the effects of malicious actions [14], [15]. However, in many real-world applications, e.g., power systems, the underlying topology is designed and can not be changed. From this view, the defense mechanism proposed in this paper has an advantage compared to the previous methods in the sense that it does not rely on the connectivity level of the underlying network.

II. Graph Theory

We use $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ to denote an undirected graph where $\mathcal{V}$ is the set of vertices (or nodes) and $\mathcal{E} \subseteq \{\{v_i, v_j\} | v_i, v_j \in \mathcal{V}, i \neq j\}$ is the set of undirected edges, where $e = \{v_i, v_j\} \in \mathcal{E}$ if an only if there exists an undirected edge between v_i and

[1] Proofs of all the results of this paper are placed in the Appendix.

v_j. Let $n = |\mathcal{V}|$. The adjacency matrix of $\mathcal{G}$ is denoted A, where $A_{ij} = 1$ if there is an edge between v_j and v_i in $\mathcal{G}$ and zero otherwise. The *neighbors* of vertex $v_i \in \mathcal{V}$ in the graph $\mathcal{G}$ are denoted by the set $\mathcal{N}_i = \{v_j \in \mathcal{V} \mid \{v_j, v_i\} \in \mathcal{E}\}$. We define the degree for node v_i as $d_i = \sum_{v_j \in \mathcal{N}_i} A_{ij}$. The Laplacian matrix of an undirected graph is denoted by $L = D - A$, where $D = \text{diag}(d_1, d_2, ..., d_n)$. We use $\mathbf{e}_i$ to indicate the i-th vector of the canonical basis. The eccentricity $\epsilon(v)$ of a vertex v in a connected graph $\mathcal{G}$ is the maximum graph distance between v and any other vertex $u \in \mathcal{G}$. The center of a graph is a set of vertices with minimum eccentricity. The *effective resistance* between a pair of nodes i and j, denoted R_{ij}, is the electrical resistance measured across nodes i and j when the network represents an electrical circuit where each edge e has electrical conductance w_e [16]. The effective eccentricity $\epsilon_f(v)$ of a vertex v in a connected graph $\mathcal{G}$ is the maximum graph effective resistance between v and any other vertex u of $\mathcal{G}$. The *effective center* of a graph is a set of vertices with minimum effective eccentricity. A degree central node in the network is the node with the largest degree.

III. System Model and Preliminaries

Consider a network of agents $\mathcal{V}$ where each agent follows second-order dynamics

$$\begin{aligned} \dot{x}_i(t) &= v_i(t), \\ \dot{v}_i(t) &= u_i(t) + w_i(t), \end{aligned} \tag{1}$$

where $x_i(t)$ and $v_i(t)$ represent position (or phase) and velocity (or frequency), respectively. $u_i(t)$ and $w_i(t)$ are the control input and additive disturbance to the dynamics. The control policy can be either of the following two

$$u_i = -\sum_{j \in \mathcal{N}_i} a_{ij}(x_i - x_j) - (b_i + b_0)v_i. \tag{2a}$$

$$u_i = -\sum_{j \in \mathcal{N}_i} a_{ij}(x_i - x_j) - \sum_{j \in \mathcal{N}_i} b_{ij}(v_i - v_j) - a_0 x_i - b_0 v_i. \tag{2b}$$

Here a_{ij}, b_{ij}, a_0 and b_0 are nonnegative control gains. Control law (2a) uses the *relative position* and *absolute velocity* as feedbacks whereas (2b) uses both relative position and velocity as control feedbacks. To simplify our analysis, we assume that $a_{ij} = b_{ij} = 1$ and $a_0 = b_0 = k > 0$, where k is called the defender's control gain.[2] Note that all of the analysis in this paper can be readily extended to the weighted case. Control laws (2a) and (2b) are canonical forms of well-known second-order systems. In particular, (2a) is the linearized swing equation for a network of power generators [17], [18], and (2b) describes the formation control of autonomous agents, e.g., a platoon of connected vehicles [19].

[2]This parameter is private and only known by the system designer.

A. Attack Model

Let $\mathfrak{F}$ denote the set of nodes under attack. The state of a node which is under attack evolves as

$$\begin{aligned} \dot{x}_i(t) &= v_i(t) + \zeta_{1i}(t), \\ \dot{v}_i(t) &= u_i(t) + w_i(t) + \zeta_{2i}(t), \quad i \in \mathfrak{F}, \end{aligned} \tag{3}$$

where $\zeta_{1i}(t)$ and $\zeta_{2i}(t)$ are the effects of attack signals on the first and the second states, respectively. In vector form, (3) is given by

$$\dot{\boldsymbol{X}} = A\boldsymbol{X} + B_1\boldsymbol{w}(t) + B_2\boldsymbol{\zeta}(t), \tag{4}$$

where $\boldsymbol{X} = [\boldsymbol{x} \quad \dot{\boldsymbol{x}}]^T$, $\boldsymbol{\zeta} = [\boldsymbol{\zeta}_1 \quad \boldsymbol{\zeta}_2]^T$. Depending on whether control law (2a) or (2b) is in place, the matrix A respectively takes on the form

$$A = \begin{bmatrix} \mathbf{0}_n & I_n \\ -L & -H \end{bmatrix}, \text{ or } A = \begin{bmatrix} \mathbf{0}_n & I_n \\ -\bar{L} & -\bar{L} \end{bmatrix}, \tag{5}$$

where $\bar{L} \triangleq L + kD_y$ and $H \triangleq I_n + kD_y$ where $D_y = \text{diag}(\boldsymbol{y})$. $\boldsymbol{y}$ is a binary vector, i.e., $y_i \in \{0, 1\}$, whose i-th element is one if node i has a self feedback and zero if it does not. We assume that D_y has at least one nonzero diagonal element so that $\bar{L}$ is non-singular [20]. Here $B_1 = [\mathbf{0} \quad I_n]^T$ and $B_2 = I_2 \otimes F$, since we assume that if node i is under attack, then its both states are affected by the attack signal. Matrix B_2 encodes the decisions of the attacker. The i-th row of F has a single 1 if node i is affected by the attack, and all zeros otherwise. The set of nodes under attack and the set of nodes with feedback (defense nodes) are denoted by $\mathfrak{F}$ and $\mathfrak{D}$, respectively. An example of attacker and defender actions on a networked system is schematically shown in Fig. 1 (a).

B. Attacker-Defender Game

Because we do not have a priori knowledge of the frequency contents of the attack signal, we must choose a system norm which captures the average impact of all frequencies of the attack input. We therefore choose system $\mathcal{H}_2$ norm, which is widely used to measure the level of coherence in synchronization of coupled oscillators [21], [22]. We first calculate the $\mathcal{H}_2$ norm of (4).

Proposition 1: The $\mathcal{H}_2$ norms of (4) from the attack signal $\boldsymbol{\zeta}(t)$ to output $\boldsymbol{y} = \dot{\boldsymbol{x}}$ for control laws (2a) and (2b) are

$$\begin{aligned} ||G_1||_2^2 &= \frac{1}{2}\sum_{i \in \mathfrak{F}} H_{ii}^{-1} d_i + \frac{1}{2}\sum_{i \in \mathfrak{F}} H_{ii}^{-1}, \\ ||G_2||_2^2 &= \frac{1}{2} f + \frac{1}{2}\sum_{i \in \mathfrak{F}} \bar{L}_{ii}^{-1}, \end{aligned} \tag{6}$$

where f is the number of attacked nodes, G_1 and G_2 are transfer functions of (4) from $\boldsymbol{\zeta}(t)$ to $\dot{\boldsymbol{x}}$ for control laws (2a) and (2b), respectively. H_{ii}^{-1} and $\bar{L}_{ii}^{-1}$ are the i-th diagonal elements of H^{-1} and $\bar{L}^{-1}$, respectively.

Now, we formally define the attacker-defender game.

*Definition 1 (**Attacker-Defender Game**):* The attacker chooses a set of f nodes to attack, $\mathfrak{F} \subseteq \mathcal{V}$, in order to maximize the $\mathcal{H}_2$ norm from the attack signal $\boldsymbol{\zeta}(t)$ to the output $\boldsymbol{y} = \dot{\boldsymbol{x}}$. The defender places local feedback control

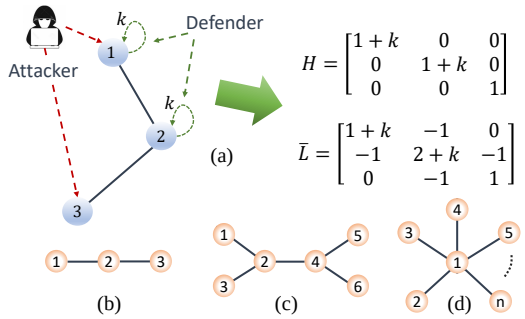

Fig. 1: (a) An example of an attacker-defender game and matrices H and $\bar{L}$, (b) graph topology discussed in example 1, (c) a graph structure where $\Delta_1 = \Delta_2$, thus does not admit NE, (d) star graph admits largest threshold for k over all connected graphs as $\Delta_1 = n-1$ and $\Delta_2 = 1$.

at f nodes, $\mathfrak{D} \subseteq \mathcal{V}$, to minimize the system $\mathcal{H}_2$ norm.[3] The result is a zero-sum game in which the payoff, based on (6), is given by

$$J_1(\mathfrak{F},\mathfrak{D}) = \frac{1}{2}\sum_{i\in\mathfrak{F}} H_{ii}^{-1} d_i + \frac{1}{2}\sum_{i\in\mathfrak{F}} H_{ii}^{-1},$$
$$J_2(\mathfrak{F},\mathfrak{D}) = \frac{1}{2} f + \frac{1}{2}\sum_{i\in\mathfrak{F}} \bar{L}_{ii}^{-1}. \tag{7}$$

The set of attacked nodes $\mathfrak{F}$ determine matrix B_2, and the set of defense nodes $\mathfrak{D}$ determines matrix D_y and consequently matrices H and $\bar{L}$ in (5). □

The actions of the attacker and the defender, when f nodes are under attack and f nodes are defended, define a matrix game $\mathcal{M}_{\binom{n}{f}\times\binom{n}{f}}$. Here $\mathcal{M}_{ij} = J(\mathfrak{F}_j, \mathfrak{D}_i)$, where $\mathfrak{F}_j$ corresponds to the set chosen by the attacker and $\mathfrak{D}_i$ corresponds to the set chosen by the defender. In other words, the attacker, the maximizer, chooses columns of matrix $\mathcal{M}$ and the defender, the minimizer, chooses the rows.

IV. Attacker-Defender Game on $J_1(\mathfrak{F},\mathfrak{D})$

In this section, we discuss equilibrium strategies for the attacker-defender game when the control law is (2a). First, consider a single attacked node and single defense node.

A. Single Attacked-Single defense Nodes

In this case, attacker's payoff is

$$J(\mathfrak{F},\mathfrak{D}) = \frac{H_{ii}^{-1}}{2}(d_i+1), \quad i \in \mathfrak{F}. \tag{8}$$

A Nash equilibrium may not exist, as discussed in the following example.

Example 1: For the path graph of length 3 shown in Fig. 1 (b), payoff matrix becomes

$$\mathcal{M} = \frac{1}{2}\begin{bmatrix} \frac{2}{k+1} & 3 & 2 \\ 2 & \frac{3}{k+1} & 2 \\ 2 & 3 & \frac{2}{k+1} \end{bmatrix}, \tag{9}$$

[3]Due to the lack of knowledge of the number of attack signals, the defender considers f as an upper bound of the number of attacked nodes and acts based on this worst-case scenario.

where the attacker (maximizer) chooses columns and the defender (minimizer) chooses the rows. For $k \leqslant \frac{1}{2}$ both the attacker and defender choose node 2 at NE, and the equilibrium payoff is $J^* = \frac{3}{2k+2}$. For k bigger than this threshold, there is no NE for the game. □

The following is a necessary and sufficient condition for the existence of an NE for the attacker-defender game.

Proposition 2: Suppose that in the game on $J_1(\mathfrak{F},\mathfrak{D})$ in (7), there are one attacked and one defense nodes. Then there exists an NE if and only if $k \leqslant \frac{\Delta_1-\Delta_2}{\Delta_2+1}$, where Δ_1 and Δ_2 are the largest and second largest degrees of nodes in graph $\mathcal{G}$. In this case, the game value is $J^* = \frac{\Delta_1+1}{2k+2}$ and the NE strategy is that both attacker and defender choose the node(s) with the largest degree.

Remark 1: According to Proposition 2, the value of k which ensures the existence of NE is limited by the gap between the largest and the second largest degrees in the network. For the cases where this does not hold, e.g., when the node with the largest degree is not unique as in Fig. 1 (c), there is no NE. Moreover, the largest possible threshold for graphs on n vertices corresponds to the star graph in which the threshold becomes $\frac{n-2}{2}$, as in Fig. 1 (d). □

When there is no NE, we instead analyze a Stackelberg game in which the defender acts as the leader. We can write $J_1(\mathfrak{F},\mathfrak{D})$ in (7) as

$$J_1(\mathfrak{F},\mathfrak{D}) = \frac{1}{2}\operatorname{tr}\left(F^T(L+I)H^{-1}F\right).$$

As leader, the defender solves the following optimization problem

$$J^*(D_y) = \min_{D_y} \frac{1}{2}\operatorname{tr}\left(F^{*^T}(D_y)(L+I)H^{-1}F^{*^T}(D_y)\right) \tag{10}$$

where D_y is chosen over all f defense nodes in $\mathcal{V}$. $F^*(D_y)$ is the best response of the attacker when the strategy of the defender is D_y, i.e., $F^*(D_y)$ is the solution of the following optimization problem

$$F^*(D_y) = \arg\max_F \frac{1}{2}\operatorname{tr}\left(F^T(L+I)H^{-1}F\right), \tag{11}$$

where F is chosen over all f attacked nodes in $\mathcal{V}$. Unlike NE, a Stackelberg game always admits an equilibrium strategy.

Remark 2: We note that for the attacker to play the Stackelberg game, i.e., find the optimal strategy (11), it is not necessary to know the exact value of the feedback gain k. According to proposition 2 and Theorem 2, which comes later, it is sufficient for the attacker to only know if k is above or below the threshold $\frac{\Delta_1-\Delta_2}{\Delta_2+1}$ in order to find its best response strategy. □

The following theorem, characterizes the equilibrium of the Stackelberg game.

Theorem 1: Consider a Stackelberg attacker-defender game on $J_1(\mathfrak{F},\mathfrak{D})$ in (7) in which there exists a single attacked node and single defense node, the defender as the game leader, and $k > \frac{\Delta_1-\Delta_2}{\Delta_2+1}$. Then the equilibrium strategy corresponds to the case where the defender chooses $v = \arg\max_{i\in\mathcal{V}} d_i$, i.e., the node with the largest degree.

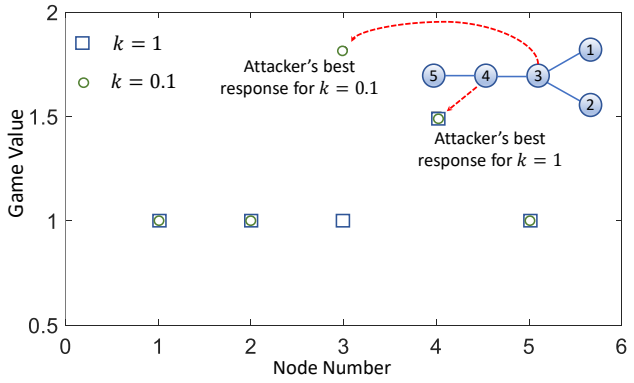

Fig. 2: The effect of the feedback value k on attacker's best response. The defender has chosen node 3 (its optimal decision).

In this case, the attacker's best response will be $\bar{v} = \arg\max_{i \in \mathcal{V} \setminus v} d_i$, i.e., the node with the second largest degree. □

The following example discusses the role of the threshold $\bar{k} = \frac{\Delta_1 - \Delta_2}{\Delta_2 + 1}$ in the attacker's strategy.

Example 2: For the graph shown in Fig. 2 we have $\Delta_1 = 3, \Delta_2 = 2$. Hence, the threshold is $\bar{k} = \frac{1}{3}$. The attacker's decisions are plotted with respect to the defender's best action, i.e., the node with the largest degree. For $k = 0.1 < \bar{k}$, the attacker's best action is the node with the largest degree (as follows from Proposition 2) and for $k = 1 > \bar{k}$, the attacker's best response is the node with the second largest degree (as follows from Theorem 1). For $k = \bar{k}$, the payoff will be the same when the attacker chooses either nodes 3 or 4. □

B. Multiple Attacked-Multiple Defense Nodes

Now consider the case that there exist f attacked nodes and f defense nodes, i.e., $|\mathfrak{F}| = |\mathfrak{D}| = f \geqslant 1$. Here we only consider a Stackelberg setup as it is more applicable to security problems [2]. We remark that if the defender is the leader, it reflects the defender's need to consider the worst case. Thus, it is more convenient to have the defender as the game leader.

The Stackelberg game is a combinatorial problem. Thus, in general, its computational cost would be high, unless it is reduced with specific assumptions. With this in mind, in our problem, finding optimal defense nodes when the defender is the game leader is burdensome, unless the control gain k is sufficiently large and the number of attacked nodes is sufficiently small.

Theorem 2: Consider a Stackelberg attacker-defender game on $J_1(\mathfrak{F}, \mathfrak{D})$ in (7) where there exists f attacked nodes and f defense nodes, $f \geqslant 1$ and $n \geqslant 2f$, with the defender as the game leader. If $k \geqslant \frac{1}{2}(f d_{\max} - 2)$ then at the equilibrium the defender chooses $v = \arg\max_{\substack{\mathfrak{D} \subseteq \mathcal{V} \\ |\mathcal{D}| = f}} \sum_{v_i \in \mathfrak{D}} d_i$, i.e., f nodes with the largest degrees in the network. The best response of the attacker is to choose $\bar{v} = \arg\max_{\substack{\mathfrak{F} \subseteq \mathcal{V} \setminus v \\ |\mathcal{F}| = f}} \sum_{v_i \in \mathfrak{F}} d_i$. □

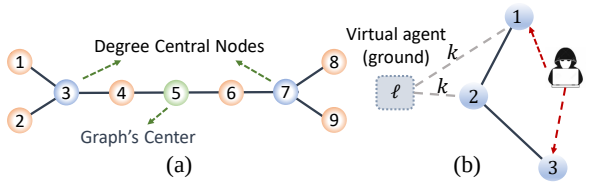

Fig. 3: (a) Optimal locations of the defense node for objective functions $J_1(\mathfrak{F}, \mathfrak{D})$ and $J_2(\mathfrak{F}, \mathfrak{D})$, (b) Extended graph and the virtual agent (ground).

V. Attacker-Defender Game on $J_2(\mathfrak{F}, \mathfrak{D})$

In this section, we discuss the dynamics with control law (2b) and objective function $J_2(\mathfrak{F}, \mathfrak{D})$ in (7).

A. Single Attacked, Single Defense Nodes

Similar to the case of $J_1(\mathfrak{F}, \mathfrak{D})$, we start with the case of single attacked and single defense nodes. We first have the following proposition.

Proposition 3: The attacker-defender game on $J_2(\mathfrak{F}, \mathfrak{D})$ in (7) with a single attack and single defense node does not admit an NE.

Similar to the attacker-defender game on $J_1(\mathfrak{F}, \mathfrak{D})$, in the absence of NE, an optimal defense strategy can be determined by finding the solution of the Stackelberg game. Recalling the notion of the *effective center of a graph*, from Section II, we have the following theorem which is proven in Appendix F.

Theorem 3: Consider the Stackelberg attacker-defender game on $J_2(\mathfrak{F}, \mathfrak{D})$ in (7) on graph $\mathcal{G}$ with the defender as the game leader. Then, a solution of the game corresponds to the case when the defender chooses the effective center of $\mathcal{G}$, i.e., $\mathfrak{D}^* = \arg\min_{v \in \mathcal{V}} \epsilon_f(v)$. In this case, the best response of the attacker will be $\mathcal{B}^*(\mathfrak{D}) = \arg\max_{j \in \mathcal{V}} R_{\mathfrak{D}^* j}$, i.e., a node with the maximum effective resistance from $\mathfrak{D}^*$. □

For the case of acyclic networks, Theorem 3 reduces to the following corollary.

*Corollary 1 (**Acyclic Networks**):* Consider the Stackelberg attacker-defender game on $J_2(\mathfrak{F}, \mathfrak{D})$ in (7), with the defender as the game leader, over the connected undirected tree $\mathcal{G}$. At equilibrium, the defender chooses the *center* of the graph and the attacker chooses the node with the greatest distance from the center.

*Remark 3 (**Game Equilibriums and Network Centrality**):* As mentioned before, the optimal location of the defense node for the objective function $J_1(\mathfrak{F}, \mathfrak{D})$ is the degree central node (Theorem 1) and for $J_2(\mathfrak{F}, \mathfrak{D})$ is the graph's center for acyclic networks (Corollary 1) or effective center for general graphs (Theorem 3). These network centralities (and consequently optimal defense node placements) can differ substantially from each other. One of such examples is the graph shown in Fig. 3 (a), in which by increasing the length of the path, the two centralities become far apart. □

B. Multiple Attacked, Multiple defense Nodes

In order to tackle this problem, we interpret the self-feedback loops in the form of connections to some virtual

agent (or grounded node) as shown in Fig. 3 (b). In this case, matrix $\bar{L}$ would be a submatrix of the Laplacian matrix $L_{(n+1)\times(n+1)}$ of the extended graph (including ℓ) where the row and the column corresponding to ℓ are removed. Such submatrices are called grounded Laplacian in the literature [20]. With this is mind, it is known that the i-th diagonal element of $\bar{L}^{-1}$ is $R_{i\ell}$, i.e., the effective resistance between node v_i and the virtual node ℓ [16].[4] As an example, consider nodes 1 and 2 in Fig. 3 (b) which are chosen as defenders and nodes 1 and 3 which are under attack. In this case, we have $J_2(\mathfrak{F},\mathfrak{D}) = 1 + \frac{1}{2}\sum_{i\in\mathfrak{F}} \bar{L}_{ii}^{-1} = 1 + \frac{1}{2}(R_{1\ell} + R_{3\ell})$. Based on this fact, the proof of the following theorem is straightforward.

Theorem 4: Consider the Stackelberg attacker-defender game on $J_2(\mathfrak{F},\mathfrak{D})$ in (7) with f defense nodes and f attack nodes, $f \geqslant 1$, with the defender as the game leader, over the connected undirected graph $\mathcal{G}$. Denote the virtual agent corresponding to a set of f defense nodes $\mathfrak{D}$ by $\ell_{\mathfrak{D}}$. Then, a solution of the game is when the defender chooses set $\mathfrak{D}$ in which the maximum sum of effective resistances between $\ell_{\mathfrak{D}}$ and all combinations of f nodes in the network is minimized, i.e., $\mathfrak{D}^* = \arg\min_{\mathfrak{D}\subseteq\mathcal{V}} \max_{\mathfrak{F}\subseteq\mathcal{V}} \sum_{j\in\mathfrak{F}} R_{\ell_{\mathfrak{D}} j}$. In this case, the attacker chooses the set of f attacked nodes as $\mathfrak{F}^* = \arg\max_{\mathfrak{F}\subseteq\mathcal{V}} \sum_{j\in\mathfrak{F}} R_{\ell_{\mathfrak{D}^*} j}$. □

As it is seen from Theorem 4, finding the optimal set of defense nodes requires a high level of computation.

Remark 4: **(The Effect of Increasing Connectivity)**: Since the effective resistance between two nodes in the graph is an increasing function of edge weights [20], adding extra edges to the network (or increasing the weight of edges) decreases the diagonal elements of $\bar{L}^{-1}$ and consequently decreases the system $\mathcal{H}_2$ norm. Hence, unlike control law (2a), increasing connectivity is beneficial from the defender's perspective for (2b). □

APPENDIX

A. Proof of Proposition 1

Proof: We prove for the first case, the second case (2b) follows a similar procedure. We compute the $\mathcal{H}_2$ norm using the trace formula $||G||_2^2 = \mathrm{tr}(B_2^\mathsf{T}\mathcal{W}_o B_2)$, where $\mathcal{W}_o$ is the observability Gramian $\mathcal{W}_o = \int_0^\infty e^{A^\mathsf{T} t} C^\mathsf{T} C e^{At}$ and it is uniquely obtained from the Lyapunov equation $\mathcal{W}_o A + A^\mathsf{T}\mathcal{W}_o = -C^\mathsf{T}C$ with an additional constraint $\mathcal{W}_o v = \mathbf{0}_{2n}$ where $v = [\mathbf{1}_n^\mathsf{T}\ \mathbf{0}_n^\mathsf{T}]^\mathsf{T}$ is the mode corresponding to the marginally stable eigenvalue of A. It is due to the fact that the marginally stable mode v is not detectable, i.e., $Ce^{At}v = Cv = \mathbf{0}_{2n}$ for all $t \geqslant 0$, and since the rest of the eigenvalues are stable, the indefinite integral exists. The proof of the uniqueness of $\mathcal{W}_o$ is the same as [23, Lemma 1] and is omitted here. To calculate the observability Gramian, we have

$$\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} A + A^\mathsf{T} \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_n & \mathbf{0}_n \\ \mathbf{0}_n & -I_n \end{bmatrix} \quad (12)$$

[4]When the graph is a tree, the effective resistance and physical distance become the same.

By solving (12) we get $W_{11} = \frac{1}{2}LH^{-1}$, $W_{22} = \frac{1}{2}H^{-1}$ and $W_{12} = W_{21} = \mathbf{0}$. Hence we have $||G||_2^2 = \mathrm{tr}(B_2^\mathsf{T}\mathcal{W}_o B_2) = \mathrm{tr}(F^T W_{11} F + F^T W_{22} F)$ which yields the result. ■

B. Proof of Proposition 2

Proof: It is easy to verify that each element of the matrix game $\mathcal{M}$ is

$$\mathcal{M}_{ij} = \begin{cases} \frac{d_j+1}{2k+2} & i = j, \\ \frac{d_j+1}{2} & i \neq j. \end{cases} \quad (13)$$

We first prove the sufficient condition, i.e., assume $k \leqslant \frac{\Delta_1-\Delta_2}{\Delta_2+1}$. Then, if the attacker changes its strategy (unilaterally) from the node with the maximum degree to some node v_i, according to (13) and the upper bound for k, the game value becomes $J = \frac{d_i+1}{2} \leqslant \frac{\Delta_1+1}{2k+2}$. Moreover, if the defender wants to change its strategy to another node v_i, based on (13) since the smallest element of each column is its diagonal element, it will get $J = \frac{d_i+1}{2} \geqslant \frac{\Delta_1+1}{2k+2}$. Hence, neither the attacker nor the defender are willing to change their strategy unilaterally.

Now suppose that having both attacker and defender choose the node with the largest degree is NE. Then we have to have $\frac{\Delta_1+1}{2k+2} \geqslant \frac{d_j+1}{2}$ for all $j = 1, 2, ..., n$ which results in $k \leqslant \frac{\Delta_1-d_i}{d_i+1}$ for all $j = 1, 2, ..., n$ and this proves the claim. ■

C. Proof of Theorem 1

Proof: When $k > \frac{\Delta_1-\Delta_2}{\Delta_2+1}$, for each row (defender's action) of matrix $\mathcal{M}$, the largest element (the best response of the attacker) will be $\frac{1}{2}(\Delta_1+1)$, except the row corresponding to the node with the largest degree. In that row, the largest element will be $\frac{1}{2}(\Delta_2+1)$. Since $\Delta_1 \geqslant \Delta_2$, the optimal action of the defender will be $v = \arg\max_{i\in\mathcal{V}} d_i$. The best response of the attacker will be the node with the second largest degree. This solution may not be unique, however, the optimal value of this game is unique and given by $J^* = \frac{\Delta_2+1}{2}$. ■

D. Proof of Theorem 2

Proof: For multiple attacked-multiple defense nodes case, each element of the matrix game $\mathcal{M}_{ij}$ (corresponding to defender decision set $\mathfrak{D}_i$ and attacker decision set $\mathfrak{F}_j$) is

$$\mathcal{M}_{ij} = \begin{cases} \frac{\sum_{j\in\mathfrak{F}_j} d_k}{2k+2} + \frac{f}{2k+2} & i = j, \\ \frac{\sum_{k\in\mathfrak{F}_j\cap\mathfrak{D}_i} d_k}{2k+2} + \frac{\gamma_1^{ij}}{2k+2} & \\ +\frac{1}{2}(\sum_{k\in\mathfrak{F}_j\setminus\mathfrak{D}_i} d_k + \gamma_2^{ij}) & i \neq j, \end{cases} \quad (14)$$

where $\gamma_1^{ij} = |\mathfrak{F}_j \cap \mathfrak{D}_i|$ and $\gamma_2^{ij} = |\mathfrak{F}_j \setminus \mathfrak{D}_i| = f - \gamma_1^{ij}$. Since the defender is the game leader, it has to choose a row in game matrix $\mathcal{M}$ whose maximum element is minimum (over all other rows). When k is lower bounded by $k \geqslant \frac{1}{2}(fd_{\max} - 2)$, considering a fixed set $\mathfrak{D}$ (set of defense nodes), for each set of attacked nodes $\mathfrak{F}$ we have

$$\frac{\sum_{j\in\mathfrak{F}\cap\mathfrak{D}} d_j}{2k+2} + \frac{\gamma_1}{2k+2} \leqslant 2. \quad \forall\mathfrak{F}\subseteq\mathcal{V} \quad (15)$$

Inequality (15) together with (14) shows that for the row corresponding to set $\mathfrak{D}$, its largest element corresponds to the set of attackers $\bar{\mathfrak{F}}$ for which $\bar{\mathfrak{F}} \cap \mathfrak{D} = \phi$. In order for this to happen, we must have $n \geqslant 2f$. In this case, $\gamma_1 = 0$ and the maximum element in the row corresponding to set $\mathfrak{D}$ is (according to the second term in (14)) $\bar{\mathcal{M}} = \max_{\mathfrak{F} \subseteq \mathcal{V} \setminus \mathfrak{D}} \frac{1}{2} \sum_{j \in \mathfrak{F}} d_j + f$. Thus, the best action of the defender, to minimize that maximum row element, is to choose $\bar{\mathfrak{D}} = \arg\max_{\mathfrak{D} \subseteq \mathcal{V}} \sum_{j \in \mathfrak{D}} d_j$, i.e., f nodes with largest degrees in the graph. ■

E. Proof of Proposition 3

Proof: As mentioned in Section V, the j-th diagonal element of $\bar{L}^{-1}$ is the effective resistance from node v_j and the virtual node ℓ which is connected to the single defense node v_i with an edge of weight (conductance) k [20]. Thus, we have $[\bar{L}^{-1}]_{jj} = R_{\ell j}$. Hence, the value of each diagonal element of the game matrix $\mathcal{M}$ is $\mathcal{M}_{ii} = \frac{1}{2} + \frac{1}{2k}$ and each off-diagonal element is $\mathcal{M}_{ij} = \frac{1}{2} + \frac{1}{2k} + \frac{1}{2} R_{ij}$. Thus, each diagonal element is strictly less than the elements of its corresponding row and column. Now, assume that a NE exists and let (i^*, j^*) denote the equilibrium strategies of the attacker and defender. Thus, we should have $[\mathcal{M}]_{i^* j} \leqslant [\mathcal{M}]_{i^* j^*} \leqslant [\mathcal{M}]_{i j^*}$ for all $i \neq i^*$ and $j \neq j^*$. If element $[\mathcal{M}]_{i^* j^*}$ is a diagonal element, then the left inequality will be violated and if it is a non-diagonal element, the right inequality will be violated. ■

F. Proof of Theorem 3

Proof: We know that for the game matrix $\mathcal{M}$ we have $\mathcal{M}_{ij} = \frac{1}{2} + \frac{1}{2} R_{\ell j}$, where v_ℓ is the virtual agent connected to the defense node v_i with an edge with weight k and v_j is the attacked node. As the defender is the leader of the Stackelberg game, it minimizes (over all rows) the maximum element of each row of $\mathcal{M}$. Thus, the optimal place for the defender is $v^* = \arg\min_i \max_j \mathcal{R}_{\ell j}$ and this is the effective center of the graph defined in Section II. Note that this solution (strategies of the defender and attacker) may not be unique since the effective center of the network may not be a single node. However, the value of the game is unique. ■

II. Conclusion

A game-theoretic approach to the resilience of two canonical forms of second-order network control systems was discussed. For the case of a single attacked node and a single defense node, it was shown that the optimal location of the defense node for each of the two second-order systems introduces a specific network centrality measure. The extension of the game to the case of multiple attacked and defense nodes was also discussed and graph-theoretic interpretations of the equilibrium of the Stackelberg game for this case was investigated. An avenue for the future work is to extend these results to directed networks.

References

[1] Q. Zhu and T. Basar, "Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: Games-in-games principle for optimal cross-layer resilient control systems," in *IEEE control systems*, vol. 35, no. 1, 2015, pp. 45–65.

[2] M. Manshaei, Q. Zhu, T. Alpcan, T. Basar, and J. P. Hubaux, "Game theory meets network security and privacy," *ACM Computing Surveys*, vol. 45, pp. 53–73, 2013.

[3] A. Gupta, C. Langbort, and T. Basar, "Optimal control in the presence of an intelligent jammer with limited actions," *49th IEEE Conference on Decision and Control*, pp. 1096–1101, 2010.

[4] S. Amin, G. A. Schwartz, and S. S. Sastry, "Security of interdependent and identical networked control systems," *Automatica*, vol. 49, no. 1, pp. 186–192, 2013.

[5] M. Wu and S. Amin, "Securing infrastructure facilities: When does proactive defense help?" *Dynamic Games and Applications, https://doi.org/10.1007/s13235-018-0280-8*, 2018.

[6] M. Pirani, E. Nekouie, H. Sandberg, and K.H.Johansson, "A game-theoretic framework for security-aware sensor placement problem in networked control systems," *American Control Conference (to appear)*, 2019.

[7] Q. Zhu and T. Basar, "robust and resilient control design for cyber-physical systems with an application to power systems," in *in Proc. 50th IEEE Conf. Decision Control European Control*, 2011, pp. 4066–4071.

[8] J. P. H. M. Felegyhazi, *Game Theory in Wireless Networks: A Tutorial*. EPFL Technical report, 2006.

[9] J. Marden, G. Arslan, and J. S. Shamma, "Ieee transactions on systems, man, and cybernetics, part b (cybernetics)," *IEEE Trans. Smart Grid*, vol. 39, no. 6, pp. 1393–1407, 2009.

[10] P. N. Brown and H. B. N. J. R. Marden, *Security Against Impersonation Attacks in Distributed Systems*. arXiv preprint arXiv:1711.00609, 2017.

[11] S. Amin, G. A. Schwartz, and S. S. Sastry, "Security of interdependent and identical networked control systems," *Automatica*, pp. 186–192, 2013.

[12] S. M. Dibaji and H. Ishii, "Resilient consensus of second-order agent networks: Asynchronous update rules with delays," *Automatica*, vol. 81, pp. 123–132, 2017.

[13] I. Shames, A. Teixeira, H. Sandberg, and K. H. Johansson, "Distributed fault detection for interconnected second-order systems," *Automatica*, vol. 47, pp. 2757–2764, 2011.

[14] 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, pp. 766–781, 2013.

[15] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, pp. 90–104, 2012.

[16] A. Ghosh, S. Boyd, and A. Saberi, "Minimizing effective resistance of a graph," *SIAM Review*, vol. 50, no. 1, pp. 37–66, 2008.

[17] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. New York, NY, USA: Wiley, 2012.

[18] M. Pirani, J. W. Simpson-Porco, and B. Fidan, "System-theoretic performance metrics for low-inertia stability of power networks," in *IEEE Conference on Decision and Control*. IEEE, 2017.

[19] W. Ren, R. Beard, and E. Atkins, "Information consensus in multivehicle cooperative control," in *IEEE Control systems magazine*, 2007, pp. 71–82.

[20] M. Pirani, E. M. Shahrivar, B. Fidan, and S. Sundaram, "Robustness of leader - follower networked dynamical systems," *IEEE Transactions on Control of Network Systems*, 2017.

[21] B. Bamieh, M. R. Jovanovic, P. Mitra, and S. Patterson, "Coherence in large-scale networks: Dimension-dependent limitations of local feedback," *IEEE Transactions on Automatic Control*, vol. 57, pp. 2235–2249, 2012.

[22] E. Tegling, B. Bamieh, and D. F. Gayme, "The price of synchrony: Evaluating the resistive losses in synchronizing power networks," *IEEE Transactions on Control of Network Systems*, vol. 2, pp. 254–266, 2015.

[23] B. K. Poolla, S. Bolognani, and F. Dorfler, "Optimal placement of virtual inertia in power grids," *American Control Conference*, 2016.