Automatica 50 (2014) 852–857

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Security concepts for the dynamics of autonomous vehicle networks☆

Mengran Xue [a], Wei Wang [b], Sandip Roy [b,1]

[a] University of Michigan, Ann Arbor, MI 48109, United States

[b] Washington State University, Pullman, WA 99164, United States

ARTICLE INFO

Article history:
Received 23 January 2012
Received in revised form
22 September 2013
Accepted 30 October 2013
Available online 1 February 2014

Keywords:
Network security
Autonomous vehicle teams
Network dynamics
Estimation
Control theory
Graph theory

ABSTRACT

The secure operation of autonomous vehicle networks in the presence of adversarial observation is examined, in the context of a canonical double-integrator-network (DIN) model. Specifically, we study the ability of a sentient adversary to estimate the full network's state, from noisy local measurements of vehicle motions. Algebraic, spectral, and graphical characterizations are provided, which indicate the critical role of the inter-vehicle communication topology and control scheme in achieving security.

1. Introduction

Threat assessment and mitigation challenges are increasingly arising in networks with tightly intertwined physical- and cyber-components. Dynamical-network estimation and control tools appear to be well-suited to address these challenges, in that they permit modeling and design of network dynamics involving both complex physical processes and information–communication/processing capabilities. In consequence, there has been a growing focus in the network-controls community on defining and analyzing security and vulnerability notions in a range of cyber–physical networks, e.g. Pasqualetti, Dorfler, and Bullo (2011), Xue, Roy, Wan, and Das (2011) and Zhu and Martinez (2011). In this article, motivated by autonomous-vehicle-network (AVN) applications, we examine one interesting cyber–physical threat-assessment or security-analysis problem. Specifically, we focus on the canonical double-integrator-network (DIN) model for an AVN engaged in a tracking task, which provides an abstract representation of both the network's physical dynamics and its cyber- (communication/control) capabilities. Here, we enrich the DIN to represent an adversary, which can obtain noisy measurements of some vehicles' local dynamics but is not able to actuate the dynamics in any way. We then define notions of security, which capture whether or not an adversary can discover or estimate the initial state of the DIN from its measurements, and also describe the fidelity of these estimates when observations are noisy.[2] Starting from an observability analysis of the closed-loop dynamics, several spectral and graphical characterizations of these security notions are obtained (for both the noise-free and noisy cases), that indicate the role of the network's communication and control architecture in protecting the network dynamics from discovery.

The security analysis pursued here is related to several current research thrusts in the network-controls and cyber-physical-systems literature. First, our work is closely tied to several fundamental studies of network dynamics and structure estimation from local observations (including via distributed estimation), e.g. Pasqualetti, Bicchi, and Bullo (2012), Pasqualetti et al. (2011), Roy, Xue, and Das (2012), Sou, Sandberg, and Johansson (2012), Sundaram and Hadjicostis (2011) and Wan and Roy (2009). We

☆ The material in this paper was partially presented at the 51st IEEE Conference on Decision and Control (CDC), December 10–13, 2012, Maui, Hawaii. This paper was recommended for publication in revised form by Associate Editor Antonis Papachristodoulou under the direction of Editor Frank Allgöwer.

E-mail addresses: mxue@umich.edu (M. Xue), wwang1@eecs.wsu.edu (W. Wang), sroy@eecs.wsu.edu (S. Roy).

[1] Tel.: +1 509 335 2448; fax: +1 509 335 3818.

0005-1098/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.automatica.2013.12.001

[2] The term "security" is also used to describe a system's ability to thwart an active attack. In accordance with our earlier work (Xue et al., 2011), we here use the term to describe the protection of information from a measuring adversary, but acknowledge the varying definitions. It is also worth stressing that the security notion we consider here is closely connected to the emerging concept of privacy or anonymity in distributed algorithms/controls, see e.g. Le Ny and Pappas (2012) and Telerius, Varagnolo, Baquero, and Johansson (2013). From this perspective, our results can be viewed as characterizing the level of privacy among the agents/vehicles in a DIN.

particularly point out that several of these efforts are also concerned with local monitoring of network dynamics, whether by system planners to detect an attack or by adversaries to detect nominal network dynamics. As a dual to these graph-theoretic estimation/observability analyses, graph-theoretic viewpoints on controllability have also been developed that have a similar flavor (Rahmani, Ji, Mesbahi, & Egerstedt, 2010). Second, this work is complementary to control-theoretic modeling of attacks in cyber-networks and networked control systems (e.g., Alpcan & Basar, 2003, Amin, Cardenas, & Sastry, 2009, Liu, Ning, & Reiter, 2011, Mo & Sinopoli, 2009 and Texiera, Sandberg, & Johansson, 2010). While these efforts mostly are concerned with active adversaries, many of these works also obtain graphical results on adversarial conduct in networks, as in our work. We also note the connection of our work to recent efforts on fault and event detection in networks, including for systems with second-order local dynamics (Roy & Chen, 2013; Shames, Texiera, Sandberg, & Johansson, 2010). More broadly, this study is related to efforts to define security and vulnerability concepts for cyber–physical networks (e.g., Pasqualetti et al., 2011 and Xue et al., 2011); it also enriches the extensive study of AVN control in the control community (e.g., Ren & Beard, 2005 and Roy, Saberi, & Herlugson, 2004), toward performance design to achieve security. While our explorations here are connected to these research thrusts, we stress that a particular focus of our analyses is to distinguish the roles of the AVN's physical dynamics (motion), communication, and control, and of the adversary's (sparse) measurement capabilities, in network security.

In brief, the particular contributions described in this article are the following:

- We extend the DIN model to capture an adversary with local observation capabilities, and introduce attendant notions of security concerned with estimation of the full network's initial state.
- We characterize a binary notion of security (an observability notion) when the adversary's observations are not noisy, in terms of the network's graph matrix, its spectrum, and the graph topology itself.
- We characterize security levels (which codify estimation error) in the noisy-measurement case, in terms of the graph matrix and its spectrum.

2. Problem formulation

In this section, the DIN model is reviewed, the adversary is modeled, and security notions are formally defined.

The DIN. We consider a team of n vehicles, labeled as $i = 1, \ldots, n$. For convenience, we assume that each vehicle is moving in a single dimension: the multi-dimensional case can be transformed to a single-dimensional model with more vehicles, see Roy et al. (2004). The vehicles' positions satisfy the differential equation $\ddot{\mathbf{x}}(t) = \mathbf{u}(t)$, where the *full position vector* $\mathbf{x}(t) = \begin{bmatrix} x_1(t), & \ldots, & x_n(t) \end{bmatrix}^T$ contains the positions $x_i(t)$ of each vehicle, and the *full input vector* $\mathbf{u}(t) = \begin{bmatrix} u_1(t), & \ldots, & u_n(t) \end{bmatrix}^T$ specifies the control input $u_i(t)$ for each vehicle. We call $h_i(t) \triangleq \dot{x}_i(t)$ the *velocity* of vehicle i, and call $\mathbf{h}(t) = \begin{bmatrix} h_1(t), & \ldots, & h_n(t) \end{bmatrix}^T$ the *full velocity vector*.

Each vehicle uses information that is sensed and/or communicated from other vehicles, as well as (possibly) information about a target location, to set its control inputs. Formally, each vehicle i is assumed to have available a p_i-component *position-observation vector* $\mathbf{y}_{pi}(t) = G_i\mathbf{x}(t)$, where the $p_i \times n$-dimensional matrix G_i specifies agent i's capability to observe the vehicle network's full state and hence is called agent i's *observation matrix* (see Roy et al., 2004 for many illustrative examples). The vehicle is also assumed to have commensurate velocity observations—specifically, that p_i-component *velocity-observation vector* $\mathbf{y}_{vi}(t) = G_i\mathbf{h}(t)$—either obtained through direct measurement, or as the derivative of the position measurement.[3] The vehicles' internal dynamics together with their observation capabilities nominally specify the DIN. For notational convenience, we also stack the observation matrices into a single matrix: $G \triangleq \begin{bmatrix} G_1^T, & \ldots, & G_n^T \end{bmatrix}^T$, which we call the *full observation matrix*. We assemble $\mathbf{y}_{pi}$ and $\mathbf{y}_{vi}$ as: $\mathbf{y}_p \triangleq \begin{bmatrix} \mathbf{y}_{p1}^T, & \ldots, & \mathbf{y}_{pn}^T \end{bmatrix}^T$, and $\mathbf{y}_v \triangleq \begin{bmatrix} \mathbf{y}_{v1}^T, & \ldots, & \mathbf{y}_{vn}^T \end{bmatrix}^T$.

Tracking control in the DIN. Decentralized controller designs have been obtained for the DIN, for several coordinated-motion tasks. Here, we focus on a fixed-target-tracking problem. Specifically, the vehicles in the team are all tasked with moving from initial locations or *home bases* $x_i(0)$ to a specified scalar *target location* $\bar{s}$. It is assumed that the target location is distributed to the vehicles as needed prior to the tracking task. We stress here that the vehicles may not have measurements of absolute position information in the reference frame of the target, and hence they depend on sensing or communication to complete target tracking (Roy et al., 2004). It turns out that vehicles whose measurements are all relative positions/velocities (i.e., each row of G_i sums to 0) do not require knowledge of the target (Roy et al., 2004).

Here, we consider using a memoryless linear decentralized controller architecture to achieve tracking. Specifically, each vehicle i is assumed to use a controller of the form $u_i(t) = K_i\mathbf{y}_{vi}(t) + \alpha K_i(\mathbf{y}_{pi}(t) - G_i\mathbf{s}^*)$, where the 1-by-$p_i$ *gain matrix* K_i weights the observations in computing the concurrent input, α is a scalar gain factor that is common for all vehicles, and $\mathbf{s}^* \triangleq \begin{bmatrix} \bar{s}, & \ldots, & \bar{s} \end{bmatrix}^T$. It is convenient to assemble the gain matrices as the diagonal blocks of a *full gain matrix*: $K = \mathrm{diag}(K_i)$. In this notation, the controllers of all n vehicles are captured by the equation $\mathbf{u}(t) = K\mathbf{y}_v(t) + \alpha K(\mathbf{y}_p(t) - G\mathbf{s}^*)$, or equivalently $\mathbf{u}(t) = KG\mathbf{h}(t) + \alpha K(G\mathbf{x}(t) - G\mathbf{s}^*)$. A couple notes about the control architecture are worthwhile. First, we note that $G_i\mathbf{s}^* = 0$ for any vehicle i whose observations are relative positions/velocities, which verifies that these vehicles do not require information about the tracking task. Also, we stress that at least one vehicle must make measurements in an absolute frame (and hence must have information about the target location) for tracking to be possible: otherwise, G will not have full rank and stabilization is impossible.

Controllers of this architecture have been designed to achieve tracking, for a broad class of network topologies G, see Roy et al. (2004). Specifically, under broad conditions, the full gain matrix K can be designed so that KG has real, negative, and distinct eigenvalues. In turn, by choosing α to be small (specifically, less than $1/4$ of the minimum eigenvalue of KG), the tracking task is achieved. Such memoryless low-gain controllers can be designed to be robust to actuator saturation, communication delay, and topological variation (Locatelli & Schiavoni, 2012; Roy et al., 2004). Beyond this class of special low-gain controllers, an even broader family of designs has been obtained to place the eigenvalues of KG in the *open left half plane* (OLHP), and in turn to achieve tracking. Many of our results here encompass any tracking controller of this type, while others are focused on the special low-gain control.

Closed-loop dynamics: matrix representation. Let us define an *extended state vector* of dimension $2n + 1$ that stacks the positions and velocities of the vehicles as well as the target location: $\boldsymbol{\varphi}(t) = \begin{bmatrix} \mathbf{x}^T(t) & \mathbf{h}^T(t) & s(t) \end{bmatrix}^T$, where $s(t)$ is a scalar that is fixed at the target location (i.e., $s(t) = \bar{s}$ for all t). This state vector $\boldsymbol{\varphi}(t)$, in the closed loop, evolves according to:

$$\dot{\boldsymbol{\varphi}} = A\boldsymbol{\varphi}, \tag{1}$$

[3] Derivative computations magnify high-frequency noise, so must be low-pass filtered. Implementations of derivatives used in feedback have been widely studied, including for the DIN, so we omit the details and assume that an accurate derivative observation is available.

where the state matrix is $A = \begin{bmatrix} \mathbf{0} & I_n & 0 \\ \alpha KG & KG & -\alpha KG\mathbf{1}_n \\ \mathbf{0} & \mathbf{0} & 0 \end{bmatrix}$, where I_n is an identity matrix of dimension $n \times n$, and $\mathbf{1}_n$ is an n-component vector whose entries are all unity. It is easy to check that the closed-loop state matrix A has one eigenvalue equal to 0. Further, the tracking task is achieved if and only if the other eigenvalues of A are in the open left half plane. The described low-gain memoryless control design in fact places these eigenvalues on the negative real axis.

We note that the initial value of the position vector $\mathbf{x}(0)$ specifies the vehicles' home-base locations, $\mathbf{h}(0)$ specifies the initial velocities of the vehicles, and the initial value of $s(0)$ (which is maintained with time) is the target location $\bar{s}$. Since we are interested in the estimation of these quantities by an adversary, it is important to formalize that they are *nonrandom* parameters in our formulation.

Interaction graph. The matrix KG specifies the interaction topology among the agents, as determined through their communication/sensing and control. As such, it is natural to define an *interaction graph* for the vehicle team from KG. Specifically, the weighted, directed interaction graph Γ is defined to have n vertices, which represent the n vehicles (in order). An edge is drawn from vertex j to vertex i (not necessarily distinct), if and only if the entry of KG at row i and column j is nonzero. In this case, the edge weight is set to that entry. The interaction graph Γ specifies how each vehicle's position and velocity variable actuates other vehicles' dynamics.

Adversarial observations and security definitions. We envision an adversary that makes (possibly noisy) local measurements of DIN dynamics over the interval $[0, t_f]$. We focus on the case that these *adversary's measurements* are the positions of a small group of vehicles, say vehicles $1, \ldots, q$ ($q \in \{1, \ldots, n\}$) without loss of generality. In the noise-free case, the adversary's measurements are thus assumed to be $\mathbf{z}(t) = C\boldsymbol{\varphi}(t)$, where $C = \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_q \end{bmatrix}^T$ has dimension $(2n+1) \times q$, and where $\mathbf{e}_j$ ($j = 1, \ldots, q$) is an $(2n+1)$-entry indicator vector with jth entry equal to 1. In the noisy case, the adversary's measurements are given by $\mathbf{z}(t) = C\boldsymbol{\varphi}(t) + \mathbf{m}(t)$, where the noise $\mathbf{m}(t)$ is assumed to be white and Gaussian with zero mean and autocorrelation $R(\tau) = \sigma^2 I\delta(\tau)$. It is worth noting that another very plausible model is that the adversary obtains the sensed/communicated observations of a subset of the vehicles. We do not pursue this case in the interest of space.

In this study, the adversary is assumed to have full knowledge of the vehicle-team model (including the internal dynamics, the sensing/communication capabilities, and the identities of the measured vehicles). The adversary's goal is to estimate features of the state (dynamics) of the vehicle team. We will focus on the estimation of the home-base locations, initial velocities, and target location, which are the initial conditions of the state vector $\boldsymbol{\varphi}$ governed by Eq. (1), or features that are linear combinations of $\boldsymbol{\varphi}$. The circumstance considered in this paper, that the vehicle-team model is entirely known but the dynamics are unknown to the adversary, is particularly likely to arise in circumstances where a manufactured multi-vehicle or multi-agent team with fixed communications/algorithms is used repeatedly (as may be case in micro-robotics-based health monitoring, automated target-detection in hostile environments, or mobile-computing applications). One remark is worthwhile: we formally assume that the adversary commences measurement when the vehicles leave their home bases. If the measurement were commenced at a different time, it turns out that the adversary could reconstruct the vehicles' trajectories (forward and backwards in time) under the same conditions as presented below; however to identify the home base locations along these trajectories, the adversary would need an estimate of the initiation time. We omit the details to save space.

Let us now define security and security-level notions for the DIN. In the case that the adversary's measurements are noise-free, security is naturally defined as a binary notion: the initial state is defined to be *secure*, if the adversary cannot uniquely compute the initial state from its measurements. Otherwise, the initial state is *insecure*. In the case where the adversary's measurements are noisy, finite-variance estimation may or may not be possible, depending on whether or not the noise-free model is insecure or secure. When finite-variance estimation is possible (i.e., the noise-free model is insecure), we define the *security level* of the DIN to reflect the quality of the adversary's estimate of the desired nonrandom statistics. Here, we define the error covariance for a minimum-variance-unbiased (MVU) estimate of the initial-condition vector as the *security level matrix*. We also consider scalar measures defined from this error covariance matrix (e.g., its extremal eigenvalues, which capture extremal errors in computing unitary statistics of the initial condition). We refer to these measures as *security-level scalars*. Security and security-levels can be defined analogously for features of the initial state.

3. Security in the noise-free case

We characterize whether or not the DIN's initial state is secure in the noise-free case, in terms of its communication topology and control architecture. Two results are obtained: first, security is characterized in terms of the spectrum of the matrix KG, which captures the communication and control paradigm; second, this spectral characterization is used to obtain sufficient conditions for security in terms of the interaction graph. At their essence, these characterizations derive from the fact that security is equivalent to unobservability of the pair (C, A), where C is the adversary's observation matrix and A is the closed-loop state matrix. In this development, we use the notation λ_i, $i = 1, \ldots, n$ for the n eigenvalues of KG. We also denote the corresponding eigenvectors and generalized eigenvectors as $\mathbf{v}_i$, $i = 1, \ldots, n$. Let us begin with the spectral result, which indicates that the zero patterns of the eigenvectors of KG determine security.

Theorem 1. *Consider a DIN (1) whose controller has been designed to achieve a static target-tracking task. The initial state is secure if and only if KG has a right eigenvector whose first q entries are all zero.*

Proof. We note the equivalence between security and unobservability of the pair (C, A). Thus, invoking the classical spectral condition for observability, it follows immediately that the initial state is secure if and only if the closed-loop state-matrix A has an eigenvector whose first q entries are 0. Let us relate the eigenvectors of A with those of KG, noting first that the eigenvalues of KG cannot be 0 since the tracking task is achieved (Roy et al., 2004). To do so, consider the matrix $H = \begin{bmatrix} 0 & I \\ \alpha KG & KG \end{bmatrix}$. Consider an eigenvalue $\lambda_i \neq 0$ of KG, with associated (non-generalized) eigenvector $\mathbf{v}_i$. It is easy to verify algebraically that $\mathbf{p}_{i,1} = \begin{bmatrix} \mathbf{v}_i \\ \mu_{i,1}\mathbf{v}_i \end{bmatrix}$ and $\mathbf{p}_{i,2} = \begin{bmatrix} \mathbf{v}_i \\ \mu_{i,2}\mathbf{v}_i \end{bmatrix}$ are both eigenvectors of H, when $\mu_{i,1} \neq 0$ where $\mu_{i,2} \neq 0$ are the two distinct solutions of the equation $\mu_i^2 - \lambda_i\mu_i - \alpha\lambda_i = 0$. In the case that the two solutions are identical, it is easy to check that $\mathbf{p}_{i,1} = \begin{bmatrix} \mathbf{v}_i \\ \mu_i\mathbf{v}_i \end{bmatrix}$ is an eigenvector, and further that a generalized eigenvector of H with eigenvalue λ_i can be constructed (let us call it $\mathbf{p}_{i,2}$).

Meanwhile, if $\mathbf{v}_i$ is a generalized eigenvector of KG associated with λ_i, let us show that H has two associated generalized eigenvectors. First, let us assume that $KG\mathbf{v}_i = \lambda_i\mathbf{v}_i + \tilde{\mathbf{v}}_i$, where $\tilde{\mathbf{v}}_i$ is an eigenvector of KG satisfying that $KG\tilde{\mathbf{v}}_i = \lambda_i\tilde{\mathbf{v}}_i$ (i.e., $\mathbf{v}_i$ is the first generalized eigenvector in a Jordan chain). Then, the two generalized eigenvectors of H are $\mathbf{p}_{i,1} = \begin{bmatrix} \frac{2\mu_{i,1} - \lambda_i}{\alpha + \mu_{i,1}}\mathbf{v}_i \\ \frac{2\mu_{i,1}^2 - \lambda_i\mu_{i,1}}{\alpha + \mu_{i,1}}\mathbf{v}_i + \tilde{\mathbf{v}}_i \end{bmatrix}$ and $\mathbf{p}_{i,2} = \begin{bmatrix} \frac{2\mu_{i,2} - \lambda_i}{\alpha + \mu_{i,2}}\mathbf{v}_i \\ \frac{2\mu_{i,2}^2 - \lambda_i\mu_{i,2}}{\alpha + \mu_{i,2}}\mathbf{v}_i + \tilde{\mathbf{v}}_i \end{bmatrix}$, where $\mu_{i,1}$ and $\mu_{i,2}$ are again the roots

of $\mu_i^2 - \lambda_i\mu_i - \alpha\lambda_i = 0$. In a similar fashion, two eigenvectors can be constructed for each generalized eigenvector of *KG* of further depth in the Jordan chain; details are omitted. In this way, the eigenvectors and generalized eigenvectors of *H* can be constructed from those of *KG*.

Finally, we note that *A* has a zero eigenvalue with eigenvector $\begin{bmatrix}\mathbf{1}_n^T & \mathbf{0}_n^T & 1\end{bmatrix}^T$. Meanwhile, the remaining $2n$ eigenvalues of *A* are those of *H*, with associated eigenvectors and generalized eigenvectors of the form $\begin{bmatrix}\mathbf{p}_{i,1}\\ 0\end{bmatrix}$ or $\begin{bmatrix}\mathbf{p}_{i,2}\\ 0\end{bmatrix}$. From these expressions, it is immediately clear that *A* has an eigenvector in the null space of *C*, if *KG* has an eigenvector $\mathbf{v}_i$ whose first *q* entries are zero. Conversely, if the DIN is secure, the closed-loop state-matrix *A* has an eigenvector (associated with a non-zero eigenvalue) whose first *q* entries are 0. Notice that the first *n*-entry block this eigenvector is necessarily an eigenvector of *KG*, from a simple algebraic argument. Hence, *KG* has an eigenvector whose first *q* entries are nil in this case. We have thus shown necessity and sufficiency.

Remark. Very often, even if the DIN is secure (unobservable) according to the binary notion above, the adversary may be able to glean critical information about the parts of the network's dynamics. A natural way to formalize this possibility is to study whether *features* or projections of the state dynamics can be estimated. It is a consequence of our formulation that security of a feature (specifically, linear projection) of the initial state is achieved, if and only if a vector can be found that is not in the null space of the projection matrix but is in the unobservable space (i.e., the null space of the *observability matrix*). This algebraic condition can naturally be translated to a spectral condition on *A*, and in turn to spectral and graphical conditions on *KG*, see Roy et al. (2012) for a similar analysis. Two special cases are worth mentioning. First, if only the initial positions of all the vehicles (or only the initial velocities of the vehicles) are sought by adversary, identical conditions for security are obtained as for the full-state case. Meanwhile, if the feature of interest is the target location, then the estimation goal is always insecure. We leave a detailed development of spectral/graphical conditions for other features of interest to future work.

Remark. Estimation of the initial state allows recovery of the vehicles' trajectories, since the network model is assumed to be known. We omit details in the interest of space.

Next, with the aim of facilitating design of secure network dynamics, we present two graphical conditions that are sufficient for security and one that guarantees insecurity. We omit proofs for these results, which are based on constructing eigenvectors of *KG*, see Roy et al. (2012) and Xue et al. (2011) for similar analyses. We begin with a symmetry-based condition, requiring the following definition. A graph is said to have *symmetrically-interconnected components*, if two disjoint subsets V_A and V_B of the vertices can be found with the following properties (see Diestel, 2000 for a graph theory reference): (1) The induced subgraphs on these subsets are isomorphic. (2) Equivalent vertices in each subgraph connect symmetrically; that is, if $i \in V_A$ and $j \in V_B$ are equivalent, and $\widehat{i} \in V_A$ and $\widehat{j} \in V_B$ are equivalent, then the edge-weight $\Gamma_{i,\widehat{j}}$ equals $\Gamma_{j,\widehat{i}}$. (3) Equivalent vertices in each subgraph connect identically to the remaining vertices in the graph, say V_C. That is, if $i \in V_A$ and $j \in V_B$ are equivalent, and $q \in V_c$, then $\Gamma_{i,q} = \Gamma_{j,q}$. Conceptually, a graph has symmetrically-interconnected components, if it has two subgraphs that look identical and also are identically connected to the rest of the graph (see Xue et al., 2011 for an illustration). If the interaction graphs have symmetrically-interconnected components, then the vehicle team has certain indistinguishable dynamics that yield security, as formalized in the following corollary:

Corollary 1. *Consider a DIN that achieves a static target-tracking task, and whose graph has symmetrically-interconnected components. Also, assume that the adversary does not measure the positions of any vehicles corresponding to vertices in V_A and V_B. Then the initial state is secure.*

We note that our notion of symmetrically-interconnected components generalizes the notions of symmetry and equitable partitions considered in Martini, Egerstedt, and Bicchi (2010) and Rahmani et al. (2010) toward directional graphs; the relationship between such symmetry structures and unobservability are dual to the analyses of uncontrollability considered in Rahmani et al. (2010), albeit for different dynamical models.

The second condition for security is based on the presence of certain "unmeasured" bottlenecks in the graph, that prevent estimation by the adversary. In particular, we say that a subset of *m* vertices, say V_D, in the network graph constitute a *bottleneck*, if the vertices in V_D (1) are independent, (2) have identical self-loops, and (3) in total have $r < m$ downstream neighbors (i.e., there are directed edges from the vertices in V_D to $r < m$ other vertices).

Corollary 2. *Consider a DIN that achieves the target-tracking task, whose graph has a bottleneck vertex set V_D. If the adversary does not measure any vehicle positions corresponding to vertices in V_D, then the initial state is secure.*

The bottleneck-set concept is a specification of the concept of dilations considered in e.g. Siljak (1991).

Finally, let us present a sufficient condition for insecurity when the interaction graph is a strongly-connected tree, i.e. there is a unique directed path between each (ordered) pair of vertices.

Corollary 3. *Consider a DIN that achieves the target-tracking task. Assume that the interaction graph Γ is a strongly connected tree. Also, say that the graph vertices corresponding to the vehicles measured by the adversary include all the leaves of the tree. Then the initial state $\varphi(0)$ is insecure.*

4. The noisy case: security level analysis

In this section, the security level matrix and attendant scalar performance measures are characterized, in the case that the adversary's measurements are corrupted by noise. We assume throughout that the network is insecure, i.e. the adversary would be able to estimate the initial state if its measurements were not corrupted by noise. To begin, we recall that the error covariance matrix COV for the MVU estimate of the initial state (which is the security level matrix) is $\mathrm{COV} = \sigma^2 M(0,t_f)^{-1}$, where the *observability Grammian* $M(0,t_f)$ is given by $M(0,t_f) = \int_0^{t_f} (C\Phi(t,0))^T (C\Phi(t,0))\,dt$, and $\Phi(t,0) = e^{At}$ is the transition matrix of the closed-loop dynamics (e.g., Stengel, 1994). (Recall that σ is the intensity of the observation noise.)

We find it convenient to express the security level matrix in terms of the spectrum of the closed-loop state matrix *A*. Specifically, we use the notation $A = W\Lambda W^{-1}$ for the Jordan decomposition. Here, $W = \begin{bmatrix}\mathbf{w}_1 & \cdots & \mathbf{w}_{2n+1}\end{bmatrix}$, where $\mathbf{w}_k$, $k = 1,\ldots,2n+1$ are the $2n+1$ eigenvectors and generalized eigenvectors of matrix *A*. Also, Λ is the Jordan matrix of *A*. Without loss of generality, we order the eigenvalues and eigenvectors/generalized-eigenvectors as follows: for each $i = 1,\ldots,n$, diagonal entries $2i-1$ and $2i$ of Λ are assumed to be the two eigenvalues $\mu_{i,1}$ and $\mu_{i,2}$ of *A* obtained from eigenvalue λ_i of *KG* (specifically, as the two solutions of $\mu_i^2 - \lambda_i\mu_i - \alpha\lambda_i = 0$). Correspondingly, $\mathbf{w}_{2i-1}$ and $\mathbf{w}_{2i}$ are assumed to be the eigenvectors or generalized eigenvectors of *A* associated with $\mu_{i,1}$ and $\mu_{i,2}$. The bottom right diagonal entry of Λ is assumed to be the zero eigenvalue of *A*, and the corresponding eigenvector $\mathbf{w}_{2n+1}$ is $\begin{bmatrix}\mathbf{1}_n^T & \mathbf{0}_n^T & 1\end{bmatrix}^T$. In this notation, $\Phi(t,0) = We^{t\Lambda}W^{-1}$.

Let us now develop several bounds of the security level scalars, in terms of the spectrum of *KG*. We begin with a characterization of the maximum eigenvalue $\gamma_{\max}(\mathrm{COV})$ of the error covariance matrix, which indicates the adversary's worst-case performance in estimating a unitary linear projection (feature) of the initial state,

and also lower-bounds the total squared error in an initial-state estimate (Xue et al., 2011). The result shows that the security level is high, if the network has a mode that is either fast or has small components at the adversary's observation location. For ease of presentation, we focus on the case that the eigenvalues of A are real and non-defective (as results from the low-gain result presented in Section 2). We obtain the following:

Theorem 2. *Consider a DIN whose controller has been designed to achieve tracking. Specifically, assume that the gain K has been selected so that the eigenvalues of KG are real, distinct, and negative. Further assume that α is chosen sufficiently small (i.e., $0 < \alpha \leq \min_i(-\frac{\lambda_i}{4})$) so that the nonzero eigenvalues of A are all real and negative. Then, we have that $\gamma_{\max}(\mathrm{COV}) \geq \max_{i\in\{1,\ldots,n\}} \frac{-2\sigma^2\mu_i(1+\mu_i^2)}{\sum_{j=1}^{q} v_{ij}^2}$, where $\mu_i = \frac{\lambda_i - \sqrt{\lambda_i^2+4\alpha\lambda_i}}{2}$; v_{ij} is the jth entry in the eigenvector $\mathbf{v}_i$ of KG (normalized to unity length), associated with the eigenvalue λ_i.*

Proof. Invoking the relationship between eigenvalues of a matrix and its inverse, and then using the Courant–Fisher theorem for symmetric matrices, we obtain $\gamma_{\max}(\mathrm{COV}) = \frac{\sigma^2}{\gamma_{\min}(M(0,t_f))} = \frac{\sigma^2}{\min_{\mathbf{g}:\|\mathbf{g}\|=1} \mathbf{g}^T M(0,t_f)\mathbf{g}}$. Since $\|\mathbf{v}_i\| = 1$, then $\frac{1}{\sqrt{1+\mu_i^2}}\begin{bmatrix}\mathbf{v}_i^T & \mu_i\mathbf{v}_i^T & 0\end{bmatrix}$ is also a vector with unit norm (where μ_i are the roots of $\mu_i^2 - \lambda_i\mu_i - \alpha\lambda_i$, for $i = 1, \ldots, n$). Therefore, for any $i \in \{1, \ldots, n\}$, we have

$$\begin{aligned}&\min_{\mathbf{g}:\|\mathbf{g}\|=1} \mathbf{g}^T M(0,t_f)\mathbf{g}\\ &\leq \frac{1}{1+\mu_i^2}\begin{bmatrix}\mathbf{v}_i^T & \mu_i\mathbf{v}_i^T & 0\end{bmatrix} M(0,t_f) \begin{bmatrix}\mathbf{v}_i^T & \mu_i\mathbf{v}_i^T & 0\end{bmatrix}^T. \qquad (2)\end{aligned}$$

Noticing that $\begin{bmatrix}\mathbf{v}_i^T & \mu_i\mathbf{v}_i^T & 0\end{bmatrix}^T$, is an eigenvector of A (i.e., either $\mathbf{w}_{2i-1}$ or $\mathbf{w}_{2i}$ for the two possible roots μ_i), we obtain $W^{-1}\begin{bmatrix}\mathbf{v}_i^T & \mu_i\mathbf{v}_i^T & 0\end{bmatrix}^T = \mathbf{e}_k$ with $k = 2i-1$ or $2i$, where $\mathbf{e}_k$ is an indicator vector with length $2n+1$ whose kth entry is nonzero. With some algebraic effort, we have:

$$\begin{aligned}&\begin{bmatrix}\mathbf{v}_i^T & \mu_i\mathbf{v}_i^T & 0\end{bmatrix} M(0,t_f) \begin{bmatrix}\mathbf{v}_i^T & \mu_i\mathbf{v}_i^T & 0\end{bmatrix}^T\\ &= \int_0^{t_f} \left(CWe^{t\Lambda}W^{-1}\begin{bmatrix}\mathbf{v}_i\\ \mu_i\mathbf{v}_i\\ 0\end{bmatrix}\right)^T \left(CWe^{t\Lambda}W^{-1}\begin{bmatrix}\mathbf{v}_i\\ \mu_i\mathbf{v}_i\\ 0\end{bmatrix}\right) dt\\ &= \int_0^{t_f} \left(CWe^{t\Lambda}\mathbf{e}_k\right)^T \left(CWe^{t\Lambda}\mathbf{e}_k\right) dt\\ &\leq \int_0^{t_f} e^{2\mu_i t}\sum_{j=1}^{q} v_{ij}^2 dt \leq \frac{1}{-2\mu_i}\sum_{j=1}^{q} v_{ij}^2. \qquad (3)\end{aligned}$$

Finally, we obtain that $\gamma_{\max}(\mathrm{COV}) \geq \frac{-2\sigma^2\mu_i(1+\mu_i^2)}{\sum_{j=1}^{q} v_{ij}^2}$ for any i, where again we are checking the two possible roots μ_i for each i. We note that the maximum value for the right side of the inequality is achieved with $\mu_i = \frac{\lambda_i - \sqrt{\lambda_i^2+4\alpha\lambda_i}}{2}$, and the theorem thus follows.

Remark. The bound can be trivially extended to the case where A has complex or defective eigenvalues, by restricting the maximization to be over the real eigenvalues of multiplicity 1. However, stronger bounds can be obtained by also searching over complex eigenvectors and generalized eigenvectors. We omit the details due to space constraints, but note that stronger bounds also depend on real parts of the eigenvalues of KG.

The next two results are concerned with characterizing security levels of important linear projections (features) of the initial state vector, specifically the target location and the initial positions/velocities. They are important in that they distinguish the adversary's ability to estimate the target location (infinitely accurate estimates are possible, asymptotically) with his/her ability to estimate other initial states. They also complement Theorem 2, in that they are concerned with the easiest-to-estimate characteristics rather than the hardest-to-estimate ones. We omit the proofs and refer to these results as lemmas, since they are similar in flavor to Theorem 2. The first result formalizes that the target-location feature becomes less secure as the observation time horizon increases.

Lemma 1. *Consider a DIN for which the static target-tracking task is achieved. The variance in the MVU estimate of the target location approaches $\frac{\sigma^2}{qt_f}$ when t_f becomes large.*

We see that, asymptotically, the error variance is the same as if an MVU estimate from qt_f independent noisy observations of the target location itself were computed. The next result quantifies minimum security-level scalars (i.e., best possible estimation performance) associated with position and velocity estimates:

Lemma 2. *Consider a DIN for which the tracking task has been achieved. Specifically, assume that K has been selected so that the eigenvalues of KG are real, distinct, and negative. Further assume that α is chosen sufficiently small ($\alpha \leq \min_i(-\frac{\lambda_i}{4})$), so that the nonzero eigenvalues of A are all real and negative. Consider estimation of any unitary linear projection of the initial positions and velocities. For a sufficiently long observation horizon, the adversary can estimate one such feature with variance at most $\min_{i:\mu_i\neq 0} \frac{-2\sigma^2\mu_i(1+\mu_i^2)}{\sum_{j=1}^{q} v_{ij}^2}$, where $\mu_i = \frac{\lambda_i - \sqrt{\lambda_i^2+4\alpha\lambda_i}}{2}$, v_{ij} is the jth entry in the eigenvector $\mathbf{v}_i$ of KG (associated with its eigenvalue λ_i).*

The above bounds make explicit the dependence of security levels on certain dominant eigenvalues of KG, and the corresponding eigenvectors' components associated with the measurement locations.

References

Alpcan, T., & Basar, T. (2003). A game-theoretic approach to decision and analysis in network intrusion detection. In *Proceedings of the 2003 IEEE conference on decision and control*. Maui, HI, December.

Amin, S., Cardenas, A., & Sastry, S. (2009). Safe and secure networked control systems under denial-of-service attacks. In *Lecture notes in computer science series: Vol. 5469. Hybrid systems: computation and control* (pp. 31–45).

Diestel, R. (2000). *Graduate texts in mathematics: Vol. 173. Graph theory*. New York: Springer-Verlag.

Le Ny, J., & Pappas, G. J. (2012). Differentially private filtering. In *Proceedings of IEEE 51st conference on decision and control*. Maui, HI, December.

Liu, Y., Ning, P., & Reiter, M. K. (2011). False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security*, *14*(1). Art. 13.

Locatelli, A., & Schiavoni, N. L. M. (2012). Fault-tolerant pole-placement in double-integrator networks. *IEEE Transactions on Automatic Control*, *57*(11), 2912–2917.

Martini, S., Egerstedt, M., & Bicchi, A. (2010). Controllability analysis of multi-agent systems using relaxed equitable partitions. *International Journal of Systems, Control and Communications*, *2*(1–3), 100–121.

Mo, Y., & Sinopoli, B. (2009). Secure control against replay attacks. In *Proceedings of the 2009 Allerton conference on communication, control, and computing*. Urbana, IL, September 30–October 2.

Pasqualetti, F., Bicchi, A., & Bullo, F. (2012). Consensus computation in unreliable networks: a system theoretic approach. *IEEE Transactions on Automatic Control*, *57*(1), 90–104.

Pasqualetti, F., Dorfler, F., & Bullo, F. (2011). Cyber-physical attacks in power networks: models, fundamental limitations and monitor design. In *IEEE conf. on decision and control and European control conference* (pp. 2195–2201). Orlando, FL, USA, December.

Rahmani, A. R., Ji, M., Mesbahi, M., & Egerstedt, M. B. (2010). Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM Journal on Control and Optimization*.

Ren, W., & Beard, R. W. (2005). Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, *50*(5), 655–661.

Roy, S., & Chen, C.-W. (2013). State detection from local measurements in network synchronization processes. *International Journal of Control*, *86*(9).

Roy, S., Saberi, A., & Herlugson, K. (2004). Formation and alignment of distributed sensing agents with double-integrator dynamics and actuator saturation. In S. Phoha (Ed.), *Sensor network operations*.

Roy, S., Xue, M., & Das, S. K. (2012). Security and discovery of spread dynamics in cyber-physical networks. *IEEE Transactions on Parallel and Distributed Systems*, *23*(9), 1694–1707 [special issue on Cyber Physical Systems].
Shames, I., Texiera, A. H., Sandberg, H., & Johansson, K. H. (2010). Distributed fault detection for interconnected second-order systems with application to power networks. In *First workshop on secure control systems*. Stockholm, Sweden, April.
Siljak, D. D. (1991). *Decentralized control of complex systems*. Boston: Academic Press.
Sou, K. C., Sandberg, H., & Johansson, K. H. (2012). Detection and identification of cyber-attacks in power-system state estimation. In *Proceedings of the 2012 American control conference*. Montreal, Canada, June.
Stengel, R. F. (1994). *Optimal control and estimation*. Dover Publications.
Sundaram, S., & Hadjicostis, C. N. (2011). Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Transactions on Automatic Control*, *56*(7), 1495–1508.
Telerius, H., Varagnolo, D., Baquero, C., & Johansson, K. H. (2013). Fast distributed estimation of probability mass functions over anonymous networks. In *Proceedings of IEEE 52nd conference on decision and control*. Florence, Italy, December.
Texiera, A., Sandberg, H., & Johansson, K. H. (2010). Networked control systems under cyber attacks with applications to power networks. In *Proceedings of the 2010 American control conference*. Baltimore, MD, June 30–July 2.
Wan, Y., & Roy, S. (2009). On estimation of network time constants from impulse–response data. In *Proceedings of the 2009 IEEE conference on decision and control*. Shanghai, China, December.
Xue, M., Roy, S., Wan, Y., & Das, S. K. (2011). Security and vulnerability of cyber-physical infrastructure networks: a control-theoretic perspective. In S. K. Das, K. Kant, & N. Zhang (Eds.), *Handbook on securing cyber-physical critical infrastructure* (Chapter 1).
Zhu, M., & Martinez, S. (2011). Stackelberg-game analysis of correlated attacks in cyber-physical systems. In *Proceedings of the 2011 American control conference*. San Francisco, CA, June.

Mengran Xue received B.S. in Electrical Engineering from the University of Science and Technology of China and Ph.D. in Electrical Engineering from Washington State University, in 2007 and 2012 respectively. She then completed a postdoctoral research appointment at the University of Michigan. Her research is focused on uncertainty modeling and evaluation in complex networks, with applications to robotics, transportation systems, and power networks.

Wei Wang completed Master's degree in Electrical Engineering at Washington State University.

Sandip Roy received his Ph.D. degree in Electrical Engineering from Massachusetts Institute of Technology in 2003. Currently, he is an Associate Professor of Electrical Engineering, and an affiliate faculty in the School of Global Animal Health, at Washington State University. He is interested in estimation, control, and management of complex networks.