Proceedings of the 1st IFAC Workshop on Estimation and Control of Networked Systems, Venice, Italy, September 24-26, 2009

Identification and Infiltration in Consensus-type Networks ⋆

Airlie Chapman, Marzieh Nabi-Abdolyousefi and Mehran Mesbahi *

* *Department of Aeronautics and Astronautics, University of Washington, Seattle, (e-mail: airliec, mnabi@u.washington.edu; mesbahi @aa.washington.edu,)*

Abstract: This paper examines the system dynamics of a controlled networked multi-agent system, operating with a consensus-type algorithm, that is under the influence of attached node(s). We introduce an identification scheme, involving excitation and observation of the network by the attached node(s), that identifies the spectrum of the underlying system matrix - in this case the influenced network's modified graph Laplacian. Following this identification, infiltration node(s) are then attached to a set of nodes in the network with the objective of sabotaging the network by delivering a constant mean control signal. The spectrum of the modified graph Laplacian provides bounds on the convergence of the system states due to infiltration, quantifying the network's security. We also derive bounds on the effectiveness of the network infiltration in a more general setting by examining the controllability gramian of the infiltrated consensus-type coordination algorithm.

Keywords: Consensus problem; Graph theory; Coordinated control and estimation over networks, Network security

1. INTRODUCTION

Consensus-type algorithms provide effective means for distributed information-sharing and control for networked multi-agent systems, in settings such as multi-vehicle control, formation control, swarming, and distributed estimation; see for example, [Olfati-Saber et al. (2007); Tanner et al. (2004); Jadbabaie et al. (2003); Hatano and Mesbahi (2005)]. This paper addresses the problem where the dynamics of a network, which has adopted a consensus-type algorithm, can be influenced and observed by *infiltrator* nodes attaching to the network. The additional nodes, ignoring consensus rules, will influence the system dynamics compared to the *unforced* networked system. In such a setting, we first propose an *identification phase* whereby identification (ID) excitation and observing node(s) attach to the network in such a way that the resulting input-output system is controllable and observable. System identification methods are then utilized to find the spectrum of the underlying system dynamics. This spectrum is used to bound the convergence cost of the *infiltration phase* where infiltration (IF) excitation node(s) attach to an *arbitrary* set of nodes in the network. Performing identification on multiple networks, based on the discovered convergence bounds, the IF excitation node(s) can then attached to a *susceptible* network with the objective of sabotaging the network by delivering a constant mean control signal. This approach is then generalized by examining the controllability gramian of the infiltrated network and deriving bounds for the *security* of the underlying coordination algorithm.

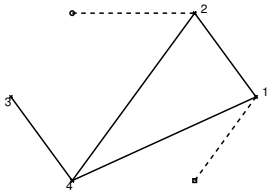
The current work is part of a more general effort that aims to identify fundamental bounds on the security of coordination algorithms for dynamic systems when infiltrated by an adversary. As such, our work is related to a number of other research works, such those in computer network security [Gueye and Walrand (2008)], spread of epidemics [Bloem et al. (2007); Wang et al. (2003)], and predator/prey swarming [Olfati-Saber (2006)]. The paper presents the problem of network infiltration from the infiltrator's perspective complementing work on infiltration detection such as Fagiolini et al. (2007). Rahmani et al. (2009) have examined the graph properties of a consensus-type network that allow system controllability where a single anchor node already within the network is controlled. One of our contributions in this work is extending the convergence models examined by Rahmani et al. (2009) with an emphasis on the minimum and maximum non-zero eigenvalues of the graph Laplacian in the context of network security. Moreover, to the authors' knowledge this is the first application of system identification techniques on consensus-type networks.

⋆ This work has been supported by AFOSR FA9550-09-1-0091.

2. BACKGROUND AND MODEL

We provide background on constructs and models that will be used in this paper, including abbreviated descriptions on graphs and the consensus protocol, in its unforced and infiltrated versions.

An undirected graph $G = (V, E)$ is defined by a vertex (or node) set V with cardinality n and an edge set E comprised of a pairs of nodes, where node i and j are adjacent if $\{i, j\} \in E$. The degree d_i of node i is the number of its adjacent nodes. The degree matrix $\Delta(G)$ is a diagonal matrix with d_i at position (i, i). The adjacency matrix is a symmetric matrix with $[\mathcal{A}(G)]_{ij} = 1$ when $\{i, j\} \in E$ and $[\mathcal{A}(G)]_{ij} = 0$ otherwise. The combinatorial Laplacian is defined as $L(G) = \Delta(G) - \mathcal{A}(G)$ which is



$$A(G) = \begin{bmatrix} -3 & 1 & 0 & 1 \\ 1 & -2 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

$$B(G) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; C(G) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^T$$

Fig. 1. Network graph with an *excitation* node attached to node 1 (square) and an *observing* node (circle) attached to node 2, leading to an altered Laplacian $A(G)$ and input and output matrices $B(G)$ and $C(G)$ of model (2).

a (symmetric) positive semi-definite. The adjacency and Laplacian representations of a graph are unique to that graph; the degree matrix is not. The analysis of this paper will be concerned with the spectrum of the graph Laplacian. That spectrum is assumed to be ordered as $0 = \lambda_1(G) \leq \lambda_2(G) \leq \cdots \leq \lambda_n(G)$, where for brevity we have used $\lambda_i(G)$ instead of $\lambda_i\left(L\left(G\right)\right)$.

Consider now x_i be the state of the i-th node of the first order dynamic system. The continuous-time consensus protocol is defined as $\dot{x}_i(t) = \sum_{\{i,j\}\in E}\left(x_j(t) - x_i(t)\right)$, which in a compact form is

$$\dot{x}(t) = -L(G)x(t), \tag{1}$$

with $L(G)$ being the Laplacian of the underlying interaction topology. From the definition of the graph Laplacian all rows of $L(G)$ sum to zero, and $\lambda_1(G) = 0$ with the corresponding eigenvector as $v_1 = \mathbf{1}^T = [1, \ldots, 1]^T$. Subsequently, when G is connected, it can be deduced that $x = \alpha\mathbf{1}$ is a unique globally exponentially stable equilibrium of the process (1) with α as the average of the initial states [Godsil and Royle (2001); Olfati-Saber et al. (2007)].

We now introduce a model of *infiltrated* consensus, whereby r *excitation* nodes each attach to a graph node in set $R \subseteq V$, $|R| = r$ using control signals $u(t)$, and s *observing* nodes each attach to a graph node in set $S \subseteq V$, $|S| = s$ monitoring output signals $y(t)$. The resulting input-output dynamics then assumes the form,

$$\dot{x}(t) = A(G)x(t) + B(G)u(t),\; y(t) = C(G)x(t), \tag{2}$$

where $A(G) = -\left(L(G) + D\right)$, $B(G) = \left[\, e_{R(1)} \;\cdots\; e_{R(r)} \,\right]$, and $C(G) = \left[\, e_{S(1)} \;\cdots\; e_{S(s)} \,\right]^T$. The matrix D is a diagonal with $[D]_{ii} = 1$ for all $i \in R$ and $[D]_{ij} = 0$, otherwise; e_i is a vector whose i-th component is 1 and all other components set to zero. We note that in discrete form the corresponding system dynamics has the form

$$x(k+1) = A_d x(k) + B_d u(k),\; y(k) = C_d x(k). \tag{3}$$

An example of this system model is shown in Figure 1. This model is applied to the identification (ID) phase with r-ID excitation nodes and s-ID observing nodes with sets $R \subseteq V$, $|R| = r$ and $S \subseteq V$, $|S| = s$, and the infiltration (IF) phase with $\tilde{r}$-IF excitation nodes and $\tilde{s}$-IF observing nodes with sets $\tilde{R} \subseteq V$, $\left|\tilde{R}\right| = \tilde{r}$ and $\tilde{S} \subseteq V$, $\left|\tilde{S}\right| = \tilde{s}$.

3. IDENTIFICATION PHASE

We now briefly describe two identification methods, namely, the *Iterative Prediction-Error Minimization Method* and *Subspace Identification Method*, each of which produce a realization of the system (3) namely

$$x(k+1) = \tilde{A}_d x(k) + \tilde{B}_d u(k),\; y(k) = \tilde{C}_d x(k) + \tilde{D}_d u(k), \tag{4}$$

with continuous relatives $\tilde{A}, \tilde{B}, \tilde{C}$ and $\tilde{D}$. This system exhibits the same input-output characteristics of the original system - and as such -a similarity transformation exists between the ID and true models and $\lambda_i(-\tilde{A}) = \lambda_i(-A(G))$ but generally differing eigenvectors (noting that as $A(G)$ is positive semidefinite $\lambda_i(-A(G)) = -\lambda_{n+1-i}(A(G))$. Both methods assume that the number of nodes in the system is known (though examining the output accuracy of models with differing n a suitable n can be approximated), that via the ID excitation node(s) the system is controllable and via the ID observing node(s) the system is observable (this will be discussed further later).

The *Iterative Prediction-Error Minimization Method* commences by restating model (3) as $A_d(q)y(k) = B_d(q)u(k)$, where $A_d(q) = 1 + a_1 q^{-1} + \cdots + a_n q^{-n}$ and $B_d(q) = b_1 q^{-1} + \cdots + b_n q^{-n}$. The model parameters are estimated by comparing the actual output $y(k)$ at time k and predicted output $\tilde{y}(k|k-1)$ at time k given $k-1$. The output predictor is constructed as

$$\tilde{y}(k|k-1) = [\, -y(k-n) \;\ldots\; -y(1) \;\; u(k-r) \;\ldots\; u(1) \,]\,\theta,$$

where $\theta = [a_1, \ldots, a_n, b_1, \ldots b_n]^T$. The optimal vector θ is then formed by minimization of the mean squared error, i.e., by the solution of

$$\min_{\theta} \sum_{k=1}^{N} \left\| y(k) - \tilde{y}(k|k-1) \right\|_2^2 .$$

Provided with the optimal θ, the characteristic equation of the system can be formed and a system realization as in model (4) may be found.

In the *Subspace Identification Method* an important role is played by Hankel matrices that can be constructed solely from input-output data. This topic is explained in more details in Ljung (1999) and Soderstrom and Stoica (1989).

As an example, identification on the network graph shown in Figure 1 using both methods found the eigenvalues of the modified system to be -4.1796, -3.4882, -1.1538, -0.1783, which is accurate to a mean squared error of 1×10^{-10}.

3.1 Identifiable Features of the Network

The ID model $\tilde{A}$ is generally insufficient to produce the original system, except when the number of ID excitation and ID observation nodes are sufficiently large i.e. $2\left(r + s\right) + 1 \geq n$. Therefore in general we have the eigenvalue of $A(G)$ without explicit information about its eigenvectors. Further the eigenvalues of the Laplacian $L(G)$, even with its distinct structure, are generally insufficient to reproduce the initial graph. For example, cospectral graphs are indistinguishable by their spectra. In the meantime, the eigenvalues of the matrix $A(G)$ provide bounds on the eigenvalues of the Laplacian which can help discern important features of the underlying network and, in particular, with respect to their *security* features.

Proposition 1. The eigenvalues of the system matrix $-A(G)$ of model (2) produced by system ID, using r-ID excitation nodes, place the following bounds on the eigenvalues of the graph Laplacian $L(G)$; $\lambda_i(L(G)) \leq \lambda_i(-A(G))$ (for $r < i$), $\lambda_{i-r}(-A(G)) \leq \lambda_i(L(G))$ (for all r) and $\lambda_i(-A(G)) \leq \lambda_i(G) + 1$ (for all r).

Proof. As $-A(G)$ is the sum of two positive semidefinite matrices $L(G)$ and D, where D has rank r and as such $-A(G)$ is positive semidefinite. Subsequently, by the Interlacing theorem [Horn and Johnson (1990)] the first two bounds follow. By Weyl's theorem [Horn and Johnson (1990)], $\lambda_i(-A(G)) = \lambda_i(L(G) + D) \leq \lambda_i(G) + \lambda_n(D) = \lambda_i(G) + 1$.

Proposition 2. Consider the eigenvalues of $-A(G)$ of model (2) using r-ID excitation nodes. Then

(a) The number of edges $|E|$ in the graph is
$$|E| = \frac{1}{2}\left(\sum_{i=1}^{n} \lambda_i(-A(G)) + r\right).$$
(b) If there is a zero eigenvalue of $A(G)$, the number of disconnected components $\chi(G)$ is
$$\textbf{co-rank}(A(G)) + r - 1.$$
For $r = 1$, the graph is disconnected if and only if $\lambda_1(-A(G)) = 0$.
(c) The ordered degrees of the graph $d_1 \leq \cdots \leq d_n$ satisfy
$$\sum_{i=1}^{k} d_i + \alpha(k) \geq \sum_{i=1}^{k} \lambda_i(-A(G))$$
with equality for $k = n$, where $\alpha(k) = k$ for $k = 1, \ldots r$ and $\alpha(k) = r$ otherwise.

Proof. Since $|E| = \frac{1}{2}\sum_{i=1}^{n} \lambda_i(G)$,
$$\sum_{i=1}^{n} \lambda_i(-A(G)) = \mathrm{tr}(-A(G)) = \mathrm{tr}(L(G) + D) = \sum_{i=1}^{n} \lambda_i(G) + r.$$
In the meantime, from Godsil and Royle (2001), $\chi(G) = \textbf{co-rank}(L(G)) - 1$; the (a)-(b) part of the proposition now follows by applying Proposition 1. For (c) ordering the diagonals of $A(G)$ as $\beta_1 \leq \cdots \leq \beta_n$, $\beta_i = d_i + 1$ for r elements, otherwise $\beta_i = d_i$. Using Theorem 4.3.26 from Horn and Johnson (1990), the vector of diagonals of $A(G)$ ordered as $\beta_1 \leq \cdots \leq \beta_n$ majorizes the vector of eigenvalues of $A(G)$. And so $\sum_{i=1}^{k} d_i + \alpha(k) \geq \sum_{i=1}^{k} \beta_i \geq \sum_{i=1}^{k} \lambda_i(-A(G))$.

We note that in the above discussion we have assumed knowledge of the number of nodes as well as controllability and observability. One can envisage consensus-type networks which are designed about a set of agents able to control all nodes in the network - for example a set of human operators controlling a robotic swarm. The identification nodes may not have the convenience of these key nodes, in which case, we propose a verification procedure to check the identification results by exploiting the structure of the matrix $A(G)$. Specifically, if the system is controllable/observable with the correct value of n, there should exist a similarity transformation T such that $\tilde{A} = TA(G)T^{-1}$, $\tilde{B} = TB(G)$ and $\tilde{C} = C(G)T^{-1}$. As such $\tilde{C}\tilde{A}\tilde{B} = C(G)A(G)B(G)$, and as the number of ID excitation and observing nodes are known, so too are the matrices $C(G)$ and $B(G)$. The matrices $C(G)$ and $B(G)$, in the meantime, serve to truncate $A_{n\times n}$ into a submatrix $\bar{A}_{r\times m}$ formed from r-excitation nodes and m-observing nodes. If $\bar{A}_{r\times m}$ is not a possible submatrix of $A(G) = -(L(G) + D)$, then the system can be assumed not controllable and/or observable, requiring increasing or modifying the excitation / observing nodes or varying the approximated value n. Other verifications methods based on the properties in Proposition 2 can also be utilized.

4. INFILTRATION PHASE

We now propose two infiltration schemes for consensus-type coordination algorithms. The first approach is based on relating the spectrum of $L(G)$ to the convergence cost of steering the network by the infiltration nodes with a signal of constant expected value, e.g., white noise with a given mean. We then provide a more general approach that examines the effectiveness of infiltration based on the controllability gramian.

4.1 Constant Mean Intrusion Protocol

The Constant Mean Intrusion Protocol adopts a naive approach to network intrusion as compared to the Identification Phase. In this case, the IF excitation nodes merely attempts to steer the system to a common state u_c, deterministically or in the mean and no IF observing node is required. We note that when the network is driven by stochastic signal with a constant expected value $\mathbf{E}(u) = u_c\mathbf{1}$, the expected value of the node states $\mathbf{E}(x)$ can be modeled as a standard consensus problem with $u = u_c\mathbf{1}$ as
$$\frac{d}{dt}\mathbf{E}(x) = \mathbf{E}(Ax + Bu) = A\mathbf{E}(x) + Bu_c\mathbf{1}.$$
Hence, intruding the network with a random signal with a constant expected value is equivalent- in the mean- with intruding the network with a constant signal.

Before continuing we state an auxiliary result.

Proposition 3. The matrix $A(G)$ of model (2) is negative definite (and so invertible) if the original graph is connected.

Proof. Consider the influence on a single node attached to the infiltration node. In this case, $D = bb^T$, where D is diagonal and $\|b\|_2 = 1$. We assume $L(G) = Q\Lambda Q^T$, with $Q^TQ = QQ^T = I$, Λ is a diagonal matrix of eigenvalues of $L(G)$. Thus,$-A(G) = L(G) + bb^T = Q\Lambda Q^T + Qzz^TQ^T = Q\left(\Lambda + zz^T\right)Q^T$, where $b = Qz$. Thus, if $\Lambda + zz^T = X\tilde{\Lambda}X^T$, then $-A(G) = \tilde{Q}\tilde{\Lambda}\tilde{Q}^T$, where $\tilde{Q} = QX$. Let $z = [\zeta_1, \ldots, \zeta_n]^T$. If $\zeta_i = 0$, then $\lambda_i(G) = \lambda_i(-A(G))$ and the corresponding eigenvector remains unchanged since $\left(\Lambda + zz^T\right)e_i = \Lambda e_i + z\zeta_i = d_ie_i$ as $\zeta_i = 0$. Examining the first eigenvector of a connected graph, $z_1 = e_1^TQ^Tb = \frac{1}{\sqrt{n}}$, therefore for a finite graph $0 = \lambda_1(G) < \lambda_1(-A(G))$. Consequently, the matrix $A(G)$ is positive definite for $r = 1$. If additional nodes are influenced to form $\bar{A}(G)$, then as each positive definite matrix is added, by the interlacing theorem [Horn and Johnson (1990)], $0 < \lambda_i(-A(G)) \leq \lambda_i(-\bar{A}(G))$ for $i = 1, \ldots, n$, and so $\bar{A}(G)$ is also positive definite.

We note that the state $u_c\mathbf{1}$ is reachable for the infiltrated system (2) as $A(G)$ is negative definite (Proposition 3).

Next, we examine the cost from uniform control u_c over an infinite horizon for steering the network to consensus on the intruders' injected signal from an arbitrary initialization. In this case, the convergence cost, as a function of the initial state x_0, is

$$J(\tilde{R}, x_0) = \frac{1}{2}\int_0^\infty x(t)^T x(t)dt = -\frac{1}{4}x_0^T A(G)^{-1}x_0,$$

where the infiltrators are connected to node set $\tilde{R}$.

In our subsequent discussion, we will only consider connected graphs as disconnected graphs can be analyzed as the union of their connected components. In order to parametrize the resilience of consensus-type networks to "constant mean infiltration," let us define the minimum $J_{\tilde{r}*}$, maximum $J^*_{\tilde{r}}$, and minimum average $\bar{J}_{\tilde{r}*}$ infiltration convergence costs over all permutations of $\tilde{r}$-network nodes attached to $\tilde{r}$-IF excitation nodes as

$$\begin{aligned} J_{\tilde{r}*} &= \min_{\tilde{R}\in P_{\tilde{r}}} \min_{\|x_0\|=1} J(\tilde{R}, x_0), \\ J^*_{\tilde{r}} &= \max_{\tilde{R}\in P_{\tilde{r}}} \max_{\|x_0\|=1} J(\tilde{R}, x_0), \\ \bar{J}_{\tilde{r}*} &= \min_{\tilde{R}\in P_{\tilde{r}}} \mathbf{E}\left(J(\tilde{R}, x_0)\right), \end{aligned}$$

where $P_{\tilde{r}}$ is the set of all $\binom{n}{\tilde{r}}$ $\tilde{r}$-tuple permutations of the n network nodes. For a graph G with *specific* $\tilde{R}$ we will simply denote J_*, J^* and $\bar{J}$ as the minimum, maximum and average cost, respectively.

Lemma 4. For connected graphs and the infiltration model (2) with $\tilde{r}$-IF excitation nodes,

(a) For $\tilde{r} < n$: $(\lambda_n(G)+1)^{-1} \le 4J_{\tilde{r}*} \le \lambda_n(G)^{-1}$, $\max(1, \lambda_{1+\tilde{r}}(G)^{-1}) \le 4J^*_{\tilde{r}} < \infty$ and

$\sum_{i=1}^{\tilde{r}} (\lambda_i(G)+1)^{-1} + \sum_{i=\tilde{r}+1}^{n} \lambda_i(G)^{-1} \le 4n\bar{J}_{\tilde{r}*}$.

(b) For $\tilde{r} = n$: $J_* = (1/4)(\lambda_n(G)+1)^{-1}$, $J^* = 1/4$, and $\bar{J} = (1/4n)\sum_{i=1}^{n}(\lambda_i(G)+1)^{-1}$.

Proof. Restraining $\|x\|_2 = 1$,$\lambda_n(-A(G))^{-1} = \lambda_1(-A(G)^{-1}) = 4J_{\tilde{r}*} \le 4J_{\tilde{R}}(x_0)$, and $4J_{\tilde{R}}(x_0) \le 4J^*_{\tilde{r}} = \lambda_n(-A(G)^{-1}) = \lambda_1(-A(G))^{-1}$. From Proposition 3 it follows that the quantity $\lambda_1(-A(G))^{-1}$ is defined. For $\tilde{r} = n$, with system matrix $A_{\tilde{r}=n}(G)$, one has $D = I$ and hence $\lambda_i(G) + 1 = \lambda_i(-A_{\tilde{r}=n}(G))$; therefore $\lambda_n(G)+1 = \lambda_n(-A_{\tilde{r}=n})$ and $\lambda_1(G)+1 = 1 = \lambda_1(-A_{\tilde{r}=n})$. For $\tilde{r} < n$, with system matrix $A_{\tilde{r}<n}$, by Interlacing Theorem [Horn and Johnson (1990)], if $\lambda_i(G) \le \lambda_i(-A_{\tilde{r}<n}(G)) \le \lambda_{i+\tilde{r}}(G)$ and $\lambda_i(-A_{\tilde{r}<n}(G)) \le \lambda_i(-A_{\tilde{r}=n}(G)) = \lambda_i(G)+1$. Therefore for $i = 1$, $\lambda_1(-A_{\tilde{r}<n}(G)) \le \lambda_1(G)+1 = 1$, and so $\max(1, \lambda_{1+\tilde{r}}(G)^{-1}) \le \lambda_1(-A_{\tilde{r}<n}(G))^{-1}$. Similarly, for the largest eigenvalue $\lambda_n(G) \le \lambda_n(-A_{\tilde{r}<n}(G)) \le \lambda_n(G)+1$, generating the bounds for $4J_{\tilde{r}*}$. The expected cost for a given set $\tilde{R}$ is $\bar{J} = (1/4n)\sum_{i=1}^{n}\lambda_i(-A(G))^{-1}$; the proposition follows after applying the above bounds.

We assume that the ID excitation node(s) are low cost and therefore more abundant than the IF excitation node(s), i.e., $\tilde{r} < r$. When the system model is not completely recoverable from the eigenvalues of $A(G)$, the spectrum can still be used to assess the network's security.

Theorem 5. Consider a connected graph, where r-ID excitation node(s) are attached to a *specific* set of nodes $R \subseteq V$ forming a modified network specified by $A(G)$ as defined in (2) and $\tilde{r}$-IF excitation node(s) are attached to an *arbitrary* set of nodes $\tilde{R} \subseteq V$. Then

(a) for $r < n$, $\tilde{r} < n$

$$\begin{aligned} &(\lambda_n(-A(G))+1)^{-1} \le 4J_{\tilde{r}*} \\ &\le \min\{(\lambda_n(-A(G))-1)^{-1}, \lambda_{n-r}(-A(G))^{-1} \text{ and} \\ &\max\{1, \lambda_{1+\tilde{r}}(-A(G))^{-1}\} \le 4J^*_{\tilde{r}} < \infty, \text{ and} \end{aligned}$$

(b) for $r + \tilde{r} \le n$

$$\sum_{i=1}(\lambda_i(-A)+1)^{-1} + \sum_{i=\tilde{r}+1}^{n-r}\lambda_i(-A)^{-1} + \sum_{i=n-r+1}^{n}(\lambda_i(-A)-1)^{-1} \le 4n\bar{J}_{\tilde{r}}.$$

(c) for $r + \tilde{r} > n$

$$\sum_{i=1}^{n-r}(\lambda_i(-A(G))+1)^{-1} + \sum_{i=n-r+1}^{\tilde{r}}\lambda_i(-A(G))^{-1} + \sum_{i=\tilde{r}+1}^{n}(\lambda_i(-A(G))-1)^{-1} \le 4n\bar{J}_{\tilde{r}}.$$

Proof. Starting with Lemma 4 and applying Proposition 1, the theorem follows.

Absolute bounds We now show that the n node complete graph K_n and path graph P_n provide, respectively, the smallest minimum cost and largest maximum cost for an n node graph under single node infiltration.

Proposition 6. For the n node complete graph the minimum, maximum and average costs associated with single node infiltration are

$$J_*, J^* = \frac{1}{2}\left(n+1 \pm \sqrt{n^2+2n-3}\right)^{-1}, \text{ and}$$
$$\bar{J} = \frac{1}{4n}\left(\frac{n-2}{n} + J_* + J^*\right).$$

Proof. For a complete graph, $0 = \lambda_1 \le \lambda_2 = \cdots = \lambda_n = n$, and the eigenvector v_1 and v_2 associated with λ_1 and λ_n are, respectively, $v_1 = (1/\sqrt{n})\mathbf{1}_n^T$ and $v_n = [-\frac{1}{\sqrt{n}}(1/2\sqrt{n})\mathbf{1}_{n-1}]^T$. Following the proof of Proposition 3, for $b = e_1$, we have $\zeta_1 = \frac{1}{\sqrt{n}}$, $\zeta_n = \sqrt{\frac{n-1}{n}}$, $\zeta_2, \ldots, \zeta_{n-1} = 0$. As such $\lambda_i(G) = \lambda_i(-A(G)) = n$ for $i = 2, \ldots, n-1$. Solving the secular equation Golub (1973), for the remaining eigenvalues, it follows that $0 = 1 + \sum_{i=1}^{n}(\zeta_i^2)(\lambda_i(L) - \lambda(-A)) = 1 - 1/(n\lambda(-A)) + (n-1)/(n(n-\lambda(-A)))$, with solution of the form,

$$\lambda_1(-A), \lambda_n(-A) = \frac{1}{2}\left(n+1 \pm \sqrt{n^2+2n-3}\right) = \frac{1}{4}J^*, \frac{1}{4}J_*.$$

Proposition 7. For an n node path graph with a single node infiltration at one of the end nodes, the minimum, maximum and average costs are

$$J_* = \frac{1}{8}\left(1+\cos\frac{2\pi}{2n+1}\right)^{-1}, \; J^* = \frac{1}{8}\left(1+\cos\frac{2\pi n}{2n+1}\right)^{-1}, \text{ and}$$
$$\bar{J} = \frac{1}{8n}\sum_{i=1}^{n}\left(1+\cos\frac{2\pi i}{2n+1}\right)^{-1}.$$

Proof. The eigenvalues of modified path graph with the infiltration node attached to node n can be found as $\lambda_i(-A(G)) = 2\left(1+\cos\frac{2\pi i}{2n+1}\right)$; see Yueh (2005). By symmetry, the scenario where the infiltration node is attached

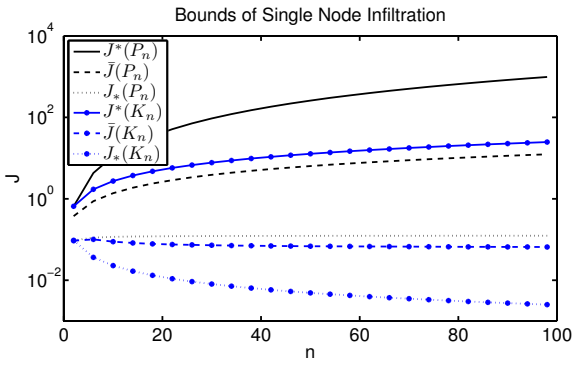


Fig. 2. Bounds of single node infiltration for a n node complete and path graph as described in Propositions 6, 7 and 8.

to the other end of the path, i.e., node 1, will also produce the same set of eigenvalues.

Proposition 8. For the n node graph the cost of single node infiltration is bounded as

$$\frac{1}{2}\left(n+1+\sqrt{n^2+2n-3}\right)^{-1} \leq J_{1*},\, J_1^* \leq \frac{1}{8}\left(1+\cos\frac{2\pi n}{2n+1}\right)^{-1}.$$

Proof. Let G_n be an arbitrary n node graph with its complement graph $\hat{G}_n$, noting that $L(G_n)+L(\hat{G}_n) = L(K_n)$ where K_n is the n node complete graph. Let D be corresponding to single node infiltration; since $L(G_n)+L(\hat{G}_n)+D = L(K_n)+D$ it follows that

$$\lambda_n(L(G_n)+D) \leq \lambda_n(L(K_n)+D) = \lambda_n(L(K_n)+\tilde{D})$$

where $\tilde{D}$ is an arbitrary single node infiltration matrix as the eigenvalues are the same for any single node infiltration of a complete graph. Therefore,

$$J_{1*}(K_n) = \frac{1}{\lambda_n(L(C_n)+\tilde{D})} \leq \frac{1}{\lambda_n(L(G_n)+D)} = J_{1*}(G_n).$$

Now consider a spanning tree $\tilde{G}_n$ of graph G_n, which can be formed by progressively removing edges, $L(\tilde{G}_n)+L(H) = L(G_n)$ where H is the graph of removed edges. Adding a single node infiltration and examining the maximum cost, from $L(\tilde{G}_n)+L(H)+D = L(G_n)+D$ we observe that $\lambda_1(L(\tilde{G}_n)+D) \leq \lambda_1(L(G_n)+D)$. Let T_n be a n node tree with one attached infiltration node, the modifying graph corresponding to D. Construct a new tree T_{2n+1} by mirroring the graph T_n about that infiltration node and treating the infiltration node as a normal node in the new graph T_{2n+1}. Then from Lemma 6 of [Biyikoglu and Leydold (2009)] it follows that $\lambda_2(L(T_{2n+1})) \leq \lambda_1(L(T_n)+D)$. The path P_n is the tree with the smallest λ_2 [Petrovic and Gutman (2002)], i.e.,$\lambda_2(L(P_{2n+1})) \leq \lambda_2(L(T_{2n+1}))$. From Proposition 7 and [Petrovic and Gutman (2002)] $\lambda_2(L(P_{2n+1})) = \lambda_1(L(P_n)+\bar{D})$ where $\bar{D}$ is the special end node infiltration matrix of a path. Combining bounds and letting $\tilde{G}_n = T_n$,

$$\lambda_2(L(P_{2n+1})) \leq \lambda_2(L(T_{2n+1})) = \lambda_1(L(T_n)+D),$$

letting the end path infiltration problem be $\bar{P}_n$ then,

$$J_1^*(G_n) = \frac{1}{\lambda_1(L(G_n)+D)} \leq \frac{1}{\lambda_1(L(P_n)+\bar{D})} = J_1^*(\bar{P}_n).$$

These bounds are plotted in Figure 2.

4.2 Controllability Gramian

The purpose of this Infiltration Phase is to cause maximum disruption to the states of the network with minimum effort on the part of the infiltration nodes. The controllability gramian defined as $P =: \int_0^\infty e^{A\tau}BB^Te^{A^T\tau}d\tau$ proves to be particularly suitable for such an analysis. In general, $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ can be used as indicators of directions that are least and most susceptible to infiltration. In the meantime, $\mathbf{tr}(P)/n$ is a particular good metric for estimating the "average controllability," as the accumulative energy of a n state system at the output from a unit intensity white noise input is $(\mathbf{tr}\,(P))^{1/2}$. The controllability gramian can be obtained by the solution to the Lyapunov equation $AP+PA^T = -BB^T$ and it is under this approach that the following lemma is formed.

Lemma 9. Consider a connected graph and the infiltration model (2) with $\tilde{r}$-IF excitation node(s); then,

(a) $\lambda_{\min}(P) \leq (1/2)\,\lambda_n(-A(G))^{-1}$,

(b) $(1/2)\,\lambda_n(-A(G))^{-1} \leq \lambda_{\max}(P) \leq (1/2)\,\lambda_1(-A(G))^{-1}$,

and (c)

$$\max\left\{\frac{\tilde{r}}{2}\lambda_n(-A(G)), \frac{\tilde{r}^2}{2}\mathbf{tr}(-A(G))^{-1}\right\} \leq \mathbf{tr}(P_{\tilde{r}}) \leq \frac{1}{2}\sum_{i=1}^{\tilde{r}}\lambda_i(-A(G))^{-1}$$

Proof. Let $Q = BB^T$. As A is positive defined and symmetric $\lambda_i^{1/2}(A(G)^TA(G)) = \lambda_i(-A(G))$. We define $Q_\epsilon = Q+\epsilon I$ for $\epsilon > 0$ and note that $\mathbf{tr}(Q) = r+\epsilon n$, $\lambda_i(Q) = \epsilon$ for $i = 1,\ldots,r-1$, and $\lambda_i(Q) = 1+\epsilon$ otherwise. Applying Theorems 2.3, 2.4, 2.12 and 2.13 from Gajic and Qureshi (1995) and allowing $\epsilon \to 0$ so that $Q_\epsilon \to Q$ the bounds follow.

The identification process produces a $\tilde{P}$ from $(\tilde{A},\tilde{B},\tilde{C},\tilde{D})$ *specific* to nodes R for a *arbitrary* set $\tilde{R}$ the following theorem can be applied.

Theorem 10. Consider a connected graph, where r-ID excitation node(s) are attached to a *specific* set of node(s) $R \subseteq V$ forming a modified graph $A(G)$ as defined in (2) and $\tilde{r}$-IF excitation node(s) are attached to an *arbitrary* set of node(s) $\tilde{R} \subseteq V$, then

$$\lambda_{\min}(P) \leq \frac{1}{2}\min\left\{\left(\lambda_n(-A(G))-1\right)^{-1}, \lambda_{n-r}(-A(G))^{-1}\right\},$$

$$(1/2)\left(\lambda_n(-A(G))+1\right)^{-1} \leq \lambda_{\max}(P),$$

$$\max\{(\frac{\tilde{r}}{2})(\lambda_n(-A(G))+1)^{-1}, (\frac{\tilde{r}^2}{2})\left(\mathbf{tr}(-A(G))-r+\tilde{r}\right)^{-1}\} \leq \mathbf{tr}(P_{\tilde{r}}).$$

Proof. Starting with Lemma 9 and applying Proposition 1, the statement of the theorem follows.

Proposition 11. For connected graphs and the infiltration model (2) with 1-IF excitation node, $\mathbf{tr}(P_1) = 1/2$.

Proof. We have $\mathbf{tr}(P_1) = \mathbf{tr}\left(\int_0^\infty e^{A(G)\tau}Qe^{A(G)^T\tau}d\tau\right) = \mathbf{tr}\left(Q\int_0^\infty e^{2A(G)\tau}d\tau\right) = -(1/2)\,\mathbf{tr}\left(QA(G)^{-1}\right)$ where $Q = \mathrm{Diag}(1,0,\ldots 0)$. Now $A(G)^{-1} = -\begin{bmatrix}1 & \mathbf{1}^T\\ \mathbf{1} & M\end{bmatrix}$, where $M = A(1,1)^{-1}\left(I+[A]_{11}\,\mathbf{1}\right)$ and $A(1,1)$ is the principal

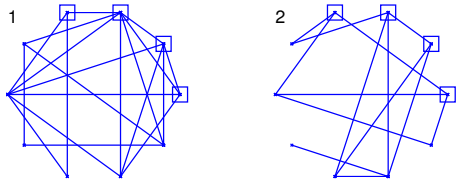


Fig. 3. Two random graphs G^1_{10} and G^2_{10}, each graph is controllable from their square nodes.

	$\underline{J_{2*}}$	J_{2*}	$\overline{J_{2*}}$	$\underline{\bar{J}_{2*}}$	$\bar{J}_{2*}$	$\underline{\mathbf{tr}(P_2)}$	$\mathbf{tr}(P_2)$
K_{10}	0.02	0.02	0.03	0.04	0.16	0.08	0.55
G^1_{10}	0.03	0.03	0.03	0.08	0.22	0.10	0.57
G^2_{10}	0.04	0.04	0.05	0.11	0.30	0.14	0.60
P_{10}	0.05	0.06	0.07	0.17	0.43	0.18	0.67

Table 1. Bounds on the J_{2*}, $\bar{J}_2$ and $\mathbf{tr}(P_2)$ for four graph types. True values were found from an exhaustive search over all 2-IF excitation node configurations.

submatrix of $A(G)$ formed from deleting its first row and column (and so $A(1,1) > 0$ and therefore invertible). Hence $\mathbf{tr}(QA(G)^{-1}) = -1$ and $\mathbf{tr}(P_1) = 1/2$.

The implication of Proposition 11 is that on the average, *single node infiltration has the same effectiveness for an n-node graph regardless of the structure of the network and where the IF excitation node is connected.*

5. SIMULATION

Consider two 10-node random graphs G^1_{10} and G^2_{10} in Figure 1. During the Identification Phase, the ID-excitation and ID-observing nodes attach to the four square nodes ($r = 4$) of each graph - the graph is considered to be controllable and observable from these nodes by using the verification processes described previously. Subsequently, the eigenvalues of $A(G)$ from model (2) are identified and applying Theorems 5 and 10 with two node infiltration ($\tilde{r} = 2$), bounds are found for J_*, $\bar{J}$ and $tr(P)$. These particular bounds were examined as they represent easiest and average cost of infiltration, notable in *risky* infiltration where the remaining conservative bounds are not as prevalent. These bounds are designated $\underline{J_{2*}} \leq J_{2*} \leq \overline{J_{2*}}$, $\underline{\bar{J}_{2*}} \leq \bar{J}_{2*}$ and $\underline{\mathbf{tr}(P)} \leq \mathbf{tr}(P)$ respectively. These bounds accompanied by the actual J_{2*}, $\bar{J}_{2*}$ and minimum $\mathbf{tr}(P)$ values are displayed in Table 1. For comparison bounds for a 9-ID and 2-ID excitation node identification (the number of nodes required for ID) of K_{10} and P_{10} are also displayed. The 2-IF excitation nodes will subsequently be sent to G^1_{10} with smaller J_{2*}, $\bar{J}_{2*}$ and $\mathbf{tr}(P_2)$ bounds compared to G^2_{10}.

6. CONCLUSION

This paper presents a formulation for identification and infiltration problems in consensus-type networks. System identification techniques are utilized to probe networks with limited initial knowledge. The system realization acquired by the identification process is then subsequently used to form analytic bounds on the cost of infiltration with a constant mean signal, and more generally, via the notion of network controllability gramian. The framework provides a setting for reasoning about security of coordination algorithms. It also identifies critical graph-theoretic parameters that can influence the *synthesis of secured network geometries* that support the operation of multi-agent networks.

REFERENCES

Biyikoglu, T. and Leydold, J. (2009). Algebraic connectivity and degree sequences of trees. *Linear Algebra and its Applications*, 430(2-3), 811–817.

Bloem, M., Alpcan, T., and Basar, T. (2007). Optimal and robust epidemic response for multiple networks. In *Proc. 46th IEEE Conference on Decision and Control*, 5074–5079.

Fagiolini, A., Valenti, G., Pallottino, L., Dini, G., and Bicchi, A. (2007). Decentralized intrusion detection for secure cooperative multi-agent systems. In *Proc. 46th IEEE Conference on Decision and Control*, 1553–1558.

Gajic, Z. and Qureshi, M.T.J. (1995). *Lyapunov Matrix Equation in System Stability and Control*. Academic Press, San Diego, CA.

Godsil, C. and Royle, G. (2001). *Algebraic Graph Theory*. Springer-Verlag New York.

Golub, G.H. (1973). Some modified matrix eigenvalue problems. *SIAM Review*, 15(2), 318–334.

Gueye, A. and Walrand, J.C. (2008). Security in networks: A game-theoretic approach. In *Proc. 47th IEEE Conference on Decision and Control CDC 2008*, 829–834.

Hatano, Y. and Mesbahi, M. (2005). Agreement over random networks. *IEEE Transactions on Automatic Control*, 50(11), 1867–1872.

Horn, R.A. and Johnson, C.R. (1990). *Matrix Analysis*. Cambridge University Press, N Y.

Jadbabaie, A., Lin, J., and Morse, A.S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6), 988–1001.

Ljung, L. (1999). *System Identification - Theory for the User*. Prentice-Hall, Upper Saddle River, N J.

Olfati-Saber, R. (2006). Flocking for multi-agent dynamic systems: algorithms and theory. *IEEE Transactions on Automatic Control*, 51(3), 401–420.

Olfati-Saber, R., Fax, J.A., and Murray, R.M. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1), 215–233.

Petrovic, M. and Gutman, I. (2002). The path is the tree with smallest greatest laplacian eigenvalue. *Kragujevac Journal of Mathematics*, 24, 67–70.

Rahmani, A., Ji, M., Mesbahi, M., and Egerstedt, M. (2009). Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM Journal on Control and Optimization*, 48(1), 162–186.

Soderstrom, T. and Stoica, P. (1989). *System Identification*. Prentice Hall, NY.

Tanner, H.G., Pappas, G.J., and Kumar, V. (2004). Leader-to-formation stability. *IEEE Transactions on Robotics and Automation*, 20(3), 443–455.

Wang, Y., Chakrabarti, D., Wang, C., and Faloutsos, C. (2003). Epidemic spreading in real networks: an eigenvalue viewpoint. In *Proc. 22nd International Symposium on Reliable Distributed Systems*, 25–34.

Yueh, W.C. (2005). Eigenvalues of several tridiagonal matrices. *Applied Mathematics E-Notes*, 5, 66–74.