Estimation of Spread and Flow Dynamics: A Network Control Theory Approach

Mengran Xue, **Sandip Roy**, Sean Warnick, and Anurag Rai; with thanks to Yan Wan and Sajal Das

4/19/2012

My Research Group's Focus

- My primary interest is in designing decision-support capabilities for strategic management of large-scale infrastructures
  - Including air traffic management, power-system uncertainty analysis, disease control
- This research requires understanding of the structure and dynamics of networks, especially ones with tightly conjoined cyber- and phyiscal components.
  - *Few observation and actuation points*

- I also am pursuing network modeling in biological applications.

My Research on Networks

- Modeling, of physical dynamics, cyber-/ human- components, and environmental uncertainties.

- **Inference of network models and dynamics from local measurements (Problem 1A?).**

- Controller, topology, and algorithm design (Problem 1b?).

Inference in Networks

- I am interested in both model inference/ parameterization and estimation of dynamics.

- Key advance: we study how the estimator and its performance depend on the network topology. *Why???*

Several Examples

Dynamics:

- Initial condition estimation in network synchronization processes *(Proc. AIAA, 2010)*
- **Security/discoverability of spreads *(IEEE TPDS, 2012)***
- Security of vehicle-team dynamics (*Automatica, subm.)*

Structure:

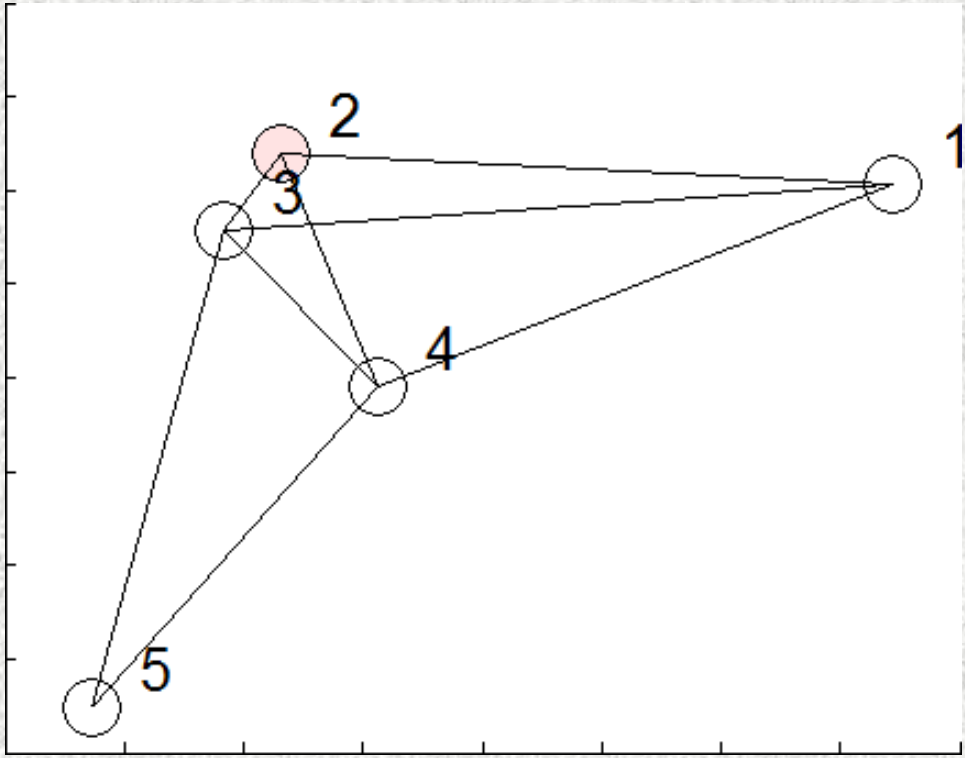
- Mode estimation for synchronization processes (*Proc. IEEE CDC)*
- **Steady-state-probability inference in Markov chains (J. Franklin Inst., 2011)**
- Influence-model parameterization from ensemble weather forecasts (Proc. AIAA 2011).

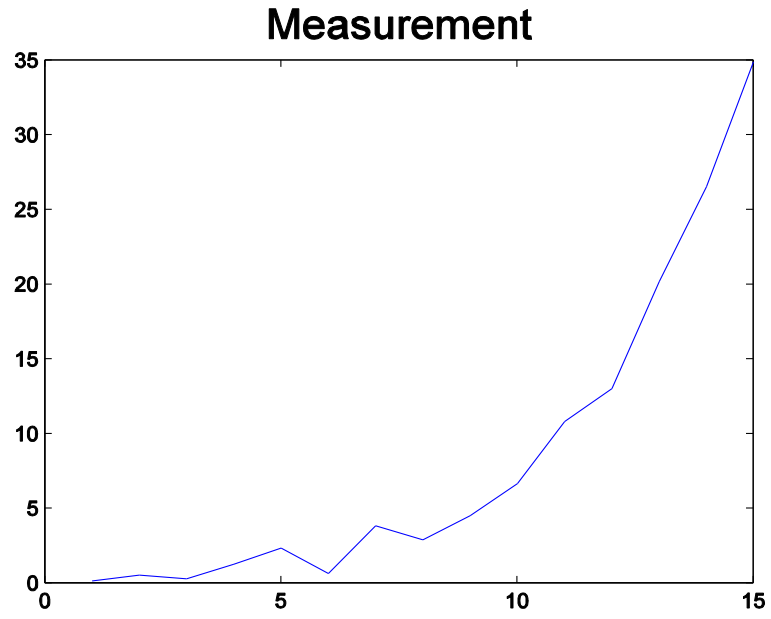
Canonical Problem 1: Spread Models

- Network spread models (for viruses, ideas, rumors, etc) often track penetration in subpopulations or individuals' infection probabilities.
- Here's a simple model for penetration of a new product, or a virus.
- An observer makes noisy measurements of the spread's penetration.

$$x[k+1] = \begin{bmatrix} 0.6 & 0.20 & 0.19 & 0.20 & 0 \\ 0.30 & 0.6 & 0.27 & 0.24 & 0 \\ 0.19 & 0.24 & 0.6 & 0.25 & 0.20 \\ 0.20 & 0.24 & 0.25 & 0.6 & 0.21 \\ 0 & 0 & 0.20 & 0.21 & 0.6 \end{bmatrix} x[k]$$

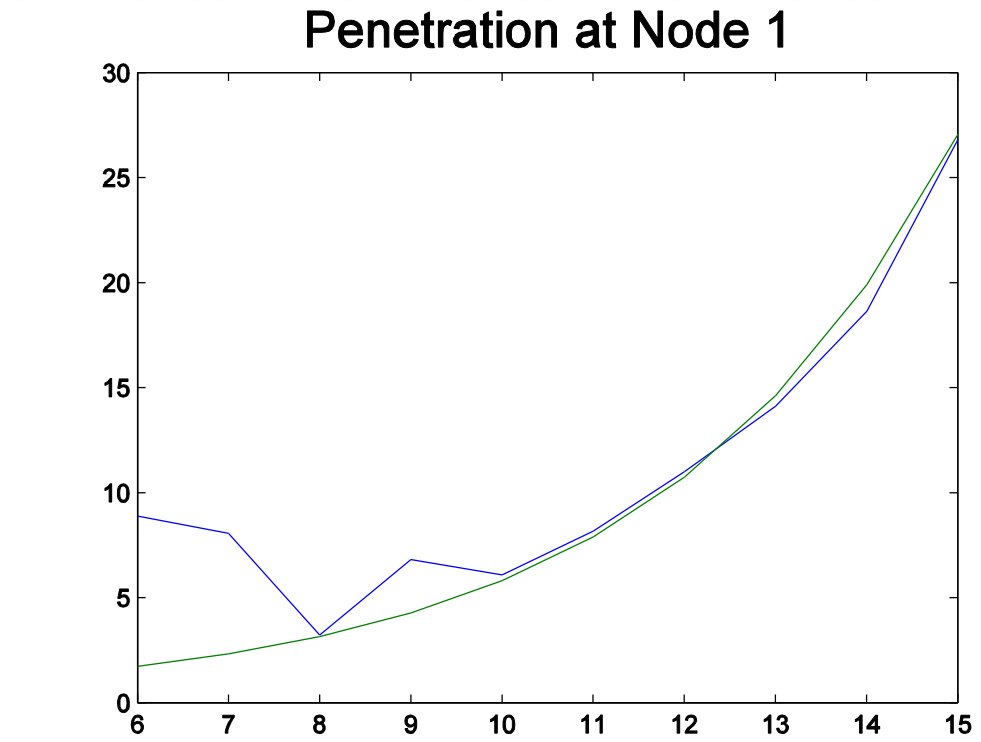
$$y[k] = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \end{bmatrix} x[k] + N[k]$$





Spread Models: Estimation Problem

- **Observer's Goal**: to identify statistics of the dynamics and/or model from the observation sequence.
  - Statistics of interest include: initial condition, **current spread state**, total spread, spread penetration at a particular node, basic reproductive ratio (dominant eigenvalue).
  - From the observer's viewpoint, these are classical filtering problems.



Measuring Security

- **Measuring Security/Discoverability**
  1. Binary problem: is estimation possible at all?
  2. If yes, we can define a measure for security, e.g. the trace of the error covariance
     - For a particular filter, or as a lower bound among estimators
  3. Interpretation depends on whether the observer is good or bad.

- **Our Goal:** to relate security/discoverability, and estimator structure, to graph structure

- **Results:** We have obtained graph-theoretic results for several goals and security measures, but let us focus on the Cramer-Rao lower bound (here, ML estimator error covariance) for current-state estimation.
  - Security Measures: Tr(err_cov) and maximum eigenvalue of err_cov

Security: Classical Algebraic Results

- Algebraic condition: resolves to an observability question.

$$\mathcal{Q} = \begin{bmatrix} E \\ EA \\ \vdots \\ EA^{k_f-1} \end{bmatrix}$$

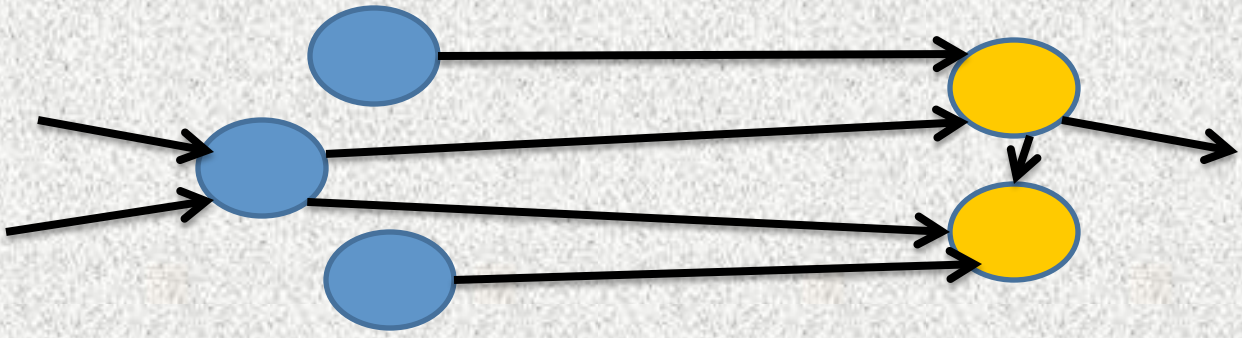
- Spectral condition: iff some right eigenvector of A has zero entries at all the observation locations

- Estimator: $\hat{\mathbf{x}}[0] = (\mathcal{Q}^T \mathcal{Q})^{-1} \mathcal{Q}^T \mathbf{y}.$

- Performance: $S = A^k (\mathcal{Q}^T \mathcal{Q})^{-1} (A^T)^k$

Security: Graph Theoretic Results

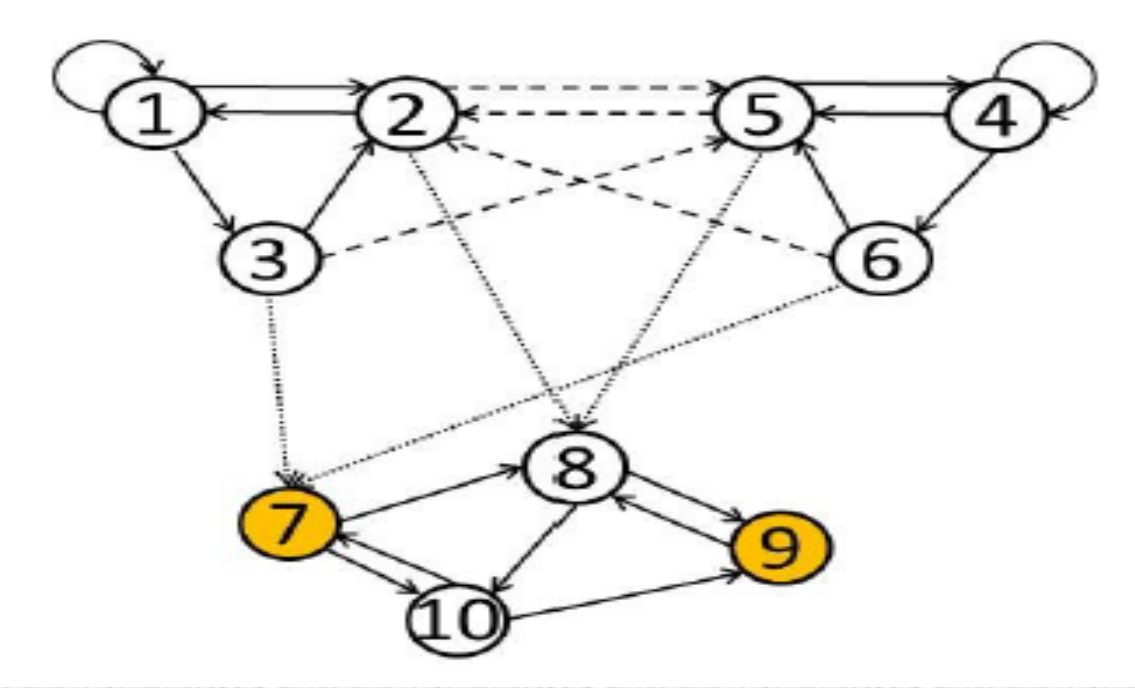
- Approach: Matrix → Spectrum → Graph
- The spread model's security can be classified into three cases:
  1. **Structurally secure:** iff there is a group of *m* unobserved nodes that are 1) mutually fully independent (including self loops) and 2) have edges to *q*<*m* other nodes.



  2. **Specifically secure:** spectral observability conditions allow us to identify symmetries that lead to unobservability---see next slide
  3. **Insecure:** Let us relate the security measure to the graph---on later slides!

Specifically Secure: Graphical Conditions

- Symmetry:



- Insufficient Measurement:

Theorem 11. *Consider a partial-knowledge scenario, and assume that the initial spread state is the estimation goal. If the measurement time horizon k_f is either 1) less than or equal to the maximum distance from any spread-graph vertex to the nearest observation location or 2) less than the ratio between the number of vertices and the number of observation locations ($\frac{n}{m}$), then the observation goal is secure.*

Spectral Results, Insecure Case

- Many of our graphical characterizations of the security level are built from the spectral characterizations.
  - Stable eigenvalues don't impact security measures asymptotically.
  - A bound if there are unstable eigenvalues:

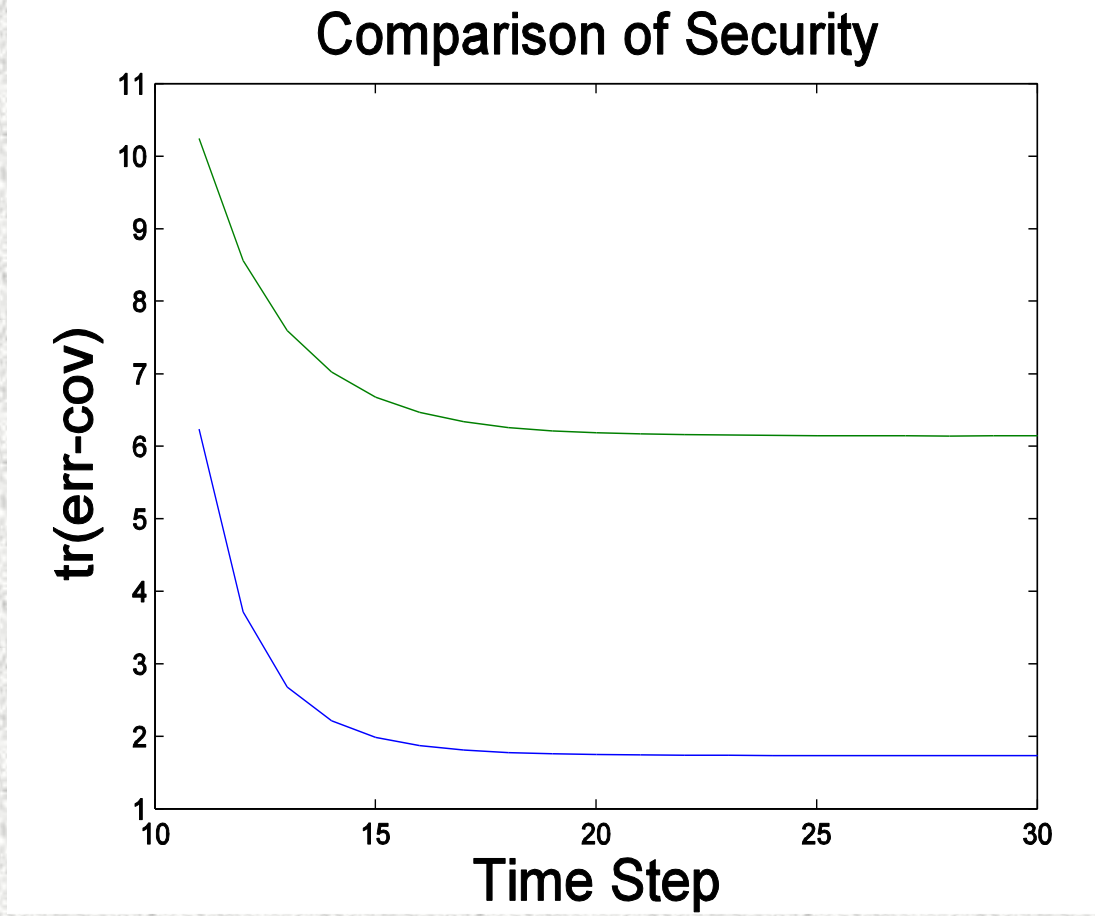
$$T(S) \geq \gamma_{max}(S) \geq \max_{i \in F} \mu_i^{k-k_f+1} (1 - \frac{1}{\mu_i^2})(\frac{1}{\sum_{j \in \mathscr{U}} |v_{i,j}|^2}).$$

  - The above is asymptotically tight if there's one unstable eigenvalue
  - An uglier bound that accounts for complex eigenvalues, and interactions between eigenvalues, can be developed. This bound makes clear that nearby eigenvalues yield security.

  - BTW, security of the initial conditions depend on the stable eigenvalues

Example…

- Eigenvalues=
  1.36, .68, .38, .33,.25
- Dominant Eigenvector=
  [.39,.50,.51,.51,.27]
- Security:
  Observer at 3: S=1.73
  Observer at 5: S=6.14

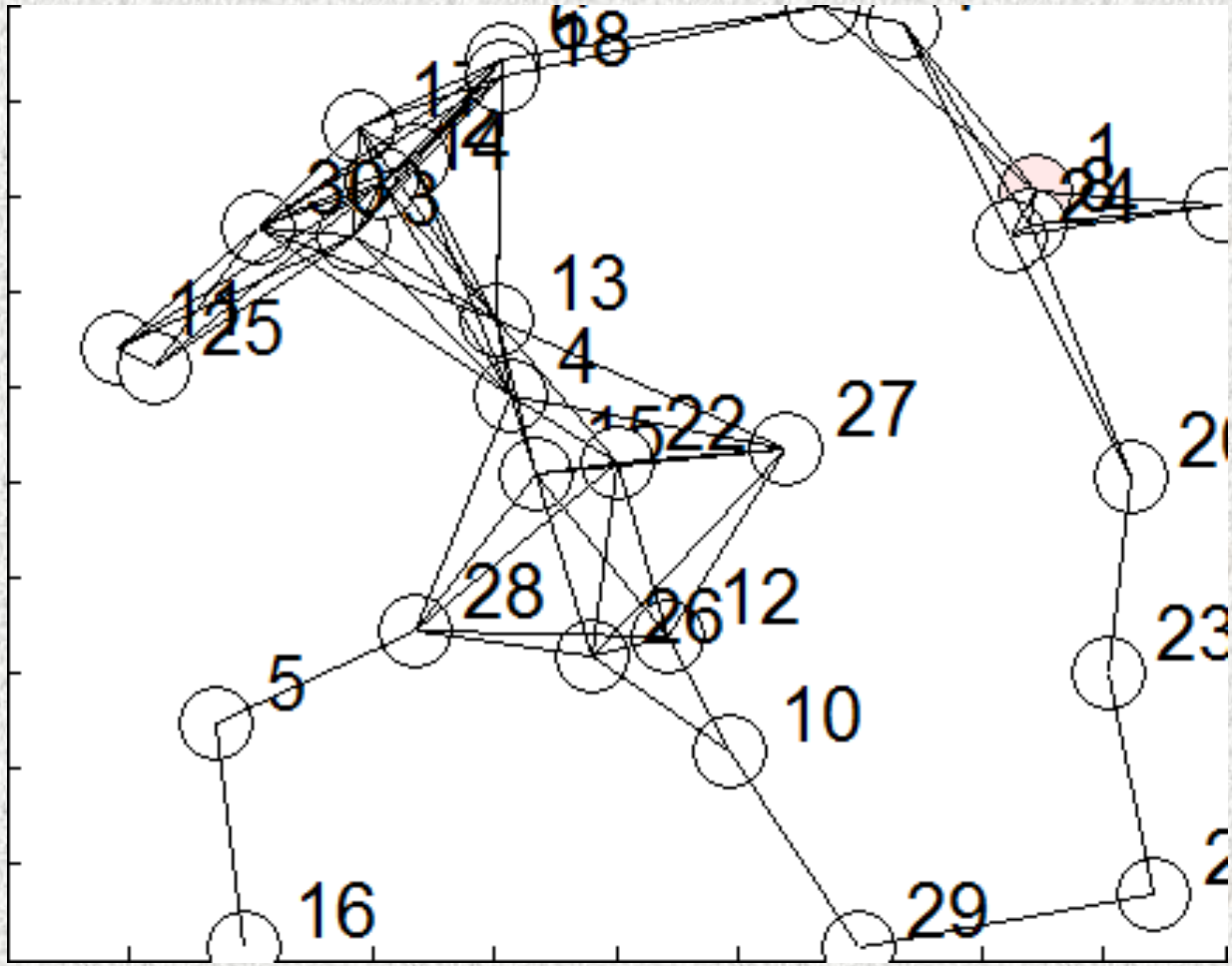


From Spectral to Graphical Results…

- The spectral analysis highlights the key role of the dominant eigenvalue and corresponding eigenvector on security.
- Many results on the spectra of adjacency matrices and other nonnegative matrices can be brought to bear.
- **For specific graphs:** the dominant eigenvalue and corresponding eigenvector can be found easily, allowing sensor placement.
- **For classes of graphs:**
  - For large random graphs, a wide range of transmission rates yield 1) a single unstable eigenvalue (with magnitude decided by wiring probability and transmission rate); 2) a dominant eigenvector whose components are almost equal.
  - Meanwhile, for slow-coherency structures, eigenvector localization oten makes estimation hard.
  - Performance can be bounded in terms of graph statistics

More on the random graph case…

- For a broad class of random graphs:
- T(S)= $\left(1 - \frac{1}{(np)^2}\right)\frac{n}{m}$



Even the simple spectral bound shows that this slow-coherent graph is 50 times more secure than a comparable random graph!

More Graphical Results

- **Graphical Insights on Sensor Placement**: The least secure observation locations typically have above-average degree.

- **Insights Regarding Graph Changes:** Assuming one unstable eigenvalue, adding edges out from a node always 1) makes that node relatively less secure and 2) all other nodes more secure in absolute terms.
  - When an edge of 0.3 is added from node 5 to node 1:

    Observer at Node 3: Security increases from 1.73 to 1.92

    Observer at Node 5: Security decreases from 6.41 to 3.07

- **Security for Optimal Designs:** optimal spread controllers tend to equalize dominant-eigenvector components; security becomes less dependent on sensor location in this case.

Consensus Algorithms: Summary

- **Synchronization** or **consensus algorithms** on graphs can capture decision-making or voting processes in social networks.
  - Extensive literature in computer science, physics, biology, controls engineering…
- We impose an observation model on the classical LTI synchronization model, and so define notions of security/ discoverability.
  - Our focus has been on initial-condition estimation and mode estimation.
- We have obtained results that are analogous to those for the spread-estimation problem [WR10,XY11].

Markov Chain Steady-State Probability Estimation

- Several *inference* problems for MCs have been intensively studied.
- A rich literature has been specifically dedicated to the steady-state estimation problem.
- Steady-state probability estimation is also a step toward *transition-matrix inference*, which is important in network-mapping tasks.

Background & Introduction

- Characterizing the *performance* of common steady-state-probability estimators is of particular interest.

- Steady-state-probability estimator performance has been related to the MC transition matrix.

Problem Formulation: Motivation

- We revisit the MC steady-state-probability estimation problem from a spectral and graph-theoretic perspective:
  1) This approach yields good performance bounds;
  2) Useful for design and algorithm-evaluation tasks.

Problem Formulation: Overview

- Consider a discrete-time ergodic MC with m states:
  1) state transition matrix D: transition probability from state i to state j is d_{ij}
  2) observation sequence: $\mathbf{s}[k]$, $k=0,1,\ldots,N$ (We use an indicator-vector notation for the state observation.)
  3) probability vector $\mathbf{p}[k]=E[\mathbf{s}[k]]$
  4) Steady state probability $\boldsymbol{\pi}=[\pi_1,\ldots,\pi_m]^T=\lim_{k\to\infty}\mathbf{p}[k]$
  5) Weighted, directed graph defined from D.

- Markov chain steady state estimation (MCSS) problem: estimate the steady-state probability vector from observation sequence.

Problem Formulation: Estimator

- The sample-frequency-based estimator:

$$\mathbf{p} = \frac{1}{N} \sum_{k=1}^{N} \mathbf{s}[k]$$

- Let's characterize the performance of the sample-frequency-based estimator, assuming that the Markov chain begins in steady-state.

Problem Formulation: Estimator (2)

- Two classical performance measures for the estimate $\mathbf{p}$: *error covariance matrix* cov (defined as $E[(\mathbf{p}-\boldsymbol{\pi})(\mathbf{p}-\boldsymbol{\pi})^T]$) and *expected total squared error* (defined as $E[(\mathbf{p}-\boldsymbol{\pi})^T(\mathbf{p}-\boldsymbol{\pi})]$).
  - Total squared error equals tr(COV)

Results: Preliminary Analysis

$$COV = \frac{1}{N^2}\left[\sum_{r=1}^{N-1}(N-r)\Delta D^r + \sum_{r=1}^{N-1}(N-r)(D^T)^r\Delta + N\Delta\right] - \pi\pi^T$$

$$tr(COV) = \frac{1}{N^2}\left(2\sum_{r=1}^{N-1}(N-r)\sum_{i=1}^{m}e_i D^r e_i \pi_i + N\right) - \pi^T\pi$$

Results: Spectral Analysis (1)

- Performance measures, simple-eigenvalue case:

$$COV = \sum_{i=2}^{m} \gamma_i (\Delta \mathbf{v}_i \mathbf{w}_i^T + \mathbf{w}_i \mathbf{v}_i^T \Delta)$$

$$tr(COV) = 2 \sum_{i=2}^{m} \gamma_i \left(\sum_{j=1}^{m} v_{ij} w_{ij} \pi_j \right)$$

where $\gamma_i = \frac{\lambda_i^{N+1} - N\lambda_i^2 + (N-1)\lambda_i}{N^2(1-\lambda_i)^2} + \frac{1}{2N}, i = 2, \cdots, m.$

- Jordan-form case is similar…

Results: Performance Bounds (1)

- General lower and upper bounds for tr(COV), in terms of eigenvalues of the state transition matrix:

$$2\pi_{min} \sum_{i=2}^{m} \gamma_i \leq tr(COV) \leq 2\pi_{max} \sum_{i=2}^{m} \gamma_i,$$

Results: Performance Bounds (2)

- Upper bounds on $\mathrm{tr}(\mathrm{COV})$ that are phrased in terms of the eigenvalue λ_2 (the non-unity eigenvalue that is closest to $1+j0$ in the complex plane):

$$tr(COV) < \frac{2(m-1)}{N|1-\lambda_2|} + \frac{4(m-1)}{N^2|1-\lambda_2|^2}.$$

Results: Performance Bounds (3)

- Special case, eigenvalues real:

$$tr(COV) < \frac{2(m-1)}{N(1-\lambda_2)} - \frac{m-1}{N} + \frac{4(m-1)}{N^2(1-\lambda_2)^2}.$$

- Other special cases: symmetric, two-state

Results: Graphical Characterization

- Relating tr(COV) to the MC's graph in limiting cases, i.e., for strong or weak connectivities.

Lemma 1 *For a stochastic matrix $D = \{d_{ij}\}$, the non-unity eigenvalues can be bounded as*

$$|\lambda| \leq \frac{1}{2} \sum_{k=1}^{m} \max_{i,j} |d_{ik} - d_{jk}|. \tag{23}$$

[Zenger 1972]

Theorem 6 *Consider an ergodic Markov chain described as in the MCSS estimation problem. We assume that, in the underlying directed graph G defined by the state transitions, the product of edge weights along a shortest path between any pair of vertices is lower bounded by some positive constant q. Then, the non-unity eigenvalues of D can not be close to 1 if q is large.*

Results: Other Graphical Analyses

- The bounds on λ_2 and $\mathrm{tr}(\mathrm{COV})$ can be related to a variety of other graph features using results from algebraic theory.

- Some other graphical analyses:
  1) Using Cheeger-type bounds, λ_2 and $\mathrm{tr}(\mathrm{COV})$ can be bounded in terms of the diameter and various degree measures.
  2) Using the geometric evaluations of λ_2, we can show that increasing edge weights in a reversible MC always moves λ_2 further from 1.
  3) We conjecture that increasing edge weights between vertices in a reversible MC leads to a monotonic decrease in $\mathrm{tr}(\mathrm{COV})$.

Where in the Network?

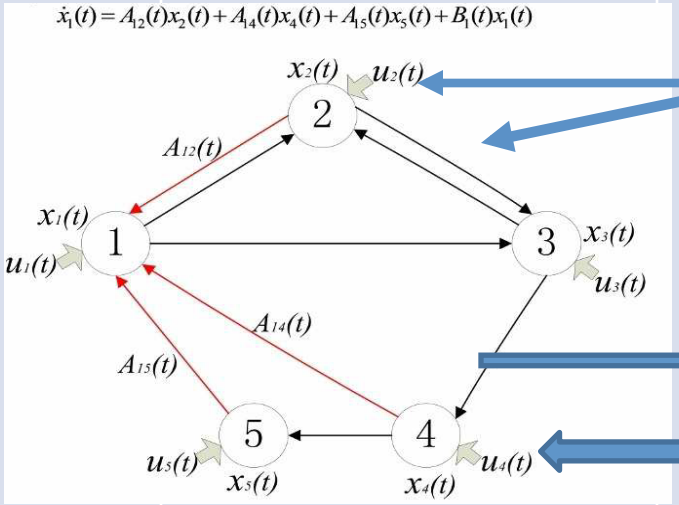
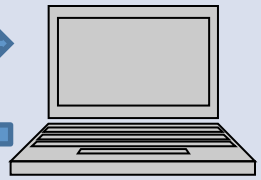
- In absolute terms, higher-probability states also have higher estimation-standard-deviation.
- In relative terms, however, unlikely states are harder to estimate.

- *Participation factors* of states in eigenvalues near 1+j0 are also important.

Estimating Transition Frequencies/ Probabilities

- *Transition frequencies* can be estimated by determining the fraction of time steps that a particular transition is taken.
  - Noting that the transitions themselves are governed by a Markov chain, the estimation error covariance can be analyzed as before.
  - This Markov chain has exactly the same eigenvalues as the original chain, except some extra zero eigenvalues.
- *Transition probabilities* estimates are harder to characterize, except in the symmetric case.
- How about hidden Markov models?

A Broader Security and Robustness Framework

The Framework	Models for CPNs	Adversaries	Security/Robustness Definitions
Prelim. Work	   Natural Sentient	*Robustness:* Controllability, dynamics within safe set, low variability *Security:* Observability, estimator performance,
Further Research	Use moment-linear models to capture cyber- and hybrid dynamics, time varying interactions	Represent multi-layer dynamics, trust, and intelligence of adversaries.	Define heterogeneous costs, pursue risk analysis by capturing temporal progression of threats, closing the loop

Many Thanks!

- Questions?

- We are appreciative of the support from the NSF, FAA, and Mitre Corporation, for this work.