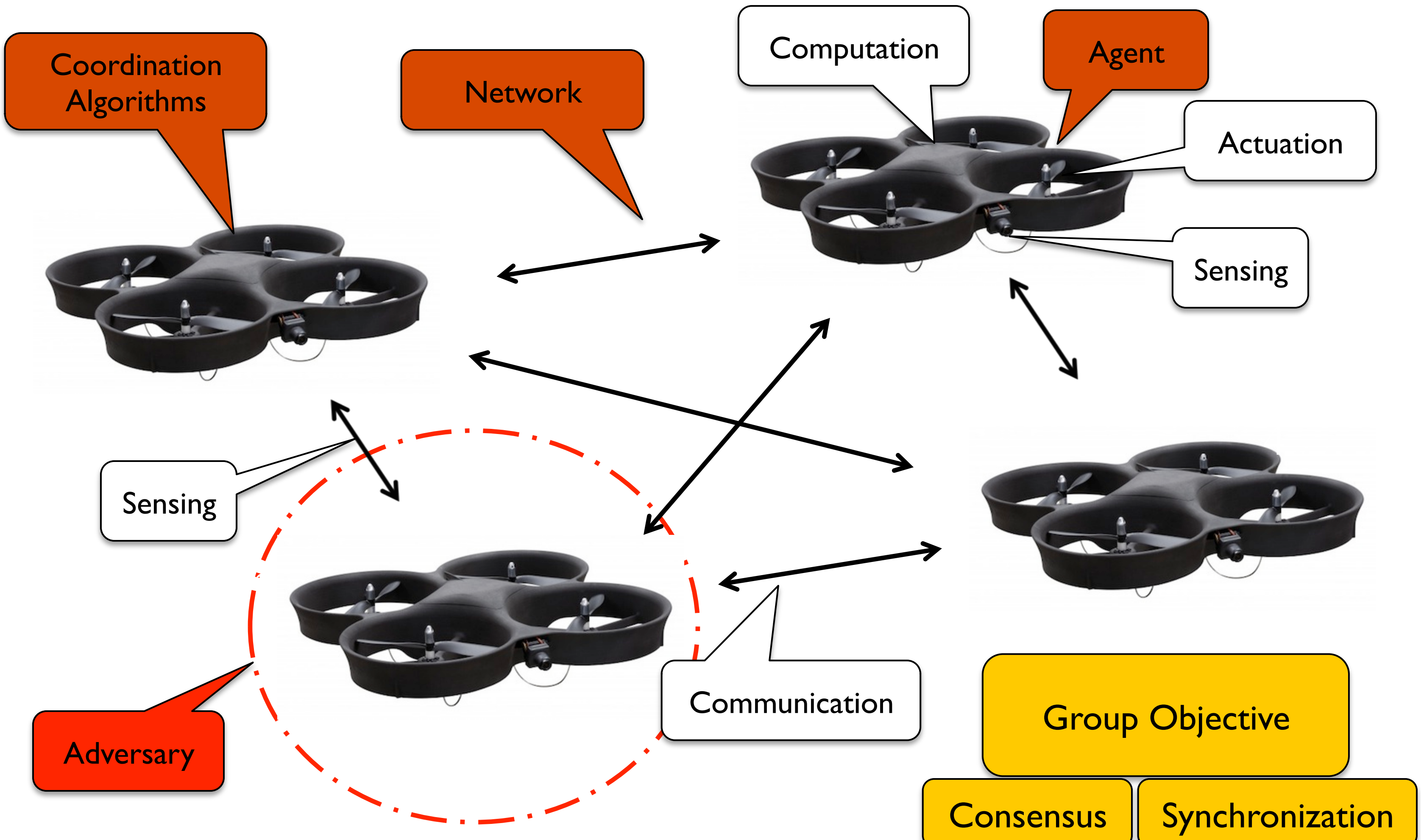
Resilient Cooperative Control of Cyber-Physical Systems

Xenofon Koutsoukos
(work with Heath LeBlanc)

Resilient CPS

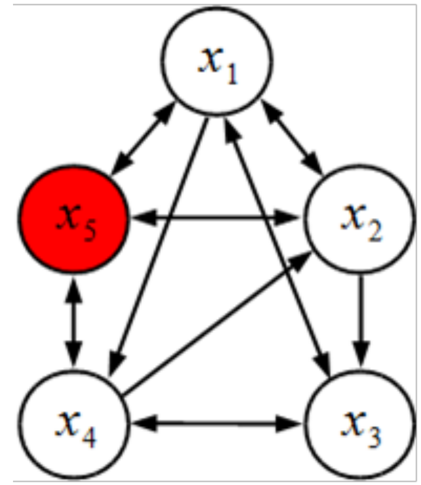


Outline

- Adversary models
- Resilient consensus
  - Complete networks
  - High-degree networks
  - Robust networks
- Resilient synchronization
- Conclusions and future work

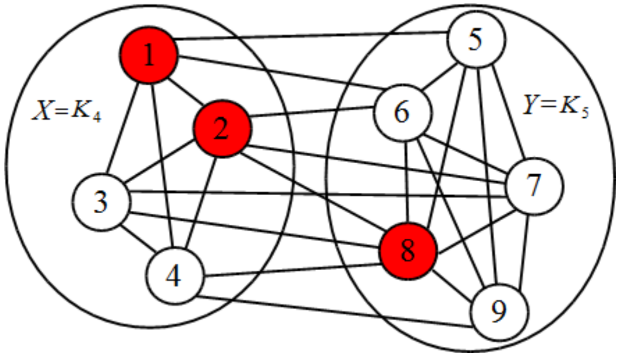
Adversary Models

- Crash Adversary
  - Choose a time to "crash" the node
    - States of the node remain unchanged after the "crash" event
- Malicious Adversary
  - Can change the state values arbitrarily
    - Continuous trajectory in continuous time
    - No limits in discrete time
  - Must convey the same information to all neighbors
    - Local broadcast model
- Byzantine Adversary
  - Can convey different information to different neighbors
- All adversaries are omniscient; i.e., know
  - Topology of the network
  - States and algorithms of the other nodes
  - Other adversaries (can collude)

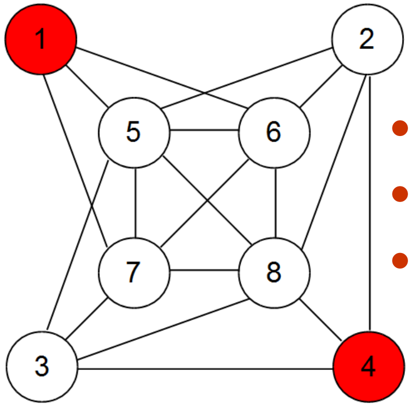


Scope of Threat Models

- *F*-Total Model
  - Assumes at most *F* adversaries in the entire network
- *F*-Local Model
  - Assumes at most *F* adversaries in the neighborhood of any normal node
- *f*-Fraction Local Model
  - Assumes at most a fraction *f* of adversaries in the neighborhood of any normal node



- 3-Total
- 3-Local
- (3/5)-Fraction Local



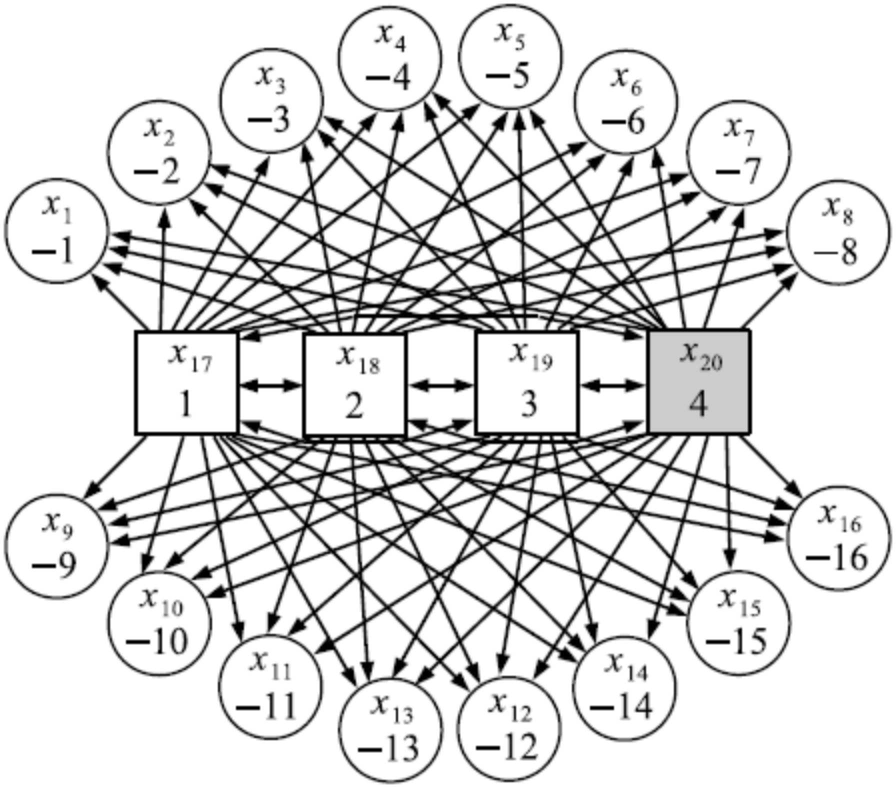
- 2-Total
- 1-Local
- (1/4)-Fraction Local

Resilient Consensus

- Consensus protocols are fundamental for multi-agent coordination
  - Time synchronization, rendezvous, formation control, distributed estimation
- In distributed computing, consensus protocols robust to faulty (Byzantine) processors have studied extensively
- Approximate Agreement with Byzantine processors
  - *Agreement:* Decision values of any two processes within ε each other
  - *Validity:* Any decision value for a nonfaulty process is within the range of initial values of the nonfaulty processes
  - *Termination:* All nonfaulty processes eventually decide
- ConvergeApproxAgreement algorithm [D. Dolev et al.]
  - Uses sorting, reduction, and selection functions on multisets

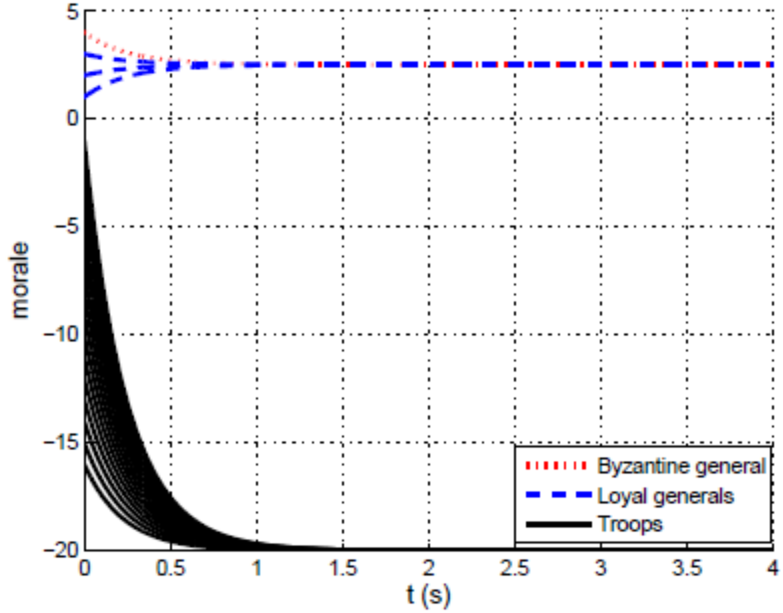
Variation of Byzantine Generals Problem

- Morale modeled by single real value x_i for troop i
  - $x_i > 0$, good morale
  - $x_i < 0$, bad morale
- Loyal generals attempt to improve troop morale and reach consensus on the level of morale despite Byzantine generals



Simulation Results

- Linear Consensus

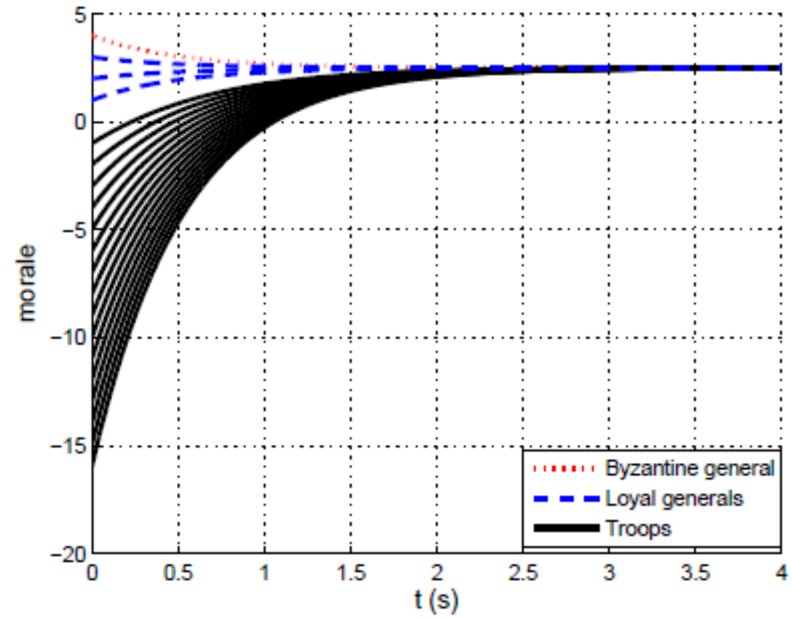


(a) LCP.

$$\dot{x}_i(t) = \sum_{j \in \{\text{Generals}\}} (x_j(t) - x_i(t)),$$

$$\forall i \in \{\text{Troops}\} \cup \{\text{Loyal Generals}\}$$

- Resilient Consensus

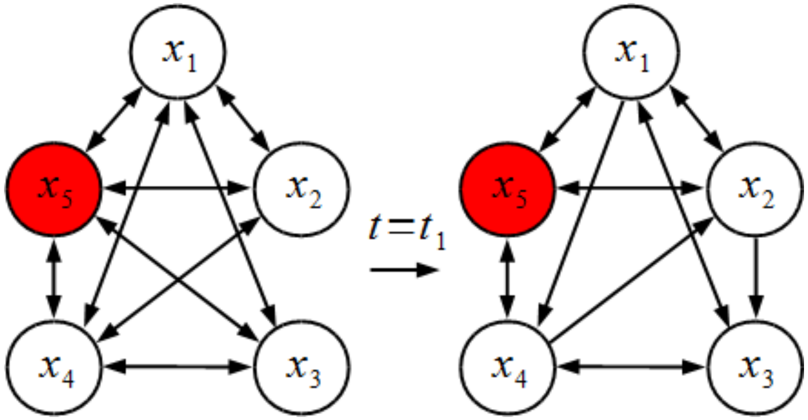


(b) ARC-P.

$$\dot{x}_i(t) = \sum_{j \in \{\text{Generals}\} \setminus \{\text{Extreme Morale}\}} (x_j(t) - x_i(t)),$$

$$\forall i \in \{\text{Troops}\} \cup \{\text{Loyal Generals}\}$$

Networked Multi-Agent System



- Switched System
  - Ordinary Differential Equations (ODEs)
  - Switching network topology
- Normal nodes have scalar state & integrator dynamics

$$\dot{x}_i = u_i = f_{i,\sigma(t)}(t, x_{\mathcal{N}}, x_{(\mathcal{A},i)})$$

- Switched system model

$$\dot{x}_{\mathcal{N}} = f_{\sigma(t)}(t, x_{\mathcal{N}}, x_{(\mathcal{A},\mathcal{N})}),\ x_{\mathcal{N}}(0) \in \mathbb{R}^N, \mathcal{D}_{\sigma(t)} \in \Gamma_n$$

Continuous-Time Resilient Asymptotic Consensus (CTRAC)

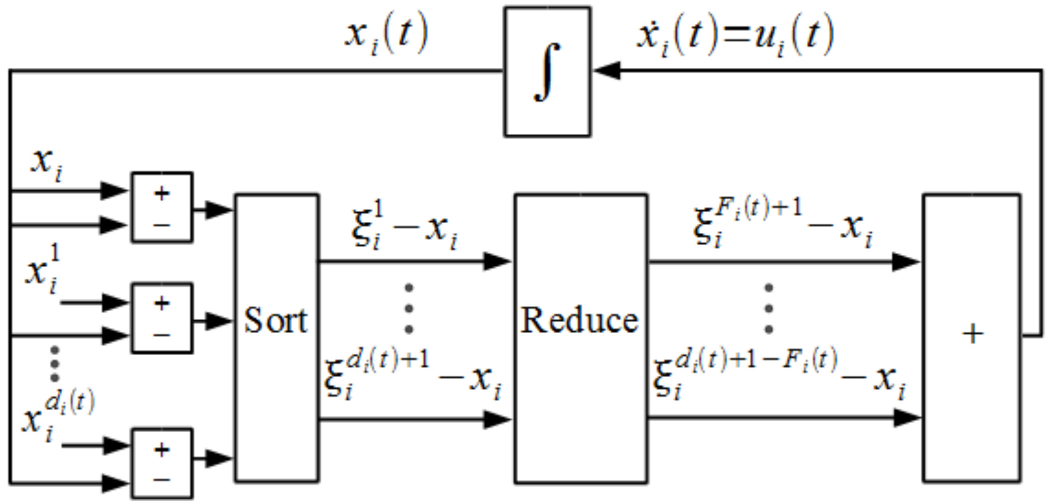
- Design a continuous-time consensus algorithm (control law) that is resilient to adversaries:
  - Agreement Condition: States of the normal nodes asymptotically align to a common limit

$$\exists L \in \mathbb{R} \text{ such that } \lim_{t\to\infty} x_i(t) = L, \quad \forall i \in \mathcal{N}$$

  - Safety Condition: The minimal interval containing the initial values of the normal nodes is an invariant set

$$x_i(t) \in \mathcal{I}_0 = [m_{\mathcal{N}}(0), M_{\mathcal{N}}(0)], \quad \forall t \geq 0, \forall i \in \mathcal{N}$$

Adversarial Resilient Consensus Protocol (ARC-P)



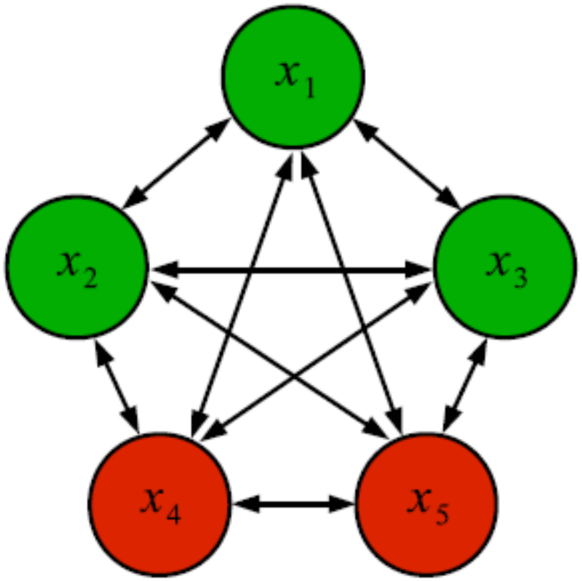
- ARC-P with parameter F (or f)
  - If $d_i(t) \geq 2F_i(t)$
    - $F_i(t) = F$ if the parameter is F
    - $F_i(t) = \lfloor f d_i(t) \rfloor$ if the parameter is f
  - Otherwise, do nothing
- Only local information
- Low complexity

ARC-P2

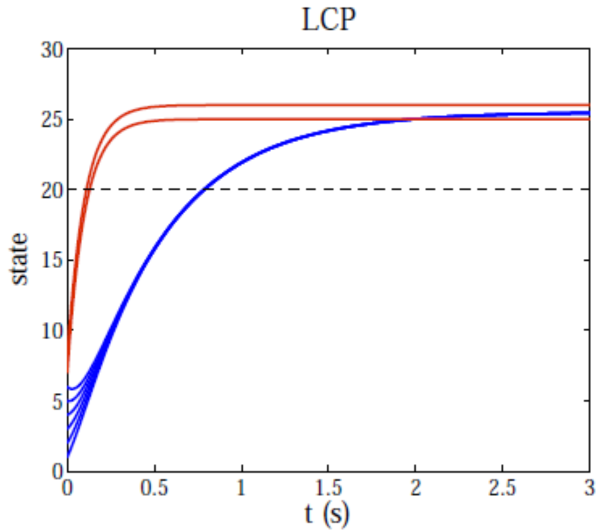
- Weighted ARC-P with selective reduce (ARC-P2)
  - Parameter F (or f)
    - $F_i(t) = F$ if the parameter is F
    - $F_i(t) = \lfloor f d_i(t) \rfloor$ if the parameter is f
  - Nonnegative, piecewise continuous, bounded weights
    - $0 < \alpha \leq w_{(j,i)}(t) \leq \beta$ if j is a neighbor at time t
    - $w_{(j,i)}(t) = 0$ otherwise
  - Compare values of neighbors with own value $x_i(t)$
    - Remove (up to) $F_i(t)$ values strictly larger than $x_i(t)$
    - Remove (up to) $F_i(t)$ values strictly smaller than $x_i(t)$
  - Let $\mathcal{R}_i(t)$ denote the set of nodes whose values are removed
  - Update as $\dot{x}_i(t) = \sum_{j \in \mathcal{N}_i^{\text{in}}(t) \backslash \mathcal{R}_i(t)} w_{(j,i)}(t) \left(x_{(j,i)}(t) - x_i(t) \right)$

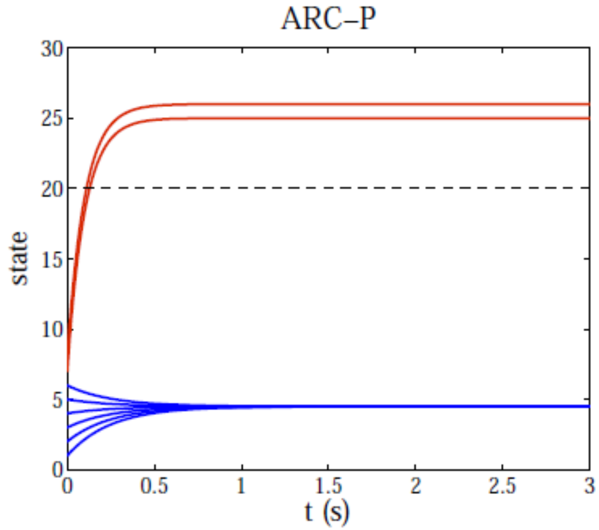
Complete Networks

- ARC-P satisfies the agreement condition
- The convergence to the agreement space is exponential with rate $m = n - 2F$
  - Symmetry of the complete network
- ARC-P satisfies the safety (validity) condition
  - The minimal hypercube containing the initial values is positively invariant

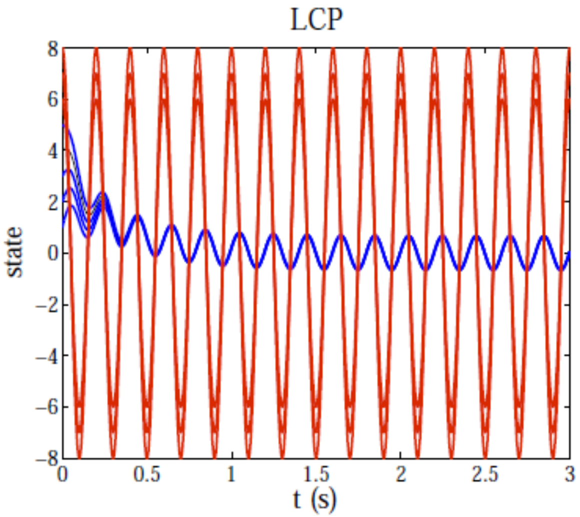


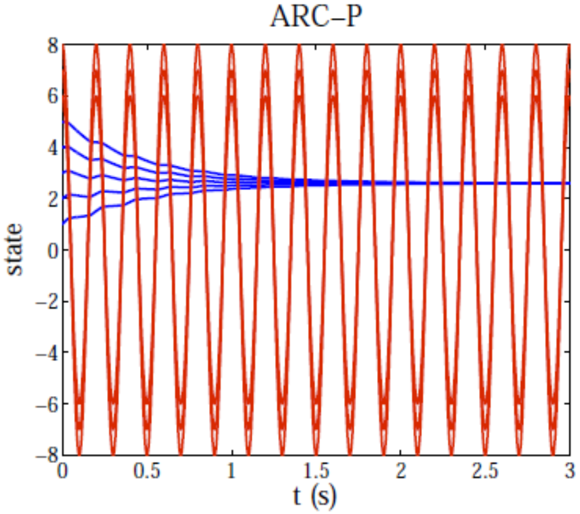
Simulation Results





Unsafe Region: 8-agent network, 2 adversaries





Oscillations: 8-agent network, 3 adversaries

High-Degree Networks

- $D_s \in \Gamma_{M,F} \subset \Gamma_n$ if adversaries are *malicious*

$$\Gamma_{M,F} = \{D_k \in \Gamma_n | M1_F \text{ OR } M2_F \text{ holds}\}$$

where

$$M1_F : \delta^{\text{in}}(D_k) \geq \lfloor n/2 \rfloor + F$$

$$M2_F : \exists S \subseteq V, |S| \geq 2F + 1, \quad \text{such that } d_i^{out} = n - 1, \forall i \in S$$

- $D_s \in \Gamma_{B,F} \subset \Gamma_n$ if adversaries are *Byzantine*

$$\Gamma_{B,F} = \{D_k \in \Gamma_n | B1_F \text{ OR } B2_F \text{ holds}\}$$

where

$$B1_F : \delta^{\text{in}}(D_k) \geq \begin{cases} n/2 + \lfloor 3F/2 \rfloor & n \text{ is even, } F \text{ odd;} \\ \lfloor n/2 \rfloor + \lceil 3F/2 \rceil & \text{otherwise.} \end{cases}$$

$$B2_F : \exists S \subseteq V, |S| \geq 3F + 1, \quad \text{such that } d_i^{out} = n - 1, \forall i \in S$$

Safety and Agreement

- Suppose each cooperative agent uses ARC-P with parameter F and there are at most
  - F malicious agents with $D_{\sigma(t)} \in \Gamma_{M,F}$
  - F Byzantine agents with $D_{\sigma(t)} \in \Gamma_{B,F}$
- Then the safety condition is satisfied
- Then x_c globally exponentially converges to the agreement space.
- The rate of convergence is bounded by

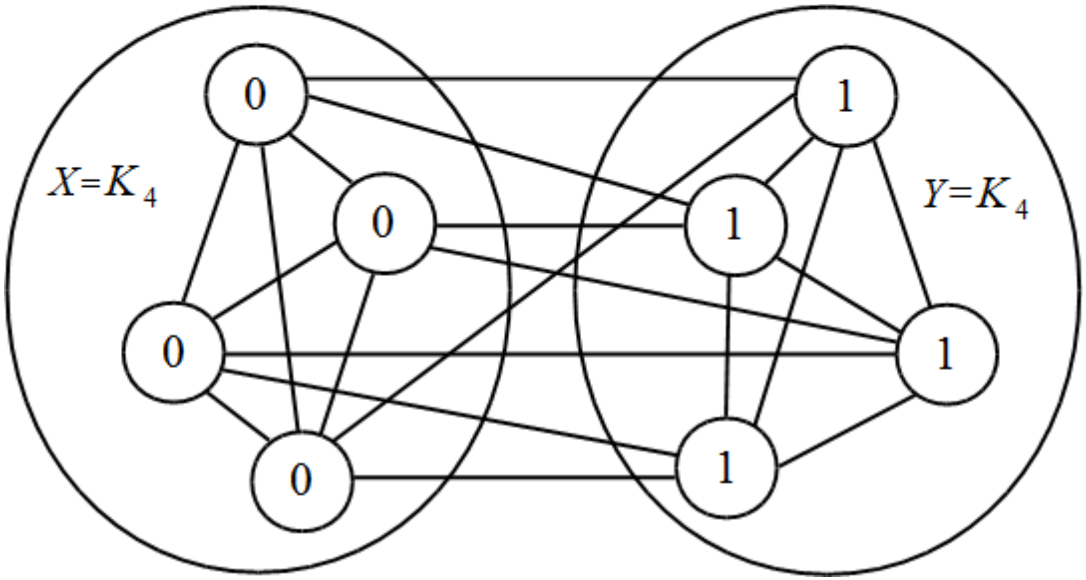
$$\mathrm{dist}(x_c(t), A) \leq 2\sqrt{p}\,\mathrm{dist}(x_c(0), A)e^{-t}$$

Lyapunov Analysis

- Properties of $\Psi(x_c) = \max_{k \in V_c}\{x_k\} - \min_{j \in V_c}\{x_j\}$
  - $\Psi \geq 0$ with $(x_c) = 0$ for $x_c \in A$; $(x_c) > 0$ otherwise
  - Globally Lipschitz;
  - Strictly increasing away from A:
    - $\Psi(y_1) > \Psi(y_2)$ whenever $\mathrm{dist}(y_1, A) > \mathrm{dist}(y_2, A)$
  - Radially unbounded away from A:
    - $\Psi(y) \to \infty$ as $\mathrm{dist}(y, A) \to \infty$
  - Not everywhere differentiable
- *Upper-directional derivative*

$$D^+\Psi(x_c, x_a) = \limsup_{h \to 0^+} \frac{\Psi(x_c + h f_{c,\sigma(t)}(x_c, x_a)) - \Psi(x_c)}{h}$$

Robust Network Topologies



- Nodes in X have value 0 and nodes in Y have value 1
- ARC-P2 with parameter F=2
- No consensus, even with no adversaries
- ($\lfloor n/2 \rfloor + F - 1$)-connected, (in this case, 5-connected)
- We need a new graph theoretic property to capture **local redundancy**

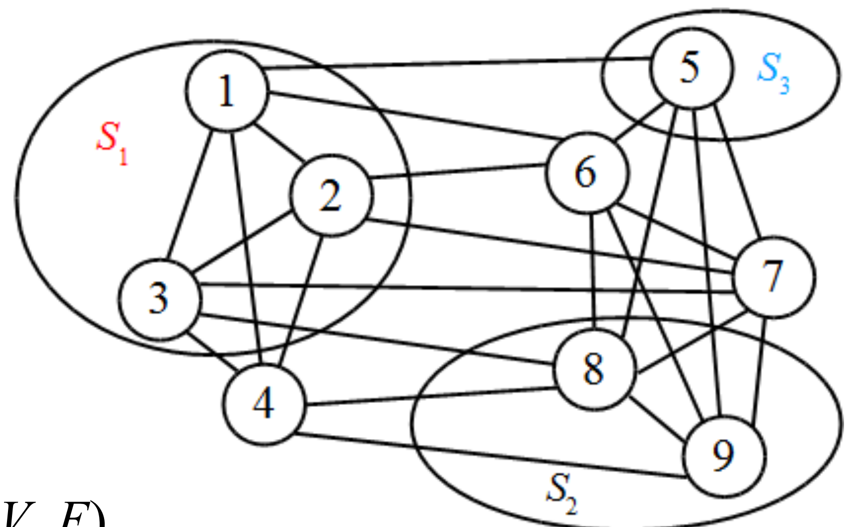
[Collaboration with Haotian Zhang and Shreyas Sundaram]

r-Edge Reachable & r-Robust

- A nonempty subset S of nodes of a nonempty digraph is ***r*-edge reachable** if there exists $i \in S$ such that

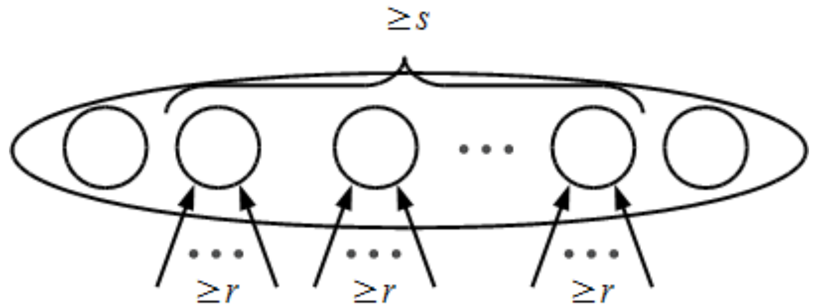
$$|\mathcal{N}_i^{\text{in}} \setminus S| \geq r$$

  - S_1 is 3-edge reachable
  - S_2 is 5-edge reachable
  - S_3 is 5-edge reachable



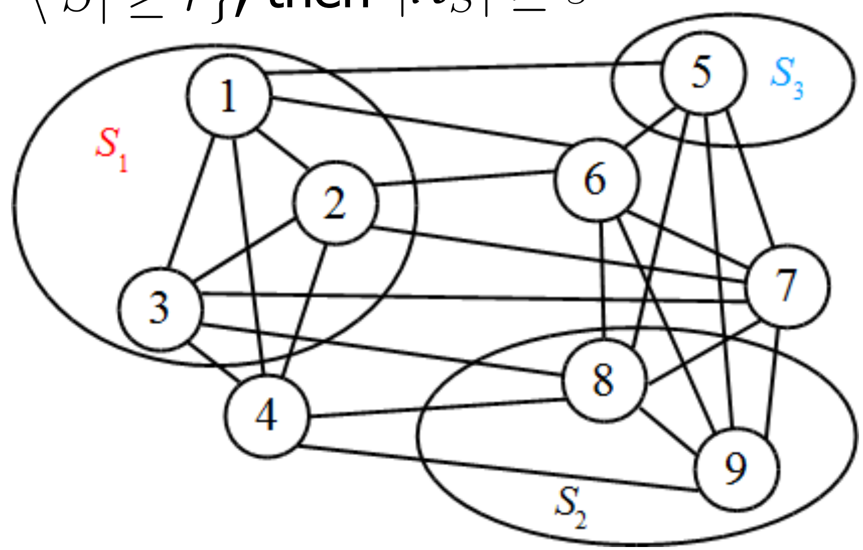
- A nonempty, nontrivial digraph $D=(V, E)$ is ***r*-robust** if for every pair of nonempty, disjoint subsets of V, at least one of the subsets is r-edge reachable

(r,s)-Edge Reachable



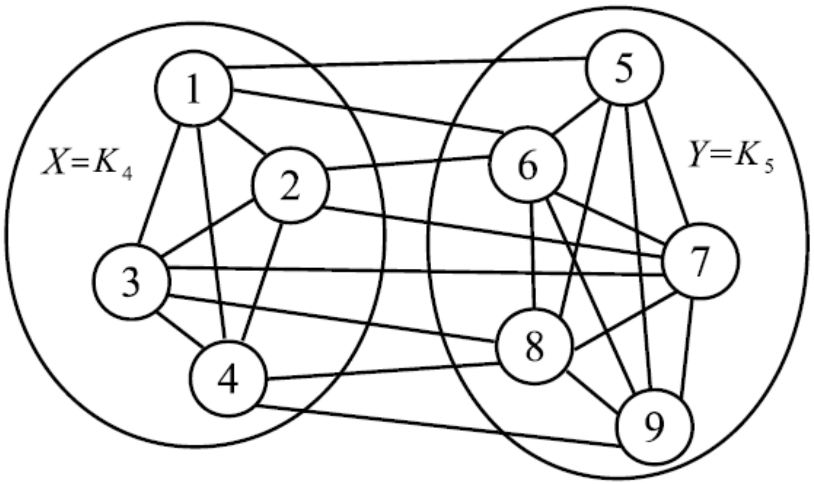
- A nonempty subset S of nodes of a nonempty digraph is **(r,s)-edge reachable** if there are *at least* s nodes in S with *at least* r neighbors outside of S, where $r,s \geq 0$
  - Given $\mathcal{X}_S = \{i \in S : |\mathcal{N}_i^{\text{in}} \setminus S| \geq r\}$, then $|\mathcal{X}_S| \geq s$

- S_1 is (3,3)-edge reachable
- S_2 is (4,2)-edge reachable
- S_2 is (5,1)-edge reachable
- S_3 is (5,1)-edge reachable



(r,s)-Robustness

- A nonempty, nontrivial digraph is $D=(V, E)$ on n nodes is (r,s)-robust with $r \geq 0$, $n \geq s \geq 1$, if for every pair of nonempty, disjoint subsets S_1 and S_2 of V, such that S_k is $(r,s_{r,k})$-edge reachable with $s_{r,k}$ maximal for $k \in \{1,2\}$, then at least one of the following holds
  - $s_{r,1} + s_{r,2} \geq s$
  - $s_{r,1} = |S_1|$
  - $s_{r,2} = |S_2|$



$(2,s)$-robust for $n=9 \geq s \geq 1$

CTRAC Time-Invariant Network: ARC-P2 with parameter F (or f)

Threat	Scope	Necessary	Sufficient
Crash & Malicious	F-Total	(F+1,F+1)-robust	(F+1,F+1)-robust[1]
Crash & Malicious	F-Local	(F+1,F+1)-robust	(2F+1)-robust
Crash & Malicious	f-Fraction local	f-fraction robust	p-fraction robust, where $2f < p \leq 1$
Byzantine	F-Total & F-Local	Normal Network is (F+1)-robust	Normal Network is (F+1)-robust
Byzantine	f-Fraction local	Normal Network is f-robust	Normal Network is p-robust where $p > f$

- Normal network is the network induced by the normal nodes

[1] Requires additional assumption of uniformly continuous malicious agent trajectories

CTRAC Time-Varying Network: ARC-P2 and parameter F (or f)

- Assume there exists a minimum dwell time τ
- Assume there exists time t_0 after which the network topologies always belong to the class of robust networks given below

Threat	Scope	Sufficient
Crash & Malicious	F-Total	(F+1,F+1)-robust
Crash & Malicious	F-Local	(2F+1)-robust
Crash & Malicious	f-Fraction local	p-fraction robust, where $2f < p \leq 1$
Byzantine	F-Total & F-Local	Normal Network is (F+1)-robust
Byzantine	f-Fraction local	Normal Network is p-robust where $p > f$

Resilient Synchronization in the Presence of Adversaries

- Synchronization is a generalization of consensus
- Assume identical LTI systems (agents)

$$\dot{x}_i(t) = Ax_i(t) + Bu_i(t)$$
$$y_i(t) = Cx_i(t).$$

  - A weakly stable, (A,B) stabilizable, (A,C) detectable

- **Problem:** Design distributed control law so that there exists open-loop trajectory

$$\dot{x}_0(t) = Ax_0(t)$$

  such that

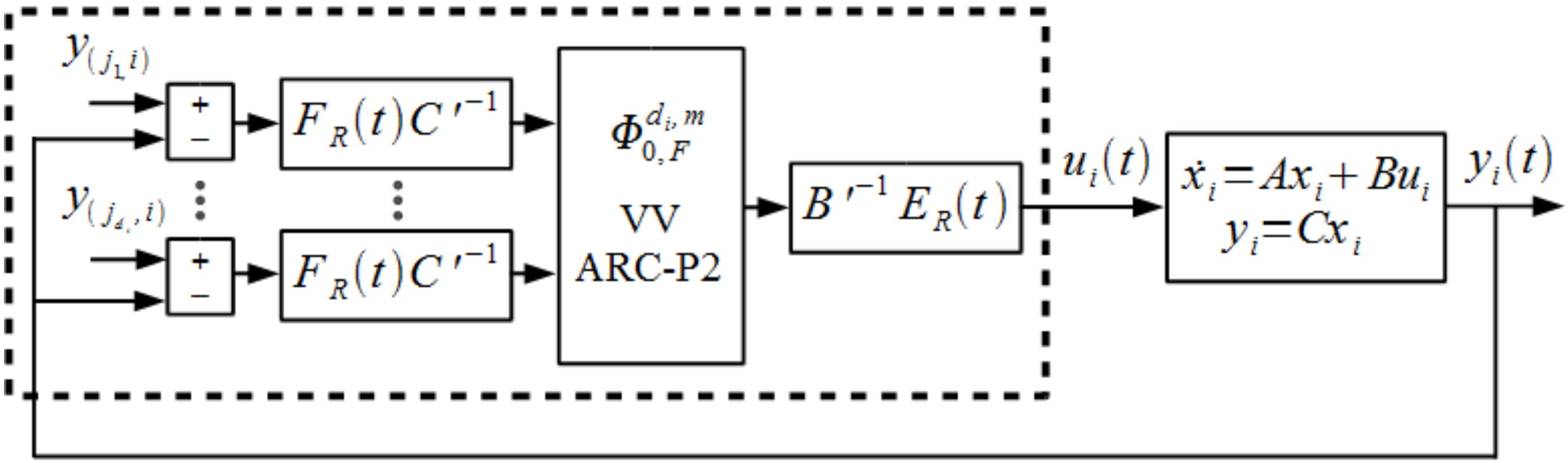
  - $x_0(0) \in S_{0,\mathcal{N}}$, where $S_{0,\mathcal{N}}$ is a known safe set that contains the $H_{0,\mathcal{N}}$ hyperrectangle
  - $||x_i(t) - x_0(t)|| \to 0$ as $t \to \infty$, for all normal agents

Resilient Synchronization Control Protocol

Assumptions

- B, C invertible
- Uniformly cts malicious outputs
- A weakly stable
- F-total malicious model
- Network $(F+1,F+1)$-robust

$$u_i(t) = B'^{-1} E_R(t) \Phi_{0,F}^{d_i,m} \left(\tilde{N}_i [I_n \otimes (F_R(t) C'^{-1})] y_{(\mathcal{V},i)}(t) - [(F_R(t) C'^{-1} y_i(t)) \otimes 1_{d_i}], w_i(t) \right)$$



$$R = Q^{-1} A Q \qquad C' = CQ \qquad B' = Q^{-1} B$$

RAS with Full State Feedback

Assumptions

- (A,B) stabilizable
- Full state feedback
- K stabilizing matrix such that $A+BK$ is stable

- A weakly stable
- F-total malicious model
- Network $(F+1,F+1)$-robust
- Uniformly cts malicious states & controller states

Then, the dynamic control law with initially relaxed controller state

$$\dot{\eta}_i = (A+BK)\eta_i - QE_R\Phi_{0,F}^{d_i,m}\left(\tilde{N}_i[I_n \otimes F_R Q^{-1}]s_{(\mathcal{V},i)} - [(F_R Q^{-1}s_i)\otimes 1_{d_i^{\text{in}}}], w_i\right)$$

$$u_i = K\eta_i,$$

where $s_j = x_j - \eta_j$ achieves RAS

RAS with Output Feedback

Assumptions

- (A,B) stabilizable
- (A,C) detectable
- K and H are stabilizing and observer matrices, resp., such that $A+BK$ and $A+HC$ are stable
- A weakly stable
- F-total malicious model
- Network $(F+1,\mathrm{IF}+1)$-robust
- Uniformly cts malicious observer states & controller states

Then, the dynamic control law with initially relaxed controller state and Luenberger observer states in some hyper-rectangle within the safe set given by

$$\begin{aligned}\dot{\eta}_i &= (A+BK)\eta_i + H(\hat{y}_i - y_i) \\ &\quad - QE_R(t)\Phi_{0,F}^{d_i,m}\Big(\tilde{N}_i[I_n \otimes F_R(t)Q^{-1}]\hat{s}_{(\mathcal{V},i)}(t) - [(F_R(t)Q^{-1}\hat{s}_i(t)) \otimes 1_{d_i}], w_i(t)\Big)\end{aligned}$$

$$\dot{\hat{x}}_i = A\hat{x}_i + Bu_i + H(\hat{y}_i - y_i) \qquad u_i = K\eta_i \qquad \hat{y}_i = C\hat{x}_i \qquad \hat{s}_j = \hat{x}_j - \eta_j$$

achieves RAS.

Algorithms to Determine Robustness

- There are $R(n)$ pairs of subsets to check, where

$$R(n) = \sum_{k=2}^{n} \binom{n}{k} \left(2^{k-1} - 1\right),$$

- $n = |\mathcal{V}|$;
- each $k = 2, 3, \ldots, n$ in the sum is the size of the k-subsets of $\mathcal{V} = \{1, 2, \ldots, n\}$. Each k-subset of $\mathcal{V}$ is partitioned into exactly two nonempty parts, $\mathcal{S}_1$ and $\mathcal{S}_2$;
- $\binom{n}{k}$ is the number of k-subsets of $\{1, 2, \ldots, n\}$;
- $2^{k-1} - 1 = S(k, 2)$ is a Stirling number of the 2nd kind, and is the number of ways to partition a k-set into 2 nonempty unlabelled subsets (swapping the labels $\mathcal{S}_1$ and $\mathcal{S}_2$ results in the same pair).

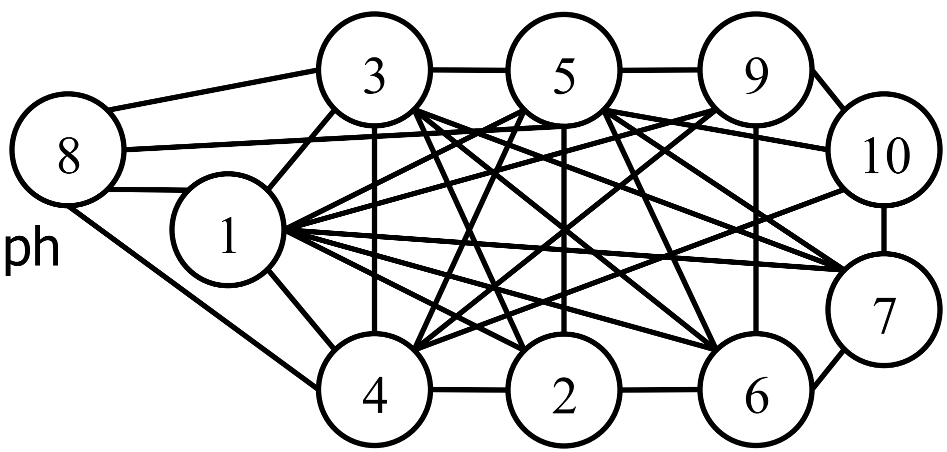
Construction of Robust Digraphs

- Let $D=(V, E)$ be a nontrivial (r,s)-robust digraph . Then, $D'=(V \cup \{v_{\text{new}}\}, E \cup E_{new})$, where v_{new} is a new node added to D and E_{new} is the directed edge set related to v_{new}, is (r,s)-robust if

$$d^{\text{in}}_{v_{\text{new}}} \geq r + s - 1$$

Preferential-attachment model

- Initial graph: K_5
- K_5 is (3,2)-robust
- Num edges / round: 4
- End with (3,2)-robust graph
- In fact, it is also 4-robust



Conclusions and Future Work

- Resilient Asymptotic Consensus
  - Continuous-Time
  - Discrete-Time
    - Synchronous
    - Asynchronous
- Resilient Asymptotic Synchronization
  - Continuous-time LTI systems
- Network robustness
- Algorithms for determining robustness

- Broader distributed control and estimation problems
- Hierarchical multi-tier networks comprised of agents with various security protections and privileges
- Optimize the action of cooperative agents using attack models that represent adversary strategies

Publications

- Asynchronous robust networks
  - Heath J. LeBlanc, Xenofon Koutsoukos: Resilient Asymptotic Consensus in Asynchronous Robust Networks. Allerton Conference on Communication, Control, and Computing. Monticello, IL. October, 2012.
- Discrete-time robust networks
  - Heath J. LeBlanc, Haotian Zhang, Shreyas Sundaram, Xenofon Koutsoukos: Consensus of Multi-Agent Networks in the Presence of Adversaries Using Only Local Information. Conference on High Confidence Networked Systems (*HiCoNS 2012*), Beijing. China. April, 2012. pp. 1–10.
- High-degree networks
  - Heath J. LeBlanc, Xenofon Koutsoukos: Low Complexity Resilient Consensus in Networked Multi-Agent Systems with Adversaries. Hybrid Systems: Computation and Control (*HSCC 2012*). Beijing, China. April, 2012. pp. 5–14. **Honorable Mention for Best Paper Award**.
- Complete networks
  - Heath J. LeBlanc, Xenofon Koutsoukos: Consensus in Networked Multi-Agent Systems with Adversaries. Hybrid Systems: Computation and Control (*HSCC 2011*), Chicago, IL. April, 2011. pp. 281–290.
- Overall approach
  - Heath J. LeBlanc, Resilient Cooperative Control of Networked Multi-Agent Systems, PhD Thesis, Department of EECS, Vanderbilt University, August 2012.

Acknowledgements: National Science Foundation (CNS-1035655, CCF-0820088),
the U.S. Army Research Office (ARO W911NF-10-1-0005), and Lockheed Martin.