Dimitra Panagou
Assistant Professor, Aerospace Engineering

Multi-Agent Planning and Control
Ground, marine, aerial, space vehicles

Safety and Resilience under Uncertainty
Towards advancing autonomy

Nonlinear Control and Estimation
Robust control, estimation and learning
Dimitra Panagou
Assistant Professor, Aerospace Engineering

7 PhD Students
2 Postdoctoral Researchers
4 MS Students
2 Undergraduate Students

Dynamic Coverage / Exploration

Human-Robot Interaction
Adversaries may compromise information and safety in manned/unmanned teams

- Communication must be resilient
- Coordination must be safe

Task 1: Resilience in multi-agent networks against unknown adversaries

Task 2: Safety in multi-agent networks against unknown adversaries

Task 3: Integration of safety and resilience; adversary detection
Safety & Resilience Architecture

- Safety Control
  - Agent: \( \dot{x}_i = f(t, x_i, u_i) \)
  - Barrier Control
  - Resilient Estimation
    - Information Filtering
      - Reference Estimate
      - State & Goal Estimate
  - Secure Communication
    - (Time- and spatially-varying k-circulant graphs)

Physical Adversaries

Cyber Adversaries
Earlier Resilience Results

- Network as a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ $\mathcal{V} = \{1, \ldots, n\}$ $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$

- Up to F adversaries
  - Share malicious information and/or do not play consensus

- Resilient Communication Graphs
  - $r$-robustness and $(r,s)$-robustness

- Resilient Filtering: W-MSR algorithm
  - Principle: Each agent
    - sorts received information
    - filters out the F highest and F lowest values
  - Guaranteed consensus for normal agents
    - The network must be $(2F+1)$-robust, or $(F+1,F+1)$-robust

- Challenges:
  - Checking $r$-robustness and $(r,s)$-robustness is NP-hard
  - Consensus to arbitrary reference values is not guaranteed
Our Resilience Results

- **[Our earlier work]:** Proved that $k$-circulant graphs have $r$-robustness and $(r,s)$-robustness as functions of $k$
  - Resilient, scalable network topologies [CDC17]

- **[Year 1]:** Resilient consensus to arbitrary reference values in time-invariant and time-varying graphs
  - Resilient Leader-Follower consensus [ACC18]

- **[Year 1]:** **Resilient formation control**
  - In finite time under bounded control inputs [CDC18]

- **[Year 1]:** Graph $r$-robustness and $(r,s)$-robustness as a MILP
  - More efficient than state-of-the-art methods [ACC19]
  - Approximate lower bounds of $r$- and $(r,s)$-robustness

- **[Year 2]:** Resilient Barriers for Undirected Networks

- J. Usevitch and D. Panagou (*TAC, Automatica*)
Resilient Formation Control

- Time invariant digraph $D = (V, E)$, $V = \{1, \ldots, n\}$
- Agent states $p_i \in \mathbb{R}^n$, $i \in V$
- $\xi_i \in \mathbb{R}^n \ \forall i \in V$: Formation vectors
- $\tau_i = p_i(t) - \xi_i \ \forall i \in V$: Center of formation

Algorithm 1 Continuous-Time Filtering

```
procedure UPDATEFILTEREDLIST
    Calculate $\tau_{ij} = \|\tau_j - \tau_i\| \ \forall j \in V_i$
    if $t = m\epsilon_d$, $m \in \mathbb{Z}_{\geq 0}$, $\epsilon_d > 0$ then
        Sort $\tau_{ij}$ values such that $\tau_{ij_1} \geq \ldots \geq \tau_{ij_{|V_i|}}$
        $R_i(t) \leftarrow \{j : \tau_{ij} \in \{\tau_{ij_{F+1}}, \ldots, \tau_{ij_{|V_i|}}\}\}$
```

$R_i(t)$: Filtered list after removing $F$ maximum elements

Update on discrete instances $t = m\epsilon_d$
Closed loop system:

\[ \dot{\tau}_i = u_i, \]
\[ u_i(t) = \gamma_i(t) \sum_{j \in R_i(t)} w_{ij}(t)(\tau_j - \tau_i)\|\tau_j - \tau_i\|^{\alpha-1}, \quad 0 < \alpha < 1 \]

where

- \( \gamma_i(t) = \frac{\sigma_i(t)}{\|u_i^p(t)\|} \)
- Saturation function:
  \[ \sigma_i(t) = \min\{\|u_i^p(t)\|, u_M\}, \]
  \[ u_i^p(t) = \sum_{j \in R_i(t)} w_{ij}(t)(\tau_j(t) - \tau_i(t))\|\tau_j(t) - \tau_i(t)\|^{\alpha-1}, \quad 0 < \alpha < 1 \]
- Input satisfies bounds \( \|u_i\| \leq u_M \forall i \in \mathcal{V} \)

**Theorem 2**

Consider a digraph \( \mathcal{D} \) which is an RDAG with parameter \( 3F + 1 \), where \( S_0 = \mathcal{L} \) and \( \mathcal{A} \) is an \( F \)-total set. Under the proposed closed loop dynamics, \( \tau_i \) will converge to \( \tau_L \) in finite time for all normal agents \( i \in \mathcal{N} \).
• k-circulant communication graph
• Leaders
  - Transmit center-of-formation to followers
• Followers
  - Filter out malicious agents via W-MSR
  - Reconstruct center-of-formation and use formational offset to compute target waypoint
• [Year 2]: Resilient Reference Generation
  – Resilient agreement on reference trajectories via the MS-RPA algorithm
    [*CDC19*]
• [Year 2]: Resilient Barrier Functions
  – Safe trajectories and network robustness
    [Ongoing work]
• [Year 2]: Simulations and Experiments
  – ROS/Gazebo and with Ground/Aerial Robots
**Multi-Source Resilient Propagation Algorithm**

- Strongly \((2F+1)\)-robust network w.r.t. leaders
- Up to \(F\) misbehaving agents
  - including misbehaving leaders

**Cyber state (reference) dynamics:**

**Leaders:**

\[
x_l(t+1) = \begin{cases} 
 f_r(\tau) & \text{if } t - t_0 = \tau \eta - 1, \ \tau \geq 1, \ \tau, \eta \in \mathbb{Z}_+ \\
 x_l(t) & \text{otherwise}
\end{cases}
\]

**Followers:**

\[
x_j(t+1) = x_j(t) + \alpha(t)u_j(t)
\]

\[
\alpha(t) = \begin{cases} 
 1 & \text{if } t - t_0 = \tau \eta - 1, \ \tau \geq 1, \ \tau, \eta \in \mathbb{Z}_+ \\
 0 & \text{otherwise}
\end{cases}
\]
Multi-Source Resilient Propagation Algorithm

- Leaders transmit state to out-neighbors
Multi-Source Resilient Propagation Algorithm

- Leaders transmit state to out-neighbors
- Followers accept message if identically received from at least \((F+1)\) in-neighbors
Multi-Source Resilient Propagation Algorithm

- Leaders transmit state to out-neighbors
- Followers accept message if identically received from at least \((F+1)\) in-neighbors
- Accepted messages retransmitted
- Repeated for \(\eta\) time steps, then process restarts
- Allows for leaders’ signal to be time-varying
Resilient Reference Generation

Leaders:
- Determine trajectory for center of formation (COF)
- Encode COF trajectory into unique parameters
- Resiliently transmit parameters to out-neighbors

Followers:
- Receive and accept parameters only if resilience criteria satisfied
- Reconstruct unique trajectory of COF
- Add local formation offset to obtain local desired trajectory
- Track local trajectory
Resilient Barriers

- Multi-Task Barrier Functions [TAC18]
  - Safe trajectories
  - Network connectivity

- Prior work assumes no malicious agents. **However:**
  - Even one misbehaving agent can disrupt operation of normal agents since BF's prevent normal agents to disconnect from malicious agents

- **Resilient** Barrier Functions:
  - Normal agents filter out the influence of malicious agents
  - **Without knowing who the malicious agents are!!**
  - Safety and network robustness
• **Resilient Filtering:** At each time $t$, each normal agent $i$:

  - Forms a sorted list of the norm between its own state and the state of its in-neighbors $V_i(t)$.
  - Filters out the in-neighbors $\mathcal{R}_i(t)$ with the highest norm difference.

• **Resilient Barrier Controller:**

  $u_i(t) = - \sum_{j \in V_i \setminus \mathcal{R}_i(t)} \nabla y_j \Psi_{ij}^e(\|y_{ij}\|) = \sum_{j \in V_i} \nabla x_j \Psi_{ij}^c(\|x_{ij}\|)$

where

- **Edge Maintenance**
  \[ \Psi_{ij}^e(\|y_{ij}\|) = \frac{\|y_{ij}\|^2}{r s - \|\tau_{ij}\| + (r s - \|\tau_{ij}\|)^2} \]

- **Collision Avoidance**
  \[ \Psi_{ij}^c(\|x_{ij}\|) = \begin{cases} \frac{(\|x_{ij}\| - \|\tau_{ij}\|)^2}{\|x_{ij}\| - d_s + \frac{(ds - \|\tau_{ij}\|)^2}{\mu_2}} & \text{if } \|x_{ij}\| \leq \|\tau_{ij}\| \\ 0 & \text{otherwise} \end{cases} \]

Agent state: \( x_{ij} = x_i - x_j, \quad x_i, x_j \in \mathbb{R}^n \)
Formation offset: \( \tau_{ij} = \tau_i - \tau_j, \quad \tau_i, \tau_j \in \mathbb{R}^n \)
\[ y_i = x_i - \tau_i \quad y_{ij} = y_i - y_j \]
\( y_{ij} = 0 \quad \forall i, j \quad \Rightarrow \text{formation achieved} \)
AION R1 Rovers (7 robots)
  • Pixhawk 2.1 with Ardupilot
  • Nvidia Jetson TX2 architecture with Ubuntu and ROS
  • Capable of skid steering or differential drive

Tests in M-Air (April - May 2019)
  • Resilient trajectory tracking
Physical Adversaries

Resilient Estimation

Secure Communication
(Time-and spatially-varying k-circulant graphs)
**Problem statement:** Design control input $a_i$ so that

- $|r_i(t) - r_j(t)| \geq d_m$ for all $t \geq 0$ (safety)
- $|r_i(t) - r_{g_i}| \to 0$ as $t \to T < \infty$ (reach goal)

**Class-A** (controlled agents):

\[
\begin{align*}
\dot{r}_i &= u_i + w(t, r_i), \\
\dot{u}_i &= a_i \\
y_i &= r_i
\end{align*}
\]

where $r_i, y_i, u_i, a_i \in \mathbb{R}^2$, $w: \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}^2$

**Class-B** (uncontrolled agents)

**Assumptions:**

- Limited and erroneous communication
  - Communication radius $R_c$ ($N_i = \{j \mid \|r_i - r_j\| \leq R_c\}$)
  - Position sensed with bounded error $\varepsilon_s$
- Wind is unknown but bounded:
  - Mean value $\bar{w}$ is known
  - Maximum deviation $\delta_w = \max w - \bar{w}$ is known
- Class-B agents (dynamic obstacles)
  - Upper bounded velocity, arbitrary direction
Main idea:
- Finite-time state observer
- Finite-time feedback controller
  - Tracking of safe desired velocity $u_{id}$:
- Design input $a_i$ so that $\hat{u}_i \to u_{id}$ in finite time.

Outcomes:
- Robustness against
  - Wind disturbance
  - Sensing error and
  - Velocity tracking error
- Robust safety region $d_s = d_m + \epsilon$
Finite-Time Estimator (State Observer)

Finite-Time State Estimator

\[
\dot{\hat{r}}_i = \dot{\hat{u}}_i + k_1 (y_i - \hat{y}_i)y_i - \hat{y}_i|^{\alpha_1 - 1} + \bar{w}, \\
\dot{\hat{u}}_i = a_i + k_2 (y_i - \hat{y}_i)y_i - \hat{y}_i|^{\alpha_2 - 1},
\]

where \( \hat{\dot{y}}_i = \hat{\dot{r}}_i, \ k_1, k_2 > 0, \ 0 < \alpha_1 < 1, \ \alpha_2 = 2\alpha_1 - 1 \)

Theorem 1

The state estimation error is bounded as

\[
\| [r_{ie} \ u_{ie}]^T \| \leq \begin{cases} 
\| u_{ie}(0) \|, & 0 \leq t \leq T_{1i} \\
\ c \delta_w^\gamma, & t \geq T_{1i} 
\end{cases}
\]

where \( c > 0, \gamma > 1 \) and \( T_{1i} < \infty \).
Finite-Time Control (State Feedback)

**Finite-Time State-Feedback Controller**

\[
a_i = \dot{u}_{id} - \lambda (\hat{u}_i - u_{id}) ||\hat{u}_i - u_{id}||^{\alpha_3 - 1} - k_2 (y_i - \hat{y}_i) |y_i - \hat{y}_i|^{\alpha_2 - 1}
\]

\[
u_{id} = v_{VF} - k_1 (y_i - \hat{y}_i) |y_i - \hat{y}_i|^{\alpha_2 - 1} - \bar{w}
\]

where
- \(\lambda > 0, 0 < \alpha_3 < 1\)
- \(v_{VF}\): Desired velocity aligned with a vector field
- Safety via repulsive field \(F_{ij}\):
  - Acts along \((r_i - r_j)\)
  - Active only if \(d_{ij} = ||r_i - r_j|| \leq R_c\)
- Convergence via attractive field \(F_{gi}\)
  - Acts along \((r_{gi} - r_i)\)
  - Active only if \(d_{ij} = ||r_i - r_j|| > d_r > d_m\) for all \(j \in N_i\)
- Blending of \(F_{ij}\) and \(F_{gi}\) via bump function
  - \(F_i = \sum_{j \in N_i} \sigma_{ij} F_{ij} + \prod_{j} (1 - \sigma_{ij}) F_{gi}\)

\[
\sigma_{ij}(d_{ij}) = \begin{cases} 
1, & d_m \leq d_{ij} < d_r; \\
a d_{ij}^3 + b d_{ij}^2 + c d_{ij} + d, & d_r \leq d_{ij} \leq R_c; \\
0, & d_{ij} > R_c;
\end{cases}
\]
Theorem 2 (Safety)

If the safe distance is taken as $d_s = d_m + r_e + \delta_e + \epsilon_s$, where
- $d_m$ is the minimum required separation
- $r_e$ is the maximum overshoot in position error in transient period
- $\delta_e$ is the state-estimation error
- $\epsilon_s$ is the sensing error
then, the closed-loop trajectories are collision-free, i.e., $\|r_i - r_j\| \geq d_m$ for all $t \geq 0$, where agent $j$ can belong to either class-A or class-B.

Theorem 3 (Convergence)

Under the effect of designed control input $a_i$, the closed-loop trajectories of each class-A agent $i$ reach a $\delta_{ie}$-neighborhood around the goal location $r_{gi}$ in finite time, i.e., $\exists T_i < \infty$, such that $\|r_i(t) - r_{gi}\| \leq \delta_{ie}(t)$ for all time $t \geq T_i$. 
Human-Robot Interaction

Year 1

Year 2

Year 3

Human

Swarm

Modeling

Prediction

Motion Planning

Control

Online learning of human activity and robot policies

Prediction of human activity based on prior knowledge/data (Offline + Online)

Prediction of human activity and future motion based on observed motion

Selection of Robot Action under Uncertainty

Model any Environment Constraints or other geometric task based constraints

Collision Avoidance

Maximum Cost
Human Robot Interaction (HRI) in IVA

- Algorithmic developments in:
  - Multiple task servicing and/or inventorying
  - Environmental monitoring
  - Mobile camera scenario

- NASA Astrobot

Autonomous Co-Robots in EVA

- Algorithmic contributions in:
  - Autonomous exterior inspection

- Hypothetical NextGen Mini-AERCam

Image: NASA
Earlier scenario for Astrobée (Bualat et al., 2015):
- Astrobée streams mobile video images of crew to ground controllers
- Astrobée is teleoperated by ground controller.
- The astronaut is out of the loop.

Our scenario (W. Bentz et al, ICRA 2019):
- The Astrobée learns assistive camera views online
- Broadcasts them to the astronauts via an Augmented Reality headset to assist them in multitasking.

NASA has already begun initial evaluations of Augmented Reality aboard the ISS via Microsoft HoloLens.


Kelly, Scott (StationCDRKelly): “This #saturdaymorning checked out the @Microsoft #HoloLens aboard @Space_Station! Wow! #YearInSpace” 20 February 2016, 11:01 AM. Tweet.
• Domain $\mathcal{D} \subset \mathbb{R}^3$ with visual regions of interest $\mathcal{I}_i, \forall i \in \{1, ..., N\}$.

• A human, with visual field $\mathcal{S}_H$, splits their attention between the visual regions.
  – Due to multi-tasking activity

• An aerial robot streams camera views of $\mathcal{S}_R$ to the human’s augmented reality display.

• Objectives:
  – Determine the locations and relative importance of each member of $\mathcal{I}$ via unsupervised learning.
  – Direct the robot to view elements of $\mathcal{I}$ for which the human is not currently observing.

• We aim to assess:
  – perceived efficacy, comfort, and naturalness of interactions,
  – quantitative reductions in task completion time and head motion.
• Human visual acuity:
  – Modeled as a truncated Gaussian with $\alpha_{H} = 60^\circ$ and $\sigma_{H} = 8^\circ$.
  – Good fit for clinically-drawn visual acuity plots common in sensory physiology textbooks (Schmidt, 1986).

• Robot sensing model:
  – Anisotropic spherical sector which degrades in quality both towards the periphery as well as with distance.
  • $S_R$ defined in terms of sensing constraint functions:
    $$c_1 = \beta R_R^2 - (\hat{x} - x_R)^2 - (\hat{y} - y_R)^2 - (\hat{z} - z_R)^2,$$
    $$c_2 = \alpha_R - \phi_R(\hat{p}),$$
    \[
    S_R(\hat{q}_R, \hat{p}) = \begin{cases} 
    \frac{C_1 C_2}{C_1 + C_2}, & \text{if } \text{card}(\hat{C}) < 2 \land \|\hat{p} - p_R\| > 0; \\
    0, & \text{otherwise},
    \end{cases}
    \]
- The human shifts their attention among $I$
- The volume occupied by $S_H$ varies over time
- Human’s visual activity:
  \[ Q(t, \bar{p}) = \int_0^t S_H(\bar{q}_H(\tau), \bar{p}) d\tau, \]

  This raw data is fit to a Gauss mixture model denoted $\psi_H$ via expectation maximization (EM-GMM).

- Scale the output of EM-GMM via a logistic function so that the region of interest $I_i$ currently observed by the human decays to zero:
  \[ \psi_H = \frac{\psi_H}{1 + \exp(-k(\phi_H - \alpha_H))}. \]

- A gradient-following controller sweeps $S_R$ across $\psi_H$ to minimize:
  \[ J(t) = \int_D \max(0, S_R - \psi_H)^2 d\bar{p} \]

  - Taking first derivative we have:
    \[ \dot{J}(t) = \int_D 2 \max(0, S_R - \psi_H) \left( \frac{\partial S_R}{\partial t} - \frac{\partial \psi_H}{\partial t} \right) d\bar{p} \]

  - Design controllers such that $\dot{J} \leq 0$. 

TABLE I
ASYMPTOM AND REACTION TIME STATISTICS

<table>
<thead>
<tr>
<th></th>
<th>Mean Assembly Time (sec)</th>
<th>Std. Dev. (sec)</th>
<th>Mean Reaction Time (sec)</th>
<th>Std. Dev. (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boat Exp.</td>
<td>161.1</td>
<td>38.6</td>
<td>5.7</td>
<td>4.5</td>
</tr>
<tr>
<td>Boat Ctrl.</td>
<td>151.0</td>
<td>49.3</td>
<td>6.0</td>
<td>4.6</td>
</tr>
<tr>
<td>Cloud Exp.</td>
<td>190.6</td>
<td>46.1</td>
<td>5.4</td>
<td>4.0</td>
</tr>
<tr>
<td>Cloud Ctrl.</td>
<td>174.0</td>
<td>45.7</td>
<td>5.9</td>
<td>4.1</td>
</tr>
</tbody>
</table>
• From the preliminary experiment, we learned that:
  – the Vufine Wearable Display was not ergonomic,
  – localization of $\mu_i$ was frequently inaccurate without eye tracking and gaze depth data,
  – subjects were generally dissatisfied with the seemingly random manner in which the robot selected from $\mathcal{T}$

• In our ongoing experiments, we have:
  – replaced the Vufine with a Microsoft Hololens,
  – added Pupil Labs eye trackers and an Intel RealSense depth camera to more accurately localize $\mu_i$,

• Also, we allow for feedback from the human in terms of:
  – the human’s task execution sequence,
  – the human’s level of satisfaction with the robot’s chosen task.
• We now consider the most recent tasks in sequence.
  – This provides context as to how tasks may be related within the environment.

• Markov Decision Process (MDP) to model the human-robot interaction:
  – States represent the most recent $K$ visual interests of the human in sequence.
  – Actions represent the visual interest observed by the robot for any human state.
  – Given visual interests A, B, and C, the MDP formulation for memory length $K = 3$ may be represented by:

  ![MDP Diagram]

  • Note that:
    – each state has two defined transitions with the latter two characters transitioning to the former two,
    – each transition can occur given either robot action A, B, or C.
    – **We shall learn the robot’s action policy via Q-learning.**
Algorithm for Q-learning of Assistive Views:

- Initialize $Q(s, a) = 0 \forall s, a$.
- Observe the human’s visual interest until the outputs of sequential EM-GMM executions have converged.
- Launch assistive robot.
- Repeat:
  - Choose $a_t$ that maximizes $Q(s_t, a_t)$ with probability $(1 - \varepsilon)$, otherwise choose a random $a_t$. ($\varepsilon$-greedy)
  - Take action $a_t$ and observe reward $r_t$ and state $s_{t+1}$.
  - Update: $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left( r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right), s_t \leftarrow s_{t+1}$.

- The reward consists of two components: $r_t = h_t + f_t$.
  - Human input: $h_t \in \{-1, 0, 1\}$ is requested of the user after each transition by asking whether the robot’s camera view is bad, neutral, or good.
  - Transition time reward: $f_t = \eta \text{ sat} \left( \frac{t_o - t_t}{t_t \eta} \right)$, where $t_t$ and $t_o$ respectively refer to the time spent transitioning to and then observing a visual interest region.
    - $f_t$ penalizes $a_t$ that result in $t_t > t_o$.
    - $\eta > 0$ may be increased to place greater training weight on the transition time reward.
Physical Adversaries

Resilient Estimation

Secure Communication
(Time- and spatially-varying k-circulant graphs)
Two approaches:
1) Capture/interception
2) Herding

Objectives:
• **Attackers:** Reach the protected area $P$ as a team
• **Defenders:** Herd the attacking team to $S$
• Avoid static obstacles and collisions

Assumptions:
1) Risk-averse attackers avoid defenders
2) Defenders know state of the attackers

General Approach:
1) Formation of the defenders around the attackers
2) Guide the attackers to the safe area $S$
3) Perform 1) and 2) while avoiding obstacles
Herding Single Attacker

Assumptions:
• Attacker navigates to $P$ under bounded speed
• Attacker avoids defenders

Strategy:
• Defenders form a circular arc around the attacker
• Formation is safely guided to a predefined safe area
Herding Multiple Attackers

Strategy:
• **Open Formation:** Place the defenders on an arc in the way of the attackers
  - Use virtual $\beta$-agents moving along the boundary of the obstacles and potential functions to avoid collisions
• **StringNet Formation:** Close around the attackers
• Guiding StringNet to safe area: StringNet as a rigid formation

Assumptions:
• Risk-averse attackers attack as a flock with double integrator dynamics and bounded acceleration
• Strings between the defenders do not allow attackers to pass through
Resilient network architectures
- Resilient information filtering
- Resilient safety control
- In the presence of adversaries

- Human-Robot Assistive Interaction in Augmented Reality Environments

- Next steps
  - Directed networks
  - Complex dynamics
  - Adversary detection
  - Human-Swarm Interactions

Physical Adversaries

Summary