Learning-Based Actuator Placement and Receding Horizon Control for Security against Actuation Attacks

KYRIAKOS G. VAMVOUDAKIS

THE DANIEL GUGGENHEIM SCHOOL OF AEROSPACE ENGINEERING
Affiliate
Institute for Robotics and Intelligent Machines
Decision and Control Laboratory
The Center for Machine Learning
Institute of Information Security and Privacy
Center for the Development and Application of Internet of Things Technologies
I. Motivation

II. Receding Horizon Control for Security against Stealthy Actuation Attacks

III. Learning-Based Actuator Placement for Security against Actuation Attacks

IV. Conclusion and Future Ideas
#1 Priority: Cyber-Physical Systems
Our lives depend on them.

An airplane is a network of computers.
Cyber-physical attacks: A reality we need to face.

Here's a heart stopper: On March 21, the Department of Health and patients that hundreds of thousands of implants, "potentially impacting product functionality."

While the FDA noted that some company's devices for the interception of patient data, there have been reports they are working to patch the vulnerabilities. The full extent of the damage are still unknown.

A power cut that hit part of Ukraine's energy grid has been judged a cyber-attack by researchers investigating the incident.
Receding Horizon Control for Security against Stealthy Actuation Attacks
Types of attacks

- **Sensor Attacks.**
  - Injection of faulty sensor measurements.

- **Actuator attacks.**
  - Injection of faulty control inputs.

- **Denial-of-Service attacks.**
  - Jamming of sensors/actuators.

- **Stealthy attacks.**
  - Attacks that cannot be detected.

Introduction

System + State Estimation

Uncertainty makes attack stealthy.

Detection

System

State Estimation

Actuator attack

Attack leads to unsafe regions.
How to mitigate?

➤ **Game Theory.**
Secure decision making by considering worst-case scenarios.
Cooperation between operators using equilibrium-based concepts.

➤ **Receding Horizon Control (RHC).**
Devises stable, optimal control laws.
Allows constraints that characterize stealthy attacks.

➤ **Moving Horizon Estimation.**
Combines RHC and game-theory for secure state estimation.
**Framework description**

### Plant Dynamics

Defender’s Input

\[ \dot{x}(t) = Ax(t) + Bu(t) + d(t) + \sum_{l=1}^{N_a} a_l(t), \]

Output (known by everyone)

\[ y(t) = Cx(t), \quad x(t_0) = x_0. \]

Disturbance (known by the attackers)

\[ \text{Disturbance} \]

### Attack-Free Dynamics

Defender’s goals

- Estimate the initial condition.
- Regulate the system optimally.

\[ \dot{x}_h(t) = Ax_h(t) + B(u(t) + d_h(t)), \]

\[ y_h(t) = Cx_h(t), \quad x_h(t_0) = x_0. \]

**Assumptions:**

\[ \|d(t)\| < \Delta. \]
\[ \|a_l(t)\| < \bar{a}, \quad l \in N_a. \]

**Stealthiness:**

\[ y = y_h. \]

**Atack-Free Dynamics**

**Atackers’ goals**

- Cooperate to achieve stealthiness.
- Deteriorate the plant’s performance.

**Defender’s goals**

- Estimate the initial condition.
- Regulate the system optimally.
State estimation and game-theoretic RHC

- **Non-Stealthy attacks**: Feedback control can be stopped after detection.
- **Stealthy attacks**: How to deal with those?

Game theoretic RHC & state estimation

- **Estimate** a worst-case initial condition based on output history.
- **Predict** worst-case stealthy attacks using the attack-free model.
- **Control** by computing the zero sum Nash equilibrium.

Lead to safe mitigation policies.

Control allocation guaranteeing secure stabilization.
State estimation and game-theoretic RHC

**Problem**

- Output data is collected over a past horizon $[t_j - T, t_j], j \in \mathbb{N}$.
- State estimator uses the past plant model:

  $\dot{x}(t) = Ax(t) + B \left( u(t) + d(t) + \sum_{l=1}^{N_a} a_l(t) \right),$

  $y(t) = Cx(t) \quad t \in [t_j - T, t_j], j \in \mathbb{N}.$

**BUT!!**

RHC optimizes the future,

$t \in [t_j, t_j + T], j \in \mathbb{N}.$

Incompatible
Solution

Reverse the past model!

**Reversed past model**

\[
\dot{x}_p^d(t) = -Ax_p^d(t) - B\left(u(2t_j - t) + d(2t_j - t) + \sum_{l=1}^{N_0} a_l(2t_j - t)\right), \quad t \in [t_j, t_j + T], \quad j \in \mathbb{N}.
\]

**Theorem**

\[
x_p^d(t_j) = x(t_j), \quad j \in \mathbb{N}, \text{ if and only if:}
\]

\[
Cx_p^d(t) = y(2t_j - t), \quad \forall t \in [t_j, t_j + T].
\]
State estimation and game-theoretic RHC

Game-theoretic RHC

$$\min_{u_i} \max_{A_i} J^d(t_j) = \int_{t_j}^{t_j+T} \left( \|x^d(\tau)\|_Q^d + \|u(\tau)\|_R^d + \sum_{l=1}^{N_a} \|a^d_l(\tau)\|_{K_{li}, a_i} \right) d\tau + \|x^d(t_j + T)\|_F^d.$$

Ideally, $$T \to \infty$$

- Penalizes attacks with unlikely high values.
- Needed for stability

Constraints

Dynamic

$$\dot{x}^d(t) = Ax^d(t) + B \left( u(t) + \sum_{l=1}^{N_a} a^d_l(t) + d^d(t) \right),$$

$$\dot{x}_p^d(t) = -Ax_p^d(t) - B \left( u_p(t) + \sum_{l=1}^{N_a} a_{p,l}^d(t) + d_p^d(t) \right),$$

$$\dot{x}_h^d(t) = Ax_h^d(t) + B \left( u(t) + d_h^d(t) \right).$$

Output Compatibility

Path

$$Cx_h^d(t) = Cx(t),$$

$$Cx_p^d(t) = y_p(t).$$

Disturbances bounded by $$\Delta$$

Worst-case state estimation

Boundary

$$x^d(t_0) = x_h^d(t_0) = x_p^d(t_0).$$
State estimation and game-theoretic RHC

Problem: Cost not concave w.r.t. maximizers

Solution: Concavification of the cost

Relaxed game

\[
\min_{u \in \mathcal{F}_u^j} \max_{A \in \mathcal{F}_A^j} \tilde{J}^d (u, A; t_j) = \int_{t_j}^{t_j+T} \left( \| x^d(\tau) \|_{Q^d} + \| u(\tau) \|_{R^d} - \sum_{l=1}^{N_a} \| a^d_l(\tau) \|_{K_i^d, \bar{a}} \right) \text{d}\tau + \| x^d(t_j + T) \|_{F^d},
\]

Approaches original game as induced term weights → 0

Induced concave terms
State estimation and game-theoretic RHC

Enforce path constraints by demanding:

\[ \epsilon_h^d(t_j + T) = \epsilon_p^d(t_j + T) = 0 \]

where:

\[ \epsilon_h^d(t) = \int_{t_j}^{t} \| Cx_p^d(\tau) - y_p(\tau) \|_I \ d\tau \]
\[ \epsilon_p^d(t) = \int_{t_j}^{t} \| Cx_h^d(\tau) - C\dot{x}^d(\tau) \| \ d\tau \]

**Theorem**

The Nash defender’s policy is given, for all, \( i = 1, \ldots, N \), by

\[ u^* \left( t \right) = -R^{-1} B_T \left( \lambda(t) + \lambda_h(t) \right), \]
\[ a_{i}^{d*,}(t) = \bar{a} \times \tanh \left( K_l^{d^{*}} B_T \lambda(t) \right), \quad \forall l \in N_a, \]
\[ d_{i}^{d*,}(t) = \Delta \times \tanh \left( D^{d^{*}} B_T \lambda(t) \right), \]
\[ d_{h}^{d*,}(t) = \Delta \times \tanh \left( D^{d^{*}} B_T \lambda_h(t) \right), \]
\[ a_{p,i}(t) = -\bar{a} \times \tanh \left( K_l^{d} B_T \lambda_p(t) \right), \quad \forall l \in N_a, \]
\[ d_{p,i}(t) = -\Delta \times \tanh \left( D^{d} B_T \lambda_p(t) \right), \]
\[ \dot{\lambda}(t) = -A^T \lambda(t) - Q^d x^d(t) - C^T C \left( x^d(t) - x_h(t) \right) \rho_h(t) \]
\[ \dot{\lambda}_p(t) = A^T \lambda_p - C^T \left( C x_p^d(t) - y^2(2t_j - t) \right) \rho_p(t), \]
\[ \dot{\lambda}_h(t) = -A^T \lambda_h(t) - C^T C \left( x_h^d(t) - x^d(t) \right) \rho_h(t), \]
\[ \dot{\rho}_p(t) = \dot{\rho}_h(t) = 0, \]
\[ \lambda_p(t_j + T) = \lambda_h(t_j + T) = 0, \]
\[ \lambda(t_j + T) = F^d x^d(t_j + T), \]
\[ \lambda(t_j) + \lambda_h(t_j) + \lambda_p(t_j) = 0, \]
State estimation and game-theoretic RHC

Stability Guarantees

Assumption:

The terminal cost \( F(x^d) = \|x^d\|_{F^d} \) is proper: there exists a controller \( \psi(x^d(t)) \) and weighting matrices such that

\[
\frac{dF(x^d(t))}{dt} \bigg|_{u=\psi(x^d)} \leq -\|x^d(t)\|Q^d - \|\psi(x^d(t))\|_{R^d} \\
+ \sum_{l=1}^{N_a} \|a_l^d(t)\|_{K_{l,a}^d} + \|d^d(t)\|_{D^d,\Delta},
\]

Interpretation: Do not underestimate the attackers!

Terminal cost is an ISS Lyapunov

Theorem:

Closed trajectories are bounded for sufficiently small weighting matrices on maximizers.
What if the past output is available only intermittently? **Assume** the output is available every $\delta$ seconds.

**Theorem**

The worst-case initial condition can be uniquely estimated, given worst case past disturbances and attacks, as long as,

$$\text{rank} \left( \begin{bmatrix} C^T & (Ce^{-\delta \cdot 1})^T & \cdots & (Ce^{-N\delta \cdot 1})^T \end{bmatrix} \right) = n.$$
State estimation and game-theoretic RHC

But... what could the attackers do?
Stealthiness

The attackers $i \in N$ can remain undetected over the interval $t \in [t_j, t_j + T]$ if

$$C \left( x^i(t) - x^i_h(t) \right) = 0, \quad \forall t \in [t_j, t_j + T],$$

where

$$\dot{x}^i(t) = Ax^i(t) + B \left( u^i(t) + d(t) + \sum_{l=1}^{N_a} a_l(t) \right),$$

$$\dot{x}^i_h(t) = Ax^i_h(t) + B \left( u^i(t) + d^i(t) \right),$$

$x^i(t_j) = x^i_h(t_j) = x_j, \ t \in [t_j, t_j + T].$

Constrained RHC allows to impose this.

Knowledge of disturbance necessary for stealthiness.

Not equal to actual defender input!
The attackers’ point of view: equilibrium

Cooperative game

\[
\min_{u^i \in \mathcal{F}_u^i} \max_{a_i \in \mathcal{F}_a^i, d_h \in \mathcal{F}_d^j} \left( \sum_{j}^{t_j+T} \left( \|x^i(\tau)\|_{Q^a_i} + \|u^i(\tau)\|_{R^a_i} - \|a_i(\tau)\|_{K^a_i} - \|d^h(\tau)\|_{D^h_i} \right) d\tau \right)
\]

Equilibrium assumed

Dynamic

\[
\dot{x}^i(t) = Ax^i(t) + B \left( u^i(t) + a_i(t) + \sum_{l=1, l \neq i}^{N_a} a^*_l(t) + d(t) \right),
\]

\[
\dot{x}^i_h(t) = Ax^i_h(t) + B \left( u^i(t) + d^h_i(t) \right),
\]

Stealthiness constraints

\[
\eta^i(t) = \|Cx^i_h(t) - Cx^i(t)\|_{L_2}
\]

Constraints

Nominal model

Attack-free model

Path

Disturbances bounded by \( \Delta \)

Boundary

\[
x^i(t_j) = x^i_h(t_j) = x_j,
\]

\[
\eta^i(t_j) = \eta^i(t_j + T) = 0.
\]
The attackers' point of view: equilibrium

Theorem

The attackers' Nash policy over $t \in [t_j, t_j + T], i \in \mathcal{N}_a$, is given by:

$$u^i_*(t) = \frac{1}{T(t_j+T-t)} \int_{t_j}^{t_j+T} u^i(t) \, dt$$

But, are attackers always rational?
Experimental evidence suggests that non-equilibrium games based on level-k thinking and cognitive hierarchy can often out-predict equilibrium (Camerer, Ho, 2004, Stahl, Wilson, 1995, Nagel, 1995).

Co-states

$$\hat{\mu}^i_h(t) = -A^T \mu^i_h(t) - C^T C (x^i_h(t) - x^i(t)) \xi^i(t),$$

$$\dot{\xi}^i(t) = 0,$$

$$\mu^i(t_j + T) = F_i^0 x^i(t_j + T), \quad \mu^i_h(t_j + T) = 0,$$
The attackers’ point of view: non-equilibrium

Level-k thinking

- A suitable framework to model boundedly rational agents (Camerer, Ho, 2004).
- A level-0 agent is assumed to follow a naïve pattern.
- A more intelligent, level-1, agent derives his best response assuming the rest are level-0.
- A more intelligent, level-2, agent assumes the rest are level-1, and so on.
- The model grows up to level-k, where it is possible that $k \rightarrow \infty$. 

Level 0  Level 1  Level 2  Level 3  ...

Π₀  Π₁  Π₂  Π₃
The attackers’ point of view: non-equilibrium

Level-k thinking

Level 0

Solves the zero-sum game

\[
\min_{u_0 \in \mathcal{F}_a} \max_{a_{i,0} \in \mathcal{F}_a} J_{i,0}^a
\]

assuming no one else attacks.

Level k

Solves the following zero-sum game, by assuming everyone is level k-1:

\[
\min_{u_k^i \in \mathcal{F}_a^i} \max_{a_{i,k} \in \mathcal{F}_a^i} J_{i,k}^a (u_k^i, a_{i,k}; t_j) = \|x_k^i(t_j + T)\|_{F_a^i} + \int_{t_j}^{t_j + T} \left( \|x_k^i(\tau)\|_{Q_a^i} + \|u_k^i(\tau)\|_{R_a^i} - \|a_{i,k}(\tau)\|_{K_{i,a}^k} \right) d\tau,
\]

subject to the dynamics, \( \forall t \in [t_j, t_j + T] \),

\[
\dot{x}_k^i(t) = Ax_k^i(t) + B(u_k^i(t) + (N_a - 1)a_{i,k-1}^a(t) + a_{i,k}(t) + d(t)), \quad x_k^i(t_j) = x_j,
\]

with the constraint

\[
\|a_{i,k}(t)\| \leq \kappa(k,j), \quad \forall t \in [t_j, t_j + T],
\]

Attacker may **overthink**! Stealthiness at risk.

Attacker’s capability at level k.
The attackers’ point of view: non-equilibrium

Minimize risk: Estimate cognitive level of other attackers.

Solved via quadratic programming.

- Probability: attacker is level i.
- Database policies.
- Distance between attack and database policies.
- Bias to initial beliefs.
Simulations

Power System

\[
\frac{d}{dt} \begin{bmatrix}
\Delta \tilde{\alpha} \\
\Delta P_m \\
\Delta f_G
\end{bmatrix} = \begin{bmatrix}
-\frac{1}{T_g} & 0 & \frac{1}{R_g T_g} \\
\frac{K_p}{T_e} & -\frac{1}{T_e} & 0 \\
0 & \frac{K_p}{T_p} & -\frac{1}{T_p}
\end{bmatrix} \begin{bmatrix}
\Delta \tilde{\alpha} \\
\Delta P_m \\
\Delta f_G
\end{bmatrix} + \begin{bmatrix}
\frac{1}{T_g} \\
0 \\
0
\end{bmatrix} \bar{u},
\]

\[
x := [\Delta \tilde{\alpha} \quad \Delta P_m \quad \Delta f_G]^T, \quad y = \Delta f_G,
\]

- 1 defender.
- 2 stealthy attackers.
- Disturbance \( d = 0.05 \sin(\pi t) \).
- Disturbance and attack bound \( \bar{a} = \Delta = 0.5 \).
- Prediction horizon of 1 [sec], control horizon of 0.2 [sec].
Simulations

**Fig. 1:** Evolution of the states and the predicted cost when the system is under infinitely rational attacks for $t \in [0, 6]$.

**Fig. 2:** Evolution of the control policies when the system is under infinitely rational attacks for $t \in [0, 6]$.

Effect of non-stealthy attack

Stability maintained
Simulations

Boundedly rational case: a level 3 & a level 1 attacker.

The level 3 attacker successfully identifies the other attacker’s level.

Cost lower than in the infinite rationality case.

Fig. 3: Evolution of the states and the predicted cost when the system is under boundedly rational attacks for $t \in [0, 6]$.

Fig. 4: Evolution of the beliefs of the first attacker and the attack policies when the system is under boundedly rational attacks for $t \in [0, 6]$. 
Learning-Based Actuator Placement for Security against Actuation Attacks
Motivation

Designing a control system involves the selection of its actuators.

• Actuators need to optimize controllability.
• The number of actuators cannot be arbitrarily large.

But!!

CPS are:
• Vulnerable to actuation attacks.
• Subject to unknown dynamics.

Solution: Learning-based actuator placement.
Problem formulation

Continuous-time system:

\[
\dot{x}(t) = Ax(t) + B(u(t) + a(t)), \quad x(0) = x_0,
\]

- \(x : \mathbb{R}_+ \rightarrow \mathbb{R}^n\) is the state.
- \(u, a : \mathbb{R}_+ \rightarrow \mathbb{R}^m\) are the control and the actuation attack.
- \(A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}\) are the plant and input matrices.

The input matrix is such that \(B = [B_0 \ B_Y]\), where:

- \(B_0 = [\beta_1, \beta_2, \ldots, \beta_{m-k}] \in \mathbb{R}^{n \times (m-k)}\) are actuators already incorporated in the system.
- \(B_Y = [v_1, v_2, \ldots, v_k] \in \mathbb{R}^{k}\) are actuators that need to be selected.
Let $\mathcal{B} = \{b_1, \ldots, b_N\}$ be a set of available actuators. Problem: choose a set of actuators $\mathcal{V} = \{v_1, v_2, \ldots, v_k\}$ so that:

\[
\max_{\mathcal{V} \subseteq \mathcal{B}} f(\mathcal{V}),
\]

s.t. $\text{card}(\mathcal{V}) = k$,

$A$ is uncertain,

- $f$ quantifies controllability & attack resilience.
- $k$ is less than $N$.
- Only an upper bound of the spectral abscissa of $A$ is known.

**Problem 1:** Find a metric $f$ that quantifies both controllability and actuation attack resilience, and which can be tractably estimated in a partially model-free manner.
Since $A$ is not known, the metric $f$ needs to be estimated.

**Problem 2:** Estimate the metric $f$ without knowledge of $A$.

Accordingly, the actuators will need to be placed adaptively:

**Problem 3:** Let $t_j \in \mathbb{R}_+$, $j \in \mathbb{N}$, be time instants such that $t_{j+1} - t_j > 0$, $\forall j \in \mathbb{N}$, $t_0 = 0$ and $\lim_{j \to \infty} t_j = \infty$. Design an actuator-scheduling procedure which will place actuators at each time instant $t_j$, $j \in \mathbb{N}$, while guaranteeing that $\mathcal{V}_j$ eventually converges to the optimal set of actuators.

As a result of Problem 3, the closed-loop system is:

$$\ddot{x}(t) = A\dot{x}(t) + \bar{B}(t)(u(t) + a(t)), \ x(0) = x_0, \ \forall t \geq 0,$$

where $\dot{x} : \mathbb{R}_+ \to \mathbb{R}^n$ are the new state trajectories, and

$$\bar{B}(t) = [B_0 \ B_{\mathcal{V}_j}], \ \forall t \in [t_j, \ t_{j+1}), \ j \in \mathbb{N}.$$
Actuator evaluation metric

Solve a constrained zero-sum game:

\[
\min_u \max_a \ J(u, a; t_f) = \frac{1}{2} \int_0^{t_f} \left( u^T(\tau) Ru(\tau) - a^T(\tau) K a(\tau) \right) \, d\tau
\]

s.t.

\[
\dot{x}(t) = Ax(t) + B(u(t) + a(t)),
\]

\[
x(0) = x_0, \quad x(t_f) = 0,
\]

\[
u, \ a : [0, \ t_f] \to \mathbb{R}^m,
\]

- A defender wants to regulate the CPS with minimum energy.
- An attacker wants to disrupt the regulation.
- \( K = K(V) = \text{diag}_v([k_1 \ k_2 \ \ldots \ k_m \ k_{v_1} \ k_{v_2} \ \ldots \ k_{v_k}]^T) \succ 0. \)
- \( R = R(V) = \text{diag}_v([r_{\beta_1} \ r_{\beta_2} \ \ldots \ r_{\beta_m} \ r_{v_1} \ r_{v_2} \ \ldots \ r_{v_k}]^T) \succ 0. \)
Actuator evaluation metric

Theorem: Let $(A, B)$ be a controllable pair, and $K \succ R$. Then, the zero-sum game admits a saddle-point solution $(u^*, a^*)$, for all $x_0 \in \mathbb{R}^n$, with value

$$J^*(t_f) = J(u^*, a^*; t_f) = \frac{1}{2} x_0^T e^{A^T t_f} Q^{-1}(t_f) e^{A t_f} x_0,$$

where $Q(t_f)$ is the robust controllability Gramian:

$$Q(t_f) = \int_0^{t_f} e^{A \tau} B (R^{-1} - K^{-1}) B^T e^{A^T \tau} d\tau.$$
Actuator evaluation metric

In the specific case that the state matrix $A$ is Hurwitz and $t_f = \infty$:

$$AQ + QA^T + B(R^{-1} - K^{-1})B^T = 0, \quad Q = Q(\infty).$$

If $A$ is not Hurwitz, then define the discounted Gramian:

$$Q_\gamma = \int_0^\infty e^{-2\gamma \tau} e^{A\tau} B(R^{-1} - K^{-1})B^T e^{A^T\tau} \, d\tau$$

$$= \int_0^\infty e^{(A-\gamma I)\tau} B(R^{-1} - K^{-1})B^T e^{(A-\gamma I)^T\tau} \, d\tau, $$

**Lemma:** Given $\gamma > \alpha(A)$, the Gramian $Q_\gamma$ is well defined and uniquely solves the Lyapunov equation

$$(A - \gamma I) Q_\gamma + Q_\gamma (A - \gamma I)^T + B(R^{-1} - K^{-1})B^T = 0.$$
Actuator evaluation metric

To minimize the average robust controllability, choose:

\[ f(\mathcal{V}) = \text{tr}(Q_\gamma). \]

**Lemma:** Let \( \gamma > \alpha(A) \). Then

\[ \text{tr}(Q_\gamma) = \text{tr}\left( (R^{-1} - K^{-1})B^T P_\gamma B \right), \]

where \( P_\gamma \) is the unique solution of the dual Lyapunov equation

\[ (A - \gamma I)^TP_\gamma + P_\gamma (A - \gamma I) + I = 0. \]

Only one Lyapunov equation need be solved to evaluate \( f \)!
Second advantage: $f$ can be optimized in polynomial time!

$$\text{tr}((R^{-1} - K^{-1})B^TP_\gamma B) = \sum_{v \in \mathcal{V}} (r_v^{-1} - k_v^{-1})v^TP_\gamma v + \sum_{i=1}^{m-k} (r_{\beta_i}^{-1} - k_{\beta_i}^{-1})\beta_i^TP_\gamma \beta_i.$$ 

This decoupling of the effect of the actuators in $f$, aids at significantly reducing computational complexity.

**Reformulated problem:** Given that the state matrix $A$ is uncertain, find the set of actuators $\mathcal{V}^*$ that optimally solves the optimization problem

$$\max_{\mathcal{V} \subseteq \mathcal{B}} f(\mathcal{V}) = \sum_{v \in \mathcal{V}} (r_v^{-1} - k_v^{-1})v^TP_\gamma v,$$

s.t. $\text{card}(\mathcal{V}) = k$. 
Learning-based estimation of $f$

The metric $f$ can be described in a data-based fashion, given *persistence of excitation*.

**Definition:** A signal $\phi : [t_0, \infty) \to \mathbb{R}^q$, $t_0 \geq 0$, is persistently exciting if there exist constants $\gamma_1, \gamma_2, T_f > 0$ such that

$$\gamma_1 I \leq \int_{t}^{t+T_f} \phi(\tau)\phi^T(\tau) d\tau \leq \gamma_2 I, \quad \forall t \geq t_0.$$
Learning-based estimation of $f$

**Theorem:** Consider the state trajectories $\bar{x}$ in the absence of attacks, $\forall t \geq 0$, and let $\gamma > a(A)$. Then, the data-based, time-dependent equation

$$\Psi^T(t)\text{vech} (P_\gamma) + \int_{t-T}^{t} \|\bar{x}(\tau)\|^2 d\tau = 0, \forall t \geq T,$$

where $T > 0$, and $\Psi(t) \in \mathbb{R}^{n(n+1)/2}$ is the regression vector

$$\Psi(t) = \text{vech} \left( W(t) + W^T(t) - \text{diagm}(W(t)) \right),$$

$$W(t) = \bar{x}^T(t) \otimes \bar{x}(t) - \bar{x}^T(t-T) \otimes \bar{x}(t-T)$$

$$- \int_{t-T}^{t} (2\gamma \bar{x}^T(\tau) \otimes \bar{x}(\tau) + 2\bar{x}^T(\tau) \otimes (\bar{B}(\tau)u(\tau))) d\tau,$$

admits a constant solution with respect to $P_\gamma$, which satisfies the model-based Lyapunov equation. In addition, if $\Psi$ is persistently exciting, this solution is unique.
Learning-based estimation of $f$

Estimate $P_\gamma$ with $\hat{P}_\gamma$ by minimizing the error

$$E(t) = \frac{1}{2} e^2(t),$$

where $e(t)$ is the databased LE:

$$e(t) = \Psi^T(t) \text{vech} \left( \hat{P}_\gamma \right) + \int_{t-T}^{t} \| \tilde{x}(\tau) \|^2 \, d\tau, \ \forall t \geq T.$$  

Gradient descent learning law:

$$\text{vech}(\dot{\hat{P}}_\gamma) = -\beta \frac{\Psi(t)}{1 + \| \Psi(t) \|^2} \left( \Psi^T(t) \text{vech} \left( \hat{P}_\gamma \right) + \int_{t-T}^{t} \| \tilde{x}(\tau) \|^2 \, d\tau \right).$$
Define the estimation error $\tilde{P}_\gamma = P_\gamma - \hat{P}_\gamma$.

**Theorem:** The gradient descent learning law guarantees that:

1. The norm of the estimation error vector $\left\|\text{vech}(\tilde{P}_\gamma)\right\|$ is non-increasing with respect to time.

2. Given that $\bar{\Psi}(\cdot) := \frac{\Psi(\cdot)}{\sqrt{1 + \left\|\Psi(\cdot)\right\|^2}}$ is persistently exciting, the norm of the estimation error vector $\left\|\text{vech}(\tilde{P}_\gamma)\right\|$ decays to zero exponentially fast.
Having an estimate of $P_{\gamma}$, we can solve the approximate optimization:

$$\max_{V \subseteq \mathcal{B}} \hat{f}(V; t_j) = \sum_{v \in V} (r_v^{-1} - k_v^{-1})v^T \hat{P}_\gamma(t_j)v,$$

s.t. card($V$) = $k$,

where $\hat{f}(\cdot; t_j) : 2^B \rightarrow \mathbb{R}$ is an approximation of $f(\cdot)$ at $t = t_j$.

**Complexity:** $\mathcal{O}(N\log N)$. 
Online actuator placement

Algorithm 1: Data-Based Actuator Placement

Input: Constants $k_{bi} > r_{bi} > 0$, $i = 1, \ldots, N$, and $\gamma > \alpha(A)$.
Output: Actuator Selection Sequence $\{\mathcal{V}_j\}_{j \in \mathbb{N}}$.

1: procedure
2: for $t \geq 0$ do
3: Tune $\hat{P}_\gamma$ according to learning law (25).
4: if $t = t_j$, $j \in \mathbb{N}$, then
5: Sort the values $(r_{bi}^{-1} - k_{bi}^{-1}) b_i^T \hat{P}_\gamma(t_j) b_i$, $\forall i = 1, \ldots, N$.
6: Choose the set of actuators $\mathcal{V}_j \subseteq \mathcal{B}$ corresponding the $k$ largest such values.
7: Set $\bar{B}(t) = [B_0 \mid B_{\mathcal{V}_j}]$, $\forall t \in [t_j, t_{j+1})$.
8: end if
9: end for
10: end procedure
Theorem: Let $\mathcal{V}^*$ be an optimal solution to the actual problem, and $\mathcal{V}^*$ an optimal solution to the perturbed problem. Then:

- The perturbed admits a unique optimal solution with probability 1; and
- it holds that $|f(\mathcal{V}^*) - f(\mathcal{V}^*)| \leq k\bar{\eta}$.

\[
\max_{\mathcal{V} \subseteq \mathcal{B}} f_p(\mathcal{V}) = \sum_{v \in \mathcal{V}} \left( (r^{-1}_v - k_v^{-1})v^T P_{\gamma} v + \eta_v \right),
\]

s.t. $\text{card}(\mathcal{V}) = k$,

where $\eta_i, i = 1, \ldots, N$, are independent random variables, each following a uniform distribution over the interval $[0, \bar{\eta}]$, for some $\bar{\eta} > 0$. 
Partially model-free attack detection

In the steady state of the learning law, one can detect actuation attacks in a partially model-free manner.

Consider the filter:

\[
\text{vech}(\hat{\mathbf{P}}_{\gamma,s}) = -\beta \frac{\Psi_s(t)}{1 + \|\Psi_s(t)\|^2} \left( \Psi_s^T(t) \text{vech} \left( \hat{\mathbf{P}}_{\gamma,s} \right) \right) + \int_{t-T}^{t} \|x(\tau)\|^2 \, d\tau - k_d(\text{vech}(\hat{\mathbf{P}}_{\gamma,s} - \hat{\mathbf{P}}_{\gamma,s}(T))),
\]

\[
\text{vech}(\hat{\mathbf{P}}_{\gamma,s}(T)) = \text{vech}(\mathbf{P}_{\gamma}), \quad t \geq T,
\]

with \(k_d > 0\), and \(\Psi_s\) being \(\Psi\) at the steady state.

**Theorem:** An attacker can remain undetected only if:

- C1: \(\int_{t-T}^{t} x^T(\tau) \mathbf{P}_{\gamma} B a(\tau) \, d\tau = 0\), or
- C2: \(\Psi_s(t) = 0\).

**Lemma:** Let \(d(t) := \left\| \hat{\mathbf{P}}_{\gamma,s}(t) - \hat{\mathbf{P}}_{\gamma,s}(T) \right\|_F\), \(\forall t \geq T\). If \(d(t) \neq 0\) for some \(t \geq T\), then \(a(t) \neq 0\).
Simulations

We consider the Innovative Control Effectors (ICE) eight-state aircraft, flying at an altitude of 15,000 ft.

- $N = 11$ available actuators.
- $k = 4$ actuators picked.
- The learning parameters are $\beta = 100$, $T = 0.05$.
- The optimal set of actuators is $\mathcal{V}^* = \{b_1, b_3, b_7, b_{10}\}$
- The actuators are switched every 20 seconds.

For this algorithm, we choose the control weighting terms as $r_{b_i} = 1$, for all $i = 1, \ldots, N$. In addition, the attack weighting terms are chosen as $k_{b_1} = k_{b_3} = k_{b_5} = k_{b_6} = k_{b_7} = 10$, $k_{b_4} = k_{b_8} = k_{b_{10}} = 2$ and $k_{b_2} = k_{b_{11}} = 1.01$. 
The matrix $P_\gamma$ is successfully identified!
The optimal set of actuators is found after 300 s!
After 500 s, the detection is employed with $k_d = 2$. An attack takes place over $t \in [535, 565]$ s.

The attack is successfully detected.
Conclusion & future ideas

- Constructed an equilibrium and a non-equilibrium based decision making mechanism for stealthy attackers.
- Developed a level-k thinking model for the attackers, along with a level estimator.
- Constructed a metric that evaluated both controllability and attack resilience.
- Estimated this metric in a partially model-free manner.
- Designed a learning-based actuator-placement algorithm, with optimality guarantees.
- Constructed a partially model-free attack detection scheme.

Future work

- Extension to cases of joint sensor-actuator attacks.
- Consideration of cases where more statistics of the disturbance are available.
- Implementation in a networked control setting with decentralized information.
- Development of resilient RHC to deal with possible DoS attacks.
- Extension to sensor placement.
- Extension to completely unknown systems.
QUESTIONS?

THANK YOU

For papers please see: kyriakos.ae.gatech.edu/

Filippos Fotiadis