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

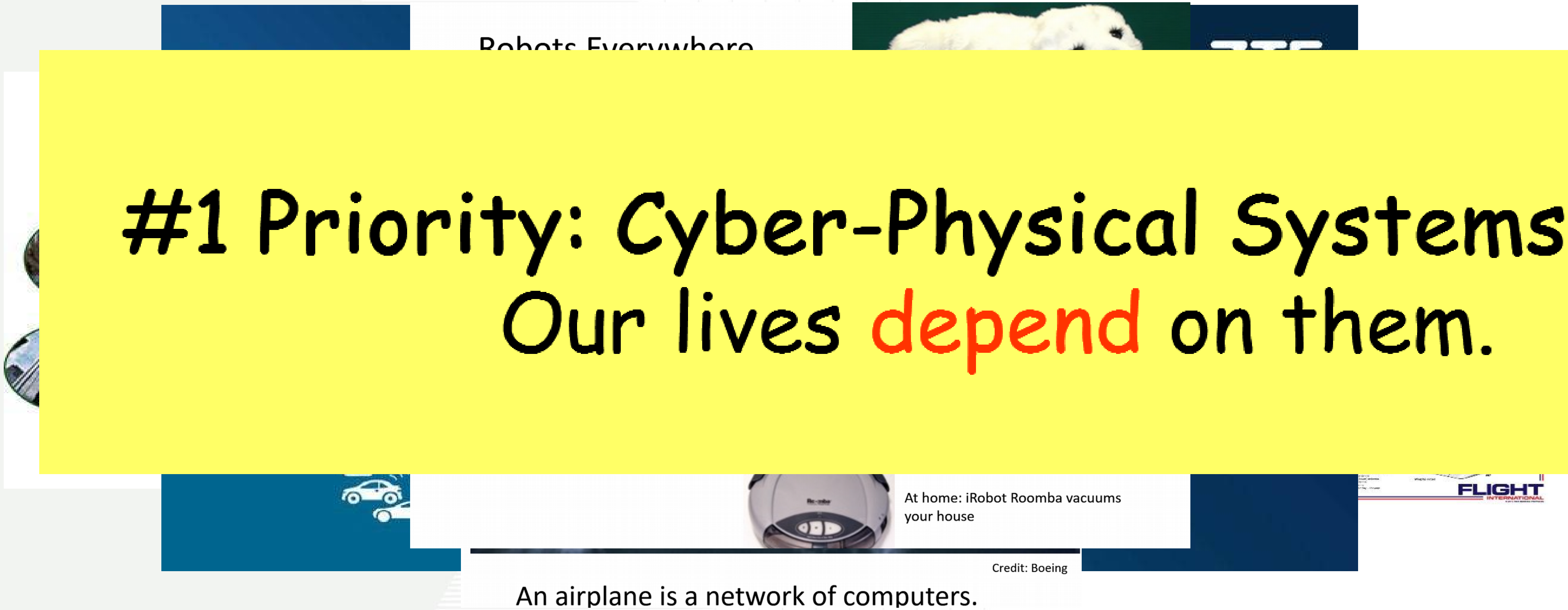
Center for the Development and Application of Internet of Things Technologies

CREATING THE NEXT®

Outline

- 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**

Cyber-physical systems



#1 Priority: Cyber-Physical Systems

Our lives depend on them.

Cyber-physical atta…

German Steel Mill Meltdown: Rising

Cyber-physical attacks:
A reality we need to face.

Stuxnet w…
assets'

By Jonathan Fildes
Technology reporter, …

23 September 2010

One of the most s… pieces of malware eve… was probably targetin… value" infrastructure … experts have told the…

Stuxnet's complexity su… only have been written … state", some researche… claimed.

It is believed to be the f… worm designed to target real-world infrastruc… plants and industrial units.

…ly for

… experts say

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

While the FDA noted that some company's device… the interception of patient data, there have been n… they are working to patch the vulnerabilities.

…full extent of the damage are still unknow…

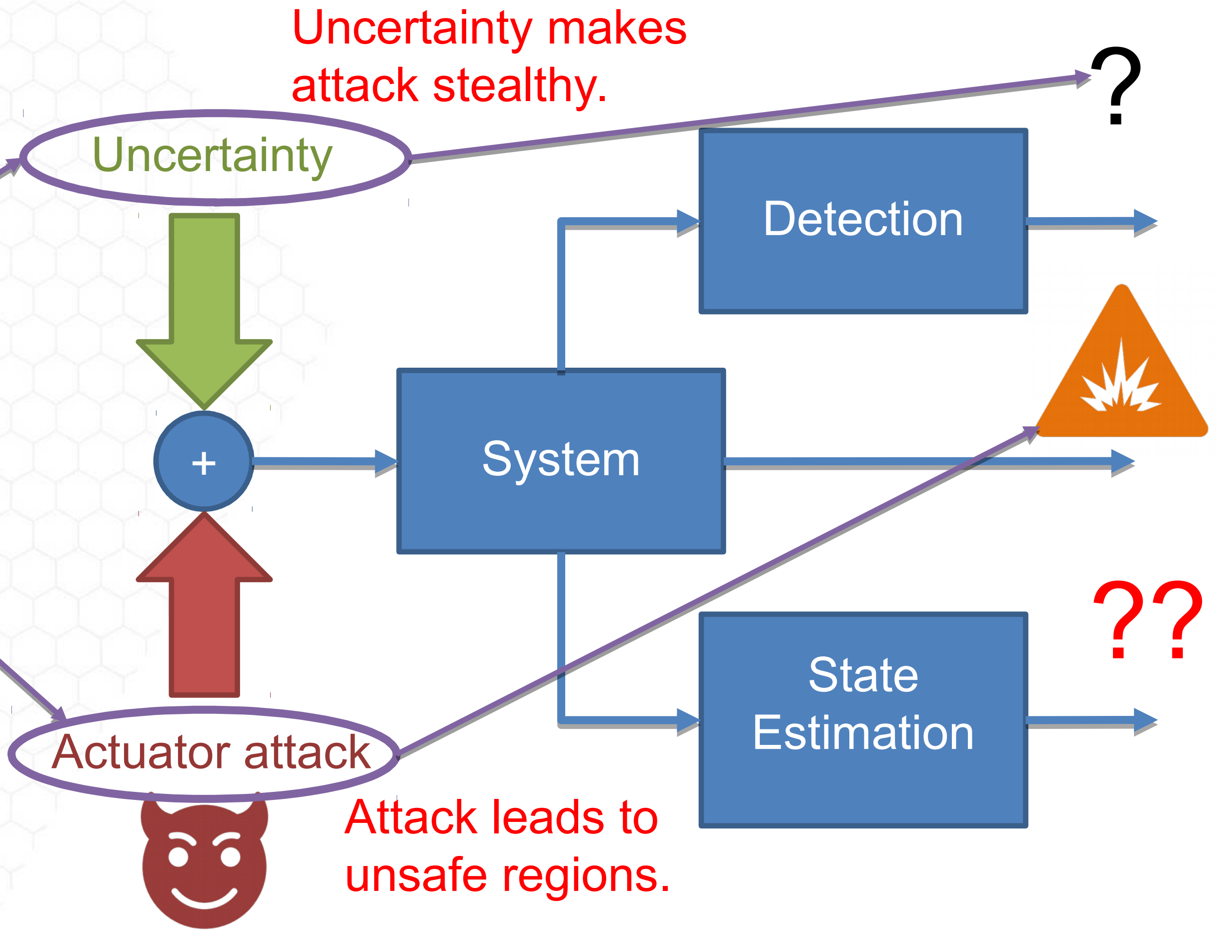
Ukraine's energy grid has been attac…

A power cut that hit part of th… judged a cyber-attack by resea… investigating the incident.

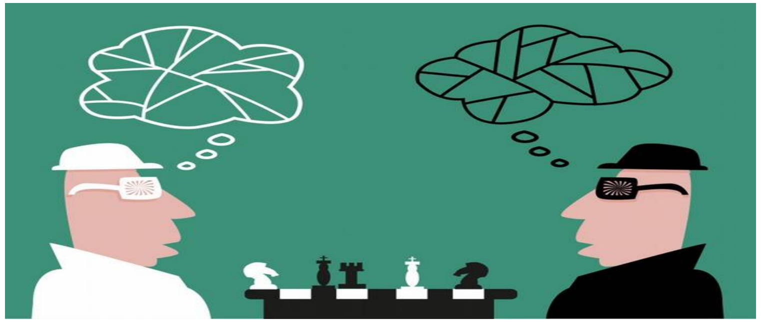
Receding Horizon Control for Security against Stealthy Actuation Attacks

Types of attacks

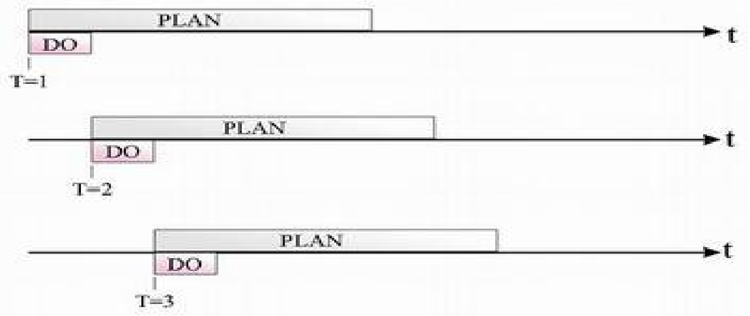
- **Sensor Attacks.**
  -Injection of faulty sensor measurements.
- **Actuator attacks.**
  -Injection of faulty control inputs.
- **Denial-of-Service attacks.**
  -Jamming of sensors/actuators.
- **Stealthy attacks.**
  -Attacks that cannot be detected.



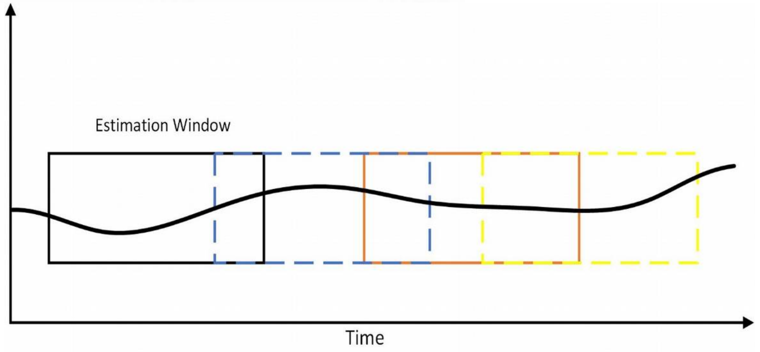
How to mitigate?



- **Game Theory.**
  - Secure decision making by considering worst-case scenaria.
  - 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

Defender's Input

Plant Dynamics

Attackers' Inputs

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

$$y(t) = Cx(t), \; x(t_0) = x_0.$$

Output (known by everyone)

Disturbance (known by the attackers)

Attack-Free Dynamics

$$\dot{x}_h(t) = Ax_h(t) + B\left(u(t) + d_h(t)\right),$$
$$y_h(t) = Cx_h(t), \; x_h(t_0) = x_0.$$

Assumptions: $\|d(t)\| < \Delta.$ $\|a_l(t)\| < \bar{a}, \; l \in \mathcal{N}_a.$

Stealthiness: $y = y_h.$

Attackers' 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

Georgia Tech

- 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.

Georgia Tech

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:

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

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

Incompatible

BUT!!

RHC optimizes the future,

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

Solution

Reverse the past model!

Reversed past model

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

Theorem

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

$$Cx_p^d(t) = y(2t_j - t),\ \forall t \in [t_j,\ t_j + T].$$

State estimation and game-theoretic RHC

Georgia Tech

Ideally, $T \to \infty$

Penalizes attacks with unlikely high values.

Needed for stability

Game-theoretic RHC

$$\min_{u_i} \max_{\mathcal{A}_i} J^d(t_j) = \int_{t_j}^{t_j+T} \left(\|x^d(\tau)\|_{Q^d} + \|u(\tau)\|_{R^d} + \sum_{l=1}^{N_a} \|a_l^d(\tau)\|_{K_l^d, \bar{a}} \right) \mathrm{d}\tau + \|x^d(t_j+T)\|_{F^d}.$$

$\{a_1^i, \ldots, a_{N_a}^i, a_{p,1}^i, \ldots, a_{p,N_a}^i, d^i, d_p^i, d_h^i\}$

Constraints

Dynamic

Future

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

Past

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

Future (attack free)

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

Path

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

Output Compatibility

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

Worst-case state estimation

Disturbances bounded by Δ

Boundary

$$x^d(t_0) = x_h^d(t_0) = x_p^d(t_0).$$

State estimation and game-theoretic RHC

Georgia Tech

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

Solution: Concavification of the cost

Relaxed game

$$\min_{u\in\mathcal{F}_u^j}\ \max_{\mathcal{A}\in\mathcal{F}_\mathcal{A}^j}\ \tilde{J}^d(u,\ \mathcal{A};\ t_j) = \int_{t_j}^{t_j+T}\Bigg(\|x^d(\tau)\|_{Q^d}+\|u(\tau)\|_{R^d}-\sum_{l=1}^{N_a}\|a_l^d(\tau)\|_{K_l^d,\bar{a}}$$
$$-\sum_{l=1}^{N_a}\|a_{p,l}^d(\tau)\|_{K_l^d,\bar{a}}-\|d^d(\tau)\|_{D^d,\Delta}-\|d_p^d(\tau)\|_{D^d,\Delta}-\|d_h^d(\tau)\|_{D^d,\Delta}\Bigg)\mathrm{d}\tau+\|x^d(t_j+T)\|_{F^d},$$

Approaches original game as induced term weights $\rightarrow 0$.

Induced concave terms

State estimation and game-theoretic RHC

Georgia Tech

Enforce path constraints by demanding:

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

where:

$$\epsilon_p^d(t) = \int_{t_j}^{t} \left\| C x_p^d(\tau) - y_p(\tau) \right\|_I \mathrm{d}\tau$$

$$\epsilon_h^d(t) = \int_{t_j}^{t} \left\| C x_h^d(\tau) - C x^d(\tau) \right\|_I \mathrm{d}\tau$$

Costates from min-max optimization

Nash controller

Theorem

The Nash defender's policy is given, for all, $i = 1, \dots, N_a$ by

$$u^\star(t) = -{R^d}^{-1} B^{\mathrm{T}} \left(\lambda(t) + \lambda_h(t)\right),$$

$$a_l^{d\star}(t) = \bar{a} \times \tanh\left({K_l^d}^{-1} B^{\mathrm{T}} \lambda(t)\right), \ \forall l \in \mathcal{N}_a,$$

$$d^{d\star}(t) = \Delta \times \tanh\left({D^d}^{-1} B^{\mathrm{T}} \lambda(t)\right),$$

$$d_h^{d\star}(t) = \Delta \times \tanh\left({D^d}^{-1} B^{\mathrm{T}} \lambda_h(t)\right),$$

$$a_{p,l}^{d\star}(t) = -\bar{a} \times \tanh\left({K_l^d}^{-1} B^{\mathrm{T}} \lambda_p(t)\right), \ \forall l \in \mathcal{N}_a,$$

$$d_p^{d\star}(t) = -\Delta \times \tanh\left({D^d}^{-1} B^{\mathrm{T}} \lambda_p(t)\right),$$

$$\dot{\lambda}(t) = -A^{\mathrm{T}}\lambda(t) - Q^d x^d(t) - C^{\mathrm{T}} C \left(x^d(t) - x_h^d(t)\right) \rho_h(t),$$

$$\dot{\lambda}_p(t) = A^{\mathrm{T}}\lambda_p - C^{\mathrm{T}} \left(C x_p^d(t) - y(2t_j - t)\right) \rho_p(t),$$

$$\dot{\lambda}_h(t) = -A^{\mathrm{T}}\lambda_h(t) - C^{\mathrm{T}} C \left(x_h^d(t) - x^d(t)\right) \rho_h(t),$$

$$\dot{\rho}_p(t) = \dot{\rho}_h(t) = 0,$$

$$\lambda_p(t_j+T) = \lambda_h(t_j+T) = 0,$$

$$\lambda(t_j+T) = F^d x^d(t_j+T),$$

$$\lambda(t_j) + \lambda_h(t_j) + \lambda_p(t_j) = 0,$$

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

$$\frac{\mathrm{d}F(x^d(t))}{\mathrm{d}t}\Big|_{u=\psi(x^d)} \leq -\|x^d(t)\|_{Q^d} - \|\psi(x^d(t))\|_{R^d} + \sum_{l=1}^{N_a} \|a_l^d(t)\|_{K_l^d,\bar{a}} + \|d^d(t)\|_{D^d,\Delta}\,.$$

Terminal cost is an ISS Lyapunov

Theorem:

Closed trajectories are bounded for sufficiently small weighting matrices on maximizers.

Interpretation: Do not underestimate the attackers!

State estimation and game-theoretic RHC

What if the past output is available only intermittently?
Assume the output is available every δ seconds.

Theorem

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

$$\operatorname{rank}\left(\begin{bmatrix} C^{\mathrm{T}} & (Ce^{-\delta A})^{\mathrm{T}} \dots & (Ce^{-N\delta A})^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}\right) = n.$$

Strengthened observability condition for the discretized continuous dynamics.

But… what could the attackers do?

The attackers' point of view: equilibrium

Stealthiness

Theorem

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

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

where

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

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

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

Constrained RHC allows to impose this.

Knowledge of disturbance necessary for stealthiness.

Not equal to actual defender input!

The attackers' point of view: equilibrium

Concavification term

Cooperative game

$$\min_{u^i\in\mathcal{F}_u^j}\ \max_{a_i\in\mathcal{F}_a^j,\ d_h^i\in\mathcal{F}_d^j} \tilde{J}_i^a\left(u^i, a_i, d_h^i;\ t_j\right) = \int_{t_j}^{t_j+T} \left(\left\|x^i(\tau)\right\|_{Q_i^a} + \left\|u^i(\tau)\right\|_{R_i^a} - \left\|a_i(\tau)\right\|_{K_i^a,\bar{a}} - \left\|d_h^i(\tau)\right\|_{D_i^a,\Delta}\right) \mathrm{d}\tau$$

Equilibrium assumed

Constraints

Dynamic

Nominal model

$$\dot{x}^i(t) = Ax^i(t) + B\Big(u^i(t) + a_i(t) + \sum_{\substack{l=1,\\ l\neq i}}^{N_a} a_l^\star(t) + d(t)\Big),$$

Attack-free model

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

$$\eta^i(t) = \left\|Cx_h^i(t) - Cx^i(t)\right\|_{I_p},$$

Stealthiness constraints

Path

Disturbances bounded by Δ

Boundary

$$x^i(t_j) = x_h^i(t_j) = x_j,$$

$$\eta^i(t_j) = \eta^i(t_j + T) = 0.$$

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:

$u^{i\star}(t$ ...

$a_i^\star(t$ ...

$d_h^{i\star}(t$ ...

$\dot{\mu}^i(t$ ...

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).

$$\dot{\mu}_h^i(t) = -A^{\mathrm{T}}\mu_h^i(t) - C^{\mathrm{T}}C\left(x_h^i(t) - x^i(t)\right)\xi^i(t),$$

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

$$\mu^i(t_j + T) = F_i^a x^i(t_j + T),\ \mu_h^i(t_j + T) = 0,$$

Co-states

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 $k \to \infty$.

Georgia Tech

Level-k thinking

Level 0

Solves the **zero-sum** game $\min_{u_0^i \in \mathcal{F}_u^j} \max_{a_{i,0} \in \mathcal{F}_a^j} J_{i,0}^a$, assuming no one else attacks.

Level k

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

$$\min_{u_k^i \in \mathcal{F}_u^j} \max_{a_{i,k} \in \mathcal{F}_a^j} J_{i,k}^a \left(u_k^i, a_{i,k};\ t_j\right) = \left\|x_k^i(t_j+T)\right\|_{F_i^a} + \int_{t_j}^{t_j+T} \left(\left\|x_k^i(\tau)\right\|_{Q_i^a} + \left\|u_k^i(\tau)\right\|_{R_i^a} - \left\|a_{i,k}(\tau)\right\|_{K_i^a,\bar{a}}\right) \mathrm{d}\tau,$$

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

$$\dot{x}_k^i(t) = Ax_k^i(t) + B\big(u_k^i(t) + (N_a - 1)a_{i,k-1}^\star(t) + a_{i,k}(t) + d(t)\big),\ x_k^i(t_j) = x_j,$$

with the constraint

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

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.

ias matrix

Solved via quadratic programming.

Probability: attacker is level i.

Database policies.

Distance between attack and database policies.

Bias to initial beliefs.

Simulations

Power System

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\Delta\tilde{\alpha}\\ \Delta P_m\\ \Delta f_G\end{bmatrix} = \begin{bmatrix}-\frac{1}{T_g} & 0 & \frac{1}{R_g T_g}\\ \frac{K_t}{T_t} & -\frac{1}{T_t} & 0\\ 0 & \frac{K_p}{T_p} & -\frac{1}{T_p}\end{bmatrix}\begin{bmatrix}\Delta\tilde{\alpha}\\ \Delta P_m\\ \Delta f_G\end{bmatrix} + \begin{bmatrix}\frac{1}{T_g}\\ 0\\ 0\end{bmatrix}\bar{u},$$

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

- 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

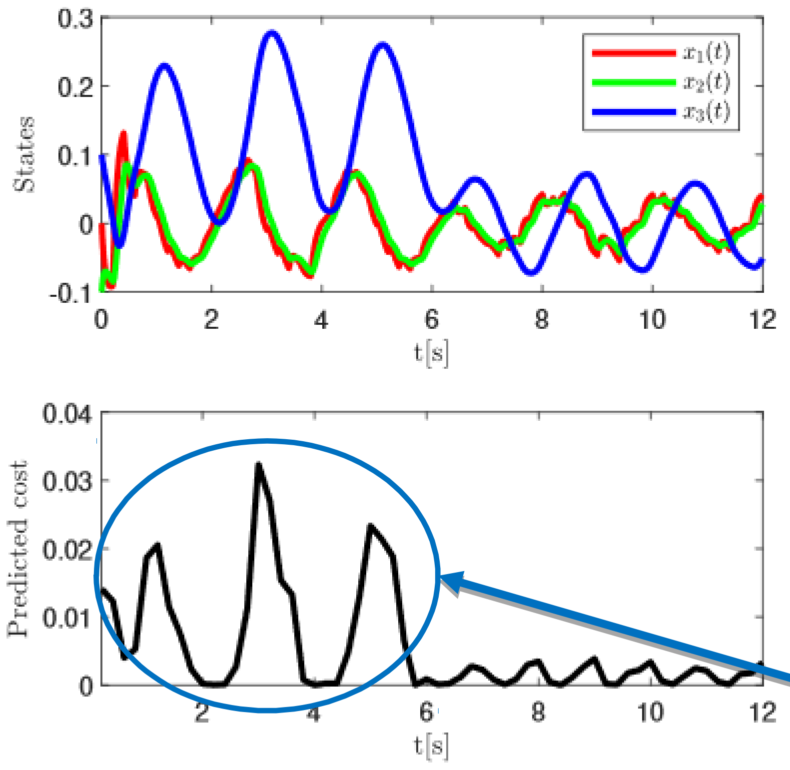


Fig. 1: Evolution of the states and the predicted cost when the system is under infinitely rational attacks for $t \in [0,\ 6]$.

Stability maintained

Effect of non-stealthy attack

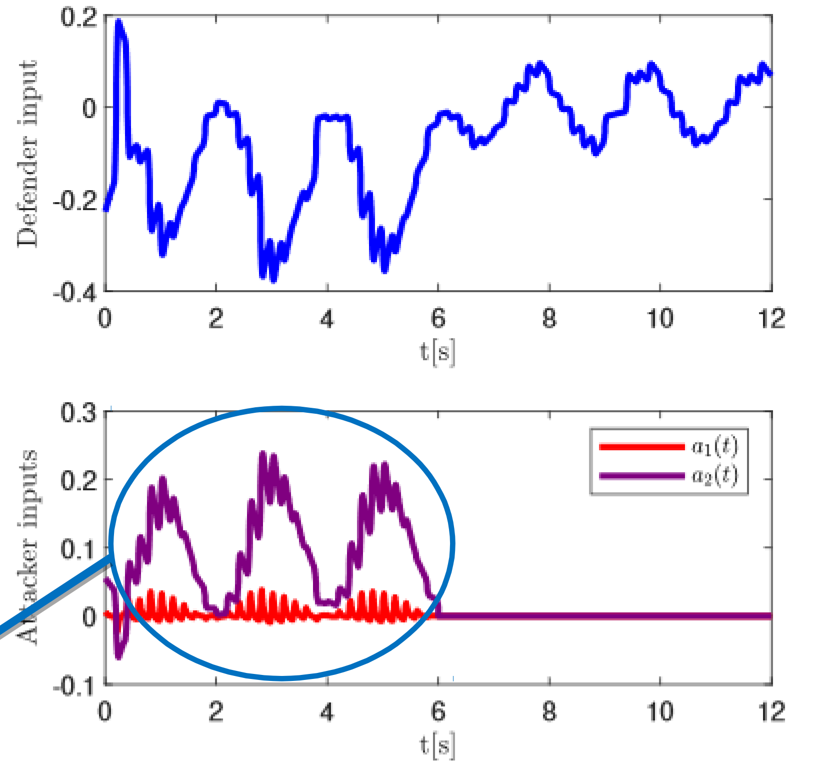


Fig. 2: Evolution of the control policies when the system is under infinitely rational attacks for $t \in [0,\ 6]$.

Simulations

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

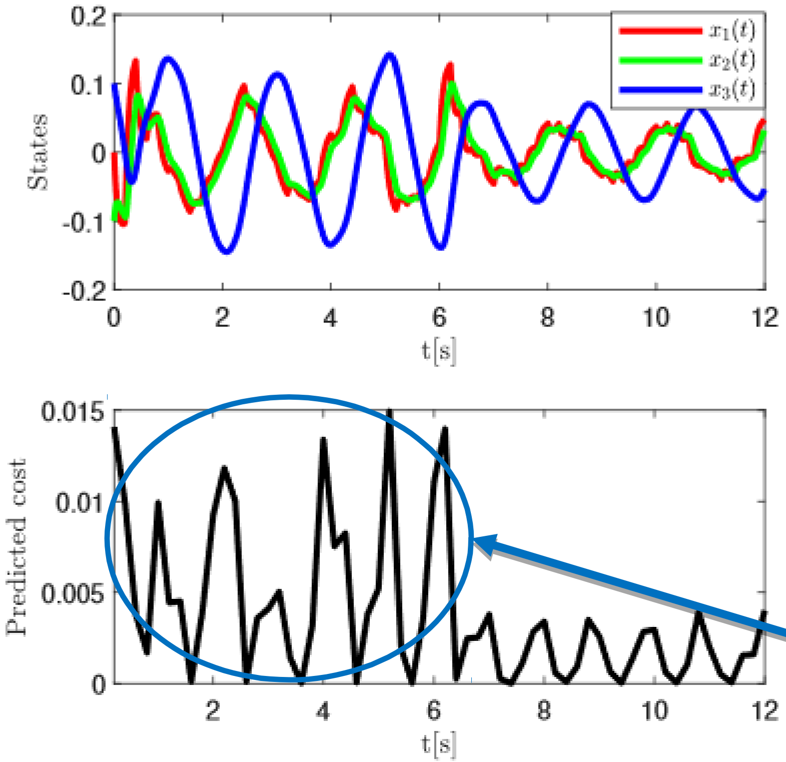


Fig. 3: Evolution of the states and the predicted cost when the system is under boundedly rational attacks for $t \in [0,\ 6]$.

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

Cost lower than in the infinite rationality case.

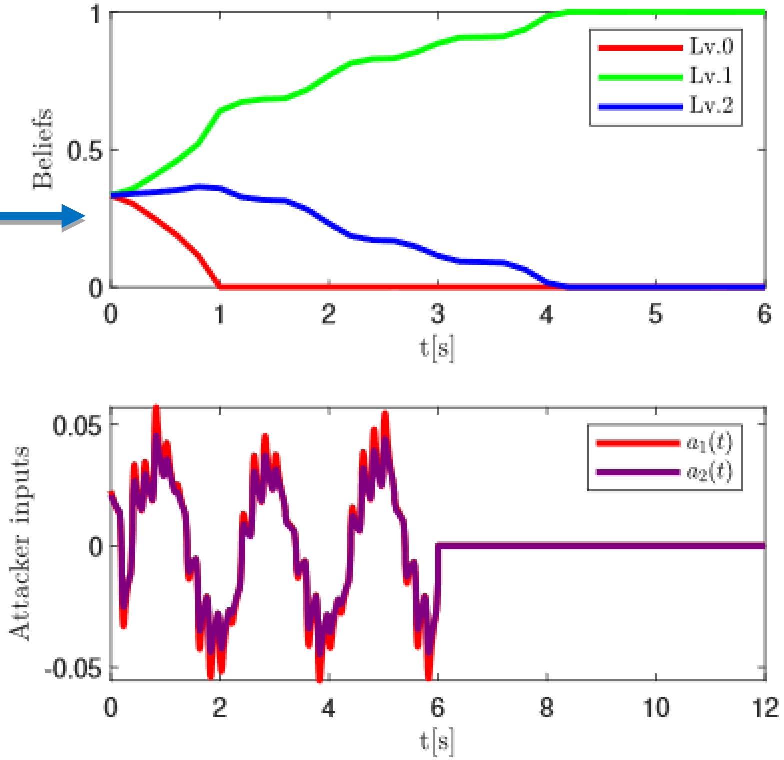


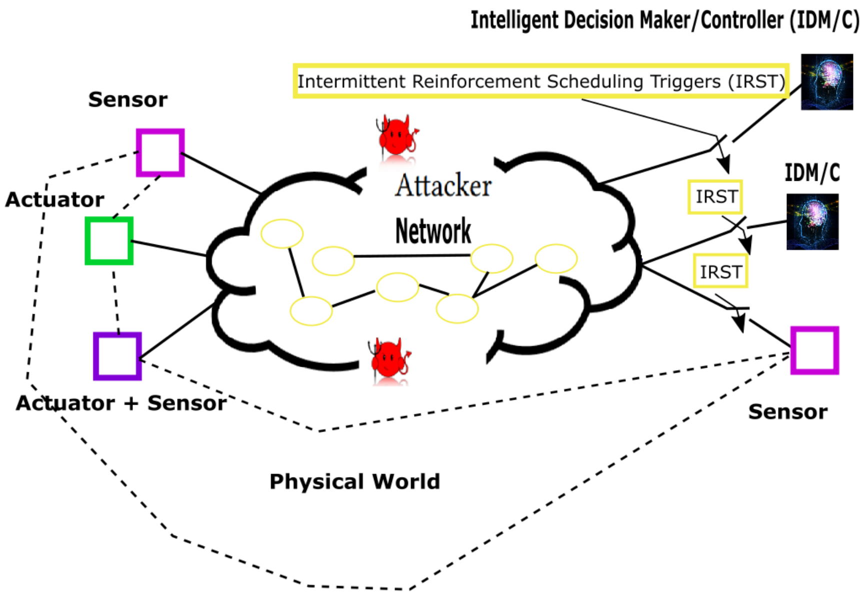
Fig. 4: Evolution of the beliefs of the first attacker and the attack policies when the system is under boundedly rational attacks for $t \in [0,\ 6]$.

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:

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

- $x : \mathbb{R}_+ \to \mathbb{R}^n$ is the state.
- $u,\ a : \mathbb{R}_+ \to \mathbb{R}^m$ 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 $B = [B_0\ B_{\mathcal{V}}]$, 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.
- $B_{\mathcal{V}} = [v_1,\ v_2,\ \ldots,\ v_k] \in \mathbb{R}^k$ are actuators that need to be selected.

Problem formulation

- 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:

$$\begin{aligned} \max_{\mathcal{V}\subseteq\mathcal{B}} \quad & f(\mathcal{V}), \\ \text{s.t.} \quad & \text{card}(\mathcal{V}) = k, \\ & A \text{ is uncertain}, \end{aligned}$$

- *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.

Problem formulation

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:

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

where $\bar{x} : \mathbb{R}_+ \to \mathbb{R}^n$ 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:

$$\begin{aligned} \min_{u} \max_{a} \quad & J(u,a;t_f) = \frac{1}{2}\int_0^{t_f} \big(u^{\mathrm{T}}(\tau)Ru(\tau) - a^{\mathrm{T}}(\tau)Ka(\tau)\big)\,\mathrm{d}\tau \\ \text{s.t.} \quad & \dot{x}(t) = Ax(t) + B(u(t)+a(t)), \\ & x(0) = x_0,\ x(t_f) = 0, \\ & u,\ a : [0,\ t_f] \to \mathbb{R}^m, \end{aligned}$$

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

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 $(u^\star,\ a^\star)$, for all $x_0 \in \mathbb{R}^n$, with value

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

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

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

Actuator evaluation metric

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

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

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

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

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

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

Actuator evaluation metric

To minimize the average robust controllability, choose:

$$f(\mathcal{V}) = \mathrm{tr}(Q_\gamma).$$

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

$$\mathrm{tr}(Q_\gamma) = \mathrm{tr}\big((R^{-1} - K^{-1})B^{\mathrm{T}}P_\gamma B\big),$$

where P_γ is the unique solution of the dual Lyapunov equation

$$(A - \gamma I)^{\mathrm{T}}P_\gamma + P_\gamma(A - \gamma I) + I = 0.$$

Only one Lyapunov equation need be solved to evaluate f!

Actuator evaluation metric

Second advantage: f can be optimized in polynomial time!

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

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 $\mathcal{V}^\star$ that optimally solves the optimization problem

$$\begin{aligned} \max_{\mathcal{V}\subseteq\mathcal{B}} \quad & f(\mathcal{V}) = \sum_{v\in\mathcal{V}} (r_v^{-1} - k_v^{-1})v^{\mathrm{T}}P_\gamma v, \\ \text{s.t.} \quad & \mathrm{card}(\mathcal{V}) = k. \end{aligned}$$

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

$$\gamma_1 I \leq \int_t^{t+T_f} \phi(\tau)\phi^{\mathrm{T}}(\tau)\mathrm{d}\tau \leq \gamma_2 I, \quad \forall t \geq t_0.$$

Learning-based estimation of *f*

Georgia Tech

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

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

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

$$\begin{aligned}\Psi(t) &= \mathrm{vech}\left(W(t) + W^{\mathrm{T}}(t) - \mathrm{diagm}(W(t))\right), \\ W(t) &= \bar{x}^{\mathrm{T}}(t) \otimes \bar{x}(t) - \bar{x}^{\mathrm{T}}(t-T) \otimes \bar{x}(t-T) \\ &- \int_{t-T}^{t} \left(2\gamma\bar{x}^{\mathrm{T}}(\tau) \otimes \bar{x}(\tau) + 2\bar{x}^{\mathrm{T}}(\tau) \otimes \left(\bar{B}(\tau)u(\tau)\right)\right)\mathrm{d}\tau,\end{aligned}$$

admits a constant solution with respect to P_{γ}, which satisfies the model-based Lyapunov equation. In addition, if Ψ is persistently exciting, this solution is unique.

CREATING THE NEXT®

Learning-based estimation of *f*

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

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

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

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

Gradient descent learning law:

$$\mathrm{vech}(\dot{\hat{P}}_\gamma) = -\beta \frac{\Psi(t)}{1 + \|\Psi(t)\|^2}\left(\Psi^{\mathrm{T}}(t)\mathrm{vech}\left(\hat{P}_\gamma\right) + \int_{t-T}^{t} \|\bar{x}(\tau)\|^2 \,\mathrm{d}\tau\right).$$

Learning-based estimation of *f*

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 $\bar{\Psi}(\cdot) := \frac{\Psi(\cdot)}{\sqrt{1+\|\Psi(\cdot)\|^2}}$ is persistently exciting, the norm of the estimation error vector $\left\|\text{vech}(\tilde{P}_\gamma)\right\|$ decays to zero exponentially fast.

Online actuator placement

Having an estimate of P_γ, we can solve the approximate optimization:

$$\max_{\mathcal{V}\subseteq\mathcal{B}} \ \hat{f}(\mathcal{V};\ t_j) = \sum_{v\in\mathcal{V}} (r_v^{-1} - k_v^{-1}) v^{\mathrm{T}} \hat{P}_\gamma(t_j) v,$$
$$\text{s.t.} \quad \mathrm{card}(\mathcal{V}) = k,$$

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

Complexity: $\mathcal{O}(N\log N)$.

Online actuator placement

Online Placement Algorithm:

Gradient Learning Law:

Approximate Optimization:

Switch Actuators:

Algorithm 1: Data-Based Actuator Placement

Input: Constants $k_{b_i} > r_{b_i} > 0$, $i = 1, \ldots, N$, and $\gamma > \alpha(A)$.

Output: Actuator Selection Sequence $\{\mathcal{V}_j\}_{j \in \mathbb{N}}$.

```
procedure
    for t ⩾ 0 do
        Tune P̂_γ according to learning law (25).
        if t = t_j, j ∈ ℕ, then
            Sort the values (r_{b_i}^{-1} − k_{b_i}^{-1}) b_i^T P̂_γ(t_j) b_i, ∀i = 1, ..., N.
            Choose the set of actuators V_j ⊆ B corresponding the k largest such values.
            Set B̄(t) = [B_0 | B_{V_j}], ∀t ∈ [t_j, t_{j+1}).
        end if
    end for
end procedure
```

In mathematical notation: Tune $\hat{P}_\gamma$ according to learning law (25); Sort the values $(r_{b_i}^{-1} - k_{b_i}^{-1}) b_i^{\mathrm{T}} \hat{P}_\gamma(t_j) b_i$, $\forall i = 1, \ldots, N$; Choose the set of actuators $\mathcal{V}_j \subseteq \mathcal{B}$ corresponding the k largest such values; Set $\bar{B}(t) = [B_0 \mid B_{\mathcal{V}_j}]$, $\forall t \in [t_j, t_{j+1})$.

Online actuator placement

Theorem: Let $\mathcal{V}^{\star'}$ be an optimal solution to the actual problem, and $\mathcal{V}^{\star}$ an optimal solution to the purturbed problem. Then:

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

$$\max_{\mathcal{V}\subseteq\mathcal{B}} \quad f_p(\mathcal{V}) = \sum_{v\in\mathcal{V}} \left((r_v^{-1} - k_v^{-1})v^{\mathrm{T}} P_\gamma v + \eta_v\right),$$
$$\text{s.t.} \quad \mathrm{card}(\mathcal{V}) = k,$$

where η_{b_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:

$$\operatorname{vech}(\dot{\hat{P}}_{\gamma,s}) = -\beta \frac{\Psi_s(t)}{1+\|\Psi_s(t)\|^2}\left(\Psi_s^{\mathrm{T}}(t)\operatorname{vech}\left(\hat{P}_{\gamma,s}\right) + \int_{t-T}^{t}\|x(\tau)\|^2\,\mathrm{d}\tau\right) - k_d(\operatorname{vech}(\hat{P}_{\gamma,s} - \hat{P}_{\gamma,s}(T)),$$

$$\operatorname{vech}(\hat{P}_{\gamma,s}(T)) = \operatorname{vech}(P_\gamma),\ t \geq T,$$

with $k_d > 0$, and Ψ_s being Ψ at the steady state.

Theorem: An attacker can remain undetected only if:

$$\mathbf{C1}: \int_{t-T}^{t} x^{\mathrm{T}}(\tau)P_\gamma B a(\tau)\mathrm{d}\tau = 0, \text{or}$$

$$\mathbf{C2}: \Psi_s(t) = 0.$$

Lemma: Let $d(t) := \left\|\hat{P}_{\gamma,s}(t) - \hat{P}_{\gamma,s}(T)\right\|_F$, $\forall t \geq T$. 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 $\mathcal{V}^\star = \{b_1,\ b_3,\ b_7,\ b_{10}\}$
- 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 $k_{b_1} = k_{b_3} = k_{b_5} = k_{b_6} = k_{b_7} = 10$, $k_{b_4} = k_{b_8} = k_{b_9} = 2$ and $k_{b_2} = k_{b_{11}} = 1.01$.

Simulations

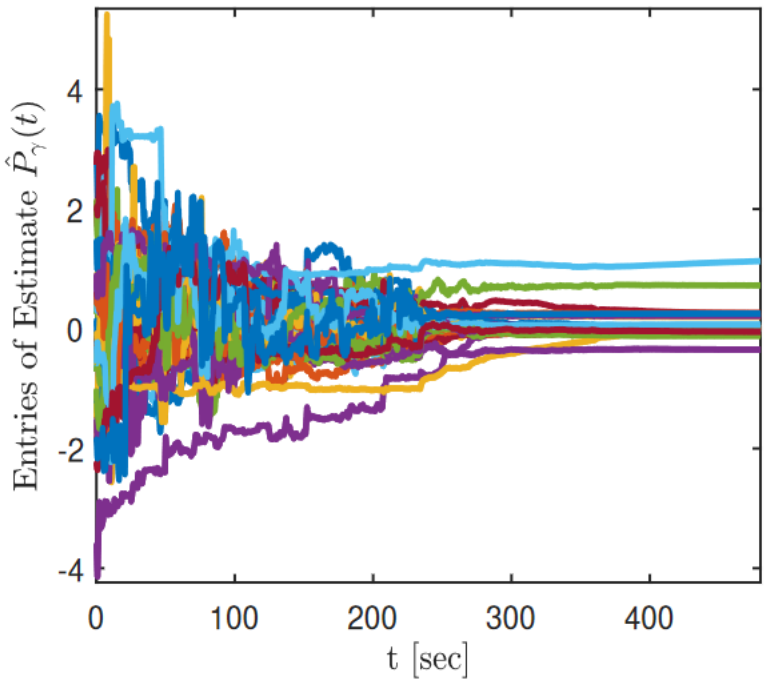


Figure 1. The evolution of the estimate of P_γ resulting from the application of the learning law (25).

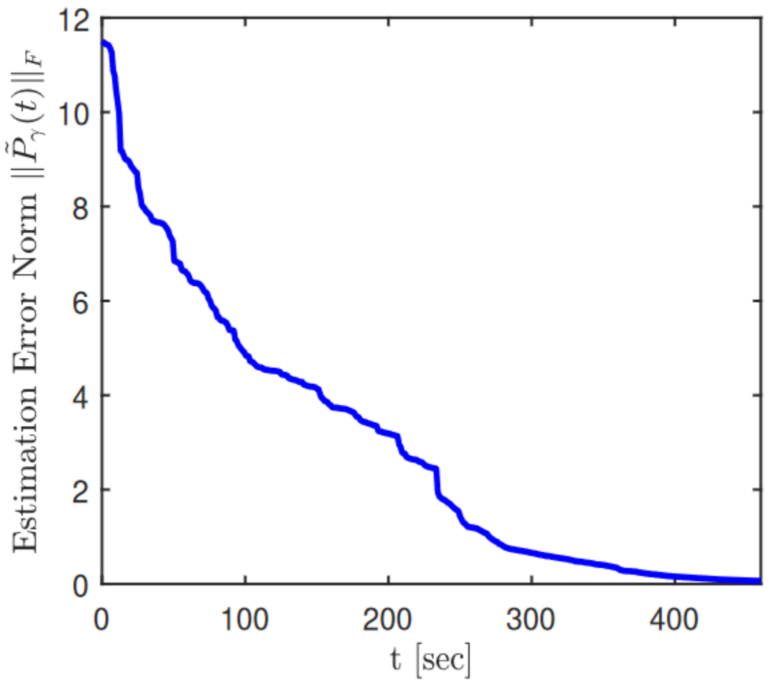


Figure 2. The evolution of the Frobenius norm of the estimation error $\left\|\tilde{P}_\gamma\right\|_F$.

The matrix P_γ is successfully identified!

Simulations

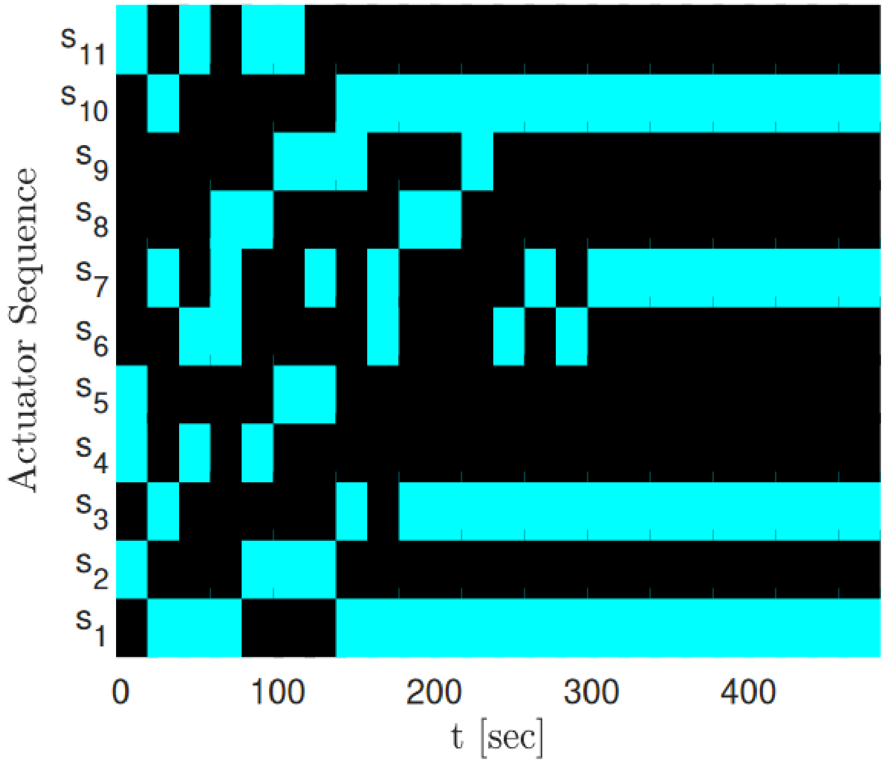


Figure 3. The evolution of the actuator sequence $\mathcal{V}_j$ that was created due to Algorithm 1. Cyan color denotes that an actuator is chosen, while black color denotes that an actuator is not being used.

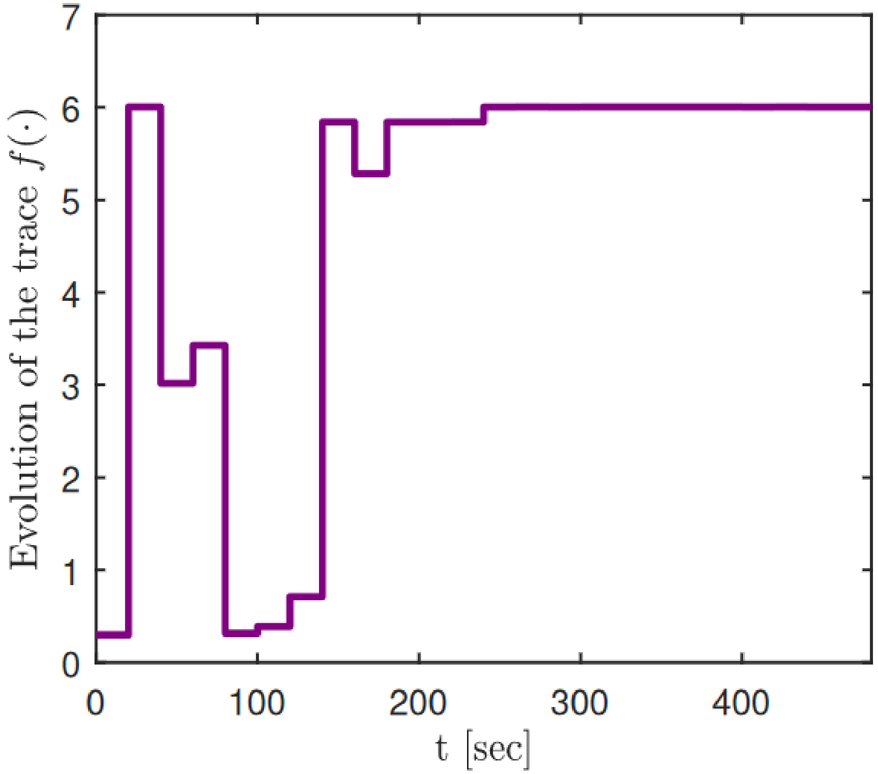


Figure 4. The evolution of the actuator evaluation metric (17) for the sequence $\bar{B}(t)$ of actuators generated by Algorithm 1.

The optimal set of actuators is found after 300 s!

Simulations

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.

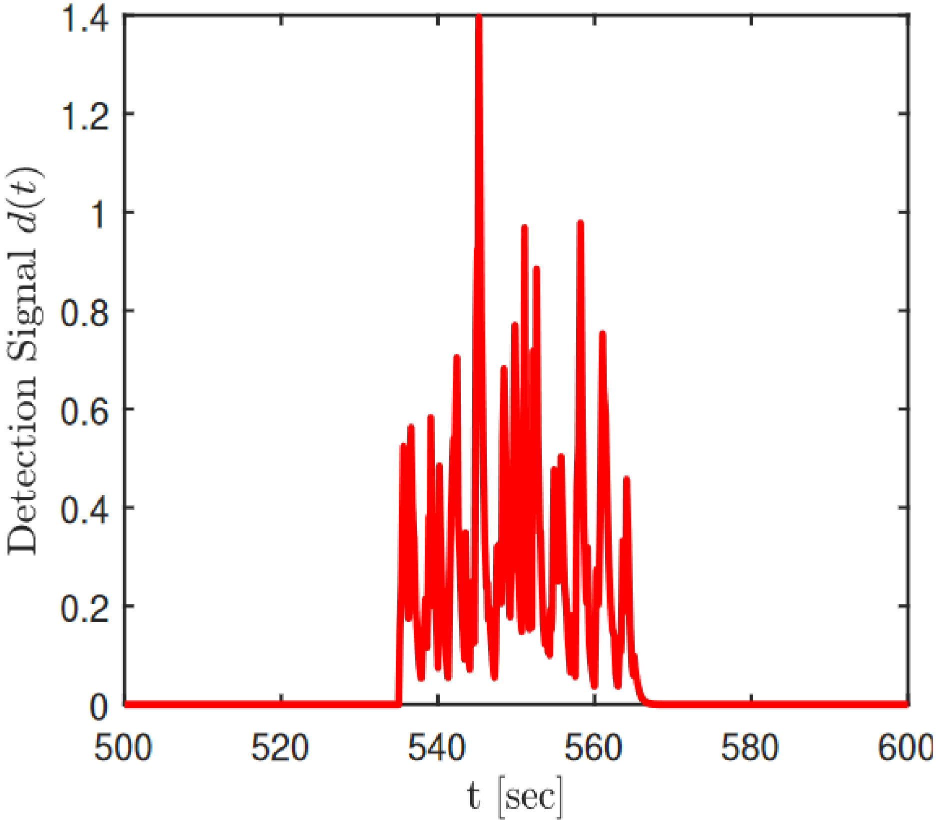


Figure 5. The evolution of the detection signal $d(t)$ over $t \in [500,\ 600]$ seconds.

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.

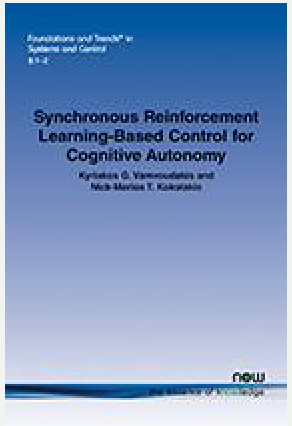
Future work

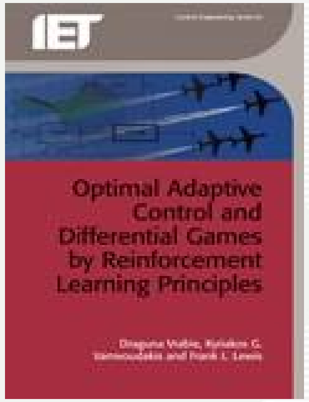
- Extension to cases of joint sensor-actuator attacks.
- Consideration of cases where more statistics of the disturbance are available.
- Implementation in a networked control setting with decentralized information.
- Development of resilient RHC to deal with possible DoS attacks.
- Extension to sensor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

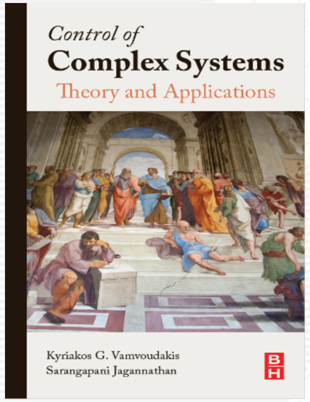
QUESTIONS?

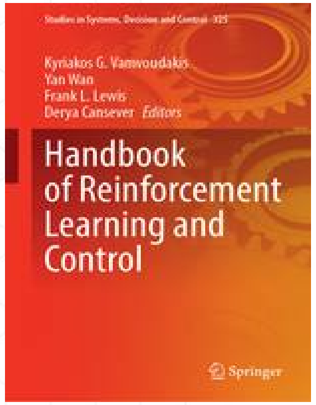
THANK YOU

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









Filippos Fotiadis

Georgia Tech

CREATING THE NEXT®