Assuring Autonomy in Contested Environments

Attack-Resilient Design

Miroslav Pajic

Cyber-Physical Systems Lab (CPSL)

Pratt School of Engineering

Duke University

UF UNIVERSITY of FLORIDA

Duke UNIVERSITY

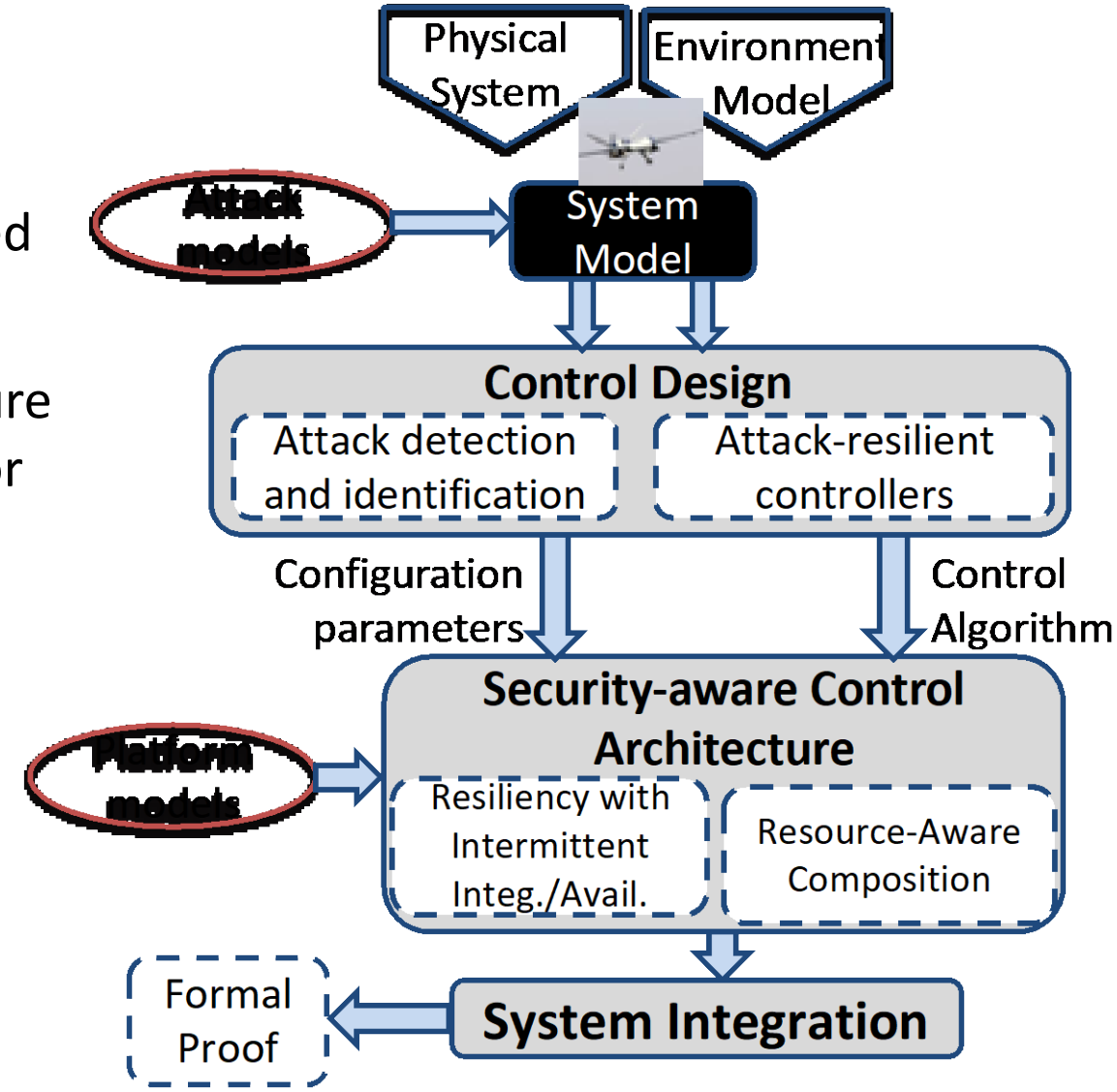
TEXAS The University of Texas at Austin

UC SANTA CRUZ

Attack Resilient Design

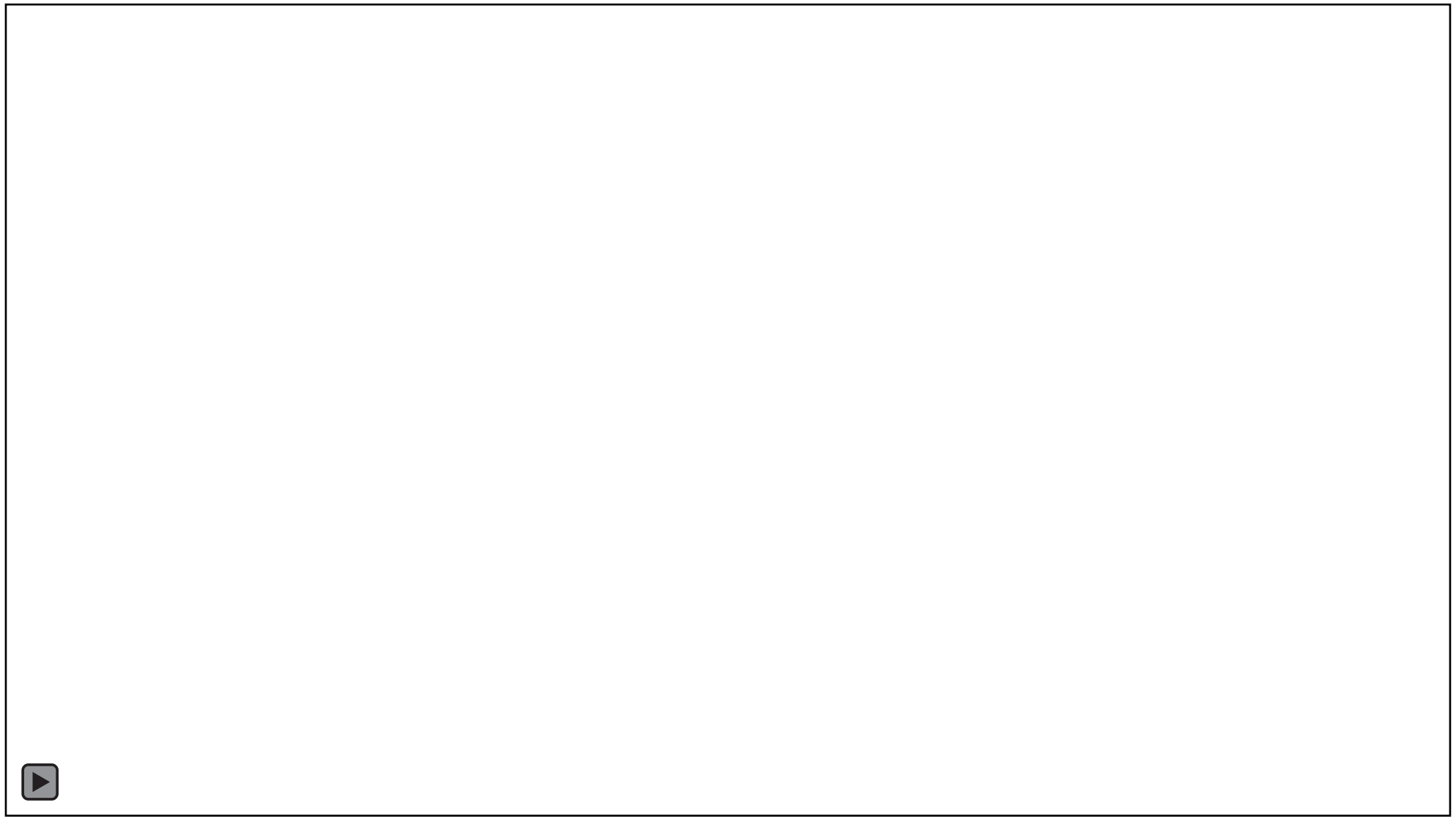
Duke
PRATT SCHOOL of ENGINEERING

- Distributed methods for attack-detection & identification (ADI) and security-aware mission planning by exploiting dynamics of the environment and agents, as well as our knowledge of the expected mission evolution
- Develop a platform-aware attack-resilient architecture integrating the control and estimation techniques for resource-constrained autonomous systems
- Develop methods to assure desired QoC despite communication and computation limitations (optimal balance between QoC and security guarantees)
- Investigate MDPs, PTAs, and stochastic hybrid automata (SHAs, to model the interaction between the IDS and the controller/environment



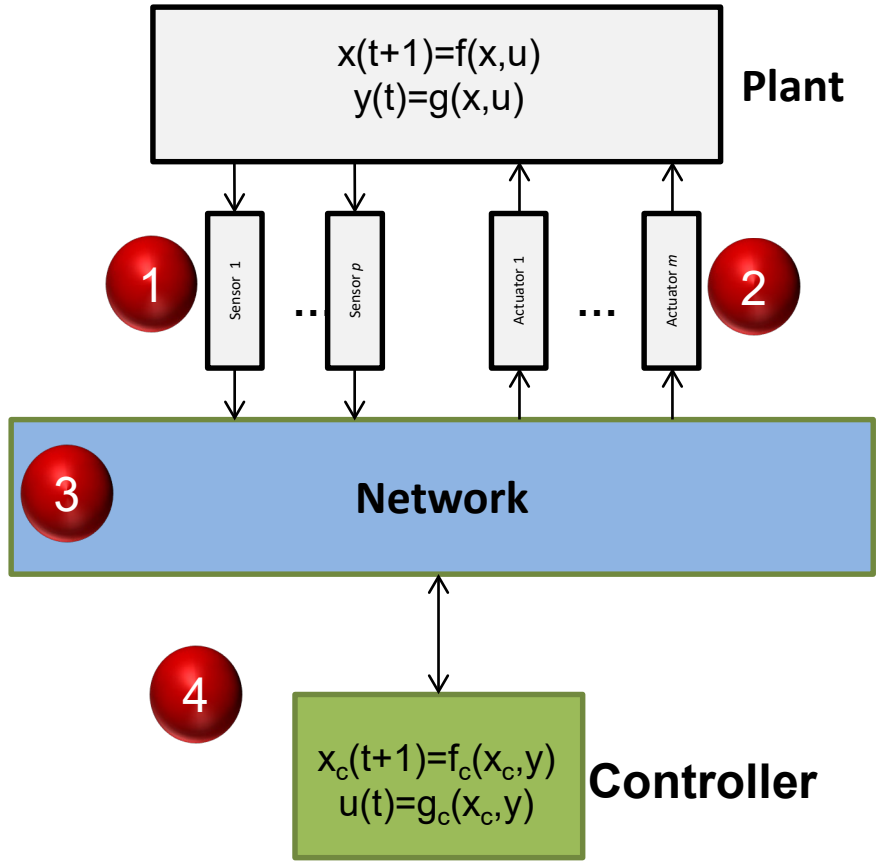
Vehicle Trajectory Following

Duke
PRATT SCHOOL of ENGINEERING



Attacks on Autonomous Control

1. Sensor attacks
   - The attacker can arbitrarily change sensor measurements
2. Actuator attacks
   - The attacker can arbitrarily change actuator values
3. Communication attacks
   - The attacker can change messages between sensors and controllers, and messages between controllers and actuators.
4. Controller attacks
   - The attacker can change the controllers' parameters, resources (e.g., execution model) or even the controllers' code.



Security-Aware Design of Autonomous Systems

Duke
PRATT SCHOOL of ENGINEERING

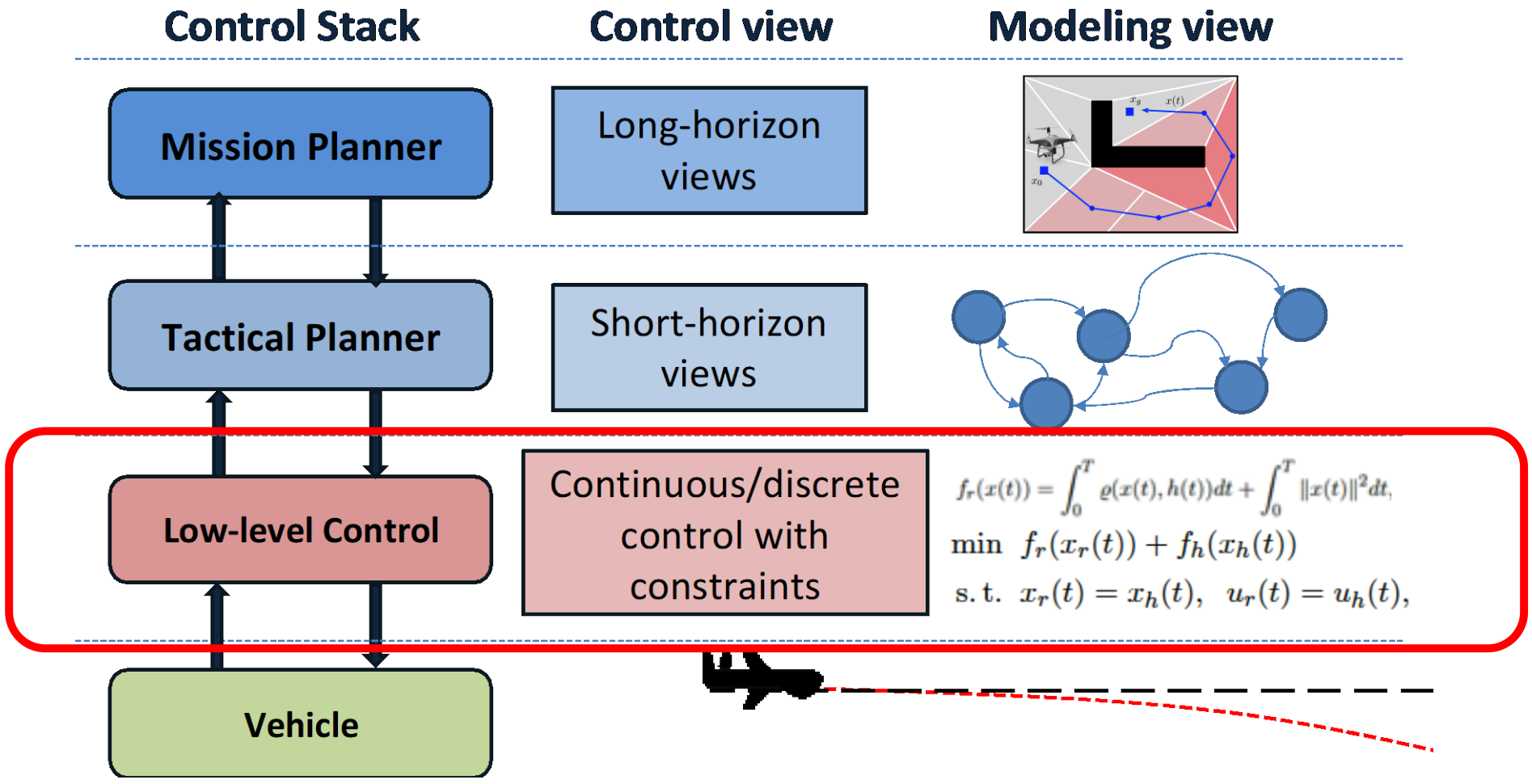
- Physical world abides by the laws of physics!
- Physical interfaces introduce new attack vectors!

- How can we exploit ***limited*** knowledge of laws of physics (system model) for control and attack detection/identification

- Attack-Resilient design with *uncertainty, resource/platform constraints*, as well as varying (especially high) levels of autonomy
  - How much can the attacker exploit modeling limitation?
  - How can we effectively exploit physics to improve guarantees in the presence of attacks?

Security-Aware Control for Autonomous Systems

Duke
PRATT SCHOOL of ENGINEERING

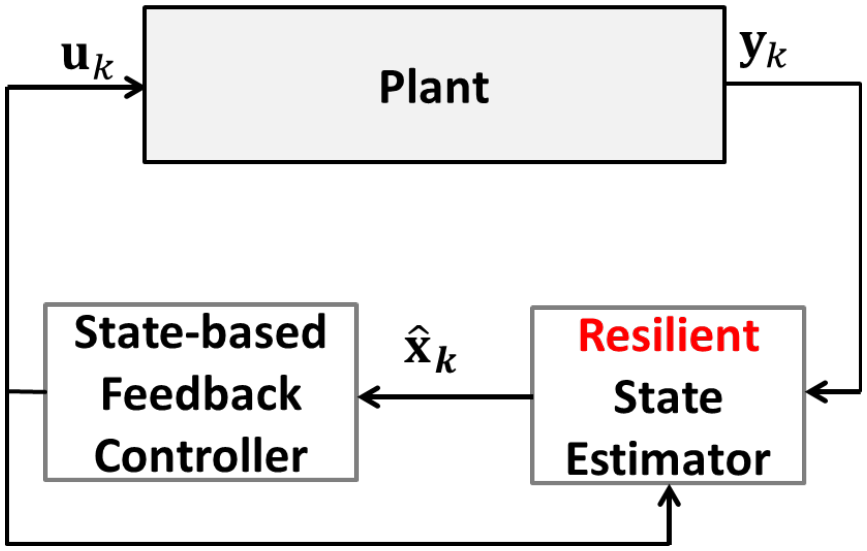


Our Goal: Add resiliency to controls across different/all levels of control stack

Attack-resilient State Estimation

Duke
PRATT SCHOOL of ENGINEERING

- Attack-resilient control of Cyber-Physical Systems
  - Idea: Design attack-resilient state estimators



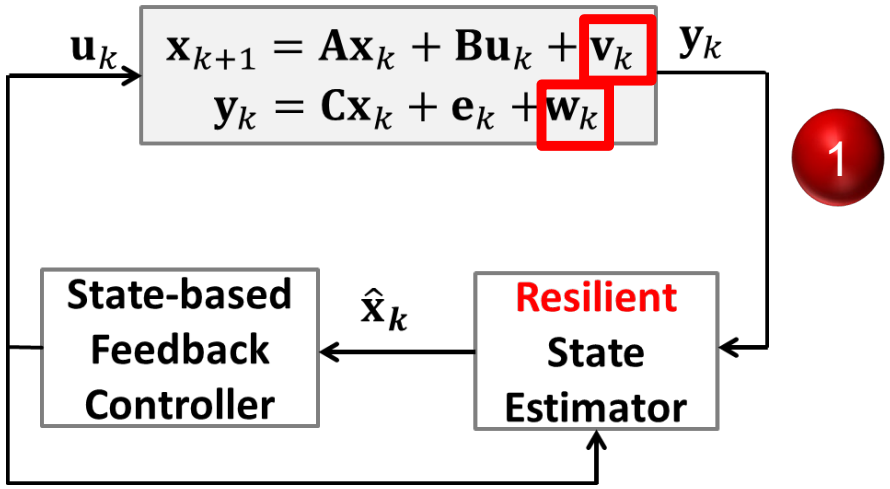
- Initially required an accurate LTI system model
  - Fawzi *et al.* 2012
  - Pasqualetti *et al.* 2013

- If the number of attacked sensors is below a threshold, state can be reconstructed from a history of sensor readings [Fawzi *et al.* 2012]
  - Also identifies the sensors under attack

Modeling attacks on sensors and *actuators*

- Consider an LTI system
  - $\mathbf{x}_k \in \mathbb{R}^n$ plant's state at time k
  - $\mathbf{u}_k \in \mathbb{R}^m$ plant input at time k
  - $\mathbf{y}_k \in \mathbb{R}^p$ plant output
    - state information is availably only via sensors measurements

$$\mathcal{S} = \{s_1, s_2, \dots, s_p\}$$



Attacker wants to [illegible] by creating a discrepancy between [illegible] and the estimates

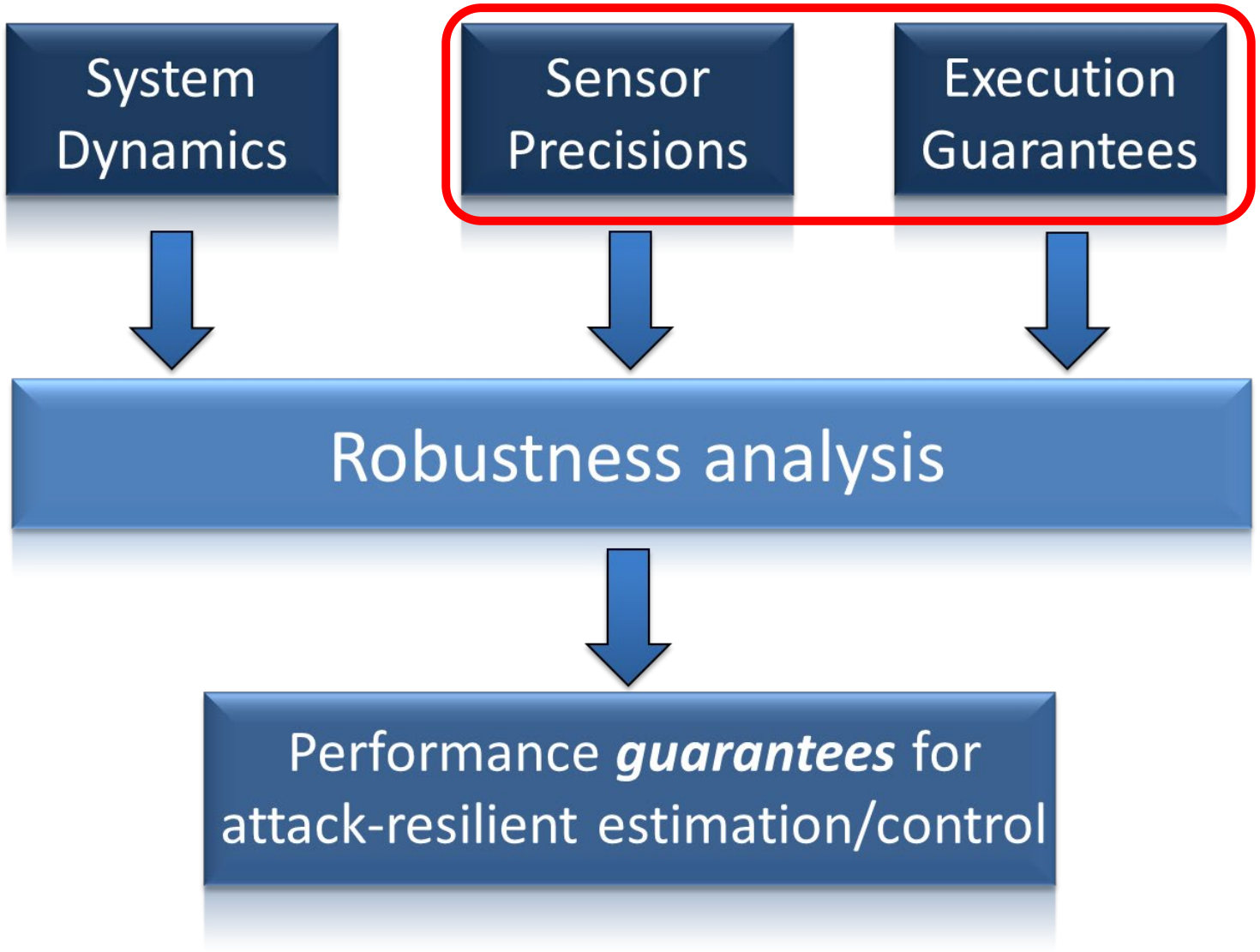
- Attacks on sensors in $\mathcal{K} = \{s_{i_1}, \dots, s_{i_q}\} \subseteq \mathcal{S}$
  - modeled with attack vector $\mathbf{e}_k$
  - $\mathbf{e}_{k,i} \neq 0 \iff$ sensor s_i is under attack at time k

$$\mathcal{K} = \{s_2, s_5\}$$

$$\mathbf{e}_k = \begin{bmatrix} 0 \\ 1.7 \\ 0 \\ 0 \\ -9 \end{bmatrix}$$

Goal: Attack-resilient state-estimation with performance guarantees

Duke
PRATT SCHOOL of ENGINEERING



Problem Description

- We consider an LTI system, with state $\mathbf{x} \in \mathbb{R}^n$ and output $\mathbf{y} \in \mathbb{R}^p$ measurements from the set of sensors $\mathcal{S} = \{s_1, s_2, \ldots, s_p\}$

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{w}_k + \mathbf{e}_k$$

  - bounded size noise $|\mathbf{w}_k| \leq \delta_{w_k}$
  - sparse attack vector $\mathbf{e} \in \mathbb{R}^p$ captures attacks on a subset of sensors

$$\mathcal{K} = \{s_{i_1}, s_{i_2}, \ldots, s_{i_q}\} \subseteq \mathcal{S}$$

$$\mathcal{K} = \{s_2, s_5\}$$

$$\mathbf{e}_k = \begin{bmatrix} 0 \\ 1.7 \\ 0 \\ 0 \\ -9 \end{bmatrix}$$

- **Goal**: *Reconstruction* of the initial system state $\mathbf{x}_0$ from a set of *N* output observations $\mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_{N-1}$

$$\mathbf{y}_k = \mathbf{C}\mathbf{A}^k\mathbf{x}_0 + \mathbf{e}_k + \mathbf{w}_k$$

Representation via Block Vectors

Duke
PRATT SCHOOL of ENGINEERING

- System evolution observed via a single sensor

$$\tilde{\mathbf{y}}_i = \begin{bmatrix} \mathbf{y}_{0,i} \\ \mathbf{y}_{1,i} \\ \vdots \\ \mathbf{y}_{N-1,i} \end{bmatrix} \in \mathbb{R}^N \qquad \tilde{\mathbf{e}}_i = \begin{bmatrix} \mathbf{e}_{0,i} \\ \mathbf{e}_{1,i} \\ \vdots \\ \mathbf{e}_{N-1,i} \end{bmatrix} \in \mathbb{R}^N \qquad \tilde{\mathbf{w}}_i = \begin{bmatrix} \mathbf{w}_{0,i} \\ \mathbf{w}_{1,i} \\ \vdots \\ \mathbf{w}_{N-1,i} \end{bmatrix} \in \mathbb{R}^N \qquad 1 \leq i \leq p$$

$$\tilde{\mathbf{y}}_i = \mathbf{O}_i \mathbf{x}_0 + \tilde{\mathbf{e}}_i + \tilde{\mathbf{w}}_i$$

$$\mathbf{O}_{s_i} = \begin{bmatrix} P_{\{s_i\}}\mathbf{C} \\ P_{\{s_i\}}\mathbf{C}\mathbf{A} \\ \vdots \\ P_{\{s_i\}}\mathbf{C}\mathbf{A}^{N-1} \end{bmatrix} \qquad \mathbf{O}_{\mathcal{K}} = \begin{bmatrix} \mathbf{O}_{s_{i_1}} \\ \mathbf{O}_{s_{i_2}} \\ \vdots \\ \mathbf{O}_{s_{i_{|\mathcal{K}|}}} \end{bmatrix}$$

- System evolution observed from all sensors

$$\tilde{\mathbf{y}} = \begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \vdots \\ \tilde{\mathbf{y}}_p \end{bmatrix}, \tilde{\mathbf{e}} = \begin{bmatrix} \tilde{\mathbf{e}}_1 \\ \vdots \\ \tilde{\mathbf{e}}_p \end{bmatrix}, \tilde{\mathbf{w}} = \begin{bmatrix} \tilde{\mathbf{w}}_1 \\ \vdots \\ \tilde{\mathbf{w}}_p \end{bmatrix}, \qquad \mathbf{O} = \begin{bmatrix} \mathbf{O}_1 \\ \vdots \\ \mathbf{O}_p \end{bmatrix}$$

$$\tilde{\mathbf{y}} = \mathbf{O}\mathbf{x}_0 + \tilde{\mathbf{e}} + \tilde{\mathbf{w}}$$

Representation via Block Vectors

- System evolution observed via a single sensor

$$\tilde{\mathbf{y}}_i = \begin{bmatrix} \mathbf{y}_{0,i} \\ \mathbf{y}_{1,i} \\ \vdots \\ \mathbf{y}_{N-1,i} \end{bmatrix} \in \mathbb{R}^N \qquad \tilde{\mathbf{e}}_i = \begin{bmatrix} \mathbf{e}_{0,i} \\ \mathbf{e}_{1,i} \\ \vdots \\ \mathbf{e}_{N-1,i} \end{bmatrix} \in \mathbb{R}^N \qquad \tilde{\mathbf{w}}_i = \begin{bmatrix} \mathbf{w}_{0,i} \\ \mathbf{w}_{1,i} \\ \vdots \\ \mathbf{w}_{N-1,i} \end{bmatrix} \in \mathbb{R}^N \qquad 1 \le i \le p$$

$$\tilde{\mathbf{y}}_i = \mathbf{O}_i \mathbf{x}_0 + \tilde{\mathbf{e}}_i + \tilde{\mathbf{w}}_i$$

$$\mathbf{O}_{s_i} = \begin{bmatrix} P_{\{s_i\}}\mathbf{C} \\ P_{\{s_i\}}\mathbf{C}\mathbf{A} \\ \vdots \\ P_{\{s_i\}}\mathbf{C}\mathbf{A}^{N-1} \end{bmatrix} \qquad \mathbf{O}_{\mathcal{K}} = \begin{bmatrix} \mathbf{O}_{s_{i_1}} \\ \mathbf{O}_{s_{i_2}} \\ \vdots \\ \mathbf{O}_{s_{i_{|\mathcal{K}|}}} \end{bmatrix}$$

- System evolution observed from all sensors

$$\tilde{\mathbf{y}} = \begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \vdots \\ \tilde{\mathbf{y}}_p \end{bmatrix}, \tilde{\mathbf{e}} = \begin{bmatrix} \tilde{\mathbf{e}}_1 \\ \vdots \\ \tilde{\mathbf{e}}_p \end{bmatrix}, \tilde{\mathbf{w}} = \begin{bmatrix} \tilde{\mathbf{w}}_1 \\ \vdots \\ \tilde{\mathbf{w}}_p \end{bmatrix}, \qquad \mathbf{O} = \begin{bmatrix} \mathbf{O}_1 \\ \vdots \\ \mathbf{O}_p \end{bmatrix}$$

q-block sparse vector

$$\tilde{\mathbf{y}} = \mathbf{O}\mathbf{x}_0 + \tilde{\mathbf{e}} + \tilde{\mathbf{w}}$$

$$\|\tilde{\mathbf{e}}\|_{l_2,l_0} = \sum_{i=1}^{p} \mathbb{I}(\|\tilde{\mathbf{e}}_i\|_{l_2} > 0) \qquad \|\tilde{\mathbf{e}}\|_{l_2,l_1} = \sum_{i=1}^{p} \|\tilde{\mathbf{e}}_i\|_{l_2}$$

$$\|\tilde{\mathbf{e}}\|_{l_2,l_0} = q$$

Attack-Resilient State Estimation for Noisy Dynamical Systems

Duke
PRATT SCHOOL of ENGINEERING

- Consider an initial state $\mathbf{x}_0$ and attack vectors from $\tilde{\mathbf{e}}$

$$P_0: \quad \min_{\tilde{\mathbf{e}},\mathbf{x}} \|\tilde{\mathbf{e}}\|_{l_2,l_0} \quad s.t. \quad \tilde{\mathbf{y}} - \mathbf{O}\mathbf{x}_0 - \tilde{\mathbf{e}} = \mathbf{0}$$

$$P_{0,\omega}: \quad \min_{\tilde{\mathbf{e}},\mathbf{x}} \|\tilde{\mathbf{e}}\|_{l_2,l_0} \quad s.t. \quad \tilde{\mathbf{y}} - \mathbf{O}\,\mathbf{x}_0 - \tilde{\mathbf{e}} = \tilde{\mathbf{w}}, \ \tilde{\mathbf{w}} \in \Omega$$

$$P_{1,\omega}: \quad \min_{\tilde{\mathbf{e}},\mathbf{x}} \|\tilde{\mathbf{e}}\|_{l_2,l_1} \quad s.t. \quad \tilde{\mathbf{y}} - \mathbf{O}\,\mathbf{x}_0 - \tilde{\mathbf{e}} = \tilde{\mathbf{w}}, \ \tilde{\mathbf{w}} \in \Omega$$

- Goal: guarantees for $P_{0,\omega}$ and $P_{1,\omega}$ based estimators
  - Bounds on the state estimation errors
  - Sound attacked sensor identification

[ICCPS'14 – **Best paper award**, CDC15, IEEE CSM'17, IEEE TCNS'17]

Performance Guarantees for $P_{0,\omega}$ Estimator

Duke
PRATT SCHOOL of ENGINEERING

$$(\mathbf{x}_{l_0,\omega}, \tilde{\mathbf{e}}^{l_0}) = arg\min P_{0,\omega}, \quad q_{0,\omega} = \|\tilde{\mathbf{e}}^{l_0}\|_{l_2,l_0}$$
$$\Delta\mathbf{x}^{l_0} = \mathbf{x}_{l_0,\omega} - \mathbf{x}_0, \quad \Delta\tilde{\mathbf{e}}^{l_0} = \tilde{\mathbf{e}}^{l_0} - \tilde{\mathbf{e}}^*$$

- Definition [Shoukry *et al*., 2013]: An LTI system is *s*-sparse observable if for every set $\mathcal{K} \subseteq \mathcal{S}$ of size *s*, the pair $(\mathbf{A}, P_{\mathcal{K}^\complement}\mathbf{C})$ is observable

- Lemma: q_{max} is equal to the maximal *s* for which the system is 2*s*-sparse observable

- Theorem: If $q \leq q_{max}$ sensors have been attacked, then

$$\|\Delta\mathbf{x}^{l_0}\|_{l_2} \leq 2 \cdot \max_{\substack{\mathcal{R}\subset S,\\ |\mathcal{R}|=p-2q_{max}}} \left(\|\mathbf{O}_{\mathcal{R}}^{\dagger}\|_{l_2} \cdot \max_{\tilde{\mathbf{w}}_{\mathcal{R}} \in \Omega_{\mathcal{R}}} \|\tilde{\mathbf{w}}_{\mathcal{R}}\|_{l_2} \right)$$

Performance Guarantees for $P_{1,\omega}$ Estimator

Duke
PRATT SCHOOL of ENGINEERING

$$(\mathbf{x}_{l_1,\omega}, \tilde{\mathbf{e}}^{l_1}) = arg\min P_{1,\omega}$$
$$\Delta\mathbf{x}^{l_1} = \mathbf{x}_{l_1,\omega} - \mathbf{x}_0, \qquad \Delta\tilde{\mathbf{e}}^{l_1} = \tilde{\mathbf{e}}^{l_1} - \tilde{\mathbf{e}}^*$$

- Theorem: If *q* sensors from the set $\mathcal{K} \subseteq \mathcal{S}$ have been attacked, then

$$\sum_{s_i \in \mathcal{K}^{\complement}} \|\mathbf{O}_i \Delta\mathbf{x}^{l_1}\|_{l_2} \leq \sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i \Delta\mathbf{x}^{l_1}\|_{l_2} + 2\sigma_\Omega$$

where

$$\sigma_\Omega = \max_{\tilde{\mathbf{w}} \in \Omega} \|\tilde{\mathbf{w}}\|_{l_2,l_1}$$

- Proposition: If $P_{1,\omega}$ correctly estimates the state for a noiseless system, then the error is either zero or for all $\mathcal{K} \subseteq \mathcal{S}$ with *q* elements

$$\sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i \Delta\mathbf{x}^{l_1}\|_{l_2} < \sum_{s_i \in \mathcal{K}^{\complement}} \|\mathbf{O}_i \Delta\mathbf{x}^{l_1}\|_{l_2} \leq \sum_{s_i \in \mathcal{K}} \|\mathbf{O}_i \Delta\mathbf{x}^{l_1}\|_{l_2} + 2\sigma_\Omega$$

Performance Guarantees for $P_{1,\omega}$ Estimator

Duke
PRATT SCHOOL of ENGINEERING

$$(\mathbf{x}_{l_1,\omega}, \tilde{\mathbf{e}}^{l_1}) = arg\min P_{1,\omega}$$
$$\Delta\mathbf{x}^{l_1} = \mathbf{x}_{l_1,\omega} - \mathbf{x}_0, \qquad \Delta\tilde{\mathbf{e}}^{l_1} = \tilde{\mathbf{e}}^{l_1} - \tilde{\mathbf{e}}^*$$

- Theorem: Suppose that for all $\mathcal{K} \subseteq \mathcal{S}$ with *q* elements

$$\mathbf{O}_{\mathcal{K}^{\mathbb{C}}}^T \mathbf{O}_{\mathcal{K}^{\mathbb{C}}} - qN^2 \mathbf{O}_{\mathcal{K}}^T \mathbf{O}_{\mathcal{K}} \succeq \lambda \mathbf{I}_n$$

for some $\lambda > 0$. Then if most *q* sensors are compromised it holds

$$\|\Delta\mathbf{x}^{l_1}\|_{l_2} \leq \frac{2\sqrt{N}\sigma_\Omega}{\lambda} \cdot \max_{\mathcal{K}\subset\mathcal{S},|\mathcal{K}|=q} (\|\mathbf{O}_{\mathcal{K}^{\mathbb{C}}}\|_{l_2} + \sqrt{q}N\|\mathbf{O}_{\mathcal{K}}\|_{l_2})$$

- For N=1, a static state estimation problem $\mathbf{y} = \mathbf{C}\mathbf{X} + \mathbf{e} + \mathbf{w}$
  - The above condition is not as conservative

Attack Identification with Noise and Modeling Errors

Duke
PRATT SCHOOL of ENGINEERING

- Goal: Sound identification of compromised sensors
  - One candidate

$$\mathbb{I}(\tilde{\mathbf{e}}_i^{l_t} \neq \mathbf{0})$$

- Thus, we use the state estimation guarantees

$$Attacked^{l_t}(s_i) = \mathbb{I}(\|\tilde{\mathbf{e}}_i^{l_t}\|_{l_2} > D_i^{\tilde{\mathbf{e}}^{l_t}}), \qquad i = 1, \ldots, p.$$

  - Sound
  - Identifies all attack vectors that satisfy $\|\tilde{\mathbf{e}}_i^*\|_{l_2} > 2D_i^{\tilde{\mathbf{e}}^{l_t}}$

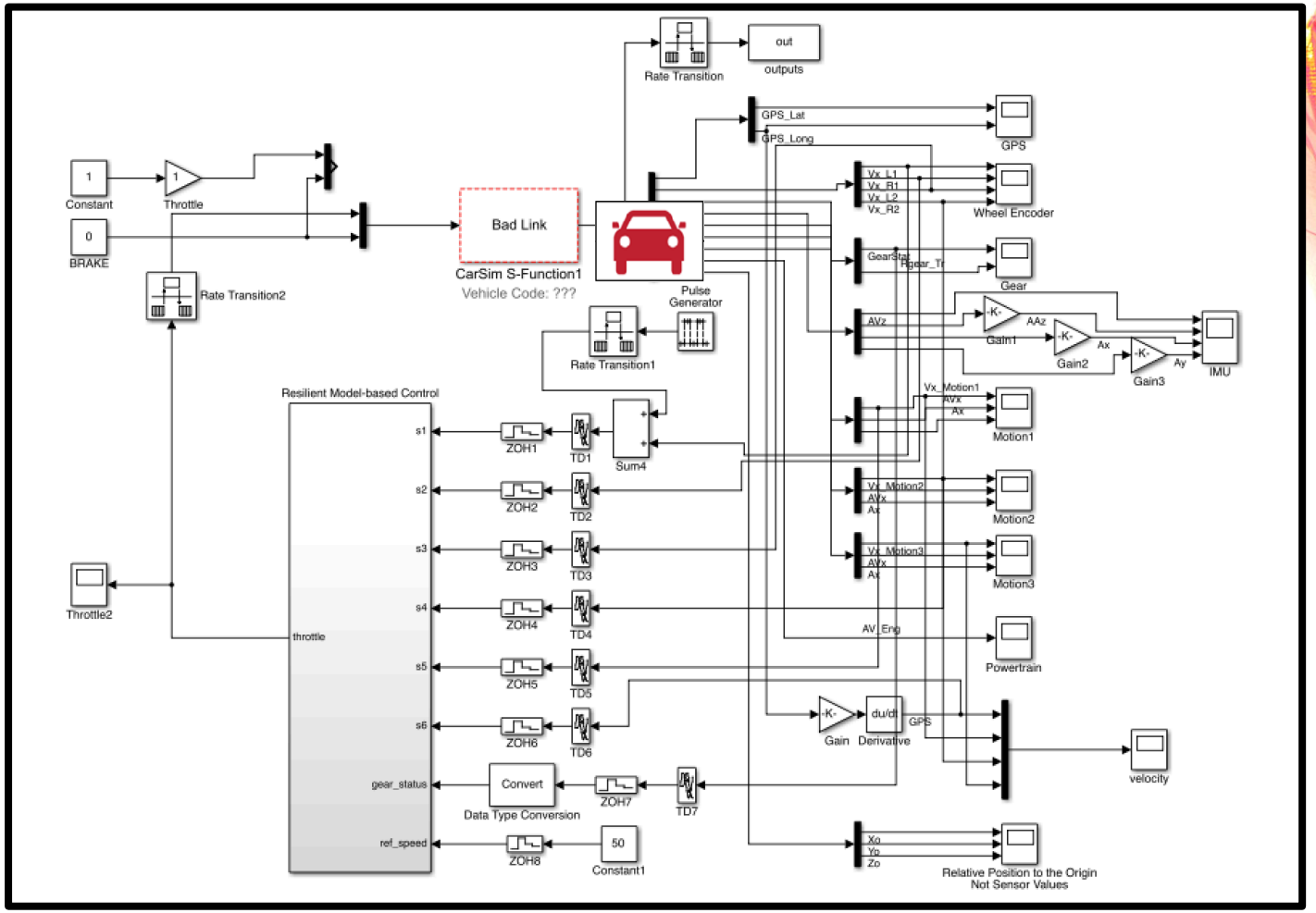
Attack-Resilient Cruise Control Demo

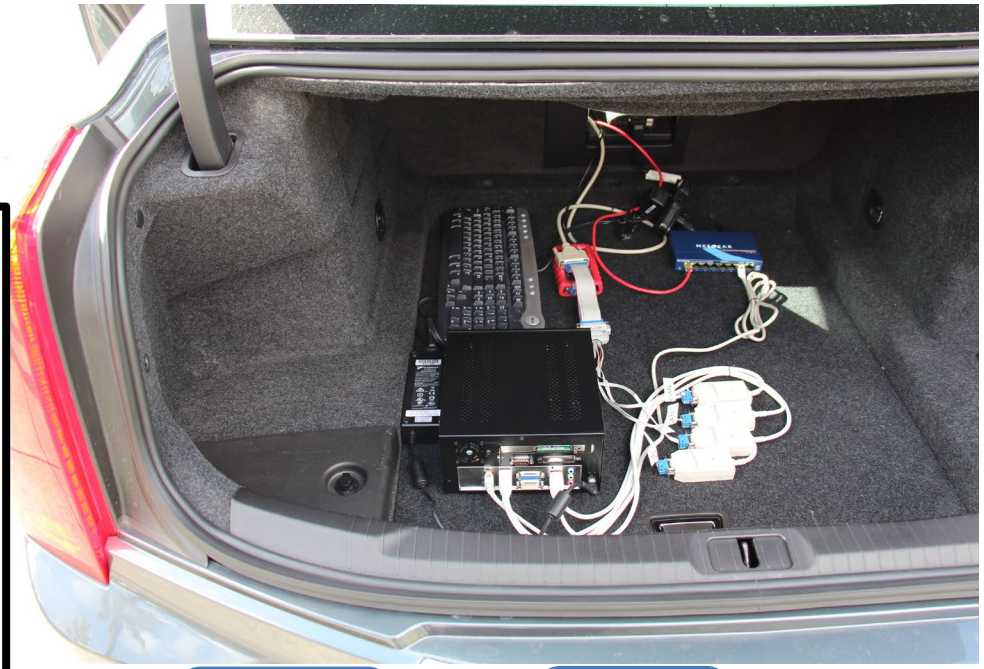
Duke
PRATT SCHOOL of ENGINEERING

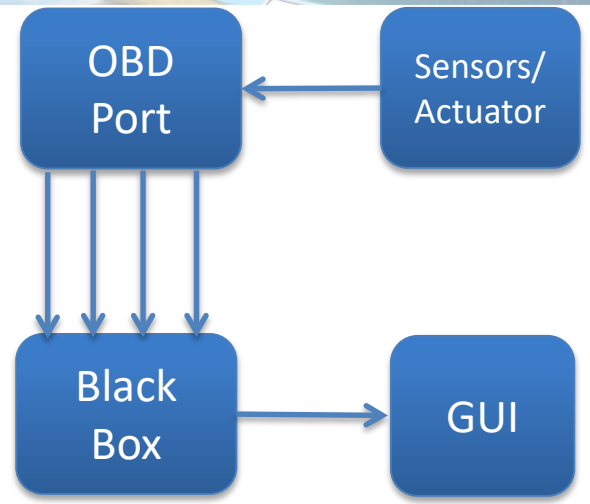
Attack-resilient state estimator for American Built Car

Duke
PRATT SCHOOL of ENGINEERING

- CarSim Simulation
- In-Car Implementation





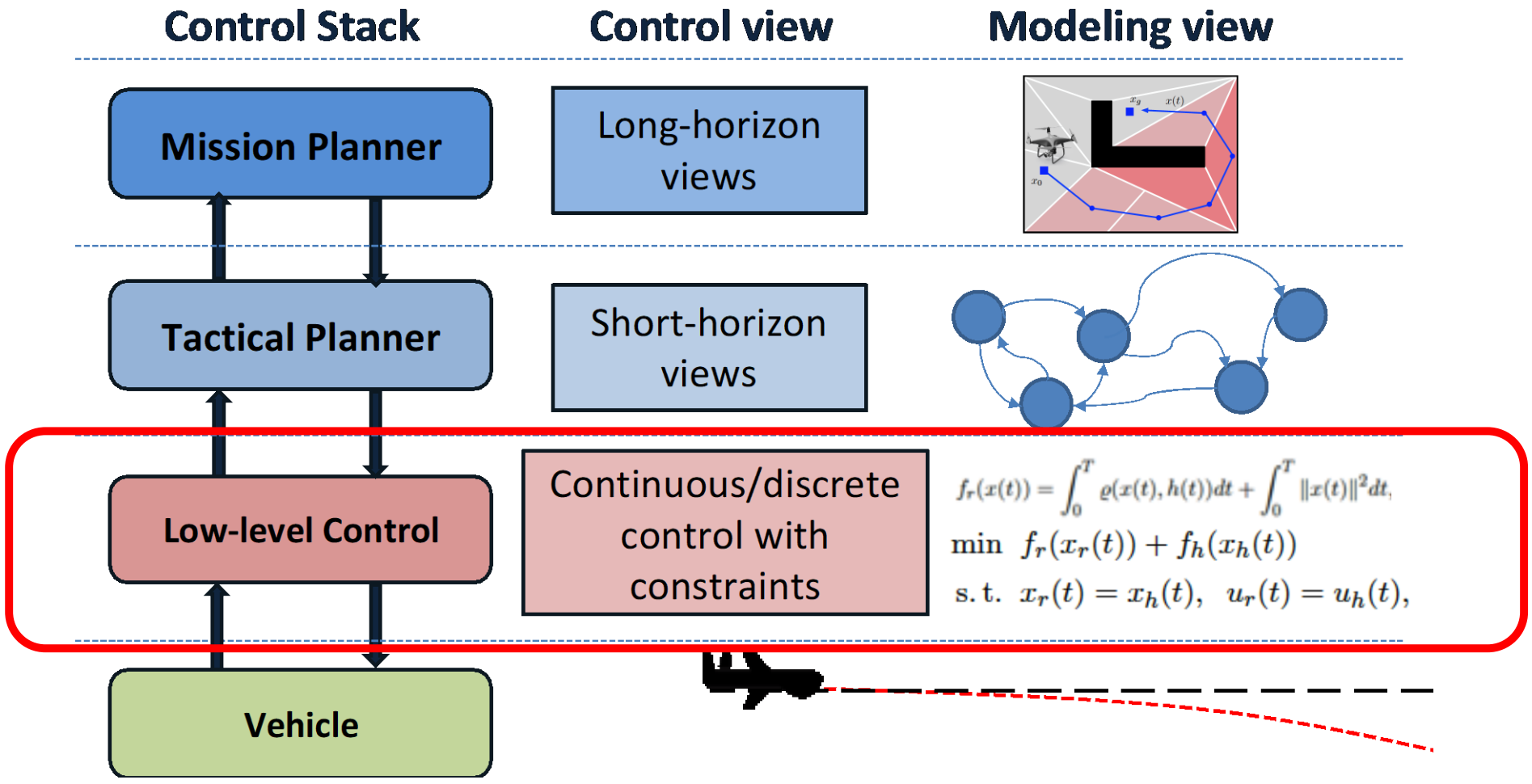


Attack-resilient state estimator for an American Built Car

Duke
PRATT SCHOOL of ENGINEERING

Security-Aware Control for Autonomous Systems

Duke
PRATT SCHOOL of ENGINEERING



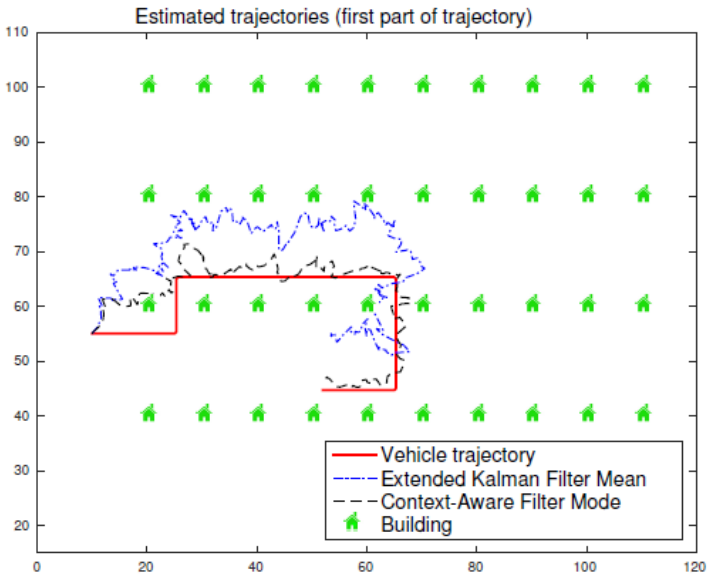
Our Goal: Add resiliency to controls across different/all levels of control stack

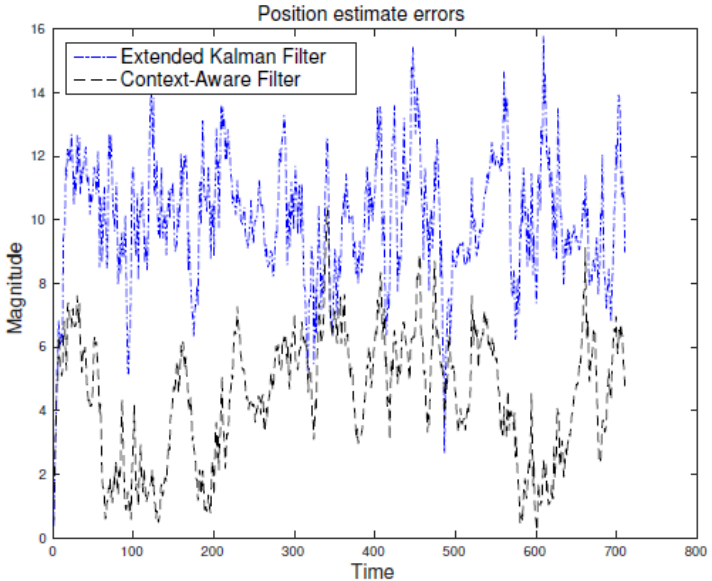
Distributed Control and Estimation in Adversarial Conditions

Duke
PRATT SCHOOL of ENGINEERING

- Guarantees for complex system dynamics
  - Requirements from attack-resilient supervisory control (**IEEE TAC'19, CDC'19a, CDC'19b)**
  - Probabilistic sensor models – Connection to privacy guarantees
- Combining with data-driven methods for attack-detection and identification
- How to add context-based sensing (**IEEE TAC 2018**)

- Using GPS – high variance and bias
- Camera-based landmark recognition





Active Attack Monitoring

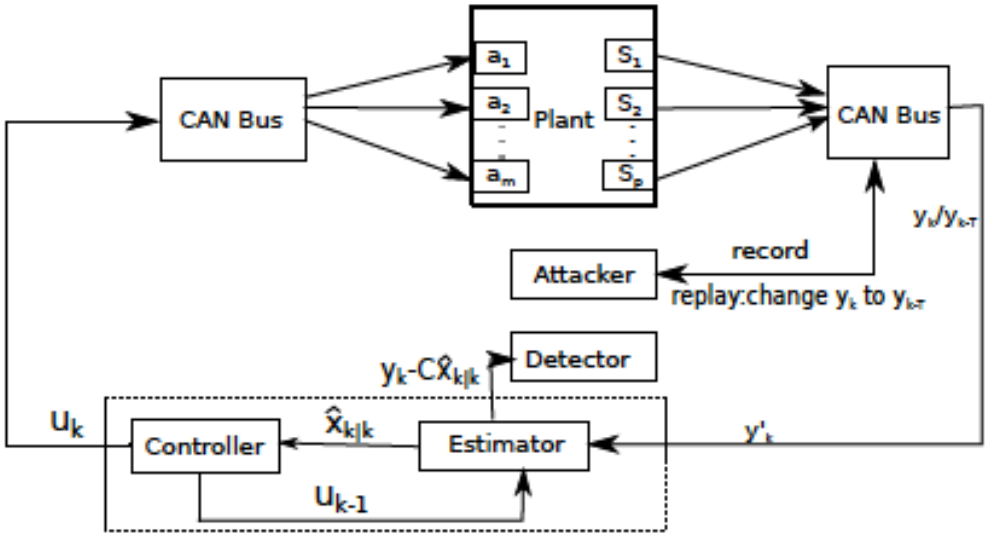
Duke
PRATT SCHOOL of ENGINEERING

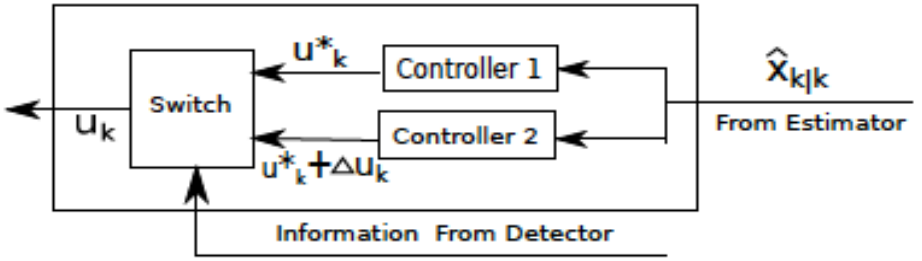
- Available actuation signals are not only used to optimally control physical process, but to also increase confidence that the system has not been compromised
  - For replay attack detection [Automatika'18]

- Challenge: resiliency and performance objective may conflict

- Proposed work: derive framework for optimal use of active monitoring that balances performance and attack detection requirements

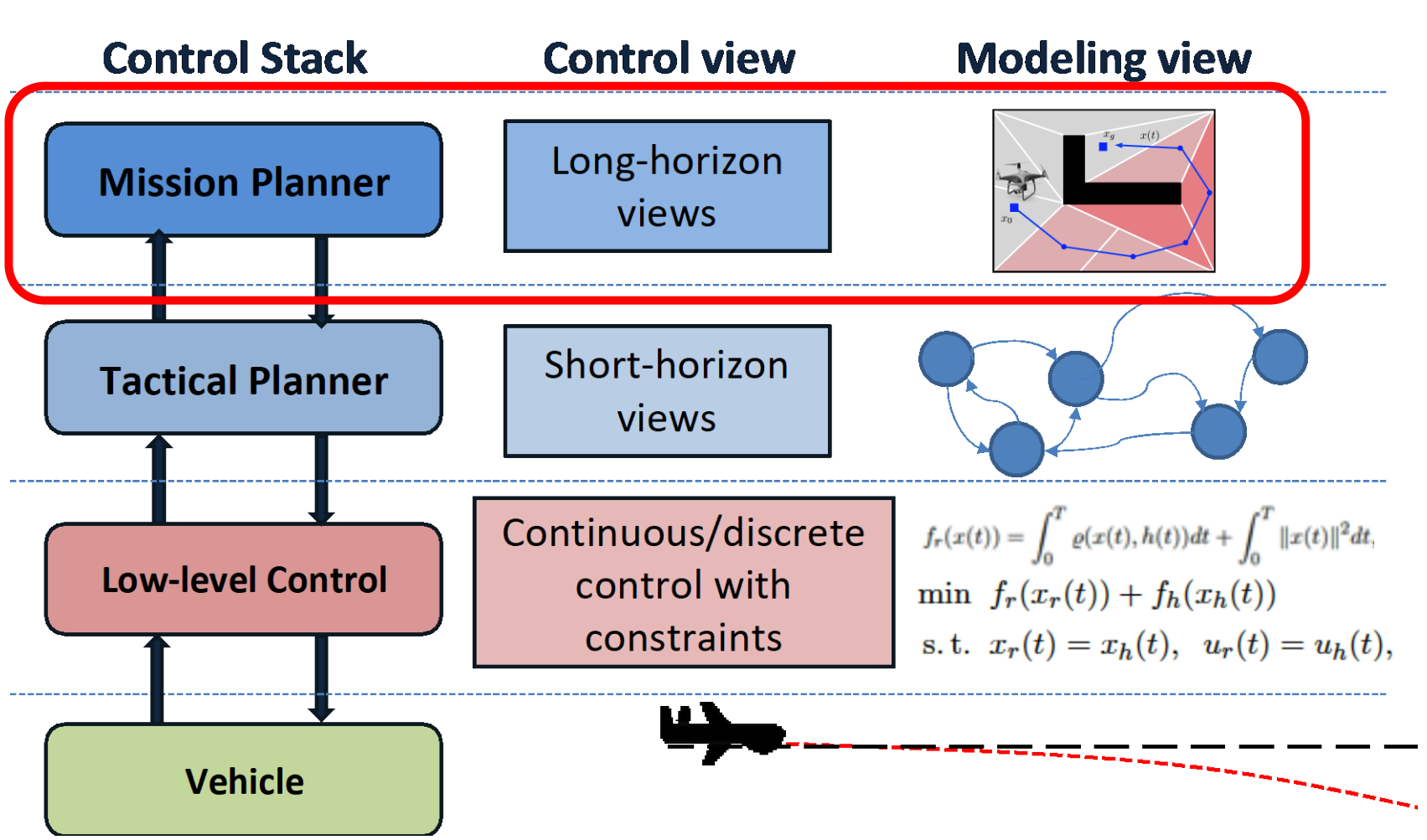
- **Nonlinearity is our friend!**

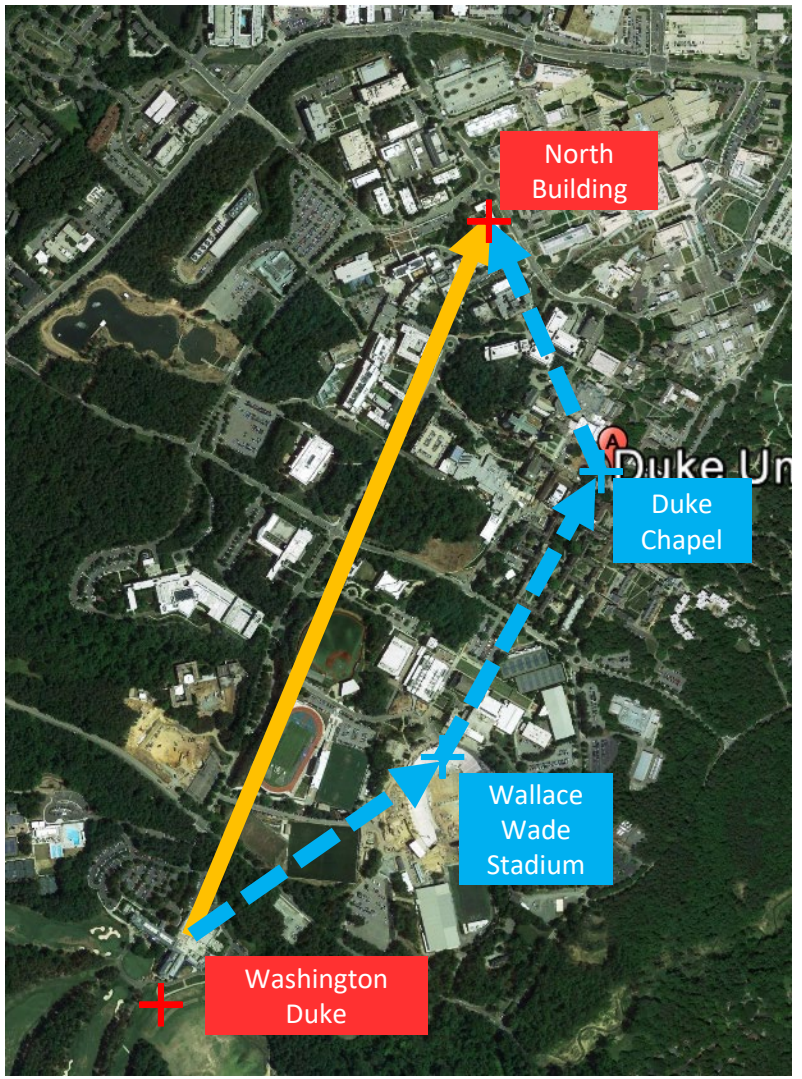




Security-Aware Control for Autonomous Systems

Duke
PRATT SCHOOL of ENGINEERING

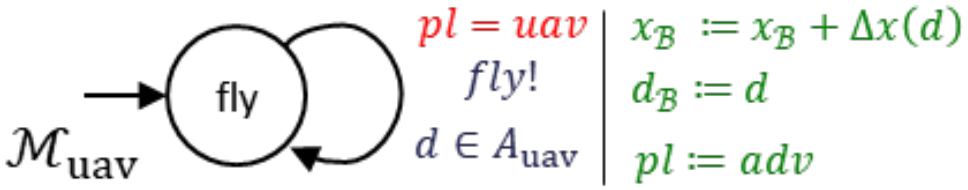




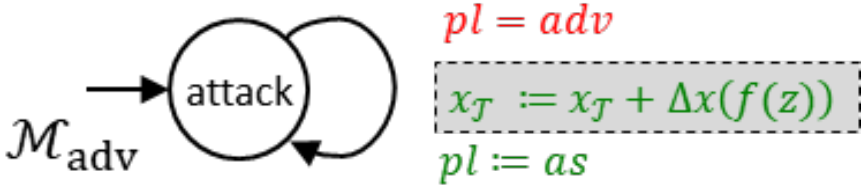
DAG | Hidden-Information Semantics

Duke
PRATT SCHOOL of ENGINEERING

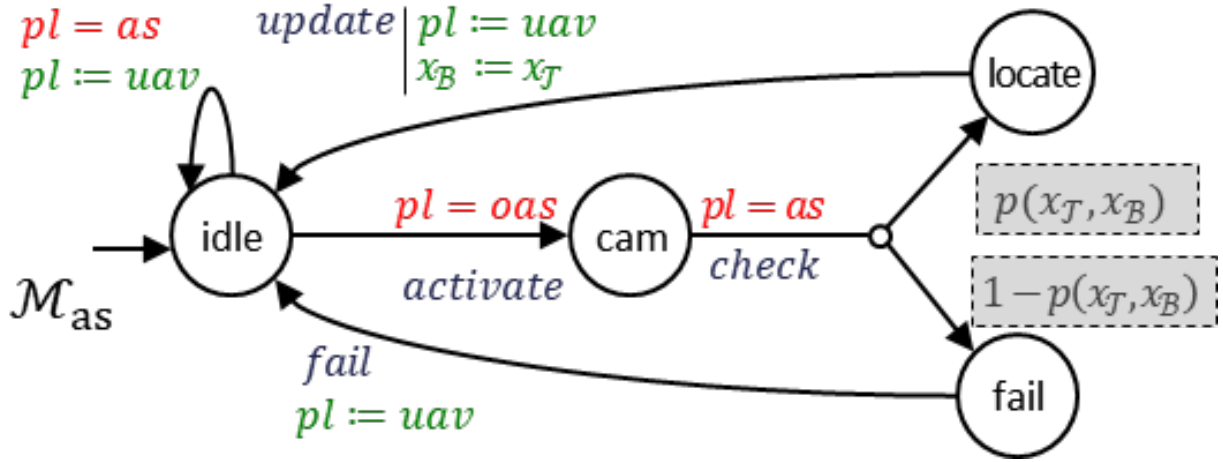
UAV Model



Adversary Model



Advisory System Model

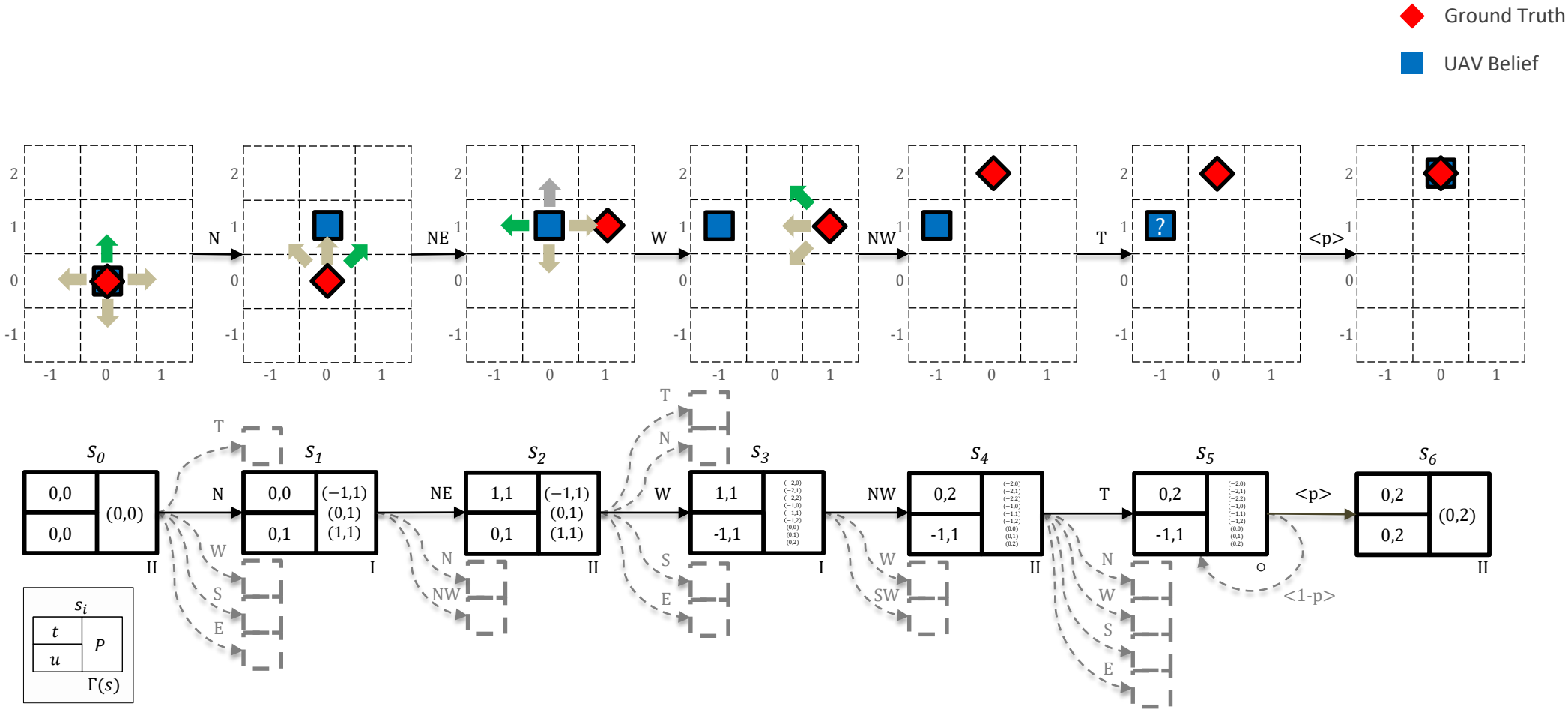


Information inside this box is oftentimes unknown, i.e., **hidden**

Off-the-shelf model checkers do NOT support hidden variables
Strategies CANNOT be synthesized based on hidden information

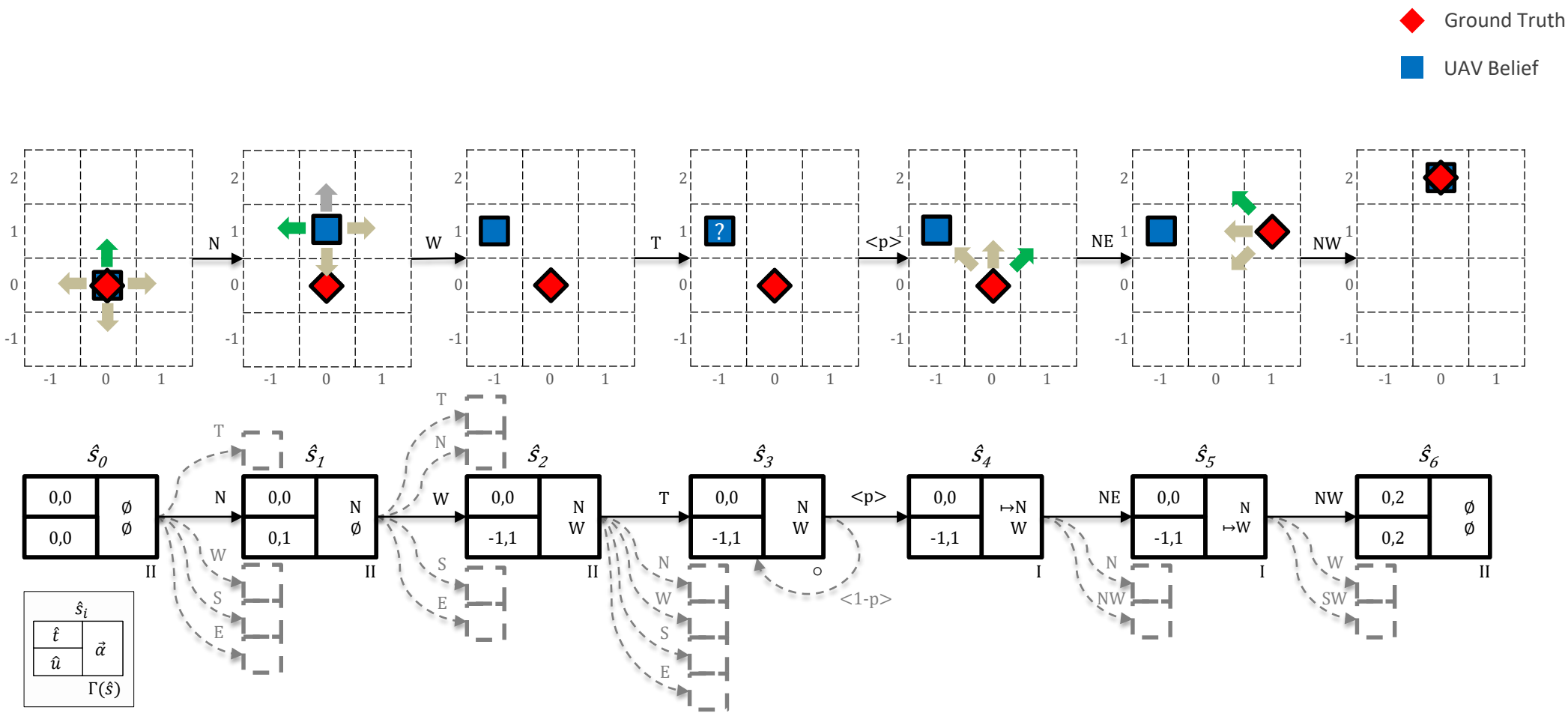
Private Variables Representation

Duke
PRATT SCHOOL of ENGINEERING

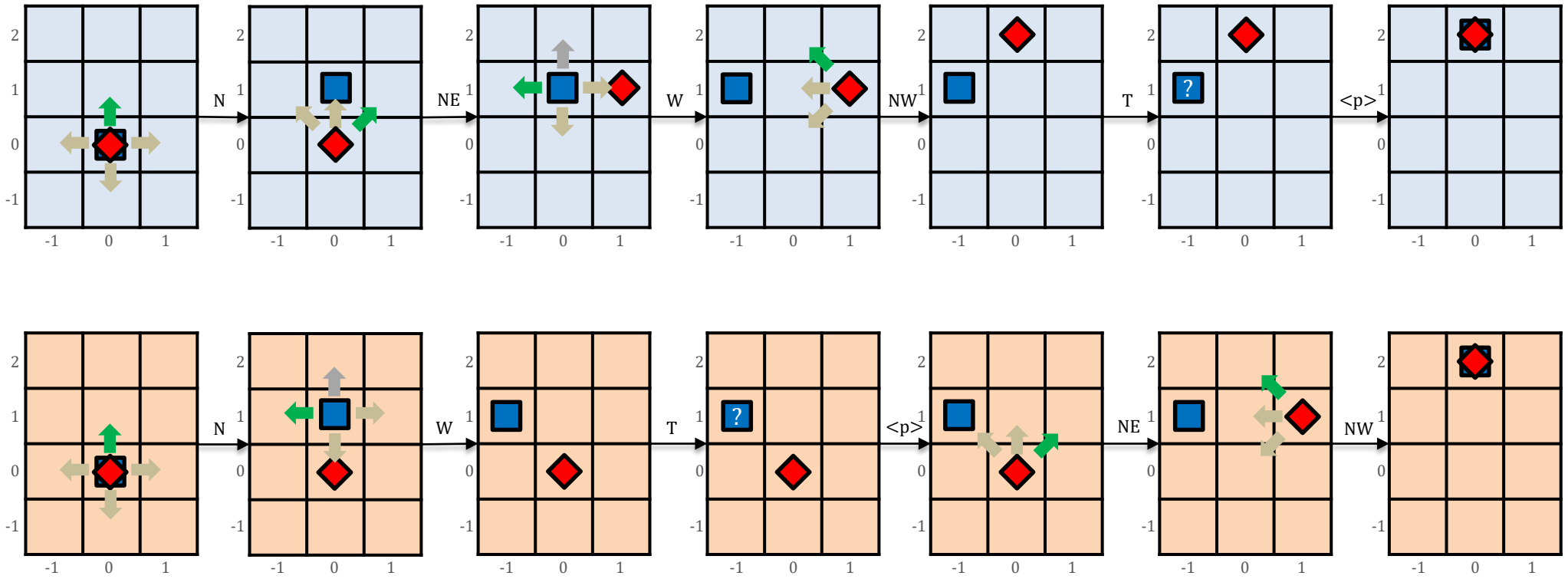


Delayed Actions Representation [CAV19]

Duke
PRATT SCHOOL of ENGINEERING

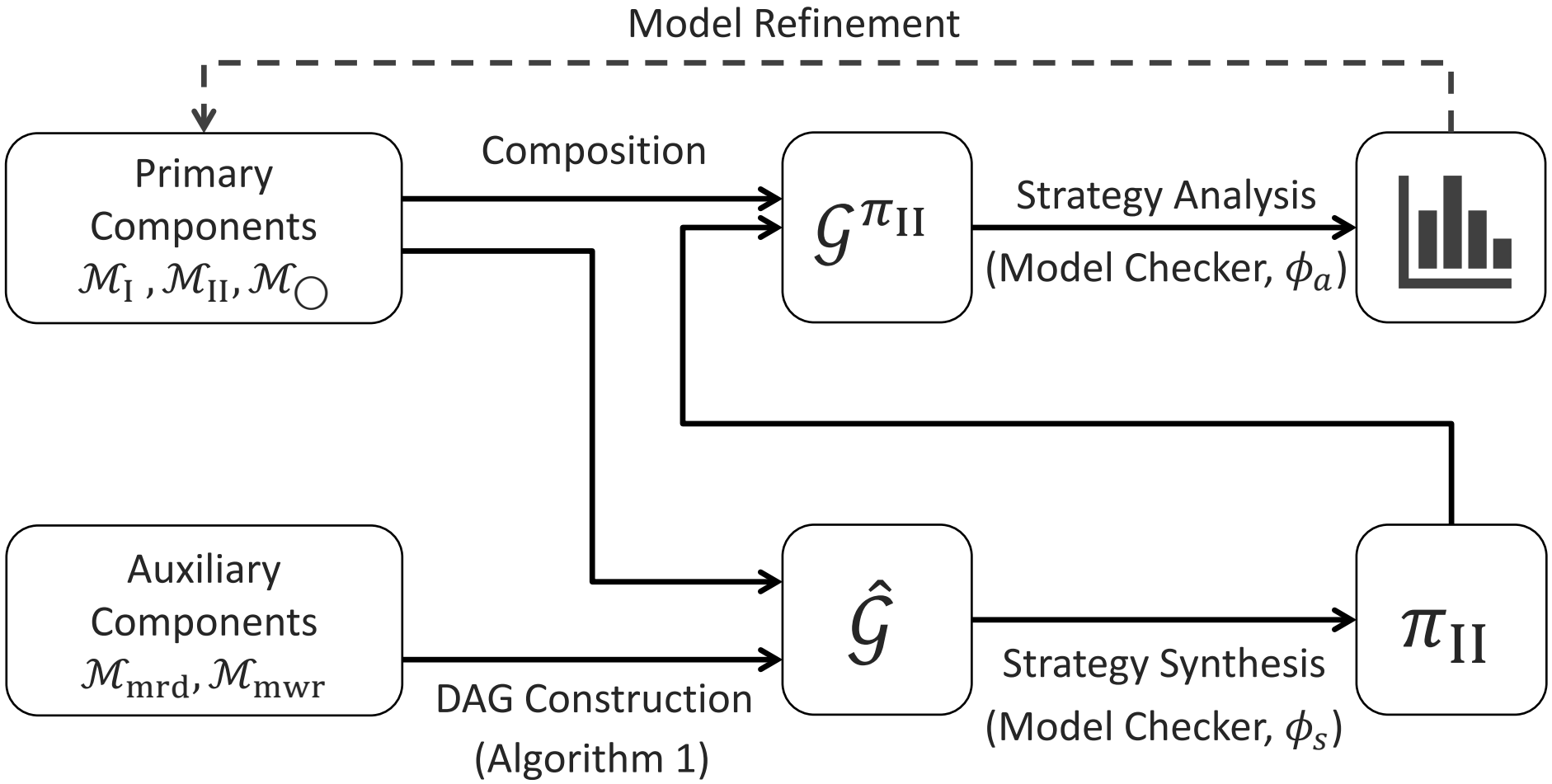


Proper simulation [CAV19]

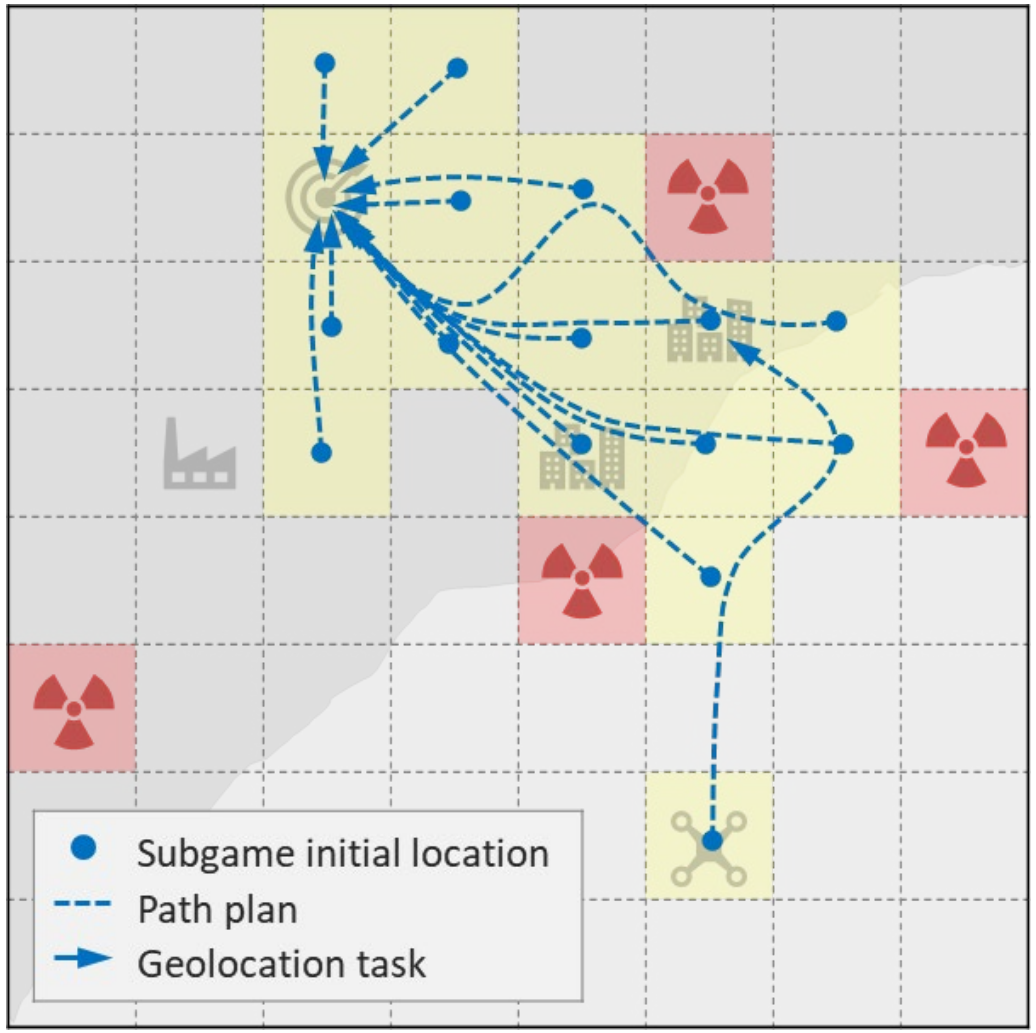


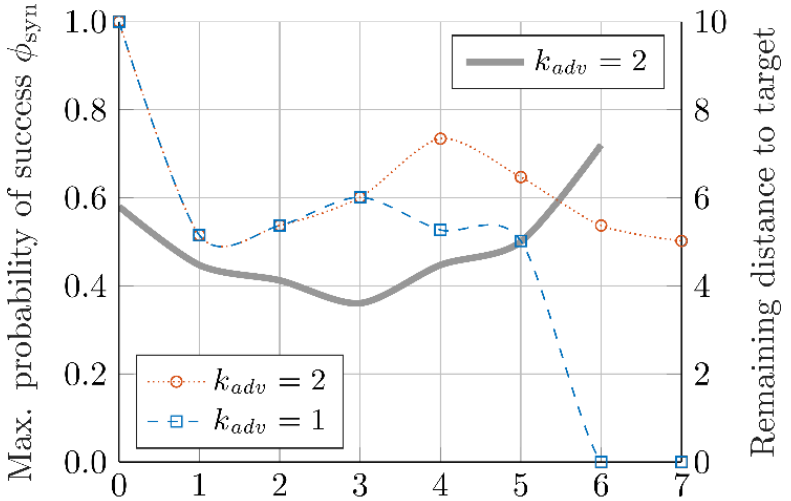
Bisimulation allows model checking and strategy synthesis using standard tools

Synthesis Framework

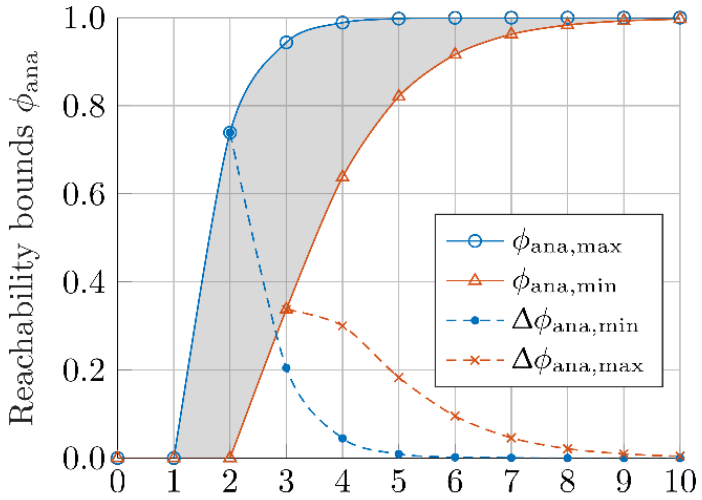


Case Study | Results





(a) Geolocation task at stage k



(b) Max. no. of geolocation tasks h

Attack-Resilient Mission Design

Duke
PRATT SCHOOL of ENGINEERING

- Develop planning methods that will improve attack-detection guarantees by allowing the deployed intrusion detection system to interact with the controller and the rest of the system

- How to model such interactions? – MDPs, PTAs, SHAs

- Optimization based on solving stochastic games
  - How to incorporate learning?
  - How to incorporate formal guarantees?

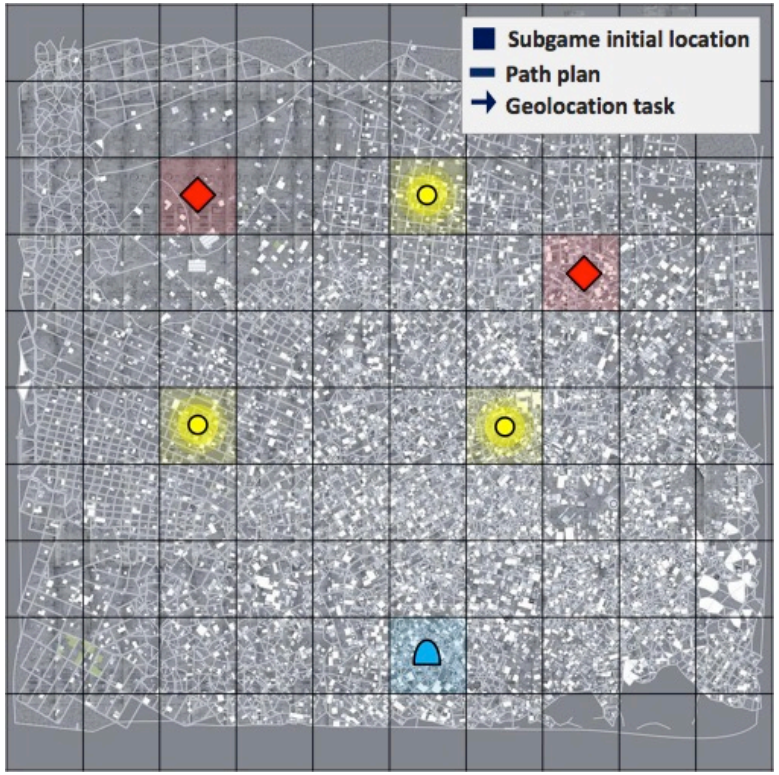
Security-aware Human-on-the-Loop Planning

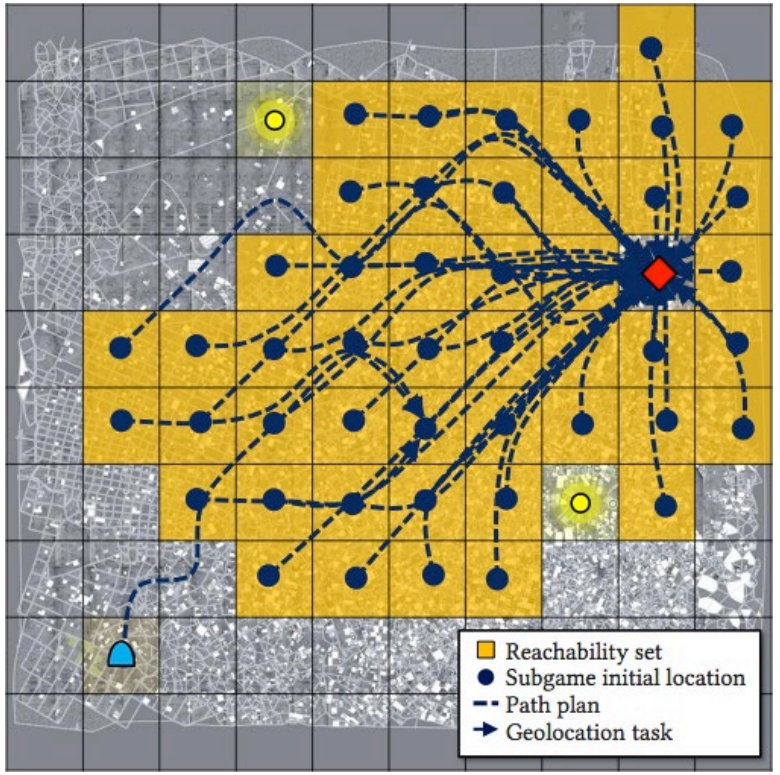
Duke
PRATT SCHOOL of ENGINEERING

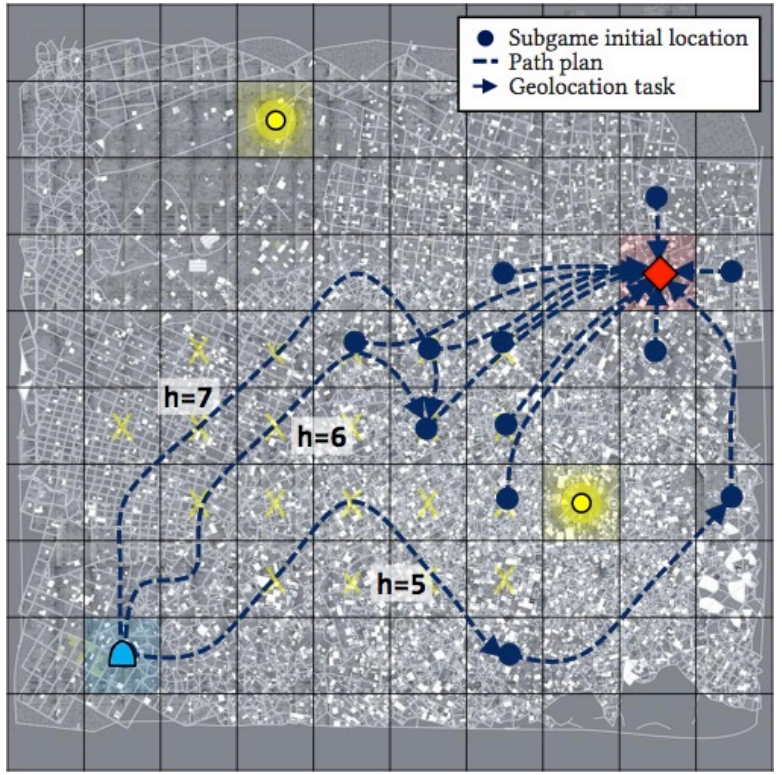
[ICRA'19,
iEEE THMS'19]

Security-aware Human-on-the-Loop Planning [ICRA'19]

Duke
PRATT SCHOOL of ENGINEERING

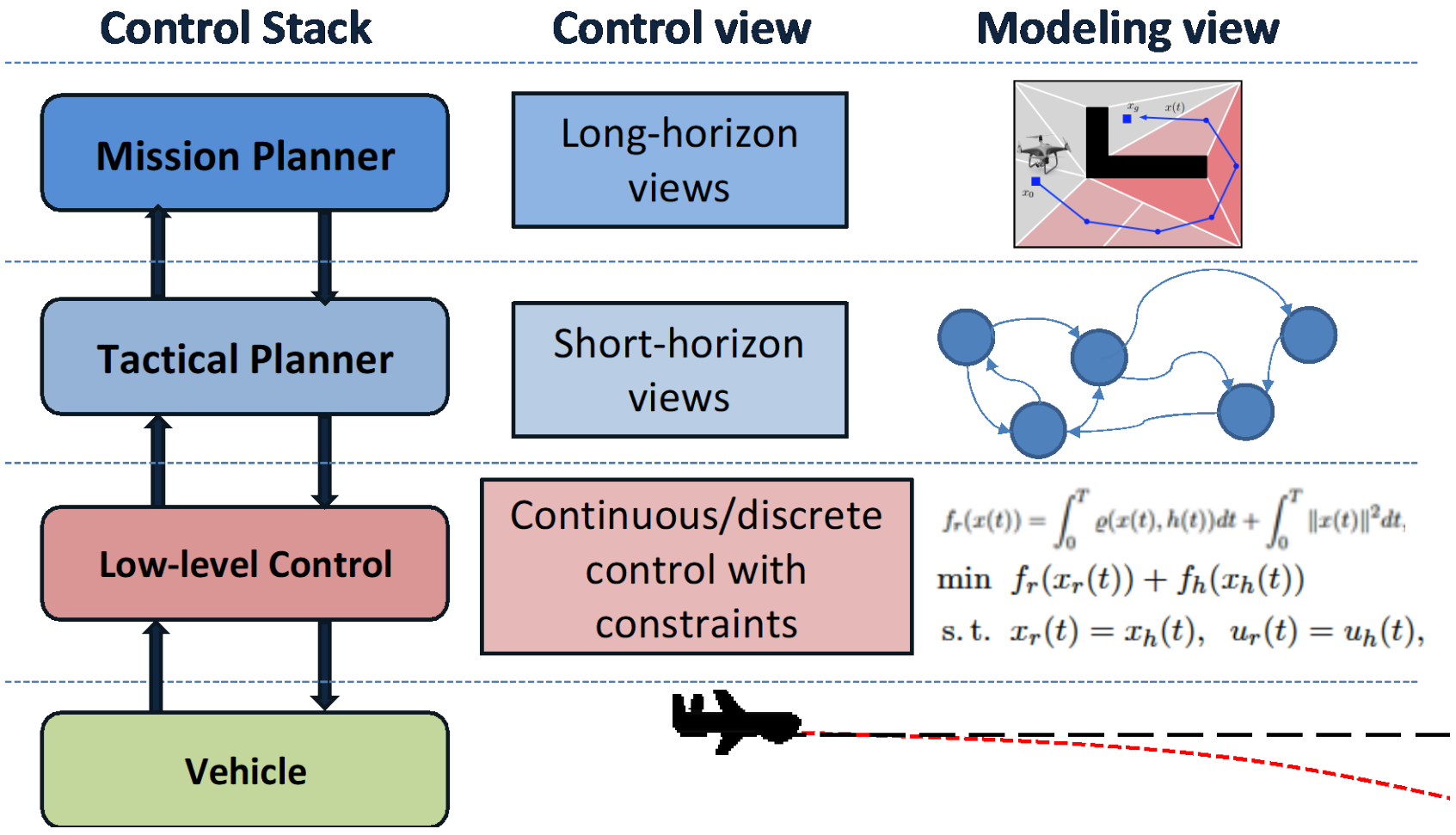






Security-Aware Control for Autonomous Systems
Are we done?

Duke
PRATT SCHOOL of ENGINEERING

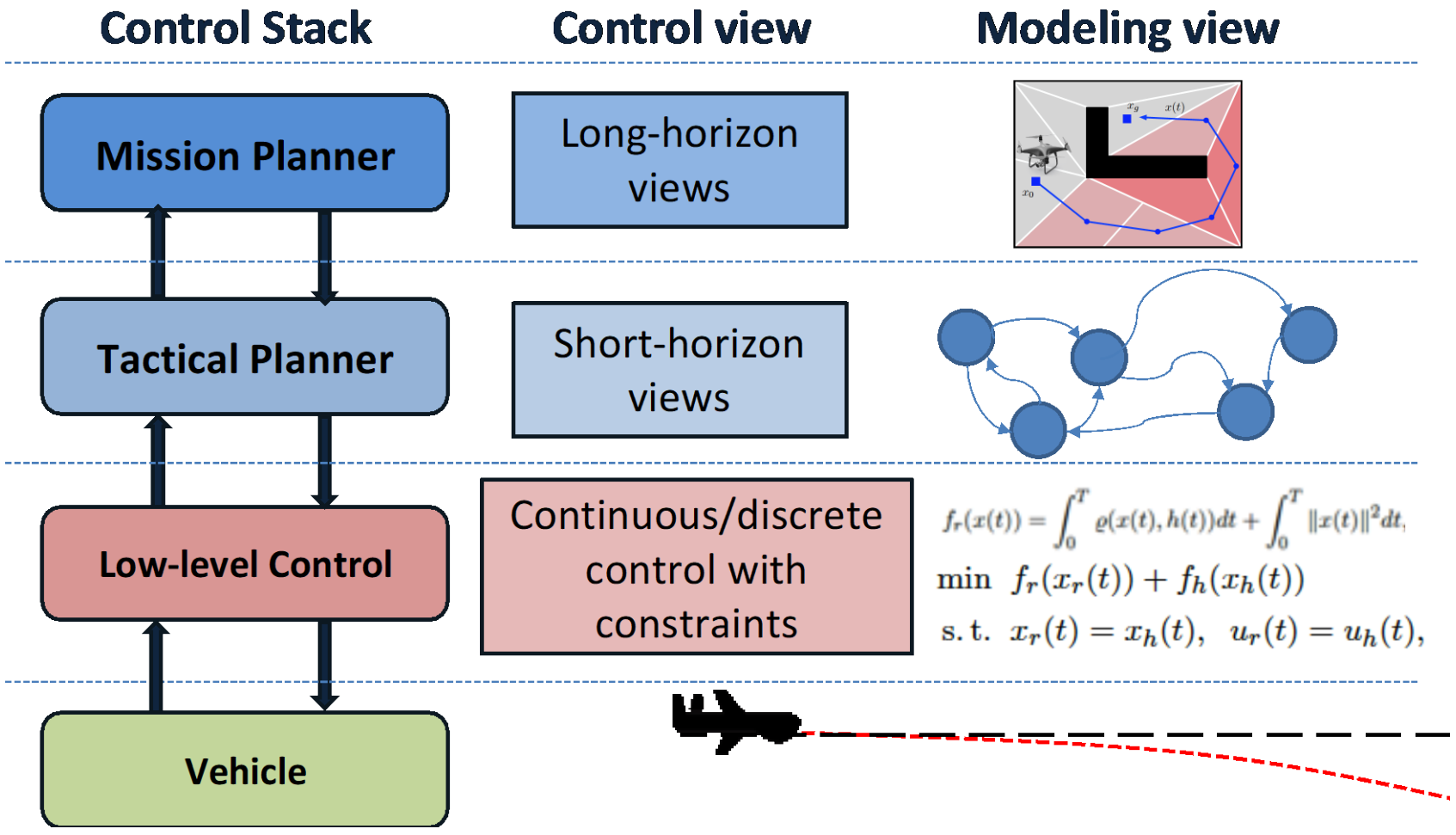


Our Goal: Add resiliency to controls across different/all levels of control stack

Security-Aware Control for Autonomous Systems

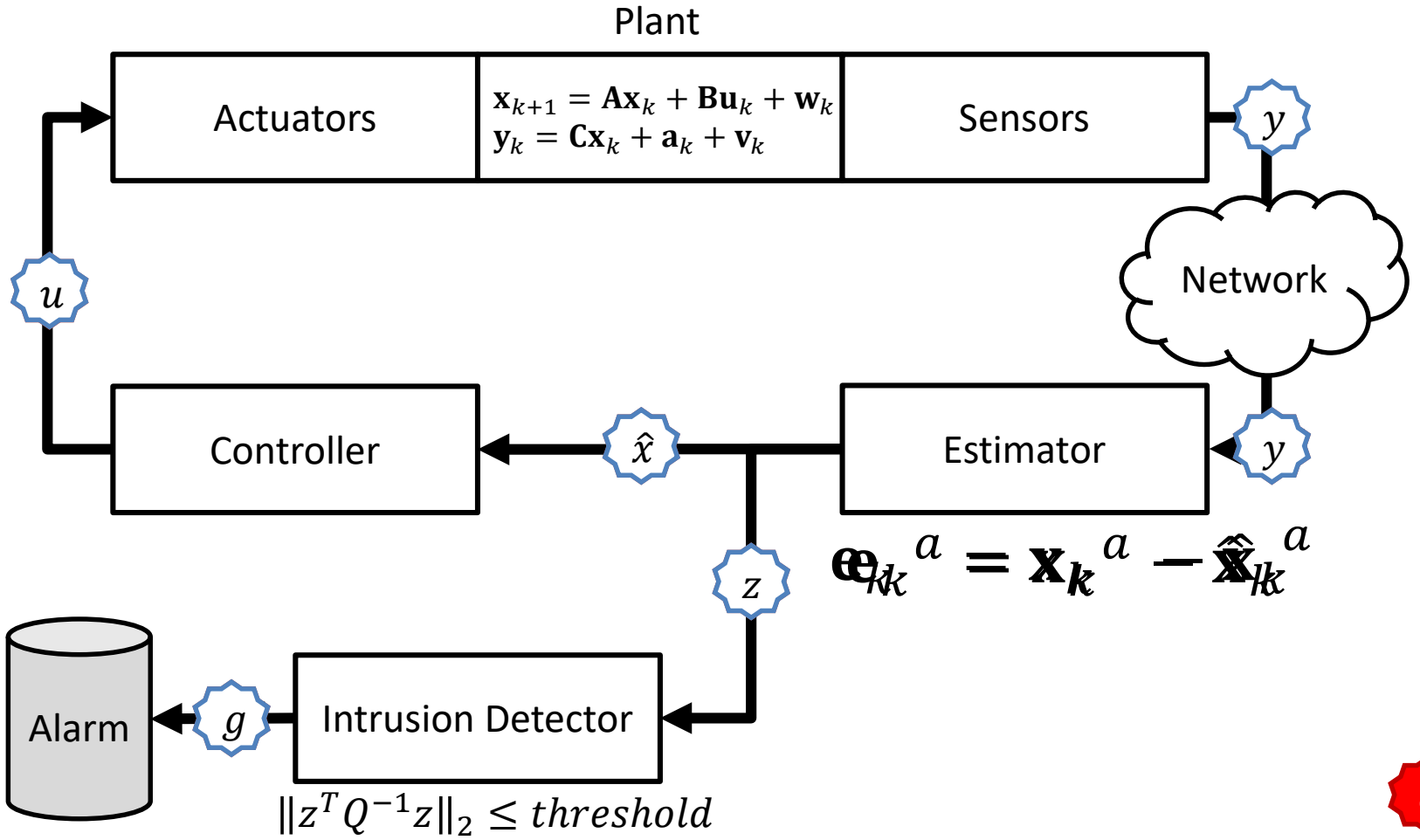
Are we done? No – conservative assumptions!

Duke
PRATT SCHOOL of ENGINEERING



Our Goal: Add resiliency to controls across different/all levels of control stack

System Model With Attacks
Especially legacy systems



Can Attacker Reach Any State?

Duke
PRATT SCHOOL of ENGINEERING

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{a}_k + \mathbf{v}_k$$

$$supp(\mathbf{a}_k) = \mathcal{K}$$
$$\mathbf{a}_{k,i} = 0, \forall i \in \mathcal{K}^C$$

Theorem 1 [1,2,3]:
A system presented above is perfectly attackable if and only if the matrix **A** is unstable, and at least one eigenvector **v** corresponding to an unstable mode satisfies $supp(\mathbf{C}\mathbf{v}) \subseteq \mathcal{K}$ and **v** is a reachable state of the dynamic system.

Physical detectors cannot always protect us from an intelligent attacker...

Can data authentication help?

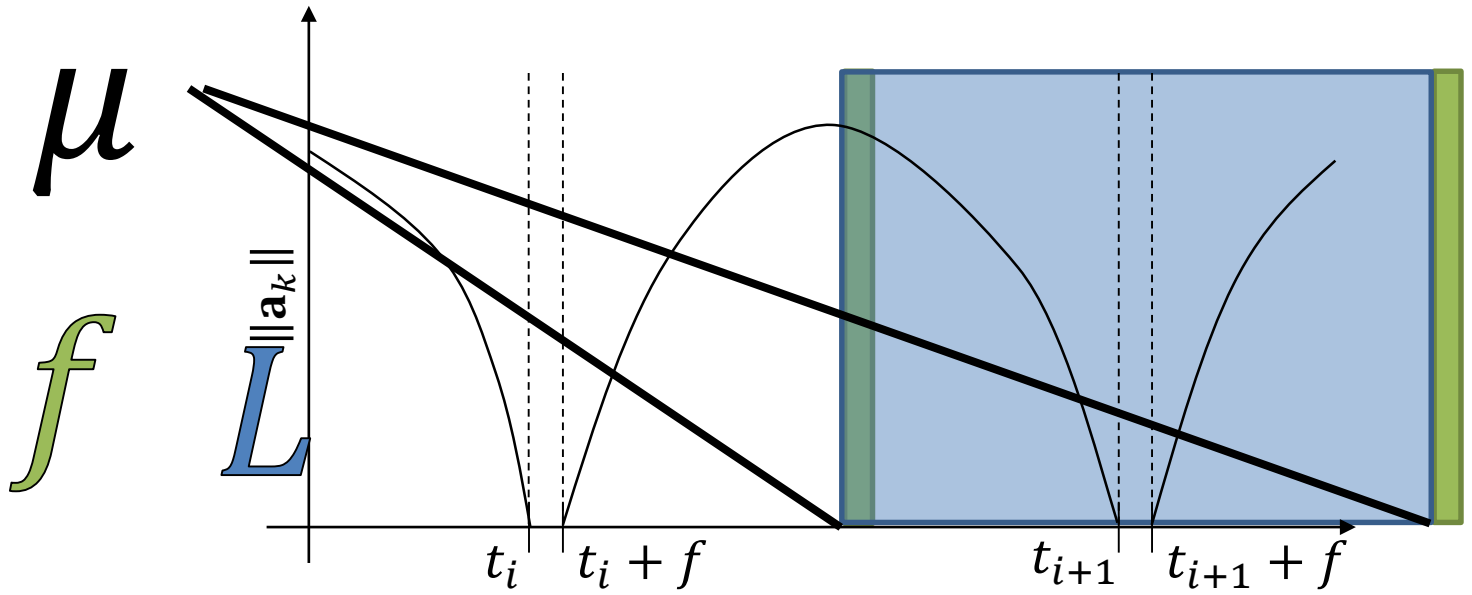
[1] Y. Mo and B. Sinopoli, "*False data injection attacks in control systems*," in First Workshop on Secure Control Systems, 2010
[2] C. Kwon, W. Liu, and I. Hwang, "*Analysis and design of stealthy cyber attacks on unmanned aerial systems",* Journal of Aerospace Information Systems, 1(8), 2014
[3] I. Jovanov and M. Pajic, "*Relaxing Integrity Requirements for Attack-Resilient Cyber-Physical Systems*", IEEE Trans. on Automatic Control, 2019

Integrity Enforcement Policy - Definition

Duke
PRATT SCHOOL of ENGINEERING

Definition 2: Intermittent data integrity enforcement policy (μ, f, L), where $\mu = \{t_k\}_{k=0}^{\infty}$, such that for all $k > 0$, $t_{k-1} < t_k$ and $L = \sup_{k>0} t_k - t_{k-1}$ ensures that

$$\mathbf{a}_{t_k} = \mathbf{a}_{t_k+1} = \cdots = \mathbf{a}_{t_k+f-1} = \mathbf{0}, \forall k \geq 0$$



Integrity Enforcement Policy

Duke
PRATT SCHOOL of ENGINEERING

Definition: Intermittent data integrity enforcement policy (μ, f, L), where $\mu = \{t_k\}_{k=0}^{\infty}$, such that for all $k > 0$, $t_{k-1} < t_k$ and $L = \sup_{k>0} t_k - t_{k-1}$ ensures that

$$\mathbf{a}_{t_k} = \mathbf{a}_{t_k+1} = \cdots = \mathbf{a}_{t_k+f-1} = \mathbf{0}, \forall k \geq 0$$

Theorem: A system Σ with a global data integrity policy (μ, f, L), where

$$f = \min(nullity(\mathbf{C}) + 1, q_{un})$$

and q_{un} is the number of distinct unstable eigenvalues of **A**, is not perfectly attackable.

[1] I. Jovanov and M. Pajic, "Sporadic Data Integrity for Secure State Estimation", IEEE Conference on Decision and Control (CDC), 2017
[2] I. Jovanov, and M. Pajic, "Secure State Estimation with Cumulative Message Authentication", IEEE Conference on Decision and Control (CDC), 2018
[3] I. Jovanov and M. Pajic, Relaxing Integrity Requirements for Resilient Control Systems, (2017). IEEE Transactions on Automatic Control, 2019

State Estimation Error
In the Presence of Stealthy Attacks

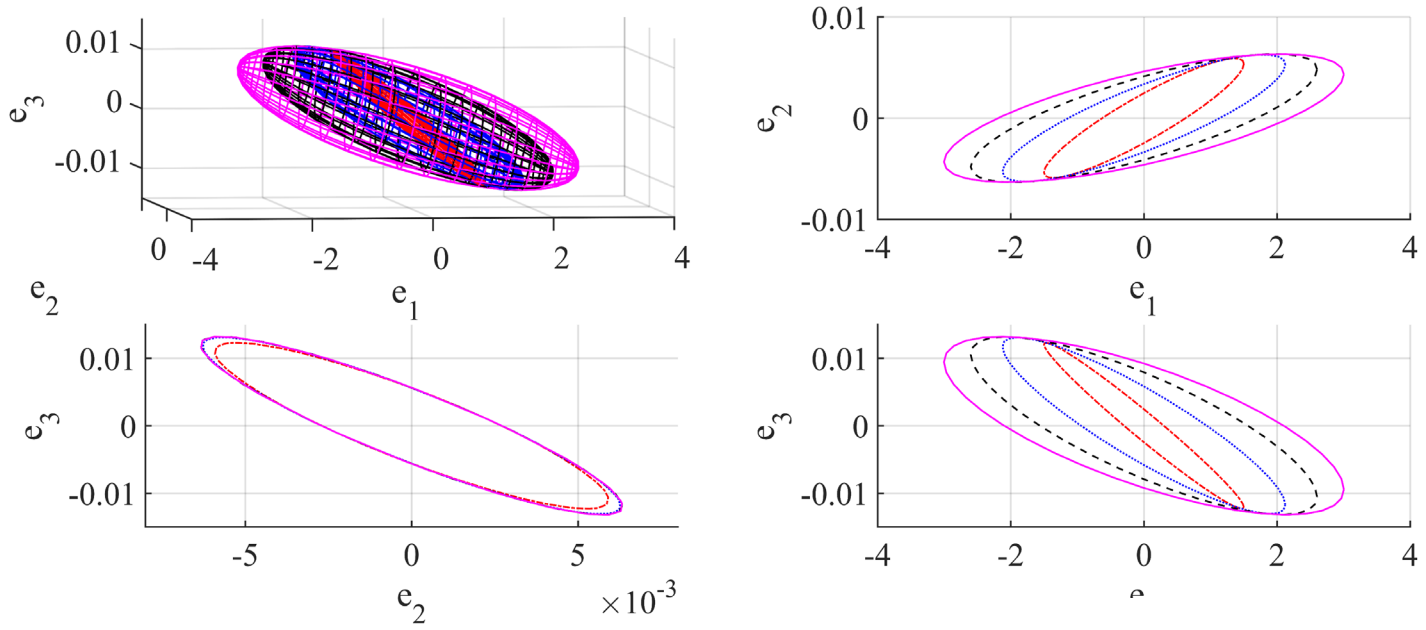
Duke
PRATT SCHOOL of ENGINEERING

Reachable region of the state estimation error under attack [1,2,3]

$$\mathcal{R}[k] = \left\{ \boldsymbol{e} \in \mathbb{R}^{\boldsymbol{n}} \middle| \begin{array}{c} \boldsymbol{e}\boldsymbol{e}^{\mathrm{T}} \preccurlyeq E\left[\boldsymbol{e}^{a}[k]\right] E\left[\boldsymbol{e}^{a}[k]\right]^{\mathrm{T}} + \gamma Cov(\boldsymbol{e}_k^{a}) \\ \boldsymbol{e}^{a}[k] = \boldsymbol{e}_k^{a}(\mathbf{a}_{1\ldots k}), \mathbf{a}_{1\ldots k} \in \mathcal{A}_k \end{array} \right\}$$

$\mathbf{a}_{1\ldots k} = [\mathbf{a}[1]^{\mathrm{T}} \ldots \mathbf{a}[k]^{\mathrm{T}}]^{\mathrm{T}}$
$\mathcal{A}_k$ is the set of all stealthy attacks

$\boldsymbol{e}_k^{a}(\mathbf{a}_{1\ldots k})$ is the estimation error evolution due to attack $\mathbf{a}_{1\ldots k}$



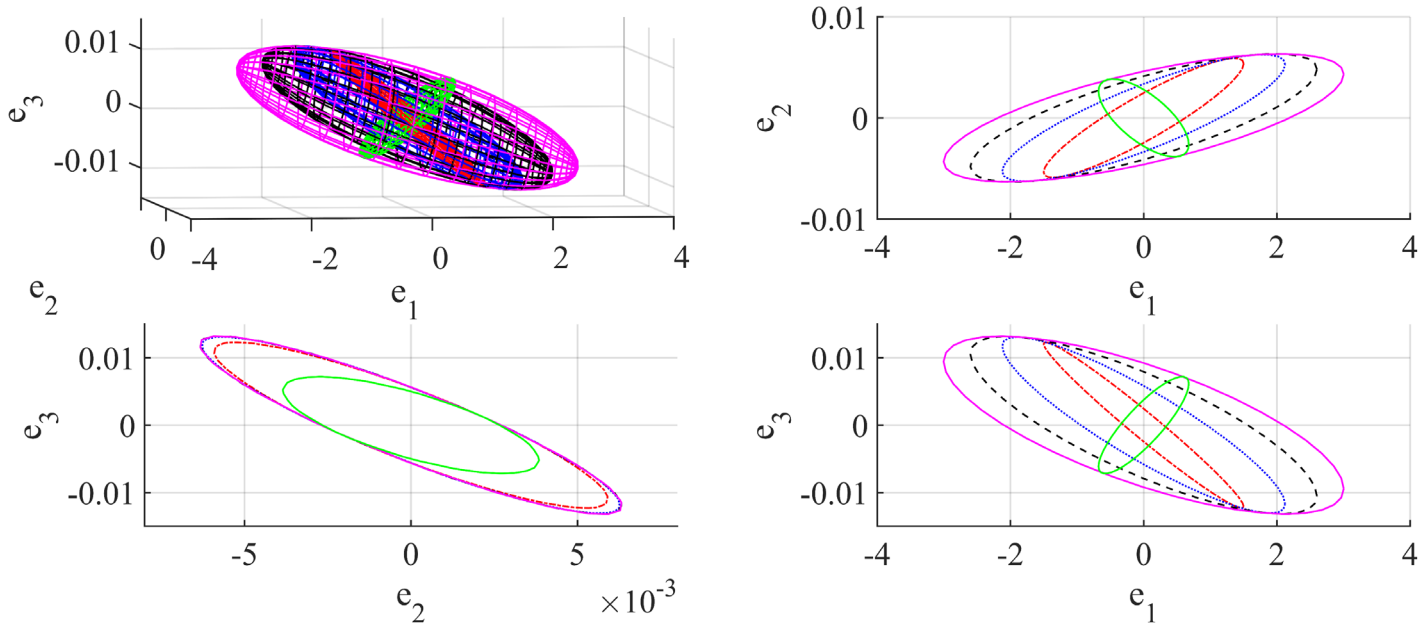
Integrity Enforcement Policy

Duke
PRATT SCHOOL of ENGINEERING

Integrity enforcement policy ensures attacker's influence is zeroed at enforcement points

Data integrity enforcement policy (μ, l) where $\mu = \{t_k\}_{k=0}^{\infty}$, with $t_{k-1} < t_k, \forall k > 0$
and $l = \sup_{k>0} t_k - t_{k-1}$ ensures that $\mathbf{a}_{1\ldots k} = 0, \forall k \geq 0$

This means that at points of authentication $\boldsymbol{y}_i^{net,a}[k] = \boldsymbol{y}_i^{a}[k]$

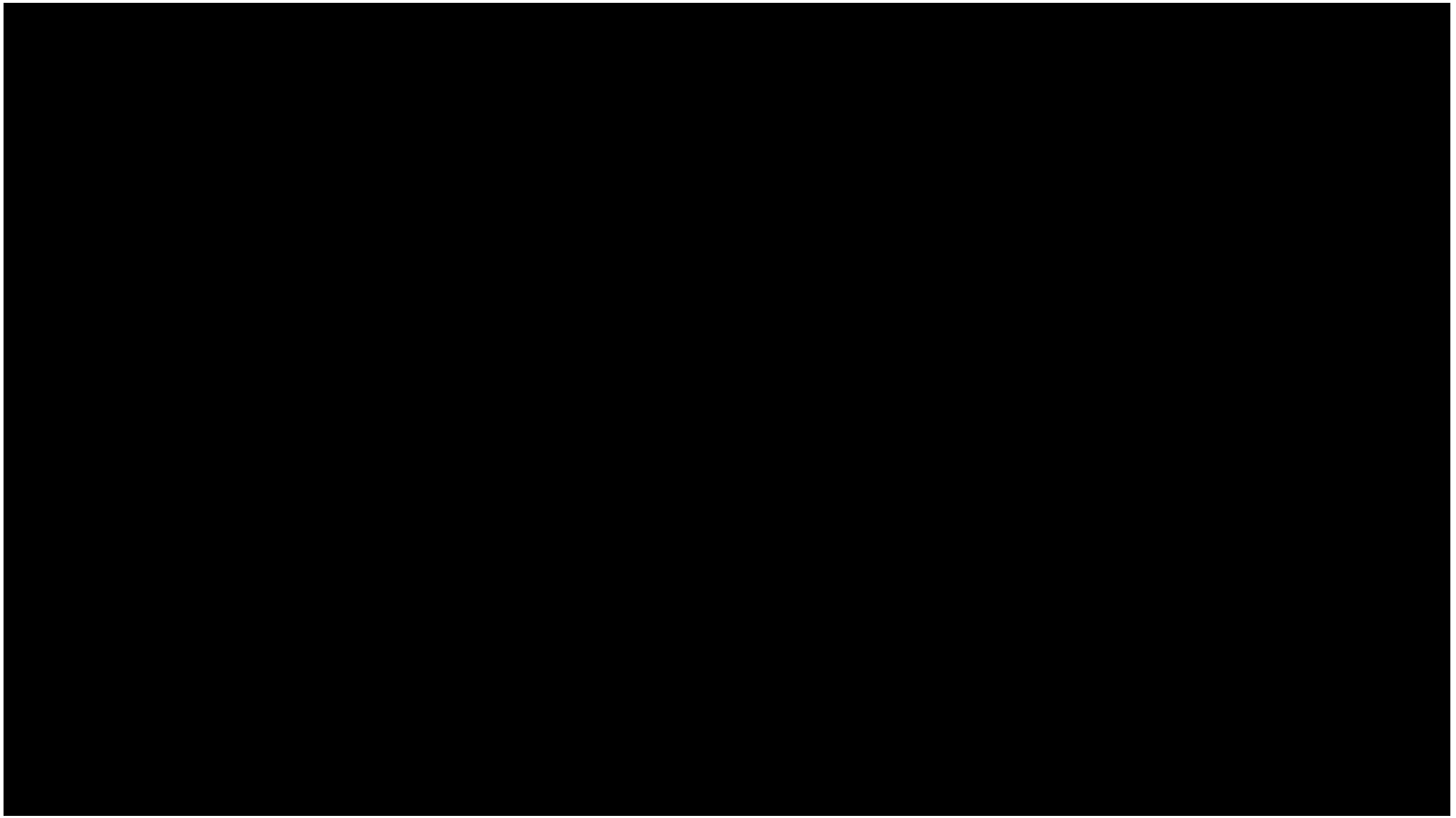


Non-Secure Vehicle Platooning

Duke
PRATT SCHOOL of ENGINEERING

Secure Vehicle Platooning
With Intermittent Integrity Guarantees

Duke
PRATT SCHOOL of ENGINEERING



System Performance Metric

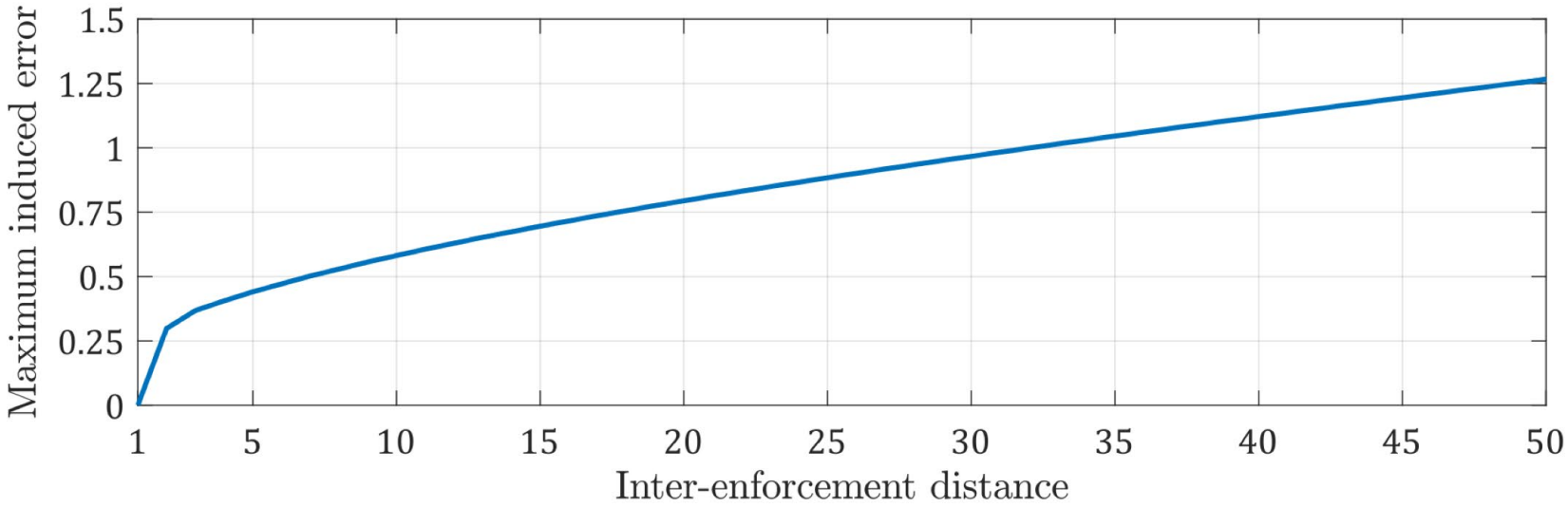
Quality-of-Control (QoC) under Attack

Duke
PRATT SCHOOL of ENGINEERING

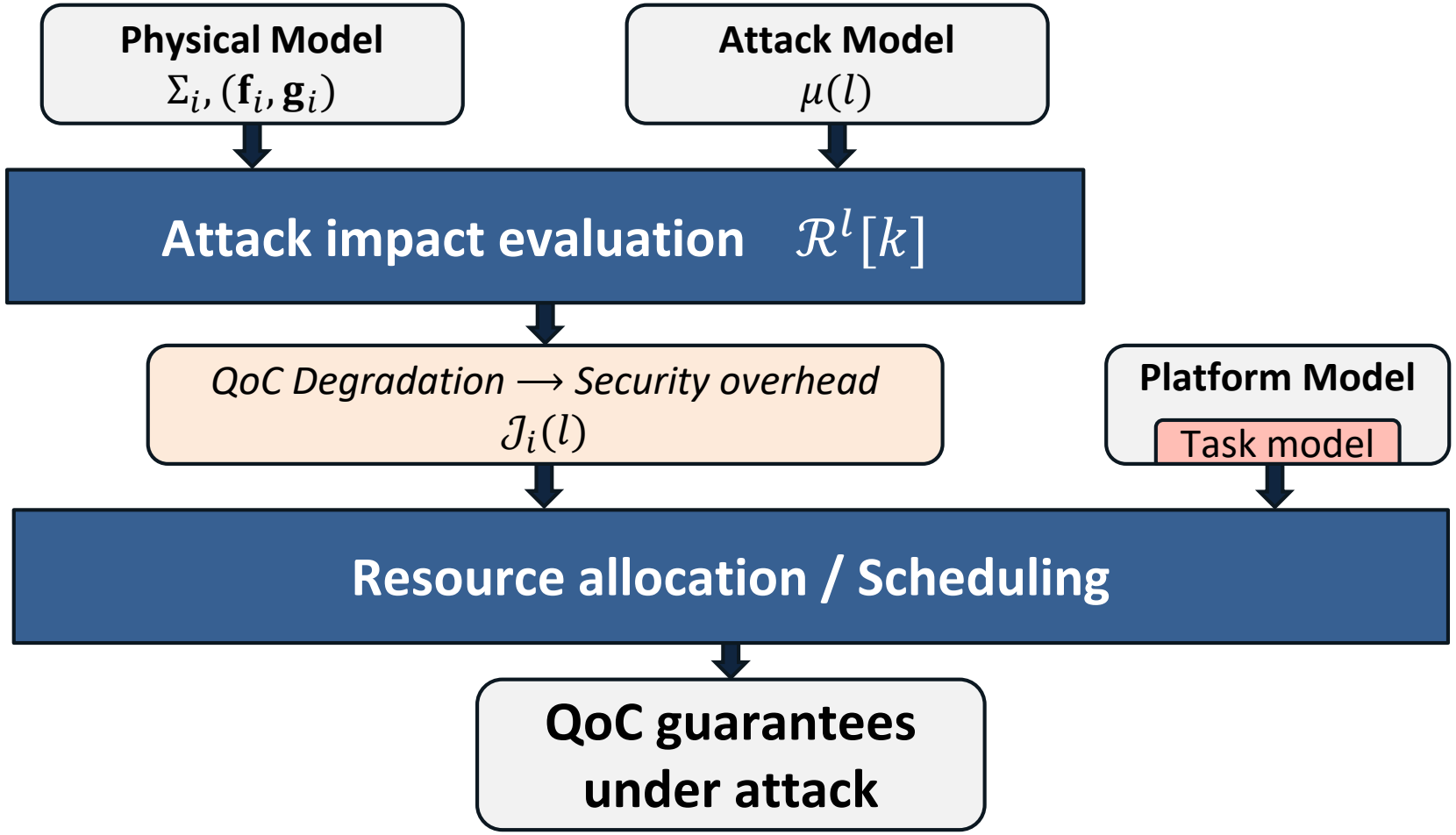
Evolution of the state-estimation error due to attack is a sound performance metric

$$\mathcal{J}(l) = sup\{\|\boldsymbol{e}^a\|_2 \big| \boldsymbol{e}^a \in \mathcal{R}^l\} \qquad \mathcal{R}^l = \bigcup_{k=0}^{\infty} \mathcal{R}^l[k]$$

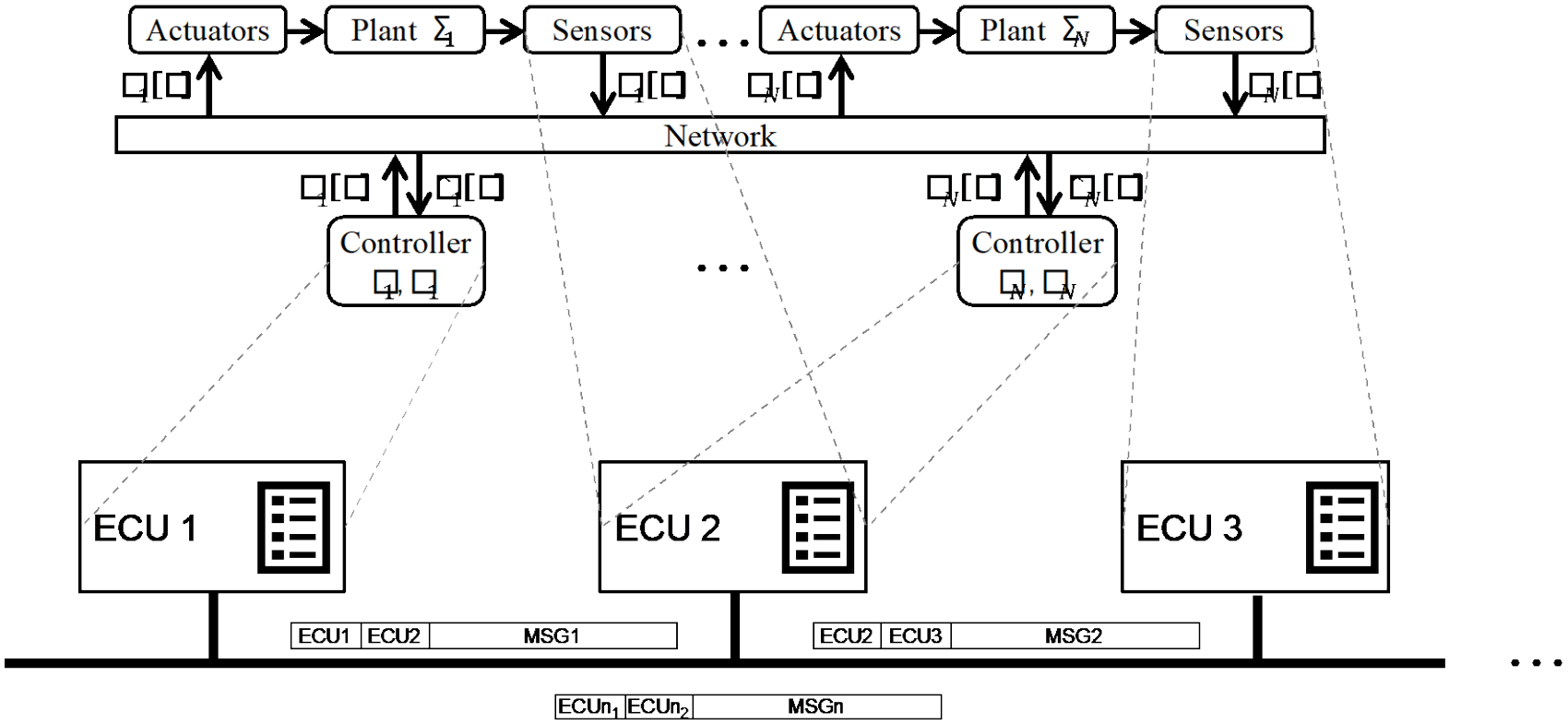
where $\mathcal{R}^l[k]$ denotes $\mathcal{R}[k]$ computed over all integrity enforcement policies with parameter l



Security-Aware Design Framework



Three Scheduling Problems



- ECUs are the bottleneck [ACM TECS/EMSOFT 17] – **Best Paper Award!**
- Network is the bottleneck [RTSS 17]
- Both ECU time and network bandwidth are a concern – ACM TCPS*

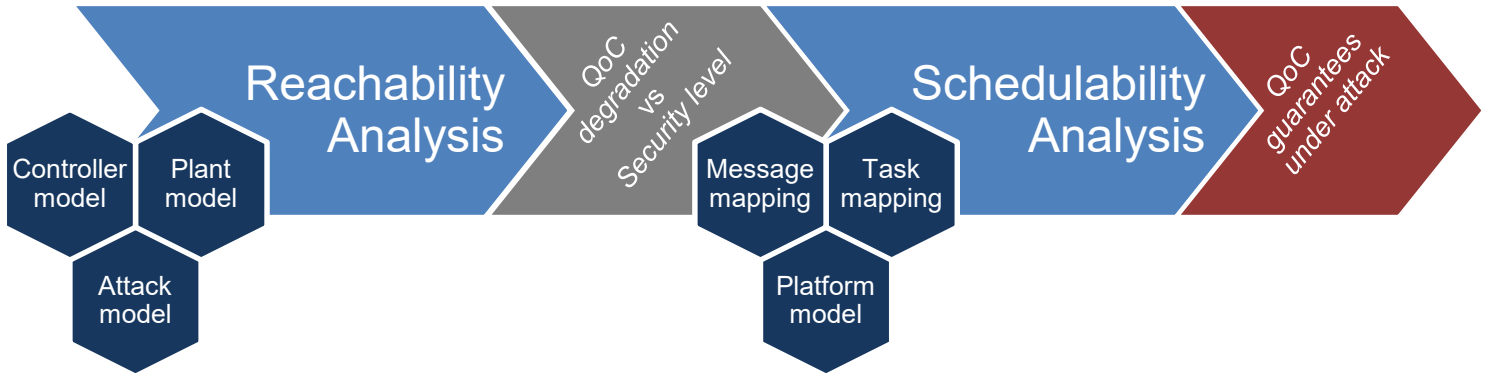
Resiliency with Intermittent Data Integrity & Availability

Duke
PRATT SCHOOL of ENGINEERING

- Goal: Derive distributed, event-based resilient control/estimation for cases when only intermittent communication between agents can be achieved

- When such communication should occur, how often, and between which agents?
- How we can opportunistically use available communication links to increase mission resilience against attacks.
- These requirements will be used as part of the design specifications for intermittent wireless communication in RT3

- Rich security models
  - Cumulative authentication
  - Forgery attacks
  - Probabilistic guarantees



Thank you


UF
UNIVERSITY of
FLORIDA

SPACE VEHICLES DIRECTORATE
TECHNOLOGY FOR THE FINAL FRONTIER

Duke
UNIVERSITY

AFRL SENSORS DIRECTORATE

TEXAS
The University of Texas at Austin

MUNITIONS DIRECTORATE

UC SANTA CRUZ