Secure Operations of Trustworthy Autonomous Systems

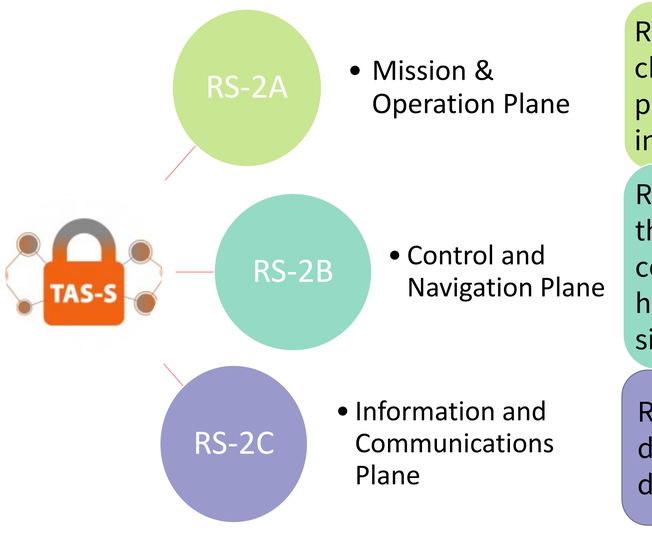
Cranfield University, Lancaster University

Investigators: Weisi Guo, Gokhan Inalhan, Plamen Angelov, Antonios Tsourdos, Vasileios Giotsas, David Hutchison
Research Fellows: Zhuangkun Wei, Oscar Villarreal, Burak Yuksek, Pierre Ciholas

Autonomous systems face numerous challenges in their operation, due to the uncertain and dynamic multi-layer attack surfaces

TAS-RS2 aims to solve following challenges

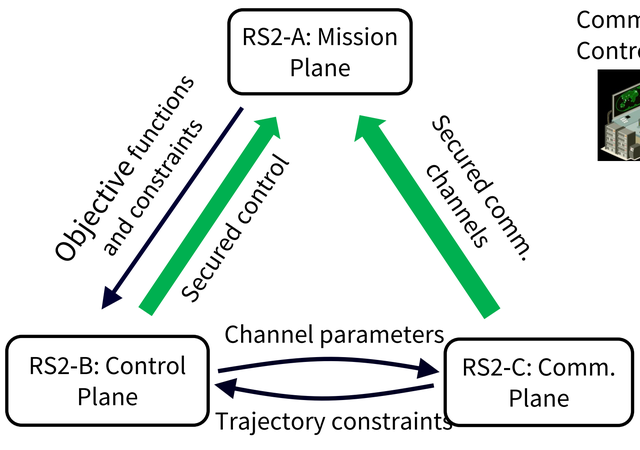
- Modelling & addressing potential attacks in discrete mission, control and communication layers
- Study & address hybrid cascaded cross-layer threats in the dynamic AS space

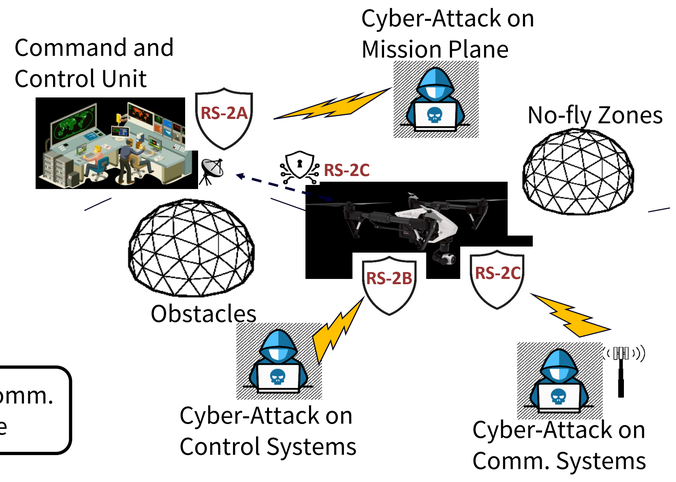


RS-2A: Exposure to cyber-physical attacks by characterizing the attack surfaces, i.e., entry points and likelihoods across the mission surface in a technology & mission-invariant manner.

RS-2B: Provide quantifiable safety and feedback to the mission surface when the limits of secure controllability are compromised within a time horizon under current policies and adversarial situations.

RS-2C: Provide secure communications across the different layers in the informatics plane from detection of signals to networking.





Mission Control for Secure Trustworthy Autonomous Systems requires flexible but reliable real-time optimal decision making and monitoring to handle a wide range of attacks

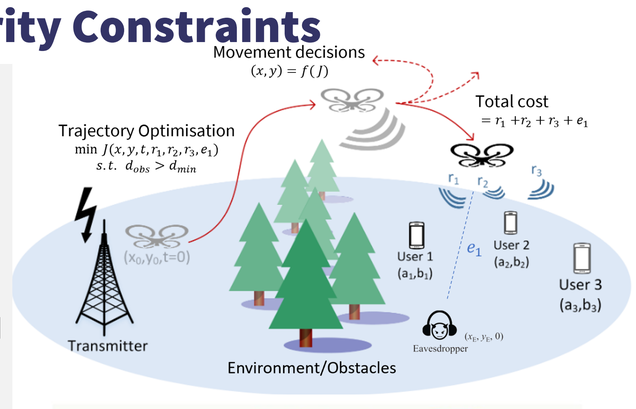
Critical Tasks:

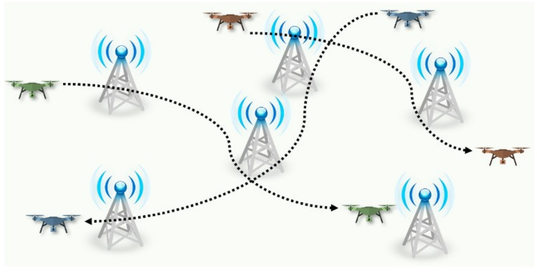
- Handling Communication Errors and Security Constraints
- Assessing Control Faults and Performance Limitations
- Handling Environmental Limitations (Uncertainty/Dynamic Obstacles/No-Fly zones)
- Avoiding and Handling Electronic/Electro-magnetically induced attacks
- Achieving Deterministic/Real-Time Performance for the Optimal Decision Making
- Handling and Detecting Security Threats under Learning-based Scenarios

Key Challenges for Trustworthy Learning-based Mission Control under Security Constraints

Methods and Focus:

- Real-Time Non-Convex Trajectory Optimisation for Path Planning under Uncertainty, Power Consumption, Dynamic Obstacle Avoidance and Communication Security Constraints
- Adaptive and Fault-Tolerant Learning-based Design for Mission Control to improve reliability of safety critical systems
- Supervisory Control for Anomaly Detection and Isolation Systems
- Intelligent Resource Allocation for Multi-UAV Design under Security Threats
- Reliable Self-Assessment under Learning-based Scenarios





Autonomous Systems rely on the ability to conduct run time adaptations of control decisions over attacks or "perceived" attacks:

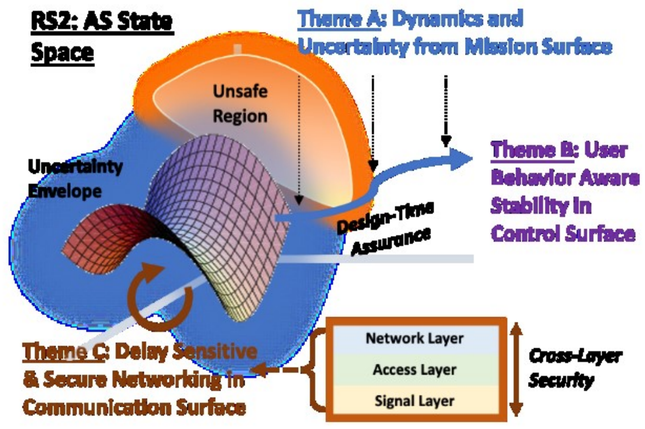
- Adversaries
- Environment uncertainties
- Degraded performance

How to do this in a "trustworthy" fashion?

- Safe, Secure and Reliable

Attack Definitions

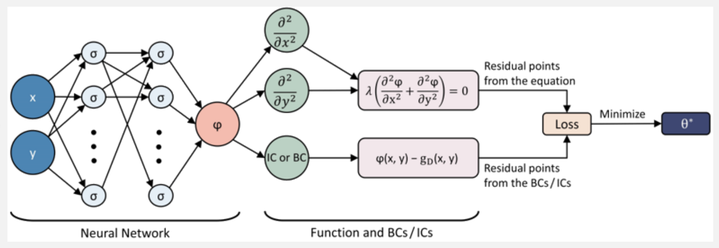
- Sensing and Communication Errors
- Loss of an actuator
- Environmental conditions
- Electronic attacks
- Electromagnetic deception
- Injecting false pattern into data



Key Solution Cornerstones in Learning-Enabled Context

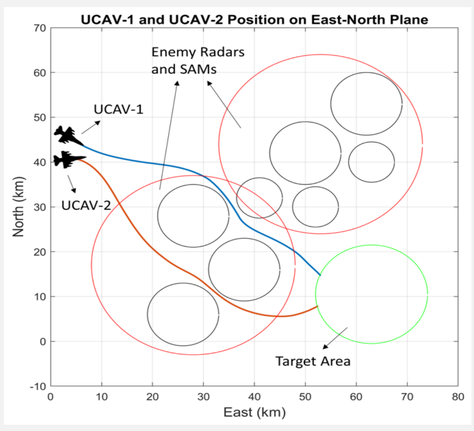
Interpretability => Explainable and Trustworthy AI

- Physics Informed Deep Learning
- Ability to identify system behaviour
- Generalization capability
- Anomaly detection/classification



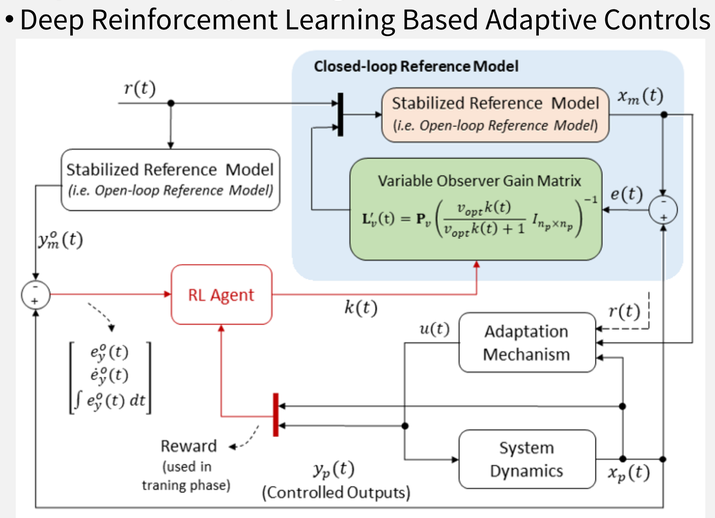
Continual Assurance

- Detect and avoid
- Learning enabled context



Adaptive Security Strategies

- Deep Reinforcement Learning Based Adaptive Controls



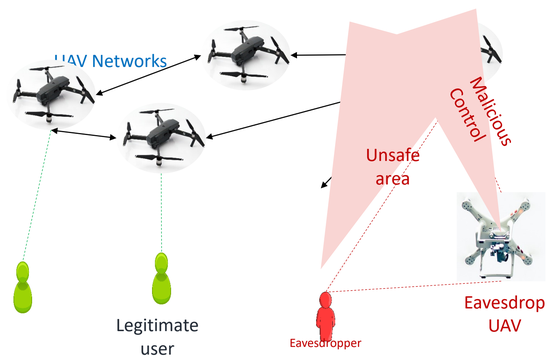
Secure and real-time communications serve as the fundamentals for Autonomous Systems to achieve reliable control and mission delivery.

Attack vectors:

- Key intercept
- Active interference (jamming)
- Passive eavesdropping
- Erode secrecy rate

Traditional Cryptography method:

- Complex key generation/distribution
- High computational complexity
- High latency
- No secrecy guaranteed by brute force



Physical Layer Security: using RECIPRO radio environment

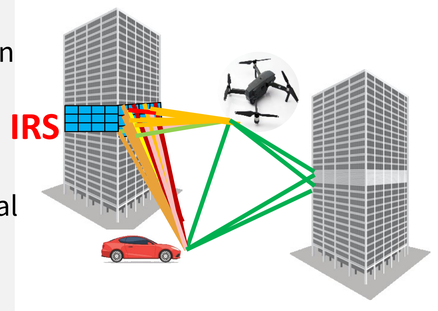
Innovation: exploit unique, dynamic, recipro CSIs between entities due to radio propagation nature

Advantages: low latency & complexity, using only physical channel properties:

- randomness of wireless channel
- Superiority of legitimate over wiretap channels

Steps:

- Generate randomness using intelligent reflecting surface (IRS)
- Channel Probing between legitimate users
- Generate cipher keys using recipro CSIs

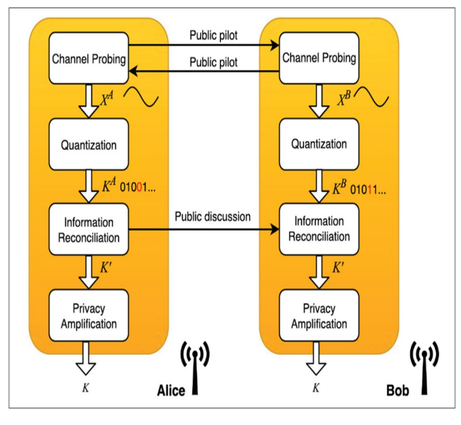


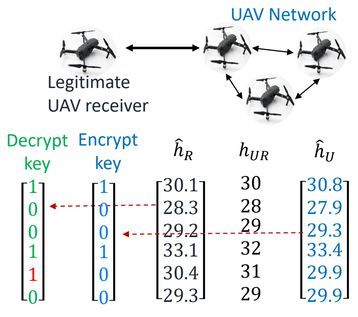
IRS randomly control phase

Recipro CSI $\begin{cases}\hat{h}_A = h_{AB} + \mathbf{h}_{rB}^H \mathbf{W} \mathbf{h}_{Ar} + n_A \\ \hat{h}_B = h_{BA} + \mathbf{h}_{rA}^H \mathbf{W} \mathbf{h}_{Br} + n_B\end{cases}$

Eve estimated CSI $\hat{h}_E = h_{EA} + \mathbf{h}_{rE}^H \mathbf{W} \mathbf{h}_{Ar} + n_E$

$\mathbb{E}[\hat{h}_A] = \mathbb{E}[\hat{h}_B] \neq \mathbb{E}[\hat{h}_E]$





This work is supported, in part, by the Engineering and Physical Sciences Research Council [grant number: EP/V026763/1]