Butterfly Attack: Adversarial Manipulation of Temporal Properties of Cyber-Physical Systems

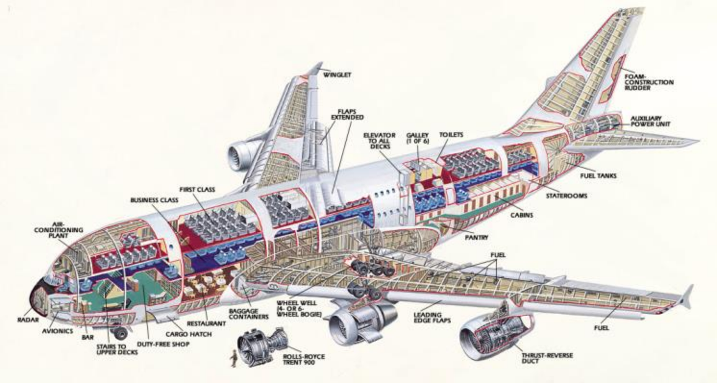
Rouhollah Mahfouzi[1], Amir Aminifar[2], Soheil Samii[1,3], Mathias Payer[2], Petru Eles[1], Zebo Peng[1]

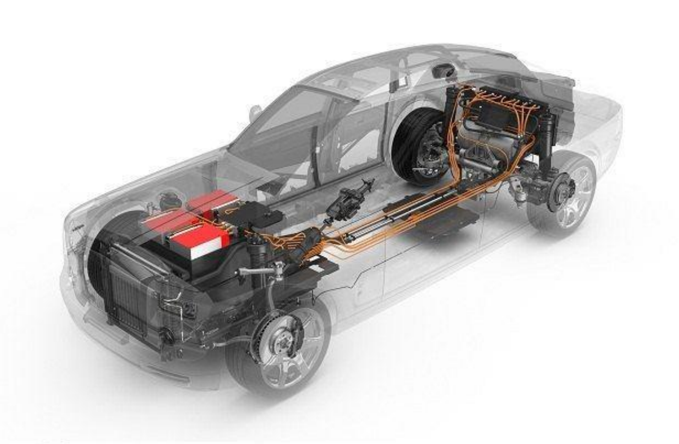
1) Linköping University, Sweden

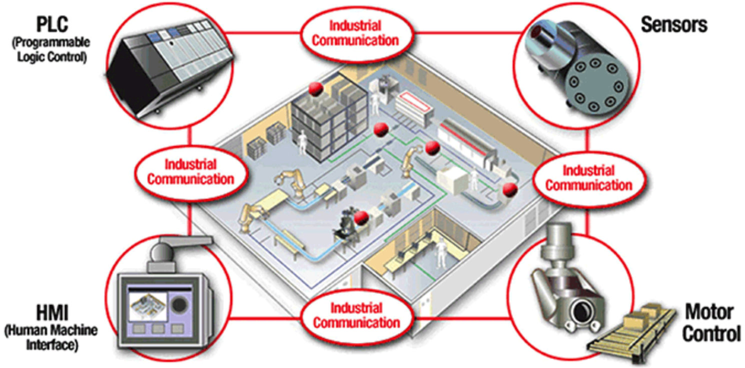
2) École Polytechnique Fédérale de Lausanne, Switzerland

3) General Motors R&D, US

Cyber-Physical Systems

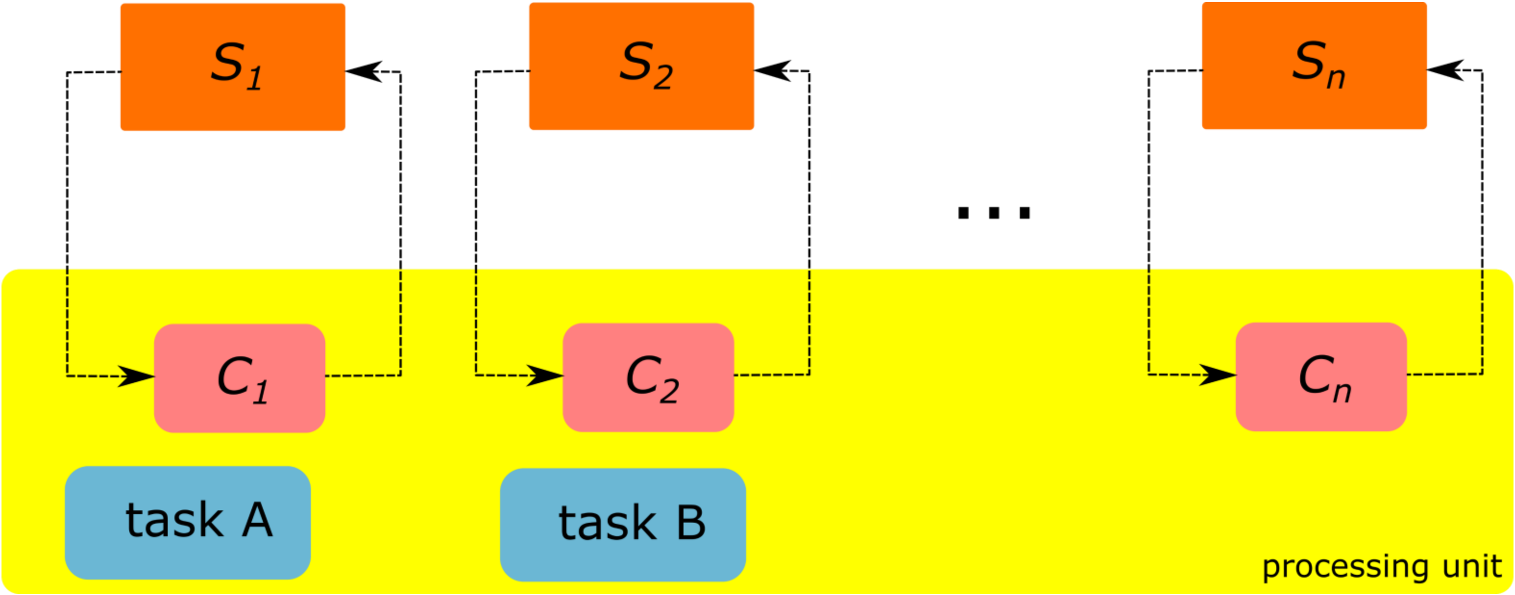






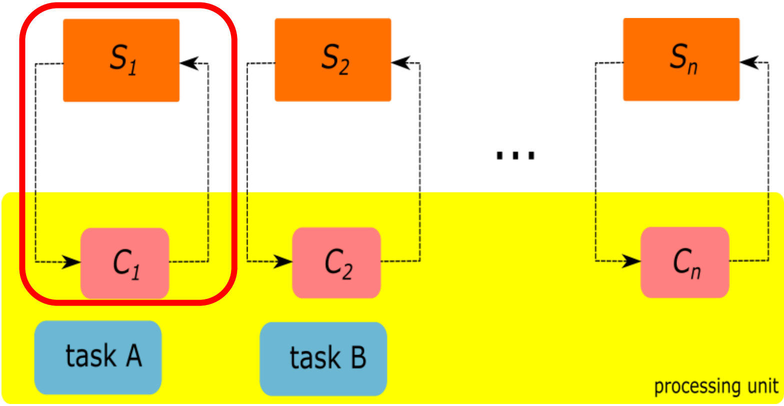
Images from: Electricity Forum, ACS solutions, eeNews

Shared Processor



- ❖ **Shared platform**
- ❖ **Mixed-critical tasks**
- ❖ **Control applications**

Latency, Jitter

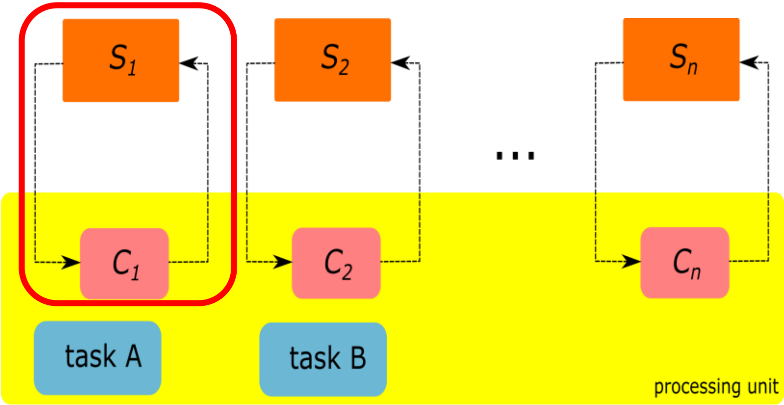


Latency, Jitter

$h =$ sampling period

$L = R^b$ (Latency)

$J = R^w - R^b$ (Jitter)

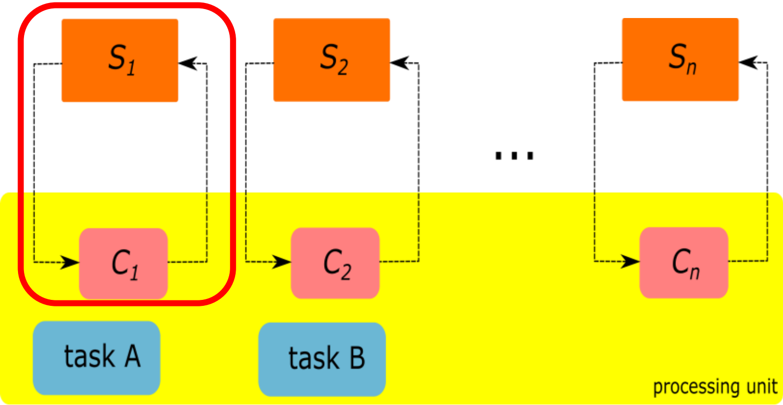


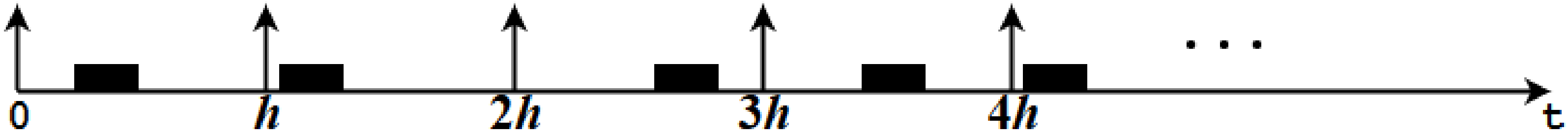
Latency, Jitter

$h =$ sampling period

$L = R^b$ (Latency)

$J = R^w - R^b$ (Jitter)



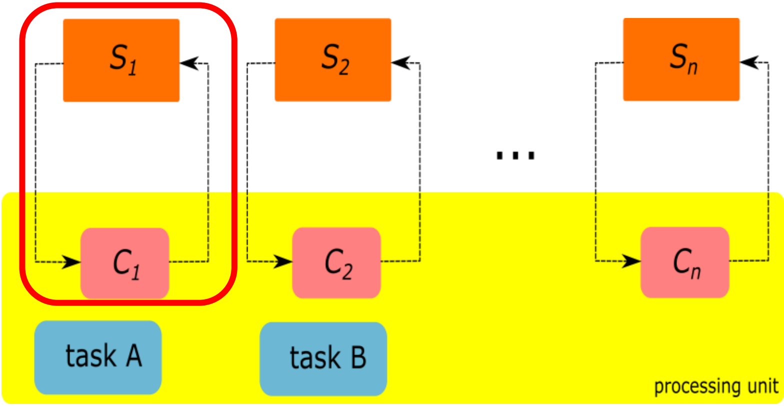


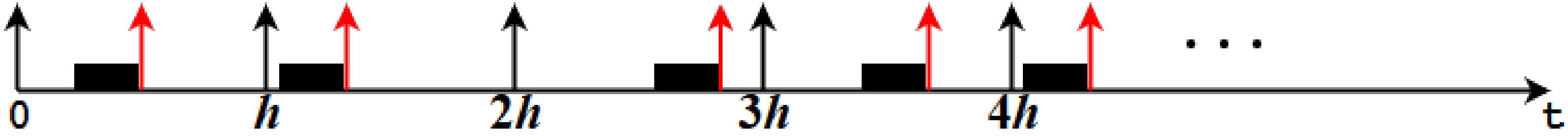
Latency, Jitter

$h =$ sampling period

$L = R^b$ (Latency)

$J = R^w - R^b$ (Jitter)



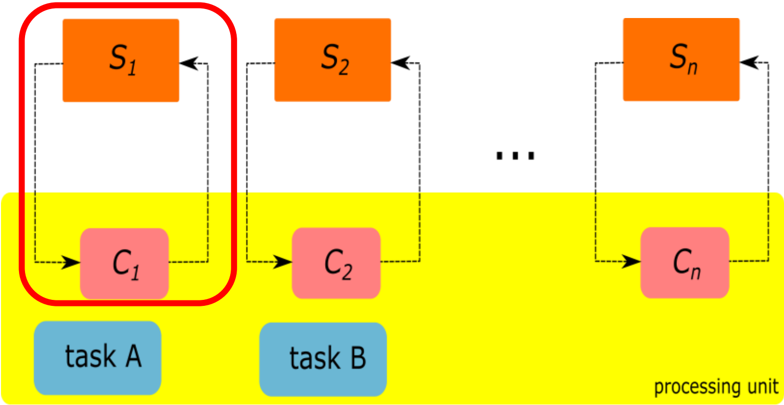


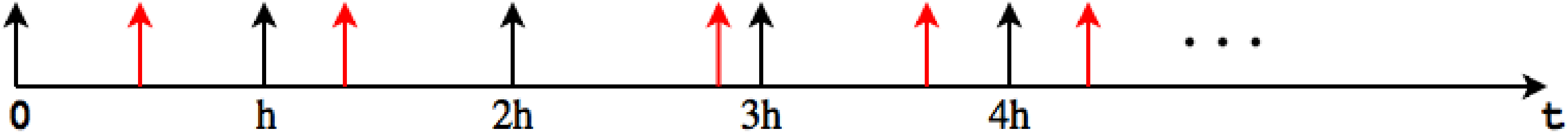
Latency, Jitter

$h = \text{sampling period}$

$L = R^b$ (Latency)

$J = R^w - R^b$ (Jitter)



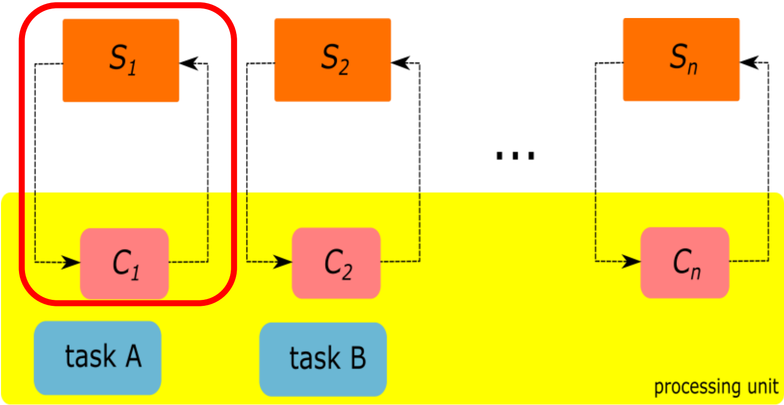


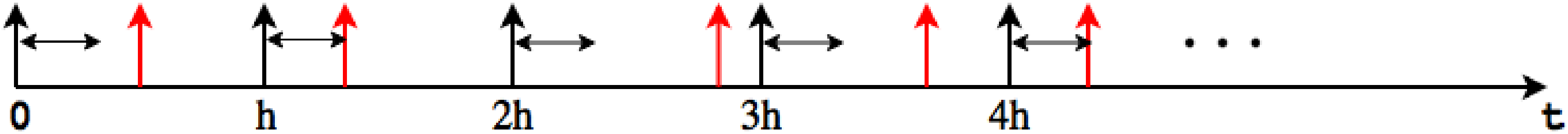
Latency, Jitter

$h = \text{sampling period}$

$L = R^b$ (Latency)

$J = R^w - R^b$ (Jitter)



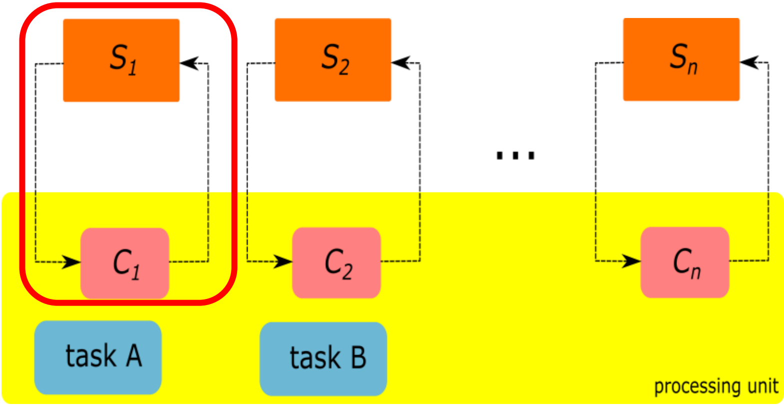


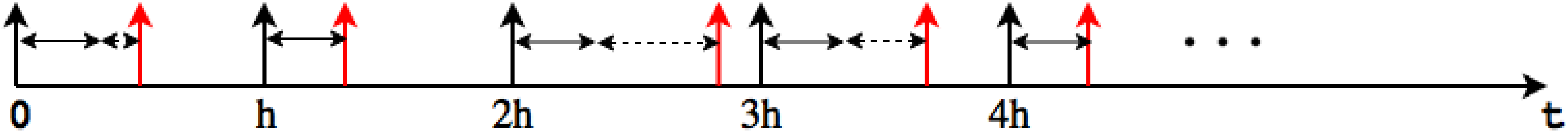
Latency, Jitter

$h =$ sampling period

$L = R^b$ (Latency)

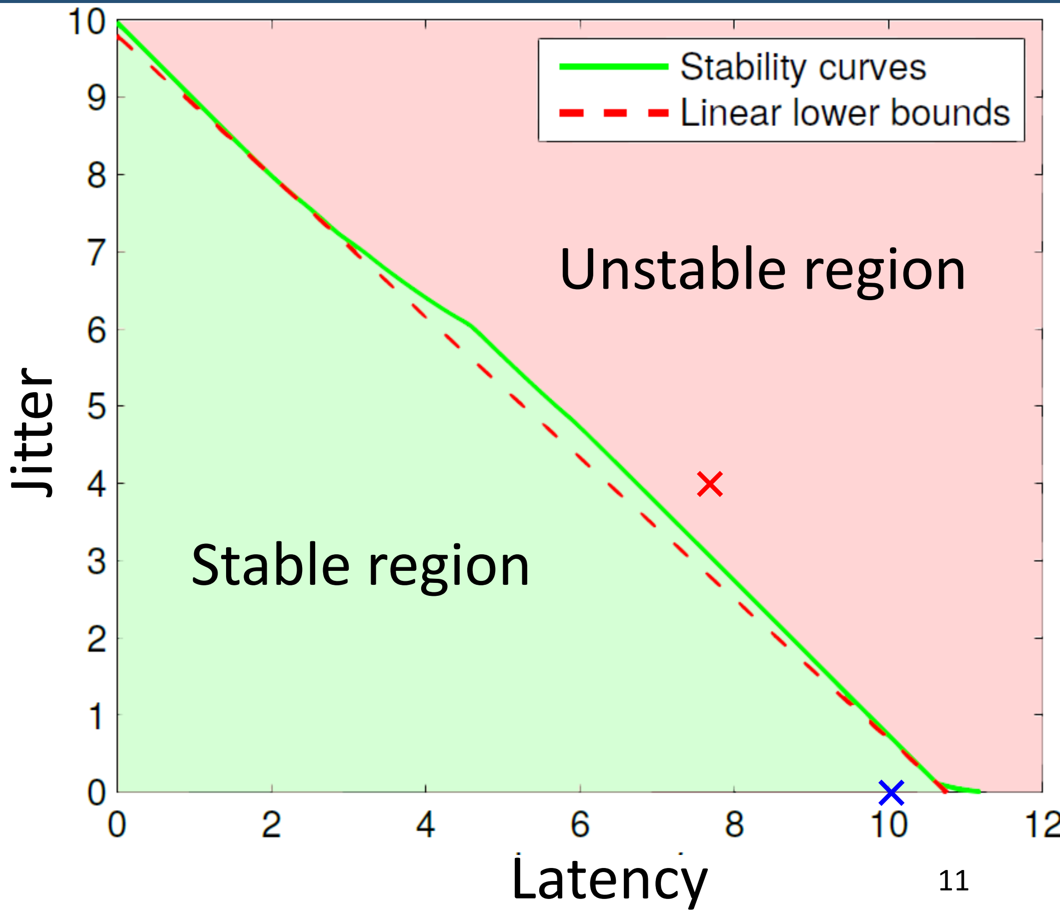
$J = R^w - R^b$ (Jitter)



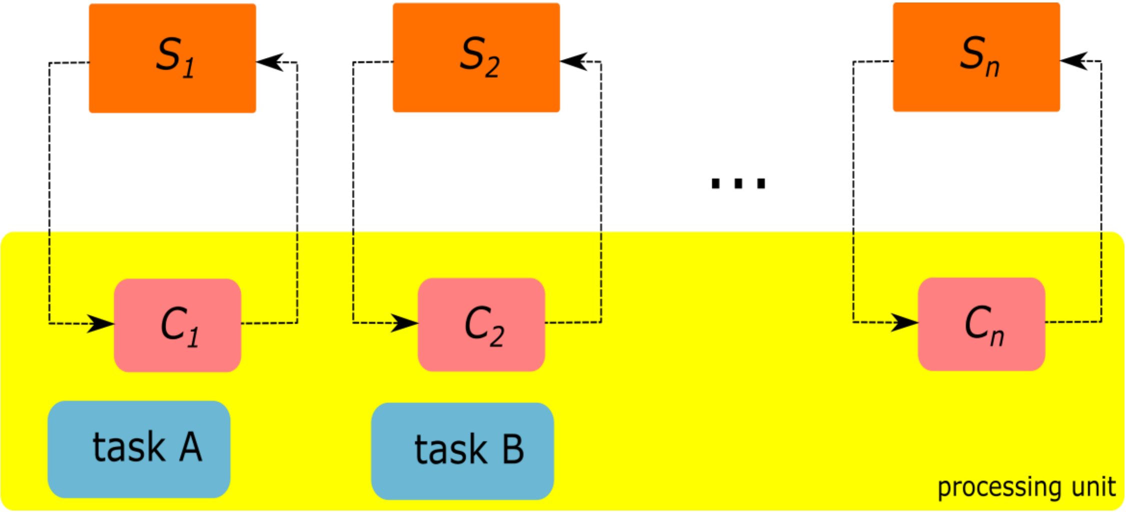


Stability

Stability curve derived with Jitter Margin toolbox



Common Belief*

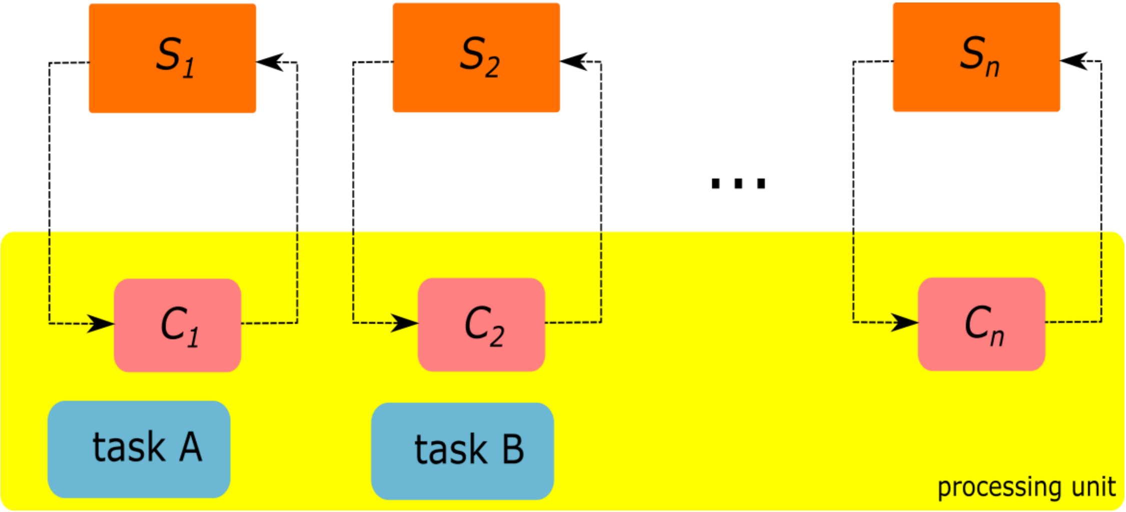


- Main thread: Exceeding a certain worst-case execution/computation time
- Giving more resource to a controller leads to better control quality
- Security measures: prevent tasks from consuming more resources

*Vestal S. Preemptive scheduling of multi-criticality systems with varying degrees of execution time assurance. In28th IEEE International Real-Time Systems Symposium (RTSS 2007)

Common Belief*

Our Contribution!



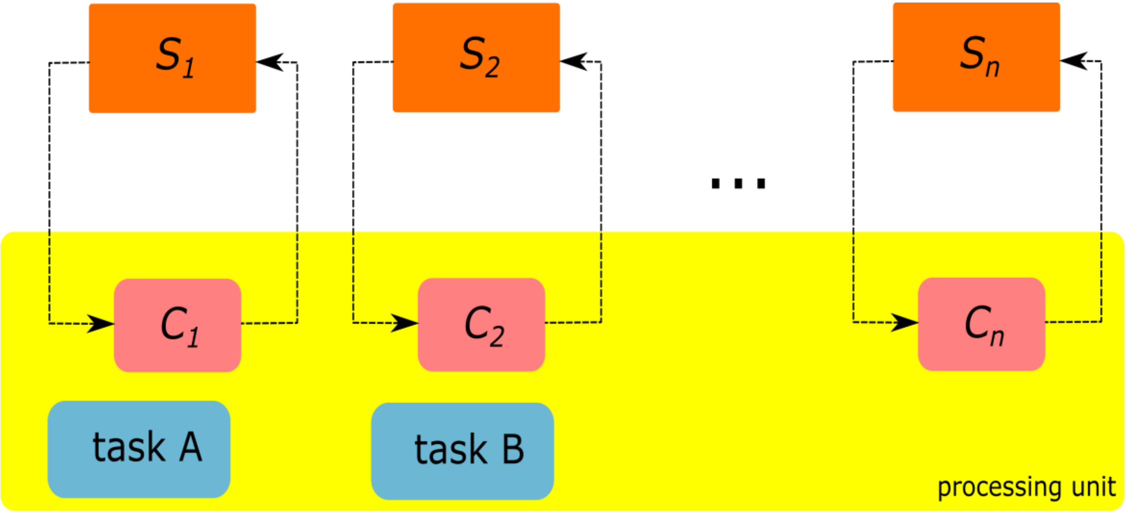
- ❖ **Main thread: Exceeding a certain worst-case execution/computation time**
- ❖ **Giving more resource to a controller leads to better control quality**
- ❖ **Security measures: prevent tasks from consuming more resources**

*Vestal S. Preemptive scheduling of multi-criticality systems with varying degrees of execution time assurance. In28th IEEE International Real-Time Systems Symposium (RTSS 2007)

Common Belief*

Our Contribution!

Wrong!

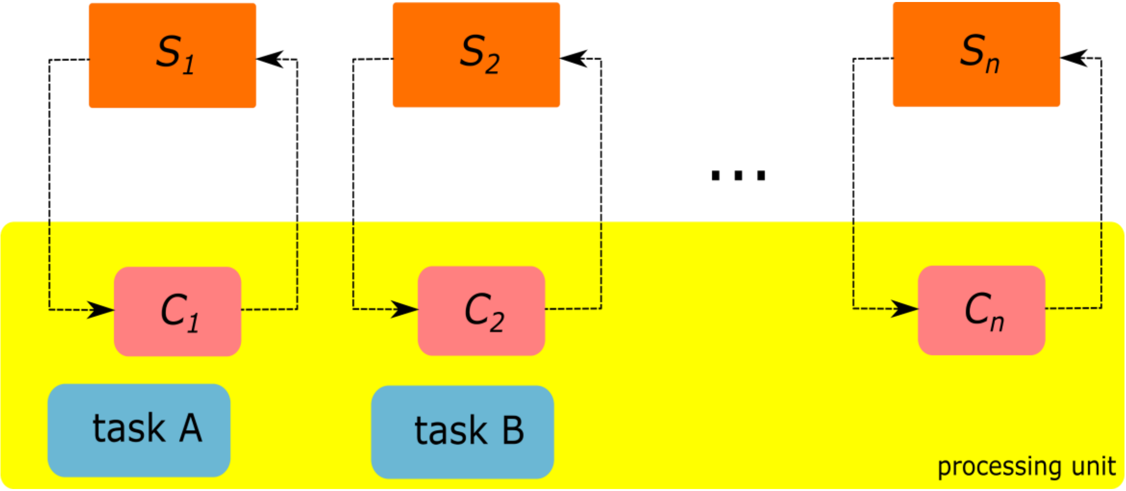


- ❖ **Main thread: Exceeding a certain worst-case execution/computation time**
- ❖ **Giving more resource to a controller leads to better control quality**
- ❖ **Security measures: prevent tasks from consuming more resources**

*Vestal S. Preemptive scheduling of multi-criticality systems with varying degrees of execution time assurance. In28th IEEE International Real-Time Systems Symposium (RTSS 2007)

Control Tasks Characteristics

- **Inter-dependency**
- **Non-monotonicity**



Control Tasks Characteristics

- **Inter-dependency**

- **Non-monotonicity**

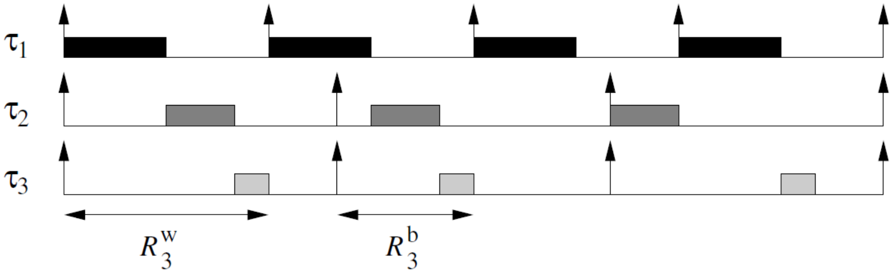
$$\boldsymbol{\tau}_i = (\rho_i, \boldsymbol{c}_i, \boldsymbol{h}_i)$$
$$\boldsymbol{\tau}_1 = (H, 3,6)$$
$$\boldsymbol{\tau}_2 = (M, 2,8)$$
$$\boldsymbol{\tau}_3 = (L, 1,8)$$

Control Tasks Characteristics

- **Inter-dependency**

- **Non-monotonicity**

$$\tau_i = (\rho_i, \boldsymbol{c_i}, \boldsymbol{h_i})$$
$$\tau_1 = (H, 3{,}6)$$
$$\tau_2 = (M, 2{,}8)$$
$$\tau_3 = (L, 1{,}8)$$

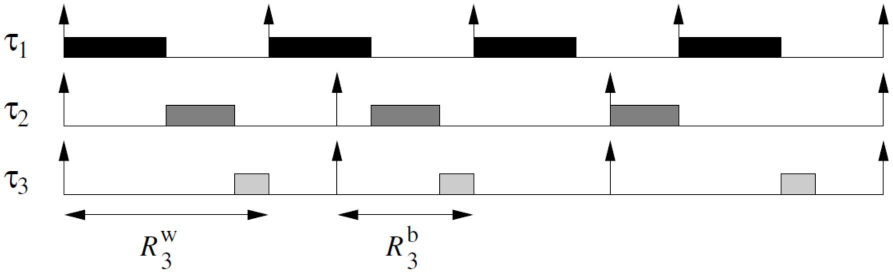


(a) Original task set

Control Tasks Characteristics

- Inter-dependency
- Non-monotonicity

$$\tau_i = (\rho_i, c_i, h_i)$$
$$\tau_1 = (H, 3{,}6)$$
$$\tau_2 = (M, 2{,}8)$$
$$\tau_3 = (L, 1{,}8)$$

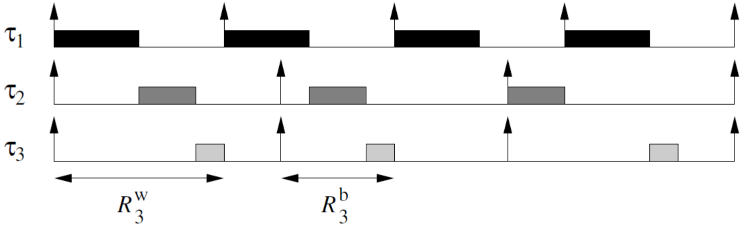


(a) Original task set

$$J_3 = R_3^w - R_3^b = 2$$

Control Tasks Characteristics

- **Inter-dependency**
- **Non-monotonicity**

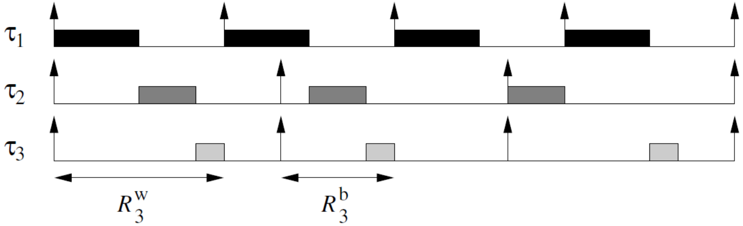


(a) Original task set

$$J_3 = R_3^w - R_3^b = 2$$

Control Tasks Characteristics

- ❖ **Inter-dependency**



(a) Original task set

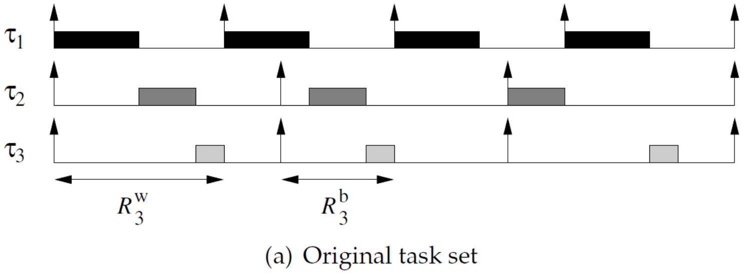
$$J_3 = R_3^w - R_3^b = 2$$

- ❖ **Non-monotonicity**

$$\tau_i = (\rho_i, c_i, h_i)$$
$$\tau_1 = (H, 1, 6)$$
$$\tau_2 = (M, 2, 8)$$
$$\tau_3 = (L, 1, 8)$$

Control Tasks Characteristics

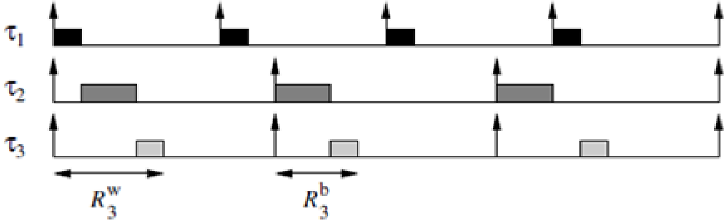
- Inter-dependency
- Non-monotonicity



(a) Original task set

$$J_3 = R_3^w - R_3^b = 2$$

$$\tau_i = (\rho_i, c_i, h_i)$$
$$\tau_1 = (H, 1, 6)$$
$$\tau_2 = (M, 2, 8)$$
$$\tau_3 = (L, 1, 8)$$

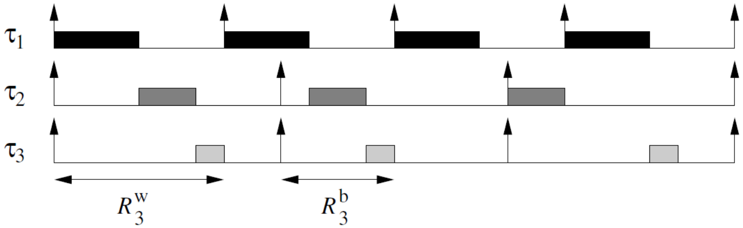


(e) Decreasing computation time c_1

$$J_3 = R_3^w - R_3^b = 1$$

Control Tasks Characteristics

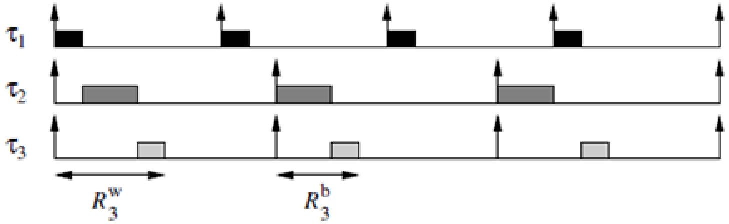
- ❖ **Inter-dependency**
- ❖ **Non-monotonicity**



(a) Original task set

$$J_3 = R_3^w - R_3^b = 2$$

$$\tau_i = (\rho_i, c_i, h_i)$$
$$\tau_1 = (H, 3, 6)$$
$$\tau_2 = (M, 1, 8)$$
$$\tau_3 = (L, 1, 8)$$



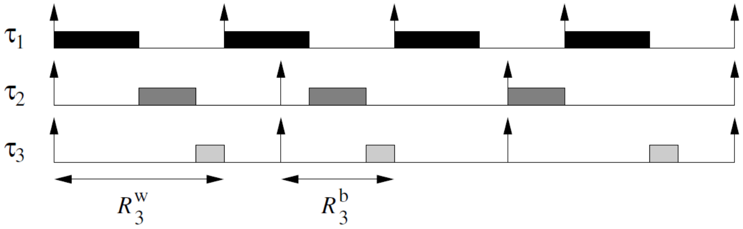
(e) Decreasing computation time c_1

$$J_3 = R_3^w - R_3^b = 1$$

Control Tasks Characteristics

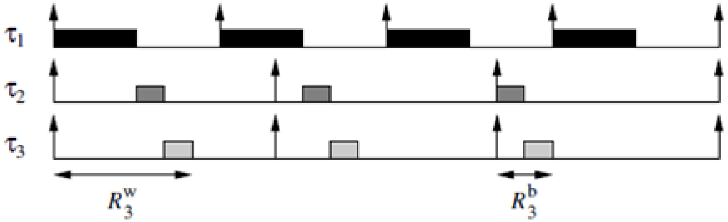
- ❖ **Inter-dependency**
- ❖ **Non-monotonicity**

$$\tau_i = (\rho_i, c_i, h_i)$$
$$\tau_1 = (H, 3, 6)$$
$$\tau_2 = (M, 1, 8)$$
$$\tau_3 = (L, 1, 8)$$



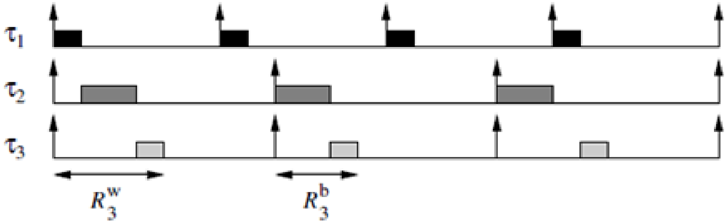
(a) Original task set

$$J_3 = R_3^w - R_3^b = 2$$



(d) Decreasing computation time c_2

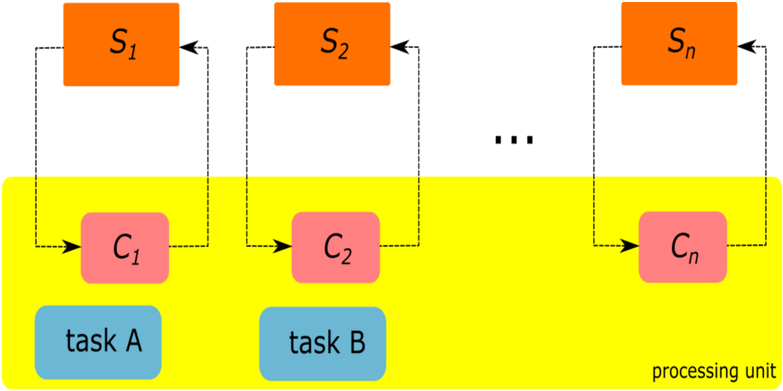
$$J_3 = R_3^w - R_3^b = 3$$



(e) Decreasing computation time c_1

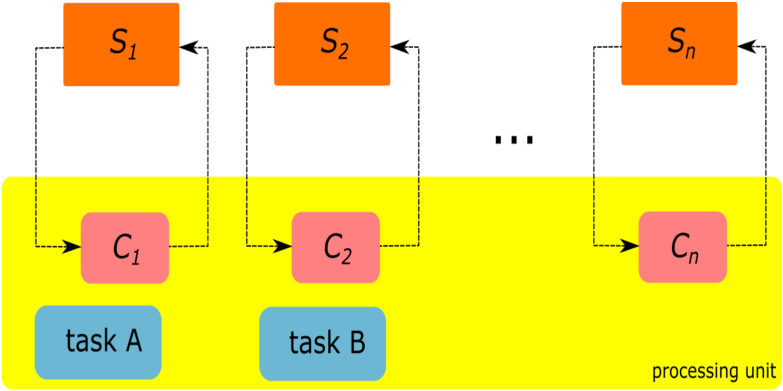
$$J_3 = R_3^w - R_3^b = 1$$

Butterfly Attack



❖ **Indirectly manipulate** **less critical tasks to** **increase jitter** **of a critical task and destabilize the physical system**

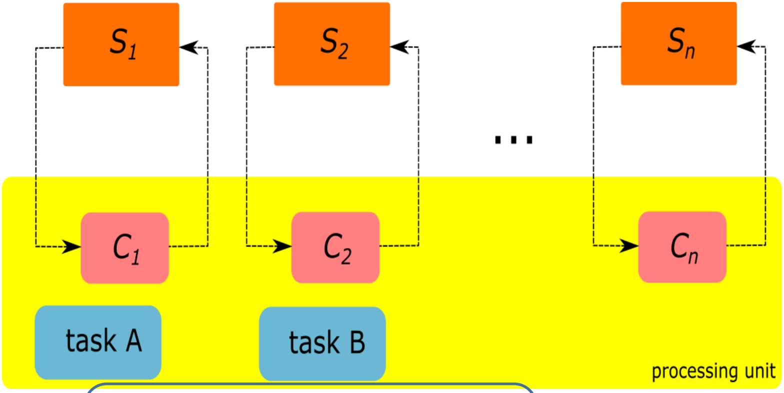
Butterfly Attack



Inter-dependency

- **Indirectly manipulate less critical tasks to increase jitter of a critical task and destabilize the physical system**

Butterfly Attack

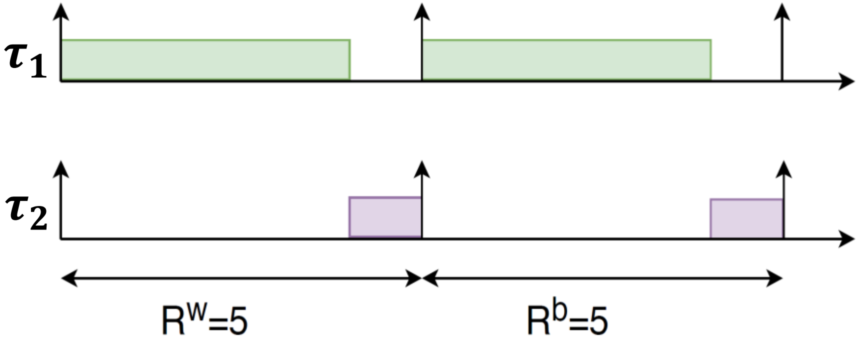


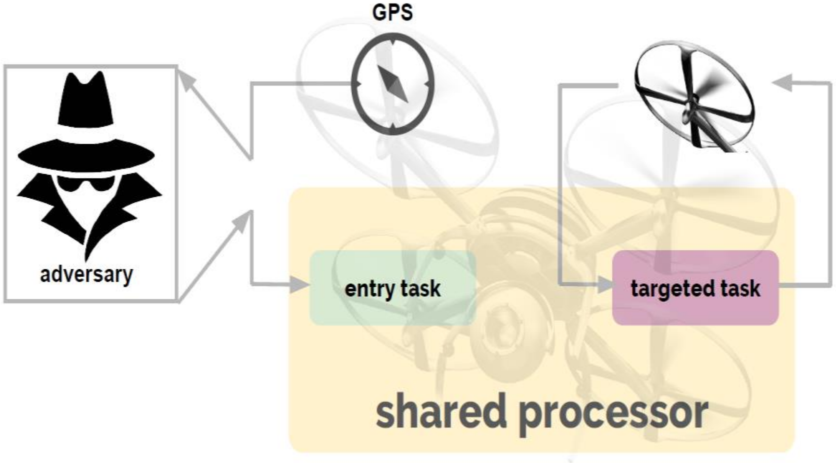
Inter-dependency

Non-monotonicity

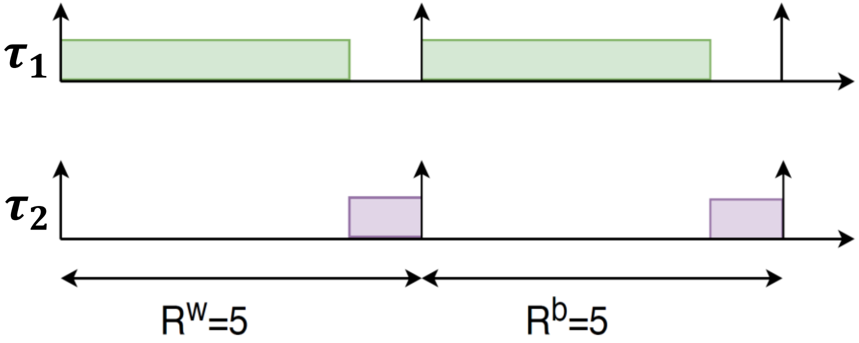
- **Indirectly manipulate less critical tasks to increase jitter of a critical task and destabilize the physical system**

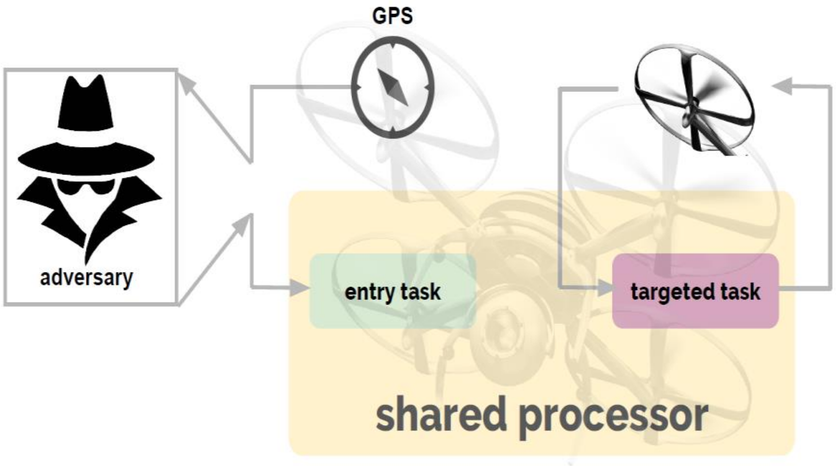
Experimental Results





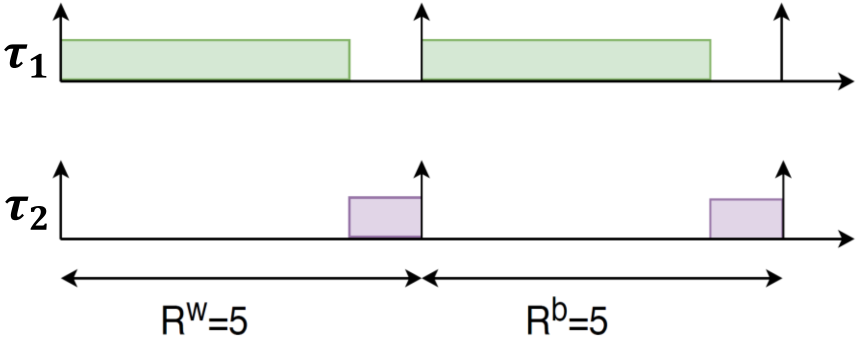
Experimental Results



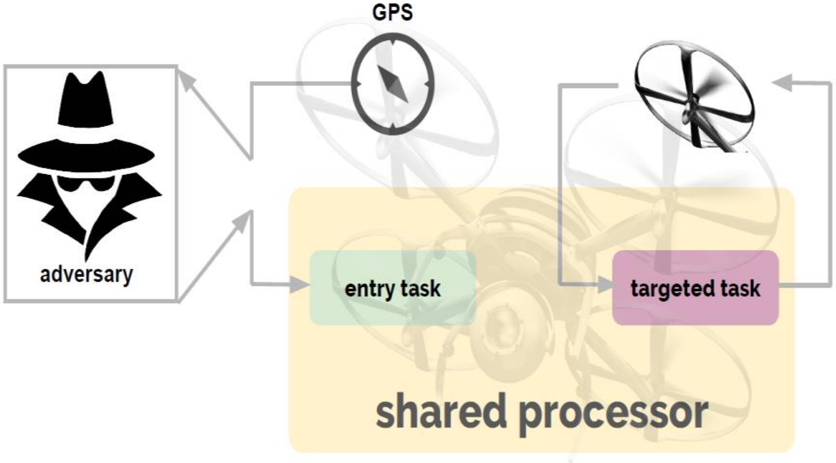


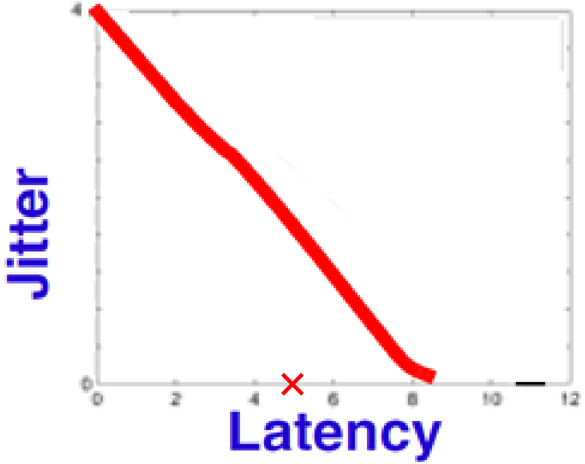
$$L_2 = 5,\ J_2 = R_2^w - R_2^b = 0$$

Experimental Results

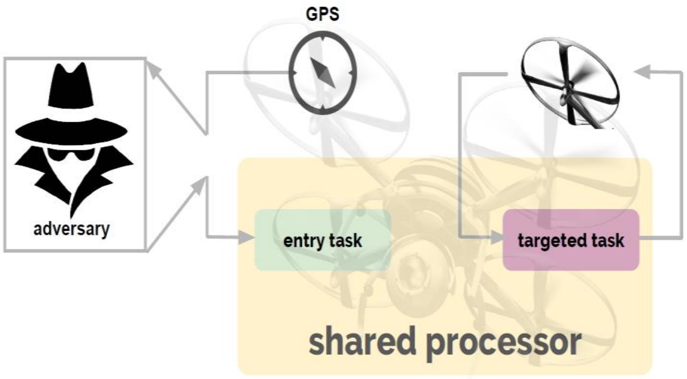


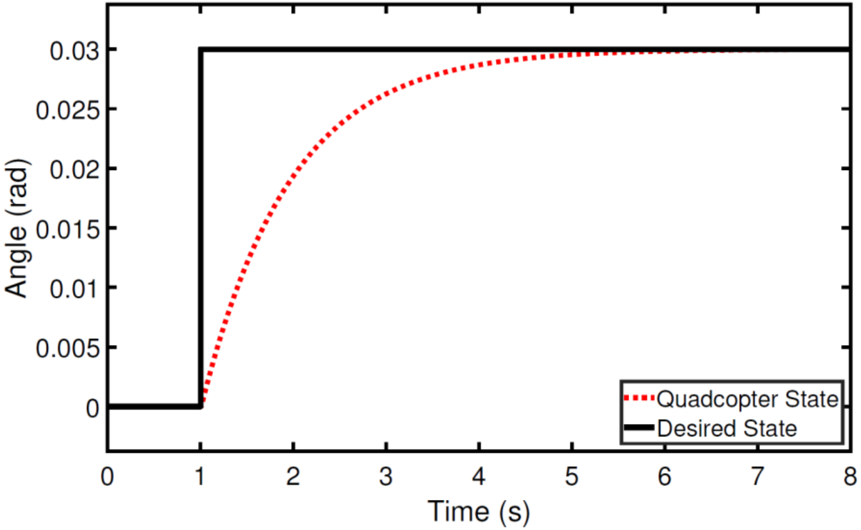
$$L_2 = 5,\ J_2 = R_2^w - R_2^b = 0$$



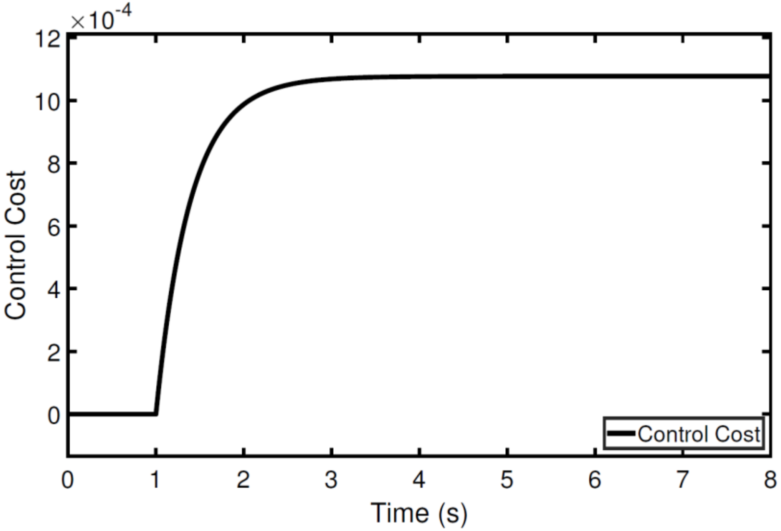


Experimental Results



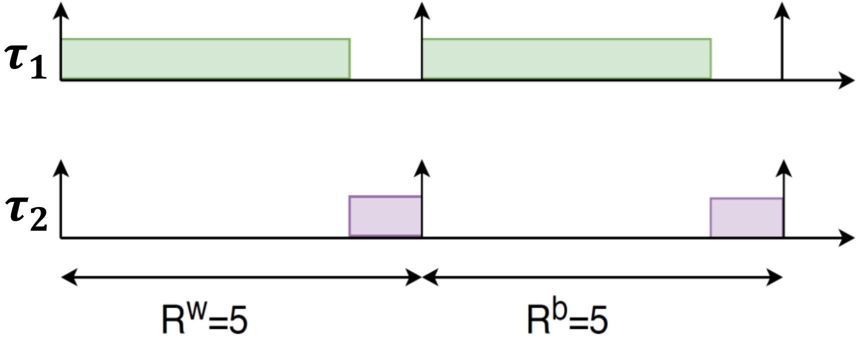


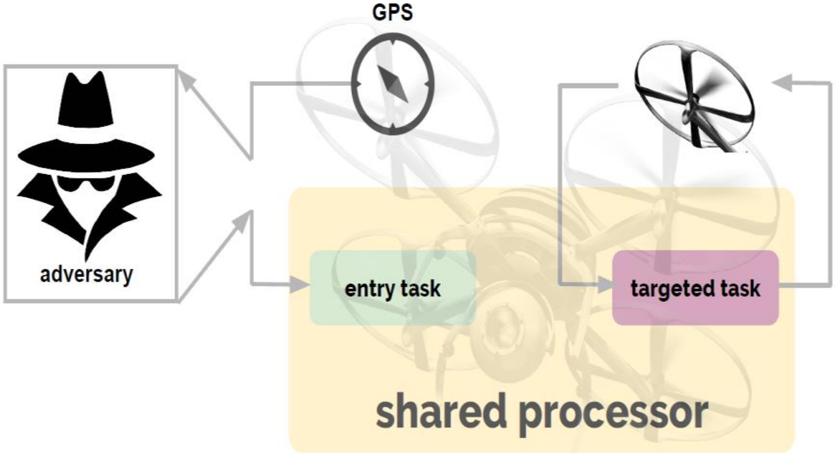
(a) Quadcopter vertical angle (stable).



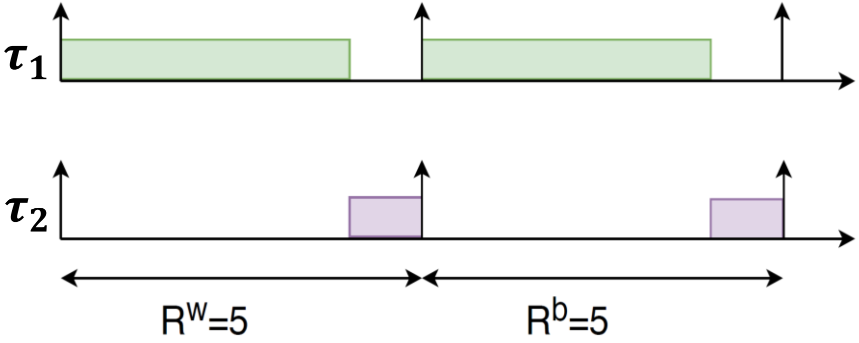
(b) Quadcopter control cost (stable).

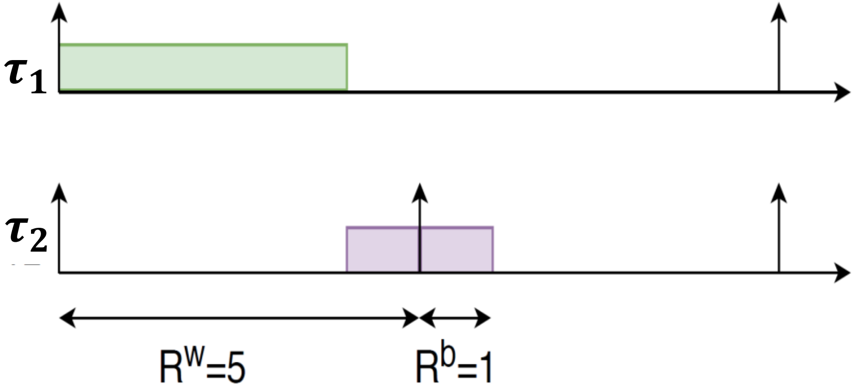
Experimental Results

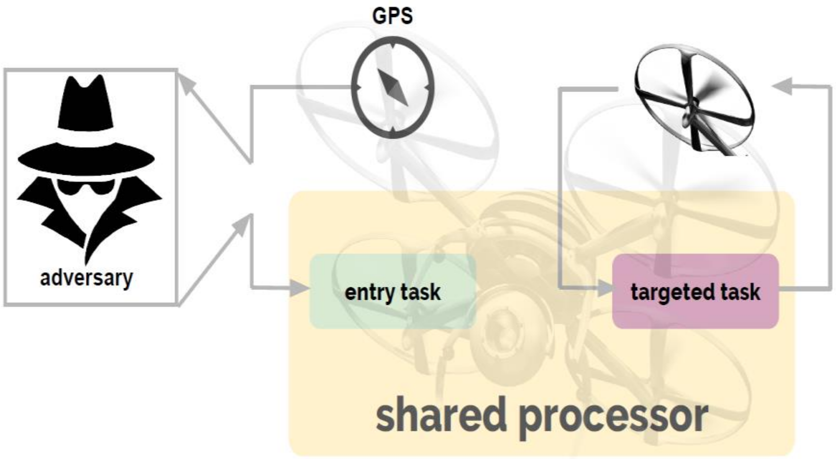




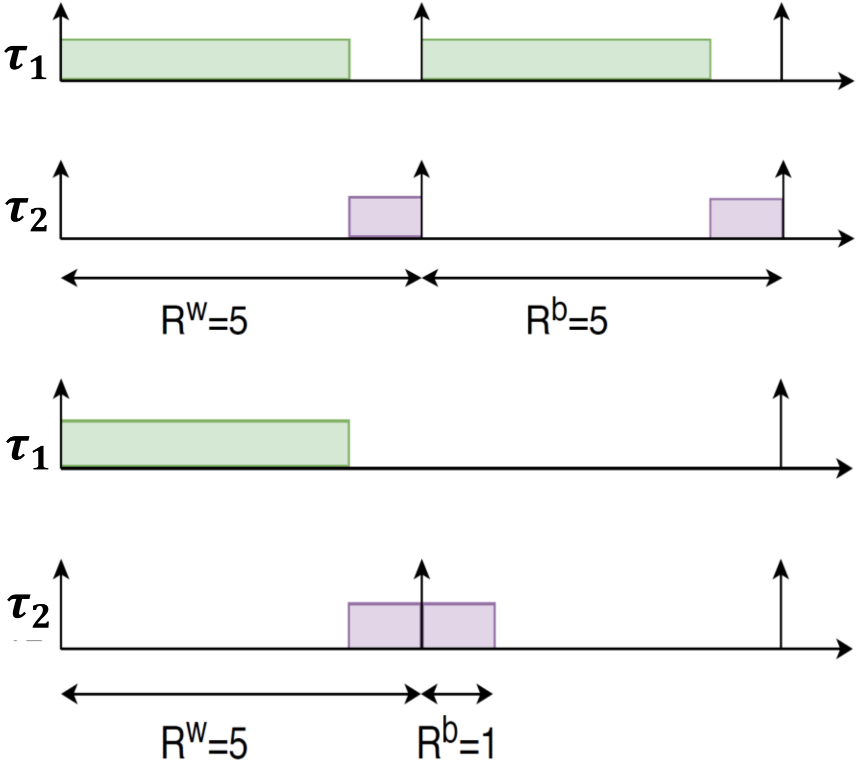
Experimental Results

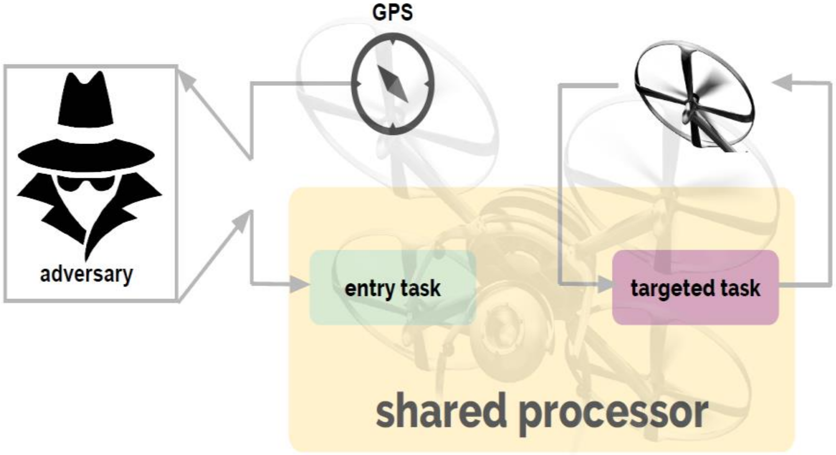






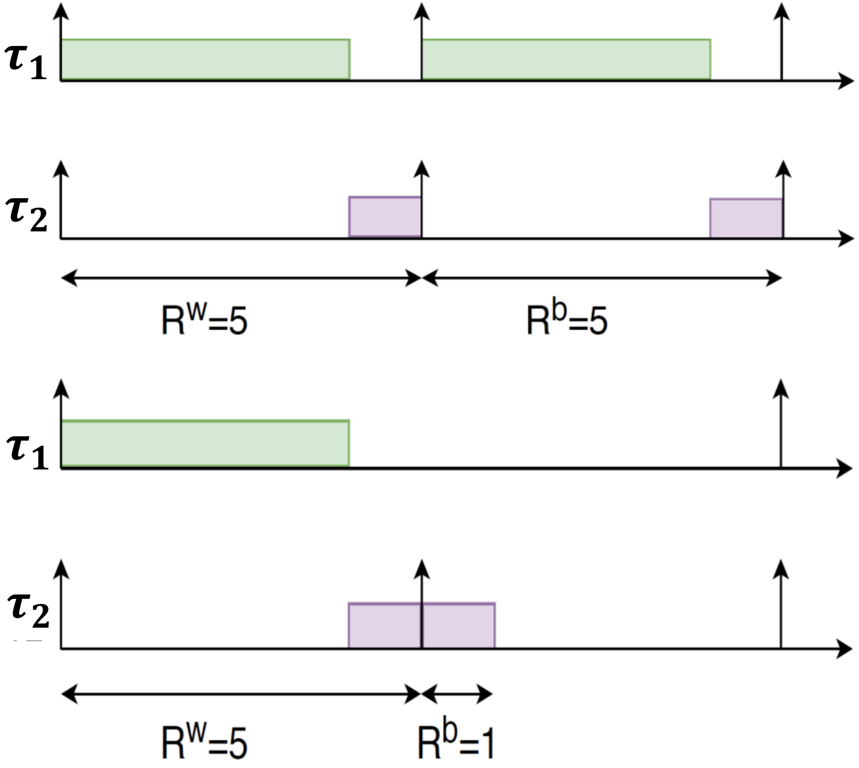
Experimental Results



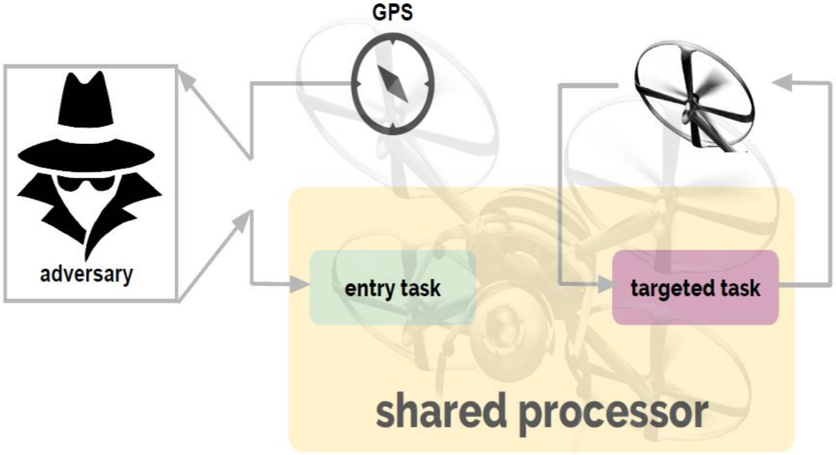


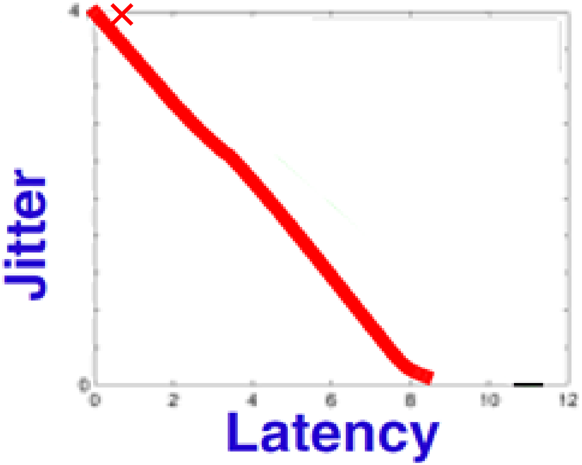
$$L_2 = 1,\ J_2 = R_2^w - R_2^b = 4$$

Experimental Results

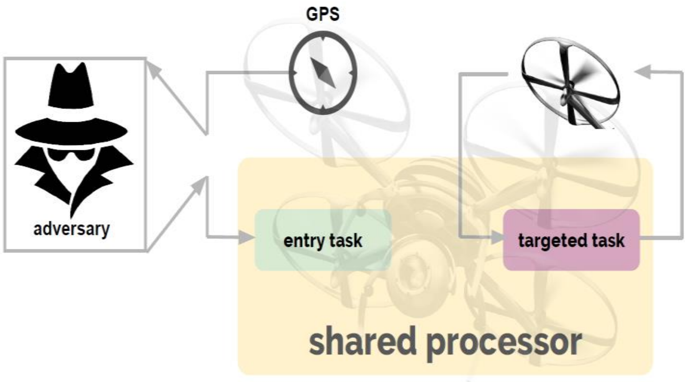


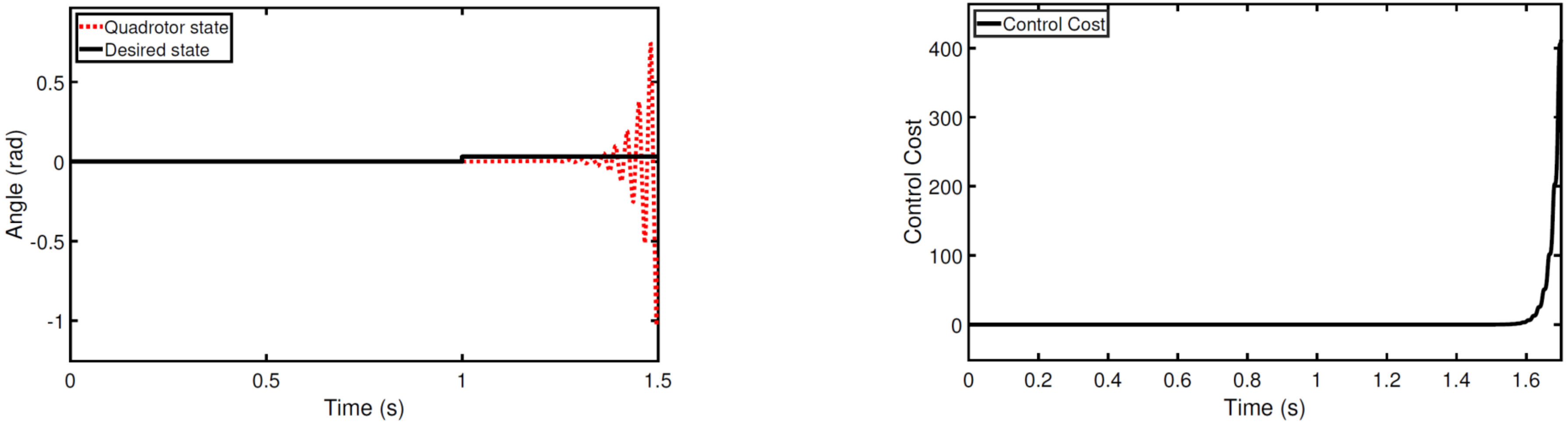
$$L_2 = 1,\ J_2 = R_2^w - R_2^b = 4$$





Experimental Results





(a) Quadcopter vertical angle (unstable).

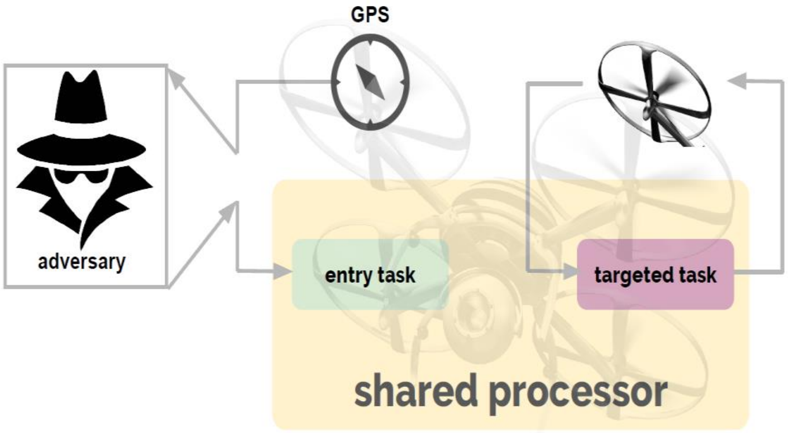
(b) Quadcopter control cost (unstable).

Butterfly attack

- Manipulating less critical (protected) task

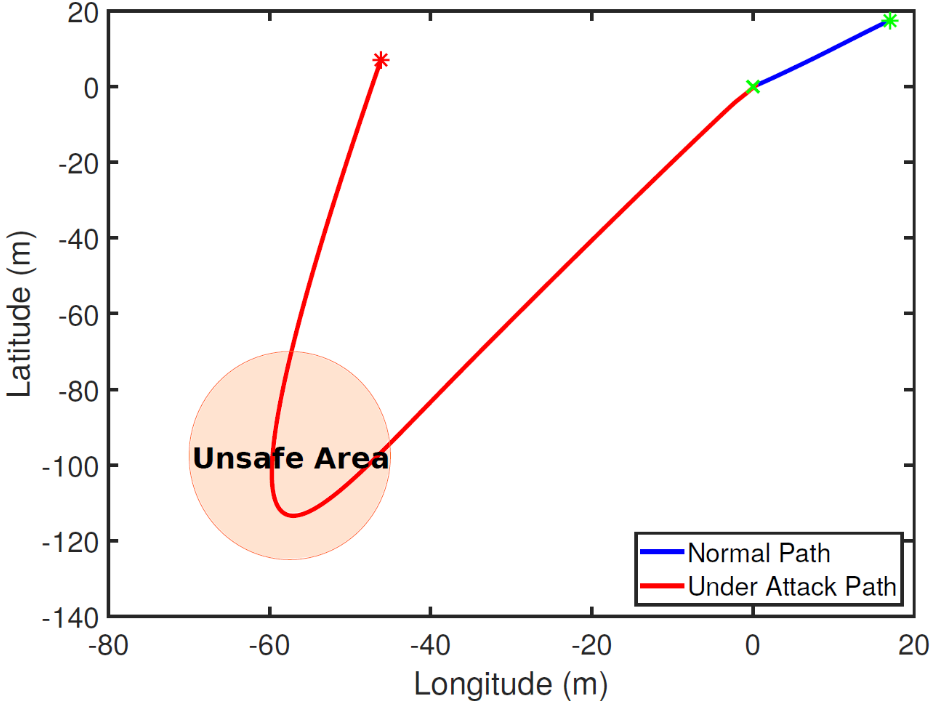
- Increase the available resource

- Difficult to detect attacker



Beyond Butterfly attack

- **Hijack the drone using Butterfly attack**
- **Launch the attack for short time**
- **Needs some extra knowledge**
- **Unpredicted results**



Mitigation

- Run a dummy task to compensate
- Design a robust controller
- Ensure temporal isolation using servers

Conclusion

- **Introduction to Butterfly attack**
- **Identify Inter-dependency and Non-monotonicity**
- **Demonstrate the possibility of attack experimentally**

Questions