Stealthy Delay Attacks for Cyber-Physical Systems

Talitha Nauta
Lund University, Department of Automatic Control
Supervisor: Martina Maggio

LUND UNIVERSITY

Motivation & Research Goals

Control systems are vulnerable to cyber attacks which silently alter control signals or output measurements. However, these attacks are not always practical and their implementation is usually challenging, as it requires overcoming security defences. A more feasible attack manipulates the timing of controller operations, subtly impacting system performance without altering data or control signals directly. This research explores the worst-case impacts of these timing attacks on control systems.

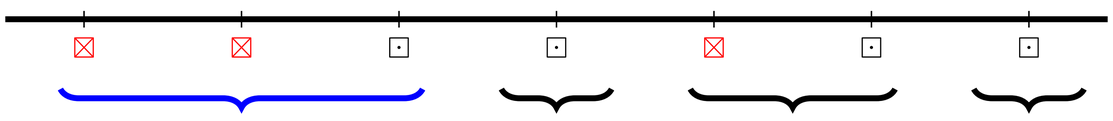
Attack Model

We control the discrete-time linear time-invariant system $\mathcal{P}$, using a one-step delay linear feedback controller $\mathcal{C}$.

$$\mathcal{P}: \begin{cases} x_{k+1} &= A\,x_k + B\,u_k \\ y_k &= C\,x_k + D\,u_k \end{cases} \qquad \mathcal{C}: \begin{cases} z_{k+1} &= F\,z_k + G\,y_k \\ u_{k+1} &= H\,z_k + K\,y_k \end{cases}$$

The attacker introduces a higher-priority (with respect to the controller) task that causes a delay of a_k time units in the computation of the control signal. The attack is computed as a sequence of delays.

- $(a_i)_{i=1}^{\ell} = \begin{pmatrix} 3 & 1 & 2 & 1 \end{pmatrix}$
- $(a_{\mathrm{bi},j})_{j=1}^{\bar{k}} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$



Deadline hit ⊡ Deadline miss ⊠

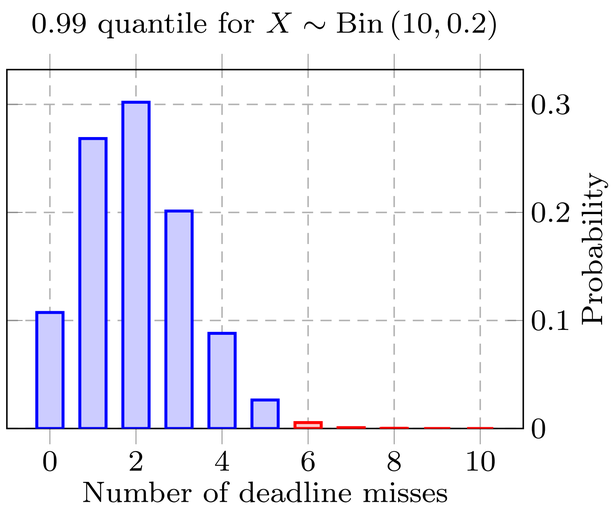
The objective of the attacker is to calculate an attack sequence $(a_i)_{i=1}^{\ell}$ that maximises the output displacement at time $\bar{k}$. This gives the following optimisation problem.

$$\begin{aligned} (a_i^*)_{i=1}^{\ell} \quad = \quad & \arg\max \quad y_{\bar{k}}^T\, Q\, y_{\bar{k}} \\ & \text{s.t.} \quad 1 \le a_i \le \bar{a} \\ & \quad\; a_i \in \mathbb{N}^+ \\ & \quad\; \textstyle\sum_{i=1}^{\ell} a_i = \bar{k} \\ & \quad\; \text{closed-loop system eq.} \\ & \quad\; \text{remain undetected} \end{aligned}$$

Detector Model

Detector #1: Number of deadline misses X is distributed by $X \sim \mathrm{Bin}(w, \hat{p})$. Attack is detected if H_0 can be rejected.

$$H_0 : p \le \hat{p}$$
$$H_1 : p > \hat{p}$$



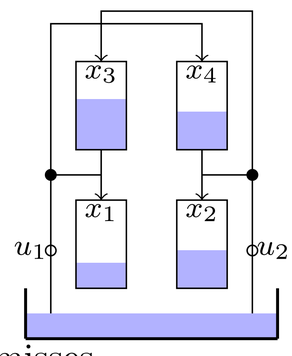
Detector #2: Sequence of delays Y is distributed by $Y \sim \mathrm{Geo}(\hat{p})$

$$Q_{\chi^2} = \sum_{j=1}^{\bar{a}} \frac{(n\,P(Y=j) - O_j)^2}{n\,P(Y=j)} \qquad Q_{\chi^2} \le \chi^2_{1-\alpha,\bar{a}-1}$$

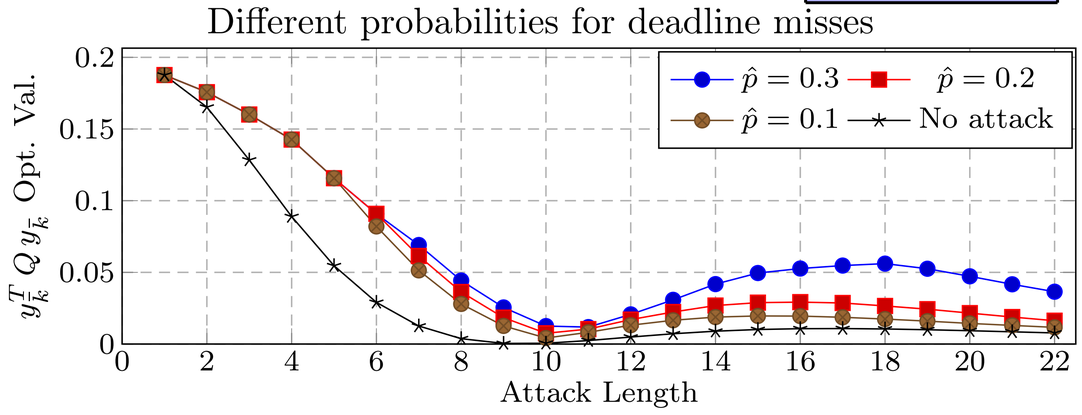
O_j is the number of delays of length j in $(a_i^*)_{i=1}^{\ell}$, α is the false alarm rate, and $\bar{a}-1$ are the degrees of freedom.

Selected Results: Detector #1

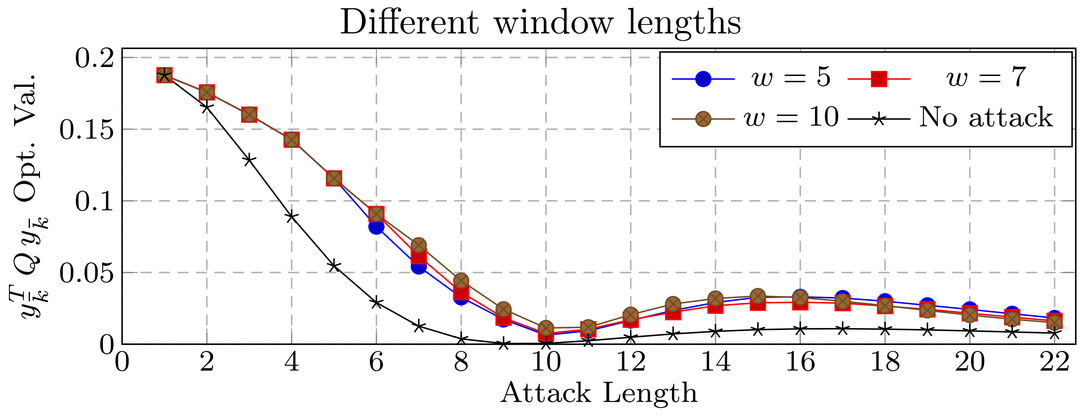
Quadruple Tank (stable system): the control objective is to control of the water level in the lower two tanks.



Different probabilities for deadline misses

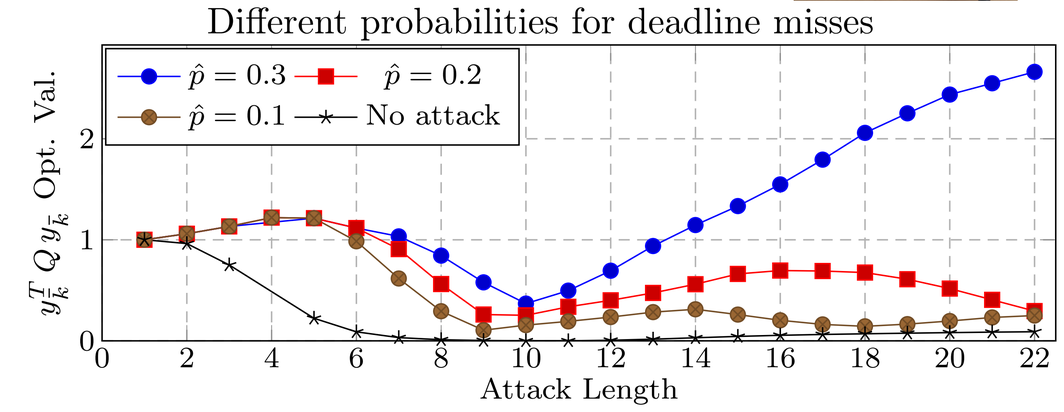


Different window lengths

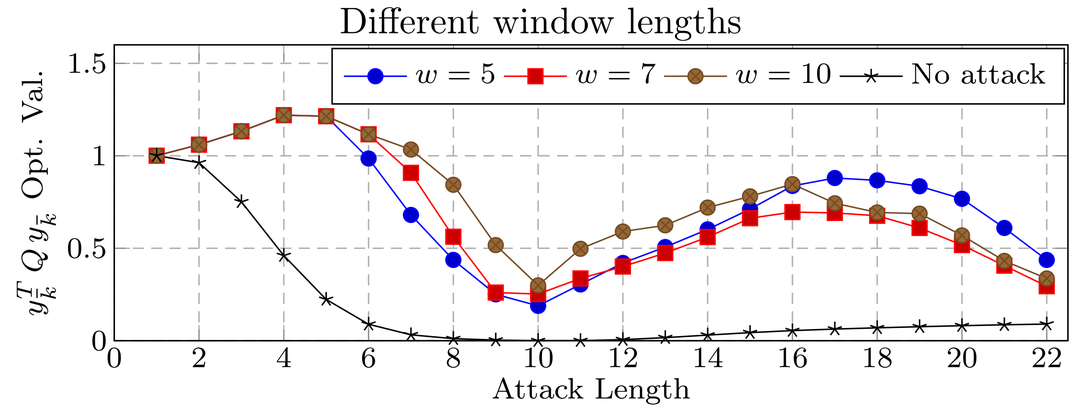


Furuta Pendulum (unstable system): the control objective is to control the pendulum stick in the upright position.

Different probabilities for deadline misses



Different window lengths



WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM