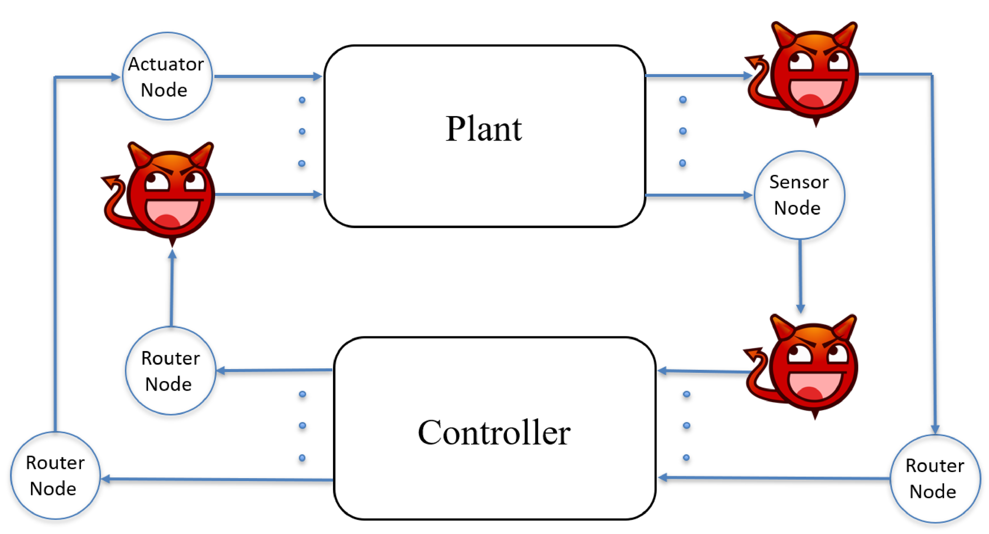
Authentication of cyber-physical systems under learning-based attacks

M. J. Khojasteh, A. Khina, M. Franceschetti, T. Javidi

Attacks on CPS

In network control systems, sensor observations and control signals can be hijacked.



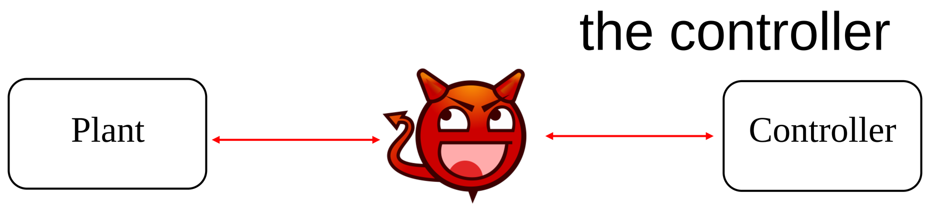
Computer virus Stuxnet a 'game changer,' DHS official tells Senate

"Stuxnet has changed the way we view the security threat"

Cloud robotics

The man in the middle

A fictitious plant for the controller



A malicious controller for the plant

MITM attack types

Replay

[redacted]

Y. Mo, B. Sinopoli (2009)

Statistical-duplicate

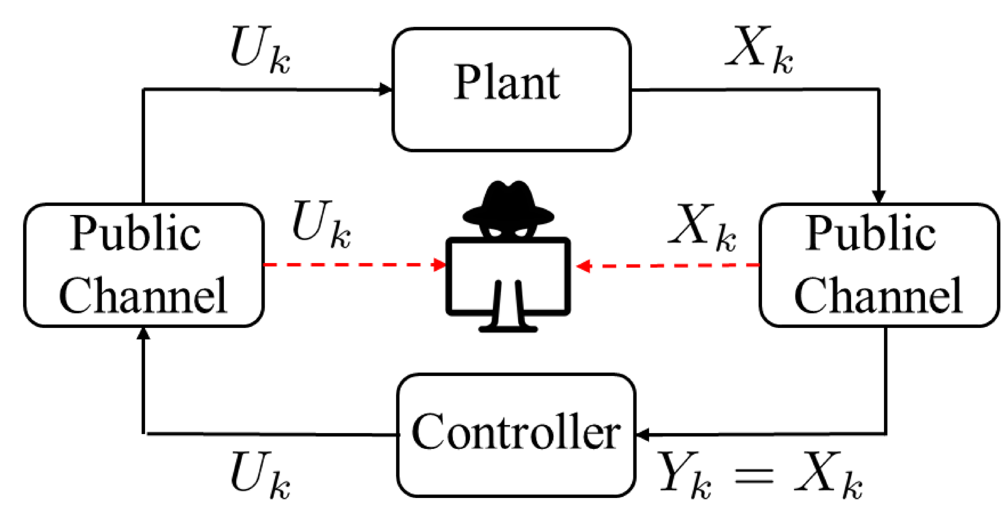
$X_{k+1} = aX_k + U_k + W_k$

B. Satchidanandan, P. R. Kumar (2017)
R. S. Smith (2011)

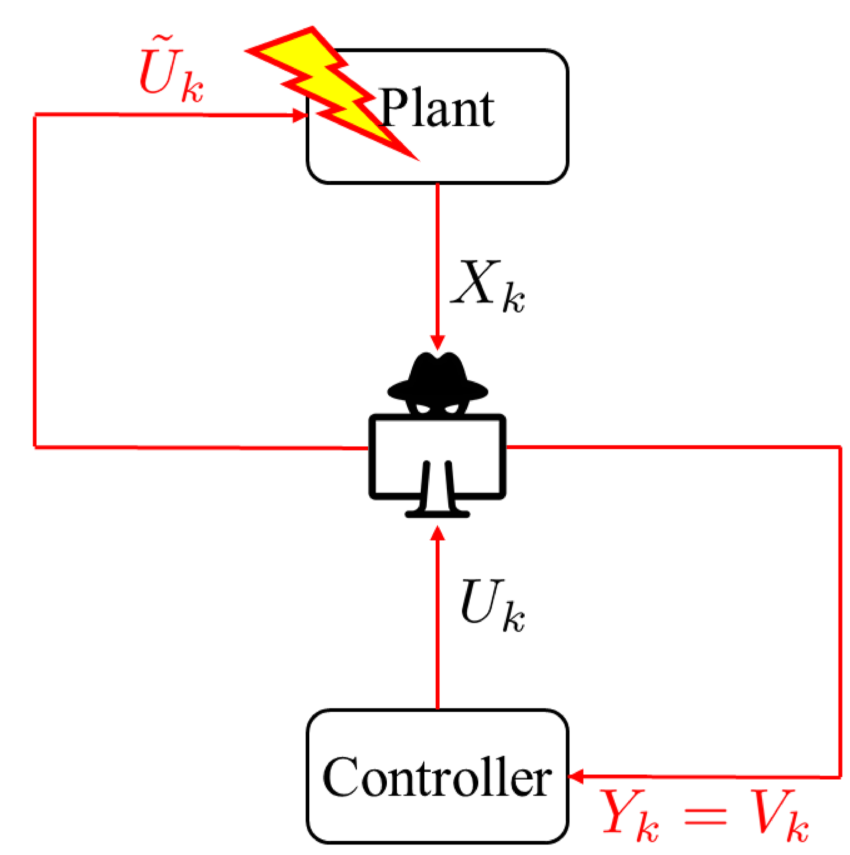
Learning-based

$X_{k+1} = aX_k + U_k + W_k$

Exploration phase



Exploitation phase



Mathematical formulation

Linear scalar dynamical system

$$X_{k+1} = aX_k + U_k + W_k$$
$$W_k \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}$$

Under legitimate system operation we expect

$$Y_{k+1} - aY_k - U_k(Y_1^k) \sim \text{i.i.d. } \mathcal{N}(0, \sigma^2)$$

Anomaly detector

$$\frac{1}{T}\sum_{k=1}^{T}\left[Y_{k+1} - aY_k - U_k(Y_1^k)\right]^2 \in (Var[W] - \delta, Var[W] + \delta)$$

Fictitious sensor reading

$$V_{k+1} = \hat{A}V_k + U_k + \tilde{W}_k$$
$$\tilde{W}_k \sim \mathcal{N}(0, \sigma^2) \quad \text{i.i.d.}$$

Assumption on the power of the fictitious sensor reading

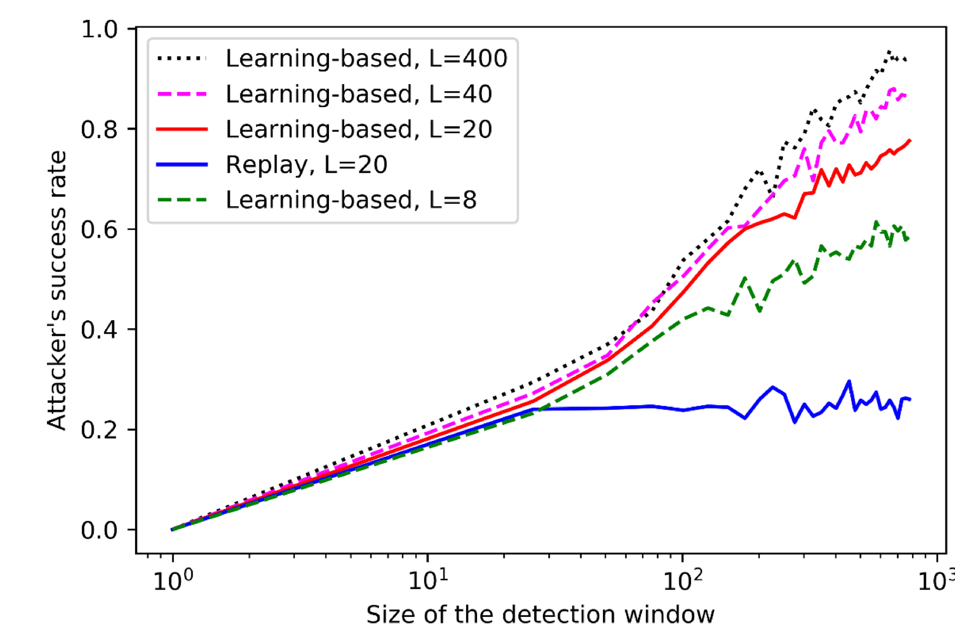
$$\lim_{T\to\infty} \frac{1}{T}\sum_{k=L+1}^{T} V_k^2 = 1/\beta < \infty$$

The deception probability, lower bound

Least-square learning algorithm

$$\hat{A} = \frac{\sum_{k=1}^{L-1}(X_{k+1} - U_k)X_k}{\sum_{k=1}^{L-1} X_k^2}$$

$$\lim_{T\to\infty} P_{dec}^{a} \geq 1 - \frac{2}{(1+\delta\beta)^{L/2}}$$



The deception probability, upper bound

Assume the open-loop gain of the plant is a random variable

$$A \sim \text{Unif.}[-R, R]$$

whose distribution is known to the attacker, and whose realization is known to the controller. Then letting

$$Z_1^k = (X_1^k, U_1^k)$$

we have

$$\lim_{T\to\infty} P_{dec} \leq \frac{I(A; Z_1^L) + 1}{\log(R/\sqrt{\delta\beta})}$$

The denominator represents the intrinsic uncertainty of A when this is observed at resolution

$$\epsilon = \sqrt{\delta\beta}$$

corresponding to the entropy of the quantized random variable $H(A_\epsilon)$

The numerator represents the information revealed about A from the observation of the random vector Z_1^L

Privacy-enhancing signal

To impede the learning process of the attacker

$$U_k = \bar{U}_k + \Gamma_k$$