Randomness Monitor for Stealthy Attacks on Autonomous Systems

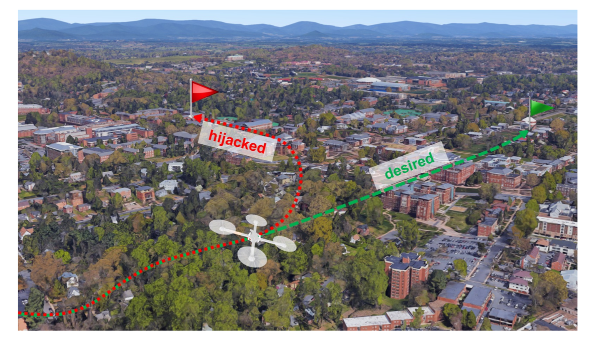
UVA ENGINEERING
LINK LAB

Paul J Bonczek[1], Shijie Gao[1], and Nicola Bezzo[1,2]

[1]Department of Electrical and Computer Engineering, [2]Department of Engineering Systems and Environment
University of Virginia

MOTIVATION

- Autonomous vehicles are susceptible to malicious attacks with the intent to hijack the vehicle to an undesired state.



OBJECTIVE

- Monitor sensor measurements for **inconsistent** and **non-random** behavior
- **Tune** detectors for specific false alarm rates under normal operation.

While guaranteeing:
✓ **Detection:** A range of attacks that are always detected

PROBLEM FORMULATION

PROBLEM: **Measuring randomness:**

- Monitor randomness at run-time for each sensor
- Send an alarm when non-randomness is detected
- Develop a tuning method allowing a desired false alarm rate according to the system model.

Detection \ Attacks	Yes	No
Yes	True/ Positive	False/ Positive
No	False/ Negative	True/ Negative

APPROACH

State Estimation: A standard steady state Kalman Filter is implemented to provide the system a prediction of the state evolution.

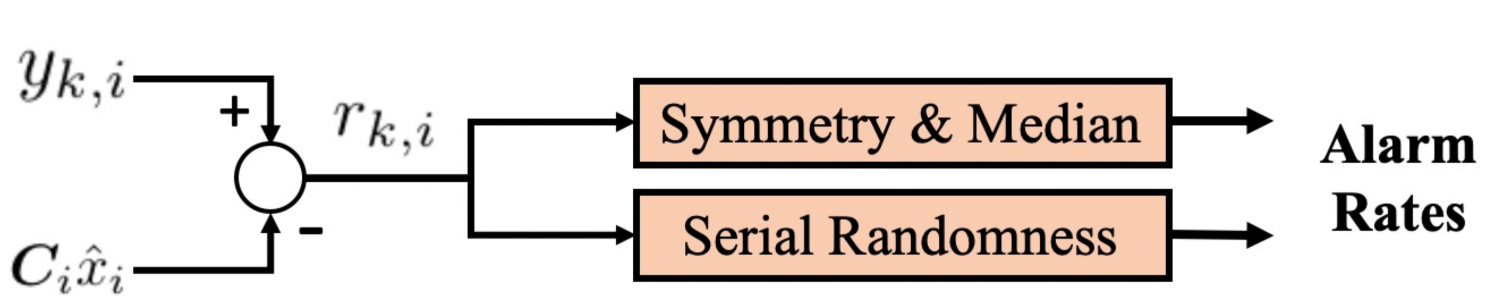
- A comparison between the sensor *measurement* and the state *prediction* from the Kalman Filter, known as the ***residual***, is the value to be analyzed.

DEFINITION: **Randomness –** A measurement is considered random if:

1) Corresponding residual is symmetric over its expected value.
2) An incoming residual sequence is being received free of patterns, sequence should be impossible to predict.

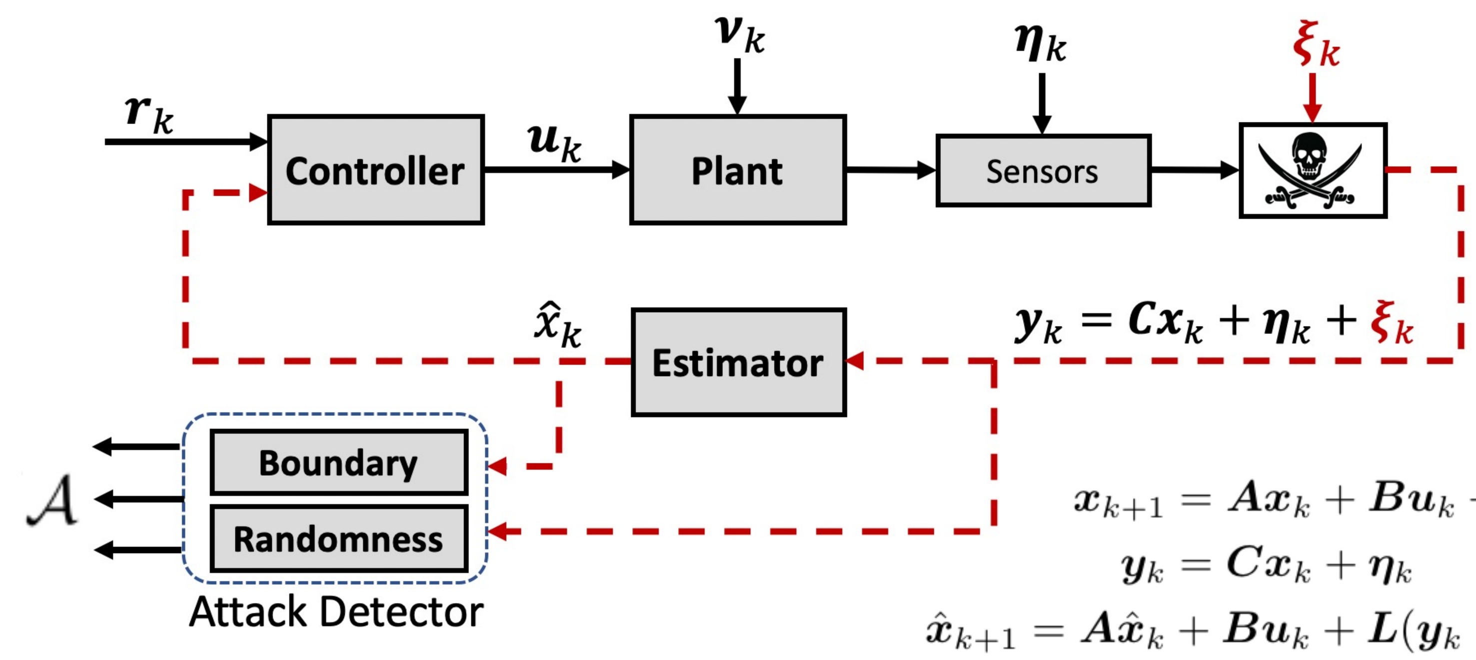
Statistical Analysis: A series of tests determining whether a residual sequence is truly random with alarm rates.

- Three different tests are utilized to search for attacks:
  1) Symmetric distribution over the residual expected value
  2) Determine whether the measurement sequence is received randomly



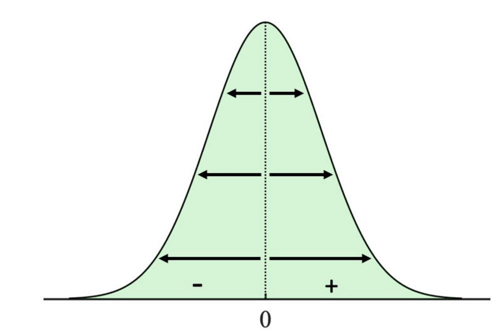
3) Measurements should remain within bounds

Cyber-Physical Systems under the effects of attacks



$$x_{k+1} = Ax_k + Bu_k + \nu_k$$
$$y_k = Cx_k + \eta_k$$
$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + L(y_k - C\hat{x}_k)$$

Statistical-based monitoring: Leverage various statistical tests on the residual to determine the randomness of incoming measurements.



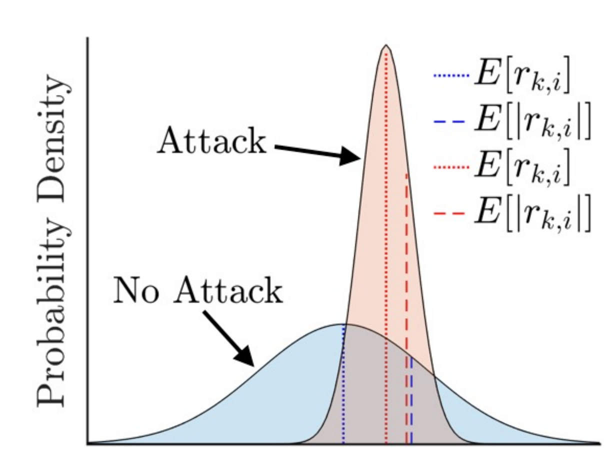
- **Wilcoxon Signed-Rank Test** compares the absolute value of positive and negative residuals by providing a ranking to determine symmetry.
- **Serial Independence Runs Test** is a comparison of the current residual to the previous over a given sequence to determine random behavior.

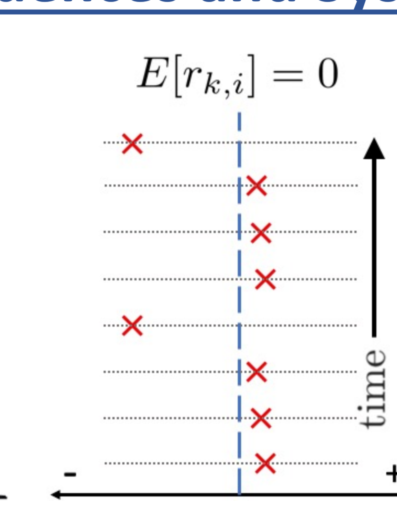
- **Boundary Detectors:** We can augment our statistical framework for Randomness Monitoring to any state-of-the-art boundary detector.

Why the need for an extra layer of security?

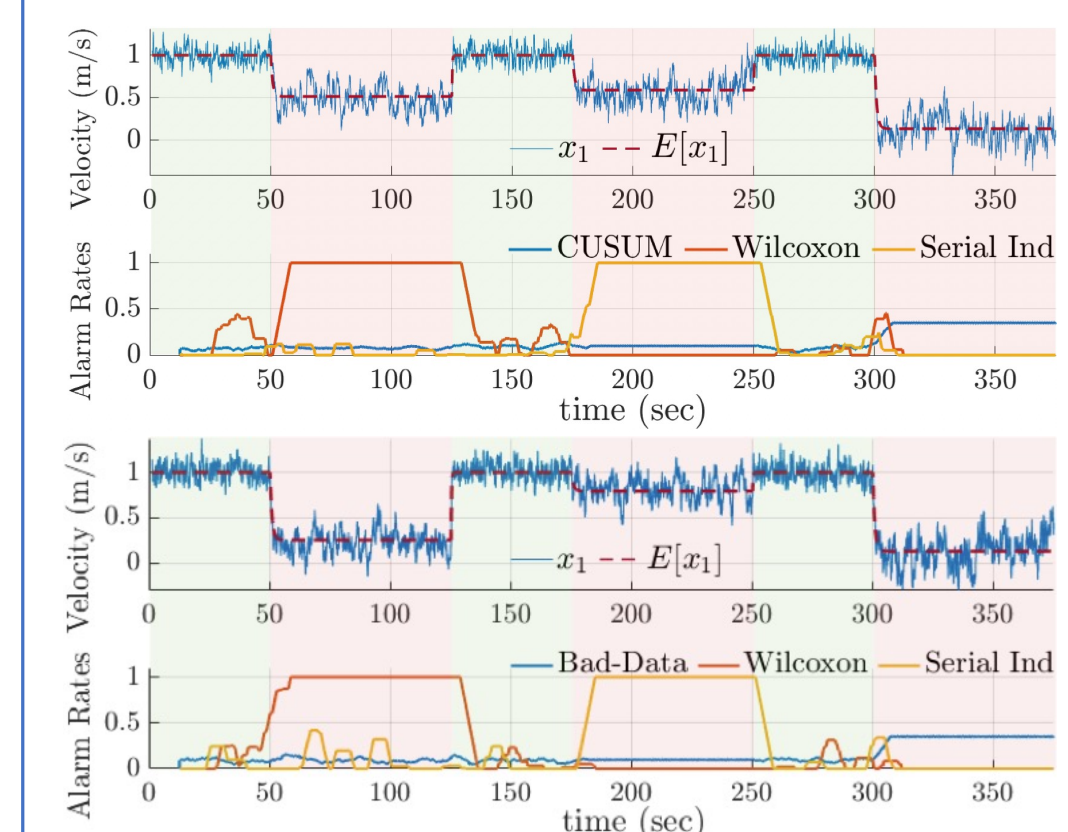
- Stealthy attackers may be able to learn system and noise dynamics, on-board controllers, and state estimators.
- Attacks can create an attack signal to completely fool state-of-the-art detection schemes

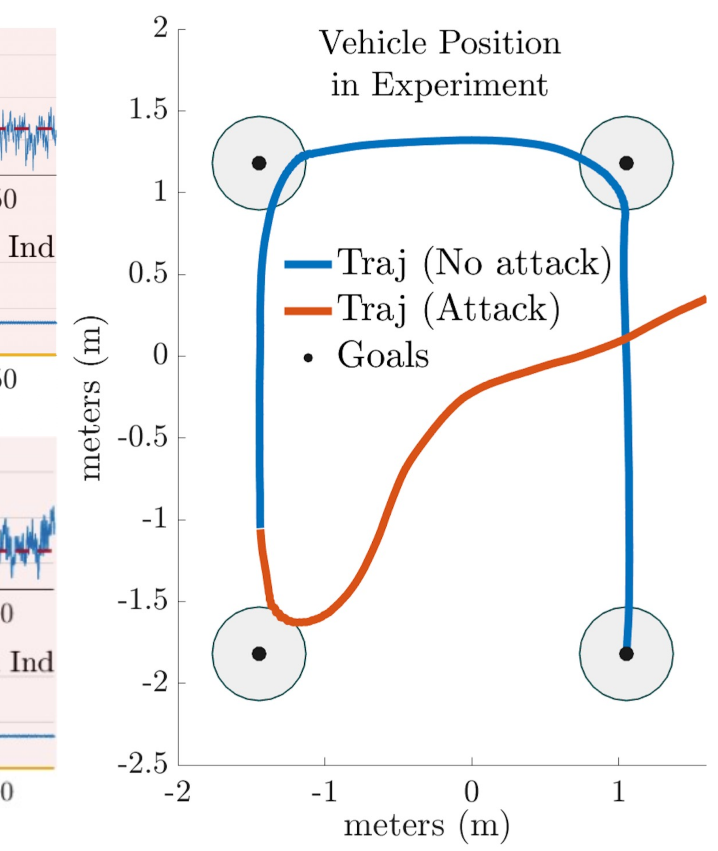
Attack Sequences and System Platform

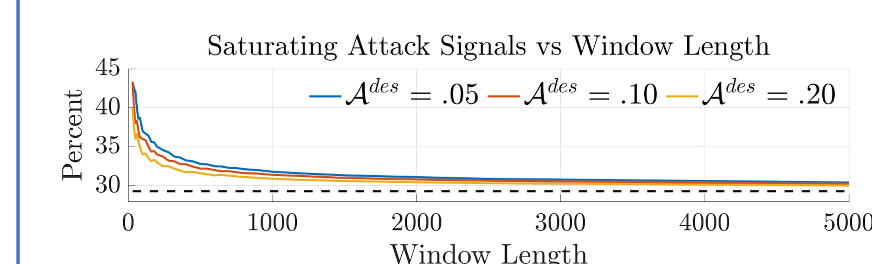


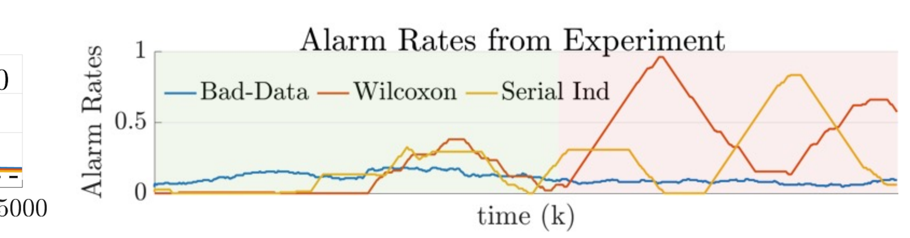


SIMULATION/EXPERIMENT RESULTS









CONCLUSION & FUTURE WORK

- **Outcome:** Detection of compromised sensors by monitoring the residual randomness..
- **Future Efforts:**

➢ Include disturbances in the model.
➢ Formalize a method of filtering out compromised sensors with sensor redundancy for resilient state estimation.
➢ Improve statistical methodology to find different classes of stealthy attacks