CERIAS

The Center for Education and Research in Information Assurance and Security

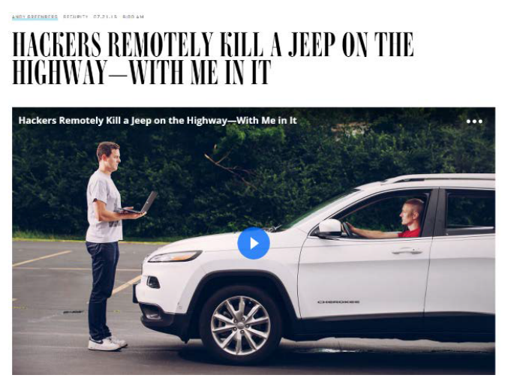
A Model-based Cyber Attack Detection and Identification for Networked Vehicle Systems

Sangjun Lee and Byung-Cheol Min

SMART Lab, Department of Computer and Information Technology, Purdue University

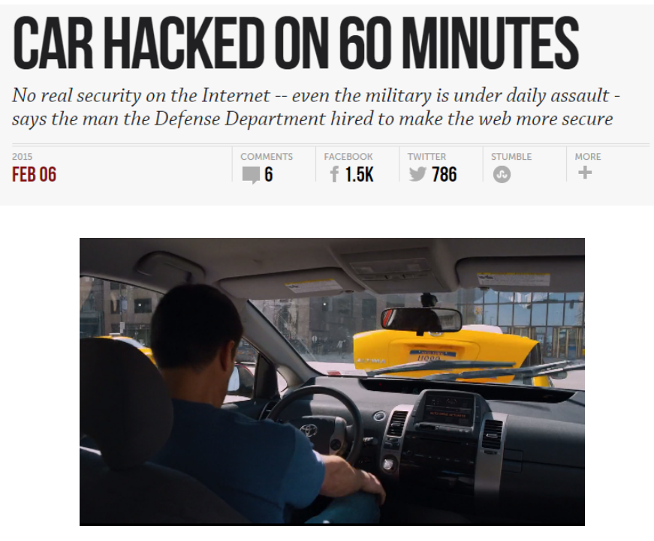
Background

- The reliance on communication networks to transmit measurements and control signals increases the possibility of cyber attacks against networked vehicle systems.



HACKERS REMOTELY KILL A JEEP ON THE HIGHWAY—WITH ME IN IT

Hackers Remotely Kill a Jeep on the Highway—With Me in It



CAR HACKED ON 60 MINUTES

No real security on the Internet -- even the military is under daily assault - says the man the Defense Department hired to make the web more secure

2015 FEB 06 COMMENTS 6 FACEBOOK 1.5K TWITTER 786 STUMBLE MORE

Feds Preach Cybersecurity to Carmakers

By Philip E. Ross
Posted 31 Oct 2016 | 13:00 GMT

Researcher Hacks Self-driving Car Sensors

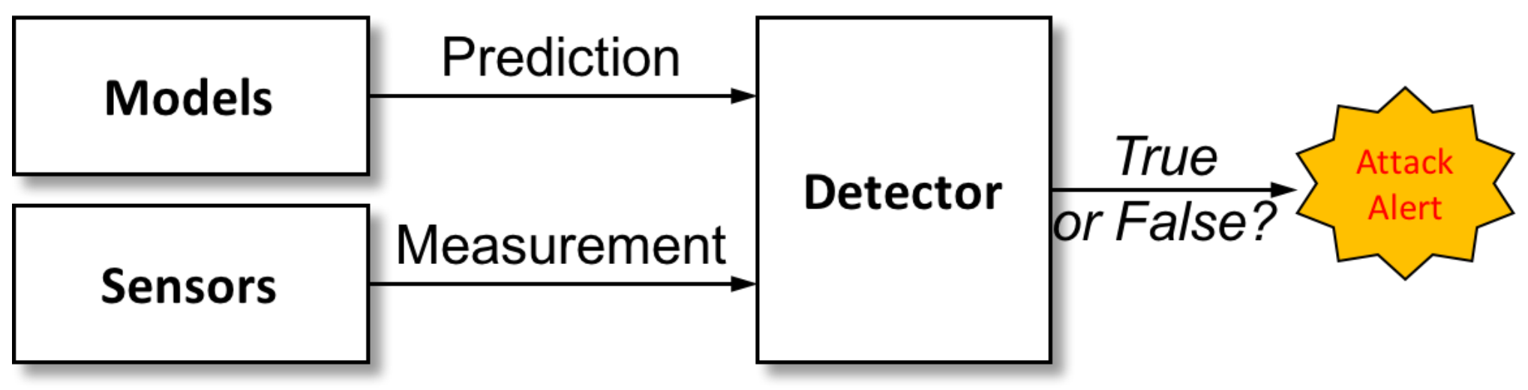
By Mark Harris
Posted 4 Sep 2015 | 19:00 GMT

> ***Research Questions***
> *What if the units that govern safety are being compromised?*
> *Is each agent able to detect any unexpected behavior?*

- **Cyber Attack in This Study**
  An action which undermines the security of robotic systems for malicious purposes. (e.g. possible sensor failures mainly caused by false data injection attacks, corrupted signal reading, or any combination of these.)

Approach

- **Model-based Detection Scheme**
  Focusing on a vehicle's physical behavior, rather than traditional computer-based approaches such as firewalls, encryptions, etc.



> *Models allow us to predict vehicle's normal behaviors and sensors provide actual behaviors.*

- **System Models under Attacks:** Considers a discrete time stochastic LTI system

$$x(k+1) = Ax(k) + Bu(k) + v(k)$$
$$y_{\alpha}(k) = Cx(k) + \alpha(k) + \omega(k)$$

where α represents additive attacks on sensors.

- **Sensor Measurements under Normal Operation**

$$y(k) = [y_1 \quad y_2 \quad \cdots \quad y_n]^T$$

- **State Estimation:** Consider the system stays the steady-state before attacks.
- **Residual Generation:** The difference between the actual measurement and the measurement estimate

$$r(k+1) = Ce(k+1) + \alpha(k+1)$$

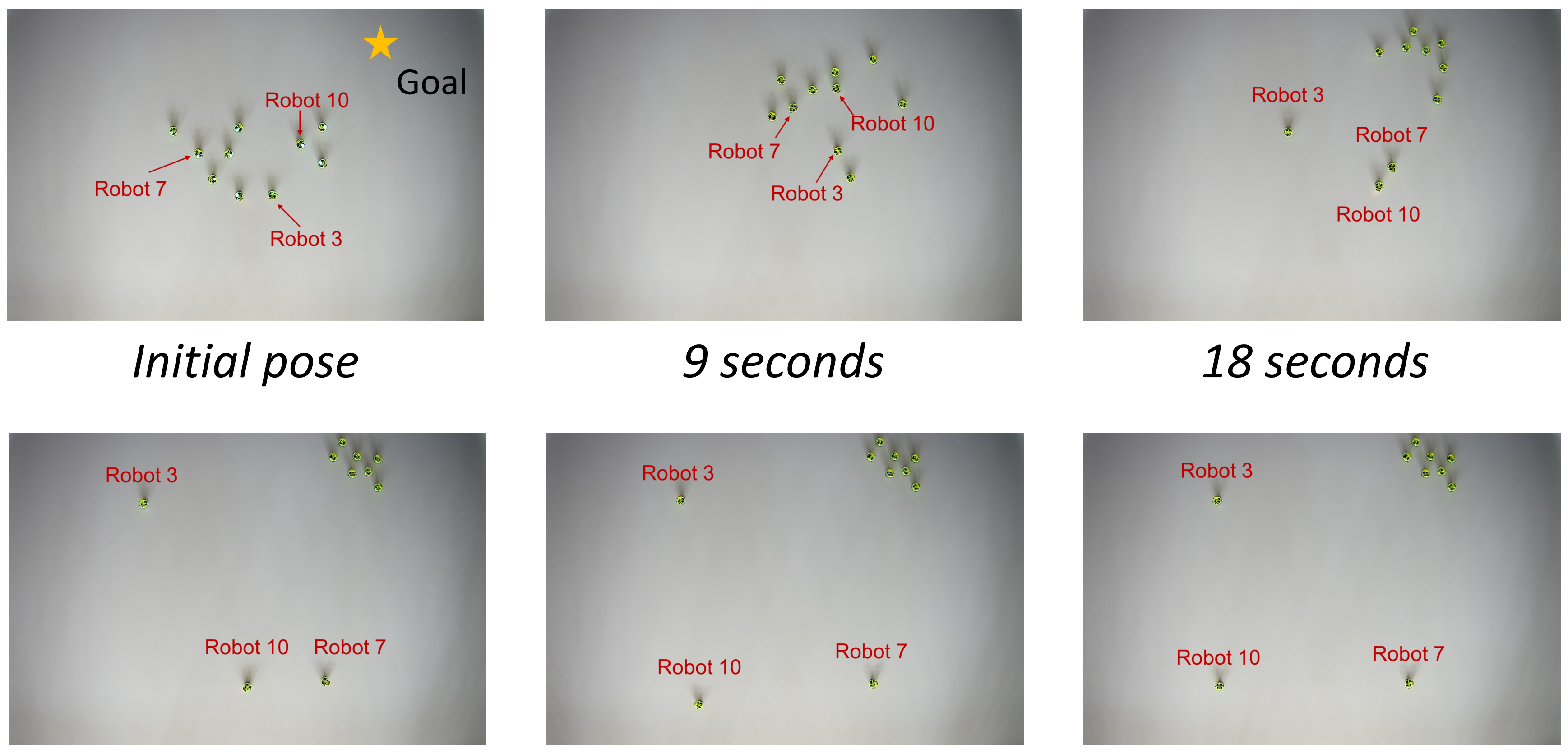
- **Hypothesis Testing**

$$\mathcal{H}_0\colon r(k+1) \backsim \mathcal{N}(0,\Sigma)$$
$$\mathcal{H}_1\colon r(k+1) \nsim \mathcal{N}(0,\Sigma)$$

> *If the system is under normal operation, the mean of the residuals will be* ***zero*** *otherwise* ***non-zero****.*

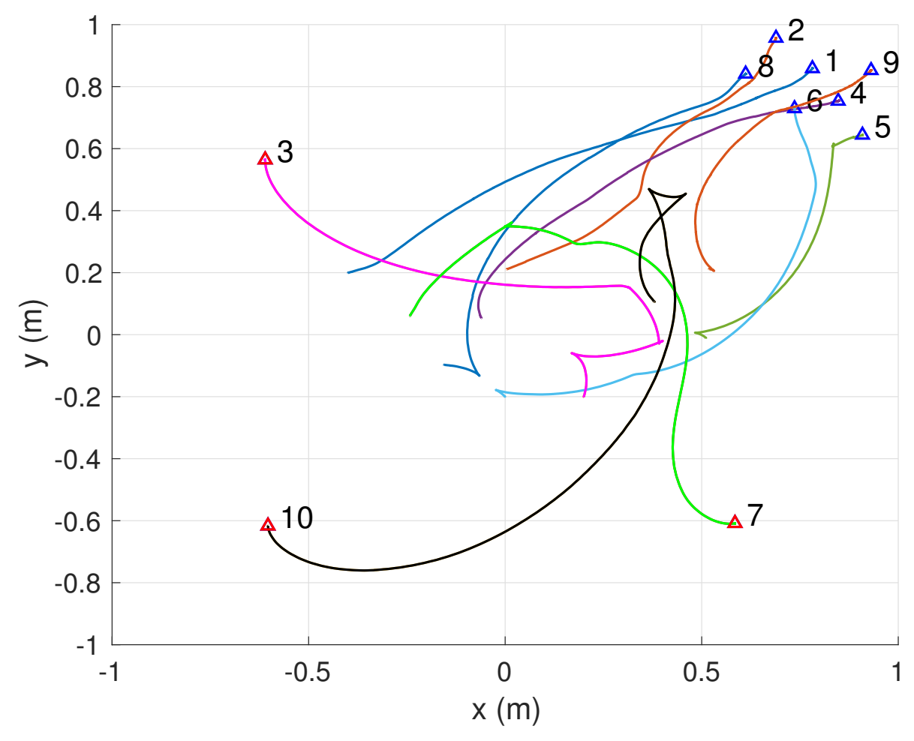
Experimental Results

- 10 robots are spawned at random pose with a global goal of reaching one common destination (top right corner).
- Attacks are injected into the pose measurements of 3 arbitrarily selected agents when the global clock reached 7 seconds.



Initial pose *9 seconds* *18 seconds*

27 seconds *36 seconds* *Final pose*



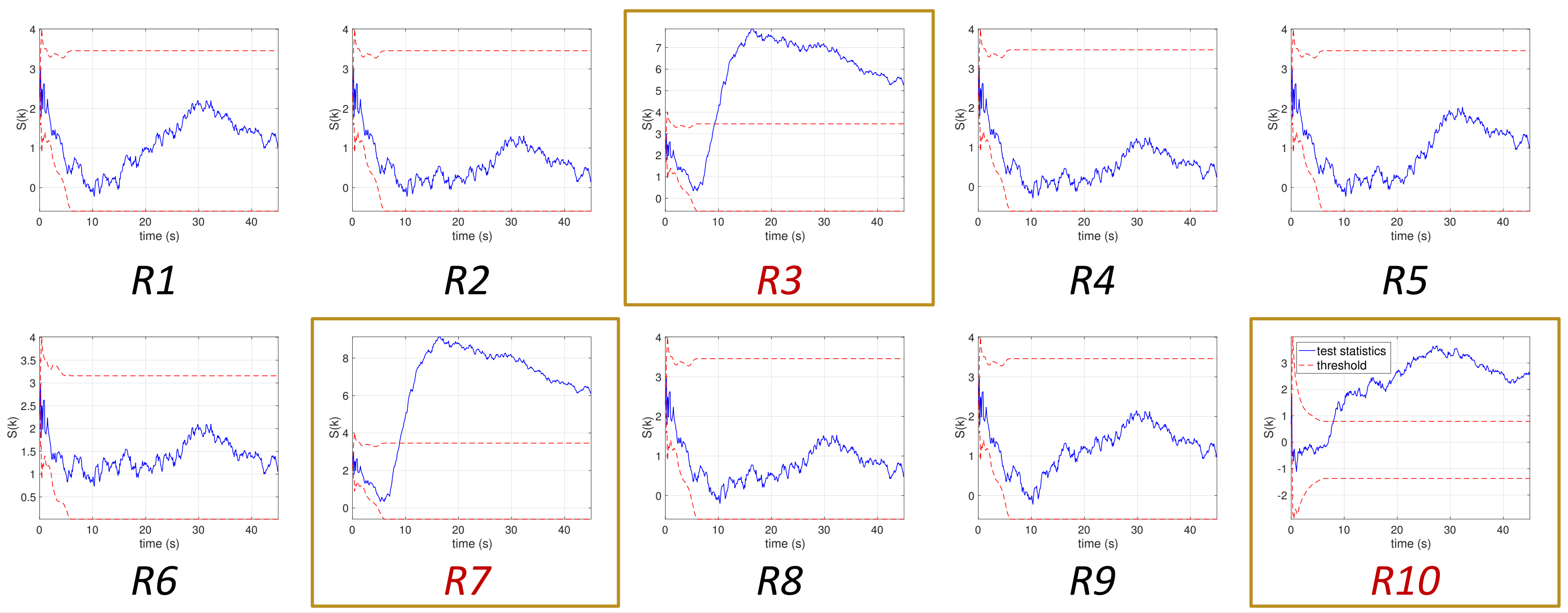
Actual trajectories

- Injected attacks caused the failure of Robots 3, 7, and 10 while other robots succeeded.

> *How does a compromised agent identify any abnormal behavior itself?*

Discussion

- The proposed detector in each robot was able to determine that there was an attack when the residual went above the upper limit of the threshold.



References

Lee, S., Cho, Y., & Min, B. C. (2017). Attack-aware multi-sensor integration algorithm for autonomous vehicle navigation systems. In *Systems, Man, and Cybernetics (SMC), 2017 IEEE International Conference on* (pp. 3739-3744). IEEE.

Lee, S. & Min, B. C. (2018). Distributed Direction of Arrival Estimation-aided Cyberattack Detection in Networked Multi-Robot Systems. In *Intelligent Robots and Systems (IROS), 2018 IEEE/RSJ International Conference on*. IEEE. (under review)

*This work is sponsored by fellowship 2017-R2-CX-0001 from the U.S. National Institute of Justice.

PURDUE UNIVERSITY Discovery Park

CERIAS