# A Demonstration of Mitigation of Sensor-based Deception Attacks on Cyber-Physical Systems

Brien Croteau, Aksel Thomas, Palkpoom Pongsuphat, and Deepak Krishnankutty
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Email: {croteau1, aksel1, palkpon1, deepakk1}@umbc.edu

***Abstract*—Networked control systems improve the efficiency and availability of Cyber-Physical Systems (CPS). However, this also means the safety and stability of the entire plant are more vulnerable to malicious agents. This demonstration showcases how CPS systems can be designed or retrofitted to have resilience to cyber attacks. Specifically, this demonstration employs replay attacks launched from compromised sensors. The closed-loop control feedback signal is constructed by weighted consensus of estimates of the process state gathered from other interconnected processes. Observers are used to generate the state estimates from the data of connected processes. Side-channel monitors are attached to each primary sensor in order to assess proper code execution. These monitors provide estimates of the trust assigned to each observer output that are used as weights in the consensus algorithm. As a result, the augmented system can be shown to be more resilient to attacks and still operate safely, even if a majority of the sensors have been compromised.**

## I. Introduction

Cyber-Physical Systems (CPS) rely on the tight integration of computing, communication, and control. Enabled by advancements in all three areas, greater efficiencies can be obtained by eliminating multiple interfaces of traditional control systems. This does open a larger attack surface with more security vulnerabilities.

This demonstration is intended to showcase the results of joint research between U.S. Naval Academy (USNA) and UMBC; the latest results [1] will be published in the upcoming American Controls Conference (ACC) this June in Milwaukee, WI. Since there is likely not much overlap between attendees of ACC and HOST, our group wanted the chance to also share their findings with the hardware security community. The plots shown in this proposal are taken from data gathered as part of that paper. The demonstration proposed for HOST will be a stand-alone free-running system that utilizes the framework outlined in [1].

## II. Previous Research

Fig. 1 demonstrates the concept and an overview of the approach. Using data from its paired sensor, the function of each observer is to estimate the state of other processes connected to its own. Additionally, side-channel monitors

This work was supported by the U.S. Office of Naval Research under Award N00014-15-1-2179.

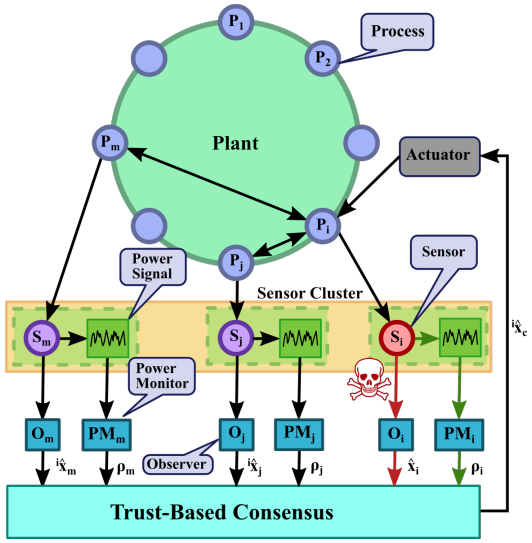


Fig. 1. Diagram of a Trust-Based Framework. Each sensor's processing unit is surveilled by a side-channel power monitor device that determines a trust quality metric. Both this trust metric and the observed state estimates are simultaneous periodic inputs to a trust-based consensus algorithm that determines the feedback to the control process to keep the plant in a safe operating environment despite sensor $S_i$ being attacked. [1]

($PM_x$) analyze the power consumption profile of the sensors' processing units [2]. By monitoring the code execution at the microprocessor level of sensors $S_x$ using this side-channel, the trustworthiness of each sensor can be evaluated and their outputs weighted appropriately. This is done by comparing how different a measured power signal is from a reference good waveform using a correlation measure shown in (1).

$$\rho_i = \frac{\sum_{n=1}^{N_{\max}} a(n) \cdot b(n)}{\sqrt{\sum_{n=1}^{N_{\max}} a^2(n) \cdot \sum_{n=1}^{N_{\max}} b^2(n)}} \tag{1}$$

These weights are used with the local estimates in a trust-based consensus algorithm to feed back the control input ${}^i\hat{x}_c$ and keep the system in a safe operating range. Using such a cross-layer approach (micro-level side-channel analysis and process-level consensus), previous work by the authors and collaborators have shown that single-point failures can be avoided when a compromised sensor injects false data [3]–[5].

## III. Experiment Design

The temperature control testbed, similar to the one shown in Fig. 2, is comprised of six TI LM34 temperature sensors spaced diagonally on an aluminum plate. Sensor $S_6$ lies directly on top of the thermoelectric heater, which simulates the correcting element or actuator. The remaining sensors are diagonally spread on the plate. Each sensor is connected to a TI MSP430 which communicates the temperature measurements along an SPI bus at 20 Hz. An additional mbed NXP LPC1768 micro-controller acts as the SPI master and communicates the temperatures over a serial connection to a PC running MATLAB to close the control loop in near real-time.

Two different code sequences will be run in each sensor micro-controller, and the power supply current traces will be captured. In the baseline code sequence, the actual ADC output are used for all the measurements. For the *replay attack*, prior ADC values are sampled and stored in memory, and seven out of every ten measurements use the stored values instead of the correct ADC readings.

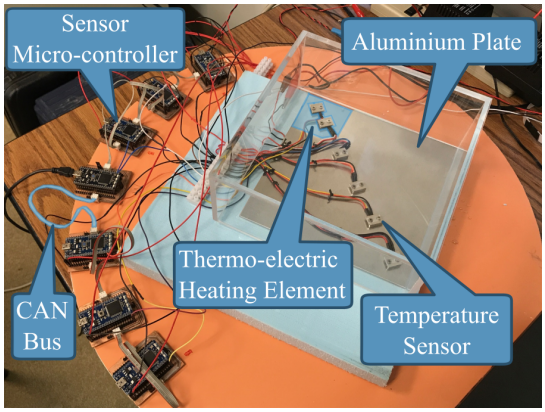


Fig. 2. Testbed consisting of six LM34 temperature sensors mounted on an aluminum plate connected to six micro-controllers. The heating element was located in the upper left corner of the plate. [5]

In this demonstration, a simplified control scheme running at 0.4Hz will be employed that will toggle power on and off of the heating element to maintain the temperature of the sixth node at 90°F.

Different timed, multi-attack scenarios will be conducted. $S_6$, the sensor right on top of the heating element, will always be the first attacked; then other sensors chosen at random will be attacked at regular intervals.

## IV. Demonstration

As a live and interactive demonstration, this testbed will show the closed loop temperature experiment in near real-time. A new more tightly integrated testbed is planned to be brought to the conference that is evolved from the one used in [1]. It will allow interested conference attendees to see how side-channel power measurements can be used to determine if code running on a micro-controller has been modified. It will also show a visualization of the temperature

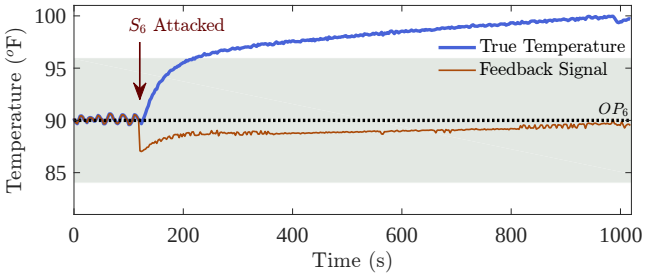


Fig. 3. Temperature Control using a Single Sensor to Close the Loop. The feedback signal for the controller was the output of the process's sensor $S_6$, which was attacked after 120 s. The lower temperature reported by the attacked sensor (red) caused the heater to remain on and drive the node's true temperature (blue) outside its operational normalcy region of 90°F±6°F. [1]

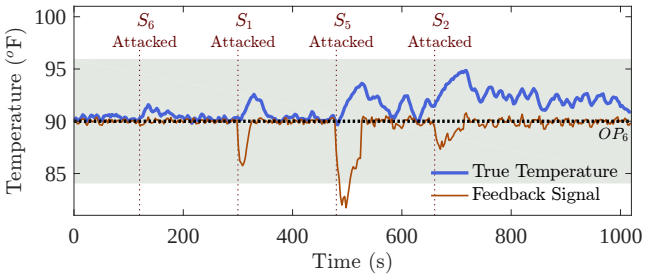


Fig. 4. Temperature Control using the Trust-Based Weighted Consensus Algorithm to Close the Loop. After each sensor-based attack, there was a small performance degradation yet the system maintained operational normalcy. [1]

measurements of this simplified HVAC system and the results of two different control laws, one that closes the loop based on the information passed directly from the sensors and one that uses the weighted consensus scheme to provide cyber resilience, resulting in responses similar to the graphs shown in figures 3 and 4. Due to heating/cooling dynamics of the plate being on the order of ten minutes, the demonstration will plan to also be able to display previously recorded data so that both control laws can been exhibited in a few minutes.

## References

[1] T. Severson, E. Rodriguez-Seda, B. Croteau, D. Krishnankutty, R. Robucci, C. Patel, N. Banerjee, and K. Kiriakidis, "Trust-based framework for resilience to sensor-targeted attacks in cyber-physical systems," in *2018 American Control Conference (ACC)*, June 2018, (to appear).

[2] D. Krishnankutty, R. Robucci, N. Banerjee, and C. Patel, "Fiscal: Firmware identification using side-channel power analysis," in *2017 IEEE 35th VLSI Test Symposium (VTS)*, April 2017, pp. 1–6.

[3] E. J. Rodriguez-Seda, T. Severson, and K. Kiriakidis, "Recovery after attacks of deception on Networked Control Systems," in *Proceedings of the 9th International Symposium on Resilient Control Systems*, 2016, pp. 109–114.

[4] B. Croteau, D. Krishnankutty, R. Robucci, C. Patel, N. Banerjee, K. Kiriakidis, T. Severson, and E. Rodriguez-Seda, "Cross-level Detection of Sensor-based Deception Attacks on Cyber-Physical Systems," in *Proc of the 7th Annu IEEE Int Conf on CYBER Technol in Auton, Control, and Intell Syst*, August 2017.

[5] B. Croteau, D. Krishnankutty, K. Kiriakidis, T. Severson, C. Patel, R. Robucci, E. Rodriguez-Seda, and N. Banerjee, "Cross-level detection framework for attacks on cyber-physical systems," *Journal of Hardware and Systems Security*, vol. 1, no. 4, pp. 356–369, Dec 2017. [Online]. Available: https://doi.org/10.1007/s41635-017-0027-9