Safeguarding Sensor Device Drivers Using Physical Constraints

Gregory Brooks*, Youchao Wang†, and Phillip Stanley-Marbell

University of Cambridge

Introduction: Software that interfaces with hardware peripherals often makes assumptions ranging from the formatting of peripheral interface protocol packets to assumptions on the values in packet fields. When these assumptions are false, malicious or faulty peripherals can compromise system integrity.

For sensor peripherals, physics imposes constraints on plausible protocol fields values. Software interfacing with sensors can use such information on physical constraints to check the validity of sensor output. We present examples of constraints on sensor signals along with a method for inferring the likelihood of a transduction attack in device drivers.

Sensors can generate physically-implausible values and such erroneous sensor outputs can cause software to make implausible inferences. Transduction attacks [2] are repeatable methods for forcing erroneous sensor output. Despite their increasing importance, transduction attacks are not well understood [5]. The implications of these challenges are broad and range from autonomous vehicles to medical devices and mobile / wearable devices. Recent work by Payer [4] shows how malicious USB devices are the cause of several zero-day in-kernel buffer overflow attacks. Transduction attacks may similarly cause such vulnerabilities. New techniques for validating the output of sensors could improve safety in computing systems in the presence of both benign or malicious erroneous sensor output. We present preliminary work on using information about the physical constraints on signals to validate sensor output for improved device driver security.

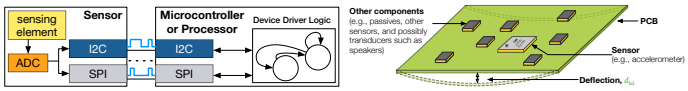
Contributions: Figure 1(a) shows an example of a system comprising a sensor and a microcontroller connected by one of the common serial interfaces such as I2C and SPI. The microcontroller obtains measurements from the sensor by reading device registers in the sensor and makes control decisions based on the values read. Figure 1(b) shows an example of a physical structure on which an accelerometer is deployed based on a recent transduction attack on accelerometers [1].

For the structure in Figure 1(b) , let Q_{bd} be the quality factor of a structure, let d_{bd} be the deflection of the structure (annotated in the figure) and let ω be the frequency of a pure sinusoidal audio signal being used in a transduction attack. Then, the acceleration a at resonance is given [1] by:

$$a = Q_{bd} \cdot d_{bd} \cdot \omega^2 \cdot \sin(\omega t). \quad (1)$$

Figure 1(c) shows Equation 1 encoded in Newton [3] (using a Taylor series to represent $\sin(2\pi ft)$).

Threat model: We assume an operating system collects samples from sensors as needed (e.g., for determining screen orientation) and that it stores the values it reads over time. As a result, the collected samples from a sensor such as an accelerometer are random samples of the time-varying measurand. Because signals sampled on a Fourier basis can be represented as a sum of sinusoids, we assume without loss of




```
include "NewtonBaseSignals.nt"
board : invariant(Q:dimensionless, d:distance, f:frequency, a:acceleration, t:time) = {
  a ~ Q*d*(2*kNewtonUnitfree_pi*f)**2*((2*kNewtonUnitfree_pi*f*t)
      - (2*kNewtonUnitfree_pi*f*t)**3/6 + (2*kNewtonUnitfree_pi*f*t)**5/120)
}
```

Figure 1: (a, top left): Sensors provide data consumed by device drivers. (b, top right): Physics imposes constraints on sensor signals. (c, bottom): We can represent physical constraints in the Newton physics specification language [3].

generality that the transduction attack signal is a pure sinusoid.

Approach: Random samples of a pure sinusoidal signal can be modeled with a bimodal Beta distribution. Random samples of the signal a in Equation 1, derived from the sinusoidal transduction attack signal, can be modeled with a unimodal Laplace distribution, $L(\mu, b)$ where μ and b are functions of the structural properties in Figure 1(b).

We can use this information about the distribution induced on random samples by the physics of structures, for formulating a Bayesian likelihood function. Given a batch of N sensor readings obtained by the operating system over time, we can compute the likelihood, $\Pr\{X = x \mid \Theta\}$, of the samples given the assumption that a transduction attack has occurred:

$$\Pr\{X = x \mid \Theta\} = \prod_{i=1}^{N} \Pr\{X = x_i\} = \prod_{i=1}^{N} \frac{1}{2b} e^{-\frac{|x_i - \mu|}{b}}. \quad (2)$$

Our goal is to automate inferring this likelihood in drivers.

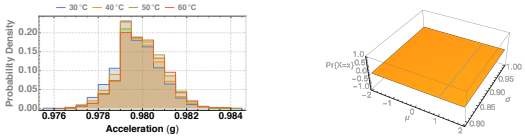


Figure 2: Left: Acceleration measurements (the x_is in Equation 2). Right: The likelihood function ($Pr\{X = x \mid \Theta\}$).

Preliminary Results: Figure 2(a) shows noise characteristics of the output of an MMA8451Q accelerometer under controlled temperature operating conditions and in a vibration-isolated environment. Using the data in Figure 2(b) and a probability distribution for the likelihood derived from Equation 1 in Equation 2, Figure 2(b) shows the likelihood function. This function quantifies the credibility that the output of a sensor is the result of a transduction attack.

References

[1] Analog Devices. Analog devices advisory to ICS ALERT-17-073-01. Technical report, 2017.

[2] K. Fu and W. Xu. Risks of trusting the physics of sensors. *Commun. ACM*, 61(2):20–23, Jan. 2018.

[3] J. Lim and P. Stanley-Marbell. Newton: A language for describing physics. http://arxiv.org/abs/1811.04626, 2018.

[4] M. Payer and G. Kroah-Hartman. USB: check `usb_get_extra_descriptor` for proper size. Linux kernel mailing list.

[5] G. Zhang et al. Dolphinattack: Inaudible voice commands. CCS '17, pages 103–117. ACM, 2017.

*M.Eng. student, will be presenting.

†M.Phil. student.