The massive scale and decentralized nature of the IoTT provide attackers with a large attack surface for exploitation.

Raising Awareness of Security Challenges for the Internet of Trillions of Things

John A. Stankovic and Jack Davidson

John A. Stankovic

Jack Davidson

Computer security problems have evolved over the last 50 years from a minor concern to major operational risks. Every day new devices are added to the Internet of Things (IoT). Conservative projections have 50 billion devices on the internet by 2020, but—with autonomous vehicles, smart phones, smart wearables, smart cities, numerous other smart applications, and nanotechnology—we foresee an "Internet of Trillions of Things (IoTT)" before long. If computer security problems are formidable now, consider when there is an IoTT!

The proliferation of devices and applications will give rise to many new complications and research challenges, especially in cyberphysical system (CPS) security. Because the smart devices of IoTT systems will be so numerous and easily accessible, and will interact directly with the physical world (including humans), they will exhibit tremendously large attack surfaces with increasing types and numbers of vulnerabilities. Attacks on these systems may cause the inoperability of major infrastructures such as transportation or energy, great financial losses, and many other negative impacts, even death.

CPS technology is required to build IoTT systems. It is therefore necessary to increase awareness of CPS security problems and to address them

John Stankovic is the BP America Professor and Jack Davidson is a professor, both in the Department of Computer Science at the University of Virginia.

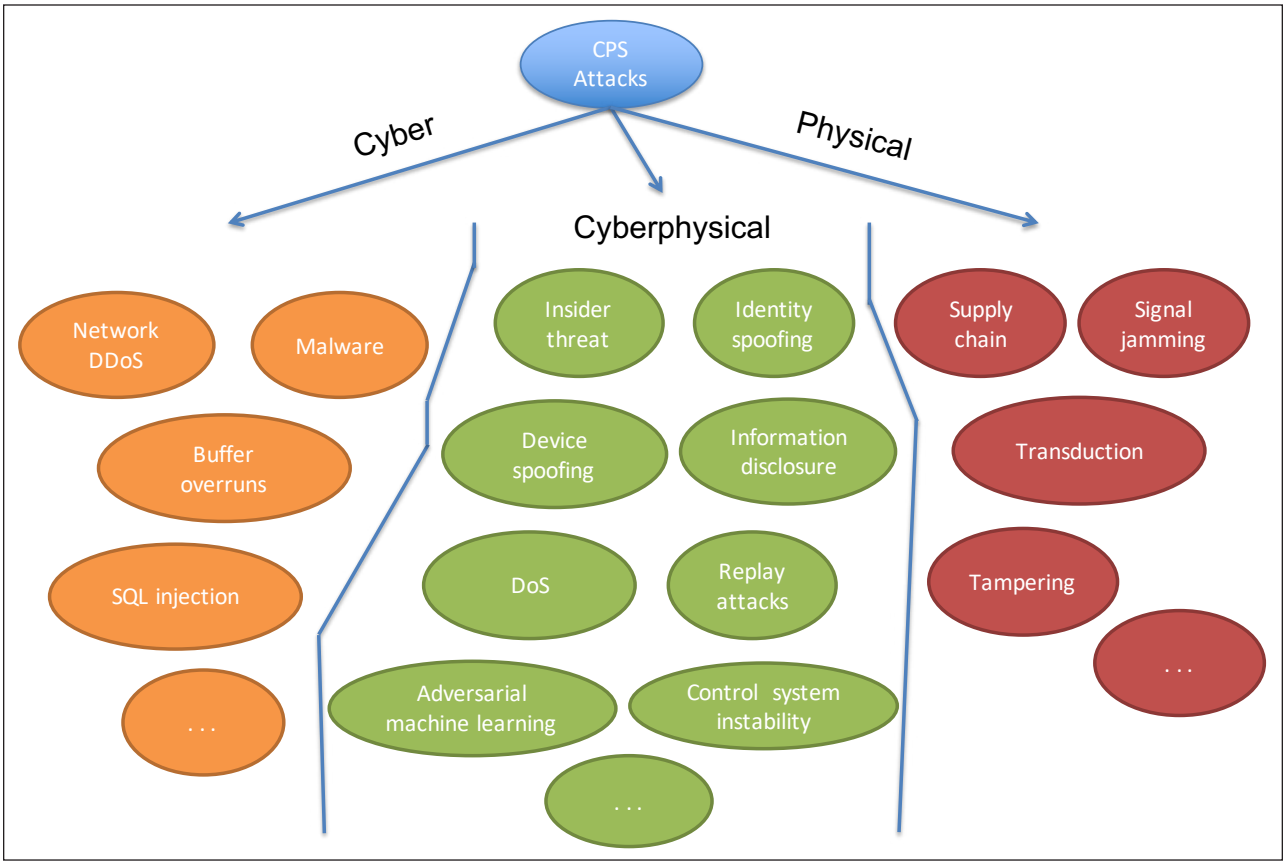

FIGURE 1 Taxonomy of sample attacks on cyberphysical systems (CPS). DDoS = distributed denial of service; DoS = denial of service; SQL = structured query language.

before trillions of devices are deployed with insufficient security protections.

In this paper we define a CPS security attack taxonomy based on new problems for the cyber, physical, and cyberphysical aspects of the IoTT, and discuss examples of new security problems for the physical and cyberphysical areas. For each example, we explain potential consequences and present approaches to solutions. We do not discuss cyberattacks, such as network distributed denial of service and malware, as they are well covered by existing literature.

CPS Attack Taxonomy

Cyberphysical systems are engineered systems that are built from, and depend on, the seamless integration of computational and physical components. They are often connected to the internet, with applications in many domains, such as smart cities, health care, transportation, energy, emergency response, agriculture, and defense. CPS attacks can thus have serious and harmful physical results (e.g., an air bag being activated during normal driving). This characteristic sets them apart from traditional IT systems, where confidentiality is usually the most important property.

Figure 1 illustrates a taxonomy of CPS attacks. It does not show a complete range of attacks but rather provides examples in three categories: cyber, physical, and cyberphysical.

At the cyber layer, the IoT is susceptible to many common types of attacks, such as malware and distributed denial of service. A particularly serious concern with the IoTT will be the timely application of critical patches and system updates, which often require temporarily disabling system security protections. Given the massive scale and decentralized nature of the IoTT, this update process provides an attacker with a window of vulnerability and a large attack surface for exploitation.

Cyberphysical attacks combine software intrusion/alteration with effects on the physical aspects of a system. Attacks by insiders, information disclosure, replay, and denial of service (DoS) have been common and can now also be applied to the physical aspects of

the IoT. Identity and device spoofing and control system instability are newer attacks, created—or significantly increased in frequency—with the emergence of the IoT.

Many novel types of attacks have appeared at the physical layer of systems. These include ways to disrupt systems by attacking different steps in the supply chain. Attacks based on transduction (discussed below) are particularly debilitating. Physical tampering is a risk because smart devices often operate in open environments. And because most smart devices communicate wirelessly, they are vulnerable to jamming.

Physical Attacks on the IoT

Supply Chain and Other Tampering

Many security solutions assume that hardware is trustworthy, but with the common practice of outsourcing the construction of hardware platforms, the supply chain can be a source of security attacks (Ray et al. 2018). Trojan horse circuits and embedded software might be included in delivered products to cause harm or surreptitiously transmit data to an adversary. For example, an internet router might not only transmit packets to the intended destination but also send copies to the adversary.

Many companies automatically collect data for maintenance and performance control. Such data may be a source of security attacks.

Compounding the problem is that many companies, such as those involved in communications and the manufacture of printers, automobiles, and aircraft, want to (or already do) automatically collect data for maintenance and performance control. Such data may be a source of security attacks.

Supply chain attacks can affect all parts of society and almost all applications: financial records and company secrets can be stolen, automobiles and planes can be made to crash. New tools are needed to validate that delivered hardware/software platforms do not include hidden circuits or embedded software that can cause attacks.

With billions (or more) of IoT devices and easy access to them, other types of tampering are also possible. For example, an attacker can physically move sensors to an unwanted location, point a fixed-direction camera in the wrong direction, impede an actuator from its full range of motion, or jam wireless communications. These changes may result in fires not detected, missed detection of a serious crime at a previously monitored location, safety doors that don't close properly, or complete inaccessibility of an IoT application system.

Solutions for such attacks must be developed or improved. Unwanted movement of devices could be detected with additional motion sensors or accelerometers. Correlation among data from a set of sensors could reveal that one sensor has been moved away from the others. Related sensing modalities (e.g., temperature, pressure, and volume sensors relate to each other by physics) could be used to detect attacks on one modality. And frequency hopping, spread spectrum, and other techniques can provide some resilience to the jamming of wireless communications.

Transduction Attacks

One especially complex class of attacks for smart devices is transduction attacks, which exploit the physics and unintended functions of circuits and sensors to alter a sensor's output (Fu and Xu 2018).

Manipulation through Voice Recognition

The Dolphin Attack (Zhang et al. 2017) uses an inaudible sound wave to trick a speech recognition system such as Siri, Google Now, or Alexa into taking action that was not requested. It takes advantage of the fact that, while microphones are built to primarily hear the human voice, they can also detect (unintended) inaudible sounds, making them vulnerable to transduction attacks (Roy et al. 2017).

Because more and more IoT systems have or are developing voice interfaces, the consequences of transduction attacks are unbounded. The attacks may permit illegal entry to a location, open the door locks of a home or business, or provide harmful advice via a medical cognitive assistant.

Solutions must include better frequency filters and signal processing to avoid the appearance of ultrasound resonances in the voice frequency range.

Backdoor Coupling

Another type of transduction attack is called backdoor coupling: signals enter a system indirectly via coupling between wires or components, so a sensor designed to detect one modality may be activated by another. For example, it was shown that playing sounds embedded in a YouTube video allowed an adversary to control a smartphone's accelerometer, thanks to a mechanical coupling between the speaker and the resonant frequency of the sensor (Fu and Xu 2018). This can have negative consequences for any application that relies on the accelerometer, such as a step counter, dead reckoning location estimators, or a sensor that monitors an elderly person's level of activity.

Most security solutions are ineffective for transduction attacks because they are designed for digital risks, not analog. Circuits must be manufactured to reduce the effects of resonance, increase the frequency of checking sensor output by software, and enhance the layout of components on system boards to minimize unwanted coupling.

Cyberphysical Attacks

Types of security attacks that are exacerbated in the IoT are denial of service, spoofing, adversarial machine learning, and control system attacks.

Denial of Service

A DoS attack diminishes or eliminates a system's capacity to perform its expected function. Hardware failures, software bugs, resource exhaustion, environmental conditions, or any complicated interaction between these factors can cause a DoS. When a bug is detected attackers often use the internet to initiate DoS attacks.

As an example of a DoS in communications, attackers may induce a collision in only one byte of a transmission to disrupt an entire packet. A naïve link-layer implementation may attempt retransmission repeatedly, culminating in the exhaustion of batteries in one or more connected smart devices and disrupting system function, such as transportation monitoring in a smart city.

Solutions must detect such attacks and create power-aware smart devices that avoid using all the power when under attack.

Spoofing

In a spoofing attack, a person or program successfully masquerades as another to gain an illegitimate advantage. With the IoTT it could also be a smart device.

The attacking device can act as a data source, presenting fake data streams to the system, or even pretend to be a particular person and issue commands as if it were that person. Spoofing can cause an IoT system to produce wrong, often safety-critical, results. For example, an IoT-based process control plant may be erroneously informed that chemical vats are overheating and shut them down, causing loss of revenue, or indicate that they are operating well when they are not, resulting in overheating and even explosion.

Adversarial machine learning can cause misidentification of a stop sign, with severe consequences for a self-driving car.

Solutions require standard techniques such as authentication, encryption, and anomaly detection, as well as IoT-specific measures that coordinate among properly acting smart devices. The use of redundant sensors that are not accessible through the internet may also prove useful.

Adversarial Machine Learning

Machine learning (ML) is essential to the functionality of many cyberphysical systems. For example, autonomous vehicles use deep neural networks (DNNs) to detect, identify, and locate objects in the environment and navigate the vehicle. Deep learning algorithms are also used to analyze and recognize speech for voice-activated cyberphysical systems and to recognize people and objects in security systems.

The extensive use of ML algorithms has enabled a new type of CPS attack: adversarial machine learning. Attackers can develop adversarial inputs with small (even imperceptible) perturbations that cause a trained ML model to misclassify an object, such as a road sign (Eykholt et al. 2018); misidentification of a stop sign, for example, could have very serious consequences for a self-driving car.

Adversarial ML attacks can also target speech recognition systems (Carlini and Wagner 2018) and image recognition systems (Kurakin et al. 2017). A targeted

attack causes the ML system to assign a specific (i.e., targeted) label to the object; for example, an attack on a voice-activated command and control system would enable a command of the attacker's choosing. Similarly, a targeted attack on a facial recognition system that matches faces against a whitelist of approved people (e.g., to control access for large events) could admit an unknown, potentially malicious person.

Many cyberphysical systems (e.g., self-driving cars, voice-activated command and control systems) can easily be acquired legally or illegally by attackers to carry out black box attacks. A black box attack is defined as one that does not have direct access to the underlying ML model (as in a white box attack) to develop sophisticated attacks; the attacker can only deduce the CPS operation by providing specific inputs and observing the output. Black box attacks on complex systems are more difficult than white box attacks, but they have been demonstrated in a number of domains.

A chemical plant control system that becomes unstable might rapidly open and close a critical valve, causing it to fail.

Adversarial ML attacks are a relatively new CPS threat. Research is needed to understand them and to build robust ML models that are not susceptible to adversarial manipulation of the physical artifacts (e.g., signs, images, audio) that are inputs to the system.

Control System Attacks

Control systems are a critical component of many types of cyberphysical systems. Examples include industrial control systems, supervisory control and data acquisition systems, autonomous vehicles, and medical devices.

Instability and Physical Damage

In a control system attack, the adversary seeks to move the system from a region of stability to one of instability where control outputs may fluctuate arbitrarily and exceed normal operating parameters. For example, gain scheduling is often used to control nonlinear systems. Essentially, the system is controlled by a family of linear controllers or gains, each of which is designed for a particular operating region. Gain scheduling attacks can be effected using techniques such as sensor spoofing and denial of service.

Consider an autonomous aerial vehicle (UAV). It uses carefully constructed gains for operating modes such as takeoff, landing, cruising, or hovering. By tampering with the inputs, an attacker can cause a transition from a gain that is appropriate for the UAV's operating mode to one that is inappropriate. For example, when the UAV is hovering, a change in the gain computed for cruising could result in loss of the UAV.

The ability to spoof a sensor reading or to delay receipt of a signal opens the possibility of a control system instability attack. Many control systems are designed assuming that sensor readings are within certain operating thresholds and that the communication channel to send both data and control signals operates as intended. By sending carefully constructed inputs using a replay attack, the adversary may make the control system unstable—and the system could enter a state where returning to a stable state is not possible.

Similarly, using DoS techniques, an attacker could cause a control system to become unstable by delaying packets that contain control information needed to stabilize a system. The instability could result in severe oscillations that could cause physical damage. For example, a chemical plant control system that becomes unstable might begin rapidly opening and closing a critical valve, causing it to fail.

Unlike purely cyber systems, cyberphysical systems open the door to attacks that cause severe physical damage—on par with the damage caused by kinetic weapons. However, unlike kinetic attacks, CPS attacks can be stealthy and precise identification of the attacker is often difficult. These attacks require a deep understanding of the physics of both the process being controlled and the logic that controls the equipment. A well-publicized example of a CPS control system attack is the Stuxnet attack on the Iranian Natanz enrichment facility (Langner 2013).

Sophisticated stealthy attacks on physical infrastructure can also exploit the physics of the process being controlled. A compelling example is an attack on an industrial pump in which the attack payload is a stream of cavitation bubbles created via malicious control of an upstream valve. Over time the stream of

bubbles will pit the pump's impellers and eventually cause the pump to fail (Krotofil 2017).

Protective Approaches

The attacks described in this section rely on malicious inputs that disrupt or hijack the control system. Standard software engineering techniques, such as rigorous testing, can and do reduce the threat surface, but they often do not provide the necessary coverage—especially for a complex system with a variety of sensors and actuators.

An approach called *fuzzing* shows promise in uncovering CPS vulnerabilities not found by traditional techniques. In fuzzing, inputs are automatically generated to force coverage of unexplored code (Miller et al. 1990). The technique is particularly applicable to cyberphysical systems where the range of possible inputs is difficult to enumerate or bound.

For stealthy physical attacks, solutions include redundant sensors and consistency checks to detect a deteriorating system.

Summary

The CPS-based IoTT presents an enormous increase in potential attack surfaces. Many of these systems will interact with humans, further expanding the attack surface and resulting in significantly more vulnerabilities and potential negative impacts on society. If the past is any harbinger of the future, the security attack-solution competition will continue. Highly inventive attackers will exploit the physical, cyberphysical, and cyber layers to their advantage. Developers of smart devices and smart applications must be aware of new classes of potential attacks and build solutions for them as first principles, not only after problems are uncovered.

Some new solutions are promising, but of course they will be useful only if actually implemented. Diverse techniques have proven helpful in computer security and should also help in the IoTT. However, too often speed to market, cost in dollars, or the benefits of homogeneity keep the development of devices and systems from incorporating known security solutions. If these conditions persist and there are trillions of smart devices, there may be widespread chaos and increased risks of physical danger.

References

Carlini N, Wagner D. 2018. Audio adversarial examples: Targeted attacks on speech-to-text. Proceedings, 2018 IEEE Security and Privacy Workshops, May 21–24, San Francisco.

Eykholt K, Evtimov I, Fernandes E, Li B, Rahmati A, Xao C, Prakash A, Kohno T, Song D. 2018. Robust physical-world attacks on deep learning visual classification. Proceedings, 2018 IEEE/CVF Conf on Computer Vision and Pattern Recognition, Jun 18–22, Salt Lake City.

Fu K, Xu W. 2018. Inside risks: Risks of trusting the physics of sensors. Communications of the ACM 61(2):20–23.

Krotofil M. 2017. Evil bubbles or how to deliver attack payload via the physics of the process. Black Hat USA, Jul 22–27, Las Vegas.

Kurakin A, Goodfellow IJ, Bengio S. 2017. Adversarial examples in the physical world. 5th International Conf on Learning Representations, Apr 24–26, Toulon.

Langner R. 2013. To Kill a Centrifuge: A Technical Analysis of What Stuxnet's Creators Tried to Achieve. Dover DE: Langner Group.

Miller B, Fredriksen L, So B. 1990. An empirical study of the reliability of UNIX utilities. Communications of the ACM 33(12):32–44.

Ray S, Peeters E, Tehranipoor MM, Bhunia S. 2018. System-on-chip platform security assurance: Architecture and validation. IEEE Proceedings 106(1):21–37.

Roy N, Hassanieh H, Choudhury RR. 2017. BackDoor: Making microphones hear inaudible sounds. MobiSys, Jun 19–23, Niagara Falls.

Zhang G, Chen Y, Ji X, Zhang T, Zhang T, Xu W. 2017. DolphinAttack: Inaudible voice commands. Proceedings, ACM Conf on Computer and Communications Security, Oct 30–Nov 3, Dallas.