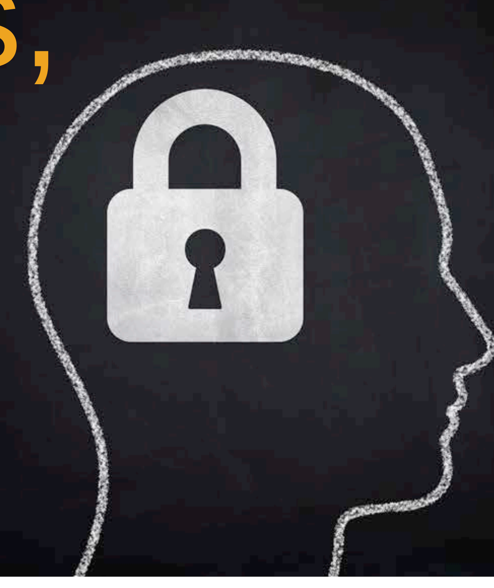
Safety-critical Cyber-physical Attacks: Analysis, Detection, and Mitigation

Hui Lin, Homa Alemzadeh, Daniel Chen, Zbigniew Kalbarczyk, Ravishankar K. Iyer

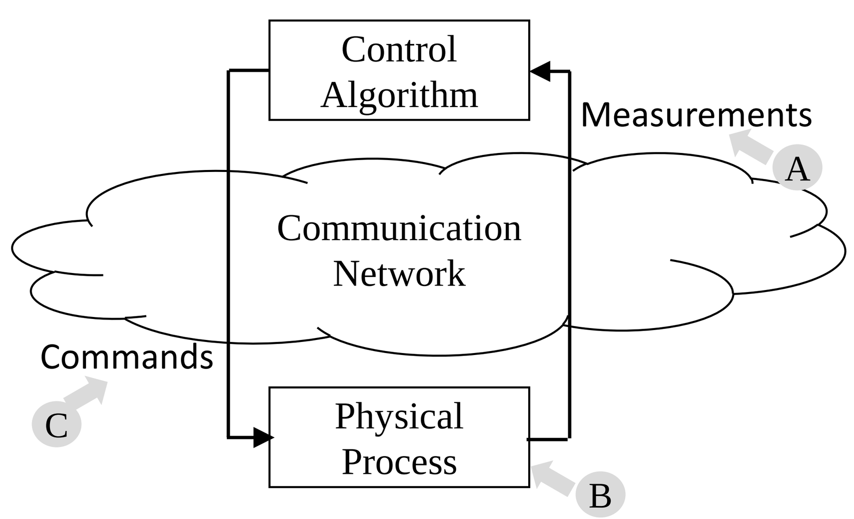


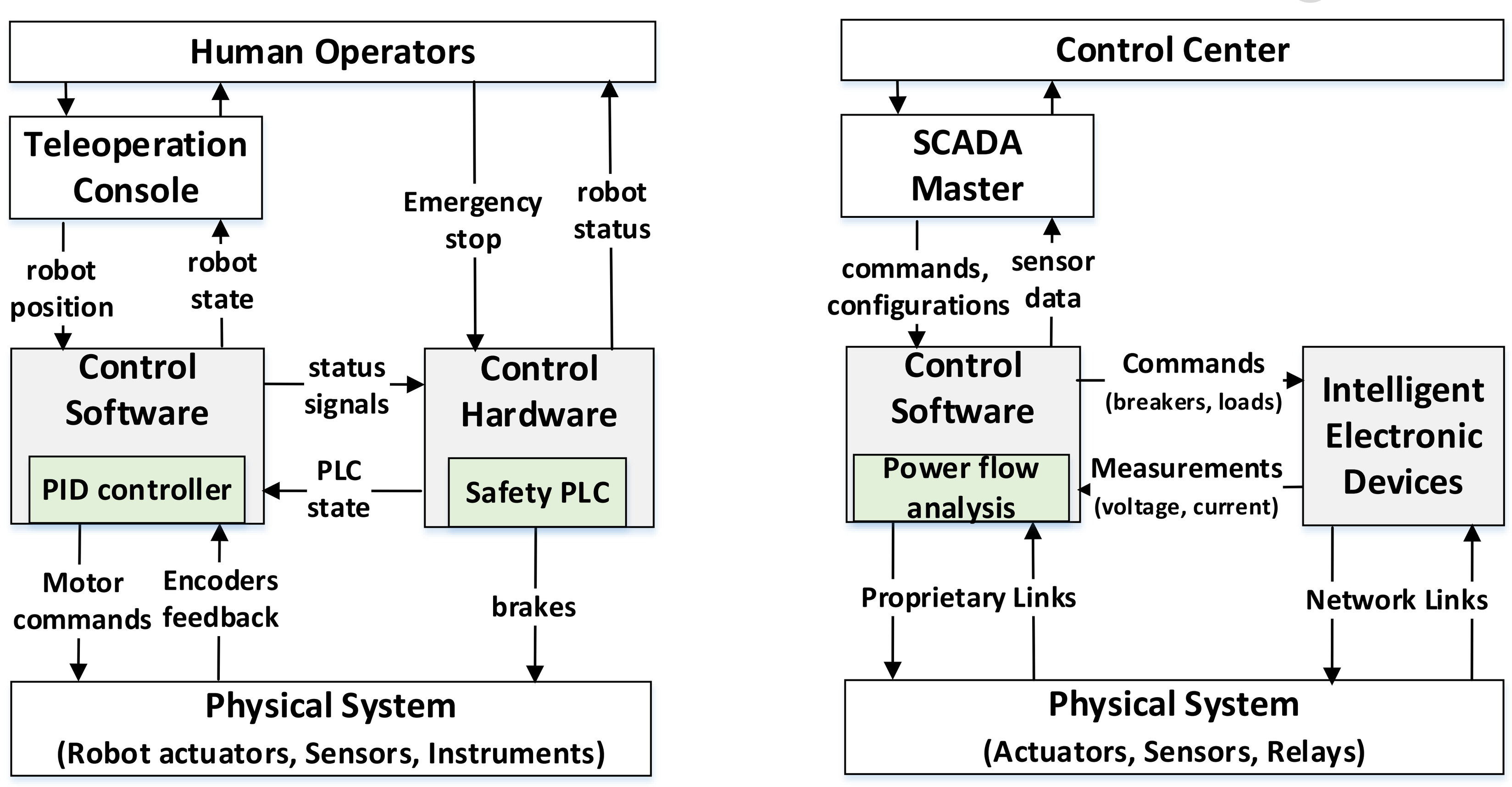
Goal

- Analyze common characteristics of safety-critical attacks for cyber-physical systems
  - Introduce safety violations in physical processes without introducing anomalies in cyber domain
  - Exemplify attacks on two cyber-physical systems: (i) robotic surgical systems and (ii) power grid infrastructures
- Propose a general principle to detect the cyber-physical attacks
  - Integrate the knowledge from both cyber and physical domains

Cyber-Physical Systems

- Feedback control loops
  - Measurements from physical processes used as an input to control algorithms
  - The control algorithms use the estimation of physical state of physical processes to decide the control actions





Example control structures for robotic surgical systems (left) and power grid infrastructures (right)

	Robotic Surgical Systems	Power Grids
Measurements	Robot state, e.g., positions of robotic arms	Current, voltage, and power usage at substations
Commands	Adjust robot positions	Adjust configuration of transmission network
Safety Procedure	PLC monitors system state and controls fail-safe brakes	"*N*-1" contingency analysis ensures power system operation when one device is out of service

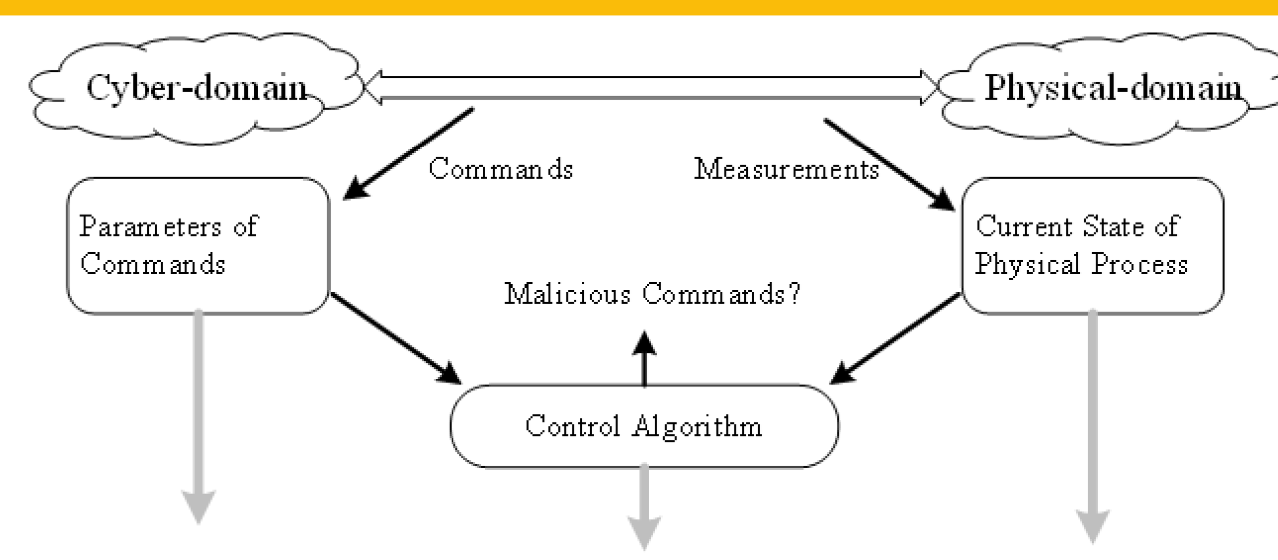
Attack Targets

- Type *A*, false or bad data injection attacks
  - Attackers try to mislead the control algorithms by corrupting the cyber system state
  - Indirectly disrupt control operations or cause economic losses
- Type *B*, perturbations of physical components
  - Identify and rank the attack patterns, to reveal vulnerabilities
  - Require physical access to actual CPS devices (may not be practical in reality)
- Type *C*, malicious modifications of control fields of commands delivered over communication channels to CPS devices
  - Require same privilege as Type A
  - Directly disrupt control operations to perturb physical state
  - Can introduce no anomalies in the control flow and communication protocols

Challenge of Detecting Attacks

	Challenges	Example Cyber-Physical Systems: Power Grids	Example Cyber-Physical Systems: Surgical Robots
Cyber domain	Lack of encryption and authentication mechanisms for legacy devices	Communication is in a plain text.	Leaking of user commands and state information from the unencrypted data transferred through network and serial links.
Cyber domain	Malicious and unsafe commands can be encoded in legitimate formats	Modification of a few bits in network traffic can maintain the correct communication syntax.	TOCTTOU (time of check to time of use) vulnerability allowing malicious modification of the control commands after they are checked by the software and before are communicated to the hardware.
Cyber domain	Inconsistency between the state estimation in the cyber domain and the actual state in physical process.	False data injection attacks on measurements	Lack of complex models for accurate estimation of the system dynamics and behavior of robotic joints in real-time.
Cyber domain	Real-time constraints on control systems	Control operations should be delivered in a few hundred milliseconds.	Real-time constraint of 1 millisecond per control iteration.
Physical domain	Attacks are hard to distinguish from incidental failures and human induced safety hazards.	Contingency analysis evaluates the consequence of incidents, in which one or two physical components are out of service.	Similar safety-critical impact might occur due to unexpected physical failures or unintentional human errors.
Physical domain	Inadequate knowledge of the global system state.	Periodically performing state estimation can detect the consequence of attacks based on the collected measurements. However, it is difficult for each substation to decide the impact of a command on the whole power grid.	There are limited hardware resources on the embedded computational units in the interface and the physical layer of the robot to perform sophisticated computations for estimating system state.

Detection Principle



	Commands	Control Algorithm	Measurements
Common Principle	Increase visibility in the cyber domain	Estimate (ahead of time) the consequence of command execution	Increase the integrity of measurements
Surgical Robotic Systems	Intercept commands sent by control software	Model the robot manipulator dynamics with selected degrees of freedoms (e.g., three degrees of freedoms in our case)	Retrofit hardware interface board (custom USB board) in order to deliver measurements to the detection module.
Power Grids	Integrate network monitors (e.g., Bro) with SCADA protocol (e.g., DNP3 or Modbus) analyzers Classify critical and noncritical commands	Use "*N*-1" contingency analysis to decide relative severity of an attack Dynamically adjust the number of iterations in AC power flow analysis to balance detection accuracy and latency	Compare measurements observed by network monitors placed at different locations on the network in order to validate the integrity of measurements

- Increase the visibility in the cyber-domain, to better understand the interactions between the cyber and physical components
- Use the knowledge of physical domain to estimate the real impact of attacks on the CPSs.
- Integrate control algorithms and estimation techniques to look-ahead the changes in states and dynamics of physical system upon execution of control commands.
- Combine the information on the activities observed in the cyber domain (e.g., the network activities) with multiple estimated measurements from the physical domain, to further optimize the computation and reduce the detection latency.

Attack Targets

- H. Alemzadeh, D. Chen, X. Li, T. Kesavadas, Z. T. Kalbarczyk, R. K. Iyer, "Targeted Attacks on Teleoperated Surgical Robots: Dynamic Model-based Detection and Mitigation," to appear in the *46th IEEE/IFIP International Conference on Dependable Systems and Networks (DSN).*
- Hui Lin, Adam Slagell, Zbigniew Kalbarczyk, Peter W. Sauer, and Ravishankar K. Iyer, "Runtime Semantic Security Analysis to Detect and Mitigate Control-related Attacks in Power Grids," in *IEEE Transactions on Smart Grid*, vol.PP, no.99, pp.1-1.

SCIENCE OF SECURITY
VIRTUAL ORGANIZATION
Funded by the National Security Agency.

INFORMATION TRUST INSTITUTE