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

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

Detection Principle

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

Cyber-Physical Systems

Control Center

PLC monitors system state and controls fail-safe brakes

Physical System

Robot (Actuators, Sensors, Relays)

Human Operators

Robot position

Robot state

Emergency stop

Safety PLC

Physical layer

Communication Network

Commands

Physical Process

Cyber-domain

Physical-domain

Challenges

Example Cyber-Physical Systems

Power Grids

Surgical Robots

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.

Malicious and unsafe commands can be encoded in legitimate formats

Modification of a few bits in network traffic can maintain the correct communication syntax.

Malicious commands (e.g., TOCTOU (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.)

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.

Real-time constraints on control systems

Control operations should be delivered in a few hundred milliseconds.

Real-time constraint of 1 milliseconds per control iteration.

Attacks are hard to distinguish from incidental failures and human induced safety hazards.

Contingency analysis evaluates the consequences of incidents, in which one or two physical components are out of service.

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 attacker to decide which one to be injected to the state to cause the impact of a command in the whole power grid.

Physical domain

There are limited hardware resources on the embedded computer units in the interface and the physical layer of the robot to perform sophisticated computations for estimating state system.

Challenges

Detection of Attacks

Example Cyber-Physical Systems

Power Grids

Surgical Robots

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.

Malicious and unsafe commands can be encoded in legitimate formats

Modification of a few bits in network traffic can maintain the correct communication syntax.

Malicious commands (e.g., TOCTOU (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.)

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.

Real-time constraints on control systems

Control operations should be delivered in a few hundred milliseconds.

Real-time constraint of 1 milliseconds per control iteration.

Attacks are hard to distinguish from incidental failures and human induced safety hazards.

Contingency analysis evaluates the consequences of incidents, in which one or two physical components are out of service.

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 attacker to decide which one to be injected to the state to cause the impact of a command in the whole power grid.

Physical domain

There are limited hardware resources on the embedded computer units in the interface and the physical layer of the robot to perform sophisticated computations for estimating state system.