DDDAS for Attack Detection, Isolation, and Reconfiguration of Control Systems

Luis Francisco Combita, Jairo Giraldo, Alvaro A. Cardenas, Nicanor Quijano

University of Texas at Dallas
Universidad de Los Andes, Colombia

InfoSymbiotics/DDDAS
August 11, 2016
Control Systems

- **Attacks to Regulatory Control**
  - A1 and A3 are deception attacks: the integrity of the signal is compromised
  - A2 and A4 are DoS attacks
  - A5 is a physical attack to the plant
The Threat is not Hypothetical

Cyberattack on German steel factory causes 'massive damage'

IN Inside the cunning, unprecedented hack of Ukraine's Power Grid

Countdown to Zero Day

KrebsonSecurity

FBI: Smart Meter Hacks Likely to Spread

A series of hacks perpetrated against so-called "smart meter" installations over the past several years may have cost a single U.S. electric utility hundreds of millions of dollars annually, the FBI said in a cyber intelligence bulletin obtained by
Defense in Depth

• Security is not only about keeping attackers out
• It is also about
  – Mitigating
  – Detecting
  – Responding
• to adversaries that have partial access to your system
DDDAS Anomaly Detection and Response

- **Actuators**
  - $u_k$

- **Physical Process (Plant)**
  - $z_k$

- **Sensors**
  - $y_k$

- **Simulation**
  - $y_{k-1}$
  - $\hat{y}_k$

- **Controller**
  - $u_k$
  - $T_k$

- **Anomaly Detection**
  - (ignore bad sensors, reconfigure simulation)

- **Reconcile Data**

- **Reconfigure Controller**
  - (account for bad actuators)

- **Dynamically Request More Data from Other Systems**
Network Intrusion Detection

[Urbina et al. ACM CCS 2016]
Extracting Sensor and Control Commands from Network Traffic

- Protocol specification correct but false info

Scapy parser for Modbus
Detection = Simulation + Statistics
LDS Model for Raw Water Tank

\[
\frac{dV_i}{dt} = A_i \frac{dh_i}{dt} = Q_{i,\text{in}} - Q_{i,\text{out}} \\

h_{k+1} = h_k + \frac{Q_{i,k} - Q_{o,k}}{A}
\]
Implementing the Attack
Problem: We Can Always Create Attacks That Are Detected
Attackers are More Cunning than Failures (they try to avoid being detected)

Anomaly Detection Statistic

threshold for raising an alarm
Undetected Attacks to Water Testbed
Our Proposed Metric

Security Metric: Impact of undetected attacks

Usability Metric: Time between false alarms

Detector 2 is better than Detector 1: For the same level of false alarms, undetected attackers can cause less damage to the system.
Trade-off Curves Can Help us Identify Which Detectors are Better than Others
What Happens After Detection?

• Alert to operator
• Automatic Response
  • Identify compromised device
  • Isolate it
  • Reconfigure the control system
Three Tank Example for Isolation and Response
Luenberger vs. Unknown Input Observer (UIO) Estimators

Luenberger Detects Attacks Faster with Little False Alarms, but difficult to identify source of attack.

UIO identify source of anomaly but have higher false alarms / detection delay.
Detection (Luenberger) + Isolation (UIO) + Reconfiguration

![Graph showing level 1 (m) over time with and without reconfiguration.]
Other DDDAS-Inspired Architectures for Secure/Private Control

Risk-Aware Operation
Privacy-Preserving Control
Safe Control Under DoS Attacks

For constrained linear systems

$$x_{k+1} = Ax_k + Bu_k + w_k, \quad k = 1, \ldots, N - 1$$
$$x_k^a = \gamma_k x_k, u_k^a = \nu_k u_k, \quad (\gamma_k, \nu_k) \in \{0, 1\}^2$$

find causal feedback policies \( u_k = \mu_k(x_0^a, \ldots, x_k^a) \), that
minimize \( J(x_0, u, w) = \sum_{k=1}^N x_k^T Q^{xx} x_k + \sum_{k=1}^{N-1} \nu_k u_k^T Q^{uu} u_k, \)
subject to power constraints

$$\begin{pmatrix} x_k^a \\ u_k^a \end{pmatrix}^T \begin{pmatrix} H_{i}^{xx} & 0 \\ 0 & H_{i}^{uu} \end{pmatrix} \begin{pmatrix} x_k^a \\ u_k^a \end{pmatrix} \leq \beta_i, \quad i = 1, \ldots, L_1,$$

and safety constraints

$$\begin{pmatrix} x_k^a \\ u_k^a \end{pmatrix} \in T_j, \quad j = 1, \ldots, L_2,$$

for all disturbances \( w \in W_\alpha \) OR \( w \sim \mathcal{N}(0, W) \) and a given set of
\( (\gamma_0^{N-1}, \nu_0^{N-1}) \in A_{pq} \) attack signatures.
DDDAS-Inspired Risk-Operation

- If there is any indicator “cyber or physical” of potential future attack, then predict attack and operate conservatively
Privacy Guidelines for Smart Grid

• Collect “only … necessary [data] for Smart Grid operations, including planning and management”
  – Perhaps plan and manage better with more data?

• Retain data “only for as long as necessary”
  – Data for a longer time presumably means better forecasting?
Microgrid Synchronization with Privacy Sampling

For example: Frequency variations of a four generators system (top) and the transmitted information (bottom) for $\delta = 20$ s.

Frequency synchronization and stability is maintained, while the private event is hidden.

[Giraldo et al. IEEE CDC 2014]
Reaching consensus independent of sampling rate and time delays

**Theorem 2**

The consensus algorithm with sampled and delayed information only from its neighbors reaches a consensus state independent of the sampling period and the time-delay if the graph is connected (exists a spanning tree).

**Four generators model**

![Diagram of four generators model](image)

**Demand profile $DG_1$**

![Graph showing demand profile](image)
New Sampling Policy: Discretionary Sampling

- **Periodic sampling**
  - Do not require knowledge about the events
  - Less computational costs
  - Do not assure total privacy

- **Discretionary sampling**
  - Require knowledge about the events
  - Need to determine the transmitted “false information”
  - May assure total privacy
  - Using a flexible sampling period, may improve convergence time, but it requires real-time sampling-period modifications.

---

**Graphs**

- **Synchronization settling time (sec)**
  - **Periodic sampling**
  - **Discretionary sampling**

- **# of detected events from MG**
  - **Periodic sampling**
  - **Discretionary sampling**
Questions?

Alvaro A. Cardenas
alvaro.cardenas@utdallas.edu