A Control-Theoretic Approach for Optimization and Security of Automated Traffic Networks

Gianluca Bianchin and Fabio Pasqualetti

Department of Mechanical Engineering
University of California, Riverside

Cyber-Physical Security Education Workshop
July 17-19, 2017, Paris, France

Emerging Threats for Cyber-Physical Systems

Cyber-Physical systems are prone to failures and attacks against their physical, communication, and computational layers

Stuxnet worm (Iran, 2010)
New York Times 15jan2011: replay attack as if “out of the movies.”
- records normal operations and plays them back to operators
- spins centrifuges at damaging speeds

Cars vulnerable to cyber attacks
Hackers take control of cars: start/stop the engine, shut off the lights, hit the brakes...

Smart cities: security of automated cars, roads, and intersections?
A Growing Research Area

Analysis of vulnerabilities and detection schemes:


Design of remedial actions:


Cyber-Physical Security ≠ Cyber Security, Fault Tolerance

Cyber-Physical security complements cyber security

- Cyber security
  - does not verify “data compatible with physics/dynamics”
  - is ineffective against direct attacks on the physics/dynamics
  - is never foolproof (e.g., insider attacks)

Cyber-physical security complements cyber security

- Physical dynamics: classical generator model & DC load flow
- Measurements: angle and frequency of generator $g_1$
- Attack: modify real power injections at buses $b_4$ & $b_5$


The attack affects the second and third generators while remaining undetected from measurements at the first generator
Models of Cyber-Physical Systems: Power Networks

Small-signal structure-preserving power network model:
- transmission network: generators ■, buses ●, DC load flow assumptions, and network susceptance matrix $Y = Y^T$
- generators ■ modeled by swing equations:
  $$M_i \ddot{\theta}_i + D_i \dot{\theta}_i = P_{\text{mech.in}},i - \sum_j Y_{ij} \cdot (\theta_i - \theta_j)$$
- buses ● with constant real power demand:
  $$0 = P_{\text{load},i} - \sum_j Y_{ij} \cdot (\theta_i - \theta_j)$$

⇒ Linear differential-algebraic dynamics: $E \dot{x} = Ax$

Models of Cyber-Physical Systems: Water Networks

Linearized municipal water supply network model:
- reservoirs with constant pressure heads: $h_i(t) = h_i^{\text{reservoir}} = \text{const}$.
- pipe flows obey linearized Hazen-Williams eq:
  $$Q_{ij} = g_{ij} \cdot (h_i - h_j)$$
- balance at tank:
  $$A_i \dot{h}_i = \sum_{j \to i} Q_{ji} - \sum_{i \to k} Q_{ik}$$
- demand = balance at junction:
  $$d_i = \sum_{j \to i} Q_{ji} - \sum_{i \to k} Q_{ik}$$
- pumps & valves:
  $$h_j - h_i = +\Delta h_{\text{pump/valves}} = \text{const.}$$

⇒ Linear differential-algebraic dynamics: $E \dot{x} = Ax$

Models of Networks, Attackers, and Monitors

Actuator Attack
State Attack
Data Attack

Plant
Sensors
Control Center
State Attack Data Attack
Actuator Attack

Previously recorded measurements

$E \dot{x}(t) = Ax(t) + Bu(t)$ (state and actuator attack)
y(t) = Cx(t) + Du(t) (data substitution attack)

- attackers are colluding and omniscient (model, params, state)
- attackers aim to change physical state and mislead monitors
- monitors aim to detect/identify attacks via measurements

Modeling Stuxnet as Unknown Inputs

Actuator Attack
State Attack
Data Attack

Plant
Sensors

$Bu_3(t)$

$Du_1(t) + Du_2(t)$

System dynamics:

$E \dot{x}(t) = Ax(t) + Bu_3(t)$
y(t) = Cx(t) + Du_1(t) + Du_2(t)$
Undetectable Attacks

The attack \( u \) is undetectable if its effect on measurements is undistinguishable from the effect of some nominal operating condition

\[
y(x_1, 0, t) = y(x_2, u, t)
\]

The attack \( u \) is undetectable if

\[
y(x_1, 0, t) = y(x_2, u, t)
\]

Detectability of Attacks

Equivalent characterizations of undetectable attacks:

- **Vulnerability:** undetectable attack \( y(x_1, 0, t) = y(x_2, u, t) \)
- **System theory:** intruder/monitor system has invariant zeros
- **Graph theory** \# attack signals \( > \) size of input-output linking

By linearity, an undetectable attack is such that \( y(x_2 - x_1, u, t) = 0 \).

\[
\begin{bmatrix}
  sE - A & -B \\
  C & D
\end{bmatrix}
\begin{bmatrix}
  x_0 \\
  u_0
\end{bmatrix} = 0
\]

Detectability of Attacks

Equivalent characterizations of undetectable attacks:

- **Vulnerability:** undetectable attack \( y(x_1, 0, t) = y(x_2, u, t) \)
- **System theory:** intruder/monitor system has invariant zeros
- **Graph theory** \# attack signals \( > \) size of input-output linking

Attack \((Bu(t), Du(t))\) is not detectable by measurements \( y(t) \) & destabilizes the system.

\[
\begin{align*}
\omega_1(t) &= y(t) \\
\omega_2(t) &= y(t) \\
\omega_3(t) &= y(t)
\end{align*}
\]
### Unidentifiable Attacks

The attack $B_1u_1$ is unidentifiable if its effect on measurements is undistinguishable from the effect of the attack $B_2u_2$.

### Design of Targeted/Undetectable Attacks

- Malicious coalition: $K = \{1, 9\}$ (PacNW) with sacrificial machine $K^* = \{9\}$
- Attack input minimizes $\|\omega_9(t)\|_{L_\infty}$ subject to $\|\omega_{16}(t)\|_{L_\infty} \geq 1$ (Utah)

$\Rightarrow$ Non-colluding generators will be damaged

### Identifiability of Attacks

Equivalent characterizations of unidentifiable attacks:

- **Vulnerability**: unidentifiable attack $y(x_1, u_1, t) = y(x_2, u_2, t)$
  - $y(x_1 - x_2, u_1 - u_2, t) = 0$

- **System theory**: extended intruder/monitor system has invariant zero
  - $(E, A, [B_1, B_2], C, [D_1, D_2])$

- **Graph theory**: # attack signals > (size of input-output linking) / 2

So far we have shown:

- Fundamental detection/identification limitations
- System-theoretic conditions for undetectable/unidentifiable attacks
- Graph-theoretic conditions for undetectable/unidentifiable attacks
- Secure-by-design criteria: zero dynamics $\iff$ vulnerabilities

---

**Physical dynamics**: classical generator model & DC load flow

**Measurements**: angle and frequency of generator $g_1$

**Attack**: modified real power injections at buses $b_4$ & $b_5$

The attack through $b_4$ and $b_5$ excites only zero dynamics for the measurements at the first generator

---

F. Pasqualetti  
Optimization and security of automated traffic networks  
07/17-19/17 14 / 33

---

F. Pasqualetti  
Optimization and security of automated traffic networks  
07/17-19/17 15 / 33

---

F. Pasqualetti  
Optimization and security of automated traffic networks  
07/17-19/17 16 / 33
Other Results Leveraging Control-Theoretic Framework

Other results:
- Design of centralized and distributed monitors
- Detection and identification of attacks in stochastic control systems
- Security tradeoffs in resource-constrained real-time systems
- Design and operation methods for secure systems

Security issues in automated traffic networks...
- Vulnerability of automated traffic networks?
- Models? Accuracy vs complexity vs scale
- Are control-theoretic methods still useful?

Models of Cyber-Physical Systems: Traffic Networks

Traffic network: directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
- Vertices $\mathcal{V} = \{1, \ldots, n\}$ represent roads
- Edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ allow cars from one road to another
- Edges vary over time (see intersection phases)

Roads/traffic flow: conservation + Lighthill-Whitham-Richards model

Intersections and phases: automata, lead to switching graph topology

Existing Results on Traffic Optimization and Security

Traffic optimization and cyber security:

- most works focus on isolated problems, e.g., intersection scheduling
- strong assumptions, e.g., no delays on roads or simple topologies
- network-wide approach address no security
- most security approaches focus on “cyber” issues

Cyber-physical security of traffic networks remains to be characterized

Roads and Traffic Flow #1

Flow is continuous and it obeys a conservation law:
$$\frac{\partial \rho(s,t)}{\partial t} + \frac{\partial f(s,t)}{\partial s} = 0$$
- density variation at $(s,t)$
- flow variation at $(s,t)$

Flow as a function of density (Lighthill-Whitham-Richards):
$$f(s,t) = f(\rho(s,t)) = v(\rho(s,t)) \rho(s,t)$$

Speed also depends on density...
### Static density-speed relation: low density $\approx$ constant speed

\[
\frac{\partial \rho(s,t)}{\partial t} + \left( v(\rho(s,t)) + \rho(s,t) \frac{\partial v(\rho(s,t))}{\partial \rho} \right) \frac{\partial \rho(s,t)}{\partial s} = 0
\]

\[\gamma = \text{average speed}\]

### Traffic flow across road:

\[
\frac{\partial \rho(s,t)}{\partial t} + \gamma \frac{\partial \rho(s,t)}{\partial s} = 0
\]

### Intersections and Phases

- Intersections control the right of way to coordinate traffic flows
- Admissible transitions $\mathcal{M} = \{(1,6), (1,8), (5,2), (5,4), (7,2), (3,6), (3,8)(3,2), (7,4), (7,6), (5,8), (1,4)\}$
- Phase (simultaneous transitions) $\mathcal{P}_j = \{(1,6), (1,8), (5,2), (5,4)\}$
- $\mathcal{P}(t) = \mathcal{P}_1 \cup \cdots \cup \mathcal{P}_m$ are the edges of the traffic network at time $t$

### Traffic Flows, Graphs, and Dynamical Models #1

- Inflow into node (road) $i$: $f_{i}^{\text{in}}(t) = \sum_{(i,j) \in \mathcal{P}(t)} f_{ij}^{\text{out}}(t)$
- Outflow from node (road) $j$ into $i$: $f_{ij}^{\text{out}} = c_{ij} \rho_j(t)$
- External flows may enter/exit the network at certain locations

**Roads and Traffic Flow #2**

- Congestion-Free Flow

**Roads and Traffic Flow #3**

- Discretize spatial derivative
  \[
  \frac{\partial \rho(s,t)}{\partial s} = \frac{\rho(s_{k-1},t) - \rho(s_k,t)}{h}
  \]

- Road length is captured by dimension of $A_i$. Accuracy depends on $h$. 

**Traffic Flows, Graphs, and Dynamical Models #1**

- (a) Network with 3 intersections
- (b) Traffic graph and flows
At time $t$, the traffic network evolves as
\[
\begin{pmatrix}
\dot{x}_1 \\
\vdots \\
\dot{x}_n
\end{pmatrix} =
\begin{bmatrix}
A_1 & \cdots & B_{1n} \\
\vdots & \ddots & \vdots \\
B_{n1} & \cdots & A_n
\end{bmatrix}
\begin{pmatrix}
x_1 \\
\vdots \\
x_n
\end{pmatrix} +
\begin{pmatrix}
u_1 \\
\vdots \\
u_n
\end{pmatrix}
\]

Densities/queues    Roads model and phase $P(t)$    External flows

Over time, the traffic network evolves as a switching system:
\[
\dot{x} = A_P x + u
\]

- How to schedule the phases $P(t)$ to optimize traffic?
- How to model/analyze/remedy malicious attacks and failures?

Optimization problem to minimize queues length over time:
\[
\min_P \int_0^\infty \|x(t)\|^2 dt \\
\text{s. t. } \dot{x}(t) = \tilde{A}_P x(t) \\
x_0 \text{ (initial queues)}
\]

Equivalent optimization problem:
\[
\min_P \text{Trace}(W_P) \\
\text{s. t. } \tilde{A}_P W_P + W_P \tilde{A}_P^T = -x_0 x_0^T
\]

- $W_P$ is the controllability Gramian of the pair $(\tilde{A}_P, x_0)$
Comparison With Existing Policies

Optimization done by minimizing smoothed spectral abscissa...

Max-pressure creates “overshoots”

Effect of Falsifying Initial Flows

Estimated traffic density:

Assume the true state is

False data injection deteriorates performance; higher overshoots in distributed policies