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

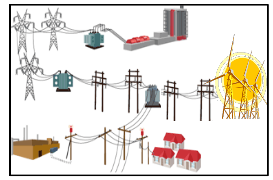
Gianluca Bianchin and Fabio Pasqualetti

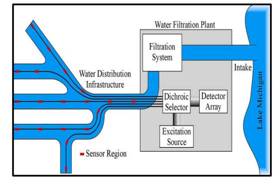
Department of Mechanical Engineering
University of California, Riverside

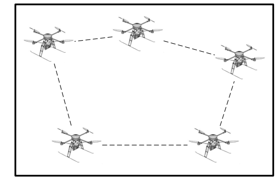
Cyber-Physical Security Education Workshop

July 17-19, 2017, Paris, France

Cyber-Physical Systems



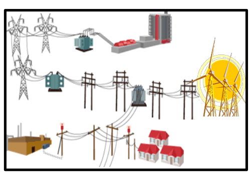


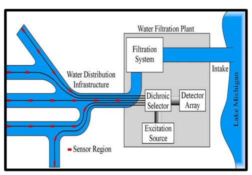


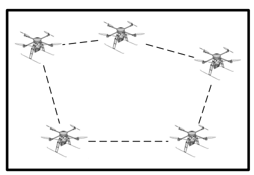
- Power generation, transportation, distribution networks
- Water, oil, gas and mass transportation systems
- Sensor networks and multi-agent systems
- Process control and industrial automation systems (metallurgical process plants, oil refining, chemical plants, pharmaceutical manufacturing ... ubiquitous SCADA/PLC systems)

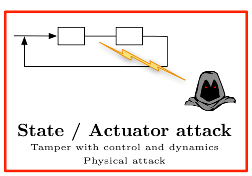
Security of these systems is critically important

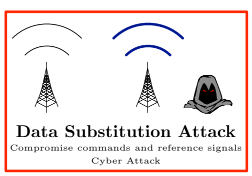
Emerging Threats for Cyber-Physical Systems

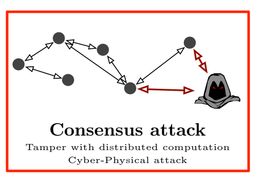












Cyber-Physical systems are prone to failures and attacks against their physical, communication, and computational layers

Documented Security Incidents

Stuxnet worm (Iran, 2010)
New York Times 15jan2011: replay attack as if "out of the movies:"

1. records normal operations and plays them back to operators
2. 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...


BBC NEWS
TECHNOLOGY

21 February 2014 Last updated at 08:46 ET

South Korea to develop Stuxnet-like cyberweapons

South Korea is to develop cyber-attack tools in an attempt to damage North Korean nuclear facilities.

Cockrell School Researchers Demonstrate First Successful "Spoofing" of UAVs

Wednesday, 27 June 2012 09:16

Obama: Cyber attack serious threat to economy, national security

Summary: U.S. President Barack Obama is urging the Senate to pass the Cybersecurity Act of 2012. He believes legislation will help the U.S. fight "the cyber threat to our nation," which he calls "one of the most serious economic and national security challenges we face."

Pacific Northwest National Laboratory Report Reveals Dramatic Increase in Cyber Threats and Sabotage on Critical Infrastructure and Key Resources
June 2012 – US Dept of Energy

Smart cities: security of automated cars, roads, and intersections?

A Growing Research Area

Analysis of vulnerabilities and detection schemes:

- Y. Mo, J. Hespanha and B. Sinopoli "Resilient Detection in the Presence of Integrity Attacks," in *IEEE Transactions on Signal Processing*, 62(1):31-43, 2014.
- C. Bai and F. Pasqualetti and V. Gupta. "Security in Stochastic Control Systems: Fundamental Limitations and Performance Bounds," in *IEEE American Control Conference*, 2015.
- R. Smith. "Covert Misappropriation of Networked Control Systems: Presenting a Feedback Structure," in *IEEE Control Systems Magazine*, 35(1):82-92, 2015.
- F. Hamza, P. Tabuada, and S. Diggavi "Secure state-estimation for dynamical systems under active adversaries," in *Allerton Conf. on Communications, Control and Computing*, Sep. 2011.
- H. Nishino and H. Ishii. "Distributed detection of cyber attacks and faults for power systems," in *Proceedings of the 19th IFAC World Congress.*, 2014.
- Y. H. Chang, Q. Hu, C. J. Tomlin "Secure Estimation based Kalman Filter for Cyber-Physical Systems against Adversarial Attacks," in *Arxiv*, Dec. 15 2015.

Design of remedial actions:

- V. Dolk, P. Tesi, C. De Persis, and W. Heemels "Event-triggered control systems under denial-of-service attacks," in *IEEE Transactions on Control of Network Systems*, 4(1):93-105, 2017.
- M. Zhu and S. Martiínez. "Stackelberg-game analysis of correlated attacks in cyber-physical systems," in *American Control Conference.*, 2011.

Cyber-Physical Security $\neq$ 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 extends fault tolerance

- fault detection considers *accidental/generic failures*
- cyber-physical security models *worst-case attacks*

Cyber-Physical Security $\neq$ Cyber Security, Fault Tolerance

Cyber-physical security complements cyber security

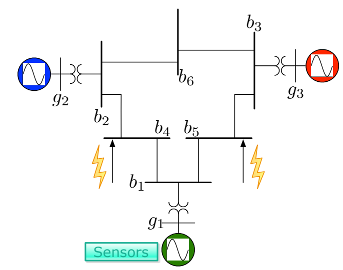
Cyber security

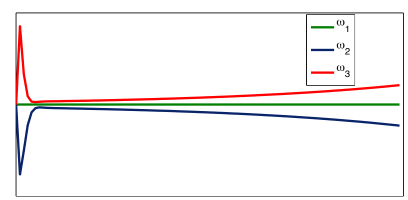
- does not verify "data compatible with physics/dynamics"
- is ineffective against direct attacks on the physics/dynamics
- is never foolproof (e.g., insider attacks)

Cyber-physical security extends fault tolerance

- fault detection considers *accidental/generic failures*
- cyber-physical security models *worst-case attacks*

An Example of Cyber-Physical Attack





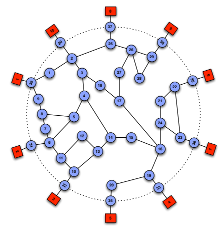
- **Physical dynamics:** classical generator model & DC load flow
- **Measurements:** angle and frequency of generator g_1
- **Attack:** modify real power injections at buses b_4 & b_5

- A. H. Mohsenian-Rad and A. Leon-Garcia "Distributed internet-based load altering attacks against smart power grids" *IEEE Transactions on Smart Grid*, 2011

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:



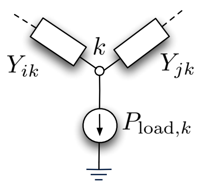
1. transmission network: generators ■, buses ●, DC load flow assumptions, and network susceptance matrix $Y = Y^T$

2. generators ■ modeled by swing equations:

$$M_i \ddot{\theta}_i + D_i \dot{\theta}_i = P_{\text{mech.in},i} - \sum\nolimits_j Y_{ij} \cdot (\theta_i - \theta_j)$$

3. buses ● with constant real power demand:

$$0 = P_{\text{load},i} - \sum\nolimits_j Y_{ij} \cdot (\theta_i - \theta_j)$$



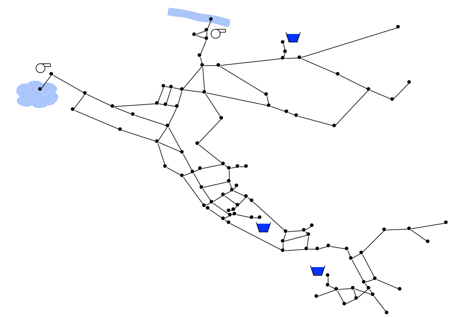
⇒ Linear differential-algebraic dynamics: $E\dot{x} = Ax$

Models of Cyber-Physical Systems: Water Networks

Linearized municipal water supply network model:

1. reservoirs with constant pressure heads: $h_i(t) = h_i^{\text{reservoir}} = const.$

2. pipe flows obey linearized Hazen-Williams eq: $Q_{ij} = g_{ij} \cdot (h_i - h_j)$



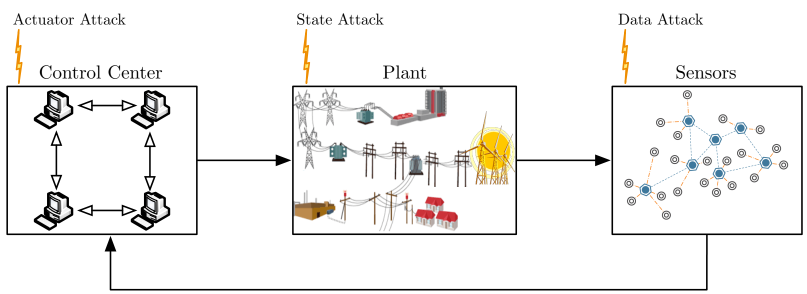
3. balance at tank:
$A_i \dot{h}_i = \sum_{j \to i} Q_{ji} - \sum_{i \to k} Q_{ik}$

4. demand = balance at junction:
$d_i = \sum_{j \to i} Q_{ji} - \sum_{i \to k} Q_{ik}$

5. pumps & valves:
$h_j - h_i = +\Delta h_{ij}^{\text{pump/valves}} = \text{const.}$

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

Models of Networks, Attackers, and Monitors

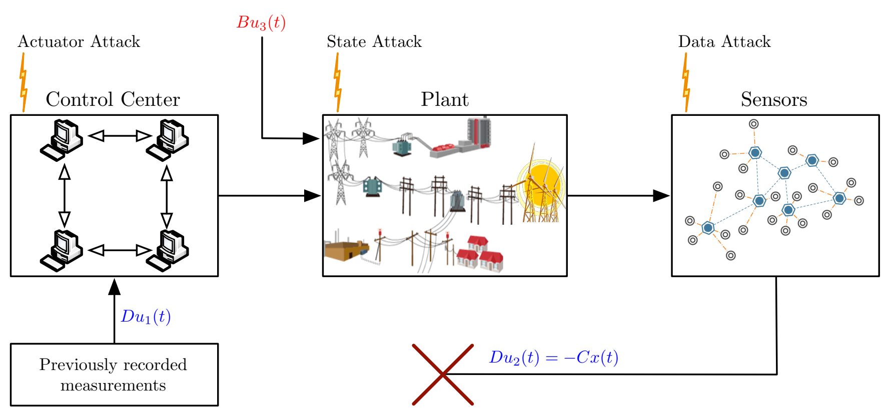


$$E\dot{x}(t) = Ax(t) + Bu(t) \qquad \text{(state and actuator attack)}$$

$$y(t) = Cx(t) + Du(t) \qquad \text{(data substitution attack)}$$

- 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



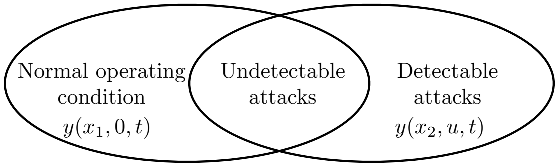
System dynamics:

$$E\dot{x}(t) = Ax(t) + Bu_3(t)$$

$$y(t) = Cx(t) + Du_1(t) + Du_2(t)$$

Undetectable Attacks

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

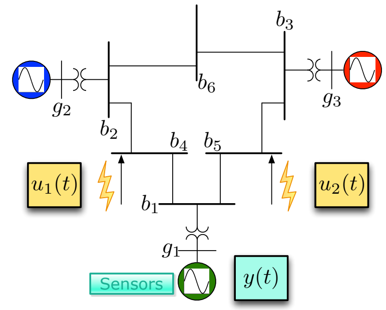


The attack u is undetectable if
$y(x_1, 0, t) = y(x_2, u, t)$

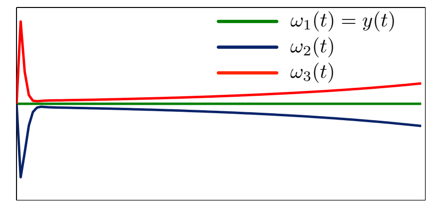
Detectability of Attacks

Equivalent characterizations of undetectable attacks:

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



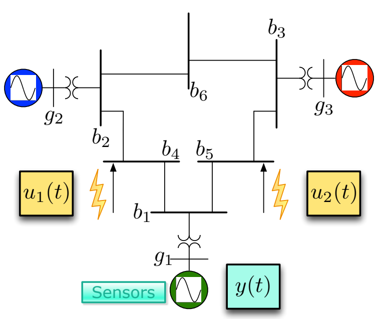
Attack $(Bu(t), Du(t))$ is not detectable by measurements $y(t)$ & destabilizes the system



Detectability of Attacks

Equivalent characterizations of undetectable attacks:

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



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

$\Leftrightarrow$ invariant zeros for system

$$E\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$

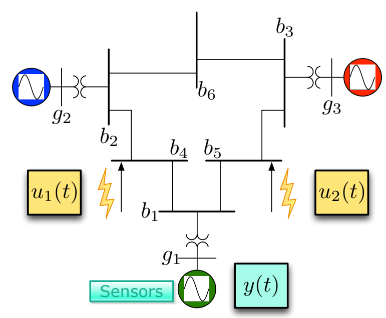
$\Leftrightarrow$ nontrivial solution to

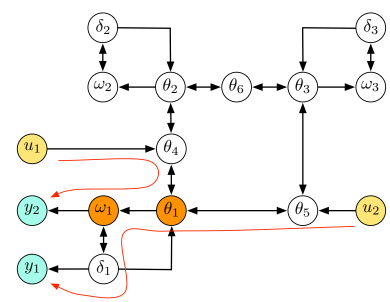
$$\begin{bmatrix} sE - A & -B \\ C & D \end{bmatrix} \begin{bmatrix} x_0 \\ u_0 \end{bmatrix} = 0$$

Detectability of Attacks

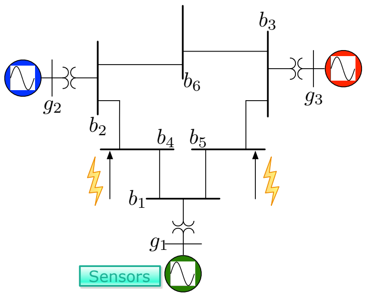
Equivalent characterizations of undetectable attacks:

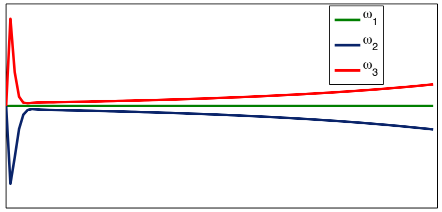
1. **Vulnerability:** undetectable attack $y(x_1, 0, t) = y(x_2, u, t)$
2. **System theory:** intruder/monitor system has invariant zeros
3. **Graph theory** # attack signals $>$ size of input-output linking





An Example fo Undetectable Attack



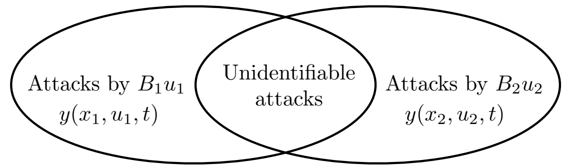


1. **Physical dynamics:** classical generator model & DC load flow
2. **Measurements:** angle and frequency of generator g_1
3. **Attack:** modified real power injections at buses b_4 & b_5

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

Unidentifiable Attacks

The attack $B_1 u_1$ is unidentifiable if its effect on measurements is undistinguishable from the effect of the attack $B_2 u_2$



The attack u_1 is unidentifiable if
$$y(x_1, u_1, t) = y(x_2, u_2, t)$$

- an undetectable attack is also unidentifiable

Identifiability of Attacks

Equivalent characterizations of unidentifiable attacks:

1. **Vulnerability:** unidentifiable attack $y(x_1, u_1, t) = y(x_2, u_2, t)$
   - $y(x_1 - x_2, u_1 - u_2, t) = 0$
2. **System theory:** extended intruder/monitor system has invariant zero
   - $(E, A, [B_1\ B_2], C, [D_1\ D_2])$
3. **Graph theory** # attack signals $>$ (size of input-output linking) / 2

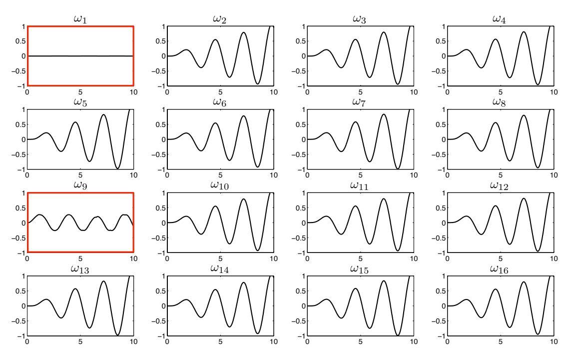
So far we have shown:

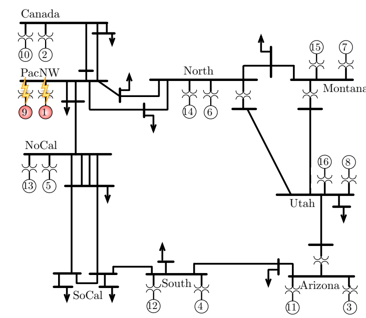
- fundamental detection/identification limitations
- system-theoretic conditions for undetectable/unidentifiable attacks
- graph-theoretic conditions for undetectable/unidentifiable attacks
- secure-by-design criteria: zero dynamics $\Leftrightarrow$ vulnerabilities

Design of Targeted/Undetectable Attacks

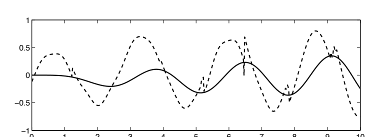
- malicious coalition: $K = \{1, 9\}$ (PacNW) with sacrificial machine $K^* = \{9\}$
- attack input minimizes $\|\omega_9(t)\|_{\mathcal{L}_\infty}$ subject to $\|\omega_{16}(t)\|_{\mathcal{L}_\infty} \geq 1$ (Utah)

$\Rightarrow$ non-colluding generators will be damaged





Reduced WECC grid



governor control input

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?

Existing Results on Traffic Optimization and Security

Traffic optimization and cyber security:

C. Diakaki, M. Papageorgiou, and K. Aboudolas, "A multivariable regulator approach to traffic-responsive network-wide signal control," in *Control Engineering Practice* , vol. 10, no. 2, pp. 183-195, 2002.

L. B. de Oliveira and E. Camponogara, "Multi-agent model predictive control of signaling split in urban traffic networks" in *Transportation Research Part C: Emerging Technologies* , vol. 18, no. 1, pp. 120-139, 2010.

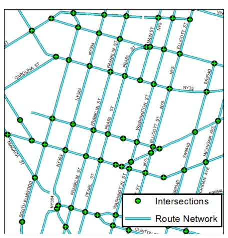
P. Varaiya, "Max pressure control of a network of signalized intersections,"' in *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 177-195, 2013.

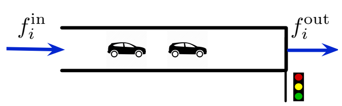
A. Ghafouri, W. Abbas, Y. Vorobeychik, and X. Koutsoukos, "Vulnerability of fixed-time control of signalized intersections to cyber-tampering," in *ResilienceWeek(RWS)* ,2016. IEEE,2016, pp.130-135.

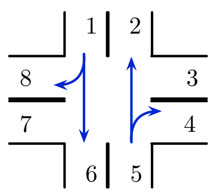
- most works focus on isolated problems, e.g., intersection scheduling
- strong assumptions, e.g., no delays on roads or simple topologies
- network-wide approach address no security
- most security approaches focus on "cyber" issues

Cyber-physical security of traffic networks remains to be characterized

Models of Cyber-Physical Systems: Traffic Networks

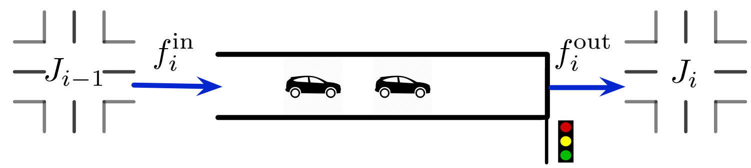






- Traffic network: directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
  - Vertices $\mathcal{V} = \{1, \dots, n\}$ represent roads
  - Edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ allow cars from one rode to another
  - Edges vary over time (see intersection phases)
- Roads/traffic flow: conservation + Lighthill-Whitham-Richards model
- Intersections and phases: automata, lead to switching graph topology

Roads and Traffic Flow #1



- Flow is continuous and it obeys a conservation law:

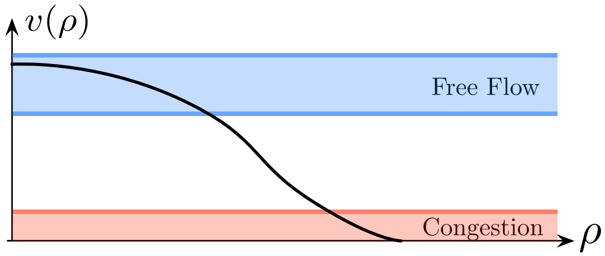
$$\underbrace{\frac{\partial \rho(s,t)}{\partial t}}_{\text{density variation at } (s,t)} + \underbrace{\frac{\partial f(s,t)}{\partial s}}_{\text{flow variation at } (s,t)} = 0$$

- 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...

Roads and Traffic Flow #2



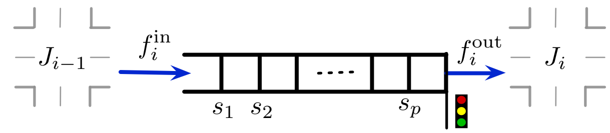
- Static density-speed relation: low density $\approx$ constant speed

$$\frac{\partial \rho(s,t)}{\partial t} + \underbrace{\left(v(\rho(s,t)) + \rho(s,t)\frac{\partial v(\rho(s,t))}{\partial \rho} \right)}_{\gamma\,=\,\text{average speed}} \frac{\partial \rho(s,t)}{\partial s} = 0$$

Traffic flow across road:

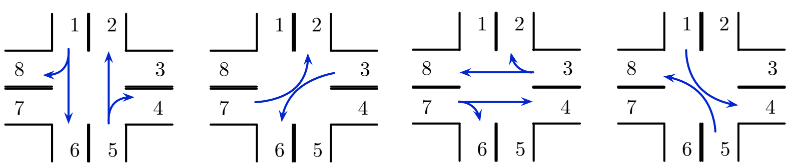
$$\frac{\partial \rho(s,t)}{\partial t} + \gamma \frac{\partial \rho(s,t)}{\partial s} = 0$$

Roads and Traffic Flow #3



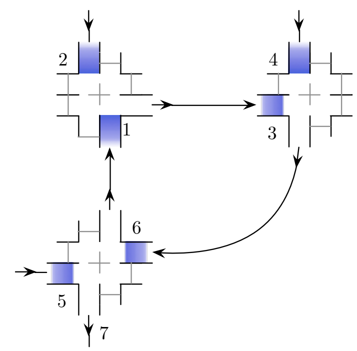
- Discretize spatial derivative
- $\frac{\partial \rho(s,t)}{\partial s} = \frac{\rho(s_{k-1},t) - \rho(s_k,t)}{h}$
- $$\underbrace{\begin{bmatrix}\dot\rho_i^p\\ \vdots\\ \dot\rho_i^1\end{bmatrix}}_{\dot x_i} = \frac{\gamma_i}{h}\underbrace{\begin{bmatrix}0&1&&\\&-1&\ddots&\\&&\ddots&1\\&&&-1\end{bmatrix}}_{A_i} \underbrace{\begin{bmatrix}\rho_i^p\\ \vdots\\ \rho_i^1\end{bmatrix}}_{x_i} - \begin{bmatrix}1\\ \vdots\\ 0\\ 0\end{bmatrix} f_i^{\text{out}} + \begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} f_i^{\text{in}}$$
- Road length is captured by dimension of A_i. Accuracy depends on h.

Intersections and Phases

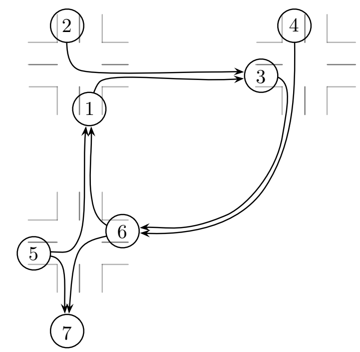


- Intersections control the right of way to coordinate traffic flows
- Admissible transitions $\mathcal{M} = \{(1,6),(1,8),(5,2),(5,4),(7,2),(3,6),(3,8)(3,2),(7,4),(7,6),(5,8),(1,4)\}$
- Phase (simultaneous transitions) $\mathcal{P}_j = \{(1,6),(1,8),(5,2),(5,4)\}$
- $\mathcal{P}(t) = \mathcal{P}_1 \cup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

Traffic Flows, Graphs, and Dynamical Models #1



(a) Network with 3 intersections



(b) Traffic graph and flows

- 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

Traffic Flows, Graphs, and Dynamical Models #2

At time t, the traffic network evolves as

$$\underbrace{\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix}}_{\text{Densities/queues}} = \underbrace{\begin{bmatrix} A_1 & \dots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{n1} & \dots & A_n \end{bmatrix}}_{\text{Roads model and phase } \mathcal{P}(t)} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \underbrace{\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}}_{\text{External flows}}$$

Over time, the traffic network evolves as a switching system:

$$\dot{x} = A_{\mathcal{P}} x + u$$

1. How to schedule the phases $\mathcal{P}(t)$ to optimize traffic?
2. How to model/analyze/remedy malicious attacks and failures?

Traffic Optimization #1

- Periodic phases $\rightarrow$ system becomes time-invariant every period
- Average system described by network matrix $\bar{A}_{\mathcal{P}}$
- Optimal periodic phases to minimize queues length given initial flows

Optimization problem to minimize queues length over time:

$$\begin{aligned} \min_{\mathcal{P}} \quad & \int_0^\infty \|x(t)\|^2 dt \\ \text{s. t.} \quad & \dot{x}(t) = \bar{A}_{\mathcal{P}} x(t) \\ & x_0 \text{ (initial queues)} \end{aligned}$$

Traffic Optimization #2

Optimization problem to minimize queues length over time:

$$\begin{aligned} \min_{\mathcal{P}} \quad & \int_0^\infty \|x(t)\|^2 dt \\ \text{s. t.} \quad & \dot{x}(t) = \bar{A}_{\mathcal{P}} x(t) \\ & x_0 \text{ (initial queues)} \end{aligned}$$

$$\int_0^\infty \|x(t)\|^2 dt = \int_0^\infty x_o^\mathsf{T} e^{A^\mathsf{T} t} e^{At} x_0 dt = \int_0^\infty \text{Trace}\left(x_o^\mathsf{T} e^{A^\mathsf{T} t} e^{At} x_0\right) dt$$

$$= \int_0^\infty \text{Trace}\left(e^{At} x_0 x_o^\mathsf{T} e^{A^\mathsf{T} t}\right) dt = \text{Trace}\left(\underbrace{\int_0^\infty e^{At} x_0 x_o^\mathsf{T} e^{A^\mathsf{T} t} dt}_{\text{Controllability Gramian!}}\right)$$

Traffic Optimization #3

Optimization problem to minimize queues length over time:

$$\begin{aligned} \min_{\mathcal{P}} \quad & \int_0^\infty \|x(t)\|^2 dt \\ \text{s. t.} \quad & \dot{x}(t) = \bar{A}_{\mathcal{P}} x(t) \\ & x_0 \text{ (initial queues)} \end{aligned}$$

Equivalent optimization problem:

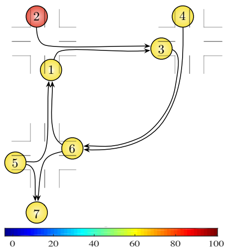
$$\begin{aligned} \min_{\mathcal{P}} \quad & \text{Trace}(W_{\mathcal{P}}) \\ \text{s. t.} \quad & \bar{A}_{\mathcal{P}} W_{\mathcal{P}} + W_{\mathcal{P}} \bar{A}_{\mathcal{P}}^\mathsf{T} = -x_0 x_0^\mathsf{T} \end{aligned}$$

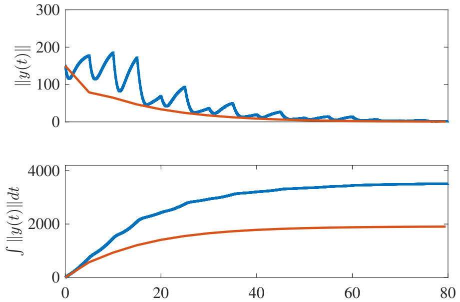
- $W_{\mathcal{P}}$ is the controllability Gramian of the pair $(\bar{A}_{\mathcal{P}}, x_0)$

Comparison With Existing Policies

Optimization done by minimizing smoothed spectral abscissa...

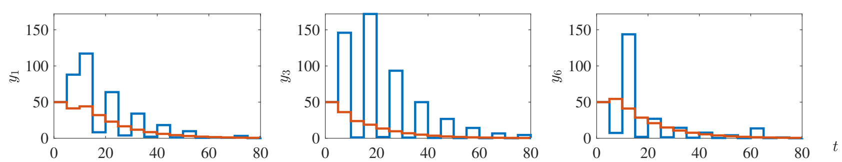
Traffic density:





Red: Gramian – Blue: Max-Pressure

Max-pressure creates "overshoots"



Security?

Malicious attacks:

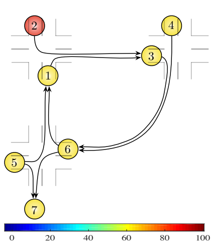
- Manipulation of phases
- False data injection to estimate initial flows x_0
- Road obstruction

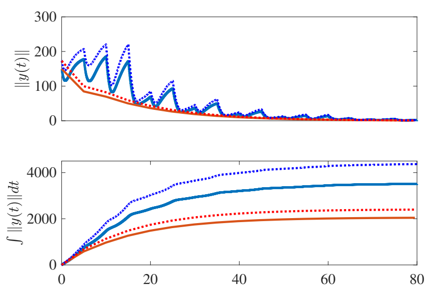
Research questions:

- Effect of attacks?
- Detectability/identifiability of attacks?
- Remedial actions?
- Can we design more robust networks?

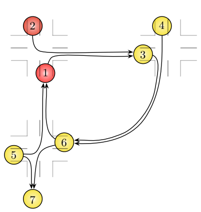
Effect of Falsifying Initial Flows

Estimated traffic density:





Assume the true state is



False data injection deteriorates performance; higher overshoots in distributed policies

Summary

Control-theoretic methods for cyber-physical security:

1. Algebraic and graphical conditions for detectability of attacks
2. Control-theoretic security is complementary to cyber security
3. Models and algorithms for optimization of traffic networks
4. Challenges and opportunities to secure traffic networks

F. Pasqualetti, F. Dörfler, F. Bullo. "Attack Detection and Identification in Cyber-Physical Systems," *IEEE Transactions on Automatic Control*, 58(11):2715-2729, 2013.

C. Bai, F. Pasqualetti, and V. Gupta. "Data-injection attacks in stochastic control systems: Detectability and performance tradeoffs," *Automatica*, 82(0):251-260, 2017.

F. Pasqualetti and Q. Zhu. "Design and Operation of Secure Cyber-Physical Systems," *Embedded Systems Letters*, 7(1):3-6, 2015.

Students: Gianluca Bianchin, Rajasekhar Anguluri, Vaibhav Katewa