FORCES
FOUNDATIONS OF RESILIENT
CYBER-PHYSICAL SYSTEMS

Network interdiction and inspection models for cyber-physical security

Saurabh Amin
MIT

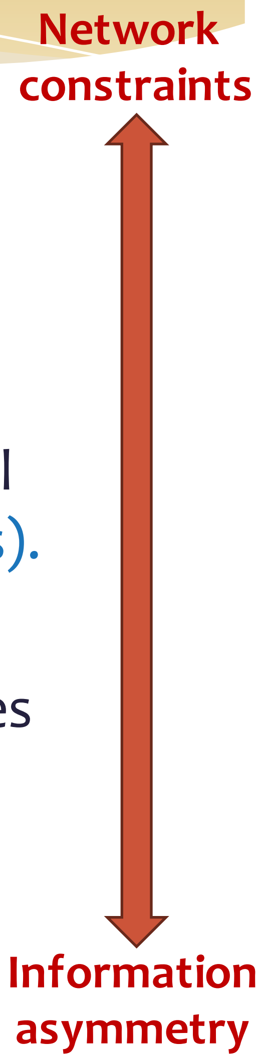
NSF Review Meeting, January 25-26, 2017

Focus: Three problems in CPS security

1. **Network interdiction models**: Vulnerability assessment of electricity networks, and design of protection strategies against disruptions (with D. Shelar, S. Zonouz, P. Sun)

2. **Network Security games**: Optimal monitoring of water networks to achieve detection of strategic failures using small # of sensors (with M. Dahan, L. Sela, W. Abbas, X. Koutsoukos).

3. **Inspection games**: Evaluating incentives of distribution utilities to employ analytics tools and inspection technologies to limit losses, including theft (with A. Sethi, G. Schwartz, S. Sastry)



Network constraints

Information asymmetry

Model 1: Network interdiction (literature)

Sequential interaction between players:

- Bilevel (maximin), Stackelberg, and multi-level optimization models

(+) Outer and inner levels may have different objectives; network constraints can be imposed; allow both integer and continuous variables

(+) Computational approaches: Benders decomposition and KKT-based reformulation to single level MILP (scalable to medium-sized problems)

(+) Useful for vulnerability assessment, typically for physical failures (e.g., papers by D. Bienstock, S. Wright, R. Baldick)

(-) Limited emphasis on the structural insights on player strategies

(-) Adversary model is conservative and defender assumed to have perfect observability of the attack (e.g., line / generator disconnects)

(-) Cyber aspects are typically not included in these models.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

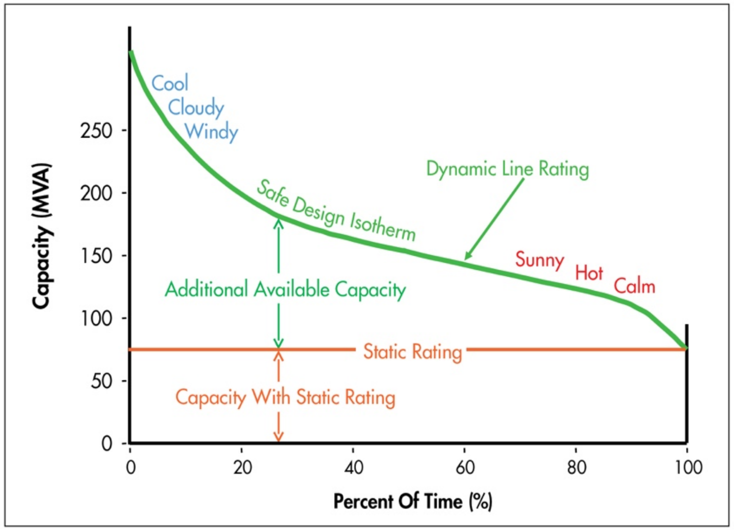
Model 1: Attacks to Dynamic Line Ratings (DLRs)

Bilevel problem

- Leader (Attacker): Compromises DLRs
- Follower (Defender): Economic dispatch problem for the compromised system

- **Problem**: Find optimal attack plan to maximinimize line rating violations

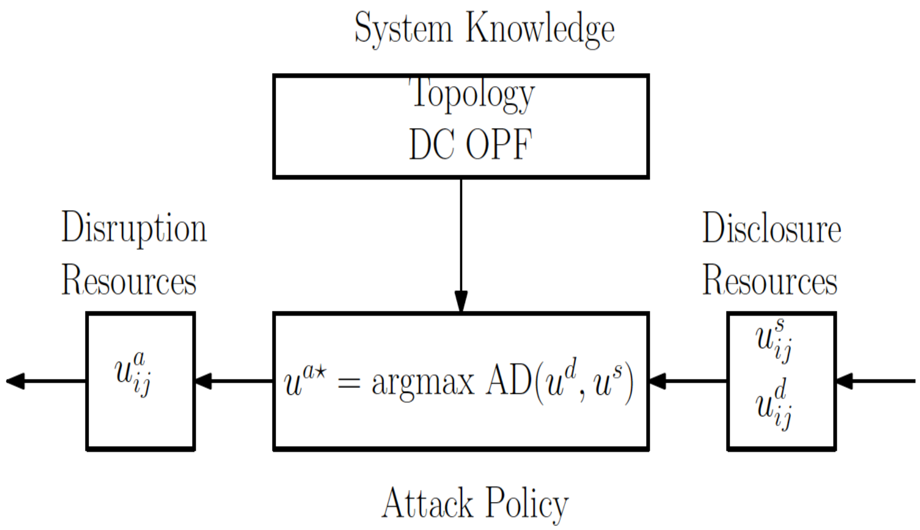
Figure 1: Tapping into existing capacity above the static rating

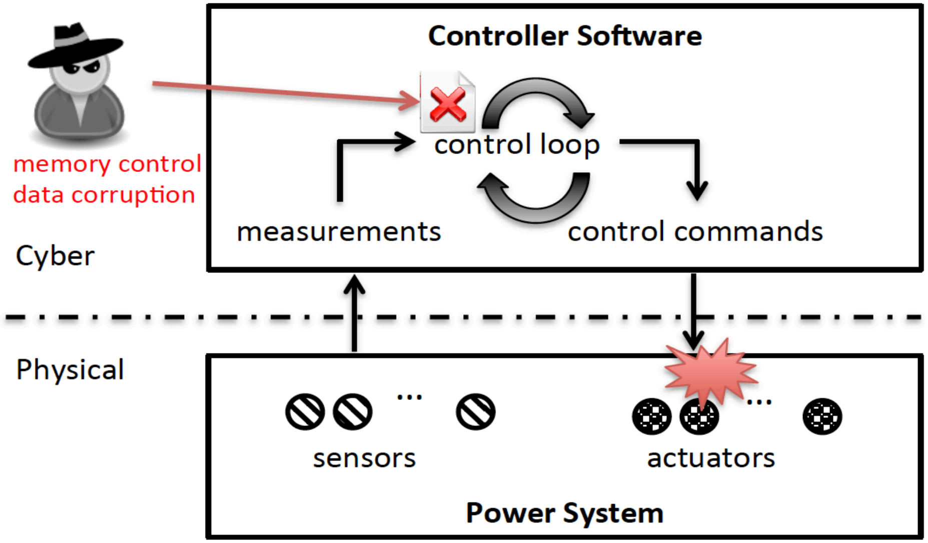


Source: Valley Group

- **Solution approach**: Apply KKT optimality conditions for the inner problem, and reformulate complementarity constraints
- Use branch-and bound techniques to solve the resulting MILP

1: Underlying CPS security problem



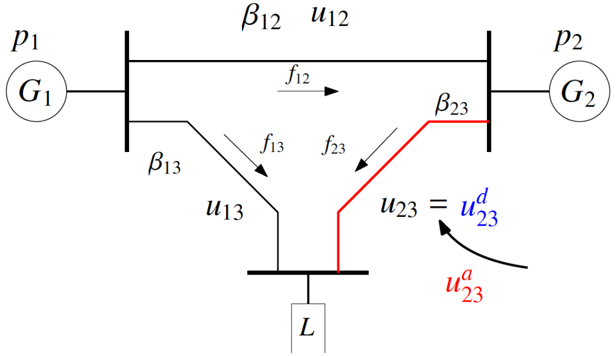


Control systems security viewpoint (Teixeira, Johansson, Sandberg ‘15)

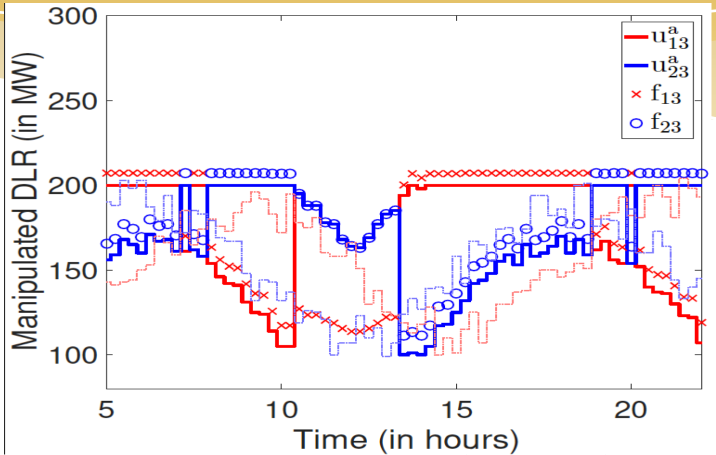
?

Software systems security viewpoint

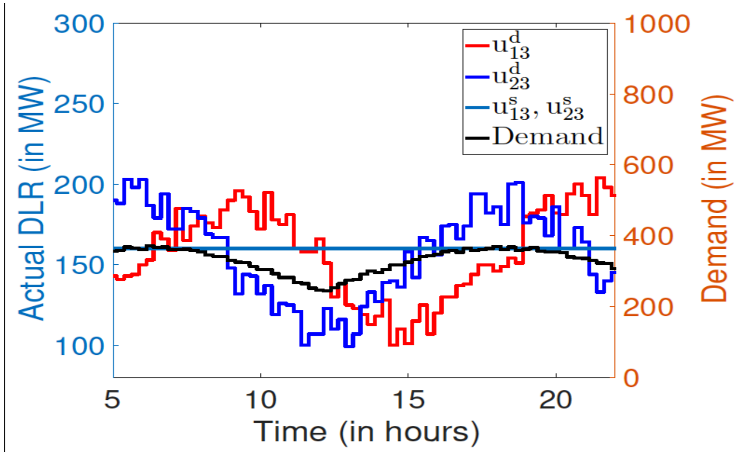
1: Optimal attack strategy



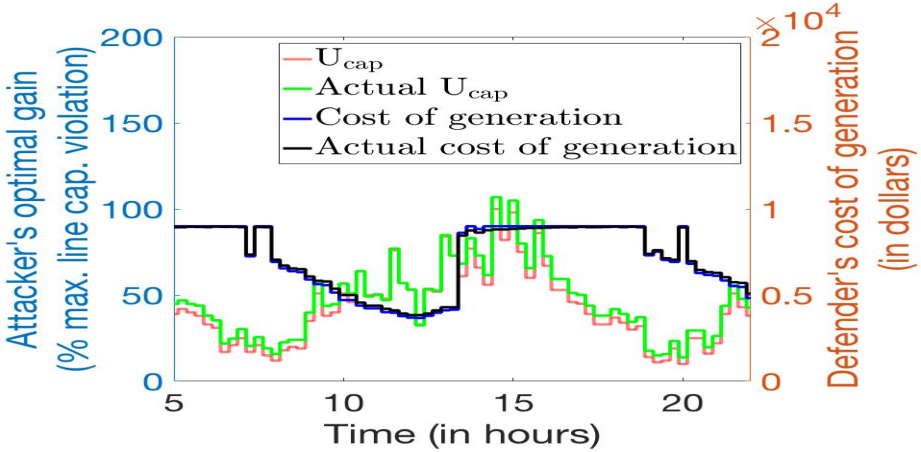
3-bus system



Attacker strategy (largely) exhibits a *bang-bang* policy.



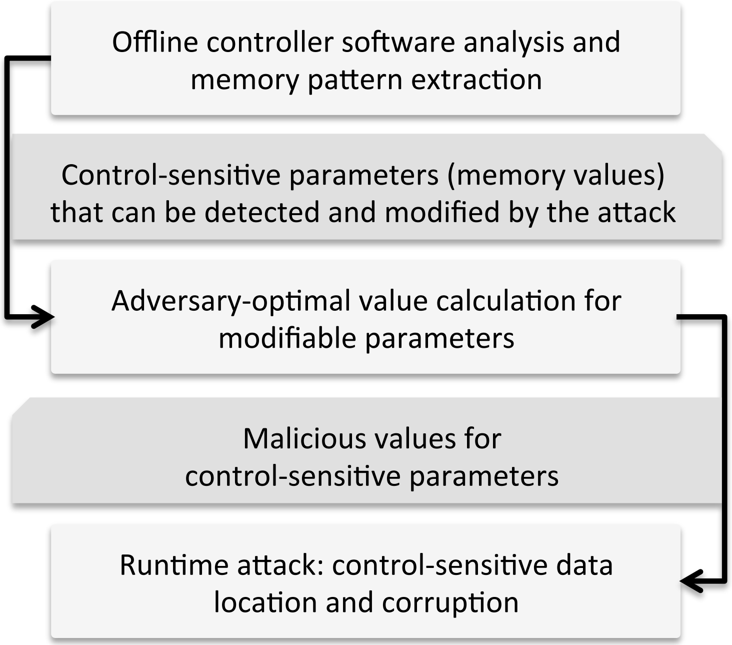
True line capacity ratings and demand over 24 hour horizon



Attacker's gain and operator's cost

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

1: Semantics-aware ``optimal” memory attack



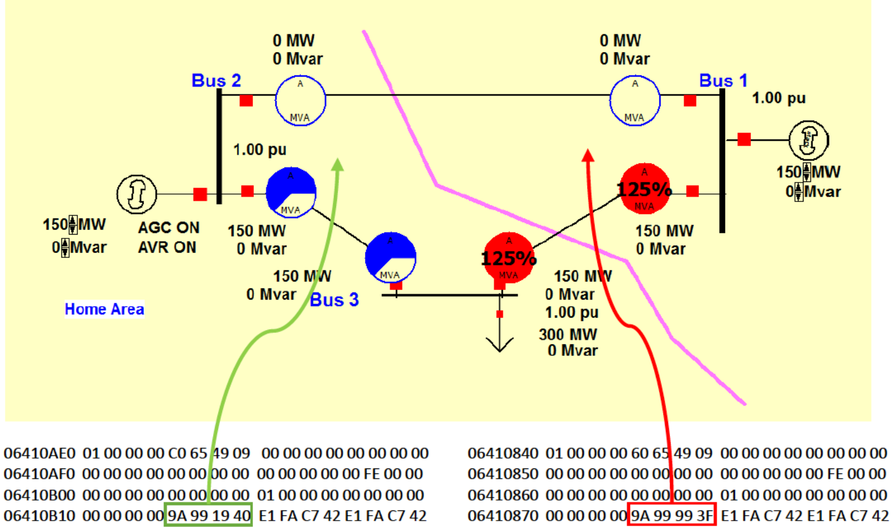
Offline controller software analysis and memory pattern extraction

Control-sensitive parameters (memory values) that can be detected and modified by the attack

Adversary-optimal value calculation for modifiable parameters

Malicious values for control-sensitive parameters

Runtime attack: control-sensitive data location and corruption



```
06410AE0  01 00 00 00 C0 65 49 09   00 00 00 00 00 00 00 00
06410AF0  00 00 00 00 00 00 00 00  00 00 00 00 00 FE 00 00
06410B00  00 00 00 00 00 00 00 00  01 00 00 00 00 00 00 00
06410B10  00 00 00 00 9A 99 19 40  E1 FA C7 42 E1 FA C7 42

06410840  01 00 00 00 60 65 49 09  00 00 00 00 00 00 00 00
06410850  00 00 00 00 00 00 00 00  00 00 00 00 00 FE 00 00
06410860  00 00 00 00 00 00 00 00  01 00 00 00 00 00 00 00
06410870  00 00 00 00 9A 99 99 3F  E1 FA C7 42 E1 FA C7 42
```

Post-attack power system state

Potential mitigations

- Fine-grained data isolation mechanisms
- Controller output verification
- N-version programming
- Attack-aware distributed control

Model 2: Network security games (literature)

Simultaneous interaction between players:

- Attacker targets one or more network components; defender chooses a protection strategy. Both players subject to resource constraints
- Typically formulated as [Strategically equivalent / non-] zero-sum game

(i) Theory of zero-sum games is well-developed:

  - Computational approaches to solve small-to-mid sized problems
  - Modeling flexibility: CPS Network structure, budget constraints

(ii) Theory of dynamic games of incomplete info: Aumann, Maschler, Renault,…

Key issues (more in Demos's talk):

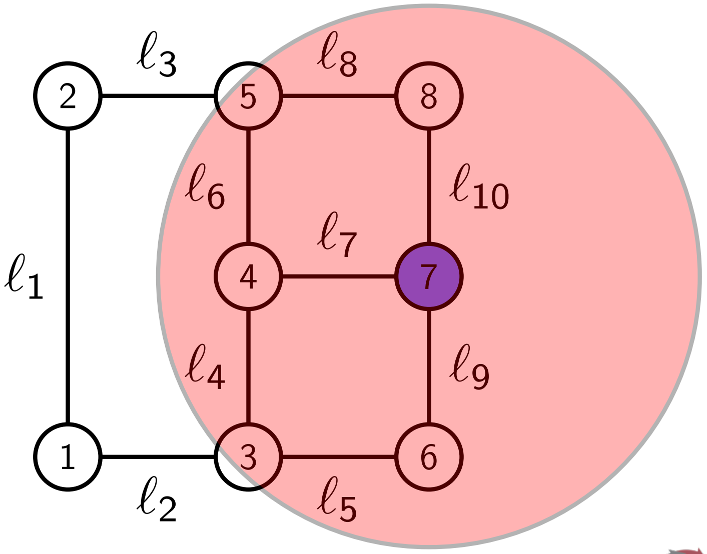
(?) How to compute equilibria in scalable manner for large-scale CPS

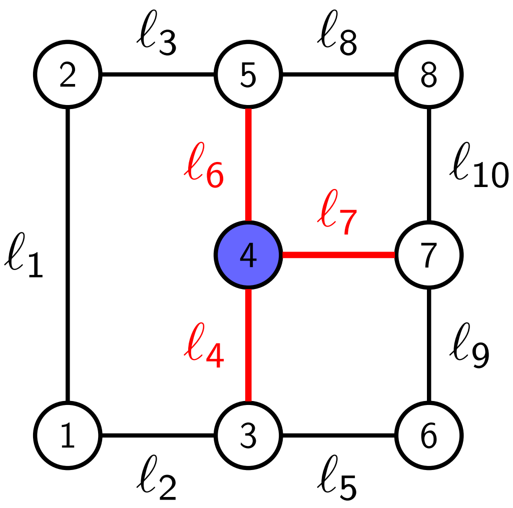
(?) How to characterize the equilibrium structure and what are the practical implications of player strategies, e.g., monitoring, protection, investment

Let us consider the problem of water network sensing under strategic attacks.

2. Sensing Model

- Undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{L})$.
  - $\mathcal{N}$: Set of nodes.
  - $\mathcal{L}$: Set of links.
- For every node $i \in \mathcal{N}$, $\mathcal{C}_i \subseteq \mathcal{L}$ represents the subset of links monitored by a sensor placed in node i. For example, $\mathcal{C}_i$ may represent:
  - The links that are within a certain distance from i.
  - The adjacent links of node i.



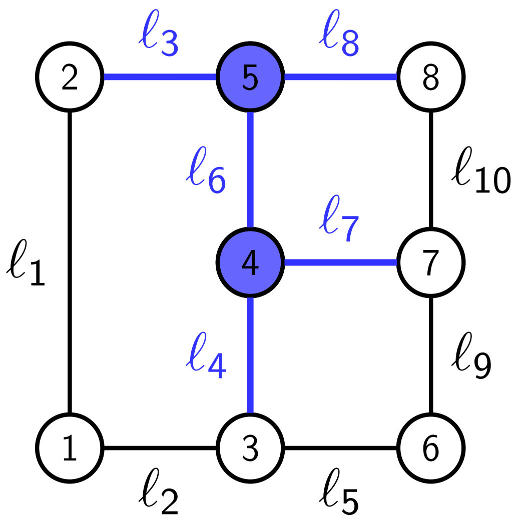


2. Sensor Placement

- For a given sensor placement $S \subseteq \mathcal{N}$, the subset of links that are monitored by at least one sensor in S is given by:

$$\mathcal{C}_S := \bigcup_{i \in S} \mathcal{C}_i$$

- Example: For one-hop sensing, a sensor node can detect only adjacent edges:



- $S = \{4, 5\}$
- $\mathcal{C}_4 = \{\ell_4, \ell_7, \ell_6\}, \quad \mathcal{C}_5 = \{\ell_3, \ell_6, \ell_8\}$
- $\mathcal{C}_S = \{\ell_3, \ell_4, \ell_6, \ell_7, \ell_8\}$

2. Game

$\Gamma := \langle \{1,2\}, (\mathcal{A}_1, \mathcal{A}_2), (u_1, u_2) \rangle$

- Player 1 (Defender) chooses a sensor placement $S \in \mathcal{A}_1$.

$$\mathcal{A}_1 = \left\{ S \subseteq \mathcal{N} \;\middle|\; |S| \leq b_1 \right\}$$

- Player 2 (Attacker) chooses a subset of links $L \in \mathcal{A}_2$ to target:

$$\mathcal{A}_2 = \left\{ L \subseteq \mathcal{L} \;\middle|\; |L| \leq b_2 \right\}$$

Resource limitations:

- b_1: maximum number of simultaneously deployable sensors.
- b_2: maximum number of simultaneous link failures.

2. Game

$$\Gamma := \langle \{1,2\}, (\mathcal{A}_1, \mathcal{A}_2), (u_1, u_2) \rangle$$

- Player 1 (Defender) payoff: number of **detected** failure events

$$u_1(S, L) := |\mathcal{C}_S \cap L|$$

- Player 2 (Attacker) payoff: number of **undetected** failure events

$$u_2(S, L) := |L| - |\mathcal{C}_S \cap L|$$

- Mixed-extension: for $(\sigma^1, \sigma^2) \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2)$:

$$U_1(\sigma^1, \sigma^2) = \mathbb{E}[u_1(S, L)], \qquad U_2(\sigma^1, \sigma^2) = \mathbb{E}[u_2(S, L)]$$

- $\mathcal{S}_\Gamma$ is the set of Nash Equilibria.

2. Computation of Nash Equilibria

Challenge: How to compute Nash equilibria of this game in a scalable manner?

- **Option 1:** Strategic equivalence to a zero-sum game:
  - $\mathcal{S}_\Gamma$ can be obtained by solving:

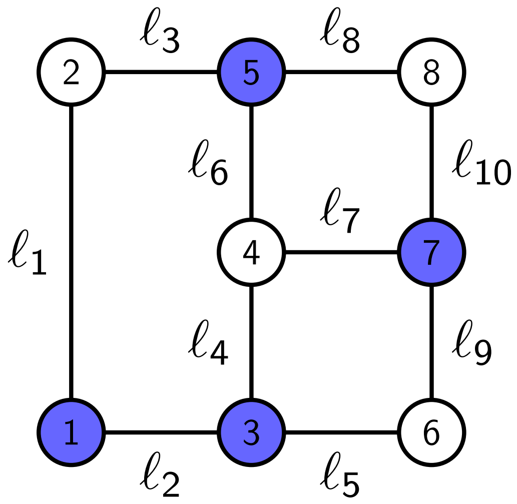
$$
\begin{array}{llll}
(\mathrm{LP}_1) & \max & z & \\
 & \text{s.t.} & \widetilde{U}_1(\sigma^1, L) \geq z, & \forall L \in \mathcal{A}_2 \\
 & & \sigma^1 \in \Delta(\mathcal{A}_1) &
\end{array}
\qquad
\begin{array}{llll}
(\mathrm{LP}_2) & \max & z' & \\
 & \text{s.t.} & \widetilde{U}_2(S, \sigma^2) \geq z' & \forall S \in \mathcal{A}_1 \\
 & & \sigma^2 \in \Delta(\mathcal{A}_2) &
\end{array}
$$

  - Issue: If $|\mathcal{N}| = |\mathcal{L}| \approx 1000$ and $b_1 = b_2 = 10$, then the number of variables and constraints in both LPs is $\mathbf{2.66 \cdot 10^{23}}$!!!
- **Option 2:** Computing approximate Nash equilibria using solutions of two classical combinatorial optimization problems.

2. Minimum Set Cover

Defender minimizes the number of sensors needed to monitor all the links:

$$(\text{MSC}) \quad \underset{S \subseteq \mathcal{N}}{\text{minimize}} \quad |S|$$
$$\text{subject to} \quad \bigcup_{i \in S} \mathcal{C}_i = \mathcal{L} \quad \text{(full coverage)}$$



Example: For one-hop sensing:

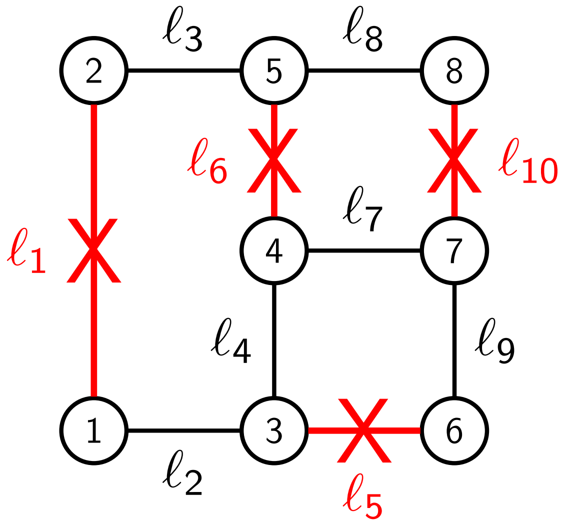
- $\{1, 3, 5, 7\}$ is a sensor placement over an MSC.

- n^*: size of MSC.

2. Maximum Set Packing

Attacker maximizes the number of link failures such that a sensor placed in any node cannot detect more than one failure:

$$\text{(MSP)} \quad \underset{L \subseteq \mathcal{L}}{\text{maximize}} \quad |L|$$
$$\text{subject to} \quad |\mathcal{C}_i \cap L| \leq 1, \quad \forall i \in \mathcal{N} \qquad \text{(at most one detection)}$$



Example: For one-hop sensing:

- $\{\ell_1, \ell_5, \ell_6, \ell_{10}\}$ is the attack of an MSP.

- m^*: size of MSP.

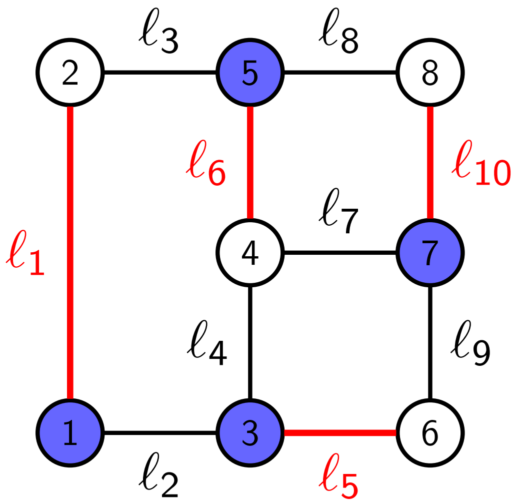
- Weak Duality: $m^* \leq n^*$

2. Resource Limitations

Case of Interest

The network is sufficiently large relative to players' resources, in the sense that Player 1 (resp. Player 2) has less resources than the size of the MSC (resp. the MSP),

$$b_1 < n^*, \qquad b_2 < m^*$$



- MSC: $\{1, 3, 5, 7\}$
- MSP: $\{\ell_1, \ell_5, \ell_6, \ell_{10}\}$
- Let $b_1 < 4$ and $b_2 < 4$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

2. Main Theorem

Theorem

In the case where $b_1 < n^*$ and $b_2 < m^*$, we have:

- *In any equilibrium $({\sigma^1}^*, {\sigma^2}^*) \in \mathcal{S}_\Gamma$, the players' payoffs are constant and can be bounded as follows:*

$$\frac{b_1 b_2}{n^*} \leq U_1({\sigma^1}^*, {\sigma^2}^*) \leq \frac{b_1 b_2}{m^*}$$

$$b_2\left(1 - \frac{b_1}{m^*}\right) \leq U_2({\sigma^1}^*, {\sigma^2}^*) \leq b_2\left(1 - \frac{b_1}{n^*}\right)$$

- *In any equilibrium $\sigma^* \in \mathcal{S}_\Gamma$, the expected detection rate is constant and bounded as follows:*

$$\frac{b_1}{n^*} \leq \mathbb{E}_{\sigma^*}\left[\frac{|\mathcal{C}_S \cap L|}{|L|}\right] \leq \frac{b_1}{m^*}$$

FORCES

FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

2. Number of Sensors

(Q) How many sensors are required to limit the losses in the network?

- Detection rate in equilibrium:

$$\frac{b_1}{n^*} \leq \mathbb{E}_{\sigma^*}\left[\frac{|\mathcal{C}_S \cap L|}{|L|}\right] \leq \frac{b_1}{m^*}$$

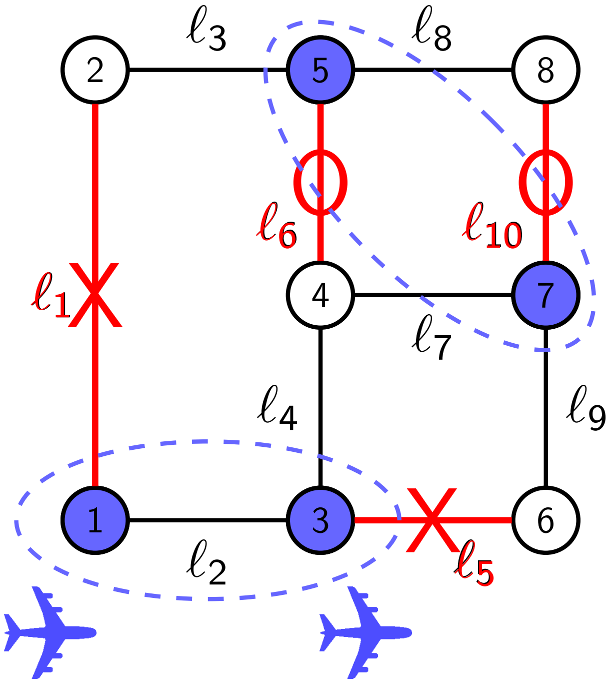
- If α is the target detection rate:
  - Necessary condition: $b_1 = \lceil \alpha m^* \rceil$
  - Sufficient condition: $b_1 = \lceil \alpha n^* \rceil$

2. MSC-MSP Based Strategies

Theorem

- *For any MSC S^{min} and any MSP L^{max}, $(\sigma^{S^{min}}, \sigma^{L^{max}})$ defined as follows is an $\epsilon-$NE, where $\epsilon = b_1 b_2 \frac{n^*-m^*}{n^* m^*}$.*

Example:



- $b_1 = b_2 = 2$
- $\sigma^{1^*}_{\{1,3\}} = \frac{1}{2}, \quad \sigma^{1^*}_{\{5,7\}} = \frac{1}{2}$
- $\sigma^{2^*}_{\{\ell_1,\ell_5\}} = \frac{1}{2}, \quad \sigma^{2^*}_{\{\ell_6,\ell_{10}\}} = \frac{1}{2}$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

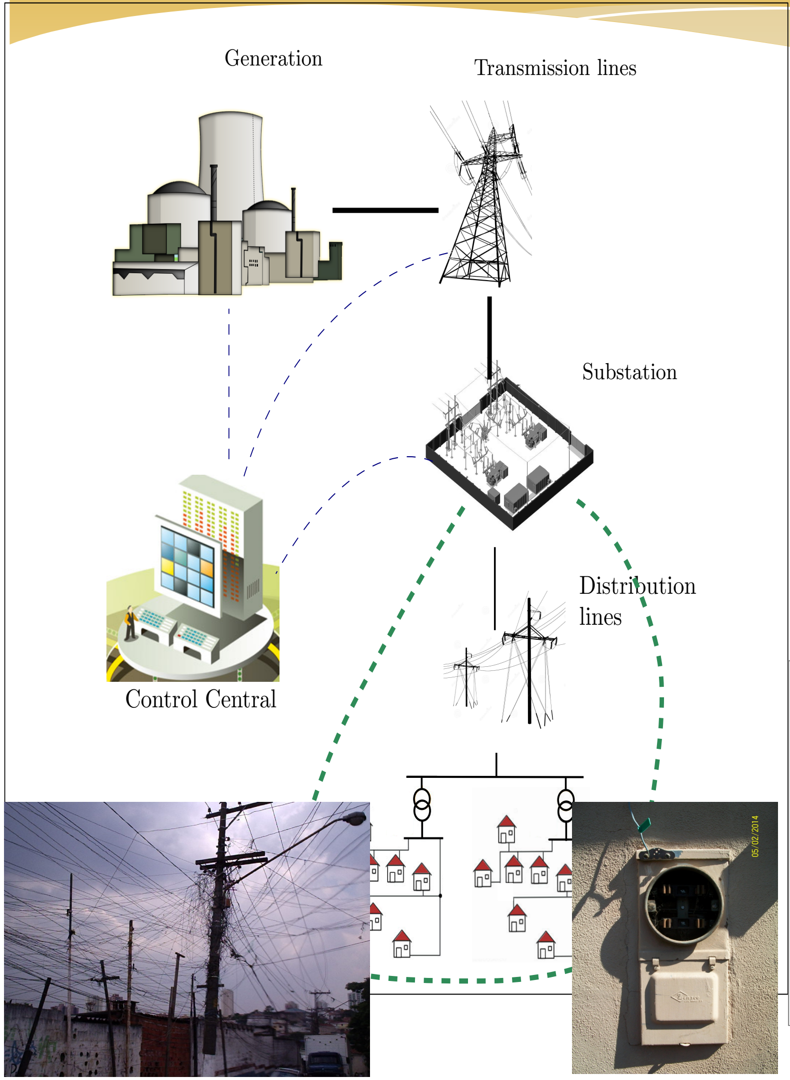
2. Examples

Table: Network data

Network	Length	No. of	No. of	m^*	n^*	Running	time [s]	Optimality
	[km]	pipes	nodes			(MSP)	(MSC)	Gap
Net1	37.56	168	126	7	7	0.05	0.11	0 %
Net2	91.29	366	269	15	15	0.01	0.03	0 %
Net3	96.58	496	420	18	19	0.02	0.05	5.3 %
Net4	137.05	603	481	28	28	0.09	0.08	0 %
Net5	123.20	644	543	24	24	0.08	0.06	0 %
Net6	166.60	907	791	31	31	0.03	0.08	0 %
Net7	153.30	940	778	28	30	0.06	0.08	6.7 %
Net8	152.25	1124	811	18	19	0.39	0.41	5.3 %
Net9	260.24	1156	959	62	64	0.03	0.05	3.1 %
Net10	247.34	1614	1325	45	45	0.14	0.22	0 %
Net11	779.86	14965	16000	119	121	4.34	8.36	1.7 %
Net12	1,844.04	12523	14822	352	361	0.77	4.06	2.5 %
Net13	476.67	24681	25484	50	52	58.89	68.67	3.8 %

- MSC/MSP-based strategies provide solutions that are **accurate** and **scalable**.
- Results are still applicable if MSC and MSP are computed with approximation algorithms.

3. Diagnostics for Energy Diversion Attacks

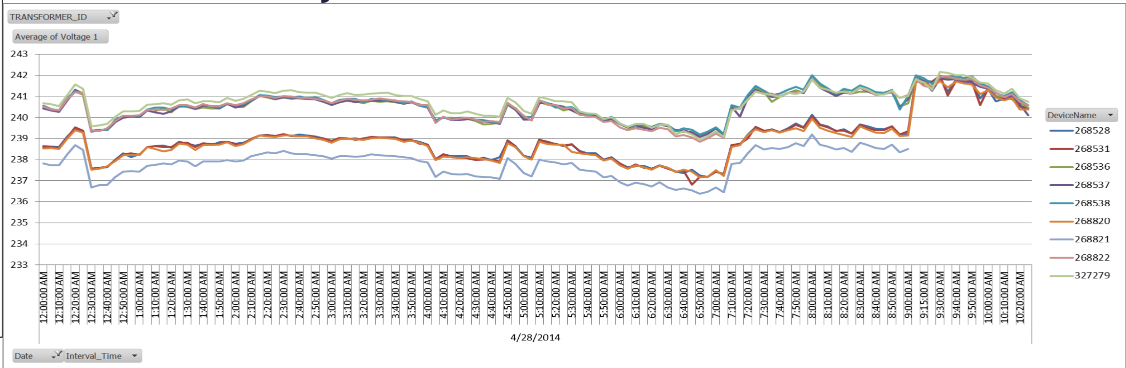


Distribution utility:

- Fighting theft: identify fraudulent consumers
- Smart meter data: statistical detection tools to fight theft and impose fines
- Choose investment in theft prevention

Consumers:

- Compromise meters and inject false data
- New ways to steal and evade detection!



Modesto irrigation district: energy diversion data

Model 3: Energy theft by fraudulent customers

- Inspection games are a class of incomplete information games that account for illegal actions by a strategic inspectee who wants to evade detection by an inspector
- Distribution utility (inspector):
  - faces two “types” of consumers: genuine and fraudulent
  - uses an intrusion detection system (IDS) for monitoring
  - IDS is a classifier with tradeoff between detection and false alarm rates
- Fraudulent consumers (inspectee):
  - Manipulate meter readings to under-report actual consumption

Two questions:

- How should the utility evaluate the benefit of IDS?
- Should the utility use the fixed configuration or optimally tune an IDS?

Model 3: Our contribution

Bayesian inspection game: models incomplete info for utility in both
(a) type of customer: fraudulent or genuine
(b) average level of theft: high or low

- We study the tradeoff in terms of gain from fraud investigations (and deterrence) versus cost of investigation and false alarms
- We evaluate the value of IDS with incomplete information on both type of customer and level of theft and analyze its properties.
  - Value of IDS: Monetary benefit in using a tuned IDS relative to a fixed configuration (or even no IDS) case
  - Value of information on theft level: Additional benefit if an accurate estimate of the average theft level is used to configure IDS

Summary: Game-theoretic models for CPS security

- **Network interdiction model**
  (+) Computational tools for bilevel problems with network constraints
  (-) Do not capture information asymmetries and repeated interactions
- **Network security games**
  (+) Computational tools for (non) zero-sum games on networks
  (-) Interpretation of equilibrium (and off-equilibrium) strategies
- **Inspection games**
  (+) Relevant for monitoring and enforcement problems
  (+) Capture the present of asymmetric information in these problems
  (-) Do not [naturally] capture the underlying network structure

*** Introduction of info. asymmetry & CPS dynamics requires care!**

- Computational tractability
- Interpretation of equilibria

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS