On Bounded Rationality in Cyber-Physical Systems Security: Game-Theoretic Analysis with Application to Smart Grid Protection

CPSR-SG 2016 – CPS Week 2016

Anibal Sanjab and Walid Saad
April 12, 2016
Vienna, Austria
Outline

• CPS Security and its Challenges
• CPS Security Games
• Bounded Rationality in CPS Security Games
• CPS Security Model
  • Attack diffusion model
  • Game-theoretic formulation
  • Bounded rationality
  • Case analysis: smart grid wide area protection
Cyber Physical Systems (CPS)

- **Cyber-physical systems**
  - Physical system
  - Cyber layer
- **Purpose:** smart systems

Smart Grid: “an electricity network that can intelligently integrate the actions of all users connected to it - generators, consumers and those that do both - in order to efficiently deliver sustainable, economic and secure electricity supplies” – European Technology Platform for Smart Grids 2035 Strategic Research Agenda
CPS Security - Exposed Vulnerability

• Stuxnet (2010):
  • Target: 14 industrial systems in Iran – plant for Uranium enrichment
  • Computer worm targeting control of industrial systems – Programmable Logic Controllers (PLC)

• Water (2000):
  • Target: Maroochy Water Services in Queensland, Australia
  • Block communication links with waste water pumping stations
  • 1 million liters of sewage water spill

• Transportation (2001):
  • Target: Port of Houston, TX, USA
  • Denial-of-Service over its ship assistance system
CPS Security

• Cyber-Physical Systems:
  • Beneficial but vulnerable!
  • Solution: devise solutions to make systems less vulnerable, more robust, and more resilient to attacks.

• CPS Security Research:
  • Type of threats/attacks: data injection, DoS, time synchronization …
  • Purpose: prevention, detection, mitigation …
  • Area of research: smart grids, transportation systems, water distribution, smart cities …
  • Attack vs. Defense: Game Theory!
CPS Security Games

• What is Game Theory?
Set of mathematical tools to analyze strategic interaction and decision making between entities with interconnected interests.

• CPS Security Games:
  • Players: Attacker(s) and Defender(s)
  • Actions: set of attack strategies | set of defense strategies
  • Objective:
    • Attackers: optimize a payoff function reflecting
      • Level of caused damage to the system
      • Financial benefit that it can earn through attack, etc.
    • Defenders: optimize a payoff function reflecting
      • Deviation from normal operating state (minimize)
      • Operation performance level and/or amount of uncompromised resources (maximize)
Bounded Rationality

• Game theory requires *rationality*:
  • Players are purposeful: optimize an objective function
  • Players do not make mistakes
• Risk, stress, incomplete information, extreme complexity, constraints (time, etc.) → limited rationality
• Prone to making mistakes
  • 1965 Northeast blackout → 30 million people affected in Ontario and 8 U.S states → cause: Human mistake: incorrect setting of a protective relay by maintenance personnel!
Cognitive Hierarchy

- Perception over the skill levels of opponents
- Multiple thinking steps
  - Example: the beauty contest
- In criminology studies, attackers carry out a reconnaissance phase
  - Accuracy of their perception?
- $k$ thinking steps:
  - Player assumes having most sophisticated strategy (level $k$)
  - Presume a probability distribution over the skill levels of opponents
    - Proportion of opponents at each thinking step $0 \rightarrow k - 1$
- CPS security application:
  - Example: defensive has highest knowledge of system model $\rightarrow$ uses perception of adversaries skill levels distribution to design a defense strategy.
Beauty Contest – Cognitive Hierarchy

*General Theory of Employment, Interest, and Money (1936) – John Keynes*

**Beauty contest game:** A number of participants are asked to choose a number from 0 – 100. The player whose number is closest to 2/3 of the average of all chosen numbers wins.

Experimental group avg: 20 - 35
Cognitive Hierarchy

• Perception over the skill levels of opponents

• Multiple thinking steps
  • Example: the beauty contest

• In criminology studies, attackers carry out a reconnaissance phase
  • Accuracy of their perception?

• $k$ thinking steps:
  • Player assumes having most sophisticated strategy (level $k$)
  • Presume a probability distribution over the skill levels of opponents
    • Proportion of opponents at each thinking step $0 \rightarrow k - 1$

• CPS security application:
  • Example: defensive has highest knowledge of system model $\rightarrow$ uses perception of adversaries skill levels distribution to design a defense strategy.
CPS SECURITY MODEL

Game-Theoretic Formulation | Application to Smart Grids’ Security
Attack Diffusion Model

• CPS Model:
  - $N_c$ cyber nodes, $N_p$ physical nodes
  - $r_{c,p}$: weight interconnection between physical node $p$ and cyber node $c$.
    - Weight of data sent by $c$ on control action over $p$
    - $r_{c,p} = \Pr(p \text{ fails} | c \text{ has failed})$
  - Failure of $c$:
    - Implication: send corrupt data
    - Reason: cyber attack, misconfiguration,…
  - $\pi_p$: probability of failure of $p$ due to failures in the cyber layer
    \[
    \pi_p = \sum_{c=1}^{N_c} r_{c,p} \kappa_c
    \]
Attack Diffusion Model

- \( R = [r_{c,p}]_{N_c \times N_p} \): cyber-physical interconnection matrix
- \( \pi = [\pi_1, \ldots, \pi_{N_p}] \in [0,1]^{N_p} \): failure probability vector of physical nodes
- \( \kappa = [\kappa_1, \ldots, \kappa_{N_c}] \in [0,1]^{N_c} \): failure probability vector of cyber nodes

\[
\pi_p = \sum_{c=1}^{N_c} r_{c,p} \kappa_c
\]

\[
\pi = \kappa R
\]

- \( f_p \): cost of failure of physical node \( p \)
- Expected total loss to system:

\[
E_f = \sum_{p=1}^{N_p} \pi_p f_p
\]
Game Formulation

• Under no attack:
  • \( \kappa = [\kappa_1, \ldots, \kappa_{N_c}] \) small \( \rightarrow \pi = [\pi_1, \ldots, \pi_{N_p}] \) small
  • Minimize \( E_f = \sum_{p=1}^{N_p} \pi_p f_p \) is a reliability evaluation problem.

• Under cyber-attack
  • \( c \) is attacked \( \rightarrow \kappa_c = 1 \rightarrow \pi \uparrow \rightarrow E_f \uparrow \)
  • \( c \) is defended \( \rightarrow \kappa_c = 0 \rightarrow \pi \downarrow \rightarrow E_f \downarrow \)

• Attacker vs. Defender Game
  • Players: defender \( d \), attacker \( a \)
  • Set of actions: which cyber nodes to attack/defend \( |S_d| = \binom{N_c}{n_d}, |S_a| = \binom{N_c}{n_a} \)
  • Utility function: \( U_d(s_d, s_a) = -U_a(s_d, s_a) = -E_f \)
  • \( s_i \in S_i, n_i = \) number of concurrently attacked/defended nodes
Bounded Rationality

• Under full rationality (standard game theory)
  • Defender and attacker play best response strategies
    → Full knowledge of $E_f = \sum_{p=1}^{N_p} \pi_p f_p$ is needed

• Defender vs Attacker
  • CPS are very complex → obtaining $f$ (vector of $f_p$’s) is challenging
  • Solution: build own perception of $f \rightarrow \hat{f}_{i \in (a,d)}$
  • Level of thinking of $i \in (a, d)$: how close is $\hat{f}_i$ to $f$
  • Higher level thinkers (Analogous to Cognitive Hierarchy Theory)
    • More intelligent
    • Have better knowledge of the system
    • Better computational capabilities
      Generate better $\hat{f}_i$
Application to Wide Area Protection

• Wide area monitoring and protection schemes:
  • Rely on global data
  • React to disturbances
    • Ex: disconnection of a transmission line, generator, or load shedding
  → Availability of a physical node dependent on spread cyber nodes
• Dependability vs Security
  • Dependability: successful isolation of a fault when it happens
  • Security: take protection actions only when disturbance occurs.

PJM 5-bus system:
Application to Wide Area Protection

• Application:
  • 1 attacker vs. 1 defender
  • # concurrent attacks = 2
  • # secured nodes = 2
  • Energy markets implications
    • Optimal Power Flow
Application to Wide Area Protection

- $V^O$: value function of the original OPF ($/h$
- $V^{pi}$: value function of the OPF with loss of $p_i$ ($/h$
- $T^{pi}$: time needed to bring $p_i$ back to operation (h)
- $CR^{pi}$: cost of repair of $p_i$

\[ f_{pi} = (V^{pi} - V^O)T^{pi} + CR^{pi} \]

Nash equilibrium strategies:

\[
\gamma_d^* = [0.2931, 0.3034, 0.3107, 0.0842, 0.0047, 0.0040]
\]
\[
\gamma_a^* = [0.1276, 0.1244, 0.1222, 0.1922, 0.2167, 0.2169]
\]

\[ \bar{U}_d = -\bar{U}_a = -$110,240 \]
Application to Wide Area Protection

- **Bounded rationality:**
  - Requirements for the solution of the OPF:
    - knowledge of the full system
    - high computational capabilities
  - 3 types of attackers:
    - Level 0 \( (l_0) \): chooses attack randomly
    - Level 1 \( (l_1) \): attacks line with highest power flow
    - Level 2 \( (l_2) \): can solve OPF, attacks line with highest \( f_{p_i} \)
  - Defender:
    - Can solve OPF
    - Strategic thinker
    - Highest level thinker
Application to Wide Area Protection

• Defender faced with an attacker of type $k$ with probability:

$$\alpha(k) = \frac{e^{-\lambda} \lambda^k}{k!} \rightarrow \text{Poisson distribution}$$

$$\begin{align*}
\alpha(0) &= \alpha(1) = \alpha(2) = \tau \\
\alpha(0) + \alpha(1) + \alpha(2) &= 1
\end{align*}$$

• $\tau < 1 \rightarrow$ a low level attacker is most probable

• $\tau > 1 \rightarrow$ a high level attacker is most probable

• Optimal defense against: $\alpha(0)l_0 + \alpha(1)l_1 + \alpha(2)l_2$
Application to Wide Area Protection

- Advantage for accounting for multiple possible types of attackers
- Probability of facing a higher attacker *increases*, the gain from deviating from the NE *decreases*
Conclusions

• Introduced general CPS security model showing propagation of attacks from cyber to physical
• Attacker vs. Defender Game
• Bounded rationality
• Application to wide area protection
  • Optimal defense strategy accounting for multiple attacker types
  • Beneficial deviation from the NE defense strategy
ACKNOWLEDGEMENT

National Science Foundation

Contact information: anibals@vt.edu, walids@vt.edu

Thank you!

Questions…