FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

A Game-Theoretic Approach for Alert Prioritization in Cyber-Physical Systems

Aron Laszka

Vanderbilt University

Berkeley UNIVERSITY OF CALIFORNIA

Massachusetts Institute of Technology

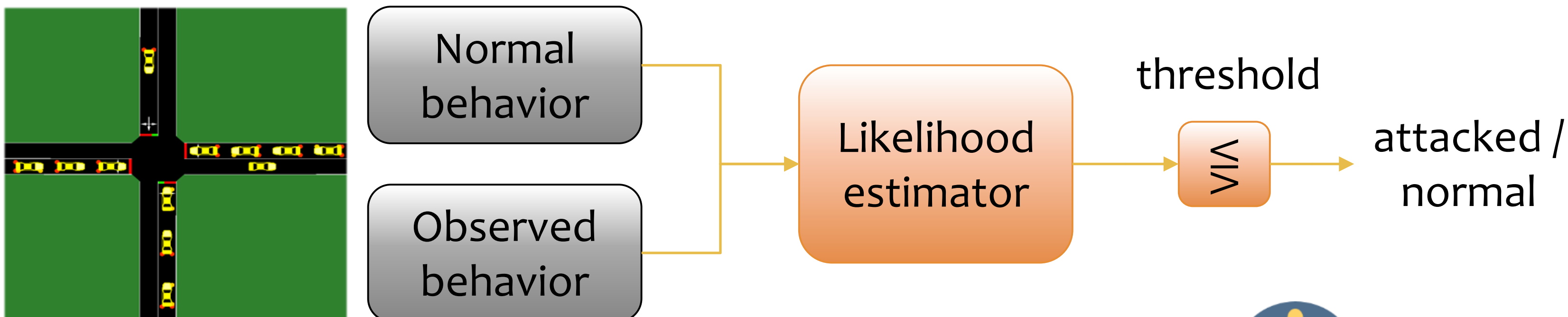
NSF

MICHIGAN

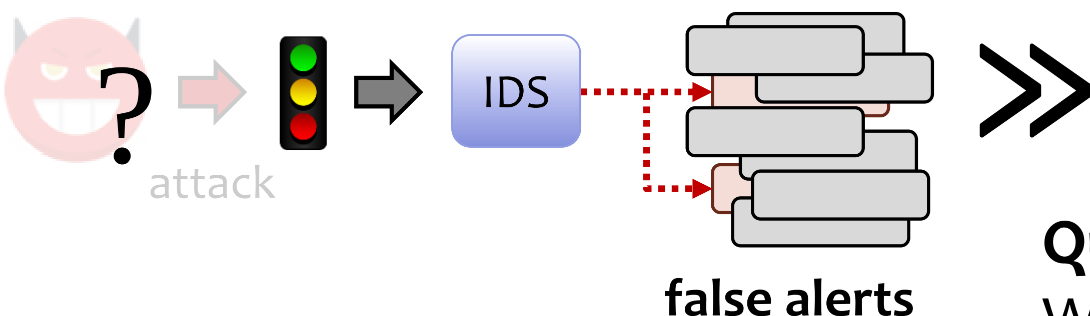
VANDERBILT UNIVERSITY

Physical Anomaly Based Intrusion Detection

* Resilience to novel threats through anomaly-based detection
* Previously: traffic-anomaly based attack detection



* Problem

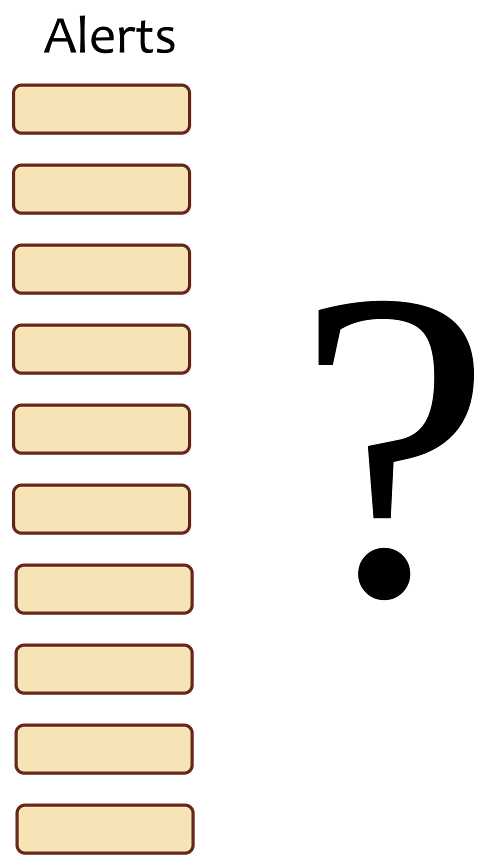


Alert-investigation budget *B*
(available manpower, …)

Question:
Which alerts to investigate?

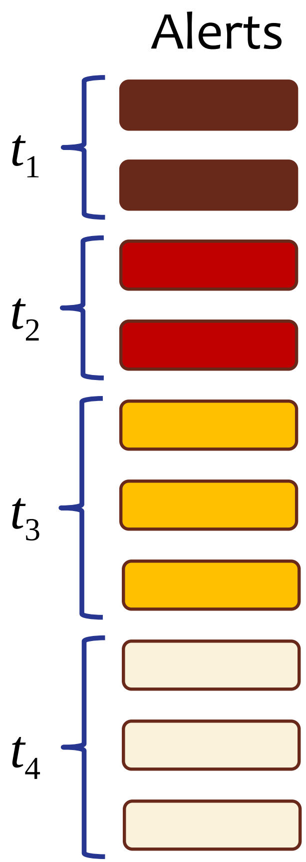
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Alert Prioritization



Alert Prioritization

Alert types T



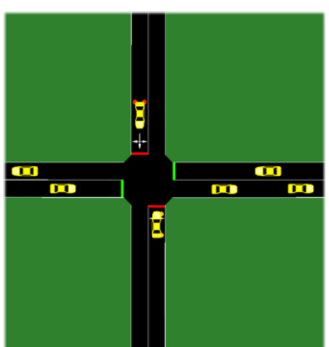
Alert Types

- Alert types T
  - example:

anomalous system		anomaly magnitude: low	anomaly magnitude: high
	sensing	t_4	t_2
	control	t_3	t_1

  - **cumulative distribution** F_t of the number of false alerts of type t:

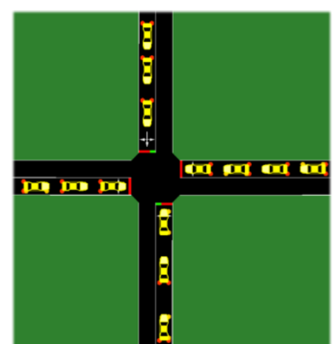
simulate normal behavior



⇨ IDS ⇨ false alerts

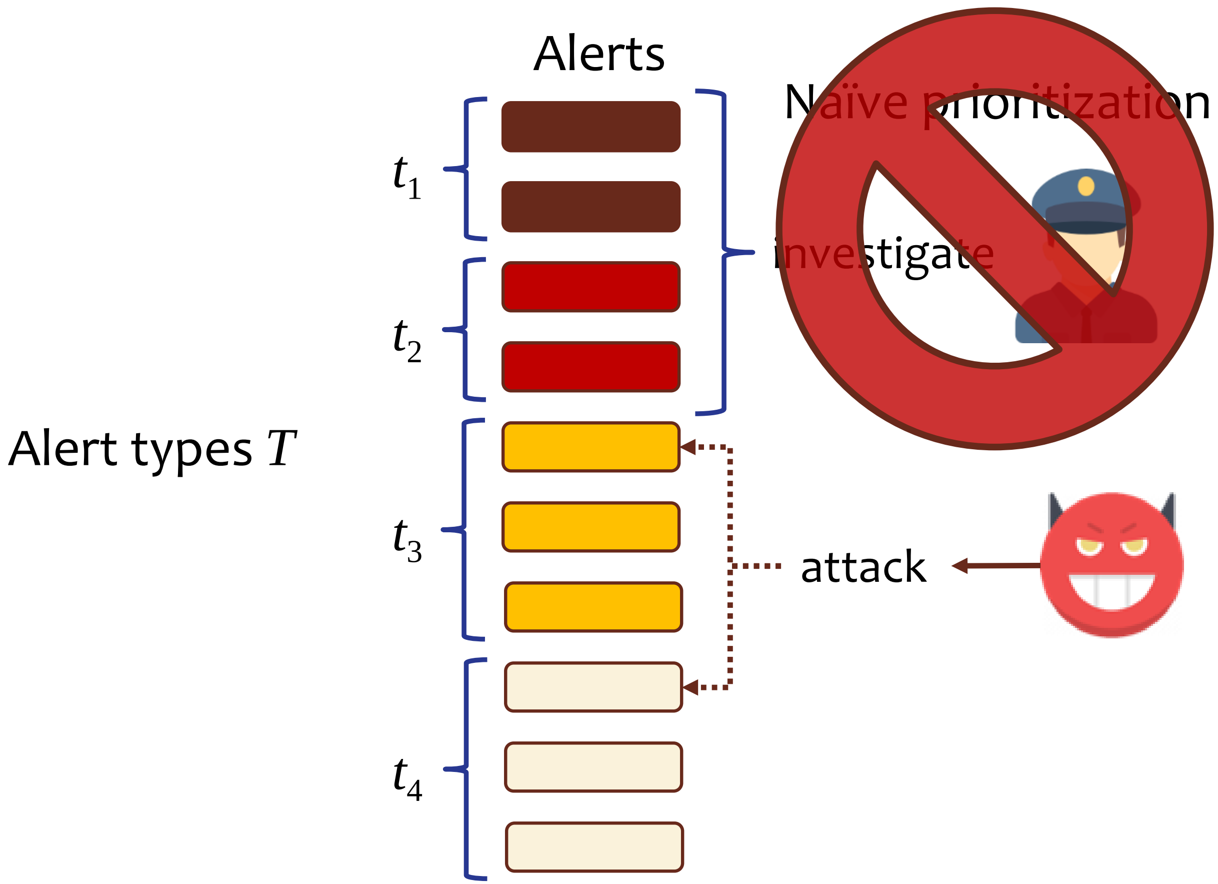
- Attacks A
  - example: attacked devices, integrity / availability
  - **impact** L_a **of attack** a: our previous work (ICCPS'16)
  - **probability** $R_{a,t}$ **of raising an alert** of type t for attack a:

simulate attack a



⇨ IDS ⇨ alerts

Alert Prioritization Problem



FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Stackelberg Security Game

* Players

1. Defender

- selects alert prioritization strategy $\boldsymbol{p}$, which is a probability distribution over orderings of T

2. Adversary

- selects an attack a from the set of possible attacks A

* Payoffs
  * let $PD(\boldsymbol{o}, a)$ be probability of investigating attack a using ordering $\boldsymbol{o}$
  * defender's expected loss = $\sum_{\boldsymbol{o} \in O} p_{\boldsymbol{o}} \cdot (1 - PD(\boldsymbol{o}, a)) \cdot G_a - K_a$
  * adversary's expected payoff = $\sum_{\boldsymbol{o} \in O} p_{\boldsymbol{o}} \cdot (1 - PD(\boldsymbol{o}, a)) \cdot L_a$
* Solution concept
  * adversary's best response: $BR(\boldsymbol{p}) = \operatorname{argmax}_{a \in A} \sum_{\boldsymbol{o} \in O} p_{\boldsymbol{o}} \cdot (1 - PD(\boldsymbol{o}, a)) \cdot L_a$
  * **optimal prioritization strategy:** $\min_{\boldsymbol{p}, a \in BR(\boldsymbol{p})} \sum_{\boldsymbol{o} \in O} p_{\boldsymbol{o}} \cdot (1 - PD(\boldsymbol{o}, a)) \cdot G_a - K_a$

(details can be found in our paper published at AAAI-17 AICS)

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Computational Results

- Computing probabilities

exponential number of terms

  - definition: $PD(\boldsymbol{o},a) = \sum_{\boxed{\hat{T} \subseteq T}} \prod_{t \in \hat{T}} R_{a,t} \prod_{t \in T \setminus \hat{T}} (1 - R_{a,t})\, PI\left(\boldsymbol{o}, \min\{i \mid o_i \in \hat{T}\}\right)$

$$PI(\boldsymbol{o}, k) = \sum_{\boxed{\substack{\boldsymbol{n}: \\ C_{o_k} + \sum_{i=1}^{k} n_i \cdot C_{o_i} \leq B}}} \left[\left(F^*_{o_k}(n_k) - F^*_{o_k}(n_k - 1)\right) \cdot \prod_{i=1}^{k-1} \left(F_{o_i}(n_i) - F_{o_i}(n_i - 1)\right) \right]$$

where C_t is investigation cost for type t

  - **dynamic programming algorithm:**

Algorithm 1 Computing $PD(\boldsymbol{o}, a)$

Input: prioritization game, prioritization $\boldsymbol{o}$, attack a
for $b = 0, 1, \ldots, B$ **do**
  $PD(\boldsymbol{o}, a, |T|, b) \leftarrow R_{a,o_{|T|}} \cdot F^*_{o_{|T|}}\left(\lfloor b/C_{o_{|T|}} \rfloor - 1\right)$
end for
for $i = |T| - 1, \ldots, 2, 1$ **do**
  for $b = 0, 1, \ldots, B$ **do**
    $PD(\boldsymbol{o}, a, i, b) \leftarrow R_{a,o_i} \cdot F^*_{o_i}(\lfloor b/C_{o_i} \rfloor - 1) + (1 - R_{a,o_i}) \sum_{j=0}^{\lfloor b/C_{o_i} \rfloor} \left[(F_{o_i}(j) - F_{o_i}(j-1)) \cdot PD(\boldsymbol{o}, a, b - j \cdot C_{o_i}, i+1) \right]$
  end for
end for
Return $PD(\boldsymbol{o}, a) := PD(\boldsymbol{o}, a, 1, B)$

polynomial time

Theorem: Finding an **optimal prioritization strategy** is an **NP-hard** problem.

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Column Generation Algorithm

* Optimal strategy: linear programming based formulation
  * for each attack $a \in A$:

$$\max_{\boldsymbol{p}} \sum_{\boldsymbol{o} \in O} p_{\boldsymbol{o}} \cdot PD(\boldsymbol{o}, a)$$

subject to

$$\forall\, a' \in A: \quad \sum_{\boxed{\boldsymbol{o} \in O}} p_{\boldsymbol{o}} \cdot D(\boldsymbol{o}, a') \geq \Delta\left(K_{a'}\right)$$

exponential number of possible orderings

where

$$D(o, a') = [(1 - PD(\boldsymbol{o}, a))G_a - (1 - PD(\boldsymbol{o}, a'))G_{a'}]$$

$$\Delta\left(K_{a'}\right) = K_a - K_{a'}$$

* Polynomial-time column-generation approach

Algorithm 2 Greedy Column Generation

Input: prioritization game, reduced cost function $\bar{c}$

1: $\boldsymbol{o} \leftarrow \emptyset$
2: **while** $\exists t \in T \setminus \boldsymbol{o}$ **do**
3: $\quad \boldsymbol{o} \leftarrow \boldsymbol{o} + \mathrm{argmax}_{t \in T \setminus \boldsymbol{o}}\, \bar{c}(\boldsymbol{o} + t)$
4: **end while**
5: Return $\boldsymbol{o}$

---

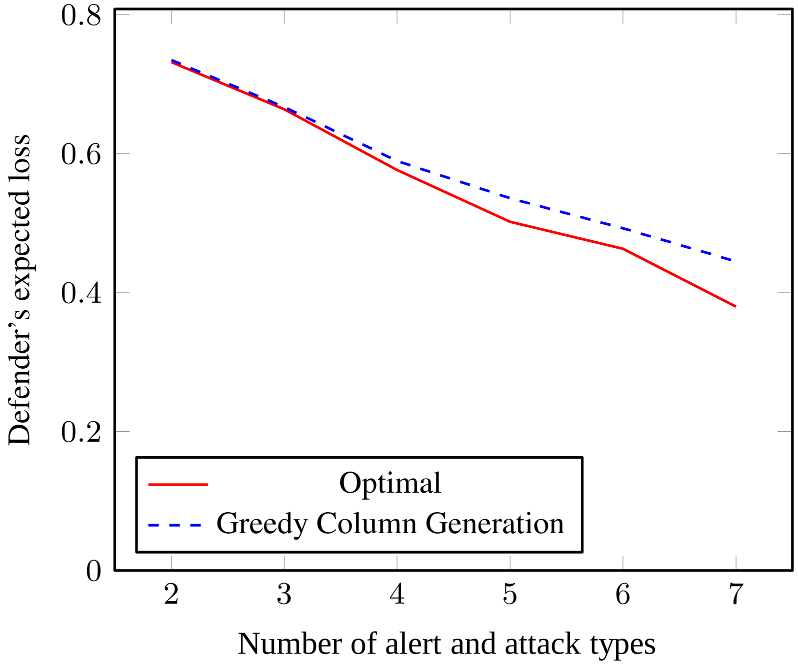
reduced cost function:

$$\bar{c}(\boldsymbol{o}) = PD(\boldsymbol{o}, a) + \sum_{a' \in A} y(\bar{O}, a') D(\boldsymbol{o}, a')$$

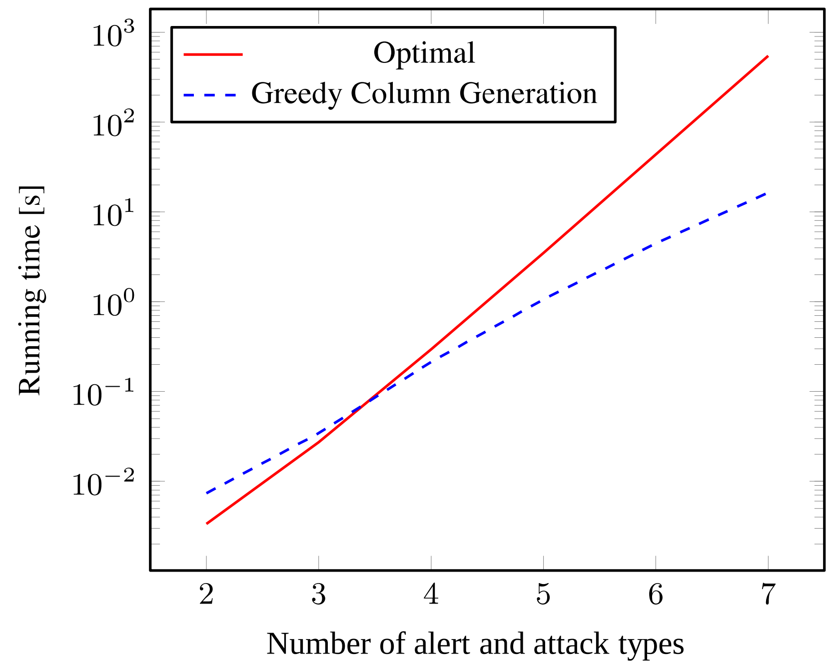
FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Numerical Results

Expected Loss



Running Time



$K_a = 0$, $C_t = 1$, D_a and G_a are drawn at random from [0.5, 1], $B = 5|T|$, each $R_{a,t}$ is either 0 (with probability 1/3) or drawn at random from [0, 1], every F_t is a Poisson whose mean is drawn at random from [5, 15]

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Conclusion

- Anomaly-based intrusion detection for cyber-physical systems
  $\longrightarrow$ prohibitively high number of false alerts
- Alert prioritization:
  deciding which alerts to investigate with a limited budget
  - naïve strategies are very vulnerable
- Game-theoretic formulation and analysis
  - finding optimal prioritization strategy is computationally hard
  - polynomial-time dynamic-programming algorithm for computing detection probabilities
  - polynomial-time column-generation approach for alert prioritization

Thank you for your attention!

Questions?