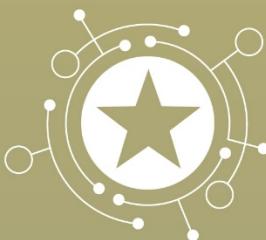


Internet of Battlefield Things

**COLLABORATIVE
RESEARCH
ALLIANCE**



IoBT
REIGN



Task 2.3: Fast and Adaptive Learning in Self-Aware IoBTs

- B. Jalaian - ARL
- S. Jha – SRI
- P. Tabuada – UCLA
- P. Thomas – U. Massachusetts
- V. Veeravalli – UIUC
- G. Verma – ARL
- S. You – ARL
- J. Smith - ARL

Team Members



Jalaian, ARL

Adversarial
Machine Learning,
Uncertainty
Quantification for
Machine Learning

Jha, SRI

Adversarial
machine
learning,
formal
methods

Thomas, UMass

Reinforcement
Learning

Tabuada, UCLA

Cyber-physical
systems, control,
formal methods,
security

Veeravalli, UIUC

Statistical inference,
stochastic
optimization,
information theory

Verma, ARL

statistics,
wireless
networks,
machine
learning

You, ARL

Deep
learning,
image and
signal
processing

Smith, ARL

Mathematical
Modeling

Task Goal

- **Goal:**

- Enable improved **reliability, survivability** and intelligent **interoperability** of IoBTs through a **“principled fast and safe change detection and adaptation”**.
- Supporting “command-by-intent”, “situational awareness”, and “timely, decisive action”.
- Relevant to Network C3I for Expeditionary Operations in line with Army’s Modernization Priorities



Notional Example:

Contested urban environment with multimodal sensing (cameras, road sensors, mobile sensing drones) with network communication to obtain situational awareness for achieving mission goal.

Technical Approach

- **Goal:** Develop safe autonomic reflexes in network that can quickly detect and adapt to uncertainty and adversarial perturbations
- **IoBT Challenges:**
 - Scale and heterogeneity of IoBT
 - Rapidly evolving fast-tempo environment
 - Need to operate in uncertainty and risk-sensitive decision-making
 - Adversarial cyber and physical perturbations in operationally-contested environment
 - Need to ensure safety of adaptation decisions
- **Potential:** Fast and high-assurance adversarially-resilient risk-sensitive learning paradigm.

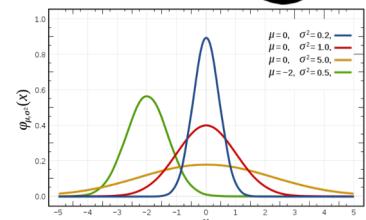
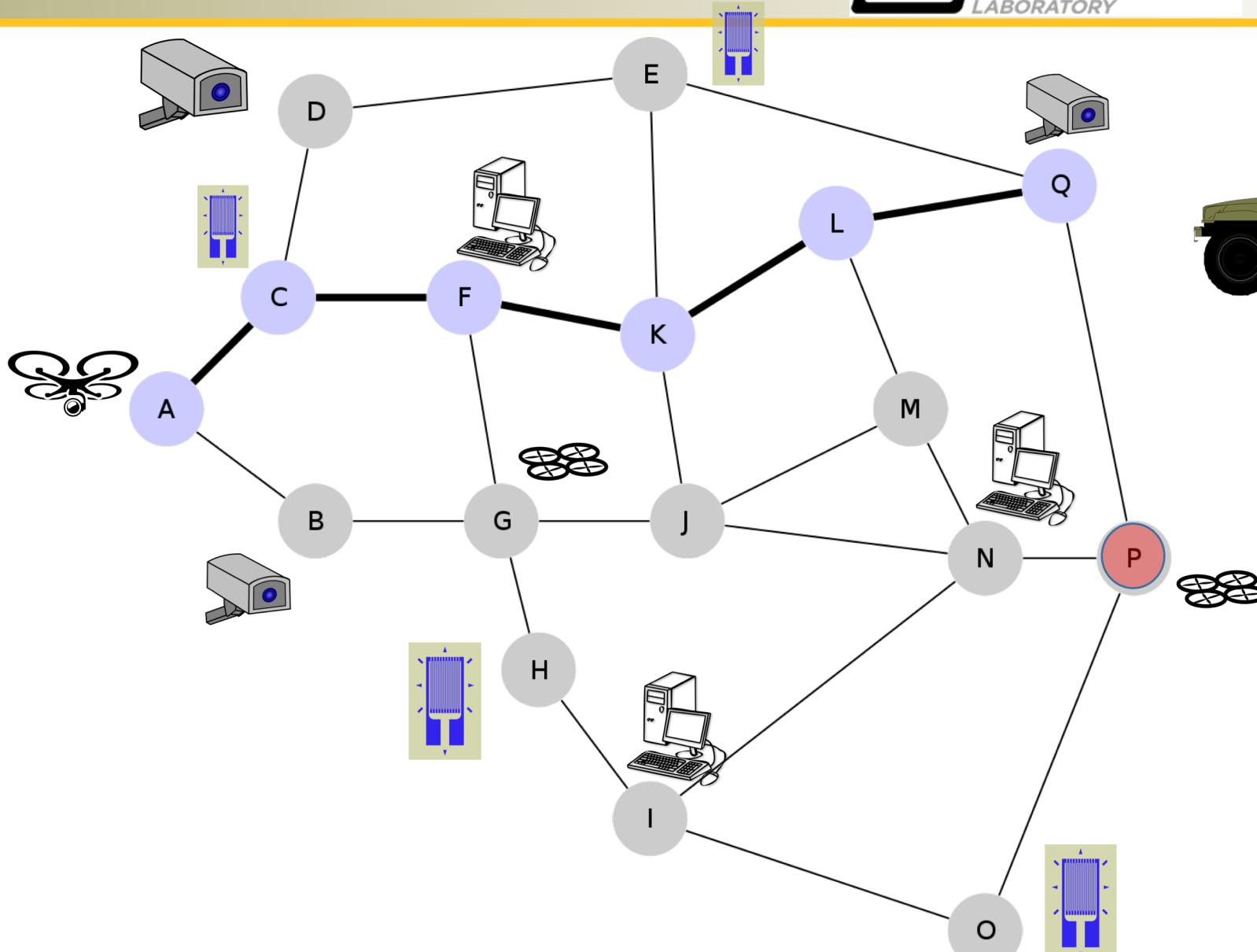


Notional Example:
Contested urban environment with multimodal sensing (cameras, road sensors, mobile sensing drones) with network communication to obtain situational awareness for achieving mission goal.

Task 2.3: Fast and Adaptive Learning



IoBT REIGN



Model-based change detection

Concept Drift Model Change

Task 2.3: Fast and Adaptive Learning



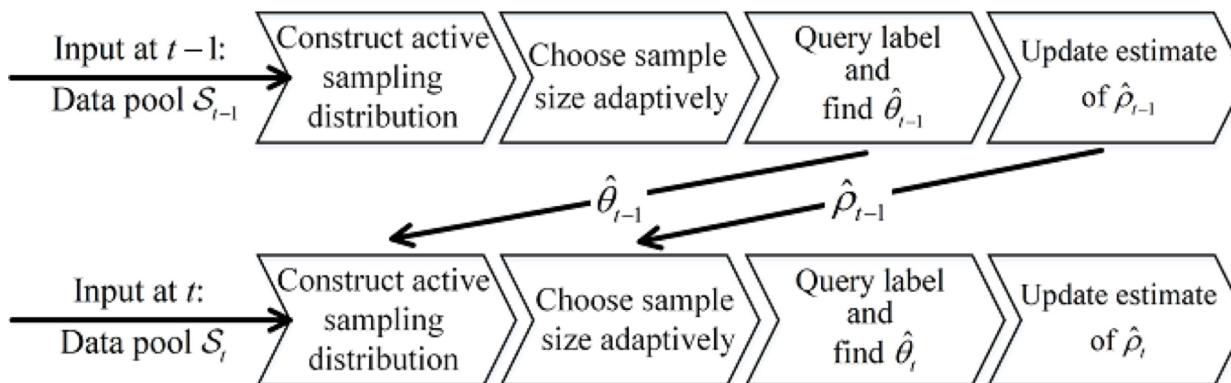
Addressed Army Need: Supporting machine learning functions in rapidly changing battlefield environments

- Active and adaptive sequential learning strategies for learning tasks that change in a bounded manner:

$$\ell(y|x, \theta) \triangleq -\log p(y|x, \theta)$$

$$L_{U_t}(\theta) \triangleq \mathbb{E}_{X \sim U_t, Y \sim p(Y|X, \theta^*)} [\ell(Y|X, \theta)]$$

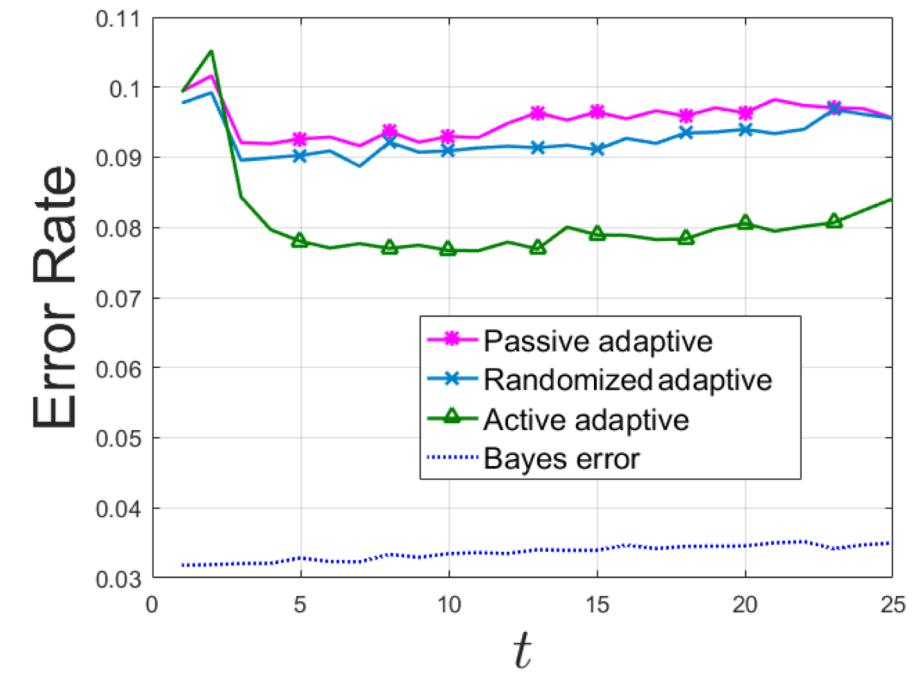
$$\|\theta_t^* - \theta_{t-1}^*\|_2 \leq \rho$$



Passive means drawing new samples using a uniform distribution

Random means replacing the estimate of θ_{t-1} with a random point from Θ

Active and adaptive learning framework can significantly improve accuracy while being efficient in the use of training samples



Tracking user preferences (Yelp 2017 dataset)

Task 2.3: Fast and Adaptive Learning



IoBT REIGN

Addressed Army Need: Supporting machine learning functions in rapidly changing battlefield environments

- Model change detection strategies for detecting abrupt changes in tasks:
 - Generalized likelihood ratio test, test based on [Jensen-Shannon divergence](#), and test based on maximum mean discrepancy (MMD)

$$\ell(y|x, \theta) \triangleq -\log p(y|x, \theta)$$

$$L_{U_t}(\theta) \triangleq \mathbb{E}_{X \sim U_t, Y \sim p(Y|X, \theta^*)} [\ell(Y|X, \theta)] \quad \|\theta_t^* - \theta_{t-1}^*\|_2 \leq \rho$$

$$H_0 : (\theta, \theta') \in \chi_0 \triangleq \{(\theta, \theta') | \|\theta - \theta'\|_2 \leq \rho\},$$

$$H_1 : (\theta, \theta') \in \chi_1 \triangleq \{(\theta, \theta') | \|\theta - \theta'\|_2 > \rho\},$$

$$2JS(P, P') \triangleq D(P||\bar{P}) + D(P'||\bar{P}),$$

where $\bar{P} = \frac{P+P'}{2}$ and $D(\cdot||\cdot)$ denotes the KL divergence.

$$\delta_{ED} = \begin{cases} 1, & \text{if } \hat{JS}(\mathcal{S}, \mathcal{S}') \geq \eta \\ 0, & \text{if } \hat{JS}(\mathcal{S}, \mathcal{S}') < \eta \end{cases}$$

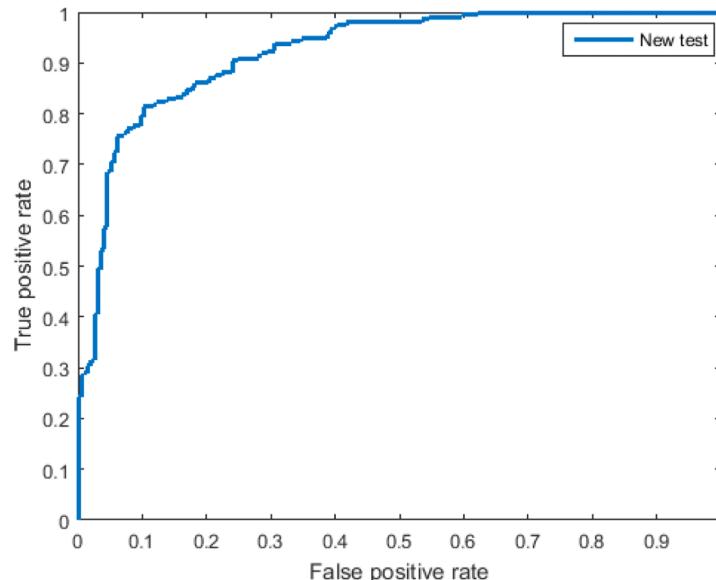
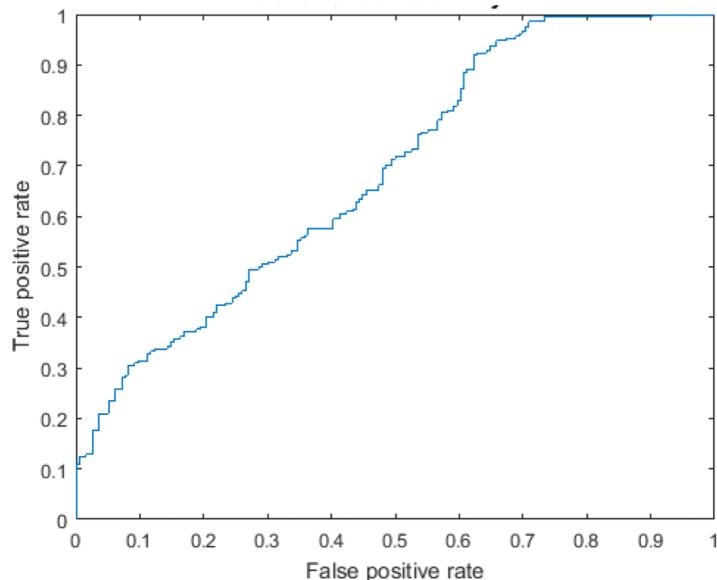
$$\begin{aligned} \hat{JS}(\mathcal{S}, \mathcal{S}') &= \sum_{i=1}^n \log \frac{2f_{\hat{w}_{ERM}}(x_i, y_i)}{f_{\hat{w}_{ERM}}(x_i, y_i) + f_{\hat{w}'_{ERM}}(x_i, y_i)} \\ &\quad + \sum_{i=1}^n \log \frac{2f_{\hat{w}'_{ERM}}(x'_i, y'_i)}{f_{\hat{w}_{ERM}}(x'_i, y'_i) + f_{\hat{w}'_{ERM}}(x'_i, y'_i)} \end{aligned}$$

$$\begin{aligned} \hat{w}_{ERM} &\triangleq \arg \min L(w), & \hat{w}'_{ERM} &\triangleq \arg \min L'(w) \\ L(w) &\triangleq -\sum_{i=1}^n \log f_w(X_i, Y_i), & L'(w) &\triangleq -\sum_{i=1}^{n'} \log f_w(X'_i, Y'_i). \end{aligned}$$

Addressed Army Need: Supporting machine learning functions in rapidly changing battlefield environments

Experiments and Results:

- Model change detection: **Landmines dataset**
 - The goal is to detect landmines in specific regions. Overall 29 binary classification tasks
 - Each datum is a 9-dimensional feature vector from radar images that capture a single region of landmine fields
 - Tasks 1-15 correspond to regions that are relatively highly foliated. Other 14 tasks are bare earth or desert
 - ${}^{29}C_2 = 406$ number of pairs to test whether the distribution is same.

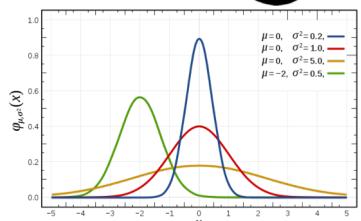
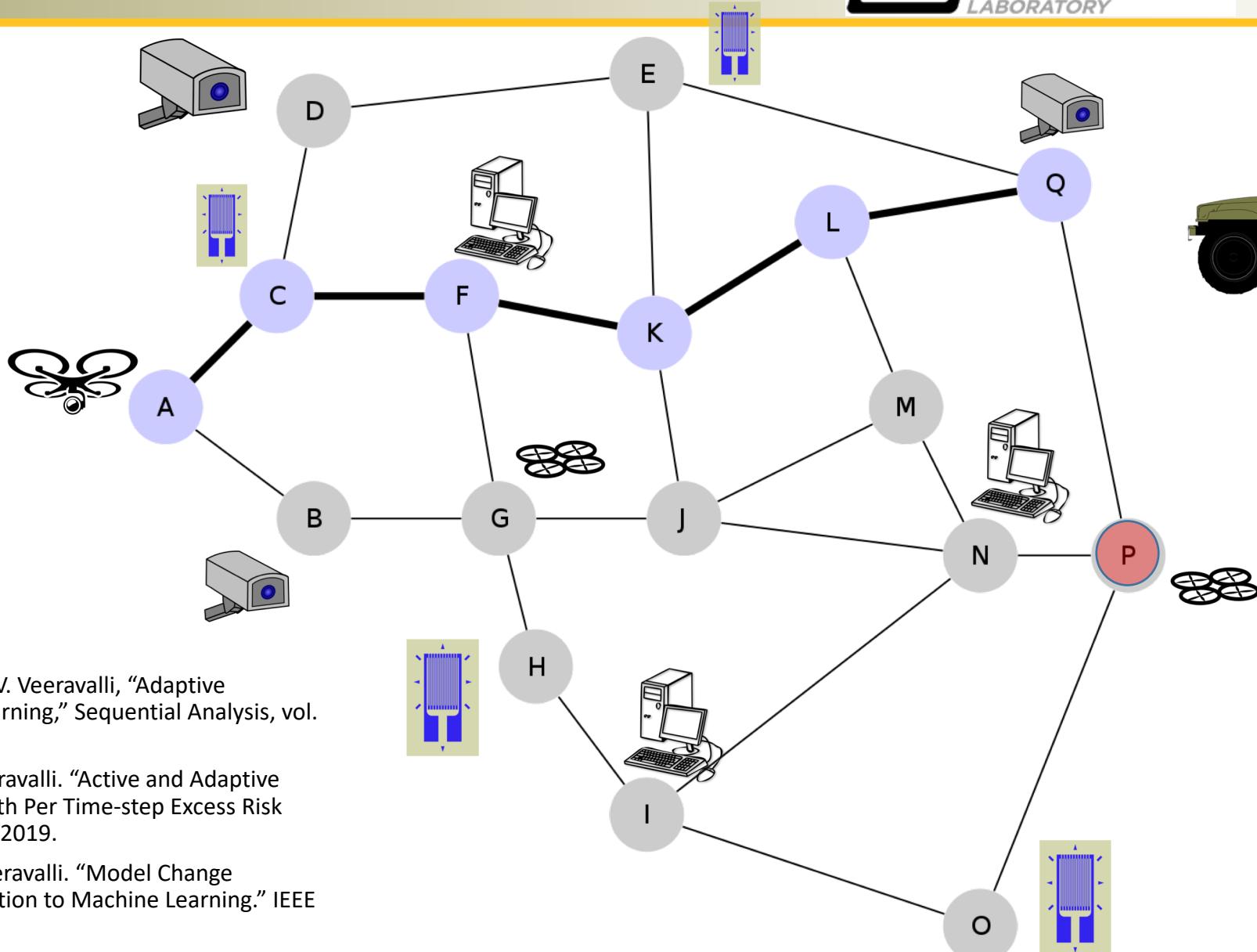


Tests based on Jensen-Shannon divergence and MMD effective in detecting model changes

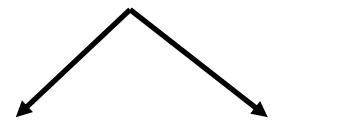
Task 2.3: Fast and Adaptive Learning



IoBT REIGN



Model-based change detection



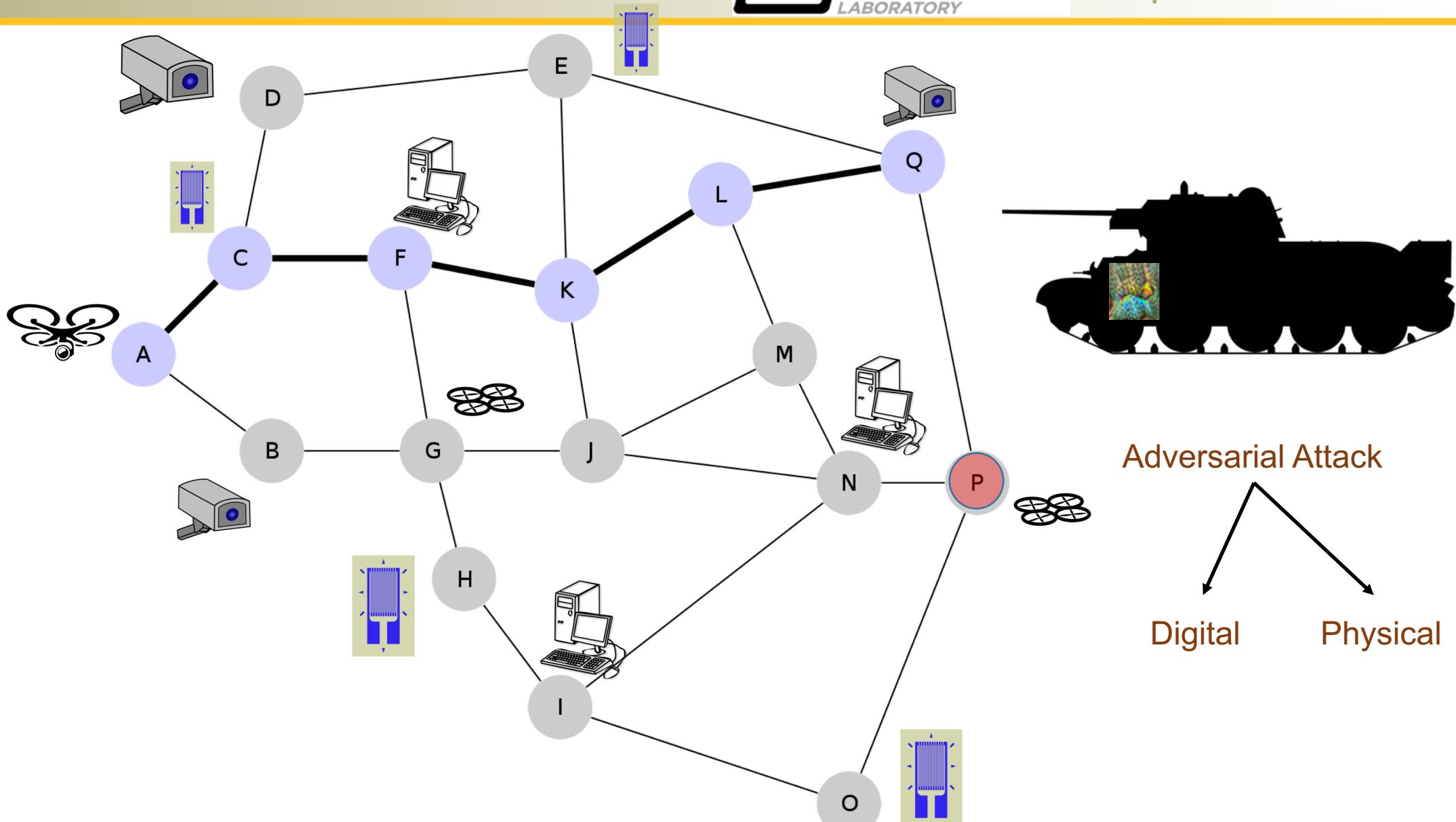
Key Publications:

- C. Wilson, Y. Bu, and V.V. Veeravalli, "Adaptive sequential machine learning," *Sequential Analysis*, vol. (to appear), 2020
- Y. Bu, J. Lu and V.V. Veeravalli. "Active and Adaptive Sequential Learning with Per Time-step Excess Risk Guarantees," *Asilomar* 2019.
- Y. Bu, J. Lu, and V.V. Veeravalli. "Model Change Detection with Application to Machine Learning." *IEEE ICASSP* 2019.

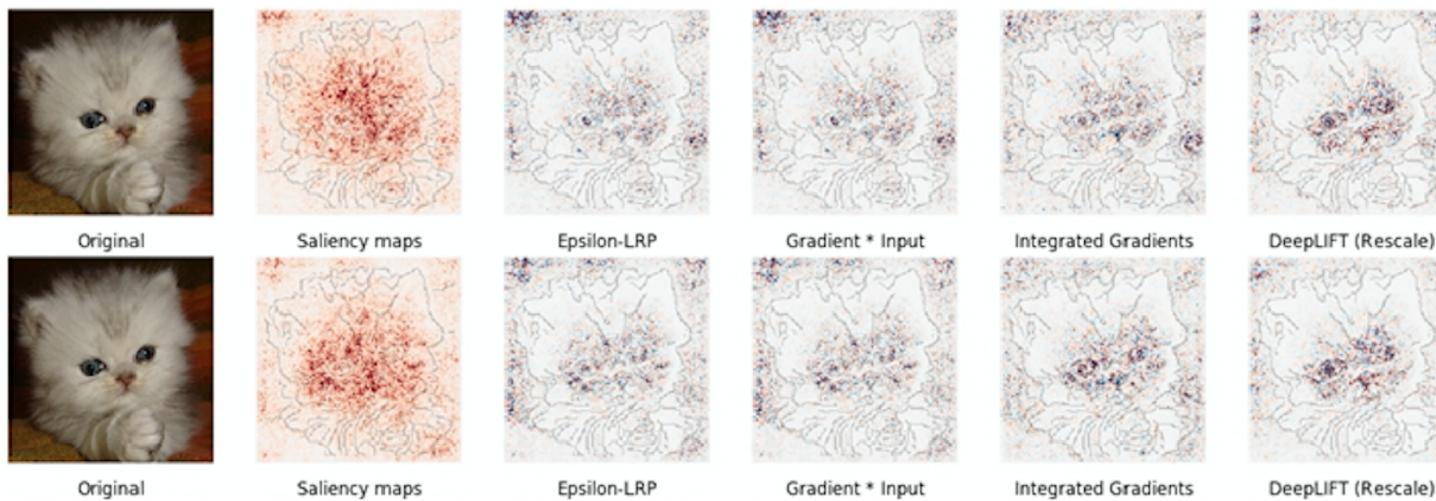
Task 2.3: Fast and Adaptive Learning



IoBT REIGN



Task 2.3: Fast and Adaptive Learning



Young (1985) demonstrated that Shapley values are the only set of values that satisfy these properties.

$$a_i = \sum_{z \subseteq x} \frac{|z|! (M - |z| - 1)!}{M!} [f_x(z) - f_x(z' \setminus \{x^i\})]$$

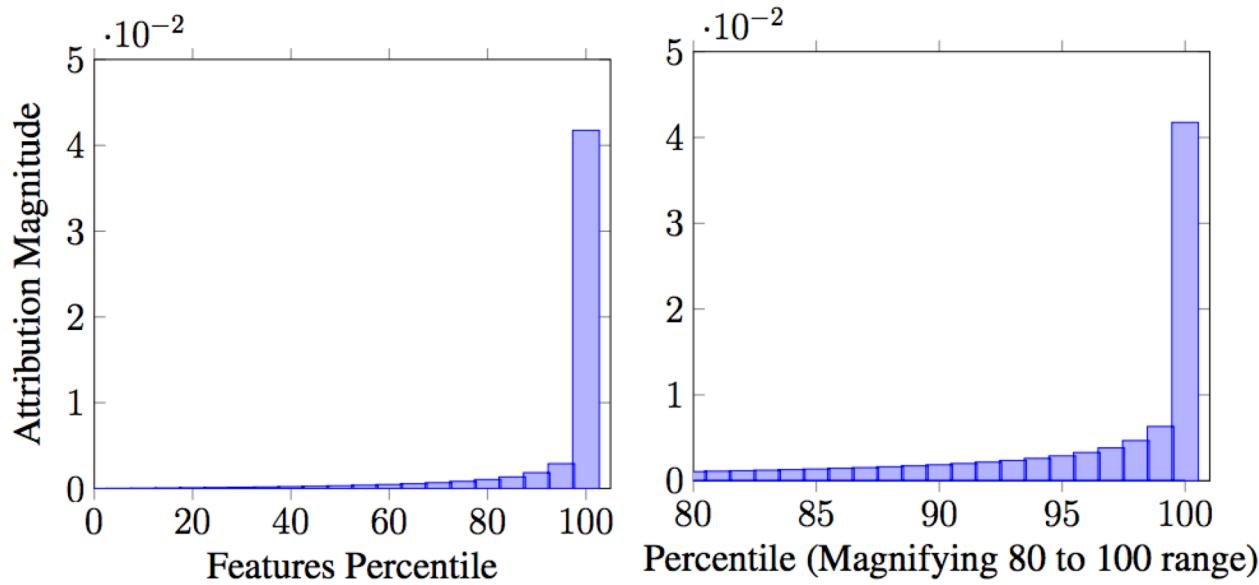
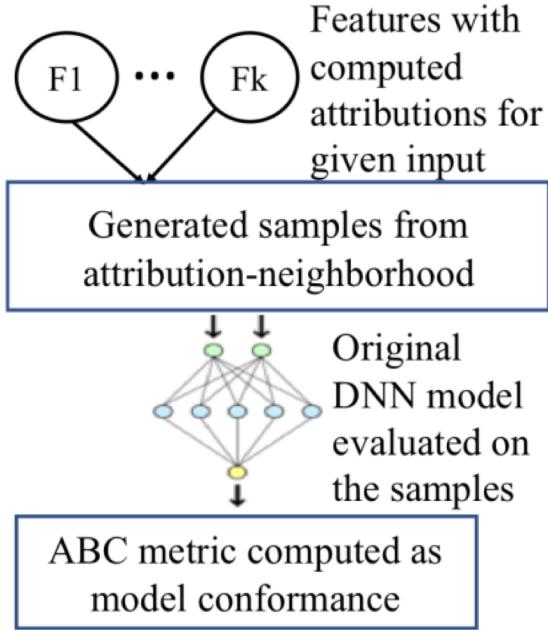
Apply sampling approximations to above equation and approximate the effect of removing a variable from the model by integrating over samples.

Baseline and path based methods.

Friedman, Eric J. Paths and consistency in additive cost sharing. *International Journal of Game Theory*, 32(4): 501–518, 2004.

Given $\gamma = (\gamma_1, \dots, \gamma_n): [0,1] \rightarrow R^n$ be a smooth function specifying a path in R^n from baseline x^b to input x , that is, $\gamma(0) = x^b, \gamma(1) = x$.

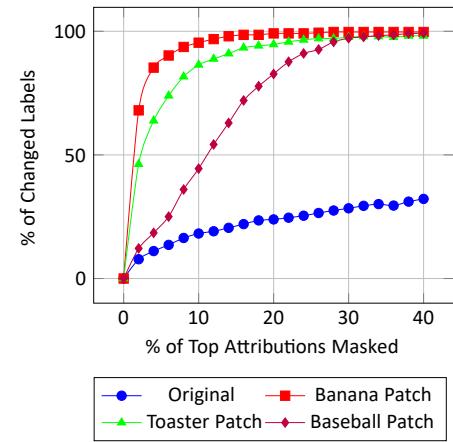
Task 2.3: Fast and Adaptive Learning



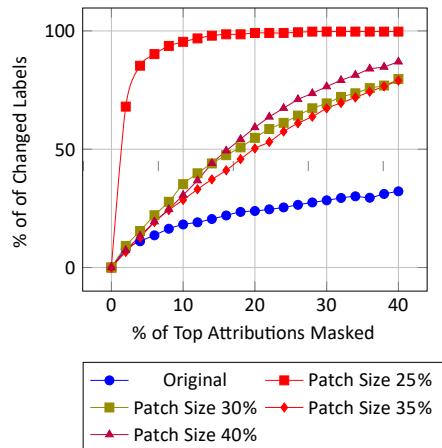
Feature Concentration in well-trained models

Theorem 1. *The sensitivity of the output $\mathcal{F}(\mathbf{x})$ with respect to an input feature \mathbf{x}_j in the neighborhood of \mathbf{x} is approximately the ratio of the attribution $\mathcal{A}_j(\mathbf{x})$ to the value of that feature \mathbf{x}_j , that is, $\frac{\mathcal{A}_j(\mathbf{x})}{\mathbf{x}_j}$.*

Task 2.3: Fast and Adaptive Learning



Dropping 0.4% of the attribution causes 99.71% of the attacks based on banana patches, 98.14% of the attacks based on toaster patches, and 99.20% of the attacks based on baseball patches to be detected.



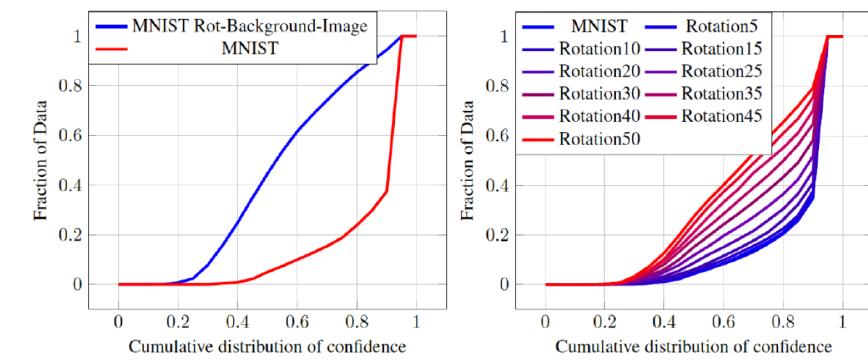
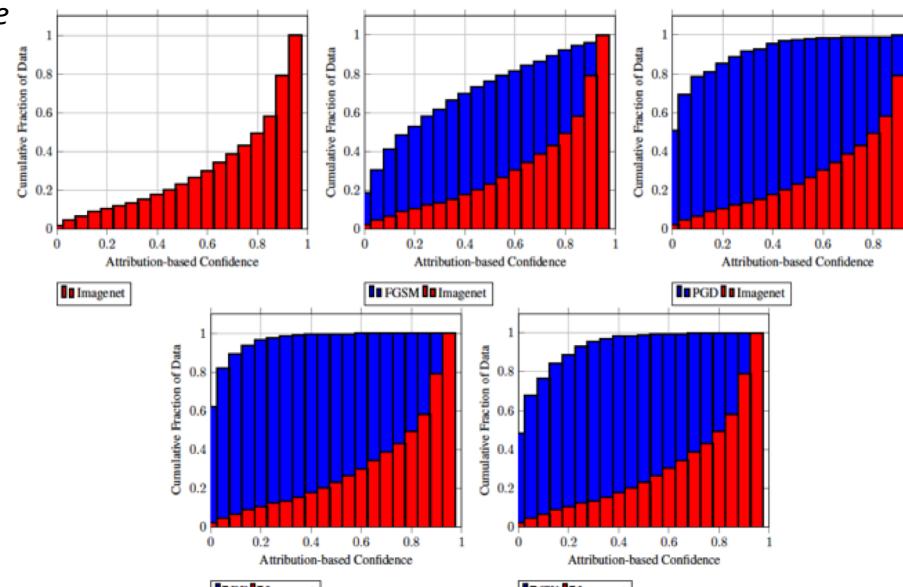
Masking 0.4% of attributions caused nearly 80% of labels to change for images with adversarial patches.



Image with a banana patch generated using adversarial patch method

Masking its top 0.2% of attribution

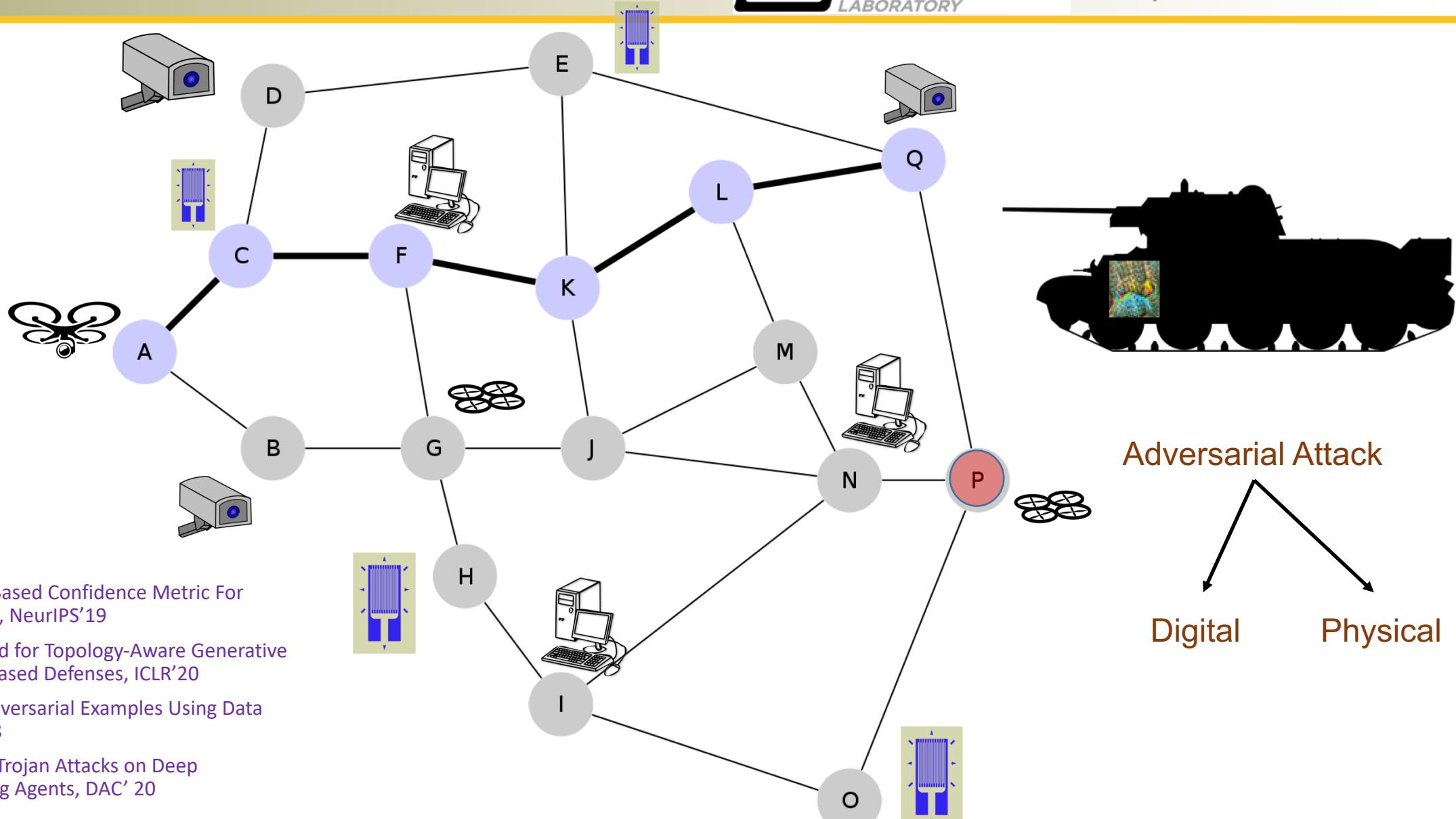
Masking its top 0.4% of attribution



Task 2.3: Fast and Adaptive Learning



IoBT REIGN



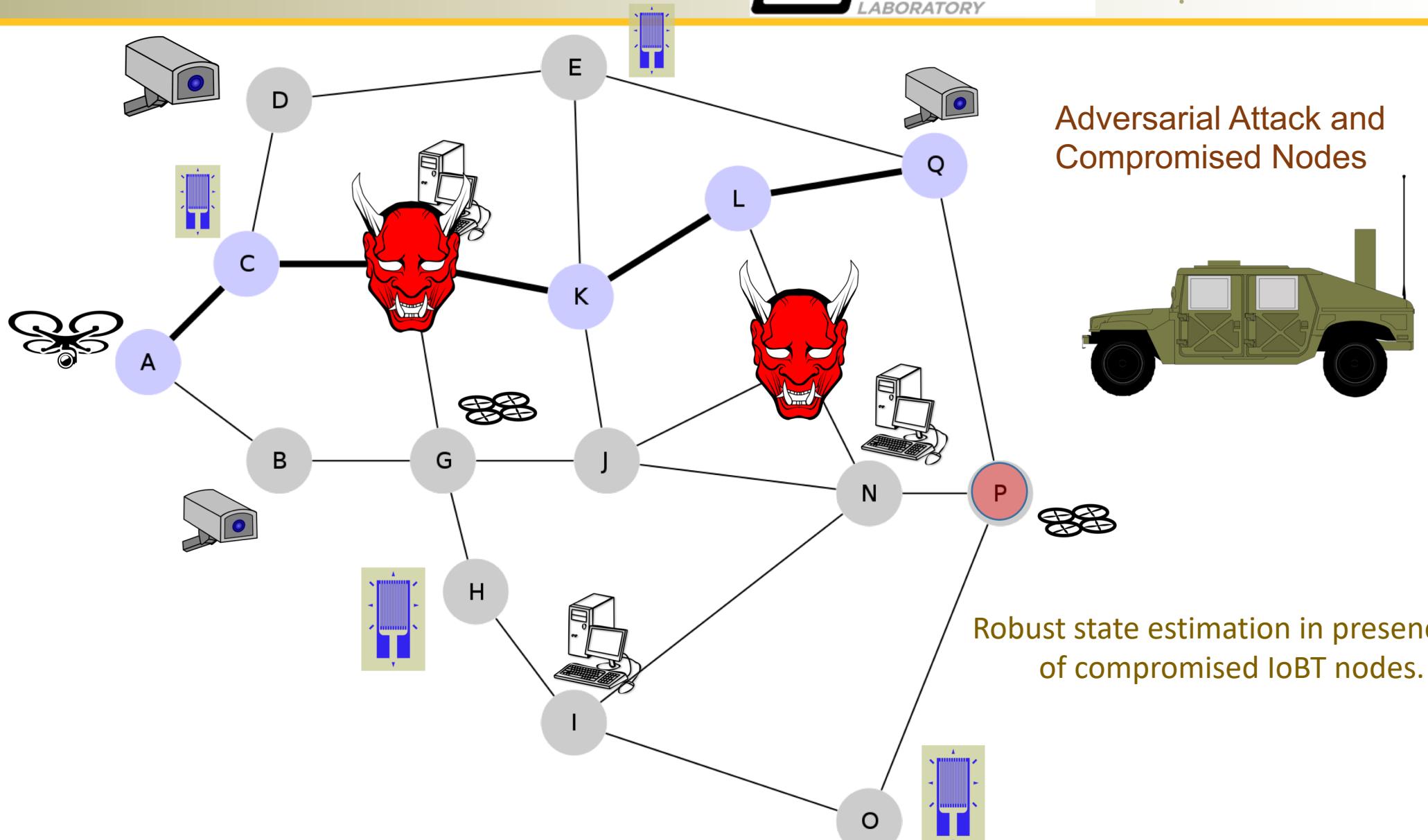
Key Publications:

- Jha et. al. Attribution-Based Confidence Metric For Deep Neural Networks, NeurIPS'19
- Jang et. al. On the Need for Topology-Aware Generative Models for Manifold-Based Defenses, ICLR'20
- Jha et. al. Detecting Adversarial Examples Using Data Manifolds, MILCOM'18
- Kiourtis et. al. TrojDRL: Trojan Attacks on Deep Reinforcement Learning Agents, DAC' 20

Task 2.3: Fast and Adaptive Learning



IoBT REIGN



Task 2.3: Fast and Adaptive Learning



- Model the dynamics of sensed quantities as a linear dynamical system: $x(k+1)=Ax(k)$, with measurement equation: $y(k)=Cx(k)+e(k)$, subject to the adversarially injected signal e .
- We assume the adversary can attack at most s sensors (at most s entries of e are non-zero)
- **Objective:** estimate the state x despite the presence of the injected signal e .

Given a sequence of measurements $y(0), y(1), \dots, y(k)$ we need to compute a subset of (malicious) sensors so that the data provided by the non-malicious sensors obeys the linear dynamics.

If there are p sensors and at most s are under attack, we need to search over all possible “ p choose s ” subsets of sensors.

Task 2.3: Fast and Adaptive Learning



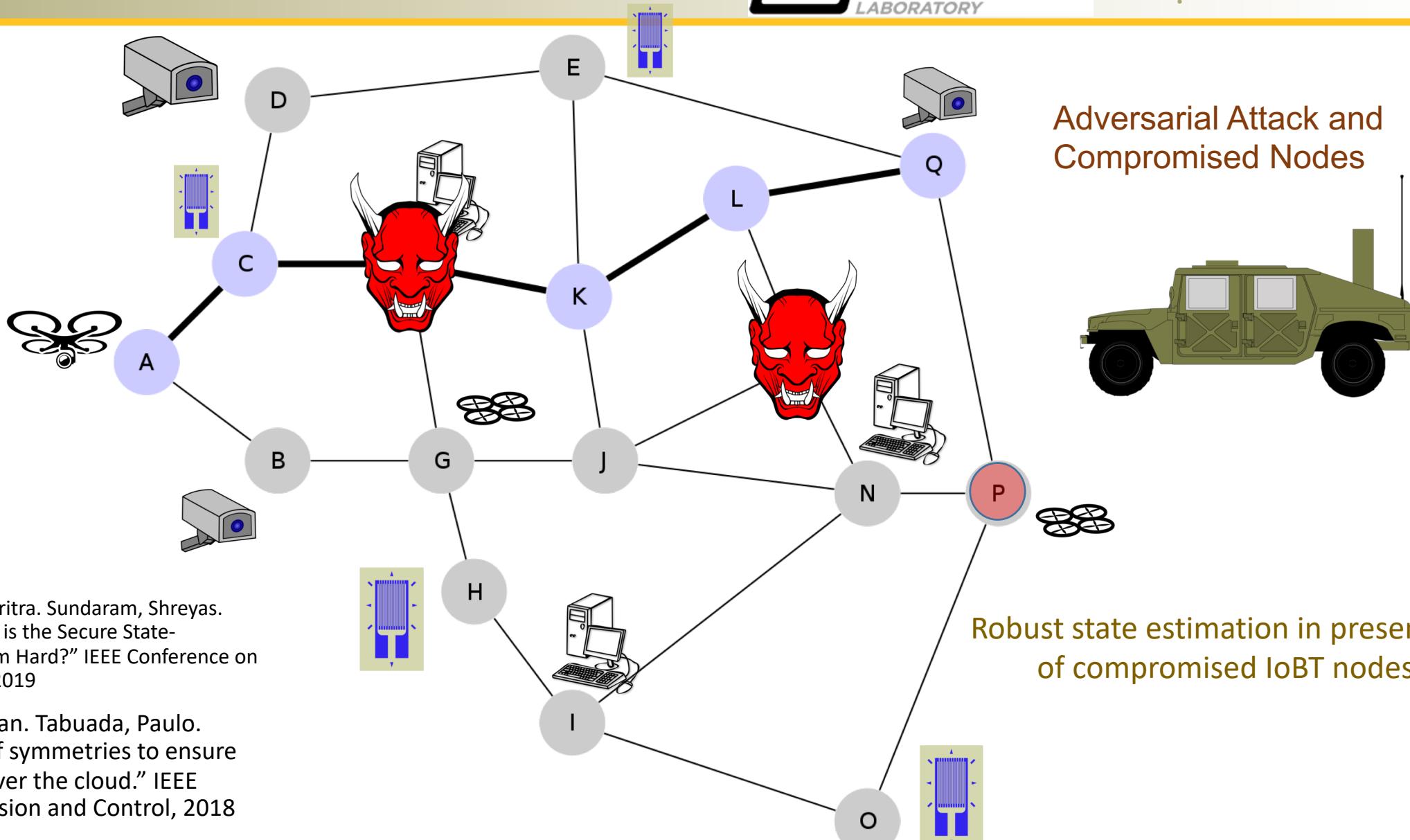
Given a sequence of measurements $y(0), y(1), \dots, y(k)$ we need to compute a subset of (malicious) sensors so that the data provided by the non-malicious sensors obeys the linear dynamics:

$$x(t+1) = Ax(t), y(t) = Cx(t)$$

If there are p sensors and at most s are under attack, we need to search over all possible “ p choose s ” subsets of sensors.

- We showed this problem is, in general, NP-Hard.
- We identified a large class of problems that can be solved in polynomial time: the eigenvalues of the matrix A have geometric multiplicity 1.
- Investigated a more challenging version of this problem where sensor data is routed through a network with attacked nodes and links. Identified necessary and sufficient conditions to solve this problem, relating (A,C) to the number of sensors and network nodes/links under attack.

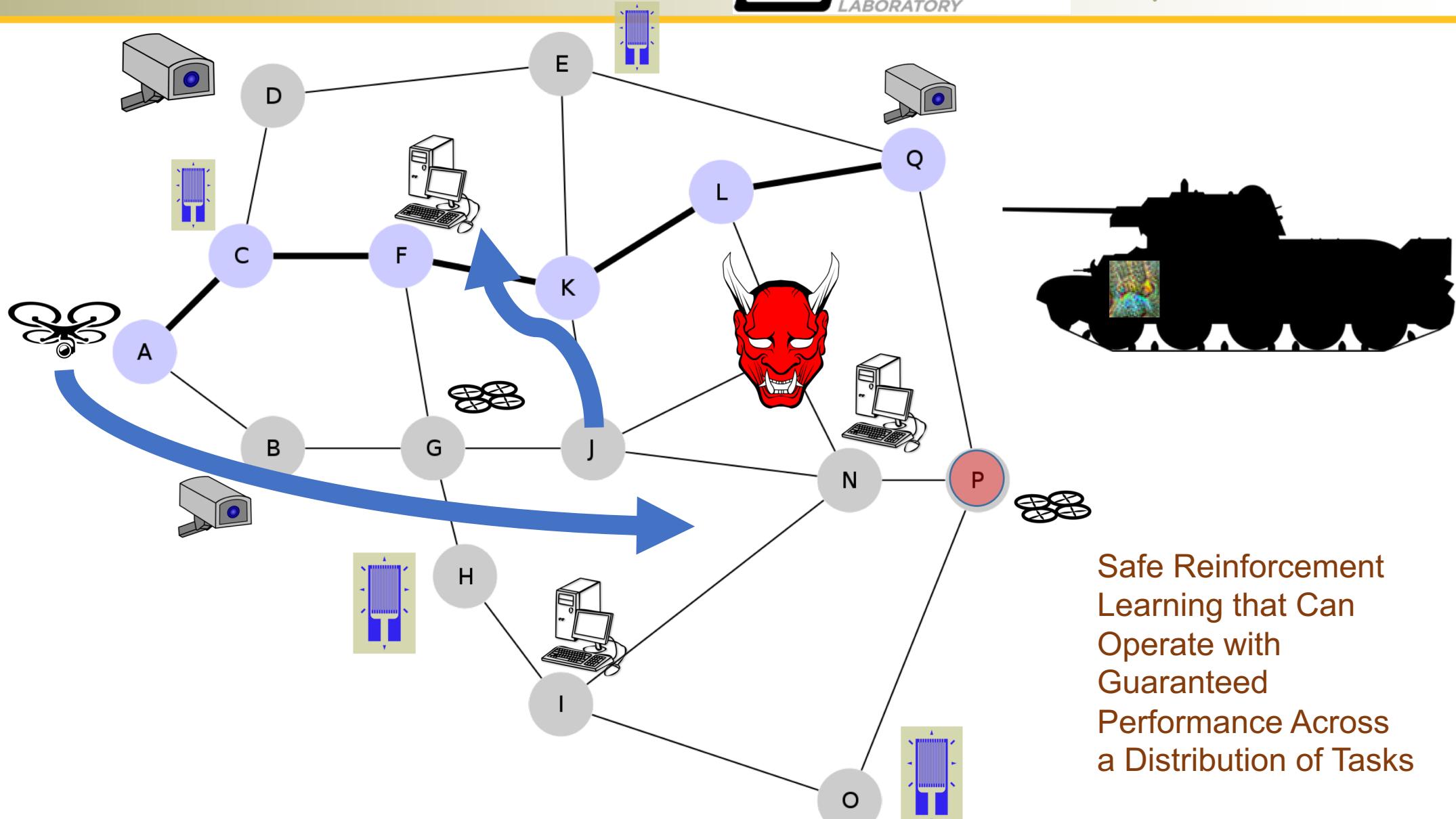
Task 2.3: Fast and Adaptive Learning



Key Publications:

- Mao, Yanwen. Mitra, Aritra. Sundaram, Shreyas. Tabuada, Paulo "When is the Secure State-Reconstruction Problem Hard?" IEEE Conference on Decision and Control, 2019
- Sultangazin, Alimzhan. Tabuada, Paulo. "Towards the use of symmetries to ensure privacy in control over the cloud." IEEE Conference on Decision and Control, 2018

Task 2.3: Fast and Adaptive Learning

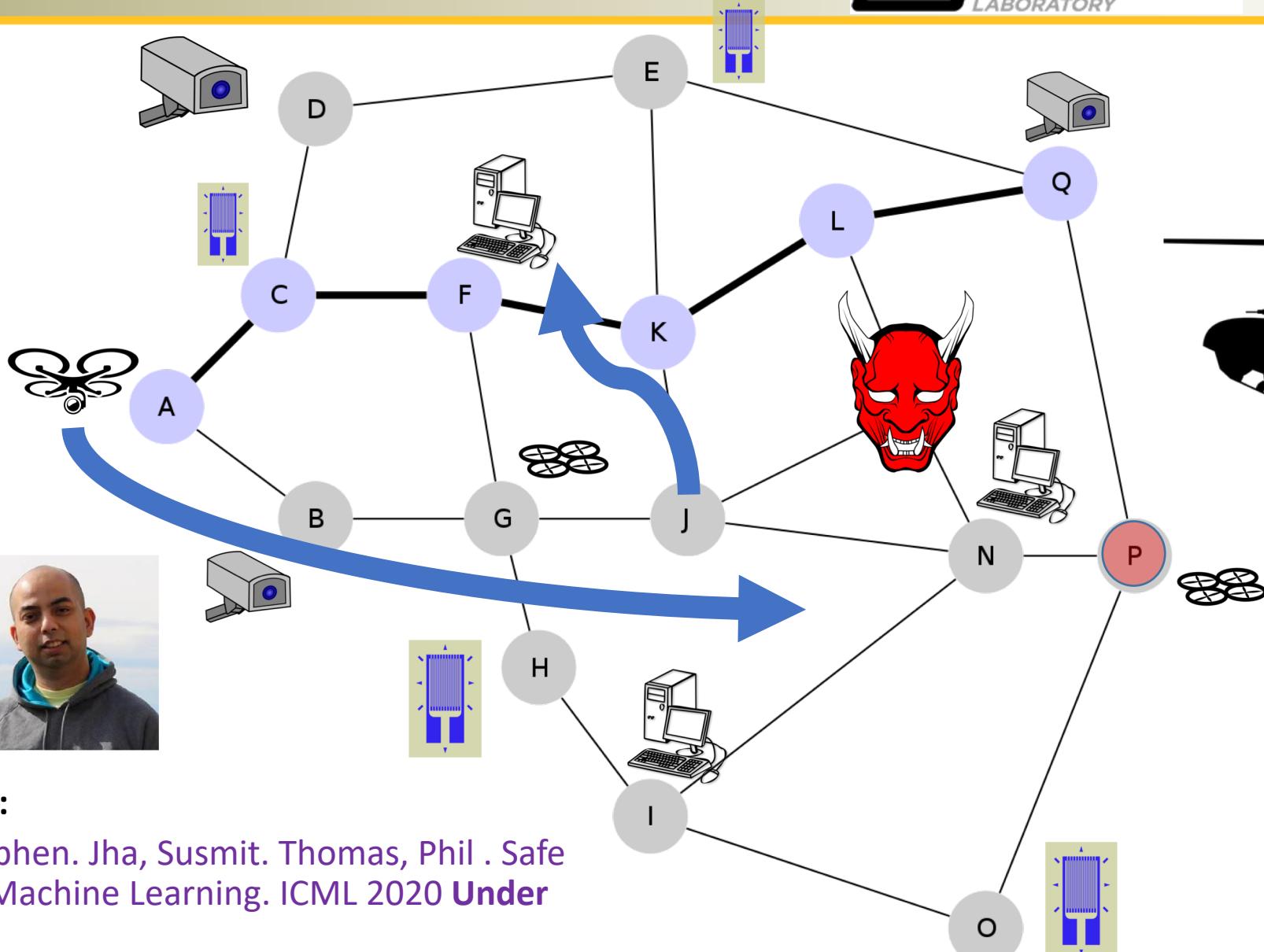


Task 2.3: Fast and Adaptive Learning



- Reinforcement Learning (RL) Algorithms that Can Operate with Guaranteed Performance Across a Distribution of Tasks.
 - E.g. Resource Allocation for Tracking multiple vehicles.
- The user defines a minimum safe measure of performance for the RL task and specifies the probability with which the algorithm must achieve this measure of performance.
- **Train on a distribution of tasks, and returns:**
 - a solution that is guaranteed to be safe across all tasks from the same distribution with the specified probability, or
 - if a safe solution cannot be found with the specified probability, the algorithm returns “NO SOLUTION FOUND”.
- **Robust to new distribution (bounded from training):** By modifying the types of bounds we use, we can guarantee performance even when the target task does not come from the training distribution (assuming other assumptions are met, and that sufficient data exists).

Task 2.3: Fast and Adaptive Learning

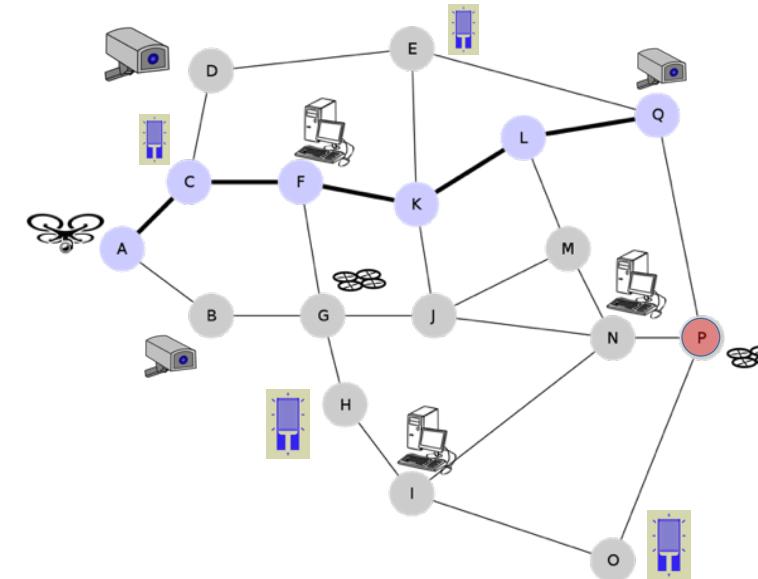
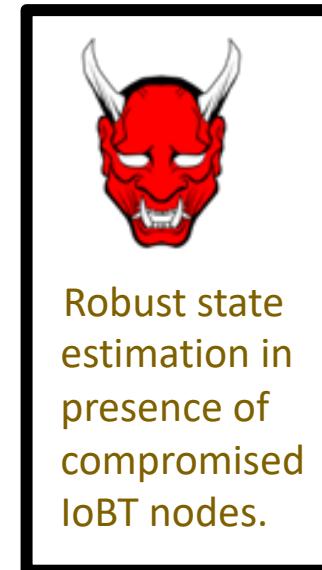
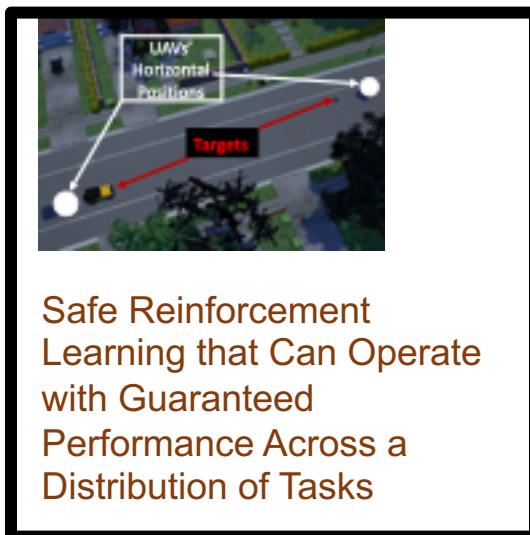
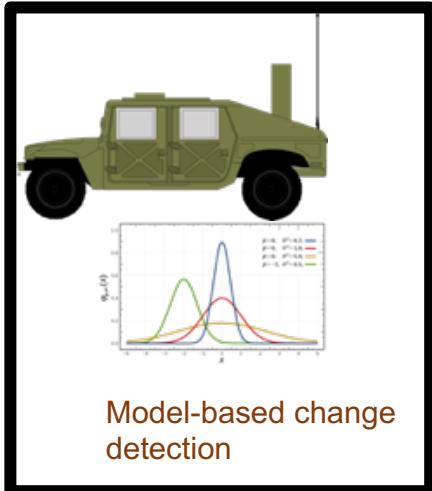


Safe Reinforcement
Learning that Can
Operate with
Guaranteed
Performance Across
a Distribution of Tasks

Key Publications:

- Giguere, Stephen. Jha, Susmit. Thomas, Phil . Safe and Robust Machine Learning. ICML 2020 **Under Submission**

Military Impact



- Enable Assured Adaptive IoBT network in a contested and congested environment for Network C3I.
- Enable Rapid Change Detection in U.S. Army C3I Systems and Networks
- Enable Adversarial Attack Detection for AI/ML in U.S. Army Network C3I System and Network
- Enable Command by intent through autonomic reflex capability.

Academic Impact

- **Publications:**

- **30 peer-reviewed publications** including prestigious venues such as **AAAI, NeurIPS, ICLR, ICRA, ACC, HSCC, DAC and MILCOM** : 27 Conference/Symposium papers + 3 Journal papers
- Each thrust has multiple PIs working collaboratively

- Services, Awards and Books

Veeravalli, UIUC

- Book: P. Moulin and V.V. Veeravalli. Statistical Inference for Engineers and Data Scientists. Cambridge University Press, 2019.
- Book: V.V. Veeravalli and A. ElGamal. Interference Management in Wireless Networks: Fundamental Bounds and the Role of Cooperation. Cambridge University Press, 2018
- Technical Program Committee Co-Chair, IEEE International Symposium on Information Theory, Paris, France, 2019.
- Area Editor for IEEE Open Journal of Signal Processing, 2019 - present

Tabuada, UCLA

- Guest editor for the special issue of Acta Informatica on Synthesis, 2020.
- Steering committee member for CPS-IoTWeek
- Chair of steering committee for the International Conference on Hybrid Systems: Computation and Control
- General co-chair for the 10th ACM/IEEE International Conference on Cyber-Physical Systems, 2019
- Program committee co-chair for the 9th ACM/IEEE International Conference on Cyber-Physical Systems, 2019.
- Chair of the IEEE Transactions on Control of Network Systems Outstanding Paper Award Committee.
- Vice-chair of the IFAC technical committee on networked systems.
- Visits/rotations to/from or with ARL and duration: 3 day visit to ARL in Adelphi, September 2019.

Jha, SRI

- 10 year Most-influential Paper Award at 42nd International Conference on Software Engineering (ICSE)
- Program co-chair for 12th NASA Symposium on Formal Methods
- Best Demo Award at HSCC 2019 for Neural Network Verification Tool Sherlock
- Visits/rotations: Visit from Brian Jalain, ARL for one week at SRI, Menlo Park in February

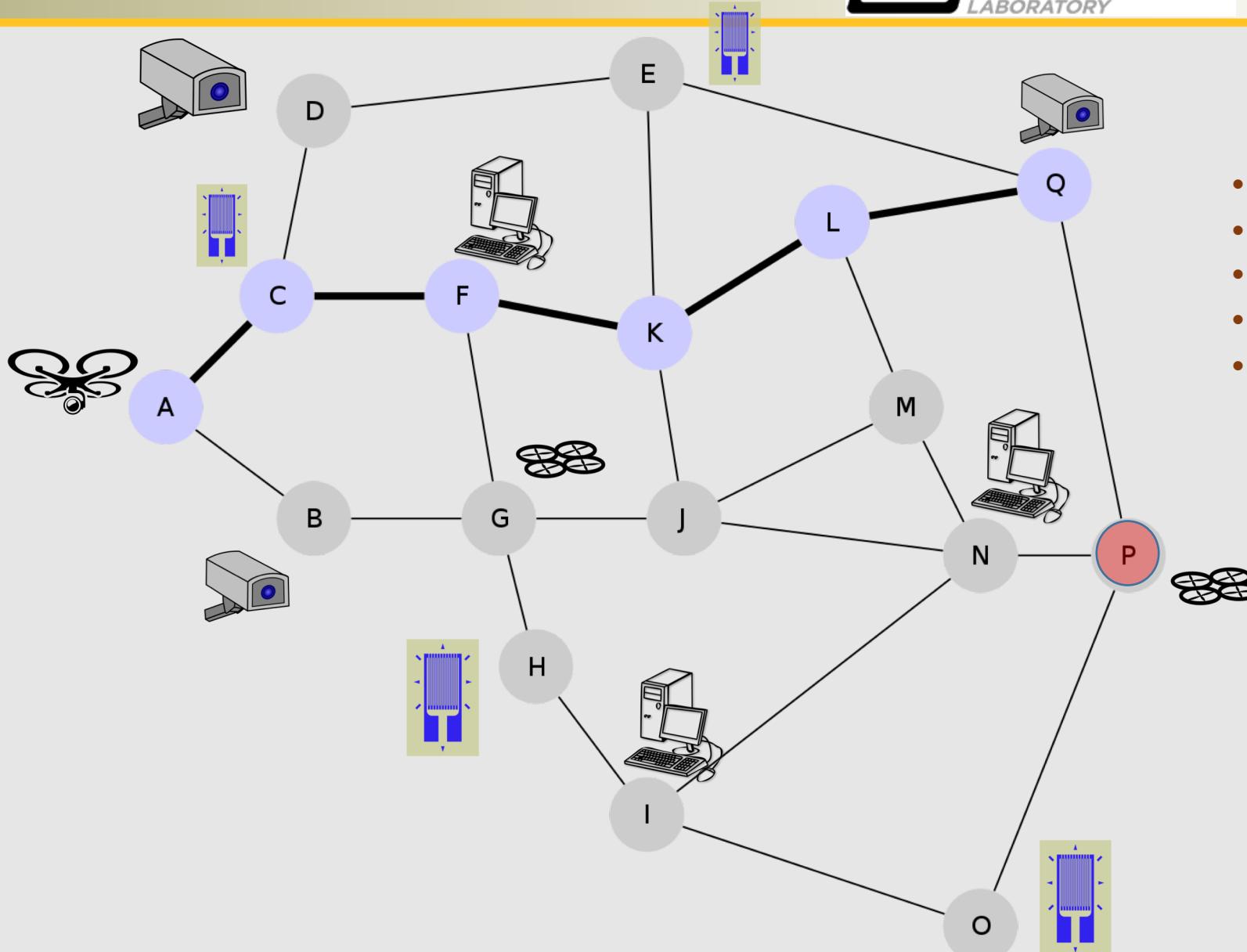
Task 2.3: Fast and Adaptive Learning



Thanks!

Back-up Slides

Task 2.3: Fast and Adaptive Learning



- multimodality,
- heterogeneous nodes,
- rapid dynamics,
- presence of adversaries,
- contested predominantly wireless environment

Task 2.3: Fast and Adaptive Learning



IoBT REIGN

Addressed Army Need: Supporting machine learning functions in rapidly changing battlefield environments

- Model change detection strategies for detecting abrupt changes in tasks:
 - Generalized likelihood ratio test, test based on Jensen-Shannon divergence, and test based on maximum mean discrepancy (MMD)

$$\ell(y|x, \theta) \triangleq -\log p(y|x, \theta)$$

$$L_{U_t}(\theta) \triangleq \mathbb{E}_{X \sim U_t, Y \sim p(Y|X, \theta_t^*)} [\ell(Y|X, \theta)] \quad \|\theta_t^* - \theta_{t-1}^*\|_2 \leq \rho$$

$$H_0 : (\theta, \theta') \in \chi_0 \triangleq \{(\theta, \theta') | \|\theta - \theta'\|_2 \leq \rho\},$$

$$H_1 : (\theta, \theta') \in \chi_1 \triangleq \{(\theta, \theta') | \|\theta - \theta'\|_2 > \rho\},$$

$$L_G(\mathcal{S}, \mathcal{S}') \triangleq \log \frac{\max_{(\theta, \theta') \in \chi_1} \prod_{i=1}^n p(\mathbf{z}_i | \theta) \prod_{i=1}^{n'} p(\mathbf{z}'_i | \theta)}{\max_{(\theta, \theta') \in \chi_0} \prod_{i=1}^n p(\mathbf{z}_i | \theta) \prod_{i=1}^{n'} p(\mathbf{z}'_i | \theta)}$$

$$\delta_{GL}(\mathcal{S}, \mathcal{S}') = \begin{cases} 1, & \text{if } L_G(\mathcal{S}, \mathcal{S}') \geq \tau \\ 0, & \text{if } L_G(\mathcal{S}, \mathcal{S}') < \tau \end{cases}$$

Additive feature attribution locally: Boolean features - present or absent

$$g(x) = a_0 + \sum_i^M a_i x^i$$

From cooperative game theory, we have classic equations to compute Shapley values

$$a_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S)]$$

- **Local accuracy:** $g(x) = f(x)$ [explanation model matches original model on the input]
- **Sensitivity:** $x^i = 0 \Rightarrow a_i = 0$
- **Consistency:** For any two models f and f' , $f'(x) - f'(\{x^i\}) \geq f(x) - f(\{x^i\})$ for all presence/absence of features x in , then $a_i^{f'} \geq a_i^f$

[Young \(1985\)](#) demonstrated that Shapley values are the only set of values that satisfy these properties.

$$a_i = \sum_{z \subseteq x} \frac{|z|! (M - |z| - 1)!}{M!} [f_x(z) - f_x(z \setminus \{x^i\})]$$

Apply sampling approximations to above equation and approximate the effect of removing a variable from the model by integrating over samples

Young (1985) demonstrated that Shapley values are the only set of values that satisfy these properties.

$$a_i = \sum_{z \subseteq x} \frac{|z|! (M - |z| - 1)!}{M!} [f_x(z) - f_x(z' \setminus \{x^i\})]$$

Apply sampling approximations to above equation and approximate the effect of removing a variable from the model by integrating over samples.

Baseline and path based methods.

Friedman, Eric J. Paths and consistency in additive cost sharing. *International Journal of Game Theory*, 32(4): 501–518, 2004.

Given $\gamma = (\gamma_1, \dots, \gamma_n): [0,1] \rightarrow R^n$ be a smooth function specifying a path in R^n from baseline x^b to input x , that is, $\gamma(0) = x^b, \gamma(1) = x$.

$$\int_{\alpha=0}^1 \frac{\partial F(\gamma(\alpha))}{\partial \gamma_i(\alpha)} \frac{\partial \gamma_i(\alpha)}{\partial \alpha} d\alpha$$

$$\mathcal{A}_j^i(x) = (x_j - x_j^b) \times \int_{\alpha=0}^1 \partial_j \mathcal{F}^i(x^b + \alpha(x - x^b)) d\alpha$$

IG. Sundararajan et. al.'17

Compositional

$$f(x) = \sum_{j=1}^M w_j x_j + b$$

$$a_i = w_i (x_i - E[x_i])$$

1. If the model makes a prediction on an input robustly in the causal neighborhood, that is, there is redundancy of features on an input, then it is more confident.
2. For out-of-distribution or adversarial examples, the model's prediction is not robust in causal space.

Given an input \mathbf{x} for a model \mathcal{F} where \mathcal{F}_i denotes the i -th logit output of the model, we can compute attribution of feature \mathbf{x}_j of \mathbf{x} for label i as $\mathcal{A}_j^i(\mathbf{x})$. We can then obtain confidence in two steps:

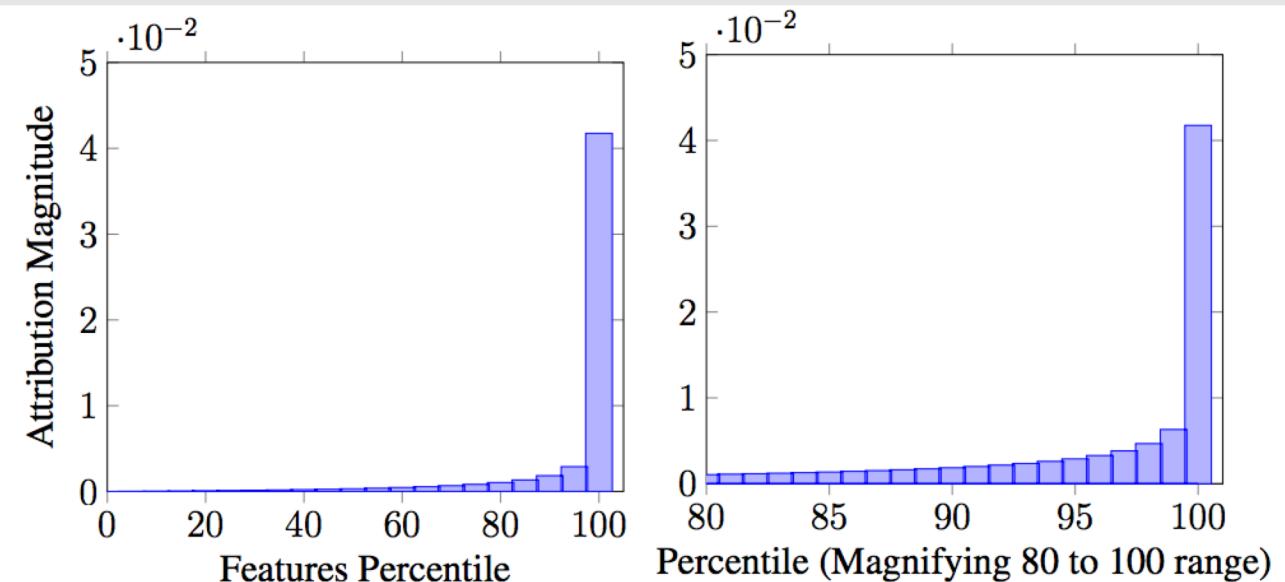
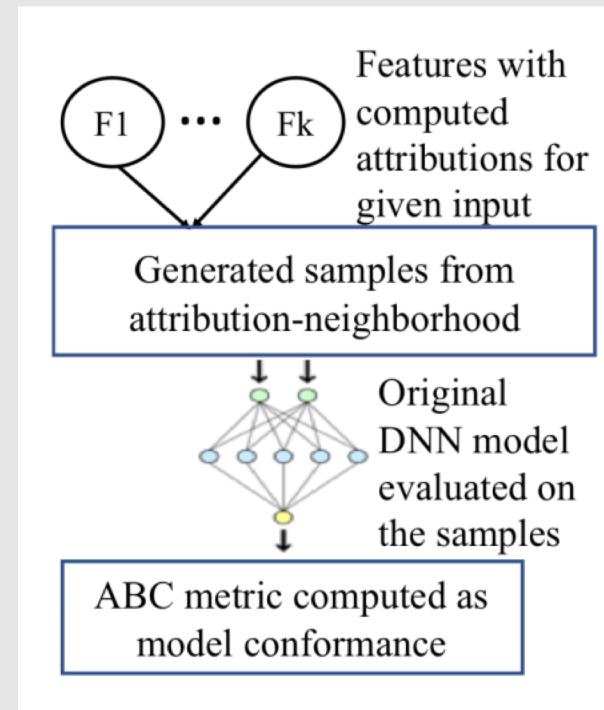
- Sample in neighborhood of \mathbf{x} by mutating each feature \mathbf{x}_j with probability $\frac{|\mathcal{A}_j^i(\mathbf{x})/\mathbf{x}_j|}{\sum_j |\mathcal{A}_j^i(\mathbf{x})/\mathbf{x}_j|}$ where the feature \mathbf{x}_j is changed to flip the label away from i .
- Report the fraction of samples points in the neighborhood of input \mathbf{x} for which the decision of the model conforms to the original decision as the conservatively estimated confidence measure.

Algorithm 1 Evaluate confidence $c(\mathcal{F}, \mathbf{x})$ of machine learning model \mathcal{F} on input \mathbf{x}

Input: Model \mathcal{F} , Input \mathbf{x} with features $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, Sample size S

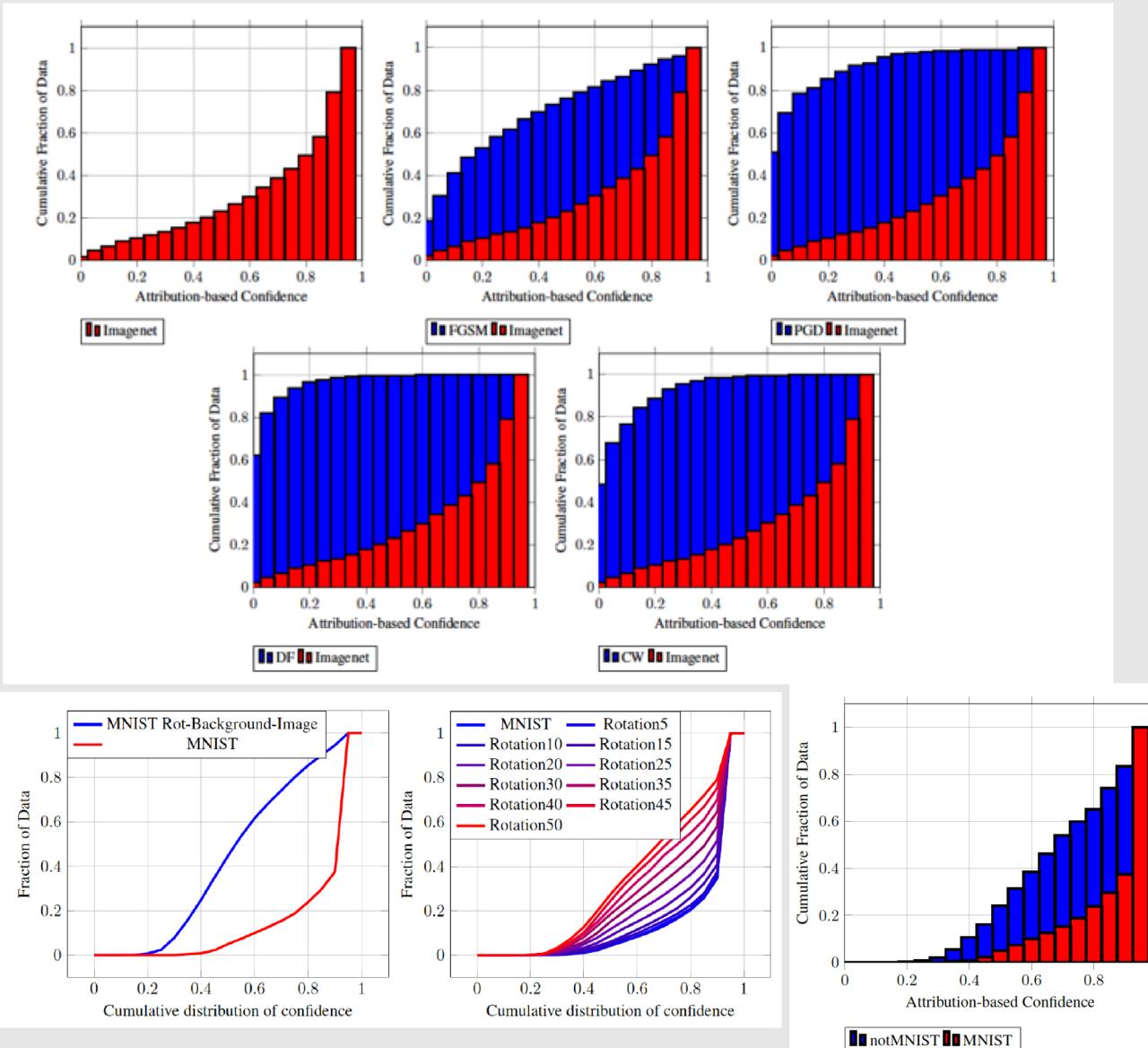
Output: Confidence metric $c(\mathcal{F}, \mathbf{x})$

- 1: $\mathcal{A}_1, \dots, \mathcal{A}_n \leftarrow$ IG attributions of features $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ from input \mathbf{x}
 - 2: $i \leftarrow \mathcal{F}(\mathbf{x})$ {Obtain model prediction}
 - 3: **for** $j = 1$ to n **do**
 - 4: $P(\mathbf{x}_j) \leftarrow \frac{|\mathcal{A}_j/\mathbf{x}_j|}{\sum_{k=1}^n |\mathcal{A}_k/\mathbf{x}_k|}$
 - 5: **end for**
 - 6: Generate S samples by mutating feature \mathbf{x}_j of input \mathbf{x} to baseline with probability $P(\mathbf{x}_j)$
 - 7: Obtain the output of the model on the S samples.
 - 8: $c(\mathcal{F}, \mathbf{x}) \leftarrow S_{conform}/S$ where model's output on $S_{conform}$ samples is i
 - 9: **return** $c(\mathcal{F}, \mathbf{x})$ as confidence of prediction by the model \mathcal{F} on the input \mathbf{x}
-

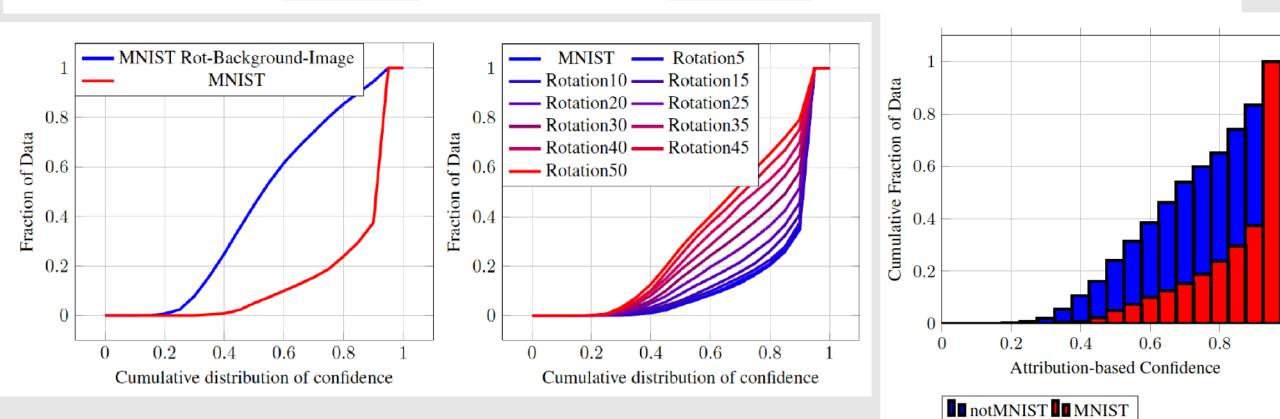


Feature Concentration in well-trained models

Theorem 1. *The sensitivity of the output $\mathcal{F}(\mathbf{x})$ with respect to an input feature \mathbf{x}_j in the neighborhood of \mathbf{x} is approximately the ratio of the attribution $\mathcal{A}_j(\mathbf{x})$ to the value of that feature \mathbf{x}_j , that is, $\frac{\mathcal{A}_j(\mathbf{x})}{\mathbf{x}_j}$.*

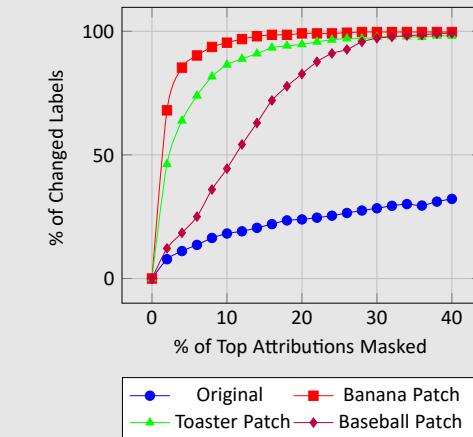
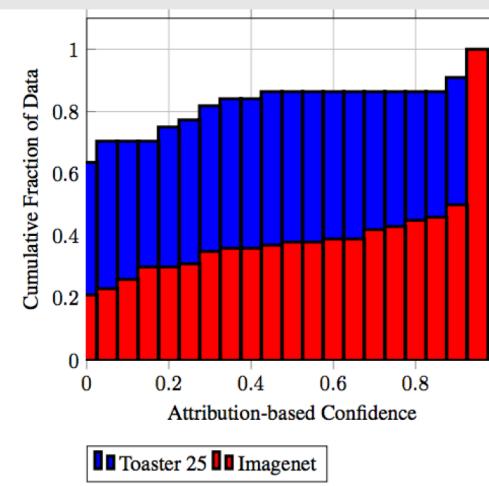
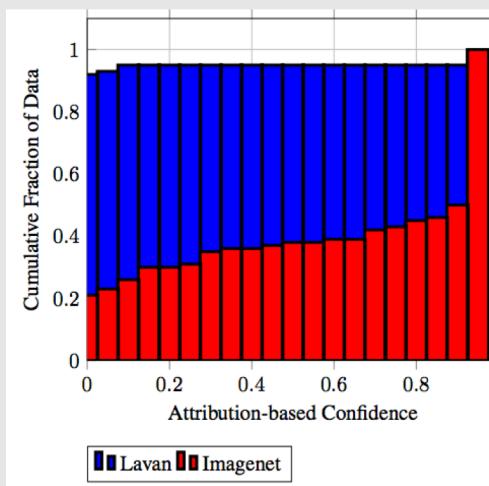
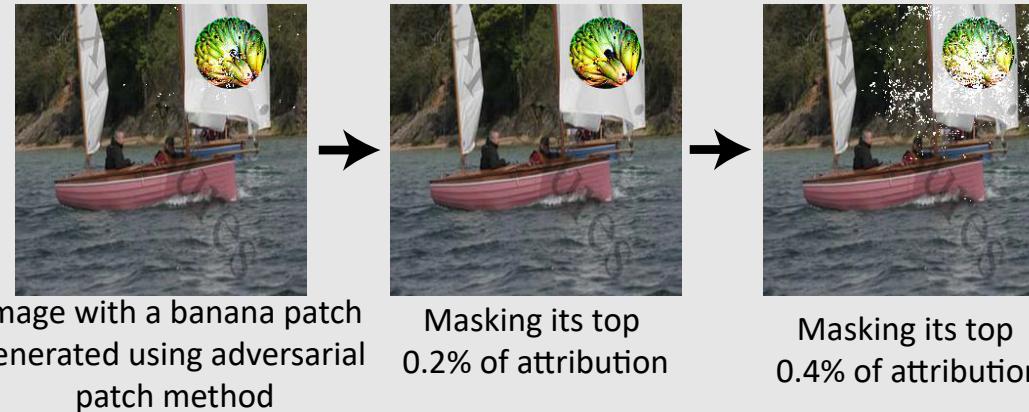


Adversarial attack leads
to low confidence
prediction

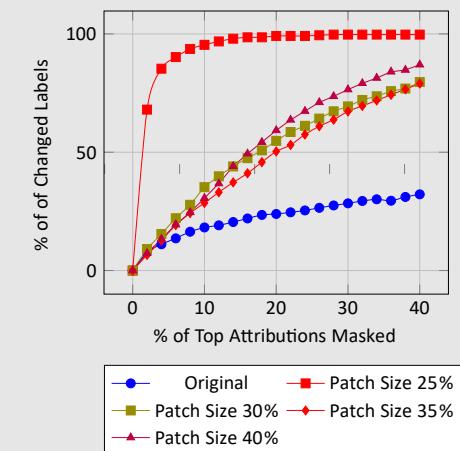


Out of distribution data
leads to low confidence
prediction

Adversarial Patch Attacks

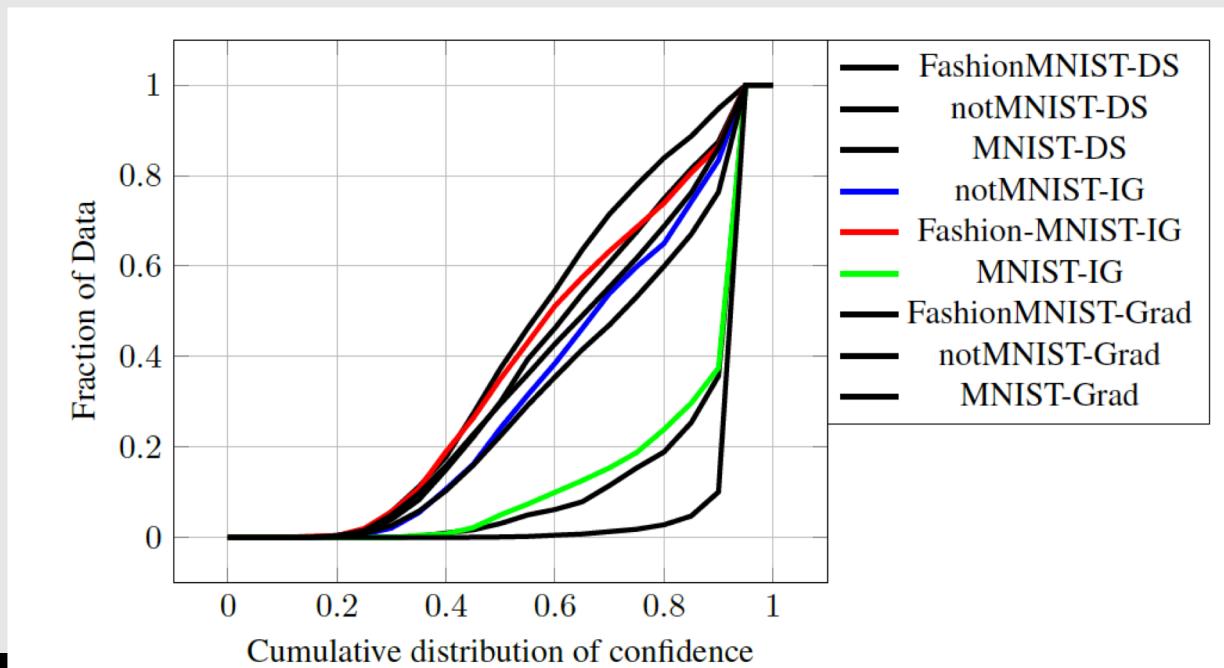
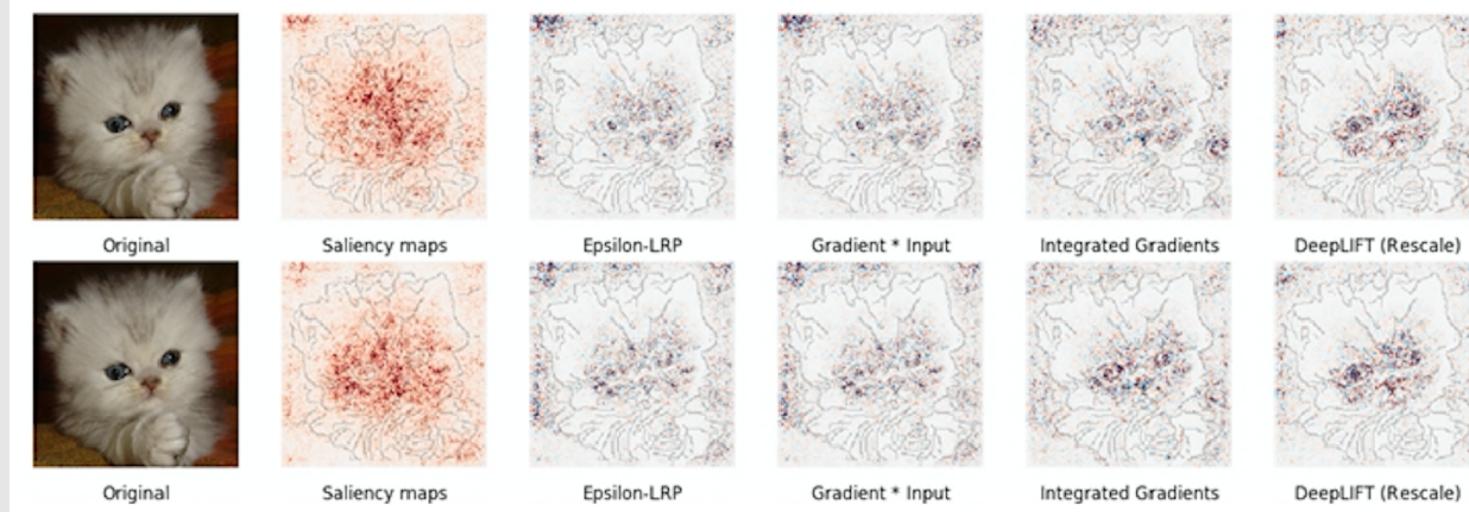


Dropping 0.4% of the attribution causes 99.71% of the attacks based on banana patches, 98.14% of the attacks based on toaster patches, and 99.20% of the attacks based on baseball patches to be detected.



Masking 0.4% of attributions caused nearly 80% of labels to change for images with adversarial patches.

ABC: Different Attribution Methods



Task 2.3: Fast and Adaptive Learning

D is set of MDPs θ corresponds to policy

Cumulative J is expected over MDPs in D_1

Input:

- Feasible set Θ , data set D , and probability $1 - \delta$.
- Objective function J such that $J(\theta, D) \in [0, 1]$ is an estimate of the utility of the solution θ , computed using data D .
- A satisfactory expected return j that our algorithm should achieve with at least probability $1 - \delta$.

Output: A solution, $\theta \in \Theta$, or NO SOLUTION FOUND.

Partition D into two data sets, D_1 and D_2 ;

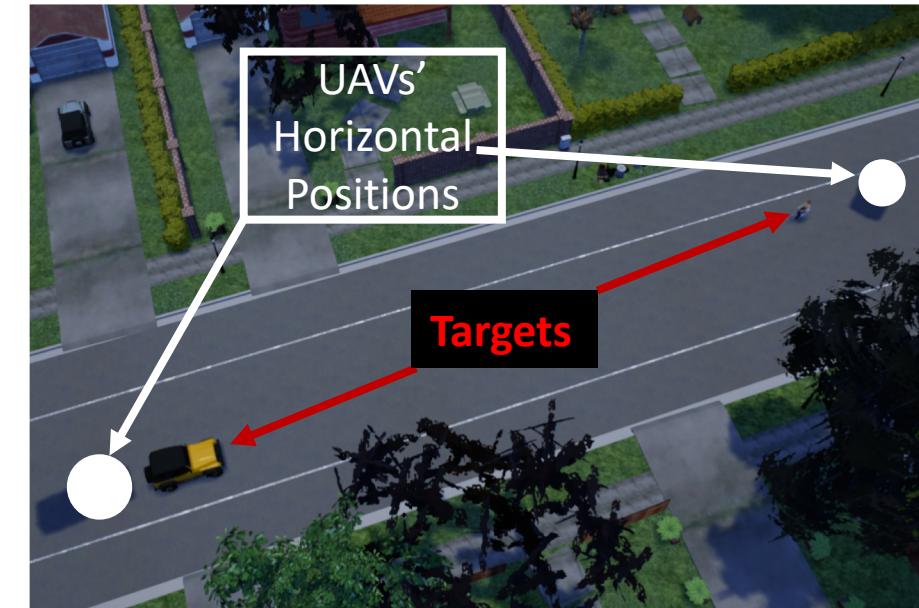
$$\theta_c = \operatorname{argmax}_{\theta \in \Theta} J(\theta, D_1);$$

$$\text{if } J(\theta_c, D_2) - \sqrt{\frac{\ln(1/\delta)}{2|D_2|}} \geq j \text{ then return } \theta_c;$$

return NO SOLUTION FOUND;

7 settings: 2 with insufficient data for safe adaptation

- Algorithm detected insufficient data in these 2 cases, and refused to return a solution in 100% of the trials.
- For other 5 settings, returned solution in 99.6% trials and all returned solution exceeded safety requirement.



Environment: Using the CARLA framework, we are developing a suite of tracking/detecting, identifying, and locating/tracking tasks for RL agents (more on next slide).

Task 2.3: Fast and Adaptive Learning



Addressed Army Need: Supporting machine learning functions in rapidly changing battlefield environments

- Active and adaptive sequential learning strategies for learning tasks that change in a bounded manner:
 - Key idea involves using a **provably efficient estimator** for the change in learning problem from one time-step to the next in conjunction with the **active sampling strategy** and optimization at each stage.

$$\ell(y|x, \theta) \triangleq -\log p(y|x, \theta)$$

$$L_{U_t}(\theta) \triangleq \mathbb{E}_{X \sim U_t, Y \sim p(Y|X, \theta_t^*)} [\ell(Y|X, \theta)] \quad \|\theta_t^* - \hat{\theta}_{t-1}\|_2 \leq \rho$$

$$\hat{\theta}_{\Gamma_t} \triangleq \operatorname{argmin}_{\theta \in \Theta} \frac{1}{K_t} \sum_{k=1}^{K_t} \ell(Y_{k,t}|X_{k,t}, \theta) \quad X_{k,t} \sim \Gamma_t, Y_{k,t} \sim p(Y|X_{k,t}, \theta_t^*) \quad \mathbb{E}[L_{U_t}(\hat{\theta}_t) - L_{U_t}(\theta_t^*)] \leq \varepsilon$$

