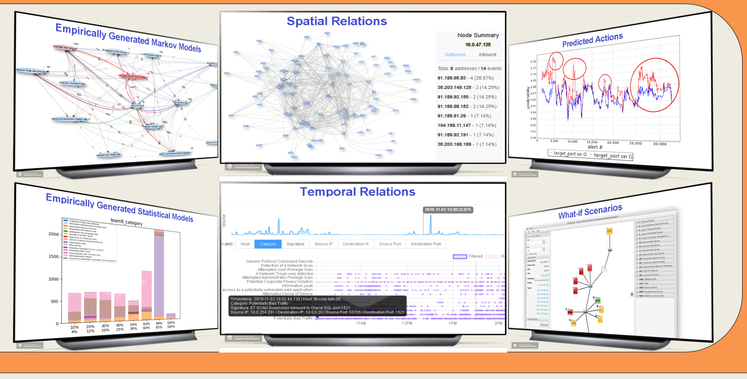
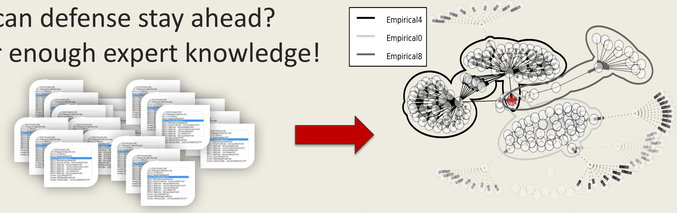
Extracting and Synthesizing Cyberattack Behavior Models for Predictive Intelligence

S. Jay Yang, M. Kuhl, B. Stackpole, and D. Johnson
Rochester Institute of Technology



Motivation and Goal

- Too many alerts and too little time!
- How can defense stay ahead?
- Never enough expert knowledge!



Transforming passive, reactive cyber defense into one with ***actionable real-time predictive intelligence***.

Challenges

- Not enough up-to-date cyberattack scenarios for the community to build knowledge & solutions.
  - Cyberattacks are diverse & fast-evolving w/ little ground truth.
  - Observables with categorical features are heterogeneous, noisy, incomplete, and deceptive.
- No sufficient theoretical grounding connecting adversary behavior to computational techniques.
- No efficient learning algorithm to create interpretable summary of temporal and spatial characteristics from categorical features.

Innovations and Publications

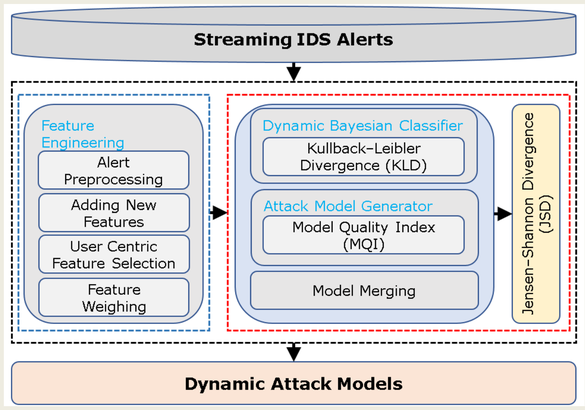
ASSERT: semi-supervised and dynamic generation of unique attack behavior models without expert knowledge.

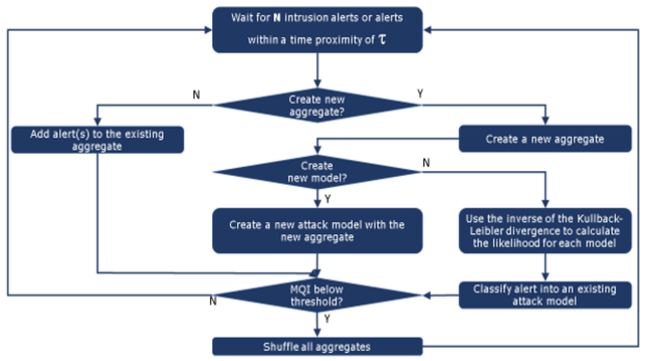
**A. Okutan and S. J. Yang, "ASSERT: Attack Synthesis and Separation with Entropy Redistribution towards Predictive Cyber Defense", Springer Journal on Cybersecurity, 2:15, May 2019.*
**A. Okutan, F.-Y. Cheng, S.-H. Su, and S. J. Yang, "Dynamic Generation of Empirical Cyberattack Models with Engineered Alert Features," in Proceedings of 2019 IEEE MILCOM, USA, Nov 12-14, 2019, Norfolk, VA..*

CASCADES: generation of synthetic attack scenarios extrapolated from extracted attack behavior models.

**S. Moskal, S. J. Yang, & M. Kuhl, "Cyber Threat Assessment via Attack Scenario Simulation using an Integrated Adversary and Network Modeling Approach," Journal of Defense Modeling and Simulation, 15.1, pp.13-29, 2018.*
**C. Sweet, S. Moskal, and S. J. Yang, "On the Veracity of Cyber Intrusion Alerts Synthesized by Generative Adversarial Networks," arXiv:1908.01219 [cs.LG].*

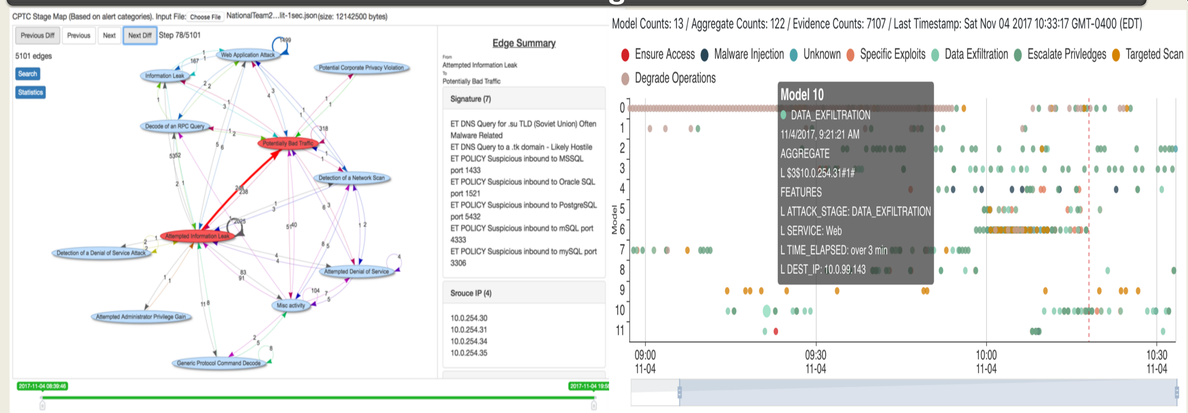
ASSERT: Architecture and Process Flow





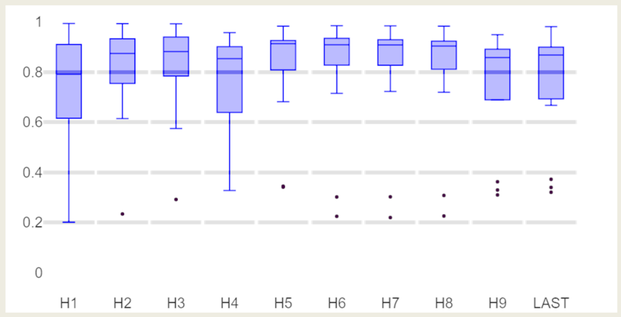
ASSERT: a semi-supervised Bayesian learning approach with information theoretical measures over non-parametric, categorical feature space to generate and update empirical attack models in near real-time.

ASSERT: Visualizing Attack Models

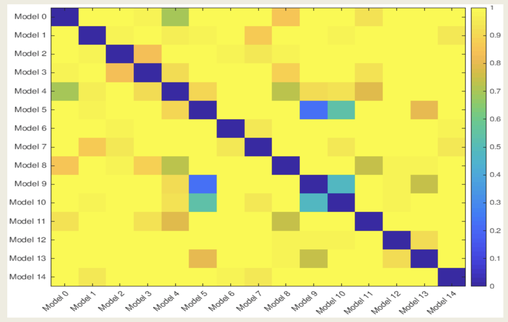


Interactive visualization that explores the spatial and temporal characteristics exhibited within each extracted attack model and compare the unique ones to reveal critical attack tactics.

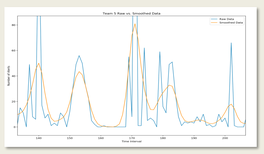
Unique Attack Models



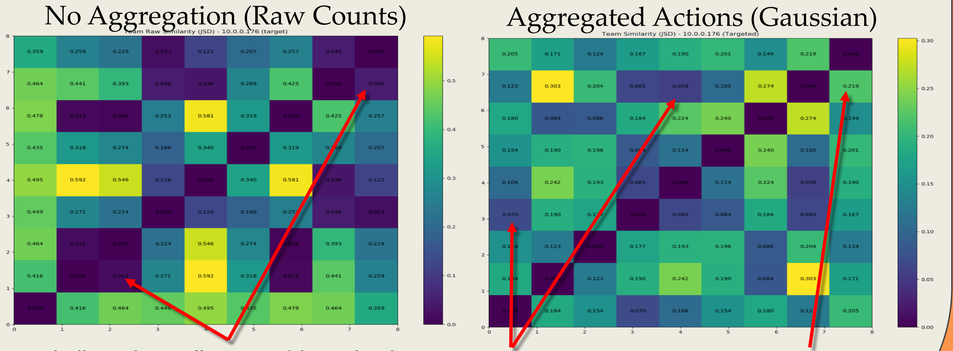
The JSD box plots between each extracted attack model and the distribution of all Suricata alerts (CPTC'17) show high uniqueness of the models over the hours throughout the competition.



The JSD heat map between extracted attack models (CPTC'17) shows excellent separation of attack behaviors (yellow indicates high divergence).



JSD comparing the usage of attack stages between competing teams (CPTC'18) against a domain controller
(Lower Score = **High** Similarity)

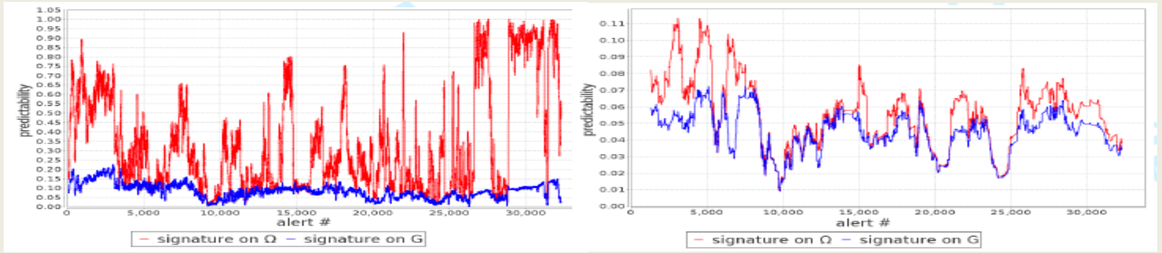


High "similarity" caused by a high freq. of one alert.

Found new similarities!

Different behaviors after aggregation

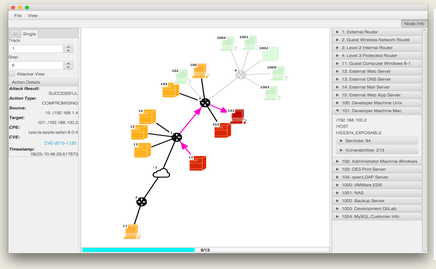
Attack Models Enhance Predictability of Future Actions



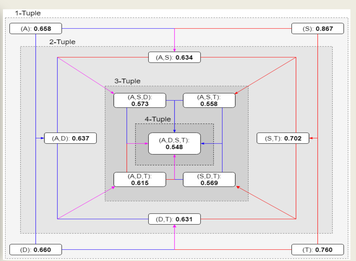
Predictability of "how" is enhanced (from blue to red) significantly with the extracted attack models using ASSERT.

Even the predictability of "unseen signatures" could be enhanced (from blue to red) using the extracted attack models.

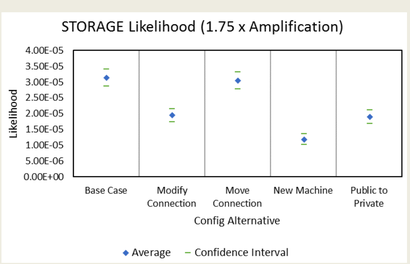
Simulated Multistage Attacks Help Evaluate What-if Scenarios



Monte-Carlo simulation based on adversary capability, opportunity, intent, and preference, driven by extracted characteristics in attack models.



GAN-based generation of synthetic attack scenarios shows promising recovery of feature distribution at various levels.



Rare-event simulations reveal high-fidelity albeit extremely small likelihoods under various changes to system configurations.

The 4th NSF Secure and Trustworthy Cyberspace Principal Investigator Meeting
October 28-29, 2019 I Alexandria, Virginia

Supported by NSF Awards # 1526383 and 1742789.