Today’s webinar topic is “Anticipatory Cyber Defense via Predictive Analytics, Machine Learning and Simulation" with Shanchieh (Jay) Yang. Our host is Jeannette Dopheide.

The meeting will begin shortly. Participants are muted. Click the Chat button to open the chat view and ask a question.

**This meeting will be recorded.**

The Trusted CI Webinar Series is supported by National Science Foundation grant #1547272.

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the NSF.
S. Jay Yang
Professor and Department Head
Computer Engineering
Director of Global Outreach
Center of Cybersecurity
Rochester Institute of Technology
Acknowledgement

- Special thanks to Florence Hudson for enabling Transition-to-Practice for the projects!

- Contributing students / alumni / collaborators
  - Ahmet Okutan, Stephen Moskal, Gordon Werner, Shao-Hsuan Su, Fu-Yuan Cheng, To Chen Chang, Christopher Sweet, Lu-Tzu Li, Ashley Resuta & Kyle Ki.

- Research described in this talk is/was supported by NSF (Awards 1526383 and 1742789), IARPA (FA875016C0114), and NSA (H98230-15-1-0277).
Let’s face it – We will be attacked!

Just a matter of

*When, Where, How, and What?*

So what can we do about it?
Cyber Defense w/ Anticipatory Intelligence

• Drawing analogy from human cognition: Perception – Comprehension – *Anticipation*

• Aiming at predicting or recognizing cyber attacks/threats *early on* to deploy (resilient) cyber defense before *critical* lost of information or mission impact!
Examples of Previous Work on Prediction

Use of Variable Length Markov Models (VLMM) to predict next attack actions [Fava08, Du10].

* CyGraph [Noel16].
Key Challenges

• Forecast of cyberattacks based on seemingly unrelated signals.
  • Incomplete, insignificant, imbalanced, and non-stationary data.

• Rapid recognition of critical and related attack behaviors.
  • Large volume, noisy, and potentially misleading observables.

• Simulation of potential scenarios based on limited usable data.
  • Evolving attack behaviors requires an intelligent way to generate scenarios that have not yet observed.
Diverse and Evolving Attack Behaviors

Non-stationary and limited data makes prediction very challenging!

* Data collected during 2017 CPTC National.
Capturing Cyber-Attack Prediction of Threats from Unconventional Resources

Early Indicators

Critical Scenarios

Salient Features

Asserting Attack Strategy Synthesis & Ensemble Prediction

Cascades: Cyber Attack Scenario & Network Defense Simulator
CAPTURE Forecasts Plausible Cyberattacks

…using incomplete, insignificant, lagged, & non-stationary signals w.r.t. the cyber incidents.

Examples of Unconventional Signals:

**Raw Data Sources** – Twitter, GDELT, Dark Web, News, OTX, Hackmageddon

**Unconventional Signals** – Spike Profile in Negative Sentiment, State/Organizational Instability, Cyber Vulnerability Mentions, Special-interest Events, Underground Economy, Malware Product Cycle, etc.
**CAPTURE v1: Prototype w/ ELFS and CDMT**

- **ELFS** (Entropy-based Lagged Feature Selection): finds the least set of lagged signals that maximize information gain / are most correlated.
- **CDMT** (Concept Drift for Model Training): concatenates the periods of signal-GT data that exhibits similar predictive power.
- **Provenance**: offers insights on what and how unconventional lagged signals contribute to the warnings generated.
- **Dial-Function**: allows user to tune the confidence threshold to generate different true/false positives/negatives.
- **N-Day Ahead Forecast**: predicts likelihood of having an attack up to N-day before incident.
CAPTURE v1 Demo against Baseline BayesNet

- **Target Organization:** Knox
- **Attack Type:** Malicious Email
- **Training Period:** Nov 2016 – Feb 2018
- **Test Period:** March 2018
- **Max Signal Lags:** 15 days for 400+ signals
- **Forecasting for:** Next 7 days
• Trained using cumulatively updated GT incidents up to the test month for two attack types (EM & ME) for Knox.
• Used 146 signals from 8 data sources with up to 60-day lags.
• Sufficiently high F-Measures (mostly 0.7~0.8) with CAPTURE (CC) against baseline BayesNet Forecaster (B30, B15, B3), except the “unpredictable” cases.
• Relevant lagged signals shown for EM – similar but changing lags.

[Okutan19a]
ASSERT Recognizes Emerging Attack Behaviors

- Semi-supervised Bayesian learning integrated with clustering to *generate* and *update* empirical attack models, and classify streaming observables to these models w/o ground truth [Okutan19b].

- Configurable spatial, temporal, and contextual features.
- Wemmert-Gancarski Index & J-S Divergence for model quality.
- K-L Divergence and entropy redistribution to account for *unseen* features.
**ASSERT Generates and Updates Empirical Models**

- Maximum a posteriori to “group” streaming evidences of malicious activities to generate empirical models — recognize nontrivial behaviors.

\[
M_{MAP} = \arg\max_M P(M|X) = \arg\max_M P(X|M) P(M)
\]

- WGI to continuously assess the quality of models (treating models as clusters) — re-cluster using DBSCAN when needed.

---

**Statistical Attack Models**

Define a set of features, derived from the packet header information available:

- Protocol
- Port Distribution
- Graph Position
- ICMP Type

Feature distributions are learned in a "lazy", non-parametric way:

- Discrete Feature - Histogram
- Continuous Feature - Kernel Density Estimator

Feature distributions determine the likelihood of a sample, under the hypothesis of a particular attack model:

\[
M_{MAP} = \arg\max_M P(M|X) = \arg\max_M P(X|M) P(M)
\]

---

**Feature Space Visualization**

How to validate that the segmentation result is intuitive in more complex cases?

Parallel Coordinate plots can visualize high dimensional feature spaces.

Consider an example with slightly more complex graph structure:

- Distributed Scan on TCP Ports 17826 & 24040
- Backscatter on TCP Port 10322 Unassigned
- Scan on TCP Port 80 HTTP Traffic
- Suspicious Repeat Collaborators
ASSERT Uses Novel and Unseen Features

- Configurable features (modeled as non-parameterized histograms) extracted from intrusion alerts to reflect attack characteristics of interests:
  - E.g., time elapsed (when), services targeted (what), change of targets (where), exploit/scan types (how), attack stages (how), etc.
- Use KLD and entropy redistribution to accommodate general non-parameterized feature distribution, where *unseen* features do not mean absolutely impossible.
ASSERT Recognizes Emerging Attack Behaviors

**ASSERT:**
Attack Synthesis & Separation with Entropy Redistribution
ASSERT Enhances Predictability

Predictability of “how” could be enhanced (from blue to red) significantly with the intelligent grouping of observables by ASSERT.

[Okutan19b]

Even the predictability of "unseen signatures" could be enhanced (from blue to red) with the intelligent grouping of observables by ASSERT.
CASCADeS Generates Synthetic Attack Scenarios


Sequences of Actions ➔ Behavior Model

Aggregate alerts to uncover potential attack action sequences.

Extract key attack characteristics from attack sequences using machine learning techniques, e.g., Fuzzy-Logic, GAN/RNN, Markov model, etc.

Simulate attack scenarios and its impact on networks with various configurations.
CASCADeS Generates Synthetic Attack Scenarios

- Making of a Cyber Attacker:
  - *Capabilities, Opportunity, Intent, Preferences.*
  - Attacker’s knowledge develops over time to influence choice of actions.
- Coupled with *Virtual Terrain*— a model of connected systems w CVEs.

<table>
<thead>
<tr>
<th>Intent</th>
<th>Opportunities</th>
<th>Capabilities</th>
<th>Preferences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation objectives, attack goals, attack plan</td>
<td>What can be done; how the attacker can achieve their goal</td>
<td>Skill set of the attacker: tools, techniques, vulnerabilities</td>
<td>Preferences over types of attacks, services, techniques, etc.</td>
</tr>
</tbody>
</table>

Decreasing size of possible attack actions through each stage

- Use *Important Sampling* to estimate likelihood of rare attack scenarios!

<table>
<thead>
<tr>
<th>Estimate Probabilistic Attack Parameters</th>
<th>Assess Candidates for Amplification</th>
<th>Perform Amplification</th>
<th>Simulate using the Amplified Network</th>
<th>Scale the Output to the Original Network</th>
</tr>
</thead>
</table>

[Krall19] [Moskal18]
CASCADES Generates Synthetic Attack Scenarios
CASCADeS Enables Analysis of What-if’s

<table>
<thead>
<tr>
<th></th>
<th>Amateur</th>
<th>Expert</th>
<th>Comprehensive</th>
<th>Random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Steps</td>
<td>28.76</td>
<td>20.36</td>
<td>36.54</td>
<td>33.57</td>
</tr>
<tr>
<td>Minimum</td>
<td>10</td>
<td>4</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Maximum</td>
<td>113</td>
<td>72</td>
<td>168</td>
<td>106</td>
</tr>
<tr>
<td>Failure Rate</td>
<td>0.16</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>% change</td>
<td>4.35</td>
<td>-26.33</td>
<td>-33.41</td>
<td>-29.14</td>
</tr>
</tbody>
</table>

*A small misconfiguration leads to change in likelihood and steps needed to succeed the attack.*

* Rare-event simulation enables what-if analysis even for unlikely events.
Use of Generative Adversarial Networks (GAN) to learn and recreate intrusion alerts of similar “behaviors”.

Intersection/Dependency of Alert Features:
- Alert Signature (A)
- Destination Port (D)
- Source IP (S)
- Time bin (T)

Intersection score / JSD indicating similarity between original and generated alert sequences.

<table>
<thead>
<tr>
<th>Features</th>
<th>Ground Truth</th>
<th>Generated</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+S+D</td>
<td>0.591</td>
<td>0.337</td>
</tr>
<tr>
<td>A+D+T</td>
<td>0.652</td>
<td>0.607</td>
</tr>
<tr>
<td>S</td>
<td>A,D</td>
<td>0.799</td>
</tr>
<tr>
<td>T</td>
<td>A,D</td>
<td>0.657</td>
</tr>
<tr>
<td>A</td>
<td>D,T</td>
<td>0.050</td>
</tr>
<tr>
<td>D</td>
<td>A,S</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Joint & Conditional Entropy of Feature Distributions:

Q: How can GAN capture higher order feature dependencies to compensate lower-order rare occurrences.

[Sweet19]
How can visualization assist to **recognize** critical attack “behaviors” and to **rationalize** predictions from CAPTURE, ASSERT, and CASCADES?
Anticipatory Intelligence for Cyber Defense - Challenges and Opportunities

- Forecast plausible attacks before they happen (CAPTURE).
  - Challenge: unreliable forecasting due to non-stationary predictive power from individually insignificant signals.
  - Unconventional sensor development w/ social media analysis, criminology, management and economics, international and societal conflicts.

- Recognize emerging non-trivial and critical attacks (ASSERT).
  - Challenge: finding unknown needles from stacks of hays in a timely manner.
  - Data collection and trusted visual analytics by security professionals. Non-pattern finding AI problems (e.g., critical feature derivation).

- Synthesize attack scenarios for what-if analysis (CASCADES).
  - Challenge: revealing the rare and nontrivial scenarios by extracting evolving attack behaviors and cyber attack/defense surfaces.
  - Criminology and psychology grounding to learn and simulate adversary behavior. Simulation as a service for business and government partners.
Discussion

Feedback

Q&A

S. Jay Yang – jay.yang@rit.edu

See you at TTP Workshop in Chicago on June 19, 2019

(https://trustedci.org/2019-ttp-workshop/)
Questions?

Please take our survey
2019 Cybersecurity Technology Transition to Practice (TTP) workshop - June 19th, Chicago

For more info and to request an invitation, go to our TTP page: https://trustedci.org/ttp
Trusted CI Engagement Applications Now Being Accepted!

Today: Apply at https://trustedci.org/application
April 3, 2019: Application deadline.
May 1, 2019: Applicants notified of decisions.
July, 2019: Engagement kick-off

For more information:
"An Overview of the CCoE Engagement Application with Von Welch" Video
https://youtu.be/g-3ZaUjsJP8
Trusted CI Open Science Cybersecurity Fellows program

For more info and to apply to the program:
https://trustedci.org/fellows
About the Trusted CI Webinar series

To view presentations, join the announcements mailing list, or submit requests to present, visit: https://trustedci.org/webinars

The next webinar is March 25th at 11am Eastern.
Topic: SecureCloud
Speaker: Casimer DeCusatis