Anomaly Detection in Cyber Networks using Graph-node Role-dynamics and NetFlow Bayesian Normalcy Modeling

Anthony Palladino, PhD, Senior Research Scientist
Christopher Thissen, PhD, Research Scientist
Andrew Spisak, PhD, Senior Research Scientist

Boston Fusion Corp.
70 Westview Street, Suite 100
Lexington, MA 02421
www.bostonfusion.com

Presented at FloCon2018
9 January 2018

©2018 Boston Fusion Corp.
Agenda

• Introduction
• Advanced Persistent Threats
• Graph-node Role-dynamics
• Bayesian Normalcy Modeling
• Summary
Introduction

- **Context Aware INference for Advanced Persistent Threat (CAIN for APT)**
  - DARPA Phase II SBIR

- **Challenge**
  - Stealthy cyber attacks slip past state-of-the-art defenses, dealing crippling blows to critical US military and civilian infrastructure

- **Goal**
  - Rapid, automated, and accurate prioritization of cyber alerts provides timely and comprehensive cyber situational awareness (SA)

- **Technical Approach**
  - Novel graph-analytics makes sense of noisy IDS sensors
  - Novel Bayesian Dynamic Flow Model flags odd network traffic
  - Tests and evaluations with APT simulations
Agenda

• Introduction
• Advanced Persistent Threats
• Graph-node Role-dynamics
• Bayesian Normalcy Modeling
• Summary
Advanced Persistent Threats

- Often associated with nation-state espionage
- Targets include private organizations & nation-states
- Low and Slow: Attack campaigns may last months
- Notoriously difficult to detect


Image: https://www.secureworks.com/blog/advanced-persistent-threats-apt-a
Simulated APT Scenarios

• **Simulation attributes**
  – Approx. 1 month of data per scenario
  – Servers, laptops, switches
  – Linux & Windows machines
  – Normal & attacked behavior
  – Generates IDS alerts and NetFlow traffic
  – Detailed attack timeline

• **Hurricane Panda simulation**
  – Attack distributed over 3 days
  – Database injection to gain credentials
  – Lateral movement and firewall deactivation

• **Energetic Bear (Crouching Yeti) simulation**
  – Attack distributed over 3 hours
  – Email phishing to redirect user to malicious website
  – Lateral movement through network using a remote-desktop exploit
  – Attacker attempted to clean-up logs and other traces

---

Network topology for simulations
Agenda

• Introduction
• Advanced Persistent Threats
• Graph-node Role-dynamics
• Bayesian Normalcy Modeling
• Summary
Graph-based Approach

- Fuses disparate IDSs
- Captures alert interdependencies
- Efficiently represents many alerts
- Robust to circumvention
- Unsupervised
- Facilitates causal analysis
- Optimal parameters determined automatically
Making Sense of Noisy IDS Sensors with Graph Analytics

• Novel, graph-based analysis of IDS alerts
  – Load IDS alerts into alert graph
  – Detect graph anomalies

• Advantages of graph-based approach:
  – Captures alert interdependencies
  – Fuses disparate IDSs
  – Efficiently represents alerts
  – Robust to circumvention

Alert Graphs from Hurricane Panda Simulation

Akoglu et al. 2014
**Alert Graphs**

- Graph of alerts (Not network topology)

**OSSEC Alert (Host IDS)**

```plaintext
```

**Snort Alert (Network IDS)**

```plaintext
```

**Alert Graph**

```
ip_10.10.255.50 ——— sid_2010939

ip_10.10.255.79 ——— ip_10.10.255.77

rule_11401

log_/var/log/vsftpd.log
```
Alert Graphs

- Graph of alerts (Not network topology)
- Alert properties become nodes
- Node colors indicate property type

**OSSEC Alert (Host IDS)**

```
```

**Snort Alert (Network IDS)**

```
```
Alert Graphs

OSSEC Alert (Host IDS)

Snort Alert (Network IDS)

• Graph of alerts (Not network topology)
• Alert properties become nodes
• Node colors indicate property type
• Edges connect nodes that co-occur in alerts
• Edges weighted by frequency of co-occurrence
Alert Graphs

• Cyber attacks change IDS alert logs
• Intuition
  – Changes in alert logs modify alert graph
  – Anomalies in the graph features (properties) may indicate cyber attacks

• Quick test
  – Degree of IP nodes shows marked changes during simulated attack
  – But a single feature is likely insufficient
  – What features should we track?
  – Should we model all features for anomalies?
• Infeasible to model every feature of every node
• Instead, use graph-based anomaly detection algorithms
• Role dynamics (Rossi et al., 2012)
  – Collect features and factorize as roles
  – Roles provide a succinct, integrated summary across a large number of features
  – Output is probability of membership in each role, for each node
  – Application to IDS alerts is novel
Infeasible to model every feature of every node

Instead, use graph-based anomaly detection algorithms

Role dynamics (Rossi et al., 2012)
- Collect features and factorize as roles
- Roles provide a succinct, integrated summary across a large number of features
- Output is probability of membership in each role, for each node
- Application to IDS alerts is novel
- Track role memberships over time
Role Dynamics

- **Why role dynamics?**
  - Linear
  - Weighted
  - Dynamic
  - Attributed
  - Unsupervised
  - Explainable
  - Extensible
  - Automated parameter selection
  - Available

- **Explainable**
  - Identifies anomalous nodes
  - Helps with causal analysis

- **Automated parameter selection**
  - Recursive features
  - Optimal number of roles
  - Set during a training period
Finding Role Anomalies

- **Role anomalies**
  - Now we have roles over time for all nodes in graph
  - How to identify anomalies in the roles?
- **Aggregate changes into a few useful metrics**
  - For example, average magnitude of the rate of change in role membership:
    \[ \sum_{n=1}^{N} |P_n(t) - P_n(t-1)| / N \]
  - Monitor metrics for anomalies

APT Attack Start
Results: APT Scenario 1

- **Hurricane Panda scenario**
  - Virtual network of servers, laptops, switches, etc.
  - Linux & Windows machines
  - 9 Nov 2016 – 3 Dec 2016
  - Attack distributed 30 Nov – 2 Dec
  - Snort (NIDS) & OSSEC (HIDS)
  - Database injection to gain credentials
  - Lateral movement and firewall deactivation

- **Results**
  - Using threshold at 0.3, CAIN identified 4 anomalies
  - Second two anomalies relate to machines coming online for the first time
  - Last anomaly corresponds with the start of Hurricane Panda’s attack
Results: APT Scenario 2

• **Energetic Bear scenario**
  – Same network as Hurricane Panda
  – 1 Jan 2017 – 4 Feb 2016
  – Attack on Jan 31, 2017
  – 644,067 OSSEC (HIDS) alerts
  – Email phishing to redirect user to malicious website
  – Lateral movement through network using a remote-desktop exploit
  – Attacker attempted to clean-up logs and other traces

• **Results**
  – Using threshold at 0.3, CAIN identified 2 anomalies
  – Third anomaly corresponds with the start of the Energetic Bear attack
Conclusions: Making Sense of Noisy IDS Sensors with Graph Analytics

• **Graph-based Role-dynamics:**
  – Fuses IDS sensor alerts
  – Reduces >750k alerts to a handful of anomalies
  – Identifies anomalies in IDS alerts during APT attacks

• **Success in 2 APT scenarios demonstrates:**
  – Robust to different types of APTs and attack vectors
  – Insensitive to IDS systems
Agenda

• Introduction
• Advanced Persistent Threats
• Graph-node Role-dynamics
• Bayesian Normalcy Modeling
• Summary
Bayesian Dynamic Flow Model

- Unsupervised model of NetFlow traffic dynamics
- Assume data follows Poisson distribution
  \[ x_t \sim \text{Poisson}(\phi_t) \]
- Model temporal evolution as Gamma-Beta discount model
  - Prior: \[ x_t \sim P(\phi_t|x_{0:(t-1)}) = \Gamma(\delta_t r_{t-1}, \delta_t c_{t-1}) \]
  - Posterior: \[ x_t \sim P(\phi_t|x_{0:t}) = \Gamma(\delta_t r_t, \delta_t c_t) \]

(X. Chen, et al. 2016)
Results

Bayesian Dynamic Flow Model

- Identifies anomaly during APT attack
- Complementary to graph-based role-dynamics
- Multiple methods corroborate detection

Model Results for 10.10.255.50 <-> 10.10.255.63
APT Attack Start

Identifies anomalous change in packet flow volatility

Obs. – Pred.
Agenda

• Introduction
• Advanced Persistent Threats
• Graph-node Role-dynamics
• Bayesian Normalcy Modeling
• Summary
Summary

- Developed two complementary anomaly detection techniques
  - IDS: Graph-based Role Dynamics
  - NetFlow: Bayesian Dynamic Flow Model
- Tested on two APT scenarios
  - Hurricane Panda
  - Energetic Bear (a.k.a. Crouching Yeti)
- Successful anomaly detection in two APT scenarios suggests:
  - Robust to different types of APTs and attack vectors
  - Insensitive to IDS systems
References

Statements / Disclaimers

• Copyright 2018 Boston Fusion Corp.

• This research was developed with funding from the Defense Advanced Research Projects Agency (DARPA)

• The views, opinions and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government

• Distribution Statement "A" (Approved for Public Release, Distribution Unlimited)