True Attacks, Attack Attempts, or Benign Triggers? An Empirical Measurement of Network Alerts in a Security Operations Center

Limin Yang*, **Zhi Chen***, Chenkai Wang, Zhenning Zhang, Sushruth Booma, Phuong Cao, Constantin Adam, Alexander Withers, Zbigniew Kalbarczyk, Ravishankar K. Iyer, Gang Wang

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

IBM

NCSA

* The authors contribute equally to this paper (co-first authors).

Security Operation Centers (SOC) are Critical to Security Incident Response

Complex Networks

→

Monitoring and Detection

Analysis and Response

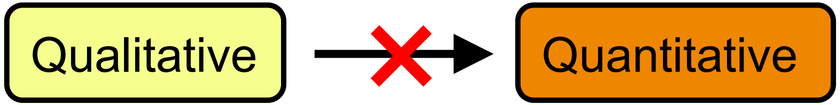
Modeling and Updating

What SOCs are doing

SOCs Face Critical Challenges

A study [1] with 2,303 IT security and SOC analysts:

- The majority (51%) feel their team is overwhelmed by the volume of alerts
- 55% admit that they aren't entirely confident in their ability to prioritize and respond to alerts
- 70% mentioned emotional impaction by their work managing IT threat alerts



Qualitative

Quantitative

Wound Up

Anxious

Grumpy

Negative Personas & Behaviors

[1] Allie Mellen, Adapt Or Die: XDR Is On A Collision Course With SIEM And SOAR (Forrester, 2021)

Our Key Research Questions for a Quantitative Study:

1. What are the key bottlenecks in the SOCs for threat detection?

2. How excessive are the security alerts, and what are the common reasons behind the alert triggering?

3. How effective are the alerts to correlate or indicate true/successful attacks?

Real-world SOC Dataset from NCSA

NCSA: National Center for Supercomputing Applications

- Located in UIUC; thousands of servers
- 17,000+ users/researchers
- Segmented network: open vs. heavily guided
- SOC: SOC-2-Type-2 certified

NCSA's Delta supercomputer

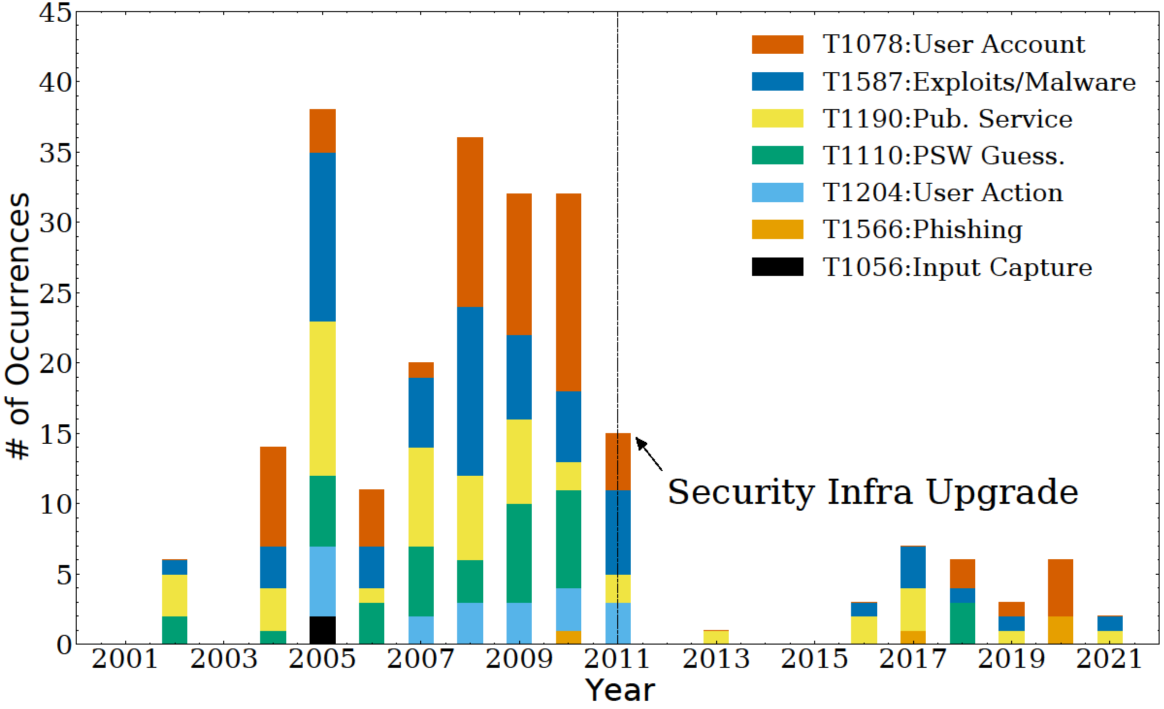
Data:

- 227 true attack incident reports from 2002 to 2022
- 115 million Zeek alerts from 2018 to 2022

zeek

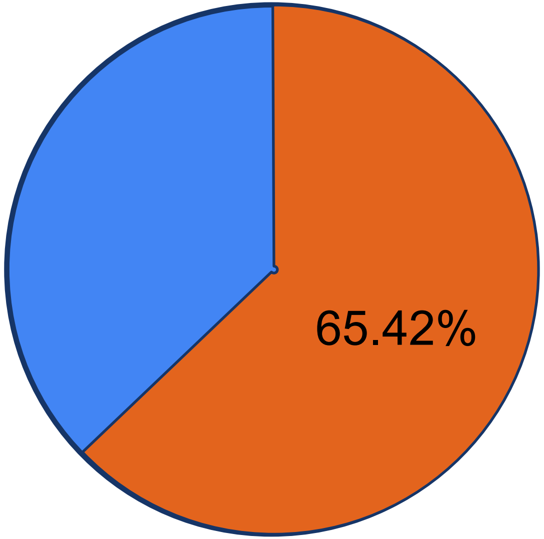
Attacks Incidents Report Analysis

Break-in Methods in the Past 20 Years



- Manually label incident reports with MITRE ATT&CK Techniques
- 178/227 (78%) break-in method identified
- Security infra (e.g., 2FA) significantly reduce compromises since 2011

Bottlenecks for Threat Detection: Humans



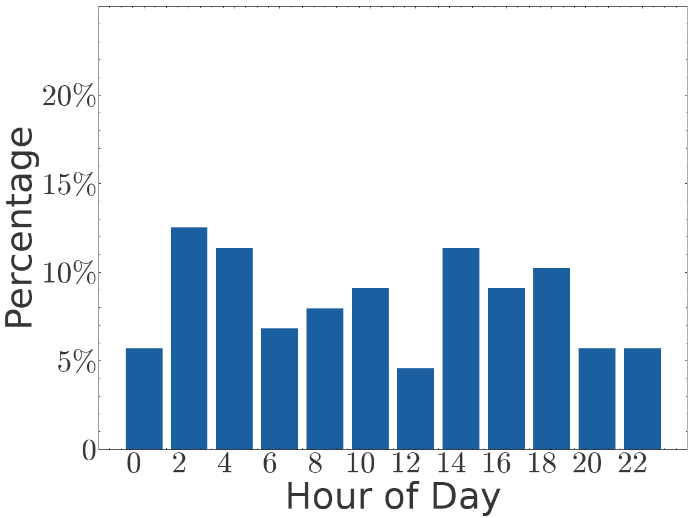
🟦 107 documented the involvement of analysts

🟧 70 need two or more analysts

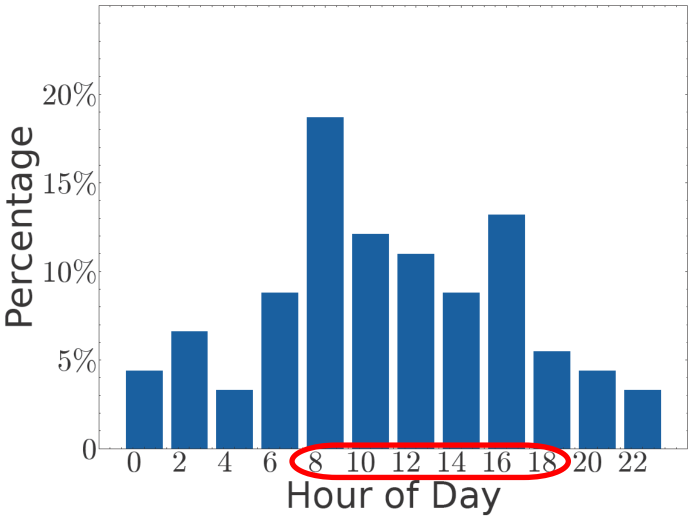
A larger number of people (e.g., 6–7) may involved. It may include people outside of the core SOC team

1. Usually takes more than one analyst to work on a single attack
2. Further, post-attack analysis takes a long time (53.2 days on average) to investigate and understand

Bottlenecks for Threat Detection: Humans



Attack Start Time

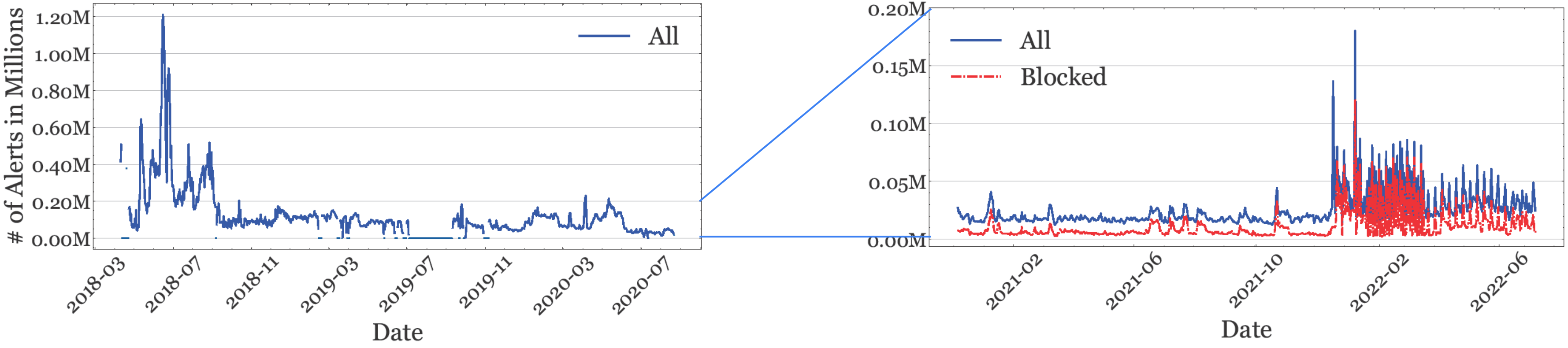


Attack Discovery Time

Attack detection is more aligned with analysts' working hours. This may cause extra delays for detection to the attacks happened during off-hours

Alerts Dataset Analysis

Excessive Volume of Security Alerts



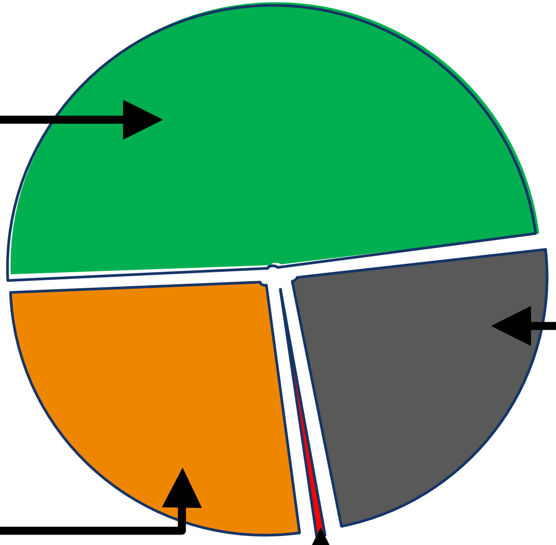
Time Range	Days	# of Alerts	Alerts Per Day	Auto-blocked (Black Hole Router)
04/2018 – 08/2020	751	101 million	134k	Not implemented
12/2020 – 07/2022	578	14 million	25k	6 million (45%)

Reasons for Triggering Security Alerts

Benign trigger

48.91%

- Correctly fired alerts
- Have business-justified reasons



Unknown

23.72%

- Cannot identify their alert reasons

Attack attempt

27.37%

- Tried to attack
- Failed to compromise the system

True attack

0.01%

- Correctly fired alerts
- Compromised the system successfully

Understanding Benign Triggers

- Manually check each type of alerts with two SOC analysts

- Benign triggers occupy a significant portion of alerts (at least **48.91%**)

Category	**# of Alerts**
Internal Scan	194,518 (2.76%)
External Pen-test	2,694,851 (38.19%)
DNS	1,462,404 (20.72%)
SSL	39,818 (0.56%)
Non-Attack	2,692,427 (38.15%)
Services w/ Exceptions	260 (0.003%)
Total	7,057,012

Not all benign triggers are well-documented by the SOC and it took significant (manual) efforts to gather evidence and craft rules to flag them

Reasons for Triggering Security Alerts

Benign trigger

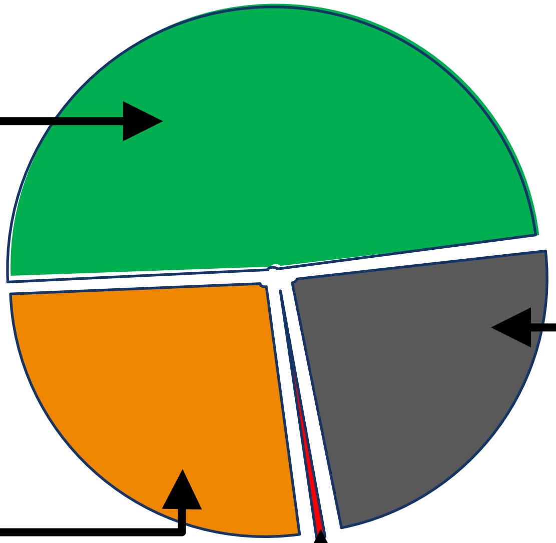
48.91%

- Correctly fired alerts
- Have business-justified reasons

Attack attempt

27.37%

- Tried to attack
- Failed to compromise the system



Unknown

23.72%

- Cannot identify their alert reasons

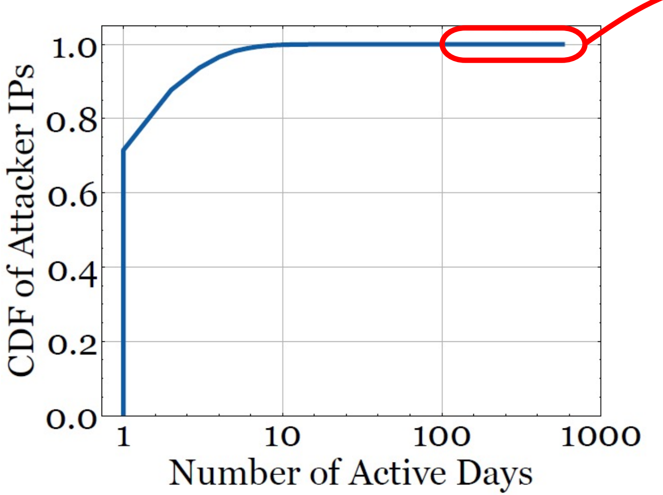
True attack

0.01%

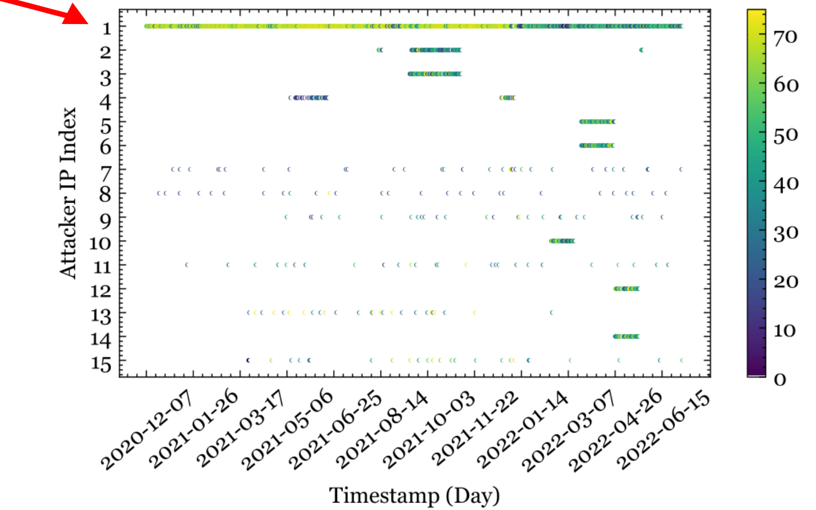
- Correctly fired alerts
- Have business-justified reasons

Labeled based on Black Hole Router

Understanding Attack Attempts



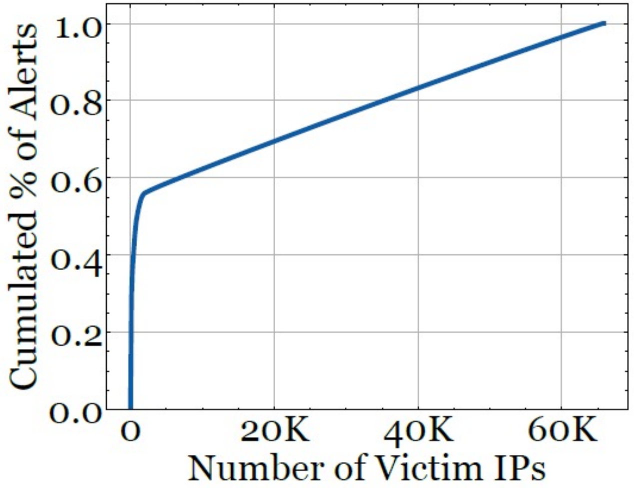
Analysis for Recurrent Attackers



Activities of the Most Active Attackers

1. Vast majority of attack attempts are short-lived
2. Persistent attack attempts exist (0.3%, 28 IPs)

Understanding Attack Attempts



Alert Distribution of Victims

- Small portion of victim IPs (2.3%) contribute 55% alerts

- The rest of the IPs (97.7%) contribute 45% alerts

Uneven distribution of security alerts among hosts may create challenges to develope **per-host** security prediction models

Linking Alerts with True Attacks

Linking Alerts with True Attacks

11 true attacks in four years. Mapping is mainly based on IPs and Time.

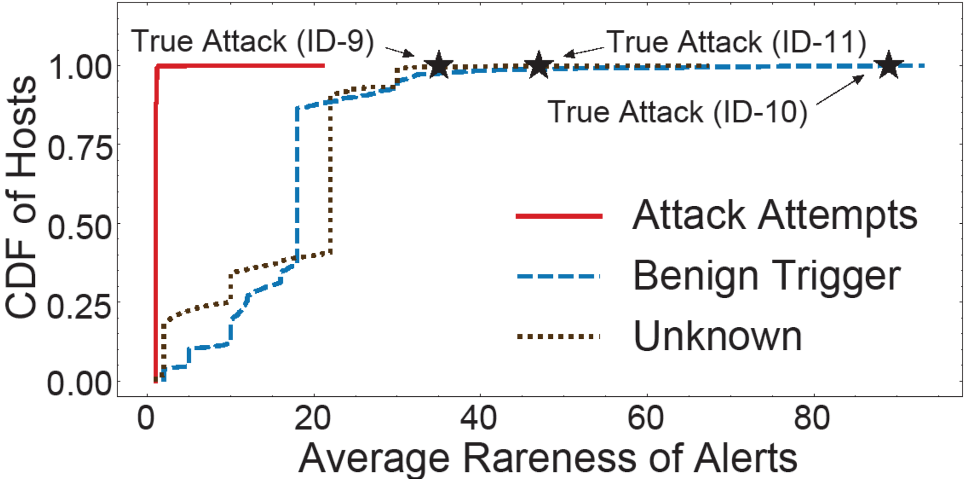
ID	Attack Info	# Related Alerts	# Total Alerts
2	Acct. compromised; scanning	22	163,878
5	0-day; scanning	18	124,154
9	Open port; DoS reflection	12	17,946
10	Postgres compromise; scanning	1101	17,839
11	Internal account crypto mining	1	24,921

Two other incident reports suggest that Zeek alerts were triggered but the logs are missing. So 7/11 of true attacks have triggered alerts.

1. The mapping between alerts and attacks took significant extra efforts
2. While excessive alerts are problematic, false negatives are concerning

Possible Direction: Abnormal Alert Patterns

- Rareness for one kind alert is its ranking of the frequency
- Final rareness score is the sum of the score of each unique alert on the host in a day
- All three true attacks are located in the outlier area



There is an opportunity to identify and prioritize those that indicate true attacks based on abnormal alert patterns (rare combinations of alerts)

Takeaways and Recommendations

Network intrusion detection:

1. “Attack attempts” and “benign triggers” should be distinguished in the benchmark dataset construction and evaluation process.

Excessive alerts in SOC:

1. To speed up the post-attack investigation and improve the efficiency of analysts, SOCs need efficient ways to store, link, and query different logs.
2. For threat response, automations are needed to handle alerts, especially during off-hours when human analysts have limited availability.