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

Supplementary Materials

Limin Yang[1*], Zhi Chen[1*], Chenkai Wang[1], Zhenning Zhang[1], Sushruth Booma [1], Phuong Cao[2],
Constantin Adam[3], Alexander Withers[2], Zbigniew Kalbarczyk[1], Ravishankar K. Iyer[1], Gang Wang[1]
[1]University of Illinois Urbana-Champaign [2]NCSA [3]IBM Research

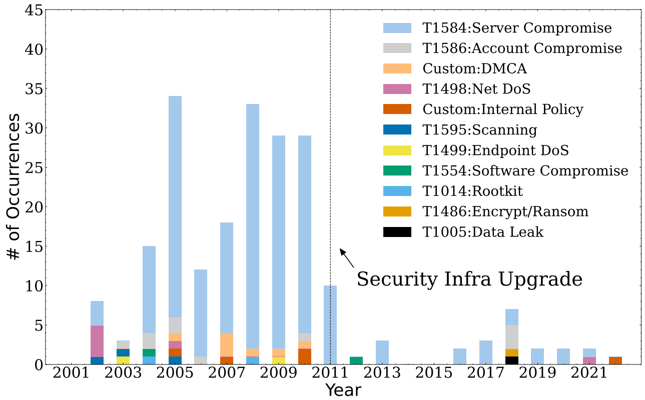


Figure 1: Occurrences of MITRE Technique IDs per year related to attack consequences.

1 Detailed Analysis of Incident Reports

Among the 227 true attacks, we can assign MITRE techniques/tags to 220 attacks (Table 1). For the 7 remaining attacks, we cannot find a MITRE technique that accurately describes them, and thus we create two custom tags, namely "DMCA hosting" and "internal policy violations." Figure 1 shows the occurrences of MITRE tags from 2002 to 2022 for the attack consequences. These attacks are related to internal servers hosting content that violates Digital Millennium Copyright Act (DMCA) [2], and actions that violate internal policies (e.g., using the computation platform for personal Bitcoin mining). We observe that most attacks (119, 67.2%) have two or more METRE tags. The detailed annotation results are presented in Table 1.

2 Geolocation and ASes of Attacker IPs

Figure 2 shows the heatmap of the attacker IP geolocation of BHR alerts (attack attempts) during 2020–2022. We looked up the IP geolocation using the Maxmind Geolocation Service [1]. While IP geolocation databases like Maxmind may contain mapping errors, the results at the country/region level are largely reliable [3, 4]. The top countries with the highest number of alerts associated with such IPs are China, the US, India, Brazil, and Germany. In total, there are 3.9 million BHR alerts triggered by external IPs that represent attack attempts. The top 5 countries have contributed to 54.83% of the total alerts. We further investigate the AS organizations of the attacker IPs and the top 10 ASes are listed in Table 2. We search the ASes based on the ASN dataset provided by Maxmind [1]. The topic 10 Ases contributed 39.60% of the external attack attempt alerts. Several large cloud providers such as Digital Ocean, Google Cloud, and Alibaba appear on the top 10 list.

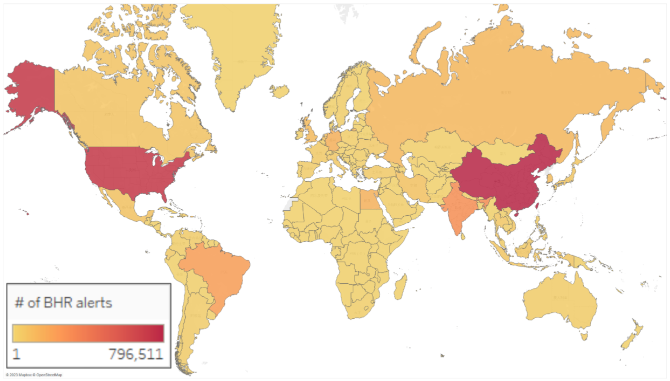


Figure 2: **Attacker IP Geolocation—** We show the geolocation distribution of attacker IPs identified from BHR alerts (attack attempts) during 2020–2022.

3 Suspicious Internal Scanning

We further analyze suspicious internal hosts that performed network scan activities. While the scanning alerts can be benign triggers (e.g., undocumented legitimate scanners), it is also possible to be scanning activities from compromised hosts. In the unknown category, we first identify internal IPs that have triggered network scanning alerts. In total, we

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

TID	Full Name	Short Name	Brief Description	Type	Attacks (%)
T1584	Compromise Infrastructure	Server Compromise	Servers compromised by the attacker	C	175 (77%)
T1078	Valid Accounts	User Account	Break-in with existing/valid user accounts	B	62 (27%)
T1587	Develop Capabilities	Exploits/Malware	Attack with exploits or malware	B	60 (26%)
T1190	Exploit Public-Facing Application	Pub. Service	Exploit public-facing services/websites	B	49 (22%)
T1110	Brute Force	PSW Guess.	Break-in with password guessing/stuffing	B	36 (16%)
T1204	User Execution	User Action	Users click on malicious links/files/images	B	19 (8%)
T1586	Compromise Accounts	Account Compromise	User accounts compromised by the attacker	C	10 (4%)
T1498	Network Denial of Service	Net DoS	Network DOS via flooding/reflection	C	6 (3%)
T1595	Active Scanning	Scanning	Attacker scans elsewhere after break-in	C	4 (2%)
T1566	Phishing	Phishing	Attacker runs phishing attacks	B	4 (2%)
T1014	Rootkit	Rootkit	Use rootkits to hide the presence of programs	C	4 (2%)
T1554	Compromise Client Software Binary	Software Compromise	Attacker modifies existing binaries after break-in	C	2 (1%)
T1499	Endpoint Denial of Service	Endpoint DoS	Block the availability of hosts/services	C	2 (1%)
T1056	Input Capture	Input Capture	Capture user input to obtain credentials	B	2 (1%)
T1005	Data from Local System	Data Leak	Attacker leaks/steals private data	C	2 (1%)
T1486	Data Encrypted for Impact	Encrypt/Ransom	Encrypt data to demand ransom	C	1 (.5%)
Custom	DMCA Hosting	DMCA	Hosting DMCA content in the target system	C	7 (3%)
Custom	Internal Policy Violation	Internal Policy	Violation of internal policies	C	5 (2%)

Table 1: **Type of Attacks**—MITRE techniques were used to annotate the true attacks' incident reports (227). For each technique, we provide the technique ID (TID) and its official name, followed by a short name and a brief description. We also classify it as either "B" (break-in method) or "C" (consequence). The bottom two custom tags are introduced by us since they cannot be annotated by the MITRE framework. The "Attacks" column does not add up to exact 227 (100%) because (a) one attack can have multiple B tags and/or multiple C tags; and (b) not all incident reports include information about break-in methods and consequences.

AS Org. Name	Country	# Alerts
DIGITALOCEAN-ASN	US	469,613
CHINANET-BACKBONE	China	260,471
CHINA UNICOM China169 Backbone	China	229,732
TE-AS	Egypt	131,007
National Internet Backbone	India	99,573
Hangzhou Alibaba Advertising Co.,Ltd.	China	96,807
CENTURYLINK-US-LEGACY-QWEST	US	87,244
TELEFONICA BRASIL S.A	Brazil	74,032
GOOGLE-CLOUD-PLATFORM	US	64,869
Shenzhen Tencent Computer Systems Company Limited	China	50,265

Table 2: **Top 10 ASes of Attacker IPs**— Top ASes based on the number of alerts of attack attempts during 2020–2022. There are 17,510 unique AS organizations in our data.

identified 384 internal IPs that triggered 1,442 scanning alerts. Table 4 shows the top 5 scanned ports that are documented in the address scanning alerts. Most scanning activities are checking on potential vulnerabilities.

To further characterize the scanning activities, we want to understand whether the internal hosts are scanning other internal hosts or hosts outside of the network. The scanning alerts are "aggregated" summaries where one alert can represent the scanning of thousands of hosts, but the list of scanned hosts is omitted from the alert. We follow the method described in the main paper to recover the scan destination IPs by looking up the connection log on the alert date. We only successfully recovered the destination IPs for 55 internal hosts. After confirming with the SOC security lead, the unrecoverable alerts were likely caused by a misconfigured router outside of the SOC. The records were filtered out from the connection logs.

Figure 3 shows a scatter plot for these 55 internal hosts in terms of their scanned internal and external IPs. The red dot represents a confirmed true attack where the scan is from a compromised host, which is used by the attacker to scan over 300K external hosts. Other internal IPs that scanned a large number of external hosts are suspicious candidates. On the other extreme, there are internal hosts that scan a large number of internal hosts. We select the top 5 hosts that scanned the most internal and external hosts respectively (10 in total) for SOC analysts for inspection. The finding is that all the top 5 external scans are from a student running a scraper for web crawling. Such behavior is within the acceptable scope, but the activity should have been communicated and documented. For the top 5 scans of internal hosts, SOC analysts confirm four of them are indeed legitimate internal scanners (set up on other hosts instead of using the dedicated IPs). The remaining one was related to a private IP that belonged to a different unit of the organization. An inquiry was sent to confirm the nature of the scan but without response yet.

4 Additional Per-Host Alert Analysis

We present two additional behavioral metrics for per-host alerts in Figure 4. Figure 4a shows the distribution of the average number of daily alerts among the internal hosts and Figure 4b shows the average number of daily unique alert types. We observe that 45% or more hosts do not have benign triggers or unknown alerts while almost all the hosts have alerts on attack attempts. In comparison, the benign trigger category has a higher number of daily alerts per host, and more diverse alert types. Finally, we notice that true attacks (i.e.,

ID	Attack Description (MITRE TID)
1	Account compromised; collected more credentials (T1078, T1584, T1586)
2	Account compromised via weak password; attacker did massive scan and password guessing against external hosts (T1110, T1584)
3	Account compromised via weak password; meeting notes stolen (T1110, T1589, T1005)
4	Account compromised; sent spam emails (T1078, T1586)
5	0-day vulnerability of internal tool; attacker run network scan (T1190, T1584, T1587)
6	Account compromised via weak password; attacker then sent a large number of outbound SSH requests (T1078, T1584)
7	Account compromised; sent phishing emails (T1078, T1566)
8	Account compromised; installed crypto mining program (T1078, T1584)
9	Port 123 was left open; attacker run DoS reflection attacks (T1498)
10	Postgres compromise; attacker did massive scanning and password guessing to external hosts (T1190, T1584, T1587)
11	Internal account performs bitcoin mining (Custom)

Table 3: **Alert-Attack Association Results**—Description of true attacks between 2018 and 2022.

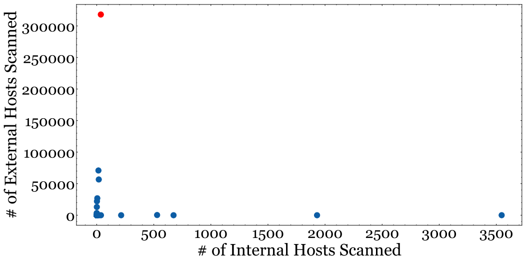


Figure 3: **Internal IPs' Scanning Activities**—We can recover the scan-destination IPs for 55 internal hosts. The red dot represents the internal host compromised by a true attack (the Postgres attack).

cases 9 and 11) may only have a small number of alerts (and alert types) if we only examine the general statistics. However, in practice, analysts may determine that certain alert types deserve a higher priority. Case 10 has an anomalously higher number of alerts and unique alert types for a given day.

References

[1] Maxmind. https://www.maxmind.com/, 2024.

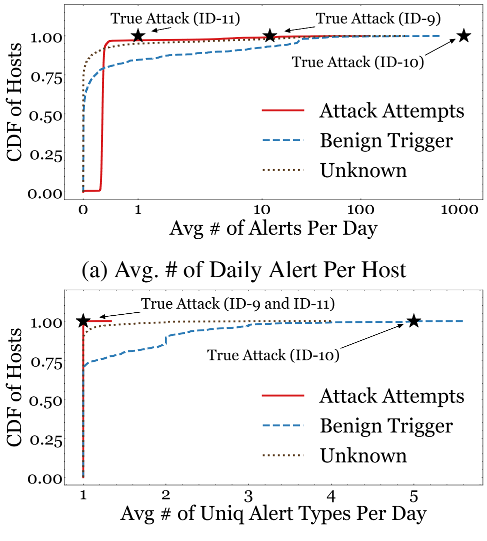
[2] DMCA. H.r.2281 - digital millennium copyright act. https://www.congress.gov/bill/105th-congress/house-bill/2281, 1998.

[3] Manaf Gharaibeh, Anant Shah, Bradley Huffaker, Han Zhang, Roya Ensafi, and Christos Papadopoulos. A look at router geolocation in public and commercial databases. In *Proc. of IMC*, 2017.

[4] Ioana Livadariu, Thomas Dreibholz, Anas Saeed Al-Selwi, Haakon Bryhni, Olav Lysne, Steinar Bjørnstad, and Ahmed Elmokashfi. On the accuracy of country-level ip geolocation. In *Proceedings of the applied networking research workshop*, 2020.

Scanned Port	Common Service	# Scanning Hosts	# Alerts
445	Microsoft-DS	264	817
80	HTTP	31	75
7	Echo service	2	64
1433	MS SQL server	24	62
22	SSH	8	46

Table 4: **Top 5 Scanned Ports by Internal Hosts**—Microsoft-DS is a network protocol used by Windows to share files, printers, and other resources on a local network. Leaving port 445 open can be vulnerable to a number of worms and trojans. Echo service can be abused by smurf/fraggle attacks.



(a) Avg. # of Daily Alert Per Host

(b) Uniq. Alert Type Per Day Per Host

Figure 4: **Per-Host Alert Characteristics**—We show the CDF plots of per-host behavioral metrics for internal hosts.