Mining Threat-intelligence from Billion-scale SSH Brute-Force Attacks

Yuming Wu¹§,
Phuong M. Cao¹§, Alexander Withers², Zbigniew T. Kalbarczyk¹, Ravishankar K. Iyer¹

¹ University of Illinois at Urbana-Champaign (UIUC)
² National Center for Supercomputing Applications (NCSA)

§ Joint first authors
Key Findings

• Over 70% are persistent attackers

• Identification of 7 SSH keys related to outdated vulnerabilities

• Globally distributed IPs massively spoofed over one million fake client versions

• Discovery of human-supervised versus fully automated botnets

Implications

• Discerning global coordination efforts in SSH key exploitation and client version spoofing

• Alerting cloud providers and IoT vendors regarding stolen SSH keys

• Deterring large-scale evasion techniques using anomaly detectors or rate limiters

• Preparing for resourceful and strategic human-supervised attacks
Analysis Workflow

Data Collection
~ 1,000 Days

Clickhouse Database

11 Billion Records

Summary Statistics

SSH Connection Forensics

Trend Anomalies

Key Matching & Client Version Investigation

Findings

1.7 M new SSH client versions burst in Aug 2018
70% of attacks were persistent

Evasion increased by 8,000x
Coordination accelerated by 50x
### Exploitation, Coordination, and Evasion - Leaked SSH Keys

<table>
<thead>
<tr>
<th>SSH Key (SHA256)</th>
<th>Key Owner</th>
<th>Appliance Type</th>
<th>Public Disclosure Year</th>
<th>1st Attack Attempt Year</th>
<th>Username</th>
</tr>
</thead>
<tbody>
<tr>
<td>1M4Rz...qu0ZA</td>
<td>Vagrant</td>
<td>Base box for development environments</td>
<td>2010</td>
<td></td>
<td>root</td>
</tr>
<tr>
<td>9prMb...Ghro4</td>
<td>F5</td>
<td>BigIP appliances</td>
<td>2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEc4H...UfTww</td>
<td>Loadbalancer</td>
<td>Virtual load balancer</td>
<td>2014</td>
<td>2018</td>
<td></td>
</tr>
<tr>
<td>VtjqZ...PiQpc</td>
<td>Quantum</td>
<td>Virtual deduplication backup appliance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/LJp6...POCc0</td>
<td>Array Networks</td>
<td>Virtual application delivery controllers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z+q4X...8kIXM</td>
<td>Ceragon</td>
<td>IP traffic router</td>
<td>2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f+1oG...zEDhc</td>
<td>VMware</td>
<td>Data Protection appliances</td>
<td>2016</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- We identified 7 keys related to outdated vulnerabilities – indicating some devices still unpatched

**Attackers had adequate details (i.e., credentials) about relevant vulnerabilities that were related with these 7 keys, when plotting the targeted attacks**
### Exploitation, Coordination, and Evasion

#### Leaked SSH Keys: Attack Origins

<table>
<thead>
<tr>
<th>Autonomous System</th>
<th>Client Version [SSH-2.0-]</th>
<th>SSH Key (SHA256) &amp; Key Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google LLC</td>
<td>libssh_0.7.0</td>
<td>1M4Rz... Vagrant</td>
</tr>
<tr>
<td>Charter Communications</td>
<td>Ruby/Net::SSH...</td>
<td>9prMb... F5</td>
</tr>
<tr>
<td>Portlane</td>
<td>libssh-0.6.1</td>
<td>MEc4H... Loadbalancer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VtjqZ... Quantum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>/JlP6... Array Networks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Z+q4X... Ceragon</td>
</tr>
<tr>
<td></td>
<td></td>
<td>f+1oG... VMware</td>
</tr>
</tbody>
</table>

**Speculation:** Attackers were rapidly switching ASes to evade detection, and possibly switching targets.

- Attackers leveraged Google LLC (Google), Charter Communications, and Portlane to exploit the 7 identified leaked keys
- Attackers from Google-registered IPs attempted all 7 keys with four other unknown keys on the same day
Exploitation, Coordination, and Evasion - Key-based Collaboration

• An SSH key was exploited by 20 countries
  • The globally coordinated botnet exploited a single SSH key 90 times within only 4 days
• The last key was persistently used one single country for 2,700 times spanning 5 months

The globally coordinated bot wrapped up its fruitless attacks and shifted targets 50× faster than the persistent, single-country botnet
Exploitation, Coordination, and Evasion

- **Client Version-based Collaboration and Evasion**

  - More than 1.7 million new client versions were spoofed in August alone
    - Only several hundred globally-distributed IPs were spoofing (e.g. SSH-2.0-OpenSSH_+qLfH)
  - Yet 90% IPs used only 1 client version
  - The top-spoofing IP advertised 400,000 unique client versions during its 200-hour attack campaign

_A globally-coordinated botnets were involved in forging a million permutations of client versions at high frequencies_

_Voids signature-based detectors_
Analysis Workflow

Data
- 11 Billion Records
- Clickhouse Database
- Data Collection ~ 1,000 Days

Methods
- Summary Statistics
- SSH Connection Forensics
- Attack Sources Analysis
- Trend Anomalies
- Key Matching & Client Version Investigation
- Tail Analysis

Findings
- 1.7 M new SSH client versions burst in Aug 2018
- 70% of attacks were persistent
- Evasion increased by 8,000x
- Coordination accelerated by 50x
- Human-supervised botnet used sophisticated credential-guessing techniques
Human-supervised Attack Techniques
- Data-driven Methodology

Purpose: identify evidence of human attackers

- Time zone and duration selection
- Ratio: average weekday to weekend attempt computation for each IP
- Tail analysis of ratio distribution
- All IPs in the tail present similar activity patterns; used the same group of credentials; came from the same /8 subnet
- Periodic variations with decreasing activities on weekends (especially Sundays)
### Human-supervised versus Fully Automated Bots

<table>
<thead>
<tr>
<th>Type</th>
<th>Illustration of Daily Attempts</th>
<th>List of Unique Client Versions</th>
<th>List of Unique Username(s)</th>
<th># Unique Passwords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human-supervised</td>
<td><img src="image" alt="Human-supervised botnet attempts graph" /></td>
<td>PuTTY</td>
<td>root</td>
<td>35,952</td>
</tr>
<tr>
<td></td>
<td>[May 27 – July 21, 2019 (8 weeks)]</td>
<td>OpenSSH_5.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>OpenSSH_6.2p2...</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>nsssh2_4.0...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fully automated</td>
<td><img src="image" alt="Fully automated botnet attempts graph" /></td>
<td>sshlib-0.1</td>
<td>root, user, admin, ubuntu,</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>usuario, pi, supervisor,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>service, mother</td>
<td></td>
</tr>
</tbody>
</table>

OpenSSH_6.2p2...refers to OpenSSH_6.2p2 Ubuntu-6;
nsssh2_4.0...refers to nsssh2_4.0 NetSarang Computer, Inc.

Human-supervised botnet is more resourceful, ambitious, and strategic than full automated one
Conclusions

• Investigated a broad scope of SSH attack strategies

• Discovered large-scale, persistent, and evasion attacks

• Contributed a scientific data-driven approach to differentiate between human-supervised and fully automated botnet

Future

• Landscape of unidentified, unknown SSH keys

• Resourceful attackers with relatively large number of legitimate client versions

• Threat intelligence sharing across peer sites with preservation of privacy
Thank you!
Acknowledgements

• SDAIA: https://wiki.ncsa.illinois.edu/display/cybersec/SDAIA

• NSF Grant: CICI: Secure Data Architecture: Shared Intelligence Platform for Protecting our National Cyberinfrastructure. Award Number: 1547249

• NSF Grant: SI2-SSE: AttackTagger: Early Threat Detection for Scientific Cyberinfrastructure. Award Number: 1535070

• DEPEND group Symphony Cluster
References