Counting Outdated Honeypots:
Legal and Useful

Alexander Vetterl*, Richard Clayton* and Ian Walden‡

*University of Cambridge, ‡Queen Mary University of London

4th International Workshop on Traffic Measurements for Cybersecurity — May 23, 2019
Introduction

Honeypot: A resource whose value is being attacked or compromised

- Honeypots have been focused for years on the monitoring of human activity
- Adversaries attempt to distinguish honeypots by executing commands
- Honeypots continuously fix commands to be “more like bash”

Cowrie - commands implemented
How we currently build SSH honeypots

1. Find a library that implements the desired protocol (e.g. TwistedConch for SSH)
2. Write the Python program to be “just like bash”
3. Fix identity strings, error messages etc. to be “just like OpenSSH”

```
def _unsupportedVersionReceived(self, remoteVersion):
    """
    Change message to be like OpenSSH
    """
    self.transport.write(b'Protocol major versions differ.\n')
```

<table>
<thead>
<tr>
<th>RFCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenSSH</td>
</tr>
<tr>
<td>TwistedConch</td>
</tr>
<tr>
<td>sshd</td>
</tr>
<tr>
<td>Cowrie</td>
</tr>
<tr>
<td>bash</td>
</tr>
</tbody>
</table>

Problem: There are lot of subtle differences between TwistedConch and OpenSSH...
Fingerprinting honeypots at internet scale

We send probes to various different implementations
- SSH honeypots (Cowrie/Kippo)
- OpenSSH, TwistedConch

We find ‘the’ probe that results in the most distinctive response across all implementations and perform Internet wide scans

<table>
<thead>
<tr>
<th></th>
<th>Date</th>
<th>#ACKs</th>
<th>Sum</th>
<th>Kippo</th>
<th>Cowrie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scan 1 (SSH)</td>
<td>2017-09</td>
<td>18,196k</td>
<td>2844</td>
<td>906</td>
<td>1938</td>
</tr>
<tr>
<td>Scan 2 (SSH)</td>
<td>2018-01</td>
<td>20,586k</td>
<td>2779</td>
<td>758</td>
<td>2021</td>
</tr>
</tbody>
</table>

Login to get more details, but...

Paper was rejected due to ethical concerns

“This paper was rejected due to ethical concerns.

[...]

It was pointed out that these attempts are likely a violation of US law, especially the Computer Fraud and Abuse Act which prohibits accessing a computer without authorization.

The PC recommends to consult with a lawyer before trying to publish this paper a different venue.”

Summary of the PC discussion
Uniformed legislation for unauthorised access

Convention on Cybercrime ("Budapest Convention")

- States must have laws that forbid access ‘without right’
- Ratified by 62 states

EU Directive 2013/40/EU Article 3

- ‘Member states [...] shall ensure that, when committed intentionally, the access without right, [...] is punishable as a criminal offence where committed by infringing a security measure, at least for cases which are not minor.’
Legislation in the UK and USA

**UK: Computer Misuse Act 1990**

Access of any kind by any person to any program or data held in a computer is unauthorised if –

a) [...] 

b) he does not have consent to access by him of the kind in question to the program or data.

**USA: Fraud and Abuse Act 1986**

‘Whoever [...] intentionally accesses a computer without authorization [...] and thereby obtains [...] information from any protected computer.’

**Factors to consider**

– No consent to access [by him] of the ‘kind in question’
– Overcome some form of security mechanism
– Offences which are not minor
Legislation in the context of honeypots

In general much authorisation is implicit

- Devices and services intentionally connected to the Internet
- Web servers/ftp servers with the username ‘anonymous’ and email address as password

Our access was not unauthorised because the controller of the honeypot has -

- intentionally made available a (vulnerable) system and
- implicitly permits the access of the ‘kind of question’
Ethical considerations

- We followed our institution’s ethical research policy
- We used the exclusion list maintained by DNS-OARC
- We notified all local CERTs of our scans/actions
- We respected requests to be excluded from further scanning
- We started and ended every SSH session with an explanation
- We notified the relevant honeypot and library developers of our findings
Results - Authentication configuration (1/2)

- We used the username root and initially 6 passwords, later 500 passwords
- We managed to successfully log in to about 70% of the honeypots

<table>
<thead>
<tr>
<th>Outcome</th>
<th>6 passwords</th>
<th>500 passwords</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scan 1: 2017-03</td>
<td>Scan 1: 2017-03</td>
</tr>
<tr>
<td>successful login</td>
<td>859 (70.9%)</td>
<td>794 (65.5%)</td>
</tr>
<tr>
<td>all passwords failed</td>
<td>110 (9.1%)</td>
<td>136 (11.2%)</td>
</tr>
<tr>
<td>connection timed out</td>
<td>49 (4.0%)</td>
<td>110 (9.1%)</td>
</tr>
<tr>
<td>other errors</td>
<td>194 (16.0%)</td>
<td>172 (14.2%)</td>
</tr>
</tbody>
</table>
Results - Authentication configuration (2/2)

- Using 500 passwords is not better than 6 passwords
- About 11% of honeypot operators do not allow logins

<table>
<thead>
<tr>
<th>Outcome</th>
<th>500 passwords Scan 2: 2017-06</th>
<th>500 passwords Scan 3: 2017-09</th>
<th>500 passwords Scan 4: 2018-01</th>
</tr>
</thead>
<tbody>
<tr>
<td>successful login</td>
<td>1165 (66.7%)</td>
<td>1347 (69.5%)</td>
<td>1578 (78.1%)</td>
</tr>
<tr>
<td>all passwords failed</td>
<td>187 (10.7%)</td>
<td>195 (10.1%)</td>
<td>223 (11.0%)</td>
</tr>
<tr>
<td>connection timed out</td>
<td>41 (2.4%)</td>
<td>43 (2.2%)</td>
<td>7 (0.3%)</td>
</tr>
<tr>
<td>other errors</td>
<td>354 (20.2%)</td>
<td>353 (18.2%)</td>
<td>213 (10.6%)</td>
</tr>
</tbody>
</table>
Revision history for command selection

- We looked for commands in the revision history (uname -a, tftp)

**Cowrie < 2016-11-02**

```
root@svr04:~# tftp
-bash: tftp: command not found
root@svr04:~#
```

**Cowrie ≥ 2016-11-02**

```
root@svr04:~# tftp
root@svr04:~#
```
Results - Counting outdated honeypots (1/2)

- High market share for Kippo, which had last been updated years earlier
- Only ~25% of honeypots were up-to-date

<table>
<thead>
<tr>
<th></th>
<th>Scan 1: 2017-03</th>
<th></th>
<th>Scan 2: 2017-06</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kippo &lt; 2014-05-28</td>
<td>1384 (42.5%)</td>
<td></td>
<td>1519 (42.8%)</td>
<td></td>
</tr>
<tr>
<td>Kippo &lt; 2015-05-24</td>
<td>659 (20.3%)</td>
<td></td>
<td>285 (8.0%)</td>
<td></td>
</tr>
<tr>
<td>Cowrie &lt; 2016-09-05</td>
<td>385 (11.8%)</td>
<td></td>
<td>392 (11.0%)</td>
<td></td>
</tr>
<tr>
<td>Cowrie &lt; 2016-11-02</td>
<td>—</td>
<td></td>
<td>556 (15.7%)</td>
<td></td>
</tr>
<tr>
<td>Cowrie &lt; 2017-06-06</td>
<td>—</td>
<td></td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Cowrie ≤ date of scan</td>
<td>827 (25.4%)</td>
<td></td>
<td>799 (22.5%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3255</td>
<td></td>
<td>3551</td>
<td></td>
</tr>
</tbody>
</table>
Results - Counting outdated honeypots (2/2)

- The number of SSH honeypots is slightly declining (-14.6%)
- Kippo is slowly being replaced by Cowrie

<table>
<thead>
<tr>
<th></th>
<th>Scan 3: 2017-09</th>
<th></th>
<th>Scan 4: 2018-01</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kippo &lt; 2014-05-28</td>
<td>695 (24.4%)</td>
<td></td>
<td>546 (19.6%)</td>
<td></td>
</tr>
<tr>
<td>Kippo &lt; 2015-05-24</td>
<td>211 (7.4%)</td>
<td></td>
<td>212 (7.6%)</td>
<td></td>
</tr>
<tr>
<td>Cowrie &lt; 2016-09-05</td>
<td>134 (4.7%)</td>
<td></td>
<td>147 (5.3%)</td>
<td></td>
</tr>
<tr>
<td>Cowrie &lt; 2016-11-02</td>
<td>360 (12.7%)</td>
<td></td>
<td>422 (15.2%)</td>
<td></td>
</tr>
<tr>
<td>Cowrie &lt; 2017-06-06</td>
<td>734 (25.8%)</td>
<td></td>
<td>381 (13.7%)</td>
<td></td>
</tr>
<tr>
<td>Cowrie ≤ date of scan</td>
<td>710 (25.0%)</td>
<td></td>
<td>1071 (38.6%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2844</td>
<td></td>
<td>2779</td>
<td></td>
</tr>
</tbody>
</table>
Results - Set-up options

SSH Version strings

- 61 different version strings
- 72% use the default - SSH-2.0-OpenSSH_6.0p1 Debian-4+deb7u2

Hostname (uname -a)

- 3.3% use the default - svr04
- debnfwmgmt-02 is used for 296 honeypots (14.6%)
  - This is the default hostname for Cowrie when it is used in T-Pot
  - T-Pot is a popular docker container and combines 16 honeypots
  - T-Pot has a significant market share
Conclusion

Many honeypots are outdated and not looked after
  – Update your honeypots!

Honeypot operators do not change default configurations
  – Usernames/passwords, hostnames, SSH version strings etc.

Our access to honeypots was not unauthorized
  – Detailed legal analysis to enable more research in this area
  – Lessons learned: Provide not only an ethical justification, but also some legal analysis
Q & A

Alexander Vetterl
alexander.vetterl@cl.cam.ac.uk