Longitudinal Analysis of SSH Honeypot Logs

Security and Privacy Lab, Department of Computer Science

Dominic Rudigier, 11832156 | Supervisor: Maximilian Hils

Motivation

Motivation

Motivation

Honeypot:

- Mimics easy target and attracts attackers
- Generates logs about connection data (commands, files uploaded)
- Gained information can be used to improve systems
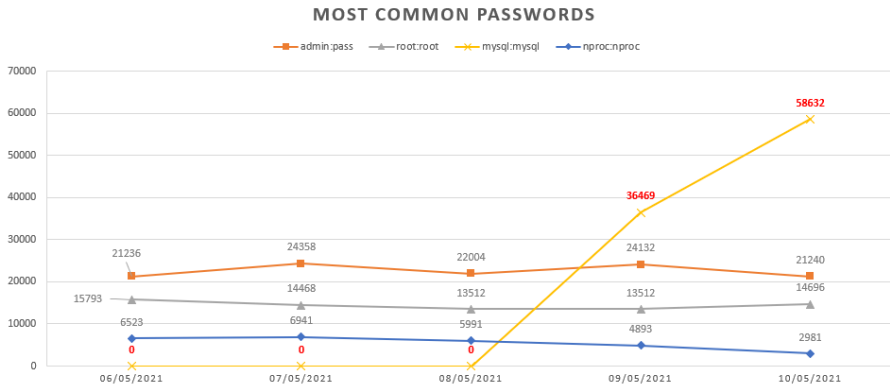
Motivation

Honeypot:

- Mimics easy target and attracts attackers
- Generates logs about connection data (commands, files uploaded)
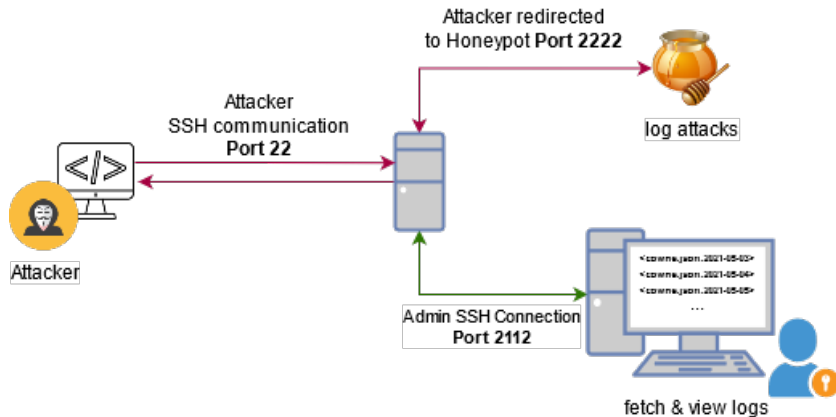- Gained information can be used to improve systems

Example

1. Hardcoded credentials in software
2. Attacker found out somehow
3. Analyzing honeypot logs shows that hacker knows
4. Operator can react to vulnerability

Motivation

MOST COMMON PASSWORDS

universität
innsbruck

Motivation

Status quo

Cowrie:
- SSH and Telnet honeypot
- Different log formats for connection data (JSON, UML, MongoDB,..)

Problem

- Malware has become more intelligent
- e.g. Aisuru detects Cowrie honeypots
  - existence of "@LocalHost:]"
  - existence of a service, started on Jun 22nd, or Jun 23rd
  - user exists on the device named "richard"
- Improve honeypot configurations
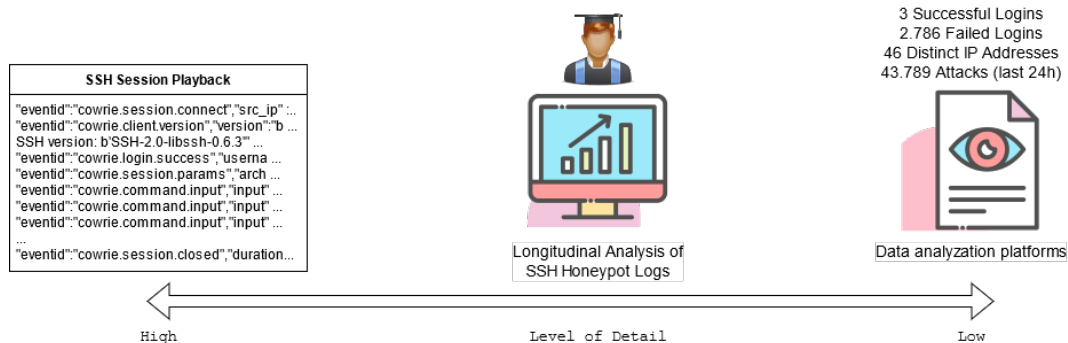
Status quo

Already existing:

- SSH Session Playback (single connection)
- Data analyzation platforms (aggregated statistics)

Problem

- Providers
  - Multiple thousands of honeypots
  - Each honeypot logs thousands of attacks per hour
  - Too much information gathered, hard to handle.

Thesis contribution

SSH Session Playback

"eventid":"cowrie.session.connect","src_ip" :...
"eventid":"cowrie.client.version","version":"b ...
SSH version: b'SSH-2.0-libssh-0.6.3'" ...
"eventid":"cowrie.login.success","userna ...
"eventid":"cowrie.session.params","arch ...
"eventid":"cowrie.command.input","input" ...
"eventid":"cowrie.command.input","input" ...
"eventid":"cowrie.command.input","input" ...
...
"eventid":"cowrie.session.closed","duration... ...

Longitudinal Analysis of
SSH Honeypot Logs

3 Successful Logins
2.786 Failed Logins
46 Distinct IP Addresses
43.789 Attacks (last 24h)

Data analyzation platforms

High   Level of Detail   Low

Thesis contribution

Longitudinal analysis:
- Research design that involves repeated observations of the same variables.
- Research attacker behaviour over time.

Batch-processing of log files

1. Cowrie generates log files **<cowrie.json.2021-05-01, 100 to 200 MB>**
   - per honeypot (instance)
   - per day (multiple)
   - 1k honeypots / 30 days = 3-6 TB log data per month
2. Batch-process using MapReduce model with Python
3. Visualize (Python or Flask + ReactJS)

Batch-processing of log files
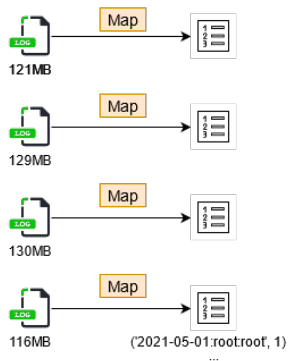
Cowrie Logs per Day

121MB

129MB

130MB

116MB

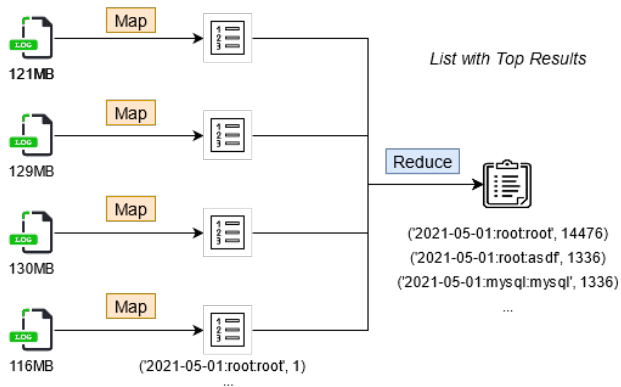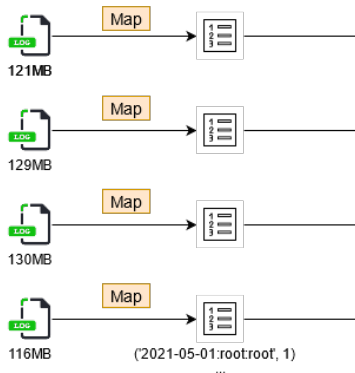| cowrie.json.2021-05-01 |
| --- |
| {"eventid":"cowrie.session.connect","src_ip.. |
| {"eventid":"cowrie.client.version","version.. |
| {"eventid":"cowrie.login.failed","usernamem.. |
| {"eventid":"cowrie.session.closed","duration.. |
| {"eventid":"cowrie.command.input","input.. |

Batch-processing of log files

Cowrie Logs per Day

121MB → Map →

129MB → Map →

130MB → Map →

116MB → Map → ('2021-05-01:root:root', 1)
...

Batch-processing of log files

Cowrie Logs per Day

121MB

Map

129MB

Map

130MB

Map

116MB

Map

('2021-05-01:root:root', 1)
...

List with Top Results

Reduce

('2021-05-01:root:root', 14476)
('2021-05-01:root:asdf', 1336)
('2021-05-01:mysql:mysql', 1336)
...

Batch-processing of log files

Cowrie Logs per Day

Map
121MB

Map
129MB

Map
130MB

Map
116MB

('2021-05-01:root:root', 1)
...

Reduce

List with Top Results

('2021-05-01:root:root', 14476)
('2021-05-01:root:asdf', 1336)
('2021-05-01:mysql:mysql', 1336)
...

Visualization

root:root up 18% (24h)
asdf:asdf up 36% (7d)
mysql:mysql up 12% (1 mo)
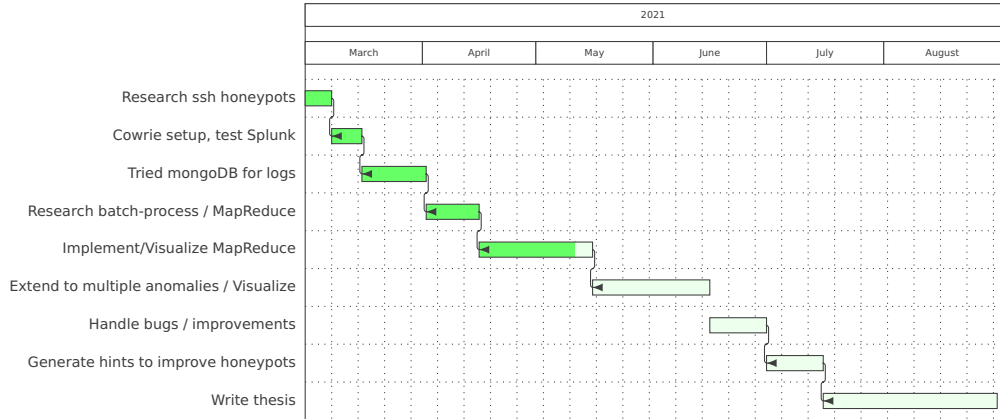
Goal

Extract information:

1 Detect changes in attacker behavior over time

Changes over time might indicate new vulnerabilities

- User:Password combination changes
- Commands executed before disconnect
- Command quantity changes
- All log anomalies not previously shown up

2 Visualize

3 Find ways to improve honeypot configurations

Timeline

References I

[Cab+19]   W. Cabral et al. "Review and Analysis of Cowrie Artefacts and Their Potential to be Used Deceptively". In: *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*. 2019, pp. 166–171. DOI: 10.1109/CSCI49370.2019.00035.

[DG04]   Jeffrey Dean and Sanjay Ghemawat. "MapReduce: Simplified Data Processing on Large Clusters". In: *OSDI'04: Sixth Symposium on Operating System Design and Implementation*. San Francisco, CA, 2004, pp. 137–150.

References II

[GGL03]   Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung. "The Google file system". In: *Proceedings of the nineteenth ACM symposium on Operating systems principles*. 2003, pp. 29–43.

[KJE19]   S. Kumar, B. Janet, and R. Eswari. "Multi Platform Honeypot for Generation of Cyber Threat Intelligence". In: *2019 IEEE 9th International Conference on Advanced Computing (IACC)*. 2019, pp. 25–29. DOI: 10.1109/IACC48062.2019.8971584.

References III

[KS18]   A. Kyriakou and N. Sklavos. "Container-Based Honeypot Deployment
           for the Analysis of Malicious Activity". In: *2018 Global Information
           Infrastructure and Networking Symposium (GIIS)*. 2018, pp. 1–4. DOI:
           `10.1109/GIIS.2018.8635778`.

[San20]   Chris Sanders. *Intrusion Detection Honeypots, Detection Through
           Deception. Detection Through Deception*. Applied Network Defense,
           2020. ISBN: 978-1735188300.

Thank you for your attention!

Dominic Rudigier, 11832156 | Supervisor: Maximilian Hils

Appendix 1

https://www.sicherheitstacho.eu/start/main
https://www.avira.com/en/blog/
new-mirai-variant-aisuru-detects-cowrie-opensource-honeypots