

INTRO TO LOG ANALYSIS

AIM:

Gain a foundational understanding of log analysis in cybersecurity by learning to investigate events using log data from various systems. This includes identifying anomalies and suspicious behavior using command-line tools, regular expressions, and platforms like CyberChef.

PROCEDURE:

1. Begin with the theory of log types, timelines, and threat indicators.
2. Use Linux CLI tools like ``cut``, ``awk``, ``grep``, and ``uniq`` for log filtering.
3. Decode obfuscated payloads with CyberChef (e.g., Base64, MACs).
4. Use regex patterns to extract specific values from logs.
5. Understand the function of Logstash Grok filters.
6. Write and understand detection rules using YARA and Sigma YAML format.

TASK 1 – INTRODUCTION

- Introduces log analysis and its role in cybersecurity operations.
- Explains how logs help detect and investigate malicious activity.
- Describes various log types like system, application, and security logs.
- Sets the foundation for working with forensic tools and log files.
- Highlights how log trails are essential in incident response

- Encourages a mindset of curiosity and pattern recognition

Answer the questions below

I'm ready to proceed!

No answer needed

✓ Correct Answer

TASK 2 – TYPES OF LOGS

- Covers various types of logs used in analysis, such as Apache, DNS, Syslog.
- Explains the structure and purpose of each log type.
- Helps identify which logs are useful for which kind of threat or anomaly.
- Emphasizes reading timestamps, IPs, and method/status fields.
- Reinforces log relevance in real-world investigations.
- Forms the basis for choosing the right log during triage.

Answer the questions below

I understand the basics of logs and I'm ready to proceed!

No answer needed

✓ Correct Answer

TASK 3 – INVESTIGATION THEORY

- Introduces the concept of timelines and event correlation.
- Defines a "Super Timeline" for cross-system analysis.
- Discusses threat indicators like file hashes (MD5).
- Covers visualizing events and identifying intrusion patterns.
- Questions help reinforce understanding of analysis theory.
- Equips users to think systematically during log review.

Answer the questions below

What's the term for a consolidated chronological view of logged events from diverse sources, often used in log analysis and digital forensics?

Super Timeline

✓ Correct Answer

Which threat intelligence indicator would `5b31f93c09ad1d065c0491b764d04933` and `763f8bdbc98d105a8e82f36157e98bbe` be classified as?

File Hashes

✓ Correct Answer

TASK 4 – DETECTION ENGINEERING

- Focuses on identifying suspicious behavior in logs.
- Highlights default log locations, like `/var/log/nginx/access.log``.
- Teaches detection of encoded attacks like path traversal.
- Shows how to decode ``%2E%2E/`` and other encoded threats.
- Builds awareness of signature-based log indicators.
- Practical examples prepare users for real detection tasks.

Answer the questions below

What is the default file path to view logs regarding HTTP requests on an Nginx server?

/var/log/nginx/access.log

✓ Correct Answer

A log entry containing `%2E%2E%2F%2E%2Fproc%2Fself%2Fenviron` was identified. What kind of attack might this infer?

Path Traversal

✓ Correct Answer

TASK 5 – AUTOMATED VS. MANUAL ANALYSIS

- Compares automated log parsing with manual investigation.
- Shows when to use tools vs. human-led judgment.
- Demonstrates strengths and limits of both approaches.
- Promotes hybrid usage of automated detection and human insight.
- Reinforces how automation saves time, but humans catch context.
- Simple Q&A makes the concept clear and applicable.

Answer the questions below

A log file is processed by a tool which returns an output. What form of analysis is this?

Automated

✓ Correct Answer

An analyst opens a log file and searches for events. What form of analysis is this?

Manual

✓ Correct Answer

TASK 6 – LOG ANALYSIS TOOLS: COMMAND LINE

- Uses CLI tools like `cut`, `awk`, `sort`, `uniq`, and `wc`.
- Extracts URLs, IPs, and counts response codes in Apache logs.
- Helps identify most active IPs or anomalies in logs.
- Tasks include timestamp extraction, pattern filtering.
- Encourages hands-on practice and efficient log handling.
- Reinforces Linux CLI as a primary skill for analysts.

Answer the questions below

Use `cut` on the `apache.log` file to return only the URLs. What is the flag that is returned in one of the unique entries?

✓ Correct Answer

💡 Hint

In the `apache.log` file, how many total HTTP 200 responses were logged?

✓ Correct Answer

💡 Hint

In the `apache.log` file, which IP address generated the most traffic?

✓ Correct Answer

💡 Hint

What is the complete timestamp of the entry where `110.122.65.76` accessed `/login.php`?

✓ Correct Answer

💡 Hint

TASK 7 – LOG ANALYSIS TOOLS: REGULAR EXPRESSIONS

- Introduces regex for log pattern extraction and filtering.
- Teaches matching ranges (e.g., `post=2[2-6]`) and wildcards.
- Shows how regex simplifies locating key data entries.
- Explains the Grok plugin for parsing unstructured logs.

Answer the questions below

How would you modify the original `grep` pattern above to match blog posts with an ID between 20-29?

`post=2[0-9]`

✓ Correct Answer

🔑 Hint

What is the name of the filter plugin used in Logstash to parse unstructured log data?

Grok

✓ Correct Answer

- Forms the base for automation in SIEM log parsing.
- Builds muscle memory in log filtering precision.

TASK 8 – LOG ANALYSIS TOOLS: CYBERCHEF

- Demonstrates use of CyberChef for IP/MAC extraction and decoding.
- Shows regex matching for IPv4 and Base64 decoding.
- Uses filters to refine large datasets into actionable data.
- Tasks include decoding embedded flags and extracting patterns.
- Reinforces visual/logical chaining of transformations.
- Makes advanced parsing accessible for beginners.

Answer the questions below

Locate the "loganalysis.zip" file under `/root/Rooms/introloganalysis/task8` and extract the contents.

No answer needed ✓ Correct Answer

Upload the log file named "access.log" to CyberChef. Use regex to list all of the IP addresses. What is the full IP address beginning in 212?

212.14.17.145 ✓ Correct Answer

Using the same log file from Question #2, a request was made that is encoded in base64. What is the decoded value?

THM(CYBERCHEF_WIZARD) ✓ Correct Answer

Using CyberChef, decode the file named "encodedflag.txt" and use regex to extract by MAC address. What is the extracted value?

08-2E-9A-4B-7F-61 ✓ Correct Answer

TASK 9 – LOG ANALYSIS TOOLS: YARA AND SIGMA

- Introduces detection rule writing with YARA (malware) and Sigma (logs).
- Explains syntax like `rule` (YARA) and `title` (Sigma YAML).
- Demonstrates how Sigma helps standardize detection across platforms.
- Teaches rule readability and structure in threat detection.
- Builds a bridge between manual detection and automated alerts.
- Finalizes the room by integrating rules into practical use.

Answer the questions below

What languages does Sigma use?

YAML

✓ Correct Answer

What keyword is used to denote the "title" of a Sigma rule?

title

✓ Correct Answer

What keyword is used to denote the "name" of a rule in YARA?

rule

✓ Correct Answer

RESULT:

Successfully understood the principles of log analysis, practiced log filtering and decoding, and applied detection rule writing using industry tools, laying a strong foundation for real-world security operations.