Log File Analyzer for Intrusion Detection - Internship Project Report

Project Title

Log File Analyzer for Intrusion Detection

Intern Details

Name: Mohammed Anas C
Internship Phase: Major Project (Phase 2)
Domain: Cybersecurity (Log Analysis / Threat Detection)
Platform: Kali Linux

Abstract

In today's threat landscape, log file analysis plays a crucial role in identifying and mitigating security incidents. This project presents a Python-based solution to detect intrusion patterns in system and web server logs, including brute-force login attempts, DoS attacks, and unauthorized scanning. The tool is designed to parse log files, extract meaningful data, and generate alert reports and visualizations, aiding security teams in real-time decision-making.

Objective

To develop a lightweight intrusion detection system that:
- Parses Apache and SSH logs
- Detects brute-force login attempts, DoS attacks, and scanning activity
- Visualizes traffic patterns
- Cross-checks IP addresses with blacklists
- Exports incident reports

Tools & Technologies Used

Operating System: Kali Linux
Language: Python 3
Libraries: pandas, matplotlib, seaborn, re (regex), collections
IDE: Nano or VS Code
Log Files: access.log, auth.log

Implementation Steps

1. Environment Setup:
- Verified Python 3 and installed pip: sudo apt install python3 python3-pip
- Installed necessary libraries: pip3 install pandas matplotlib seaborn
2. Log File Preparation:
- Sample logs were created manually using nano for testing
3. Script Development: A modular script was written in Python to:
- Parse Apache logs using regex to extract IPs, timestamps, and request status.
- Parse SSH logs for failed password attempts.
- Count request volume per IP (DoS detection).
- Count repeated login failures from the same IP (brute-force).
- Export CSV reports and generate bar charts for visualization.
4. Visualization:
- The script generated a bar chart showing the top 5 IP addresses by access volume.
- Graph was saved as access_patterns.png.
5. Report Generation:
- The script automatically saved detection results into:

Results

Brute Force Report: brute_force_report.csv: IPs with >5 failed SSH logins

DoS Report: dos_report.csv: IPs with >10 Apache requests

Graph: access_patterns.png: Top 5 IPs traffic graph

Security Insight

This project demonstrates how simple log parsing with Python can uncover major signs of intrusion like brute-force and DoS attempts. When combined with visualization and automation, such tools significantly aid a SOC team in early threat detection.

Conclusion

The Log File Analyzer project effectively showcases the practical use of Python and regex in detecting security incidents from log files. With no external costs and complete open-source tools, it serves as a powerful educational and operational tool for cybersecurity analysts, especially in resource-limited environments.

NB: All code, logs, and output reports were created and tested on a Kali Linux machine without the use of any paid tools.