**Final Project Report**

**Automated Log File Analyzer for Security Auditing**

**Course: Information Security Audit System  
Project members: Qamber raza k21-3251 || Shafey Sidiqque 21k-  
Submission Date: 4/27/2025**

**Abstract**

The Automated Log File Analyzer is a lightweight and accessible security auditing tool developed to process and analyze system logs. It identifies potential security threats by detecting unusual access patterns, failed login attempts, and frequent IP connections. The system supports multiple log types, including Apache access logs, SSH authentication logs, and firewall logs. By automatically extracting key data such as IP addresses, timestamps, and HTTP status codes, the tool generates meaningful visual analytics through bar charts, making threat detection faster and more intuitive. Implemented on Google Collab, the system requires no external dependencies, ensuring easy accessibility for cybersecurity monitoring and audit control.

**Objectives**

* To automate the analysis of log files for security auditing.
* To detect unauthorized access attempts and suspicious activity.
* To visualize attack trends through bar charts and frequency graphs.
* To create an accessible, efficient, and lightweight auditing solution without external library requirements.
* To allow easy report generation for audit documentation.

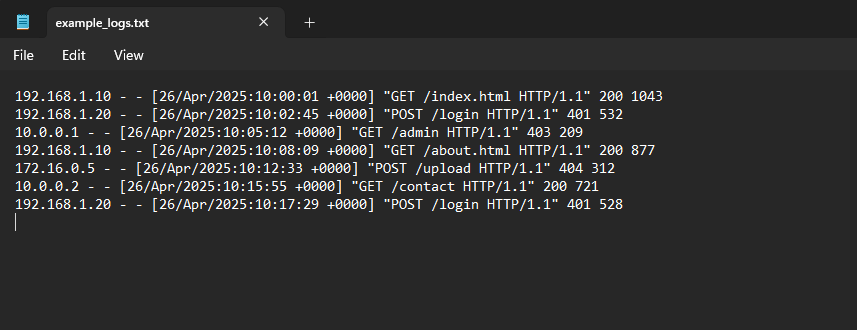
**System Architecture**

* **Input:** Log files (Apache, SSH, Firewall logs).
* **Processing:**
  + Extract IP addresses, timestamps, status codes.
  + Detect failed login attempts and unusual behavior.
* **Output:**
  + Text summary reports.
  + Bar charts visualizing top IPs and failed access attempts.
  + Downloadable audit reports (.txt and .pdf formats).

**Methodology**

1. **File Upload:**
   * Users upload log files directly into Google Colab.
2. **Log Parsing and Data Extraction:**
   * Regular Expressions (regex) are used to extract IP addresses, timestamps, and status codes.
3. **Security Threat Detection:**
   * Identifies:
     + Failed login attempts (HTTP codes 401, 403, 404).
     + Frequent IP connections indicating potential brute-force attacks.
4. **Data Analysis:**
   * IP addresses are counted to determine access patterns.
   * Top 10 frequent IPs and top 5 suspicious IPs are identified.
5. **Visualization:**
   * Bar charts display the number of requests per IP.
   * Failed login attempts visualized separately.
6. **Report Generation:**
   * Text Summary (Security\_Audit\_Report.txt).
   * PDF Report (Security\_Audit\_Report.pdf) including visual charts.
   * All reports are automatically available for download.

**Sample Results**

**Logs File:** ****

**Conclusion**

The Automated Log File Analyzer successfully streamlines security auditing by automatically detecting anomalies and presenting clear visualizations of potential threats. Its minimalistic and dependency-free implementation ensures ease of use for auditors and system administrators. The tool enhances the ability to monitor systems proactively, spot unusual patterns, and take corrective actions quickly.

**Future Improvements**

* Incorporate advanced anomaly detection algorithms (e.g., time-based access spikes).
* Extend support for Windows Event Logs and other formats.
* Integrate real-time monitoring and email alerts.
* Enhance the PDF report with detailed graphs and trend analysis over time.

**End of Report**