Ms. Tejaswini Vikas Patil, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Multi-Dimensional Cyber Threat Profiling Using System and Web Logs for Regional, Behavioural, and Risk-Based Attack Analysis

Ms. Tejaswini Vikas Patil, Professor Dr. R. N. Patil
M.Tech Defence Technology, Department of Defence Technology
D. K. T. E. Society's Textile and Engineering Institute, Ichalkaranji

Abstract- Cybersecurity monitoring is essential for protecting digital systems from web-based attacks, unauthorized access, and suspicious user activity. Traditional log monitoring usually focuses on individual events and does not provide a complete view of attacker behaviour, geographical origin, risk severity, and response requirements. This paper presents a real-time multi-dimensional cyber threat profiling system that analyses system and web logs for regional, behavioural, and risk-based attack analysis. The developed system uses Logstash for real-time log ingestion and processing, OpenSearch for storing processed security events, and OpenSearch Dashboards for visualization. The system processes simulated JSON security events, Apache web server access logs, and SSH authentication logs. It detects attack categories such as brute-force login attempts, SQL injection, cross-site scripting, directory traversal, web scanning, restricted resource access, insider suspicious access, and normal user activity. GeoIP enrichment is used for regional analysis, while a rule-based threat scoring model classifies each event as Low Risk, Medium Risk, or High Risk. High-risk events are stored in a dedicated alert index, and severe events generate response recommendations such as temporary IP blocking with pending analyst review. The implementation demonstrates real-time log ingestion, attack classification, threat scoring, risk classification, alert generation, and dashboard-based threat visualization.

Keywords— Cyber threat profiling, log analysis, OpenSearch, Logstash, SIEM, risk scoring, GeoIP enrichment, web logs, SSH logs, cyber threat visualization.

I. INTRODUCTION

Cybersecurity has become a critical requirement for organizations, academic institutions, defence systems, and online service providers. Modern attackers use different techniques such as brute force login attempts, SQL injection, cross-site scripting, directory traversal, scanning, and unauthorized access attempts. These attacks produce useful evidence in system logs, web server logs, and authentication logs. Therefore, log analysis is an important method for identifying suspicious activity and supporting cyber incident investigation.

Many conventional monitoring approaches generate a large volume of raw logs or alerts without clearly identifying the severity of each event. Security analysts may spend significant time checking low-priority logs while critical threats remain hidden among normal events. A multi-dimensional approach is required to analyse threats according to attack behaviour, geographical source, and risk level. This work proposes a cyber threat profiling system that collects logs from Apache access logs, SSH authentication logs, and JSON security events. The system converts raw logs into structured events, classifies attack categories, enriches public IP addresses with regional details, calculates threat scores, generates high-risk alerts, recommends response actions, and visualizes the results using OpenSearch Dashboards.

II. RELATED WORK

Log analysis has been widely used for intrusion detection, incident response, and forensic investigation. System logs, authentication logs,

firewall logs, and web server logs contain information about user activity, network behaviour, failed login attempts, and suspicious requests. Security Information and Event Management systems collect and correlate these logs, but they often require complex configuration and may generate many alerts.

Web application attacks such as SQL injection, cross-site scripting, and directory traversal can be identified using suspicious request patterns in HTTP logs. Brute force attacks can be detected by analysing repeated failed login attempts in authentication logs. Regional threat analysis using GeoIP information helps identify the geographical origin of public IP traffic, which supports threat investigation and security monitoring.

Existing studies indicate that combining parsing, enrichment, classification, scoring, and visualization improves cyber situational awareness. However, many systems focus only on detection or only on visualization. The proposed system integrates both by using a rule-based scoring model and dashboard-based analysis for actionable cyber threat profiling.

III. PROPOSED SYSTEM

The proposed system performs multi-dimensional cyber threat profiling using system and web logs. It collects logs from JSON security events, Apache access logs, and SSH authentication logs. Logstash is used to parse and process raw records, OpenSearch stores the structured events, and OpenSearch Dashboards visualizes attack trends, risk levels, regional patterns, alerts, and response recommendations.

The system architecture is divided into log source, processing, analysis, storage, and visualization layers. Log source components generate raw data. Logstash performs field extraction, GeoIP enrichment, attack classification, and risk calculation. OpenSearch stores all processed events, high-risk alerts, and response actions. The dashboard layer provides visual analysis for security analysts.

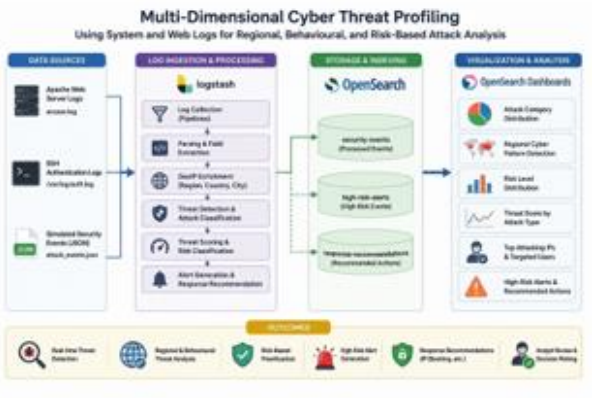

Fig. 1. System architecture of the proposed cyber threat profiling system.

IV. METHODOLOGY

1. Log Collection and Parsing

Logstash monitors configured input files and reads new events in near real time. JSON security events are decoded directly, Apache logs are parsed to extract HTTP fields, and SSH authentication logs are parsed to identify failed login attempts. Important fields include timestamp, source IP address, username, request URL, response code, login status, event type, and attack category.

2. GeoIP Enrichment

Public IP addresses are enriched using GeoIP information. The enrichment process adds country, city, country code, and geographical location fields. Localhost and private addresses are excluded from regional enrichment. This allows the system to analyse regional cyber patterns and suspicious source locations.

3. Attack Classification

The system classifies attacks using rule-based conditions. SQL injection attempts are detected using query patterns such as OR, 1=1, and union. XSS attempts are detected using script-based patterns such as <script> and alert. Directory traversal is detected using patterns such as ../, etc/passwd, and etc/shadow. Restricted access is identified when sensitive pages such as admin, wp-admin, or phpMyAdmin are requested. SSH failed login attempts are classified as brute force events.

Fig. 2. Threat profiling workflow from log ingestion to visualization.

4. Threat Scoring and Risk Classification

A rule-based scoring model calculates event severity. Factors such as failed login attempts, request frequency, unusual access time, suspicious region, restricted resource access, unique username count, and attack type contribute to the final threat score. Events with score 0-20 are classified as Low Risk, 21-50 as Medium Risk, and 51 or above as High Risk.

Fig. 3. Rule-based threat scoring and risk classification logic.

Table 1: Risk Classification Rules

Score	Level	Meaning
0-20	Low	Normal or low suspicious activity
21-50	Medium	Suspicious activity for observation
51+	High	Serious threat for analyst action

V. IMPLEMENTATION

The system was implemented on a Windows 11 host using Ubuntu WSL as the Linux development environment. Docker Desktop was used to run OpenSearch, Logstash, and OpenSearch Dashboards containers. The main project directory contained the Docker Compose configuration, Logstash pipeline configuration, input log files, and data files.

The main OpenSearch indexes used in the implementation were security-events, high-risk-alerts, and response-actions. The security-events index stores all parsed and enriched events. The high-risk-alerts index stores events classified as High Risk. The response-actions index stores recommended mitigation actions for severe events.

Table 2: Implementation Environment

Layer	Tool	Purpose
Host	Windows 11	Base OS
Linux	Ubuntu WSL	Execution shell
Container	Docker	Runs services
Pipeline	Logstash	Parse and score logs
Storage	OpenSearch	Store events
Visual	Dashboards	Charts and alerts

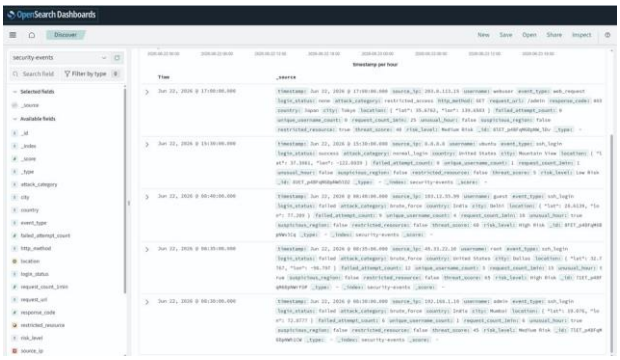

Fig. 4. Sample processed security event showing extracted fields and risk details.

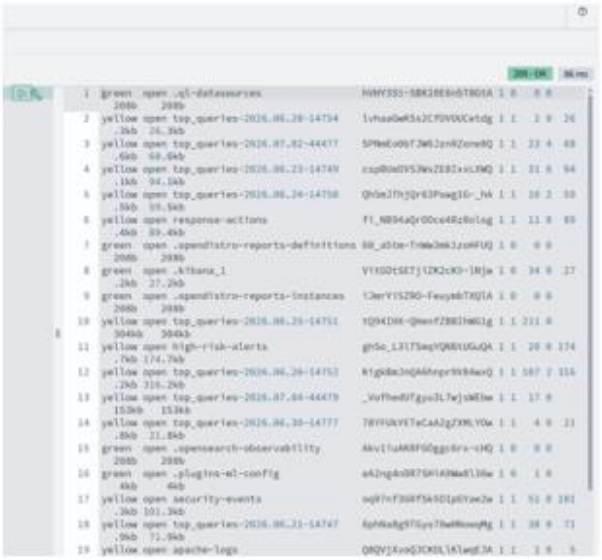

Fig. 5. OpenSearch index list showing project indexes for events, alerts, and response actions.

VI. RESULTS AND DISCUSSION

The proposed system was tested using JSON events, Apache access logs, and SSH authentication logs. The testing verified log ingestion, parsing, GeoIP enrichment, attack classification, scoring, risk classification, high-risk alert generation, response recommendation, and dashboard visualization. The system successfully processed different attacks including SQL injection, XSS, directory traversal, restricted access, and brute force login attempts.

Normal web requests were classified as Low Risk, restricted access attempts were classified as Medium Risk, and attacks such as SQL injection, XSS, directory traversal, and SSH brute force were classified as High Risk. Severe events generated response recommendations such as Recommend IP Blocking and Pending Analyst Review.

Table 3: Sample Threat Scoring Results

Attack	Score	Risk	Action
Normal web	0	Low	None
Restricted access	30	Medium	Monitor
Brute force	60	High	Alert
Traversal	70	High	Block review
SQL injection	70-75	High	Block review
XSS	65-85	High	Block review

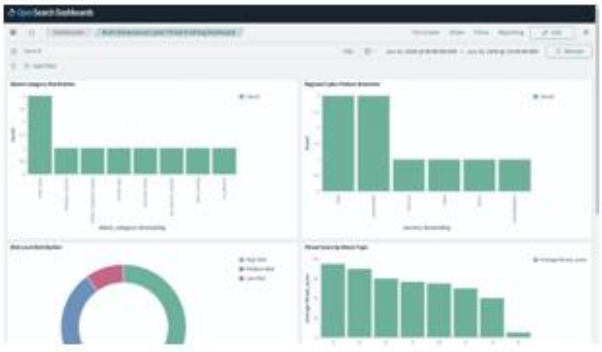

Fig. 6. Dashboard visualization of attack categories and risk level distribution.

Fig. 7. Regional cyber pattern dashboard showing geographical attack sources.

Fig. 8. High-risk alerts and recommended response actions dashboard.

Table 4: Functional Test Summary

Scenario	Result	Status
Services	All containers started	Pass
Ingestion	JSON, Apache, SSH logs processed	Pass
Detection	Attack categories detected	Pass
Risk scoring	Low, Medium, High assigned	Pass
Alerts	High-risk events stored	Pass
Response	Action suggestions generated	Pass
Dashboard	Charts and tables displayed	Pass

Advantages and Limitations

The proposed system provides multiple advantages. It supports multiple log sources, performs near real-

time log ingestion, classifies attack categories, calculates understandable threat scores, generates separate high-risk alerts, recommends response actions, and provides visual analysis using dashboards. The use of open-source tools makes the system suitable for academic and experimental cybersecurity environments.

The main limitation is that detection is rule-based. Unknown or highly advanced attacks may not be detected unless their patterns are added to the rule logic. The system was tested using limited log sources, and real enterprise deployment would require integration with firewall, DNS, endpoint, IDS/IPS, and cloud audit logs. GeoIP data can also be inaccurate when attackers use VPNs, proxies, Tor networks, or private addresses.

Future Scope

The system can be extended using machine learning algorithms for anomaly detection and unknown attack identification. More log sources such as firewall logs, DNS logs, endpoint logs, IDS/IPS alerts, and cloud audit logs can be integrated. Threat intelligence feeds can be added to compare source IP addresses with known malicious IP databases. Future work can also include automated firewall blocking after analyst approval, email or SMS alerting, role-based access control, and automated incident report generation.

VII. CONCLUSION

This paper presented a multi-dimensional cyber threat profiling system using system and web logs. The system collects logs from Apache, SSH, and JSON security events, processes them using Logstash, stores them in OpenSearch, and visualizes the results using OpenSearch Dashboards. The implementation performs attack classification, GeoIP enrichment, threat scoring, risk classification, alert generation, and response recommendation.

The experimental results show that the system can detect attacks such as brute force login attempts, SQL injection, cross-site scripting, directory traversal, and restricted access. The rule-based scoring method provides a simple and interpretable approach to risk prioritization. Overall, the proposed system demonstrates how open-source tools can be used to build a SIEM-like cyber threat profiling platform for regional, behavioural, and risk-based cyber threat analysis.

REFERENCES

1. OpenSearch Project, "OpenSearch Documentation," OpenSearch Project Documentation, 2024.
2. OpenSearch Project, "OpenSearch Dashboards Documentation," OpenSearch Project Documentation, 2024.
3. Elastic, "Logstash Reference Documentation," Elastic Documentation, 2024.
4. Docker Inc., "Docker Desktop and Docker Compose Documentation," Docker Documentation, 2024.
5. Apache Software Foundation, "Apache HTTP Server Documentation," Apache Documentation, 2024.
6. MaxMind, "GeoIP and GeoLite2 Database Documentation," MaxMind Documentation, 2024.
7. K. Scarfone and P. Mell, "Guide to Intrusion Detection and Prevention Systems," NIST Special Publication 800-94, 2007.
8. P. Cichonski, T. Millar, T. Grance, and K. Scarfone, "Computer Security Incident Handling Guide," NIST Special Publication 800-61 Revision 2, 2012.
9. OWASP Foundation, "OWASP Top 10 Web Application Security Risks," OWASP Documentation, 2021.
10. MITRE Corporation, "MITRE ATT&CK Framework for Enterprise," MITRE Documentation, 2024.
11. A. Chuvakin, K. Schmidt, and C. Phillips, Logging and Log Management, Syngress, 2012.
12. R. Bejtlich, The Practice of Network Security Monitoring, No Starch Press, 2013.