International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 06 | Jun 2026 www.irjet.net p-ISSN: 2395-0072

Intelligent Cyber Threat Hunting Using Honeypot Logs with Machine Learning and MITRE ATT&CK Mapping

S. Manikanta Sai[1], M. Sreenivasu[2]

[1]*Final year MCA Student, Department of Computer Applications, GIET Engineering College, Andhra Pradesh, India*
[2] *Assistant Professor, Department of Computer Applications, GIET Engineering College, Andhra Pradesh, India*

Abstract – *Cyber-attacks against exposed digital services are increasingly automated, repetitive, and difficult to analyse manually in real time. This project presents an intelligent cyber threat-hunting platform that combines honeypot telemetry, intrusion detection events, machine learning, and MITRE ATT&CK mapping into a unified live-monitoring workflow. The system ingests logs from Cowrie and Suricata, normalises heterogeneous event formats, extracts attack-relevant features such as command behaviour, login failures, protocol context, session activity, and network traffic characteristics, and applies ensemble-based prediction to classify suspicious behaviour. The predicted attack type is further mapped with MITRE ATT&CK tactics and techniques to provide structured adversary context for investigation and response. A FastAPI-based backend manages ingestion, preprocessing, model execution, alert generation, and storage, while a web dashboard delivers real-time visibility into alerts, attacker activity, attack trends, and mapped techniques using WebSocket updates. This platform also generates reports and supports attack-focused filtering. By integrating live log analytics, intelligent attack prediction, and mapping into a single framework, the proposed system provides a practical approach for cyber threat intelligence, security monitoring, and real-time attack analysis in research and local deployment environments.*

Key Words: Cyber Threat Hunting, Honeypot, Cowrie, Suricata, Machine Learning, MITRE ATT&CK, Intrusion Detection, Real-time Monitoring, Threat Intelligence, Attack Classification

1. INTRODUCTION

He rapid growth of internet-connected services, remote access systems, and automated attack tooling has significantly increased the exposure of modern computing environments to cyber threats. Attackers routinely use scanning tools, brute-force techniques, scripted payload delivery, and exploit attempts to identify vulnerable targets and compromise systems. In many cases, raw security logs are generated in large volumes, but it makes difficult for defenders to distinguish meaningful attack activity from ordinary background noise.

Cyber threat hunting aims to address this problem by proactively identifying hostile behaviour through continuous analysis of system, network, and security telemetry. Honeypots and intrusion detection systems are especially valuable in this context because they capture attacker interactions and suspicious traffic in a controlled environment. Honeypots can reveal how adversary services, attempt logins, and execute commands, while network-based detection platforms can identify scanning behaviour, protocol misuse, exploit signatures, and other malicious patterns. However, these sources produce heterogeneous data that must be normalised, correlated, and enriched before it becomes useful for investigation.

This project proposes an intelligent cyber threat hunting framework that integrates live metrics from Cowrie and Suricata with machine learning and MITRE ATT&CK mapping. The goal of the system is not only to collect attack data, but also to transform it into structured data, and process incoming logs in real-time and predicts likely attack categories, and associate detected activity with corresponding MITRE ATT&CK tactics and techniques. The implementation is designed as a full-stack local platform. A FastAPI backend handles data ingestion, preprocessing, model execution, alert creation, and persistence.

A live dashboard presents alerts, attack trends, attacker IP activity, and monitoring. This makes the system useful for research, demonstration, and local defensive experimentation where live attack visibility and rapid contextualization are important.

2. OBJECTIVES

The main objective of this work is to develop a live cyber threat hunting platform that can collect, process, and analyse attack-related events from security monitoring sources in real-time. The system is designed to improve visibility into malicious activity by combining honeypot logs, intrusion detection data, machine learning, and MITRE ATT&CK-based interpretation within a single workflow. The specific objectives of the project are as follows:

- To collect live security events from Cowrie and Suricata in a unified monitoring environment.
- To preprocess heterogeneous log data and extract useful attack-oriented features for analysis.
- To classify suspicious activity using machine learning models trained on a combination of self-generated datasets and publicly available datasets collected from online sources.
- To map detected attacks to MITRE ATT&CK tactics and techniques for better threat interpretation.
- To provide a real-time dashboard for monitoring alerts, attacker activity, and attack trends.
- To support faster analysis by filtering non-critical telemetry and highlighting meaningful malicious events.

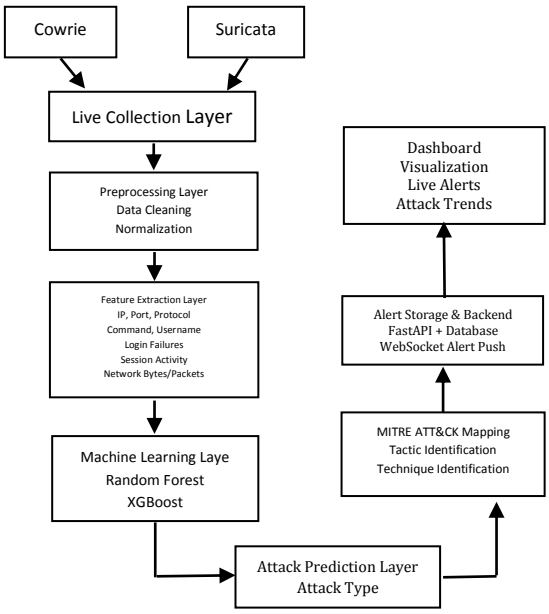
3. PROPOSED SYSTEM

The proposed system is a real-time cyber threat hunting framework that integrates log collection, event pre-processing, attack prediction, adversary mapping, and dashboard-based monitoring. The system is built to observe live events generated from honeypot and intrusion detection sources and convert them into structured threat intelligence.

In this framework, Cowrie is used to capture attacker interaction such as login attempts, command execution, and suspicious session behaviour. Suricate is used to monitor network-level events and detect patterns such as scanning activity, brute-force attempts, exploit traffic, and abnormal communication behaviour. These events are collected and normalized by the backend so that they can be processed in a common format. After normalisation, the system extracts important attributes such as attacker IP address, destination port, protocol, command content, session duration, login-failure count, signature text, and network traffic statistics. These features are then passed to machine learning models for attack prediction. The predicted attack category is enriched with MITRE ATT&CK tactic and technique mapping so that the identified activity can be interpreted in a structured security context.

The processed output is stored and displayed through a web-based dashboard. The dashboard presents real-time alerts, attack types, top attacker IPs, timeline trends, and MITRE-aligned threat context. This design allows the proposed system to act as a compact end-to-end threat hunting platform for live monitoring and security analysis.

4. SYSTEM ARCHITECTURE



5. METHODOLOGY

The methodology of the proposed work consists of five main stages: data collection, preprocessing, feature extraction, attack prediction, and visualization.

In the first stage, the system collects live log data from Cowrie and Suricata. Cowrie provides application-level attacker interaction data, including login attempts and executed commands, while Suricata provides network-level events such as alerts, protocol activity, and suspicious traffic patterns.

In the second stage, the collected events are normalised into a unified structure. Since the incoming logs are generated by different tools and use different field formats, the system standardises important attributes before further analysis. Duplicated or incomplete records are handled during this stage to improve consistency.

In the third stage, attack-related features are extracted from the normalised events. These include source information, destination port, username, command content, protocol type, command length, event frequency, session behaviour, failed-login count, and traffic-related values such as bytes and packets.

In the fourth stage, the extracted features are passed to machine learning models to predict the likely attack category behaviour analysis. The predicted result is then correlated with MITRE ATT&CK tactics and techniques

to provide meaningful threat context instead of only raw labels.

In the fifth stage, the processed results are stored and displayed on the dashboard. Real-time alerts are sent to the user interface, where analysts can observe current attack activity, recent alerts, attack distributions, and adversary mapping information. This step improves usability by transforming raw security telemetry into interpretable operational intelligence.

6. RESULTS AND DISCUSSION

The developed system was able to collect, process, and display live security events from both Cowrie and Suricata in a unified monitoring workflow. During testing, the platform successfully ingested attacker interaction logs, normalized the incoming events, extracted relevant features, and generated predicted attack labels in near real time. The processed events were then mapped to MITRE ATT&CK tactics and techniques and displayed on the dashboard for easier interpretation.

The results showed that the system was effective in identifying multiple categories of suspicious activity, including brute-force login attempts, port-scanning behaviour, command execution patterns, and other abnormal network events. Cowrie logs were especially useful for capturing attacker commands, login failures, and session behaviour, while Suricata provided broader network-level visibility such as scan-related traffic, alert signatures, and protocol-based anomalies. The combination of these two sources improved the overall visibility of attack activity.

The machine learning component helped convert raw telemetry into meaningful attack predictions. Instead of showing only unprocessed logs, the system classified events into attack-oriented categories and attached contextual information for analysis. This made the monitoring process more practical, especially when multiple events were generated within a short period of time. The inclusion of MITRE ATT&CK mapping further improved the usefulness of the output by connecting predicted attacks with recognised adversary tactics and techniques.

The dashboard provided a clear real-time view of the detected activity. Alerts, top attacker IPs, attack trends, and mapped techniques could be observed from a single interface. The attack-focused filtering mechanism also improved readability by reducing the effect of low-value or routine telemetry. As a result, the platform was more useful for highlighting potentially meaningful threats than a raw view alone.

Overall, the results indicate that the proposed system can serve as a practical threat hunting platform for live attack observation and security analysis. Although the system is intended for research and local deployment, it demonstrates that combining honeypot data, intrusion detection events, machine learning, and MITRE ATT&CK-based interpretation can improve the analysis of suspicious cyber activity in real time.

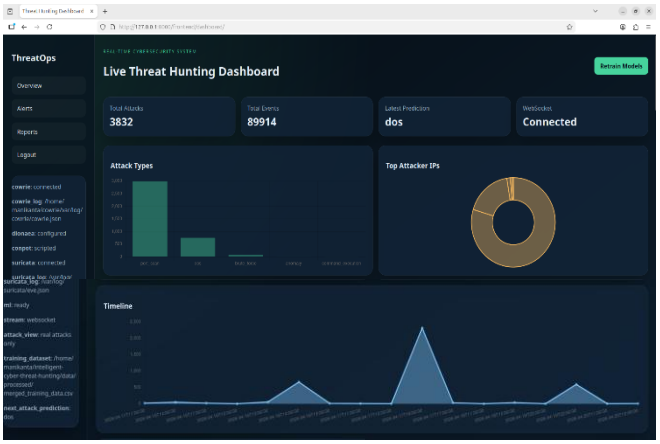


Fig.1: Live Threat Hunting Dashboard

7. CONCLUSION

This work presented a real-time cyber threat hunting platform that combines honeypot telemetry, intrusion detection events, machine learning, and MITRE ATT&CK mapping within a unified monitoring framework. The system was designed to collect live security events from Cowrie and Suricata, preprocess heterogeneous log data, extract attack-oriented features, and generate meaningful threat predictions for analysis.

The developed platform showed that raw security logs can be transformed into more useful threat intelligence when combined with classification models and structured visualization, the system improves visibility into suspicious activity such as brute-force attempts, scanning behaviour, command execution, and other abnormal events. The inclusion of MITRE ATT&CK context further strengthens the interpretability of detected incidents and supports quicker understanding of attacker behaviour.

Overall, the proposed system provides a practical and extensible approach for live cyber threat monitoring in research and local deployment environments. It demonstrates that combining honeypot logs, network alerts, and intelligent analysis can support more effective threat hunting than isolated log inspection alone.

REFERENCES

1) J. Velasquez, J. Angulo, C. Mazutti, H. Goncalves, S. Khanchi, T. Makanju, R. Rab, and U. Zakia, "Orchestrating Cyber Threat Intelligence using T-Pot and MITRE ATT&CK: From Cyber Attack Data Collection to Advanced Insights," in 2025 IEEE 16th Annual Information Technology, Electronics and Mobile Communication Conference(IEMCON),2025,doi:10.1109/IEMCON67450.2025.11381043.

2) MITRE ATT&CK, "MITRE ATT&CK Framework," MITRE.[Online].Available:https://attack.mitre.org/. Accessed: Jun. 5, 2026.

3) Cowrie Project, "Cowrie Documentation," [Online].Available:https://docs.cowrie.org/en/stable/. Accessed: Jun. 5, 2026.

4) OISF, "Suricata User Guide," [Online]. Available: https://docs.suricata.io/en/. Accessed: Jun. 5, 2026.

5) F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.

6) T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785-794.

7) Sharafaldin, A. H. Lakshkari, and A. A. Ghorbani, "Towars Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization, "in Proceedings of the 4th International Conference on Information systems Security and Privacy (ICISSP), 2018, PP. 108-116.

8) Canadian Institute for Cybersecurity, "CICIDS2017 Dataset," University of New Brunswick.[Online].Available:https://www.unb.ca/cic/datasets/ids-2017.html. Accessed: Jun. 5, 2026.