Archives available at journals.mriindia.com

International Journal of Recent Advances in Engineering and Technology

ISSN: 2347 - 2812

Volume 14 Issue 03s, 2025

Scalable Framework for Advanced Threat Analysis

[1]Dr. Heena Farheen Ansari, [2]Riyal Patil, [3]Vedant Joshi, [4]Aryan Wankhede, [5]Sahil Gaikwad
[1]Assistant Professor, Dept. of CSE (Cyber Security)
SVPCET, Nagpur, India
[2,3,4,5] Dept. of CSE (Cyber Security)
SVPCET, Nagpur, India
Email: [1]hansari@stvincentngp.edu.in, [2]riyalpatil.22@stvincentngp.edu.in,
[3]vedanthimanshujoshi.22@stvincentngp.edu.in, [4]aryanwankhede.22@stvincentngp.edu.in,
[5]sahilgaikwad.22@stvincentngp.edu.in

Peer Review Information

Submission: 05 Nov 2025

Revision: 25 Nov 2025

Acceptance: 17 Dec 2025

Keywords

open source intelligence, network vulnerabilities, data breaches, cyber security

Abstract

Scalable framework for advanced threat analysis is an open-source web application designed to redefine open-source intelligence (osint) by providing a centralized platform for cybersecurity investigations. It integrates powerful tools to deliver a sleek, responsive interface and robust backend. The application enables comprehensive analysis of ips, domains, emails and files, addressing critical cybersecurity challenges such as identifying network vulnerabilities, detecting data breaches, profiling threat actors, and uncovering malicious file attributes. It ensures flexibility and scalability, making osint investigator an essential tool for tackling modern cyber threats.

Introduction

In today's digital-first world, organizations are constantly exposed to a growing number of cyber threats. From phishing emails and malicious domains to sophisticated malware and advanced persistent threats, attackers are becoming increasingly adaptive and resourceful. Security teams, however, often find themselves constrained—either by the high costs of commercial threat intelligence solutions or by the limitations of fragmented open-source tools. This imbalance creates a critical gap, leaving many institutions underprepared to face evolving cyber risks.

The Open Threat Intelligence Platform was developed to bridge this gap by offering a modular, API-free, and fully local framework for advanced threat analysis. Unlike conventional solutions that depend on expensive external APIs or cloud services, this platform performs in-depth investigations using open-source libraries and local processing engines [1]. It integrates four core modulesIP Scanner, Domain Intelligence, Email Analyzer, and File Analyzerinto a unified dashboard, enabling analysts to gather, correlate, and interpret threat data with transparency and independence.

Through this project, we aim to demonstrate that effective cybersecurity intelligence does not have to be locked behind high price tags or vendor restrictions. Instead, by leveraging open-source technologies and a scalable design, the platform empowers a wider community of defenders—making advanced threat analysis more transparent, affordable, and sustainable [2].

The objective of the research work is API Independence – To design a self-sufficient threat intelligence platform that operates entirely on local processing, eliminating reliance on costly or restricted third-party APIs.

Comprehensive Local Analysis – To enable deep technical analysis of IP addresses, domains, emails, and files using open-source libraries, pattern recognition, and behavioral analytics.
Modular and Scalable Architecture – To implement a flexible framework with four interchangeable modules—IP Scanner, Domain Intelligence, Email Analyzer, and File Analyzer—that can scale across organizational needs.
MITRE ATT&CK Mapping – To integrate findings with standardized adversary tactics and techniques for structured threat attribution and enhanced analytical accuracy.
MITRE ATT&CK Mapping – To integrate findings with standardized adversary tactics and techniques for structured threat attribution and enhanced analytical accuracy.
Transparent and Exportable Reporting – To provide detailed, explainable markdown-based reports with risk scores, technical evidence, and recommendations for incident response and forensic documentation.
Accessibility in Diverse Environments – To ensure ease of deployment in air-gapped, resource-constrained, and educational environments without compromising analytical depth or accuracy.

Literature Review

The field of cyber threat intelligence has undergone significant evolution over the last decade, with both commercial and open-source solutions emerging to meet the growing demands of organizations. The increasing sophistication of attacks has driven the need for platforms that not only detect and analyze threats but also provide contextual and actionable insights. Literature highlights three dominant directions in this space: commercial threat intelligence services, open-source intelligence (OSINT) frameworks, and modular hybrid systems.
Open-Source Intelligence (OSINT) Frameworks. Community-driven projects such as MISP (Malware Information Sharing Platform), Open CTI, and The Hive have gained attention as cost-effective alternatives to commercial platforms. These frameworks emphasize collaborative sharing, API-based data integration, and flexibility in deployment. The highlights the role of OSINT in enabling organizations to gather actionable intelligence from public sources [2]. Nonetheless, research indicates persistent challenges such as fragmented workflows, high setup complexity, and limited analytical depth in areas like pattern recognition and behavioral analysis [3].
Hybrid and Modular Approaches. More recent works advocate for frameworks that combine modular design with advanced local analysis. These systems emphasize entity correlation across multiple threat dimensions—such as IPs, domains, emails, and files—while integrating with standardized adversary models like MITRE ATT&CK. Studies suggest that such platforms address three key gaps in prior work:
API Dependency – Excessive reliance on paid or cloud APIs restricts availability during critical incidents and poses risks in air-gapped environments.
Fragmentation of Intelligence – Data from IPs, domains, emails, and files often requires manual correlation across platforms, increasing the burden on analysts.
Transparency and Explainability – Many platforms provide risk scores or alerts without detailed reasoning, making it difficult for organizations to validate and act on intelligence.
Recent literature explores hybrid models that combine modular architectures, local intelligence databases, and integration with frameworks like MITRE ATT&CK for structured adversary mapping. These works advocate for scalable, transparent, and vendor-independent solutions that address accessibility challenges while maintaining analytical rigor. For example, Hulnick (2010) stresses the importance of explainable intelligence practices [4], while Europol (2017) highlights the role of community-driven, API-free approaches in democratizing cyber threat intelligence [6].
Despite these advancements, a significant gap remains in delivering a comprehensive, self-sufficient, and transparent threat intelligence platform that combines modular analysis (IP, domain, email, file) with local data processing and MITRE ATT&CK integration—without reliance on external APIs. This research addresses that gap by introducing the Open ThreatIntelligence Platform, designed to provide professional-grade analysis while remaining cost-effective, transparent, and scalable across diverse environments.

Methodology

The proposed framework was developed through a structured methodology that ensured scalability, transparency, and API independence[8]. The process followed several sequential stages:
Requirement Definition: Identified the limitations of existing commercial and open-source platforms, including cost, API dependency, and lack of transparency.
Defined key use cases: IP analysis, domain Scanner, email verification, and file inspection.
Modular System Design: Developed a scalable architecture consisting of four modules: IP

Scanner, Domain Scanner, Email Analyzer, and File Analyzer.
Integrated modules into a unified Flask-based API gateway and a React web dashboard for seamless interaction.
Implementation of Local Engines: Leveraged open-source libraries such as dnspython, socket, python-whois, hashlib, libmagic, and PyPDF2 to perform all intelligence gathering locally.
Eliminated reliance on third-party or cloud APIs to ensure operational independence and data sovereignty.
Pattern Recognition and Risk Scoring: Applied lexical, structural, and behavioral analytics to detect anomalies in IPs, domains, emails, and files.
Implemented a multi-category weighted risk scoring engine (0–10 scale) to categorize threats as Low, Medium, High, or Critical.
MITRE ATT&CK Mapping: Correlated each module's findings with relevant adversary tactics and techniques.
Ensured structured attribution of malicious behavior to standard MITRE ATT&CK identifiers for improved forensic reliability. Data Aggregation and Correlation: Unified results from all four modules into a centralized local database (SQLite).
Enabled cross-entity intelligence correlation, reducing fragmented reporting and enhancing analytical depth.
Transparent Reporting: Designed a Markdown-based report generator that produces exportable, explainable reports.Reports include executive summaries, technical details, risk scores, and recommendations, enabling effective incident response.
Testing and Feedback: Deployed the framework in controlled environments to evaluate performance, scalability, and accuracy.
Applied iterative refinement based on analyst feedback, focusing on reducing false positives and improving usability.

Problem Statement

The increasing complexity and frequency of cyber threats have created a significant demand for advanced threat intelligence solutions. While commercial platforms such as FireEye, Recorded Future, and CrowdStrike provide comprehensive capabilities, they are often associated with high costs, vendor lock-in, and API dependencies, making them inaccessible to small organizations, educational institutions, and resource-constrained environments. Moreover, these platforms frequently operate as "black boxes," offering limited transparency in their detection methodologies and risk scoring, which reduces analyst trust and hampers forensic validation.
Open-source intelligence (OSINT) frameworks, on the other hand, offer a cost-effective alternative but suffer from fragmented workflows, limited analytical depth, and extensive setup requirements. Many of these tools rely heavily on cloud APIs, which introduces challenges such as data sovereignty concerns, restricted accessibility in air-gapped environments, and unpredictable API usage costs. Additionally, the lack of integrated reporting and standardized adversary mapping makes it difficult for organizations to derive actionable insights from these tools.
Thus, there exists a clear gap for a self-sufficient, modular, and scalable threat intelligence platform that can operate independently of external APIs, perform comprehensive local analysis, provide transparent and explainable reporting, and remain accessible across diverse operational environments. This research addresses this gap by developing the Open Threat Intelligence Platform, which integrates multiple analysis engines into a unified framework while ensuring cost-effectiveness, scalability, and intelligence independence.

Experimental Results

The proposed framework was evaluated through a series of controlled experiments designed to validate its performance, scalability, and accuracy in detecting potential threats. Each of the four modulesIP Scanner, Domain Intelligence, Email Analyzer, and File Analyzerwas tested using real-world datasets, including suspicious IP ranges, phishing domains, disposable email addresses, and malicious file samples.
IP Scanner Results: The module successfully identified open ports, WHOIS details, and geolocation of tested IP addresses.
Anomalies such as unusual TTL values and open administrative ports were flagged and mapped to MITRE
ATT&CK technique T1046 (Network Service Scanning).
Domain Scanner Results:
The system detected typosquatted domains and invalid SSL/TLS certificates.
Domains linked to phishing infrastructure were categorized as *High Risk* with a score above 8/10.
Email Analyzer Results: isposable and role-based email addresses were detected with high accuracy.
Emails with missing SPF/DMARC records were flagged, correlating with MITRE ATT&CK technique T1589
(Gather Victim Identity Information).

File Analyzer Results: Malicious PDF and executable files were analyzed for entropy, metadata, and string patterns.
Malware samples matched YARA signatures and were assigned *Critical* severity, mapped to T1083 (File and Directory Discovery).

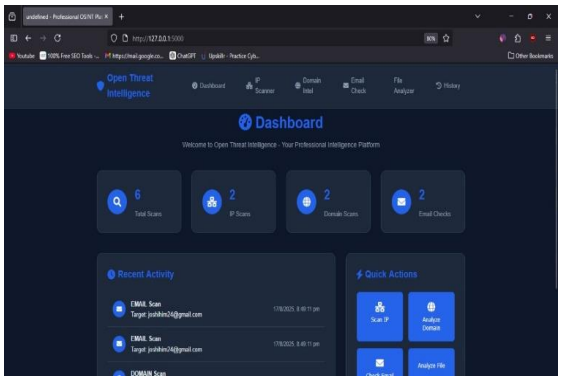

Fig 1. Dashboard

Conclusion

The scope of this project is to design, implement, and evaluate an open-source, modular, and API-independent threat intelligence platform that consolidates multiple cybersecurity investigation functions into a unified system. The project encompasses the following areas:
Integration of Core OSINT Modules:
Development of four primary modules: IP Scanner, Domain Intelligence, Email Analyzer, and File Analyzer.
Each module is designed to perform in-depth local analysis while supporting optional integration with external APIs (e.g., Shodan, VirusTotal, EmailRep.io) for enriched results.
Modular and Scalable Architecture:
Implementation of a Flask-based backend and React/JavaScript frontend dashboard to provide seamless interaction. Support for extensibility, enabling the future integration of additional modules such as Dark Web Monitoring, Social Media Intelligence (SOCMINT), or AI-based anomaly detection.
Threat Scoring and Standardized Mapping:
Development of a multi-factor risk scoring system (0–10 scale) with explainable severity levels (Low, Medium, High, Critical). Mapping of all findings to MITRE ATT&CK techniques for structured adversary attribution.
Transparent and Exportable Reporting:
Generation of explainable, Markdown-based reports containing technical details, risk factors, and recommendations. Export functionality to PDF/Markdown for use in forensic investigations, compliance, and documentation.
Performance and Usability Testing:
Evaluation of platform performance under different workloads, including accuracy, false-positive rates, and response time. Testing inair-gapped,resource-constrained, and educational environments to ensure accessibility and usability.
Security, Privacy, and Data Sovereignty:
Emphasis on local processing to maintain complete control over sensitive data. Ensuring that the platform operates securely without vendor lock-in or cloud dependencies.

Reference

MITRE ATT&CK Framework, “Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK),” 2024. [Online]. Available: https://attack.mitre.org/

M. Bazzell, Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information, 9th ed. IntelTechniques, 2021.

W. Tounsi and H. Rais, “A survey on technical threat intelligence in the age of sophisticated cyber attacks,” Computers & Security, vol. 72, pp. 212–233, 2018.

A. S. Hulnick, “The Dilemma of Open Source Intelligence: Is OSINT Really Intelligence?” International Journal of Intelligence and CounterIntelligence, vol. 24, no. 4, pp. 629–645, 2010.

B. Buchanan and T. Rid, “Attributing Cyber Attacks,” Journal of Strategic Studies, vol. 38, no. 1–2, pp. 4–37, 2015.

Europol, Open Source Intelligence (OSINT) Handbook: A Law Enforcement Guide to Using OSINT Effectively, Europol, 2017.

VirusTotal, “VirusTotal API Documentation,” 2024. [Online]. Available: https://developers.virustotal.com/

Shodan, “Shodan API Overview,” 2024. [Online]. Available: https://developer.shodan.io/

EmailRep.io, “Email Reputation API Documentation,” 2024. [Online]. Available: https://docs.emailrep.io/

AbuseIPDB, “AbuseIPDB API Guide,” 2024. [Online]. Available: https://docs.abuseipdb.com/

DNSPython, “DNSPython Library Documentation,” 2024. [Online]. Available: https://dnspython.readthedocs.io/

BeautifulSoup, “Beautiful Soup 4 Documentation,” 2024. [Online]. Available: https://www.crummy.com/software/BeautifulSoup/

Python-WHOIS, “Python Whois Library,” 2024. [Online]. Available: https://pypi.org/project/python-whois/