Cyber Triage Tool to Streamline Digital Forensic Investigation

Soham Kawadkar
Dept. of Computer Science & Engineering
MIT School of Computing
Pune, India
sohamkawadkar31@gmail.com

Prabhakar Mishra
Dept. of Computer Science & Engineering
MIT School of Computing
Pune, India
piyushprabhakar2004@gmail.com

Akhilesh Palve
Dept. of Computer Science & Engineering
MIT School of Computing
Pune, India
akhileshpalve89@gmail.com

Vedant Nagre
Dept. of Computer Science & Engineering
MIT School of Computing
Pune, India
Vedantnagre1111@gmail.com

Prof. Aman Kamble
Dept. of Computer Science & Engineering
aman.kamble@mituniversity.edu.in

Abstract—Digital forensics investigations normally take a long time because investigators have to manually collect, scan, and analyze a lot of evidence. This causes delays, especially during cyber incidents where time is very important. In this paper, We propose a cyber triage automation tool that helps investigators quickly analyze disk images and extract important indicators of compromise (IOCs). The tool uses automated scripts, open-source forensic tools, and a simple dashboard to show the findings. The main aim is to reduce investigation time and give a quick first-level understanding of the case.

Keywords: Digital Forensics, Cyber Triage, Incident Response, Automation, Disk Image Analysis, Cybersecurity.

I. Introduction

Cyber attacks are increasing every year, and organizations face different types of incidents like malware infections, data breaches, and insider threats. During these incidents, digital forensics plays a major role in understanding what happened. But doing forensics manually takes too much time because investigators have to mount images, check logs, run tools and create reports.

In many real-world cases, the first few hours are very important for decision making. This is why cyber triage tools are becoming popular. These tools don't replace full forensic analysis, but they give a quick summary to help understand the situation. However, many existing triage tools are expensive or complicated to use.

In this paper, we describe a simple and affordable cyber triage tool designed for investigators and students. It focuses on automating basic tasks like analyzing disk images, scanning memory dumps, collecting metadata, and generating a readable report.

II.Objective

1. To automate evidence collection from disk images and memory dumps.
2. To integrate open-source forensic tools like Autopsy and Volatility.
3. To detect indicators of compromise (IOCs) such as suspicious files, processes, and network artifacts.
4. automation.

4. To generate a readable dashboard that shows all important findings.

5. To reduce the time required for first-level forensic analysis.

III. Literature review

Digital forensics automation has been discussed in many research works. Most papers highlight that manual forensics is slow and requires experienced analysts. Tools like Autopsy, Volatility, Redline, and FTK provide strong features but are either expensive or require deep technical knowledge.

Previous studies show that beginners struggle with using multiple tools together. There is also a gap in tools that focus only on quick triage instead of full analysis. This motivated me to design a lightweight tool that combines commonly used techniques and automates repetitive steps.

IV. Tools and Languages

Programming Languages

- Python – Primary language for AI, machine learning, digital forensics, and backend development. Known for simplicity, rich libraries, and strong community support.
- JavaScript – Used in frontend development and browser extension components for the project.
- C, C++ – Used for performance-critical components and security-related functionalities.
- Node.js – Backend development for API and server-side logic.
- Flask – Python web framework for backend API development.
- HTML, CSS – Frontend design and interface development.

Machine Learning & AI Frameworks

- TensorFlow / Keras – For developing deep learning models that assist in anomaly detection and forensic image analysis.
5. Scikit-learn – Implementing traditional machine learning algorithms.
6. NLTK / spaCy – Natural language processing for analyzing textual artifacts and logs.

Web Scraping & URL Analysis

- BeautifulSoup / Scrapy – Extracting and parsing data from web pages for threat intelligence.
- URLNet / TLDExtract – Analyzing URLs and extracting features for malicious detection.

Cybersecurity Tools & APIs

- VirusTotal API – Checking URLs and files against known malicious databases for evidence validation.
- Google Safe Browsing API – Detecting dangerous websites during investigations.

Database & Storage

- MySQL – Primary relational database for storing evidence data and logs.
- SQLite / PostgreSQL – Alternative databases used for storing scam patterns and forensic data.

Development & Deployment

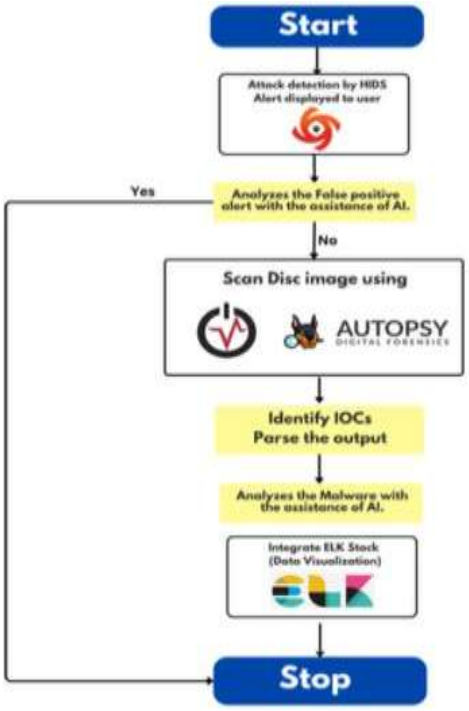
- React / Next.js – Frontend frameworks for constructing user dashboards and visualization interfaces.
- Docker – Containerization technology for deploying the application in consistent environments.
- AWS / Google Cloud – Cloud platforms used for hosting AI models, databases, and backend services.

Forensic & Security Tools Integration

- Autopsy – Open-source digital forensics tool integrated for disk

- Volatility – Memory forensics framework for live memory analysis.
- FTK Imager – Forensic imaging tool.
- OSSEC – Host-based Intrusion Detection System (HIDS).
- ELK Stack (Elasticsearch, Logstash, Kibana) – For log indexing, analysis, and real-time dashboard visualization.

V. Visual Representation



VI. Results

The automated triage tool was tested using sample disk images from open-source forensic datasets. The tool successfully extracted basic information like:

- Suspicious executables
- Browser activity
- USB connection logs
- Deleted files

VII. Process Flow

The overall working of the system follows a step-by-step flow that starts from evidence ingestion and ends with the final report. The complete process flow is shown below:

1. **Start the Investigation**
   The investigator launches the triage tool and selects the disk image or memory dump to analyze.

2. **Evidence Ingestion and Mounting**
   The tool loads the evidence in a read-only mode to maintain integrity.
   Supported formats include **E01**, **RAW**, **VMDK**, and **memory dump (.mem)** files.

3. **Automated Pre-Processing**
   Basic file system parsing, metadata reading, and directory structure mapping are done automatically.
   This prepares the image for deeper analysis.

4. **Run Automated Forensic Modules**
   The system executes several built-in analysis routines:
   - **Autopsy modules** for file scanning, hash checking, and EXIF extraction
   - **Volatility** for memory process scanning and network artifacts
   - **Custom Python scripts** for keyword search, IOC detection, and anomaly checks

5. **IOC (Indicator of Compromise) Detection**
   Suspicious files, strange processes, abnormal logs, and browser traces are flagged automatically.
   Hash comparison is done using known-bad lists.

6. Timeline Generation

Event timestamps like file creation, modification, browser activity, and process start times are combined to build a simple chronological timeline.

7. Data Aggregation and Analysis

All findings are grouped into categories such as user activity, system activity, network activity, and suspicious events.

The triage logic highlights items that require human attention.

1. **Dashboard Visualization**
   The dashboard displays:
   - Suspicious files
   - Recent user actions
   - Memory artifacts
   - Browser history
   - Network logs
2. This helps the investigator review findings quickly.
3. **Report Generation**
   After analysis, the tool automatically generates a PDF/HTML report that summarizes all the important evidence.
   The report includes tables, lists, and timestamps for easy reference.
4. **End of Triage Phase / Handover**
   The investigator uses the generated report to decide if deeper, full-scale forensic investigation is needed.

VIII. Acknowledgment

We express our gratitude to everyone who has contributed to the development of this research and overall project on Cyber Triage tool. We would like to extend our sincere and deepest gratitude to Prof. Aman Kamble for his guidance and mentorship throughout this research. His expertise in Cybersecurity allowed us to carve the direction of our study and ensure the quality of our project.

We would also like to acknowledge the various contributions of cybersecurity and AI research communities. The studies have provided a strong foundation for our work. Additionally, we also acknowledge the open-source tools and datasets that have played a crucial role in our project.

Finally, we thank our friends, college seniors and family for their encouragement and motivation.

IX. Limitations

Although helpful, the tool has some limitations:

- It does not replace full forensic analysis.
- It depends on Autopsy and Volatility installation.
- Dashboard is simple and may lack advanced visualizations.
- Some complex incidents still require manual investigation.

X. Conclusion

The cyber triage automation tool developed in this project makes the initial phase of digital forensics faster and easier. It automates evidence collection, analysis, and reporting using open-source tools. Although it is not a replacement for expert analysis, it helps investigators save time and focus

on deeper investigation. This tool is useful for students, small organizations, and first responders who need a quick understanding of incidents.

XI. References

1. categorization of digital media," *Digital Investigation*, vol. 10, no. 2, pp. 193-204, 2013. ScienceDirect+1

2. G. Hitchcock, N.-A. Le-Khac and M. Scanlon, "Tiered Forensic Methodology Model for Digital Field Triage by Non-Digital Evidence Specialists," arXiv preprint arXiv:1604.03844, 2016. arXiv

3. R. In de Braekt, N.-A. Le-Khac, J. Farina, M. Scanlon and M-T. Kechadi, "Increasing digital investigator availability through efficient workflow management and automation," arXiv preprint arXiv:1708.09053, 2017. arXiv

4. M. Gogia and others, "ML based Digital Forensics Framework," *Journal of Digital Forensics, Security and Law*, vol. 17, no. 2, 2022. Scholarly Commons

5. "Methods and Tools of Digital Triage in Forensic Context," *Symmetry*, vol. 9, no. 4, 2017. MDPI

6. "The Impact of Artificial Intelligence on Digital Forensic," Online Scientific Research (journal), published Dec 2024. Online Scientific Research

7. "Digital Forensics Triage Classification Model using Hybrid Learning Approaches," *IJIRCST*, vol. 10, 2022. IJIR CST+1

8. "A Framework for AI Generated Digital Forensic Code and Tool Testing (AutoDFBench)," ACM, published 2024. ACM Digital Library

9. "AI-Driven Digital Evidence Triage in Digital Forensics: A Comprehensive Review," ResearchGate preprint, 2025. ResearchGate

10. "Creating a Triage Tool to Streamline Digital Forensics Investigation," *IJRASET*, 2025.