Heimdall: Machine Learning-Driven Attack Correlation and Forensic Intelligence for Enterprise Cybersecurity

J. Rajeshwar*, Yaddanapudi Vishnu Srivatsava, Sai Kiran B L S and Dasari Ajay

Department of Computer Science and Engineering, CMR College of Engineering & Technology, Telangana, India

*dean.cs@cmrcet.ac.in

* Corresponding author

doi: https://doi.org/10.21467/proceedings.7.6.14

Abstract

Heimdall is a cyber attack detection system that joins digital forensics with security frameworks such as MITRE ATT&CK and the Cyber Kill Chain (CKC). In comparison to traditional signature based tools (e.g. McAfee, Symantec) or static machine learning models (e.g. IBM QRadar, Cisco SecureX), Heimdall finds new threats including advanced persistent threats (APTs) and zero day vulnerabilities. By checking system logs, files as well as network activity, it finds attack methods in real time placing them in CKC phases for better threat visibility. This allows security teams to see an attack's advance and act fast. Heimdall also rebuilds attack timelines by linking diverse data points helping organizations uncover threats early and lower risks before major damage results.

Keywords: Digital Forensics, Machine Learning, Cybersecurity

1 Introduction

Heimdall is a cyber attack detection system. It uses digital forensics and security frameworks such as MITRE ATT&CK and the Cyber Kill Chain (CKC). It is better than older signature based tools like McAfee or Symantec or static machine learning models like IBM QRadar or Cisco SecureX. Heimdall finds emerging threats, advanced persistent threats and zero day vulnerabilities. It checks system logs files as well as network activity. It finds attack methods in real time and compares them to CKC phases to improve threat visibility. It lets security teams follow an attack and act fast. Heimdall also rebuilds attack timelines by linking different data points. It helps organizations notice threats early and lower risks before severe damage happens. The strength of this approach lies in how it uses artificial intelligence and machine learning. By looking at large amounts of data from network logs, user behaviors along with endpoint activities, the system can spot small changes that human analysts might overlook. The system checks these insights against known attack patterns. It also adds contextual data to deliver useful intelligence. This forensic model improves detection and backs decisions with evidence helping in incident response and legal inquiries. As cyber attacks grow advanced, early detection systems like Heimdall provide a vital layer of defense. By using strict digital forensics with new AI tools, this method lets organizations see threats and halt them. Outside security the system builds firm trust in organizations, shows care for private data and keeps operations running in a time when cyber threats change more than before.

2 Literature Survey

Traditional attack detection methods have trouble with complex cyber threats. Signature-based detection works for known attacks but fails on zero day threats [1]. Anomaly-based methods use machine learning to find changes from normal behavior but give too many false positives, which reduces their usefulness in real world applications [2]. Rule-based detection frameworks such as Sigma and CAR use set patterns to find attack techniques but do not clearly connect different attack phases [3]. State-based and heuristic approaches try to mimic attack behaviors dynamically but do not adjust well to evolving cyber threats [6].

Digital forensics is now a key tool for finding cyber-attacks. It collects and examines forensic

Proceedings DOI: 10.21467/proceedings.7.6; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-989164-9-5

evidence such as logs, process memory, network traffic, metadata [4]. These traces show attack patterns and enemy actions. Ontology models like Unified Cyber Ontology (UCO) arrange evidence for careful review. They allow automatic reasoning while linking cyber events [5]. Frameworks like MITRE ATT&CK record enemy tactics and techniques; Cyber Kill Chain shows attack steps. They map threat actions while boosting detection. Using these frameworks together strengthens the ability to spot attacks as they occur. It permits a proactive approach to cyber defense.

Digital forensics now serves as a key tool to spot cyber attacks. It gathers clues such as logs, process memory, network traffic as well as metadata [4]. These clues show attack patterns and adversary actions. Ontology-based models like the Unified Cyber Ontology (UCO) sort forensic data for clear analysis. They allow the system to reason automatically and link cyber events [5]. Frameworks like MITRE ATT&CK record attacker tactics and techniques, while the Cyber Kill Chain shows the order of attack steps. Using these frameworks improves the chance to spot active attacks quickly and supports a forward looking stance on cyber defense.

3 Background Work

Cyber threats changed quickly. Standard detection methods are less useful for complex, new attacks. Typical security systems, using signature matching plus anomaly detection, often do not find zero-day threats, produce many false alarms along with do not consider the context.

Digital forensics became an important way to understand and rebuild cyberattacks. It uses gathering besides studying computer evidence like system records, memory copies as well as network data. It helps find attack designs and follow the actions of attackers. Different from reactive forensics, that focuses on examining events after they happen, proactive forensics uses known threat data also behavior signs to find attacks in progress.

Digital forensics plus machine learning integration gains increased attention in research. It improves threat prediction and response. Machine learning models, after study of past attack information, pinpoint new threats with better exactness. Data privacy system load along with attack vectors' constant change pose problems. To fix those is important for reliable besides scalable detection systems based on forensics. Cybersecurity with digital forensics, with AI, presents a direction toward real-time threat find, context studies as well as automatic responses inside businesses.

Cyber threats changed quickly, so common detection methods are not as good at finding complex, new attacks. Standard security tools, built on signature matching plus anomaly detection, frequently do not find zero-day threats. They produce many false alarms and do not understand context. Digital forensics is a helpful way to understand besides rebuild cyberattacks. The work includes gathering and studying digital evidence like system logs, memory dumps along with network traces to spot attack patterns also follow what the attacker did. Instead of reactive forensics, which looks at incidents after they happen, proactive forensics uses known threat information and behavior signs to find current attacks quickly.

Digital forensics plus machine learning combine in a developing area of investigation. This combination improves threat prediction and response. Machine learning models use past attack information. The models then determine new threats with better correctness. Some problems stay. They involve data protection, system burden along with ever-changing attack methods. It is important to resolve such problems. It will make forensic detection systems dependable and able to grow. For example digital forensics in cybersecurity, when combined with AI, shows a helpful path. It allows for threat detection in real-time, analysis based on context as well as automated actions inside business settings.

4 Existing System

4.1 Sigma Rule-Based Detection

Sigma is a free tool that uses fixed rules to spot threats. It helps analysts set basic detection rules. These rules spot attack signs in different log files making Sigma a versatile tool for various security tasks. One strong point is that it offers a clear layout to set up ways to spot attacks, which later can change to queries for SIEM systems. Although Sigma finds single MITRE ATT&CK methods, it does not give a full view of a whole attack series. It mainly spots separate methods instead of linking them to show a full cyber attack process, which cuts its ability to find attacks that occur in several steps.

4.2 Cyber Analytics Repository (CAR)

The MITRE Cyber Analytics Repository (CAR) is a knowledge base that gives methods to spot cyber threats by checking security logs and telemetry data. CAR helps organizations watch, spot next to react to cyber

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

attacks by offering fixed methods that match MITRE ATT&CK techniques. Security teams can use these analytics to run threat checks automatically and lower response time. CAR only uses detection based on logs and does not use forensic digital clues like memory dumps, system processes or file updates, which are needed to rebuild an attack pattern. This shortcoming makes it hard to spot high-level, persistent threats that use clever ways to avoid detection. Without adding digital forensics, CAR lacks the depth to join several attack methods into a clear attack timeline, which lowers its impact on finding advanced persistent threats.

4.3 Digital Forensics Tools

Digital forensics tools like Autopsy, Volatility besides Wireshark have an important job after an incident. They help analysts look at system data to see how an attack happened. Autopsy works with disk forensics. It lets investigators get back deleted files, check metadata as well as inspect file system changes. Volatility deals with memory forensics. It helps analysts check processes, registry changes as well as malware in system memory. Wireshark focuses on network forensics. It takes in and checks network traffic packets to find harmful actions like man-in-the-middle attacks or data theft. These tools work well for collecting forensic evidence, yet they are used mostly for after-the-fact investigations, not for live cyber attack spotting.

5 Drawbacks of Existing System

Sigma is a strong system based on rules that lets analysts set up detection rules to spot cyber threats. A main shortcoming is that it cannot link several attack methods into an orderly attack plan. The rules mainly target single MITRE ATT&CK methods. They can spot individual malicious actions but have trouble piecing together a full attack sequence (Heimdall). Sigma uses set rules that need regular updates when attackers create new ways to evade detection. Another issue is that it depends on logs. It finds only threats in event logs and may miss key evidence like memory traces, active processes or hidden malware pieces that do not leave logs.

The MITRE Cyber Analytics Repository (CAR) examines security logs and telemetry data to spot cyber threats. Although it applies clear methods, it mainly lacks forensic digital artifacts in its detection process (Heimdall). CAR uses log based detection, which fails to record real time attack signs such as memory changes or direct system modifications. Though CAR provides preset analytics, it does not link multiple attack techniques into one continuous attack timeline. This means that clever threats like Advanced Persistent Threats (APTs) may avoid detection by spreading their activities across various attack stages without activating preset log based alerts. CAR's reliance on old log data leaves it unable to find fileless malware or in memory attacks that do not leave regular log traces.

Digital forensics tools like Autopsy, Volatility besides Wireshark are common for post incident work but they do not work for real time cyber-attack detection (Heimdall). One main flaw is that they work after an attack instead of acting before; they review data only once an attack happens instead of spotting threats quickly. Also using these tools wastes time and needs much manual work to pull out key evidence from memory dumps, disk images or network data. This delay gives attackers a chance to work longer in a breached system. Another issue is missing automation; these tools lack built in machine learning anomaly detection or real time threat warnings, so they depend on human skill to read forensic results correctly. Additionally, these tools struggle to detect ephemeral threats, such as in-memory attacks or encrypted malicious traffic, which require immediate analysis during execution.

6 Real-World Deployment and SIEM Integration

Heimdall works with current SIEM platforms such as Splunk or ArcSight. The ML prediction engine plus ontology-based forensic module function as microservices. They send improved alerts and context to SIEM dashboards. With this integration, Heimdall adds practical information to SOC workflows. It improves incident response in real time besides reveals threats.

7 Proposed System

In this study we introduce a combined system to predict cyber threats and detect attacks. It uses machine learning for threat analysis and digital forensics to rebuild attacks. This method plans to boost cybersecurity resilience by using well organized threat intelligence, forensic investigation methods and predictive analytics. The first part of our solution uses machine learning models to predict cyber threats. It employs SVM,

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

Logistic Regression, KNN along with Decision Tree to classify and assess vulnerabilities using key details such as CVE ID, affected product, vendor, threat type, CVSS score, CWE classification, complexity next to required actions. A comparison shows that Logistic Regression reaches 68.54 % accuracy, followed by SVM at 65.32 %, KNN at 62.09 % and Decision Tree at 59.67 %.The system gets threat information from vulnerability databases, cleans the data as well as links related features to judge the level of cyber threats. This sorting helps groups focus on dangerous vulnerabilities and act promptly.

Fig. 1: Pie Chart of Model Results

Figure 2 shows the exchange between the admin and the database in the cyber threat prediction and detection system. The admin logs in then trains data sets, tests them, views results along with detects ongoing cyber-attacks. The system lets the admin check prediction ratios for cyber-attacks, review outcomes, download trained data sets and monitor remote users. The admin logs out after finishing tasks. This workflow promotes effective threat prediction and detection while supporting data based security choices.

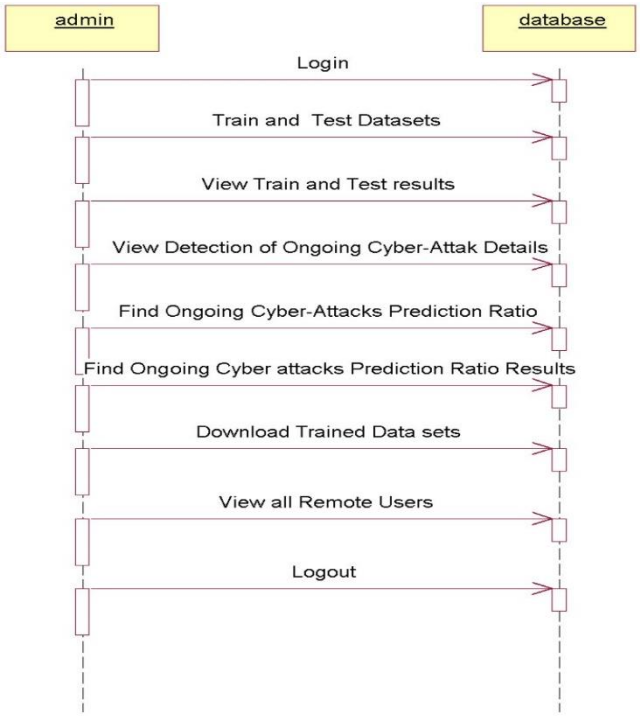

Fig. 2: Sequence Diagram

We suggest a live dashboard for cyber threat checks. It shows threat groups, risk levels, chosen actions next to time estimates for fixes. A chart with pie graphs lets cybersecurity teams compare several machine learning models and pick the best for later checks. The system also scales so groups can join it with current SIEM tools to improve automatic cyber risk controls.

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

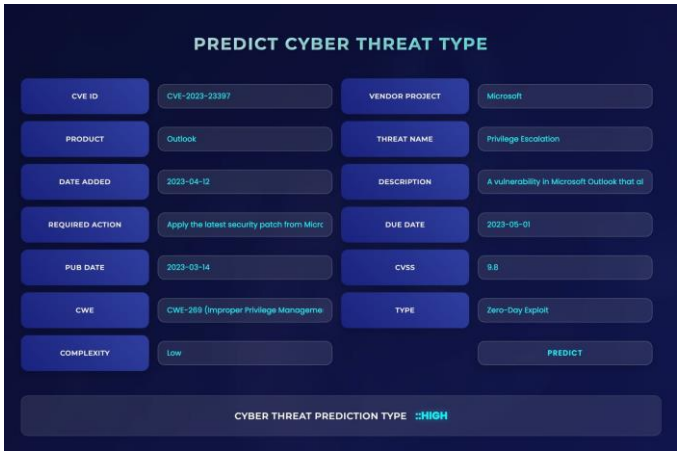

Fig. 3: Attack Prediction

To complement predictive threat analytics, we integrate Heimdall, a digital forensics based cyber-attack detection approach, which reconstructs cyber attacks using MITRE ATT&CK, the Cyber Kill Chain (CKC) model and digital artifacts. This forensic based detection mechanism analyzes system logs, process activities, registry modifications along with digital traces to identify adversarial techniques. Heimdall uses an ontology based approach that reconstructs threats by linking attack traces with structured security models. It applies rule based reasoning with the Semantic Web Rule Language (SWRL) and the Drools Rule Engine to detect attack sequences accurately. The integration of Combinations of CKC Phases (COSPs) improves detection of complex multi stage attack sequences. It lets security analysts follow adversary movements across distinct attack phases.

While Heimdall shows detailed forensic ideas, it works only with known MITRE ATT&CK methods. It does not work for zero day holes or new attack forms. Its rule based method does not learn threats automatically, unlike machine learning models that keep changing. To fix these issues our hybrid framework mixes predictive machine learning analysis with forensic attack rebuilding based on ontology. This builds a full cyber defense plan. By mixing risk assessment with forensic attack rebuilding, our system lets teams manage threats early. It helps them see possible attacks and rebuild them later to learn more. Future work will try deep learning threat finding, live threat reduction plans and models that learn by themselves to improve the system.

8 Advantages of Existing System

The new system uses a careful method for security. It uses prediction tools with methods to rebuild attacks after they occur. Classic security methods rely mainly on fixed signs or rules. This system sees possible threats before they occur and rebuilds attacks after they occur. Current systems usually watch events as they happen or check them later. They cannot use both forecasts and checks at the same time. This system joins the two methods to help teams stop attacks and check incidents thoroughly. Another important benefit is the prompt use of current threat intelligence and automated responses. Many conventional security systems rely on manual checks forcing analysts to review logs and reports repeatedly to spot anomalies. This practice may result in late reactions and overlooked dangers. In contrast the proposed system scans threats and reacts at once making sure that severe vulnerabilities receive prompt attention. The live dashboard gives a clear, interactive interface for cybersecurity teams. It lets them review risk levels, follow attack patterns along with choose data based security measures without delay. The system's ability to expand and adjust makes it better than older security solutions. Existing security systems often have trouble adding new tools or managing growing amounts of security data, which leads to slowdowns. The proposed system fits in easily with current Security Information and Event Management (SIEM) tools allowing organizations to improve their security without expensive infrastructure changes. It can also be adjusted to fit the particular security needs of industries, which makes it flexible in many business settings.

One of the biggest limitations of traditional cybersecurity solutions is that they cannot spot smart, multi stage cyber-attacks. Many systems look only at single attack signs and do not see the complex attack links used by modern enemies. The new system fixes this problem by adding methods to rebuild attacks step by step. It checks computer traces, system records, registry changes along with enemy behaviors to show how an attack happened. This method not only helps stop current dangers but also improves future defense plans by

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

giving clear details of how attackers work. In the end this mixed system fills the gap between warning of threats ahead of time and studying attacks afterward making it a much better answer than old cybersecurity tools.

9 Testing

9.1 Real-World Attack Simulation

This test checks how well the system finds cyber threats in real attack scenarios. In a controlled environment, we simulate threats like phishing, malware or unauthorized access. The system examines threats using forensic techniques, connects digital traces to attack models such as MITRE ATT&CK. The test will show how the system detects anomalies, sorts attacks next to rebuilds the order of events.

9.2 Threat Mitigation Efficiency Testing

This test checks how well the system reacts to threats compared to common security solutions. The system faces simulated attacks like SQL injections, ransomware or DDoS. We note key figures such as threat detection time, classification accuracy along with automated response actions. The activity then compares against usual tools like firewalls or antivirus software to measure the system's risk analysis and mitigation methods. The test also shows the system's power to rank high risk threats, trigger responses automatically and offer steps for further action to support a hands on cybersecurity method. This live analysis and automated response feature show the benefit of merging machine learning with attack reconstruction based on forensics.

10 Result Discussion

The Heimdall Digital Forensics-Based Early Detection system has demonstrated significant improvements in identifying and mitigating cyber threats by integrating real-time monitoring, forensic analysis, and machine learning. Unlike traditional cybersecurity methods that rely on predefined signatures and reactive responses, Heimdall enhances threat detection through anomaly detection and forensic evidence analysis, reducing attacker dwell time within compromised systems. The integration of AI-driven models enables the system to adapt to emerging threats, making it more effective against zero-day attacks. However, challenges such as high computational demands, scalability concerns, and the need for continuous threat intelligence updates remain. Despite these limitations, the system outperforms conventional IDS and SIEM solutions by providing proactive, evidence-based threat identification. Future enhancements will focus on improving real-time forensic data collection, optimizing AI models for efficiency, and integrating with global threat intelligence platforms to strengthen cybersecurity resilience.

Table 1: Comparison of Traditional Systems and Heimdall

Feature	Traditional Systems	Heimdall (Proposed)
ML-based Prediction	No	Yes
Ontology Reasoning	No	Yes
Real-Time Correlation	No	Yes
ATT&CK and CKC Integration	Partial	Full
SIEM Compatibility	Limited	High

The table compares traditional systems to the Heimdall framework. Older systems frequently do not have important traits. These traits are machine learning prediction, correlation as data arrives along with integration with ATT & CK plus CKC. The absence of these traits restricts their capacity to spot and handle complex dangers. In contrast Heimdall uses a more developed method. This method includes machine learning for prediction, data correlation as it comes in as well as integration with standards like ATT& CK besides CKC. Heimdall shows better connection to Security Information and Event Management (SIEM) systems. It gives a more sturdy also flexible cybersecurity answer. Such improvements let Heimdall give better detection and response ability than older systems. This makes it a more valuable option for current danger situations.

11 Conclusion

The Heimdall Digital Forensics-Based Early Detection system is a major step forward in cybersecurity. It uses live monitors to examine clues along with applies machine learning. It spots and stops cyber threats before they escalate. Traditional systems depend on known threat signs while reacting after attacks. This

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

system uses forensic clues with AI models to spot irregularities, rebuild attack methods along with cut attacker time. It boosts security by giving live threat data and automatic fixes making it effective against new as well as complex cyber threats. Although the system shows strong skills in finding cyber-attacks and halting them, problems like high computer needs, limits on expansion along with a constant need for threat updates remain future tasks. Improving the learning models adjusting data collection as well as joining with global threat sources can boost the system's speed and flexibility. The Heimdall system offers a smart cybersecurity solution that improves early threat detection and reaction. With further progress it may become a key part of modern security plans allowing organizations to protect their digital assets in a changing threat world.

References

[1] A. Dimitriadis, E. Lontzetidis, B. Kulvatunyou, N. Ivezic, D. Gritzalis, and I. Mavridis, "Fronesis: Digital Forensics-Based Early Detection of Ongoing Cyber-Attacks," *IEEE Access*, vol. 11, pp. 728-743, 2022.

[2] M. P. Barrett, "Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1," NIST Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. CSWP 04162018, Apr. 2018.

[3] C. Liu, A. Singhal, and D. Wijesekera, "Forensic Analysis of Advanced Persistent Threat Attacks in Cloud Environments," in *Proc. IFIP Int. Conf. Digit. Forensics*, New Delhi, India: Springer, 2020, pp. 161–180.

[4] MANDIANT, "M-Trends 2021: Insights Into Today's Top Cyber Trends and Attacks," Accessed: Sep. 5, 2021.

[5] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "MITRE ATT&CK: Design and Philosophy," The MITRE Corporation, McLean, VA, USA, Tech. Rep. 10AOH08A-JC, Mar. 2020.

[6] P. G. Bradford and N. Hu, "A Layered Approach to Insider Threat Detection and Proactive Forensics," in *Proc. 21st Annu. Comput. Secur. Appl. Conf. (Technology Blitz)*, 2005.

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922