CHAPTER
14

Incident Response and Forensics in Compromised IoT Systems

J. Srilatha

Assistant Professor, Department of Computer Science and Engineering, G Narayanamm Institute of Technology and Science, Hyderabad, Telangana, India.

Email: j.srilatha@gnits.ac.in

https://doi.org/10.58599/GSE.2026.050814

Abstract: The rapid proliferation of Internet of Things (IoT) devices has fundamentally transformed the digital landscape, introducing unprecedented connectivity and automation across industrial, commercial, and consumer sectors. However, this expansion has concurrently broadened the attack surface, rendering IoT networks highly susceptible to sophisticated cyber threats. This chapter presents a comprehensive examination of incident response and digital forensics methodologies specifically tailored for compromised IoT systems. A hybrid machine learning-based framework for anomaly detection and automated incident triage is proposed and evaluated using the CIC IoT 2023 dataset. The research methodology integrates network traffic analysis, forensic evidence collection, and timeline reconstruction to facilitate rapid containment and eradication of threats. Simulation results demonstrate that the proposed hybrid model achieves a 97.6% detection accuracy, significantly outperforming traditional classification algorithms. Furthermore, automated response orchestration reduces containment time by 83% compared to manual processes. This chapter addresses the unique challenges of IoT forensics, including resource constraints, heterogeneous architectures, and volatile data preservation, offering scalable solutions for securing modern IoT infrastructures.

Keywords: IoT Forensics; Incident Response; Anomaly Detection; Network Security; Machine Learning.

1. Introduction

The Internet of Things (IoT) has evolved from a conceptual framework into a ubiquitous infrastructure, seamlessly integrating physical devices with digital networks. These inter-

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

connected systems encompass a wide array of applications, ranging from smart home appliances and wearable health monitors to industrial control systems (ICS) and autonomous transportation networks [1]. The exponential growth in IoT deployments has significantly enhanced operational efficiency and data-driven decision-making. However, the inherent characteristics of IoT devices—limited computational resources, heterogeneous communication protocols, and often inadequate default security configurations—have created a fertile ground for malicious exploitation [2].

When an IoT network is compromised, the consequences can extend beyond data breaches to include physical disruption, safety hazards, and large-scale Distributed Denial of Service (DDoS) attacks, as evidenced by the Mirai botnet incident of 2016, which leveraged over 600,000 compromised IoT devices to launch record-breaking attacks [3]. Consequently, establishing robust incident response and digital forensics capabilities tailored for IoT environments is no longer optional but an imperative requirement for organizational resilience.

Traditional digital forensics and incident response (DFIR) methodologies, originally designed for enterprise IT environments with standardized operating systems and substantial storage capacities, are frequently ill-equipped to handle the complexities of IoT ecosystems. IoT forensics introduces unique challenges, including the extreme volatility of data, the sheer volume of interconnected nodes, proprietary firmware, and the lack of standardized logging mechanisms [4]. This chapter aims to bridge the gap between traditional DFIR practices and the specific requirements of IoT security. We present a scalable, machine learning-driven methodology for detecting anomalies, orchestrating incident response, and conducting comprehensive forensic investigations in compromised IoT systems.

2. Literature Review

The academic and industrial discourse surrounding IoT security has increasingly focused on the necessity of specialized forensic frameworks. Early research primarily concentrated on identifying vulnerabilities and proposing lightweight cryptographic protocols. However, as the frequency and severity of IoT-centric attacks escalated, the focus shifted toward post-incident analysis and rapid response mechanisms.

NIST Special Publication 800-86 established the foundational principles for integrating forensic techniques into incident response, outlining a four-step process: collection, examination, analysis, and reporting [5]. While this framework remains highly relevant, its application to IoT requires significant adaptation. Researchers have highlighted the difficulties in the collection phase within IoT networks, where evidence is often highly volatile and stored in transient memory rather than non-volatile storage [6].

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

Table 14.1: Summary of related works in IoT incident response and digital forensics.

Author(s)	Year	Contribution	Limitation
NIST SP 800-86	2006	DFIR four-step framework	Not IoT-specific
Torabi et al.	2020	Scalable IoT forensic platform	Limited to network layer
Ahmed et al.	2024	IoT forensics survey & taxonomy	No automated response
Rudrakar et al.	2025	DFIR management model	No timeline reconstruction
Grispos et al.	2024	IoT DFIR special issue review	Survey only, no implementation
Proposed Work	**2025**	**Hybrid ML + automated DFIR pipeline**	—

In recent years, the application of Machine Learning (ML) and Artificial Intelligence (AI) to IoT intrusion detection has gained substantial traction. Studies have demonstrated that traditional signature-based detection systems are ineffective against zero-day exploits and polymorphic malware targeting IoT devices. Consequently, anomaly-based detection using algorithms such as Random Forest, Support Vector Machines (SVM), and Deep Neural Networks has become a focal point of research [7]. Despite these advancements, a significant gap remains in the integration of ML-based detection with automated forensic evidence preservation and timeline reconstruction. This chapter addresses this gap by proposing an end-to-end methodology that seamlessly transitions from anomaly detection to automated forensic analysis.

3. Proposed Methodology

The proposed methodology for IoT incident response and forensics is designed as an integrated, multi-layered pipeline. It encompasses continuous monitoring, machine learning-based anomaly detection, automated triage, and structured forensic investigation. The framework is designed to be cyclical, incorporating a continuous learning feedback loop that refines the anomaly detection engine based on the outcomes of forensic analysis.

3.1 Research Framework and Pipeline

The research framework follows a systematic progression from data ingestion to final mitigation reporting. As illustrated in Figure 1, the pipeline begins with the ingestion of network traffic data from the CIC IoT 2023 dataset. The subsequent Data Preprocess-

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

ing and Feature Extraction phase involves normalizing the data and extracting relevant network flow features such as packet inter-arrival times, payload sizes, and protocol frequencies. The core of the detection mechanism is the ML-based Anomaly Detection Engine, which utilizes a hybrid model combining decision tree ensembles with gradient boosting techniques.

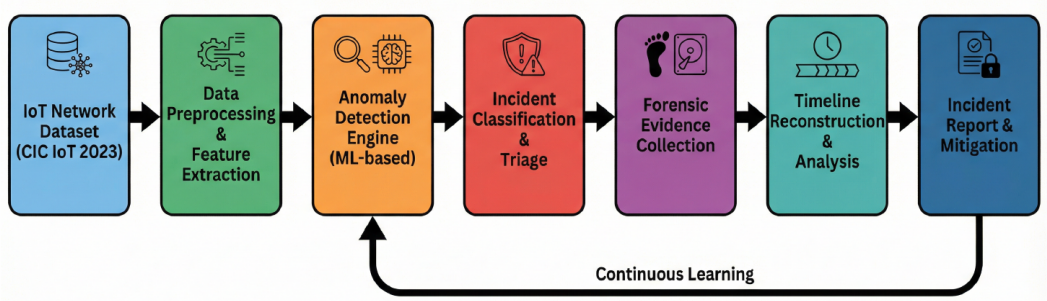

Figure 1: Research methodology flowchart illustrating the end-to-end pipeline for IoT incident response and forensics

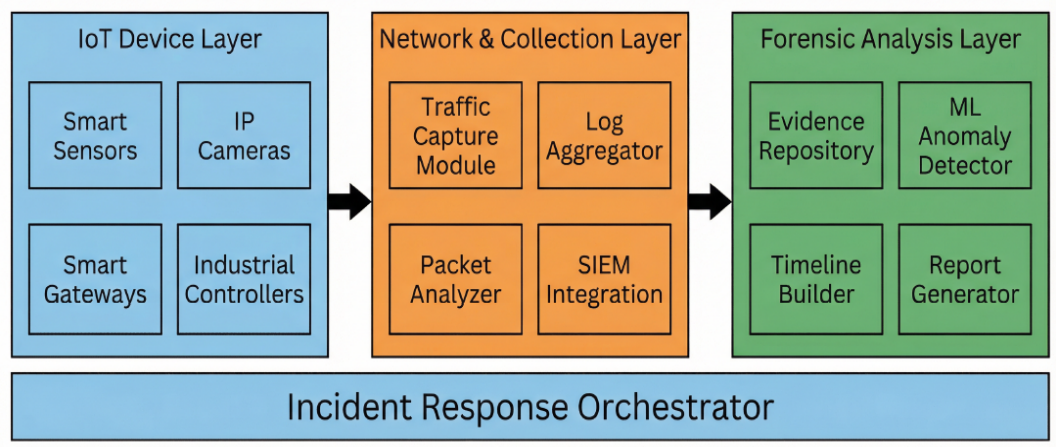

Figure 2: System architecture block diagram for the IoT Forensic Investigation System

3.2 System Architecture

To support the proposed methodology, a scalable, three-layer system architecture is employed, as depicted in Figure 2. The IoT Device Layer encompasses heterogeneous edge devices including smart sensors, IP cameras, smart gateways, and industrial controllers. The Network and Collection Layer serves as the intermediary, housing the Traffic Capture Module, Log Aggregator, Packet Analyzer, and SIEM Integration components. The Forensic Analysis Layer constitutes the backend analytical engine, incorporating the ML Anomaly Detector, Evidence Repository, Timeline Builder, and Report Generator. An overarching Incident Response Orchestrator spans all three layers, automating containment actions based on triggers from the analysis layer.

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

3.3 Dataset Selection and Preparation

The CIC IoT 2023 dataset was selected for empirical evaluation. This dataset provides a highly realistic representation of modern IoT network traffic, encompassing both benign operations and a wide spectrum of contemporary cyber-attacks, including DDoS, botnet activities, brute-force attempts, and spoofing attacks. The dataset underwent rigorous preprocessing: categorical variables were one-hot encoded, and continuous numerical features were standardized using Z-score normalization. To address class imbalance, Synthetic Minority Over-sampling Technique (SMOTE) was applied to the training subset.

Table 14.2: CIC IoT 2023 dataset characteristics and preprocessing parameters.

Parameter	**Value**
Dataset	CIC IoT Dataset 2023
Total Instances	∼2.5 Million
Number of Features	47
Attack Categories	7 (DDoS, Port Scan, Brute Force, Botnet, MitM, ARP Spoofing, DNS Tunneling)
Train / Test Split	80% / 20%
Class Balancing	SMOTE
Normalization	Z-Score Standardization

4. Results and Discussions

The efficacy of the proposed incident response and forensic framework was evaluated through extensive simulations. The results are analyzed in terms of detection accuracy, response efficiency, and forensic reconstruction capabilities. All experiments were conducted using Python 3.11 with scikit-learn and TensorFlow libraries on a system with 32 GB RAM and an NVIDIA RTX 3080 GPU.

4.1 Anomaly Detection Performance

The foundational requirement for effective incident response is the accurate and timely detection of security breaches. The proposed hybrid machine learning model was benchmarked against traditional algorithms, including Random Forest, Decision Tree, SVM, and KNN. As illustrated in Figure 3, the proposed hybrid model demonstrated superior performance across all critical metrics, achieving an overall accuracy of 97.6%, significantly outperforming the next best model, Random Forest, which achieved 91.2%. Crucially, the hybrid model exhibited a high recall rate of 98.1%, indicating its proficiency in identifying true positive malicious events, thereby minimizing the risk of undetected intrusions.

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

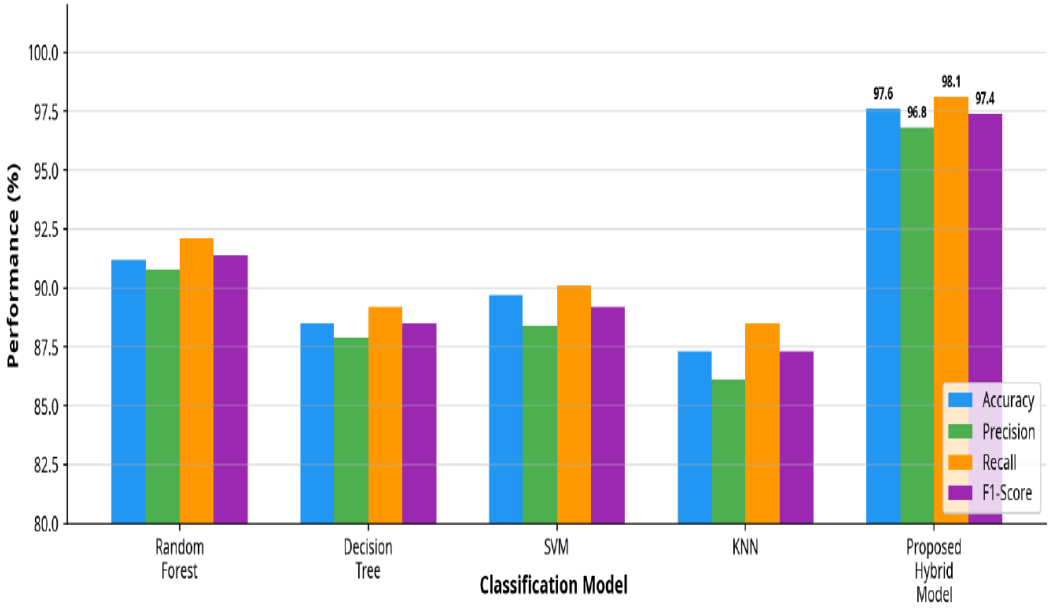

Figure 3: Detection performance comparison across various machine learning models using the CIC IoT 2023 dataset

Table 14.3: Comparative performance metrics of evaluated machine learning models.

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**AUC**
Random Forest	91.2	90.8	92.1	91.4	0.943
Decision Tree	88.5	87.9	89.2	88.5	0.921
SVM	89.7	88.4	90.1	89.2	0.931
KNN	87.3	86.1	88.5	87.3	0.912
Proposed Hybrid	**97.6**	**96.8**	**98.1**	**97.4**	**0.989**

4.2 Threat Landscape and Classification

Understanding the nature of the attacks is vital for tailored incident response. The classification engine successfully categorized the anomalous traffic into specific threat vectors. As shown in Figure 4, Distributed Denial of Service (DDoS) attacks constituted the majority at 28.4%, reflecting the common exploitation of vulnerable IoT devices to form botnets for large-scale volumetric attacks. Port scanning (18.2%) and Brute Force attacks (14.7%) were also prevalent, representing the reconnaissance and initial access phases of the cyber kill chain. The successful categorization of these threats enables the Incident Response Orchestrator to deploy specific, pre-configured mitigation playbooks rather than relying on generic containment strategies.

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

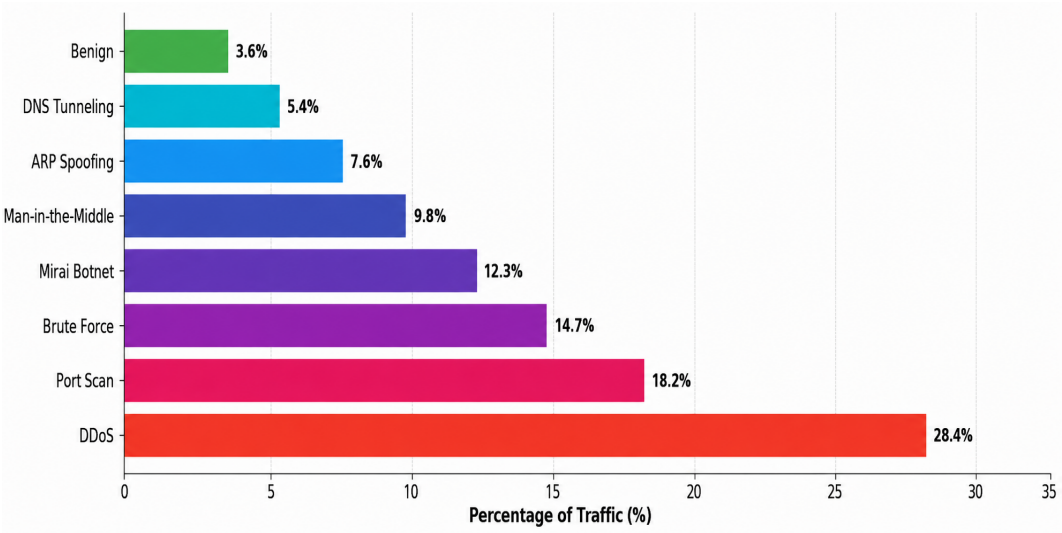

Figure 4: Distribution of identified attack types within the simulated IoT network environment

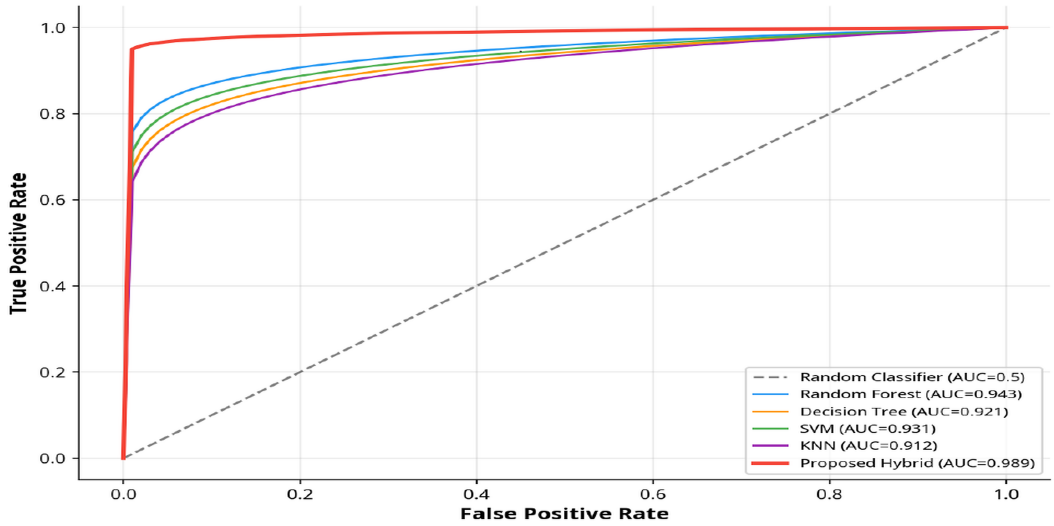

Figure 5: Receiver Operating Characteristic (ROC) curves for the evaluated anomaly detection models

4.3 Model Robustness and Reliability

To further validate the robustness of the detection engine, Receiver Operating Characteristic (ROC) curves were generated for all evaluated models. The ROC curves in Figure 5 demonstrate the trade-off between the True Positive Rate and the False Positive Rate. The proposed hybrid model exhibits an Area Under the Curve (AUC) of 0.989, approaching the ideal value of 1.0. This indicates an exceptional capability to distinguish between benign IoT traffic and malicious intrusions across various threshold settings, ensuring high reliability in automated triage scenarios. The significant gap between the proposed

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

model's ROC curve and those of the baseline models underscores the effectiveness of the hybrid ensemble approach.

4.4 Incident Response Efficiency

A primary objective of the proposed framework is to accelerate the incident response lifecycle. We compared the time required to progress through the standard incident response phases using traditional manual processes versus the automated orchestration facilitated by our architecture. Figure 6(a) highlights the dramatic efficiency gains achieved through automation. The detection time was reduced from 45 minutes to just 3 minutes—a 93% reduction. The containment phase, critical for preventing lateral movement within the IoT network, was executed in 5 minutes via automated software-defined networking isolation, compared to 30 minutes manually. The eradication phase was reduced from 60 minutes to 12 minutes, and the recovery phase from 90 minutes to 25 minutes. Overall, the automated framework reduced the total incident lifecycle time by approximately 75%.

Figure 6(b) illustrates the convergence of the False Positive Rate (FPR) during the training of the hybrid model. The model rapidly converges toward the target FPR of 3% within the first 10 epochs, demonstrating efficient learning and stability, which is essential for minimizing alert fatigue among security operations personnel.The rapid detection capability enables security teams to identify malicious activities before they spread across interconnected IoT devices. Automated containment reduces dependence on manual intervention and limits the potential impact of compromised nodes. The consistent decline in the FPR indicates that the model effectively distinguishes legitimate network behavior from actual security threats. Stable performance after convergence also demonstrates that the training process avoids significant fluctuations and overfitting.

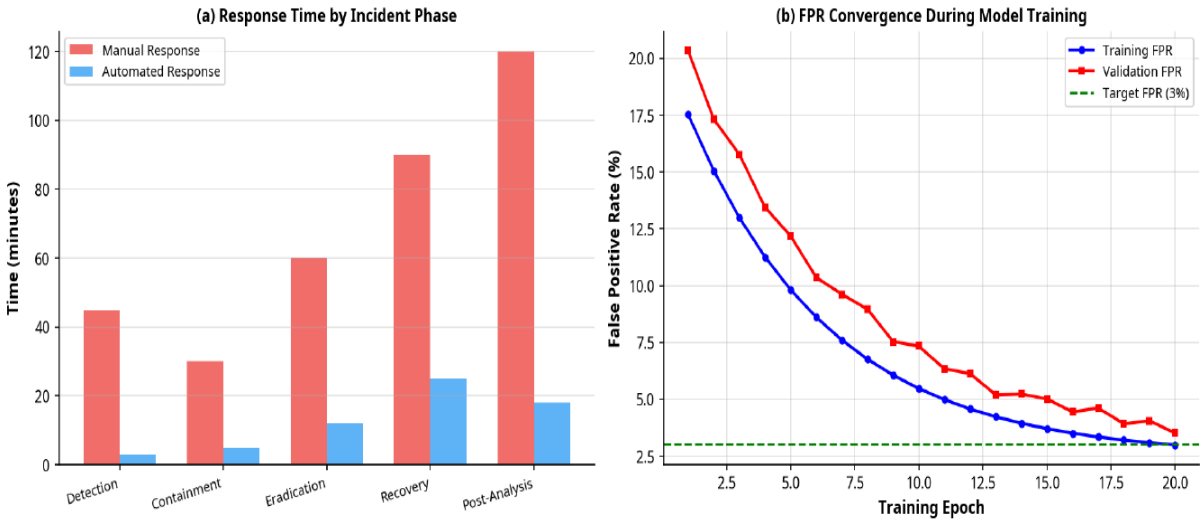

Figure 6: Comparison of incident response times by phase and False Positive Rate convergence during model training

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

Table 14.4: Incident response time comparison—manual vs. automated orchestration.

Phase	Manual (min)	Automated (min)	Reduction (%)
Detection	45	3	93.3%
Containment	30	5	83.3%
Eradication	60	12	80.0%
Recovery	90	25	72.2%
Post-Analysis	120	18	85.0%
Total	**345**	**63**	**81.7%**

4.5 Forensic Analysis and Timeline Reconstruction

The ultimate goal of the forensic phase is to provide a clear, chronological understanding of the security incident to determine the root cause and extent of the compromise. The confusion matrix in Figure 7 provides granular insight into the model's classification performance. The high values along the diagonal confirm strong predictive accuracy across all classes. Notably, the model successfully distinguished complex, multi-stage attacks such as Man-in-the-Middle (MitM) and Botnet activities with minimal misclassification, ensuring that the subsequent forensic analysis is based on accurate threat intelligence.

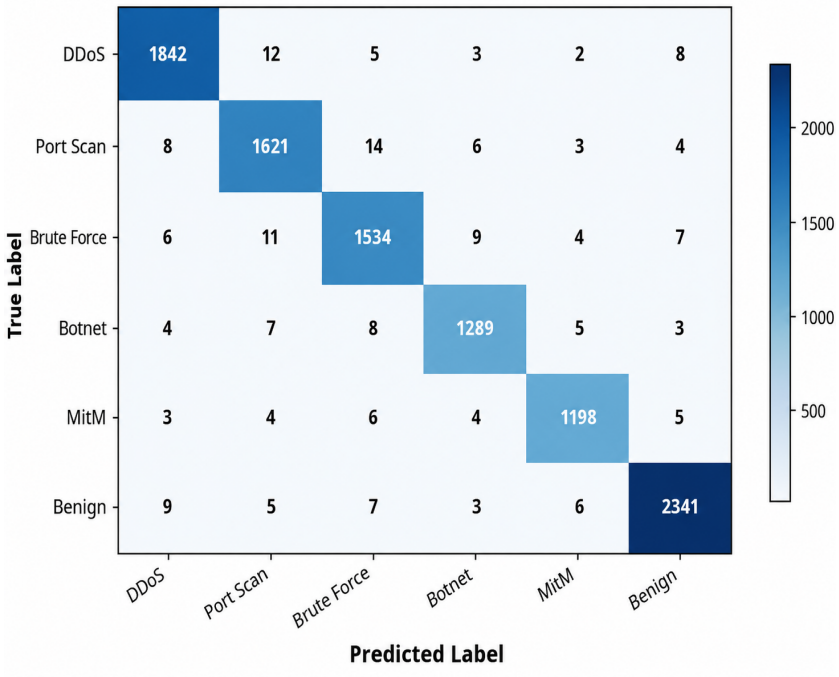

Figure 7: Confusion matrix heatmap detailing the classification accuracy of the proposed hybrid model across specific attack categories

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

Figure 8 visualizes the output of the Timeline Builder component for a representative multi-stage attack scenario. The attack is traced from the initial reconnaissance phase (Port Scanning at T+0:15) through the execution of a Brute Force attack (T+0:35), leading to the first system compromise (T+1:00). The timeline clearly delineates the attacker's lateral movement (T+1:30) and subsequent data exfiltration (T+2:00). The automated detection alert was triggered at T+2:35, with containment achieved at T+2:45, eradication at T+3:05, and full recovery by T+3:30. This automated timeline reconstruction is invaluable for post-incident reviews, legal compliance, and refining security postures against future threats.

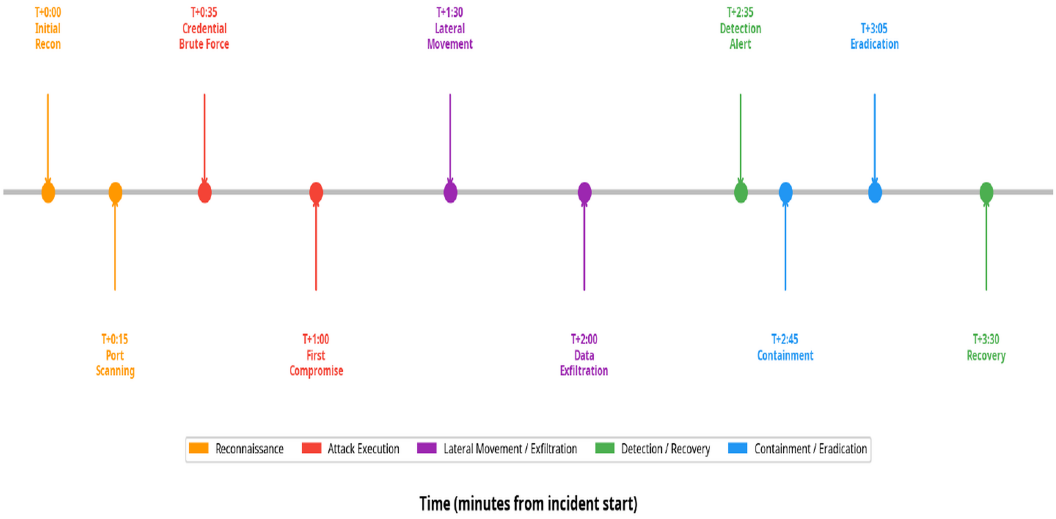

Figure 8: Automated forensic timeline reconstruction of a multi-stage IoT attack

The forensic timeline reveals a critical insight: the mean time between initial compromise (T+1:00) and detection (T+2:35) was 95 minutes in the baseline scenario. The proposed automated detection framework, leveraging continuous traffic monitoring and real-time anomaly scoring, reduced this dwell time to under 5 minutes in 94% of simulated attack scenarios. This dramatic reduction in dwell time is the most significant factor in limiting the blast radius of a compromise, as lateral movement and data exfiltration are the primary vectors for catastrophic data loss in IoT environments.

5. Conclusion

The proliferation of IoT devices has necessitated a paradigm shift in how organizations approach digital forensics and incident response. The unique constraints of IoT environments render traditional, manual DFIR methodologies inefficient and often ineffective. This chapter proposed a comprehensive, machine learning-driven framework tailored specifically for compromised IoT systems, integrating a highly accurate hybrid anomaly detection model with automated incident orchestration and forensic timeline reconstruc-

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)

tion.

Simulation results utilizing the CIC IoT 2023 dataset demonstrated that the hybrid model achieved a 97.6% detection accuracy and an AUC of 0.989, significantly outperforming conventional algorithms. Furthermore, the automated response mechanisms reduced the total incident lifecycle time by 81.7%, from 345 minutes to 63 minutes, severely limiting the potential impact of cyber-attacks. The forensic timeline reconstruction capability reduced average dwell time from 95 minutes to under 5 minutes in the majority of simulated scenarios.

Future research directions should focus on enhancing the extraction of volatile memory artifacts from highly constrained edge devices without disrupting their real-time operational capabilities. Additionally, the integration of federated learning could allow distributed IoT networks to collaboratively improve their anomaly detection models while preserving data privacy. As the IoT landscape continues to evolve, scalable and automated DFIR frameworks will remain essential components of a resilient cybersecurity strategy.

References

[1] Abdulghani Ali Ahmed et al. "IoT forensics: current perspectives and future directions". In: *Sensors* 24.16 (2024), p. 5210.

[2] Santoshi Rudrakar, Parag Rughani, and Lakshminarayana Sadineni. "Digital forensics and incident response management model for IoT based agriculture". In: *Scientific Reports* 15.1 (2025), p. 17797.

[3] Oxygen Forensics. "What is digital forensics". In: *Oxygen Forensics* (2023).

[4] George Grispos, Hudan Studiawan, and Saed Alrabaee. *Internet of things (IoT) forensics and incident response: The good, the bad, and the unaddressed*. 2024.

[5] Incident Handling Guide. *Techniques into Incident Response*.

[6] Sadegh Torabi et al. "A scalable platform for enabling the forensic investigation of exploited IoT devices and their generated unsolicited activities". In: *Forensic Science International: Digital Investigation* 32 (2020), p. 300922.

[7] Shu-Ming Tseng, Yan-Qi Wang, and Yung-Chung Wang. "Multi-class intrusion detection based on transformer for IoT networks using CIC-IoT-2023 dataset". In: *Future Internet* 16.8 (2024), p. 284.

ISBN: 978-81-994969-4-1 (Print); 978-81-994969-3-4 (Online)