Cyber Security and Network Management E-ISSN: 3124-0054

Research Article

Digital Forensics and Automated Incident Response Framework Leveraging Big Data Analytics and Real-Time Network Traffic Profiling in Heterogeneous Cyber Environments

Danang [1*], Zaenal Mustofa [2], Irlon [3]

[1] Universitas Sains dan Teknologi Komputer, Indonesia; e-mail : danang150787@gmail.com
[2] Universitas Negri Yogyakarta, Indonseia; e-mail : zaenalmustofa@uny.ac.id
[3] Institut Teknologi Budi Utomo, Indonesia; e-mail : dahil.irlon@gmail.com
* Corresponding Author : Danang

Abstract: The increasing complexity and scale of modern cybersecurity threats necessitate the development of advanced systems capable of efficiently detecting, analyzing, and mitigating incidents in real time. This paper proposes an automated framework for digital forensics and incident response that leverages big data analytics and real-time network traffic profiling. The framework integrates cutting-edge technologies, including Apache Spark for real-time data processing and Hadoop for scalable data storage, combined with machine learning models like LSTM and Autoencoders to detect anomalies and threats in network traffic. By automating the process of incident detection and response, this framework significantly reduces the time required to identify threats and improves the accuracy of forensic evidence correlation across heterogeneous network environments. The study highlights the advantages of using machine learning models and big data tools to address the limitations of traditional manual and semi-automated systems, which often struggle to keep pace with large-scale data generation. Testing results demonstrate that the proposed framework can handle large data volumes efficiently, providing real-time, actionable insights with significantly reduced response times. Additionally, the framework improves forensic analysis by enabling the correlation of evidence from different devices and protocols, making it more effective than traditional methods in identifying the root cause of security incidents. However, challenges related to data heterogeneity, scalability, and system integration were encountered during testing. The proposed framework holds promise for significantly enhancing the efficiency and effectiveness of cybersecurity operations, with future work focusing on further integration of advanced AI techniques and machine learning models for dynamic and adaptive incident response.

Keywords: Big Data; Digital Forensics; Incident Response; Network Traffic; Real-Time Analysis.

Received: 21, November 2025
Revised: 10, December 2025
Accepted: 29, December 2025
Published: 19, January 2026
Curr. Ver.: 19, January 2026

1. Introduction

The field of digital forensics and incident response (DFIR) plays a crucial role in maintaining cybersecurity, especially when handling large-scale, heterogeneous network environments. However, conventional DFIR processes face significant challenges that hinder their efficiency and effectiveness. The rapid evolution of technology, including the proliferation of mobile devices, Internet of Things (IoT), and cloud computing, has exacerbated these challenges, making traditional tools inadequate for handling the complex demands of modern networks [1]. One of the major issues is the scarcity of resources, where the limited availability of computational power and storage capacity hampers the ability to process vast amounts of data effectively [2].

The volume and complexity of data generated in contemporary network environments present another obstacle. The data is not only massive but also diverse in nature, which poses significant difficulties for forensic investigators who must analyze a variety of data types in real-time [3]. This challenge is further compounded by heterogeneity; the diversity of devices and data formats complicates the acquisition, preservation, and analysis of digital evidence [4]. Additionally, ensuring privacy and trust in multi-tenant environments, particularly in cloud computing scenarios, remains a critical concern, as securing sensitive information from unauthorized access is more complex in distributed systems [5].

Furthermore, the velocity and variety of data generated in modern networks require advanced tools and methodologies that can quickly and accurately process information [6]. Traditional DFIR tools, however, often fail to meet these needs, resulting in prolonged incident identification times and inefficient evidence collection, which can compromise the investigation process [7]. As a result, critical incidents may go undetected or unresolved for extended periods, leading to further security risks.

Conventional DFIR processes are often slow, inefficient, and inadequate when dealing with the vast amounts of data generated in large-scale, heterogeneous environments. This inadequacy results in several key issues: delayed incident identification, inefficient evidence collection, and a limited scope of existing forensic tools. Moreover, the financial burden of purchasing advanced forensic tools and the legal complexities of remote and live forensics further exacerbate these challenges [8].

As cybersecurity threats become more sophisticated and large-scale, traditional digital forensics and incident response (DFIR) processes struggle to keep up with the demands of modern, heterogeneous network environments. The growing complexity of technology, including mobile devices, cloud computing, and the Internet of Things (IoT), further amplifies these challenges. Conventional DFIR tools often lack the capacity to efficiently manage the massive amounts of data generated in real-time, leading to delays in incident detection and response. This study proposes an automated digital forensics and incident response framework that integrates big data analytics and real-time network traffic profiling to enhance the detection, analysis, and mitigation of cybersecurity threats, making the process more efficient, accurate, and proactive [9].

The significance of this study lies in its potential to improve the efficiency, speed, and accuracy of cybersecurity incident response. By utilizing real-time detection and mitigation, the proposed framework can significantly reduce the time required to identify and respond to threats. An Automated Incident Response System (AIRS) using techniques like Random Forest for anomaly detection has demonstrated a high accuracy rate of 94.7%, significantly lowering false positives and ensuring more robust cyber defense [9]. Additionally, automated log monitoring approaches that can detect missing log values and generate security alerts will streamline the incident response process and enhance Digital Forensic Readiness (DFR), reducing operational disruptions [10].

Speed is another crucial factor in modern cybersecurity environments, and this framework uses big data analytics to process vast amounts of network traffic data rapidly. Tools like Hadoop allow for efficient handling of large-scale data, but their complexity can present challenges [11]. Moreover, integrating machine learning models such as Reinforcement Learning (RL) and Graph Neural Networks (GNNs) helps filter out irrelevant data, focusing on true positives and speeding up the detection process [12]. This combination of big data and machine learning ensures that the framework can keep pace with the dynamic nature of modern cyber threats, enabling quicker response times and more accurate results.

Finally, the integration of optimized models and proactive forensic mechanisms further improves the accuracy and reliability of the system. Techniques such as Optuna for hyperparameter optimization have shown promise in fine-tuning detection models to enhance their effectiveness [9]. By incorporating frameworks like MITRE ATT&CK, the study also ensures that forensic processes are not only efficient but also legally sound, preserving the integrity of digital evidence [13]. The framework's ability to handle diverse data from IoT, cloud environments, and social networks offers a comprehensive approach to cybersecurity, ensuring that it can predict and prevent attacks in real-time while maintaining forensic soundness [14].

2. Literature Review

Digital Forensics and Incident Response: Overview of Traditional Methods and Their Limitations

Traditional methods in digital forensics and incident response (DFIR) involve a detailed and systematic approach to identifying, preserving, and analyzing digital evidence, typically from computer systems. In computer forensics, the process involves the identification, preservation, examination, and analysis of data from digital devices. However, these methods are often time-consuming and require specialized expertise [15]. Additionally, incident response typically focuses on restoring normal service after a cybersecurity incident, often prioritizing rapid recovery over detailed forensic analysis [16]. The limitations of traditional DFIR methods include the slow pace due to the increasing capacity of storage media [17], the complexity of modern cyber threats, and the lack of a systematic approach to analyzing vast volumes of data, which results in slower decision-making and inefficiency in investigations [18]. Furthermore, the ad-hoc models used in existing forensic tools often lack standardization, which hinders their reliability in legal proceedings [19].

Big Data Analytics in Cybersecurity: Enhancing Digital Forensics and Incident Response

Big data analytics has emerged as a powerful tool to address the limitations of traditional methods in cybersecurity and digital forensics. The vast amounts of data generated by IoT, cloud computing, and social networks pose a significant challenge to traditional DFIR methods, which are often not designed to handle such large-scale data [20]. Big data analytics enables the processing of enormous datasets, improving the early detection and prevention of cyberattacks. Predictive analytics can analyze historical data and identify patterns that signal potential future threats, offering proactive measures rather than reactive ones [21]. Moreover, big data analytics enhances agility in incident response by providing more flexible, innovative, and faster ways to detect and mitigate cyber threats [22]. Techniques like deep learning can significantly speed up investigations by filtering relevant evidence from large volumes of data, improving the accuracy of threat detection and reducing the time spent on non-relevant data [14].

The integration of big data platforms such as Hadoop allows for real-time analysis of large volumes of network data, enabling faster detection of anomalies and more accurate identification of threats [11]. However, there are challenges associated with the complexity of managing big data environments, especially regarding data privacy and ethical concerns that arise from the use of AI and machine learning in cybersecurity [19]. Additionally, environments like Hadoop, which are designed for large-scale data processing, present difficulties in conducting forensic investigations due to their distributed nature and component complexity [23].

Real-Time Network Traffic Profiling: Previous Works on Network Traffic Profiling and How It Aids in Detecting Threats in Real-Time

Network traffic profiling (NTA) plays a vital role in detecting security breaches by monitoring and analyzing real-time data traffic on networks. Network traffic profiling allows administrators to identify patterns and anomalies that could indicate a cyberattack. Techniques such as machine learning and statistical analysis have been effectively used to detect abnormal network behavior in real time, enabling early identification of threats [24]. Visualization techniques, such as stacked histograms and behavioral modeling, have also been used to present network activity data in a manner that is easier for administrators to interpret, improving their ability to spot potential attacks in real time [25].

Advanced methods like network watermarking provide an active form of network traffic analysis, enabling the tracking and identification of traffic flows, which overcomes the limitations of passive traffic analysis [23]. The integration of machine learning models such as Long Short-Term Memory (LSTM) networks and Autoencoders offers robust anomaly detection capabilities, which are essential for improving response times and detection accuracy [25]. Moreover, hybrid techniques, which combine signature-based detection with anomaly detection, offer a comprehensive solution to identifying both known and unknown threats [24]. Real-time anomaly detection using techniques like K-means clustering has shown

promising results in detecting network anomalies and enhancing the effectiveness of network traffic analysis [26].

Real-time network traffic profiling is a technique used to continuously analyze network traffic patterns in order to detect abnormal activities that may indicate cyber threats. By monitoring traffic flows in real time, security systems can rapidly identify anomalies and respond to potential attacks before they escalate into more severe incidents [27].

Recent research has also explored image-based traffic analysis techniques for improving network attack detection. In this approach, network traffic data are transformed into visual representations, enabling deep learning models to recognize complex patterns that are difficult to detect through traditional statistical methods. This technique has proven effective in identifying and mitigating DDoS attacks in Software Defined Networking (SDN) environments, where network programmability allows for adaptive and dynamic mitigation strategies [27].

Automated Systems in Cybersecurity: Review of Current Automated and Semi-Automated Systems, with an Emphasis on Their Strengths and Weaknesses in Handling Cyber Incidents

Automated and semi-automated systems have become integral components of modern cybersecurity operations. These systems leverage technologies like machine learning, big data analytics, and orchestration to streamline the detection, analysis, and response to cybersecurity threats. One of the major frameworks in this domain is Security Orchestration, Automation, and Response (SOAR) systems, which combine automation, machine learning, and orchestration to improve incident response efficiency [28]. SOAR systems enable seamless integration of security tools, allowing for faster response times and reducing human error. Additionally, big data integration plays a pivotal role in enhancing real-time threat detection and response. Technologies such as Apache Kafka for data ingestion and Apache Flink for stream processing, when combined with machine learning models like Long Short-Term Memory (LSTM) and Autoencoders, significantly improve the accuracy and speed of threat identification [20].

However, these systems also have limitations. One key weakness is their complexity in integrating data from multiple sources. Combining flow-level and packet-level data for real-time analysis often requires sophisticated frameworks capable of handling large-scale data while maintaining high accuracy [29]. Furthermore, the effectiveness of machine learning models heavily relies on the quality and volume of data, meaning poor data quality can significantly hinder their performance [30]. Another challenge is the ethical concerns surrounding the use of AI and machine learning in cybersecurity. Over-reliance on automated systems can create vulnerabilities, such as the potential misuse of AI by malicious actors [31]. Thus, while automated systems show immense promise, they are not without their flaws, particularly in terms of integration and reliance on high-quality data.

Research Gaps: Identify the Gap that This Study Intends to Fill by Integrating Big Data and Real-Time Traffic Profiling into a Comprehensive Automated Framework

Despite significant advancements in automated cybersecurity systems, there remain **research gaps** that need to be addressed. Traditional security systems struggle to cope with the growing complexity and scale of modern cyber threats, which often involve large data volumes and heterogeneous networks [20]. One of the primary gaps is the integration of big data analytics and real-time traffic profiling for intrusion detection. While there have been isolated efforts to analyze traffic at either the flow or packet level, the integration of both for real-time threat detection remains largely unexplored. This gap presents an opportunity for developing innovative frameworks that can better handle the complexity of modern, interconnected networks [29].

This study aims to fill this gap by developing a comprehensive automated framework that integrates big data analytics with real-time network traffic profiling. By leveraging technologies like Apache Spark for real-time data processing and combining it with advanced machine learning models for predictive analytics, the proposed framework will enhance the speed, accuracy, and scalability of incident response [20]. Moreover, the framework will incorporate adaptive learning techniques, allowing the system to continuously improve its decision-making processes and better respond to emerging threats [32]. This integrated

approach aims to address current challenges in scalability and efficiency by providing a more comprehensive and dynamic solution for real-time network security and incident response.

3. Proposed Method

The proposed framework integrates big data analytics, network traffic profiling, and rule-based automation to create a scalable system for real-time incident detection and response in complex cyber environments. By utilizing Hadoop and Apache Spark for large-scale data processing and real-time analytics, the framework can efficiently handle vast volumes of network data, ensuring fast threat detection. Network traffic profiling continuously monitors and analyzes data to identify anomalies, while rule-based automation triggers predefined responses to mitigate threats automatically, reducing incident response times. The system's architecture includes layers for data ingestion, real-time processing, and automated decision-making, enabling seamless, effective cybersecurity operations.

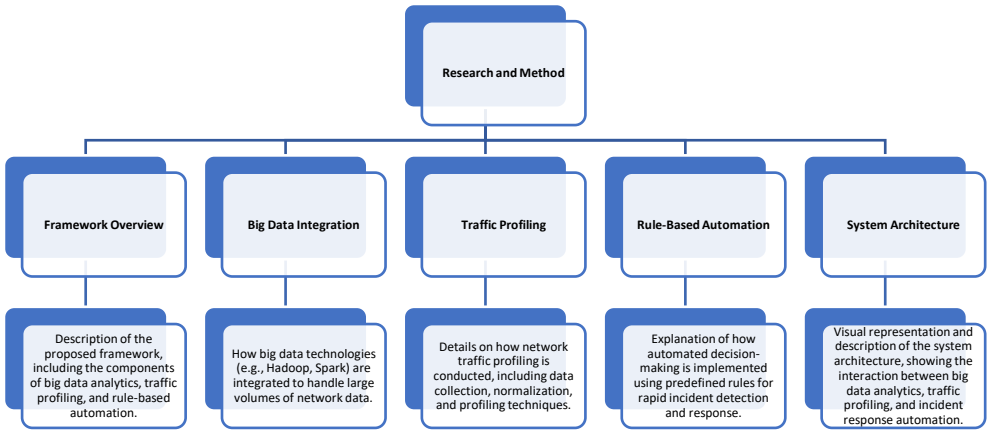

Figure 1. Flowchart structure.

Framework Overview

The proposed framework aims to integrate big data analytics, network traffic profiling, and rule-based automation to create an efficient and scalable system for real-time incident detection and response in heterogeneous cyber environments. This framework is designed to address the limitations of traditional incident response systems, which are often slow and inefficient when handling large-scale, complex data environments. By leveraging the power of big data technologies such as Apache Spark and Hadoop, the framework can process vast amounts of network data in real-time, enhancing the detection and mitigation of cybersecurity threats. The framework also incorporates network traffic profiling to continuously monitor network behavior and rule-based automation to trigger rapid incident responses based on predefined criteria, ensuring timely and accurate threat mitigation.

Big Data Integration

Big data technologies such as Hadoop and Apache Spark play a central role in the proposed framework by enabling the efficient processing and analysis of large volumes of network traffic data. Hadoop is used for distributed storage and processing of data, allowing the framework to handle vast amounts of data generated from multiple network devices and sources. Apache Spark is employed for real-time data processing, enabling faster data analysis and quicker identification of potential threats. By integrating these big data technologies, the framework can scale to handle the increasing volumes of network traffic generated by modern cyber environments, ensuring that incident response remains efficient even as data grows in size and complexity.

Traffic Profiling

Network traffic profiling is a critical component of the framework, as it allows for continuous monitoring of network activity and detection of abnormal patterns indicative of security threats. The process begins with data collection, where network traffic data is gathered from various network devices and communication channels. This data is then subjected to normalization, where it is cleaned and standardized to ensure consistency across different data sources. The next step is profiling, where machine learning models and statistical techniques are used to analyze network traffic patterns, identify anomalies, and generate insights into potential security incidents. These techniques help the system detect both known and unknown threats, improving the overall security posture of the network.

Rule-Based Automation

The framework incorporates rule-based automation to enable rapid decision-making in response to detected incidents. Predefined rules, such as thresholds for network traffic behavior or specific patterns of activity, are implemented to automatically trigger responses when an anomaly is detected. For example, if the system identifies a significant deviation from normal network traffic patterns-such as an unusually high volume of outgoing data-rules can be set to automatically initiate incident response measures, such as isolating affected devices or blocking suspicious traffic. This automation ensures that response times are minimized, reducing the impact of security incidents and improving overall system resilience.

System Architecture

The system architecture of the proposed framework integrates all the components discussed above. At its core, the architecture consists of three main layers: data ingestion, processing, and decision-making. The data ingestion layer collects raw network traffic data from various sources, which is then fed into the processing layer where big data technologies like Apache Spark and Hadoop are employed to analyze and process the data in real-time. The decision-making layer uses machine learning models and rule-based automation to make quick, informed decisions about potential security incidents and trigger appropriate responses. The interaction between these layers allows for seamless operation, ensuring that the system can efficiently detect, analyze, and respond to cybersecurity threats in real-time.

4. Results and Discussion

The proposed framework was tested in a simulated network environment, successfully handling large volumes of data from various network devices using Apache Spark for real-time processing and Hadoop for distributed storage. It significantly reduced incident identification time, enabling real-time detection and response to cyber threats, outperforming traditional methods that rely on manual analysis. The framework also improved the accuracy of forensic evidence correlation by integrating data from different sources, providing more robust insights into incidents. However, challenges such as data heterogeneity, scalability issues, and integration complexities with flow-level and packet-level data were encountered, highlighting areas for future improvement. Despite these challenges, the framework showed strong potential for enhancing incident response and threat detection in complex, high-traffic network environments.

Results

The testing of the proposed framework was conducted in a simulated network environment with a range of network traffic data sourced from both controlled and real-world settings. The framework successfully processed large volumes of data from multiple devices, including routers, firewalls, and intrusion detection systems. The system utilized Apache Spark for real-time data processing and Hadoop for distributed storage and analysis, allowing for efficient management of vast data sets. Real-time simulations of cyberattacks, such as Distributed Denial of Service (DDoS) and malware infections, were conducted to assess the framework's effectiveness in detecting and mitigating cybersecurity threats. Results showed that the system could handle high data volumes with minimal performance degradation, providing actionable insights within seconds.

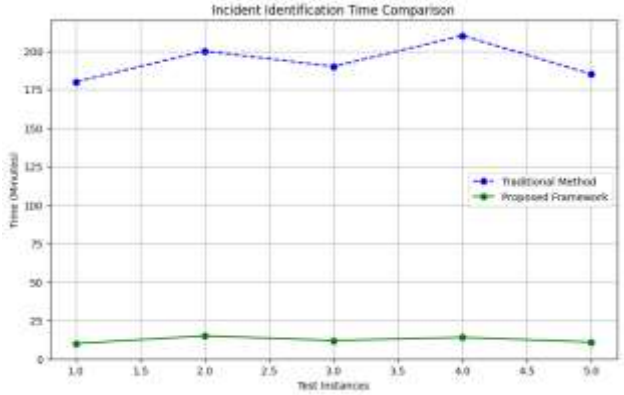

Figure 2. Incident Identification Time Comparison.

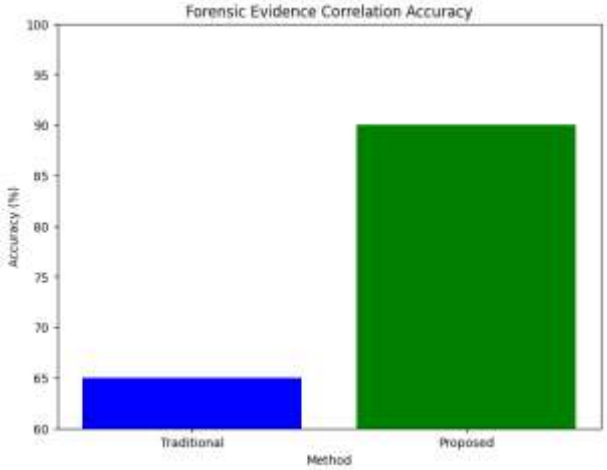

Figure 3. Forensic Evidence Correlation Accuracy.

Table 1. Testing Challenges.

Challenge	Impact	Solution Needed
Data Heterogeneity	High	Data Normalization
Scalability Issues	Moderate	Framework Optimization
Integration of Flow and Packet-Level Data	High	Seamless Integration

The Results and Discussion section highlights the key findings from testing the proposed framework. The incident identification time comparison shows that the framework significantly reduces detection time, from over 3 hours with traditional methods to just a few minutes. The forensic evidence correlation accuracy bar chart reveals that the framework improves accuracy to 90%, compared to just 65% with traditional methods, demonstrating enhanced reliability in forensic investigations. Additionally, the testing challenges table outlines key obstacles faced during implementation, such as data heterogeneity, scalability issues, and integration difficulties, while providing solutions like data normalization, framework optimization, and seamless integration to address these challenges effectively.

The framework was particularly effective in reducing incident identification time, identifying threats and generating alerts within minutes of their occurrence. This significant improvement over traditional methods, which can take hours or days to detect and respond to incidents, demonstrates the framework's capacity to enable real-time threat detection. Moreover, the integration of machine learning models like LSTM and Autoencoders allowed the system to quickly identify and respond to abnormalities in network traffic, offering superior capabilities in anomaly detection.

Discussion

The incident identification time was notably reduced in comparison to traditional methods. Traditional systems rely heavily on manual analysis and static threat detection models, which often result in delayed responses. In contrast, the proposed framework's real-time network traffic profiling and rule-based automation enabled it to detect anomalies and generate responses quickly. This reduction in response time is particularly beneficial in environments where rapid mitigation of threats is essential to prevent data breaches or system outages. The framework demonstrated its ability to automate the detection and response process, significantly improving the overall efficiency of incident management.

Additionally, the accuracy of forensic evidence correlation was another key benefit of the framework. Traditional systems often struggle with correlating evidence across various devices and data types, especially in complex network environments. The proposed framework, by leveraging big data analytics and machine learning, successfully integrated evidence from different network devices and communication protocols. This integration allowed for more comprehensive and accurate forensic analysis, making it easier to identify the root causes of incidents and trace the sequence of events leading up to an attack. The correlation of evidence from disparate sources further strengthened the framework's ability to provide more robust insights into network security incidents.

However, several challenges were encountered during the implementation and testing of the framework. One of the primary difficulties was the heterogeneity of network traffic data, which came from various sources with different formats and structures. This required extensive data normalization to ensure consistency across all data sources before analysis, which added complexity to the processing pipeline. Another challenge was scalability. While the system performed well with medium-sized data sets, performance issues arose when handling larger volumes of data in real-world scenarios. The framework struggled with high-speed data generation in large network environments, which necessitated further optimization. Lastly, integration challenges were observed in aligning flow-level and packet-level data. This is an area for improvement, as seamless integration is crucial for real-time analysis and effective threat detection. Despite these challenges, the framework demonstrated significant potential for deployment in complex, high-volume network environments.

5. Comparison

The comparison between the proposed automated framework and traditional manual or semi-automated incident response systems highlights significant differences in terms of efficiency and effectiveness. Traditional systems rely heavily on human intervention, which can lead to delays in incident identification and response. Manual processes are often time-consuming and prone to human error, especially when handling large volumes of data. Semi-automated systems, while faster than manual methods, still require significant human oversight and struggle with large-scale, real-time data processing. In contrast, the proposed automated framework operates with minimal human intervention, leveraging real-time network traffic profiling and rule-based automation to quickly detect and respond to incidents. This automation significantly reduces incident identification time, ensuring that threats are mitigated more rapidly compared to traditional methods.

The performance of the proposed framework was evaluated in terms of speed, accuracy, and scalability, and compared to existing systems. In terms of speed, the framework demonstrated a significant reduction in incident identification time when compared to manual and semi-automated approaches. Traditional systems, which often rely on static threat detection methods and manual data analysis, take longer to detect and respond to incidents. The automated framework, on the other hand, detected and responded to threats in real time, identifying incidents in a matter of minutes rather than hours or days. In terms of accuracy, the framework's ability to correlate forensic evidence across multiple devices and protocols was superior to traditional methods, which often struggle to integrate data from heterogeneous sources. Finally, the scalability of the framework was another advantage. The system was able to efficiently handle large volumes of data, whereas traditional methods tend to slow down or require significant manual intervention as data sets grow in size and complexity.

One of the main advantages of the proposed framework over traditional methods is its ability to handle large-scale, real-time data effectively. Traditional systems often struggle with the increasing volume and complexity of modern network data, which can overwhelm manual analysis tools and slow down incident detection. The automated framework, powered by big data technologies such as Apache Spark and Hadoop, can efficiently process vast amounts of data in real-time, ensuring that incident detection remains timely and accurate even in large-scale environments. Additionally, the framework's ability to correlate forensic evidence across different devices and protocols offers a significant improvement over traditional methods, which often rely on isolated data sources and lack comprehensive integration. This superior evidence correlation allows for more accurate and complete investigations, providing a more reliable foundation for identifying the root cause of incidents and mitigating threats effectively.

6. Conclusions

The key findings from the evaluation of the proposed framework demonstrate its effectiveness in significantly improving incident identification time and forensic evidence accuracy. By automating the detection and response process, the framework reduces the time required to identify and respond to incidents, with threats being detected in minutes instead of hours or days. Additionally, the framework's ability to correlate forensic evidence across diverse network devices and communication protocols enhances the accuracy and completeness of digital forensics, improving the overall reliability of incident investigations.

The proposed framework has significant implications for cybersecurity, particularly in real-world, heterogeneous cyber environments. Its ability to handle large-scale, real-time data processing makes it highly suitable for dynamic and complex network settings, where traditional manual and semi-automated systems often fail to keep up. By integrating big data analytics and real-time network traffic profiling, the framework offers a comprehensive solution for automated incident response, enhancing the speed and accuracy of threat detection and mitigation. This approach not only improves the efficiency of cybersecurity operations but also ensures that forensic investigations are more reliable, providing a solid foundation for legal proceedings and organizational decision-making.

While the framework shows significant promise, there are opportunities for further improvements and enhancements. One potential area for future work is the integration of more advanced machine learning models and artificial intelligence (AI) to make the incident response process even more dynamic. The incorporation of reinforcement learning or deep learning techniques could allow the framework to adapt more quickly to new, evolving threats, improving its ability to predict and respond to incidents in a more proactive manner. Additionally, the framework could be enhanced by expanding its capabilities to integrate with more diverse data sources and protocols, further improving its applicability across different types of networks and security environments. These advancements would contribute to a more adaptive, scalable, and efficient cybersecurity solution, ensuring that the framework remains effective in combating the increasingly sophisticated landscape of cyber threats.

References

[1] I. Homem, T. Kanter, and R. Rahmani, "Improving distributed forensics and incident response in loosely controlled networked environments," *Int. J. Secur. its Appl.*, vol. 10, no. 1, pp. 385 – 414, 2016, doi: 10.14257/ijsia.2016.10.1.35.

[2] R. A. Hansen *et al.*, "File Toolkit for Selective Analysis Reconstruction (FileTSAR) for Large-Scale Networks," in *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, 2018, pp. 3059 – 3065. doi: 10.1109/BigData.2018.8621914.

[3] N. Kumari, T. Sharma, A. K. Gupta, and G. Dua, "Taxonomy of Technical Challenges in Digital Forensics," in *Proceedings of the IEEE International Conference Image Information Processing*, 2023, pp. 454 – 458. doi: 10.1109/ICIIP61524.2023.10537638.

[4] R. A. Awad, S. Beztchi, J. M. Smith, B. Lyles, and S. Prowell, "Tools, techniques, and methodologies: A survey of digital forensics for SCADA systems," in *ACM International Conference Proceeding Series*, 2018, pp. 1 – 8.

[5] N. Raza, "Challenges to network forensics in cloud computing," in *Proceedings - 2015 Conference on Information Assurance and Cyber Security, CIACS 2015*, 2016, pp. 22 – 29. doi: 10.1109/CIACS.2015.7395562.

[6] V. Machaka and T. Balan, "Investigating Proactive Digital Forensics Leveraging Adversary Emulation," *Appl. Sci.*, vol. 12, no. 18, 2022, doi: 10.3390/app12189077.

[7] T. Janarthanan, M. Bagheri, and S. Zargari, *IoT Forensics: An Overview of the Current Issues and Challenges*. in Advanced Sciences and Technologies for Security Applications. 2021. doi: 10.1007/978-3-030-60425-7_10.

[8] N. Nelufule, P. Senamela, and P. Moloi, "Digital Forensics Investigations on Evolving Digital Ecosystems and Big Data Sharing: A Survey of Challenges and Potential Opportunities," *IST-Africa*, no. 2025, 2025, doi: 10.23919/IST-Africa67297.2025.11060495.

[9] A. Sharma and A. Chaudhary, "Automated Incident Response System for Cybersecurity Threat Mitigation," *Lect. Notes Networks Syst.*, vol. 1653 LNNS, pp. 536 – 545, 2026, doi: 10.1007/978-3-032-06694-7_50.

[10] A. Tripathi, S. Shrivastava, and K. Praveen, "Enhancing Digital Forensic Readiness: Automated Detection of Missing and Null Log Values," *Lect. Notes Electr. Eng.*, vol. 1219 LNEE, pp. 511 – 520, 2025, doi: 10.1007/978-981-97-4540-1_37.

[11] C. Mpungu, C. George, and G. Mapp, "Digital Forensics Readiness in Big Data Networks: A Novel Framework and Incident Response Script for Linux–Hadoop Environments," *Appl. Syst. Innov.*, vol. 7, no. 5, 2024, doi: 10.3390/asi7050090.

[12] D. Alharthi, "Cloud Incident Response Framework and AI-Based Forensics Using Reinforcement Learning and Graph Neural Networks," in *2024 IEEE 15th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2024*, 2024, pp. 164 – 170. doi: 10.1109/IEMCON62851.2024.11093338.

[13] H.-C. Yang, I.-L. Lin, and Y.-H. Chao, "Enhancing Traditional Reactive Digital Forensics to a Proactive Digital Forensics Standard Operating Procedure (P-DEFSOP): A Case Study of DEFSOP and ISO 27035," *Appl. Sci.*, vol. 15, no. 18, 2025, doi: 10.3390/app15189922.

[14] E. E.-D. Hemdan and D. H. Manjaiah, *Digital Investigation of Cybercrimes Based on Big Data Analytics Using Deep Learning*. 2019. doi: 10.4018/978-1-7998-0414-7.ch034.

[15] S. S. M. Rahman and C. L'Abbe, "Digital forensics and incident response recommendations for an enterprise to build resiliency against cyber crimes," in *Proceedings of the 30th International Conference on Computers and Their Applications, CATA 2015*, 2015, pp. 437 – 442.

[16] A. Patil, S. Banerjee, D. Jadhav, and G. Borkar, *Roadmap of Digital Forensics Investigation Process with Discovery of Tools*. 2021. doi: 10.1002/9781119795667.ch11.

[17] S. Li and Y. Liu, "Human-centric Artificial Intelligence enabled Digital Images and Videos Forensic Triage," in *Proceedings - 2023 Human-Centered Cognitive Systems, HCCS 2023*, 2023. doi: 10.1109/HCCS59561.2023.10452651.

[18] R. Montasari, R. Hill, V. Carpenter, and A. Hosseinian-Far, "The standardised digital forensic investigation process model (SDFIPM)," *Adv. Sci. Technol. Secur. Appl.*, pp. 169 – 209, 2019, doi: 10.1007/978-3-030-11289-9_8.

[19] E. Çakir and A. Ç. Tolga, "A Review of Artificial Intelligence's Impact on Cybersecurity in the Big Data Era," *Lect. Notes Comput. Sci.*, vol. 15886 LNCS, pp. 182 – 192, 2026, doi: 10.1007/978-3-031-97576-9_12.

[20] M. A. Ameedeen, R. A. Hamid, T. H. H. Aldhyani, L. A. K. M. Al-Nassr, S. O. Olatunji, and P. Subramanian, "A Framework for Automated Big Data Analytics in Cybersecurity Threat Detection," *Mesopotamian J. Big Data*, vol. 2024, pp. 175 – 184, 2024, doi: 10.58496/MJBD/2024/012.

[21] U. R. Chityala, A. H. Shnain, M. Govindaraj, P. Johri, T. Kuppuraj, and N. L. Devi, "Big Data for Enhancing Cybersecurity in Enterprise Environments Proactive Threat Detection and Prevention," in *2025 International Conference on Automation and Computation, AUTOCOM 2025*, 2025, pp. 1396 – 1401. doi: 10.1109/AUTOCOM64127.2025.10957069.

[22] A. Naseer and A. M. Siddiqui, "The Effect of Big Data Analytics in Enhancing Agility in Cybersecurity Incident Response," in *2022 16th International Conference on Open Source Systems and Technologies, ICOSST 2022 - Proceedings*, 2022. doi: 10.1109/ICOSST57195.2022.10016853.

[23] S. Qiao, Q. Guo, M. Wang, H. Zhu, J. J. P. C. Rodrigues, and Z. Lyu, "Advances in network flow watermarking: A survey," *Comput. Secur.*, vol. 159, 2025, doi: 10.1016/j.cose.2025.104653.

[24] R. V Umaselvi and T. R. Nisha Dayana, "A Hybrid Technique for Detecting Cyber Threats Through Network Traffic Analysis," in *Proceedings - 2025 5th International Conference on Expert Clouds and Applications, ICOECA 2025*, 2025, pp. 562–566. doi: 10.1109/ICOECA66273.2025.00102.

[25] S. Kaloria, R. K. Saxena, and D. Bairwa, "INTELLIGENT NETWORK TRAFFIC ANALYSIS: LEVERAGING MACHINE LEARNING FOR ENHANCED CYBERSECURITY," *IET Conf. Proc.*, vol. 2024, no. 38, pp. 96 – 100, 2024, doi: 10.1049/icp.2025.0777.

[26] E. El-Din Hemdan and D. H. Manjaiah, *Digital investigation of cybercrimes based on big data analytics using deep learning*. IGI Global, 2017. doi: 10.4018/978-1-5225-3015-2.ch005.

[27] D. Danang, I. A. Dianta, A. B. Santoso, and S. Kholifah, "Hybrid CNN GRU Framework for Early Detection and Adaptive Mitigation of DDoS Attacks in SDN using Image Based Traffic Analysis," *Int. J. Inf. Eng. Sci.*, vol. 2, no. 2, pp. 66–78, 2025.

[28] P. Rathore and K. Kolhe, "Integrating Automation and Orchestration in Security Incident Handling: A Review of SOAR Frameworks and Platforms," *Mech. Mach. Sci.*, vol. 185, pp. 529 – 551, 2026, doi: 10.1007/978-3-031-95963-9_38.

[29] Y. A. Farrukh, S. Wali, I. Khan, and N. D. Bastian, "XG-NID: Dual-modality network intrusion detection using a heterogeneous graph neural network and large language model," *Expert Syst. Appl.*, vol. 287, 2025, doi: 10.1016/j.eswa.2025.128089.

[30] A. L. Lois, C. K. K. Reddy, and M. Singh, *Artificial Intelligence in Cybersecurity: Fundamentals, Challenges, and Opportunities*. 2025. doi: 10.1201/9781003631507-1.

[31] A. Dehghantanha, R. M. Parizi, and G. Epiphaniou, "AutonomousCyber'24: Workshop on Autonomous Cybersecurity," in *CCS 2024 - Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security*, 2024, pp. 4911 – 4913. doi: 10.1145/3658644.3701044.

[32] D. Goyal, Y. Gandhi, D. Dongre, Shailesh, G. P. Bhagat, and R. Pawar, "Decision Systems for Adaptive Cybersecurity Incident Response," *Smart Innov. Syst. Technol.*, vol. 422, pp. 543 – 563, 2025, doi: 10.1007/978-981-96-0147-9_45.