Mesopotamian journal of Big Data
Vol. **(2024)**, 2024, **pp**. 175–184
DOI: https://doi.org/10.58496/MJBD/2024/012 ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BigData

Review Article

A Framework for Automated Big Data Analytics in Cybersecurity Threat Detection

Mohamed Ariff Ameedeen[1,*], Rula A. Hamid[2], Theyazn H H Aldhyani[3], Laith Abdul Khaliq Mohammed Al-Nassr[4], Sunday Olusanya Olatunji[5], Priyavahani Subramanian[6]

[1] *Faculty of Computing, Universiti Malaysia Pahang, Malaysia.*
[2] *College of Business Informatics, University of Information Technology and Communications (UOITC), Baghdad, Iraq.*
[3] *King Faisal university, Saudi Arabia.*
[4] *Queen Mary University of London,UK.*
[5] *Faculty of Computing, Adekunle Ajasin University Akungba Akoko, Ondo State, Nigeria.*
[6] *Iframe Network, Johor, Malaysia.*

ARTICLE INFO

Article History
Received 21 Jul 2024
Accepted 22 Aug 2024
Published 25 Sep 2024

Keywords
Big Data
Cybersecurity
Big Data Analytics
Threat Detection

ABSTRACT

This research presents a novel framework designed to enhance cybersecurity through the integration of Big Data analytics, addressing the critical need for scalable and real-time threat detection in large-scale environments. Utilizing technologies such as Apache Kafka for efficient data ingestion, Apache Flink for stream processing, and advanced machine learning models like LSTM and Autoencoders, the framework offers robust anomaly detection capabilities. It also includes automated response mechanisms using SOAR and XDR systems, significantly improving response times and accuracy in threat mitigation. The proposed solution not only addresses current challenges in handling vast and complex data but also paves the way for future advancements, such as the integration of more sophisticated AI techniques and application across various domains, including IoT and cloud security. This research contributes to the field by providing a comprehensive, adaptive, and scalable framework that meets the demands of modern cybersecurity landscapes.

1. INTRODUCTION

The rapid evolution of technology has significantly impacted the cybersecurity landscape, making it increasingly challenging for organizations to protect their digital assets[1]. In recent years, the complexity and volume of cybersecurity threats have grown at an unprecedented rate, driven by the proliferation of Internet of Things (IoT) devices, the widespread adoption of cloud computing, and the increasing sophistication of cyberattack techniques. These developments have expanded the attack surface, making traditional security measures insufficient to combat modern cyber threats effectively[2].

Recent research has shown that cybercriminals are leveraging advanced technologies, such as artificial intelligence (AI) and machine learning, to develop more sophisticated and targeted attacks[3]. For instance, AI-driven malware can adapt to security defenses in real-time, making it particularly challenging to detect and neutralize. Similarly, ransomware attacks have become more coordinated and complex, often involving multiple stages and targeting critical infrastructure with devastating consequences[4].

Given this escalating threat environment, the role of Big Data analytics in cybersecurity has become increasingly crucial. Big Data enables the processing and analysis of vast amounts of information in real-time, allowing organizations to detect patterns and anomalies that may indicate a potential cyber threat[5]. By integrating machine learning and predictive analytics

*Corresponding author. Email: xxx@gmail.com

into cybersecurity frameworks, organizations can enhance their ability to identify and respond to threats proactively, reducing the risk of significant damage[6].

However, the sheer volume of data generated by modern digital environments presents its challenges. Manual analysis is no longer feasible, as it is both time-consuming and prone to errors. This has led to the development of automated frameworks for threat detection, which leverage Big Data analytics to provide continuous monitoring and real-time response capabilities[7]. These frameworks are designed to reduce the reliance on human intervention, ensuring that threats are identified and mitigated swiftly and accurately[8].

Automated threat detection frameworks are particularly valuable in today's cybersecurity landscape, where the speed and sophistication of attacks require a rapid and dynamic response. By incorporating predictive analytics, these frameworks can also anticipate future threats, enabling organizations to take preventative measures before an attack occurs[9]. This proactive approach is essential for maintaining robust cybersecurity defenses in an environment where threats are constantly evolving[10].

The increasing complexity and volume of cybersecurity threats necessitate a shift towards more advanced and automated defense mechanisms. Big Data analytics, when integrated with machine learning and predictive analytics, offers a powerful solution to the challenges posed by modern cyber threats[11]. As organizations continue to adopt new technologies, the development and implementation of automated threat detection frameworks will be critical to ensuring the security and integrity of their digital assets[12].

1.1 Problem Statement

The rapidly evolving landscape of cybersecurity presents significant challenges for organizations attempting to safeguard their digital assets. Traditional security measures, which often rely on manual processes and reactive responses, are increasingly inadequate in the face of the growing complexity and sophistication of cyber threats[13]. The advent of advanced technologies such as artificial intelligence (AI) and machine learning has enabled cybercriminals to develop more sophisticated attack vectors, including AI-driven malware and highly coordinated ransomware campaigns[14]. These threats can adapt to and bypass conventional security defenses, leaving organizations vulnerable to potentially devastating breaches[15].

Moreover, the proliferation of IoT devices and the widespread adoption of cloud computing have expanded the attack surface, introducing new vulnerabilities that can be exploited by malicious actors. As a result, the volume of data generated by modern digital environments has surged, making it increasingly difficult for security teams to monitor, analyze, and respond to potential threats manually[16]. The sheer scale of this data, combined with the speed at which cyber threats can manifest, necessitates the development of automated solutions that can provide real-time threat detection and response[17].

Despite the critical need for these advanced capabilities, many organizations struggle to implement effective automated frameworks for cybersecurity. Existing systems often lack the necessary integration of Big Data analytics and machine learning, which are essential for processing large volumes of data and identifying complex patterns indicative of cyber threats[18]. Additionally, the absence of predictive analytics in current security frameworks limits the ability to anticipate and prevent future attacks, leaving organizations reactive rather than proactive in their cybersecurity posture[19].

Therefore, the primary challenge facing modern cybersecurity is the need to develop and implement robust, automated frameworks that leverage Big Data analytics, machine learning, and predictive analytics. These frameworks must be capable of continuous monitoring and real-time response, allowing organizations to stay ahead of increasingly sophisticated cyber threats. Without such advancements, organizations will remain vulnerable to the ever-evolving tactics of cybercriminals, risking significant financial, operational, and reputational damage.

1.2 Objectives

- **Develop an Innovative Framework for Automated Big Data Analytics in Cybersecurity:**

This research aims to create a cutting-edge framework that integrates Big Data analytics, machine learning, and predictive technologies to enhance cybersecurity practices. The framework will address the current limitations in detecting and responding to sophisticated cyber threats by automating the analysis and interpretation of vast amounts of data in real-time. By incorporating these advanced capabilities, the framework seeks to significantly improve the detection, prediction, and mitigation of complex cyber threats.

- **Demonstrate the Framework’s Efficiency and Scalability:**

Another goal is to validate the proposed framework’s ability to efficiently process large datasets without compromising performance. The research will explore the framework's scalability, ensuring it can adapt to the growing data demands of modern organizations. The framework will be evaluated for its capacity to deliver fast, accurate threat detection and response across diverse and expansive digital environments, proving its practicality for organizations of various sizes and industries.

These objectives directly address the identified gaps in current cybersecurity solutions, emphasizing the need for advanced, automated approaches that can effectively manage the increasing complexity and scale of cyber threats.

2. LITERATURE REVIEW

2.1 Overview of Cybersecurity Threat Detection Techniques

2.1.1 Traditional vs. Modern Approaches

Cybersecurity threat detection has evolved significantly over the past few decades. Traditional approaches to threat detection relied heavily on signature-based methods, where known threats were identified based on pre-defined patterns or "signatures" in the data. These methods, while effective against known threats, were limited in their ability to detect new or evolving threats, particularly those that employed sophisticated techniques to evade detection[20].

In contrast, modern approaches to cybersecurity emphasize the use of behavior-based and anomaly detection methods. These techniques do not rely solely on known signatures but instead focus on identifying unusual patterns or behaviors that may indicate a potential threat[21]. This shift has been driven by the increasing complexity of cyber threats, which often involve sophisticated tactics such as polymorphic malware, zero-day exploits, and advanced persistent threats (APTs) that can bypass traditional defenses[14].

One of the key advancements in modern cybersecurity is the integration of machine learning and artificial intelligence (AI) into threat detection frameworks. Unlike traditional methods, which require constant updates to signature databases, machine learning models can be trained to recognize patterns and detect anomalies in real-time, even for previously unseen threats. This ability to "learn" from new data and improve over time makes AI-based approaches particularly powerful in the ever-evolving landscape of cyber threats[11].

2.1.2 Role of Machine Learning and AI in Cybersecurity

Machine learning and AI have become integral components of modern cybersecurity strategies, offering significant improvements over traditional methods. These technologies enable the development of predictive models that can analyze vast amounts of data to identify potential threats before they materialize[22]. By leveraging algorithms that can detect patterns and correlations in large datasets, AI and machine learning provide a proactive approach to cybersecurity, allowing organizations to anticipate and prevent attacks rather than merely responding to them[23].

Recent research has demonstrated the effectiveness of machine learning in various cybersecurity applications, including intrusion detection, malware classification, and threat intelligence[24]. For instance, supervised learning algorithms have been used to train models on labeled datasets, enabling them to classify new data points accurately. Unsupervised learning, on the other hand, has proven valuable in detecting anomalies that may indicate novel threats. Additionally, reinforcement learning is increasingly being explored for its potential to adaptively respond to dynamic threat environments[25].

AI-driven cybersecurity tools are also enhancing the ability to detect and mitigate threats in real-time. These tools can automatically analyze and correlate data from multiple sources, providing a comprehensive view of the threat landscape[26]. Moreover, AI can automate many aspects of threat detection and response, reducing the need for human intervention and allowing security teams to focus on more complex tasks[27].

In conclusion, the evolution from traditional signature-based methods to modern, AI-driven approaches has significantly enhanced the capabilities of cybersecurity threat detection. Machine learning and AI offer the scalability, adaptability, and real-time processing power needed to address the growing complexity and sophistication of cyber threats. As these technologies continue to advance, they will play an increasingly critical role in securing digital environments against emerging threats.

2.2. Big Data in Cybersecurity

Data Sources: Logs, Network Traffic, User Behavior, etc.

Big Data has become a cornerstone of modern cybersecurity strategies, drawing from various sources such as system logs, network traffic, and user behavior analytics (UBA)[28]. Tabel1 shows data source comparisons in cyber security and their applications. System logs, generated by devices and applications, provide detailed records of events, which are crucial for identifying unauthorized access and other security incidents. Network traffic data captures the flow of information across a network, enabling the detection of anomalies such as unusual spikes in data transfers or communications with known malicious IP addresses. UBA focuses on monitoring user activities to identify deviations from typical behavior, which could indicate insider threats or compromised accounts[29].

TABLE I. COMPARISON OF DATA SOURCES IN CYBERSECURITY PROBIDES A DETAILED COMPARISON OF THESE KEY DATA SOURCES AND THEIR APPLICATIONS IN CYBERSECURITY [30].

Data Source	Application
System Logs	Tracking events, identifying unauthorized access, auditing
Network Traffic	Monitoring data flow, detecting anomalies, preventing data exfiltration
User Behavior	Identifying deviations from normal activity, detecting insider threats

2.2.1 Challenges in Processing and Analyzing Big Data for Cybersecurity

One of the primary challenges in utilizing Big Data for cybersecurity is the sheer volume and velocity of data generated. Modern networks produce enormous amounts of data every second, making real-time processing and analysis difficult[31]. Additionally, integrating data from diverse sources, each with its own format and structure, adds to the complexity of the task. Ensuring data quality and consistency across these varied sources is another major hurdle. Inaccurate or incomplete data can lead to false positives or missed threats, undermining the effectiveness of cybersecurity efforts[32].

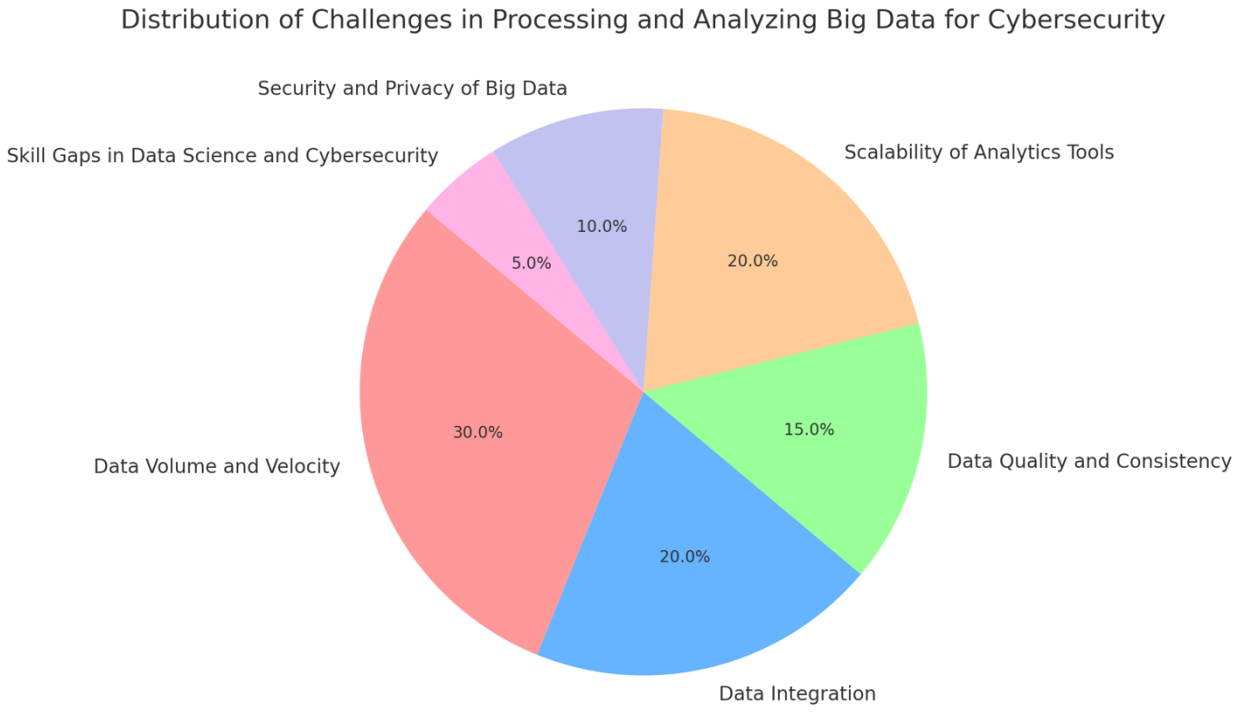

Fig .1. (Distribution of Challenges in Processing and Analyzing Big Data for Cybersecurity) illustrates the key challenges organizations face when dealing with Big Data in cybersecurity, such as data volume and velocity, data integration, and scalability[۳].

Scalability is also a critical issue, as existing tools and frameworks often struggle to keep up with the growing data volumes, especially as more organizations adopt IoT devices and cloud services. The challenge of scaling these tools without compromising performance is significant, and it often requires substantial investment in infrastructure and advanced analytics capabilities[33].

2.3Existing Frameworks and Their Limitations

2.3.1 Comparative Analysis of Current Frameworks

Several frameworks have been developed to address the challenges of Big Data in cybersecurity, each with its strengths and weaknesses. The ELK Stack, for example, is known for its centralized logging and real-time data analysis capabilities. However, it can struggle with handling large-scale network data in real-time scenarios[34]. Apache Hadoop excels in batch processing of large datasets, making it cost-effective for historical analysis, but its latency issues make it less suitable for real-time threat detection[35]. Apache Spark offers real-time analytics through in-memory processing, but it requires significant computational resources, making it complex to integrate with existing systems[36]. Splunk provides comprehensive analytics with robust alerting features but comes with a high cost and complexity[37].

2.3.2 Sophisticated Comparison of Cybersecurity Frameworks

TABLE II. PROVIDES A DETAILED COMPARISON OF THESE FRAMEWORKS, HIGHLIGHTING THEIR STRENGTHS, WEAKNESSES, BEST USE CASES, DEPLOYMENT COMPLEXITY, AND COST.

Framework	Strengths	Weaknesses	Best Use Case	Deployment Complexity	Cost
ELK Stack[38]	Centralized logging, real-time data analysis, open-source, scalable	Complex to set up and manage at scale, struggles with high data throughput	Centralized log management and monitoring for small to medium enterprises	Moderate to High	Free (Open Source), but costs can increase with scaling
Apache Hadoop [35].	Large-scale batch processing, cost-effective, handles vast amounts of data	Not ideal for real-time threat detection due to batch processing model, latency	Batch processing of large-scale datasets for historical analysis and data mining	Moderate	Free (Open Source), with costs associated with infrastructure
Apache Spark[36].	In-memory processing, real-time analytics, supports machine learning workloads	Requires significant computational resources, complex integration	Real-time analytics and processing for large datasets, particularly in AI and machine learning	High	Free (Open Source), but high infrastructure costs
Splunk[37]	Comprehensive analytics, real-time processing, robust alerting and reporting features	High cost, steep learning curve, can be resource-intensive	Comprehensive, enterprise-level security analytics with real-time monitoring and alerting	High	High, Enterprise-level pricing
Graylog[35]	Efficient log management, flexible, open-source, customizable pipelines	Limited community support compared to others, can require custom development	Flexible, open-source log management and analysis for organizations needing customization	Moderate	Free (Open Source), cost associated with infrastructure

Each framework is best suited for specific use cases, but none are without limitations. As organizations face increasingly sophisticated threats, the need for more advanced and adaptable frameworks becomes evident.

2.4 Identification of Gaps in Automation and Data Handling

Despite the capabilities of these frameworks, there are still significant gaps in automation and data handling. Many current frameworks require manual intervention for data correlation, threat identification, and response initiation, which can slow down reaction times and increase the risk of human error[39]. Automation is critical for ensuring that threats are detected and addressed as quickly as possible, without the delays inherent in manual processes[40].

Moreover, scalability remains a significant challenge. As data volumes continue to grow, many frameworks struggle to scale effectively, leading to performance bottlenecks and incomplete analyses. This is particularly problematic as organizations increasingly rely on IoT devices, which generate vast amounts of data that need to be processed and analyzed in real-time[41].

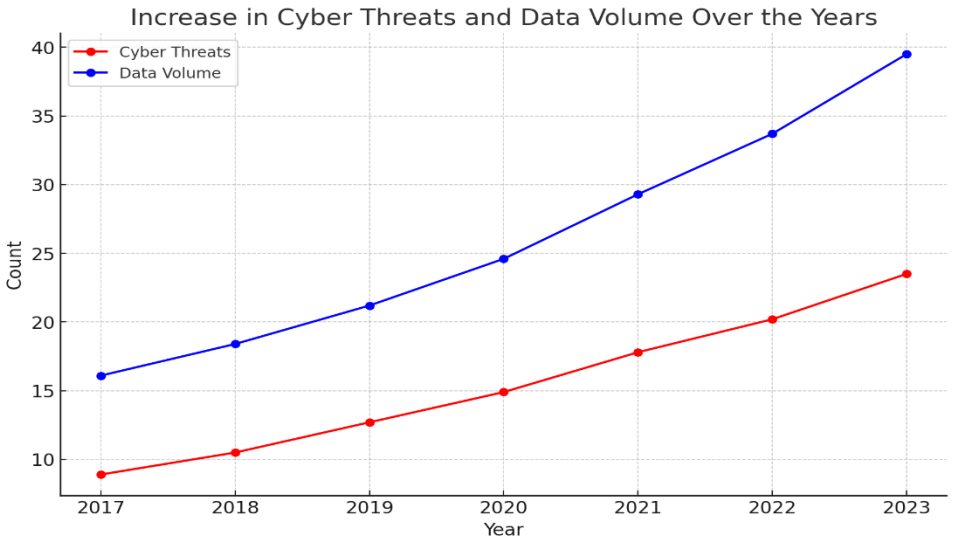

Fig .2. (Increase in Cyber Threats and Data Volume Over the Years) visually depicts the growing challenge of handling increasing data volumes alongside the rising number of cyber threats. This figure emphasizes the importance of developing scalable solutions to manage these challenges effectively[2].

Finally, the integration of diverse and distributed data sources is another area where current frameworks fall short[42]. Most frameworks are optimized for specific data types or environments, leading to fragmentation in the analysis process. This can result in missed threats or delays in identifying and mitigating security incidents, as organizations may not have a unified view of their security posture across all data sources[43].

Addressing these gaps is crucial for developing more effective and responsive cybersecurity solutions, capable of protecting organizations against increasingly sophisticated threats. The integration of advanced Big Data analytics, machine learning, and AI into cybersecurity frameworks represents a promising direction for future developments, ensuring that organizations can stay ahead of the rapidly evolving threat landscape.

3. PROPODED FRAMEWORK

3.1 Framework Overview

3.1.1 Introduction to the Proposed Framework

The proposed framework is a comprehensive, state-of-the-art system designed to meet the cybersecurity needs of large enterprises handling vast amounts of data. This framework integrates the latest advancements in data ingestion, real-time processing, machine learning, and automated threat detection to provide a scalable and efficient solution. It specifically addresses the challenges posed by environments with high data volumes, such as those generated by IoT devices, cloud services, and large enterprise networks.

3.1.2 Key Components and Their Interactions

1. Data Ingestion Layer:
    - Components: Apache Kafka, Fluentd
    - Function: This layer is responsible for efficiently collecting and streaming large volumes of data from diverse sources, including IoT devices, cloud platforms (like AWS and Azure), and on-premise systems. Apache Kafka is employed for its ability to handle real-time data streams with low latency, ensuring smooth and scalable data ingestion. Fluentd aggregates logs from various sources, normalizing and formatting the data for consistent processing.
    - Interaction: Data from sources such as cloud logs, IoT device telemetry, and user activity records is streamed into Kafka topics. Fluentd handles the aggregation and pre-processing of these logs before they are forwarded to the data processing layer.
2. Data Processing Layer:
    - Components: Apache Flink, Elasticsearch

- Function: This layer processes the ingested data in real-time using Apache Flink, which is optimized for stream processing and handling stateful computations. Flink's capabilities are crucial for processing events like network intrusions or abnormal user behavior as they happen. The processed data is then indexed in Elasticsearch, which allows for fast querying and comprehensive analysis.
- Interaction: Data processed by Flink is immediately stored in Elasticsearch, making it available for real-time querying and visualization. This interaction ensures that the framework can respond quickly to security incidents.

3. Threat Detection Layer:
   - Components: Deep Learning Models (LSTM, Autoencoders), Graph-Based Anomaly Detection
   - Function: This layer applies advanced machine learning models, such as Long Short-Term Memory (LSTM) networks for detecting sequential anomalies and Autoencoders for identifying deviations in network traffic patterns. Additionally, graph-based anomaly detection algorithms are used to detect unusual relationships within network graphs, indicating potential lateral movements by attackers.
   - Interaction: The processed data from the previous layer is analyzed by these models to detect any signs of anomalies or threats. Detected anomalies trigger alerts that are passed on to the decision-making layer.
4. Decision-Making Layer:
   - Components: SOAR (Security Orchestration, Automation, and Response), XDR (Extended Detection and Response)
   - Function: This layer automates the response to detected threats using SOAR platforms, which execute predefined playbooks to take actions like isolating compromised systems or blocking malicious IP addresses. XDR is employed to correlate threat data across endpoints, networks, and cloud environments, providing a unified response mechanism.
   - Interaction: Upon detection of a threat, SOAR systems automatically implement the appropriate response, while XDR ensures that data from various sources is correlated to provide a comprehensive threat response.
5. Visualization and Reporting Layer:
   - Components: Kibana, Grafana
   - Function: This layer provides security analysts with real-time dashboards and detailed reports. Kibana is used to visualize data stored in Elasticsearch, offering interactive dashboards for monitoring security incidents, system health, and threat trends. Grafana complements Kibana by providing advanced analytics and visualization capabilities, especially useful for tracking long-term trends and performance metrics.
   - Interaction: Data indexed in Elasticsearch is visualized in Kibana and Grafana dashboards, enabling analysts to quickly assess and respond to the organization's security posture.

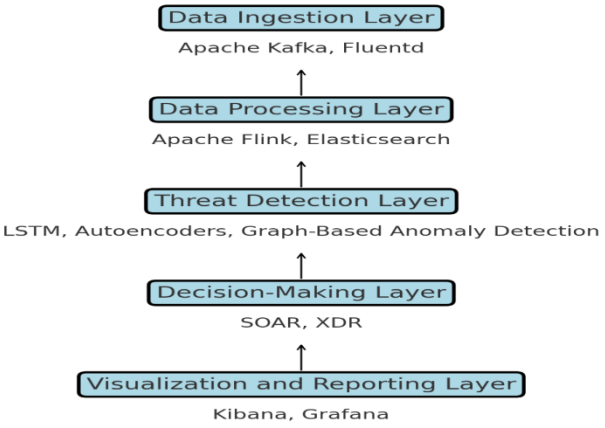

Fig .3. illustrates the entire architecture, showing the flow of data from the Data Ingestion Layer, through the Data Processing and Threat Detection Layers, to the Decision-Making and Visualization Layers. This visual representation helps in understanding how each component interacts with the others to provide a seamless, automated cybersecurity solution.

The proposed framework is built on proven technologies and methodologies, specifically tailored to the needs of large-scale, data-intensive environments. By integrating real-time analytics, machine learning, and automated response mechanisms, this framework offers a robust solution for detecting and responding to complex cyber threats with minimal delay and high accuracy.

4. CONCLUSION AND FUTURE WORK

4.1 Summary of Contributions

The proposed framework presented in this research makes significant contributions to the field of cybersecurity, particularly in the context of large-scale, data-intensive environments. The framework is specifically designed to address the challenges associated with the increasing volume and complexity of cyber threats. By integrating state-of-the-art technologies such as Apache Kafka for real-time data ingestion, Apache Flink for stream processing, advanced deep learning models like LSTM and Autoencoders for threat detection, and automated response systems like SOAR and XDR, the framework offers a comprehensive, scalable solution for modern cybersecurity needs.

Key contributions include:

- Real-time Data Processing and Analytics: The use of Apache Flink and Elasticsearch allows for the efficient processing and indexing of vast amounts of data in real-time, ensuring that threats can be detected and acted upon without delay.
- Advanced Threat Detection: By incorporating deep learning models and graph-based anomaly detection, the framework enhances the accuracy and speed of threat identification, particularly for sophisticated attacks that traditional methods might miss.
- Automated Response Mechanisms: The integration of SOAR and XDR systems within the framework enables a rapid, automated response to detected threats, reducing the reliance on manual interventions and minimizing the potential for human error.

This framework not only addresses current challenges in cybersecurity but also sets a foundation for more resilient and responsive security practices in the future.

4.2 Implications for Cybersecurity

The implementation of this framework has the potential to significantly impact the way organizations approach cybersecurity. Its ability to process and analyze large volumes of data in real-time, coupled with advanced threat detection capabilities, will enable organizations to respond more effectively to the growing sophistication of cyber threats. The automated nature of the framework ensures that security measures are both timely and accurate, reducing the window of opportunity for attackers and enhancing overall security posture.

Moreover, the scalability of the framework means it can be adopted by organizations of various sizes, from small businesses to large enterprises, making it a versatile solution in the ever-evolving landscape of cybersecurity. As more organizations integrate such frameworks, the collective security of the digital ecosystem will improve, leading to more robust defenses against emerging threats.

4.3 Future Enhancements

While the proposed framework represents a significant advancement in cybersecurity, there are several areas where future improvements could further enhance its effectiveness:

- Integration of More Advanced AI Techniques: Future iterations of the framework could incorporate more sophisticated AI models, such as reinforcement learning or generative adversarial networks (GANs), to improve the adaptability and resilience of threat detection mechanisms. These models could enable the system to better anticipate and react to novel threats.
- Expansion to Other Domains: The principles and technologies used in this framework could be adapted for use in other domains, such as fraud detection in finance, predictive maintenance in manufacturing, or anomaly detection in healthcare. Expanding the framework's application could provide similar benefits in these fields, enhancing security and operational efficiency.
- Enhanced Privacy and Compliance Features: As data privacy regulations continue to evolve, future versions of the framework could include more robust privacy-preserving techniques, such as differential privacy or federated learning, to ensure compliance with legal requirements while maintaining high levels of security.

In conclusion, this framework not only addresses the immediate challenges of cybersecurity but also provides a strong foundation for future advancements. By continuing to evolve and integrate cutting-edge technologies, this framework can remain at the forefront of cybersecurity innovation, ensuring that organizations are well-equipped to face the threats of tomorrow.

Conflicts Of Interest

The paper explicitly states that there are no conflicts of interest to disclose.

Funding

No grant or sponsorship is mentioned in the paper, suggesting that the author received no financial assistance.

Acknowledgment

The author acknowledges the support and resources provided by the institution in facilitating the execution of this study.

References

[1] M. Abdel-Rahman and others, "Advanced cybersecurity measures in IT service operations and their crucial role in safeguarding enterprise data in a connected world," Eig. Rev. Sci. Technol., vol. 7, no. 1, pp. 138–158, 2023.

[2] M. Thakur, "Cyber security threats and countermeasures in digital age," J. Appl. Sci. Educ., vol. 4, no. 1, pp. 1–20, 2024.

[3] R. Kaur, D. Gabrijelčič, and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," Inf. Fusion, vol. 97, p. 101804, 2023.

[4] A. H. Salem, S. M. Azzam, O. E. Emam, and A. A. Abohany, "Advancing cybersecurity: a comprehensive review of AI-driven detection techniques," J. Big Data, vol. 11, no. 1, p. 105, 2024.

[5] D. B. Rawat, R. Doku, and M. Garuba, "Cybersecurity in big data era: From securing big data to data-driven security," IEEE Trans. Serv. Comput., vol. 14, no. 6, pp. 2055–2072, 2019.

[6] A. Nassar and M. Kamal, "Machine Learning and Big Data analytics for Cybersecurity Threat Detection: A Holistic review of techniques and case studies," J. Artif. Intell. Mach. Learn. Manag., vol. 5, no. 1, pp. 51–63, 2021.

[7] Burhanuddin , M. (2023). Secure and Scalable Quantum Cryptographic Algorithms for Next-Generation Computer Networks. KHWARIZMIA, 2023, 95-102. https://doi.org/10.70470/KHWARIZMIA/2023/009

[8] B. Manyena, F. Machingura, and P. O'keefe, "Disaster Resilience Integrated Framework for Transformation (DRIFT): A new approach to theorising and operationalising resilience," World Dev., vol. 123, p. 104587, 2019.

[9] A. Modi et al., "Towards automated threat intelligence fusion," in 2016 IEEE 2nd International Conference on Collaboration and Internet Computing (CIC), 2016, pp. 408–416.

[10] P. Amthor, D. Fischer, W. E. Kühnhauser, and D. Stelzer, "Automated cyber threat sensing and responding: integrating threat intelligence into security-policy-controlled systems," in Proceedings of the 14th International Conference on Availability, Reliability and Security, 2019, pp. 1–10.

[11] I. Jada and T. O. Mayayise, "The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review," Data Inf. Manag., p. 100063, 2023.

[12] S. A. M. Authority, "Cyber security framework," Saudi Arab. Monet. Auth. Riyadh, Saudi Arab., 2017.

[13] V. Tzavara and S. Vassiliadis, "Tracing the evolution of cyber resilience: a historical and conceptual," 2024.

[14] J. Ferdous, R. Islam, A. Mahboubi, and M. Z. Islam, "A State-of-the-Art Review of Malware Attack Trends and Defense Mechanism.," IEEE Access, 2023.

[15] S. N. Tambe-Jagtap, "Human-Centric Cybersecurity: Understanding and Mitigating the Role of Human Error in Cyber Incidents", SHIFRA, vol. 2023, pp. 53–59, Jul. 2023, doi: 10.70470/SHIFRA/2023/007.

[16] W. Zhou, Y. Jia, A. Peng, Y. Zhang, and P. Liu, "The effect of IoT new features on security and privacy: New threats, existing solutions, and challenges yet to be solved," IEEE Internet things J., vol. 6, no. 2, pp. 1606–1616, 2018.

[17] R. Mills, A. K. Marnerides, M. Broadbent, and N. Race, "Practical intrusion detection of emerging threats," IEEE Trans. Netw. Serv. Manag., vol. 19, no. 1, pp. 582–600, 2021.

[18] A. Lheureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: Challenges and approaches," Ieee Access, vol. 5, pp. 7776–7797, 2017.

[19] A. Marab and A. Bhadrashetty, "A Novel Approach Using Deep Convolutional Neural Networks for Automated Dementia Detection and Classification", EDRAAK, vol. 2023, pp. 16–20, Mar. 2023, doi: 10.70470/EDRAAK/2023/004.

[20] M. Ahsan, K. E. Nygard, R. Gomes, M. M. Chowdhury, N. Rifat, and J. F. Connolly, "Cybersecurity threats and their mitigation approaches using Machine Learning. A Review," J. Cybersecurity Priv., vol. 2, no. 3, pp. 527–555, 2022.

[21] H. A. Salman and A. Alsajri, "The Evolution of Cybersecurity Threats and Strategies for Effective Protection. A review", SHIFRA, vol. 2023, pp. 73–85, Aug. 2023, doi: 10.70470/SHIFRA/2023/009.

[22] G. Richins, A. Stapleton, T. C. Stratopoulos, and C. Wong, “Big data analytics: opportunity or threat for the accounting profession?,” J. Inf. Syst., vol. 31, no. 3, pp. 63–79, 2017.
[23] Q. Liu and M. A. Vasarhelyi, “Big questions in AIS research: Measurement, information processing, data analysis, and reporting,” J. Inf. Syst., vol. 28, no. 1, pp. 1–17, 2014.
[24] I. F. Kilincer, F. Ertam, and A. Sengur, “Machine learning methods for cyber security intrusion detection: Datasets and comparative study,” Comput. Networks, vol. 188, p. 107840, 2021.
[25] A. S. Albahri et al., “A trustworthy and explainable framework for benchmarking hybrid deep learning models based on chest X-ray analysis in CAD systems,” Int. J. Inf. Technol. Decis. Mak., 2024.
[26] S. Rangaraju, “Secure by intelligence: enhancing products with AI-driven security measures,” EPH-International J. Sci. Eng., vol. 9, no. 3, pp. 36–41, 2023.
[27] A. K. Tyagi, T. F. Fernandez, S. Mishra, and S. Kumari, “Intelligent automation systems at the core of industry 4.0,” in International conference on intelligent systems design and applications, 2020, pp. 1–18.
[28] A. K. Bhardwaj, P. Dutta, and P. Chintale , Trans., “Securing Container Images through Automated Vulnerability Detection in Shift-Left CI/CD Pipelines ”, BJN, vol. 2024, pp. 162–170, Aug. 2024, doi: 10.58496/BJN/2024/016.
[29] E. Muhati and D. Rawat, “Data-Driven Network Anomaly Detection with Cyber Attack and Defense Visualization,” J. Cybersecurity Priv., vol. 4, no. 2, pp. 241–263, 2024.
[30] A. D. Kent, “Cyber security data sources for dynamic network research,” in Dynamic Networks and Cyber-Security, World Scientific, 2016, pp. 37–65.
[31] F. Iglesias, D. C. Ferreira, G. Vormayr, M. Bachl, and T. Zseby, “NTARC: a data model for the systematic review of network traffic analysis research,” Appl. Sci., vol. 10, no. 12, p. 4307, 2020.
[32] A. Alshaibi, M. Al-Ani, A. Al-Azzawi, A. Konev, and A. Shelupanov, “The comparison of cybersecurity datasets,” Data, vol. 7, no. 2, p. 22, 2022.
[33] E. Johnson, O. B. Seyi-Lande, G. S. Adeleke, C. P. Amajuoyi, and B. D. Simpson, “Developing scalable data solutions for small and medium enterprises: Challenges and best practices,” Int. J. Manag. \& Entrep. Res., vol. 6, no. 6, pp. 1910–1935, 2024.
[34] A. K. Goel et al., “Towards scalable real-time analytics: An architecture for scale-out of OLxP workloads,” Proc. VLDB Endow., vol. 8, no. 12, pp. 1716–1727, 2015.
[35] P. Raj et al., “Real-Time Analytics Using High-Performance Computing,” High-Performance Big-Data Anal. Comput. Syst. Approaches, pp. 161–185, 2015.
[36] E. Shaikh, I. Mohiuddin, Y. Alufaisan, and I. Nahvi, “Apache spark: A big data processing engine,” in 2019 2nd IEEE Middle East and North Africa COMMunications Conference (MENACOMM), 2019, pp. 1–6.
[37] R. Barker, “The uses and benefits of Splunk in continuous integration,” 2020.
[38] S. J. Son and Y. Kwon, “Performance of ELK stack and commercial system in security log analysis,” in 2017 IEEE 13th Malaysia international conference on communications (MICC), 2017, pp. 187–190.
[39] [39] J. Mineraud, O. Mazhelis, X. Su, and S. Tarkoma, “A gap analysis of Internet-of-Things platforms,” Comput. Commun., vol. 89, pp. 5–16, 2016.
[40] J. Prinsloo, S. Sinha, and B. Von Solms, “A review of industry 4.0 manufacturing process security risks,” Appl. Sci., vol. 9, no. 23, p. 5105, 2019.
[41] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, “Big data: Issues and challenges moving forward,” in 2013 46th Hawaii international conference on system sciences, 2013, pp. 995–1004.
[42] D. Ali, “A Comparative Review of Sustainable Enterprise System Frameworks: Integrating Web Technology, Clouding, AI, IoT, and Security for Green Business Transformation,” J. Inf. Technol. Informatics, vol. 3, no. 2, 2024.
[43] Y. Jiang, M. A. Jeusfeld, M. Mosaad, and N. Oo, “Enterprise architecture modeling for cybersecurity analysis in critical infrastructures-A systematic literature review,” Int. J. Crit. Infrastruct. Prot., p. 100700, 2024.