SPECTRUM OF ENGINEERING SCIENCES
ISSN (E) 3007-3138 (P) 3007-312X VOLUME 4, ISSUE 3, 2026

AUTOMATED CYBER THREAT INTELLIGENCE ANALYSIS AND DETECTION USING INTEGRATED GENERATIVE AI AND RETRIEVAL-AUGMENTED FRAMEWORK

Naila Anwar

PhD Scholar, Department of Information Technology, University of Gujrat, Pakistan

naila.anwwar@uog.edu.pk

DOI: https://doi.org/10.5281/zenodo.24028365

Keywords
Generative AI, Cyber Threat Intelligence, Large Language Models, Retrieval-Augmented Generation, Threat Detection, Cybersecurity.

Article History
Received: 01 January 2026
Accepted: 16 March 2026
Published: 31 March 2026

Copyright @Author
Corresponding Author: *
Naila Anwar

Abstract
Due to increase in the spread of cyber threats, there is a requirement of creating an intelligent cyber threat mechanism to combat threats. Existing methods are typically rule-based and manual, which makes them insufficient for detecting new threats. This research aims to create a Generative AI-based framework for automated threat analysis and detection using the combination of Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG). The framework receives cybersecurity data from various sources, identifies IoCs, fetches threat intelligence and classifies the threats in this context. The framework also produces enriched threat intelligence reports which include identified threats, attack techniques, supporting evidence and defensive recommendations. Its effectiveness will be measured through the accuracy, precision, recall, F1-score, hallucination rate, evidence relevance, and processing time.

1. INTRODUCTION

Digital technologies and their rapid development, together with cloud computing, Internet of Things and interconnectedness of communication, have made individuals, businesses and government organizations increasingly reliant on digital infrastructures. While these technologies offer many advantages, they have also increased the cyber-attack surface, presenting opportunities for cybercriminals to identify vulnerabilities, steal sensitive data, disrupt services and compromise network security. Cyber threats, including ransomware, phishing, malware, distributed denial-of-service attacks and advanced persistent threats (APTs), have become more sophisticated and hard-to-detect. As a result, traditional cybersecurity measures have significant difficulties when it comes to identifying and countering new threats. [2].

In today's cybersecurity landscape, Cyber Threat Intelligence (CTI) has become a significant component, dedicated to gathering, handling, analyzing, and comprehending data related to potential and actual cyber threats. CTI helps security professionals gain insight into the attackers' actions, the signs of a malicious indicator, the pattern of an attack, and what to do about it. Sources of threat intelligence information are network traffic, security logs, malware reports, vulnerabilities databases, phishing Web sites, security blogs, and publicly available threat intelligence platforms. [1].

Artificial Intelligence, specifically Machine Learning and Deep Learning, have brought new and powerful features to the detection of suspicious activity, classification of malicious activity and detection of network anomalies. But, ML methods, in general, require structured data,

feature engineering, and adequate labeled training samples. When attack characteristics shift, or when new threats are encountered, their performance decreases. Thus, Generative Artificial Intelligence (GenAI) has become a potential technology to learn sophisticated data distributions and create new data using learned pattern [3].

Large Language Models stand out among these technologies due to their capability to comprehend and craft human-readable text, summarise intricate information, and spot relationships in unstructured data. In the field of cybersecurity, LLMs can help security analysts analyze incident reports, identify Indicators of Compromise, detect suspicious activities, and provide explanations for potential threats. For instance, an LLM can be used to ingest a malware analysis report and glean potentially important IP addresses, file hashes, malicious domains, and techniques used in an attack [4].

While these benefits are promising, the direct use of generative AI in cybersecurity comes with a set of challenges. A significant challenge with LLM is hallucination, where the LLM produces information that is false, misleading, or not supported but appears convincing. In cybersecurity, inaccurate threat intelligence can result in the right decisions being made with the wrong threat intelligence, an unnecessary incident response, and failing to see the real threats. Moreover, generative models can have problems with obsolete information, incomplete threat data, adversarial inputs to the models, privacy issues, and the interpretability of the generative output. One of the key challenges noted by Zhang et al. [1] in applying LLMs to cybersecurity is model vulnerabilities and reliability.

Retrieval-Augmented Generation (RAG) offers a viable solution for incorporating external knowledge sources with generative AI to overcome these limitations. RAG was first proposed in Lewis et al. [5] and is the merging of the knowledge-generating ability of the language model with information obtained from external databases or a collection of documents. The model does not rely on only the previously learned information but retrieves relevant documents and then generates the contextually informed responses based on the information from the retrieved documents.

In this context, this research aims at the development of Generative AI-Driven Framework for Automated Cyber Threat Intelligence Analysis and Detection. The proposed framework focuses on using retrieval-based systems to gather, process, and analyze cybersecurity information from various sources, including LLMs. It will concentrate on identifying pertinent threat indicators, categorizing cyber threats, and obtaining useful information from an external cyber security knowledge base and creating evidence-based threat intelligence reports. The framework will explore the manner in which generative AI can be assessed for reliability and effectiveness through experimental evaluations with suitable cybersecurity datasets and performance measures.

2. Literature Review

One of the key technologies in cybersecurity, Artificial Intelligence (AI), plays a crucial role in the detection of suspicious activity from massive amounts of data. Most traditional cybersecurity solutions are based on preprogrammed rules and signatures, and they might not be effective against attacks that are not yet known or are constantly changing. Intrusion Detection, Malware classification, Phishing Detection and Anomaly Identification are areas where machine learning and deep learning techniques have been adopted. However, traditional AI models demand structured training data and may struggle to interpret intricate cybersecurity reports and provide context-based explanations [7].

Generative Artificial Intelligence opened the door to new opportunities in cybersecurity, as it was able to understand, analyze, and generate human-readable information. LLMs that can help security professionals summarize technical reports, provide context for security alerts, extract threat indicators, and aid with incident investigations. Uddin et al. [6] offered a comprehensive overview of the use of Generative AI in cybersecurity, its implementation in threat intelligence, security operations, and threat detection. In the same manner, Ferrag et al. have explored the use of LLMs in cybersecurity, such as intrusion detection, malware analysis,

phishing detection, and cyber threat intelligence. [6].

Cyber Threat Intelligence involves collecting, processing, and analyzing information about cyber threats to support effective security decisions. Traditional CTI analysis often relies on security analysts manually searching through long reports to find IoCs (indicators of compromise), malware characteristics, attack patterns and adversary techniques. Balasubramanian et al. [8] explored the usage and difficulties of Generative AI in cyber threat intelligence.

To address this, Rahman et al. [9] suggested ALERT, which leverages active learning and Large Language Models to automatically map CTI reports with attack techniques to the MITRE ATT&CK framework. By utilizing LLMs to automatically extract threat intelligence information, their approach will turn down the need for manual annotation. Its main goal is to extract and map attack techniques, though, and not to integrate with a multi-source knowledge retrieval, threat classification, evidence verification and automated report generation process [4].

A major drawback of LLMs is that the information they have learned can be antiquated or inadequate for comprehending new cyber threats. Retrieval-Augmented Generation (RAG) offers a solution to overcome this challenge, as it is a hybrid approach that combines information retrieval and generative AI. The RAG architecture was proposed by Lewis et al. [5] which incorporates an external information retrieval module with a language generation model. They showed that retrieval of relevant information can enhance the factual specificness of the produced answers. In cybersecurity, RAG can be used to fetch pertinent threat information, vulnerability reports, and attack techniques prior to an LLM producing an analytical output [5].

Research Challenges and Gap identified

While there have been advancements in the use of Generative AI for cybersecurity, there are still some challenges to address. LLMs can hallucinate information, misinterpret technical indicators, or provide threat classifications that are not supported. Besides, cybersecurity information is spread out between different sources and are often portrayed in an unstructured way, which makes it hard to perform a comprehensive analysis. Alqahtani and Kumar [6] surveyed existing works on cybersecurity intelligence research using LLMs and presented some of the problems such as hallucination, adversarial vulnerabilities, distribution shifts, and evaluation practices. This highlights the need for evidence-based verification and systematic performance evaluation of GenAI-based security systems [2, 6].

Existing studies have made significant strides in the application of GenAI in cyber security, as revealed by the literature reviewed. There is a need, however, for further research that combines multi-source CTI collection, automated IoC extraction, knowledge retrieval from RAG, contextual threat classification, and the generation of evidence-based reports under a single framework. To fill this gap, the present study suggests a framework that integrates LLM for threat intelligence extraction and RAG for knowledge retrieval [6].

3. Research Methodology

Following sections describes the methodology adopted in the research.

3.1 Research Design

The automatic Cyber Threat Intelligence (CTI) Analysis and Detection Framework presented in this study uses Generative AI. The key goal is to leverage Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) to enhance the detection, retrieval, categorization, and understanding of cyber threats. The limitations of traditional approaches to threat intelligence, such as manual information processing, fragmented threat information, limited context and unreliable threat analysis, are addressed via the proposed framework. The framework is composed of seven key phases: Cybersecurity Data Collection, Data Preprocessing, LLM-based Threat Intelligence Extraction, RAG-based Knowledge Retrieval, Threat Analysis and Classification, Automated Report Generation, and Performance Evaluation.

3.2 Cybersecurity Data Collection

The initial one is gathering cybersecurity information from trusted and public sources. These sources offer information on emerging threats, malicious activity, vulnerabilities and attack techniques. The framework proposed relies on known tools and resources like MITRE ATT&CK, Common Vulnerabilities and Exposures (CVE) database, CISA Known Exploited Vulnerabilities (KEV) catalog, PhishTank, and cybersecurity threat reports.

3.3 Data Preprocessing

This information can include unstructured text, duplicate information, irrelevant information, and inconsistencies in terminology. So, data is preprocessed to enhance the data quality for analysis. Text extraction, data cleaning, data normalization, removal of duplicates and text segmentation are done in this stage.

Security related indicators, such as IP addresses, URLs, domain names, file hashes, and vulnerability identifiers are retained during preprocessing. To enable efficient processing and information retrieval large cybersecurity reports are broken into smaller meaningful sections. Evidence verification is also enabled with the retention of metadata, such as source information and publication dates.

3.4 Extracting Threat Intelligence with LLM

Collected information is preprocessed and then fed into a Large Language Model for automated threat intelligence extraction. The LLM automatically recognizes key security-related entities and pulls out indicators of compromise, such as malicious IP addresses, malicious domains, malicious URLs, file hashes, malware names, and vulnerability identifiers.

Where there is enough information available, the model also correlates attacks with descriptions, targeted systems and adversary techniques. Standardized outputs consist of threat indicators, attack descriptions, and relevant security information, all generated using so-called "structured prompts.

3.5 RAG-Based Knowledge Retrieval

The proposed framework uses Retrieval-Augmented Generation to enhance the reliability of the generated threat intelligence. RAG pairs information retrieval with generative AI, allowing the LLM to include the relevant information from sources in the external world before generating a response. Firstly, a cybersecurity document is turned into a numerical form, known as embedding, and is stored in a vector database (e.g., FAISS or Chroma). The retrieval component searches the database for information, which is semantically related to the threat indicator or the security-related query. The retrieved documents are provided to LLM as a set of support context.

For instance, if a malware name or attack technique is discovered, the retrieval component will look for correlating threat reports, vulnerability information, and applicable MITRE ATT&CK techniques. This process helps the model to provide context-aware analysis based on cybersecurity evidence from the outside.

3.6 Threat analysis and classification using Generative AI

During this phase, the LLM processes the threat data extracted and also the evidence found via RAG. The goal is to determine what the cyber threat is and categorize it and label it as phishing, malware, ransomware, distributed denial-of-service (DDoS), credential-based attacks, or advanced persistent threats (APTs).

Where available, there is evidence of relevant MITRE ATT&CK techniques identified by the framework. The classifications generated are compared with the information retrieved, and claims that are not supported or inconsistencies are highlighted. Unsure findings are marked as potential threats for further investigation using a confidence assessment.

3.7: Generate Automated Threat Intelligence Reports

Threat analysis is followed by the production of a structured Cyber Threat Intelligence report. Every report contains the threat category identified, threat description, IoCs extracted, attack techniques referenced, supporting evidence and potential security impact. It can also offer

protection suggestions like blocking malicious indicators, updating vulnerable software, bolstering authentication measures, and spotting suspicious network activities.

Security analysts can use the generated reports to gain insight into the threats that are identified and use that information to make intelligent security decisions. High impact or uncertain results are sent for human verification before next step.

3.8 Performance Evaluation

Suitable cybersecurity datasets and manually verified reference labels will be used to assess the proposed framework. Standard metrics such as accuracy, precision, recall, F1-score will be used to evaluate its performance. Also, the proportion of the factual claims made by the system that are not backed by evidence or that contradict evidence will be measured by the hallucination rate. The relevance of evidence and processing times will also be discussed.

The framework will compare with an LLM based approach without RAG and a traditional machine-learning classification model. This comparison will be used to assess the role retrieval-based evidence and generative AI have to play in threat intel analysis.

Overall, the proposed methodology combines LLM-based extraction, RAG-based knowledge retrieval, contextual threat classification, and automated reporting to create a more efficient, reliable, and evidence-driven cyber threat intelligence analysis.

Conclusion:

That is why this study suggests an integrated Generative AI and Retrieval-Augmented Generation framework for the automated Cyber Threat Intelligence analysis and detection. The framework seeks to integrate LLMs with external cybersecurity knowledge bases to automate the extraction of threat information, identification of IoCs, classification of threats, and generation of intelligence reports. The embedding of evidence verification and confidence assessment is designed to enhance the reliability of the results produced and minimise the number of unsupported pieces of information. Additionally, structured reporting can help security analysts to grasp new threats and to determine what security measures to take. The framework is proposed to be measured by the traditional performance parameters to judge the accuracy, reliability, and efficiency in cybersecurity applications.

REFERENCES:

[1] J. Zhang et al., "When LLMs meet cybersecurity: A systematic literature review," Cybersecurity, vol. 8, Art. no. 55, 2025. [Online]. Available: https://doi.org/10.1186/s42400-025-00361-w

[2] P. Balasubramanian et al., "Generative AI for cyber threat intelligence: Applications, challenges, and analysis of real-world case studies," Artificial Intelligence Review, vol. 58, Art. no. 336, 2025. [Online]. Available: https://doi.org/10.1007/s10462-025-11338-z

[3] M. A. Ferrag et al., "Generative AI in cybersecurity: A comprehensive review of LLM applications and vulnerabilities," Internet of Things and Cyber-Physical Systems, vol. 5, pp. 1–46, 2025. [Online]. Available: https://doi.org/10.1016/j.iotcps.2025.01.001

[4] MITRE, "MITRE ATT&CK: Adversarial Tactics, Techniques, and Common Knowledge." [Online]. Available: https://attack.mitre.org/

[5] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in Proc. Advances in Neural Information Processing Systems (NeurIPS), vol. 33, 2020, pp. 9459–9474.

[6] M. Uddin et al., “Generative AI revolution in cybersecurity: A comprehensive review of threat intelligence and operations,” Artificial Intelligence Review, vol. 58, Art. no. 236, 2025.

[7] M. A. Ferrag et al., “Generative AI in cybersecurity: A comprehensive review of LLM applications and vulnerabilities,” Internet of Things and Cyber-Physical Systems, vol. 5, pp. 1–46, 2025.

IEER
Institute for Excellence in Education & Research

[8] P. Balasubramanian et al., "Generative AI for cyber threat intelligence: Applications, challenges, and analysis of real-world case studies," Artificial Intelligence Review, vol. 58, Art. no. 336, 2025.

[9] F. I. Rahman, S. M. Halim, A. Singhal, and L. Khan, "ALERT: A framework for efficient extraction of attack techniques from cyber threat intelligence reports using active learning," in Proc. 38th IFIP Int. Conf. Data and Applications Security and Privacy (DBSec), 2024, vol. 14901, pp. 203–220.

[10] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in Advances in Neural Information Processing Systems, vol. 33, pp. 9459–9474, 2020.

[11] H. Alqahtani and G. Kumar, "Large Language Models for Cybersecurity Intelligence: A Systematic Review of Emerging Threats, Defensive Capabilities, and Security Evaluation Frameworks," Computers, Materials & Continua, vol. 87, no. 3, 2026.

IEER
Institute for Excellence in Education & Research