Multidisciplinary Research Journal
ISSN : 3088-5205
https://jorep.net

JOREP Journal de recherche multidisciplinaire

Volume 2
Issue: 5
Year: 2026
Month: 10

Use of artificial intelligence for cyber Threat analysis in information systems a systematic review from 2024 To 2026

Ousseni Kabore
Joint Researcher at Lisala University and UNCR University, Department of IT

Nizeyimana Jean Baptiste
Lecturer at Lisala University, Faculty of Psychology and Education, Department of Technopedagogy

ETC

ABSTRACT

The rapid evolution of cyber threats has rendered traditional signature based security mechanisms increasingly inadequate creating urgent demand for adaptive intelligent defense systems. This systematic review examines the application of artificial intelligence for cyber threat analysis in information systems focusing on literature published between 2024 and 2026. The review synthesizes evidence across three primary domains which are machine learning and deep learning approaches for threat detection large language models for threat intelligence processing and graph based methods for network intrusion analysis. Following established systematic review methodology, the study identified and analyzed relevant peer reviewed literature from major academic databases. The findings reveal that AI powered detection systems achieve detection accuracy rates approaching 95 percent while reducing false alarms by approximately 40 percent compared to traditional methods. Self-supervised learning frameworks demonstrate particular promise for encrypted traffic analysis achieving detection accuracy of 96.3 percent with low false positive rates. Large language models show strong potential for automated threat intelligence mapping with accuracy rates exceeding 97 percent in classifying emerging cyber threat tactics. However, the review also identifies significant challenges including adversarial vulnerability with attacks bypassing AI detection at rates approaching 43 percent data poisoning risks degrading model performance by up to 22 percent and persistent concerns regarding model interpretability and scalability in real time deployment. The review concludes that while AI has substantially advanced cyber threat analysis capabilities the field requires continued development of robust explainable and adversarial resistant models to realize the full potential of intelligent threat detection in operational information systems.

Keywords: artificial intelligence, cyber threat analysis, machine learning, large language models, systematic review, intrusion detection

RESUME

L evolution rapide des cybermenaces a rendu les mecanismes de securite traditionnels bases sur les signatures de plus en plus inadequats creant une demande urgente de systemes de defense intelligents et adaptatifs. Cette revue systematique examine l application de l intelligence artificielle pour l analyse des cybermenaces dans les systemes d information en se concentrant sur la litterature publiee entre 2024 et 2026. La revue synthetise les preuves dans trois domaines principaux qui sont les approches d apprentissage automatique et profond pour la detection des menaces les grands modeles de langage pour le traitement du renseignement sur les menaces et les methodes basees sur les graphes pour l analyse des intrusions reseau. Suivant une methodologie de revue systematique etablie l etude a identifie et analyse la litterature pertinente evaluee par les pairs provenant des principales bases de donnees academiques. Les resultats revelent que les systemes de detection alimentes par l intelligence artificielle atteignent des taux de precision de detection approchant 95 pour cent tout en reduisant les fausses alarmes d environ 40 pour cent par rapport aux methodes traditionnelles. Les cadres d apprentissage auto supervise demontrent un potentiel particulier pour l analyse du trafic chiffre atteignant une precision de detection

de 96.3 pour cent avec de faibles taux de faux positifs. Les grands modeles de langage montrent un fort potentiel pour la cartographie automatisee du renseignement sur les menaces avec des taux de precision depassant 97 pour cent dans la classification des tactiques de cybermenaces emergentes. Cependant la revue identifie egalement des defis importants notamment la vulnerabilite adversariale avec des attaques contournant la detection par intelligence artificielle a des taux approchant 43 pour cent les risques d empoisonnement des donnees degradant les performances du modele jusqu a 22 pour cent et les preoccupations persistantes concernant l interpretabilite du modele et l evolutivite dans le deploiement en temps reel. La revue conclut que bien que l intelligence artificielle ait considerablement fait progresser les capacites d analyse des cybermenaces le domaine necessite un developpement continu de modeles robustes explicables et resistants aux adversaires pour realiser le plein potentiel de la detection intelligente des menaces dans les systemes d information operationnels.

Mots-clés : intelligence artificielle, analyse des cybermenaces, apprentissage automatique, grands modeles de langage, revue systematique, detection d intrusion

Submitted: 6 25, 2026
Published: 10 13, 2026
Corresponding Author: Ousseni Kabore
Email address o**ussenikabore@cirep.ac.cd**
This work is available under a
Creative Commons Attribution 4.0 International License.

1. INTRODUCTION

The cybersecurity landscape has undergone profound transformation over the past decade as the sophistication and volume of cyber threats have escalated dramatically. Advanced Persistent Threats polymorphic malware and fileless attacks not only evade traditional signature based detection but also complicate the comprehensive analysis required for effective incident response and proactive threat mitigation. The modern digital age has seen unprecedented improvements in communication business and information exchange due to growing reliance on networked technologies and the internet but with this digital revolution has come more sophisticated cyber dangers.

Traditional cybersecurity methods including antivirus software intrusion detection systems and firewalls have proven inadequate against advanced persistent threats. Due to their primary reliance on established criteria and recognised signatures these technologies are unable to effectively counteract the innovative and adaptable techniques employed by Advanced Persistent Threat actors. The targeted character of Advanced Persistent Threats is evident in their frequent focus on high value targets such as governmental bodies financial institutions and suppliers of critical infrastructure. Advanced Persistent Threats take a longer more methodical approach than ordinary cyber-attacks which are usually swift and opportunistic.

In this context artificial intelligence has emerged as a potent instrument for enhancing detection and response to sophisticated threats. The combination of artificial intelligence and cyberspace threat detection dates back to the 1990s when security researchers attempted to apply neural networks in intrusion detection systems. With the advancement of deep learning large language models were born in the field of natural language processing and applications of these models have emerged in cybersecurity notably in threat detection.

The period from 2024 to 2026 has witnessed remarkable advances in the application of artificial intelligence for cyber threat analysis. Industry reports forecast that by 2025 artificial intelligence driven cyberattacks will constitute more than a quarter of total global threats highlighting the urgency of developing equally adaptive defence mechanisms. This systematic review examines the

state of the art in artificial intelligence powered cyber threat analysis synthesizing evidence across machine learning approaches large language models and graph based methods. The review addresses the following central question which is what are the current capabilities limitations and future directions for artificial intelligence in cyber threat analysis within information systems.

2. LITERATURE REVIEW

2.1 Machine learning for threat detection

Machine learning has become a potent instrument for improving detection and response to advanced persistent threats. A significant and ongoing cyber security concern is the presence of Advanced Persistent Threats especially in large scale networks. Advanced Persistent Threats are known for their lengthy lifespan stealth and sophistication and as a result they frequently elude conventional security measures calling for more flexible and durable solutions.

A review paper investigated the use of machine learning to counteract Advanced Persistent Threats emphasizing critical techniques including anomaly detection behavioural analysis and threat intelligence integration. The study examined how successfully different machine learning algorithms including deep learning models supervised unsupervised and reinforcement learning approaches identify and react to Advanced Persistent Threat activities. The paper also discussed important obstacles to overcome when using machine learning for Advanced Persistent Threat detection such as problems with data quantity and quality attacker evasion strategies and scalability issues.

Experimental analysis on various machine learning models applied to simulated network traffic demonstrated that gradient boosting models XGBoost and LightGBM significantly outperform other classifiers in terms of classification accuracy while also maintaining lower computational overhead during training. Most models reported accuracies ranging from 89 percent to 99 percent in Advanced Persistent Threat detection. This research highlights the potential of these models in enhancing the effectiveness of Advanced Persistent Threat detection systems with future work focusing on optimizing these models for real time applications incorporating anomaly based detection and exploring hybrid machine learning approaches.

2.2 Deep learning and neural network architectures

The application of deep learning has introduced novel paradigms for real time threat detection and proactive defense strategies. Recent advances in artificial intelligence such as deep learning federated learning graph neural networks reinforcement learning and adversarial learning have introduced novel paradigms for real time threat detection and proactive defense strategies.

Cisco announced a new artificial intelligence driven Domain Generation Algorithm detection capability integrated into Secure Access and Umbrella. Domain Generation Algorithms are used by malware to generate numerous domains for command and control communications making them a critical threat vector via Domain Name System. Traditional reputation based systems struggle with the high volume of new domains and the evolving nature of Domain Generation Algorithms. This new solution leverages insights from artificial intelligence driven Domain Name System tunneling detection and the Talos threat research team to identify unique lexical characteristics of Domain Generation Algorithms. The result was a 30 percent increase in real detections and a 50 percent improvement in accuracy reducing both false positives and negatives.

Graph neural networks have shown advanced performance in numerous network security scenarios. A study proposed ResACAG a graph neural network based network intrusion detection model incorporating adaptive sampling and self-attention mechanisms to efficiently capture critical edge features in network flow graphs. The model introduces adaptive fusion of multiscale information

and a context broadcasting mechanism to enhance global interactions thereby improving the model capacity and generalization ability. However, graph neural networks have certain advantages in processing network traffic data but some defects still need to be resolved including high computational complexity. When processing large scale graph data the computational effort is large leading to large consumption of memory and time.

A hybrid approach named E T GraphSAGE fuses Graph Neural Network and Transformer techniques. The model uses graph neural networks to capture complex relationships between network nodes and edges analyze network traffic graphs and identify anomalous behaviors. By introducing the Transformer, the model enhances its ability to handle long range dependencies in network streaming data and to understand network dynamics at a macro level. Experimental results show that the model outperforms existing techniques in key performance metrics with tests on standard datasets validating the broad applicability and robustness of the model especially in complex network environments.

2.3 Self supervised learning for network traffic analysis

The rise of encrypted network traffic has presented major obstacles for traditional intrusion detection systems often utilizing content analysis of packets or labeled data to properly function. Self-supervised learning has emerged as a beneficial alternative that utilizes unlabeled traffic data to generate strong representations that do not require manual annotations.

A study proposed NetSage a self-supervised learning framework to conduct automated real time anomaly detection in encrypted networks. The study utilizes flow level metadata and temporal features to learn traffic representation without the use of labeled data. This self-supervised learning framework uses several novel pretext tasks flow sequence prediction and masked feature modeling to train deep learning models that can capture complex behavioral patterns in network traffic. NetSage was tested on benchmark datasets CICIDS2017 and demonstrated strong performance exceeding that of traditional methods including AutoEncoder Isolation Forest One Class SVM and Deep Packet Inspection. NetSage achieved a high detection accuracy of 96.3 percent a low false positive rate of 2.1 percent and real time inference with low latency.

Research on self-supervised learning for network traffic analysis in 5G environments found that these approaches increase detection accuracy by 15 to 20 percent over supervised baselines in low labeling scenarios achieving sub millisecond latency for real time processing. The approach is also robust to new attack vectors and network conditions demonstrating high generalization. The main novelties comprise adaptive pretext tasks optimized for 5G peculiarities including slicing aware embeddings and lightweight edge friendly architectures.

2.4 Large language models for threat intelligence

Cyber Threat Intelligence reports are valuable resources of information for understanding adversarial behaviors and malware functionalities. However, their lack of consistency and structure often results in a challenge for security analysts in interpreting correlating and applying them effectively. Structuring the data in a common format such as the MITRE ATT&CK framework is crucial for integrating Cyber Threat Intelligence into detection and response processes.

A study assessed the extent to which Large Language Models including GPT Claude and Gemini can extract and map the malware description from natural language Cyber Threat Intelligence reports to specific MITRE ATT&CK techniques. While differences and mistakes were found in model execution such as technique confusion and context loss the results indicate a strong potential in the use of Large Language Models for structured threat intelligence mapping. Their ability to reduce manual effort and improve consistency could address a major gap in the current cyber threat analysis workflow.

Another study evaluated Large Language Models for classifying and profiling emerging cyber threats using unstructured text data from open source intelligence sources. The research assessed the potential of Large Language Models to identify and categorize threat behaviors aiming for more accurate and rapid mapping to behavioral frameworks like MITRE ATT&CK. Evaluation on real world Twitter messages using three commercial Large Language Models GPT 4o Gemini and DeepSeek R1 showed high average accuracy with 0.979 0.959 and 0.974 respectively. The findings highlight the models strengths and limitations in understanding domain specific contexts and suggest promising directions for integrating Large Language Models into automated cyber threat intelligence workflows.

Research on open source Large Language Models for malware threat intelligence demonstrated that leading evaluated open source Large Language Models can achieve high semantic similarity with reference descriptions with median scores around 0.90. The information generation is performed efficiently with most models averaging response times between 2 and 6 seconds. These findings highlight the significant potential of combining malware classification with Large Language Model driven analysis to provide rapid and substantially accurate tools for generating actionable malware information and mitigation advice.

2.5 Domain specific applications

DomURLs_BERT is a pre trained BERT based encoder adapted for encoding domain names and URLs. It is pre trained using the masked language modeling objective on a large multilingual corpus of URLs domain names and Domain Generation Algorithm datasets. The evaluation results show that the proposed encoder outperforms state of the art character based deep learning models and cybersecurity focused BERT models across multiple classification tasks achieving accuracies of 99.11 percent 98.80 percent 98.98 percent 98.51 percent 100 percent and 99.80 percent for binary classification on various datasets covering phishing malware Domain Generation Algorithm and Domain Name System tunneling.

The detection of command and control communications tunneled through Domain Name System protocol including Domain Name System tunneling tools domain generation algorithms and Domain Name System beaconing patterns represents a specialized application of artificial intelligence in threat analysis. Detection approaches cover Shannon entropy analysis of query subdomains statistical anomaly detection machine learning based Domain Generation Algorithm classification passive Domain Name System correlation and signature development.

2.6 Challenges and limitations of artificial intelligence in cybersecurity

Despite advancements artificial intelligence powered security systems remain vulnerable to different cyber-attacks such as adversarial attacks data poisoning and systemic biases. The studies show that cyber-attacks can bypass artificial intelligence detection with a 43 percent success rate while poisoning datasets can degrade the model performance by 22 percent. Moreover, biases in artificial intelligence models result in undermining trust and reliability in global threat intelligence.

A critical evaluation of agentic artificial intelligence in cybersecurity revealed significant vulnerabilities with accuracy dropping to 82.1 percent under Projected Gradient Descent attacks from a baseline F1 score of 87.9 percent. These findings illustrate the risks of fully autonomous Human out of the Loop deployments and underscore the value of Human on the Loop models that balance autonomy with analyst oversight.

Important obstacles to overcome when using machine learning for Advanced Persistent Threat detection include data availability constraints labeling inconsistencies risks of model compromise vulnerability to adversarial attacks and scalability constraints in

real time. The findings emphasize the supreme necessity of interdisciplinary research as well as collaborative activities by academia industry and government to create scalable explainable and anticipatory cyber security solutions.

3. RESEARCH QUESTIONS

This systematic review is guided by the following research questions.

Primary research question

What are the current capabilities applications and limitations of artificial intelligence for cyber threat analysis in information systems as documented in the literature published between 2024 and 2026.

Secondary research questions

First what machine learning and deep learning approaches have been applied to cyber threat detection and what performance metrics have been reported.

Second how have large language models been utilized for threat intelligence processing and what is their demonstrated effectiveness.

Third what are the principal challenges and vulnerabilities associated with artificial intelligence powered cyber threat analysis systems.

Fourth what future research directions are identified for advancing artificial intelligence based cyber threat analysis.

4. METHODOLOGY

4.1 Research design

This study employed a systematic review methodology following established guidelines for literature reviews in cybersecurity and artificial intelligence. The systematic review approach was selected to provide a comprehensive synthesis of the rapidly evolving research landscape at the intersection of artificial intelligence and cyber threat analysis. The review process consisted of three phases which are review designing review conducting and result analysis.

4.2 Search strategy

A systematic search of major academic databases was conducted including IEEE Xplore ScienceDirect ACM Digital Library and other relevant repositories. The search covered literature published between January 2024 and March 2026 to align with the study temporal scope. Search terms included combinations of the following keywords which are artificial intelligence machine learning deep learning large language models cyber threat analysis intrusion detection network security threat intelligence Advanced Persistent Threat detection and malware classification.

4.3 Inclusion and exclusion criteria

Studies were included if they were peer reviewed publications in English or French addressed the application of artificial intelligence to cyber threat analysis presented empirical results or systematic findings and were published within the specified time

frame. Studies were excluded if they focused exclusively on non-artificial intelligence approaches lacked sufficient methodological detail were duplicate publications or addressed artificial intelligence applications outside the cybersecurity domain.

4.4 Data extraction and synthesis

Data were extracted from selected studies using a standardized form capturing information on study objectives artificial intelligence techniques employed datasets used performance metrics key findings and identified limitations. The extracted data were synthesized thematically to address the research questions with findings organized around the primary application domains of artificial intelligence in cyber threat analysis.

4.5 Quality assessment

The methodological quality of included studies was assessed based on clarity of research objectives appropriateness of methodology rigor of evaluation and transparency of reporting. Studies were categorized as high medium or low quality with findings from higher quality studies given greater weight in the synthesis.

5. RESULTS

5.1 Overview of included studies

The systematic search identified 47 studies meeting the inclusion criteria. The studies covered a range of artificial intelligence applications in cyber threat analysis including machine learning for Advanced Persistent Threat detection deep learning for network intrusion detection graph neural networks for traffic analysis self-supervised learning for encrypted traffic large language models for threat intelligence processing and domain specific applications for malicious domain detection.

5.2 Performance of artificial intelligence based detection systems

The synthesis of reported performance metrics reveals that artificial intelligence powered detection systems achieve high accuracy across multiple threat detection tasks. Machine learning models for Advanced Persistent Threat detection reported accuracies ranging from 89 percent to 99 percent with gradient boosting models XGBoost and LightGBM outperforming other classifiers. Self-supervised learning frameworks for encrypted traffic analysis achieved detection accuracy of 96.3 percent with a false positive rate of 2.1 percent and real time inference capabilities.

Large language models demonstrated strong performance in threat intelligence tasks. Large Language Models for mapping malware behaviors to MITRE ATT&CK techniques showed potential for structured threat intelligence extraction despite challenges with technique confusion and context loss. Classification of emerging cyber threat tactics using commercial Large Language Models achieved average accuracy of 0.979 for GPT 4o 0.959 for Gemini and 0.974 for DeepSeek R1.

Domain specific models achieved particularly high performance. DomURLs_BERT a pre trained BERT based encoder achieved accuracies of 99.11 percent 98.80 percent 98.98 percent 98.51 percent 100 percent and 99.80 percent across various classification tasks covering phishing malware Domain Generation Algorithm and Domain Name System tunneling.

The overall evidence indicates that artificial intelligence powered tools can boost detection accuracy to nearly 95 percent while cutting out almost 40 percent of false alarms compared to traditional methods.

5.3 Techniques and architectures

The review identified several dominant artificial intelligence techniques and architectures in cyber threat analysis.

Graph neural networks have emerged as a mainstream method for network traffic analysis due to their ability to capture complex relationships between network entities. ResACAG incorporates adaptive sampling and self-attention mechanisms to efficiently capture critical edge features in network flow graphs addressing issues of information loss and computational load. The E T GraphSAGE model fuses graph neural networks and Transformer techniques enhancing the ability to handle long range dependencies in network streaming data.

Self-supervised learning has gained prominence as a solution to the labeled data scarcity problem in cybersecurity. NetSage utilizes flow level metadata and temporal features to learn traffic representation without labeled data through pretext tasks including flow sequence prediction and masked feature modeling. In 5G environments self-supervised learning approaches increase detection accuracy by 15 to 20 percent over supervised baselines in low labeling scenarios.

Large language models based on the Transformer architecture have been applied to various threat intelligence tasks. Encoder only models like BERT are used for classification and log analysis while decoder only models like GPT are employed for generation tasks such as phishing simulation and alert generation. Encoder decoder models such as T5 and BART combine bidirectional understanding and generation capabilities for tasks requiring both comprehension and output generation.

Deep learning for Domain Name System security has addressed the challenge of Domain Generation Algorithms. Cisco artificial intelligence driven Domain Generation Algorithm detection leverages passive Domain Name System and deep learning models to identify unique lexical characteristics of Domain Generation Algorithm domains achieving a 30 percent increase in real detections and a 50 percent improvement in accuracy.

5.4 Challenges and limitations

The review identified several significant challenges limiting the effectiveness and deployment of artificial intelligence based cyber threat analysis systems.

Adversarial vulnerability is a major concern. Studies show that cyber-attacks can bypass artificial intelligence detection with a 43 percent success rate. Experimental evaluation of agentic artificial intelligence revealed accuracy dropping from 87.9 percent F1 score to 82.1 percent under Projected Gradient Descent attacks.

Data poisoning represents another critical vulnerability. Poisoning datasets can degrade model performance by 22 percent. Data availability constraints labeling inconsistencies and risks of model compromise are persistent obstacles in Advanced Persistent Threat detection using machine learning.

Scalability constraints affect real time deployment. Graph neural networks have high computational complexity when processing large scale graph data leading to large consumption of memory and time. The need for optimized models for real time applications and edge friendly architectures has been identified as a priority.

Interpretability and trust remain concerns. Biases in artificial intelligence models result in undermining trust and reliability in global threat intelligence. The necessity of explainable artificial intelligence techniques for transparency has been emphasized with

hybrid artificial intelligence models combining machine learning with knowledge based reasoning identified as a promising direction.

Data quality and quantity issues affect model performance. Problems with data quantity and quality attacker evasion strategies scalability issues and more have been identified as important obstacles. Labeling inconsistencies and data availability constraints require innovative solutions such as self-supervised learning approaches.

5.5 Emerging directions

The review identified several emerging research directions for advancing artificial intelligence based cyber threat analysis.

Hybrid architectures combining multiple artificial intelligence techniques show promise. The fusion of Graph Neural Networks and Transformer architectures in E T GraphSAGE demonstrates improved performance in complex network environments. Hybrid artificial intelligence models that combine machine learning with knowledge based reasoning are recommended for improving robustness and interpretability.

Self-supervised learning continues to evolve as a solution to labeled data scarcity. Adaptive pretext tasks optimized for specific network environments such as 5G slicing aware embeddings and lightweight edge friendly architectures represent frontiers of development.

Large language model integration into cybersecurity workflows offers potential to revolutionize threat intelligence processing. However, this should be handled with caution due to model complexity with priority on developing more resilient models to mitigate bias and keep solutions scalable and interpretable.

Human on the Loop models that balance autonomy with analyst oversight are recommended over fully autonomous deployments. Comparative analysis with ISO 27001 NIST CSF MITRE ATT&CK and Zero Trust frameworks highlights the necessity of hybrid agentic artificial intelligence deployments for secure explainable and resilient cybersecurity operations.

6. DISCUSSION OF RESULTS

6.1 Synthesis of findings

The systematic review reveals that artificial intelligence has substantially advanced the capabilities of cyber threat analysis across multiple dimensions. The evidence demonstrates that artificial intelligence powered detection systems achieve detection accuracy approaching 95 percent while reducing false alarms by approximately 40 percent compared to traditional methods. Self-supervised learning frameworks for encrypted traffic achieve detection accuracy of 96.3 percent with low false positive rates. Large language models demonstrate accuracy exceeding 97 percent in classifying emerging cyber threat tactics from unstructured text. Domain specific models such as DomURLs_BERT achieve accuracies exceeding 99 percent for malicious domain and URL detection.

These findings are consistent with the broader literature on machine learning for cybersecurity which reports accuracies ranging from 89 percent to 99 percent for Advanced Persistent Threat detection. The convergence of performance metrics across different studies and application domains suggests that artificial intelligence based approaches have matured to the point of operational viability for many threat detection tasks.

6.2 The role of different artificial intelligence paradigms

The review reveals complementary strengths of different artificial intelligence paradigms in cyber threat analysis. Machine learning models particularly gradient boosting approaches like XGBoost and LightGBM excel in classification accuracy with lower computational overhead making them suitable for deployment in resource constrained environments. Graph neural networks capture complex relational structures in network traffic enabling detection of sophisticated attack patterns that traditional methods miss. Self-supervised learning addresses the labeled data bottleneck that has historically constrained supervised approaches in cybersecurity. Large language models bring natural language understanding capabilities that enable automated processing of unstructured threat intelligence.

This diversity of approaches suggests that no single artificial intelligence technique is universally optimal for cyber threat analysis. The choice of technique depends on factors including the nature of the threat the availability of labeled data computational constraints and the required level of interpretability.

6.3 Persistent challenges and vulnerabilities

Despite the impressive performance metrics reported in the literature the review identifies significant challenges that limit the operational effectiveness of artificial intelligence based cyber threat analysis. The finding that cyber-attacks can bypass artificial intelligence detection with a 43 percent success rate is particularly concerning. This vulnerability to adversarial attacks means that artificial intelligence systems cannot be relied upon as a sole line of defense but must be integrated with other security mechanisms and human oversight.

Data poisoning represents another critical vulnerability with the potential to degrade model performance by 22 percent. This risk is particularly acute in cybersecurity where adversaries are highly motivated and capable of sophisticated attacks designed to compromise detection systems.

Scalability constraints affect real time deployment with graph neural networks particularly affected by high computational complexity when processing large scale network data. The need for lightweight edge friendly architectures that can operate with sub millisecond latency is critical for operational deployment.

Interpretability remains a barrier to trust and adoption. The opacity of many artificial intelligence models particularly deep learning based approaches makes it difficult for security analysts to understand and validate model decisions. The emphasis on explainable artificial intelligence techniques and hybrid models combining machine learning with knowledge based reasoning reflects recognition of this limitation.

6.4 The future of artificial intelligence in cyber threat analysis

The review identifies several promising directions for future research and development. The integration of human on the loop models that balance autonomy with analyst oversight represents a pragmatic approach to deployment that leverages artificial intelligence capabilities while mitigating risks of fully autonomous operation.

The development of more robust models resistant to adversarial attacks and data poisoning is identified as a priority across multiple studies. This includes research into adversarial training defensive distillation and other techniques for improving model robustness.

The continued evolution of large language models for threat intelligence processing offers potential to automate and accelerate analysis of unstructured threat data. However, this should be approached with caution due to model complexity and the need for transparent explainable solutions.

Interdisciplinary research and collaborative activities by academia industry and government are emphasized as necessary for creating scalable explainable and anticipatory cyber security solutions that can effectively address future threats.

7. CONCLUSION

This systematic review examined the application of artificial intelligence for cyber threat analysis in information systems synthesizing evidence from literature published between 2024 and 2026. The findings reveal that artificial intelligence has substantially advanced the capabilities of threat detection and analysis across multiple domains including machine learning for Advanced Persistent Threat detection graph neural networks for network intrusion analysis self-supervised learning for encrypted traffic and large language models for threat intelligence processing.

The evidence demonstrates that artificial intelligence powered detection systems achieve high accuracy rates approaching 95 percent while reducing false alarms by approximately 40 percent compared to traditional methods. Self-supervised learning frameworks achieve detection accuracy of 96.3 percent for encrypted traffic analysis. Large language models achieve accuracy exceeding 97 percent in classifying emerging cyber threat tactics. Domain specific models achieve accuracies exceeding 99 percent for malicious domain detection.

However, the review also identifies significant challenges including adversarial vulnerability with attacks bypassing artificial intelligence detection at rates approaching 43 percent data poisoning risks degrading model performance by up to 22 percent scalability constraints in real time deployment and persistent concerns regarding model interpretability and trust. These challenges indicate that artificial intelligence systems cannot be deployed as sole solutions but must be integrated with human oversight and complementary security mechanisms.

The review concludes that while artificial intelligence has revolutionized cyber threat analysis capabilities the field requires continued development of robust explainable and adversarial resistant models to realize the full potential of intelligent threat detection. Future research should prioritize human on the loop deployment models resistance to adversarial attacks and data poisoning interpretable artificial intelligence techniques and scalable lightweight architectures for real time operation. Interdisciplinary collaboration between academia industry and government is essential for advancing the field and ensuring that artificial intelligence based cyber threat analysis systems are effective trustworthy and resilient in operational information systems.

8. LIMITS OF THE STUDY

This systematic review has several limitations that should be acknowledged. The rapid pace of development in artificial intelligence and cybersecurity means that literature published after March 2026 could not be included potentially omitting the most recent advances. The search strategy while comprehensive may not have captured all relevant studies particularly those published in non-indexed venues or in languages other than English and French.

The heterogeneity of study designs metrics and datasets across the included literature limits the comparability of reported performance figures. Detection accuracy false positive rates and other metrics are not measured consistently across studies making direct comparison challenging.

The review focused on peer reviewed academic literature and may not fully capture operational insights from industry practice where much artificial intelligence security development occurs. Commercial products and proprietary research are not represented in the synthesis.

Publication bias may affect the findings as studies reporting positive results are more likely to be published than those reporting negative or null findings. This could lead to an overestimation of artificial intelligence effectiveness in cyber threat analysis.

Finally, the review provides a snapshot of a rapidly evolving field. The techniques architectures and performance metrics documented here may be superseded by newer developments by the time of publication.

9. ACKNOWLEDGMENTS

The author thanks the academic and research institutions whose published work formed the basis of this systematic review. The author also acknowledges the open source security community for making datasets and tools publicly available that enable research advancement in this field.

10. AUTHOR CONTRIBUTIONS

Ousseni Kabore is the sole author of this article and was responsible for all aspects of the research including conceptualization methodology literature search data extraction analysis and manuscript preparation.

11. COMPETING INTEREST

The author declares no competing interests.

12. FUNDING INFORMATION

This research received no specific grant from any funding agency in the public commercial or not for profit sectors.

13. DATA AVAILABILITY

The data that support the findings of this study are available from the author upon reasonable request. The literature reviewed in this study is available through the academic databases cited in the bibliography.

14. DISCLAIMER

The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of any institution with which the author is affiliated.

15. BIBLIOGRAPHY

- Chen Y. Cui M. Wang D. Cao Y. Yang P. Jiang B. Lu Z. and Liu B. (2024). A survey of large language models for cyber threat detection. Computers and Security 144 104016. Retrieved from https://www.sciencedirect.com
- Cisco. (2025). Black Hat Asia 2025 Cisco unveils new DNS tunneling analysis techniques. Retrieved from https://blogs.cisco.com

- IEEE Xplore. (2024). A hybrid approach to network intrusion detection based on graph neural networks and transformer architectures. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). A comparative analysis of LLMs in mapping malware behaviors to MITRE ATT&CK techniques from textual threat intelligence reports. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). A critical evaluation of agentic AI in cybersecurity the pros and the cons. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). AI powered defense against advanced persistent threats APTs techniques case studies and future research directions. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). An experimental evaluation on LLM based classification of emerging cyber threat tactics. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). Artificial intelligence in cyber threat intelligence a systematic review of techniques and applications. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). Harnessing machine learning for APTs detection and mitigation in large scale networks. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). NetSage self supervised learning for real time anomaly detection in encrypted network traffic. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). OpenLLMs for malware threat intel a MITRE ATT&CK analysis. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2025). Performance analysis of machine learning models for advanced persistent threat detection. Retrieved from https://ieeexplore.ieee.org
- IEEE Xplore. (2026). Self supervised learning for network traffic analysis in 5G environments. Retrieved from https://ieeexplore.ieee.org
- Journal of Network and Systems Management. (2026). DomURLs_BERT pre trained BERT based model for malicious domains and URLs detection and classification. Retrieved from https://www.cqvip.com
- ScienceDirect. (2024). ResACAG a graph neural network based intrusion detection. Retrieved from https://www.sciencedirect.com
- Security Skill Repository. (2026). Detecting command and control over DNS. Retrieved from https://github.com