Computing&AI Connect
ISSN: 3104-4719
2026, Article ID. 2026.0031, Cite as: https://doi.org/10.69709/CAIC.2026.155073

SCIFINITI PUBLISHING

Perspective

OPEN ACCESS

Fine-Tuning vs. RAG: A Position Paper from the Perspective of LLM-Based Cybersecurity Modeling

Mohammad Amaz Uddin[✉,1,2,*] **Iqbal H. Sarker**[✉,3,*]

[1] Department of Computer Science and Engineering, International Islamic University Chittagong, Chattogram 4318, Bangladesh
[2] Department of Computer Science and Engineering, BGC Trust University Bangladesh, Chittagong 4381, Bangladesh
[3] Centre for Securing Digital Futures, Edith Cowan University, Perth, WA 6027, Australia

Article History
Received: July 04, 2025 Accepted: January 27, 2026 Published: March 26, 2026

Abstract
Large Language Models (LLMs) have revolutionized natural language processing (NLP) tasks, enabling critical capabilities across domains such as business, finance, and cybersecurity analysis. In particular, LLMs are becoming increasingly important in cybersecurity by supporting threat and vulnerability analysis, intelligent response generation, pattern recognition, and automated threat detection. Although LLMs are pre-trained with a large amount of knowledge, their application to specific domains, such as cybersecurity, is often limited by insufficient domain-specific coverage in their pre-training data. To address this limitation, domain-specific knowledge is typically incorporated through fine-tuning or retrieval-augmented generation (RAG). RAG enriches the model's responses at inference time by retrieving relevant external information, while fine-tuning embeds new knowledge directly into the model's parameters. In this position paper, we examine fine-tuning and RAG in the context of cybersecurity applications. Moreover, we discuss the hybrid approach combining fine-tuning and RAG, and highlight the most appropriate scenarios for selecting fine-tuning, RAG, or a hybrid approach based on the desired cybersecurity use case and operational constraints. Overall, this position paper aims to contribute to the ongoing discussion on LLM adaptation strategies and best practices for applying RAG and/or fine-tuning.

Keywords:
large language models; natural language processing; retrieval-augmented generation; GraphRAG; cybersecurity

1. Introduction

In recent years, the rapid advancement of LLMs has substantially broadened the applications of Artificial Intelligence (AI), particularly in domains such as cybersecurity, where they automate and expedite tasks that previously required considerable human effort. Modern LLMs, with notable examples including Google's Gemini and OpenAI's Generative Pre-trained Transformer (GPT) series, have transformed natural language processing (NLP) by achieving advanced capabilities in contextual reasoning, text generation, and comprehension [1]. During pre-training, LLMs are exposed to diverse datasets, enabling broad generalization and domain expertise. Their ability to understand, generate, and reason over complex information makes them valuable tools in cybersecurity, analyzing logs, emails, and network traffic to detect phishing attacks, malware, spam, or insider threats, often faster while reducing manual workload, thereby saving time and costs.

Language models acquire knowledge from diverse datasets [2], but struggle with uncommon, domain-specific, or inaccurate data [3], and cannot incorporate new information or changing facts after training, as their knowledge is static. This can result in temporal degradation, outdated or inaccurate responses, and hallucinations. These issues are concerning in high-stakes fields like cybersecurity, where accuracy, adaptability, and up-to-date threat intelligence are crucial for tasks such as threat de-

* **Corresponding Authors:**
Mohammad Amaz Uddin, Department of Computer Science and Engineering, International Islamic University Chittagong, Chattogram 4318, Bangladesh; Department of Computer Science and Engineering, BGC Trust University Bangladesh, Chittagong 4381, Bangladesh, amazuddin722@gmail.com;
Iqbal H. Sarker, Centre for Securing Digital Futures, Edith Cowan University, Perth 6027, WA, Australia, m.sarker@ecu.edu.au

tection, anomaly identification, and malware analysis. Researchers have focused on customizing LLMs for specific industries, such as cybersecurity, healthcare, finance, and others, while enhancing their access to current data to address these challenges. Two prominent methods have emerged: fine-tuning and RAG. RAG keeps the model up-to-date without altering its internal parameters by retrieving relevant external knowledge in real time, similar to In-Context Learning (ICL) [4], which boosts responses by adding information to the input rather than adjusting the model's weights. In contrast, fine-tuning involves retraining the model on specialized datasets to enable it to learn domain-specific knowledge.

RAG is more flexible and scalable, but fine-tuning improves accuracy within a specific range. Both approaches offer distinct advantages and trade-offs in terms of computational cost, data requirements, update frequency, and deployment complexity. In this position paper, the aim is to explore which methodologies are most suitable to enhance LLMs in low-resource, high-stakes domains, such as cybersecurity. Specifically, we focus on:

- Detailed analysis of effective fine-tuning techniques for cybersecurity modeling.
- Highlighting effective RAG strategies for cybersecurity modeling.
- Describe the key challenges and risks within the study of fine-tuning and RAG.
- Analyzing the comparative impact of retrieval-based vs. fine-tuning strategies and the hybrid approach.

2. Understanding Fine-Tuning and RAG

Fine-tuning, RAG, and their evolving hybrid approach are key paradigms in the development and use of LLMs. This section examines the background of fine-tuning, RAG, and their hybrid approaches, including the underlying techniques and procedures.

2.1. Fine-Tuning

Fine-tuning is the process of continued training on smaller, more specialized datasets and is used to specialize a pre-trained LLM for domain-specific tasks [5]. It has multiple methodologies, techniques, and applications. This paper discusses the methodologies, techniques, and applications of fine-tuning in cybersecurity.

2.1.1. Learning Paradigms in Fine-Tuning

Supervised Fine-Tuning (SFT): SFT [6] adapts large language models to cybersecurity tasks by training them on labeled datasets that capture real security threats, such as malware samples [7], spam messages [8], network intrusion annotations [9], and others. SFT improves the model's ability to map inputs to correct outputs for tasks such as threat classification and anomaly detection. Figure 1 illustrates this SFT process.

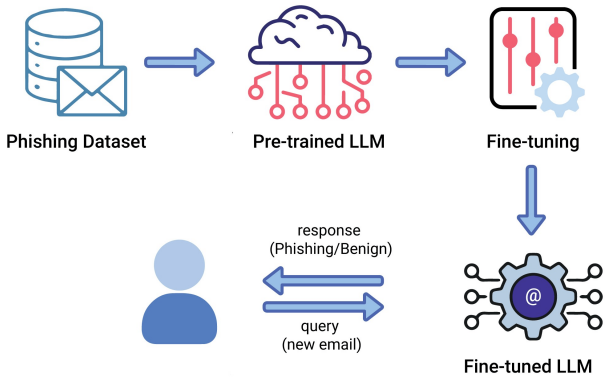

Figure 1: Supervised Fine-tuning in Phishing Email Detection.

Unsupervised Fine-Tuning (UFT): UFT [10] adapts pre-trained LLMs to the cybersecurity domain without labeled data by continuously training on large-scale unlabeled sources. This method leverages self-supervised learning and the underlying structure of the data to enable the model to learn cybersecurity terminology, patterns, and context.

Reinforcement Learning (RL): RL, particularly Reinforcement Learning from Human Feedback (RLHF) [11], offers a new frontier for ongoing improvements to LLMs in cybersecurity through iterative feedback aligned with preferences and operational goals. Different RL algorithms, such as reward modeling, Proximal Policy Optimization (PPO), and Direct Preference Optimization (DPO), enable adaptive learning in dynamic threat environments, supporting more effective security decision making.

2.1.2. Fine-Tuning Techniques for Cybersecurity Tasks

Hyperparameter Tuning: Hyperparameter tuning is an effective technique for fine-tuning LLMs to improve model performance, generalization, and training stability. Selecting optimal hyperparameters, such as batch size, epochs, learning rate, and regularization, can significantly improve the model and prevent overfitting, especially in cybersecurity tasks where data are often noisy or imbalanced.

Parameter-Efficient Fine-Tuning (PEFT): PEFT [12] adapts LLMs to various cybersecurity tasks by updating a small number of model parameters, significantly reduc-

2026, Vol. 3, Article ID. 2026.0031
https://doi.org/10.69709/CAIC.2026.155073

ing computation and memory resources while preserving pre-trained knowledge. Techniques such as Low-Rank Adaptation (LoRA), adapters, and prompt tuning enable rapid model adaptation, which is useful for keeping pace with continuously evolving cyber threats.

Task-Specific Fine-Tuning: Task-specific fine-tuning is the process of adapting a pre-trained language model to a specific cybersecurity task using domain-specific labeled data to enable the model to identify patterns relevant to the task and improve its performance. Although task-specific fine-tuning can result in catastrophic forgetting due to parameter updates, it enables the model to achieve high accuracy in targeted cybersecurity applications.

Transfer Learning: Transfer learning adapts pre-trained LLMs to cybersecurity tasks using limited data, leveraging domain-specific models, such as SecureBERT [13], achieving superior performance with fewer computational resources than out-of-the-box LLMs for malware, phishing, and vulnerability detection.

Few-Shot and Zero-Shot Fine-Tuning: Few-shot and zero-shot fine-tuning approaches leverage the generalization capabilities of large pre-trained models to perform tasks with few or no labeled examples [14]. Few-shot learning enables the model to infer patterns in new data for detection using a small number of labeled examples. In contrast, zero-shot learning relies solely on prompts and prior knowledge, without using task-specific examples.

2.2. Retrieval-Augmented Generation

Retrieval-Augmented Generation (RAG) [15] is an emerging paradigm that augments LLMs with knowledge retrieval and generation capabilities. Although LLMs demonstrate remarkable performance, they face issues such as hallucinations and outdated information, especially with Out-of-Distribution (OoD) questions. The hallucination issue arises when a model produces information that appears to be correct and real but is factually incorrect or fabricated. RAG overcomes these challenges by dynamically retrieving up-to-date, domain-specific information from structured and unstructured sources at inference time [16]. Unlike traditional fine-tuning, which encodes task-specific knowledge into model parameters, RAG—particularly in the cybersecurity domain—retrieves information from external sources such as threat databases, security logs, and technical documentation. This approach ensures adaptability and interpretability while enabling the acquisition of new information without modifying or retaining it within the model parameters.

2.2.1. RAG Core Components & Workflow

Core Components: RAG is made up of multiple components that work together to generate context-aware responses. The first component is the retriever, which identifies and collects relevant documents and knowledge snippets from external sources such as databases, search indexes, or structured knowledge bases, with key resources in cybersecurity, including MITRE ATT&CK [17], Common Vulnerabilities and Exposures (CVE) [18], and National Vulnerability Database (NVD) [19]. To support semantic retrieval, embedding models convert both queries and documents into vector representations, enabling dense retrieval of the most contextually relevant information from a knowledge base. The knowledge base is another component that serves as an external repository of domain-specific information, including texts, databases, APIs, and real-time feeds stored in vector databases. Finally, the generator, typically a language model, produces the final response by combining the original query with the retrieved context to generate accurate answers.

RAG Workflow: The RAG workflow is based on the input query, retrieval, and generation process. The workflow begins with a user query, which is converted into a vector that represents its semantic meaning. This vector is used to retrieve the most relevant document fragments from a vector database for the submitted query. After that, the retrieved document chunks are filtered and ranked to select the best matches to the input query. Finally, a language model processes a prompt formed by combining the retrieved context with the original query, generating a coherent and contextually appropriate response. The system then returns a formatted response to the user.

The RAG workflow can be understood through an example of misleading feedback or fake review detection in Small and Medium Enterprises (SMEs) [20]. These false or misleading reviews are a growing concern in the SME industry [21], and the deployment of RAG can play a significant role in this industry. Using the RAG methodology, SME analysts can identify and detect fake reviews with more accuracy. First, the analyst submits a query, such as a product review with the question, “Is this review fake?”, and the system converts the review text into a semantic vector that captures its meaning. The semantic vector is used to retrieve labeled fake reviews, expert analyses, and the most relevant flagged document chunks from the database. Relevant document chunks, such as reviews that show similar patterns, are retrieved. The retrieved context is then combined with the original query and passed to a generation model, which generates an appropriate response. An overview of the RAG workflow is shown in Figure 2 for fake review detection.

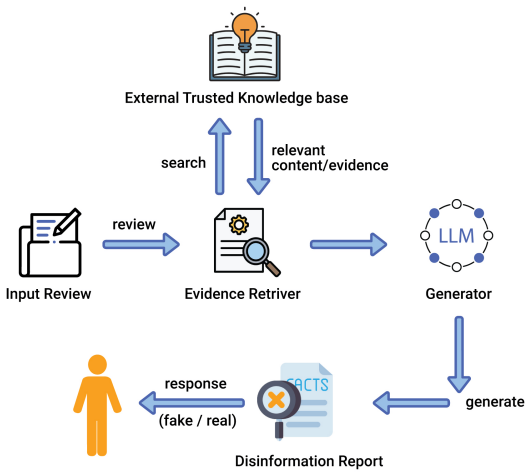

Figure 2: RAG Workflow in Disinformation Detection.

2.2.2. Graph Retrieval-Augmented Generation (GraphRAG)

GraphRAG is a graph-based extension of RAG that integrates Knowledge Graphs (KGs) into the RAG pipeline to improve reasoning, contextual understanding, and explainability in LLMs [22]. The GraphRAG framework is designed to overcome the drawbacks of the traditional RAG system, including a lack of understanding of relationships, failure in multi-hop reasoning, and a tendency to hallucinate. The GraphRAG workflow consists of several key steps. For example, in fake review detection for SMEs, when an analyst submits a query—such as a product review accompanied by the question, "Is this product review fake?"—the system first converts the review text into structured entities (e.g., user, product, review, sentiment) and their relationships. These entities and relationships are stored in a Knowledge Graph that can be constructed from both local data sources, such as private reviews, customer histories, and external data sources (public review repositories and other external sources). When the query is processed, the system performs a graph-based retrieval, rather than relying solely on text similarity. It traverses the knowledge graph to identify relational patterns, such as users posting multiple identical reviews, repeated sentiment expressions across unrelated products, or unnatural temporal correlations. The retrieved subgraph representing the most relevant entities and their relationships is aggregated with the original query and submitted to the language model. Finally, the model generates a response by reasoning over the retrieved graph context, enabling context-aware and explainable detection of fake reviews. The example workflow of GraphRAG is shown in Figure 3.

Compared to other domains, this framework is not widely explored in cybersecurity, but it is becoming increasingly popular. For example, in [23], the authors proposed the CyKG-RAG framework, which integrates KGs with the RAG approach to detect and analyze cybersecurity threats. Similarly, in other research areas, such as network security monitoring and analysis, GraphRAG drives demand for context-aware, adaptive solutions.

2.3. Hybrid Approach

The hybrid approach, which combines fine-tuning and RAG, has gained popularity because it addresses the drawbacks of each technique when used independently. Figure 4 provides a conceptual overview of the hybrid approach combining fine-tuning and RAG.

This approach is beneficial when domain-specific knowledge is required alongside a real-time update process. For instance, in cybersecurity modeling, both deep domain adaptation and real-time responsiveness may be needed, and this hybrid methodology can enable a more robust and adaptive approach to cybersecurity modeling. The hybrid approach has also been successfully applied in various other fields, including healthcare [24], multilingual question answering [25], etc.

The hybrid approach provides several key advantages, including increased robustness, lower retraining costs, enhanced explainability, and greater task flexibility. For instance, spam filtering in email, messaging, and other communication channels requires knowledge of both linguistic patterns and dynamic content structures. The fine-tuned model is optimized to detect linguistic and structural characteristics of spam, including anomalous sender behaviour and the excessive use of specific phrases. By acquiring the latest known spam samples, blacklisted IPs (Internet Protocol addresses) and domains, or publicly available spam databases, the RAG module enhances detection and enables real-time, explainable judgments based on both live data and learned patterns. This hybrid approach increases detection accuracy and facilitates better generalization to multilingual or obfuscated spam tactics and zero-day spam patterns.

3. Applications in Cybersecurity

This section presents several real-world applications within the cybersecurity domain, illustrating the potential of fine-tuning and RAG methods for LLMs to deliver substantial contributions.

Phishing Detection and Analysis: Fine-tuned LLMs effectively detect phishing content by adapting to new phishing patterns via parameter-efficient, few-shot fine-

2026, Vol. 3, Article ID. 2026.0031
https://doi.org/10.69709/CAIC.2026.155073

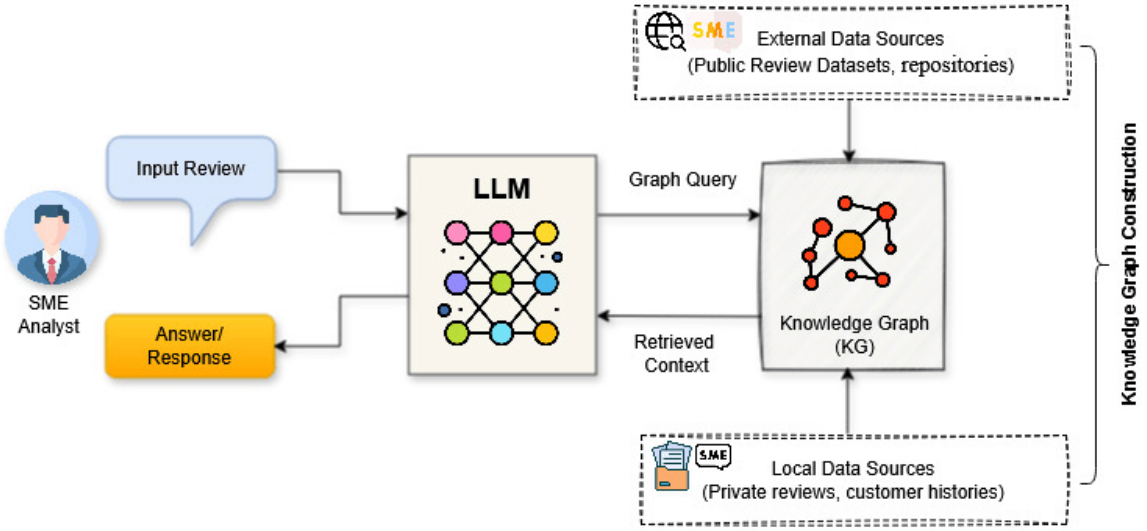

Figure 3: GraphRAG Workflow for Fake Product Review Detection in the context of SMEs.

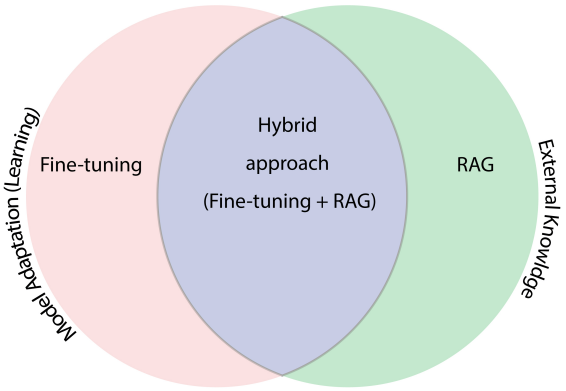

Figure 4: Hybrid Approach.

tuning, while hyperparameter optimization further enhances detection accuracy. Consequently, RAG works in scenarios where models need to consider the most recent phishing samples or attack indicators from external sources, such as threat feeds or phishing databases, to adapt to emerging threats.

Intrusion Detection and Log Analysis: Supervised fine-tuning on datasets such as UNSW-NB15, NSL-KDD, and CIC-IDS enables LLMs to detect intrusion activities in system logs or network traffic. Their contextual understanding aids in analyzing data, detecting rare attack vectors, and linking multi-step intrusions. RAG methods assess recent threats by retrieving up-to-date intelligence to enhance detection speed, accuracy, and organizational insights.

Vulnerability Detection and Analysis: Fine-tuned LLMs improve vulnerability management by summarizing complex CVE data and identifying key threats such as buffer overflows and Remote Code Execution (RCE), thereby aiding prioritization. RAG techniques, such as the knowledge-level framework, retrieve relevant information from a vulnerability knowledge base to detect issues and improve accuracy.

Fraud Detection: Fine-tuned LLMs are ideal for identifying known, stable fraud patterns by incorporating temporal and behavioural features from large volumes of transactional data. RAG-based frameworks are useful for detecting fraud or deception across transactions because they can analyze transactions in real time and retrieve relevant trends, rules, and cases from associated fraud history. This approach performs exceptionally well in few-shot, domain-agnostic settings, enabling effective detection of emerging fraud schemes.

SME Cybersecurity Advisor: SMEs are particularly vulnerable to cyber threats due to limited resources, outdated systems, and a shortage of security experts, leading to financial losses, diminished customer confidence, and operational disruptions. Fine-tuned LLMs can act as cybersecurity advisers by learning from domain policies, processes, and event data, while RAG enhances this capability by retrieving real-time threat intelligence, compliance requirement updates, and best-practice mitigations, enabling cost-effective cybersecurity support without retraining.

4. Challenges and Risks

Even though both fine-tuning and RAG are effective approaches for adapting LLMs to cybersecurity tasks, they present distinct challenges and operational risks. Under-

standing the limitations of each method is essential for selecting or designing effective LLM-based cybersecurity solutions. The key challenges and risks are summarized below:

Data Dependency: Fine-tuning requires a substantial amount of high-quality labeled data, which is often difficult to obtain in specialized domains such as cybersecurity due to security restrictions or other constraints. RAG depends heavily on the quality and reliability of external retrieval data, raising the possibility of including poisoned, out-of-date, or biased data in its responses.

Security and Adversarial Risks: Fine-tuning faces security risks, such as data poisoning and hidden triggers. Attackers can exploit vulnerabilities in the model's decision process. RAG is at risk of attacks on its retrieval system, such as the injection of malicious data or tampering with knowledge bases, which can lead to harmful outputs.

Performance and Forgetting: Fine-tuning offers rapid adaptation but may lead to overfitting or catastrophic forgetting of previously learned tasks when data is limited. RAG is slower due to external database retrieval and relies on database quality, risking performance drops if databases are unavailable or slow.

Maintenance and Resource Intensity: Fine-tuning requires periodic retraining, which is resource-intensive, especially for large models. Organizations must weigh the costs of computing and data annotation against the benefits of improved accuracy. RAG systems require continuous maintenance of retrieval databases and robust access controls to prevent data corruption and ensure data integrity. Scaling the retrieval infrastructure for high query volumes can also be challenging.

Deployment Challenges: Fine-tuning large models for commercial applications necessitates substantial computational resources and stringent data governance. Key challenges include model drift, repeatability, privacy compliance, and deployment complexities such as explainability, version control, and domain bias. The commercial deployment of RAG presents data security and governance risks, particularly when internal data sources are employed. Low latency, scalability, and retrieval accuracy are difficult with large, changing datasets. The hybrid workflow also poses challenges related to data freshness, compliance, and explainability.

5. Our Position and discussion

Based on the comparative analysis of fine-tuning and RAG, in the context of cybersecurity modeling, adapting LLMs requires working with both approaches. Although each method presents distinct advantages, the choice between fine-tuning and RAG primarily depends on the specific use case, available resources, and operational requirements. Moreover, RAG is the preferred option for most enterprise use cases, while fine-tuning serves as the most practical solution for some scenarios. However, this does not mean that RAG is always the right choice over fine-tuning. To determine which method is best for LLMs in cybersecurity modeling, it is essential to understand the different aspects and features.

Knowledge Integration: Fine-tuning absorbs and assimilates domain knowledge during training, resulting in high performance in this domain. However, it is static by nature and quickly becomes outdated without retraining, making it unsuitable for rapidly evolving fields such as cybersecurity. In contrast, RAG maintains up-to-date responses by dynamically accessing external knowledge; however, its reliance on retrieval mechanisms can introduce discrepancies and errors in the outputs.

Adaptability and Flexibility: Fine-tuning adjusts the model's parameters by retraining on domain-specific data, which incurs an additional cost for organizations. RAG offers superior flexibility by retrieving relevant data from an external knowledge base without changing the model's internal parameters, though this increases complexity because it relies on retrieval quality.

Latency: RAG is slightly slower due to the heavy data retrieval and generation phases required to ensure up-to-date data, which takes additional time. Fine-tuning is typically faster than RAG because it can generate answers almost instantly without accessing external data sources.

Scalability: RAG is highly scalable, seamlessly handling large volumes of data by updating the external knowledge base without modifying the model itself. Fine-tuning is less flexible and faces challenges when scaling to evolving domains, as it requires periodic retraining. Adapting to new information requires retraining or additional fine-tuning, thereby limiting responsiveness in fast-evolving domains.

Computational Resources: RAG has lower training costs because it leverages the base or pre-trained language model and does not require any modifications during the training phase. However, computational resources are needed for retrieval during inference. Consequently, due to its resource demands and the need for extensive dataset preparation, fine-tuning during the training phase requires substantial memory, significant time, and multiple high-performance GPUs.

Additionally, the hybrid option can be a suitable solution on some occasions. Hybrid methods leverage the fixed-domain knowledge acquired through fine-tuning while also leveraging the dynamic adaptation offered by RAG, resulting in systems that are both knowledgeable

2026, Vol. 3, Article ID. 2026.0031
https://doi.org/10.69709/CAIC.2026.155073

and context-aware. However, without careful planning, the hybrid system may potentially inherit the drawbacks of both strategies, including potential latency bottlenecks, difficult implementation, and expensive maintenance. A comparison of fine-tuning, RAG, and the Hybrid Approach is presented in Table 1, focusing on various aspects to choose the most suitable one for a particular use case.

Table 1: Comparison of fine-tuning, RAG, and hybrid approach for quick decision-making.

Aspect	Fine-Tuning (FT)	RAG	Hybrid
Core Idea	Update the model's parameters using domain/task-specific data.	Augments generation with retrieved relevant external documents.	FT + RAG: Domain-adapted model + real-time external retrieval.
Knowledge Source	Static	Dynamic	Both
Latency	Low	Medium	high
Transparency	Low	High	Medium
Scalability	Poor	Highly scalable	Good
Handles New Data	Requires retraining.	Instant: can update knowledge sources	Instant (update DB) + FT for deeper tuning
Cost	High	Moderate	High

Additionally, another technique named prompt engineering is also used to design and optimize prompts to effectively guide LLMs, particularly in NLP tasks, to generate desired outcomes or responses [26]. It can be an alternative to fine-tuning or RAG in cybersecurity-based experiments. It is a faster, cost-effective, and lightweight approach compared to fine-tuning and RAG, and it works without retaining the model with minimal computational resources. In the cybersecurity domain, prompt engineering can analyze malicious logs, detect phishing attacks, assess disinformation, and perform vulnerability assessments, among other tasks. Despite its numerous advantages, prompt engineering has certain limitations that must be considered. These include the lack of real-time detection and reliance on pre-trained knowledge, meaning the model can only respond based on previously acquired information.

6. Conclusions

This position paper highlights the role of LLMs in cybersecurity modeling by exploring two prominent approaches: fine-tuning and RAG. This study began by detailing various fine-tuning paradigms and techniques for cybersecurity modeling, highlighting their effectiveness and strengths in adapting LLMs for domain-specific tasks such as detection, classification, and analysis. The study also explored RAG with its core components, working procedure, and a variant. Along with these approaches, a hybrid approach that integrates fine-tuned LLMs with RAG mechanisms was also discussed. The hybrid integration of both paradigms introduces a promising method that combines dynamic reasoning over recently retrieved information with stable core knowledge. Moreover, this study examined the applications, challenges, and risk factors associated with the risk factors of fine-tuning and RAG. Ultimately, the choice between fine-tuning, RAG, or a hybrid approach should be guided by specific cybersecurity use cases, available data, infrastructure constraints, and desired levels of interpretability. We conclude that future cybersecurity systems will benefit most from models that combine the precision of fine-tuning with the adaptability of RAG, providing both strong baseline intelligence and real-time responsiveness to emerging threats.

List of Abbreviations

AI	Artificial Intelligence
CVE	Common Vulnerabilities and Exposures
GPT	Generative Pre-trained Transformer
KGs	Knowledge Graphs
LLM	Large Language Model
LoRA	Low-Rank Adaptation
NLP	Natural Language Processing
NVD	National Vulnerability Database
RAG	Retrieval-Augmented Generation
RLHF	Reinforcement Learning from Human Feedback

Author Contributions

Conceptualization: M.A.U., I.H.S.; writing—original draft preparation: M.A.U.; Visualization: M.A.U.; Writing—review & editing: I.H.S.; Supervision: I.H.S. All authors have read and agreed to the final version of the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

The study did not receive any external funding and was conducted using only institutional resources.

Acknowledgments

Declared none

AI Declaration

AI tool (Grammarly) is used to refine the text, improving the language precision and readability of the paper.

References

[1] I. H. Sarker, "Llm potentiality and awareness: A position paper from the perspective of trustworthy and responsible AI modeling," *Discov. Artif. Intell.*, vol. 4, no. 1, 2024, Art. no. 40. [CrossRef]

[2] L. Hu, Z. Liu, Z. Zhao, L. Hou, L. Nie, and J. Li, "A survey of knowledge enhanced pre-trained language models," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 4, pp. 1413–1430, 2023. [CrossRef]

[3] M. Arslan, S. Munawar, and C. Cruz, "Business insights using RAG–LLMs: A review and case study," *J. Decis. Syst.*, pp. 1–30, 2024. [CrossRef]

[4] T. Zhang et al., "Raft: Adapting language model to domain specific RAG," in *First Conference on Language Modeling*, 2024. Accessed: Jun. 10, 2025. [Online]. Available: https://openreview.net/forum?id=rzQGHXNReU

[5] S. Pratap, A. R. Aranha, D. Kumar, G. Malhotra, A. P. N. Iyer, and S. SS, "The fine art of fine-tuning: A structured review of advanced llm fine-tuning techniques," *Nat. Lang. Process. J.*, vol. 11, 2025, Art. no. 100144. [CrossRef]

[6] A. Pareja et al., "Unveiling the secret recipe: A guide for supervised fine-tuning small LLMs," *arXiv Preprint*, arXiv:2412.13337, 2024. [CrossRef]

[7] T. Carrier, *Detecting Obfuscated Malware Using Memory Feature Engineering*. Fredericton, NB, Canada: University of New Brunswick, 2021. Accessed: Jun. 10, 2025. [Online]. Available: https://unbscholar.lib.unb.ca/handle/1882/37374

[8] T. A. Almeida, J. M. G. Hidalgo, and A. Yamakami, "Contributions to the study of SMS spam filtering: New collection and results," in *Proceedings of the 11th ACM Symposium on Document Engineering*, Mountain View, CA, USA, Sept. 19–22, 2011, pp. 259–262. [CrossRef]

[9] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, Australia, Nov. 10–12, 2015, pp. 1–6. [CrossRef]

[10] K. Tanwisuth, S. Zhang, H. Zheng, P. He, and M. Zhou, "POUF: Prompt-oriented unsupervised fine-tuning for large pre-trained models," In *International Conference on Machine Learning, PMLR*, Honolulu, HI, USA, Jul, 23–29, 2023, pp. 33816–33832. Accessed: Jun. 15, 2025. [Online]. Available: https://proceedings.mlr.press/v202/tanwisuth23a.html

[11] N. Lambert, "Reinforcement learning from human feedback," *arXiv Preprint*, arXiv:2504.12501, 2025. [CrossRef]

[12] N. Houlsby et al., "Parameter-efficient transfer learning for NLP," in *International Conference on Machine Learning, PMLR*, Long Beach, CA, USA, Jun 9–15, 2019, pp. 2790–2799. Accessed: Jun. 10, 2025. [Online]. Available: https://proceedings.mlr.press/v97/houlsby19a.html

[13] E. Aghaei, X. Niu, W. Shadid, and E. Al-Shaer, "SecureBERT: A domain-specific language model for cybersecurity," in *International Conference on Security and Privacy in Communication Systems*. Berlin/Heidelberg, Germany: Springer, 2022, pp. 39–56. [CrossRef]

[14] T. Gao, A. Fisch, and D. Chen, "Making pre-trained language models better few-shot learners," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (volume 1: Long Papers)*, Bangkok, Thailand, August 1–6, 2021, pp. 3816–3830. [CrossRef]

[15] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 9459–9474, 2020. [View Online]

[16] S. Borgeaud et al., "Improving language models by retrieving from trillions of tokens," in *International Conference on Machine Learning, PMLR*, Baltimore, MD, USA, Jul 17–23, 2022, pp. 2206–2240. Accessed: Jun. 10, 2025. [Online]. Available: https://proceedings.mlr.press/v162/borgeaud22a.html

[17] The MITRE Corporation, "MITRE ATT&CK®," 2025. Accessed: Oct. 28, 2025. [Online]. Available: https://attack.mitre.org/

[18] The MITRE Corporation, "Common Vulnerabilities and Exposures (CVE)," 2025. Accessed: Oct. 28, 2025. [Online]. Available: https://cve.mitre.org/

[19] National Institute of Standards and Technology (NIST), "National Vulnerability Database (NVD)," 2025. Accessed: Oct. 28, 2025. [Online]. Available: https://nvd.nist.gov/

[20] I. H. Sarker, H. Janicke, A. Mohsin, and L. Maglaras, "SME-TEAM: Leveraging trust and ethics for secure

and responsible use of AI and LLMs in smes," *arXiv Preprint*, arXiv:2509.10594, 2025. [CrossRef]

[21] R. Das et al., "Towards the development of an explainable e-commerce fake review index: An attribute analytics approach," *Eur. J. Oper. Res.*, vol. 317, no. 2, pp. 382–400, 2024. [CrossRef]

[22] H. Han et al., "Retrieval-augmented generation with graphs (GraphRAG)," *arXiv Preprint*, arXiv:2501.00309, 2024. [CrossRef]

[23] K. Kurniawan, E. Kiesling, and A. Ekelhart, "CyKG-RAG: Towards Knowledge-Graph Enhanced Retrieval Augmented Generation for Cybersecurity," 2024. Accessed: Jun. 10, 2025. [Online]. Available: https://ceur-ws.org/Vol-3950/paper1.pdf

[24] B. Pingua et al., "Medical LLMs: Fine-tuning vs. retrieval-augmented generation," *Bioengineering*, vol. 12, no. 7, 2025, Art. no. 687. [CrossRef]

[25] L. Y. da Costa et al., "Adapting LLMs to new domains: A comparative study of fine-tuning and RAG strategies for portuguese QA tasks," in *Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL)*. Porto Alegre, Brazil: SBC, 2024, pp. 267–277. [CrossRef]

[26] G. Marvin, N. Hellen, D. Jjingo, and J. Nakatumba-Nabende, "Prompt engineering in large language models," in *International Conference on Data Intelligence and Cognitive Informatics*. Singapore: Springer, 2023, pp. 387–402. [CrossRef]