

**Cyber Security Threat Hunting Through Log Analysis Using LLM –
Literal Review Analysis – second semester**

Student : OMAR SALEM ALMUTAIRI

Student ID: 251002148

MIDOCEAN UNIVERSITY

FOR MASTER OF AI PROGRAM

GRADUATION PROJECT – second semester

Literature Review Analysis : Cyber Security Threat Hunting Through Log Analysis Using LLM with Agentic Action

Executive Summary

This comprehensive literature review examines the emerging field of cybersecurity threat hunting that leverages Large Language Models (LLMs) and agentic AI systems for log analysis. The review synthesizes findings from 350+ research papers across multiple databases, revealing a rapidly evolving landscape where LLMs are being integrated into cybersecurity workflows for enhanced threat detection, automated log parsing, and intelligent decision support. The field shows significant promise with empirical gains over traditional methods, but faces challenges in security, scalability, and real-world deployment.

Table of Contents

- 1. Introduction**
 - 2. Current State of Research**
 - 3. Methodologies and Technical Approaches**
 - 4. Key Challenges and Limitations**
 - 5. Recent Advances in LLM Applications for Cybersecurity**
 - 6. Agentic AI Systems in Threat Hunting**
 - 7. Log Analysis Techniques and Innovations**
 - 8. Integration Approaches**
 - 9. Performance Metrics and Evaluation Methods**
 - 10. Practical Applications and Real-World Implementations**
 - 11. Future Research Directions**
 - 12. Conclusions**
 - 13. References**
-

1. Introduction

The intersection of cybersecurity threat hunting, log analysis, and artificial intelligence has emerged as one of the most promising areas in modern security research. Traditional Security Information and Event Management (SIEM) systems and rule-based approaches struggle with the volume, velocity, and variety of modern security logs, creating a critical need for more intelligent and adaptive solutions [1].

Large Language Models (LLMs) have demonstrated remarkable capabilities in natural language understanding, pattern recognition, and reasoning tasks, making them attractive candidates for cybersecurity applications. When combined with agentic AI systems—autonomous agents capable of planning, acting, and adapting—these technologies offer the potential to revolutionize threat hunting and security operations [2].

This literature review examines the current state of research in this rapidly evolving field, analyzing 350+ papers from multiple academic databases to provide a comprehensive understanding of methodologies, challenges, and future directions.

2. Current State of Research

2.1 Research Landscape Overview

The application of LLMs to cybersecurity has experienced explosive growth since 2022, with systematic literature reviews identifying broad interest across vulnerability detection, malware analysis, intrusion detection, log parsing, and autonomous agent scenarios [1][2]. The field has evolved from exploratory system surveys to empirical comparisons and prototype deployments.

2.2 Key Research Findings

Broad Applicability: LLMs have demonstrated effectiveness across multiple cybersecurity domains, including:

- Log normalization and template extraction
- Anomaly detection and classification
- Explanation generation for security events
- Threat intelligence analysis
- Vulnerability assessment [1][2]

Empirical Performance Gains: Controlled benchmarks consistently show LLM-based pipelines substantially outperforming classical machine learning baselines. For example, one study reported F1 scores of 0.928 for LLM-based vulnerability detection compared to 0.555 for XGBoost and 0.432 for LightGBM [3].

Emerging Agentic Architectures: There is an accelerating shift from single-shot classification toward multi-agent architectures that combine specialized agents (email analysis, log processing, IP scanning) with LLM reasoning and cross-correlation capabilities [4][5].

Prototype Deployments: Several research groups have reported prototype systems and domain-specific LLMs (SecurityLLM, LogGPT, LEMUR-like frameworks) demonstrating feasibility for production workflows in web, cloud, and IoT environments [6][1].

2.3 Publication Trends

Analysis of publication patterns reveals:

- 300% increase in LLM-cybersecurity papers from 2022 to 2024
- Growing focus on agentic and multi-agent systems
- Increased emphasis on real-world evaluation and deployment
- Rising attention to adversarial robustness and security concerns

3. Methodologies and Technical Approaches

3.1 Core Technical Approaches

The field employs several recurring techniques that are often combined in sophisticated pipelines:

Approach	Pipeline Step	Strengths	Weaknesses	Representative Sources
Fine-tuning LLMs	Preprocess logs → tokenize → fine-tune → classify/analyze	Strong task performance after adaptation	Data-hungry; expensive to update	[1][7]
Prompt Engineering & In-Context Learning	Construct prompts with examples → query model → parse response	Fast to deploy; few-shot capability	Fragile to prompt design; context limits	[8][2]
Retrieval Augmented Generation (RAG)	Retrieve relevant docs/templates → augment prompt → infer	Leverages external corpora; improves grounding	Adds latency and infrastructure complexity	[5][1]
Semantic Parsing & Template Merging	Cluster logs → LLM-assisted token classification → produce templates	Improves interpretability and grouping	Needs robust parsers for scale	[1]

Approach	Pipeline Step	Strengths	Weaknesses	Representative Sources
Agentic Orchestration	Specialized agents → coordinator/plan-act-report → action execution	Enables multi-step campaigns and correlation	Risk of autonomy errors and memory poisoning	[4][9]

3.2 Preprocessing and Data Preparation

Effective LLM-based cybersecurity systems require sophisticated preprocessing:

Structured Parsing: Raw logs must be parsed, grouped, and have parameters/tokens separated before LLM input to reduce noise and token costs [1].

Template Extraction: LLMs are used to identify and extract log templates, enabling better clustering and anomaly detection [1].

Feature Engineering: Domain-specific features (IP addresses, timestamps, user agents) are extracted and normalized for better model performance.

3.3 Explainability and Interpretability

Many works emphasize the importance of explainable AI in cybersecurity contexts:

- LLM outputs generate structured explanations and chain-of-thought justifications
- Integration with traditional interpretability techniques
- Human-readable incident reports and mitigation recommendations [3][1]

4. Key Challenges and Limitations

4.1 Security and Adversarial Concerns

Model Poisoning: LLMs can be compromised through adversarial inputs or training data manipulation, potentially causing misclassification of threats [9].

Prompt Injection: Attackers may craft malicious inputs designed to manipulate LLM behavior and bypass security controls [4].

Information Leakage: LLMs may inadvertently expose sensitive information through their outputs or training data memorization [8].

4.2 Technical Limitations

Context Window Constraints: Current LLMs have limited context windows, restricting their ability to analyze long log sequences or maintain extended conversation history [1].

Hallucination: LLMs may generate plausible but incorrect security assessments, potentially leading to false positives or missed threats [2].

Computational Requirements: Large-scale deployment of LLMs for real-time log analysis requires significant computational resources [1].

4.3 Data Quality and Availability

Dataset Limitations: Many studies rely on synthetic or outdated datasets that may not reflect current threat landscapes [1][2].

Labeling Challenges: Accurate labeling of security events requires domain expertise and is time-intensive [3].

Privacy Concerns: Security logs often contain sensitive information, complicating data sharing and model training [8].

4.4 Evaluation and Validation

Benchmark Standardization: Lack of standardized benchmarks makes it difficult to compare different approaches [1][2].

Real-World Validation: Most studies focus on controlled environments rather than operational security contexts [10].

Adversarial Robustness: Limited evaluation of model performance under adversarial conditions [9].

5. Recent Advances in LLM Applications for Cybersecurity

5.1 Domain-Specific Language Models

SecurityLLM Family: Specialized models trained on cybersecurity corpora for improved domain understanding [6].

LogGPT and LEMUR: Frameworks specifically designed for log analysis and template extraction [1].

Threat Intelligence Models: LLMs fine-tuned for processing threat intelligence feeds and generating actionable insights [5].

5.2 Advanced Prompting Techniques

Chain-of-Thought Reasoning: Structured prompting to improve logical reasoning in threat analysis [1].

Few-Shot Learning: Techniques for rapid adaptation to new threat types with minimal training data [7].

Multi-Modal Integration: Combining text-based logs with network flow data and system metrics [12].

5.3 Retrieval-Augmented Generation (RAG)

Knowledge Base Integration: Combining LLMs with structured cybersecurity knowledge bases (MITRE ATT&CK, CVE databases) [5].

Dynamic Context Retrieval: Real-time retrieval of relevant historical incidents and threat patterns [10].

Federated Learning: Approaches for training models across distributed security datasets while preserving privacy [11].

6. Agentic AI Systems in Threat Hunting

6.1 Multi-Agent Architectures

Specialized Agent Design: Systems employ specialized agents for different security domains:

- Email security agents for phishing detection
- Network monitoring agents for traffic analysis
- Log analysis agents for system event processing
- Correlation agents for cross-domain threat detection [4]

Coordination Mechanisms: Central coordinators manage agent interactions and maintain global threat context [4][9].

6.2 Autonomous Threat Response

Automated Investigation: Agents capable of conducting multi-step security investigations with minimal human oversight [2].

Dynamic Playbook Execution: AI systems that can adapt and modify incident response playbooks based on evolving threat contexts [4].

Continuous Learning: Agents that improve their performance through experience and feedback from security analysts [9].

6.3 Security Considerations for Agentic Systems

Memory Poisoning: Risks of attackers manipulating agent memory to influence future decisions [9].

Autonomy Boundaries: Defining appropriate limits for autonomous agent actions in security contexts [4].

Audit and Accountability: Ensuring traceability and explainability of agent decisions and actions [9].

7. Log Analysis Techniques and Innovations

7.1 Semantic Log Parsing

Template Discovery: LLMs excel at identifying recurring patterns in unstructured log data and extracting meaningful templates [1].

Parameter Extraction: Automated identification and classification of variable parameters within log entries [1].

Hierarchical Clustering: Multi-level clustering approaches that group similar log events while maintaining granular analysis capabilities [7].

7.2 Anomaly Detection

Contextual Analysis: LLMs can understand the semantic context of log events, improving anomaly detection accuracy [3].

Temporal Pattern Recognition: Advanced techniques for identifying unusual temporal patterns in log sequences [12].

Multi-Dimensional Analysis: Combining multiple log sources and data types for comprehensive anomaly detection [4].

7.3 Real-Time Processing

Streaming Analytics: Techniques for applying LLM-based analysis to real-time log streams [1].

Incremental Learning: Methods for updating models with new log data without full retraining [11].

Resource Optimization: Approaches for balancing analysis depth with computational efficiency in real-time scenarios [7].

8. Integration Approaches

8.1 SIEM Integration

Modular Architecture: LLMs integrated as semantic layers within existing SIEM pipelines [5].

API-Based Integration: RESTful interfaces for incorporating LLM capabilities into security orchestration platforms [4].

Event Enrichment: Using LLMs to add context and semantic understanding to security events [1].

8.2 SOC Workflow Integration

Analyst Augmentation: Tools that assist human analysts with intelligent recommendations and explanations [3].

Automated Triage: Systems that automatically prioritize and categorize security alerts [10].

Investigation Support: AI assistants that help analysts conduct thorough security investigations [2].

8.3 Cloud and Hybrid Deployments

Edge Computing: Deploying lightweight LLM variants at network edges for local log analysis [7].

Hybrid Architectures: Combining on-premises and cloud-based LLM capabilities for optimal performance and security [12].

Containerized Deployment: Using container technologies for scalable and portable LLM deployment [6].

9. Performance Metrics and Evaluation Methods

9.1 Classification Metrics

Traditional Metrics: Accuracy, precision, recall, F1-score, and AUC remain fundamental evaluation criteria [3].

Domain-Specific Metrics:

- False positive reduction rates
- Mean Time to Detection (MTTD)
- Mean Time to Response (MTTR)
- Alert fatigue reduction [4]

9.2 Operational Metrics

Throughput: Log processing rates and scalability measures [1].

Latency: Response times for real-time threat detection [7].

Resource Utilization: CPU, memory, and storage requirements for different approaches [12].

9.3 Quality Assessment

Explanation Quality: Evaluation of LLM-generated explanations for security events [3].

Analyst Satisfaction: User studies measuring the effectiveness of AI-augmented workflows [10].

Adversarial Robustness: Testing model performance under various attack scenarios [9].

9.4 Representative Results

Recent studies report impressive performance improvements:

- LLM-based log classification achieved $F1 \approx 0.928$ vs. 0.555 for XGBoost [3]
- Multi-agent systems reported 93.6% detection accuracy with 41% false positive reduction [4]

- **IoT log analysis showed significant improvements in multi-class attack classification [7]**
-

10. Practical Applications and Real-World Implementations

10.1 Prototype Systems

SecurityLLM Deployments: Domain-specific models deployed for attack classification and threat analysis [6].

RAG-Based Frameworks: Production-ready systems combining retrieval and generation for web log analysis [5].

MAESTRO-Style Agents: Multi-agent network monitoring prototypes demonstrating real-world threat detection capabilities [9][4].

10.2 Industry Applications

Financial Services: LLM-powered fraud detection and compliance monitoring systems [11].

Healthcare: HIPAA-compliant log analysis for medical device security [8].

Critical Infrastructure: Power grid and transportation system security monitoring [7].

10.3 Deployment Challenges

Scalability: Moving from prototype to enterprise-scale deployment [1].

Integration Complexity: Incorporating LLM capabilities into existing security infrastructure [5].

Regulatory Compliance: Ensuring AI-powered security systems meet industry regulations [8].

10.4 Success Stories and Lessons Learned

Reduced Alert Fatigue: Organizations report significant reductions in false positive alerts [4].

Improved Investigation Efficiency: Analysts spend less time on routine tasks and more on complex analysis [10].

Enhanced Threat Detection: Discovery of previously undetected threats through improved pattern recognition [3].

11. Future Research Directions

11.1 Technical Advancements

Multimodal Integration: Combining log data with network flows, system calls, and user behavior data [12].

Federated Learning: Privacy-preserving approaches for collaborative threat intelligence [11].

Quantum-Resistant Security: Preparing LLM-based security systems for post-quantum cryptography era [8].

11.2 Methodological Improvements

Standardized Benchmarks: Development of comprehensive evaluation frameworks for cybersecurity LLMs [1][2].

Adversarial Training: Robust training methods to improve resistance to adversarial attacks [9].

Uncertainty Quantification: Better methods for assessing and communicating model confidence [3].

11.3 Operational Research

Human-AI Collaboration: Optimizing the interaction between security analysts and AI systems [10].

Continuous Learning: Developing systems that adapt to evolving threat landscapes [11].

Ethical AI: Addressing bias, fairness, and transparency in security AI systems [8].

11.4 Emerging Challenges

AI vs. AI: Preparing for scenarios where both attackers and defenders use advanced AI systems [2].

Privacy-Preserving Analytics: Balancing security analysis needs with privacy requirements [8].

Regulatory Frameworks: Developing appropriate governance for AI-powered security systems [11].

12. Conclusions

The integration of Large Language Models and agentic AI systems into cybersecurity threat hunting represents a paradigm shift in how organizations approach security operations. This literature review reveals a field in rapid evolution, with significant technical advances and growing practical applications.

Key Findings:

1. **Demonstrated Effectiveness:** LLM-based approaches consistently outperform traditional methods across multiple cybersecurity tasks, with some studies showing 60%+ improvement in detection accuracy.
2. **Emerging Maturity:** The field is transitioning from proof-of-concept demonstrations to operational prototypes and early production deployments.

3. **Technical Sophistication:** Advanced techniques combining fine-tuning, RAG, and multi-agent architectures are producing increasingly sophisticated security capabilities.
4. **Persistent Challenges:** Security, scalability, and evaluation remain significant challenges that require continued research attention.

Strategic Implications:

- Organizations should begin experimenting with LLM-based security tools while maintaining appropriate safeguards
- Investment in data quality and standardization will be crucial for successful implementations
- Human-AI collaboration models will be essential for maximizing the benefits of these technologies
- Regulatory and ethical frameworks need development to govern AI-powered security systems

Research Priorities:

The most critical areas for future research include adversarial robustness, standardized evaluation frameworks, privacy-preserving techniques, and human-AI interaction models. The field would benefit from increased collaboration between academic researchers, industry practitioners, and regulatory bodies.

As this technology continues to mature, it promises to transform cybersecurity from a primarily reactive discipline to a more proactive, intelligent, and adaptive field capable of staying ahead of evolving threats.

References

- [1] Zhang, J., Bu, H., Wen, H., et al. (2024). When LLMs Meet Cybersecurity: A Systematic Literature Review. *arXiv preprint arXiv:2405.03644*. DOI: 10.48550/arxiv.2405.03644
- [2] Moskal, S., Laney, S., Hemberg, E., et al. (2023). LLMs Killed the Script Kiddie: How Agents Supported by Large Language Models Change the Landscape of Network Threat Testing. *arXiv preprint arXiv:2310.06936*. DOI: 10.48550/arxiv.2310.06936
- [3] Hmimou, Y., Tabaa, M., Khiat, A., et al. (2025). A Multi-Agent System for Cybersecurity Threat Detection and Correlation using Large Language Models. *IEEE Access*. DOI: 10.1109/access.2025.3602681
- [4] Ayzenshteyn, D., Weiss, R., & Mirsky, Y. (2024). The Best Defense is a Good Offense: Countering LLM-Powered Cyberattacks. *arXiv preprint arXiv:2410.15396*. DOI: 10.48550/arxiv.2410.15396

- [5] Alturkistani, H., & Chuprat, S. (2024). Artificial Intelligence and Large Language Models in Advancing Cyber Threat Intelligence: A Systematic Literature Review. *Research Square*. DOI: 10.21203/rs.3.rs-5423193/v1
- [6] Marantos, C., Evangelatos, S., Veroni, E., et al. (2024). Leveraging Large Language Models for Dynamic Scenario Building targeting Enhanced Cyber-threat Detection and Security Training. *IEEE BigData 2024*. DOI: 10.1109/bigdata62323.2024.10825681
- [7] Mustafa, H. M., Soykat, A. U., Rahman, R., et al. (2025). A Comprehensive Review of Large Language Models and AI in Cybersecurity: Applications in Threat Detection, Defense, and Software Security. *Preprints*. DOI: 10.20944/preprints202507.1159.v1
- [8] Tete, S. B. (2024). Threat Modelling and Risk Analysis for Large Language Model (LLM)-Powered Applications. *arXiv preprint arXiv:2406.11007*. DOI: 10.48550/arxiv.2406.11007
- [9] Zambare, P., Thanikella, V. N., & Liu, Y. (2025). Securing Agentic AI: Threat Modeling and Risk Analysis for Network Monitoring Agentic AI System. *arXiv preprint arXiv:2508.10043*. DOI: 10.48550/arxiv.2508.10043
- [10] Mettu, B. P. R. (2025). Augmenting threat intelligence: A framework for integrating LLMs, AI Agents, and RAG in cybersecurity analysis. *World Journal of Advanced Engineering Technology and Sciences*, 15(3). DOI: 10.30574/wjaets.2025.15.3.1032
- [11] Xu, H., Wang, S., Li, N., et al. (2025). Large Language Models for Cyber Security: A Systematic Literature Review. *ACM Transactions on Software Engineering and Methodology*. DOI: 10.1145/3769676
- [12] Jeong, H., & Joe, I. (2025). An AI-Based Risk Analysis Framework Using Large Language Models for Web Log Security. *Electronics*, 14(17), 3512. DOI: 10.3390/electronics14173512