

**Cyber Security Threat Hunting Through Log Analysis Using LLM with Agentic
Action – Article**

Student : OMAR SALEM ALMUTAIRI

Student ID: 251002148

MIDOCEAN UNIVERSITY

FOR MASTER OF AI PROGRAM

GRADUATION PROJECT 1

Abstract

This literature review provides a comprehensive analysis of the application of Large Language Models (LLMs) in cybersecurity threat hunting, focusing specifically on log analysis. LLMs represent a paradigm shift from traditional, rule-based security systems to more flexible, behavior-based anomaly detection. Current research highlights the potential of LLMs to automatically parse, interpret, and identify subtle threats within vast, unstructured log data, with several frameworks demonstrating superior performance metrics over existing methods. However, the practical deployment of these models is hindered by significant challenges. These include operational hurdles like scalability, high false positive rates, and integration complexity. More fundamentally, LLMs introduce new security risks, including vulnerabilities to adversarial attacks (e.g., prompt injection, data poisoning), data privacy concerns, and the need for real-time performance. The inherent "black box" nature of these models also underscores a critical need for Explainable AI (XAI). This review concludes that the responsible and effective integration of LLMs into threat hunting requires a strategic focus on developing domain-specific architectures, establishing robust benchmarking, and prioritizing Human-in-the-Loop (HITL) systems to ensure trust, reliability, and security in real-world operations.

Introduction

The modern cyber threat landscape is defined by the relentless evolution of attack methods and the sheer volume of malicious activity. This dynamic environment renders traditional, reactive security measures insufficient, necessitating a proactive approach like threat hunting. Threat hunting actively searches for indicators of compromise that have bypassed initial defenses, aiming to neutralize threats before significant damage occurs. A cornerstone of this process has always been log analysis, but conventional methods are struggling to keep up. These techniques typically depend on pre-defined rules and manual review, making them time-consuming and ineffective against novel or polymorphic attacks that don't match known signatures.

Large Language Models (LLMs) are poised to revolutionize log analysis for threat hunting. With an unparalleled ability to process and find meaning in massive amounts of unstructured text, LLMs can discern subtle patterns and anomalies that escape traditional tools. Their natural language processing (NLP) capabilities allow them to understand the semantic context within chaotic log entries from diverse system, network, and application sources.

The shift to LLM-based analysis is more than just an incremental upgrade; it's a fundamental change in security philosophy. It marks a transition from a rigid, signature-driven posture to a more adaptive, probabilistic one. This allows for the detection of zero-day exploits and sophisticated evasive maneuvers that defy pre-defined rules. Consequently, security professionals must adapt, moving from writing explicit detection rules to fine-tuning, prompting, and interpreting the nuanced outputs of AI models. This review synthesizes the current state of LLM

applications in log analysis, examines the core methodologies and their performance, critically assesses the formidable challenges, and proposes strategic recommendations for future research and deployment.

Related Work

The application of LLMs to log analysis for threat hunting is an emerging field, with several innovative frameworks demonstrating significant potential. The research can be broadly categorized by its primary technical approach, including anomaly detection, log parsing and enrichment, and multi-agent collaboration.

Anomaly Detection Frameworks

The core task in threat hunting is often identifying anomalous behavior. Recent frameworks leverage LLMs' semantic understanding to improve this process significantly.

- **LogLLM** uses a hybrid approach, employing a pre-trained BERT model to extract semantic vectors from log messages and a Llama model to classify log sequences. This method has been shown to outperform traditional deep learning techniques in capturing the meaning embedded in logs, leading to superior anomaly detection.
- **LogRAG** is a semi-supervised framework based on Retrieval-Augmented Generation (RAG). It performs detection in phases using both log tokens and templates, which helps mitigate parsing errors. By using a single-class classifier to model normal system behavior, it effectively reduces false positives, achieving a 15% F1-score improvement on the BGL dataset and a 60% improvement on the Spirit dataset compared to previous leading algorithms.

Log Parsing and Data Enrichment

Preparing raw log data for analysis is a critical first step where LLMs offer substantial advantages over traditional REGEX-based parsers.

- **LLMeLog** is a novel approach that uses an LLM to enrich log events via in-context learning. It then fine-tunes a BERT model with this enriched data and trains a transformer-based anomaly detection model. This method proved highly efficient, achieving F1-scores exceeding 99% on public datasets while using only 10% of the labeled data for training.
- A hybrid approach (referred to as **AuditLLM-Hybrid** in some studies) uses an LLM to generate REGEX for accurate parsing while using another pre-trained LLM for event embedding. Combined with an autoencoder featuring an attention mechanism, this unsupervised framework achieved 96% accuracy on a public anomaly detection dataset.

Advanced Methodologies: Multi-Agent and System State Analysis

Beyond simple classification, researchers are exploring more sophisticated applications of LLMs for deeper analysis and reasoning.

- **Audit-LLM is a multi-agent framework designed specifically for detecting insider threats from logs. It features three collaborative agents: a Decomposer that breaks down tasks, a Tool Builder that creates reusable functions, and an Executor that generates the final conclusion. A key innovation is its Evidence-based Multi-agent Debate (EMAD) mechanism, which refines conclusions and significantly improves the faithfulness of the LLM's explanations, directly addressing the critical need for XAI.**
- **The From Log to Fingerprint concept represents a paradigm shift away from direct log analysis. It uses LLMs and multidimensional projection to transform raw logs and system metrics into a compact, information-rich "fingerprint" of the system's state at a given moment. By comparing these fingerprints over time, analysts can instantly detect subtle changes and anomalies without manually reviewing gigabytes of data. This enables real-time tracking of system drift and greatly accelerates debugging and troubleshooting.**