ICT1099CS01 - LLM-Augmented Modular Investigation Workflows

Principal Investigator

Dr. Shahnawaz Khan	
Polytechnic ID: 12010895	**Department/School**: ICT

Programme	ICT1099 PhD in Computer Science
Research Area and Specialization:	Artificial Intelligence for Cybersecurity, Large Language Models (LLMs), graph-driven modular investigation workflows, Human-in-the-Loop Automated Threat Triage
Company Endorsement	Beyon Cyber
Industry mentor	Mr. Britto Irudhayasamy

Background of the problem

Security Operations Centers (SOCs) suffer alert overload. Existing automated triage systems either run monolithic playbooks (a single, large, automated investigation or response script) for every alert or rely on heavyweight provenance systems, creating inefficiency and scale problems. This project proposes restructuring investigations into modular, reusable blocks, where each block represents specific investigative objectives such as verifying user identity, checking source IP reputation, assessing privilege misuse, or detecting lateral movement. Instead of running one long playbook for every alert, we assemble only the required blocks based on the alert's context.

Each alert maps to a graph of investigation blocks, creating a scenario-specific workflow. For example, a suspicious login on a Domain Controller triggers blocks related to privileged access and lateral movement analysis and a suspicious login on a normal user workstation triggers only basic identity and authentication checks. This approach ensures investigations are aligned with actual alert context, unnecessary steps are eliminated, workflows remain consistent across similar scenarios and new blocks can be added without redesigning entire playbooks

A local LLM assists with query generation and evidence summaries but cannot modify the investigation graph, keeping the workflow deterministic and controlled. In summary, proposed solution will provide a flexible, scalable, and structured investigation model where each alert receives the exact steps it needs.

List of publications related to the field

- Dehing, Y., Henseler, H., Meconi, T., Worring, M., & van Beek, H. (2026, March). Structured Report Generation using Local LLMs for Chat-Based Digital Forensics. In Proceedings of the Digital Forensics Doctoral Symposium (pp. 1-11).
- Sharma, B., Ghawaly, J., McCleary, K., Webb, A. M., & Baggili, I. (2025). ForensicLLM: A local large language model for digital forensics. Forensic Science International: Digital Investigation, 52, 301872.
- Jeong, S., Lee, S., & Park, J. (2025). LangurTrace: Forensic analysis of local LLM applications. Forensic Science International: Digital Investigation, 54, 301987.
- Michelet, G., & Breitinger, F. (2024). ChatGPT, Llama, can you write my report? An experiment on assisted digital forensics reports written using (local) large language models. Forensic Science International: Digital Investigation, 48, 301683.
- Liu, Y., Shu, X., Sun, Y., Jang, J., & Mittal, P. (2022, December). Rapid: real-time alert investigation with context-aware prioritization for efficient threat discovery. In Proceedings of the 38th annual computer security applications conference (pp. 827-840).
- Khan, S., Yusuf, A., Haider, M., Thirunavukkarasu, K., Nand, P., & Rahmani, M. K. I. (2022, June). A review of android and ios operating system security. In 2022 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS) (pp. 67-72). IEEE.
- Sachdeva, S., Ali, A., & Khan, S. (2022). Secure and privacy issues in telemedicine: Issues, solutions, and standards. In Telemedicine: The Computer Transformation of Healthcare (pp. 321-331). Cham: Springer International Publishing.
- Alrubaiei, M. H., Al-Saadi, M. H., Shaker, H., Sharef, B., & Khan, S. (2022). Internet of things in cyber security scope. In Blockchain Technology and Computational Excellence for Society 5.0 (pp. 146-187). IGI Global Scientific Publishing.

Research Objectives and Expected Outcomes:

Aim

The aim is to develop a context-aware framework that reduces non-determinism in large language model (LLM)-assisted incident response by employing deterministic retrieval, cybersecurity knowledge graphs, and structured contextual reasoning.

Objectives:

- To design and evaluate a deterministic, modular investigation framework that maps incoming alerts to minimal, scenario-specific graphs of reusable investigation blocks, integrating a local LLM only for safe, auditable query generation and evidence summarization, and show that this approach improves triage efficiency and analyst

effectiveness compared to monolithic playbooks and current provenance/triage systems.

Expected Outcomes

- Define formal ontology and semantics for investigation blocks (inputs, outputs, preconditions, postconditions, cost/priority metadata) and for graph composition rules.
- Build an orchestration engine and block repository that: constructs graphs from alert context, schedules block execution, enforces determinism, and instruments complete audit trails.
- Integrate a local LLM module limited to generating data-source queries and producing evidence summaries; define safety controls, prompt templates, and verification steps to prevent LLM-driven workflow changes.
- Implement example blocks for identity verification, IP reputation lookups, privilege misuse assessment, lateral movement detection, and evidence consolidation.
- Evaluate the system on realistic datasets and SOC workflows to measure improvements in investigation time, resource usage, analyst workload, and detection/false-positive rates versus baselines (monolithic playbooks, RAPID-like provenance approaches).

Research Design/Methodology

- Derive a formal model for blocks and graphs: define block interface (required inputs, produced evidence types), deterministic composition rules, and a cost/priority model for scheduling. Ground design in literature on provenance-aware triage and modular automation.
- Build a prototype orchestration engine and a block repository with versioning and metadata; implement a policy/mapping module that converts alert context (alert fields, asset criticality, recent telemetry) to a directed graph of blocks. Use an on-premises local LLM (open-source) with strict API boundaries; the orchestration engine will record LLM inputs/outputs for audit.
- Implement representative investigation blocks: identity/auth checks, IP and domain reputation, process provenance queries for lateral movement (e.g., host-to-host process hops), and privilege-escalation checks using endpoint and AD logs. Connectors will use common telemetry sources (e.g., syslog, EDR, AD logs, SIEM).
- Define prompt templates, verification checks, and rejection policies to ensure the LLM only returns queries or summaries; implement sanity-checking of LLM-generated queries against known-safe query grammar; log all LLM activity for forensic reproducibility.
- Experimental setup: compare three configurations (a) monolithic playbooks (typical SOAR), (b) provenance-aware system similar to RAPID, (c) proposed modular graph-driven system with LLM assistance. Use public datasets (e.g., DARPA Transparent Computing traces used

in RAPID evaluation; then test on the live data available from the industry if possible) and curated SOC alert streams; metrics include mean time per investigation, number of executed steps per alert, analyst time measured in human-in-the-loop experiments, detection/false-positive rates, and resource utilization. Statistical tests will assess significance.

- Conduct user studies with SOC analysts to gather qualitative feedback on workflow clarity, trust in LLM summary outputs, and maintainability of block repositories. Perform case studies on complex multi-alert campaigns to show how graphs compose and scale.

Significance and potential impact of research in Bahrain

Significance

- Operational efficiency: By executing only scenario-relevant blocks, SOCs can reduce wasted computation and analyst time, address alert fatigue and improve time-to-detect/contain.
- Maintainability and extensibility: Reusable blocks and graph composition allow new investigative capabilities to be added without redesigning entire playbooks, lowering operational overhead and increasing agility.
- Safety and auditability: Restricting LLMs to query generation and summarization preserves deterministic workflow control and supports reproducible, auditable investigations, reducing risks from model hallucination or unauthorized automation.

Potential Impact

The project contributes a formal model, orchestration algorithms, and empirical results to the areas of automated triage, SOAR research, and trustworthy LLM-assisted security operations.

Alignment with Bahrain's national research strategy

Contribution to Bahrain Society: Enhances Bahrain's cybersecurity readiness by enabling faster, more accurate, and more efficient investigation of suspicious digital activity.

Alignment with Industry: The project addresses industry needs for scalable, maintainable, and auditable security operations. Its modular investigation framework can be applied in sectors such as banking, telecommunications, healthcare, and enterprise IT.

National Priorities: Supports Bahrain's digital transformation, cybersecurity enhancement, innovation, and human capital development.

Benefits for Industry: Stronger cyber resilience, faster incident investigation, and lower operational cost.