AN AGENTIC AI-DRIVEN FRAMEWORK FOR CLOUD LOG SECURITY AND MISCONFIGURATION MITIGATION IN MULTI-CLOUD ENVIRONMENTS

Tran Thanh Truc[1]

[1] University of Information Technology
HCMC, Vietnam

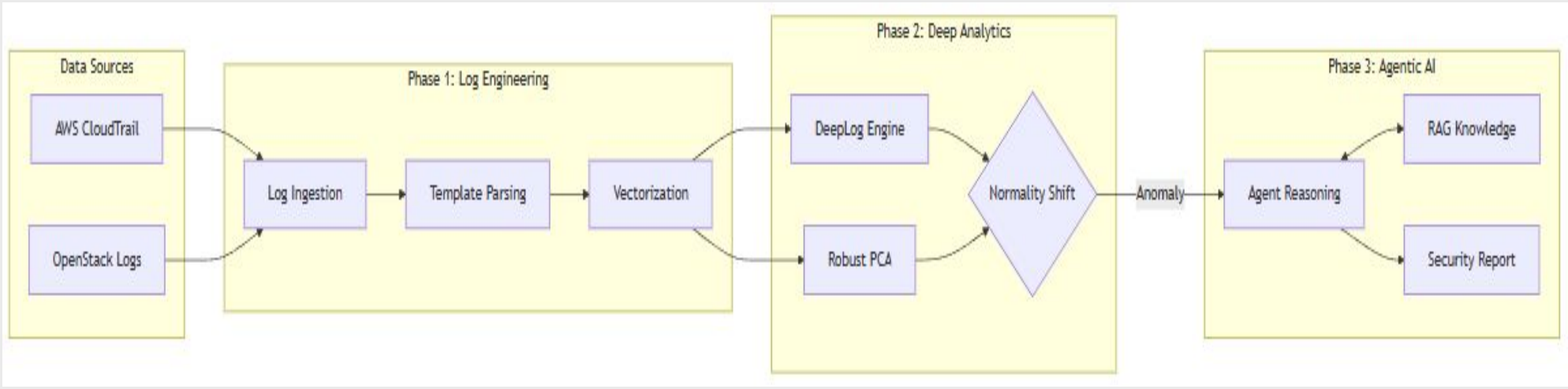
What ?

- We introduce a framework to detect and mitigate cloud misconfigurations using Agentic AI and Deep Learning.
- Built a comprehensive security knowledge base using state-of-the-art datasets: **CAShift (2025)**, **AWSCTD**, and **Loghub**.
- Evaluated an AI-driven "SOC Analyst Assistant" to reduce false positives and provide automated remediation steps.

Why ?

- Cloud misconfigurations are the primary cause of data breaches, yet remain difficult to manage in complex Multi-Cloud environments.
- Current SIEM systems struggle with massive log volumes and the evolving nature of cloud attacks, known as **Normality Shift**.
- Most studies have focused on static detection rather than **automated reasoning** and real-time response in production environments.

Overview



Description

1. Log Anomaly Detection & Normalization

- The system utilizes the **DeepLog** architecture, employing Long Short-Term Memory (LSTM) or Transformer networks to model system logs as natural language sequences.
- By learning the "language" of normal execution patterns from datasets like **AWSCTD** and **Loghub**, the engine predicts the next likely log entry in a sequence.
- Log templates are extracted via automated parsing to handle unstructured data, where any entry falling outside the high-probability prediction threshold is flagged as a potential security breach.

2. Robustness & Normality Shift Adaptation

- To address the evolving nature of cloud environments, we implement strategies from CAShift to benchmark detection performance under "Normality Shift," ensuring models remain accurate as infrastructure updates occur.
- A Robust Principal Component Analysis (RPCA) model is integrated to project high-dimensional OpenStack log data into a clean column space, effectively separating malicious anomalies from background system noise.
- This dual-layered approach allows the framework to maintain high precision even when the definition of "normal" behavior changes over time.

3. Agentic AI Reasoning & Remediation

- Detected anomalies are passed to an **Agentic AI** layer that uses **Retrieval-Augmented Generation (RAG)** to correlate log events with external security knowledge bases like OWASP or CSA.
- The agent performs autonomous reasoning to interpret the root cause of an alert, such as an overly permissive IAM policy or an exposed S3 bucket.
- Final outputs include a natural language summary for SOC analysts and actionable mitigation scripts (e.g., Terraform rollback or CLI commands) to close the security gap immediately.

NII
Trần Thanh Trực – Trường Đại học Công nghệ Thông tin
TEL : 0867852509 Email : tructt.20@grad.uit.edu.vn