

AI-Driven Detection of Malicious Codes when Integrating Data from Monitoring Technologies

Serrhini Mohamed

University Mohamed First Oujda Morocco



INTRODUCTION AND MAIN RESULTS

Integrating data from various monitoring technologies has become essential to detect and mitigate the spread of malicious code in modern digital ecosystems. However, the heterogeneity and complexity of data sources present significant challenges to ensure smooth and accurate detection. This paper presents an AI-based framework for detecting malicious code that leverages a deep learning approach to analyse and correlate data from various monitoring technologies, including network traffic analyzers, endpoint detection and application logs. The proposed system uses advanced feature extraction and fusion methods to unify disparate data streams into a coherent data set, enabling the identification of complex attack patterns that are often missed by traditional detection methods. Experimental evaluations demonstrate the framework's ability to improve detection accuracy, reduce false positives and adapt to evolving threats in real time. This work highlights the potential of AI to improve cybersecurity by providing a robust and scalable approach to detecting malicious code on multi-source monitoring platforms.



The Challenge: Data Heterogeneity in Modern Ecosystems

The Problem:

Modern digital environments generate vast amounts of data from disparate sources:

Network Traffic Analyzers (Flow data, packet capture)

Endpoint Detection and Response (EDR) (Process, file, registry events)

Application Logs (User activity, error messages)

Key Challenges:

✗ **Data Silos:** Correlating events across different formats and systems is complex.

✗ **Evolving Threats:** Traditional signature-based methods fail against zero-day and polymorphic attacks.

✗ **Alert Fatigue:** High volume of false positives from isolated systems overwhelms analysts.

Conclusion: A unified, intelligent approach is needed to see the whole picture.

Our Proposed AI Framework

A Unified AI-Powered Framework

1. Data Ingestion & Feature Extraction

Collects raw data from all monitoring technologies (Network, EDR, Logs).

AI models extract meaningful features from each heterogeneous data stream.

2. Advanced Data Fusion

Unifies the disparate features into a **coherent, single dataset**.

Creates a holistic view of system behavior across all layers.

3. Deep Learning Analysis

A deep learning model (e.g., LSTM, CNN) analyzes the fused data.

Identifies subtle, complex **attack patterns** and anomalies invisible to siloed tools.

4. Real-Time Adaptive Detection

Continuously learns and adapts to new, evolving threats.

Provides actionable alerts with high-confidence prediction

Experimental Results & Advantages

Evaluation Demonstrates Significant Improvements

📈 Key Results:

Higher Detection Accuracy: Identifies sophisticated, multi-vector attacks that evade traditional tools.

Drastic Reduction in False Positives: AI correlation provides context, separating real threats from noise.

Real-Time Adaptability: The system learns and evolves with the threat landscape.

✔ Core Advantages:

Robustness: Effective across diverse and complex IT environments.

Scalability: AI-driven automation handles massive data volumes.

Proactive Defense: Shifts from reactive to predictive security.





This chart shows the superior detection rate of the proposed AI framework compared to two traditional methods (Signature-Based and a simple Anomaly Detection system) against a known dataset of attacks.

