AUTOMATING MALWARE DETECTION: A DATA-DRIVEN APPROACH TO GENERATING HIGH-FIDELITY DETECTION RULES FOR EMERGING THREATS

Student: David Antonio Martin Escar C0326762@setu.ie
Supervisor: Michael Gleeson

Ollscoil Teicneolaíochta an Oirdheiscirt
South East Technological University

Introduction

Current malware detection heavily relies on security vendor signals and trusting those vendors to update their signatures and behavioural patterns for detection quickly after new threats are seen in the wild.

Current detection engineering practices often also rely on a reactive approach: a malicious behaviour is identified, either in logs or a forensics report, and rules are created for it after the fact.

This research aims to identify if an automated and data-driven approach to detection rule creation is comparable, in terms of complexity and capabilities, to the current manual approach.

This research also tries to identify if detection rules can be used to target general malicious behaviours from other families.

Literature Review

- **Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software** by M. Sikorski and A. Honig (2014), is generally considered a great reference manual for in-depth malware analysis. **Practical Binary Analysis: Building Your Own Linux Tools for Binary Instrumentation, Analysis, and Disassembly**, by D. Andriesse (2019). Both have educated decisions to understand and extract the key behaviours from malware samples to be able to create detection rules.
- **Reflective Beam Search for Automated TTP Extraction and Sigma Rule Generation from Cyber Threat Intelligence** by J. Fairbanks and E. Serra (2025) presented some similarities in the topic of automated rule generation and leveraged several evaluation scores that will also be used in this research.
- **CTI-REALM: Benchmark to Evaluate Agent Performance on Security Detection Rule Generation Capabilities**, by the Microsoft Security AI team (2026) presents a framework to benchmark AI agents' interpretation of CTI and generating Sigma detection rules in an automated AI-assisted manner.
- **Digital Forensics and Incident Response: Incident response tools and techniques for effective cyber threat response** (3rd Edition), by G. Johansen (2022) leveraged to understand how to securely extract indicators of compromise and logs from the sandboxed environment, analyse host and network evidence, and use tools such as a SIEM for detection engineering.

Research Questions

The operational value of security detections is determined by three partially independent factors: **authorship method, detection type** and **survivability.**

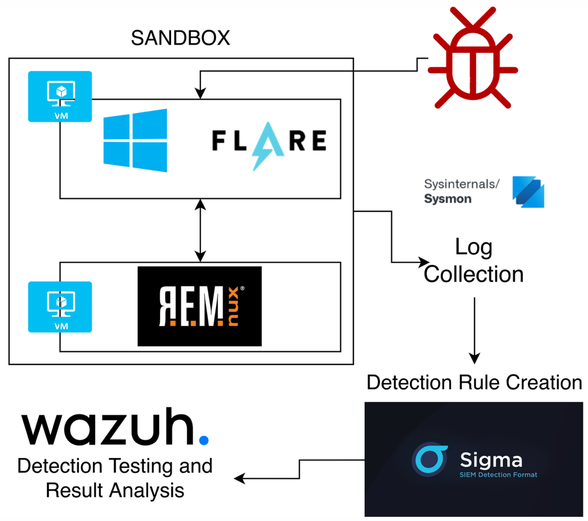
The research aims to answer the three following questions:

1. How do automated detection rule generation methods compare to expert-authored detection rules in terms of fidelity and operation efficacy across atomic and correlation rule types?
2. Do correlation rules demonstrate measurably different detection efficacy compared to atomic detection rules, regardless of how those rules were authored, and what are the associated trade-offs in complexity and difficulty of deployment?
3. Do detection rules, across structure types and authorship methods, generalize to novel malware variants or families not represented during rule generation?

Proposed Methodology: Experimental

1. Collection of 4 malware samples from 2025 after confirming newer samples of the same malware are available.
2. Every sample will be detonated in a controlled and sandboxed environment where logs and telemetry will be collected.
3. Detection rules will be created manually and programmatically to be implemented in a SIEM.
4. Substantial telemetry and logs will be fed to the SIEM to go through the detection engine, some of them containing those malware executions. The resulting data will be collected to assess the rule (precision, recall, FP rate, etc.); answering RQ1, and RQ2.
5. Detection rules will be assessed against newer versions of the same malware, against the same malware family (i.e., Ransomware A rules against Ransomware B's behaviour), and against a different malware family (i.e., Ransomware rules against a C2 implant); answering RQ3.

Technology Stack



Early Indications & Next Steps

- AI-assisted automated detection engineering is becoming more prevalent based on existing literature and research, and personal work-related experience where AI-assisted (and even fully AI autonomous) frameworks are appearing and gaining traction.
- Research around detection engineering is scarce, and generally the research on the topic seems to be related to AI improvements.
- Malware-initiated infections are still a very prevalent vector of entry into companies' IT infrastructure. Most of these infections target initially a human element via social engineering and/or phishing, and once in the environment.
- The next steps for the research are to completing the setup of the sandboxed environment and collect different malware samples for the test runs. Malware samples will be run, then the data will be collected and used to create detection rules as defined in the methodology.

setu.ie

Sigma SIEM Detection Format · wazuh. · FLARE · Sysinternals/ Sysmon · REMnux