Actionable Malware Classification in Embedded Environments using Hardware Performance Counters

Citation for published version (APA):

Document license:
Unspecified

Document status and date:
Published: 12/12/2021

Document Version:
Other version

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.

Download date: 08. Jun. 2022
Actionable Malware Classification in Embedded Environments using Hardware Performance Counters

Martin Rosso\textsuperscript{1}, Joost Renes\textsuperscript{2}, Nikita Veshchikov\textsuperscript{2}, Eduardo Alvarenga\textsuperscript{2}, Jerry den Hartog\textsuperscript{1}

SUMMARY

- **Actionable classification** of malicious activities based on CPU hardware performance counters (HPC)
- Evaluation with simulated tasks and attacks on real hardware
- **Interpretable** rules allow manual inspection of misclassifications

PROBLEM INTRODUCTION

- Connected, embedded or IoT devices are a desirable target for cyber attacks (see e.g., Mirai botnet [1])
- Detecting such attacks early allows to stop propagation and limit their impact, especially in large deployments
- Choosing a suitable mitigation strategy requires timely identification and **classification of ongoing attacks**

METHODOLOGY AND EXPERIMENT

- We detect and classify attacks using HPCs by the CPU
- We create a **labelled dataset** of more than 2.5 million traces, with and without attacks
- Data was collected on MCIMX8M-EVK development boards
- Tasks simulate background activity on the device:
  1. \texttt{ghostscript-ppm} as defined in MiBench [2]
  2. \texttt{e-book} is a combination of MiBench [2] algorithms typical for the functionalities of an e-book reader
  3. **Driver Monitoring System (DMS)** for cars, implemented using OpenCV image recognition [3]
- We implemented four different attacks:
  1. **Ransomware**: encrypting files on the target device
  2. **Page Cache** Attack: a timing side-channel attack
  3. **Cache Attack**: side-channel attack targeting CPU caches
  4. **DDoS**: the device becomes part of a DDoS botnet
- We use a **supervised machine learning**, a tree learning classifier in our case, to extract simple profiles for each type of attack

HPC-BASED ATTACK CLASSIFICATION

- Small Trusted Computing Base (TCB)
- **Hard to circumvent**, e.g., cache attacks must always interact with the CPU cache and the event counter always picks this up [4]
- HPCs are low cost, low effort, and out of the box available on many modern CPUs (e.g., x86/AMD64, Arm, RISC-V)
- Works best on simple, well-defined use cases

RESULTS

- Using HPC it is possible to detect and classify attacks
- Different attack types have different characteristics, see Figure 1
- Detailed results are depicted in Table 1 (right side)

IMPACT

- Very efficient with minimal computational overhead in software or hardware, even on low-end systems
- Timely attack classification allows **better counteractions** against ongoing attacks

FUTURE WORK

- Parameter optimization to maximize accuracy, e.g., choice of micro-architectural events and sampling periods
- Add **context awareness** to reduce false classifications e.g., by evaluating multiple consecutive values together in a window
- More types of attacks and additional use cases

REFERENCES


This work was partially funded by NWO projects INTERSECT (NWA.1160.18.301) and DEPICT (628.001.032).