VITAMIN-V: Virtual Environment and Tool-Boxing for Trustworthy Development of RISC-V-Based Cloud Services

Hardware-based Stack Buffer Overflow Attack Detection on RISC-V Architectures

Cristiano Chenet, Ziteng Zhang, Enrico Magliano, Alessandro Savino, Stefano Di Carlo, Dimitris Gizopoulos, Ramon Canal

SMILIES reSilient coMputer archItectures and LIfE Sciences

Politecnico di Torino
Department of Control and Computer Engineering

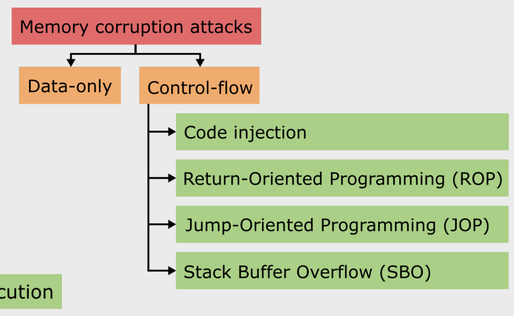
OVERVIEW

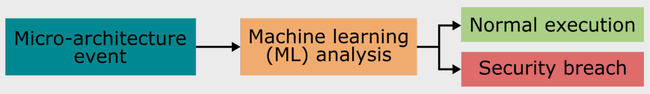
This study evaluates how well hardware-based methods detect stack buffer overflow (SBO) attacks in RISC-V systems. We conducted simulations on the PULP platform and examined micro-architecture events using semi-supervised anomaly detection techniques. The findings indicate that for a malicious function comprising 1% of the application size, detection accuracies exceeded 90% for AES, RSA (with fixed prime numbers), SHA, and Dijkstra applications. This approach presents compelling benefits that could enhance security of RISC-V-based systems.

Vitamin-V aims to develop a complete RISC-V open-source software stack for cloud services with iso-performance to the cloud-dominant x86 counterpart and a powerful virtual execution environment for software development, validation, verification, and testing.

1. INTRODUCTION

- Cyberattacks are among the top global risks [1].
- Exploiting memory-corruption vulnerabilities is one of today's most common exploitation methods [2].
- The detection of security breaches based on hardware events arose in the early 2010s [3].

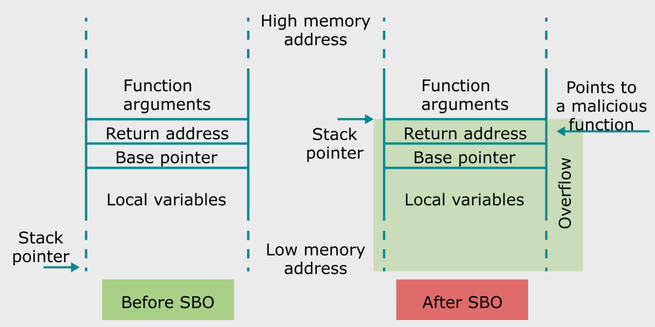




- Advantages of hardware-based SBO detection approaches [4]:
  - Possibility of runtime detection;
  - Adaptability to code variants and zero-day breaches;
  - Resilience against subverting the protection mechanism;
  - Reduced detection costs.
- **Goal**: to analyze the performance of hardware-based approaches in detecting SBO attacks in RISC-V architectures:
  - Focus on semi-supervised anomaly detection techniques;
  - Performance evaluation of four different classification models;
  - Evaluation of an autoencoder to improve anomaly detection accuracy.

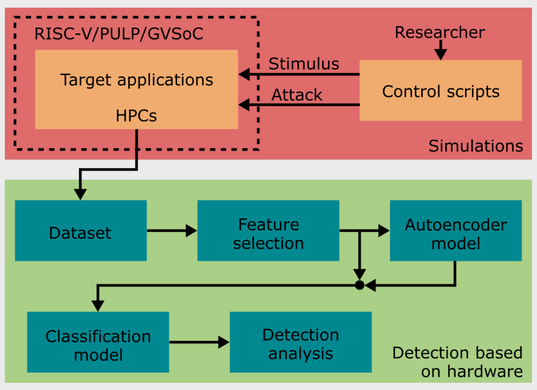
2. STACK BUFFER OVERFLOW ATTACK

- Characterized by non validated input overflowing a buffer allocated in the memory stack, deviating the program execution to a malicious function.



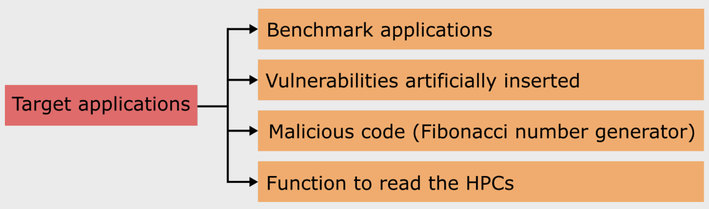
3. METHODOLOGY

- Hardware performance Counters (HPCs) are recorded at the end of the application execution.
- Training machine learning algorithms with randomized inputs allows for breach detection with any input in the benchmark applications.



3.1. Simulation environmental

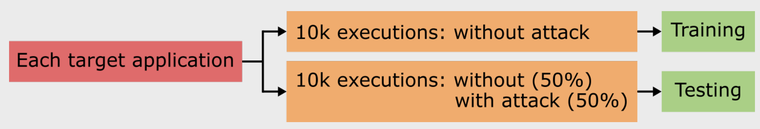
- GVSoC simulator [5].
- Target applications including:



3.2. Detection based on hardware

- Implemented in Python using the Scikit-learn library.
- Feature selection:
  - Principal Component Analysis (PCA) employed to build a ranking;
  - Cores may have strict limitations on the number of HPCs recorded at a time.
- Classification models:
  - One-class Support Vector Machines (OC-SVM)
  - Local Outlier Factor (LOF)
  - Isolation Forest
  - Elliptic Envelope

4. EXPERIMENTAL RESULTS



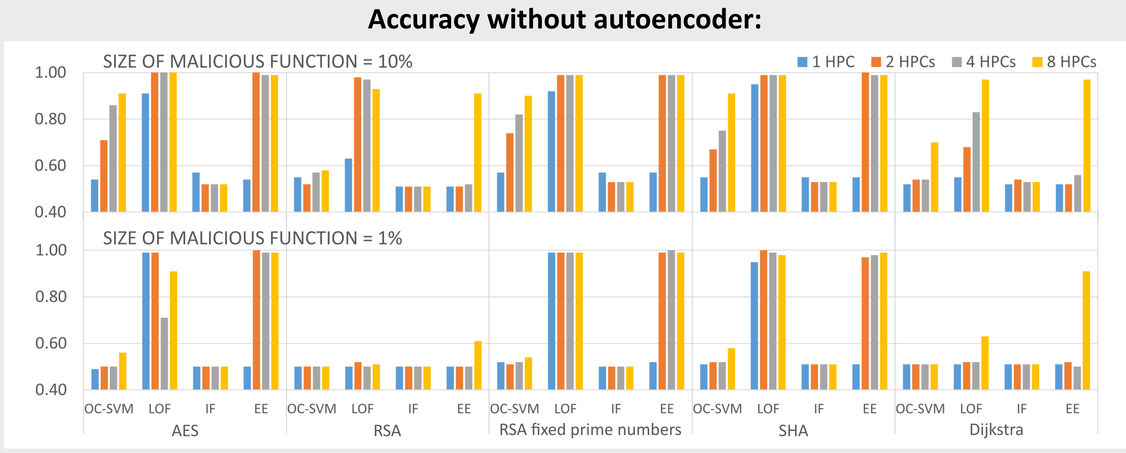
- Ranking of HPCs (1 is the most significative).

HPCs	AES	RSA	RSA fixed PN[1]	SHA	Dijkstra
INSTR	1	1	1	1	1
LD_STALL	4	7	7	4	5
LD	3	3	3	3	3
ST	5	5	5	5	6
JUMP	6	8	8	6	8
BRANCH	7	4	4	7	4
BTAKEN	8	6	6	8	7
RVC	2	2	2	2	2

[1] PN is prime numbers.

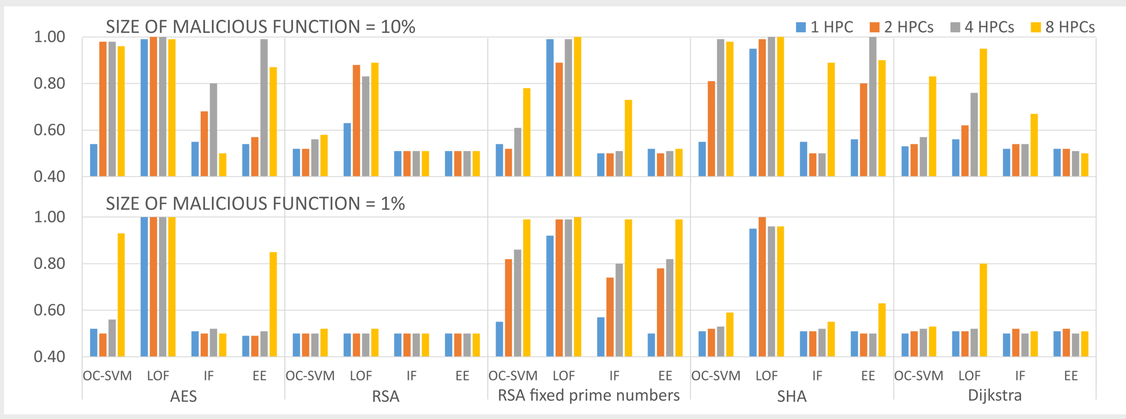
- The results are expressed as a function of the malicious function size.

Accuracy without autoencoder:



- In AES, RSA with fixed prime numbers, SHA and Dijkstra, a malicious function size of 1% can be detected with an accuracy higher than 90%.
- In complete RSA, a malicious function size of 10% is needed to an accuracy higher than 90% (randomness is a challenge).
- LOF has accuracy of at least 95% with just 1 HPC and malicious function size of 1% (AES, RSA with fixed prime numbers and SHA applications).

Accuracy with autoencoder:



- There is no significant gain with the addition of the autoencoder.
- OC-SVM shows interesting benefits with autoencoder, while Elliptic Envelope mostly decreases its performance.

5. FINAL CONSIDERATIONS

- Good performance of the Local Outlier Factor (LOF) model.
- The randomness on RSA algorithm offers a challenge for detection.
- The autoencoder is not decisive in enabling attack detection.
- The detection performance is the main challenge, a potential solution combines software and hardware-based detectors.
- The approach offers advantages that may benefit RISC-V architectures, like runtime detection, code variants and zero-day breaches detection, resilience and reduced costs.

REFERENCES

1. **Global Risk Report 2023**. Jan. 2023. url: https://www.weforum.org/publications/global-risks-report-2023/in-full/.
2. Marco Brohet and Francesco Regazzoni. "**A Survey on Thwarting Memory Corruption in RISC-V**". In: ACM Comput. Surv. 56.2 (Sept. 2023). doi: 10.1145/3604906.
3. John Demme et al. "**On the Feasibility of Online Malware Detection with Performance Counters**". In: SIGARCH Comput. Archit. News 41.3 (June 2013). doi: 10.1145/2508148.2485970.
4. Cristiano Pegoraro Chenet, Alessandro Savino, and Stefano Di Carlo. "**A Survey on Hardware-Based Malware Detection Approaches**". In: IEEE Access 12 (2024). doi: 10.1109/ACCESS.2024.3388716.
5. Nazareno Bruschi et al. "**GVSoC: A Highly Configurable, Fast and Accurate Full-Platform Simulator for RISC-V based IoT Processors**". In: 2021 IEEE 39th International Conference on Computer Design (ICCD). doi: 10.1109/ICCD53106.2021.00071.

This work was supported by Project Vitamin-V project (Project number: 101093062) funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

Cristiano Pegoraro Chenet / Enrico Magliano

DOI —10.5281/zenodo.11209298

cristiano.chenet@polito.it / enrico.magliano@polito.it