Obfuscation Revealed: Leveraging Electromagnetic Signals for Obfuscated Malware Classification

ACSAC, December 10th, 2021

Duy-Phuc Pham, Damien Marion, Matthieu Mastio, Annelie Heuser
Univ Rennes, Inria, CNRS, IRISA, France

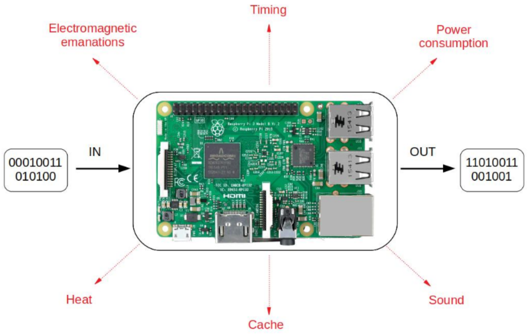
Introduction

- Trending of attacks on embedded devices.
- Limited resources of embedded devices, diversity of architectures.
- Malware analysis and bypasses: difficulties such as malware evasion techniques, packed and obfuscated samples.

$\rightarrow$ Malware analysis through side-channel signals.

Introduction (2)

- Side channel information
  - Power consumption, heat
  - Electromagnetic
  - Sound
  - Cache, HPC (software)
- Black-box dynamic execution

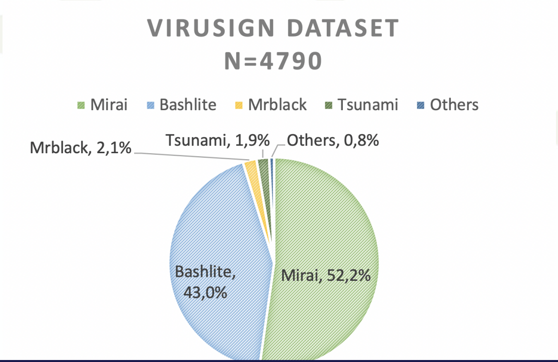


State of the art

- Anomaly detection using power consumption and EM.
- Lack of research on in-the-wild malware side-channel detection.
- No variations regarding obfuscation and packers.
- Our contributions:
  - Malware classification
  - Real-world malware
  - Malware variants

Dataset: Understanding of IoT malware epidemiology

- AVClass to classify malware labels



EMSEC embedded security & cryptography

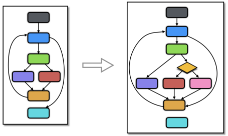
UMR IRISA

Dataset: Malware through code reviews and reverse engineering

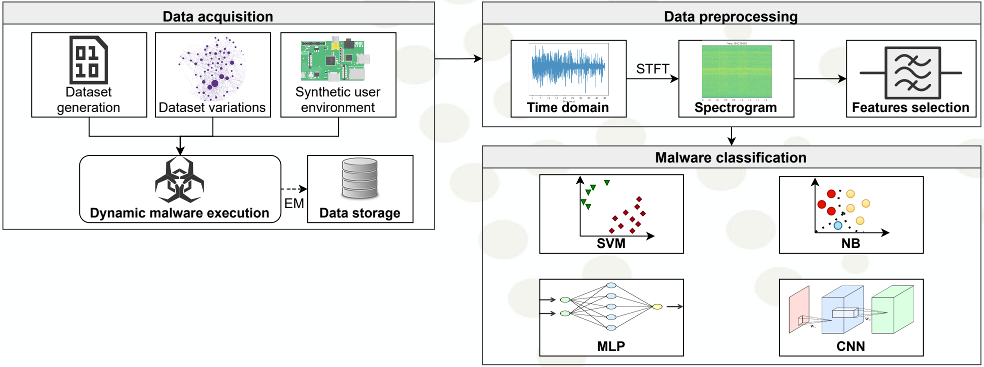
DDoS	Ransomware	Rootkits
Mirai	GonnaCry	KeySniffer
Bashlite	(AES, Blowfish, DES)	MaK_It

Dataset: Variations through Obfuscations

- UPX, Tigress, O-LLVM
- Opaque predicates, bogus control flow, instructions substitution, control-flow flattening; packer and code virtualization



Proposed framework (Open source)



Target device

Requirements

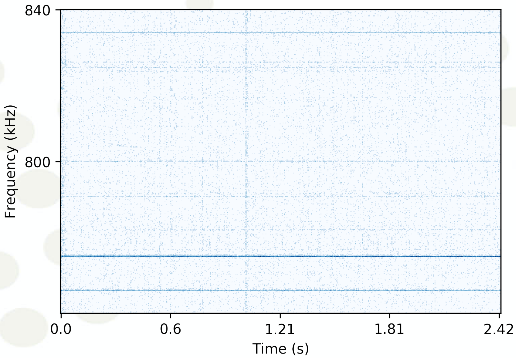
- Multi-purpose embedded device.
- Prominent architecture (ARM).
- Vulnerable to EM side-channel attack.

→ Raspberry Pi 2B

Data and pre-processing

- **Raw traces:**
  106k(traces) $\times$ 2(MS/s) $\times$ 2.5(s) [1.2To]
- **Time-frequency representation:**
  Short-time Fourier transform

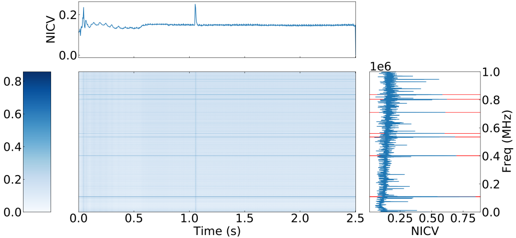
$$\begin{cases} \text{windows} & = 8192 \\ \text{overlap} & = 4096 \end{cases}$$



Features selection

$$\mathrm{NICV}(X,Y) = \frac{\mathrm{Var}[\mathbb{E}[X|Y]]}{\mathrm{Var}[X]}$$

$$F_{\text{extract}} = \mathrm{argmax}_{\epsilon}\left(\left\{\max\left[\mathrm{NICV}(X,Y)_f^D\right]\right\}_{f<F}\right)$$



EMSEC embedded security & cryptography
UMR IRISA

Machine Learning & Deep Learning models

- Linear Discriminant Analysis (LDA) + Naive Bayes (NB)
- Linear Discriminant Analysis (LDA) + Support vector machine (SVM)
- Multi-Layer Perceptron (MLP)
- Convolutional Neural Network (CNN)

Malware classification results

Scenarios	#	MLP	CNN	LDA+NB	LDA+SVM
Executables	31	73.56 [24]	**82.28** [24]	70.92 [28]	71.84 [20]
Type	4	99.75 [28]	**99.82** [28]	97.97 [24]	98.07 [24]
Family	6	98.57 [28]	**99.61** [28]	97.19 [28]	97.27 [28]
Novelty	5	88.41 [16]	**98.85** [24]	98.25 [28]	98.61 [28]
Virtualization	2	95.60 [20]	**95.83** [24]	91.29 [6]	91.25 [6]
Packer	2	93.39 [28]	**94.96** [20]	83.62 [16]	83.58 [16]
Obfuscation	7	73.79 [28]	**82.70** [24]	64.29 [10]	64.47 [10]

Table 1. Accuracy obtained with MLP, CNN, LDA + NB and LDA + SVM applied on several scenarios.

Conclusion

- Classify various malware samples in multiple in-the-wild scenarios.
- Obfuscation technique can be distinguished.
- Evaluation of both DL/ML.
- Evaluated Artifacts:
  - Code: https://github.com/ahma-hub/analysis/wiki
  - Data: https://zenodo.org/record/5414107

Thank you!