Dial 258x: play your app on your speakers to get the malware family
Guillermo Izquierdo-Moreno¹ · Pedro Peris-Lopez¹ · Carmen Camara¹
¹Department of Computer Science, Carlos III University of Madrid

Introduction
– The proliferation along with the sophistication of the malware has made the identification of malicious pieces of software a critical problem to solve.
– Inspiring work based on image-signal processing [2]

Related Work
– A wide amalgam of static and dynamic analysis solutions exist in the literature
– Main drawback: time consuming & expert knowledge

Proposed Solution
– Malware classification via audio-signal processing
– Main advantage: novel approach & high performance (time and accuracy)

Motivation

Workflow
Malware analysis via audio-processing involves four steps.

Step 1: Audio Conversion
– Binary file is converted to WAV format.
– Tool: SoX (Sound eXchange)

Step 2: Feature extraction
– MATLAB & MIRToolbox
– Parallelized implementation
– Full Feature Set: 115

Step 3: Filtering
– Remove families \( \leq 30 \) instances
– Attribute selection

\[ \square \text{Reduced Feature Set: 12} \]

Step 4: Classification
Learner:
– KNN
– Random Forest
– …

Full prototype running in a cluster (Intel Xeon E5-2630 v3, 2.4GHz, 8-cores, 378GB RAM)

Results suggest high performance!

Malware Family Classification
The proposal works like a charm …

Timing Performance
– High speed processing-time
– Linearity (sample size)

Processing rate (steps 1 to 3)

\[ \square \text{3.4 specimens/sec.} \]
\[ \square \text{5.7 specimens/sec.} \]

Conclusions
– Abstraction of the binary structure (non-expert knowledge needed)
– Rich characterization of malware pieces with reduced number of features
– High performance

\[ \square \text{Correctly classified instances \( \sim \) 90\%} \]
\[ \# \text{Specimens analyzed/day \( \sim \) 492\%} \]

Acknowledgements
This work was supported by the MINECO grant TIN2016-79095-C2-2-R (SMOG-DEV - Security mechanisms for fog computing: advanced security for devices).
We thank ARCOS research group (https://www.arcos.inf.uc3m.es) for facilitating access to the used cluster.

Bibliography