Cloud Malware

Domhnall Carlin, Dr. P. O'Kane & Prof. S. Sezer

Centre for Secure Information Technologies, Queen's University, Belfast, N. Ireland.

Research Objective

Develop a strategy for the detection of malware, which is immune to modern obfuscation methods, and which is applicable at the hypervisor level.

Motivation

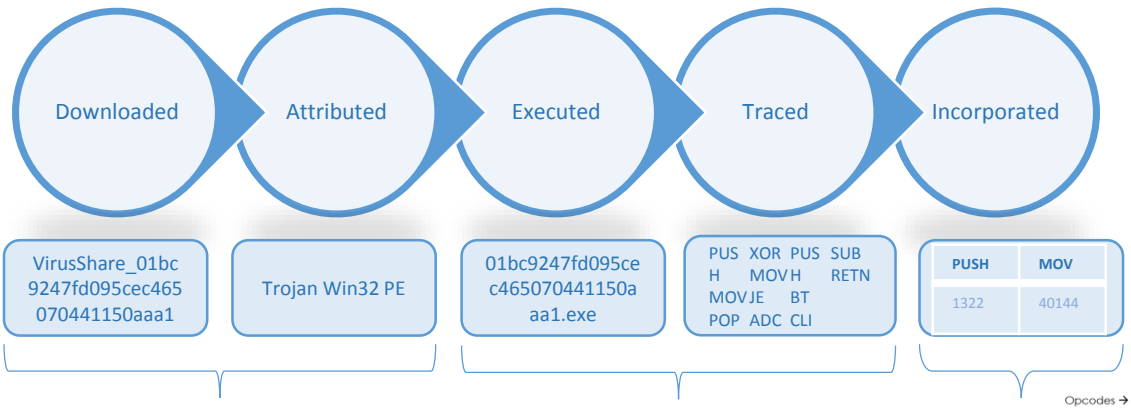
Signature-detection is the most widely used approach in commercial malware detection.
New malware must be captured & analysed for a signature, which is deployed to users.
Obfuscation techniques compound the issue- 1000s of variations generated per day.
By definition, we are constantly behind the curve in the arms race against malware.

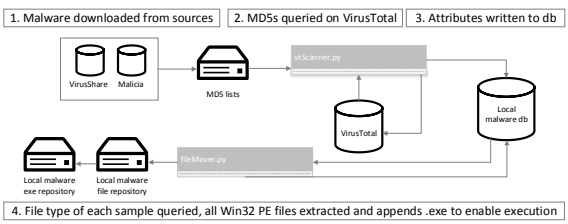
Relevant Research

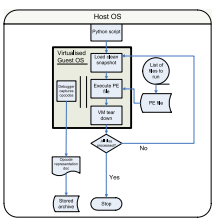
- Analysis of opcodes (assembly instructions) can detect malware, bypassing issues inherent in signature-based methods (i.e. the instance can be novel).
- Packed or encrypted malware cannot generally be investigated statically. Dynamic analysis allows the obfuscated malware to reveal itself.
- Datasets in the literature have been small, poorly sampled and prone to class imbalances.

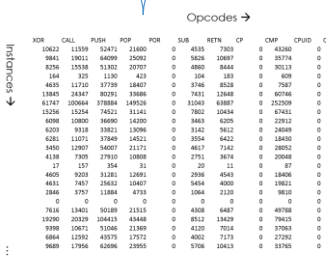
Creation of a large dynamically-generated run-trace dataset

Processing each malicious sample









Future work facilitated by new dataset:

Dataset stands at 48,252 executed, attributed and traced malware samples.

Unsupervised Deep Learning
Neural Net Classification
Relabelling based on opcode feature clusters

Optimal dataset composition
Algorithm comparisons
Novel feature selection &reduction

One-class learning
Rule-based classifiers
Dynamically uninvestigated malware types