Introduction

• Machine learning-driven malware detection systems have demonstrated potential in identifying zero-day malware.
• Existing approaches lack robustness and need more testing on different types of malware.
• AML attacks can help to determine effectiveness and robustness of a detection system.

Challenges:

• Obfuscated malware can be difficult to catch. Memory forensics is the solution. (VolMemLyzer)
• CIC-MalMem-2022 dataset only covers Spyware, Ransomware, and Trojan Horses.
• ML based malware detection systems have been tested on Windows, but further research is needed on Linux and MacOs to create unification between the systems.

Approach

Phase 1:
Develop and train machine learning based Malware Detection Model:
• Take memory snapshot and extract features.
• Data balancing using SMOTE.
• Split data and input into detection system.
• Binary output (malicious or benign).

Phase 2:
Attack the detection model using JSMA
• Collect malware binaries to execute on a VM and take memory snapshot.
• VolMemLyzer to extract features to CSV file (new dataset).
• Feed CSV files into the detection model.
• Record performance for analysis in phase 3.

Phase 3: (Future Work)
Analyze model performance and Adversarial Example Transferability
• Robustifying Techniques
  - Defensive Distillation
  - Adversarial Training
• AE Transferability
  - Provides insight into ML models

Results

• Algorithms Tested in Detection Model
  - Decision Trees, Random Forest, LGBM, XGBoost
  - Top Performers
    - XGBoost, Random Forest
• Metrics used:
  - 10-fold cross validation
  - accuracy
  - F1 Score
  - FPR
  - sensitivity
  - PPV
  - Cohen kappa
  - specificity
  - MCC

Authors

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF, DT</td>
<td>92.01, 99.00</td>
</tr>
<tr>
<td>LR</td>
<td>99.97</td>
</tr>
<tr>
<td>KNN w/ Stacked Ensemble</td>
<td>97.00</td>
</tr>
<tr>
<td>XGBoost, RF</td>
<td>99.98, 99.98</td>
</tr>
</tbody>
</table>

Table 1: Performance comparison of related works.

Conclusions

• ML based Detection systems are a viable solution to combat zero-day malware, but needs more research.
• The new dataset from Phase 2 will help researchers to robustify their models against many forms of malware.

Future work:

• Defensive Distillation, Adversarial Training
• AE Transferability Problem
• Test model on MacOS and Linux