Enhancing Malware Analysis and Detection Using Adversarial Machine Learning Techniques

BOISE STATE UNIVERSITY

COLLEGE OF ENGINEERING

Troy Tolman, Md. Mashrur Arifin, Dr. Jyh-haw Yeh

Boise State University

Introduction

- Machine learning-driven malware detection systems have demonstrated potential in identifying zero-day malware.
- Existing approaches lack robustness and needs 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).

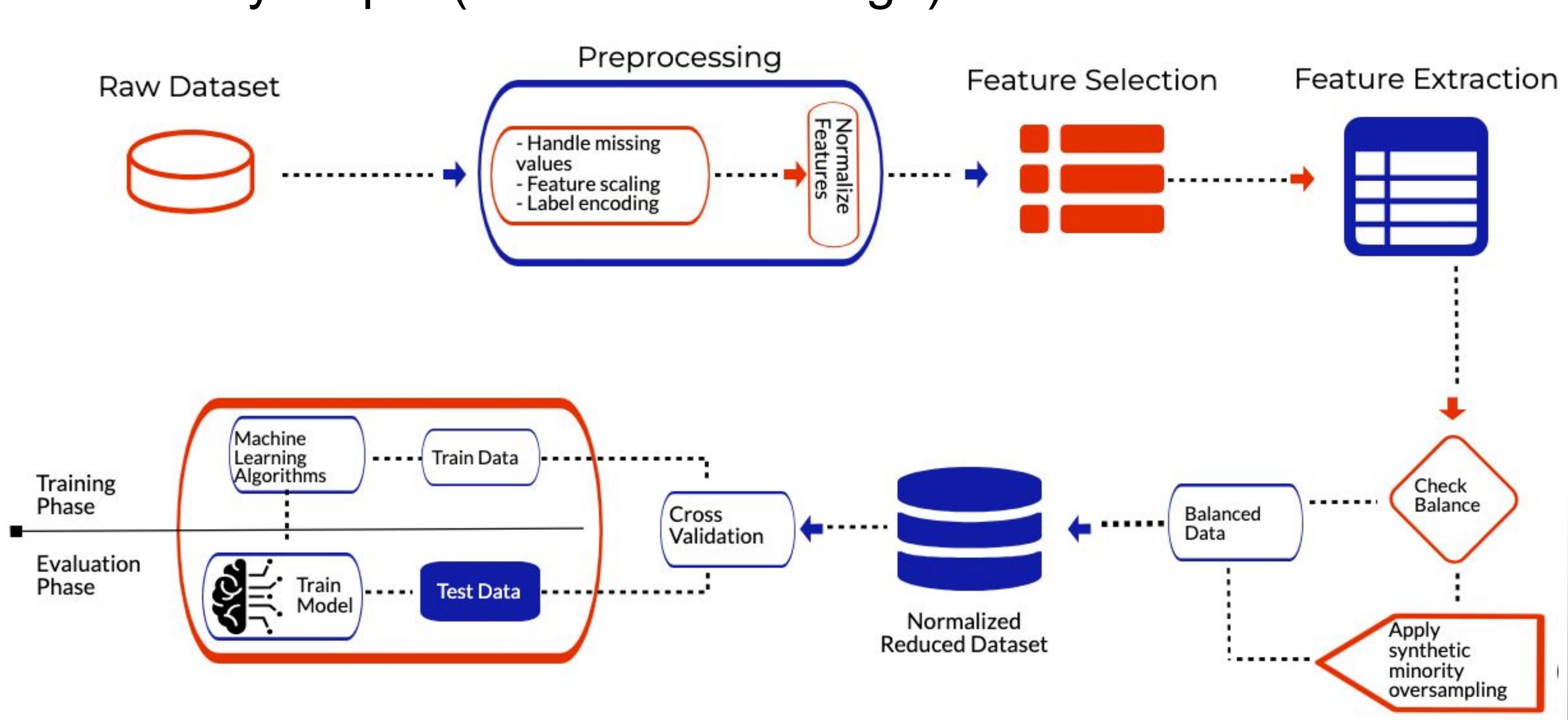


Figure 1: Basic ML based Detection System Workflow

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.

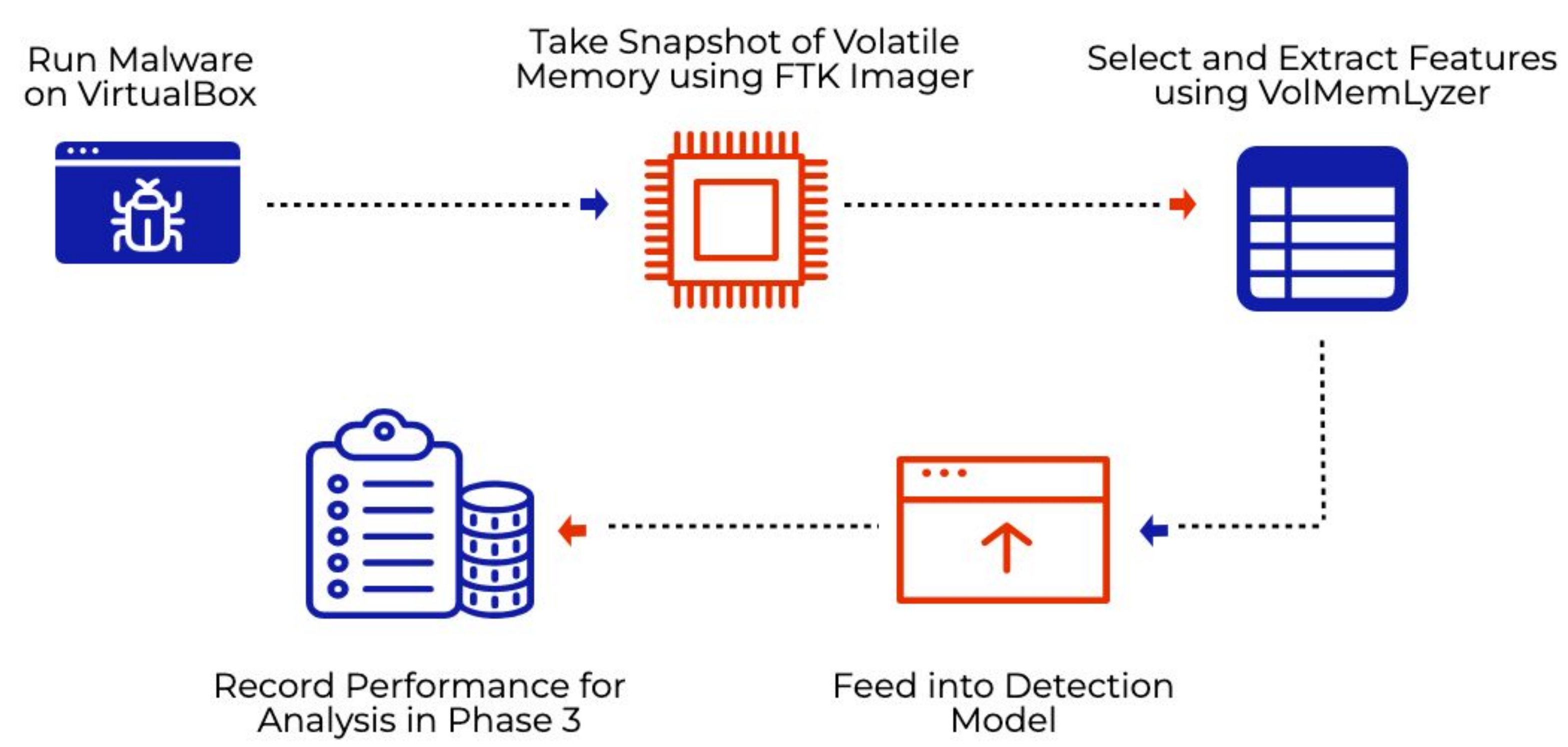


Figure 2: AML Model Workflow

Phase 3: (Future Work)
Analyze model performance and Adversarial Example Transferability

- Robustifying Techniques
  - Defensive Distillation
  - Adversarial Training
- AE Transferability
  - Provides insight into ML models

Tools

python

Volatility

VirtualBox

scikit learn

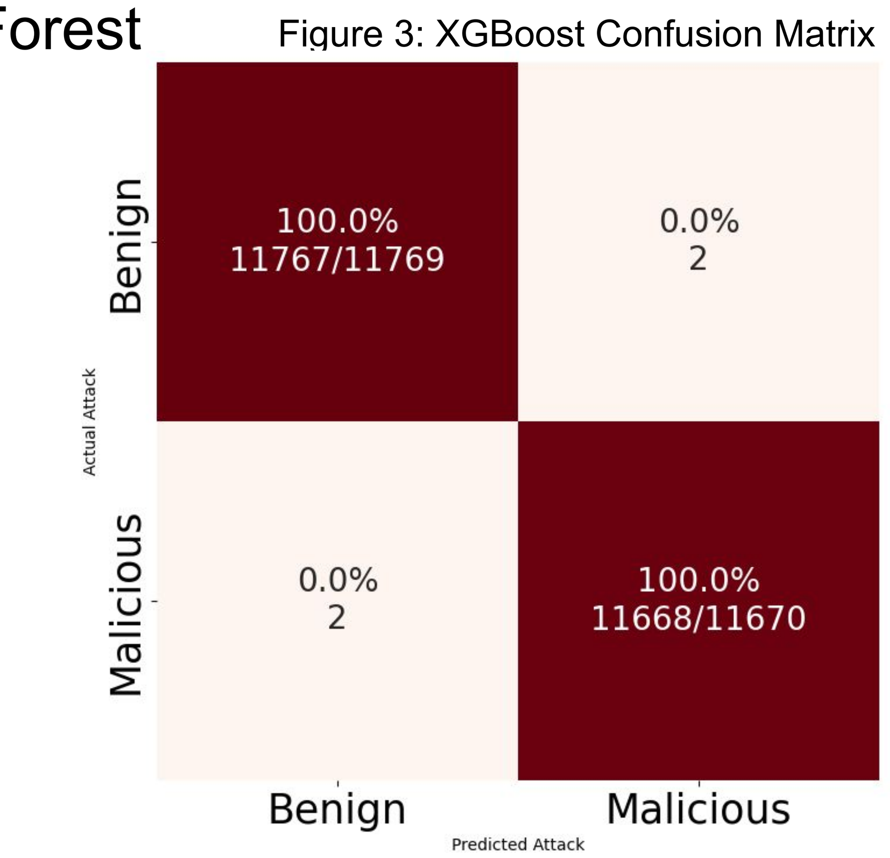
jupyter

exterro FTK® Imager

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

Figure 3: XGBoost Confusion Matrix



Authors	Algorithm	Accuracy (in %)
[1]	RF, DT	92.01, 99.00
[2]	LR	99.97
[3]	KNN w/ Stacked Ensemble	97.00
This study	XGBoost, RF	99.98, 99.98

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

References
1. Akhtar et al, Malware Analysis and Detection Using Machine Learning Algorithms. Symmetry 2022, 14, 2304.
2. Dener et al, Malware detection using memory analysis data in big data environment. Appl Sci, 12 (17) (2022), p. 8604
3. Carrier et al, Detecting obfuscated malware using memory feature engineering. ICISSP, SciTePress (2022), pp. 177-188

NSF Cloud Computing and Privacy REU Award Number: 2244596