DOI: 10.20396/revpibic2720192426

XXVII Congresso de Iniciação Científica Unicamp
16 a 18 de outubro de 2019 - Campinas | Brasil

Caracterização do Comportamento de Malware em Ambientes Linux - Linux Malware Behavior Description

Lucas B. Galante*, Marcus F. Botacin, Prof. Dr. Paulo L. de Geus.

Abstract

A major threat to system's security is malware infections, which cause financial and image losses to corporate and end-users, thus motivating the development of malware detectors. In this scenario, Machine Learning (ML) has been demonstrated to be a powerful technique to develop classifiers able to distinguish malware from goodware samples. However, many ML research work on malware detection focus only on the final detection accuracy rate and overlook other important aspects of classifier's implementation and evaluation, such as feature extraction and parameter selection. In this project, we shed light to these aspects to highlight the challenges and drawbacks of ML-based malware classifiers development. We discovered that (i) dynamic features outperforms static features; (ii) Discrete-bounded features present smaller accuracy variance; (iii) Datasets presenting distinct characteristics impose generalization challenges to ML models; and (iv) Feature analysis can be used as feedback information for malware detection and infection prevention.

Key words:

Malware, Linux, Reverse Engineering

Introduction

Machine Learning (ML) has been demonstrated to be a powerful technique to develop classifiers able to distinguish malware from goodware samples [1]. However, many ML research work on malware detection focus only on the final detection accuracy rate and overlook other important aspects of classifier's implementation and evaluation, such as feature extraction.

In this project, we shed light to these aspects to highlight the challenges and drawbacks of ML-based malware classifiers development. We trained 25 distinct classification models and applied them to 2,800 Linux ELF malware binaries. Our models considered distinct types of features, thus motivating a discussion about their implications in ML models.

Results and Discussion

Classification algorithms are supported by distinct assumptions, thus they produce different outcomes for the same datasets. In addition, each classifier presents its own tuning parameters, which can be adjusted according the classification task. Therefore, observing classifiers behavior is essential to select the best one for each task. In this work, we considered the behavior of three ML classifiers (SVM, RF, MLP) and varied their parameters such as to always achieve the best accuracy rates.

Table 1 presents accuracy rates of the best classifiers of SVM, RF and MLP considering varying analysis (e.g., static vs dynamic) and feature representation (e.g., continuous vs discrete). Whereas this accuracy-focused classifier selection step is considered by most work in the academic literature, additional reasoning steps are often overlooked.

From an ML algorithm point of view, features are understood only as a vector which is classified regardless of its source or interpretation. From a malware analysis point of view, however, features represent the behavior of the sample it was extracted from. To understand the impact of relying on distinct feature extraction procedures, we submitted the same malware samples to static and dynamic analysis procedures and considered accuracy results for the same features models and classifiers. All models based in dynamic features presented higher accuracy rates than models based on static features.

Table 1. Accuracy rates for different classifiers.

	SVM	RF	MLP
Continuous	98.62%	98.98%	96.85%
Discrete	84.48%	85.93%	85.86%
Static	98.62%	98.98%	96.85%
Dynamic	98.54%	99.36%	98.87%

ML classifiers present some advantages and drawbacks in comparison to human heuristics, but these are not often discussed. To evaluate the impact of humans and machines selecting classification boundaries, we developed two classification models for all classifiers; discrete and continuous model. Accuracy rates for continuous features are higher than those with discrete features, which are expected due to the higher capabilities of machines. On the other hand, classification variance among datasets is smaller in discrete features.

Whereas the ultimate goal of a malware classifier is to detect malware, we advocate for the need of understanding how classification decision work, thus allowing infection remediation and prevention. Evaluation of the most distinguishable features for malware classification might allow incident response and system hardening.

Conclusions

In this project, we discovered that: (i) dynamic features outperforms static features when the same classifiers are considered; (ii) Discrete-bounded features present smaller accuracy variance over time in comparison to continuous features, at the cost of some time-localized accuracy loss; (iii) Datasets presenting distinct characteristics (e.g., temporal changes) impose generalization challenges to ML models; and (iv) Feature analysis can be used as feedback information for malware detection and infection prevention.

Acknowledgement

This work was supported by the Brazilian National Counsel of Technological and Scientific Development (CNPq) - PIBIC 2016/2019, process 800295/2016-1.

[1] Imran, M., Afzal, M., and Qadir, M. A. *Journal of Intelligent & Fuzzy Systems*, **2016**, 31:837–847