MtNet: A Multi-Task Neural Network for Dynamic Malware Classification

Jeff Malavasi
Dept. of Computing Security
Rochester Institute of Technology
Email: jm3378@rit.edu

I. Review

The authors of the paper design a new malware classification system based on a deep neural network (DNN). DNNs accelerate the creation of a model, by automatically extracting features from one or more hidden layers. In their model, they compare multiple architectures that mainly extract features using dynamic analysis of the file of interest. In addition to their classification system, they also propose a deep learning network that can accurately detect the phylogeny of the sample. Additionally the authors show that they can reduce the error rate of deep neural network malware classifiers by implementing dropout and rectified linear activation functions [1].

The researchers propose both a single and multi task architecture that is trained to both classify and detect the family of the binary of interest. They utilized a dataset that comprised over 6.5 million files, and ultimately selected a ratio of 3:4 malicious to benign files. The files were ran through an anti-malware engine which aims to reduce evasion while still gathering information about the execution of the file. Specifically, the engine recorded the sequence and parameter value for each API call that was made during execution. The feature selection process ultimately produced 50,000 features, which was later reduced to 4,000 using the random projection technique [1].

During their analysis, the authors found that their single-task model performed significantly better than baseline (23%) in both classification and detection of phylogeny. Additionally, they found that through the introduction of dropout and rectified linear activation functions, feature selection improved enough to reduce the overall error rate of the model. Additionally, they found that the multi-task model also outperformed baseline in binary classification, but failed to improve accuracy of class detection [1].

In this paper, the authors successfully create a model that improves the accuracy of malware classification based on previous research. The success of their model, can likely be attributed to the extremely large sample size of their training dataset. However, there were also some limitations. First of all, the authors highlight the fact that their contributions only improve the accuracy by a small amount compared to the baseline architecture, as improving deep learning networks is extremely difficult. Additionally, while their dataset included thousands of different malware families, they only trained their model on a small subset. Future research will be needed to build classifiers for less prevalent malware.

As deep learning becomes more and more accessible, researchers will continue to use it to improve our understanding of malware. While many experiments have been able to create accurate classifiers, the authors in this study expanded on the previous work by creating a multi-task architecture. Through the use of a deep network, the authors were able to reduce the time needed for feature selection by allowing the model to directly extract features from the raw data. Additionally, they investigated methods that could be used to reduce the error rate of their system. Future work will be needed to continue to improve the accuracy of these models, while reducing the compute overhead needed to create them.

References

[1] W. Huang and J. W. Stokes, "Mtnet: A multi-task neural network for dynamic malware classification." Springer-Verlag, 2016.