A Weak Coupling of Semi-Supervised Learning with Generative Adversarial Networks for Malware Classification

Shuwei Wang[1,2], Zhengwei Jiang*[1,2], Qiuyun Wang[2], Xunren Wang[3,2], Rongqi Jing[1,2]
1. Institute of Information Engineering, Chinese Academy of Sciences.
2. University of Chinese Academy of Sciences.
3. Information Engineering College, Capital Normal University
* Corresponding Author

Introduction

We propose an improved malware image rescaling algorithm (IMIR) based on local mean algorithm. Its main goal of IMIR is to reduce the loss of information from samples during the process of converting binary files to image files.

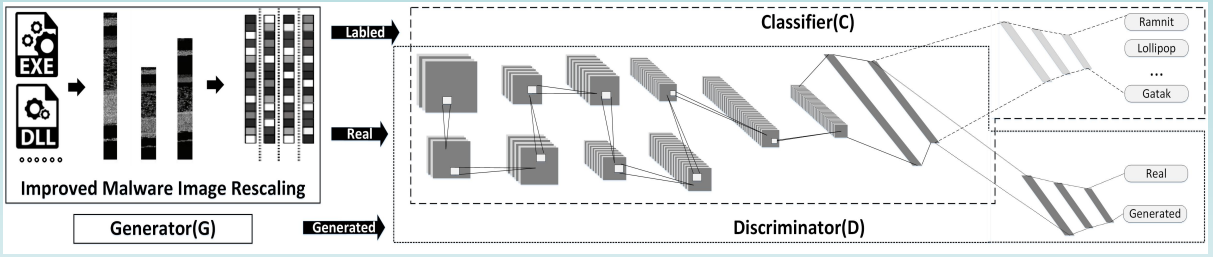
Therefore, we construct a neural network structure based on VGG model, which is suitable for image classification.

And we propose a novel method to train the deep neural network by Semi-supervised Generative Adversarial Network (SGAN), which only needs a small amount of malware that have accurate labels.

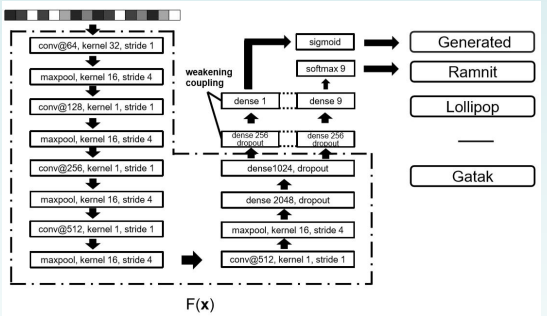
By integrating SGAN with weak coupling, we can retain the weak links of supervised part and unsupervised part of SGAN. It improves the accuracy of malware classification by making classifiers more independent of discriminators.

Proposed Method

- Improved malware image rescaling algorithm (IMIR) extends the sampling range to the edge of sliding window. Then it adjusts the calculation method of step length and adds a fill step to fit the malware images.



- A one-dimensional convolutional neural network (1D-CNN) is constructed by using the VGG model. Our malware classifier consists of five convolutional layers with max-pool layers, four dense layers, and one softmax layer as output.
- SGAN uses the supervised learning with labeled samples to train model for judging categories, and the unsupervised learning with unlabeled samples to train model for judging true or false.
- We found the separation of classifier C and discriminator D can further strengthen the effect of classifier C based on semi-supervised learning. So we split the process of classification into classification and discrimination, corresponding to the new classifier C and the new discriminator D which share a feature extractor F(x) in common.



Environment

Server Hardware:

- ✓ Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz ✕ 2.
- ✓ 94GB memory.
- ✓ Nvidia Tesla P100 graphics card ✕ 1.

Adam Optimizer:

- ➢ batch size: 64.
- ➢ epochs: 200.
- ➢ keep prob: 0.85.
- ➢ learning rate: 2e-4.
- ➢ bn momentum: 0.9.
- ➢ bn epsilon: 1e-5.

Experiment

True\Pred	Ramnit	Lolipop	Kelihoss.ver3	Vundo	Simda	Tracur	Kelihos._verl	ObfuscalorACY	Gatak
Ramnit	**97.73%**	0.00%	0.00%	0.00%	0.00%	1.29%	0.64%	0.32%	0.00%
Lolipop	0.00%	**98.80%**	0.00%	0.00%	0.00%	0.20%	0.20%	0.20%	0.60%
Kelihoss.ver3	0.00%	0.00%	**99.66%**	0.00%	0.00%	0.00%	0.17%	0.17%	0.00%
Vundo	0.00%	0.11%	0.00%	**98.94%**	0.00%	0.00%	0.00%	0.00%	0.00%
Simda	0.00%	0.00%	0.00%	0.00%	**100.00%**	0.00%	0.00%	0.00%	0.00%
Tracur	0.00%	0.00%	0.00%	0.66%	0.00%	**98.68%**	0.66%	0.00%	0.00%
Kelihos._verl	1.25%	0.00%	0.00%	0.00%	0.00%	1.25%	**93.75%**	0.00%	3.75%
ObfuscalorACY	2.03%	0.81%	0.00%	0.41%	0.41%	0.41%	0.00%	**95.12%**	0.81%
Gatak	0.00%	0.50%	0.00%	0.00%	0.00%	0.00%	0.00%	0.99%	**98.52%**

- ◆ In the experiment group, we evaluated the performance of SGAN and the weak coupling method by 5-fold cross-validation on the adjusted dataset.
- ◆ The recalls of each family are all higher than 93.75%, most of which exceed 98.00%.

Conclusion

We improve the effective with deep learning from three aspects: sample feature extraction, neural network structure, and data labeling.

After experimental verification, the three work together to reduce the time cost of model construction and use, improve update efficiency, and enhance timeliness.

This work is supported by the National Key Research and Development Program of China (Grant No.2019QY1303)