엠버 기반 악성코드 탐지모델에 대한 패커의 영향력 분석

이 지 원, 조 현 준, 정 혜 선, 김 경 백

e-mail : maum-@hanmail.net, limdugmin2@naver.com,
albaneo0724@gmail.com, kyungbaekkim@jnu.ac.kr

전남대학교 정보보안협동과정

Impact Analysis of a Packer to EMBER based Malware Detection Model

Ji-won Lee, Hyeon-jun Jo, Hye-seon Jeong, Kyungbaek Kim

Chonnam National University

요 약

최근 인공지능 기술의 발달에 따라, 정적특성 기반의 인공지능 악성코드 탐지 모델이 활발히 연구중이다. 반면, 다양한 탐지 모델을 무력화 하기 위해, 악성코드를 난독화하는 패커를 사용할 경우, AI기반 탐지모델도 제 기능을 못하고 있다. Packer는 코드 보호를 위해 개발되었으나, 공격수단으로 사용될 경우 내용분석이 어려워 매우 치명적 공격이 될 수 있다. 이 논문에서는 Packing된 악성코드가 정적특성 기반 인공지능 악성코드 탐지모델에 미치는 영향을 분석하고자 한다. 특히, 정정특성 추출 기법인 엠버 모델을 사용하는 인공지능 기반 탐지 모델에 대해 모델 생성 및 탐지 성능에 Packing된 악성코드가 미치는 영향을 분석한다. 분석을 위해 KISA에서 제공한 10여만개 악성코드를 활용하였다. 실험을 통해, Packing된 파일 비율이 올라갈수록 10% 당 3%의 탐지율이 떨어지는 것을 확인할 수 있었으며, Packing된 파일을 학습하더라도 Packer를 고려하지 않은 인공지능 악성코드 탐지 모델의 성능은 한계가 있음을 확인하였다.

I. 서론

현재 일어나는 대부분의 사이버 공격은 악성코드를 이용한 공격이며, 공공기관이나 주요기관뿐만 아니라 우리의 일상적인 환경에서도 적용되고 있다. 이에 최근에는 악성코드 백신뿐만 아니라 AI기반 악성코드 탐지-솔루션들이 나오고 있다.

최근 AI 기반 악성코드 탐지 기법을 회피하기 위한 방법이 개발되고 있다. 특히, 공격자는 Packer를 이용하여 파일을 압축, 난독화 하여 기존의 백신과 AI기반 악성코드 탐지 솔루션을 무력화 할 수 있다. 즉, Packing 과정을 통해, 기존 악성코드 탐지 기법의 정탐률을 현저히 하락시킬 수 있다.

Packer는 실행프로그램을, 실행압축상태로 만들 수 있으며, 탄생부터 현재까지 프로그래머의 지적재산에 대한 방어수단이자, 공격자에게는 자신의 공격 코드를 숨길 수 있는 공격 수단으로써 유용하게 사용되고 있다. 공격 수단으로 쓰이는 경우 그 영향은 매우 치명적이며, 해결책으로 동적 분석이 고려할 수 있으나 APT, Anti-VM 등의 기술이 접목되면서 packing된 코드의 분석이 갈수록 힘들어지고 있다.

이 논문에서는 packing된 악성코드가 정적특성 기반 인공지능 악성코드 탐지 모델에 미치는 영향을 분석하여, packing된 악성코드 탐지 기법 연구를 위한 기초자료를 제공하고자 한다.

II. 관련 연구

현재 packing에 관련된 연구로 시그니쳐 기반으로 packing여부를 탐지하는 연구가 있다

[1]. 탐지된 Packer들을 기반으로 unpacking자동화를 지원해 주는 도구가 있지만, unpacking 자동화가 가능한 경우는 매우 제한적이다. 수동으로 unpacking할 경우에는 VM 환경을 이용하여 unpacking을 진행할 수 있지만, Anti-VM 기능이 packing된 코드에 적용되어 있을 경우, unpacking이 불가능하다[2]. Packer에 관련된 연구는 대부분 unpacking 방법론에 관련된 연구가 많고, 정적상황에서 Packing된 파일에 대한 연구는 미미하다.

III. EMBER 기반 악성코드 탐지

EMBER는 ENDGAME사에서 만든 정적탐지기반 데이터셋 모델이다[3]. 크게 8가지 피처를 추출할 수 있게 해준다. 하지만 이 중 대부분은 PE파일의 헤더와 바디에서 추출한다. 하지만 EMBER는 그것 이외에도 파일 자체의 바이트에 따른 사용횟수, 그리고 JOSUA가 언급한 ENTROPY수치를 이용한 탐색 모델 기법을 가지고 있다[4].

그 모델을 기반으로 KISA에서 제공하는 악성코드 데이터셋 12만개를 가지고 학습하여 탐지율 95%의 AI모델을 만들 수 있었다.

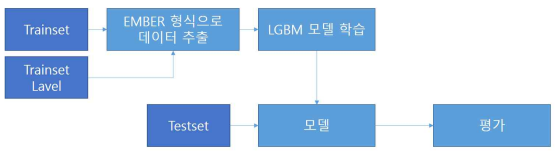

그림 1. EMBER기반 악성코드 탐지 모델

IV. Packing 영향 평가

4.1 테스트셋 Packing 비율의 영향 평가

Packer를 사용하지 않는 악성코드 데이터셋을 이용해 만든 AI기반 악성코드 탐지모델에 대하여, 테스트셋 Packing의 비율이 어떤 영향을 미치는지를 평가하였다. AI모델은 3장에서 언급한 모델을 사용했으며, 테스트셋은 KISA에서 제공하는 악성코드 데이터셋 10000개에 대하여 5%, 10%, 20%, 30% 비율로 UPX을 이용한 Packing을 수행하여 준비하였다. 각 테스트셋을 사용한 경우에 대한 AI기반 악성코드 탐지모델의 성능은 표 1과 같다. Packing 비율이 올라갈수록, 탐지율이 떨어짐을 확인하였다.

표 1. 테스트셋 패킹 비율에 따른 탐지 성능

Packing 비율	tp	tn	fp	fn	Accuracy
0%	6907	2772	228	93	96.79
5%	7088	2400	353	159	94.88
10%	7012	2305	521	162	93.17
20%	6992	2507	798	153	90.49
30%	6866	1853	1123	158	87.19

이후에는 테스트셋 전체를 패킹하였을 때에 인공지능 기반 악성코드 탐지 모델에 어떠한 영향 미치는지 살펴보았다. 그림 2는 서로 다른 패커들로 테스트셋 전체를 패킹하였을 때의 악성코드 탐지 성능을 나타낸다. 모든 결과에서 recall 비율이 상당히 높지만 accurancy 비율이 현저하게 떨어진 것을 볼 수 있다. 즉, 대부분의 패킹 파일들을 모두 악성코드로 판단하여 False positive가 급증하였다.

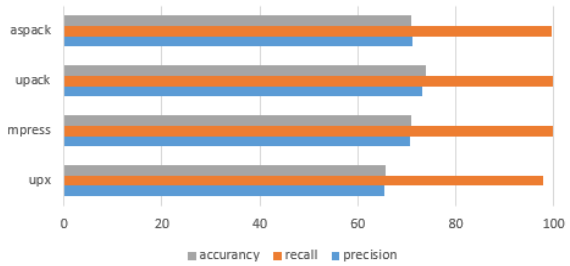

그림 2. 테스트셋 전체 패킹에 따른 탐지 성능

4.2 AI 모델 학습에 대한 Packing 영향 평가

Packing된 파일들로 EMBER 기반 악성코드 탐지 모델을 학습할 경우의 악성코드 탐지 성능에 미치는 영향을 평가하였다.

평가을 위해 KISA에서 제공하는 12만개의 악성코드 데이터셋을 8 : 3 : 0.5 : 0.5로 나누고 그림 3과 같이 각 데이터 부분을 A, B, C, D로 명명하였다. P(B) 데이터셋은 B 데이터셋을 균등하게 나누어 4개의 서로 다른 패커들(UPX, MPRESS, UPACK, ASPACK)로 Packing을 시도하여 Packing에 성공한 파일들로 구성하였다. P(D) 데이터셋은 D 데이터셋을 균등하게 나누어 5개의 서로다른 패커들(UPX, MPRESS, UPACK, ASPACK, Y0DA)로 Packing을 시도

하여, Packing에 성공하거나 실패한 파일을 모두 포함하여 구성하였다.

이렇게 분리된 데이터셋을 조합하여 학습데이터셋 2종류(Train-A,B)와 테스트셋 3종류(Test-A,B,C)를 그림 3과 같이 준비하였다. Train-A는 Packer가 미 적용된 학습데이터셋을 나타내고, Train-B는 Packer가 일부 적용된 학습데이터셋을 의미한다.

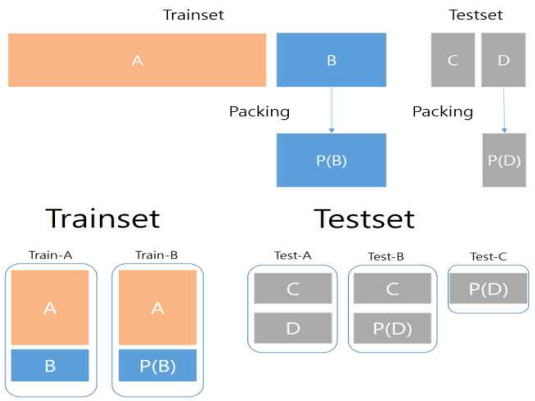

그림 3. Packing적용 데이터셋 구조

Train-A 와 Train-B로 학습한 악성코드 탐지 모델에 Test-A, B, C 순으로 각각 성능 검증을 수행하였고, 그 결과는 그림4와 그림5와 같다.

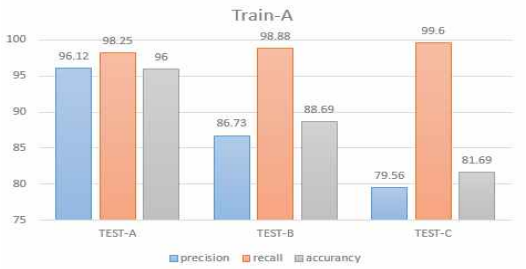

그림 4. Train-A (Packing 미적용) 결과

그림 5. Train-B (Packing 일부 적용)

Test-A에 대한 탐지 성능이 Train-A와 Train-B두 경우 비슷하게 나왔다. 하지만, Test-B와 Test-C에 대한 악성코드 탐지 성능은 Train-B로 학습한 경우 Train-A보다 Accuracy가 3%가량 높은 것을 확인할 수 있었다. 즉, 패킹을 일부 적용한 데이터를 학습한 인공지능 모델이 패킹된 악성코드를 탐지해내는데 보다 효과적이라고 평가되었다.

V. 결론

이 논문에서는 EMBER기반의 정적 정보를 활용하는 AI기반 모델에 패킹된 악성코드가 미치는 영향을 평가하였다. 악성코드가 패킹될 경우 탐지 성능이 떨어지는 것을 확인하였고, 패킹된 파일을 학습할 경우 정확도가 약간 올라가는 것을 확인하였다. 이 연구를 기반으로, 패킹된 악성코드의 정보를 고려한 AI기반 모델 설계에 대한 연구를 추진하고자 한다.

Acknowledgements

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2017R1A2B4012559). 이 논문은 정부(과학기술정토봉신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2019-0-01343, 융합보안핵심인재양성).

[참고문헌]

[1] Donghwi Shin, Chaetae Im, “The packer detection signature generation based on unpacking algorithm characteristic”, The Korean Institute of Information Scientists and Engineers 2010.6, 56-60(5 pages) , June, 2010.

[2] Sun-Kyun Kim, Hajin Kim, Mi-Jung Choi “Design and Implementation of Malware Automatic Unpacking System in Anti-VM/Debugging Environment”, The Journal of Korean Institute of Communications and Information Sciences 43(11), November, 2018.

[3] Hyrum S. Anderson, Phil Roth, "EMBER: An Open Dataset for Training Static PE Malware Machine Learning Models", Cryptography and Security (cs.CR) ,Aprill, 2018.

[4] Saxe, Joshua, and Konstantin Berlin. "Deep neural network based malware detection using two dimensional binary program features." Malicious and Unwanted Software (MALWARE), 2015 10th International Conference on. IEEE, 2015.