딥러닝 기반 악성코드 분석 프레임워크 설계

이정인*, 김지연**, 최은정*
*서울여자대학교 정보보호학과
**서울여자대학교 소프트웨어교육혁신센터
e-mail : jykim07@swu.ac.kr

Design of a Deep Learning-based Malware Analysis Framework

Jung-In Lee*, Jiyeon Kim**, Eun-Jung Choi*
*Dept of Information Security, Seoul Women's University
**Center for Software Educational Innovation, Seoul Women's University

1. 연구 필요성 및 문제점

사이버공격이 점점 진화하면서 변종 악성코드뿐 아니라, 알려지지 않은 취약점을 악용하는 제로데이(zero-day) 형태의 악성코드 또한 증가하고 있다. 기존의 시그니처(signature) 기반 백신 프로그램으로는 신·변종 악성코드를 탐지하는 데에 한계가 있기 때문에 최근에는 딥러닝(deep learning) 기반의 악성코드 탐지 연구가 활발히 수행되고 있다. 딥러닝 기반 악성코드 연구에서는 대량의 정상 및 악성코드 데이터를 학습하여 악성코드의 특징을 스스로 추출해내기 때문에 알려진 시그니처에 의존적인 전통적인 악성코드 탐지 기술보다 신·변종 악성코드 탐지에 효과적이다. [표 1]은 딥러닝 기반의 악성코드 연구를 수행한 관련 연구들의 핵심기술을 보여준다.

연구	제안된 핵심 기술	참고문헌
1	▪ 문자열 내 심볼 분석을 통한 악성코드 탐지 ▪ 심층신경망 기반 분석	[1]
2	▪ 실행파일 API 시퀀스 추출 및 이미지화 ▪ 합성곱신경망, Word2Vec 기반 분석	[2]
3	▪ Gumbel-Softmax를 활용한 전처리 ▪ 순환신경망 기반 분석	[3]

[표 1] 딥러닝 기반 악성코드 연구동향

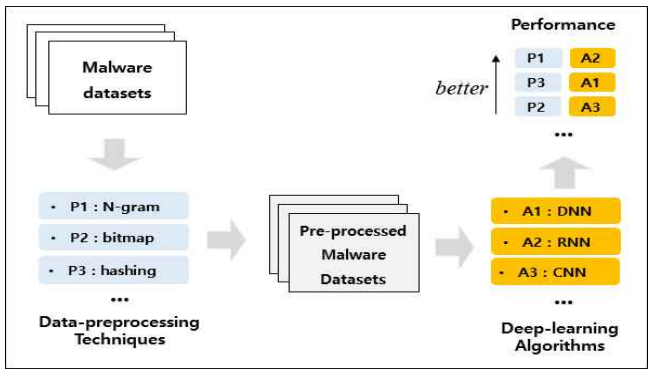
다양한 연구들에서 딥러닝 기반의 악성코드 연구를 수행하고 있지만, 악성코드 유형별로 최적의 데이터 전처리 기술 및 딥러닝 알고리즘 조합을 제시할 수 있는 프레임워크에 대한 연구는 미비한 실정이다. 본 연구는 딥러닝 기반 악성코드 연구 시 필요한 다양한 데이터 전처리 기술 및 딥러닝 알고리즘들을 탑재하고, 이를 활용하여 악성코드 데이터셋을 분석할 수 있는 프레임워크를 설계하고자 한다.

2. 연구내용과 방법

딥러닝 기반의 악성코드 분석을 위해서는 데이터 전처리 기술 및 딥러닝 알고리즘에 대한 연구가 선행되어야 한다. 데이터 전처리 기술은 악성코드 데이터셋을 어떤 방법으로 가공하여 딥러닝 알고리즘에 입력 값으로 넣을 것인지를 결정하는 기술로서 N-그램 기법, 비트맵 변환 기법, 피쳐 해싱(hashing) 기법이 대표적이다.

1) 이 논문은 2018년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No.2018R1D1A1B07050543)
2) 본 연구는 과학기술정통신부 및 정보통신기술진흥센터의 SW 중심대학지원사업의 연구결과로 수행되었음(2016-0-00022)

딥러닝 알고리즘으로는 시계열 데이터 분석에 적합한 순환신경망, 이미지 데이터 분석에 적합한 합성곱신경망, 그리고 앞의 두 알고리즘의 기반이 되는 심층신경망 등이 존재한다. 각 알고리즘 적용 시, 세부 신경망 모델설계 및 학습 파라미터 설정에 따라 정확도 및 속도가 달라지므로 최적의 성능을 얻기 위해 다양한 시나리오로 악성코드를 분석하는 것이 필요하다. 본 연구에서는 기존에 존재하는 다양한 데이터 전처리 기법 및 딥러닝 알고리즘을 적용하여 악성코드를 분석하고, 최적의 조합을 제시할 수 있는 프레임워크를 (그림 1)과 같이 설계하였다.



(그림 1) 딥러닝 기반 악성코드 분석 프레임워크 설계

3. 결론 및 향후 연구

본 논문에서는 딥러닝 기반의 악성코드 분석 연구 시 활용할 수 있는 프레임워크를 설계하였다. 제안된 프레임워크는 실제 악성코드 데이터셋을 탑재하여 딥러닝 기반으로 분석할 수 있도록 개발될 것이다. 개발된 프레임워크를 활용하면, 특정 악성코드에 대한 최적의 딥러닝 분석 방법을 제시할 수 있을 뿐 아니라, 악성코드 유형별로 최적의 전처리 기법 및 딥러닝 알고리즘 조합을 객관적인 탐지 성능을 기반으로 제시할 수 있을 것이다.

참고문헌

[1] 황선빈 외, "난독화된 악성코드 판별을 위한 2 차원 배열 기반의 기술 연구." 정보과학회논문지, 45권8호, 2018.
[2] 임태원, "실행파일 이미지화와 Word2Vec을 이용한 딥러닝 기반 악성코드 탐지 방법에 대한 연구", 2017.
[3] Hu Weiwei et el., "Black-box attacks against RNN based malware detection algorithms." Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence. 2018.