Deep Learning을 활용한 악성코드 탐지 방법 동향 분석

Analysis of Research Trend on Malware Detection Technique Utilizing Deep Learning

저자 (Authors)
이윤선, 이재웅, 장래영, 정성재, 성경, 소우영
Yoon-Seon Lee, Jae-Ung Lee, Rae-Young Jang, Sung-jae Jung, Kyung Sung, Woo-Young Soh

출처 (Source)
Proceedings of KIIT Conference , 2018.6, 166-169(4 pages)

발행처 (Publisher)
한국정보기술학회
Korean Institute of Information Technology

URL
http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07467643

APA Style
이윤선, 이재웅, 장래영, 정성재, 성경, 소우영 (2018). Deep Learning을 활용한 악성코드 탐지 방법 동향 분석. Proceedings of KIIT Conference, 166-169

Deep Learning을 활용한 악성코드 탐지 방법 동향 분석

이윤선*, 이재웅*, 장래영*, 정성재**, 성 경***, 소우영*

Analysis of Research Trend on Malware Detection Technique Utilizing Deep Learning

Yoon-Seon Lee*, Jae-Ung Lee*, Woo-Young Soh*, Rae-Young Jang*, Sung-jae Jung**, Kyung Sung***, and Woo-Young Soh*

요 약

최근 보안 동향에 따르면 신종 악성코드 보다 더 많은 기존 악성코드의 변종이 나타나고 있다. 이러한 악성코드들은 사용자의 데이터를 암호화하여 개인정보를 유출하거나 데이터를 삭제하고 금전적인 요구를 하는 등의 많은 피해를 입히고 있다. 이처럼 악의적인 목적을 가지고 급증하는 악성코드에 대응하기 위해 악성코드를 분석하기 위한 많은 연구가 진행되고 있지만 기존의 악성코드 분석 방법은 난독화, 가상 환경 우회 등의 분석 방해 기법으로 지능화된 악성코드의 등장으로 인해 분석에 많은 어려움을 겪고 있다. 최근 이러한 어려움을 극복하기 위해 기존 악성코드를 분석하여 학습하는 딥 러닝 방법이 각광을 받고 있다. 이에 따라 본 논문에서는 딥 러닝 기반의 악성코드를 분석하는 방법에 대해 소개하고 연구 동향을 살펴본다.

Abstract

According to recent security trends, there are more malware variants than new malware. These malware are damaging to the user by encrypting the user's data, leakage of personal information, deletion of data, and financial demands. Although many studies have been carried out to analyze malware in order to cope with malware that are rapidly increasing with malicious purpose, existing malware analysis methods are classified into malware such as obfuscation and virtual environment bypass, Which are difficult to analyze effectively. In order to overcome these difficulties, a deep learning method that analyzes existing malware and learns it is getting popular. Therefore, this paper introduces the method of analyzing malware based on deep learning and discusses research trends.

Key words
Deep Learning, Malware Detection, Malware Classification,

* 한남대학교 대학원 컴퓨터공학과
** (주) 엔버
*** 목원대학교 융합컴퓨터미디어 학부

Ⅰ. 서 론

악성코드는 사용자의 의사와 상관없이 컴퓨터의 데이터를 유출하거나 파일을 삭제하는 등의 악의적인 목적을 가지고 해를 끼치기 위해 제작된 모든 실행 가능한 형태의 소프트웨어를 말한다. 최근 10년간 이러한 악의적인 목적을 가진 악성코드가 크게 증가 하고 있다. 2017년에는 랜섬웨어의 일종인 wannacry가 세계적으로 30만대 이상의 컴퓨터를 감염시켜 악성코드에 대한 공포감이 확산되었다. Kaspersky Lab의 보고서에 따르면 16년 대비 17년의 신종 랜섬웨어의 수는 감소했지만 변종의 수는 약 2배 증가한 것으로 보고되었다[1]. 또한 최근의 악성코드들은 기존의 악성코드 분석을 회피하기 위해 지능화되어 가상 환경에서의 분석 과정을 탐지하고 악성행위를 중단시키거나 동작하지 않는 등의 분석 환경 회피기술이 적용되는 경우가 늘어나고 있다. 이러한 문제점을 해결하기 위해 여러 연구가 진행 중인데 그 중 기존의 수집된 악성코드 데이터들을 학습하여 변종 악성코드를 분류하고 탐지하는데 인공지능 머신 러닝을 적용하는 방법에 대한 연구가 꾸준하게 이루어지고 있다. 본 논문에서는 딥 러닝 기반의 악성코드를 분석하는 방법에 대한 동향을 소개한다.

Ⅱ. 배경 지식

2.1 악성코드 분석 방법

악성코드 분석은 방법에 따라 크게 정적 분석과 동적 분석으로 나눌 수 있다.

정적 분석은 디버거나 역공학 방법을 사용해 소스코드와 실행 없이 분석하는 방법으로 실행파일의 소스코드를 분석하여 프로그램의 기능을 파악하고 악성코드의 구조를 분석한다. 동적 분석에 비해 세부적인 분석이 가능하고 직접적으로 프로그램을 실행하지 않아 감염의 위험이 적다[2]. 그러나 자동화하기 어렵고 분석대상이 정상적 파일이라도 악성코드라고 분류되는 허위경보가 발생 할 수 있다.

동적 분석은 악성코드를 실행하여 발현되는 행위를 모니터링 함으로써 난독화, 패킹 등의 안티 디버거 영향을 받지 않는다. 신규 악성코드에 대한 탐지 가능성이 높고 분석 시간을 단축할 수 있다는 장점이 있지만, 예루살렘 바이러스와 같이 특정 조건에서만 발현되는 악성코드는 탐지하기 어렵다[2]. 동적 분석은 악성코드를 실제 실행하는 만큼 해당 프로그램이 동작을 하는 악성코드에 의해 사용자의 시스템을 감염시킬 수 있다. 그렇기 때문에 최근 많은 악성코드 분석가들은 가상머신을 기반으로 가상 분석환경을 구축한 후 동적 분석을 실행한다.

악성코드를 분석하는 방법에 따라 크게 두 개의 방법으로 나누었지만 최근의 악성코드 분석가들은 두 가지의 방법을 모두 사용한 하이브리드 분석을 활용하는 추세이다. 그림 1처럼 하이브리드 분석은 동적 분석과 정적 분석의 장점을 최대화 하고 각 분석 방식의 단점을 줄였다.

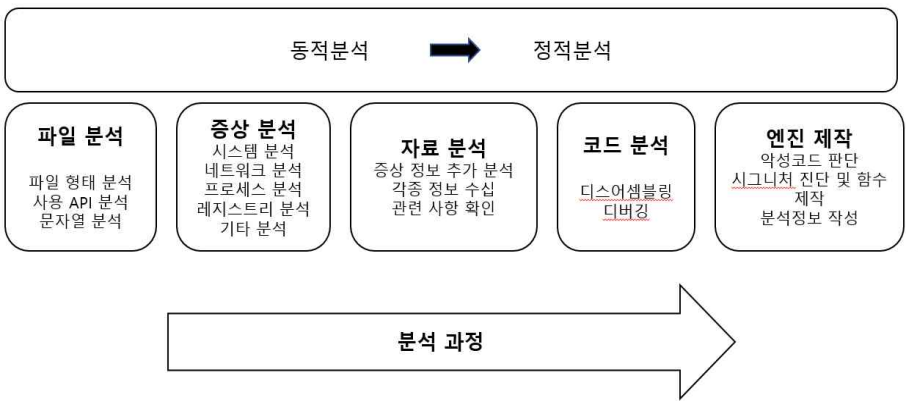

그림 1. 분석 과정의 예

2.2 딥 러닝

딥 러닝은 인간의 뇌를 모방한 머신 러닝의 한 종류로 심층학습이라고 불린다. 사람처럼 스스로 학습하는 것을 의미하는 딥 러닝은 과거 데이터 학습에 너무 많은 시간이 걸린다는 문제와 정확도가 낮다는 문제로 연구 주제에서 멀어졌었다. 하지만 최근 CPU의 개발과 GPU를 함께 사용하여 이러한 문제들을 해결했다. 또한 2012년 이미지 인식 알고리즘 대회 'ILSVRC'에서 딥 러닝의 일종인 CNN을 활용한 AlexNet이 84%의 정답률로 우승하며 딥 러닝은 다시 한 번 큰 관심을 받게 되었다. 딥 러닝의 구조는 그림 1과 같이 입력층에서 입력벡터를 입력값으로 받고 은닉층으로 주어지면 몇 개의 은닉층에서 학습을 한 후 출력층에서 결과 값을 받게 된다. 딥 러닝은 학습할 데이터가 많을수록 정확한 값을 도출해내는데 빅 데이터와 연산병렬처리를 가능케 하는 GPU의 발전으로 이를 가능하게 했다. 어떠한 데이터가 있을 때 이를 컴퓨터가 학습하여 결과를 내도록 하는 딥 러닝에 대한 연구가 계속해서 진행되고 있다.

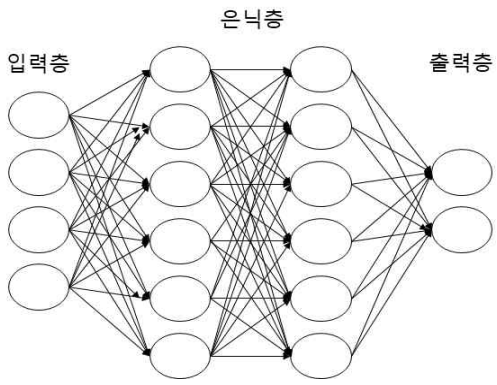

그림 2. 신경망 구조의 예

주로 이미지 인식과 음성 인식 등에서 사용되는 딥 러닝의 한 종류인 CNN(Convolutional Neural Network)은 이미지 인식 알고리즘 대회 'ILSVRC'에서 2012년 캐나다 토론토 대학팀에서 CNN을 사용하여 정답률 84%를 기록하며 압도적인 성능으로 우승한 후 딥 러닝을 연구하는 이들에게 많은 관심을 받고 있다. CNN은 입력 데이터와 필터의 값을 적용하여 서로 대응하는 원소끼리 곱한 후 그 총합을 구하는 방법으로 CNN의 구조는 그림 3과 같다. CNN은 일반적으로 3종류의 계층으로 나뉜다. 완전연결계층으로 이루어져 입력데이터와 가중치를 대응시키는 Affine계층과 입력데이터를 필터를 통해 새로운 값을 갖게 하는 합성곱 연산이 이루어진 합성곱 계층 마지막으로 대상 영역의 평균을 구해주는 풀링 계층으로 구성되어 있다. CNN은 기존 방법들에 비해 오차율이 적고 빠른 학습이 가능하다는 장점을 갖고 있다.

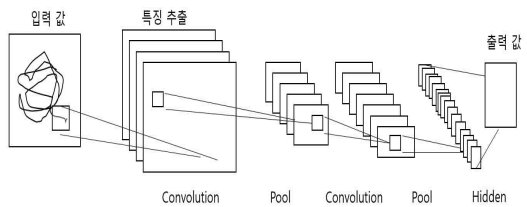

그림 3. CNN 구조

III. 딥 러닝을 활용한 악성코드 탐지 방법

본 장에서는 딥 러닝을 기반으로 악성코드를 탐지하는 연구를 소개한다.

그 중 첫 번째로 "Naive Bayes 기반 안드로이드 악성코드 분석 기술 연구" 에서는 머신 러닝 알고리즘인 Naive Bayes를 기반으로 악성코드를 분석하는 방법을 제안했다. API를 통계학적으로 접근해 정상 어플리케이션과 악성 어플리케이션을 판단하여 .dex파일의 method정보 영역에 접근 및 raw data를 추출하는 방법을 사용하였다. Naive Bayes 모델에서 검증한 결과는 85%의 정확도로 악성코드를 탐지하였다[3].

두 번째로 "악성코드로부터 빅데이터를 보호하기 위한 이미지 기반의 인공지능 딥 러닝 방법" 연구에서는 악성코드를 이미지화 하게 되면 변종 악성코드는 기존의 소스에서 일부분이 달라져도 비슷한 이미지가 형성된다는 특징을 이용하였다. 악성코드 바이너리 파일을 이미지화 하여 CNN을 통해 학습하고 분류한 결과는 91.7%의 정확도를 보이며 악성코드를 탐지하였다[4].

세 번째로 "Convolutional Neural Network 기반의

악성코드 이미지화를 통한 패밀리 분류" 연구에서도 악성코드를 실행시키지 않고 패밀리 분류를 하기 위해 8-bit gray-scale 이미지로 시각화 하였다[5]. 위 연구에서는 Microsoft 데이터 셋과 VXHeavens 데이터 셋 두 가지를 이용했다. Microsoft 데이터 셋은 9개의 패밀리 분류에 96.2%의 정확도를 VXHeavens 데이터 셋의 27개의 패밀리 분류에 정확도는 82.9%를 기록했다[5].

네 번째로 Decision Tree와 Decision Tree와 Entropy, Random Forest, SVM(Support Vector Machine) Kernel, Neural Network 등을 사용한 "A Study on the Machine Learning Algorithm for Mobile Malware Detection" 연구는 각 앱들의 특징을 나타내는 특성을 추출하여 181개의 API와 String정보를 특성으로 선정하였다. 또한 API와 String정보를 특성으로 선정하여 추출하였다. 이에 따라 평균 83%에 탐지율로 악성앱을 탐지하였다[6].

Ⅳ. 결 론

딥 러닝을 활용해 악성코드를 탐지하기 위해서는 딥 러닝 모델에 사용할 데이터를 추출해 낸 후 추출해낸 데이터를 가지고 딥 러닝 모델을 만들어 학습시켜야 한다. 이에 따라 악성코드 분석가들은 악성코드의 시그니처, api, opcode 등 다양한 데이터를 딥러닝으로 학습시키는 악성코드 탐지 기법을 연구하고 있다. 딥 러닝을 활용한 악성코드 분석은 기존의 분석 및 분류에 비해 악성코드들의 간단한 데이터 특징들로부터 복잡하고 어려운 특징들을 자동추출이 가능하고, 기하급수적으로 증가하고 있는 악성코드에 효율적으로 대처할 수 있다. 본 논문에서 분석한 딥 러닝을 활용한 악성코드 탐지 기법 동향으로 현재 연구되는 딥 러닝을 활용한 악성코드 탐지 기법을 발전시켜 다양한 악성코드 탐지에 적용할 수 있을 것으로 사료 된다.

참 고 문 헌

[1] 보안회사 - 카스퍼스키랩 보고서 2017년 12월 01일 보고서 http://news.kaspersky.co.kr/news2017/12n/171201.htm

[2] 사이토 고키, 밑바닥부터 시작하는 딥 러닝

[3] 황준호, 이태진, Naive Bayes 기반 안드로이드 악성코드 분석 기술 연구, 정보보호학회 논문지, 2017.10, pages 1087-1097.

[4] 김혜정, 윤은준, 악성코드로부터 빅데이터를 보호하기 위한 이미지 기반의 인공지능 딥 러닝 방법, 전자공학회 논문지, 2017.2 pages 76-82.

[5] 석선희, 김호원, Convolutional Neural Network 기반의 악성코드 이미지화를 통한 패밀리 분류, 정보보호학회 논문지, 2016.2, pages 197-208.

[6] 오성택, 고웅, 박준형, 모바일 악성앱 탐지를 위한 기계학습 알고리즘 연구, 한국정보과학회 학술발표논문집, 2017.12 pages 1111-1113.