Cuckoo Sandbox 기반의 악성코드 행위 분석 및 판별 연구

유은수, 하승철, 안은영, 김태훈■§, 방인규■

국립한밭대학교 지능미디어공학과, §국립한밭대학교 컴퓨터공학과

{20241103, scha}@edu.hanbat.ac.kr, {aey, thkim, ikbang}@hanbat.ac.kr

Cuckoo Sandbox-Based Malware Behavior Analysis and Classification

Eunsu You, Seungcheol Ha, Eunyoung Ahn, Taehoon Kim■§, Inkyu Bang■*

Dept. of Intelligence Media Engineering, Hanbat National University,

§Dept. of Computer Engineering, Hanbat National University

요 약

본 연구에서는 오픈 소스 기반 동적 분석 프레임워크인 Cuckoo Sandbox를 활용하여 악성코드의 행위를 체계적으로 분석하는 방법을 논의한다. Cuckoo Sandbox 기반의 분석 기법은 랜섬웨어(Ransomware)와 같이 은폐 및 회피 기술을 적극적으로 사용하는 악성코드의 공격 체인을 파악하는데 활용될 수 있으며 동적 분석을 통해 확보된 행위 기반 특징은 향후 보안 정책 수립 등에 활용될 수 있다.

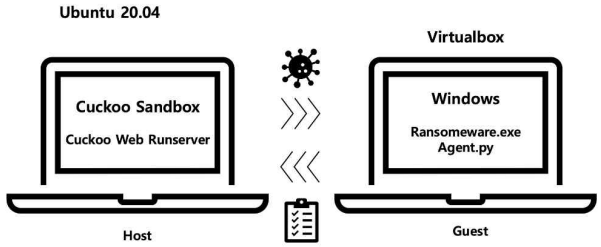
Ⅰ. 서론

사이버 보안 분야에서 랜섬웨어(Ransomware)는 기업, 정부 기관, 의료 기관 등 다양한 분야를 대상으로 심각한 위협을 초래하고 있다. 랜섬웨어는 '몸값'을 의미하는 랜섬(Ransom)과 소프트웨어(Ware)의 합성어로, 사용자의 데이터를 암호화해 인질로 삼아 금전을 요구하는 일종의 디지털 유괴 공격이다. 랜섬웨어 공격은 단순한 금전적 피해를 넘어서, 일상 업무 마비, 긴급 상황 대응 지연 등 사회적, 경제적 2차 피해를 초래할 수 있다. 랜섬웨어를 효과적으로 탐지하고 대응하기 위해서는 악성코드의 행동을 정밀하게 분석하는 과정이 필수적이다. 샌드박싱(Sandboxing)은 의심 파일을 실제 운영 환경과 유사한 격리된 네트워크에서 실행하여 안전하게 관찰하고 분석할 수 있는 기술로, 최근 보안 분야에서 널리 활용되고 있다[1].

본 논문에서는 오픈 소스 기반 동적 분석 플랫폼인 Cuckoo Sandbox를 활용하여 악성코드의 행위를 수집하고, 수집된 데이터를 기반으로 악성 여부를 점수화하는 방법을 분석하고 악정 행위 특성을 종합적으로 평가한다.

Ⅱ. Cuckoo Sandbox

본 연구에서는 랜섬웨어와 같은 악성코드 샘플의 동적 행위 분석을 위해 Cuckoo Sandbox를 활용한다. 그림 1은 Ubuntu와 Windows을 이용하는 Cuckoo Sandbox 설정 예시이다. 호스트(Host) 환경에서는 Ubuntu 20.04에 Cuckoo Sandbox를 설치하고 웹 서버(Cuckoo web runserver)를 구동하여 분석 작업을 관리·제어한다. 게스트(Guest) 환경에서는 Ubuntu 내부의 Virtualbox를 통해 Windows를 실행하고 실제 악성코드 샘플을 관찰한다.



Cuckoo Sandbox의 시스템 구조는 크게 분석 요청 관리자, 분석 에이전트(Agent), 호스트 분석 서버, 분석 결과 처리 및 보고의 네 가지 구성 요소로 이루어진다.

(1) 분석 요청 관리자: 악성코드 샘플, URL, 메모리 덤프 등의 분석 요청을 수집한 뒤 내부 큐(Queue)에 등록하고 분석 에이전트에 전달한다.

(2) 분석 에이전트: 격리된 가상머신 내에서 악성코드 샘플을 실행한다. 실행 과정에서 발생하는 시스템 호출, 파일 및 레지스트리 변경, 네트워크 트래픽 등을 실시간으로 관찰하고 기록한다.

(3) 호스트 분석 서버: 에이전트로부터 수집된 데이터를 중앙에서 통합하고 구조화하여 데이터베이스에 저장하고 최종 분석 보고서를 생성한다.

(4) 분석 결과 처리 및 보고: 수집된 로그 데이터 기반으로 악성 행위나 네트워크 패턴 등을 HTML, JSON, PDF 형식으로 제공한다.

위의 기능을 활용하여 Cuckcoo Sandbox는 게스트 Windows 환경에서 악성코드 샘플을 실행하고 파일 및 레지스트리 변경, 프로세스 인젝션, 네트워크 통신 등 다양한 악성 행위를 실시간으로 관찰하고 정략적으로 분석할 수 있다.

■ Corresponding Authors: Taehoon Kim and Inkyu Bang

Ⅲ. 실험 결과

1) 실험 환경

본 실험에서 활용하는 Cuckoo Sandbox 설정은 호스트(분석기)와 게스트(분석 대상) 간의 통신을 전제로 구성된다. 그림 2의 예시와 같이 Ubuntu에 설정된 IP, Cuckoo 내부의 resultserver가 바인딩하는 IP, 가상머신(Virtualbox)에서 사용하는 호스트 전용 네트워크의 IP, 그리고 Windows 게스트의 IP를 동일 네트워크의 주소로 설정해야 하며, 호스트 및 게스트의 통신이 방화벽에 차단되지 않도록 해당 포트의 사용을 허용해야 한다.

(a) Host IP

```
[resultserver]
ip = 192.168.219.180
```

(b) Guest IP

```
[cuckoo1]

platform = windows


ip = 192.168.219.125
```

그림 2. 호스트 및 게스트 환경의 IP 설정 예시

2) 분석 방법

본 실험에 사용된 악성코드 샘플은 Malware Bazaar에서 공개된 데이터를 활용했으며, 악성코드로 활용도가 높은 항목의 실행파일(exe 파일)을 내려받아 Cuckoo Sandbox로 분석하였다. 악성코드 샘플 실행을 통해 파일 행위, 프로세스 행위, 네트워크 통신 패턴, 외부 HTTP 통신, 파일 드랍, 프로세스 인젝션, 레지스트리 자동실행 등록 여부 등의 데이터를 수집할 수 있다. Python 기반 스크립트를 이용하여 수집 데이터를 정형화된 지표로 변환한 뒤 정상 샘플과 비교 및 통계 분석을 수행한다.

Cuckoo Sandbox는 기본적으로 웹 인터페이스에서 수동으로 report.json을 내려받아야 한다. 본 실험에서는 Python 코드를 활용하여 내부 저장소에서 최신 분석 결과를 자동으로 출력할 수 있도록 구현하였다[2].

3) 결과

그림 3은 본 실험에서 수행한 두 가지 Task에 대한 동적 분석 결과이다. 표 1에 정리된 5가지 특징을 바탕으로 분석 데이터의 악성 여부를 판단하였다. 본 연구에서는 Cuckoo Sandbox에서 제공하는 기본 점수 (info.score)를 기준으로 삼고, 안티디버깅(anti_debug), 드롭된 실행 파일(dropped_pe), 외부 네트워크 통신(external_http), 프로세스 인젝션(injection), 자동 실행 흔적(autorun)과 같은 주요 악성 행위가 관찰될 경우 각각 1~2점의 가중치를 부여하여 최종 점수를 산정한다. 최종 점수가 7.0 이상이면 악성, 4.0 이상 6.99 이하면 의심, 4.0 미만이면 정상으로 판정되도록 기준을 설정하였다. 이 기준을 적용한 결과, Task 93의 최종 점수는 2.80으로 산정되었다. 이는 주요 악성 행위가 전혀 관찰되지 않아 가중치가 추가되지 않았기 때문이며, 따라서 정상 파일로 분류되었다. 반면 Task 92는 여러 악성 행위가 동시에 관찰되었다. 이러한 요소들이 모두 가중치로 반영되면서 최종 점수가 10.00으로 누적되었고, 이에 따라 명확히 악성으로 판정되었다.

```
=== Task 93 결과 ===
기본 점수: 1.80  + 보너스 1.00 → 최종 2.80
판정: Benign
- anti_debug: no
- dropped_pe: no
- external_http: YES
- injection: no
- autorun: no

=== Task 92 결과 ===
기본 점수: 6.40  + 보너스 5.00 → 최종 10.00
판정: Malicious
- anti_debug: YES
- dropped_pe: YES
- external_http: YES
- injection: YES
- autorun: no
```

그림 3. 최종 분석 결과 예시

표 1. 분석 결과 상세 특징들

분석 특징	행위 설명
anti_debug	분석 회피 시도
dropped_pe	추가 악성 파일 생성 여부
external_http	외부 http 통신
injection	악성 코드 삽입 시도
autorism	자동 실행 등록

Ⅳ. 결론

본 연구에서는 Cuckoo Sandbox를 활용하여 악성코드의 행위를 동적으로 분석하고, 상세 특징을 기반으로 판별 결과를 수치화하여 논의하였다. 실험 결과를 통해 다양한 악성 행위의 특성과 점수 분포를 파악할 수 있었으며, 이를 통해 랜섬웨어 등 신종 악성코드에 대한 보다 정밀한 분석이 가능함을 확인하였다. 본 연구의 동적 분석을 통해 수집된 악성코드의 행위 데이터를 바탕으로 공격자가 보안 탐지를 우회할 때의 특성을 파악할 있으며, 이러한 분석 결과는 향후 보안 솔루션을 설계하거나 효과적인 방어 전략을 세우는 데 활용될 수 있을 것이다.

ACKNOWLEDGMENT

본 연구는 국립한밭대학교 공학교육혁신센터 “창의융합형 공학인재양성지원사업”, 2025년 과학기술정통신부 및 정보통신기획평가원의 SW중심대학사업의 지원으로 수행되었음 (2022-0-01068)

참 고 문 헌

[1] 전혜민, 최두섭, 임을규. (2025). 정적 특징 기반 랜섬웨어 탐지를 위한 특징 중요도 알고리즘 비교 및 특징 선정 연구. 정보처리학회 논문지 (KTSDE), 14(8), 576-587.

[2] 김지웅,정용규, "쿠쿠 샌드박스 기술을 이용한 악성코드 행위분석 기법," The Journal of Information Technology and Management, 1권, 1호, 16-22, 2013, .