Mozi Botnet의 분산 구조와 트래픽 특징에 기반한 YARA와 RNN의 통합적인 탐지 및 대응 시스템

권민아[1], 이정은[2], 여유림[3], 전성환[4], 유동영[5]
[1 2 3 4]홍익대학교 소프트웨어융합학과 학부생
[5]홍익대학교 소프트웨어융합학과 교수
alsdk871y@gmail.com, lee222kkk@naver.com, y020515@naver.com,
jsh12302183@gmail.com, ydy@hongik.ac.kr

An Integrated Detection and Response System Using YARA and RNN Based on the Distributed Structure and Traffic Patterns of the Mozi Botnet

Min-AH Kwon[1], Jung-Eun Lee[2], Yu-Rim Yoe[3], Sung-Hwan Jeon[4],
Dong-Young Yoo[5]
[1 2 3 4 5]Dept. of Software and Communications Engineering, Hongik University

요 약

이 연구에서는 IoT 보안을 강화하기 위해 Mozi 봇넷의 분산 구조와 트래픽 특징을 기반으로 YARA와 RNN을 통합한 탐지 및 대응 시스템을 제안한다. Mozi 봇넷의 분산 구조와 트래픽 특징을 분석한 후, 이를 기반으로 YARA 규칙과 RNN을 결합하여 악성 코드를 탐지하는 시스템을 설계한다. 실험 결과를 통해 이 시스템이 높은 정확도와 효율성을 보일 것으로 예상되며, 향후 연구에서는 다양한 딥러닝 기술을 활용하여 보다 효과적인 보안 대응 시스템을 개발할 것으로 기대된다.

1. 서론

사물인터넷 산업은 점점 성장하는 추세이며, 여전히 우리 사회에 자리 잡고 있다. 하지만 이를 공격하는 위협도 부상하고 있는데, 이에 대한 대응은 아직 취약한 상황이며, IoT 기기는 다양한 취약점을 가지고 있다. 사물인터넷을 이용한 봇넷은 2016년 mirai 봇넷이 최초 보고되었으며, 이후 공격 소스코드가 노출되어 여러 변종 봇넷이 발생하고 있다[1] 그 중 하나인 Mozi Botnet에 대해, 본 논문에서는 분산 구조와 트래픽 특징에 기반한 YARA와 RNN의 통합적인 탐지 및 대응 시스템 제안한다.

2. Mozi Botnet 분석

2.1 분산형

분산형 봇넷인 Mozi는 p2p(peer-to-peer) 프로토콜을 사용하여 여러 개의 분산된 컨트롤 서버를 제어하는 봇넷이다. 각 봇은 C&C서버와 통신하여 명령을 수신하고 작업을 수행한다. 분산된 컨트롤 서버를 사용하기 때문에 중앙 서버에 의한 단일 장악이 어렵다. 이는 봇넷의 탐지와 분석을 어렵게 만든다. 분산된 구조는 봇이 서로 통신하여 정보를 교환하고, 새로운 명령을 발견하는 등의 기능을 수행할 수 있도록 한다. 분산형 봇넷은 중앙 서버가 발견되더라도 전체 네트워크의 기능을 영향을 덜 받을 수 있다. 봇 마스터가 쉽게 탐지되어 서버가 종료될 수 있는 중앙집중형 봇넷과 다르게 탐지가 어렵다는 말이다.[2]

2.2 특정 포트 사용

Mozi 봇넷은 주로 다음과 같은 포트들을 이용하여 통신한다.

TCP 포트 8291: MikroTik 라우터의 Winbox 관리 도구에 대한 기본 포트로 알려져 있다. Mozi 봇넷은 이 포트를 이용하여 감염된 장치들을 제어하거나 명령을 전달한다.

UDP 포트 5555: Android 디바이스에서 사용되는 ADB (Android Debug Bridge) 서비스에 대한 포트이다. Mozi 봇넷은 이 포트를 이용하여 안드로이드 기기를 감염시키거나 제어한다.

2.3 DHT 프로토콜 통신

Mozi Botnet은 표준 프로토콜을 기반으로 하는 사용자 지정 확장 분산 해시 테이블(DHT) 프로토콜을 사용하여 구현된다.[3] DHT를 사용하여 감염된 호스트 간의 통신을 구성하는데, 이를 통해 여러 봇들이 서로 통신하고 명령 및 제어 정보를 교환할 수 있다. C&C(Command and Control) 서버와의 통신 대역으로 DHT를 사용하여 명령을 전달하고, 감염된 호스트의 동작을 제어한다. 각 봇은 DHT를 통해 다른 봇과 정보를 공유하고, 감염된 시스템에 대한 정보를 수집한다. Mozi Botnet은 노드 ID 반환과 구성 파일 반환이 랜덤하게 이루어지므로, 구성 파일을 동기화하기 위해서는 일반적인 DHT 노드보다 더 많은 이웃 리스트 요청을 송신해야 한다. 이는 정상 DHT 노드와 Mozi DHT 노드를 구분할 수 있는 중요한 특징으로, 현재 노드의 감염 여부 및 통신하는 상대 노드의 감염 여부를 파악할 수 있는 지표가 된다.[4]

3. YARA 규칙

YARA는 악성 코드나 악의적인 파일, 네트워크 트래픽 등을 탐지하기 위해 사용되는 패턴 매칭 도구이며, 악성 코드나 악의적인 행위를 탐지하는 데 사용된다. 이 규칙들은 간단한 텍스트 파일로 작성되며, 특정 문자열, 파일 특성, 패턴 등을 정의하여 악성 코드를 식별할 수 있다.

```
rule Mozi_Botnet {
    strings:
        $mozi_port = "8291"
        $mozi_udp_port = "5555"
        $mozi_signature = "Mozi"
        $dht_signature = "find_nodes" nocase
        $udp_protocol = "UDP" nocase
    condition:
        any of ($mozi_port*, $mozi_udp_port*,
        $mozi_signature*, $dht_signature, $udp_protocol)
}
```

그림 1 YARA 규칙 코드

(그림 1)은 Mozi 봇넷이 사용하는 특정 포트 (8291 및 5555) 및 특정 시그니처 (예: "Mozi")를 감지하는 YARA 규칙을 보여준다.

4. YARA규칙과 RNN의 병합

RNN은 순환 신경망으로, 순차적인 데이터를 처리하고 이전 상태의 정보를 기억하는 데 사용된다.[5] 본 연구에서는 YARA 규칙을 사용하여 악성 코드의 기본적인 패턴을 먼저 식별하고, 그 후 RNN 모델을 사용하여 보다 복잡한 패턴을 학습하고 탐지한다. 이는 시퀀스 데이터를 처리하고 이전 상태의 정보를 기억하여 패턴을 학습하는 데 사용된다. Mozi 봇넷의 특성을 고려할 때, RNN은 Mozi 봇넷의 트래픽 패턴을 학습하고 이를 기반으로 악성 코드를 탐지하는 데 사용될 수 있다. 특히, Mozi 노드가 더 많은 통신을 하고 특정 노드에 대해 더 많은 DHT 이웃 요청을 송신한다는 점을 고려할 때, RNN은 이러한 트래픽 패턴을 학습하여 Mozi 봇넷의 활동을 탐지하는 데 도움이 될 수 있다.

5. 결론 및 향후 연구

실험 결과를 통해 YARA 규칙과 RNN을 병합한 시스템이 높은 정확도와 효율성을 보이는지를 확인해야하며, Mozi 봇넷의 다양한 변종이 있을 수 있으므로, 여러 가지 트래픽 특징을 조사하여 규칙을 보강할 필요가 있다. Mozi트래픽의 특정한 포트 뿐만 아니라 특정 패킷 길이, node필드의 hash값 등이 필요하다. 이를 통해 YARA 규칙과 RNN을 병합한 시스템은 다양한 유형의 악성 코드를 탐지하는 데 뛰어난 성능을 보여줄 것이라 예상된다. YARA 규칙은 악성 코드의 기본적인 패턴을 빠르게 탐지할 수 있으며, RNN은 복잡한 패턴을 학습하여 정확도를 향상시킨다. 향후 많은 실제 데이터셋을 사용하여 시스템을 테스트하고 성능을 평가해야 하며, 더 나아가서는 다양한 딥러닝 기술을 활용하여 효과적인 탐지 및 대응 시스템을 구축하는 데 기여할 수 있을 것으로 기대된다.

참고문헌

[1] Stiv Kupchik, “Mirai-based NoaBot Emerges to Instantly Infect as soon as 'Hi' is seen”, Akamai, January 10, 2024.

[2] June Moon, “What does botnet mean and what are the types of botnet hacking?“, NordVPN, Mar 27, 2023.

[3] Sergiu Gatlan, “New Mozi P2P Botnet Takes Over Netgear, D-Link, Huawei Routers”, BleepingComputer, December 23, 2019.

[4] 김대현 외 4명, "DHT 프로토콜 트래픽을 활용한 Mozi 봇넷 탐지 모델에 관한 연구", 한국정보처리학회 학술대회논문집, 30권 1호, 147-148쪽, 2019.

[5] 김종화 외 2명, “순환신경망 모형을 활용한 시계열 비교예측”, 한국자료분석학회. 21권 4호, 1771-1779쪽, 2019.