엔드포인트의 네트워크 접속 기록을 활용한 공통 행위 탐지 기술 연구

서정훈, 염철민, 연성화, 박종상, 원유재*
충남대학교 컴퓨터공학과
e-mail : sjh9309@cnu.ac.kr, ymc12377@naver.com, yan243380606@gmail.com,
whdtkd546@gmail.com, yjwon@cnu.ac.kr
*Corresponding author

A Study of common behavior detection technology using endpoint's network connection record

Jeonghoon Seo, Cheolmin Yeom, Seonghwa Yeon, Jongsang Park, Yoojae Won *
Dept. of Computer Science, Chung-Nam University

요 약

금전적 이득을 극대화하기 위해 특정 기업, 기관을 대상으로 하는 공격이 증가하고 있다. 공격에 사용되는 일반적인 악성코드의 특징은 기존 시그니처 탐지 방법으로는 탐지하기 어렵다는 것과 공격자의 C&C 서버와의 통신이 일어난다는 점이다. 기업, 기관을 대상으로 이러한 악성코드를 이용한 공격이 시도된다면 감염된 모든 PC 로부터 공통적인 네트워크 접속 기록이 발견될 수 있다. 따라서 본 논문에서는 이러한 특징에 중점을 두고 라이브 포렌식 오픈 소스를 활용하여 엔드포인트의 네트워크 접속 기록을 활용해 공통 행위를 탐지하는 기법을 제시하고자 한다.

1. 서론

최근 악성코드는 불특정 다수를 대상으로 무차별적으로 대량 배포하는 기존의 전략과 달리 특정 목표에 최적화하고 점점 더 지능적으로 진화하고 있다. 또한 악성코드의 목적이 해커 본인의 기량 과시가 아닌 금전적 이득으로 변화되고 있다[1]. 이에 따라 기업, 기관에 대한 공격이 증가하고 있는 추세이며 2016 년 대기업 전산망 해킹 사건, 2017 년 국내 웹 호스팅 업체의 랜섬웨어 감염 사건, 2018 년 가상화폐 거래소 해킹 사건 등 매년 관련 사고가 발생하고 있다.

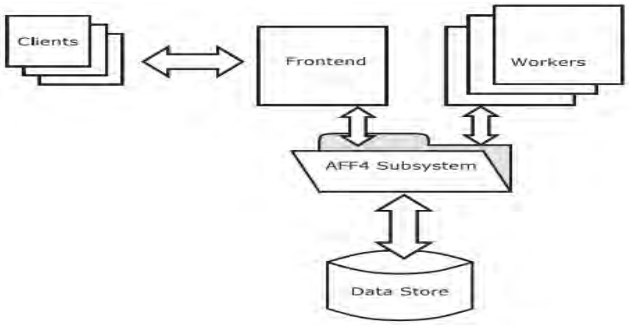
대표적으로 2016 년 대기업 전산망 해킹 사고에 사용된 악성코드는 Gh0st RAT 이라는 악성코드로 대상 PC 를 감염 시킨 후 공격자의 C&C 서버와 통신을 하여 키로깅, 원격 제어, 추가 다운로드 기능 등을 제공한다. 위와 같은 악성코드가 기업, 기관 환경을 감염시킨다면 감염된 PC 모두 공격자의 C&C 서버와 통신하게 되므로 이들로부터 공통적인 네트워크 행위가 발견될 수 있다.

또한 이러한 공격은 제로 데이 취약점을 이용하거나 변종 악성코드를 사용하는 경우가 많기 때문에 시그니처 기반 탐지 기법으로는 탐지에 한계가 있어 행위 기반 탐지의 필요성이 대두되고 있다[2].

따라서 본 논문에서 Google Rapid Response(이하 GRR)라는 오픈 소스 프레임워크를 통해 기업, 기관 환경 등에서 각 엔드포인트의 네트워크 접속 기록을 이용하여 공통적인 행위를 탐지할 수 있는 시스템을 제안하고자 한다.

2. GRR

GRR 은 Google 에서 개발한 침해사고 대응 원격 포렌식 도구로 다수의 PC 에서 동시에 아티팩트를 수집하는 기능을 제공한다.



(그림 1) GRR 의 구조

(그림 1)은 GRR 의 구조를 보여준다. GRR 은 크게 클라이언트, 프론트엔드 서버, 워커, AFF4 서브시스템, 데이터 스토어로 구성되어 있다. 클라이언트는 HTTP 프로토콜을 이용하여 프론트엔드 서버와 메시지를 교환한다. 프론트엔드 서버는 AFF4 서브시스템과 통신하여 데이터를 데이터 스토어에 저장한다. 워커는 데이터를 분석하거나 클라이언트에게 새로운 작업을 부여하기 위하여 AFF4 서브시스템과 통신한다[3].

워커가 단일 클라이언트로부터 아티팩트를 수집하

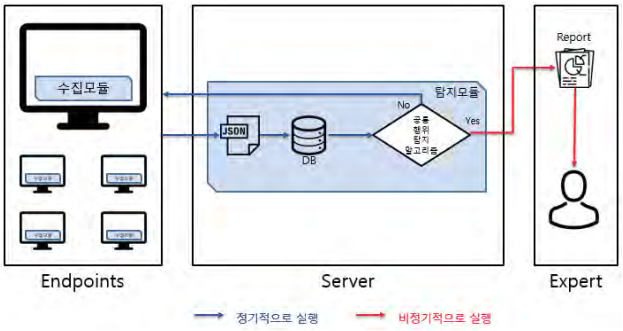
는 작업을 플로우라고 한다. 워커가 수집하고자 하는 아티팩트를 선택하여 실행 명령을 내리면 명령은 클라이언트에게 전달되고, 클라이언트는 해당하는 아티팩트를 추출하여 HTTP 프로토콜을 통해 프론트엔드 서버에 전송한다. 프론트엔드 서버는 수집한 아티팩트를 데이터스토어에 저장하고 워커는 데이터 스토어에서 아티팩트를 얻을 수 있다[4].

플로우는 단일 클라이언트로부터 아티팩트를 수집할 수 있는데, 플로우를 프론트엔드 서버와 연결된 모든 클라이언트에게 전송하여 그들로부터 아티팩트를 동시에 수집하는 기능을 헌트라고 한다[5].

3. 공통 행위 탐지 시스템

악성코드는 스팸 메일, 웹 방문 등 네트워크 활동을 통해 유포되는 경우가 많으며 대상 PC 를 감염 시킨 후에 공격자의 C&C 서버와 통신하여 명령 및 제어를 받는 경우가 많다[6]. 감염된 PC 들의 네트워크 활동 기록을 조사하면 악성코드의 유포 경로나 공격자의 C&C 서버와의 통신 기록이 공통적으로 나타날 수 있다. 만약 다수의 PC 가 같은 망에 존재하는 기업 환경에 위와 같은 악성코드를 이용한 공격이 시도되었다면 감염된 PC 들에서 악성코드 유포지 또는 공격자의 C&C 서버와의 공통적인 네트워크 통신이 감지될 것이다. 이는 공격을 탐지하고 빠른 대응을 위한 중요한 단서가 될 수 있다.

따라서 본 장에서는 GRR 을 통해 다수의 엔드포인트로부터 네트워크 접속 기록을 수집한 후 그들의 공통점을 탐지하여 해당 정보를 보안 전문가에게 제공하는 시스템을 제안한다.



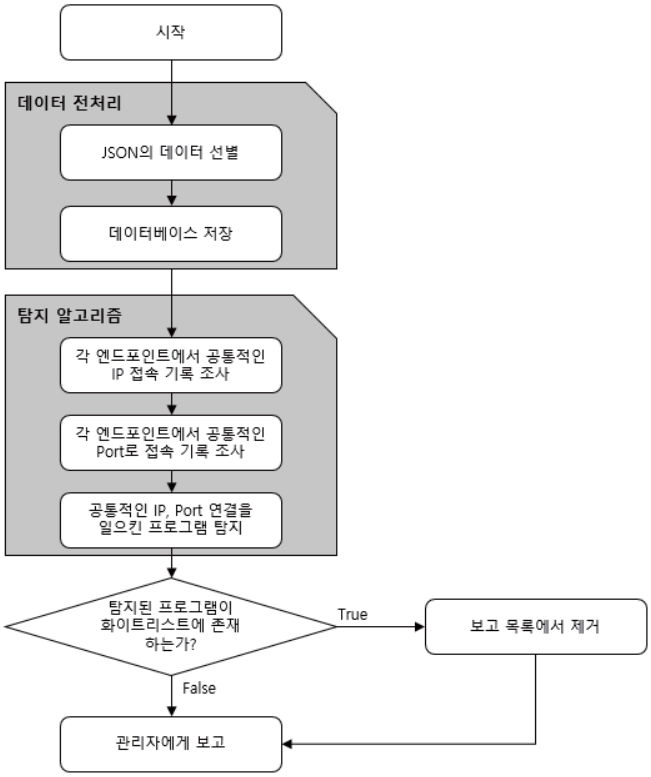
(그림 2)공통 행위 탐지 시스템의 동작 과정

(그림 2)는 제안하는 시스템의 동작 과정이다. 본 논문이 제안하는 시스템은 각 엔드포인트로부터 네트워크 접속 기록을 수집하는 수집 모듈과 데이터 전처리 후 알고리즘을 통해 공통 행위를 탐지하는 탐지 모듈로 구성된다. 탐지 모듈은 탐지된 공통 네트워크 행위의 주체 프로세스의 정보를 보안 전문가에게 제공하여 적절한 조치를 취할 수 있도록 한다.

3.1 수집 모듈

수집 모듈은 각 엔드포인트에 설치되며 정기적으로 각 엔드포인트 네트워크 접속 기록을 추출하여 서버로 전송하는 역할을 한다. 다수의 엔드포인트로부터 동시에 네트워크 접속 기록을 수집하기 위해서 GRR 의 헌트 기능을 사용한다. 수집 모듈의 결과물은 모든 엔드포인트의 네트워크 접속 기록 데이터를 가진 JSON 파일이며 .이를 탐지 모듈로 전송한다.

3.2 탐지 모듈



(그림 3) 탐지 모듈의 동작 흐름도

(그림 3)은 탐지 모듈의 동작 흐름도이다. 탐지 모듈은 데이터 전처리 과정과 알고리즘을 통한 탐지 과정으로 나눌 수 있다. 데이터 전처리 과정은 수집 모듈로부터 받은 JSON 파일을 파싱하여 데이터베이스에 저장하는 과정이다. 수집 모듈로부터 받은 JSON 파일은 각 엔드포인트 별로 네트워크와 관련된 여러 가지 데이터를 포함하고 있다. 그 중 값이 존재하지 않거나 공통 행위 탐지에 도움이 되지 않는 데이터들이 많이 존재한다. 전처리 과정은 이와 같은 데이터들을 제거하고 공통 행위 탐지에 도움이 되는 데이터들을 선정하여 데이터베이스에 보관한다. 데이터베이스에 저장되는 데이터는 전송 받은 클라이언트의 ID, 네트워크 행위를 하는 프로세스의 이름, Local IP, Local Port, Remote IP, Remote Port 이다.

데이터 전처리가 완료되면 데이터베이스에 저장된 데이터들을 이용하여 공통 행위를 탐지한다. 공통 행위는 전체 50% 이상의 엔드포인트에서 같은 IP 에 접속 기록이 존재하거나 또는 같은 Port 로 통신을 하는 것으로 정의한다. 이 조건에 맞는 공통 행위 중 발견된 엔드포인트의 수에 따라 비중을 다르게 두어 우선순위를 둔다.

탐지된 공통 행위의 주체 프로세스에는 각 엔드포인트가 실행한 프로세스와 시스템 프로세스가 모두 존재한다. 시스템 프로세스의 경우 시스템을 구성하

는 데 필수적이고 안전성이 검증되었기 때문에 모든 탐지 결과에 포함될 수밖에 없다. 안전성이 검증된 시스템 프로세스 같은 경우에는 탐지 대상에서 제외하는 것이 속도 향상 및 부하를 줄일 수 있는 방법이다. 이를 위해 관리자는 보고받은 공통 행위 중 시스템 프로세스를 구별하여 화이트리스트를 구성하고 이를 탐지 대상에 적용하는 과정이 필요하다.

4. 공통 행위 탐지 실험

4.1 실험 환경

공통 행위 탐지 실험을 위해 4 대의 엔드포인트를 이용한 실험 환경을 구성하였다. 각 엔드포인트는 1 대의 리얼 머신과 3 대의 가상 머신이며 운영체제는 모두 Windows10 으로 설정하였고 각 엔드포인트마다 수집모듈을 설치하였다. 또한 엔드포인트들과 연결된 Ubuntu 서버를 구축하고 탐지 모듈을 설치하였다.

4.2 실험 시나리오

우선 공통 행위를 확인하기 위해 인위적으로 각 엔드포인트에서 공통적으로 프로그램을 실행 시킨다. 실행하는 프로그램은 다운로드 폴더 내에 존재하는 파일을 특정 서버로 전송하는 client.exe, 원격 제어를 지원하는 TeamViewer.exe, 웹 브라우저인 Chrome.exe 이다. client.exe 는 실험을 위해 임의로 제작한 프로그램이다.

각 엔드포인트에서 프로그램을 실행하고 나면 수집 모듈을 가동시킨다. 수집 모듈은 헌트를 실행하여 각 엔드포인트로부터 네트워크 접속 기록을 추출한다. 수집 주기는 60 분으로 설정하였다.

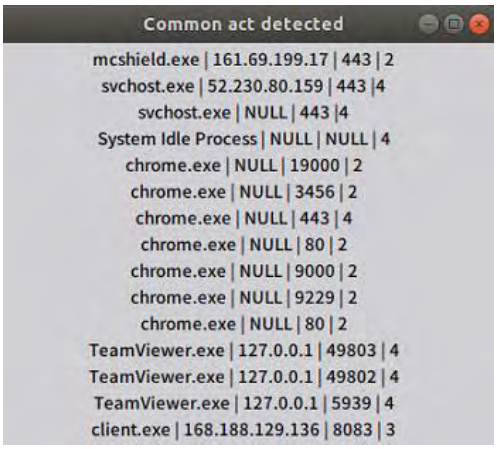
수집이 완료되면 데이터는 탐지 모듈로 전송되어 데이터베이스에 저장되고 이 데이터를 알고리즘에 입력하여 실행한 프로그램들이 공통 행위로 탐지되는지 확인해 보았다.

4.3 실험 결과

```
time : 09/14-18:38, new hunt H:73FDD745 is created
time : 09/14-19:38, download complete
time : 09/14-19:38, result save
time : 09/14-19:38, new hunt H:7D0D9AD7 is created
time : 09/14-20:38, download complete
time : 09/14-20:38, result save
time : 09/14-20:38, new hunt H:ACF66BDE is created
time : 09/14-21:38, download complete
time : 09/14-21:38, result save
time : 09/14-21:38, new hunt H:3F3401EA is created
time : 09/14-22:38, download complete
time : 09/14-22:38, result save
time : 09/14-22:38, new hunt H:891F3CBD is created
time : 09/14-23:38, download complete
time : 09/14-23:38, result save
time : 09/14-23:38, new hunt H:AEC1E77A is created
time : 09/15-00:38, download complete
time : 09/15-00:38, result save
```

(그림 4) 헌트 수행 화면

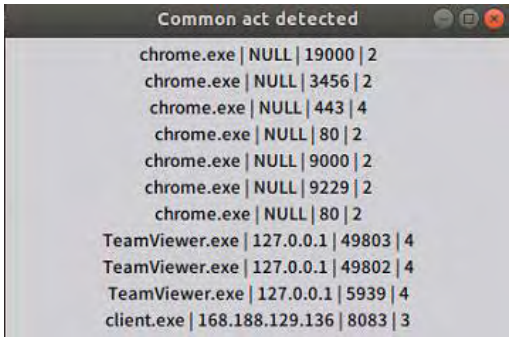
(그림 4)는 수집 모듈에서 주기적으로 헌트를 실행하여 각 엔드포인트로부터 네트워크 접속 기록을 수집하는 화면이다. 60 분을 주기로 헌트를 실행하고 데이터를 다운로드하여 데이터베이스에 저장하는 것을 확인 할 수 있다.



(그림 5) 화이트리스트 적용 전 결과 화면

(그림 5)는 화이트리스트 적용 전 결과 화면이다. 결과 화면은 각각 프로세스의 이름, 접속하는 IP 주소, 통신에 사용하는 Port 번호와 해당 행위를 하는 클라이언트의 수를 출력해주고 있다.

결과를 보면 각종 시스템 프로그램이 탐지된 모습을 볼 수 있다. 안전성이 검증된 시스템 프로그램의 경우 보안 전문가에게 보고 시 시간적 자원 낭비가 될 수 있다. 따라서 보안 전문가는 최초의 시스템 프로그램 탐지 시 안전성을 검증한 후 화이트리스트에 해당 프로그램을 추가하여 분석 과정에 시간을 절약할 수 있다. 하지만 화이트리스트는 공통 행위 탐지 시스템의 정확도를 떨어뜨리는 수 있는 요소이므로 신중히 운용할 필요가 있다.



(그림 6) 화이트 리스트 적용 후 결과 화면

(그림 6)은 화이트리스트 적용한 결과 화면이다. 시스템 프로세스가 화이트리스트를 통해 필터링 되고 사용자가 실행한 일반 프로세스들의 공통 행위가 탐지되는 것을 확인할 수 있다. 보안 전문가는 이 프로세스들의 행위의 악성 여부를 판단하여 차단 여부를 결정한다.

5. 결론

본 논문에서는 다수를 대상으로 네트워크를 이용한 공격을 탐지하기 위한 첫 단계로 공통 행위 탐지 시스템에 대해 제안하였다.

제안하는 시스템은 오픈소스인 GRR 을 활용하여 다수의 PC 로부터 네트워크 접속 기록을 수집한다. 수집한 데이터는 전처리를 통해 데이터베이스에 저장하고 알고리즘을 통해 공통점을 찾아 전문가에게 보고한다.

이 시스템은 기업, 기관 등 다수의 엔드포인트를 감염시키고 C&C 서버를 이용하여 엔드포인트들을 조종하는 공격에 대해 감염된 엔드포인트들이 공격자의 C&C 서버와 통신하는 과정을 탐지하므로 빠른 탐지가 가능할 것으로 기대된다.

본 연구에서는 공통적인 네트워크 행위를 탐지하였지만 향후에는 추가적인 연구를 통해 단순 네트워크에 국한되지 않고 시스템 행위까지 포함할 수 있는 단서를 수집하여 공통 행위를 탐지할 수 있도록 할 예정이다. 또한 탐지한 프로세스에 대해 동적, 정적 분석을 진행하고 결과를 전문가에게 제공하여 전문가가 보고받은 공통 행위가 악성 행위인지 여부를 판단하는 데 도움이 되는 시스템을 연구할 예정이다.

[Acknowledgement]
"본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 중심대학지원사업의 연구결과로 수행되었음(2015-0-00930)"

참고문헌

[1] Manuel Egele, Theodoor Scholte, Engin Kirda, Christopher Kruegel. "A Survey on Automated Dynamic Malware-Analysis Techniques and Tools"

[2] Choi Bo Min, Kang Hong Koo, Lee Tae Jin. "A Study on the Method Variants Prediction using Malware API Information"

[3] Flavio Cruz a, Andreas Moser, Michael Cohen. "A scalable file based data store for forensic analysis"

[4] M.I. Cohen, D. Bilby, G. Caronni. "Distributed forensics and incident response in the enterprise"

[5] Andreas Moser, Michael I. Cohen. "Hunting in the enterprise: Forensic triage and incident response"

[6] Kyungho Son, Taijin Lee, Dongho Won. "Design for Zombie PCs and APT Attack Detection beased on traffic analysis"