한국컴퓨터정보학회 하계학술대회 논문집 제19권 제2호 (2011. 6)

병렬 시그니처를 이용한 능동적인 해킹툴 대응방법

이세훈°, 전상표*, 한주영**, 신영진**, 박전진**
°인하공업전문대학 컴퓨터시스템과
*남서울대학교
**웰비아닷컴
e-mail: seihoon@inhatc.ac.kr, spjun7129@dreamwiz.com, neohan21@wellbia.com, codewiz@wellbia.com, apuro@wellbia.com

Active hacking tool countermeasure using parallel signature

Se-Hoon Lee°, Sang-Pyo Jeon*, Ju-Young Han**, Young-Jin Shin**, Jeon-Jin Park**
°Dept. of Computer Systems & Engineering, Inha Technical College
*Nam-Seoul University
**Wellbia.co, Co.,Ltd.

● 요 약 ●

본 논문에서는 해킹툴에 대해 선 발견, 후 차단할 수밖에 없는 보안 솔루션의 구조적인 한계를 극복하기 위해 시그니처의 역할에 대해 새로운 방식으로 접근한다. 특히 추적 시그니처를 이용하여 해킹툴을 능동적으로 대응하기 위한 방법을 모색하고 그 효과에 대해서 실험한다.

키워드: 시그니처(signature), 해킹툴(Hacking Tool), 해킹 방어(Hacking protection)

I. 서론

근래 점점 많은 정보가 디지털 화 되어감에 따라 이에 불법적인 목적으로 접근하고자 하는 시도도 지속적으로 늘어나고 있다.

바이러스, 해킹툴을 필두로 다양한 형태의 악성코드들이 하루에 수십, 수백 건씩 발견되고 있고, 이에 대응하기 위해 많은 시간과 노력이 필요한 것이 현재의 상황이다.

더구나 이미 노출이 되어있는 보안 프로그램의 특성상 보안위협을 사전에 차단하는 것은 쉽지 않다.

해킹툴은 이미 보안 프로그램을 회피한 상태로 유포되며, 이로 인한 문제가 확인된 이후에야 보안 솔루션은 그 대응이 가능하기 때문이다.

결국 수동적인 대응밖에 불가능한 구조적인 한계점을 극복하기 위해 해킹툴을 추적하여 능동적인 대응을 가능케 하는 방법을 모색하고자 한다.

II. 본론

일반적인 보안 솔루션들은 패턴을 기반으로 해킹툴을 진단한다. 새로운 공격이 발견될 때마다 솔루션 자체를 새로 빌드 하는 것보다 진단에 필요한 항목만 따로 분리하여 빌드 하는 것이 제작과 테스트, 배포 등에 유리하기 때문이다.

이러한 패턴은 해킹툴에 1:1 혹은 1:N으로 대응되는 진단 시그니처들의 집합으로 이루어져 있으며, 이 시그니처는 특정 해킹툴의 고유한 특징을 이야기 한다.

이 때 시그니처의 형태는 진단하고자 하는 대상의 형태에 따라, 각기 달라질 수 있다.

프로세스의 정보, 파일의 정보, 메모리에 적재된 실행코드의 정보뿐만이 아니라, 해킹툴이 사용하는 윈도우의 API 정보, 네트워크를 통해 전송하는 패킷의 정보 등의 행위정보, 대상까지 시그니처의 범위가 될 수 있다.

1. 시그니처가 가지는 한계점

보안 솔루션이란 검사 대상이 시그니처와 동일한 값을 포함하고 있는지 확인하는 프로그램이라고 정의할 수 있을 정도로 시그니처의 중요성은 매우 높다.

하지만 시그니처는 몇 가지 구조적인 한계점을 가지고 있다.

첫째, 해킹툴과 1:1로 매치되는 시그니처

진단 대상으로 하는 대상 해킹툴에는 효과적이다.

하지만 이를 회피하는 것이 상대적으로 쉽기 때문에 변종 해킹툴이 제작되는 주기가 짧다.

또한 해킹툴의 수량에 비례하여 진단 시그니처가 증가하게 되어, 패턴의 크기가 지속적으로 커지게 된다.

둘째, 해킹툴과 1:N으로 매치되는 시그니처

한 개의 시그니처로 비슷한 형태의 해킹툴을 다수 진단하도록 했을 경우, 실제로 수집하지 못한 해킹툴까지 진단이 가능하다는 장점과 변종 해킹툴의 제작을 어렵게 한다는 장점, 패턴의 크기를 상대적으로 줄일 수 있다는 장점을 가진다.

하지만 이러한 방법은 정상파일에 대한 오진 가능성이 높아진다는 치명적인 단점을 함께 가지고 있다.

또한 진단 시그니처를 회피하기 위한 해킹툴의 진화가 분석, 수집을 더 까다롭게 만들고 이는 대응 시간을 지연시키는 악순환을 유발시킨다는 점에도 주목해야 한다.

현재의 보안 솔루션들은 이러한 문제점을 해결하기 위해 두 가지 형태의 시그니처를 적절히 혼합하여 사용하고 있다.

하지만 비록 그 단점을 약화시켰을 뿐, 완전히 해소된 것은 아니며, 1:1로 매치되는 시그니처들을 1:N으로 변환하는 과정에서 유사 샘플의 필터링과 공통 시그니처 추출을 위한 작업이 중복으로 필요하다는 점, 이를 통한 결과물이 오진 등의 문제를 야기할 수 있다는 점을 유의해야 한다.

2. 병렬 시그니처의 정의

이에 병렬 시그니처라는 새로운 방식의 진단방법을 제안하고자 한다.

2.1 병렬 시그니처의 형태

병렬 시그니처는 1:1로 매치되는 시그니처와 1:N으로 매치되는 시그니처를 함께 사용한다.

이 때, 기존의 1:1로 매치되는 시그니처나 1:N으로 매치되는 시그니처가 모두 진단을 목적으로 하는 시그니처였던 것에 반하여, 병렬 시그니처에서는 1:1로 매치되는 시그니처를 진단 시그니처로, 1:N으로 매치되는 시그니처를 추적 시그니처로 사용한다는 차이점을 가진다.

즉 두 가지 형태의 시그니처를 모두 사용하되, 시그니처의 역할을 분리하여 그 장점을 모두 취하고자 함이다.

2.2 특이사항

병렬 시그니처의 경우 그 매치 결과를 분석자에게 전달하는 기능을 가진다.

진단정보를 솔루션의 제작사에 전달하고 이를 통계데이터로 활용하는 것은 일반적으로도 널리 사용되는 방법이다.

병렬 시그니처의 진단 시그니처 역시 통계데이터 생성을 위한 목적으로 그 진단결과를 이용한다.

하지만 추적 시그니처의 경우 단순 통계데이터를 위한 정보가 아닌 대응작업에 직접 활용이 가능한 데이터를 전달한다는 특징을 가진다.

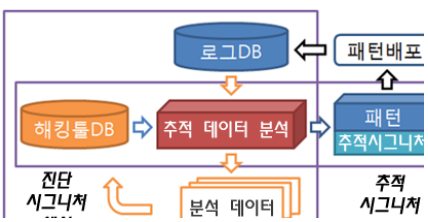


그림 1 추적 시그니처를 이용한 진단 시그니처 생성

3 병렬 시그니처의 구현

3.1 시그니처 기능

병렬 시그니처가 기존의 시그니처가 가지는 장점을 최대화 하면서 그 단점을 배제하기 위해서는 그 기능이 다음과 같이 정의되어야만 한다.

① 진단 시그니처와 추적 시그니처는 서로 다른 형태의 정보를 이용해야만 한다.

② 추적 시그니처의 동작여부를 외부에 노출시키지 않아야 하며, 분석자에게 전달하는 메시지는 반드시 암호화 시켜야만 한다.

③ 추적 시그니처는 그 정확도에 따라 복수의 시그니처를 동시에 사용할 수 있도록 한다.

3.2 추적 시그니처의 전달 메시지

추적 시그니처는 진단 시그니처로 차단되지 않는 해킹툴을 발견했을 때, 이를 즉시 진단하기 위한 목적으로 사용된다.

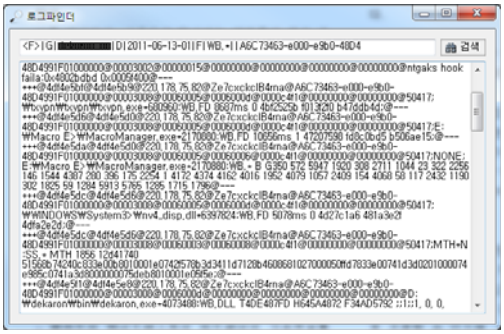


그림 2. 추적 시그니처에 의해 전달된 상태정보들

이를 위해 추적 시그니처는 분석자에게 네트워크를 통해 추적 샘플의 정보를 전달해야하며, 그 정보는 다음의 값들을 포함해야 한다.

① 발견 시스템의 고유정보

② 매치된 추적 시그니처

③ 진단 시그니처로 사용할 데이터

④ 기타 파일 정보

4. 병렬 시그니처의 효과

실제로 병렬 시그니처를 이용한 온라인게임 해킹툴 추적시스템을 구현하여 장시간 그 효과를 확인하였다.

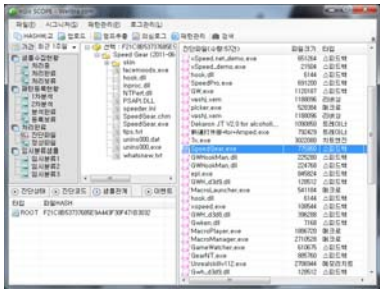


그림 3 추적 시그니처를 통해 발견한 샘플목록

4.1 솔루션 운영에 미치는 효과

① 능동적 대응이 가능하다.
② 미수집 샘플에 대한 대응력이 높아졌다.
③ 진단범위의 유연성
④ 해킹툴의 유사성 파악
⑤ 패턴 크기의 안정화

병렬 시그니처의 추적 시그니처는 오랜 시간동안 실제 운영 시스템에서 동작을 하면서 오진등의 발생여부에 대한 검증을 마친 시그니처이기 때문에 이를 진단 시그니처로 변경할 때에도 그 안전성에 대한 신뢰도는 높다고 볼 수 있다.

이처럼 추적 시그니처를 진단 시그니처로 변경하는 것은 패턴의 크기를 안정화시키는 데에도 도움을 줄 뿐만이 아니라, 솔루션 자체의 안전성에도 영향을 미치게 된다.

4.2 부가효과

병렬 시그니처는 능동적인 대응을 그 직접적인 목적으로 하지만, 시그니처들의 진단정보를 활용하여 다양한 부가효과를 이끌어 낼 수 있다.

① 진단 통계 정보

진단 시그니처의 매치정보를 이용하여 다양한 형태의 통계 데이터를 생성할 수 있다.

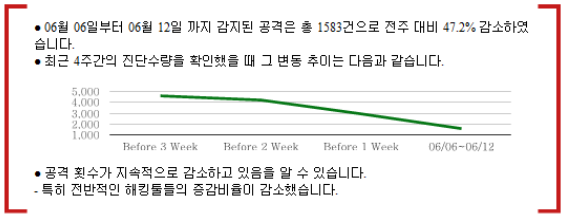


그림 5 진단추이를 이용한 해킹툴 동향

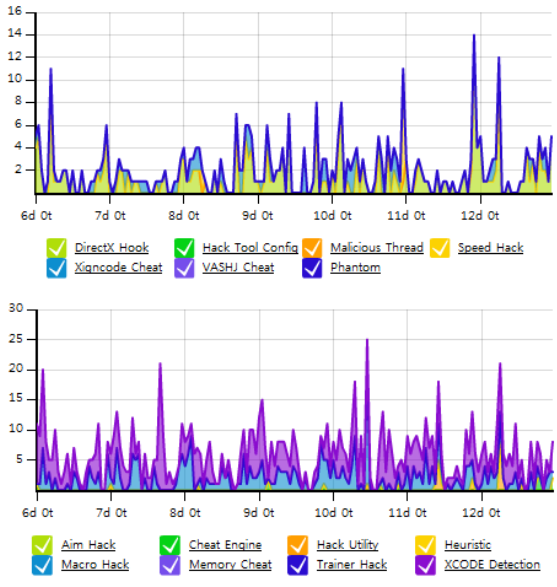


그림 6 진단 데이터를 이용한 다양한 통계 자료

② 추적 샘플의 매치결과 활용

병렬 시그니처가 동작 시스템의 고유정보를 구하고 이를 전송 데이터에 포함시키게 된다면, 최초로 변종 해킹툴이 실행되는 시스템을 알 수 있다.

이를 근거로 해킹툴의 제작 시스템의 정보를 획득하여 이들의 접근을 통제하거나 지속적인 모니터링을 통해 대응력을 높이는 데 사용할 수 있다.

③ 악성유저 모니터링

온라인게임 해킹툴의 경우 암암리에 유통이 되어 샘플의 수집이 어려운 경우가 있다.

또한 개인이 취미로 활용하는 것이 아닌, 온라인 게임의 게임머니를 획득하기 위해 기업적인 규모로 해킹툴을 사용하는 악성유저가 존재한다.

병렬 시그니처를 이용할 경우 이러한 악성유저에 대한 정보의 수집이 가능하며 이를 이용해 적절한 차단이 가능하다.

III. 결론

병렬 시그니처를 활용할 경우 보안위협에 대한 능동적인 대처가 가능하다. 실제로 구현하여 그 효과를 확인하는 동안 실제 샘플 수집 량의 2배 이상을 웃도는 수집력을 보여주어 그 효과 역시 나쁘지 않다고 판단이 된다.

하지만 현재의 논문에서 설명한 병렬 시그니처는 그 동작방식에 대한 이론을 설명한 것일 뿐, 실제 그 효과는 구현방법에 따라, 진단방법에 따라, 시그니처의 조합방법에 따라 틀려질 수밖에 없다.

또한 현재 제시한 방법은 능동적인 대응을 목적으로 한다. 이때 능동적인 대응은 변종 등의 새로운 해킹툴의 제작여부를 빠르게 판단하여 진단할 수 있게끔 하는 것을 이야기 하며, 이를 위해서는 적절한 자동화 시스템이 함께 구현되어야만 한다.

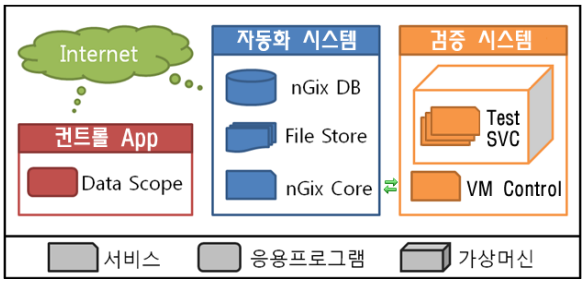
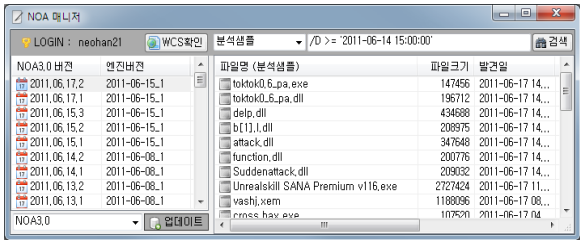


그림 7 자동화 시스템의 간략 구성도



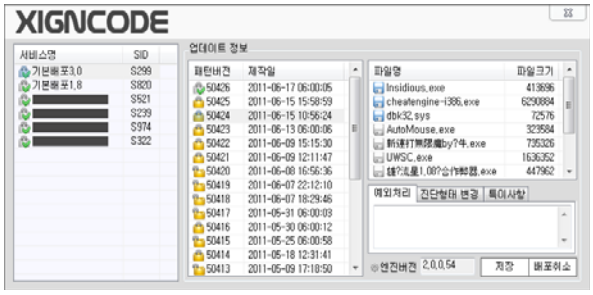


그림 8 패턴 생성/배포를 위한 자동화 프로그램

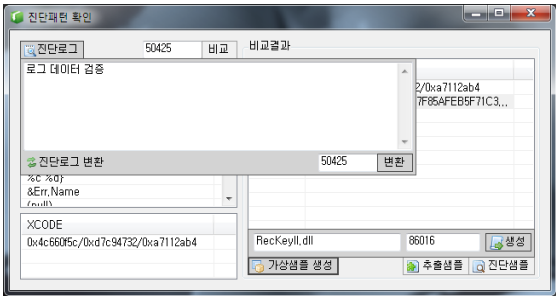


그림 9 데이터 검증을 위한 시스템

실제 병렬 시스템을 운영 중인 시스템에서 하루에 수집되는 로그의 양만해도 수십만 건에 이른다. 이 로그들을 텍스트파일로 저장했을 때, 그 크기가 기가바이트단위를 넘어가니 이를 분석자가 직접 처리하는 것은 무리가 따르는 것이 사실이다.

또한 수많은 데이터를 일일이 분석하고 진단데이터를 수집하고 또 패턴에 적용하는 것은 '능동적인 대응'이라는 목적에 반기하도 할뿐더러, 실수에 의한 각종 장애를 유발시킬 수 있다는 점을 염두에 두어야 한다.

하지만, 자동화 시스템을 구현하기에 앞서 반드시 아래의 사항에 대한 충분한 고민과 연구가 선행되어야만 한다.

① 시그니처의 높은 신뢰도

하나의 잘못된 시그니처는 심각한 오진을 불러일으킬 수 있다. 자동화 시스템은 대응이 빠른 장점만큼 오진의 전파속도도 빠르다는 이중성을 지닌다.

이 때문에 신뢰도가 높은 시그니처의 생성이 그 무엇보다도 중요하다.

② 철저한 검증 시스템

이 역시 오진에 대한 대비를 위해 반드시 필요하다. 다양한 환경을 컨트롤하여 오진을 대비해야 하며, 오진의 발견시 롤백 등의 후속조치에 손쉽도록 설계되어야만 한다.

③ 분석자의 적절한 개입

비록 많은 데이터를 처리하기 위한 자동화 시스템이지만, 적재적소에서 분석자가 시스템의 운영에 개입할 수 있어야 한다. 어떠한 훌륭한 알고리즘도 분석자의 판단력이 더해져야만 비로소 완벽해 진다는 점이 오랜 시간 자동화시스템의 운영을 통해 얻게된 교훈이다.

이처럼 자동화 시스템을 비롯하여 시그니처 추출등에 대한 충분한 연구가 병행된다면 병렬 시그니처가 보다 나은 보안 서비스를 구현하는데 큰 도움이 될 수 있을 것이다.

참고문헌

[1] 조성언, "온라인 게임에서의 보안 이슈들" 보안공학연구논문지 Vol.4 No.3, 7-14쪽, 2007년 8월

[2] 김동균, "온라인 게임 보안 강화, 게임 산업 분야별 현실과 방향 세미나, 2003.3

[3] 웰비아닷컴 특허(출원번호:1020080046659). "마이크로소프트 윈도우 운영체제를 사용하는 컴퓨터시스템에서의 악성코드 탐지 및 처리 시스템 및 방법"