바이너리 취약점의 분석 및 대응을 위한 자동화 기술 연구

김태은*, 장대일*, 손기종*, 김종기*, 나건배*, 김도원*
*한국인터넷진흥원
e-mail:tekim31@kisa.or.kr

A Study on Automation Technology for the Analysis and Response of Binary Vulnerabilities

Taeeun Kim*, Daeil Jang*, Gi-Jong Son*, Jong-Ki Kim*, Geon-Bae Na*, Dowon Kim*
*Korea Internet & Security Agency (KISA)

요 약

최근 사이버 공격은 자동화된 분석 및 해킹 기술의 발전으로 수많은 소프트웨어의 보안 취약점이 발표되고, 대응하지 못해 공격을 당하고 있다. 취약점 데이터베이스인 NVD(National Vulnerability Database)에는 2010년부터 2015년까지 보안 취약점(CVE: Common Vulnerability Enumeration) 약 8만 건이 등록되었으며, 빠르게 증가하고 있는 추세이다. 반면, 보안 취약점을 탐지하고 분석 및 대응하는 방법은 전문 분석가의 수동 분석에 의존하고 있어 대응 속도가 느리다. 이와 같은 현상을 해결하기 위해 기존에 알려진 보안 취약점의 영역은 자동화된 방식의 분석 및 패치를 통해 악의적인 공격자에게 공격 기회를 줄 수 있는 보안 취약점을 사전에 대응 할 수 있는 기술이 필요하다. 본 논문에서는 소스코드가 없는 레가시 환경의 바이너리를 대상으로 취약점 탐색을 위한 퍼징 기술 및 원인 분석, 바이너리 기반의 패치 생성 및 적용을 위한 시스템 구조와 고려사항 및 중요 성능 측정 대상에 대한 연구 내용을 소개한다.

1. 서론

최근 자동화된 분석 도구 및 기술의 발전으로 발견되는 취약점 수가 지속적으로 증가하고 있다. 취약점 정보를 공유하는 대표적인 DB인 CVE(Common Vulnerability Enumeration)에는 2010년부터 2015년까지 약 8만 건의 취약점이 등록 되었으며 2019년에만 1.1만 건 이상 등록되었다[1]. 또한 최근에는 자동화된 방식의 증가로 해결되지 않은 제로데이 취약점 건수가 늘어나고 있는 추세이다. 현재 보안 취약점에 대응하는 방식은 취약점 탐색을 위해 분석가가 직접 여러 가지 도구를 사용하여 코드를 분석하고 발견한 취약점을 패치 하고 있다. 그러나 소스코드 분석 및 취약점 탐색에 많은 시간이 소요되고 전문가에 따라 취약점을 분석하는 속도의 차이가 있으며, 오래된 운영 시스템에서 동작하는 소프트웨어는 소스코드가 존재하지 않는 경우가 많아 보안 취약점의 빠른 대응이 어려운 현실이다[2].

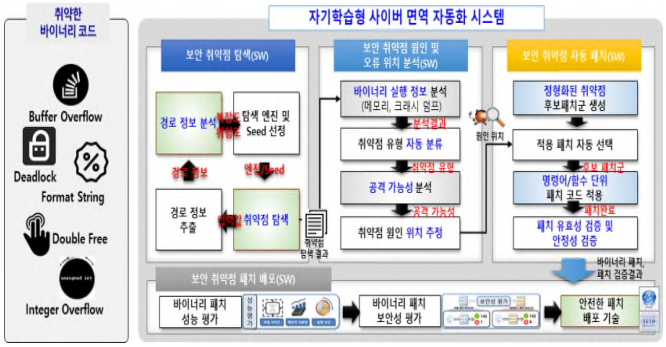
이러한 분석가의 기술 의존도를 해결하고 취약점 탐색에 들어가는 비용을 줄이기 위해 자동 취약점 탐색 및 분석 기술과 관련한 도구들이 등장하고 있다. 또한, 미국 DARPA에서 자동화 해킹 방어 대회인 Cyber Grand Challenge(CGC)를 개최하였다. 이 대회에서는 주로 취약점 탐색을 위한 퍼징 및 기호실행 기술과, 취약점을 자동으로 공격하기 위한 Exploit 생성 등의 기술을 자동화 한 기술이 활용되어 주목을 받았다. 퍼징(Fuzzing)은 대상 프로그램에 임의의 입력 값을 반복적으로 생성하고 응용 프로그램이 충돌을 발생하도록 하는 기술이다. 기호실행(Symbolic Execution)은 입력 값을 기호로 지정하고 프로그램의 취약점이 있는 경로에 도달할 수 있는 값을 찾는 방법이다. 그러나 CGC의 경우 보안 취약점을 찾아 공격을 실행하는 측면의 자동화 중점 이였다면, 보안 취약점 대응 기술은 취약점을 찾아 패치를 통해 시스템을 안전한 상태로 운영할 수 있게 하는 측면이 중요하다. 본 논문에서는 레가시 시스템에서 동작하는 바이너리의 보안 취약점 대응을 위해 필요한 요소 기술과 제안 기술이 구현된 개발 시스템의 동작 절차 설명한다.

2. 바이너리 취약점 자동 대응 기술

2.1 시스템 개요

제안 시스템은 리눅스 바이너리를 대상으로 보안 취약점을 자동으로 검출하고 발생한 기능 오류, 취약점의 발현지점에서 원인이 되는 위치를 추적하여 취약점의 유형을 식별하고, 바이너리 기반의 취약점을 패치하는 코드를 삽입하여 안전성

을 검증하는 사이버면역 자동화 기술이다. 제안 시스템의 기술 개요도는 [그림 1]과 같다.



[그림 1] 사이버면역 자동화 시스템 개요도

2.2 보안 취약점 자동 탐지 기술

바이너리 SW에 내재된 취약점, 기능오류를 외부 공격으로부터 사전 대응하기 위해서는 취약점을 찾아내는 과정인 탐색 과정을 선행해야한다. 제안 기술에서는 취약점 탐지를 위한 대상 바이너리를 2가지 분류로 나누어 각각에 특징에 맞는 탐색 전략을 수립한다.

[표 1] 취약점 탐지 대상

유형	입력	탐색 전략
바이너리	인자 입력	바이너리 위험도 기반 퍼징
	파일 입력	취약점 유발 파일 변이 퍼징

[표 1]과 같이 바이너리의 유형을 인자를 입력 받는 타입과, 파일을 입력 받는 2가지 형태로 분류하고, 인자를 입력 받는 바이너리는 보안 취약점을 빠르게 찾고 높은 커버리지를 확보하기 위해 바이너리 내에 포함된 취약요소(함수, 사용자 정의 변수)를 추출하여 위험도를 산정한 후 위험도가 높은 구간을 우선적으로 탐색하는 전략을 수립한다. 또한 파일을 입력 받는 바이너리는 취약점을 유발하는 입력 값이 파일이므로 파일의 구조(헤더, 바디 등)를 정의하고 취약점을 유발하는 코드가 삽입되는 영역을 변이하는 방식으로 기존의 방식보다 보안 취약점을 빠르게 많이 발견하는 전략을 활용한다.

2.3 보안 취약점 원인 분석 기술

보안 취약점 원인 분석 기술은 분석가가 도구를 사용하여 취약한 바이너리를 분석하는 과정을 자동화 하고자한 기술로 바이너리의 정보를 수집하고, 앞선 탐지 기술에서 탐지한 정보 등을 분석하는 과정을 거쳐 취약점의 유형을 파악하는 과정을 수행한다. 바이너리의 정보를 수집하는 과정은 정적, 동적 방식으로 명령어, 레지스터, 함수, 크래시, 변수정보 등을 추출하여 분석·식별하는 과정을 수행한다. 이렇게 수집한 실행정보는 탐지 기술에서 찾아낸 크래시 정보를 기반으로 크래시를 유발하는 값을 역추적 하고, 외부 입력 값을 분석하여 취약점의 유형을 식별 한 후 원인 정보를 생성한다.

2.4 바이너리 패치코드 자동 생성 및 적용 기술

앞선 과정에서 보안 취약점을 탐지하고 그 원인을 분석한 결과가 생성되면, 알려진 보안 취약점 정보를 기반으로 대응할 수 있는 패치 템플릿을 선정한다. 이때 생성·적용하는 패치는 소스코드 기반이 아닌 머신코드를 기반으로 바이너리에 직접 패치 하는 기술로 명령어, 함수 단위의 패치를 수행하며 바이너리가 동작하는 시스템의 런타임 메모리를 패치 하는 3가지 종류의 패치 방식을 선택적으로 적용한다.

이렇게 패치를 적용하고 난 후, 소스코드 기반의 패치를 적용한 결과보다는 조금 더 많은 검증이 필요하다. 현재 검증은 패치 성공 여부, 성능-기능성-보안성 3개 분야 검증 결과 스코어를 통해 패치를 적용한 바이너리의 유효성을 검증한다.

3. 결론

본 논문에서는 레가시 시스템 운영에 사용되는 소프트웨어(바이너리)의 내재된 취약점을 자동화 된 방식으로 탐지하고, 취약점에 원인 및 근원지를 추적 분석하여 취약점 유형별 패치를 적용하고 검증하는 과정을 자동화하는 기술을 제안하였다. 제안한 자동 취약점 분석, 대응 기술은 현재 전문 분석가 개인의 기술 역량에 따라 시간, 취약점 탐지 유무 등 결과에 많은 편차가 존재하며 분석할 수 있는 대상의 수가 한정적인 취약점 대응 문제점에 기여할 수 있으리라 생각된다. 추후에는 개발한 기술을 통해 실제 사용되는 바이너리의 취약점을 자동으로 탐지하고 대응하는 테스트를 수행하고, 일정 규모 이상의 소프트웨어에 적응하기 위한 코드 최적화 등의 고도화에 대한 연구를 할 예정이다.

ACKNOWLEDGEMENT

본 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2017-0-00184, 자기학습형 사이버 면역 기술 개발)

참고문헌

[1] U.S. National Vulnerability Database. Available online: http://cve.mitre.org/cve/ (accessed April 30, 2019).

[2] S.H. Oh, T.E. Kim, H.W. Kim, "Technology Analysis on Automatic Detection and Defense of SW Vulnerabilities", Journal of the Korea Academia-Industrial cooperation Society, Vol. 18, No. 11, pp. 94-103, 2017.