한국정보과학회
KOREAN INSTITUTE OF INFORMATION SCIENTISTS AND ENGINEERS
http://www.kiise.or.kr

Korea Computer Congress 2021

한국정보과학회 KCC2021

2021. 6. 23.(수) ~ 25.(금)

ICC 제주 & 씨에스호텔 & 온라인

Unimagined Futures: Healing and Recovery with K-SW

주요행사

6.23

[특별세션]
- Green AI를 위한 뉴로모픽 컴퓨팅 세션

[Top Conference 특별세션]

[워크샵]
- 디지털 헬스 플랫폼 워크샵
- 빅 데이터 기술의 산업 동향 워크샵
- 머신러닝을 활용한 네트워크 기술 워크샵
- 지능형 자율주행 최신기술 워크샵
- AI 기반 도시 교통 최적화 기술 교류 워크샵
- 자율성장 휴먼증강 인지컴퓨팅 기술 교류 워크샵

6.24

[기조강연]
- Gabriele Kotsis 회장(ACM)
- 박종현 부원장(ETRI)

[특별세션]
- 비대면 시대와 보안 세션
- IT 정책 세션

[특별콜로키움]
- SW · AI 교육 혁신

[워크샵]
- 휴먼증강을 위한 단일 감각 정보의 변환 및 대외 지각 검증 워크샵
- 차세대 AI 예타 후속방안 회의
- 지속가능한 디지털 경제 실현을 위한 기술 R&D 정책 워크샵
- 준지도학습 과제 3차년도 워크샵
- 암흑데이터극한활용연구센터 하계 워크샵
- 여성연구자들이 함께 할 K-SW 미래 워크샵

6.25

[특별세션]
- 추천시스템 최신 기술 동향 세션
- 신진연구자 세션

[특별콜로키움]
- 초우수 SW 육성

[워크샵]
- 제주대 SW융합 컨퍼런스
- 인공지능연구소 기술 융합, 자율주행서비스 전략 워크샵
- 현대자동차의 데이터와 AI기술 워크샵

후원 OKESTRO NAVER Jeju CVB Jeju Convention & Visitors Bureau (사)제주컨벤션뷰로 LINE NHN SAMSUNG 삼성SDS SUNJESOFT SUNJESOFT Inc. HYUNDAI kakao Microsoft DMOA DREAMMARK1 TmaxSoft (주)티맥스소프트 Gridwiz SK하이닉스 쌍용정보통신 CUBRID AIIA 지능정보산업협회 Artificial Intelligence Industry Association 올포랜드 We Believe INFRANICS SK broadband ToBeCom 주식회사 투비콤 LG히다찌

MITRE ATT&CK™ 기반 익스플로잇 코드 아카이빙 자동화 시스템 설계

주재경[1] 박기웅[2†]
[1]세종대학교 시스템보안연구실, [2†]세종대학교 정보보호학과
jj01020905906@gmail.com, woongbak@sejong.ac.kr

Design of Exploit Code Archiving Automation System based on MITRE ATT&CK™

JaeGyeong Ju[1] Ki-Woong Park[2†]
[1]Sejong Univ, Syscore Lab.
[2]Sejong Univ, of Computer and Information Security

요 약

공격자의 익스플로잇에 대응하기 위한 보안 취약점을 식별 및 관리하는 다양한 보안 프레임워크가 존재하고, 해당 취약점을 테스트할 수 있는 익스플로잇 코드를 제공하기도 한다. 하지만 다양한 방어자 환경에 맞는 익스플로잇 코드를 확보하기 위해서는 지속적인 업데이트와 수많은 취약점을 다루는 프레임워크 간 구조를 파악하여 수동으로 찾아야 하므로 수집에 어려움이 따른다. 본 논문에서는 대표적인 보안 프레임워크인 MITRE ATT&CK™(Adversarial Tactics, Techniques and Common Knowledge)을 기반으로 CVE(Common Vulnerabilities and Exposures) 리스트에 따른 익스플로잇 코드를 자동으로 아카이빙하여 조건을 만족하는 익스플로잇 코드를 추출하는 시스템을 디자인한다. 이를 통해 지속적으로 업데이트되는 보안 프레임워크에 대응하며, 익스플로잇 코드 수집에 따른 워크로드를 최소화하고자 한다.

1. 서 론

보안 취약점을 식별 및 관리하는 다양한 보안 프레임워크가 존재한다. 특히, 보안 취약점 분류체계 중 하나인 MITRE ATT&CK™[1]과 취약점 정보를 식별 및 표준화하여 관리하는 CVE[2] 그리고 취약점 진단 기준 중 하나인 CWE[3] 등은 많은 곳에서 기반으로 삼고 있는 보안 표준 프레임워크라 할 수 있다.

현재까지 식별된 수많은 보안 취약점 공격에 대해 MITRE ATT&CK™은 14개의 전술 및 약 500개의 기술로 분류하여 관리하고 있으며, CVE는 약 15만개의 취약점을 리스팅하여 관련 정보를 제공하고 있다. 이와 같은 보안 프레임워크를 제공하는 이유는 취약점의 특성(공격방식, 동작 환경 및 조건 등)을 파악하여 그에 대한 방어 체계를 구축함으로써 보호 대상 시스템을 안전하게 보호하기 위함이다[4]. 또한, 단순히 취약점을 식별 및 분류하는데 그치지 않고, 해당 취약점을 테스트할 수 있는 익스플로잇 코드를 제공하는 Exploit-DB[5]와 Metasploit[6] 같은 프레임워크도 존재한다.

Exploit-DB의 경우 CVE 리스트와 매핑하여 취약점 정보에 해당하는 익스플로잇 코드를 제공하며 모의 침투 테스트에 활용하는 도구인 Metasploit에서 바로 테스팅 가능한 코드를 배포해 해당 취약점을 실제로 테스트 및 점검할 수 있다. 이 같은 취약점 점검을 통해 보호 대상 시스템에 대한 효과적인 개선방안을 마련하고 보안성과 안전성을 확보할 수 있다[7, 8].

하지만 보안 취약점은 점점 정교해지고 다양해짐에 따라[9], 방어자 환경에 맞는 익스플로잇 코드를 확보하기 위한 예를 들면 타겟 시스템(ex. Windows Server 2018, Ubuntu 18.04 LTS)에 해당하는 약점을 CWE를 통해 체크한 뒤 매핑된 CVE를 찾아서 타겟 시스템이 해당 취약점에 노출되어 있는지, 노출되어 있다면 대응방안은 어떻게 되는지에 대한 추가 정보는 Exploit-DB 등의 프레임워크 참고하는 등 프레임워크 간 구조를 파악할 필요가 있는데, 프레임워크 간 매핑을 위한 여러 접근법[10-12]은 있으나, 해당 환경에 부합하는 익스플로잇 코드를 확보하기 위해서는 결국 수동으로 해당 익스플로잇 코드를 찾아서 수집해야 한다는 한계점이 있다.

따라서 본 논문에서는 MITRE ATT&CK™를 기반으로 하여 취약점 정보를 리스팅한 CVE와 익스플로잇 코드를 제공하는 프레임워크(Exploit-DB, Metasploit 등)를 매핑하며, 자동화된 크롤러를 통해 익스플로잇 코드를 수집 및 관리하는 아카이빙 시스템을 제안한다. 본 논문에서 제안한 시스템은 2가지 특징이 있다. 첫째, 보안 취약점 공격을 전술과 기술에 대해 분류한 체계인 MITRE ATT&CK™을 기반으로 다양한 프레임워크로부터 익스플로잇 코드 수집을 위한 기준을 제공한다. 둘째, 자동화된 크롤러를 통해 익스플로잇 코드를 수집 및 관리하므로 수집에 따른 워크로드를 최소화한다.

---

† 교신저자: 박기웅(세종대학교 정보보호학과 교수)

* 이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원 (No.2019-0-00426, IoT 기반 이식-침습형 고위험 의료장치를 위한 능동형 킬 스위치 및 바이오 마커 활용 방어 시스템 개발,30%)과 2020년도 한국연구재단 연구과제의 지원을 받아 수행된 연구임(NRF-2020R1A2C400273,IoT 침해사고 대응을 위한 지능형 분석 플랫폼 기술 연구,70%)

본 논문의 구성은 다음과 같다. 먼저 2장에서는 본 논문에서 제안하는 익스플로잇 코드 아카이빙 자동화 시스템의 요구사항을 도출한다. 이어서 3장에서는 제안하는 시스템의 설계에 대하여 설명한다. 마지막으로 4장에는 결론 및 향후 연구를 기술한다.

2. 익스플로잇 코드 아카이빙 자동화 시스템 요구사항

본 장에서는 1장에서 제안한 MITRE ATT&CK™을 기반으로, 익스플로잇 코드를 크롤러를 통해 자동으로 수집 및 관리함에서 요구되는 사항에 대해 정리한다.

2.1 보안 프레임워크 간 매핑 관련 요구사항

일반적으로 보안 프레임워크에서 제공하는 보안 취약점 정보는 CVE ID(예:CVE-2020-0796)와 매핑하여 관리하고 있다. 하지만 MITRE ATT&CK™의 경우 보안 취약점 공격에 대한 전술과 기술에 대해 분류한 체계이므로 CVE 리스트와 명시적인 연결고리는 없으므로 조건에 부합하는 익스플로잇 코드를 불러오기 위해서는 다음과 같은 요구사항을 도출할 수 있다.

- MITRE ATT&CK™과 CVE 매핑을 위한 구조 분석 : 조건에 부합하는 익스플로잇 코드를 불러오기 위해서는 조건을 제시할 기준일 필요하다. 이에 따라 취약점 분류의 기준이 되는 MITRE ATT&CK™와 취약점 정보의 기준이 되는 CVE의 구조 분석이 필요하다.
- 확장성을 고려한 매핑 설계 : 점점 정교해지고 다양해지는 보안 취약점에 유연하게 대응하기 위해서는 기존에 고려하지 않은 새로운 분류 기준이 필요할 수 있으므로 새로운 분류 셋 추가의 가능성을 고려한 확장성 있는 매핑 구조 설계가 필요하다.

2.2 익스플로잇 코드 아카이빙 관련 요구사항

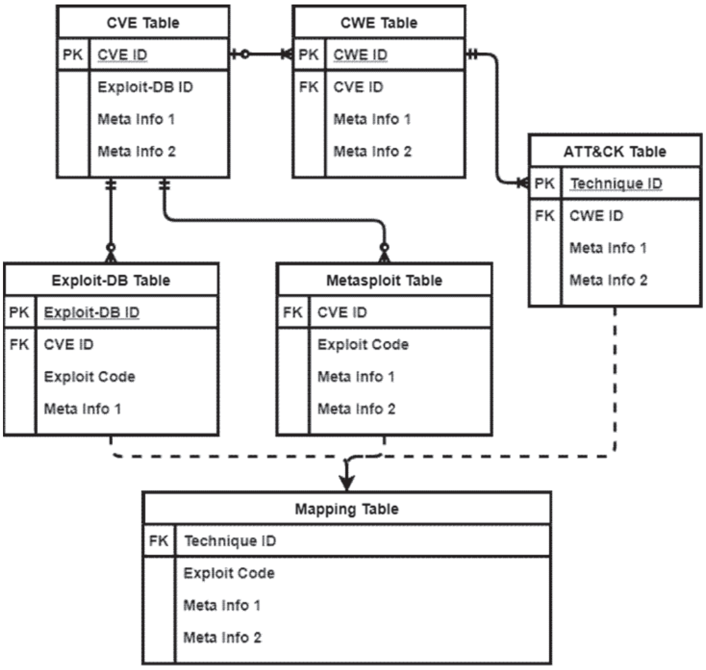
취약점 공격을 테스트하기 위해 제공되는 익스플로잇 코드의 경우 적용되는 플랫폼이 다양하므로 코드의 크기 또한 다양할 수밖에 없다. 따라서 아카이빙을 위한 데이터베이스 설계 관련하여 다음과 같은 요구사항을 도출할 수 있다.

- 프레임워크 간 관계와 익스플로잇 코드 사이즈를 고려한 데이터베이스 설계 : 대다수의 취약점 정보 제공 프레임워크의 경우 CVE ID를 기반으로 관계를 형성하고 있으므로 일반적으로 관계형 데이터베이스가 적합하지만 익스플로잇 코드 사이즈가 가변적인 점을 고려할 경우 NoSQL 타입의 데이터베이스가 효율적이므로 적절한 교차점이 필요하다.
- 크롤링 및 자동화를 고려한 설계 : 기본적으로 웹 기반 프레임워크에서 익스플로잇 코드를 크롤링하며 추후 프레임워크에 업데이트되는 익스플로잇 코드를 고려하여 자동화된 시스템 구축이 필요하다.

3. 익스플로잇 코드 아카이빙 자동화 시스템 설계

본 장에서는 분류한 취약점 별로 수집한 익스플로잇 코드를 자동으로 아카이빙 및 관리하는 시스템 설계를 위해 다음과 같이 두 부분으로 나누어 설명한다.

3.1 익스플로잇 코드 아카이빙 데이터베이스 설계

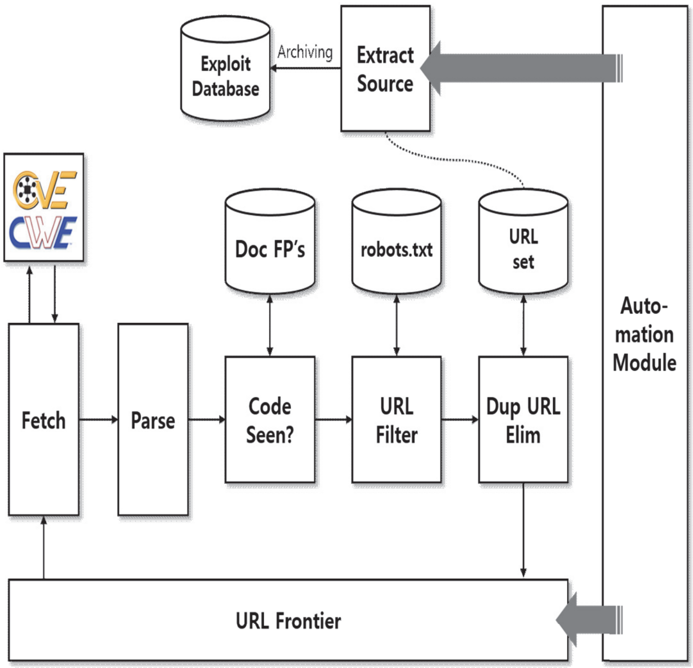


※ Meta Info 1, 2 is the additional assignment column
※ A dotted lines are to be determined according to future research

그림 1 익스플로잇 코드 아카이빙 데이터베이스 구조

본 절에서는 자동화된 크롤러를 통해 수집한 데이터를 저장 및 관리하기 위한 데이터베이스 구축을 위한 설계로 구조는 그림 1과 같다. 각 테이블은 CVE ID를 키로 하여 테이블을 구성한다. 그리고 익스플로잇 코드를 제공하는 프레임워크(Exploit-DB, Metasploit)에서 크롤링을 통해 수집한 데이터를 지정된(Exploit-DB, Metasploit) 테이블에 저장하고, CWE ID를 MITRE ATT&CK™에서 분류한 기술 ID와 매핑시키는 기준을 마련해 CWE 테이블과 ATT&CK 테이블을 매핑시킨다. 추가적으로 MITRE ATT&CK™의 기술 항목을 리스팅하는 테이블을 별도로 구성해 해당 취약점과 매칭되는 CVE ID, ATT&CK 기술 ID, 익스플잇 코드를 서로 연결하는 Mapping 테이블을 구성한다.

3.2 익스플로잇 코드 크롤러 설계



※ Doc FP's : Document Finger Prints, Dup URL Elim : Duplication URL Eliminate

그림 2 익스플로잇 코드 크롤러 구조

본 절은 익스플로잇 코드 크롤링을 위한 크롤러 설계로 구조는 그림 2와 같다. 웹 크롤러의 기본 아키텍처[13]를 기반으로 URL을 관리하는 URL Frontier에서 Fetch로 URL 수집을 요청하면 Fetch는 프레임워크(CVE,CWE)를 통해 해당 URL의 html을 요청하게 되고 가져온 html을 Parse로 보내 파싱하여 URL내 익스플로잇 코드를 수집하게 되는 일련의 구조를 따른다. 여기서 Doc FP's 는 해당 익스플로잇 코드의 중복 여부를 점검하고 robots.txt는 크롤링 정책을 점검하며 Dup URL Elim는 중복 URL을 검사하여 크롤링을 수행한다. 이를 기반으로 익스플로잇 코드를 추출 및 아카이빙하여 익스플로잇 아키이빙 데이터베이스의 해당 테이블에 자료를 저장한다. 앞에서 언급한 익스플로잇 코드 크롤링 및 데이터베이스에 저장하는 일련의 과정을 BeautifulSoup[14], Selenium[15] 등의 라이브러리를 활용해 자동으로 수행하여 익스플로잇 코드 수집에 따른 워크로드를 최소화하고자 한다. 추가로 해당 사이트에서 악성행위로 간주해 필터링되지 않도록 적절한 정책을 적용해 크롤러를 설계한다.

4. 결론 및 향후 연구

본 논문은 보안 취약점을 식별 및 관리하는 대표적인 보안 프레임워크를 통해 방어자 환경에 맞는 익스플로잇 코드를 확보하기 위한 익스플로잇 아카이빙 데이터베이스 및 크롤러를 설계하였다. 앞에서 살펴본 바와 같이 익스플로잇 아카아빙 자동화 시스템 설계 시 MITRE ATT&CK™의 분류체계를 기반으로 각 프레임워크 간의 매핑 관계를 도출하여야 최종적으로 해당 조건에 부합하는 익스플로잇 코드를 도출할 수 있다. 향후 연구로는 설계한 구조를 토대로 시스템을 구현 및 확장할 예정이다.

참고문헌

[1] https://attack.mitre.org
[2] https://cve.mitre.org
[3] https://cwe.mitre.org
[4] Ma, Qingxiong, Allen C. Johnston, and J. Michael Pearson, “Information security management objectives and practices: a parsimonious framework”, Information Management & Computer Security, 2008.
[5] https://www.exploit-db.com
[6] https://www.metasploit.com
[7] Tang, Andrew, “A guide to penetration testing”, Network Security, 8-11, 2014.
[8] 김종배. “모의 침투 테스트 방법 및 절차의 평가 방법에 관한 연구.“ 한국정보통신 종합학술대회 2014 춘계. Vol. 18. No. 1, 230-233, 2014.
[9] 과학기술정보통신부, “지능형 사이버공격 증가에 대비한 보안 강화 당부”, 2019.
[10] Legoy, Valentine, et al., “Automated Retrieval of ATT&CK Tactics and Techniques for Cyber Threat Reports”, arXiv preprint arXiv:2004.14322, 2020.
[11] Hemberg, Erik, et al. “BRON--Linking Attack Tactics, Techniques, and Patterns with Defensive Weaknesses, Vulnerabilities and Affected Platform Configurations”, arXiv preprint arXiv:2010.00533, 2020.
[12] Kwon, Roger, et al. “Cyber Threat Dictionary Using MITRE ATT&CK Matrix and NIST Cybersecurity Framework Mapping”, 2020 Resilience Week (RWS). IEEE, 2020.
[13] https://nlp.stanford.edu/IR-book/html/htmledition/crawler-architecture-1.html
[14] https://www.crummy.com/software/BeautifulSoup/
[15] https://selenium-python.readthedocs.io/