화이트햇 스쿨
WhiteHat School

팀 프로젝트 최종 발표 자료

팀 세스코

StaticAnalyze

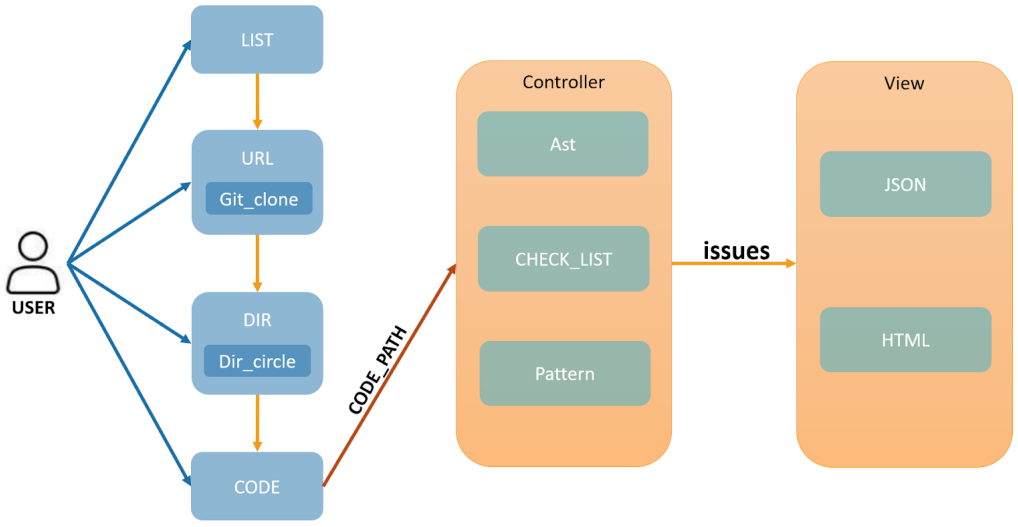
- 프로젝트 : 정적 분석기 제작
- 프로젝트 기간 : 2023.11.01 ~
- 개발 언어 : Python

팀명 : 세스코

Project Member

- Project Manager : 유재헌 SoteriaJ
- Project Advisor : 김진영 멘토님 PinguJace
- Project Leader: 홍승표 Phantomn
- 유건우 Ryu-GeonWoo
- 한동혁 OverDlive
- 여경민 GyeongminY
- 김동연 bbong0069
- 유선재 mameul
- 오윤석 lux-02

분석기 구조도



취약점 패턴화 목록

취약점 이름	CWE 넘버	취약점 설명
Incomplete regular expression for hostnames	**CWE-020**	악의적인 url로 리다이렉션되는 취약점
Path Traversal	**CWE-022**	사용자의 입력을 통해 파일 시스템에 접근하는 취약점
OS Command Injection	**CWE-078**	외부 입력을 통해 운영체제 명령을 실행하는 취약점
Improper Neutralization of Input During Web Page Generation	**CWE-079**	웹 페이지 생성 시 입력값이 적절히 중화되지 않아 발생하는 XSS 취약점
SQL Injection	**CWE-089**	데이터베이스 쿼리에 사용자 입력이 삽입되어 발생하는 취약점
Code Injection	**CWE-094**	악의적인 코드가 실행될 수 있도록 하는 취약점
Improper Encoding or Escaping of Output	**CWE-116**	출력 데이터가 적절하게 인코딩 또는 이스케이핑 되지 않는 취약점
Improper Output Neutralization for Logs	**CWE-117**	로그에 대한 출력이 적절하게 중화되지 않는 취약점
Information Exposure	**CWE-200**	정보 노출 취약점
Information Exposure Through an Error Message	**CWE-209**	에러 메시지를 통한 정보 노출 취약점
Flask in debug mode	**CWE-215**	디버그 정보를 통한 정보 노출 취약점
Hardcoded Password	**CWE-259**	소프트웨어에 하드코딩된 비밀번호 사용
Improper Privilege Management	**CWE-269**	잘못된 권한 관리로 인한 취약점

취약점 이름	CWE 넘버	취약점 설명
Improper Authorization	**CWE-285**	부적절한 인증 절차로 인한 취약점
Cleartext Storage of Sensitive Information	**CWE-312**	민감한 정보를 평문으로 저장하는 취약점
Inadequate Encryption Strength	**CWE-326**	부족한 암호화 강도
Use of a Broken or Risky Cryptographic Algorithm	**CWE-327**	취약하거나 위험한 암호화 알고리즘 사용
Insecure Temporary File	**CWE-377**	임시 파일이 보안에 취약한 방식으로 사용되는 경우.
Uncontrolled Resource Consumption	**CWE-400**	자원 소비를 제어하지 못해 발생하는 취약점
Untrusted Search Path	**CWE-426**	신뢰할 수 없는 검색 경로를 통한 라이브러리 로딩 등의 취약점
NULL Pointer Dereference	**CWE-476**	NULL 포인터 역참조로 인한 취약점
Deserialization of Untrusted Data	**CWE-502**	신뢰할 수 없는 데이터의 역직렬화로 인한 취약점
URL Redirection to Untrusted Site	**CWE-601**	신뢰할 수 없는 사이트로의 URL 리디렉션 취약점
Improper Restriction of XML External Entity Reference	**CWE-611**	XML 파서에서 외부 엔터티 참조를 제대로 제한하지 않아 발생하는 취약점
Insecure XPath Expression	**CWE-643**	보안이 취약한 XPath 표현식 사용
Improper Check or Handling of Exceptional Conditions	**CWE-703**	예외 상황에 대한 부적절한 체크나 처리.
Improper Restriction of Recursive Entity References in DTDs	**CWE-776**	DTD에서 재귀적 엔티티 참조에 대한 제한이 부적절한 경우.
Server-Side Request Forgery (SSRF)	**CWE-918**	서버 측 요청 위조 취약점
Improper Neutralization of Special Elements in Data Query Logic	**CWE-943**	데이터 쿼리 로직에서 특수 요소가 적절히 중화되지 않아 발생하는 취약점.

검출 대상 목록 (총 993개)

Repository 이름	Repository 설명
pandas	데이터 조작 및 분석을 위한 파이썬 라이브러리, 특히 테이블 형태의 데이터를 다루는 데 유용
youtube-dl	다양한 웹사이트에서 비디오를 다운로드하기 위한 커맨드라인 유틸리티
openpilot	Comma.ai에 의해 개발된 오픈소스 자동차 자율 주행 소프트웨어
whisper	고성능 음성 인식을 위한 오픈소스 소프트웨어
numpy	과학 계산을 위한 기본적인 라이브러리, 특히 다차원 배열을 처리하는 데 강력
transformers	자연어 처리를 위한 LLM 모델을 쉽게 사용할 수 있도록 하는 Hugging Face 라이브러리
langchain	자연어 처리를 위한 체인 기반 도구, 여러 AI 모델을 결합하여 새로운 애플리케이션 구축 가능
keras	딥러닝 모델을 쉽게 구축하고 훈련할 수 있는 고수준의 신경망 API
scrapy	웹 스크래핑 및 웹 크롤링을 위한 파이썬 프레임워크
pytorch	텐서 연산과 딥러닝을 위한 오픈소스 라이브러리
TTS	텍스트-음성 변환을 위한 딥러닝 기반의 오픈소스 소프트웨어
yolov5	빠르고 정확한 객체 탐지를 위한 딥러닝 모델
clova-cek-sdk	네이버의 CLOVA Extension Kit
ncloud-sdk	네이버 클라우드 플랫폼을 위한 SDK
pinpoint	네이버 대규모 분산 시스템을 위한 APM
egjs-view360	네이버 360도 이미지/비디오 뷰어
whale-browser	네이버 웨일 브라우저
toss-sdk	토스 결제 서비스 SDK
tossface	토스 얼굴 인식 기술
line-blockchain-developers-sdk	LINE 블록체인 개발자 SDK

패턴 검출 중간 결과

모든 레포지토리의 function name을 추출한 결과,

CWE 476, CWE 020, CWE 703 순으로 많이 발견된 것을 확인할 수 있었습니다.

- CWE 476 (Null Pointer Dereference): 9343개
- CWE 020 (Improper Input Validation): 4063개
- CWE 703 (Improper Check or Handling of Exceptional Conditions): 3343개
- CWE 117 (Improper Output Neutralization for Logs): 2466개
- CWE 079 (Cross-Site Scripting): 963개
- CWE 798 (Use of Hard-coded Credentials): 910개
- CWE 918 (Server-Side Request Forgery): 891개
- CWE 400 (Resource Exhaustion): 692개
- CWE 502 (Deserialization of Untrusted Data): 420개
- CWE 209 (Information Exposure Through an Error Message): 235개
- CWE 327 (Use of a Broken or Risky Cryptographic Algorithm): 227개
- CWE 078 (OS Command Injection): 172개
- CWE 776 (Improper Restriction of Recursive Entity References in DTDs): 108개
- CWE 094 (Improper Control of Generation of Code): 49개
- CWE 200 (Information Exposure): 30개
- CWE 116 (Improper Encoding or Escaping of Output): 26개
- CWE 377 (Insecure Temporary File): 18개
- CWE 215 (Information Exposure Through Debug Information): 8개
- CWE 312 (Cleartext Storage of Sensitive Information): 3개

국내 주요 빅테크 기업 레포지토리 분석

국내 주요 IT 기업들의 깃허브 내 모든 레포지토리를 대상으로 분석 진행하였습니다.

각 기업별 페이지를 클릭하시면 각 레포지토리 리스트, 검출 내용, 이슈 리포트를 확인할 수 있습니다.

Table

Aa 이름	전체 레포지토리 개수	검출된 이슈 발생 횟수
네이버(NAVER)	302	1452
카카오(KAKAO)	130	535
라인(LINE)	114	250
쿠팡(COUPANG)	18	116
우아한형제들(WOOWABROS)	22	7
당근마켓(DANGGN)	107	106
비바리퍼블리카(TOSS)	26	0

+ New

패턴 검출 최종 결과 예시

```python
import ast
import sys

def check_tarfile_import(node):
    if isinstance(node, ast.Import):
        for alias in node.names:
            if alias.name == 'tarfile':
                return True
    elif isinstance(node, ast.ImportFrom):
        if node.module == 'tarfile':
            for alias in node.names:
                if alias.name in ['extract', 'extractall']:
                    return True
    return False

def check_tarfile_method_call(node):
    if isinstance(node, ast.Call):
        if isinstance(node.func, ast.Attribute) and node.func.attr in ['extract', 'extractall']:
            return True
    return False

def detect_tarfile_issues(tree):
    issues = []

    for node in ast.walk(tree):
        if check_tarfile_import(node):
            issues.append({
                "line": node.lineno,
                "severity": "High",
                "content": "Unsafe tarfile module import detected",
                "url": "https://cwe.mitre.org/data/definitions/22.html"
            })
        elif check_tarfile_method_call(node):
            issues.append({
                "line": node.lineno,
                "severity": "High",
                "content": "Potentially unsafe tarfile extract/extractall call detected",
                "url": "https://cwe.mitre.org/data/definitions/22.html"
            })

    return issues
```

master ▾ calibre / setup / win-ci.py ↑ Top

Code Blame 106 lines (84 loc) · 2.66 KB Raw

```
3
4
5    import io
6    import os
7    import subprocess
8    import sys
9    import tarfile
10   import time
11
12
13   def printf(*args, **kw):
14       print(*args, **kw)
15       sys.stdout.flush()
16
17
18   def download_file(url):
19       from urllib.request import urlopen
20       count = 5
21       while count > 0:
22           count -= 1
23           try:
24               printf('Downloading', url)
25               return urlopen(url).read()
26           except Exception:
27               if count <= 0:
28                   raise
29               print('Download failed retrying...')
30               time.sleep(1)
31
32
33   def sw():
34       sw = os.environ['SW']
35       os.chdir(sw)
36       url = 'https://download.calibre-ebook.com/ci/calibre7/windows-64.tar.xz'
37       tarball = download_file(url)
38       with tarfile.open(fileobj=io.BytesIO(tarball)) as tf:
39           tf.extractall()
40       printf('Download complete')
41
```

clone_repo/calibre\setup\win-ci.py	win-ci.py	Cwe 022	9	High	Unsafe tarfile module import detected	https://cwe.mitre.org/data/definitions/22.html
clone_repo/calibre\setup\win-ci.py	win-ci.py	Cwe 022	39	High	Potentially unsafe tarfile extract/extractall call detected	https://cwe.mitre.org/data/definitions/22.html