51 난독화에 강한 AI 기반 웹 쉘 탐지

소속 정보컴퓨터공학부 분과 D 팀명 문지기 참여학생 문정윤, 구지원, 차기은 지도교수 최윤호

과제 개요

웹쉘이란 ?

웹쉘(WebShell)은 업로드 취약점을 통하여 시스템에 명령을 내릴 수 있는 코드를 말합니다. Webshell은 대부분 서버 스크립트 (PHP, ASP, JSP)로 만들어지며, 이 스크립트들은 웹 서버의 취약점을 통해 업로드 됩니다. 웹쉘이 서버에 업로드 될 시 해커들은 보안 시스템을 피하여 별도의 인증없이 시스템에 쉽게 접속이 가능합니다.

난독화란 ?

코드 난독화는 실행 가능한 코드를 수정하여 이해, 해석, 실행에 사용할 수 없도록 하는 것을 말한다. 공격자들은 시스템 관리자들이 웹쉘을 탐지하고 대응하는데 어려움을 겪도록 웹쉘에 난독화를 적용해 파일을 웹쉘이라고 판단하게 될 확률이 높아진다.

이 과제에서 우리는

난독화된 소스 코드를 비난독화하고 그 소스 코드로 부터 Opcode, AST 시퀀스를 생성하여 TextRank 알고리즘을 적용하여 특징을 추출한 뒤 RF, SVM, XGBoost 머신러닝 알고리즘을 활용하여 성능 평가를 수행한다. 최종적으로 해당 파일이 웹쉘 파일인지를 판단한다.

과제 내용

난독화 된 웹 쉘 파일을 **비 난독화** 시킵니다.

데이터 전처리

비난독화 된 파일에서 **OPCODE, AST** 추출한 후 특징 행렬을 만듭니다.

특징 추출

AI 모델이 **특징행렬을 바탕**으로 **웹 쉘 여부를 판단**합니다.

분류

웹 개발 도구

과제 결과

XGBoost	Precision	Recall	F1-score	Support
Normal file	0.98	0.93	0.95	355
Webshell file	0.88	0.96	0.92	191
Accuracy			0.94	546

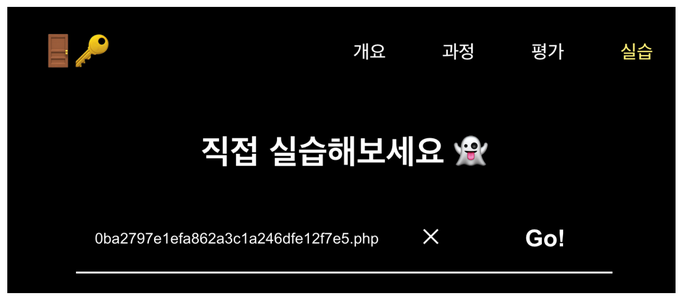
Normal과 Webshell 두 클래스의 파일에서 우수한 성능을 보였습니다.
특히 웹 쉘 탐지에서는 **정밀도(88%)**와 **재현율(96%)**이 균형 있게 높아, 거의 모든 악성 웹 쉘을 정확하게 탐지할 수 있었습니다.
두 클래스의 **F1-score도 95%, 92%**로 매우 높아, 다른 모델들에 비해 전반적으로 **가장 안정적인 성능을** 나타냈습니다.

Random Forest	Precision	Recall	F1-score	Support
Normal file	0.96	0.90	0.93	355
Webshell file	0.83	0.94	0.88	191
Accuracy			0.91	546

Random Forest

SVM	Precision	Recall	F1-score	Support
Normal file	0.84	0.97	0.90	355
Webshell file	0.91	0.65	0.76	191
Accuracy			0.86	546

SVM



원하는 PHP파일을 첨부하여 **웹쉘 여부**를 확인할 수 있습니다.

서버의 AI 모델이 **웹쉘 여부**를 판단합니다.