캡스톤 주제 제안서

교수님저희에이쁠 2 조

팀명: 교수님저희에이쁠 2 조

팀원:

소프트웨어학과 32207508 안석현

소프트웨어학과 32190393 김다은

소프트웨어학과 32184210 정지헌

발표일: 2023.03.16

목차

주제: Detection of Malware application through Machine Learning

1. 프로젝트 개요

1.1. 배경:

- 2017 년 7 월, google play store 에서 google play protect 기능이 포함되었다. protect 는 안드로이드 OS Ver 4.0(kitkat)이상, google play service Ver 11.5 이상에서 사용이 가능하다. 해당 기능은 악성 코드에 application 이 감염되어 있는 지를 확인후 google play store 에 등록되며, 이미 사용자 안드로이드 폰에 설치된 악성 application 을 스캔하여 제거해준다. 또한, 사용자의 개인정보를 훔치거나 사용자에게 기능을 속이는 앱을 감지하여 경고 메시지를 표시해준다. protect 는 머신러닝 기법과 application 사용 분석 기법을 사용하여 매일 500 억 개가 넘는 application 을 스캐닝 하고 있다. [1]
- "2018 년 한 해 동안 안드로이드 플랫폼에서 26 만개의 malware application sample 이 발견되었다." [2]
- 2022 년 11 월 1 일 malwarebyte 랩의 네이선 콜리어(Nathan Collier)는 google play store 에서 승인 받은 application 중 악성 행위를 하는 application 4 개에 대해 보고했다. 4 가지 application 모두 광고를 이용한 트로이 목마가 발견되었고, 악성 코드 탐지를 피하기 위해 일반적으로 사용되는 광고 지연 프로그램을 포함하고 있었다.[3]
- "안드로이드 시스템의 인기가 증가하면서 그에 따라, malware application 의 증가가 동반되고 있다." [4]
- "Android application 이 빠르게 발전하고 있지만, Android Malware Application 도 계속해서 발전하고 나타나고 있다." [5]
- Malware Applications 으로 부터 이용자를 보호하기 위해 카카오뱅크는 이용자 휴대폰에 설치된 Malware Application 이나 원격 제어 application 을 탐지해 피싱을 예방할 수 있게 하는 서비스를 제공한다고 한다.[6]

1.2. 목적 및 기대효과:

사람들은 google play store에서 많은 application을 사용하고 있다. 하지만, 이러한 application은 안전하지 않으며 완전한 보안 기법은 존재하지 않는다. 우리는 google에서 제공하는 protect 기능으로 보호받고 있지만, 머신러닝 기법을 사용한 Malware Application Detection Service를 제공하여 Malware Application을 자체적으로 탐지함으로써 공격자로부터 사용자를 더욱 안전하게 보호하는 것을 목표로 한다.

2. 사용 기술

2.1. 사용 언어:

- Python 및 기타 웹/앱 개발 언어

2.2. 데이터셋:

- APK from AndrooZoo
- 2019 ~ 2022 각 연도별 3000 개
  - Malware APK: 1500
  - Benign APK: 1500

2.3. 핵심 기술:

2.3.1. 머신러닝

- Random Forest
- Ada Boost
- LightGBM
- Decision Tree
- SVM
- Naive Bayes
- etc..

2.3.2. 성능 향상 기법

- K-fold
- Grid Search
- Voting System
- Model Stacking
- etc..

3. 서비스 구현:

- 웹, 앱 등으로 구현 예정

4. 진행 계획:

	3 월			4 월					5 월					6 월			
	3 주	4 주	5 주	1 주	2 주	3 주	4 주	5 주	1 주	2 주	3 주	4 주	5 주	1 주	2 주	3 주	4 주
기획																	
주제 선정																	
목적 및 기대효과																	
자료 찾기																	
시장조사																	
서비스 아키텍쳐 및 플로우																	
모델																	
알고리즘 선정																	
데이터 생성 및 전처리 알고리즘																	
모델 검증																	
정확도 향상																	
서비스 구현																	
UI/UX																	
디자인																	
주요 기능 구현																	
부가 기능 구현																	
유지보수																	
발표																	
최종 보고서 작성																	
질의응답 대비																	

5. 출처:

[1]https://source.android.com/static/docs/security/overview/reports/Google_Android_Security_2018_Report_Final.pdf

[2] AndroDialysis : Analysis of Android Intent Effectiveness in Malware Detection(2017 ELSEVIER)

[3] https://www.boannews.com/media/view.asp?idx=111385

[4] IPDroid : Android Malware Detection using Intents and Permissions(2020 IEEE)

[5] A Review of Android Malware Detection Approaches Based on Machine Learning(2020 IEEE)

[6] https://www.kakaobank.com/bank-story/234