



FACULTY
OF INFORMATICS

Masaryk University

Analýza inštalačných APK súborov pre OS Android

Bakalárska práca

Martin Styk

Ciele práce

- Vytvoriť rozsiahlu databázu inštalačných súborov pre OS Android (APK súborov)
- Popísať štruktúru a formát inštalačných súborov
- Vytvoriť nástroj na analýzu a získavanie metadát o inštalačných súboroch
- Určiť štatistické informácie o aplikáciách vo vytvorenej databáze
- Implementovať mechanizmus detekcie potenciálne modifikovaných APK súborov

Databáza inštalačných súborov

Uvažované zdroje APK súborov

- Oficiálne zdroje
 - Obchod s aplikáciami Google Play Store
- Neoficiálne zdroje
 - Neoficiálne obchody s aplikáciami
 - » SlideMe
 - » Amazon Appstore
 - Stránky na zdieľanie obsahu
 - » ZippyShare
 - » UlozTo
 - Torrenty

Databáza inštalačných súborov

Automatizácia hromadného preberania APK súborov

- Implementovaný nástroj ApkDownloader[2]
- Založený na princípe analýzy HTML kódu
 - Vyhľadanie priamych odkazov na APK súbory
 - Prevzatie APK súborov
- Implementovaný pre nasledujúce lokality
 - archív projektu Playdrone
 - www.appsapk.com
 - www.apkmaniafull.com
 - ww.androidapksfree.com
- Jednoducho rozšíriteľný, open source

Databáza inštalačných súborov

Výsledná databáza APK súborov

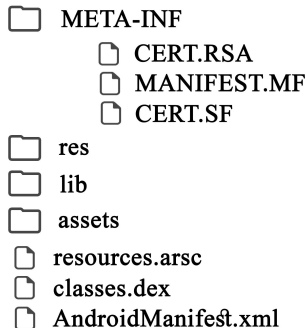
- Celkovo získaných 20 060 APK súborov, z toho viac ako 90 % pomocou ApkDownloader
- Celková veľkosť prebraných APK súborov 192 GB

| Zdroj | Počet aplikácií | % |
|-------------------------|-----------------|------|
| Playdrone | 8 200 | 40,9 |
| www.appsapk.com | 6 470 | 32,3 |
| www.apkmaniafull.com | 2 870 | 14,3 |
| www.androidapksfree.com | 1 030 | 5,1 |
| www.zippyshare.com | 750 | 3,7 |
| torrenty | 550 | 2,7 |
| www.uloz.to | 190 | 0,9 |
| Spolu | 20 060 | |

Analýza inštalačných APK súborov

Štruktúra a formát APK súborov

- Slúžia na distribúciu aplikácií na platforme Android
- Android application package file
- Štruktúra vychádza z formátu JAR
- Archívne súbory vo formáte ZIP
- Typická štruktúra s povinnými súbormi v pevne stanovenom formáte
- Niektore súbory v nečitateľnej skompilovanej podobe



Analýza APK súborov

Implementácia analýzy

- Cieľom získanie detailných metadát o rôznych aspektoch APK súborov
- Automatizované pomocou vyvinutej aplikácie ApkAnalyzer[3]
- Využitie nástroja ApkTool na dekompiláciu [4]
- Použitá knižnica AXML na konverziu binárnych XML súborov do čitateľnej formy
- Celkovo získaných 63 atribútov každého APK súboru

Analýza APK súborov

Príklady zbieraných informácií

- Základné informácie
 - veľkosť APK súboru, veľkosti dôležitých súborov
- Dáta zo súboru AndroidManifest.xml
 - komponenty aplikácie
 - prístupové oprávnenia
 - verzia Android SDK
- Informácie o certifikáte
 - algoritmus podpisu
- Informácie o zdrojových súboroch
 - formáty obrázkov
 - lokalizácia
- Hashe jednotlivých súborov

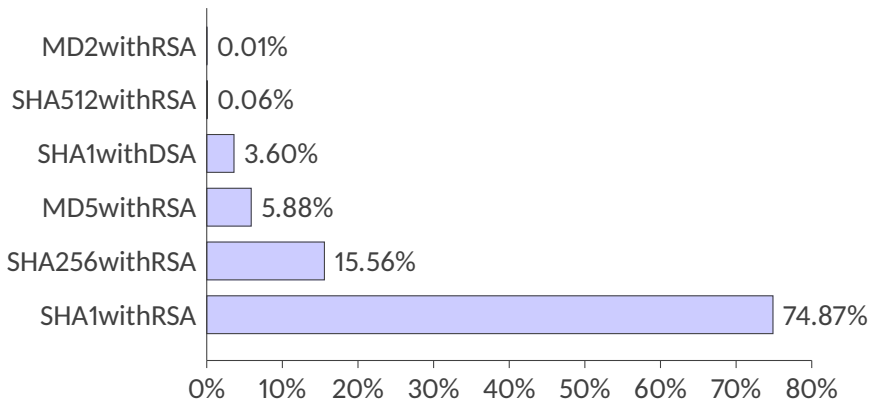
Štatistiky nad databázou APK súborov

Najčastejšie prístupové oprávnenia

| Názov | % |
|---|------|
| android.permission.internet | 92,9 |
| android.permission.access_network_state | 87,9 |
| android.permission.write_external_storage | 75,2 |
| android.permission.wake_lock | 49,5 |
| android.permission.read_phone_state | 49,4 |
| android.permission.access_wifi_state | 44,7 |
| android.permission.vibrate | 43,6 |
| android.permission.get_accounts | 31,3 |
| android.permission.receive_boot_completed | 30,5 |

Štatistiky nad databázou APK súborov

Algoritmus podpisu APK balíčka



Štatistiky nad databázou APK súborov

Lokalizácie aplikácií

| Kód | Jazyk | % |
|--------|-----------------------|------|
| es | španielsky | 61,7 |
| de | nemecký | 59,6 |
| fr | francúzsky | 59,4 |
| ru | ruský | 58,1 |
| ja | japonský | 57,6 |
| it | taliansky | 57,4 |
| ko | korejský | 56,9 |
| zh-rcn | čínsky (zjednodušený) | 55,6 |

- Anglický jazyk považovaný za základ, nevyskytuje sa v tabuľke.
- V českom jazyku je lokalizovaných 49 % aplikácií, v slovenčine 46 %

Detekcia prebalených APK súborov

Možnosti a hrozby modifikácie APK súborov

- Apk súbory je možné rozbaľiť, modifikovať a zabaliť do pôvodnej podoby
- Android obsahuje ochranný mechanizmus
 - súbor MANIFEST.MF obsahuje hash každého súboru - ochrana integrity
- Jednoduché vyhnutie ochrannému mechanizmu
 - po každej zmene nutné balíček podpísať
- Spôsob šírenia malvéru

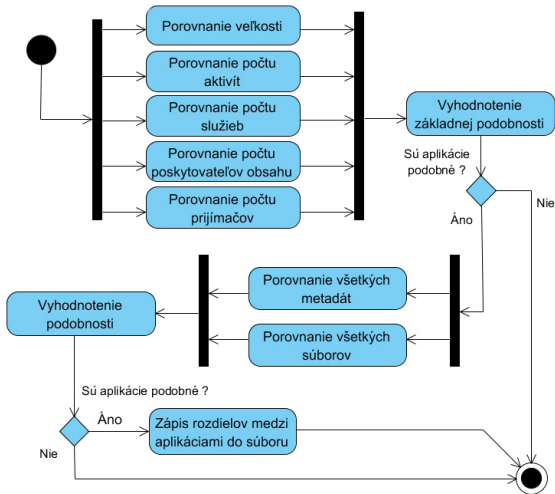
Detekcia prebalených APK súborov

Detekcia prebalených APK súborov

- Implementovaná metóda detekcie nadmieru podobných APK súborov pomocou podobnosti súborov a metadát o aplikáciách
- Malvérové aplikácie zachovávajú look & feel pôvodných
- Funkcionalita obsiahnutá v aplikácii ApkAnalyzer
- Využitie metadát získaných analýzou APK súborov
- Párové porovnanie APK súborov

Detekcia prebalených APK súborov

Navrhnutá metóda detekcie prebalených APK súborov



Detekcia prebalených APK súborov

Výstup

- Výstup párového porovnania obsahuje detailné rozdiely medzi jednotlivými aplikáciami
- Na základe zhody certifikátu a verzie aplikácie sú vyhodnotené potenciálne škodlivé aplikácie
- Identifikovaných 161 nadmieru podobných APK súborov s rovnakou verziou podpísaných rôznymi certifikátmi

Bibliografia

Analýza inštalačných APK súborov pre OS Android

- [1] WESTENBERG, Jimmy. Gartner: Android and iOS dominate smartphone market with 98 percent marketshare [online]. 2015 [cit. 2016-03-23]. Dostupný z: <http://www.androidauthority.com/android-ios-hold-98-percent-marketshare-656624/>.
- [2] MartinStyk/ApkDownloader [online]. 2012 [cit. 2016-04-26]. Dostupný z: <http://github.com/MartinStyk/ApkDownloader>.
- [3] MartinStyk/ApkAnalyzer [online]. 2012 [cit. 2016-04-26]. Dostupný z: <http://github.com/MartinStyk/ApkAnalyzer>.
- [4] Apktool [online]. 2015 [cit. 2016-03-26]. Dostupný z: <http://ibotpeaches.github.io/Apktool/>.