Masarykova univerzita
Fakulta informatiky

Analýza inštalačných APK súborov pre OS Android

Bakalárska práca

Martin Styk

Brno, jar 2016

Masarykova univerzita
Fakulta informatiky

Analýza inštalačných APK súborov pre OS Android

Bakalárska práca

Martin Styk

Brno, jar 2016

Namiesto tejto stránky vložte kópiu oficiálneho podpísaného zadania práce a prehlásenie autora školského diela.

Prehlásenie

Prehlasujem, že táto bakalárska práca je mojím pôvodným autorským dielom, ktoré som vypracoval samostatne. Všetky zdroje, pramene a literatúru, ktoré som pri vypracovaní používal alebo z nich čerpal, v práci riadne citujem s uvedením úplného odkazu na príslušný zdroj.

Martin Styk

Vedúci práce: Ing. Mgr. et Mgr. Zdeněk Říha, Ph.D.

Poďakovanie

Rád by som sa poďakoval vedúcemu práce Ing. Mgr. et Mgr. Zdeňkovi Říhovi, Ph.D. za venovaný čas, ochotu a cenné pripomienky, ktoré mi pomohli pri tvorbe tejto práce.

Zhrnutie

Práca sa zaoberá získavaním metadát o inštalačných APK súboroch pre mobilný operačný systém Android. V rámci práce je vytvorená rozsiahla databáza APK balíčkov. Na základe analýzy týchto súborov sú určené štatistické vlastnosti APK súborov a príslušných aplikácií. Ako súčasť tejto práce je implementovaný nástroj na hromadné sťahovanie APK súborov, ich analýzu a výpočet štatistických dát nad množinou APK súborov. Práca sa zaoberá aj bezpečnosťou aplikácií a detekciou modifikovaných APK súborov. V práci je navrhnutá metóda detekcie upravených a prebalených APK balíčkov, ktorá je aj prakticky implementovaná. V teoretickej časti je popísaná štruktúra APK balíčkov a súborov v nich obsiahnutych.

Kľúčové slová

APK súbor, Android, Apktool, malvér, analýza aplikácií, AndroidManifest.xml

Obsah

Zoznam tabuliek

Zoznam obrázkov

1 Úvod

TBD

2 Analýza APK súborov

Hlavnou úlohou práce je získať informácie o APK súboroch ich detailnou analýzou. APK súbory majú pevnú štruktúru a jednoduchý formát, vďaka čomu je možná ich analýza a reverzné inžinierstvo. Reverzné inžinierstvo je proces analýzy funkcionality a obsahu aplikácie. Keďže APK súbory využívajú ZIP formát, mnohé informácie je možné získať jednoduchým rozbalením. Základnou úlohou analýzy a reverzného inžinierstva APK súborov v tejto práci je získanie metadát o APK súbore, ktoré sú využívané v kapitole **??** a **??**.

2.1 Nástroje reverzného inžinierstva

Existuje viacero nástrojov poskytujúcich funkcionalitu pre reverzné inžinierstvo Android aplikácií. Okrem aplikácií tretích strán je možné vo veľkej miere použiť aj nástroje obsiahnuté v *Android Software Development Kit (SDK).Android SDK* je kolekcia štandardných nástrojov používaných pri vývoji a zostavení Android aplikácií.

2.1.1 ApkTool

Nástroj na reverzné inžinierstvo Android aplikácií. Dokáže dekódovať zdroje aplikácie do takmer originálnej podoby. Do čitateľnej podoby prevádza súbory *resources.arsc*, *classes.dex* aj binárne XML súbory. Z dekódovaných súborov umožňuje opätovné zostavenie APK súboru. Súbor *classes.dex* je dekompilovaný do súborov vo formáte SMALI. Smali súbory obsahujú nízkoúrovňový kód na úrovni asembleru. ApkTool podporuje debugovanie smali kódu [**apkTool**].

2.1.2 Dex2Jar

Nástroj podporujúci dekódovanie DEX súborov do formátu skompilovaných CLASS súborov .Výsledné CLASS súbory môžu byť prevedené do čitateľného kódu v jazyku Java pomocou dekompilátoru *JD-GUI*. Pracuje výhradne so súborom *classes.dex* a nepodporuje prevod binárnych XML do čitateľnej podoby.

2.1.3 AXML

AXML je knižnica navrhnutá na prácu s binárnymi XML súbormi, ktoré vznikajú počas zostavenia Android aplikácie pomocou nástroja *AAPT*. Knižnica umožňuje prevod takýchto XML súborov do čitateľného XML formátu, je implementovaná v jazyku Java.

2.1.4 AAPT

Android Asset Packaging Tool (*AAPT*) je štandardný nástroj obsiahnutý v *Android SDK*. Nástroj *AAPT* umožňuje vytvorenie, aktualizovanie a prezeranie súborov vo formáte APK. Dokáže skompilovať zdrojové súbory do binárnej formy a umožňuje aj ich dekompiláciu[**aapt**].

2.2 Implementácia analýzy

Analýza APK súborov je implementovaná v rámci programu *ApkAnalyzer* a môže byť spustená pomocou argumentu *–analyze*. Zároveň je potrebné špecifikovať analyzovaný APK súbor alebo priečinok obsahujúci takéto súbory pomocou argumentu *–in* a priečinok do ktorého bude zapísaný výstup analýzy pomocou argumentu *–out*. *ApkAnalyzer* je aplikácia prispôsobená na prácu s veľkým počtom APK súborov, proces analýzy je preto paralelizovaný a každé dostupné procesorové jadro analyzuje inú aplikáciu. Pre každú analyzovanú aplikáciu je vygenerovaný výstupný súbor vo formáte JSON obsahujúci získané metadáta o danej aplikácií.

Zbierané metadáta je možné rozdeliť do piatich kategórií:

- Základné informácie o APK súbore – v tejto kategórií sa nachádzajú informácie ako je veľkosť APK súboru alebo veľkosti súborov *classes.dex* a *resources.arsc*. Pre získanie veľkosti súborov obsiahnutých v APK balíčku je balíček rozbalený do dočasného adresára

- Informácie zo súboru *AndroidManifest.xml* – *AndroidManifest.xml* predstavuje hlavný zdroj meta informácií o aplikácii pre systém Android(viď **??**). Dáta nachádzajúce sa v tomto súbore tvoria

významnú časť dát získaných našou analýzou. Na prevod z binárneho XML formátu je primárne použitá knižnica *AXML* (viď 2.1.3), v prípade zlyhanie konverzie sa použije nástroj *ApkTool* (viď 2.1.1). Dáta získané analýzou tohto súboru zahŕňajú napríklad verziu aplikácie, použité prístupové oprávnenia alebo komponenty z ktorých sa aplikácia skladá

- Informácie o certifikáte – dáta získané analýzou súboru *CERT.RSA* v priečinku *META-INF*(viď **??**). Obsahujú napríklad použitý algoritmus podpisovania, názov vydavateľa alebo MD5 hash celého certifikátu. Pred prístupom k súboru *CERT.RSA* je nutné APK balíček rozbaliť

- Informácie o zdrojových súboroch[1] – informácie o zdrojoch aplikácie, napríklad formát alebo veľkosť obrázkových súborov, počet lokalizácií aplikácie alebo počet surových neskomprimovaných zdrojových súborov

- Súbory obsiahnuté v APK balíčku – zoznam všetkých súborov rozdelený do kategórií: obrázky(súbory z priečinku *res/drawable*), návrhy obrazoviek(súbory z priečinku *res/layout*), *classes.dex*, *resources.arsc* a ostatné. O každom súbore si uchovávame jeho relatívnu cestu v APK balíčku a SHA1 hash. Ako zdroj informácií slúži súbor *MANIFEST.MF* (viď **??**)

Kompletný zoznam zbieraných metadát sa nachádza v prílohe **??**.

1. angl. resources

Literatúra

1. Freed, Ned; Kucherawy, Murray; Baker, Mark; Hoehrmann, Bjoern. *Media Types* [online]. 2016 [visited on 2016-03-23]. Available from WWW: ⟨http://www.iana.org/assignments/media-types/media-types.xhtml⟩.
2. *Building and Running Overview* [online]. 2016 [visited on 2016-03-23]. Available from WWW: ⟨http://developer.android.com/tools/building/index.html⟩.
3. Yang, Herong. *META-INF Files - Digests, Signature and Certificate* [online]. 2015 [visited on 2016-03-24]. Available from WWW: ⟨http://www.herongyang.com/Android/Project-META-INF-Files-Digest-Signature-and-Certificate.html⟩.
4. *Accessing Resources* [online]. 2015 [visited on 2016-03-24]. Available from WWW: ⟨http://developer.android.com/guide/topics/resources/accessing-resources.html⟩.
5. *Providing Resources* [online]. 2015 [visited on 2016-03-24]. Available from WWW: ⟨http://developer.android.com/guide/topics/resources/providing-resources.html⟩.
6. *Providing Alternative Resources* [online]. 2015 [visited on 2016-03-24]. Available from WWW: ⟨http://developer.android.com/guide/topics/resources/providing-resources.html%5C#AlternativeResources⟩.
7. Reddy, Satheesh. *Android Application Build Process or Compilation Process* [online]. 2014 [visited on 2016-03-24]. Available from WWW: ⟨http://www.c-sharpcorner.com/UploadFile/34ef56/android-application-build-process-or-compilation-process/⟩.
8. *ART and Dalvik* [online]. 2015 [visited on 2016-03-23]. Available from WWW: ⟨https://source.android.com/devices/tech/dalvik/⟩.
9. *App Manifest* [online]. 2015 [visited on 2016-03-24]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/manifest-intro.html⟩.
10. *Manifest element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/manifest-element.html⟩.
11. *Uses-permission element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/uses-permission-element.html⟩.

12. *Permission element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/permission-element.html⟩.
13. *Uses-sdk element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/uses-sdk-element.html⟩.
14. *Uses-feature element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/uses-feature-element.html⟩.
15. *Supports-screens element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/supports-screens-element.html⟩.
16. *Activity* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/reference/android/app/Activity.html⟩.
17. *Service* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/reference/android/app/Service.html⟩.
18. *ContentProvider* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/reference/android/content/ContentProvider.html⟩.
19. *Receiver element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/receiver-element.html⟩.
20. *Uses-library element* [online]. 2015 [visited on 2016-03-26]. Available from WWW: ⟨http://developer.android.com/guide/topics/manifest/uses-library-element.html⟩.
21. Westenberg, Jimmy. *Gartner: Android and iOS dominate smartphone market with 98 percent marketshare* [online]. 2015 [visited on 2016-03-23]. Available from WWW: ⟨http://www.androidauthority.com/android-ios-hold-98-percent-marketshare-656624/⟩.
22. Thomas, Daniel R.; Beresford, Alastair R.; Rice, Andrew. Security Metrics for the Android Ecosystem. *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices - SPSM '15*. 2015, pp. 87–98. Available also from WWW: ⟨http://dl.acm.org/citation.cfm?doid=2808117.2808118⟩.

23. *Industry Leaders Announce Open Platform for Mobile Devices* [online]. 2007 [visited on 2016-03-23]. Available from WWW: ⟨http://www.openhandsetalliance.com/press_110507.html⟩.
24. Rosoff, Matt. *Google's Biggest Acquisitions So Far, And What They Became* [online]. 2011 [visited on 2016-03-23]. Available from WWW: ⟨http://www.gizmodo.com.au/2011/08/googles-16-biggest-acquisitions-so-far-and-what-happened-to-them/⟩.
25. Beavis, Gareth. *A complete history of Android* [online]. 2008 [visited on 2016-03-23]. Available from WWW: ⟨http://www.techradar.com/news/phone-and-communications/mobile-phones/a-complete-history-of-android-470327⟩.
26. *An Overview of the Android Architecture* [online]. 2013 [visited on 2016-03-23]. Available from WWW: ⟨http://www.techotopia.com/index.php/An_Overview_of_the_Android_Architecture⟩.
27. Parmar, Ketan. *In Depth: Android Package Manager and Package Installer*. 2013. Available also from WWW: ⟨https://dzone.com/articles/depth-android-package-manager⟩.
28. Elenkov, Nikolay. *Android security internals: an in-depth guide to android's security architecture*. San Francisco: No Starch Press, 2015. ISBN 978-1-59327-641-6.

Register

A An appendix

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
zh-rtw	čínsky (tradičný)	54,0
pt	portugalský	52,6

Tabuľka A.1: Lokalizácia aplikácií

Názov atribútu	Dátový typ	popis
fileName	String	Názov analyzovaného APK súboru
sourceOfFile	String	Zdroj súboru
fileSize	Long	Veľkosť APK súboru v bajtoch
dexSize	Long	Veľkosť súboru *classes.dex* v bajtoch
arscSize	Long	Veľkosť súboru *arscSize.dex* v bajtoch
packageName	String	Hodnota atribútu *package* v elemente *man*
versionCode	String	Hodnota atribútu *android:versionCode* v el
installLocation	String	Hodnota atribútu *android:installLocation* v
numberOfActivities	Integer	Počet aktivít definovaných aplikáciou
numberOfServices	Integer	Počet služieb definovaných aplikáciou
numberOfContentProviders	Integer	Počet poskytovateľov obsahu definovaný
numberOfBroadcastReceivers	Integer	Počet komponent typu *BroadcastReceiver*
namesOfActivities	List<String>	Názvy aktivít definovaných aplikáciou

namesOfServices	List<String>	Názvy služieb definovaných aplikáciou
namesOfContentProviders	List<String>	Názvy poskytovateľov obsahu definovaných a
namesOfBroadcastReceivers	List<String>	Názvy komponent typu *BroadcastReceiver* defi
usesPermissions	List<String>	Názvy povolení využívaných aplikáciou
usesLibrary	List<String>	Názvy knižníc využívaných aplikáciou
permissions	List<String>	Názvy povolení definovaných aplikáciou
permissionsProtectionLevel	List<String>	Level ochrany povolení definovaných aplikáci
usesFeature	List<String>	Názvy vlastností využívaných aplikáciou
usesMinSdkVersion	String	Hodnota atibútu *android:minSdkVersion* elemer
usesTargetSdkVersion	String	Hodnota atribútu *android:targetSdkVersion* elen
usesMaxSdkVersion	String	Hodnota atribútu *android:maxSdkVersion* eleme
supportsScreensResizeable	Boolean	Hodnota atribútu *android:resizeable* elementu *s*
supportsScreensSmall	Boolean	Hodnota atribútu *android:smallScreens* element
supportsScreensNormal	Boolean	Hodnota atribútu *android:normalScreens* eleme
supportsScreensLarge	Boolean	Hodnota atribútu *android:largeScreens* element
supportsScreensXlarge	Boolean	Hodnota atribútu *android:xlargeScreens* elemer
supportsScreensAnyDensity	Boolean	Hodnota atribútu *android:anyDensity* elementu
fileName	String	Názov súboru s certifikátom
signAlgorithm	String	Algoritmus použitý na podpis
signAlgorithmOID	String	OID algoritmu použitého na podpis
startDate	Date	začiatok platnosti certifikátu
endDate	Date	koniec platnosti certifikátu
publicKeyMd5	String	MD5 hash verejného klúča
certBase64Md5	String	Base64 MD5 hash certifikátu
certMd5	String	MD5 hash certifikátu
version	Integer	Verzia cetifikátu
issuerName	String	Názov vydávateľa vo formáte definovanom R
subjectName	String	Názov subjektu vo formáte definovanom RFC
locale	List<String>	Lokalizácie súboru *string.xml*
numberOfStringResource	Integer	Počet záznamov v súbore *string.xml*

pngDrawables	Integer	Počet PNG obrázkov
ninePatchDrawables	Integer	Počet 9.PNG obrázkov
jpgDrawables	Integer	Počet JPG obrázkov
gifDrawables	Integer	Počet GIF obrázkov
xmlDrawables	Integer	Počet XML obrázkov
ldpiDrawables	Integer	Počet obrázkov v ldpi priečinku
mdpiDrawables	Integer	Počet obrázkov v mdpi priečinku
hdpiDrawables	Integer	Počet obrázkov v hdpi priečinku
xhdpiDrawables	Integer	Počet obrázkov v xhdpi priečinku
xxhdpiDrawables	Integer	Počet obrázkov v xxhdpi priečinku
xxxhdpiDrawables	Integer	Počet obrázkov v xxxhdpi priečinku
tvdpiDrawables	Integer	Počet obrázkov v tvdpi priečinku
nodpiDrawables	Integer	Počet obrázkov v nodpi priečinku
unspecifiedDpiDrawables	Integer	Počet obrázkov nezaradených v *dpi prie
rawResources	Integer	Počet súborov v *raw/* priečinku
layouts	Integer	Počet súborov v *res/layout** priečinkoch
differentLayouts	Integer	Počet rôznych súborov v *res/layou*
menu	Integer	Počet súborov v *res/menu* priečink
dexHash	String	Hash súboru *classes.dex* prevzatý
arscHash	String	Hash súboru *arscHash.dex* prevza
drawableHash	Map<String,String>	Hashe a cesty k súborov z prečink
layoutHash	Map<String,String>	Hashe a cesty k súborov z prečink
otherHash	Map<String,String>	Hashe a cesty k všetkým ostatným

Tabuľka A.6: Zbierané metadata o AP

Názov	%
android.hardware.camera	18,1
android.hardware.touchscreen	16,1
android.hardware.telephony	14,8
android.hardware.camera.autofocus	10,6
android.hardware.location.gps	10,2
android.hardware.location	8,8
android.hardware.wifi	8,4
android.hardware.location.network	7,0
android.hardware.bluetooth	6,6
android.hardware.touchscreen.multitouch	6,0

Tabuľka A.2: Najpoužívanejšie vlastnosti

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
android.permission.vending.billing	27,1

Tabuľka A.3: Najpoužívanejšie prístupové oprávnenia

Verzia Android SDK	%
9	21,3
8	18,4
7	14,2
14	10,5
10	8,1
4	7,0
3	5,6
15	3,7
5	3,7
11	2,1

Tabuľka A.4: Hodnoty najnižsej vyžadovanej verzie Android SDK

Verzia Android SDK	%
19	25,6
17	11,8
21	11,7
15	6,8
14	6,3
22	6,0
16	5,7
18	5,6
20	3,8
8	2,9

Tabuľka A.5: Hodnoty cieľovej verzie Android SDK