Kódolás és IT biztonság
(VIHIBB01)
Laboratóriumi gyakorlat

Programvisszafejtés

Szerző:
Nagy Roland

2023. november 13.

Tartalomjegyzék

1. Mérés célja

A *reverse engineering* kifejezés tágabb értelemben azon folyamat neve, amikor valamilyen mérnöki munka eredményeként előállt termék működését próbáljuk megérteni azáltal, hogy szétszereljük, darabokban vizsgáljuk azt. Az informatikában az esetek túlnyomó részében ez, a mérnöki munka eredményeként előállt termék valamilyen szoftver. Több okból is szükséges lehet szoftverek visszafejtése, hogy valamilyen információhoz jussunk velük kapcsolatban: lehet, hogy a szoftver megfelelő üzemeltetéséhez szükséges megtudnunk róla valamit, vagy egy nyílt forráskódú projekthez akarunk hozzáfejleszteni valamit és ezért szükséges megértenünk a működését, vagy éppen azt gyanítjuk egy programról, hogy valamilyen kártékony kódot tartalmaz.

A programvisszafejtés, különösen bináris programok visszafejtése, hosszadalmas és nagy szakértelmet igénylő munka, még az alapok elsajátítása is bőven túlmutat azon, ami beleférne a laborgyakorlat kereteibe. A labor célja olyan technikák és technológiák felszínes ismertetése, amik segítségével könnyen és gyorsan tudunk információhoz jutni, ha azt gyanítjuk, hogy olyan fájllal találkoztunk, ami valamilyen kártékony kódot rejt.

2. Háttér

2.1. Hash függvények

A kriptográfiában gyakran alkalmazunk úgynevezett hash függvényeket, hogy előállítsunk egy kompakt bájtsorozatot, ami, mint egy ujjlenyomat, egyedi azonosítóként szolgálhat egy fájlhoz. Elsődlegesen arra használjuk őket, hogy megbizonyosodhassunk arról, hogy egy fájl, vagy üzenet tartalma nem változott meg: ha ismerjük egy fájl hash-ét, tetszőleges időpillanatban újra kiszámolva azt, és a két értéket összehasonlítva megbizonyosodhatunk arról, hogy a fájl tartalma nem változott. Ugyanezen technika kiterjesztéseként, ha két fájlt odaadunk egy hash függvénynek inputként, és azt látjuk, hogy a két hash megegyezik, elkönyvelhetjük, hogy a két fájl tartalma is megegyezik.

A labor során az MD5, SHA1, SHA256 és SHA512 hash függvényeket fogjuk használni. A feladatok elvégzéséhez nem szükséges ismerni magukat az algoritmusokat, amik kiszámolnak egy hash-t az adott inputból, elég annyit tudni, hogy Linuxon az `md5sum`, `sha1sum`, `sha256sum` és `sha512sum` parancsok segítségével számolhatjuk ki a fent említett hash-eket bizonyos

fájlokra.

2.2. VirusTotal

A VirusTotal egy, a Google által fejlesztett és karbantartott szolgáltatás, ami mára központi szerepet tölt be minden malware elemzéssel foglalkozó szakember munkájában. Lehetőségünk van malware minták feltöltésére, amiket a VT több tucat antivírus motoron kiértékel (**NE töltse fel a feladat során elemzendő mintát!**), illetve kereshetünk mintákra hash-ek alapján, vagy kereshetünk IP címekre és domain nevekre is, hogy megtudjuk, felhasználták-e azokat valamilyen malware-ben. Ezen keresések eredményei igen hasznosak lehetnek annak meghatározásában, pontosan mivel vagy kivel állunk szemben.

A VirusTotal ezeken felül számos szolgáltatást nyújt enterprise felhasználóknak, a labor során mi viszont csak az ingyenesen elérhető funkciókat fogjuk használni.

2.3. `strings`

A `strings` parancs egy olyan alkalmazás, ami minden Linux rendszeren alapból fel van telepítve, és ahogy arra a neve is utal, képes stringeket megtalálni bináris fájlokban. Ha lefuttatjuk egy fájlra, végigolvassa azt, és ha talál egymás után legalább 4 nyomtatható karaktert, kiírja azokat (ez a limit természetesen állítható, a `-n` kapcsoló segítségével).

Amennyire egyszerű, olyan hasznos lehet ez az eszköz, amikor bináris programokat akarunk elemezni, azok ugyanis gyakran tartalmaznak ilyen karaktersorozatokat, amikből meglepően sok információt nyerhetünk a program működésével kapcsolatban. A labor során ezt is felhasználjuk majd, hogy többet tudjunk meg egy binárisról.

3. Feladatok

3.1. feladat

Töltse le a Moodle-ből a zip fájlt és csomagolja ki. A kicsomagoláshoz szükséges jelszó a következő: `infected`.

- Futtassa le a `file` parancsot a kicsomagolt fájlra, hogy megtudjuk, pontosan milyen típusú fájllal állunk szemben.

- Számolja ki a fájlra a korábban említett hash-eket (MD5, SHA1, SHA256, SHA512)

Beadandó: a `file` parancs outputja és a 4 hash érték.

3.2. feladat

Nyisson meg egy böngészőt és látogasson el a VirusTotal oldalára (`www.virustotal.com`)! **NE töltse fel a mintát!** Kattintson a `search` fülre és keressen rá a mintára a korábban kiszámolt SHA1 hash alapján!

- A bal felső sarokban láthatja, hány antivíruson futtatta le a VirusTotal a mintát, amikor legutóbb volt elemezve, és ezek közül mennyi ismerte fel malware-ként. Hány antivírus szerint malware ez a fájl?
- Lejjebb görgetve táblázatos formában láthatja, milyen eredményeket adtak a különböző antivírus termékek. Érdemes lehet megnézni, mit mond néhány ismertebb termék: milyen eredményt adott az Avast és a Kaspersky?
- A táblázat fölött látható egy összefoglaló a táblázatban prezentált eredményekről. Mi az első Family label, amit a mintának adott a VT, amelyik vírus családba tartozik nagy valószínűséggel a minta?
- Kattintson a deatils tabra és olvassa végig, milyen információkat gyűjtött össze a VT a mintáról! Mikor töltötték fel először ezt a mintát (First submission date)?

3.3. feladat

Futtassa le a `strings` parancsot a fájlra! Keressen IP címeket a kapott eredmények között!

Beadandó: 4 db IPv4 cím (`104.*.*.*`, `206.*.*.*`, `2 x 8.*.*.*`)

3.4. feladat

Keressen rá VT-n először a `104` kezdetű címre!

- Melyik országban van szerver, amihez az IP cím tartozik? A két betűs országazonosítót adja meg válaszként!

- Hány „security vendor” értékelte gyanúsnak ezt az IP címet?
- A VT-ről azt is megtudhatjuk, hogy az adott IP cím milyen szervezethez tartozik (Autonomous System Label a details tabon); kihez tartozik ez az IP cím?

3.5. feladat

Keressen rá VT-n a `206` kezdetű címre!

- Melyik országban van szerver, amihez az IP cím tartozik? A két betűs országazonosítót adja meg válaszként!
- Hány „security vendor” értékelte gyanúsnak ezt az IP címet?
- A VT-ről azt is megtudhatjuk, hogy az adott IP cím milyen szervezethez tartozik (Autonomous System Label a detail tabon); kihez tartozik ez az IP cím?

3.6. feladat

Keressen rá VT-n a két `8` kezdetű címre!

- Melyik országban vannak a szerverek, amikhez az IP cím tartozik? A két betűs országazonosítót adja meg válaszként!
- A VT-ről azt is megtudhatjuk, hogy az adott IP cím milyen szervezethez tartozik (Autonomous System Label a detail tabon); kihez tartozik ez az IP cím?
- Ez a két cím egy széles körben használt szolgáltatáshoz tartozik. Keressen rá az egyikre a Google segítségével. Milyen protokoll tartozik ehhez a szolgáltatáshoz?