Esteganografia (Previ)

1.-a) Suposant que tens una foto anomenada "foto.jpg" a la teva carpeta de treball i un arxiu anomenat "topsecret.txt" amb un contingut com "Hola caracola" , ¿què creus que mostrarà per pantalla la comanda *cat foto.jpg topsecret.txt* ? I si ara generem un nou fitxer anomenat "imatge.jpg" amb la comanda *cat foto.jpg topsecret.txt > imatge.jpg* , ¿aquest és visible amb algun visor d'imatges (com *eog*)? La resposta al per què es troba a l'apartat següent

b) ¿Quina informació mostra la llista https://en.wikipedia.org/wiki/List_of_file_signatures (o també https://www.garykessler.net/library/file_sigs.html) ? En relació a això, ¿per què la comanda *file imatge.jpg* mostra el que mostra? ¿I per què, si canviessis l'extensió del fitxer (per a què fos, per exemple, "imatge.png"), la comanda *file* seguiria mostrant el mateix resultat?

> NOTA: Els "magic numbers" són seqüències de bits estandaritzades que estan integrades dins de tot fitxer (concretament, al seu inici) per identificar el tipus de fitxer del qual es tracta. Aquesta informació és molt més fiable que les extensions dels fitxers, completament manipulables per l'usuari. Hi ha una llista dels "magic numbers" més habitual a https://en.wikipedia.org/wiki/List_of_file_signatures o també https://www.garykessler.net/library/file_sigs.html Noteu que, en concret el tipus JPG, a més de tenir definit uns "magic numbers" d'inici, també els tenen per marcar el seu final ("trailers"), cosa que no és gaire habitual

c) Executa *hexdump -C imatge.jpg*. ¿Què veus al final de la seva sortida, a la columna de la dreta? ¿Per què? D'altra banda, ¿què mostra la comanda *strings imatge.jpg* ?

> NOTA: La comanda *strings* (del paquet "binutils") mostra les eventuals cadenes incrustades dins del fitxer binari indicat (és a dir, les seqüències de caràcters imprimibles que són dins d'un fitxer binari). El paràmetre *-t d* serveix per mostrar, al costat de cada cadena trobada, la seva posició (en bytes) des de l'inici del fitxer on es troba (es mostra en decimal amb el valor "d"; si es vol en hexadecimal es pot indicar el valor"x")

d) Si ara executem *cat topsecret.txt foto.jpg > imatge2.jpg*, ¿per què la comanda *file imatge2.jpg* ara mostra una altra cosa? ¿Per què, de fet, ni tan sols ara es pot visualitzar "imatge2.jpg"? Si observes el contingut binari/ASCII del fitxer "imatge2.jpg" amb *hexdump -C*, ¿en quina posició apareix la signatura del format de la foto original? (La resposta a totes aquestes preguntes és la mateixa).

e) ¿Què fan les següents comandes...

> *zip -r carpetasecreta.zip $HOME*
> *cat foto.jpg carpetasecreta.zip > imatge3.jpg*
> *rm carpetasecreta.zip*

...i aquesta altra: *unzip imatge3.jpg* ?

> NOTA: La comanda anterior funciona perquè la comanda *unzip* té el comportament particular de no deixar de buscar "el nombre màgic" corresponent al format ZIP en tot el fitxer des del començament fins el final. La majoria de comandes no fan això, només el busquen al començament
> NOTA: Fixa't que aquest sistema tan senzill ens permet enviar qualsevol tipus de contingut de forma embeguda en una simple foto. L'única "cosa" que a simple vista podria ser "sospitosa" és la mida del fitxer, que òbviament serà més gran com més gran sigui l'arxiu Zip embegut.

2.-a) Instal.la (en una màquina virtual qualsevol) el paquet "binwalk" i consulta la seva descripció (amb *apt show* o *dnf info*, per exemple) per esbrinar per a què serveix la comanda homònima.

> NOTA: Per més informació sobre les opcions de *binwalk*, llegiu https://github.com/ReFirmLabs/binwalk/wiki/Usage

b) Descarrega la fotografia https://ctfs.github.io/resources/topics/steganography/file-in-image/example.jpg i tot seguit executa la comanda *binwalk example.jpg* ¿Què veus? Pista: llegeix l'ajuda bàsica de la comanda *binwalk* mostrada a continuació

binwalk fitxer.raw : Llista les estructures binàries internes reconegudes. És equivalent a *binwalk -B fitxer.raw* A les columnes "Decimal" i "Hexadecimal" es mostra (cadascuna en el seu format) la posició del primer byte a partir del qual comença l'estructura concreta, assenyalada a la columna "Description"

binwalk -e fitxer.raw : Extreu a la carpeta actual (dins d'una carpeta anomenada "_nomfitxer.raw.extracted") el/s fitxer/s binari/s que automàticament s'hagin reconegut dins el fitxer indicat

> NOTA: Aquests fitxers binaris, per a què siguin reconeguts, cal que la seva estructura s'hagi definit prèviament dins de l'arxiu de configuració "~/.config/binwalk/config/extract.conf". Afortunadament, moltes d'elles ja s'hi troben per defecte (zlib, lzma, cpio, xz, gzip, ...), així que aquesta comanda funcionarà correctament en la majoria d'ocasions sense haver d'editar cap configuració.
> NOTA: Si dins del fitxer "raw" indicat hi haguessin més estructures d'empaquetació incloses (és a dir, un arxiu zip, gzip, xz... dins d'ell), és possible realitzar una extracció recursiva de totes aquestes estructures d'empaquetació internes afegint al paràmetre *-e* el paràmetre ***-M***

binwalk -D "part_descripcio:ext" fitxer.raw : Extreu a la carpeta actual (dins d'una carpeta anomenada "_nomfitxer.raw.extracted") només el/s fitxer/s binari/s amb la descripció de la signatura del/s qual/s indicada. Aquesta descripció apareix a la columna de més a la dreta en fer el llistat d'estructures binàries internes existents. Es pot escriure qualsevol part de la descripció a partir del principi i pot ser una expressió regular, però en tot cas ha de en minúscules. L'extensió indicada serà la que es s'afegirà al/s fitxer/s binari/s extret/s (per defecte és cap); el nom d'aquests serà l'"offset" en hexadecimal on la signatura corresponent s'ha trobat dins del fitxer "raw".

> NOTA: Una alternativa a la comanda *binwalk -D* anterior és fer servir directament la comanda ***dd*** amb el paràmetre ***bs=1*** , el paràmetre ***skip=nº*** indicant la posició a partir d'on extreure (aquest número es mostra al llistat oferit per *binwalk*) i el paràmetre ***count=nº*** valent la diferència entre la posició del següent element binari que apareix a la llista menys el valor indicat a *skip=* . És a dir, així: *dd if=fitxer.raw of=imatge.png bs=1 skip=68264 count=2760* , per exemple. Un exemple d'això es troba explicat a l'article https://ctfs.github.io/resources/topics/steganography/file-in-image/README.html

c) Executa la comanda *dd if=example.jpg of=arxiu.zip skip=1972141 bs=1* ¿Què has fet? Pista: repassa el significat dels paràmetres *skip=* i *bs=* de la comanda *dd* Prova de descomprimir l'arxiu "arxiu.zip" resultant d'haver executat la comanda anterior. ¿Quin és el missatge ocult?

d) Fes el mateix que l'apartat anterior (és a dir, extreure els possibles fitxers inclosos dins de la fotografia per tal de trobar-hi el missatge ocult) simplement executant la comanda *binwalk -e example.jpg* ¿Què obtens?