Pràctica 2a: Seguretat lògica - Algorismes hash i integritat de les dades

1- Objectius de la pràctica

Aquesta activitat té com objectius adquirir els següents resultats d'aprenentatge i continguts marcats al currículum de la UF a DOGC:

- RA1 i C1 →
  - RA1.7 --> Aplica tècniques criptogràfiques en l'emmagatzematge i la transmissió de la informació.
  - C1.6 → Seguretat Lògica

Un altre objectiu de l'activitat és avaluar si s'han assolit correctament aquests coneixements per mitjà d'una pràctica.

Per desenvolupar aquests objectius:

- L'alumne haurà de llegir i entendre la documentació. Per resoldre qualsevol dubte sobre la documentació, es durà a terme una sessió a classe per explicar els conceptes més importants i resoldre els dubtes que hagin aparegut durant la lectura de la documentació.
- L'alumne haurà de lliurar una pràctica funcionant d'acord a allò que es demani a l'enunciat.

2- Documentació: algorismes hash (resum) i integritat

a) Què fa un programa que implementa un algorisme hash?
Crea a partir d'unes dades d'entrada (un fitxer de texte, una imatge, un fitxer comprimit, una base de dades, imatges ISO,etc....) un codi numèric de longitud fixa (128 bits, 160 bits, 256 bits, 512 bits,....) que generalment es mostra per pantalla com una combinació de números en format hexadecimal. El codi resultant s'anomena el valor o codi hash (resum) de les dades.

El propòsit d'un programa que implement un algorisme hash no és protegir les dades donant confidencialitat a les dades. El propòsit és verificar que les dades no han estat alterades i que les dades que estem llegint són autèntiques.

b) Característiques d'un codi hash

- El codi resultant té una mida molt més petita que les dades d'entrada originals, o sigui, en general és de mida més petita que el fitxer de texte, ISO o imatge original.
- El valor hash representa de manera únivoca a les dades d'entrada.
- La probabilitat de col·lisió, és a dir, generar el mateix valor hash a partir de 2 conjunts de dades diferents de manera accidental, és extremadament petit per no dir pràcticament 0.

c) Característiques d'un algorisme hash

- No és reversible. A partit del valor hash no és possible aconseguir les dades originals d'entrada. Per tant, el valor hash no és una encriptació de les dades (la encriptació és reversible si conec la clau)
- La probabilitat de col·lisió, és a dir, generar el mateix valor hash a partir de 2 conjunts de dades diferents de manera accidental, és extremadament petit per no dir pràcticament 0.
- El cost de provocar una col·lisió expressament ha de ser molt alta en termes de temps dedicat, diners gastats i persones dedicades a aconseguir-ho.
- Un canvi molt petit de les dades d'entrada, per exemple un simple canvi d'una lletra dins d'un fitxer amb 1000 paraules o un canvi d'un pixel dins d'una imatge, pot general un valor hash completament diferent i impredictible al valor obtingut abans del canvi.
- El cost computacional i de memòria ha de ser molt baix. Això vol dir que s'ha de trobar el valor hash molt ràpidament i utilitzant poca memòria i poc temps d'utilització de la CPU.
- Ha de produir una compressió en el sentit que generalment la mida de les dades originals (fitxer de texte, imatge, etc..) és molt méd gran que la mida del valor hash obtingut.
- Les mateixes dades (el mateix fitxer de texte, imatge, etc...) sempre donen com a resultat el mateix valor hash.

d) Utilització
- Verificació de la integritat de les dades emmagatzemades o transferides.
- Detecció i correcció d'errors durant la transmissió de dades per mitjà de sistemes telemàtics
- Signatura digital.

e) Algorismes hash típics
- **MD5 (Message Digest algorithm 5)** → 128 bits, menys temps, menys potència de computació, més probabilita de col·lisió. Suficient per verificar integritat de fitxers descarregats des d'una web.
- **SHA-256 (Secure Hash Algorithm 2 - 256)** → 256 bits, més temps (20-30%), més potència decomputació, menys probabilitat de col·lisió. Utilitzat per generació de bitcoins o per verificar la integritat i la signatura digital de les distribucions de Debian.

f) Comprovació d'integritat de fitxers amb algorismes hash
Les fases de comprovació de la integritat d'un fitxer amb l'ajut d'algorismes hash són les següents:
- Primera fase:
  - Creació d'un fitxer del qual es voldrà assegurar en el futur la seva integritat després d'emmagatzemar-lo en un disc o transmetre-ho per una xarxa local o internet.
  - Utilització d'un programa que implementi un algorismes hash per calcular el codi hash del fitxer.
  - Publicació del valor hash del fitxer i de l'algoritme hash utilitzat per l'obtenció del codi.
- Segona fase:
  - Recuperació del fitxer des del dispositiu d'emmagatzematge o descarrega del fitxer via xarxa local o internet.
  - Obtenció del codi hash del fitxer del lloc a on s'ha publicat i comprovació de quin algorisme hash s'ha utilitzat.
  - Càlcul del codi hash a partir del fitxer obtingut via xarxa local, internet o del dispositiu d'emmagatzematge
  - Comparar els codis. Si són difierents podrem afirmar que no s'ha mantingut la integritat i que el fitxer obtingut ha patit canvis respecte de l'original.

3- Realització de la pràctica - part 1: Publicació d'un fitxer i el seu codi hash per mitjà del servei web

a) Comprova que els paquets **apache2** que permet convertir el teu sistema en un servidor web, **coreutils** -dins del qual hi ha els programes **md5sum** i **sha256sum**- i **wget** que permet descarregar fitxers via HTTP han estat instal·lats dins del teu sistema. Executa:

```
aptitude search ^apache2$ ^coreutils$ ^wget$
```

i el resultat ha de ser:

```
i  A  apache2                    - Apache HTTP Server
i  A  coreutils                  - GNU core utilities
i  A  wget                       - retrieves files from the web
```

b) Crea una carpeta de nom **m11uf1pr2a** dins del teu directori personal. A continuació accedeix a la carpeta **m11uf1pr2a** i des de dins de la carpeta, descarrega amb **wget** els següents fitxers:
- **nota.c** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/nota.c**
- **calcul.c** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/calcul.c**
- **dades.c** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/dades.c**
- **pantalla.c** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/pantalla.c**
- **nota.h** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/nota.h**
- **Makefile** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/Makefile**
- **README** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/README**

c) Comprova la integritat dels fitxers descarregats amb el programa **sha256sum**. Els codis hash dels fitxers utilitzant l'algorisme **SHA-256** són els següents:

```
b05bc1995878846a638c653234ba87cb8b07a9ed34980dd06dfdb0c729d1694d  calcul.c
3842885a4bca630d1a07691999693dc8125d0c461e13198beedadb12ed0fbe7b  dades.c
4a02f0b3ab995b4a5c2ba8f5b1198cad7288a97fc8e6fd1ee2bce65deb22d1b3  pantalla.c
7bc5b20a883f84aa1118cd84dbb4d88d765e83ef538dd1c3cc2f7c1072ae6210  nota.h
7fe2ca521fef8d28bc7af99ef8e6c76e65cd994343d885f8d4ed132b576866b9  nota.c
3401297f7258198f1184248525d6fc8a8d1d16d57593853452be925806aea34b  Makefile
09753914cb6fac6fa318cbc677fe4a28204ea8f7fab9db031ee1bac209540c31  README
```

d) Si has verificat la integritat dels fitxers, modifica el fitxer **README** i canvia la **data**, que és el punt 3, per la data **2021-10-25**. A continuació, troba el nou valor hash i compara-ho amb el del fitxer **README** original. Comprova si el nou valor hash és molt semblant o molt diferent a l'original. El resultat és normal a partir d'allò que has llegit a la documentació?.

e) Canvia també el nom de l'autor del fitxer README i escriu el teu nom i cognoms.

f) Crea un paquet de nom **nota.tar** amb tots els fitxers que hi ha dins de la carpeta **m11uf1pr2a** (però no la carpeta). A continuació, comprimeix el paquet i crea un paquet comprimit de nom **nota.tar.gz**.

g) Crea el codi hash del fitxer **nota.tar.gz** utilitzant l'algorisme hash **MD5** i desa-ho (per no perdre el codi) dins d'un fitxer de la carpeta **m11uf1pr2a** de nom **nota.md5** executant l'ordre:

```
md5sum nota.tar.gz > nota.md5
```

NOTA: Despres pots tornar a llegir el fitxer executant l'ordre → `cat nota.md5`

h) Descarrega amb **wget** el següent fitxer:
- **nota.html** que es troba a **http://www.collados.org/asix1/m11/uf1/m11uf1pr2a/nota/nota.html**
- Hash SHA-256: **76b2bba213b9380e8994c8ba3acc1d84da1f71d91ff6d90a2e9b02724c9f840e**
- Comprova la integritat del fitxer **nota.html** descarregat.

i) Modifica **nota.html** i fes que **mostri** el codi **hash md5** del teu fitxer **nota.tar.gz** i els teu **noms** i **cognoms** reals.

j) Canvia a usuari **root** utilitzant l'ordre `su -` .

k) **Copia** els fitxers **nota.html** i **nota.tar.gz** dins del directori **/var/www/html**.

l) Obre el navegador i a la barra d'adreces escriu `http://localhost/nota.html` per connectar-te al teu servidor web local **apache2**. Comprova que es visualitza la teva pàgina **nota.html**.

m) Atura el teu equip virtual. A la secció "Xarxa" de la màquina virtual canvia "Connectat a" a "Adaptador pont". A la secció "Nom" selecciona la teva targeta de xarxa wifi (o ethernet si treballes amb cable). A continuació reinicia l'equip virtual novament.

n) Obre un terminal i executa l'ordre: `ip -4 -br add show dev enp0s3` . Comprova si la interfície de la màquina virtual està activada i la seva adreça IP.

o) Comprova que també pots connectar-te al teu servidor web **apache2** i visualitzar **nota.html** utilitzant ara la teva **adreça IP** en comptes de **localhost**.

4- Realització de la pràctica - part 2: Descarregant el fitxer nota.tar.gz de la web del company de grup i comprovant la seva integritat

a) Estableix una connexió al servidor web **apache2** de l'equip virtual del teu company de grup a partir de la seva adreça **IP** i demana la seva pàgina **nota.html**.

b) A continuació:
- Crea una carpeta de nom **nota** dins del teu directori personal i després entra dins de la carpeta.
- Descarrega amb **wget** el fitxer **nota.tar.gz** del teu company de grup.
- Calcula el codi **hash MD5** del fitxer **nota.tar.gz** descarregat.
- Comprova la integritat del fitxer descarregat durant la transmissió. Has pogut verificar la integritat de nota.tar.gz.? Per què?

c) Finalment, descomprimeix i desempaqueta el fitxer **nota.tar.gz** i comprova que tens els fitxers **nota.c**, **calcul.c**, **dades.c**, **pantalla.c**, **nota.h**, **Makefile** i **README**. Comprova també que dins de **README** hi ha la nova data i el nom d'autor del teu company de grup.

Lliurament de l'activitat

1- Activitat en grups de 2 persones o individual.

2- S'ha de mostrar el funcionament de:

a) **Part 1 - Apartat d** → Comprovaré el nou fitxer README, el seu valor hash SHA-256, es compararà amb l'original i es demanarà respondre a la pregunta de l'apartat.
b) **Part 1 - Apartat f**
c) **Part 1 - Apartat g**
d) **Part 1 - Apartat o**
e) **Part 2 - Apartat b** → Comprovaré la descarrega, càlcul del hash i contestació de les preguntes
f) **Part 2 - Apartat c** → Comprovare descompressió, desempaquetament i README

3- Dates de lliurament: Setmana del **8-11-21** al **12-11-21** dins de l'horari escolar per aconseguir el **100%** de la nota. La setmana del **15-11-21** al **19-11-21** per aconseguir el **70%** de la nota. Després d'aquesta setmana, no s'acceptarà aquesta pràctica.