Seguretat i Protecció de dades - Pràctica 1

Marc Sànchez Pifarré

21 de setembre de 2019

Contents

Informe de la pràctica 1

Apartat a, Cèsar

Què s'ha tingut en compte :

- Donat n com a desplaçament, n >= 0.

Què no s'ha tingut en compte :

- Donat n com a desplaçament, n < 0
- Només es desplacen els caràcters "a"..."z" alfabet anglés.

Proves

```
entreu un nombre natural corresponent al desplacament: 0
entra el text que vols xifrar: a
TEXT XIFRAT:  a
TEXT ORIGINAL:  a
```

```
entreu un nombre natural corresponent al desplacament: 25
entra el text que vols xifrar: a
TEXT XIFRAT:  z
TEXT ORIGINAL:  a
```

```
entreu un nombre natural corresponent al desplacament: 26
entra el text que vols xifrar: a
TEXT XIFRAT:  a
TEXT ORIGINAL:  a
```

Amb aquestes proves l'algoritme queda testejat i estressat. Suficients per controlar els extrems.

Fem la prova amb un text del lorem Ipsum.

```
entreu un nombre natural corresponent al desplacament: 56253
```

```
entra el text que vols xifrar: Lorem Ipsum is simply dummy text of
↪  the printing and typesetting industry. Lorem Ipsum has been the
↪  industry's standard dummy text ever since the 1500s, when an
↪  unknown printer took a galley of type and scrambled it to make a
↪  type specimen book. It has survived not only five centuries, but
↪  also the leap into electronic typesetting, remaining essentially
↪  unchanged. It was popularised in the 1960s with the release of
↪  Letraset sheets containing Lorem Ipsum passages, and more
↪  recently with desktop publishing software like Aldus PageMaker
↪  including versions of Lorem Ipsum
```

```
TEXT XIFRAT:  Ldgtb Iehjb xh hxbean sjbbn itmi du iwt egxcixcv pcs
↪  inethtiixcv xcsjhign. Ldgtb Iehjb wph qttc iwt xcsjhign'h
↪  hipcspgs sjbbn itmi tktg hxcrt iwt 1500h, lwtc pc jczcdlc egxcitg
↪  iddz p vpaatn du inet pcs hrgpbqats xi id bpzt p inet hetrxbtc
↪  qddz. Ii wph hjgkxkts cdi dcan uxkt rtcijgxth, qji pahd iwt atpe
↪  xcid tatrigdcxr inethtiixcv, gtbpxcxcv thhtcixpaan jcrwpcvts. Ii
↪  lph edejapgxhts xc iwt 1960h lxiw iwt gtatpht du Ltigphti hwttih
↪  rdcipxcxcv Ldgtb Iehjb ephhpvth, pcs bdgt gtrtcian lxiw sthzide
↪  ejqaxhwxcv hduilpgt axzt Aasjh PpvtMpztg xcrajsxcv ktghxdch du
↪  Ldgtb Iehjb
```

```
TEXT ORIGINAL:  Lorem Ipsum is simply dummy text of the printing and
↪  typesetting industry. Lorem Ipsum has been the industry's
↪  standard dummy text ever since the 1500s, when an unknown printer
↪  took a galley of type and scrambled it to make a type specimen
↪  book. It has survived not only five centuries, but also the leap
↪  into electronic typesetting, remaining essentially unchanged. It
↪  was popularised in the 1960s with the release of Letraset sheets
↪  containing Lorem Ipsum passages, and more recently with desktop
↪  publishing software like Aldus PageMaker including versions of
↪  Lorem Ipsum
```

Aquest algoritme permet l'encriptament amb més d'una volta.

Per exemple :

- Encriptem "hola" amb desplaçament = 12 -> taxm.

- Encriptem “taxm” amb desplaçament = 5 -> yfcr
- Desencriptem “yfcr” amb desplaçament = 5 -> taxm
- Desencriptem “taxm” amb desplaçament = 12 -> hola

Apartat b, PolyBios

En aquest apartat hi havia més llibertat a l’hora d’implementar l’algoritme, en el meu cas he optat per fer-lo senzill per poder fer-lo el més genèric possible.

Es representa la taula com un vector v de n posicions on n és el nombre de caràcters de l’alfabet. Per cada posició del vector v s’hi insereix una tupla que conté 3 valors.

- lletra
- caràcter corresponent a la Fila de la taula.
- caràcter corresponent a la columna de la taula.

Definim un text t amb nombre de caràcters m. Llavors la complexitat de l’algoritme és O(n) * m per encriptar i desencriptar.

Propietats negatives :

- Complexitat, es podria haver actuat amb complexitat O(1) i accés directe peró complica el codi i embrut la genericitat a l’hora de generar la taula. (S’ha de controlar si les files son > columnes, si files < columnes o si files == columnes)

Propietats positives :

- Manteniment del codi.
- Escalabilitat en nombre de signes.
- Genericitat en funció del nombre de files i columnes al crear la taula.
- Alteració de la taula de manera senzilla.

S’ha implementat així per simplicitat i per poder adaptar la matriu de les transparències a la pràctica de manera fàcil i fent el menor “marranades” hardcoded possible.

Proves

```
Entreu el nombre de files i columnes, recorda dimensionar correctament
↪  la matriu => files >= 5 and columnes >= 5!
Entra el nombre de files : 5
```

```
Entra el nombre de columnes : 5
entra el text que vols xifrar: a
TEXT XIFRAT:  AA
TEXT ORIGINAL:  a
```

```
Entreu el nombre de files i columnes, recorda dimensionar correctament
↪  la matriu => files >= 5 and columnes >= 5!
Entra el nombre de files : 5
Entra el nombre de columnes : 5
entra el text que vols xifrar: z
TEXT XIFRAT:  EE
TEXT ORIGINAL:  z
```

```
Entreu el nombre de files i columnes, recorda dimensionar correctament
↪  la matriu => files >= 5 and columnes >= 5!
Entra el nombre de files : 6
Entra el nombre de columnes : 5
entra el text que vols xifrar: z
TEXT XIFRAT:  FA
TEXT ORIGINAL:  z
```

```
Entreu el nombre de files i columnes, recorda dimensionar correctament
↪  la matriu => files >= 5 and columnes >= 5!
Entra el nombre de files : 5
Entra el nombre de columnes : 6
entra el text que vols xifrar: z
TEXT XIFRAT:  EB
TEXT ORIGINAL:  z
```

Fins aquí hem testejat l'algoritme en els seus extrems, ara podem provar alguna possible col·lisió. L'únic cas en que colisionen és amb la i i la j sobre la configuració de 5 files i 5 columnes.

```
Entreu el nombre de files i columnes, recorda dimensionar correctament
↪  la matriu => files >= 5 and columnes >= 5!
Entra el nombre de files : 5
```

```
Entra el nombre de columnes : 5
entra el text que vols xifrar: i
TEXT XIFRAT:  BD
TEXT ORIGINAL:  i
```

```
Entreu el nombre de files i columnes, recorda dimensionar correctament
 ↪  la matriu => files >= 5 and columnes >= 5!
Entra el nombre de files : 5
Entra el nombre de columnes : 5
entra el text que vols xifrar: j
TEXT XIFRAT:  BD
TEXT ORIGINAL:  i
```

Fem la prova amb un text del lorem Ipsum.

```
Entreu el nombre de files i columnes, recorda dimensionar correctament
 ↪  la matriu => files >= 5 and columnes >= 5!
Entra el nombre de files : 5
Entra el nombre de columnes : 5
```

```
entra el text que vols xifrar: Lorem Ipsum is simply dummy text of
 ↪  the printing and typesetting industry. Lorem Ipsum has been the
 ↪  industry's standard dummy text ever since the 1500s, when an
 ↪  unknown printer took a galley of type and scrambled it to make a
 ↪  type specimen book. It has survived not only five centuries, but
 ↪  also the leap into electronic typesetting, remaining essentially
 ↪  unchanged. It was popularised in the 1960s with the release of
 ↪  Letraset sheets containing Lorem Ipsum passages, and more
 ↪  recently with desktop publishing software like Aldus PageMaker
 ↪  including versions of Lorem Ipsum.
```

```
TEXT XIFRAT:
```

```
CACDDBAECB BDCEDCDECB BDDC DCBDCBCECAED ADDECBCBED DDAEECDD CDBA
↪ DDBCAE CEDBBDCCDDBDCCBB AACCAD DDEDCEAEDCAEDDDDBDCCBB
↪ BDCCADDEDCDDDBED. CACDDBAECB BDCEDCDECB BCAADC ABAEAECC DDBCAE
↪ BDCCADDEDCDDDBED'DC DCDDAACCADAADBAD ADDECBCBED DDAEECDD AEEAAEDB
↪ DCBDCCACAE DDBCAE 1500DC, EBBCAECC AACC DECCBECCCDEBCC
↪ CEDBBDCCDDAEDB DDCDCDBE AA BBAACACAAEED CDBA DDEDCEAE AACCAD
↪ DCACDBAACBABCAAEAD BDDD DDCD CBAABEAE AA DDEDCEAE
↪ DCCEAEACBDCBAECC ABCDCDBE. BDDD BCAADC DCDEDBEABDEAAEAD CCCDDD
↪ CDCCCAED BABDEAAE ACAECCDDDEDBBDAEDC, ABDEDD AACADCCD DDBCAE
↪ CAAEAACE BDCCDDCD AECAAEACDDDBCDCCBDAC DDEDCEAEDCAEDDDDBDCCBB,
↪ DBAECBAABDCCBDCCBB AEDCDCAECCDDBDAACACAED DECCACBCAACCBBAEAD.
↪ BDDD EBAADC CECDCEDECAAADBBDDCAEAD BDCC DDBCAE 1960DC EBBDDDBC
↪ DDBCAE DBAECAAEAADCAE CDBA CAAEDDDBAADCAEDD DCBCAEAEDDDC
↪ ACCDCCDDAABDCCBDCCBB CACDDBAECB BDCEDCDECB CEAADCDCAABBAEDC,
↪ AACCAD CBCDDBAE DBAEACAECCDDCAED EBBDDDBC ADAEDCBEDDCDCE
↪ CEDEABCABDDCBCBDCCBB DCCDBADDEBAADBAE CABDBEAE AACAADDEDC
↪ CEAABBAECBAABEAEDB BDCCACCADEADBDCCBB EAAEDBDCBDCDCCDC CDBA
↪ CACDDBAECB BDCEDCDECB.
```

```
TEXT ORIGINAL:  lorem ipsum is simply dummy text of the printing and
↪ typesetting industry. lorem ipsum has been the industry's
↪ standard dummy text ever since the 1500s, when an unknown printer
↪ took a galley of type and scrambled it to make a type specimen
↪ book. it has survived not only five centuries, but also the leap
↪ into electronic typesetting, remaining essentially unchanged. it
↪ was popularised in the 1960s with the release of letraset sheets
↪ containing lorem ipsum passages, and more recently with desktop
↪ publishing software like aldus pagemaker including versions of
↪ lorem ipsum
```

Aquest algoritme tal i com està muntat no permet més d'una volta d'encriptat.

Apartat c, RailFence‘‘‘

Algoritme per transformació, no cal substituïr caràcters sinó simplement desordenar-los. En aquest cas s'ha optat per actuar amb tots els caràcters sense tenir en copmte si son o no lletres que pertanyen a l'alfabet anglés, es desordena tot!

S'utilitza un vector de list(). Es realitza així degut a que és molt més fàcil de codificar, a l'hora de crear el vector mitjançant l'entrada es fan appends a les llistes de dins del vector. Cada list() simbolitza un rail.

A l'hora de codificar és molt senzill i el mètode té una complexitat de O(n) sent n el nombre de caràcters a processar en funció de l'entrada. (Hi ha un doble bucle peró les iteracions acaben sumant n)

A l'hora de descodificar el nombre de iteracions segueix sent n tot i el doble bucle que hi ha. I és que es genera la taula de raïls a l'inversa de com es va construïr. Per poder fer-ho es requereix el nombre de rails amb el que es va codificar i l'habilitat de veure que n % rails ens donarà el nombre de rails que pot ser que tinguin una lletra de més.

Proves

```
entreu un nombre natural corresponent al Nombre de rails: 1
entra el text que vols xifrar: hola
TEXT XIFRAT: hola
TEXT ORIGINAL: hola
```

```
entreu un nombre natural corresponent al Nombre de rails: 4
entra el text que vols xifrar: hola
TEXT XIFRAT: hola
TEXT ORIGINAL: hola
```

```
entreu un nombre natural corresponent al Nombre de rails: 4
entra el text que vols xifrar: hol
TEXT XIFRAT: hol
TEXT ORIGINAL: hol
```

```
entreu un nombre natural corresponent al Nombre de rails: 4
entra el text que vols xifrar: hola hola
TEXT XIFRAT:  h aohloal
TEXT ORIGINAL:  hola hola
```

Aquí ja tenim testejat l'algoritme en els extrems. Quan hi ha 1 sol carril, quan hi ha tants carrils com caràcters, quan hi ha menys caràcters que carrils i finalment quan hi ha més caràcters que carrils. Provem amb un text de lorem ipsum.

```
entreu un nombre natural corresponent al Nombre de rails: 10
entra el text que vols xifrar: Lorem Ipsum is simply dummy text of
 ↪  the printing and typesetting industry. Lorem Ipsum has been the
 ↪  industry's standard dummy text ever since the 1500s, when an
 ↪  unknown printer took a galley of type and scrambled it to make a
 ↪  type specimen book. It has survived not only five centuries, but
 ↪  also the leap into electronic typesetting, remaining essentially
 ↪  unchanged. It was popularised in the 1960s with the release of
 ↪  Letraset sheets containing Lorem Ipsum passages, and more
 ↪  recently with desktop publishing software like Aldus PageMaker
 ↪  including versions of Lorem Ipsum.
```

```
TEXT XIFRAT: Lmyxpntt  i d es rk nea bad u le niiawrhileeimsde
 ↪  usiPivfso  trdirIbnsue ,ui oddksos fraeltgnanaieteten s ndbokane
 ↪  urid i nypedtmvt nnaf  epo niilaey,glgss
 ↪  hartiIamtelfegcrLmesuontg.seuamehwkt  si eksovespcp  le e1
 ↪  sasnpgolsit elso.m mfty  unsnyrehnegtctac.uteso tereydpd9tes
 ↪  gserykswAMuir sm ipiLm td   eorayr  i r  , irses .o 6h ec use
 ↪  thaladoeIiytneno trats1nw lpattmIvoc tnoemsu pi0eotoLm,
 ↪  woirdkinmpm hgsdrhhyrei5 ntlemoyetinebhtntaenIuns f no  ripneuens
 ↪  spte eueae'dxn0a oe b pn vlnueoitinctl  r strpaet g srg Iule
 ↪  atsms s tc0npoyalme heytt  cinth atweLhaeanchp l   op
```

```
TEXT ORIGINAL: Lorem Ipsum is simply dummy text of the printing and
 ↪  typesetting industry. Lorem Ipsum has been the industry's
 ↪  standard dummy text ever since the 1500s, when an unknown printer
 ↪  took a galley of type and scrambled it to make a type specimen
 ↪  book. It has survived not only five centuries, but also the leap
 ↪  into electronic typesetting, remaining essentially unchanged. It
 ↪  was popularised in the 1960s with the release of Letraset sheets
 ↪  containing Lorem Ipsum passages, and more recently with desktop
 ↪  publishing software like Aldus PageMaker including versions of
 ↪  Lorem Ipsum.
```

Apartat d, Main

No té gaire complicació, ara és utilitzar els algoritmes que s'han presentat anteriorment.

```
Entra el text que vols codificar i descodificar,
Si apretes INTRO fara automaticament el de l'apartat d.
text :
Amb el metode Cesar :
zk'j kyv yfebp kfeb nfdve kyrk xzddv, xzddv, xzddv kyv yfebp kfeb
 ↪  sclvj (yfebp kfeb nfdve, sp kyv ifcczex jkfevj)

-----------------------------------
Amb el metode PolyBios :
BDDD'DC DDBCAE BCCDCCBEED DDCDCCBE EBCDCBAECC DDBCAADD BBBDCBCBAE,
 ↪  BBBDCBCBAE, BBBDCBCBAE DDBCAE BCCDCCBEED DDCDCCBE ABCADEAEDC
 ↪  (BCCDCCBEED DDCDCCBE EBCDCBAECC, ABED DDBCAE DBCDCACABDCCBB
 ↪  DCDDCDCCAEDC)

-----------------------------------
Amb el metode RailFence :
ie oaeeeonsywbr )t tmt,, nk  oyos'hoe   tk (tm ltsonnggghybhoetlo nk
 ↪  iiie lonnhintk tmmm tunk,enehywhmmmhoek   gs

-----------------------------------
```

Anàlisi de desencriptació.

El primer que farem serà observar el text i cercar les propietats que ens permeti descartar certs algoritmes.

Text Proposat :

```
xfimr litvxl. patm tkx px ebobgz yhk? tutgwhgxw ietvxl. b znxll px dghp max lvhkx
hp fnlm zh hg, ur jnxxg)
```

Com a propietats interessants podem dir que trobem tant paraules parell com imparell en nombre de caràcters. Aquest fet ja **ens descarta l'algoritme PolyBios**.

Tot seguit si ens fixem en la distribució dels espais i dels signes de puntuació així com l'ordre dels parèntesis també ens n'adonem que si fos un algoritme de transformació hi hauria poques possibilitats que el text tingués tanta qüerència en relació a aquest fet i per tant també descartem l'algoritme RailFence.

Així doncs em quedaria parlar sobre l'algoritme cesar.

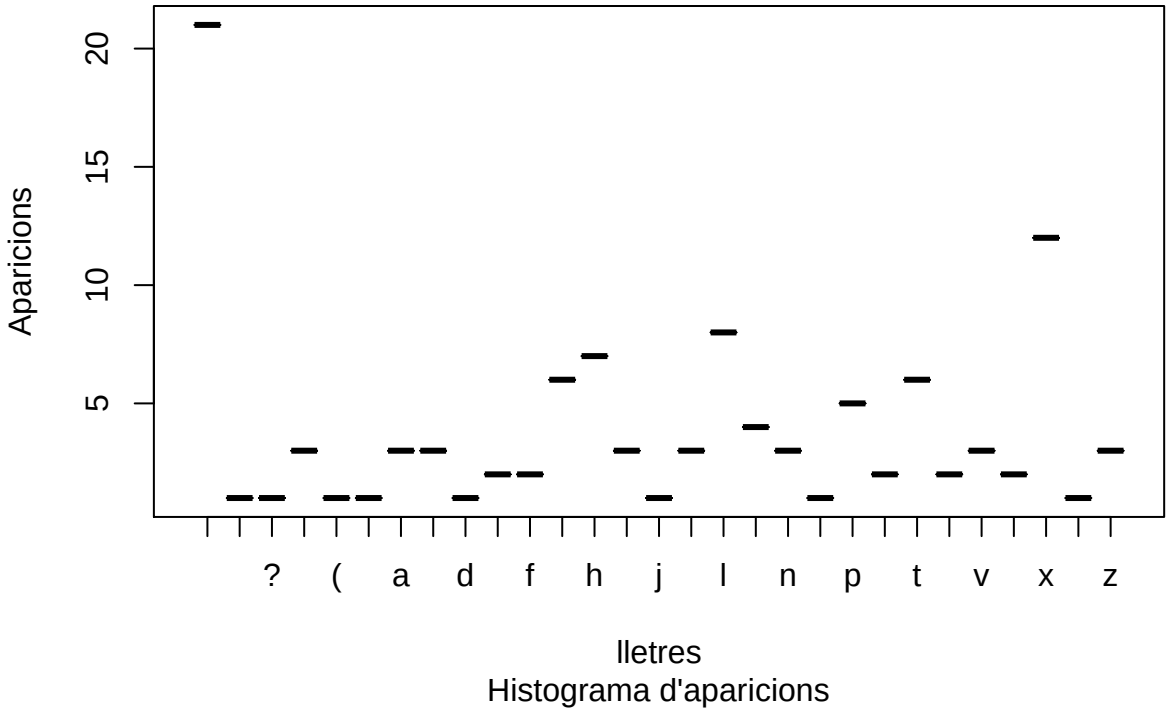
Mirem primerament el nombre de caràcters diferents amb els que ens enfrontem i fem l'histograma.

```
splitted <- as.data.frame(table(strsplit("xfimr litvxl. patm tkx px
↪  ebobgz yhk? tutgwhgxw ietvxl. b znxll px dghp max lvhkx. (la hp
↪  fnlm zh hg, ur jnxxg)", "")))
splitted
```

```
##    Var1 Freq
## 1         21
## 2     ,    1
## 3     ?    1
## 4     .    3
## 5     (    1
## 6     )    1
## 7     a    3
## 8     b    3
## 9     d    1
## 10    e    2
## 11    f    2
## 12    g    6
## 13    h    7
## 14    i    3
## 15    j    1
## 16    k    3
## 17    l    8
## 18    m    4
## 19    n    3
## 20    o    1
## 21    p    5
## 22    r    2
## 23    t    6
## 24    u    2
```

```
## 25     v     3
## 26     w     2
## 27     x    12
## 28     y     1
## 29     z     3
```

```
plot(splitted, type="h", ylab="Aparicions", xlab="lletres",
↪  sub="Histograma d'aparicions")
```



Veient aquest histograma es pot observar els caràcters més utiltizats en aquest text son l'espai (que no el tenim en compte), la lletra x i tot seguit els caràcters l, h, g i t per aquest ordre. Si observem l'histograma de les transparències de classe veiem que el caràcter més utilitzat és la e. Per tant la primera deducció que farem és igualar el desplaçament a la distància entre la e i la x.

Fem la prova doncs amb l'algoritme cèsar amb 19 de desplaçament :

```
$ python3
Python 3.6.8 (default, Aug 20 2019, 17:12:48)
[GCC 8.3.0] on linux
Type "help", "copyright", "credits" or "license" for more
↪  information.
>>> import Cesar
```

```
>>> Cesar.descodificaText("xfimr litvxl. patm tkx px ebobgz yhk?
↪  tutgwhgxw ietvxl. b znxll px dghp max lvhkx. (la hp fnlm zh hg,
↪  ur jnxxg)", 19)
'empty spaces. what are we living for? abandoned places. i guess we
↪  know the score. (sh ow must go on, by queen)'
```

The show must go on -> https://www.youtube.com/watch?v=qMUle97yGE4

Referències

- **Transparències Tema 1 - Seguretat i protecció de dades** : https://moodle2.udg.edu/pluginfile.php/1175534/mod_resource/content/3/SPD_T1_intro.pdf
- **Enunciat de la pràctica** : https://moodle2.udg.edu/pluginfile.php/1175524/mod_resource/content/3/exercici1.pdf
- **Tutorial de python** : https://moodle2.udg.edu/mod/url/view.php?id=779085
- **R Studio Split fuction** : https://www.rdocumentation.org/packages/base/versions/3.6.1/topics/strsplit
- **R Studio plot function** : https://www.rdocumentation.org/packages/graphics/versions/3.6.1/topics/plot