Mathématiques-Cryptographie

Anne Garcia-Sanchez

M2i cyber dev - CFA CCI Avignon

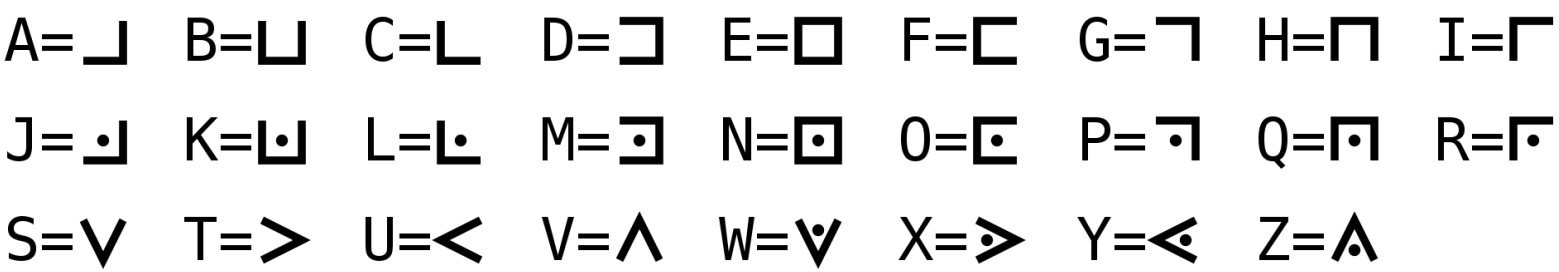
16 novembre 2023

Chiffrement par substitution mono-alphabétique

lettre de l'alphabet du message $\rightarrow$ autre lettre du même alphabet

clair	A	B	C	D	E	F	G	H	I	J	K	L	M
chiffré	D	E	F	G	H	I	J	K	L	M	N	O	P

lettre de l'alphabet du message $\rightarrow$ lettre d'un autre alphabet



Chiffrement par substitution mono-alphabétique

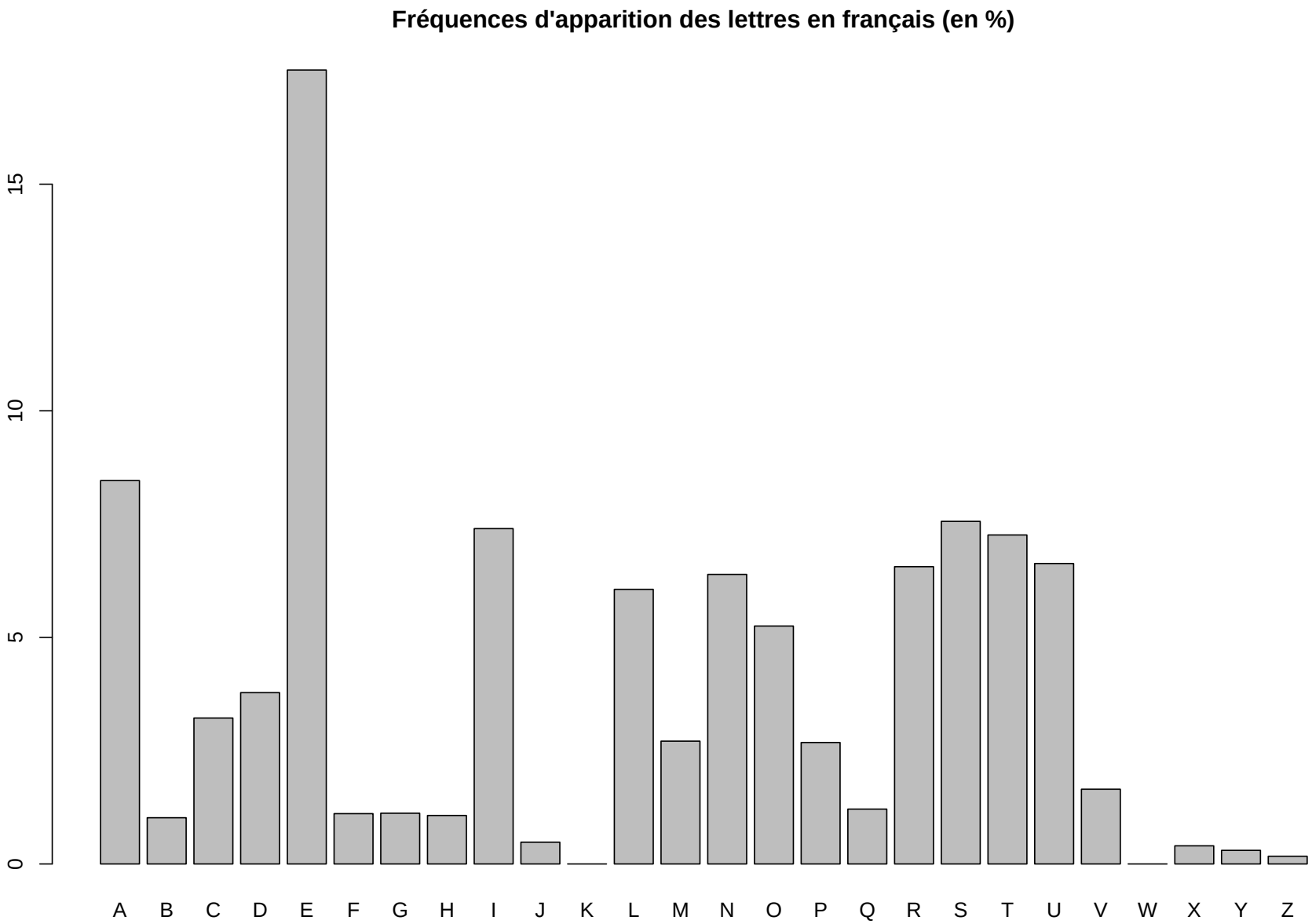
- Chiffrement par décalage - César - ROT13 - ROT47
- Chiffre Atbash, Atbah, Wolseley
- Chiffre Pigpen
- Carré de Polybe
- Chiffrement affine
- Cas général

Analyse fréquentielle

méthode de cryptanalyse décrite par le savant Al Kindi au 9ième siècle.

observation: dans chaque langue, certaines lettres ou combinaisons de lettres apparaissent plus fréquemment que d'autres

Analyse fréquentielle



Analyse fréquentielle

calcul des fréquences dépend des textes étudiés

ordre des lettres les plus fréquentes non défini

Fréquence des lettres

E	A	S	I	N	T	R	L	U	O	D	C	M
17.35%	8.2%	7.93%	7.53%	7.17%	6.99%	6.65%	5.92%	5.73%	5.53%	4.01%	3.33%	2.97%
P	V	G	F	Q	H	B	X	J	Y	Z	K	W
2.92%	1.39%	1.09%	1.08%	1.04%	0.93%	0.92%	0.47%	0.34%	0.31%	0.1%	0.06%	0.03%

Analyse fréquentielle

Fréquence des bigrammes

ES 3.05%	**LE** 2.22%	**DE** 2.17%	**RE** 2.1%	**EN** 2.08%	**ON** 1.64%	**NT** 1.62%	**ER** 1.53%	**TE** 1.52%	**ET** 1.43%	**EL** 1.42%	**AN** 1.37%	**SE** 1.32%
LA 1.29%	**AI** 1.24%	**NE** 1.14%	**OU** 1.12%	**QU** 1.11%	**ME** 1.08%	**IT** 1.06%	**IE** 1.05%	**EM** 1.01%	**ED** 1.01%	**UR** 1.01%	**IS** 0.99%	**EC** 0.95%
UE 0.92%	**TI** 0.9%	**RA** 0.86%	**NS** 0.84%	**IN** 0.84%	**TA** 0.82%	**CE** 0.81%	**AR** 0.8%	**EE** 0.79%	**EU** 0.78%	**SA** 0.76%	**CO** 0.74%	**EP** 0.71%
ND 0.7%	**IL** 0.7%	**SS** 0.68%	**ST** 0.66%	**SI** 0.65%	**TR** 0.64%	**AL** 0.64%	**UN** 0.63%	**PA** 0.62%	**AU** 0.61%	**EA** 0.6%	**AT** 0.58%	**MA** 0.58%
RI 0.58%	**SD** 0.57%	**SO** 0.57%	**US** 0.57%	**UI** 0.56%	**LL** 0.53%	**NC** 0.53%	**VE** 0.53%	**LI** 0.52%	**RO** 0.51%	**IO** 0.51%	**OR** 0.5%	**PE** 0.48%
OI 0.48%	**PR** 0.47%	**PO** 0.46%	**IR** 0.46%	**NA** 0.45%	**UT** 0.44%	**TD** 0.44%	**CH** 0.44%	**OM** 0.43%	**SP** 0.43%	**SL** 0.42%	**DA** 0.42%	**AS** 0.42%
MO 0.41%	**AC** 0.4%	**DI** 0.4%	**RS** 0.39%	**DU** 0.39%	**TL** 0.38%	**TO** 0.38%	**TS** 0.38%	**RT** 0.37%	**AM** 0.37%	**AP** 0.37%	**SC** 0.36%	**LO** 0.36%
AV 0.35%	**SU** 0.35%	**EV** 0.34%	**NO** 0.33%	**RL** 0.33%	**NI** 0.32%	**GE** 0.31%	**RD** 0.31%	**LU** 0.31%	**NN** 0.3%	**HE** 0.29%	**PL** 0.28%	**IQ** 0.28%
EF 0.28%	**MI** 0.27%	**VA** 0.27%	**TU** 0.27%	**VI** 0.27%	**CA** 0.27%	**EQ** 0.26%	**CI** 0.26%	**TT** 0.26%	**IC** 0.25%	**UX** 0.25%	**MM** 0.25%	**OL** 0.24%
AG 0.24%	**VO** 0.24%	**EI** 0.24%	**MP** 0.23%	**TP** 0.23%	**SM** 0.23%	**UL** 0.22%	**HA** 0.22%	**FI** 0.21%	**FA** 0.21%	**IM** 0.21%	**EG** 0.21%	**ID** 0.2%
DO 0.2%	**AD** 0.2%	**GR** 0.19%	**SQ** 0.19%	**AB** 0.19%	**BL** 0.18%	**UV** 0.18%	**IV** 0.18%	**NG** 0.18%	**TC** 0.17%	**IA** 0.17%	**OT** 0.17%	**CL** 0.17%
RC 0.17%	**RM** 0.17%	**OS** 0.17%	**OP** 0.16%	**CT** 0.16%	**FO** 0.16%	**UC** 0.16%	**UP** 0.16%	**RR** 0.16%	**JE** 0.16%	**HO** 0.16%	**UD** 0.15%	**CR** 0.15%
EB 0.15%	**EO** 0.15%	**IF** 0.15%	**FR** 0.14%	**RU** 0.14%	**UA** 0.14%	**NP** 0.14%	**IG** 0.14%	**BA** 0.14%	**BR** 0.14%	**OC** 0.14%	**CU** 0.14%	**FE** 0.13%
UM 0.13%	**EX** 0.13%	**BI** 0.13%	**BE** 0.13%	**GN** 0.13%	**MB** 0.13%	**AF** 0.12%	**HI** 0.12%	**EJ** 0.12%	**NF** 0.12%	**GI** 0.12%	**PP** 0.12%	**GA** 0.11%
FF 0.11%	**PU** 0.11%	**BO** 0.11%	**SF** 0.11%	**SR** 0.11%	**LS** 0.11%	**TQ** 0.11%	**OD** 0.1%	**PH** 0.1%	**TM** 0.1%	**DR** 0.1%	**NU** 0.1%	**NV** 0.1%

Analyse fréquentielle

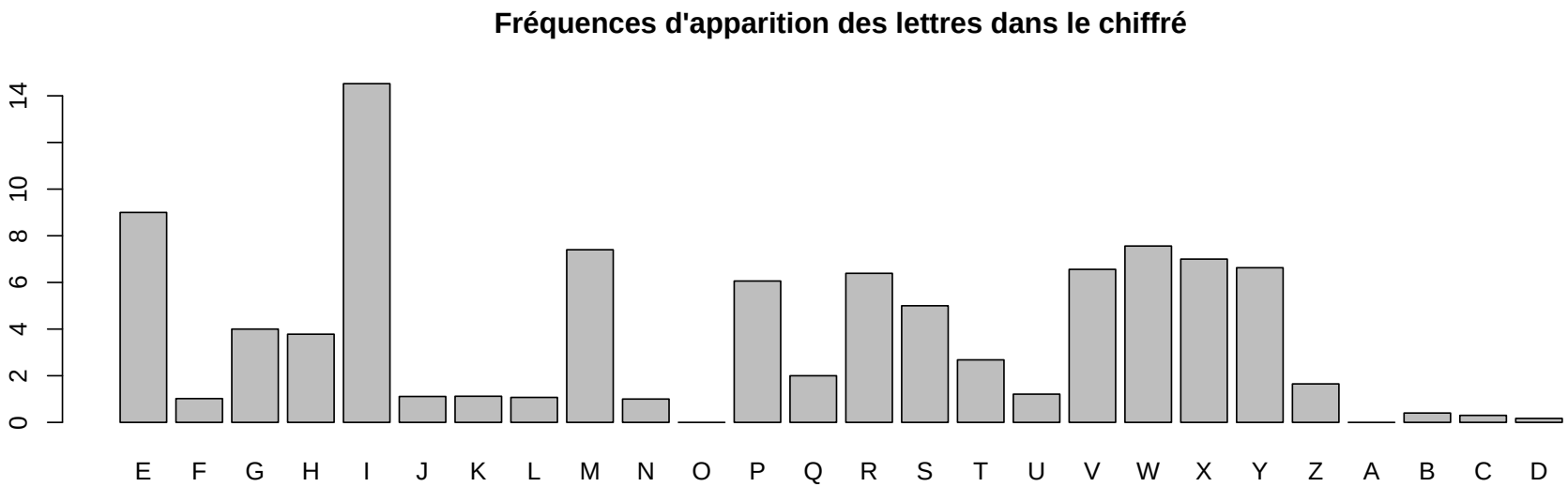
substitution mono-alphabétique : conservation des fréquences d'apparition des lettres

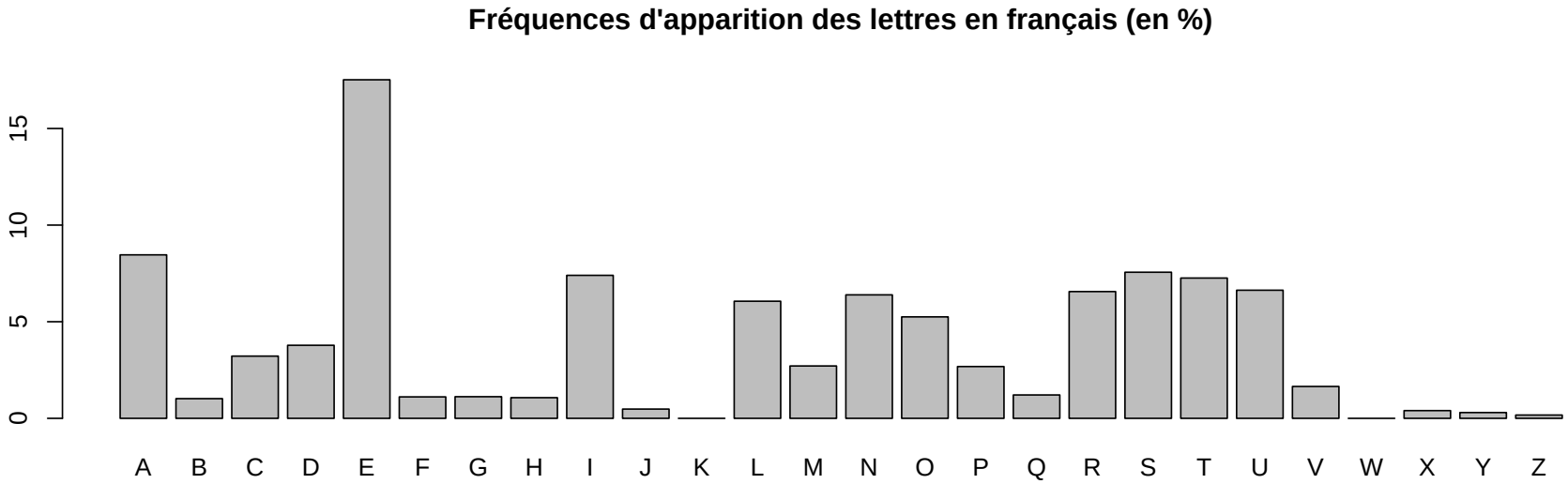
si texte chiffré suffisamment long :

recherche des lettres (ou symboles) ayant une fréquence élevée

permet de retrouver tout ou une partie du message clair

Analyse fréquentielle





Cryptanalyse du chiffre de César par analyse fréquentielle

recherche de la lettre la plus fréquente dans le chiffré

décalage entre le «e» et la lettre la plus fréquente révèle le décalage (la clé)

exemple:

texte chiffré comprenant des «i» en proportion supérieure aux autres lettres:

clé = rang de «i» - rang de «e» = 4

Cryptanalyse du chiffrement affine par analyse fréquentielle

2 clés à trouver

$$ax_1 + b \equiv y_1 \pmod{26}$$
$$ax_2 + b \equiv y_2 \pmod{26}$$

$$a(x_1 - x_2) \equiv (y_1 - y_2) \pmod{26}$$

$$a \equiv (x_1 - x_2)^{-1}(y_1 - y_2) \pmod{26}$$

éventuelle nécessité de tester plusieurs possibilités

Cryptanalyse d'une substitution quelconque par analyse fréquentielle

26! clés

brute force impossible

Cryptanalyse d'une substitution quelconque par analyse fréquentielle

calcul de fréquence de chaque caractère du texte

tâtonnement pour reconstruire le texte clair