Chapitre 3

Cryptanalyse de la cryptographie classique

3éme licence informatique

F.Boudjerida
Département Informatique
Université de Jijel

Plan

- ✓ Introduction
- ✓ Cryptanalyse par recherche des clés(Force brute, recherche exhaustive)
  - Cryptanalyse du chiffrement césar,
- ✓ Cryptanalyse par analyse de fréquences
  - Idée
  - Cryptanalyse du chiffrement arbitraire
  - Cryptanalyse du chiffrement de Vigenère

❑Cryptanalyse

- **Objectif** : Attaquer un système cryptographique. Un cryptosystème est dit **vulnérable** s'il est possible de :
  - décrypter des messages sans connaître la clé
  - encrypter des messages sans connaître la clé
  - trouver la clé

❑Classification des attaques

- **Attaques** : Elles peuvent être classifiées selon les informations disponibles aux cryptanalystes.

➢ **Attaque à texte chiffré**: l'analyste dispose de textes chiffrés *c1, … , cn* et cherche à trouver leurs correspondants en clair.

➢ **Attaque à texte clair**: l'analyste dispose de textes en clair *m1, ..., mn* et de leurs chiffrements *c1, ..., cn* respectifs et essaye de trouver la clé du cryptage ou de décrypter d'autres textes.

➢ **Attaque à texte clair choisi** : L'analyste peut choisir des textes clairs et obtenir leurs textes chiffrés correspondants.
En ayant ces connaissances, il essaye de trouver la clé du cryptage ou de décrypter d'autres textes.

➢ **Etc…**

❑Méthodologie de cryptanalyse

Techniques :

- Bien comprendre le système cryptographique en question
- Dégager ses propriétés
- Exploiter ses propriétés pour en déduire ses faiblesses

➢Recherche exhaustive de la clé (Brute force attack)

Idées :

- Un système cryptographique manipule un ensemble fini de clés (espace de clés)

- Si l'espace de clés est petit alors un adversaire peut les essayer une par une jusqu'à ce qu'il trouve la bonne

Recherche exhaustive de la clé

Exemple : Le chiffrement de César

Rappel : $e_k(x) = x + k \bmod 26$ *et* $d_k(x) = x - k \bmod 26$

Remarque : Il y'a 26 clés possibles == on peut rapidement les parcourir.

Exemple : Le message crypté est $C = JZCBM\ NWZKM$

$k = 0 \Rightarrow D_0(C) = JZCBM\ NWZKM$

$k = 1 \Rightarrow D_1(C) = IYBALMVYJL$

$k = 2 \Rightarrow D_2(C) = HXAZK\ LUXIK$

$k = 3 \Rightarrow D_3(C) = GWZYJ\ KTWHJ$

$k = 4 \Rightarrow D_4(C) = FVYXI\ JSVGI$

$k = 5 \Rightarrow D_5(C) = EUXWH\ IRUFH$

$k = 6 \Rightarrow D_6(C) = DTWVG\ HQTEG$

$k = 7 \Rightarrow D_7(C) = CSVUF\ GPSDF$

$k = 8 \Rightarrow D_8(C) = BRUTE\ FORCE$

Aha!!!··· j'ai trouvé la clé $k = 8$

Niveau de sécurité *théorique* :

– Alphabet à 26 lettres : 26! alphabets possibles.

Clairement, le chiffrement de César n'est pas sécuritaire.

➢Recherche exhaustive de la clé

Limites : Pour que cette technique soit réalisable, il faut que l'espace de clé ait une taille raisonnable.

Question : Peut-on appliquer cette technique, par exemple, sur le chiffrement du Vigenère avec une clé de taille 20 ?

Réponse : Non, il y a trop de clés à explorer 26^{20}

Question : C'est peut être trop pour un humain, mais est ce que c'est trop pour un ordinateur qui peut faire 2 milliard d'opérations par seconde (*2 GHz = 2X10^9* o.p.s.)?

NB: Pour le chiffrement arbitraire, le nombre de clés = 26**!**.
26**!** > 4X1026 (La recherche *exhaustive* est impossible).

➢Recherche exhaustive de la clé

Limites (suite) :

Réponse : Oui, c'est encore trop pour un ordinateur. En effet, avec un ordinateur de cette puissance (*2 GHz*), il faut :

$$\frac{26^{20}}{2\times10^{9}} \approx 9964074447604704576 \quad \text{sec}$$

$$\approx 1660679074600784 09 \quad \text{min}$$

$$\approx 27677984576679 73 \quad \textit{heures}$$

$$\approx 1153249357361 65 \quad \textit{jours}$$

$$\approx 3150954528 31 \quad \textit{années}$$

$$\approx 315 \quad \textit{milliard d'années}$$

➢Cryptanalyse **par analyse de fréquence**

Origine : Approche introduite par Abu Youssif Al-Kindi (**9 ème siècle)**

Idée :

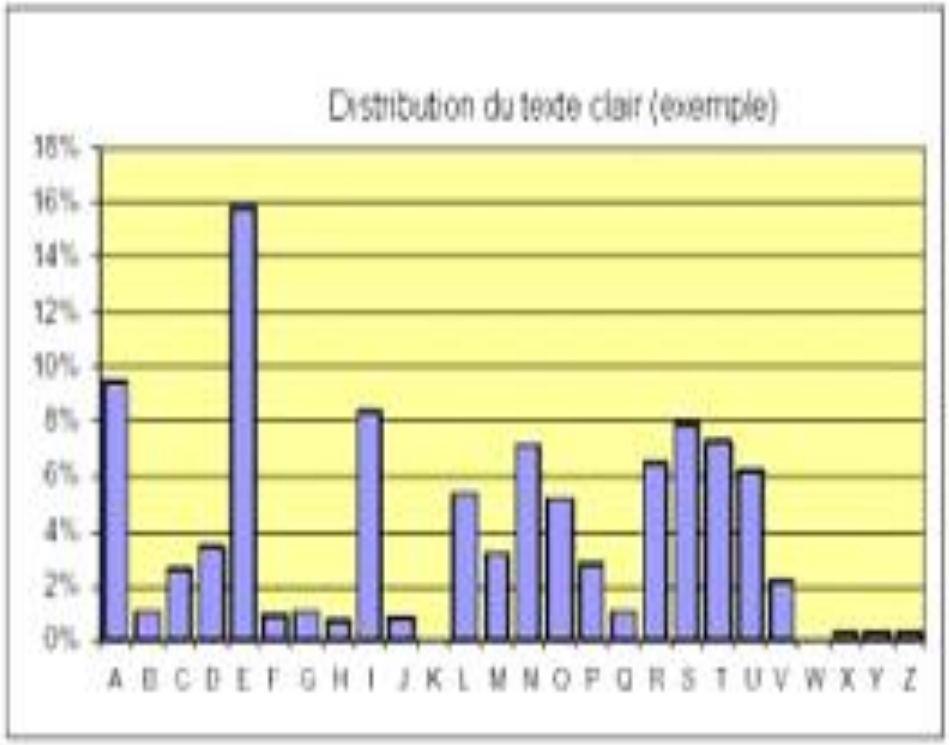
1. Établir la fréquence de chaque lettre de l'alphabet. En français la lettre la plus fréquente est « e » suivie par « a » puis par « i », etc
2. Examiner les fréquences des caractères dans le texte chiffré.
3. Remplacer les caractères les plus fréquents du texte chiffré par les caractères les plus fréquents du langage.
4. Si par exemple la lettre la plus fréquente du texte chiffré est « j », suivie par « m », suivi par « k », alors on fait un premier essai en remplaçant « j » par « e », « m » par « a » et « k » par « i ».

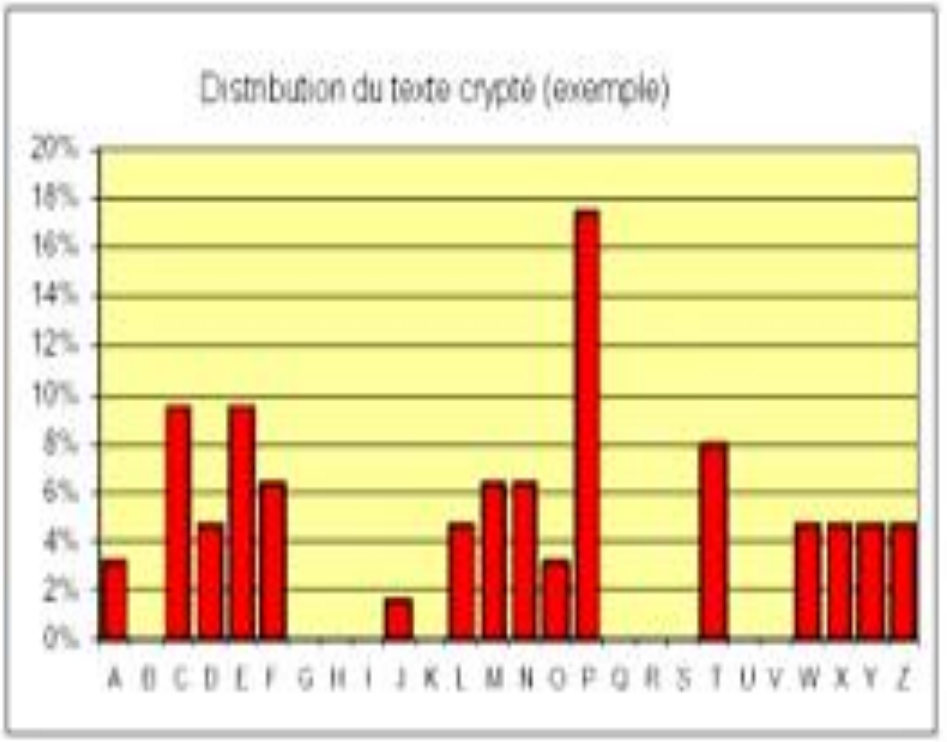
➢Cryptanalyse **par analyse de fréquence**

Remarques :

- Dans un texte quelconque chaque lettre a une fréquence d'apparition.
- Les probabilités de toutes les lettres définissent la distribution de ce texte.
- Soit D_C la distribution d'un texte clair.
- Soit D_E la distribution de son correspondant crypté.
- Le lien entre D_E et D_C dépend du cryptosystème.
- Un bon cryptosystème doit détruire le lien entre D_E et D_C.

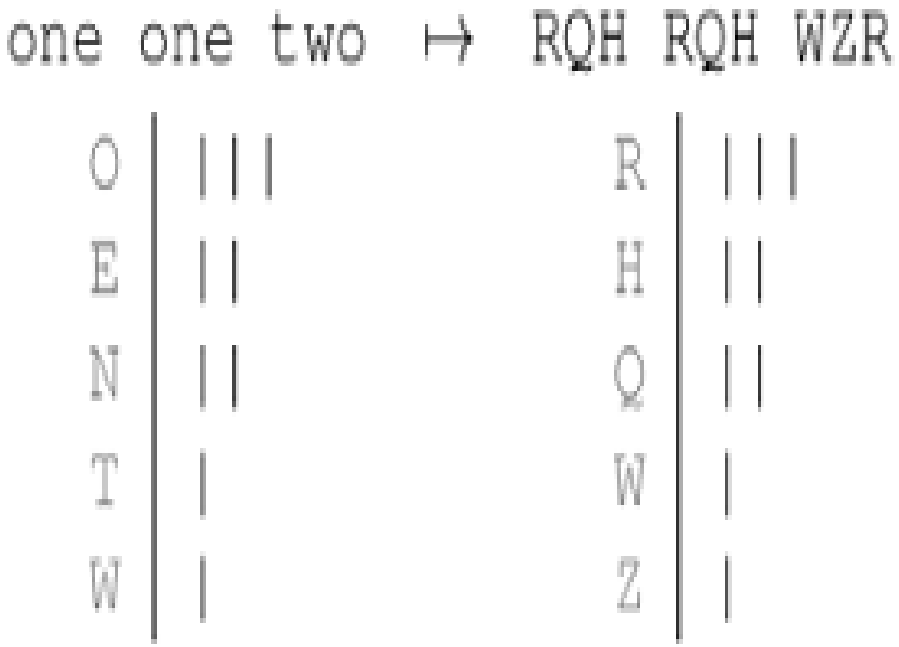
➢Cryptanalyse **par analyse de fréquence**





➢Cryptanalyse **par analyse de fréquence**

- Si le cryptosystème est une substitution monoalphabétique alors DE est une permutation de DC



➢Cryptanalyse **par analyse de fréquence**

Exemple: Substitution arbitraire

BQPSNRSJXJNJXLDPCLDLPQBE_QRKJXHNKPKSJPJIKSPUNBDKIQRB
KPQPBQPZITEJQDQBTSKPELNIUNPHNKPBKPCKSSQWKPSLXJPSNV
VXSQCCKDJPBLDWPXBPSNVVXJPGKPJKDXIPZLCEJKPGKSPSJQJXSJX
HNKSPGPLZZNIIKDZKPGKSPGXVVKIKDJKSPBKJJIKS

➢Cryptanalyse **par analyse de fréquence**

Exemple: Substitution arbitraire

En français

_	19.3	L	4.7	H	0.8
E	13.9	O	4.1	G	0.8
A	6.7	D	2.9	B	0.6
S	6.3	P	2.5	X	0.4
I	6.1	C	2.4	Y	0.3
T	6.1	M	2.1	J	0.3
N	5.6	V	1.3	Z	0.1
R	5.3	Q	1.3	K	0.0
U	5.2	F	0.9	W	0.0

Dans le cryptogramme

P	14.3	D	4.6	W	1.0
K	12.8	L	4.1	U	1.0
S	9.2	V	3.1	T	1.0
J	9.2	Z	2.6	_	0.5
X	5.6	G	2.6	O	0.0
Q	5.6	C	2.6	M	0.0
N	5.6	E	2.0	F	0.0
B	5.1	R	1.5	A	0.0
I	4.6	H	1.5	Y	0.0

➢Cryptanalyse **par analyse de fréquence**

Exemple: Substitution arbitraire

Remplaçons P par _ et K par E

BQ_SNRSJXJNJXLD_CLDL_QBE_QREJXHNE_ESJ_JIE
S_UNBDEIQRBE_Q_BQ_ZITEJQDQBTSE_ELNIUN_HN
E_BE_CESSQWE_SLXJ_SNVVXSQCCEDJ_BLDW_XB_
SNVVXJ_GE_JEDXI_ZLCEJE_GES_SJQJXSJXHNES_G
_LZZNIIEDZE_GES_GXVVEIEDJES_BEJJIES

Remplaçons Q par A et B par L

LA_SNRSJXJNJXLD_CLDL_ALE_AREJXHNE_ESJ_JIES
_UNLDEIARLE_A_LA_ZITEJADALTSE_ELNIUN_HNE_L
E_CESSAWE_SLXJ_SNVVXSACCEDJ_LLDW_XL_SNV
VXJ_GE_JEDXI_ZLCEJE_GES_SJAJXSJXHNES_G_LZ
ZNIIEDZE_GES_GXVVEIEDJES_LEJJIES

➢Cryptanalyse **par analyse de fréquence**

Exemple: Substitution arbitraire

Remplaçons S par S et G par D

LA_SNRSJXJNJXLD_CLDL_ALE_AREJXHNE_ESJ_JIES
_UNLDEIARLE_A_LA_ZITEJADALTSE_ELNIUN_HNE_L
E_CESSAWE_SLXJ_SNVVXSACCEDJ_LLDW_XL_SNV
VXJ_DE_JEDXI_ZLCEJE_DES_SJAJXSJXHNES_D_LZZ
NIIEDZE_DES_DXVVEIEDJES_LEJJIES

Remplaçons J par T et I par R

LA_SNRSTXTNTXLD_CLDL_ALE_ARETXHNE_EST_TR
ES_UNLDERARLE_A_LA_ZRTETADALTSE_ELNRUN_H
NE_LE_CESSAWE_SLXT_SNVVXSACCEDT_LLDW_XL_
SNVVXT_DE_TEDXR_ZLCETE_DES_STATXSTXHNES_
D_LZZNRREDZE_DES_DXVVEREDTES_LETTRES

➢Cryptanalyse **par analyse de fréquence**

Exemple: Substitution arbitraire

Remplaçons X par I, H par Q et N par U

LA_SU**R**STITUTI**LD**_**CLDL**_AL**E**_A**R**ETIQUE_EST_TRES
_**U**UL**D**ERA**R**LE_A_LA_**Z**R**TE**TA**D**AL**T**SE_**EL**UR**U**U_QUE
_LE_**C**ESSA**W**E_S**L**IT_SU**VV**ISA**CC**E**D**T_L**LDW**_IL_SU**V**
VIT_DE_TE**D**IR_**ZLCE**TE_DES_STATISTIQUES_D_**LZZ**U
RRE**DZ**E_DES_DI**VV**ERE**D**TES_LETTRES

Remplaçons V par F et D par N

LA_SU**RS**TITUTI**LN**_**CLNL**_AL**E**_A**R**ETIQUE_EST_TRES
_**U**ULNERA**R**LE_A_LA_**Z**R**TE**TANAL**T**SE_**EL**UR**U**U_QUE
_LE_**C**ESSA**W**E_S**L**IT_SUFFISA**CC**ENT_L**L**N**W**_IL_SUFF
IT_DE_TENIR_**ZLCE**TE_DES_STATISTIQUES_D_**LZZ**UR
REN**Z**E_DES_DIFFERENTES_LETTRES

➢Cryptanalyse **par analyse de fréquence**

Exemple: Substitution arbitraire

Remplaçons R par B et L par O

LA_SUBSTITUTION_**C**ONO_AL**E**_A**R**ETIQUE_EST_TRE
S_**U**ULNERABLE_A_LA_**Z**R**TE**TANAL**T**SE_**E**OUR**U**U_QU
E_LE_**C**ESSA**W**E_SOIT_SUFFISA**CC**ENT_LON**W**_IL_SU
FFIT_DE_TENIR_**Z**O**CE**TE_DES_STATISTIQUES_D_O**Z**
ZURREN**Z**E_DES_DIFFERENTES_LETTRES

Finalement

LA_SUBSTITUTION_MONO_ALPHABETIQUE_EST_TRE
S_VULNERABLE_A_LA_CRYPTANALYSE_POURVU_QU
E_LE_MESSAGE_SOIT_SUFFISAMMENT_LONG_IL_SU
FFIT_DE_TENIR_COMPTE_DES_STATISTIQUES_D_OC
CURRENCE_DES_DIFFERENTES_LETTRES

Cryptanalyse **par analyse de fréquence**

- Si le cryptosystème est une substitution polyalphabétique alors D_E est différent de D_C

```
clueclueclueclueclueclue
teleconference ↦  VPFIEZHJGCYREP
```

E	4
C	2
N	2
F	1
L	1
O	1
T	1

E	2
P	2
C	2
F	1
G	1
H	1
I	1
J	1
R	1
V	1
Y	1
Z	1

➢Cryptanalyse d'un chiffre de Vigenère

KQOWE FVJPU JUUNU KGLME KJINM WUXFQ MKJBG WRLFN FGHUD
WUUMB SVLPS NCMUE KQCTE SWREE KOYSS IWCTU AXYOT APXPL
WPNTC GOJBG FQHTD WXIZA YGFFN SXCSE YNCTS SPNTU JNYTG
GWZGR WUUNE JUUQE APYME KQHUI DUXFP GUYTS MTFFS HNUOC
ZGMRU WEYTR GKMEE DCTVR ECFBD JQCUS WVBPN LGOYL SKMTE
FVJJT WWMFM WPNME MTMHR SPXFS SKFFS TNUOC ZGMDO EOYEE
KCPJR GPMUR SKHFR SEIUE VGOYC WXIZA YGOSA ANYDO EOYJL WUNHA
MEBFE LXYVL WNOJN SIOFR WUCCE SWKVI DGMUC GOCRU WGNMA
AFFVN SIUDE KQHCE UCPFC MPVSU DGAVE MNYMA MVLFM AOYFN
TQCUA FVFJN XKLNE IWCWO DCCUL WRIFT WGMUS WOVMA TNYBU
HTCOC WFYTN MGYTQ MKBBN LGFBT WOJFT WGNTE JKNEE DCLDH
WTVBU VGFBI JG

➢Cryptanalyse d'un chiffre de Vigenère

Phase 1 : Trouver la longueur de la clé

Étape 1 : Soulignez les répétitions de 3 caractères ou plus :

KQOWE FVJPU JUUNU KGLME KJINM WUXFQ MKJBG WRLFN FGHUD
WUUMB SVLPS NCMUE KQCTE SWREE KOYSS IWCTU AXYOT APXPL
WPNTC GOJBG FQHTD WXIZA YGFFN SXCSE YNCTS SPNTU JNYTG
GWZGR WUUNE JUUQE APYME KQHUI DUXFP GUYTS MTFFS HNUOC
ZGMRU WEYTR GKMEE DCTVR ECFBD JQCUS WVBPN LGOYL SKMTE
FVJJT WWMFM WPNME MTMHR SPXFS SKFFS TNUOC ZGMDO EOYEE
KCPJR GPMUR SKHFR SEIUE VGOYC WXIZA YGOSA ANYDO EOYJL
WUNHA MEBFE LXYVL WNOJN SIOFR WUCCE SWKVI DGMUC GOCRU
WGNMA AFFVN SIUDE KQHCE UCPFC MPVSU DGAVE MNYMA MVLFM
AOYFN TQCUA FVFJN XKLNE IWCWO DCCUL WRIFT WGMUS WOVMA
TNYBU HTCOC WFYTN MGYTQ MKBBN LGFBT WOJFT WGNTE JKNEE
DCLDH WTVBU VGFBI JG

➢Cryptanalyse d'un chiffre de Vigenère

Étape 2 : Pour chaque répétition, mesurer la période

Séquence répétée	Distance
WUU	95
EEK	200
WXIZAYG	190
NUOCZGM	80
DOEOY	45
GMU	90

➢Cryptanalyse d'un chiffre de Vigenère

Étape 3 : Pour chaque période, décomposer en facteurs premiers et regarder quel facteur est commun à tous :

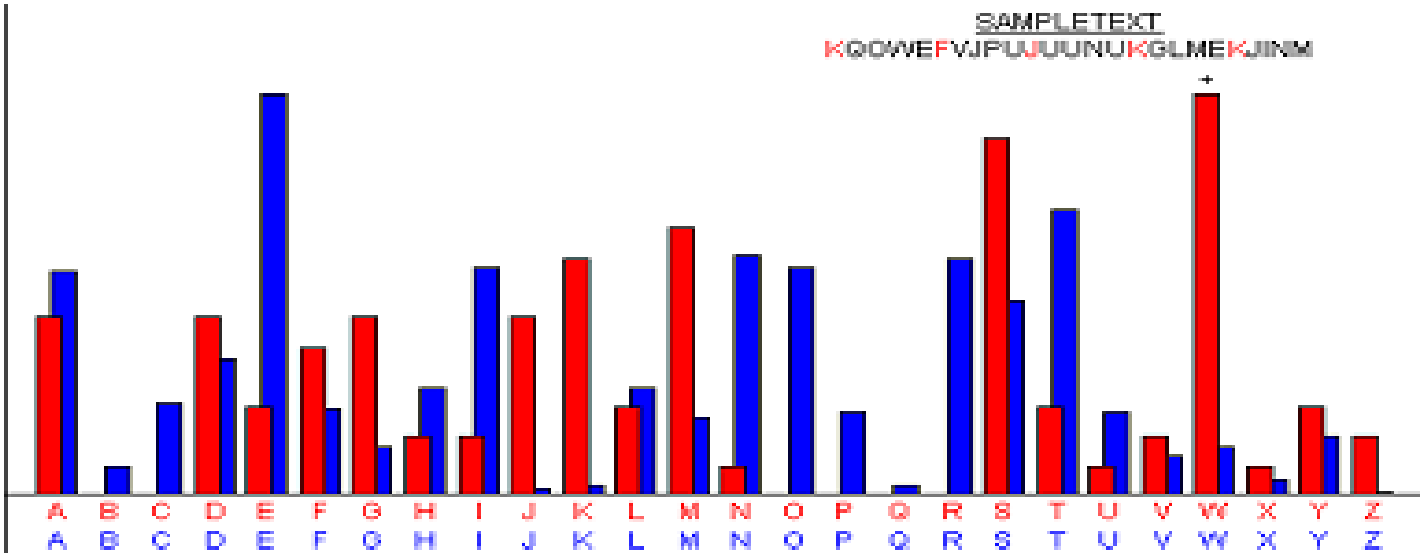
La clé est ici longue de 5 caractères.

		Longueurs de clef possibles			
Séquence répétée	Espace de répétition	2	3	5	19
WUU	95			x	x
EEK	200	x		x	
WXIZAYG	190	x		x	x
NUOCZGM	80	x		x	
DOEOY	45		x	x	
GMU	90	x	x	x	

➢Cryptanalyse d'un chiffre de Vigenère

Phase 2 : Trouver la 1er lettre du mot clé

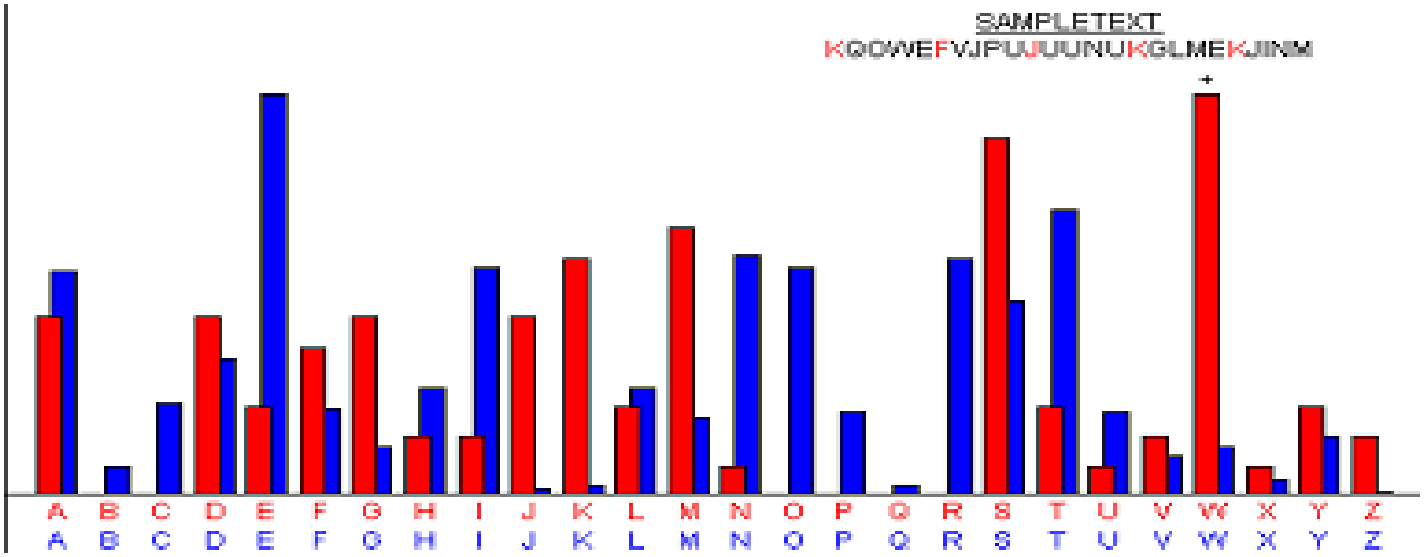
Étape 1 : Faire une analyse de fréquence
seulement sur les caractères 1, 6, 11, ...
On obtient ici :



➢Cryptanalyse d'un chiffre de Vigenère

En rouge, l'analyse de fréquence « modulo 5 »
En bleu le diagramme de fréquence des lettres en français.

Étape 2 : On décale pour faire correspondre
On décale les diagrammes pour mettre le pic du W sur le E
... L'ensemble correspond à peut prés : On à la première lettre de la clé.

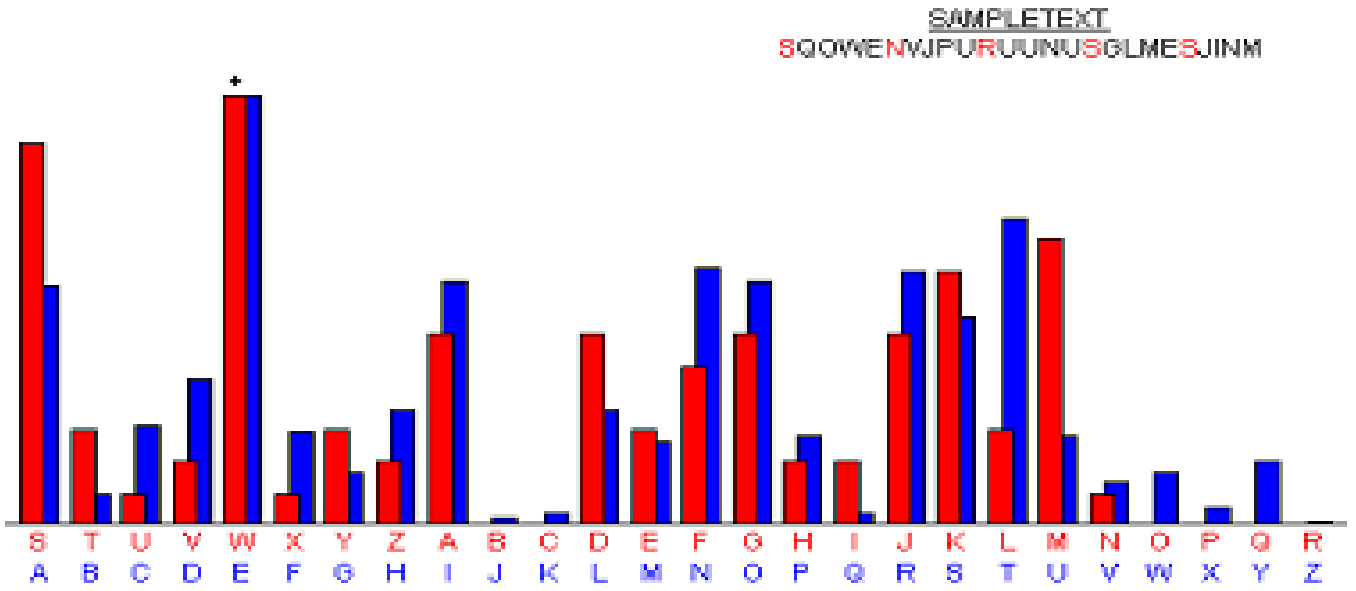


➢Cryptanalyse d'un chiffre de Vigenère

Avec W = 23 et E = 5, c'est la (23 – 5 + 1= 19_{eme}) soit S

Phase 3, 4, 5 et 6 : On recommence pour avoir les 5 lettres du mot clé.

Le mot clé est SCUBA



➢Cryptanalyse d'un chiffre de Vigenère

On peut déchiffrer le cryptogramme :
Soit encore :
Souvent pour s'amuser les hommes d'équipage prennent des albatros, vastes oiseaux des mers, qui suivent, indolents compagnons de voyage, le navire glissant sur les gouffres amers.
A peine les ont-ils déposés sur les planches que ces rois de l'azur, maladroits et honteux, laissent piteusement leurs grandes ailes blanches, comme des avirons, traîner à côté d'eux.
Ce voyageur ailé, comme il est gauche et veule, lui naguère si beau, qu'il est comique et laid. L'un agace son bec avec un brûle-gueule, l'autre mime en boitant l'infirme qui volait.
Le poète est semblable au prince des nuées, qui hante la tempête et se rit de l'archer.
Charles Baudelaire

➢Limites de l'analyse de fréquences

Échantillon représentatif

- On a besoin d'un texte assez long pour que les statistiques soient représentatives.
- Certains textes n'obéissent pas aux lois des fréquences.

- Dans le texte suivant la lettre Z, généralement la moins utilisée dans les textes en français, est la plus utilisée.

«*De Zanzibar à la Zambie et au Zaïre, des zones d'ozone font courir les zèbres en zigzags zinzins* »

- Georges Perec a écrit tout un livre *La dispariation* sans utiliser la lettre **e**