Criptologia clásica

Roberto Gómez Cárdenas
rogomez@itesm.mx
http://webdia.cem.itesm.mx/ac/rogomez

1

La criptografia en la historia

Escrito
Sistemas a mano
Mecanico
Electro-Mecanico
Electrónica
Programa almacenado
Llave pública

Fuente: Xerox Palo Alto Research Center

2

Procedimientos clásicos de encripción

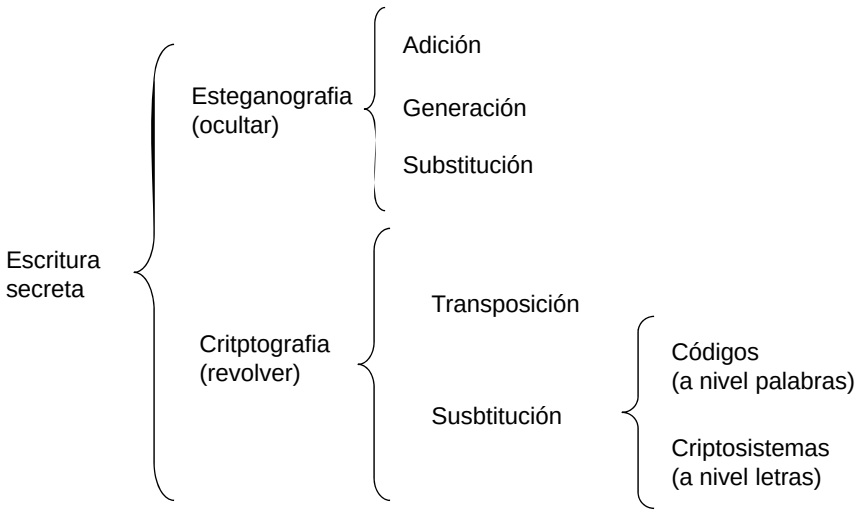
- Primeros metodos criptograficos
  - epoca romana hasta siglo XX
- Basados en dos técnicas
  - transposición
  - substitución

3

Procedimientos clásicos de encripción

- Transposición
  - “barajar” los símbolos del mensaje original colocandolos en un orden distinto, de manera que el criptograma contenga los mismos elementos del texto claro, pero colocados de tal forma que resulten incomprensibles.
- Sustitución
  - establecer correspondencia entre las letras del alfabeto en el que está escrito el mensaje original y los elementos de otro conjunto que puede ser el mismo o distinto alfabeto.

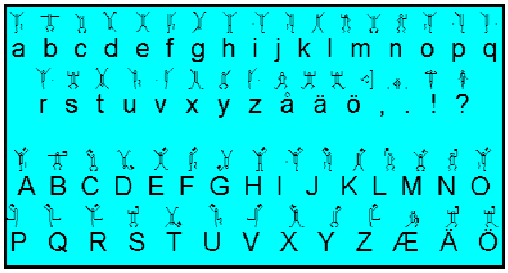
Una primera clasificación



5

Criptosistemas basados en sustitución

Establecer correspondencia entre las letras del alfabeto en el que está escrito el mensaje original y los elementos de otro conjunto que puede ser el mismo o distinto alfabeto.



6

Primer criptosistema sustitución

- Una de las más viejas descripciones de encripción por sustitución aparece en el Kama-Sutra
  - texto escrito en el siglo IV D.C. por Brahmin scholar Vatsyayana, basado en manuscritos que datan del siglo IV A.C.
- Kama-Sutra recomienda que la mujer debe estudiar 64 artes,
  - cocina, vestido, masaje y preparación perfumes
  - tambien incluye: ajedrez, carpinteria
  - número 45 de la lista: *mlecchita-vikalpa*, el arte de escritura secreta

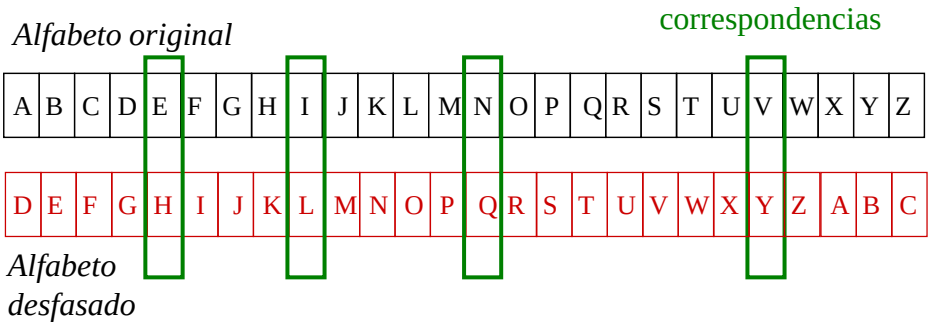
7

Criptosistema de Cesar

- Sustituye primera letra del alfabeto A, por la cuarta D; la segunda, B, por la quinta E, etc.
- También conocido como Ceaser shift cipher
- Dos alfabetos:
  - alfabeto texto plano/claro: el alfabeto usado para escribir el mensaje original (texto claro)
  - alfabeto criptosistema (o de encripción): las letras de este alfabeto substituyen a las letras del alfabeto texto claro

8

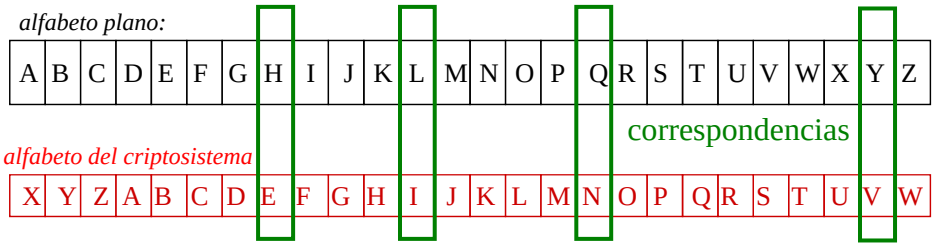
Ejemplo encripción Cesar



Mensaje:	VENI	VIDI	VICI
Llave:	DDDD	DDDD	DDDD
Criptograma:	YHQL	YLGL	YLFL

9

Ejemplo decripción Cesar



Criptograma:	YHQL	YLGL	YLFL
Llave:	DDDD	DDDD	DDDD
Mensaje:	VENI	VIDI	VICI

10

Una variante de Cesar

- Usar como llave una palabra que no tenga letras repetidas.
- La palabra es el inicio del alfabeto de encripción y el resto son las letras del alfabeto en orden creciente.

Alfabeto texto claro:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Llave: Murcielago

Alfabeto de encripción

M	**U**	**R**	**C**	**I**	**E**	**L**	**A**	**G**	**O**	B	D	F	H	J	K	N	P	Q	S	T	V	W	X	Y	Z

11

Criptosistemas monoalfabeticos

- Administradores arabes usaban criptosistemas parecido al de Cesar.
- También usaban criptosistemas que contenían otros tipos de símbolos:
  - a puede ser reemplazada por #
  - b puede ser reemplazada por +
- Un *criptosistema de substitución monoalfabético* es el nombre que se le da a cualquier criptosistema en el cual el alfabeto del criptosistema consiste de letras o símbolos, o una combinación de los dos.

12

Otros criptosistemas monoalfabeticos

- El criptosistema de Bacon
- El Polybius square
- Checker board
- Pigpen
- Atbash
- Ejemplos en la literatura
  - Arthur Conan Doyle
  - Edgar Allan Poe

13

El criptosistema de Bacon

- Francis Bacon (1521-1626) y su fraternidad “Rosicrucian” hizo uso de diferentes tipos de criposistemas.
- Uno de los más simples utiliza grupos de cinco letras y el resultado es un criptosistema monoalfabetico de caracteres
- Utiliza un criptosistema de 24 letras con I, J, U y W usados intercambiadamente

14

Ejemplo criptosistema Bacon

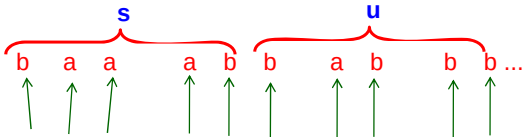
A = aaaaa	I/J= abaaa	R = baaaa
B = aaaab	K= abaab	S = baaab
C= aaaba	L = ababa	T = baaba
D= aaabb	M = bbabb	U/V= babbb
E= aabaa	N = bbbaa	W = babba
F= aabab	O = bbbab	X = babab
G= aabba	P = bbbba	Y = babba
H= aabbb	Q= bbbbb	Z = babbb

Texto claro: s u c c e s s

baaab babbb aaaba aaaba aabaa baaab baaab

Para encriptar: letra inicial de cada letra indica a o b A - M=a , N-Z=b

s = baaab
u = babbb



Criptograma We are looking for you where is possible to take ...

15

El criptosistema Polybius Square

- Polybius era un escritor anciano Griego que propuso substituir las letras con números de dos dígitos.
- El alfabeto es escrito en una matriz de 5 x 5 con los renglones y columnas numeradas:

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

16

Encriptando en Polybius Square

- Para encriptar sustituir cada letra con las coordenadas de la letra en la matriz.
- Por ejemplo:

F: es igual a 21

R E N A C I M I E N T O
42 15 33 11 13 24 32 24 15 33 44 34

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I/J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

- Es más un código que un criposistema
- Fue usado para transmitir mensajes de larga distancia en la noche usando antorchas.

17

Checker Board

- Similar al de Polybius
- La llave esta constituida por tres palabras
  - dos están en la columna y renglón de la matriz de 5 x 5 y forman las coordenadas de las entradas
  - estas palabras no pueden contener palabras similares para proporcionar coordenadas únicas
- Adentro de la matriz se escribe el alfabeto
  - la tercera palabra se escribe adentro de la matriz y se rellenan las otras celdas con el resto del alfabeto
- La encripción es igual que en Polybius

18

Ejemplo de Checker board

	G	R	O	U	P
W	B	O	A	R	D
H	C	E	F	G	H
I	I/J	K	L	M	N
T	P	Q	S	T	U
E	V	W	X	Y	Z

Texto claro: p r o t o c o l
Criptograma: TGWUWRTUWRHGWRIO

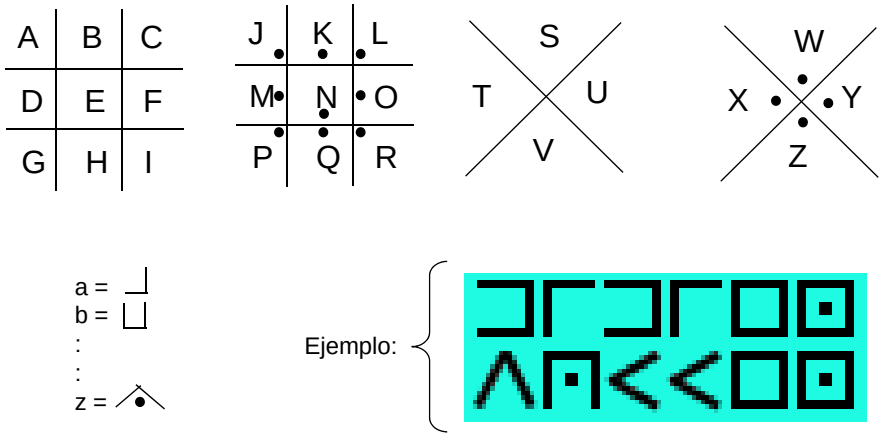
19

El criptosistema de Pigpen

- Criptosistema de tipo monoalfabetico
- Usado por los Freemasons en los 1700s para mantener sus registros secretos y aun es usado hoy en día por los niños escolares
- No sustituye una letra por otra, en lugar de estos substituye cada letra por un símbolo de acuerdo a un patrón determinado.

20

El alfabeto



21

Criptosistema Atbash

- Se toma cada letra, se calcula el número de lugares que lo separan de la primera letra del alfabeto.
- Se reemplaza con una letra que se encuentra en la misma distancia del final del alfabeto.
- En alfabeto español equivale a reemplazar:
  - la letra a, al principio alfabeto, por la letra z
  - la letra b se cambia por la letra y

22

Ejemplo Atbash

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Z	Y	X	W	V	U	T	S	R	Q	P	O	N	M	L	K	J	I	H	G	F	E	D	C	B	A

Texto claro: MENSAJE DE PRUEBA

Criptograma: NVMHZQV WV KIFVYZ

23

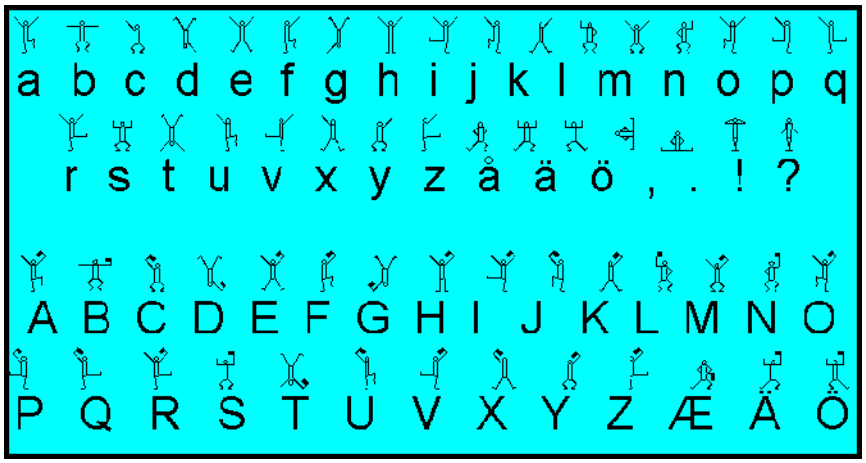
Primer ejemplo literatura: Arthur Conan Doyle

- Conan Doyle 1859 1930 es el autor de las aventuras de Sherlock Holmes
- Sherlock era un experto criptoanalista y le dice a Watson:

  I am the author of a trifling monograph upon the subject in wich I analyze one hundred and sixty separate ciphers.

- Su más fomoso criptoanalisis: Adventure of the Dancing Men

Criptosistema de Adventures Dancing Men



25

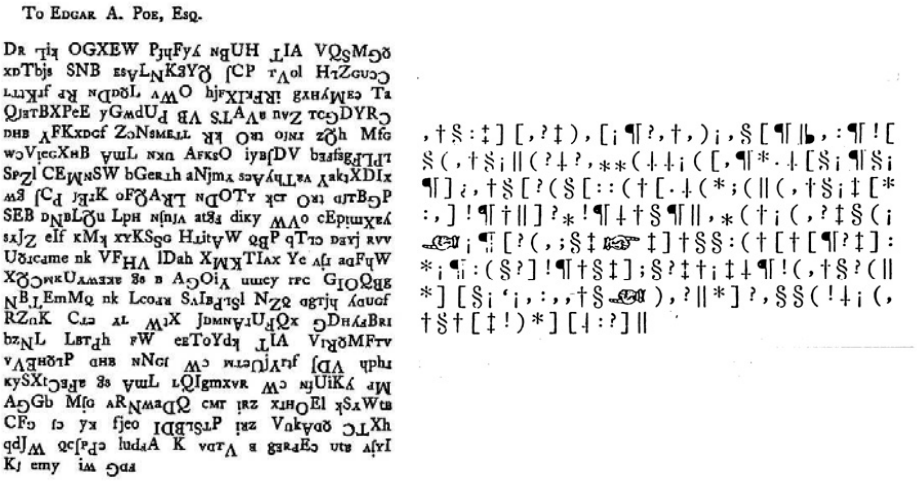
Segundo ejemplo: Edgar Allan Poe

- Poe (1809-1849) tenía interes en la cripotografía.
- Escribe para el Alexander Weekly Messenger de Filadelfia desafía a los lectores diciendo que puede resolver cualquier criptograma.
- En 1843 escribe un historia corta acerca de criptosistemas: The Gold Bug

53++!305))6*;4826)4+.)4+);806*;48!8`60))85;]8*:+*8!83(88)5*!;
46(;88*96*?;8)*+(;485);5*!2:*+(;4956*2(5*-4)8`8*; 4069285);)6
!8)4++;1(+9;48081;8:8+1;48!85;4)485!528806*81(+9;48;(88;4(+?3
4;48)4+;161;:188;+?;

26

Los retos de Allan Poe



To Edgar A. Poe, Esq.

27

Atacando los criptosistemas monoalfabéticos

El criptoanálisis arábe

28

Criptoanálisis arabe

- Estudiantes arábes capaces destruir criptosistemas.
  - en realidad ellos inventaron el criptoanálisis
- Disciplinas para criptoanálisis: matemáticas, la estadística y la linguística.
- Teólogos interesados en la cronología de las revelaciones.
  - contaban las frecuencias de las palabras contenidas en cada revelación,
  - la teoría era que algunas palabras habían evolucionado recientemente
  - si la revelación contiene un alto número de estas palabras la revelación era “joven”

29

Frecuencia letras

- Se estudiaba el Corán
  - probar si determinados textos eran consistentes con los patrones lingüísticos del Profeta
- Estudiantes no solo estudiaban palabras
  - se iban al nivel de letras.
- Descubren que algunas letras eran más comunes que otras.
  - letras a y l son las más comunes en el árabe
  - letra j aparece con muy poca frecuencia
- Esto representa primera herramienta criptoanálisis.

30

El filósofo de los árabes

- No se puede decir quien fue el primero en darse cuenta que la variación de la frecuencia de letras se podía usar para romper criptosistemas.
- La descripción más antigua de la técnica data del siglo IX del científico *Abu YusufYaqub ibn Is-haq ibn as-Sbbah ibn ‘omran ibn Ismail al-Kindi* conocido como el filósofo de los arabes.
- Escribe: Un manuscrito sobre la descripción de mensajes criptográficos.

31

Análisis de Frecuencia

- Una forma de solucionar un mensaje encriptado, si se conoce su lenguaje, es:
  - encontrar un texto claro del mismo lenguaje lo suficientemente grande para llenar una hoja
  - contar las ocurrencias de cada letra
  - la letra de mas ocurrencias es la primera, la siguiente con más ocurrencia es la segunda, la siguiente la tercera y así hasta contar todas las letras del alfabeto
- Después se ve el criptograma:
  - se clasifican los símbolos de la misma forma
  - se substituye el símbolo que más repetido por la primera letra, el siguiente más repetido es reemplazado por la segunda letra y así hasta tener todos los símbolos.

32

Factores importantes a tomar en cuenta para ataques a criptosistemas monoalfabeticos

- Análisis de frecuencia
  - distribución frecuencia: distribución general de caracteres en una escritura regular del lenguaje
- Repetición de patrones
  - patrones que se repiten con frecuencia
  - español: de, que, cion

33

Frecuencias de letras

- English:
  - E,T,A,O,N,R,I,S,H,D,L,F,C,M,U,G,Y,P,W,B,V,K,X,J,Q,Z
- Francais:
  - E,A,S,I,T,N,R,U,L,O,D,C,M,P,V,Q,G,F,B,H,J,X,Y,Z,K,W
- Español:
  - E,A,O,S,R,N,I,D,L,C,T,U,M,P,B,G,Y,V,Q,H,F,Z,J,X,W,K
- Aleman:
  - E,N,I,S,T,R,A,D,H,U,G,M,C,L,B,O,F,K,W,V,Z,P,J,Q,Y,X

34

Digraphs, Bigrams y Trigrams

- Digraphs
  - grupo indivisible de dos letras y con un solo sonido
  - ejemplo: ch en español o th en inglés
- Bigram
  - es un grupo de dos letras
- Trigram
  - es un grupo de tres letras

35

Ejemplo en inglés

- Double Letters (1000 words):
  - LL 19, SS 15, EE 14, OO 12, TT 9
- Bigram (1000 words):
  - TH 168, HE 132, AN 92, RE 91
- Trigram (20000 words):
  - the 1054, ing 317, ent 234, ion 232, tio 177
- Common Words (10000 words):
  - the 420, and 142, to 132, in 111, a 108,

36

Ejemplo en español

D. Spanish Letter Frequency Data

1–a.—*Absolute frequencies of single letters of Spanish plain text, arranged alphabetically, based on 60,115 letters of text.*

A	6,681	G	823	L	2,174	Q	346	V	602
B	799	H	367	M	1,740	R	4,628	W *	36
C	3,137	I	4,920	N [1]	4,823	S	4,140	X	127
D	2,687	J	190	O	5,859	T	3,180	Y	413
E	7,801	K	22	P	1,755	U	2,172	Z	182
F	481								60,115

1–b.—*Monographic kappa plain, Spanish language*=.0747 (I. C. =1.94)

1–c.—*Frequency distribution of single letters based on 60,115 letters in Spanish plain text, reduced to 1,000 letters, arranged according to their relative frequencies.*

E	130	S	69	U	36	V	10	J	3
A	111	T	53	P	30	F	8	Z	3
O	97	C	52	M	29	Y	7	X	2
I	82	D	45	G	14	H	6	W	1
N	80	L	36	B	13	Q	6	K	
R	77								1,000

1–d.—*Percentage of vowels, high-frequency consonants, medium-frequency consonants, and low-frequency consonants in 60,115 letters of Spanish plain text. Percentage of 7 most frequent letters in Spanish plain text.*

Vowels A, E, I, O, U, and Y=46.3%
High-Frequency Consonants N, R, and S=22.6%
Medium-Frequency Consonants C, D, L, M, P, and T=24.5%
Low-Frequency Consonants B, F, G, H, J, K, Q, V, W, X, and Z=6.6%

7 most frequent letters (in descending order of frequency) E, A, O, I, N, R, and S=64.6%

1–e.—*Absolute frequencies of single letters as initial letters of 10,129 words in Spanish plain text, arranged according to their frequencies. (One-letter words have been omitted.)*

P	1,128	L	435	Q	286	V	183	Y	27
C	1,081	R	425	I	281	F	177	W	19
D	1,012	M	403	H	230	O	169	Z	2
E	989	N	346	U	219	B	124	K	1
S	789	T	298	G	206	J	47	X	
A	761								10,129

[1] Includes Ñ throughout all tables.
* From foreign words appearing in Spanish plain text.

2–c.—*The 87 digraphs comprising 75% of Spanish plain text, based on the table of 5,000 digraphs (Item 2–a), arranged according to their relative frequencies.*

EN	126	TE	67	IN	50	NA	41	MA	32	IS	27	EA	20
ES	119	AN	64	EC	47	IE	40	SA	32	EM	26	OA	19
ON	104	[1]	1,287	RI	45	[2]	2,513	PO	31	SP	26	PU	19
ER	94	AD	64	EL	44	CA	39	MI	30	ED	26	SC	18
RE	94	AS	62	LA	44	ND	37	PA	30	OD	26	AT	18
NT	91	TA	60	RO	43	TI	35	AD	30	AP	24	CU	18
DE	84	DO	59	NO	43	LE	35	DI	30	IT	24	EE	17
AR	81	OR	58	IA	43	TR	34	ID	29	EP	23	OB	17
CI	80	SE	57	IC	42	UN	34	QU	29	SU	23	CE	17
RA	74	ST	57	ME	42	PR	34	OP	29	SO	22	ET	17
OS	73	TO	56	AL	41	OM	33	LI	28	OL	22	LO	17
CO	69	AC	54	SI	41	NC	33	NI	28	NS	21		3,753
IO	67	UE	52	NE	41	DA	32	OC	28	EG	21		

2–d.—*Frequent digraphs in Spanish plain text whose reversals are also frequent, accompanied by their frequencies from the table of 5,000 digraphs (Item 2–a).*

EN	126	NE	41	AR	81	RA	74	AS	62	SA	32	LA	44	AL	41
ES	119	SE	57	CI	80	IC	42	OR	58	RO	43	EL	44	LE	35
ON	104	NO	43	AN	64	NA	41	AC	54	CA	30	MA	32	AM	30
ER	94	RE	94	AD	64	DA	32								

2–e.—*Frequent digraphs in Spanish plain text whose reversals are rare or infrequent, accompanied by their frequencies from the table of 5,000 digraphs (Item 2–a).*

NT	91	TN	0	ST	57	TS	0	ND	37	DN	1	NC	33	CN	0
IO	67	OI	4												

2–f.—*Doublets occurring in Spanish plain text, arranged according to their frequencies from the table of 5,000 digraphs (Item 2–a).*

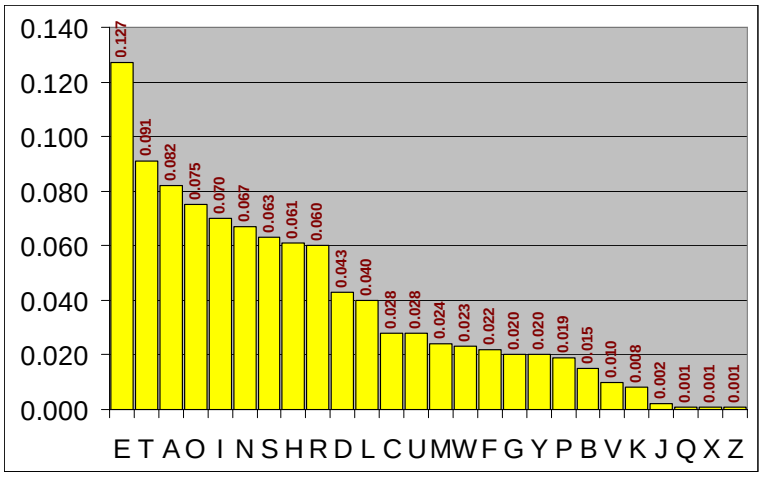
EE	17	AA	12	RR	10	SS	10	LL	9	CC	5	OO	4	NN	3	DD	2

2–g.—*The 21 digraphs appearing 100 or more times as beginnings of words in 10,189 words in Spanish plain text, arranged according to their absolute frequencies.*

CO	684	PR	307	PA	263	SE	189	CA	151	PE	111	MA	101
RE	335	ES	286	PO	247	DI	175	SI	137	UN	109	CU	100
DE	323	QU	286	IN	235	PU	157	MI	117	HA	108	SO	100

37

Frecuencias de letras en el idioma inglés



38

Ejemplo criptoanalisis de frecuencia

Criptograma:

yifqfmzrwqfyvecfmdzpcvmrzwnmdzvejbtxcddumjndife
fmdzcdmqzkceyfcjmyrncwjcszrexchzunmxznzucdrjxyy
smrtmeyifzwdyvzvyfzumrzcrwnzdzjjxzwgchsmrnmdhnc
mfqchzjmxjzwiejyucfwdjnzdir

Repetición:

a 0	h 4	o 0	v 5
b 1	I 5	p 1	w 8
c 15	j 11	q 4	x 6
d 13	k 1	r 10	y 10
e 7	l 0	s 3	z 20
f 11	m 16	t 2	
g 1	n 9	u 5	

39

Analizando las frecuencias

- Acomodando los elementos de más frecuencia:
  - z 20 ocurrencias
  - mc,d,f,j,r,y al menos 10 ocurrencias
- z tiene 20 ocurrencias, podemos conjeturar que
  - $E_K(E) = z$ o ($D_K(z) = E$)
- D_K mapea {m,c,d,f,j,r,y} a {T,A,O, I,N,S,H,R}
  - como las frecuencias no varian mucho, la probabilidad de que lo anterior sea verdad es pequeña
- Sugerencia: observar los digramas
  - dz, zw 4 ocurrencias
  - nz, zu 3 ocurrencias
  - rz, hz, xz, fz, zr, zv, zc, zd, zj 2 ocurrencias

40

¿Entonces?

- zw ocurre 4 veces y wz no ocurre
- se compara la misma propiedad con E- y –E
  - ED se presenta con frecuencia
  - DE no se presenta con frecuencia
- Entonces se deduce que
  - $D_K(w) = D$
- rw se presenta frecuentemente, tanto como ND y ya que $D_K(w) = D$, entonces se puede deducir que:
  - $D_K(r) = N$
- Lo cual se describe en la tabla siguiente.

41

Primer resultado parcial

Yifqfmzrwqfyvecfmdz
pcvmrzwnmdzvejbtxc
ddumjndifefmdzcdmq
zkceyfcjmyrncwjcszre
xchzunmxznzucdrjxyy
smrtmeyifzwdyvzvyfzu
mrzcrwnzdzjjxzwgchsm
rnmdhncmfqchzjmxjzwi
ejyucfwdjnzdir

* * * * * * e n d * * * * * * * * * e
y i f q f m z r w q f y v e c f m d z
* * * * n e d * * * e * * * * * *
p c v m r z w n m d z v e j b t x c
* * * * * * * * * * * * e * * * * *
d d u m j n d i f e f m d z c d m q
e * * * * * * * * n * * * * * e n * -
z k c e y f c j m y r n c w j c s z r e x
- - E - - - - E - E - - - N - - - - -
c h z u n m x z n z u c d r j x y y s
- N - - - - - - E D - - - E - - - E -
m r t m e y i f z w d y v z v y f z u
- N E - N D - E - E - - - E D - - - -
m r z c r w n z d z j j x z w g c h s
- N - - - - - - - - - - - - E - - - - E
m r n m d h n c m f q c h z j m x j z
D - - - - - - - - D - - - E - - N - - -
w i e j y u c f w d j n z d i r

42

Analizando más

- Digrama nz es un digrama común en el criptograma
- Digrama zn no aparece frecuentemente
- Lo mismo para HE y EH respectivamente
- Se puede deducir que
  - $D_K(n) = H$
- Entonces el segmento:

N E – N D H E
r z c r w n z

- La frecuencia del trigrama AND nos conduce a deducir que:
  - $D_K(c) = A$

43

Segundo resultado parcial

```
- - - - - - E N D - - - - - - - - E
y i f q f  m z r w  q f y v  e c f m  d z
- - - - N  E D - -  - E - -  - - - -  - -
p c v m r  z w n m  d z v e  j b t x  c d
- - - - - - - - - - - - E - - - -  E -
d u m j n  d i f e  f m d z  c d m q  z k
- - - - - - - - N - -  D - - -  E N - -
c e y f c  j m y r  n c w j  c s z r  e x
- - E - -  - - E -  E - - -  N - - -  - -
c h z u n  m x z n  z u c d  r j x y  y s
- N - - -  - - - E  D - - -  E - - -  E -
m r t m e  y i f z  w d y v  z v y f  z u
- N E - N  D - E -  E - - -  E D - -  - -
m r z c r  w n z d  z j j x  z w g c  h s
- N - - -  - - - -  - - - -  E - - -  - E
m r n m d  h n c m  f q c h  z j m x  j z
D - - - -  - - - D  - - - E  - - N -  - -
w i e j y  u c f w  d j n z  d i r -  - -
```

```
- - - - - - E N D - - - -  - A - -  - E
y i f q f  m z r w  q f y v  e c f m  d z
- A - -  N E D - -  - E - -  - - - -  A -
p c v m r  z w n m  d z v e  j b t x  c d
- - - - - - - - - - - - E  A - - -  E -
d u m j n  d i f e  f m d z  c d m q  z k
A - - -  A - - -  N H A D -  A - E N  - -
c e y f c  j m y r  n c w j  c s z r  e x
A - E - H  - - E H  E - A -  N - - -  - -
c h z u n  m x z n  z u c d  r j x y  y s
- N - - -  - - - E  D - - -  E - - -  E -
m r t m e  y i f z  w d y v  z v y f  z u
- N E A N  D H E -  E - - -  E D - A  - -
m r z c r  w n z d  z j j x  z w g c  h s
- N H - -  - - A -  - - A -  E - - -  - E
m r n m d  h n c m  f q c h  z j m x  j z
D - - - -  - A - D  - - H E  - - N -  - -
w i e j y  u c f w  d j n z  d i r -  - -
```

44

y mas

- Ahora se considera m
- Criptograma: rnm se decripta a NH-
- Lo anterior sugiere que H empieza una nueva palabra,
- Entonces m debe ser una vocal (A,E,I,O,U)
- Ya se dedujo A y E, por lo que debe ser:
  - $D_K(m) = I$
  - $D_K(m) = O$
- Ya que cm se encuentra en el criptograma, puede ser AI o AO, se puede deducir que
  - $D_K(m) = I$

45

Tercer resultado parcial

- - - - - I E N D - - - - - A - I - E
y i f q f m z r w q f y v e c f m d z
- A - I N E D - I - E - - - - - A -
p c v m r z w n m d z v e j b t x c d
- - I - - - - - - - I - E A - I - E -
d u m j n d i f e f m d z c d m q z k
A - - - A - I - N H A D - A - E N - -
c e y f c j m y r n c w j c s z r e x
A - E - H I - E H E - A - N - - - - -
c h z u n m x z n z u c d r j x y y s
I N - I - - - - E D - - - E - - - E -
m r t m e y i f z w d y v z v y f z u
I N E A N D H E - E - - - E D - A - -
m r z c r w n z d z j j x z w g c h s
I N H I - - - A I - - A - E - I - - E
m r n m d h n c m f q c h z j m x j z
D - - - - - A - D - - H E - - N - - -
w i e j y u c f w d j n z d i r - - -

46

Buscando ...

- Ahora se intenta encontrar $E_K(O)$
- O es un carácter común en ingles
  - se puede deducir que $E_K(O)$ = d,f,j o y
  - los caracteres más frecuentes en el cripograma que no se han contado
- Ocurrencia de cfm, cdm y cjm, en criptograma, sugieren que $E_K(O)$ no debe ser una vocal, por lo que se deduce que
  - $E_K(O) = Y$ o $D_K(y) = O$

47

Buscando más

- Los más frecuentes caracteres en el criptograma, sin deducir, son d,f,j, por lo que se conjetura:
  - $D_K\{d,f,j\} = \{ r,s,t\}$
- Las dos ocurrencias de nmd (decriptadas a HI-) sugieren que
  - $D_K(d)$ = A (ya que HIS es un trigrama frecuente)
- El segmento hncmf puede decriptarse como CHAIR (HIS CHAIR), por lo que:
  - $D_K(f)$ = R
  - $D_K(h)$ = C, entonces
  - $D_K(j)$ = T

48

Cuarto resultado parcial

O -R-R IEND -RO- -A-I SE
y ifqf mzrw qfyv ecfm dz
- A-IN ED-I SE-- T--- AS
p cvmr zwnm dzve jbtx cd
S -IT- S-R- RISE ASI- E-
d umjn dife fmdz cdmq zk
A -ORA TION HADT A-EN --
c eyfc jmyr ncwj cszr ex
A CE-H I-EH E-AS NT-O O-
c hzun mxzn zucd rjxy ys
I N-I- O-RE DSO- E-OR E-
m rtme yifz wdyv zvyf zu
I NEAN DHES ETT- ED-A C-
m rzcr wnzd zjjx zwgc hs
I NHIS C-AI R-AC ETI- TE
m rnmd hncm fqch zjmx jz
D --TO -ARD STHE S-N- --
w iejy ucfw djnz dir- --

49

El paso final

- Es fácil encontrar el resto del texto claro y el resto de la llave

  DK(p) = X, DK(e) = P, DK(b) = Y, DK(t) = G,
  DK(x) = L, DK(u) = W, DK(k) = V, DK(s) = K.

- El mensaje es:

OURFRIENDFROMPARISEXAMINEDHISEMPTYGLASSWITH
SURPRISEASIFEVAPORATIONHADTAKENPLACEWHILEHE
WASNTLOOKINGIPOUREDSOMEMOREWINEANDHESETTLED
BACKINHISCHAIRFACETILTEDUPTOWARDSTHESUN

50

Detalles acerca de análisis frecuencia

- La tabla de frecuencias solo es un porcentaje y no corresponde a las frecuencias de todos los textos.
- Por ejemplo:
  - From Zanzibar to Zambia and Zaire ozone zones make zebras run zany zigzags
- En general textos cortos son candidatos para desviarse de las frecuencias promedios
  - si hay menos de 100 letras, el decriptado será difícil de lleva a cabo

51

Variantes criptosistemas monoalfabeticos

combatiendo el análisis de frecuencia

52

Antecendentes

- Siglos XIV - XVI
- El uso de criptografia se ha difundido.
- Usado por los alquimistas y científicos para mantener en secreto sus descubrimientos.
- El renacimiento hace que la criptografia avance.
- Empiezan a surgir estados independientes que necesitaban comunicarse de forma secreta
  - Giovanni Sro, secretario criptosistemas veneciano 1506
  - Philipert Babou y Francois Viete (Francia)
- Batalla entre criptografos y criptoanalistas

53

Técnicas combatir análisis frecuencia

- Incorporación de nulos
- Escritura incorrecta del mensaje
- Incorporar codewords: nomenclators
- Los criptosistemas digráficos o poligráficos
  - Porta
  - Playfair
  - Hill
- Los criptosistemas homofonicos
  - Numéricos
  - Grandpre
  - Beale

54

Incorporando nulos

- Uno de las mejoras más simples fue la incorporación de nulos:
  - símbolos o letras que no substituían ninguna letra del alfabeto de texto plano
  - blancos que no representan nada
- Ejemplo:
  - sustituir cada letra del alfabeto claro por un número entre 1 y 99,
  - hay 73 números que no representan nada y pueden repartirse en el criptograma con frecuencias variadas

55

Ejemplo incorporación nulos

A 10	I 18	Q 26	Y 34
B 11	J 19	R 27	Z 35
C 12	K 20	S 28	
D 13	L 21	T 29	
E 14	M 22	U 30	
F 15	N 23	V 31	
G 16	O 24	W 32	
H 17	P 25	X 33	

Nulos: { 36,37,38,39, … 99}

Mensaje: HOLA MUNDO
Criptograma: 17 36 24 40 21 53 10 62 22 81 30 72 23 92 13 46 24 40

56

Escritura incorrecta del mensaje

- Nulos pueden confundir cualquier ataque basado en un análisis de frecuencia.
- Una variante es escribir incorrectamente algunas palabras antes de encriptar el mensaje, haciendo difícil aplicar el análisis de frecuencia
  - Thys haz thi ifetkkt off diztaughting thi ballans off frilwenseas
- El destinatario, que conoce la llave, puede descifrar el mensaje y después interpretar el mensaje.

57

Los nomenclators

- Introducir codewords, códigos de palabras
- La sustitución se realiza a un nivel más alto: palabras.
- Ejemplo:

asesinar	D	general	Σ	inmediatamente	08
correo	P	rey	Ω	hoy	73
capturar	J	ministro	Ψ	noche	28
proteger	Z	principe	Θ	mañana	43

Mensaje : asesinar al rey hoy

Criptograma : D - Ω - 73

58

Los nomenclators

- Sistema de encripción basado en:
  - sustitución de letras, usado encriptar la mayor parte mensaje,
  - una lista limitada de palabras codificadas.
- Un nomenclator puede consistir de
  - una página que contiene el alfabeto de encripción y
  - una segunda página que contiene una lista de palabras codificadas.
- No es más seguro que un criptograma de letras,
  - parte principal mensaje se decripta usando análisis frecuencia
  - las palabras pueden adivinarse a partir del contexto.

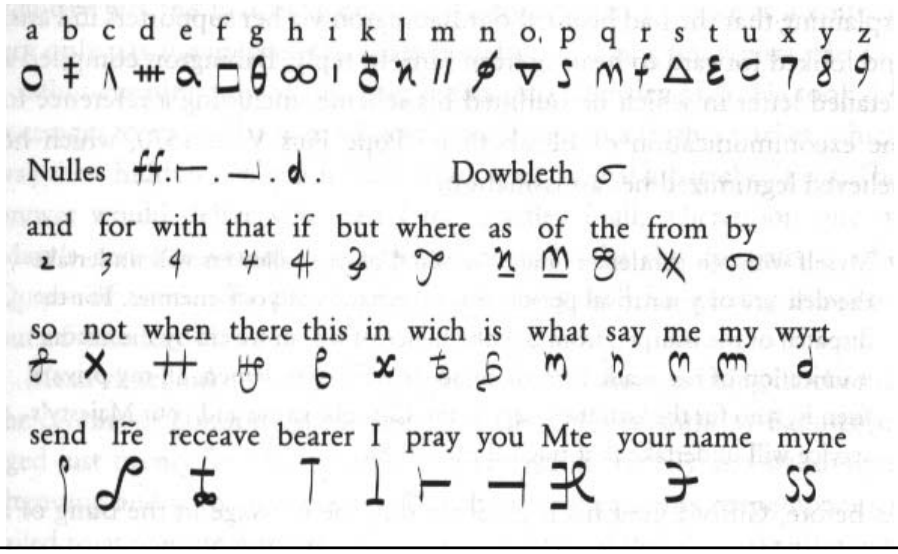
59

Ventajas y desventajas

- Palabras son menos vulnerables al análisis de frecuencias que letras.
  - en lugar de “atacar” 26 letras, se necesita identificar el valor de cientos o miles de palabras codificadas
- El tamaño de la llave cambia
  - necesario definir un código de palabras, para los cientos o miles de palabras,
  - el libro de código sería de cientos de páginas y podría verse como un diccionario
- Las consecuencias de capturar el libro de códigos son devastadoras
  - todas las comunicaciones serían transparentes

60

Ejemplo nomenclator: Mary Queen of Scots

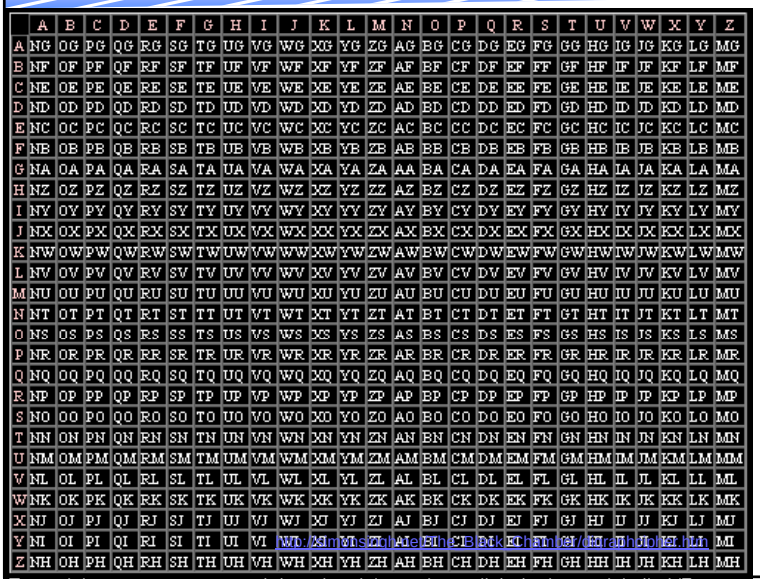


Los criptosistemas digráficos o poligráficos

- Símbolos del alfabeto en claro son encriptados en pares (digráfos)
- El análisis de frecuencia es más duro
- Ejemplo de este tipo de criptosistemas:
  - Digráfico simple
  - Porta
  - Playfair
  - Hill

62

Ejemplo criptosistema digráfico simple



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	NG	OG	PG	QG	RG	SG	TG	UG	VG	WG	XG	YG	ZG	AG	BG	CG	DG	EG	FG	GG	HG	IG	JG	KG	LG	MG
B	NF	OF	PF	QF	RF	SF	TF	UF	VF	WF	XF	YF	ZF	AF	BF	CF	DF	EF	FF	GF	HF	IF	JF	KF	LF	MF
C	NE	OE	PE	QE	RE	SE	TE	UE	VE	WE	XE	YE	ZE	AE	BE	CE	DE	EE	FE	GE	HE	IE	JE	KE	LE	ME
D	ND	OD	PD	QD	RD	SD	TD	UD	VD	WD	XD	YD	ZD	AD	BD	CD	DD	ED	FD	GD	HD	ID	JD	KD	LD	MD
E	NC	OC	PC	QC	RC	SC	TC	UC	VC	WC	XC	YC	ZC	AC	BC	CC	DC	EC	FC	GC	HC	IC	JC	KC	LC	MC
F	NB	OB	PB	QB	RB	SB	TB	UB	VB	WB	XB	YB	ZB	AB	BB	CB	DB	EB	FB	GB	HB	IB	JB	KB	LB	MB
G	NA	OA	PA	QA	RA	SA	TA	UA	VA	WA	XA	YA	ZA	AA	BA	CA	DA	EA	FA	GA	HA	IA	JA	KA	LA	MA
H	NZ	OZ	PZ	QZ	RZ	SZ	TZ	UZ	VZ	WZ	XZ	YZ	ZZ	AZ	BZ	CZ	DZ	EZ	FZ	GZ	HZ	IZ	JZ	KZ	LZ	MZ
I	NY	OY	PY	QY	RY	SY	TY	UY	VY	WY	XY	YY	ZY	AY	BY	CY	DY	EY	FY	GY	HY	IY	JY	KY	LY	MY
J	NX	OX	PX	QX	RX	SX	TX	UX	VX	WX	XX	YX	ZX	AX	BX	CX	DX	EX	FX	GX	HX	IX	JX	KX	LX	MX
K	NW	OW	PW	QW	RW	SW	TW	UW	VW	WW	XW	YW	ZW	AW	BW	CW	DW	EW	FW	GW	HW	IW	JW	KW	LW	MW
L	NV	OV	PV	QV	RV	SV	TV	UV	VV	WV	XV	YV	ZV	AV	BV	CV	DV	EV	FV	GV	HV	IV	JV	KV	LV	MV
M	NU	OU	PU	QU	RU	SU	TU	UU	VU	WU	XU	YU	ZU	AU	BU	CU	DU	EU	FU	GU	HU	IU	JU	KU	LU	MU
N	NT	OT	PT	QT	RT	ST	TT	UT	VT	WT	XT	YT	ZT	AT	BT	CT	DT	ET	FT	GT	HT	IT	JT	KT	LT	MT
O	NS	OS	PS	QS	RS	SS	TS	US	VS	WS	XS	YS	ZS	AS	BS	CS	DS	ES	FS	GS	HS	IS	JS	KS	LS	MS
P	NR	OR	PR	QR	RR	SR	TR	UR	VR	WR	XR	YR	ZR	AR	BR	CR	DR	ER	FR	GR	HR	IR	JR	KR	LR	MR
Q	NQ	OQ	PQ	QQ	RQ	SQ	TQ	UQ	VQ	WQ	XQ	YQ	ZQ	AQ	BQ	CQ	DQ	EQ	FQ	GQ	HQ	IQ	JQ	KQ	LQ	MQ
R	NP	OP	PP	QP	RP	SP	TP	UP	VP	WP	XP	YP	ZP	AP	BP	CP	DP	EP	FP	GP	HP	IP	JP	KP	LP	MP
S	NO	OO	PO	QO	RO	SO	TO	UO	VO	WO	XO	YO	ZO	AO	BO	CO	DO	EO	FO	GO	HO	IO	JO	KO	LO	MO
T	NN	ON	PN	QN	RN	SN	TN	UN	VN	WN	XN	YN	ZN	AN	BN	CN	DN	EN	FN	GN	HN	IN	JN	KN	LN	MN
U	NM	OM	PM	QM	RM	SM	TM	UM	VM	WM	XM	YM	ZM	AM	BM	CM	DM	EM	FM	GM	HM	IM	JM	KM	LM	MM
V	NL	OL	PL	QL	RL	SL	TL	UL	VL	WL	XL	YL	ZL	AL	BL	CL	DL	EL	FL	GL	HL	IL	JL	KL	LL	ML
W	NK	OK	PK	QK	RK	SK	TK	UK	VK	WK	XK	YK	ZK	AK	BK	CK	DK	EK	FK	GK	HK	IK	JK	KK	LK	MK
X	NJ	OJ	PJ	QJ	RJ	SJ	TJ	UJ	VJ	WJ	XJ	YJ	ZJ	AJ	BJ	CJ	DJ	EJ	FJ	GJ	HJ	IJ	JJ	KJ	LJ	MJ
Y	NI	OI	PI	QI	RI	SI	TI	UI	VI	WI	XI	YI	ZI	AI	BI	CI	DI	EI	FI	GI	HI	II	JI	KI	LI	MI
Z	NH	OH	PH	QH	RH	SH	TH	UH	VH	WH	XH	YH	ZH	AH	BH	CH	DH	EH	FH	GH	HH	IH	JH	KH	LH	MH

Texto claro:
HOLA

Criptograma:
BZNV

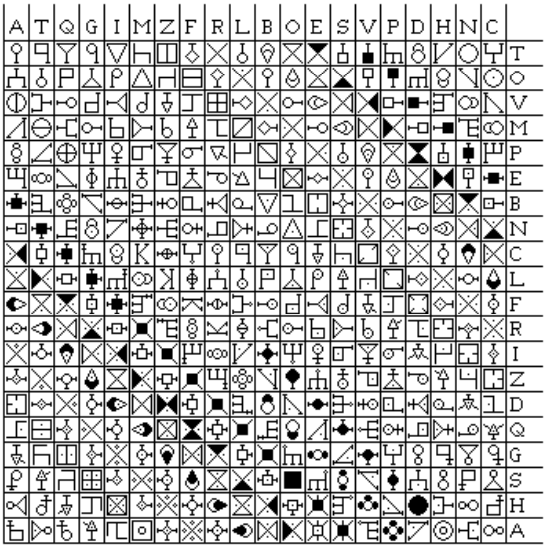
63

Criptosistema de Porta

- Un símbolo es sustituido por un par de letras en el texto plano.
- Se necesita una matriz de 26 x 26
  - las entradas de la matriz deben llenarse con símbolos diferentes (pueden ser solo números)
- El alfabeto es escrito en el borde externo de la matriz
  - uno de izquierda a derecha y otro de arriba a abajo
- Para encriptar se toma el símbolo que corresponde al par de letras del texto claro y se sustituye.

64

Tabla digrafica de Porta



65

Versión “moderna” de Porta

- La tabla se puede sustituir por números (1 … 676), acomodados al azar en la tabla.
- Suponiendo que se quiera encriptar la palabra “LETTER”.
  - se toma el primer par de letras LE, la letra L se toma como la coordenada x y la letra E se toma como la coordenada y
  - si la entrada es 291, LE se substituye por 291
  - las letras TT y ER son tratadas de la misma forma

66

Ejemplo versión moderna

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	A
B	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	B
C	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	C
D	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	D
E	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	E
F	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	F
G	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	G
H	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	H
I	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	I
J	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	J
K	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	K
L	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	L
M	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	M
N	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	N
O	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	O
P	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	P
Q	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	Q
R	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	R
S	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	S
T	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	T
U	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	U
V	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	V
W	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	W
X	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	X
Y	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	Y
Z	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	Z
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	

Texto claro: LETTER Criptograma: 291, 514, 122

67

Ventajas Porta

- La distribución de frecuencia es mucho más plana.
  - la variación se vuelve más grande ya que el alfabeto de encripción es de 676 letras
- La longitud del criptograma es dos veces más corta que la del texto claro

68

Desventajas Porta

- La tabla original de Porta cuenta con una estructura.
  - la intención de esta estructura es ayudar en la decripción
  - la estructura también ofrece una ayuda considerable a los criptoanalistas
- Para dificultar Porta
  - dejar un conjunto de símbolos ordenados y fijos (por ejemplo números) en la tabla
  - reordenar el alfabeto que indexa la tabla

69

Playfair

- Es un criptosistema más compacto que el de Porta.
- Sistema basado en una matriz de 5 x 5.
- En el primer renglón de la matriz escribimos una llave,
  - la llave solo contiene letras diferentes
- El resto de la celdas se llenarán con el resto de las letras en algún orden que sea fácil de recordar.
  - como en el criptograma de Porta, el mensaje es encriptado tomando pares de letras

70

Las reglas básicas de Playfair

- Si ambas letras se encuentran en el mismo renglón
  - usar las letras que se encuentran inmediatamente a la derecha de cada letra
  - imaginar que el final de la derecha de cada renglón esta unido con el final izquierdo
- Si ambas letras están en la misma columna
  - usar las letras que se encuentran inmediatamente abajo de cada letra
  - imaginarse que la parte baja de la columna esta conectada con su parte superior

71

Las reglas básicas de Playfair

- Si dos letras se encuentran en diferentes renglones y diferentes columnas
  - cada letra es reemplazada por la letra en el mismo renglón que se encuentra en la columna ocupada por la otra letra
- En el caso de dos letras iguales en un diagrama:
  - insertar una x entre ellas (choose = cho**x**ose)

72

Ejemplo de Playfair

	1	2	3	4	5
1	P	I/J	A	N	O
2	B	C	D	E	F
3	G	H	K	L	M
4	Q	R	S	T	U
5	V	W	X	Y	Z

Texto claro: cipher
Llave: piano
Encripción: ci se reemplaza por hc
ph se reemplaza por ig
er se reemplaza por ct
Criptograma: hcigct

http://www.simonsingh.net/The_Black_Chamber/playfaircipher.htm

73

El criptosistema Hill

- Criptosistema poligráfico de substitución basado en algebra lineal.
- Inventado por Lester S Hill en 1929
- Fue el primer criptosistema poligráfico en el cual era posible operar en más de tres símbolos a la vez
- Cada letra es tratada como un dígito en base 26
  - A = 0
  - B = 1
  - C = 2
  - :
  - :

74

Encriptando con Hill

- Consideremos el mensaje ACT y la llave

$$\begin{pmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{pmatrix}$$

- Ya que A=0, C=2 y T=19, entonces el mensaje es el vector
- El vector de encripción es

$$\begin{pmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 19 \end{pmatrix} \equiv \begin{pmatrix} 67 \\ 222 \\ 319 \end{pmatrix} \equiv \begin{pmatrix} 15 \\ 14 \\ 7 \end{pmatrix} \pmod{26}$$

- que corresponde al criptograma POH

75

Decriptando con Hill

- El criptograma se convierte en vector
- Se multiplica por la inversa de la matriz llave
- En el ejemplo, la inversa es:

$$\begin{pmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{pmatrix}$$

- Tomando el criptograma del ejemplo, POH:

$$\begin{pmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{pmatrix} \begin{pmatrix} 15 \\ 14 \\ 7 \end{pmatrix} \equiv \begin{pmatrix} 260 \\ 574 \\ 539 \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 2 \\ 19 \end{pmatrix} \pmod{26}$$

- lo cual produce el texto original ACT

76

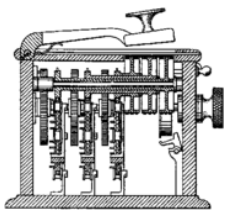
A tomar en cuenta

- Posible complicación
  - No todas las matrices tienen una inversa
- Suceptible a un ataque de tipo known-plaintext
  - es completamente lineal
  - atacante intercepta n^2 pares de texto claro/criptograma puede establecer un sistema lineal que puede ser resuelto facilmente
- Dado un factor de n que define el tamaño de la matriz, solo el 30% de estas matrices serán útiles
  - determinantes no están distribuidos de forma uniforme
  - keyspace es cerca de $4.64\ n^2 - 1.7$
  - para una llave de 5 x 5 esto es cerca de 114 bits

77

Implementación mecánica

- Hill y un socio tienen una patente
  - US patent 1,845,947
- Dispositivo multiplicación matriz 6 x 6 modulo 26 usando un sistema de cadenas y engranes
- La llave es fija para una maquina
  - se recomienda una triple encripción
    1. paso secreto no lineal
    2. se aplica hill
    3. paso secreto no lineal



78

Criptosistema de sustitución homofónico

- Cada letra es reemplazado con una variedad de substitutos.
- El número de posibles substitutos es proporcional a la frecuencia de la letra.
- Por ejemplo:
  - letra a representa el 8% de las letras, y se puede asignar 8 símbolos para representarla
  - cada letra en a en el texto claro será reemplazada por ocho símbolos elegidos al azar
- De esta forma cada símbolo constituirá el 1% del texto en total

79

Números usados como símbolos

- Una opción es usar números de dos dígitos
- Se pueden ver que todos los números de dos dígitos que representan a la ***a***, tienen el mismo *sonido* en el criptograma
- Es el origen del nombre del criptosistema
  - homos: mismo
  - fonico: phonos: sonido en griego
- El objetivo es balancear las frecuencias de símbolos en el criptograma.

80

Ejemplo criptosistema homofónico

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
09	48	13	01	14	10	06	23	32	15	04	26	22	18	00	38	94	29	11	17	08	34	60	28	21	02
12	81	41	3	16	31	25	39	70			37	27	58	05	95		35	19	20	61		89		52	
33		62	45	24			50	73			51		59	07			40	36	30	63					
47			79	44			56	83			84		66	54			42	76	43						
53				46			65	88					71	72			77	86	49						
67				55			68	93					91	90			80	96	69						
78				57										99					75						
92				64															85						
				74															97						
				82																					
				87																					
				98																					

http://simonsingh.net/The_Black_Chamber/homophoniccipher.htm

81

Otros sistema homofonicos

- Criptosistema de Grandpre
- Criptosistema de Beale
- El telegrama de Zimmerman

82

Criptosistema de Grandpre

- Extiende alfabeto texto claro y a través de alguna regla debe asegurarse que una letra del alfabeto tenga más de una representación (homophone)
- Se define una matriz de 8x8 y se llena de llaves (keywords)
  - puede ser una secuencia de palabras
  - se busca la primera ocurrencia, de la primera letra del texto claro, en base al orden lexicográfico de las coordenadas
  - para la segunda letra se hace la búsqueda a partir de donde nos quedamos en la búsqueda anterior

83

Ejemplo Grandpre

	1	2	3	4	5	6	7	8
1	a	b	a	s	h	l	n	g
2	y	o	k	o	h	a	m	a
3	c	o	e	x	l	s	t	s
4	d	e	a	t	h	f	u	l
5	j	a	c	k	p	o	t	s
6	q	u	l	v	e	r	e	d
7	w	l	t	c	h	i	n	g
8	z	o	d	i	a	c	a	l

Texto claro: T H l S l S A P L A I N T E X T
Criptograma: 37 45 76 14 84 36 28

84

El criptosistema de Beale (criptosistema de libro)

- Inventado por T.J. Beale en 1820 en USA
- La llave es la declaración de la independencia,
  - las palabras de la llave son numeradas consecutivamente
  - cada letra del texto claro se encripta substituyendo el número de la palabra con que empieza dicha letra
  - por ejemplo la letra W se puede encriptar con los números 1,19,49,66,72, 90 ó 459
- Da origen a lo que se conoce como los criptosistemas de libros (code books)
  - En lugar de considerar a la declaración de la independencia como la llave se puede tomar el texto y/o libro convenido por las dos partes.

85

Ejemplo criptosistema de Beale

When, in the course of human events, it becomes 10necessary for one people to dissolve the political bands which 20have connected them with another, and to assume among the 30powers of the earth, the separate and equal station to 40which the laws of nature and of nature's God entitle 50them, a decent respect to the opinions of mankind requires 60that they should declare the causes which impel them to 70the separation.

We hold these truths to be self-evident, 80that all men are created equal, that they are endowed 90by their Creator with certain inalienable rights, that among these 100are life, liberty and the pursuit of happiness; That to 110secure these rights, governments are instituted among men, deriving their 120just powers from the consent of the governed; That whenever 130any form of government becomes destructive of these ends, it 140is the right of the people to alter or to 150abolish it, and to institute a new government, laying its 160foundation on such principles and organizing its powers in such 170form, as to them shall seem most likely to effect 180their safety and happiness. Prudence, indeed, will dictate that governments 190long established should not be changed for light and transient 200causes; and accordingly all experience hath shewn, that mankind are 210more disposed to suffer, while evils are sufferable, than to 220right themselves by abolishing the forms to which they are 230accustomed.

Texto claro:
PROBANDO

Criptograma:
184, 96, 31, 134,
171, 156, 15, 143

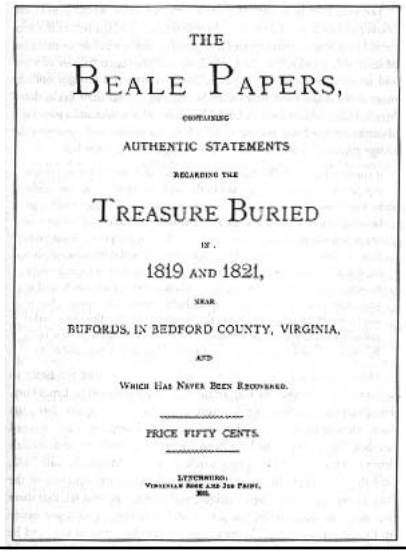
86

The Beale Papers

- Vaquero amasa una fortuna de 20 millones de dólares.
  - un misterioso conjunto de papeles encriptados describe dicha fortuna y como obtenerla
- Todo lo que se conoce acerca de esta historia, incluyendo los papeles encriptados se encuentra en un panfleto publicado en 1885
  - a pesar de que solo tiene 23 páginas, el panfleto ha cautivado a varias generaciones de criptoanalistas.
- Tres criptogramas:
  - el primer da a conocer la ubicación del tesoro
  - la segunda describe el contenido del tesoro
  - la tercera explica quien debe recibir parte del tesoro

87

El panfleto


THE

BEALE PAPERS,

CONTAINING

AUTHENTIC STATEMENTS

REGARDING THE

TREASURE BURIED

IN

1819 AND 1821,

NEAR

BUFORDS, IN BEDFORD COUNTY, VIRGINIA,

AND

WHICH HAS NEVER BEEN RECOVERED.

PRICE FIFTY CENTS.

LYNCHBURG:

88

1er. criptograma de Beale

71,194,38,1701,89,76,11,83,1629,48,94,63,132,16,111,95,84,341
975,14,40,64,27,81,139,213,63,90,1120,8,15,3,126,2018,40,74
758,485,604,230,436,664,582,150,251,284,308,231,124,211,486,225
401,370,11,101,305,139,189,17,33,88,208,193,145,1,94,73,416
918,263,28,500,538,356,117,136,219,27,176,130,10,460,25,485,18
436,65,84,200,283,118,320,138,36,416,280,15,71,224,961,44,16,401
39,88,61,304,12,21,24,283,134,92,63,246,486,682,7,219,184,360,780
18,64,463,474,131,160,79,73,440,95,18,64,581,34,69,128,367,460,17
81,12,103,820,62,110,97,103,862,70,60,1317,471,540,208,121,890
346,36,150,59,568,614,13,120,63,219,812,2160,1780,99,35,18,21,136
872,15,28,170,88,4,30,44,112,18,147,436,195,320,37,122,113,6,140
8,120,305,42,58,461,44,106,301,13,408,680,93,86,116,530,82,568,9
102,38,416,89,71,216,728,965,818,2,38,121,195,14,326,148,234,18
55,131,234,361,824,5,81,623,48,961,19,26,33,10,1101,365,92,88,181
275,346,201,206,86,36,219,324,829,840,64,326,19,48,122,85,216,284
919,861,326,985,233,64,68,232,431,960,50,29,81,216,321,603,14,612

89

1er. criptograma de Beale (cont)

81,360,36,51,62,194,78,60,200,314,676,112,4,28,18,61,136,247,819
921,1060,464,895,10,6,66,119,38,41,49,602,423,962,302,294,875,78
14,23,111,109,62,31,501,823,216,280,34,24,150,1000,162,286,19,21
17,340,19,242,31,86,234,140,607,115,33,191,67,104,86,52,88,16,80
121,67,95,122,216,548,96,11,201,77,364,218,65,667,890,236,154,211
10,98,34,119,56,216,119,71,218,1164,1496,1817,51,39,210,36,3,19
540,232,22,141,617,84,290,80,46,207,411,150,29,38,46,172,85,194
39,261,543,897,624,18,212,416,127,931,19,4,63,96,12,101,418,16,140
230,460,538,19,27,88,612,1431,90,716,275,74,83,11,426,89,72,84
1300,1706,814,221,132,40,102,34,868,975,1101,84,16,79,23,16,81,122
324,403,912,227,936,447,55,86,34,43,212,107,96,314,264,1065,323
428,601,203,124,95,216,814,2906,654,820,2,301,112,176,213,71,87,96
202,35,10,2,41,17,84,221,736,820,214,11,60,760

90

El segundo criptograma

115,73,24,807,37,52,49,17,31,62,647,22,7,15,140,47,29,107,79,84
56,239,10,26,811,5,196,308,85,52,160,136,59,211,36,9,46,316,554
122,106,95,53,58,2,42,7,35,122,53,31,82,77,250,196,56,96,118,71
140,287,28,353,37,1005,65,147,807,24,3,8,12,47,43,59,807,45,316
101,41,78,154,1005,122,138,191,16,77,49,102,57,72,34,73,85,35,371
59,196,81,92,191,106,273,60,394,620,270,220,106,388,287,63,3,6
191,122,43,234,400,106,290,314,47,48,81,96,26,115,92,158,191,110
77,85,197,46,10,113,140,353,48,120,106,2,607,61,420,811,29,125,14
20,37,105,28,248,16,159,7,35,19,301,125,110,486,287,98,117,511,62
51,220,37,113,140,807,138,540,8,44,287,388,117,18,79,344,34,20,59
511,548,107,603,220,7,66,154,41,20,50,6,575,122,154,248,110,61,52,33
30,5,38,8,14,84,57,540,217,115,71,29,84,63,43,131,29,138,47,73,239
540,52,53,79,118,51,44,63,196,12,239,112,3,49,79,353,105,56,371,557
211,505,125,360,133,143,101,15,284,540,252,14,205,140,344,26,811,138
115,48,73,34,205,316,607,63,220,7,52,150,44,52,16,40,37,158,807,37

91

El segundo criptograma (cont)

121,12,95,10,15,35,12,131,62,115,102,807,49,53,135,138,30,31,62,67,41
85,63,10,106,807,138,8,113,20,32,33,37,353,287,140,47,85,50,37,49,47
64,6,7,71,33,4,43,47,63,1,27,600,208,230,15,191,246,85,94,511,2,270
20,39,7,33,44,22,40,7,10,3,811,106,44,486,230,353,211,200,31,10,38
140,297,61,603,320,302,666,287,2,44,33,32,511,548,10,6,250,557,246
53,37,52,83,47,320,38,33,807,7,44,30,31,250,10,15,35,106,160,113,31
102,406,230,540,320,29,66,33,101,807,138,301,316,353,320,220,37,52
28,540,320,33,8,48,107,50,811,7,2,113,73,16,125,11,110,67,102,807,33
59,81,158,38,43,581,138,19,85,400,38,43,77,14,27,8,47,138,63,140,44
35,22,177,106,250,314,217,2,10,7,1005,4,20,25,44,48,7,26,46,110,230
807,191,34,112,147,44,110,121,125,96,41,51,50,140,56,47,152,540
63,807,28,42,250,138,582,98,643,32,107,140,112,26,85,138,540,53,20
125,371,38,36,10,52,118,136,102,420,150,112,71,14,20,7,24,18,12,807
37,67,110,62,33,21,95,220,511,102,811,30,83,84,305,620,15,2,108,220
106,353,105,106,60,275,72,8,50,205,185,112,125,540,65,106,807,188,96,110

92

El segundo criptograma (final)

16,73,32,807,150,409,400,50,154,285,96,106,316,270,205,101,811,400,8
44,37,52,40,241,34,205,38,16,46,47,85,24,44,15,64,73,138,807,85,78,110
33,420,505,53,37,38,22,31,10,110,106,101,140,15,38,3,5,44,7,98,287
135,150,96,33,84,125,807,191,96,511,118,440,370,643,466,106,41,107
603,220,275,30,150,105,49,53,287,250,208,134,7,53,12,47,85,63,138,110
21,112,140,485,486,505,14,73,84,575,1005,150,200,16,42,5,4,25,42
8,16,811,125,160,32,205,603,807,81,96,405,41,600,136,14,20,28,26
353,302,246,8,131,160,140,84,440,42,16,811,40,67,101,102,194,138
205,51,63,241,540,122,8,10,63,140,47,48,140,288

93

El mensaje del segundo criptograma

I have deposited in the county of Bedford about four miles from Bufords in an excavation or vault six feet below the surface of the ground the following articles belonging jointly to the parties whose names are given in number three herewith. The first deposit consisted of ten hundred and fourteen pounds of gold and thirty eight hundred and twelve pounds of silver deposited Nov eighteen nineteen. The second was made Dec eighteen twenty one and consisted of nineteen hundred and seven pounds of gold and twelve hundred and eighty eight of silver, also jewels obtained in St. Louis in exchange to save transportation and valued at thirteen [t]housand dollars. The above is securely packed i[n] [i]ron pots with iron cov[e]rs. Th[e] vault is roughly lined with stone and the vessels rest on solid stone and are covered [w]ith others. Paper number one describes th[e] exact locality of the va[u]lt so that no difficulty will be had in finding it.

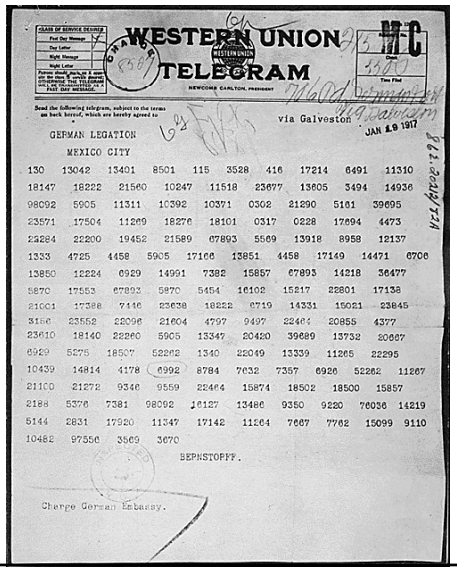
94

El telegrama de Zimmerman

- Alemania usa submarinos para hundir barcos americanos.
  - lo anterior podía provocar el que USA entre a la guerra.
- Arthur Zimmermann, ministro exterior, se le ocurre como retrasar la entrada de USA en la guerra
  - propone alianza con México y persuadir que el Presidente invada USA reclamando Texas, New Mexico y Arizona
  - Alemania soportaría a México militar y financieramente
  - Presidente México debía persuadir Japón para que ataque USA

95

El telegrama de Zimmerman



WESTERN UNION TELEGRAM

via Galveston

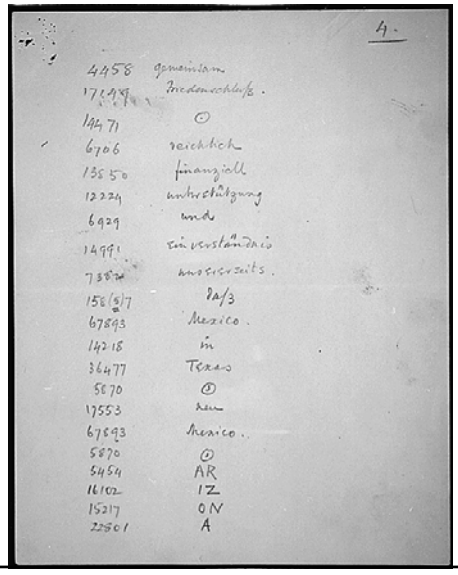
JAN 19 1917

GERMAN LEGATION
MEXICO CITY

130 13042 13401 8501 115 3528 416 17214 6491 11310
18147 18222 21560 10247 11518 23677 13605 3494 14936
98092 5905 11311 10392 10371 0302 21290 5101 39695
23571 17504 11269 18276 18101 0317 0228 17694 4473
22284 22200 19452 21589 67893 5569 13918 8958 12137
1333 4725 4458 5905 17166 13851 4458 17149 14471 6706
13850 12224 6929 14991 7382 15857 67893 14218 36477
5870 17553 67893 5870 5454 16102 15217 22801 17138
21001 17388 7110 23638 18822 6719 14331 15021 23845
3184 23552 22096 21604 4797 9497 22464 20855 4377
23610 18140 22260 5905 13347 20420 39689 13732 20667
6929 5275 18507 52262 1340 22049 13339 11265 22295
10439 14814 4178 6992 8784 7632 7357 6926 52262 11267
21100 21272 9346 9559 22464 15874 18502 18500 15857
2188 5376 7381 98092 16127 13486 9350 9220 76036 14219
5144 2831 17920 11347 17142 11264 7667 7762 15099 9110
10482 97556 3569 3670

BERNSTOFFF.

Charge German Embassy.



4458 gemeinsam
17149 Freundschaft.
14471 ⊙
6706 reichlich
13550 finanziell
12224 unterstützung
6929 und
14991 ein verständnis
7382 unsererseits.
15857 Pa/3
67893 Mexico.
14218 in
36477 Texas
5870 ⊙
17553 Neu
67893 Mexico.
5870 ⊙
5454 AR
16102 IZ
15217 ON
22801 A

96

Decriptando el mensaje

- Telegrama enviado vía el cable del Atlántico
  - el resto de las medios estaban destruidos
- Mensaje del 17 enero 1917
- Criptoanalistas britanicos Montgomety & de Grey descifran parte del mensaje.
  - usan el code book de un diplomático aleman
  - el libro se obtuvó de los bienes confiscados a Wihelm Wassmus el viceconsul alemán en Persia, el cual dejo sus oficinas apresudaramente cuando las fuerzas británicas avanzaron

97

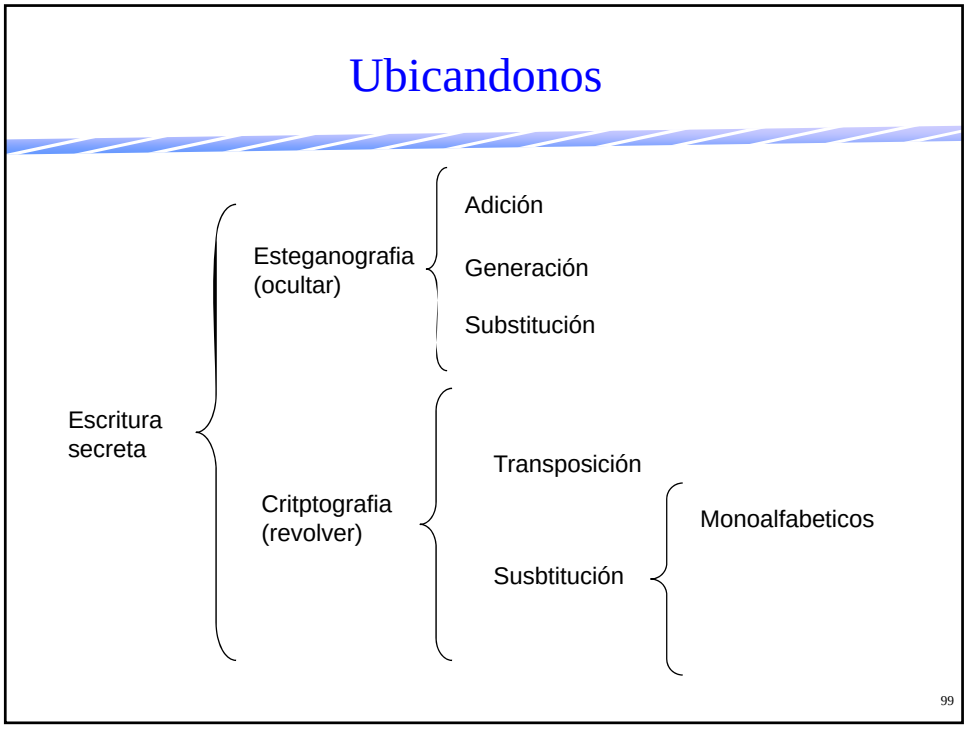
El mensaje de Zimmerman

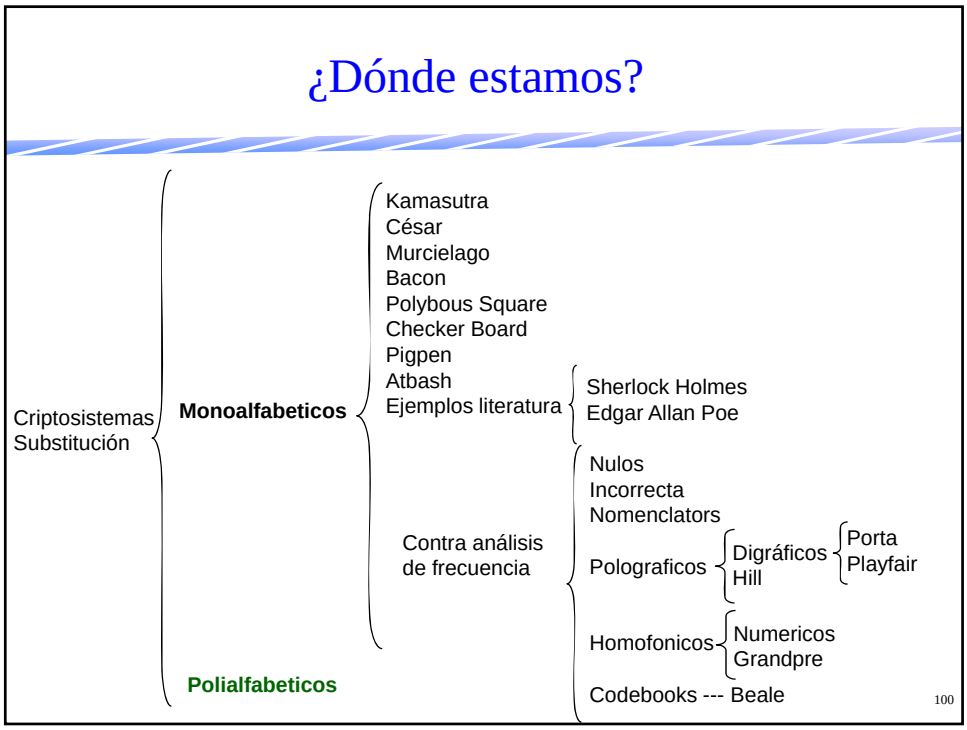
Berlin, January 19, 1917

We intend to begin on the first of February unrestricted submarine warfare. We shall endeavor in spite of this to keep the United States of America neutral. In the event of this not succeeding, we make Mexico a proposal or alliance on the following basis: make war together, make peace together, generous financial support and an understanding on our part that Mexico is to reconquer the lost territory in Texas, New Mexico, and Arizona. The settlement in detail is left to you. You will inform the President of the above most secretly as soon as the outbreak of war with the United States of America is certain and add the suggestion that he should, on his own initiative, invite Japan to immediate adherence and at the same time mediate between Japan and ourselves. Please call the President's attention to the fact that the ruthless employment of our submarines now offers the prospect of compelling England in a few months to make peace.

Zimmermann

98


Ubicandonos
Escritura secreta
Esteganografia (ocultar)
Adición
Generación
Substitución
Critptografia (revolver)
Transposición
Susbtitución
Monoalfabeticos
99


¿Dónde estamos?
Criptosistemas Substitución
Monoalfabeticos
Polialfabeticos
Kamasutra
César
Murcielago
Bacon
Polybous Square
Checker Board
Pigpen
Atbash
Ejemplos literatura
Sherlock Holmes
Edgar Allan Poe
Contra análisis de frecuencia
Nulos
Incorrecta
Nomenclators
Pologràficos
Digráficos
Hill
Porta
Playfair
Homofonicos
Numericos
Grandpre
Codebooks --- Beale
100

Hacia un criptograma más fuerte

los cripsistemas polialfabéticos

101

Hacia un criptograma más fuerte

- Ejecución Mary Queen of Scots fue una dramática ilustración de la debilidad de la sustitución monoalfabetica.
  - criptoanalistas ganaban batalla a criptografos
- Necesario que los criptografos desarrollarán un criptosistema más fuerte.
- Por 1460 florentino matemático Leon Battista Alberti, escribe un ensayo sobre lo que creía que era un nuevo criptosistema.

102

La propuesta de Alberti

- Propone usar dos o más alfabetos de encripción, alternándolos durante la encripción, y confundiendo al posible criptoanalista.

Alfabeto claro	a b c d e f g h i j k l m n o p q r s t u v w x y z
Alfabeto encripción 1:	f z b k i x a y m e p l s d h j o r g n q c u t w
Alfabeto encripción 2:	g o x b f w t h q i l a p z j d e s v y c r k u h n

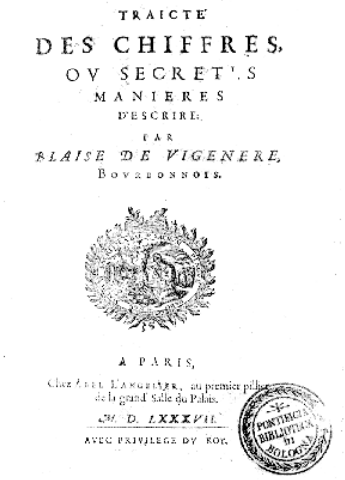
- Sin embargo, Alberti nunca pudo desarrollar su concepto y crear un criptosistema a partir de sus ideas.

103

Mejorando la idea de Albertini

- Johannes Trithmius,
  - alemán (1462-1516)
  - tratado de esteganografia y de poligrafia
- Giambattista della Porta,
  - italiano (1535-1615)
  - forma la *Accademia dei Segreti*, dedicada a estudiar secretos naturaleza, (cerrada por la Inquisición en 1578)
  - escribe *De furtivis litterarum notis vulgo de ziferis libri IV* (1563), un tratado sobre tecnicas criptograficas

Blaise de Vigenere



TRAICTE
DES CHIFFRES,
OV SECRETS
MANIERES
D'ESCRIRE.
PAR
BLAISE DE VIGENERE,
BOVRBONNOIS.

A PARIS,
Chez ABEL L'ANGELIER, au premier pilier de la grand'Salle du Palais.
M. D. LXXXVII.
AVEC PRIVILEGE DV ROY.

publicado en 1586

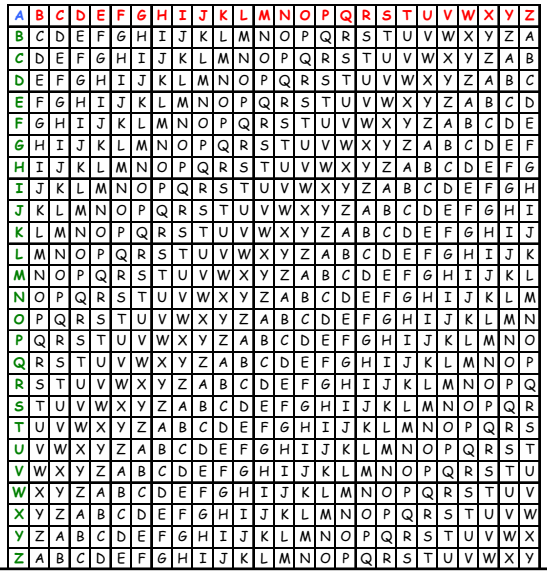
105

La idea de Viginere

- Albertini, Trithemius y Porta contribuyen, pero es Vigenere el que desarrolla el criptosistema hasta su forma final.
- Utiliza 26 alfabetos de encripción diferentes.
- Primer paso: escribir tabla Vigenere
  - cada línea es un alfabeto recorrido una letra con respecto a la línea anterior
  - hasta arriba esta el alfabeto en claro
- Es posible utilizar cualquiera de los 26 alfabetos para encriptar el mensaje, en el orden que se desee.

106

La tabla de Viginere



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

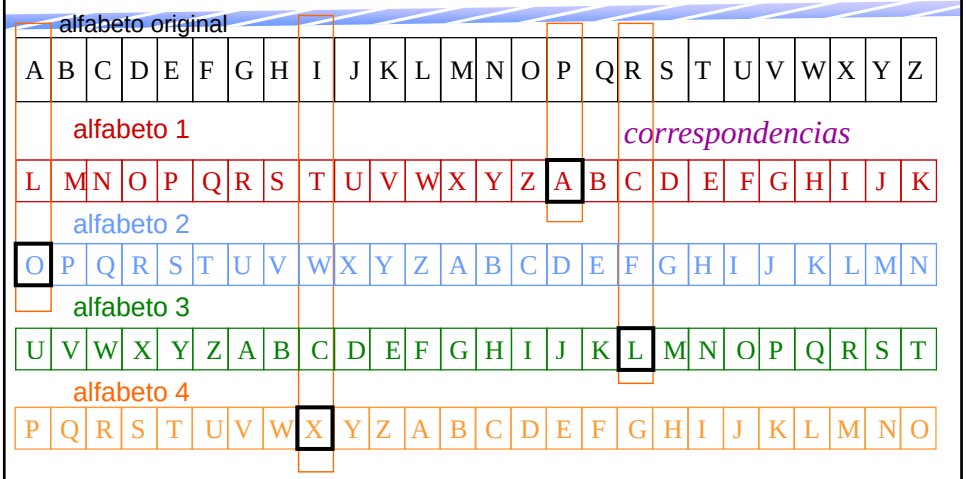
107

Criptosistema de Vigenére

- La llave toma sucesivamente diferentes valores
- Termino matemático: $Y_i = X_i \oplus Z_i \pmod{26}$
- Una misma letra en el texto claro le pueden corresponder diferentes letras en el texto cifrado
- Recuperación mensaje es análoga al procedimiento de Cesar

108

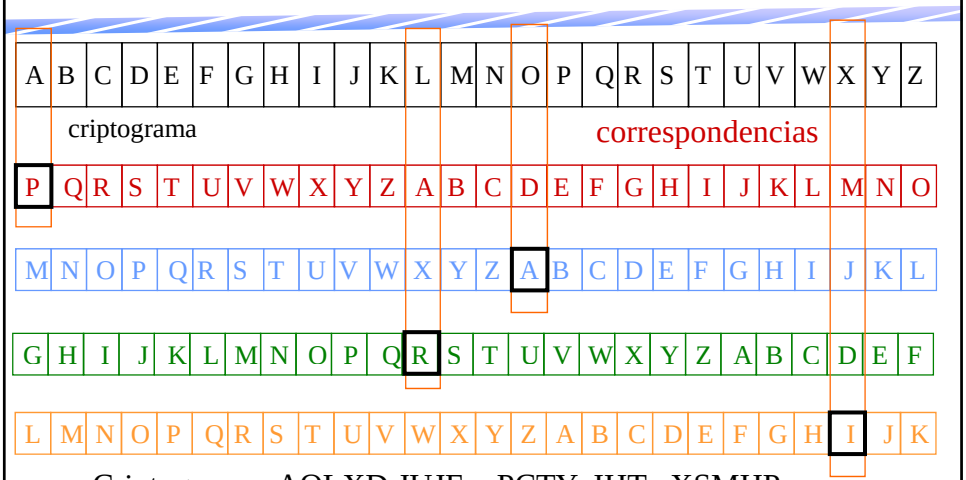
Enviando el mensaje



Mensaje: PARIS VAUT BIEN UNE MESSE
Llave: LOUPL OUPL OUPL OUP LOUPL
Criptograma: AOLXD JUJE PCTY IHT XSMHP

109

Recuperando el mensaje



Criptograma: AOLXD JUJE PCTY IHT XSMHP
Llave: LOUPL OUPL OUPL OUP LOUPL
Mensaje: PARIS VAUT BIEN UNE MESSE

110

Características Vigenere

- Resistía a ataques de análisis de frecuencia.
  - letras más comunes no se repiten con la misma frecuencia
- Emisor y receptor se ponen de acuerdo en la llave: palabra diccionario, combinación palabras o fabricarlas.
- Sistema pertenece a un criptosistema conocido como *polialfabetico*
  - utiliza varios alfabetos por mensaje

111

Algunos puntos a tomar en cuenta

- El uso de una llave (keyword) para indicar los alfabetos a usar fue idea original de Giovan Batista Bekasi en 1553
  - en 1563 Giovanni Battista Porta añade el uso de alfabetos revueltos en el sistema
  - el sistema se conoce como Viginere
- Un primer sistema de autollave, no utilizable, fue propuesto en primer lugar por Girolamo Cardano
  - fue Blaise de Vigenère quien propuso la actual forma del criptosistema de autollave, en 1585

112

Desventajas Vigenere

- Sistema no fue muy adoptado,
  - no fue muy aceptado en los próximos dos siglos
- La naturaleza polialfabetica del criptosistema de Vigenere es lo que le da su fuerza, pero lo hace muy complicado de usar.
  - el esfuerzo adicional para usarlo desalentó a mucha gente para emplearlo
- Para muchos propósitos del siglo XVII, los criptosistemas monoalfabeticos fueron adecuados.

113

Atacando criptosistema Viginere

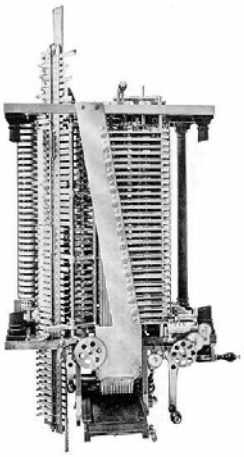
- Debido a su fortaleza al criptosistema de Vigenere se le conocio como *le chiffre indéchiffrable*.
  - los criptografos tenían una ventaja sobre los criptoanalistas (siglo XVI, XVII y XVIII)
- La figura más importante de los criptoanalistas del siglo XIX fue Charles Babbage (1791-1871)
  - primera persona intenta desarrollar máquina universal para resolver problemas
  - interesado en criptogramas desde muy pequeño
  - planea escribir libro *The Philosophy of Decephering* (no publicado)

114

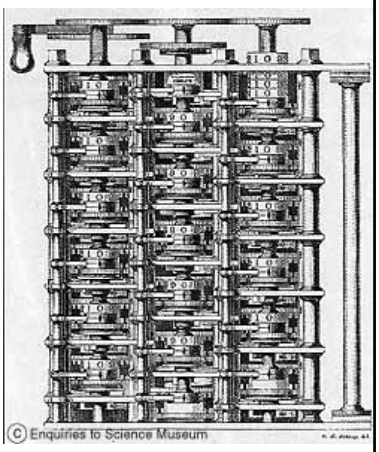
Charles Babbage

Charles Babbage

Ada Augusta,
Countess of Lovelace



Máquina Analítica



Máquina Diferencial

115

Aportaciones Babbage

- Aplicación de las matemáticas (álgebra) a la criptología.
  - sus papeles están repletos de formulas que usa para resolver criptosistemas
- Se cree que en 1854 logra resolver Viginere.
  - su descubrimiento nunca fue reconocido porque nunca lo publica
  - el descubrimiento se conoció en el siglo XX cuando un grupo de estudiantes estudió las notas de Babbage

116

Kasiski

- En la misma epoca que Babbage, su técnica fue descubierta por Friedrich Wilhelm Kasiski (1805-1881)
  - oficial prusiano retirado
- En 1863 publica Die Geheimschriften und die Dechiffrir-kunst (“Escritura secreta y el arte de decripción”)
- La técnica se conoció como la prueba de Kasiski y la contribución de Babbage fue ignorada por mucho tiempo.

117

¿Porqué Babbage no publico su trabajo?

- Babbage tenía el hábito de no terminar sus proyectos y no publicar sus descubrimientos.
- Otra alternativa:
  - su descubrimiento ocurrió después del inicio de la guerra de Crimea
  - los británicos tenían una clara ventaja sobre los enemigos rusos
  - es posible que la Inteligencia Británica le pidiera a Babbage mantener en secreto su trabajo

118

¿Y cómo lo hizo?

- La principal debilidad de Vigenere es su naturaleza cíclica.
- Si la llave es de 5 letras de largo, entonces cada cinco letras del texto claro es encriptada de acuerdo al mismo alfabeto de encripción.
- Si el criptoanalista puede identificar la longitud de la llave, el criptograma puede ser tratado como un serie de criptosistemas monoalfabéticos.

119

Ejemplo llaves en Vigenere

- hola
- furia
- invocar
- loup
- wind

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

120

La llave en Vigenere y su efecto (i)

- Si la llave es KING, la letra e puede ser encriptada como
  - la letra Q si la K de KING es usada para encriptar
  - la letra M si la I de KING es usada para encriptar
  - la letra R si la N de KING es usada para encriptar
  - la letra K si la G de KING es usada para encriptar

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F

121

Llave en Vigenere y su efecto (ii)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F

- De la misma forma palabras enteras pueden ser encriptadas de forma diferente, por ejemplo la palabra the puede ser encriptada como
  - DPR
  - BUK
  - GNO
  - ZRM

122

Llave en Vigenere y su efecto (iii)

Llave:	K I N G K I N G K I N G K I N G K I N G K I N G K I N G
Texto claro:	t h e s u n a n d t h e m a n i n t h e m o o n
Criptograma:	D P R Y E V N T N B U K W I A O X B U K W W B T

the: encriptado como DPR y después como BUK
segundo the esta desplazado 8 letras respecto al tercer the
8 es múltiplo de 4 (longitud llave)

Llave:	R U N R U N R U N R U N R U N R U N R U N R U N R U N
Texto claro:	T O B E O R N O T T O B E T H A T I S T H E Q U E S T I O N
Criptograma:	K I O V I E E I G K I O V N U R N V J N U V K H V M G Z I A

123

Patrones y tamaño llave Vigenere (i)

- Es posible que en diferentes lugares de un criptograma se cuente con secuencias idénticas de letras aparezcan.
- Estos patrones pueden proporcionar información de periodicidad dentro del texto.
- Por ejemplo:

Texto claro:	R E Q U E S T S A D D I T I O N A L T E S T
Llave:	T E L E X T E L E X T E L E X T E L E X T E
Criptograma:	C A V K T B L T E U Q W S W J G E A L T B L

124

Patrones y tamaño llave en Vigenere (ii)

Texto claro:	R E Q U E S T S A D D I T I O N A L T E S T
Llave:	T E L E X T E L E X T E L E X T E L E X T E
Criptograma:	C A V K T B L T E U Q W S W J G E A L T B L

- El texto en claro contiene la secuencia de letras EST dos veces
- En ambos casos, la misma sección de la llave es usada para encriptar
  - la secuencia TBL en el criptograma es la misma para ambos casos
- Lo anterior se debe a que la secuencia EST esta posicionada exactamente en un *número múltiplo de la longitud de la llave*

125

Algunas observaciones

- El número de letras entre dos repeticiones del criptograma esta asociado con el numero de veces que la llave se repite.
- Ejemplo:
  - en el primer ejemplo la repetición de *BUK* se debe a que el segundo *the* esta desplazado ocho letras con respecto al tercer *the*, y ocho es un múltiplo de la longitud de la llave.
- El análisis de los intervalos entre las repeticiones puede dar a conocer la longitud de la llave.

126

Un criptograma resultado de Vigenere

```
W U B E F I Q L Z U R M V O F E H M Y M W T
I X C G T M P I F K R Z U P M V O I R Q M M
W O Z M P U L M B N Y V Q Q Q M V M V J L E
Y M H F E F N Z P S D L P P S D L P E V Q M
W C X Y M D A V Q E E F I Q C A Y T Q O W C
X Y M W M S E M E F C F W Y E Y Q E T R L I
Q Y C G M T W C W F B S M Y F P L R X T Q Y
E E X M R U L U K S G W F P T L R Q A E R L
U V P M V Y Q Y C X T W F Q L M T E L S F J
P Q E H M O Z C I W C I W F P Z S L M A E Z
I Q V L Q M Z V P P X A W C S M Z M O R V G
V V Q S Z E T R L Q Z P B J A Z V Q I Y X E
W W O I C C G D W H Q M M V O W S G N T J P
F P P A Y B I Y B J U T W R L Q K L L L M D
P Y V A C D C F Q N Z P I F P P K S D V P T
I D G X M Q Q V E B M Q A L K E Z M G C V K
U Z K I Z B Z L I U A M M V Z
```

127

Rompiendo Vigenere

- Primer paso: ver las secuencias de letras que aparecen más de una vez en el criptograma.
- Dos formas de que se haya producido lo anterior:
  - la misma secuencia de palabras fue encriptada usando la misma parte de la llave
  - pequeña probabilidad de que dos diferentes secuencias de letras en el texto claro hayan sido encriptadas usando diferentes partes de llave, y que por coincidencia haya llevado a la misma secuencia en el criptosistema

128

Encontrando repeticiones/patrones en el critptograma

W U B E F I Q L Z U R M V O F E H M Y M W T
I X C G T M P I F K R Z U P M V O I R Q M M
W O Z M P U L M B N Y V Q Q Q M V M V J L E
Y M H F E F N Z P S D L P P S D L P E V Q M
W C X Y M D A V Q E E F I Q C A Y T Q O W C
X Y M W M S E M E F C F W Y E Y Q E T R L I
Q Y C G M T W C W F B S M Y F P L R X T Q Y
E E X M R U L U K S G W F P T L R Q A E R L
U V P M V Y Q Y C X T W F Q L M T E L S F J
P Q E H M O Z C I W C I W F P Z S L M A E Z
I Q V L Q M Z V P P X A W C S M Z M O R V G
V V Q S Z E T R L Q Z P B J A Z V Q I Y X E
W W O I C C G D W H Q M M V O W S G N T J P
F P P A Y B I Y B J U T W R L Q K L L L M D
P Y V A C D C F Q N Z P I F P P K S D V P T
I D G X M Q Q V E B M Q A L K E Z M G C V K
U Z K I Z B Z L I U A M M V Z

129

Repeticiones y espacios entre ellas

Secuencia repetida	Espacio repetido	Posible longitud de la llave (o factores)																	
		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
E-F-I-Q	95				x														x	
P-S-D-L-P	5				x															
W-C-X-Y-M	20	x		x	x					x										x
E-T-R-L	120	x	x	x	x	x		x		x		x			x					x

En el caso de la repetición **W-C-X-Y-M**

1. Llave es 1 letra de largo y es reciclada 20 veces en la encripción
2. Llave es 2 letras de largo y es reciclada 10 veces en la encripción
3. Llave es 4 letras de largo y es reciclada 5 veces en la encripción
4. Llave es 5 letras de largo y es reciclada 4 veces en la encripción
5. Llave es 10 letras de largo y es reciclada 2 veces en la encripción
6. Llave es 20 letras de largo y es reciclada 1 veces en la encripción

130

La longitud de la llave

- Se asume que la longitud de la llave es de 5
  - sea la llave: L_1-L_2-L_3-L_4-L_5
- En base a la naturaleza de Viginere podemos deducir que
  - L_1 encripta la 1,6,11,16, ... letras
  - L_2 encripta la 2,7,12,17, ... letras
  - L_3 encripta la 3,8,13,18, ... letras
  - L_4 encripta la 4,9,14,19, ... letras
  - L_5 encripta la 5,10,15,20, ... letras

131

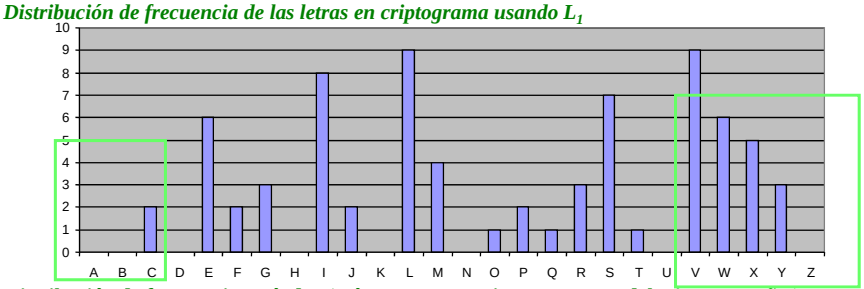
Letras cifradas con L_1

W U B E F I Q L Z U R M V O F E H M Y M W T
I X C G T M P I F K R Z U P M V O I R Q M M
W O Z M P U L M B N Y V Q Q Q M V M V J L E
Y M H F E F N Z P S D L P P S D L P E V Q M
W C X Y M D A V Q E E F I Q C A Y T Q O W C
X Y M W M S E M E F C F W Y E Y Q E T R L I
Q Y C G M T W C W F B S M Y F P L R X T Q Y
E E X M R U L U K S G W F P T L R Q A E R L
U V P M V Y Q Y C X T W F Q L M T E L S F J
P Q E H M O Z C I W C I W F P Z S L M A E Z
I Q V L Q M Z V P P X A W C S M Z M O R V G
V V Q S Z E T R L Q Z P B J A Z V Q I Y X E
W W O I C C G D W H Q M M V O W S G N T J P
F P P A Y B I Y B J U T W R L Q K L L L M D
P Y V A C D C F Q N Z P I F P P K S D V P T
I D G X M Q Q V E B M Q A L K E Z M G C V K
U Z K I Z B Z L I U A M M V Z

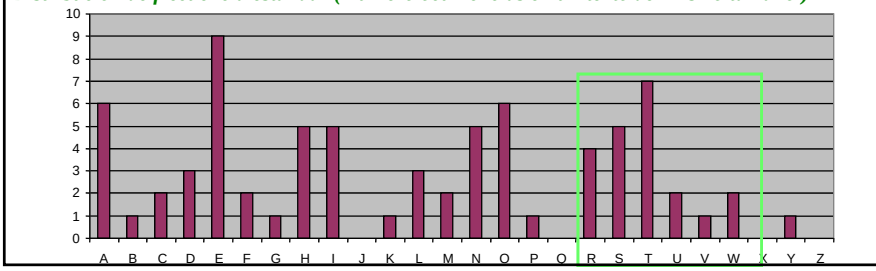
132

Análisis del criptograma con L_1

Distribución de frecuencia de las letras en criptograma usando L_1



Distribución de frecuencia estándar (número ocurrencias en un texto del mismo tamaño)



133

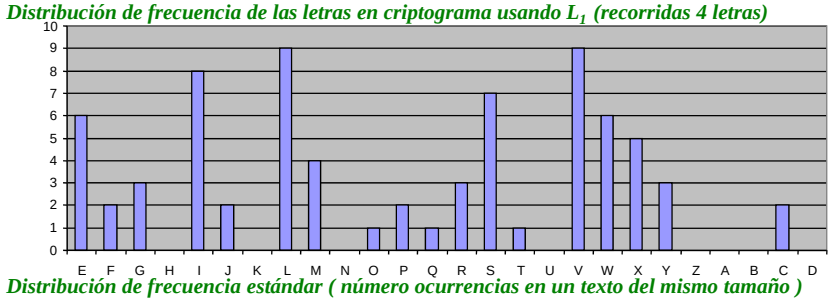
Deducciones del análisis

- En esquema distribución estándar:
  - tres picos RST y una gran depresión después U a Z
- En esquema distribución criptograma
  - lo más parecido es VWX seguido por depresión de la Y a D
- Creencia:
  - todas las letras encriptadas con L_1 fueron desplazadas cuatro lugares.
  - L_1 define un alfabeto que empieza en E F G H ...
  - muy probable que L_1 = E

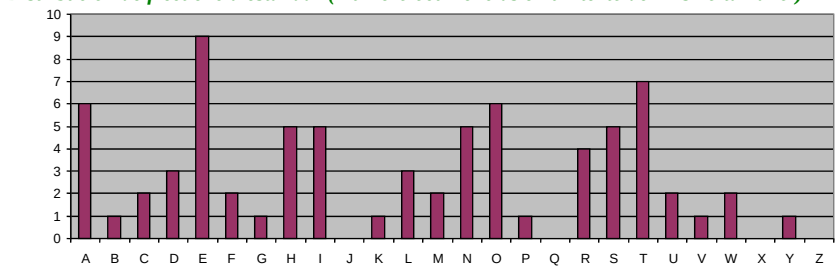
134

Comparando distribución de las letras cifradas con L1, pero recorriendo 4 letras.

Distribución de frecuencia de las letras en criptograma usando L_1 (recorridas 4 letras)



Distribución de frecuencia estándar (número ocurrencias en un texto del mismo tamaño)



135

¿Qué se ha hecho?

1. Repeticiones en el criptograma permitieron identificar la longitud de la llave
   - llave de cinco letras de largo
2. Criptograma dividido en cinco partes
   - cada una encriptada de acuerdo a un criptosistema monoalfabético
3. Analizando la parte encriptada con L_1 se pudo deducir que esta letra es E
4. Necesario aplicar metodología del punto 3 con las otras cuatro partes del criptograma.

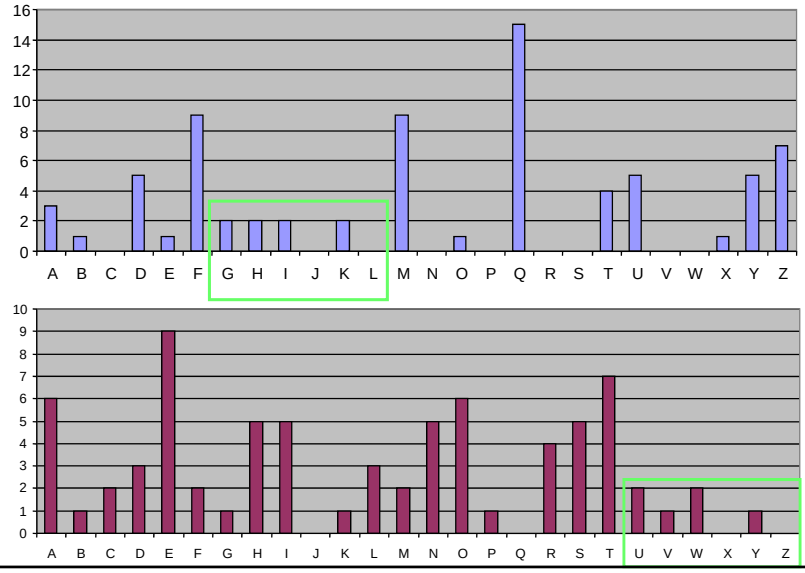
136

Letras cifradas con L_2

```
W U B E F  I Q L  Z U R M V O F E H M Y M W T
 I X C G T M P  I  F K R Z U P M V O  I  R Q M M
W O Z M P U L M B N Y V Q Q Q M V M V J L E
Y  M H F E F N Z  P S D L P  P S D L  P E V Q M
W C X Y M D A V Q E E F I Q C A Y T Q O W C
X  Y M W M S E M E F C F W Y E Y Q E T R L I
Q  Y C G M T W C W F B S M Y F P L R X T Q Y
E  E X M R U L U K S G W F P T L R Q A E R L
U  V P M V Y Q Y C X T W F Q L M T E L S F J
P  Q E H M O Z C I W C I W F P Z S L M A E Z
 I Q V L Q M Z V P P X A W C S M Z M O R V G
V  V Q S Z E T R L Q Z P B J A Z V Q I Y X E
W W O I C C G D W H Q M M V O W S G N T J P
F  P P A Y B I Y B J U T W R L Q K L L L M D
P Y V A C D C F Q  N Z P I F P P K S D V P T
I  D G X M Q Q V E B M Q A  L K E Z M G C V K
U Z K  I  Z  B Z  L I  U A M M V Z
```

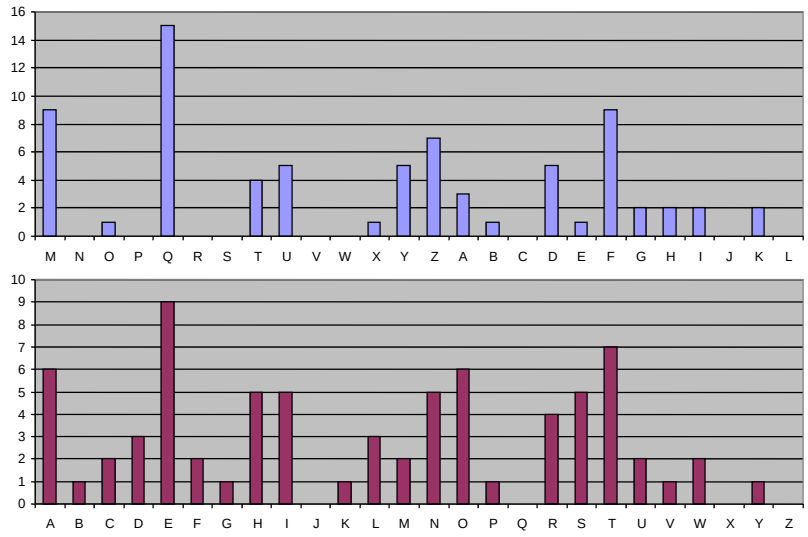
137

Análisis del criptograma con L_2



138

Análisis del criptograma recorrido 12 lugares con el estándar



139

El texto claro

Llave:
EMILY

Sit the down, and have no shame,
Cheek by jowl, and knee by knee:
What care I for any name?
What for order or degree?

Let me screw thee up a peg:
Let me loose thy tongue with wine:
Callest thou that thing a leg?
Which is thinnest? thine or mine?

Thou shalt not be saved by works;
Thou has been a sinner too:
Ruined trunks on withered forks,
Empty scarecrows, I and you!

Fill the cup, and fill the can:
Have a rouse before the morn:
Every moment dies a man,
Every moment one is born

140