

HISTORIA DE LA CRIPTOGRAFÍA

ALAN REYES-FIGUEROA

CRIPTOGRAFÍA Y CIFRADO DE INFORMACIÓN

(AULA 02) 13.JULIO.2021

Algunas definiciones

- **Cifrado:** Algoritmo utilizado para transformar texto plano en texto cifrado.
- **Texto cifrado:** El mensaje codificado o cifrado.
- **Texto plano:** El mensaje original o información que se desea mantener segura.
- **Criptanálisis:** También conocido como descifrado de código; es el estudio de principios y métodos para descifrar texto cifrado sin conocer la clave.
- **Criptografía.** El estudio de cómo alterar un mensaje para que al interceptarlo alguien no pueda leerlo sin el algoritmo y la clave adecuados.
- **Criptología:** Campo de estudio más completo e incluye tanto criptografía como criptanálisis.
- **Descifrar (decriptar):** Convertir el texto cifrado en texto sin formato.
- **Cifrar (encriptar):** Convertir el texto sin formato en texto cifrado.
- **Clave (llave):** Información, generalmente algún tipo de número, que se utiliza con el algoritmo para cifrar o descifrar el mensaje.
- **Espacio de clave:** El número total de claves posibles que se podrían usar. Por ejemplo, DES usa una clave de 56 bits; por lo tanto, el número total de claves posibles, o la clave es 2^{56} .

Cifrados de sustitución:

Los primeros cifrados registrados de la historia registrado son cifrados de sustitución. Con este método, cada letra del texto plano se sustituye por alguna letra del texto cifrado de acuerdo con algún algoritmo.

$$f : \mathbf{a} \mapsto \mathbf{k}, \quad \mathbf{b} \mapsto \mathbf{n}, \quad \mathbf{c} \mapsto \mathbf{u}, \dots$$

Hay dos tipos de cifrados de sustitución:

- *Monoalfabéticos*: Una letra dada del texto plano siempre se sustituye por la misma letra correspondiente del texto cifrado. Por ejemplo, una **a** en el texto original siempre sería una **k** en el texto cifrado.
- *Polialfabéticos*: La sustitución de polialfabética utiliza sustituciones múltiples, de modo que, por ejemplo, una **a** en el texto plano es a veces una **k** ya veces una **w** en el texto cifrado.

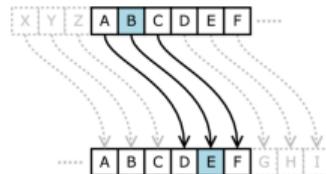
Historia

Cifrado Caesar:

De acuerdo al historiador Gaius Suetonius Tranquillus (*circa* 70–130 a.D.), Julio César usaba este cifrado para sus mensajes militares.

El cifrado consistía en correr todas las letras del alfabeto tres posiciones a la derecha:

$$f : a \mapsto d, \quad b \mapsto e, \quad c \mapsto f, \quad \dots, \quad z \mapsto c.$$



Ejemplo: *ATACAR AL AMANECER* se convierte en *DWDFDU DO DPDQHFHU*.

0	1	2	3	4 ...	23	24	25
A	B	C	D	E ...	X	Y	Z

En una notación moderna, el cifrado *Caesar* sería una función matemática de la forma

$$E(\mathbf{x}) = \mathbf{x} + 3 \pmod{26} \quad \text{y} \quad D(\mathbf{x}) = \mathbf{x} - 3 \pmod{26}.$$

En el contexto moderno, podemos pensar que existen varios cifrados Caesar, en donde en lugar de 3 recorremos las letras por una cantidad k . Elegimos como llave k alguno de los números del 0 al 25:

$$E(k, \mathbf{x}) = \mathbf{x} + k \pmod{26} \quad \text{y} \quad D(k, \mathbf{x}) = \mathbf{x} - k \pmod{26}.$$

El espacio clave del cifrado Caesar es 26, demasiado pequeño. Por ejemplo AES usa una llave de 128 bits, con un espacio clave de $2^{128} \approx 3.4 \times 10^{28}$.

Cifrado Atbash:

Los escribas hebreos que copiaron el libro bíblico de Jeremías usaron la sustitución de cifrado *Atbash*. Aplicar el cifrado Atbash es bastante simple: simplemente invierte el orden de las letras del abecedario.

Por ejemplo, en inglés, **a** se convierte en **z**, **b** se convierte en **y**, **c** se convierte en **x** y así sucesivamente.

Ejemplo: *ATACAR AL AMANECER* se convierte en *ZGZXZI ZO ZNZMVXVI*.

En notación matemática tendríamos:

$$E(\mathbf{x}) = 25 - \mathbf{x} \pmod{26} \quad \text{y} \quad D(\mathbf{x}) = 25 - \mathbf{x} \pmod{26},$$

o más generalmente

$$E(k, \mathbf{x}) = k - \mathbf{x} \pmod{26} \quad \text{y} \quad D(k, \mathbf{x}) = k - \mathbf{x} \pmod{26},$$

Cifrados Afines: Los cifrados afines son cualquier cifrado alfabético de sustitución única de la forma

$$E(a, b, \mathbf{x}) = a\mathbf{x} + b \pmod{M}, \quad \text{y} \quad D(a, b, \mathbf{x}) = a^{-1}(\mathbf{x} - b) \pmod{M},$$

donde M es el tamaño del alfabeto, y a es un entero que no posee factor común con M . La llave es el par $k = (a, b)$

Ejemplo: Con $a = 3$, $b = 8$, el mensaje *ATACAR AL AMANECER* se convierte en *INIOIH IP ISIVUOUH*.

Ejemplo: Con $a = 13$, $b = 8$, el mensaje *ATACAR AL AMANECER* se convierte en *IVIIIV IV IIIVIIIIV*.

Para $M = 26$, en este caso el espacio clave es $12 \times 26 = 286 < 26^2$.

ROT 13:

ROT 13 es un cifrado trivial de sustitución única. ROT es la abreviatura de rotar: cada letra se rota 13 a la derecha:

$$E(k, \mathbf{x}) = \mathbf{x} + 13 \pmod{26} \quad \text{y} \quad D(k, \mathbf{x}) = \mathbf{x} + 13 \pmod{26},$$

Como es de imaginar, este cifrado no es seguro. Sin embargo, en realidad se usa en algunas situaciones. Por ejemplo

- algunas de las claves del registro de Microsoft Windows están cifradas con ROT 13.
- a finales de la década de 1990, *Netscape Communicator* utilizó ROT 13 para almacenar contraseñas de correo electrónico.

Historia

En su forma más general, un cifrado de sustitución está dado por una clave k que es una permutación del conjunto de letras en el alfabeto. Por ejemplo, para las 26 letras del inglés

$$k = \begin{pmatrix} 0 & 1 & 2 & \dots & 23 & 24 & 25 \\ 16 & 7 & 23 & \dots & 4 & 19 & 11 \end{pmatrix}.$$

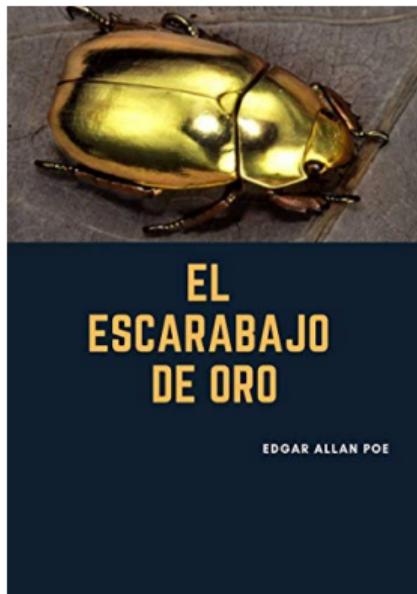
En este caso,

$$E(k, \mathbf{x}) = k(\mathbf{x}) \pmod{M}, \quad \text{y} \quad D(k, \mathbf{x}) = k^{-1}(\mathbf{x}) \pmod{M},$$

En general, tenemos un espacio de clave de $M!$.

Para $M = 26$, el espacio de claves de $26! = 4.033 \times 10^{26} \approx 2^{88}$ claves.

Obs! Aunque el espacio sea grande, esto no garantiza seguridad.

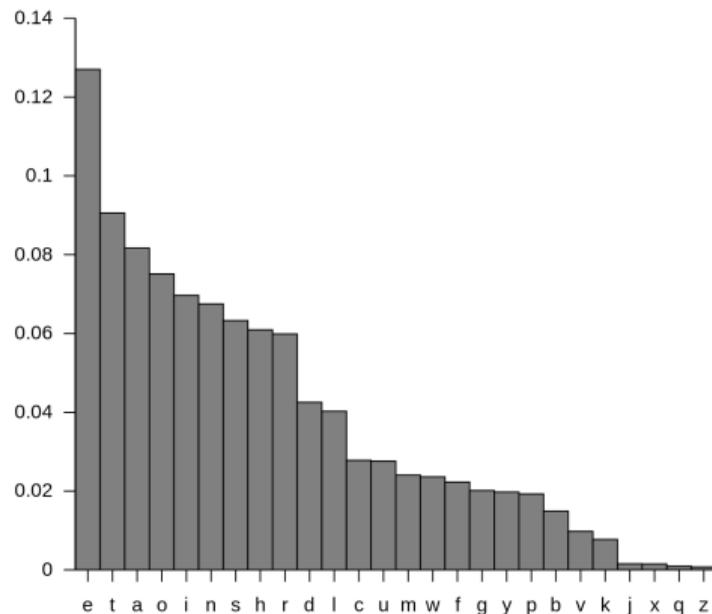
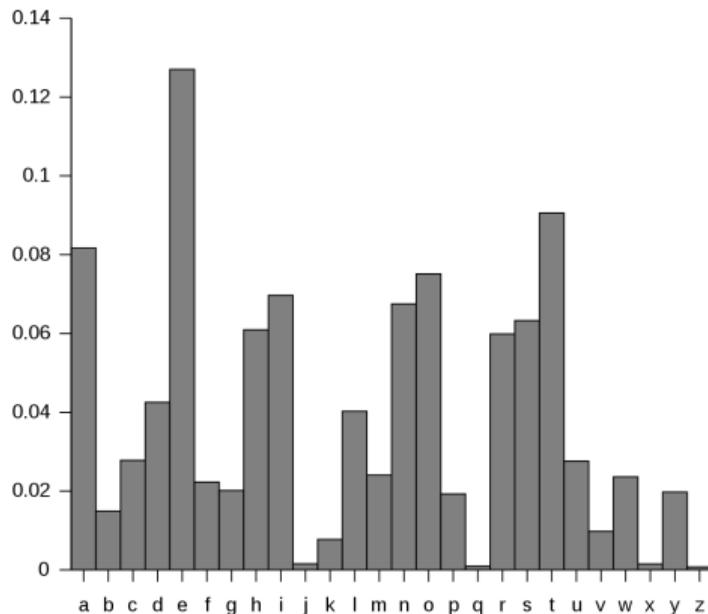


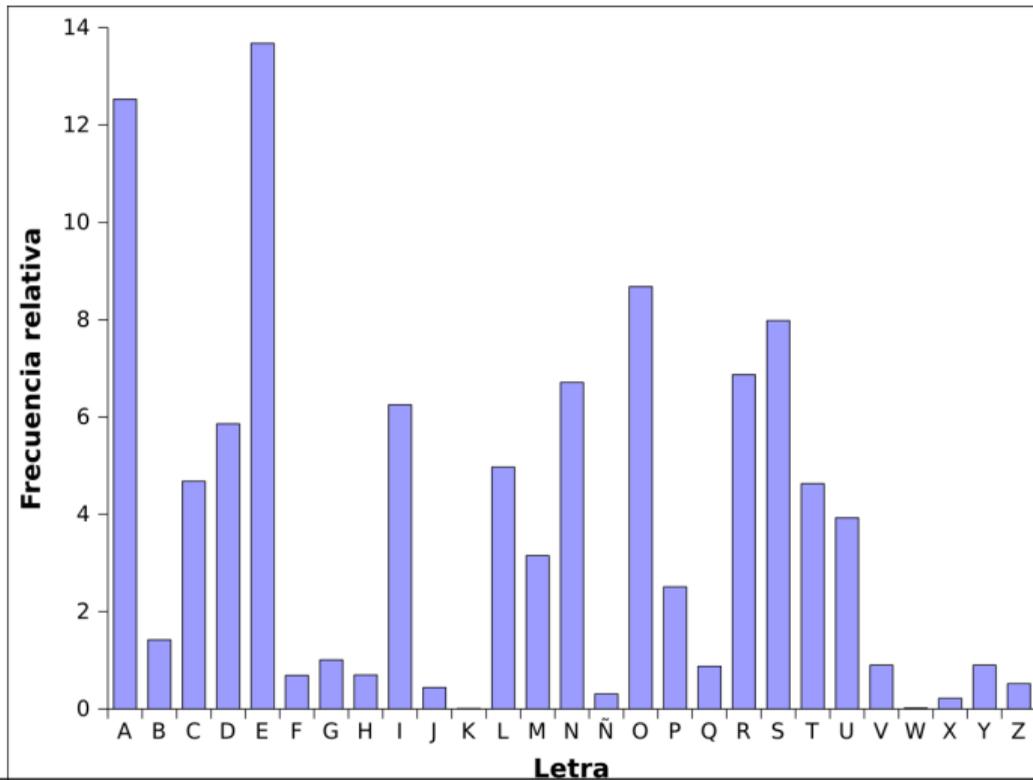
53‡‡+305))6*;4826)4‡.)4‡);806*;
48+8¶60))85;1‡(;:‡*8+83(88)
5*+;46(;88*96*?;8)*‡(;485);5*
+2:*‡(;4956*2(5*-4)8¶8*;
4069285);)6+8)4‡‡;1(‡9;48081;8:
8‡1;48+85;4)485+528806*81(‡9;48;(88;
4(‡?34;48)4‡;161;:188;‡?;

Criptograma incluido en el relato *El escarabajo de oro* de Edgar Allan Poe.

Historia

Frecuencias de letras: La falla principal de los cifrados de sustitución es que preservan la distribución de las frecuencias. *Etaoin shrdlu*.





¿Cómo quebrar un cifrado por sustitución?:

- Usamos la frecuencia de letras o símbolos en el alfabeto. Por ejemplo, en inglés: $e : 12.7\%$, $t : 9.1\%$, $a : 8.1\%$, ...
Comparamos la frecuencia de las letras en el texto cifrado (muestra) contra la distribución teórica del idioma (población).
- Usamos la frecuencia de pares de letras o **bigramas**: *he, an, in, th, ...*
- Usamos la frecuencia de tripletas o **trigramas, 4-gramas, ...** en general **n -gramas**.

En Python existen librerías para procesamiento de lenguaje (NLP), como NLTK, gensim, ...

Tabla de Polibius:

Encerraba las 26 letras en un tablero de 5×5 . Sustituía cada caracter por sus coordenadas en la tabla.

	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

Tabla de Polibius.

Otros mecanismos:



(a) Disco de cifrado. (b) Disco de Jefferson.

Cifrado Vigenère:

Un cifrado de Vigenère (Roma, siglo XV) utiliza una tabla que consta de diferentes cifrados Caesar en secuencia. Hace que el análisis de frecuencia sea más difícil.

	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

Ejemplo: El mensaje *WHATANICEDAYTODAY*, y la clave *CRYPTO*.

W	H	A	T	A	N	I	C	E	D	A	Y	T	O	D	A	Y
C	R	Y	P	T	O	C	R	Y	P	T	O	C	R	Y	P	T

Sumamos módulo 26.

W	H	A	T	A	N	I	C	E	D	A	Y	T	O	D	A	Y
C	R	Y	P	T	O	C	R	Y	P	T	O	C	R	Y	P	T
Z	Z	Z	J	U	C	L	U	D	T	U	N	W	G	C	Q	S

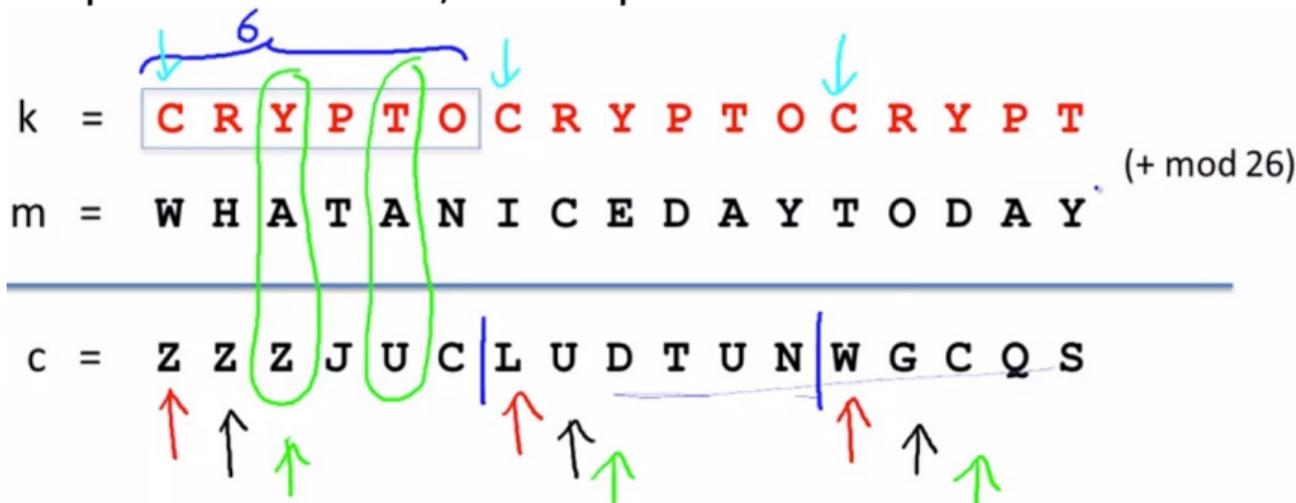
Para decodificar se sigue el mismo proceso, pero se resta $text - key$ (mod 26).

Historia

Una formulación matemática del cifrado de Vigenère sería: $M =$ tamaño del alfabeto. $N =$ tamaño de la clave.

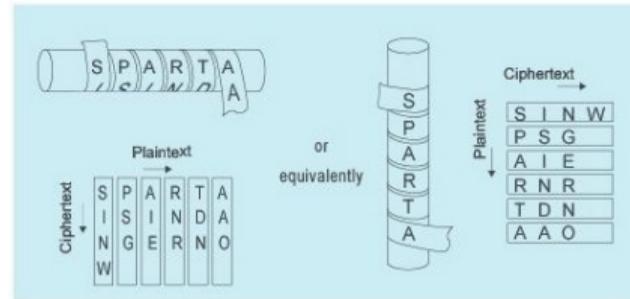
$$E(k, \mathbf{x}) = \mathbf{x}_i + k_{i(\text{mod } N)} \pmod{M}, \quad D(k, \mathbf{x}) = \mathbf{x}_i - k_{i(\text{mod } N)} \pmod{M}.$$

Para decriptar este cifrado, aún es posible hacer análisis de frecuencias.



Cifrados de transposición:

Los cifrados de sustitución son aquellos en los que cada letra se reemplaza por otra letra (o símbolo) de alguna manera sistemática. Sin embargo, el orden en el que aparecen las letras sigue siendo el mismo. En contraste, los cifrados de transposición son aquellos para los que las letras siguen siendo las mismas, pero el orden está todo mezclado. Fueron usados desde muchos siglos atrás. Por ejemplo, los antiguos egipcios y griegos utilizaron un artefacto llamado escítala *scytale*.



Cifrados de transposición matricial:

Son cifrados donde el texto plano se escribe dentro de una matriz de dimensiones predefinidas $m \times n$, fila por fila. El texto cifrado consiste en tomar las letras, columna por columna.

T	R	O	O	P	S	H	E	A	D
I	N	G	W	E	S	T	N	E	E
D	M	O	R	E	S	U	P	P	L
I	E	S	S	E	N	D	G	E	N
E	R	A	L	D	U	B	O	I	S
M	E	N	T	O	A	I	D		

Por ejemplo el texto *TROOPS HEADING WEST NEED MORE SUPPLIES*, codificado en una matriz de 6×10 , resulta en *TIDIE MRNME REOGO SANOW RSLTP EEEDO SSSNU AHTUD BIENP GODAE PEIDE LNS.*

Cifrados de transposición por rearreglo: Similar a los de transposición matricial, pero ahora las columnas de la matriz se reordenan siguiendo alguna permutación.

```
NSIAN LIGRP EEBIU LALCL HPTEW ROEUR SRNEE
DEBRD TEYHE PPOES LIUMS BTEIT TTDOE DHAI I
MNTAS RINTO FSOAN LIGGV EORME NNAET DHNTU
UESPT RAOSI NRGAE ADURD GEAIS ANBAT YIIDV
INSOF HOTGV EORME NNIET TDNOS IITCA NTDEN
SAAPR EETDA TPRET MNITS NEOHC PUMOD ENRUL
PBCFI OMRAE CTIAE OHPES WRREU RDRNE DYEBH
PTEOL EPIFE SRTIS IIDVE BDDTE EWNWE TDSOI
ICTNG VTORM ENNSE TNTAD ETHNE OHPTO RIALN
LTEOT TEDOC SAHBI UDIEV DAODM GINDT NSITN
CASPD ERTAA DPEER MATNS ETECH NAOED BEULE
USCIY RTRSA ISOET HRTEG TIHOT SFEEH PPEOL
HDTEF EIFET RNOEG VNERM TWNSL CILNR OTLAO
EHTCO EAHRT ETHAE SMIET MHTTA AHECI LWLEO
BCTON RLDLE YTBIE FSLRM FOHFT EDREE LSAIP
PTARU ENBRM EITFF OEYN
```

Cifrado de Valla (*rail fence*):

El cifrado de la valla de riel codifica un mensaje escribiéndolo hacia arriba y hacia abajo en diagonal o *zig-zag* sobre “rieles” sucesivos, o filas, en una valla imaginaria.

Por ejemplo, el mensaje secreto *THIS IS A SECRET MESSAGE* codificado en 4 carriles es

```
T . . . . A . . . . T . . . . G .  
. H . . . S . S . . . E . M . . . A . E  
. . I . I . . . E . R . . . E . S . . .  
. . . S . . . . . C . . . . . S . . . .
```

Cifrado de libro o cifrado de Ottendorf: Tanto el remitente como el destinatario de un mensaje secreto deben tener la misma copia de un libro. El emisor codifica el mensaje secreto palabra por palabra reemplazando la palabra con el mapa de coordenadas a la ubicación dentro del libro elegido: página - párrafo/línea - letra. Por ejemplo: página 39, párrafo 7, palabra 12, las coordenadas del texto cifrado son 39 : 7 : 12.

Cifrado de Playfair:

Sustituye bigramas. En lugar de codificar un mensaje reemplazando caracteres individuales, los reemplaza en pares. Para codificar un mensaje, el cifrado de Playfair usa una palabra clave para generar una tabla de codificación de 5×5 y luego sigue 4 reglas para codificar los bigramas.

Para crear una tabla de Playfair, use una palabra clave para llenar espacios en una tabla de 5 por 5 de arriba hacia abajo, de izquierda a derecha. Las letras I y J se colocan en el mismo espacio para reducir el alfabeto de 26 caracteres a 25.

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

Tabla de Playfair con la clave *PICTURE FRAME*

Historia

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

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

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

El texto se separa en bigramas. Por ejemplo *KILL THE SPY* se convierte en *KI LL TH ES PY*. Codificamos éstos siguiendo 4 reglas simples:

- Si ambas letras en el digrama son iguales, agregar una X después de la primera letra para dividir las en un nuevo bigrama y continúe codificando el mensaje.
- Si aparecen en la misma fila, mover los caracteres una posición a la derecha.
- Si aparecen en la misma columna, mover los caracteres hacia abajo una posición.
- Si forman las esquinas de un rectángulo, reemplazar con las esquinas opuestas.

Ejemplo: texto *KILL THE SPY*, bigramas *KI LX LT HE SP YX*, cifrado *DU OV QP DA LU ZY*.

El Gran Cifrado:

Utilizado por el gobierno francés hasta principios del siglo XIX. Inventado por la familia Rossignol, una familia francesa con varias generaciones de criptógrafos, todos los cuales sirvieron a la corte francesa.

Usaba 587 símbolos diferentes que representaban sílabas. Para evitar el análisis de frecuencia, el texto cifrado incluía nulos o números que no significaba nada. También había trampas o códigos que indicaban que el destinatario debía ignorar el mensaje codificado anterior.

Cifrado Copiale:

Cifrado homofónico. Era un libro de 105 páginas y 75,000 caracteres, y no se quebró durante muchos años. El cifrado de Copiale usó un código de sustitución complejo que usaba símbolos y letras, de varios alfabetos. Se cree que el documento data del 1700 de una sociedad secreta llamada la *Alta orden ocultista ilustrada de Wolfenbüttel*. Se rompió en 2011 mediante mecanismos computacionales.

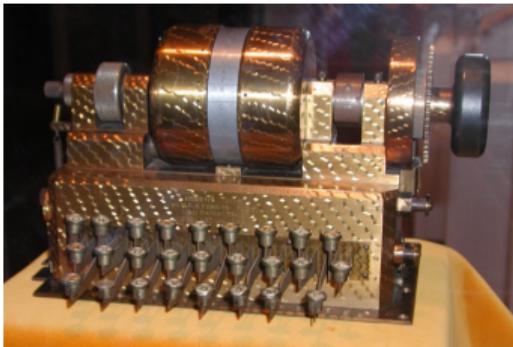
Otros cifrados históricos: Siglos XVII a XIX

- Cifrados de 2-cuadrados y de 4-cuadrados,
- Cifrado de Hill, arregla el texto en bloques de $m \times m$. Multiplica por una matriz de enteros, invertible, módulo 26.
- Cifrado ADFGVX,
- Cifrados *Bifid* y *Trifid*,
- Cifrado de Gronsfeld,
- Cifrado de Vernam,
- Cifrado de d'Agapeyeff.

Cifrados Siglo XX

Máquinas de rotor: (1870 - 1944).

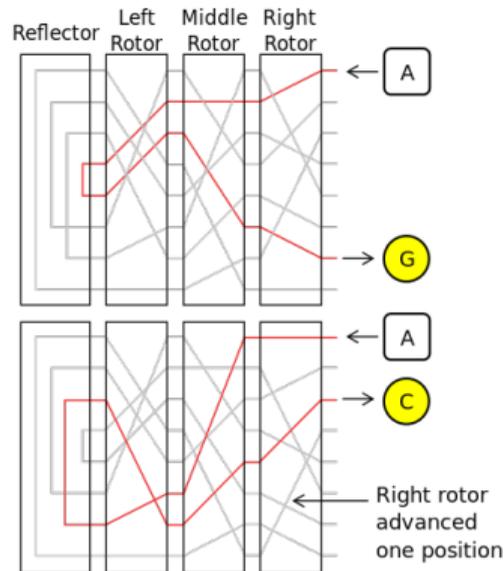
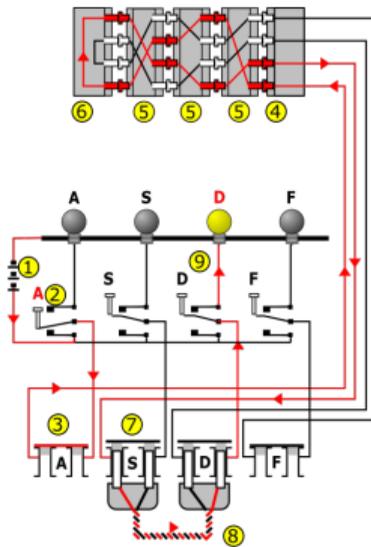
Cifrado con base en mecanismos electromecánicos. Por lo general máquinas a base de rotores. Muchos desarrollos: BID/60, Enigma, Fialka, Hagelin's family C-36, C-52, CD-57, M-209; Hebern, HX-63, KL-7, Lacida, Lorenz SZ 40/42, M-325, Mercury, NEMA, OMI, Portex, RED, Siemens and Halske T52, SIGABA, SIGCUM, Typex.



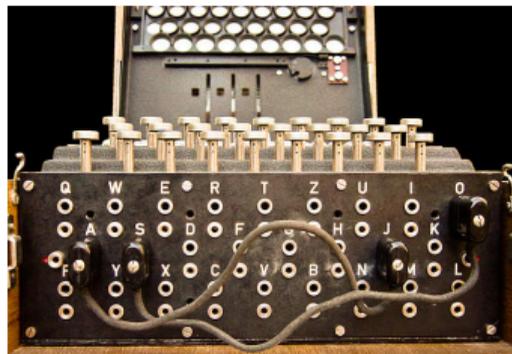
Algunas máquinas de cifrado: (a) Hebern, (b) Typex, (c) SIGCUM.

Cifrados Siglo XX

De estas, la más popular fue la Enigma, utilizada por las fuerzas alemanas en la Segunda Guerra.



Cifrados Siglo XX



Geheime Kommandosache! Jede einzelne Ingeßchlüssel ist geheim. Mitne' 2 im Flugzeug verboten' Nr. 00190

Luftwaffen-Maschinen-Schlüssel Nr. 649

Achtung! Schlüsselmittel dürfen nicht unverseht in Feindeshand fallen. Bei Gefahr restlos und frühzeitig vernichten.

Maschinen-Nr.	Walzenlage			Ringstellung	Stellereverbindungen										Anzugsgruppen							
	I	II	III		an der Umkehrtafel		am Stecherbrett															
				1	2	3	4	5	6	7	8	9	10	1	2	3						
649	31	I	V	III	14	06	24		SZ	GT	DV	KU	FO	MY	EW	JN	IX	LQ	wny	dgy	eXB	rZg
649	30	IV	III	II	05	26	02		IS	EV	MX	RW	DT	UZ	JQ	AO	CH	NY	ktl	acw	zsj	wao
649	29	III	II	I	12	24	03	KM AX PZ 00	DJ	AT	CV	IO	ER	QS	LW	PZ	PH	BH	loc	zen	ovw	wvd
649	28	II	III	V	06	08	16	DI CN BR PV	CR	PV	AI	DK	OT	MQ	EU	BX	LP	GJ	lrb	cid	ude	rzh
649	27	III	I	IV	11	03	07	LT EQ HS UW	DY	IN	BV	GR	AM	LO	PP	HT	EX	UW	woj	fbh	vct	uis
									VZ	AL	RT	KO	GO	EI	BJ	DU	PS	HP	xle	gbo	uev	rxp

Otros esquemas de cifrado históricos:

- **Cifrado Lorentz:** grupo de máquinas de rotor SZ40, SZ42A, SZ42B, usadas por los alemanes en la Segunda Guerra.
- **SIGABA**, Mercury M-325 (U.S. Patent 2877, 565) usada por Estados Unidos en los 1940s. M228 o SIGCUM.
- **HX-63:** (1950s) Desarrollada por los franceses.
- **Navajo Codetalkers:** Soldados Navajo (1930s-1950s).
- **Cifrado VIC:** Usado por los soviéticos en la Guerra Fría.
- **Sistemas IFF:** (1939–) IFF = *Identify Friend or Foe*. Envío de señales preestablecidas a intervalos específicos.
- **NSA:** (1952–) Agencia de gobierno dedicada a la seguridad. Precursores: *Cipher Bureau* (1929-1942), Signal Intelligence Service (1942-1952). Desarrollo de muchos protocolos actuales: DES, AES.

Para leer más sobre historia de la criptografía:

- d'Agapeyeff, A. (2016). *Codes and ciphers: A history of cryptography*. Read Books Ltd.
- Dooley, J. F. (2018). *History of cryptography and cryptanalysis: Codes, Ciphers, and their algorithms*. Springer.
- Mollin, R. A. (2000). *An introduction to cryptography*. CRC Press.
- Singh, S. (2000). *The code book: the science of secrecy from ancient Egypt to quantum cryptography*. Anchor.