Cifrado de Vernam

José Galaviz

Telecomunicaciones

- Samuel Morse, telégrafo eléctrico + código, 1844 (demostración pública).
- Émile Baudot, código de cinco bits, 1877.
- Donald Murray, teclado + código de Baudot modificado, 1901.
  - Adoptado por Western Union.
  - Adoptado como código internacional (1924).

Código de Baudot-Murray

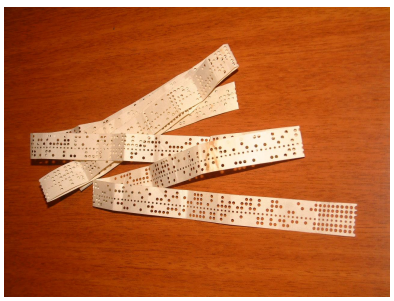
LETTERS / FIGURES	A –	B ?	C :	D WHO ARE YOU	E 3	F %	G @	H £	I 8	J BELL	K (	L)	M .	N ,	O 9	P 0	Q 1	R 4	S ’	T 5	U 7	V =	W 2	X /	Y 6	Z +	CARRIAGE RETURN	LINE FEED	LETTERS	FIGURES	SPACE	ALL-SPACE NOT IN USE
CODE ELEMENTS 1	●	●		●	●	●				●	●						●		●		●		●	●	●	●			●	●		
CODE ELEMENTS 2	●	○	●	○	○	○	●	○	●	●	●	○	○	○	○	●	●	○	○	○	●	●	●	○	○	○	○	●	●	●	○	○
CODE ELEMENTS 3			●			●	○	●	●	○			●	●		●	●		●		○	●		●	●				●	●		
CODE ELEMENTS 4		●	●	●		●	●			●	●		●	●	●			●				●		●	●		●		●	●		
CODE ELEMENTS 5		●					●	●			●	●	●		●	●	●			●		●	●	●	●	●			●	●	○	

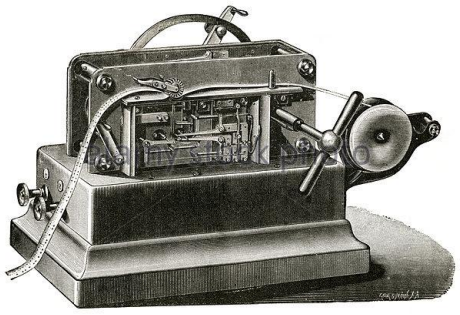
The International Telegraph Alphabet

● INDICATES A MARK ELEMENT (A HOLE PUNCHED IN THE TAPE)
○ INDICATES POSITION OF A SPROCKET HOLE IN THE TAPE

- Cinco lugares para poner hueco (1) o nada (0).
- Teclado similar al de máquina de escribir.

Emisor



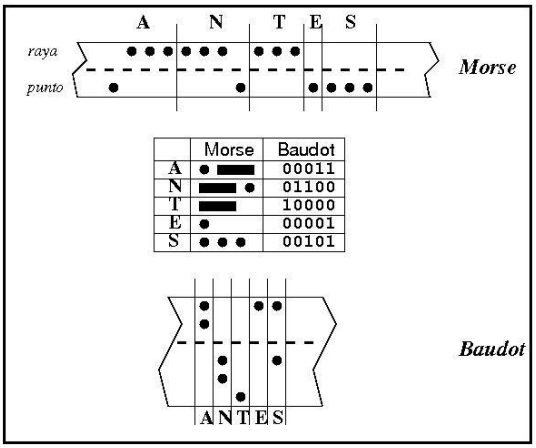


www.alamy.com - FGN480

- Teclea el texto a transmitir, la máquina perfora una cinta de papel con el código de cada letra tecleada, se genera la cinta del mensaje.
- La cinta del mensaje es llevada al aparato transmisor, que lee la cinta, la traduce en señales eléctricas y las envía por las líneas de transmisión

Receptor

- Un aparato receptor recibe las señales de la línea y las perfora en una cinta de papel.
- La cinta del mensaje es llevada al aparato impresor, que lee la cinta, traduce los códigos perforados en ella a caracteres y los imprime en una hoja de papel para meter el sobre y enviar.



Código de Baudot-Murray							
#	Código	Letters	Figures	#	Código	Letters	Figures
0	000·00	Null		16	100·00	T	5
1	000·01	E	3	17	100·01	Z	"
2	000·10	Line feed		18	100·10	L	)
3	000·11	A	-	19	100·11	W	2
4	001·00	Space		20	101·00	H	#
5	001·01	S	Bell	21	101·01	Y	6
6	001·10	I	8	22	101·10	P	0
7	001·11	U	7	23	101·11	Q	1
8	010·00	Carriage return		24	110·00	O	9
9	010·01	D	$	25	110·01	B	?
10	010·10	R	4	26	110·10	G	&
11	010·11	J	'	27	110·11	Figures (FS)	Letters (LS)
12	011·00	N	,	28	111·00	M	.
13	011·01	F	!	29	111·01	X	/
14	011·10	C	:	30	111·10	V	;
15	011·11	K	(	31	111·11	Letters (LS)	Figures (FS)

Cifrado en línea de Vernam

- 1917, en plena 1a guerra mundial se le encomienda a Razelmond Parker de AT&T que elabore un sistema seguro para transmisión telegráfica.
- Parker asigna el trabajo a Gilbert S. Vernam.
- El problema consiste en encontrar una función de conmutación capaz de cifrar y que sea invertible.

Ejemplos

Mensaje	Clave	AND
0	0	0
0	1	0
1	0	0
1	1	1

Si recibimos como mensaje cifrado un 0 y sabemos que la clave tiene un 0 en ese lugar ¿cuál es el bit de texto claro?

Sin ambigüedad

Mensaje	Clave	XOR
0	0	0
0	1	1
1	0	1
1	1	0

¿A qué es equivalente?

Msj→ Cla↓	00	01	10	11
00	00	01	10	11
01	01	00	11	10
10	10	11	00	01
11	11	10	01	00

Es un cifrado polialfabético con un alfabeto mezclado arbitrario distinto en cada renglón.

Realización

- Perforar una cinta con la clave.
- Construir un dispositivo (U.S. Patent 1,310,719; presentada 1918, concedida 1919) que tomara la cinta del mensaje y la de la clave y aplicara el xor con ambos argumentos caracter por caracter. *Mixer*.
- Cifrar y descifrar es aplicar la misma operación con la misma clave.
- Sólo hay que garantizar que emisor y receptor tengan sincronizada la clave.

Aportaciones de otros

- Clave larga (Lyman Morehouse): combinar cintas en banda cerrada.
  - Una larga cifrada (821 cars) por otra corta (150 cars).
  - Ciclo del tamaño de 821*150 = 123,150.
  - Suficiente para poco más de un día de tráfico normal (100 mil cars).
  - Deben ser primos relativos para que el ciclo (mcm) sea máximo.

C1	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0
L1	0	0	1	1	0	1	0	0	1	1	0	1	0	0	1	1	0	1	0	0	1	1	0	1
C1 XOR L1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1
C2	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0
L2	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0	0	0	1	1	0
C2 XOR L2	0	1	1	1	1	0	0	0	1	0	1	0	1	0	0	0	1	1	1	1	0	0	0	1	0

- Aleatoria además de larga (Joseph Mauborgne): “endless and senseless”.

Ejemplo

Si cualquier secuencia de clave es válida, cualquier secuencia de criptograma es posible.

Clave	C					L					A					V					E				
Código clave	0	1	1	1	0	1	0	0	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	1
Mensaje	a					n					t					e					s				
Código mensaje	0	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1
Criptograma	F					V					W					inv Shift					Space				
Código criptograma	0	1	1	0	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	0	1	0	0

Criptograma	F					V					W					inv Shift					Space				
Código criptograma	0	1	1	0	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	0	1	0	0
Clave posible	inv Shift					B					L					S					M				
Código clave	1	1	1	1	1	1	1	0	0	1	1	0	0	1	0	0	0	1	0	1	1	1	1	0	0
Texto claro posible	l					u					e					g					o				
Código texto claro	1	0	0	1	0	0	0	1	1	1	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0

Coincidencias

El XOR de dos textos cifrados con la misma clave coincide con el de sus criptogramas. Se puede recuperar el índice de coincidencias. Hay que usar la clave sólo una vez

Clave	C					L					A					V					E				
Código clave	0	1	1	1	0	1	0	0	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	1
M1	a					n					t					e					s				
Código M1	0	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1
C1	F					V					W					inv Shift					Space				
Código C1	0	1	1	0	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	0	1	0	0
M2	t					o					d					o					s				
Código M2	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	0	1
C2	V					R					R					I					Space				
Código C2	1	1	1	1	0	0	1	0	1	0	0	1	0	1	0	0	0	1	1	0	0	0	1	0	0
M1 XOR M2	1	0	0	1	1	1	0	1	0	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0
C1 XOR C2	1	0	0	1	1	1	0	1	0	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0

¿Qué tan seguro es el sistema?

- $\mathcal{M}$. Conjunto de todos los posibles mensajes claros.
- $\mathcal{C}$. Conjunto de todos los posibles criptogramas.
- $\mathcal{K}$. Conjunto de todas las posibles claves.
- f. Función de cifrado.
- Si fijamos una $k \in \mathcal{K}$, sabemos que se debe poder descifrar. Es decir, cada criptograma $c \in \mathcal{C}$ proviene de, a lo más, un texto claro $m \in \mathcal{M}$.
  - Hay al menos tantos criptogramas como textos claros: $|\mathcal{M}| \leq |\mathcal{C}|$.

Ejemplo 1.

- Mensajes: el resultado de tirar un dado menos uno: $\mathcal{M}$ = {0, 1, 2, 3, 4, 5}.
- Función de cifrado: sumar una cantidad k al resultado, módulo 6.
- Criptogramas: $\mathcal{C}$ = {0, 1, 2, 3, 4, 5}.
- Cifrar: c = (m+k) (mod 6). Descifrar: m = (c-k) (mod 6).

Probabilidades

- Probabilidad a posteriori (las del criptoanalista) P(se haya escogido m | se observó c) = P(m | c).
- No sabemos nada P(m | c) = P(m).

Ejemplo 2.

- Mensajes: texto en español de 1995 letras con alfabeto de 26 letras.
- Función de cifrado: sumar una cantidad de posiciones k a cada letra del alfabeto módulo 26.
- Criptogramas: textos con alfabeto de 26 letras de tamaño 1995.
- Criptoanálisis:
  - ¿Cuantos textos en español miden 1995 letras? **P(m) ≈ 0**.
  - ¿Cuantos textos en español tienen la misma letra en las posiciones en las que el criptograma tiene, digamos… “H” (la más frecuente)?
  - ¿Cuantos textos en español tienen a la “e” en las posiciones en las que el criptograma tiene “H”?
  - ¿Cuantos que al mismo tiempo tengan “a” donde el criptograma tiene “D”?
- Cada vez reducimos más el espacio de búsqueda.
- **P(m | c) ≫ P(m)**. Es lo que permite el criptoanálisis.

Seguridad perfecta

Un sistema simétrico $\mathcal{S} = (\mathcal{M}, \mathcal{C}, \mathcal{K})$ es perfectamente seguro si, para todo criptograma $c \in \mathcal{C}$ se tiene que:

$$P(m \mid c) = P(m)$$

para cualquier texto claro $m \in \mathcal{M}$.

Condición necesaria

Si S es un sistema criptográfico perfectamente seguro entonces cualquier mensaje claro m puede ser mapeado a cualquier criptograma c usando alguna clave k.

i.e. hay al menos tantas claves como criptogramas

$$|\mathcal{M}| \leq |\mathcal{C}| \leq |\mathcal{K}|$$

Condición suficiente

- Todas las claves son equiprobables.
- Para toda pareja (c, m), existe una k tal que m se cifra como c usando k.
- las cardinalidades de los conjuntos de claves, mensajes y criptogramas son iguales.