Resumen Teórica 4 : Criptografía

Tomás F. Melli

August 2025

Índice

1 Definiciones

1.1 Criptografía

Rama de las matemáticas y de la informática que se ocupa de cifrar/descifrar información utilizando métodos y técnicas que permitan el intercambio de mensajes de manera que sólo puedan ser leídos por las personas a quienes van dirigidos.
El objetivo de la criptografía es mantener la información cifrada secreta.

1.2 Criptoanálisis

Estudio de los métodos que se utilizan para quebrar textos cifrados con objeto de recuperar la información original en ausencia de la clave.

1.3 Criptología

Ciencia que estudia las técnicas criptográficas y de criptoanálisis.

1.4 Cifra

Método o técnica que protege a un mensaje al aplicar un algoritmo criptográfico. Sin conocer una clave específica, no será posible descifrarlo o recuperarlo.

1.5 Esteganografía

Es la comunicación secreta lograda mediante la ocultación de la existencia de un mensaje.

1.6 Criptosistema

Un **criptosistema** es un modelo matemático formal que describe el proceso de cifrado y descifrado de información. Se define como una **tupla**:

$$(E, D, M, K, C)$$

donde:

- M: conjunto de **mensajes en claro** (plaintexts).
- K: conjunto de **claves**.
- C: conjunto de **textos cifrados** (ciphertexts).
- E: conjunto de **funciones de cifrado**. Cada función está definida como:

  $$e_k : M \to C \quad \text{con } e_k(m) = c$$

  Formalmente:

  $$E : M \times K \to C$$

- D: conjunto de **funciones de descifrado**. Cada función está definida como:

  $$d_k : C \to M \quad \text{con } d_k(c) = m$$

  Formalmente:

  $$D : C \times K \to M$$

1.7 Atacante

Alguien cuya meta es quebrar un criptosistema. Se asume que conoce el algoritmo pero no la clave.

1.8 Tipos de ataques

1. Fuerza bruta

- El atacante prueba todas las claves posibles.
- Su meta es encontrar el texto en claro y, eventualmente, la clave.

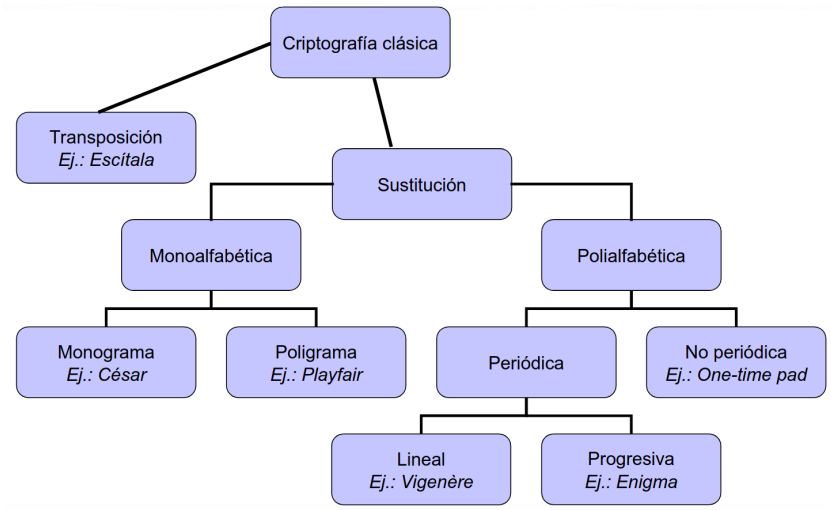
2. **Ataque de solo texto cifrado**
   - El atacante conoce únicamente el texto cifrado.
   - Su meta es deducir el texto en claro y, posiblemente, la clave.
3. **Ataque de texto en claro conocido**
   - El atacante conoce pares de texto en claro y su correspondiente texto cifrado.
   - Su meta es encontrar la clave utilizada.
4. **Ataque de texto en claro elegido**
   - El atacante puede elegir ciertos textos en claro y obtener sus versiones cifradas.
   - Su meta es descubrir la clave.
5. **Ataques matemáticos**
   - Basados en el análisis matemático del algoritmo de cifrado.
6. **Ataques estadísticos**
   - Basados en la distribución de símbolos:
     - Monogramas (letras individuales).
     - Digramas (pares de letras).
     - Trigramas (ternas de letras).
   - El atacante examina el texto cifrado y lo relaciona con estas propiedades estadísticas para intentar romper el cifrado.

2 Criptogafía clásica

La criptografía clásica abarca los métodos de cifrado anteriores a la computación moderna, es decir, técnicas manuales o mecánicas usadas durante siglos para proteger mensajes. Se basan en transformar un texto legible (texto claro) en uno ilegible (texto cifrado) mediante el uso de una clave compartida.
Las características principales son :

- Clave compartida : Tanto el emisor como el receptor poseen la misma clave, o bien una que pueda derivarse fácilmente. Ejemplo: si el emisor cifra con un desplazamiento de +3, el receptor descifra restando -3.
- Criptografía simétrica, porque la misma clave (o una equivalente) se usa para cifrar y descifrar. El desafío principal es la distribución de claves: ¿cómo se entrega la clave de forma segura antes de comunicarse?

El siguiente cuadro esquematiza los tipos :



2.1 Transposición

En los cifrados por transposición, los caracteres originales del texto claro no se sustituyen por otros, sino que se reordenan siguiendo un patrón definido por una clave. Esto implica que la frecuencia de letras no cambia (algo distinto a los cifrados por sustitución). La seguridad se basa en lo complicado que resulte reconstruir el orden correcto sin conocer la pauta.
Texto en plano:

`HELLO WORLD`

Eliminar espacios:

`HELLOWORLD`

Posiciones y caracteres:

posición:	1	2	3	4	5	6	7	8	9	10
carácter:	*H*	*E*	*L*	*L*	*O*	*W*	*O*	*R*	*L*	*D*

Tomamos las posiciones impares (1,3,5,7,9) y las pares (2,4,6,8,10):

$\underbrace{\texttt{H L O O L}}_{\text{impares}}$ $\underbrace{\texttt{E L W R D}}_{\text{pares}}$

Escribiendo en dos filas (alternando) queda:

```
H  L  O  O  L
E  L  W  R  D
```

Texto cifrado (fila 1 seguida de fila 2):

`HLOOL ELWRD`

Si recibimos el texto cifrado:

`HLOOL ELWRD`

y sabemos que se utilizó este método (impares/pares) y que la longitud es par, procedemos así:

1. Separar en dos bloques de igual longitud:

   $s_1 = \texttt{HLOOL}, \qquad s_2 = \texttt{ELWRD}$

2. Reconstruir alternando caracteres: tomar el primero de s_1, luego el primero de s_2, después el segundo de s_1, el segundo de s_2, y así sucesivamente:

   $H\ E\ L\ L\ O\ W\ O\ R\ L\ D$

3. Unir los caracteres obtenidos:

   `HELLOWORLD`

4. Insertar espacios si corresponde, recuperando el mensaje original:

   `HELLO WORLD`

2.1.1 La cifra de la Escítala

Utilizada en Grecia en el Siglo V a.c. La escítala es una vara de madera sobre la que se enrosca una tira de cuero. El emisor escribe el mensaje a lo largo de la vara y luego desenrosca la tira. La tira solía ocultarse como cinturón. Para recuperar el mensaje el receptor necesita una vara igual a la usada para crearlo. Texto en claro:

`sendmoretroopstosouthernflankand`

Longitud $n = 32$. Si usamos 8 columnas, el número de filas es $n/8 = 4$. Escribimos el texto por filas en una matriz 4×8:

s	*e*	*n*	*d*	*m*	*o*	*r*	*e*
t	*r*	*o*	*o*	*p*	*s*	*t*	*o*
s	*o*	*u*	*t*	*h*	*e*	*r*	*n*
f	*l*	*a*	*n*	*k*	*a*	*n*	*d*

Al desenrollar la tira (leer columna por columna, de izquierda a derecha, y dentro de cada columna de arriba abajo) se obtiene el texto cifrado:

`STSFEROLNOUADOTNMPHKOSEARTRNEOND`

2.2 Sustitución

En un cifrado por sustitución, cada carácter (normalmente cada letra) del texto en claro se reemplaza por otro carácter según una permutación del alfabeto. Si el alfabeto es de 26 letras, hay 26! posibles sustituciones. Las letras cambian, pero la posición relativa se conserva.
Un ejemplo de este tipo de cifrado es la **Cifra del César**

2.2.1 Cifra del César

Utilizada por Julio César para comunicarse con sus tropas. Miremos el ejemplo.
Texto en plano:

```
HELLO WORLD
```

Aplicando la regla: cada letra se reemplaza por la tercera letra que le sigue en el alfabeto ($A \mapsto D, B \mapsto E, C \mapsto F, \ldots$):

Plano:	H	E	L	L	O	W	O	R	L	D
Cifrado:	K	H	O	O	R	Z	R	U	O	G

Decimos

- $M = \{\text{secuencia de letras del alfabeto}\}$, el conjunto de mensajes posibles.
- $K = \{i \mid i \in Z,\ 0 \leq i \leq 26\}$, el conjunto de claves posibles.
- $E = \{E_k \mid k \in K\}$, el conjunto de funciones de cifrado definidas por:

  $$E_k(m) = (m + k) \bmod 27, \quad \forall m \in M$$

- $D = \{D_k \mid k \in K\}$, el conjunto de funciones de descifrado definidas por:

  $$D_k(c) = (c - k + 27) \bmod 27, \quad \forall c \in C$$

- $C = M$, el conjunto de mensajes cifrados.

Tenemos dos tipos de sustitución

2.2.2 Sustitución monoalfabética

Cada letra del mensaje se reemplaza siempre por la misma letra en el texto cifrado. Es la forma más simple de sustitución. Se representa con una permutación fija del alfabeto. Por tanto, es vulnerable a análisis de frecuencia porque las letras más frecuentes permanecen reconocibles.
Un ejemplo de esta es el cifrado del césar como vimos antes, es de tipo monoalfabética monograma.

2.2.3 Sustitución polialfabética

A la misma letra del mensaje original le pueden corresponder distintas letras en el texto cifrado, dependiendo de su posición o de una clave. Se usan varias permutaciones del alfabeto. El objetivo es dificultar el **análisis de frecuencia**, ya que la frecuencia de una letra se distribuye entre varias letras cifradas. Un ejemplo es **Playfair**: Matriz 5x5 con clave "PLAYFAIR"

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

Preparar mensaje

- Mensaje: `HELLO WORLD`
- Sin espacios: `HELLOWORLD`
- Dividido en digramas: `HE LL OW OR LD`
- Ajuste de pares repetidos: `HE LX LO WO RL DX`

Aplicar reglas de Playfair

Digrama	Cifrado
HE	KM
LX	FX
LO	YQ
WO	QS
RL	PR
DX	DZ

Texto cifrado final

`KMFXYQQSPRDZ`

Y de qué se trata esto del análisis de frencuencia ?

Análisis de frecuencia

El análisis de frecuencia es una técnica de criptoanálisis que estudia la frecuencia de aparición de los símbolos en un texto cifrado (criptograma) con el objetivo de descubrir patrones y deducir el mensaje original. Se basa en el análisis de la frecuencia de aparición de los símbolos del texto cifrado y su intento de correlación con los símbolos del lenguaje en el cual está escrito el mensaje. Se buscan los caracteres más frecuentes en el criptograma y se los asocia a las letras de mayor aparición en el idioma original. Se prueban distintas alternativas hasta alcanzar un texto coherente. Es una herramienta criptoanalítica de base. Miremos la tabla de frecuencias del español :

a	10.60	j	0.25	r	0.74
b	1.16	k	0.11	s	8.47
c	4.85	l	4.42	t	5,40
d	5.87	m	3.11	u	4,34
e	13.11	n	7.14	v	0.82
f	1.13	ñ	0.10	w	0.12
g	1.40	o	8.23	x	0.15
h	0.60	p	2.71	y	0.79
i	7.16	q	0.74	z	0.26

Otro ejemplo de cifrado polialfabético periódico lineal es el **cifrado Vingnere**. Se dice **periódico** ya que la clave se repite periódicamente sobre el mensaje, el período es igual a la clave. Decimos que es **lineal** ya que el cifrado se aplica letra por letra, en orden, siguiendo la secuencia lineal del mensaje. Esto contrasta con cifrados **no periódicos** (como el One-Time Pad), donde cada letra usa un alfabeto distinto, no repetitivo, eliminando patrones.
Veamos un ejemplo del cifrado de Vignere :

- Texto en plano: `THE BOY HAS THE BALL`
- Clave: `VIG`

1. Eliminamos espacios: `THEBOYHASTHEBALL`
2. Repetimos la clave para que coincida con la longitud del mensaje:

   `VIGVIGVIGVIGVIGVIGV`

Cifrado paso a paso :

1. Cada letra del texto plano se cifra usando un desplazamiento tipo César según la letra correspondiente de la clave.
2. Regla:

   $$C_i = (P_i + K_i) \bmod 26$$

   donde P_i y K_i representan la posición y el valor numérico de la letra (A=0, B=1, ..., Z=25).

Posición	Plano	Clave	Cifrado
1	T	V	O
2	H	I	P
3	E	G	K
4	B	V	W
5	O	I	W
6	Y	G	E
7	H	V	C
8	A	I	I
9	S	G	Y
10	T	V	O
11	H	I	P
12	E	G	K
13	B	V	W
14	A	I	I
15	L	G	R
16	L	V	G

Texto cifrado final

`OPKWWECIYOPKWIRG`

El **tableau** es una herramienta visual utilizada en cifrados polialfabéticos periódicos, especialmente en Vigenère.

- **Columnas:** letras de la clave.
- **Filas:** letras del texto en claro.
- Cada celda contiene la letra cifrada que resulta de cruzar la letra del mensaje con la letra de la clave.

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

Para cifrar una letra:

1. Tomar la letra del mensaje plano y la letra de la clave.
2. Seguir la columna correspondiente a la letra de la clave hasta la fila de la letra del mensaje.
3. La celda donde se cruzan da la letra cifrada.

- Clave: V, letra del mensaje: T $\Rightarrow$ letra cifrada: O
- Clave: I, letra del mensaje: H $\Rightarrow$ letra cifrada: P

3 Criptoanálisis

3.1 Ataque de Kasiski(1863)

a cifra de Vigenère puede ser atacada con éxito usando un análisis de frecuencia, gracias a la periodicidad de la clave. La técnica de Kasiski aprovecha repeticiones en el texto cifrado que se producen cuando los mismos caracteres de la clave se alinean sobre los mismos caracteres del mensaje en claro. Estas repeticiones permiten estimar la longitud de la clave. La técnica consiste en :

- Buscar cadenas repetidas en el criptograma: Se identifican secuencias de letras que se repiten en distintos lugares del texto cifrado.
- Calcular el período de la clave: Para cada repetición, se anotan las posiciones donde aparece. Se obtiene el máximo común divisor (MCD) de estas distancias. Este MCD es un candidato probable para el longitud del período de la clave.
- Dividir el criptograma en N sistemas monoalfabéticos: N = longitud de la clave estimada. Cada sistema contiene las letras cifradas que fueron cifradas con la misma letra de la clave.
- Analizar cada sistema por separado: Cada uno se trata como un cifrado monoalfabético clásico. Se aplica análisis de frecuencia para deducir la letra correspondiente de la clave.

Las repeticiones no son al azar, sino que ocurren porque la clave se repite periódicamente. Esto es lo que hace posible la reducción del problema polialfabético a varios problemas monoalfabéticos, que son mucho más fáciles de atacar.
En un cifrado Vigenère, el mensaje se cifra letra por letra usando la clave repetida periódicamente. Cada letra del mensaje se desplaza según la letra correspondiente de la clave. Cuando la clave se alinea nuevamente con las mismas letras del mensaje, **puede generar repeticiones** en el texto cifrado, especialmente si las letras del mensaje se repiten. Veamos

- **Mensaje plano:** `THEBOYHASTHEBALL`
- **Clave repetida:** `VIGVIGVIGVIGVIGV`
- **Texto cifrado:** `OPKWWECIYOPKWIRG`

1. **Identificar repeticiones:** Buscamos secuencias de letras que se repitan en el criptograma. En este ejemplo, la secuencia `OPK` se repite.

Posición	Cifrado	Comentario
1–3	OPK	Primera aparición
10–12	OPK	Segunda aparición

2. **Calcular la distancia:** La distancia entre repeticiones es:

   $$\text{Distancia} = 10 - 1 = 9$$

3. **Determinar posibles períodos:** El período de la clave debe ser un **factor de la distancia**, ya que la repetición ocurre cuando la clave se alinea nuevamente con las mismas letras del mensaje plano.

   `Factores de 9:` $1, 3, 9$

   Por lo tanto, posibles longitudes de la clave: 1, 3 o 9.

Queremos quebrar el siguiente texto cifrado

ADQYS	MIUSB	OXKKT	MIBHK	IZOOO
EQOOG	IFBAG	KAUMF	VVTAA	CIDTW
MOCIO	EQOOG	BMBFV	ZGGWP	CIEKQ
HSNEW	VECNE	DLAAV	RWKXS	VNSVP
HCEUT	QOIOF	MEGJS	WTPCH	AJMOC
HIUIX				

Donde la tabla de repeticiones es..

Letras	*Inicio*	*Fin*	*Distancia*	*Factores*
MI	5	15	10	2, 5
OO	22	27	5	5
OEQOOG	24	54	30	2, 3, 5
FV	39	63	24	2, 2, 2, 3
AA	43	87	44	2, 2, 11
MOC	50	122	72	2, 2, 2, 3, 3
QO	56	105	49	7, 7
PC	69	117	48	2, 2, 2, 2, 3
NE	77	83	6	2, 3
SV	94	97	3	3
CH	118	124	6	2, 3

Queremos estimar el período

1. Identificar una repetición significativa en el criptograma:
   - Secuencia repetida: `OEQOOG`
   - Es una repetición muy larga.
2. Estimar posibles períodos:
   - El período puede ser: 1, 2, 3, 5, 6, 10, 15, o 30.
3. Analizar los factores de todas las repeticiones:
   - 7 de 10 repeticiones tienen el factor 2.
   - 6 de 10 repeticiones tienen el factor 3.
4. Estimar el período de la clave:
   - Se elige el período como el producto de los factores más comunes:

     $$2 \times 3 = 6$$

   - Por lo tanto, la clave probablemente tiene **longitud 6**.

El índice de coincidencia

Es la probabilidad de que dos letras de un texto cifrado elegidas al azar sean la misma. Se define como

$$IC = \frac{1}{n(n-1)} \sum_{i=0}^{25} F_i(F_i - 1)$$

donde:

- n es la longitud del texto cifrado.
- F_i es la cantidad de veces que aparece la letra i en el texto.

Se encuentra tabulado el IC para distintos períodos:

Período	IC 1	IC 2	IC 3
1	0.066	30.047	50.044
2	0.052	40.045	100.041
>10	0.038	-	-

Obtenemos los alfabetos

Alfabeto	Letras	IC
1	AIKHOIATTOBGEEERNEOSAI	0.069
2	DUKKEFUAWEMGKWDWSUFWJU	0.078
3	QSTIQBMAMQBWQVLKVTMTMI	0.078
4	YBMZOAFCOOFPHEAXPQEPOX	0.056
5	SOIOOGVICOVCSVASHOGCC	0.124
6	MXBOGKVDIGZINNVVCIJHH	0.043

Los ICs indican que los alfabetos tienen **período 1**, excepto los alfabetos #4 y #6.

Análisis de frecuencias

Se muestran las frecuencias de cada letra en los distintos alfabetos:

Alfabeto	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
1	3	1	0	0	4	0	1	1	3	0	1	0	0	1	3	0	0	1	1	2	0	0	0	0	0	0
2	1	0	0	2	2	2	1	0	0	1	3	0	1	0	0	0	0	0	1	0	4	0	4	0	0	0
3	1	2	0	0	0	0	0	2	0	1	1	4	0	0	0	4	0	1	3	0	2	1	0	0	0	0
4	2	1	1	0	2	2	0	1	0	0	0	0	1	0	4	3	1	0	0	0	0	0	0	2	1	1
5	1	0	5	0	0	0	2	1	2	0	0	0	0	5	0	0	0	0	3	0	0	2	0	0	0	0
6	0	1	1	1	0	0	2	2	3	1	1	0	1	2	1	0	0	0	0	0	3	0	1	0	1	0

Frecuencias del idioma (H: high, M: medium, L: low):

Letra	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
Frec	H	M	M	M	H	M	M	H	H	M	M	M	M	H	H	M	L	H	H	H	M	L	L	L	L	L

Descifrado

Se analiza cada alfabeto para determinar su desplazamiento :

1. El primer alfabeto queda igual.
2. El tercer alfabeto tiene la letra `I` cambiada por `A`.
3. El sexto alfabeto tiene la letra `V` cambiada por `A`.
4. Se realizan las sustituciones sobre el texto cifrado. Las letras cifradas se indican en **negrita**:

ADI**Y**S **RIUKB** O**CKKL** MI**GH**K **A**ZO**TO** EI**OOL**
IF**TA**G **PAUEF** V**ATAS** CI**IT**W **EOCNO** EI**OOL**
BM**T**FV **EGGOP** C**NE**KI HS**SE**W **N**EC**SE** D**D**AA**A**
RWCXS **ANSNP** H**HEUL** QO**NOF** **EEGOS** W**L**PC**M**
AJ**E**OC **MIUAX**

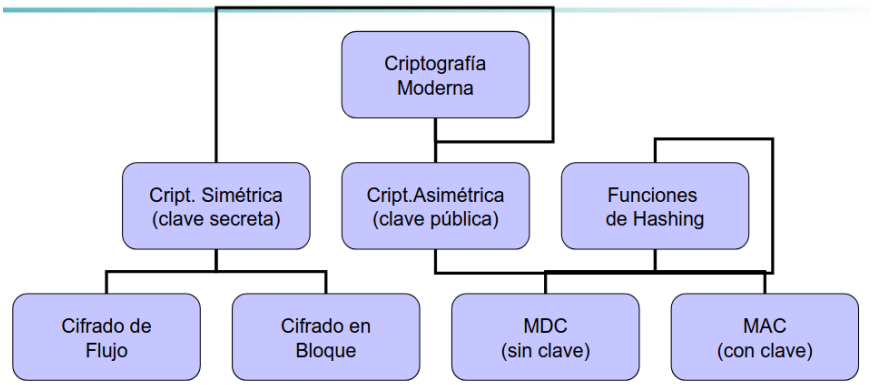
El análisis continua buscando en el texto pistas que sugieran palabras conocidas para de esta manera determinar los desplazamientos de los alfabetos restantes.
Luego de terminar el proceso tendremos que la clave: ASIMOV y el texto en claro:

ALIME	RICKP	ACKSL	AUGHS	ANATO	MICAL	INTOS
PACET	HATIS	QUITE	ECONO	MICAL	BUTTH	EGOOD
ONESI	VESEE	NSOSE	LDOMA	RECLE	ANAND	THECL
EANON	ESSOS	ELDOM	ARECO	MICAL		

One-Time Pad (cuaderno de uso único)

- Es una variante de Vigenère con una clave aleatoria tan larga como el mensaje en claro.
- Es un sistema criptográfico "perfectamente seguro". El texto cifrado `DXQR` puede corresponder al texto en claro `DOIT` (cifrado con la clave `AJIY`) y al texto en claro `DONT` (cifrado con la clave `AJDY`) y a cualquier otra combinación de 4 letras.
- Las claves deben ser aleatorias; de no serlo, se puede atacar tratando de regenerar la clave.
- Las claves se deben usar una sola vez.

4 Criptografía Moderna



4.1 Principios de Kerckhoffs (1883)

1. El sistema debe ser en la práctica imposible de criptoanalizar. Esto significa que, incluso si un atacante tiene acceso al texto cifrado y conoce el algoritmo, no debería poder descifrar el mensaje sin la clave. En la práctica, esto se logra usando claves suficientemente largas y aleatorias, o métodos criptográficos con seguridad demostrada.

2. La seguridad de un sistema criptográfico debe depender sólo de que la clave sea secreta y no de que el algoritmo de cifrado sea secreto.

3. Método de elección de claves fácil de recordar. Las claves deben ser memorables o fáciles de generar para los usuarios, sin comprometer la seguridad. Esto reduce errores humanos y evita que los usuarios reutilicen claves inseguras.

4. Transmisión del texto cifrado por telégrafo. En contextos históricos, el cifrado se aplicaba a sistemas de comunicación como el telégrafo. El diseño del sistema debía permitir que los mensajes cifrados se transmitieran sin problemas por los canales disponibles.

5. La máquina de cifrar debe ser portable. Especialmente en entornos militares o de campo, los dispositivos de cifrado debían ser transportables, ligeros y fáciles de usar sin depender de infraestructura pesada.

6. No debe existir una larga lista de reglas de uso. Sistemas complicados con muchas reglas y excepciones aumentan la probabilidad de errores humanos. Un buen sistema debe ser simple y directo, minimizando instrucciones complejas que puedan comprometer la seguridad.

4.2 Información

En 1948, Claude Shannon publicó el artículo *A Mathematical Theory of Communication*, sentando las bases de la Teoría de la Información. Donde postula que la información :

- Es el conjunto de datos o mensajes inteligibles creados con un lenguaje de representación.
- Ante varios mensajes posibles, aquel que tenga una menor probabilidad de aparición será el que contenga una mayor cantidad de información.

4.3 Seguridad perfecta

Un sistema criptográfico es perfectamente seguro si el texto cifrado no da ninguna información adicional del texto plano. Es decir que dado un texto cifrado C, cualquier posible texto plano es igualmente probable con respecto a C.

Teoremas

1. En todo sistema perfectamente seguro, la longitud de las claves es mayor o igual que la de los mensajes.

2. Existen sistemas perfectamente seguros (por ejemplo: *one-time pad*).

4.3.1 Sistema incondicionalmente seguro

Cuando es seguro frente a ataques con capacidad de cálculo ilimitada. Significa que el sistema de cifrado no puede ser quebrado aunque un atacante tenga recursos de cálculo ilimitados, tiempo infinito y acceso a todo el texto cifrado.

4.3.2 Sistema condicionalmente seguro

Cuando es seguro frente a ataques con capacidad de cálculo limitada. Significa que el sistema puede ser teóricamente quebrado, pero hacerlo requeriría una cantidad de tiempo o recursos computacionales que excede lo práctico.

4.4 Propuestas de Shannon para mejorar las operaciones de cifra

4.4.1 Difusión

Es la transformación del texto claro con el objeto de dispersar las propiedades estadísticas del lenguaje sobre el criptograma. Se logra con **Transposiciones**. La idea es que cada letra o símbolo del mensaje original afecte muchas partes del texto cifrado. Esto disminuye la posibilidad de usar patrones estadísticos del lenguaje para adivinar la clave o descifrar el mensaje.

4.4.2 Confusión

Es la transformación del texto claro con el objeto de mezclar los elementos de éste, aumentando la complejidad de la dependencia funcional entre la clave y el criptograma. Se obtiene mediante **Sustituciones**. Cada cambio en la clave debería producir cambios impredecibles en el texto cifrado, de manera que un atacante no pueda deducir la clave examinando pares de texto claro y texto cifrado.

4.5 Criptografía Simétrica

ambién conocidos como criptografía de clave secreta o clave privada. La clave utilizada en la operación de cifrado es la misma que se utilizada para el descifrado. Existen dos mecanismos de operación básicos: Sea E una función de cifrado:

- $E_k(b)$: cifrado del mensaje b con la clave k.
- $m = b_1 b_2 \ldots$, donde cada b_i es de longitud fija.

4.5.1 Cifrado de flujo

En el cifrado de flujo, el mensaje se cifra bit a bit o carácter a carácter en lugar de dividirlo en bloques. Se genera una secuencia de claves :

$$k = k_1 k_2 k_3 \ldots$$

(a veces llamada *keystream*) que se combina con el mensaje.
Supongamos que tenemos un mensaje

$$m = b_1 b_2 b_3 \ldots$$

Cada símbolo b_i se cifra usando la clave correspondiente k_i:

$$E_k(m) = E_{k_1}(b_1)\, E_{k_2}(b_2)\, E_{k_3}(b_3)\, \ldots$$

Esto permite cifrar mensajes de cualquier longitud de manera continua.

- Si la secuencia de claves $k_1 k_2 \ldots$ se repite, el cifrador se llama **periódico**.
- El período es la longitud del ciclo de la secuencia de claves.
- Los cifradores periódicos son más vulnerables a ataques de criptoanálisis porque los patrones se repiten.

Propiedades importantes :

- El espacio de las claves es igual o mayor que el espacio de los mensajes.
- Las claves son aleatorias.
- La secuencia de clave se usa sólo una vez.

Cifrador de flujo

Un cifrador de flujo necesita generar una secuencia de claves (*keystream*) para cifrar el mensaje bit a bit o carácter a carácter. En la práctica, se usan **generadores pseudoaleatorios** que producen esta secuencia a partir de una **semilla inicial**. Estos son :

- Son algoritmos determinísticos: si se conoce la semilla, siempre se produce la misma secuencia de claves.
- La semilla suele tener n bits, y el generador puede producir una secuencia muy larga con un período máximo de 2^n bits antes de repetirse.

Para transmitir la semilla :

- Para que el receptor pueda descifrar el mensaje, solo se necesita enviar la semilla, no toda la secuencia de claves.
- El receptor, usando el mismo generador pseudoaleatorio y la misma semilla, reconstruye la secuencia de claves idéntica a la que usó el emisor.

Modo de operación

Un modo de operación describe cómo aplicar el cifrado de flujo a un mensaje. Permite procesar el mensaje de manera continua, bit a bit o carácter a carácter.
Funcionamiento :

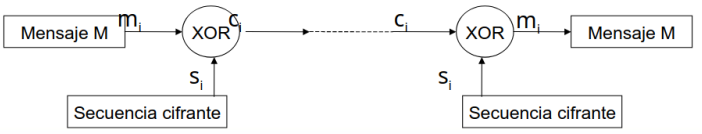
- El mensaje en claro se lee y transmite bit a bit.
- Cada bit del mensaje se combina con un bit de la secuencia cifrante usando una operación de cifra, normalmente XOR:

  $$c_i = m_i \oplus k_i$$

  donde m_i es el bit del mensaje, k_i es el bit de la secuencia cifrante y c_i es el bit cifrado resultante.

Este proceso se repite para todo el mensaje.
Requisitos de la secuencia cifrante :

- Debe tener un **período muy alto** para evitar repeticiones predecibles.
- Debe tener **propiedades pseudoaleatorias**, de manera que los patrones del mensaje no se reflejen en el criptograma.



Existen 2 mecanismos de operación de las secuencias cifrantes :

Sincrónicos

- Tanto el emisor como el receptor deben **sincronizarse previamente** antes de la transmisión.
- Cada bit del mensaje se combina con un bit de la secuencia cifrante según el orden exacto.
- **Problema:** si se pierde o altera un bit en el flujo de datos, puede inutilizar el resto de la transmisión, porque el receptor pierde la sincronización con la secuencia de claves.

Auto-Sincrónicos

- También llamados *autorrese Synchronizable Stream Ciphers*.
- La secuencia cifrante se renueva usando **parte del texto cifrado recibido**.
- **Ventaja:** si se pierde un bit, el receptor puede **recuperar la sincronización** después de un cierto número de bits.
- Esto hace que la transmisión sea más tolerante a errores en el canal de comunicación.

4.5.2 Cifrado en bloques

El cifrado en bloque consiste en dividir el mensaje en bloques de longitud fija y cifrar cada bloque por separado usando la misma clave.
Supongamos que tenemos un mensaje m que se divide en bloques $b_1, b_2, \ldots, b_n$.

- Cada bloque se cifra usando la función de cifrado E_k con la clave k:

  $$E_k(m) = E_k(b_1)\, E_k(b_2)\, \ldots\, E_k(b_n)$$

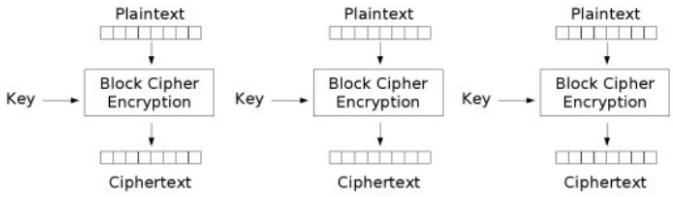
- El receptor, con la misma clave k, aplica la función de descifrado D_k a cada bloque para recuperar el mensaje original.

Modos de operación

Existen distintos modos de operación, que determinan cómo se combina la clave con el texto en claro y cómo se enlazan los bloques entre sí. Esto afecta la seguridad y la propagación de errores.

ECB (Electronic CodeBook)

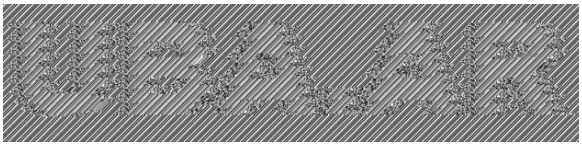
El texto se divide en bloques y cada bloque es cifrado en forma independiente utilizando la clave



Electronic Codebook (ECB) mode encryption

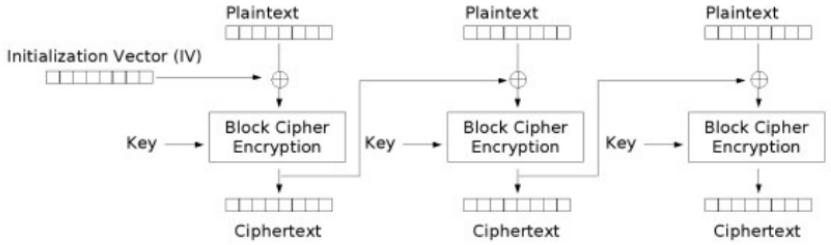
El problema es que si tengo una imagen en formato raw :

Y la cifro utilizando AES-128, con ECB obtenemos una nueva imagen raw pero puedo entender su contenido pese a estar encriptada :



CBC (Cipher Block Chaining)

El texto se divide en bloques y cada bloque es mezclado con la cifra del bloque previo, luego es cifrado utilizando la clave. Se usa un IV (vector de inicialización) para el primer bloque.

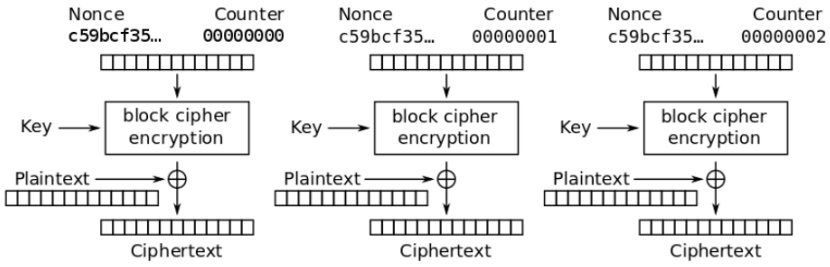


Cipher Block Chaining (CBC) mode encryption

CTR (Counter Mode)

CTR convierte un cifrador de bloques en un cifrador de flujo, usando un contador que se cifra y luego se combina con el texto plano mediante XOR.

- Nonce (Number used once) : un valor único para cada mensaje cifrado con la misma clave.
- Contador : un valor que cambia para cada bloque dentro del mismo mensaje.



Counter (CTR) mode encryption

4.6 Padding

Cuando usamos cifradores de bloques como EBC, CBC o CTR, el texto original se divide en bloques de tamaño fijo (por ejemplo, 8 bytes, 16 bytes, etc.). El problema es que muchas veces el último bloque no tiene la longitud exacta del bloque requerido. Los cifradores de bloques no pueden trabajar con bloques incompletos, así que necesitamos rellenar (padding) ese bloque para completarlo. Este texto de relleno debe ser quitado durante la operación de descifrado. Veamos :

- Relleno básico (bit 1 + ceros): Se agrega un 1 seguido de ceros hasta completar el bloque. Si el último bloque ya estaba completo, se agrega un bloque completo de relleno para poder eliminarlo luego de descifrar
- Padding según PKCS#5 / PKCS#7: Cada byte de relleno tiene el valor del número de bytes de padding. Ejemplo: Si faltan 6 bytes, cada byte de relleno vale 0x06. Si el texto ya completa exactamente un bloque, se agrega un bloque completo de padding (con valor igual al tamaño del bloque).

Primer Bloque								Segundo Bloque							
1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	0x01
A	A	A	A	A	A	A	A	B	B	B	B	B	B	0x02	0x02
...								...							
A	A	A	A	A	A	A	A	B	0x07	0x07	0x07	0x07	0x07	0x07	0x07
A	A	A	A	A	A	A	A	0x08	0x08	0x08	0x08	0x08	0x08	0x08	0x08