



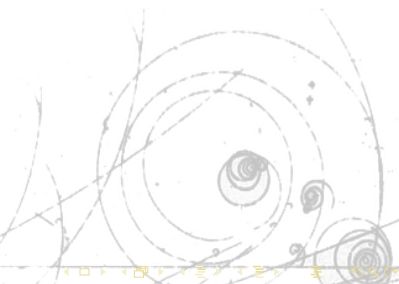
## Bloque 3. Conceptos básicos sobre Criptografía

José A. Montenegro

Dpto. Lenguajes y Ciencias de la Computación  
ETSI Informática. Universidad de Málaga  
monte@lcc.uma.es [twitter](#)

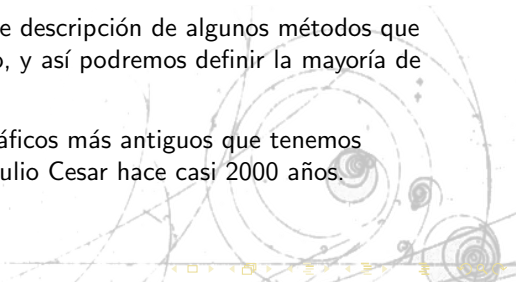
25 de noviembre de 2011

- 1 **Introducción a la Criptografía**
  - Análisis de Frecuencias
- 2 **El desarrollo de la criptografía**
  - Criptosistemas de Clave Simétrica
  - Cifrado Poli-alfabético
  - El sistema Playfair
  - Algoritmos matemáticos en criptografía
  - Métodos de ataque
- 3 **Criptografía desde el punto teórico y práctico**
  - Secreto Perfecto y One time pad
  - Métodos Iterativos
  - Estándar de Cifrado
  - El problema de distribución de claves
- 4 **Criptosistema de Clave Pública**
  - El Criptosistema RSA
    - Viabilidad de RSA
    - Fiabilidad de RSA
  - Certificados de Identidad
  - Esquema de Firma Digital y Verificación



# Criptografía

- Empezaremos discutiendo los aspectos tradicionales de la criptografía: el estudio de la codificación para el propósito de transmitir mensajes secretos escritos en un lenguaje natural.
- Posiblemente por su ya establecida historia, la criptografía ha desarrollado su propia terminología.
- Comenzaremos con una breve descripción de algunos métodos que han sido usados en el pasado, y así podremos definir la mayoría de la terminología.
- Uno de los sistemas criptográficos más antiguos que tenemos referencia fue utilizado por Julio Cesar hace casi 2000 años.



- El emisor elige un número  $k$  y reemplaza cada letra por una que está  $k$  lugares posteriores, en orden alfabético.
- Para simplificar la explicación asumiremos que el espacio  $\square$  y  $\tilde{n}$  no son modificadas.
- Por ejemplo, cuando  $k=5$  las letras son modificadas como sigue:

$\square$  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  
 $\square$  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

- Utilizando esta regla, el mensaje

MALAGA  $\square$  CAMPEON es RFQFLF  $\square$  HFRUJTS.



- En criptografía este proceso es conocido como *cifrado*, y el número  $k$  es la clave.
- En términos matemáticos, el cifrado es simplemente codificación, con la característica adicional que la función de codificación  $E_k$  depende del parámetro  $k$ .
- Por ejemplo, en nuestra versión del sistema Cesar  $E_k$  es la función  $\mathbb{A} \rightarrow \mathbb{A}$  definida mediante

$$E_k(x) = [x + k] \quad (x \neq \sqcup) , \quad E_k(\sqcup) = \sqcup,$$

donde  $[x + k]$  representa el símbolo que está  $k$  posiciones de  $x$ . La función  $E_k$  es extendida para que actúe sobre los mensajes (cadenas de símbolos) mediante concatenación:

$$E_k(DUDDA) = E_k(D)E_k(U)E_k(D)E_k(A).$$

Necesitamos que la función extendida  $E_k$  es una función inyectiva, por lo que la codificación es unívocamente decodificable. La situación general es como sigue:

### Definición 1 (Funciones Cifrado)

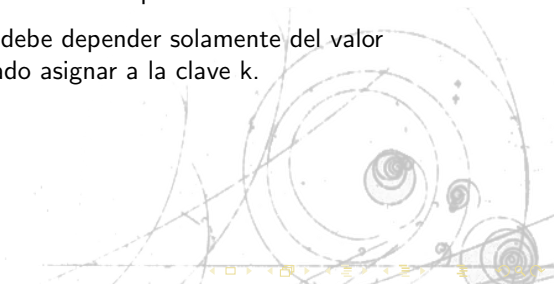
Sea  $\mathcal{M}$  y  $\mathcal{C}$  conjuntos de mensajes y sea  $\mathcal{K}$  un conjunto.

Supongase que para cada  $k \in \mathcal{K}$  hay una función inyectiva tal que  $E_k : \mathcal{M} \rightarrow \mathcal{C}$ .

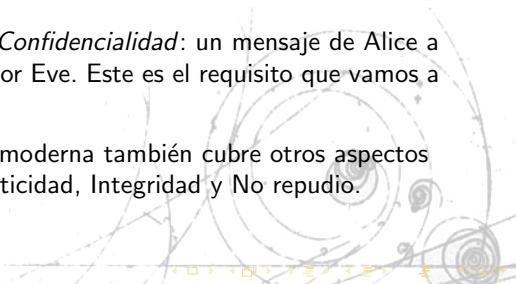
Entonces diremos que  $\{E_k\}$  es un conjunto de funciones de cifrados, y un elemento  $k \in \mathcal{K}$  es denominada una clave.

En el sistema de Cesar  $\mathcal{M}$  y  $\mathcal{C}$  son tomados del conjunto  $\mathbb{A}^*$  de cadenas de símbolos en  $\mathbb{A}$ , y  $\mathcal{K}$  es el conjunto de símbolos en el rango  $1 \leq k \leq 25$ .

- La principal lección aprendida en la historia de la criptografía es que es una mala idea ocultar el método de cifrado, la instancia de las funciones  $E_k$ .
- Si un número de personas quieren comunicarse de forma segura, no pueden basarse en ocultar el sistema que están usando.
- La seguridad del sistema debe depender solamente del valor específico que han acordado asignar a la clave  $k$ .



- Actualmente las partes involucradas en la comunicación reciben los nombres de Alice, Bob y Eve.
- Básicamente, Alice es el emisor del mensaje, Bob es el receptor, y Eve es un adversario que intenta intervenir de alguna forma en la comunicación.
- Dentro de este marco de trabajo podemos distinguir varios requisitos de seguridad.
- El requisito más obvio es la *Confidencialidad*: un mensaje de Alice a Bob no debe ser entendido por Eve. Este es el requisito que vamos a considerar inicialmente.
- Sin embargo, la criptografía moderna también cubre otros aspectos de la seguridad, como Autenticidad, Integridad y No repudio.





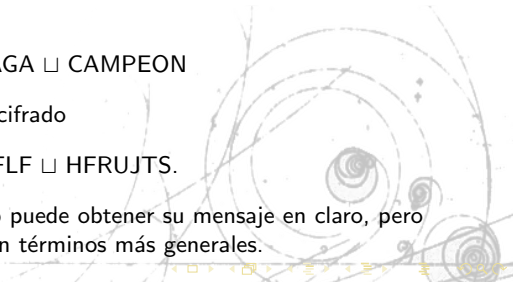
- Vamos asumir que Alice y Bob están de acuerdo con un método de cifrado y el valor de la clave  $k$ , por lo tanto ellos conocen la función de cifrado  $E_k$ .
- El mensaje que Alice quiere mandar a Bob es conocido como *texto en claro (plaintext)*, y por ahora podemos asumir que está expresado en un lenguaje natural, tal como el español.
- Alice utiliza la función  $E_k$  para transformar el texto en claro en una forma codificada, conocido como *texto cifrado (ciphertext)*, y lo envía a Bob.
- Por lo que si Alice y Bob están de acuerdo en utilizar el sistema Cesar con  $k=5$ , y el texto en claro es

MALAGA □ CAMPEON

entonces Bob recibirá el texto cifrado

RFQFLF □ HFRUJTS.

En este caso es fácil como Bob puede obtener su mensaje en claro, pero es útil considerar la situación en términos más generales.



- Suponemos que tenemos un conjunto de funciones de cifrado  $E_k : \mathcal{M} \rightarrow \mathcal{C}$ .
- Tal y como hemos establecido anteriormente cada  $E_k$  es una función inyectiva.
- Esto significa que si  $E_k(m) = c$ , entonces  $m$  es el único texto en claro que  $E_k$  cifra como  $c$ .
- Consecuentemente, existe una función  $F$ , definida como imagen de  $E_k$ , que toma  $c$  a  $m$ :

$$F(E_k(m)) = m \text{ para todo } m \in \mathcal{M}.$$

- Formalmente,  $F$  es la inversa de  $E_k$ . En la práctica no es suficiente con conocer que existe una inversa, es necesario tener una regla explícita para calcularla.

## Definición 2 (Funciones Descifrado)

Sea  $\{E_k\}$  ( $k \in \mathcal{K}$ ) sea un conjunto de funciones de cifrado  $\mathcal{M}$  y  $\mathcal{C}$ .

Entonces el conjunto de funciones  $\{D_l\}$  ( $l \in \mathcal{K}$ ) es un conjunto de funciones de descifrado para  $\{E_k\}$  si para cada  $k \in \mathcal{K}$  existe un  $l \in \mathcal{K}$  tal que  $D_l$  es la inversa de  $E_k$ :

$$D_l(E_k(m)) = m \text{ para todo } m \in \mathcal{M}$$

En el sistema Cesar, tomaremos  $l$  tal que  $l = -k(\text{mod}28)$  y la inversa de  $E_k$  es

$$D_l(x) = [x + l](x \neq \sqcup), D_l(\sqcup) = \sqcup.$$

- Ya que Bob conoce  $k$ , él también conoce  $l$ , y puede descifrar el mensaje de Alice.
- Notese que en este caso la función de descifrado  $D_l$  y la función de cifrado  $E_k$  tienen la misma forma, aunque no tiene que ser así de forma general.
- Ahora consideramos la situación cuando Eve intercepta el mensaje cifrado

RFQFLF □ HFRUJTS.

- Ella no puede simplemente aplicar una función de descifrado y recuperar el texto en claro, ya que ella no conoce cuál clave ha sido usada.
- Si ella quiere obtener el texto en claro, el método más fácil sería intentar encontrar una clave.

- El proceso para obtener el correspondiente texto en claro para algún texto cifrado, es encontrar el valor de la clave  $k$  o algún método indirecto, para romper el sistema.
- Cualquier método el cual Eve pueda llevar a cabo es conocido como un ataque.
- Hemos observado que Eve puede conocer que sistema se está utilizando, por tanto, si ella conoce la clave, entonces ella sabrá como usarla.
- Para el sistema Cesar existe un ataque simple por el método conocido como búsqueda exhaustiva.
- Debido a que solamente los posibles valores de  $k$  son  $1, 2, 3 \dots, 28$ , es fácil encontrar cual es el correcto.
- Asumiendo que el texto en claro es un mensaje expresado en un lenguaje natural, Eve conocerá cuando la clave correcta ha sido encontrada, ya que el mensaje tendrá sentido.

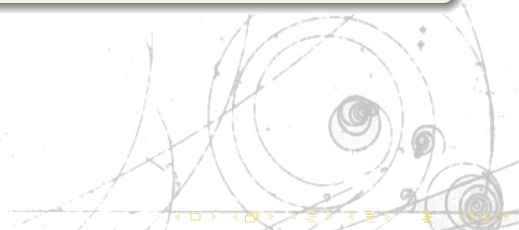


## Ejemplo 1

*Suponemos que Eve intercepta el mensaje cifrado*

*GJOBOTG □ WAK □ VUTK*

*¿Cual es el correspondiente mensaje en claro?*



Solución:

Cuando utilizamos la clave  $k=1,2,3, \dots$  en la parte inicial del texto cifrado, el resultado no tiene significado hasta que  $k=6$  es alcanzado.

$$k = 1 \text{ FINANSF} \sqcup \dots$$

$$k = 2 \text{ EHMZMRE} \sqcup \dots$$

$$k = 3 \text{ DGLYLQD} \sqcup \dots$$

$$k = 4 \text{ CFKXKPC} \sqcup \dots$$

$$k = 5 \text{ BEJWJOB} \sqcup \dots$$

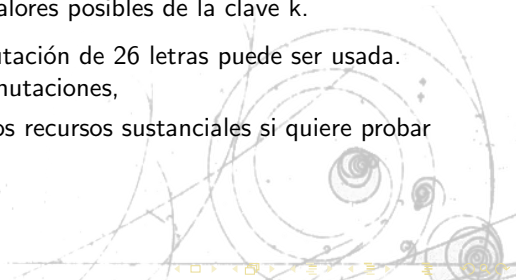
$$k = 6 \text{ ADIVINA} \sqcup \dots$$

Por lo que intentando con la clave  $k=6$  en el resto del mensaje, este produce una mensaje con significado y el texto en claro es

$$\text{ADIVINA} \sqcup \text{QUE} \sqcup \text{PONE}$$

# Análisis de Frecuencias

- ¿Como es posible que Alice y Bob se defiendan del ataque mediante una búsqueda exhaustiva?
- Una posible idea es incrementar el número de claves.
- El sistema Cesar utiliza la regla  $x \mapsto [x + k]$  para reemplazar las letras, y hay solamente 28 valores posibles de la clave  $k$ .
- Pero claramente cada permutación de 26 letras puede ser usada. Por lo que tenemos 28! permutaciones, por tanto Eve necesitará unos recursos sustanciales si quiere probar con todos ellos.





En criptografía un sistema que solamente permuta las letras del alfabeto en una forma determinada es conocido como sustitución *mono-alfabetica*.

La clave es una permutación  $\sigma$  de 28 letras y la función de cifrado puede ser definida como

$$E_{\sigma}(x) = \sigma(x) \quad (x \neq \sqcup), \quad E_{\sigma}(\sqcup) = \sqcup$$

La inversa de  $E_{\sigma}$  es la función  $D_{\tau}$  que utilizar la permutación inversa  $\tau = \sigma^{-1}$ :

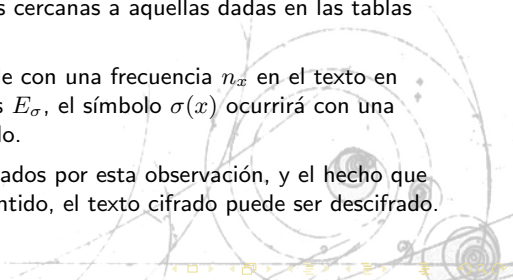
$$D_{\tau} = \tau(x) \quad (x \neq \sqcup), \quad D_{\tau}(\sqcup) = \sqcup$$

- Es conveniente utilizar una clave que puede ser memorizada, y esto es realizado escogiendo una palabra clave. Por ejemplo, si la palabra clave escogida es CAMPEON, la permutación correspondiente  $\sigma$  es definida como sigue:

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



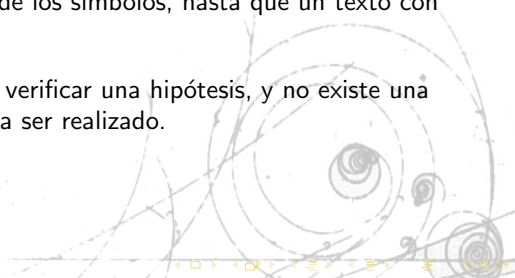
- Aunque el ataque por búsqueda exhaustiva necesita de muchos recursos, otro método de ataque, conocido como análisis de frecuencias, es usualmente más efectivo. Consta que este método fue utilizado por criptógrafos Arabes en el siglo 9.
- El método esta basado en el hecho que un mensaje en claro de un lenguaje natural tiene propiedades estadísticas. Además, es razonable asumir que el texto transporta un mensaje con sentido.
- En cualquier mensaje de longitud razonable, los símbolos, diagramas y demás, suceden con frecuencias cercanas a aquellas dadas en las tablas estándar de frecuencias.
- Por tanto si el símbolo  $x$  sucede con una frecuencia  $n_x$  en el texto en claro, y la función de cifrado es  $E_\sigma$ , el símbolo  $\sigma(x)$  ocurrirá con una frecuencia  $n_x$  en el texto cifrado.
- Utilizando los datos proporcionados por esta observación, y el hecho que el texto en claro tiene algún sentido, el texto cifrado puede ser descifrado.



En resumen, el ataque por análisis de frecuencia en una pieza de un texto cifrado producido por una sustitución monoalfabética es realizado como sigue:

- Contar las frecuencias de los símbolos en el texto cifrado;
- Comparar estas con las frecuencias estándar dadas en las tablas;
- Verificar la correspondencia de los símbolos, hasta que un texto con sentido es obtenido

El tercer paso involucra realizar y verificar una hipótesis, y no existe una regla definitiva para que eso pueda ser realizado.



## Criptosistemas de Clave Simétrica

- Anteriormente hemos descrito un marco de trabajo para la criptografía basado en un conjunto de mensajes en claro  $\mathcal{M}$ , un conjunto de mensajes cifrados  $\mathcal{C}$  y un conjunto de claves  $\mathcal{K}$ .
- Para cada  $k \in \mathcal{K}$  hay una función de cifrado  $E_k : \mathcal{M} \rightarrow \mathcal{C}$ , con su correspondiente inversa la función de descifrado  $D_l$ . En otras palabras

$$D_l(E_k(m)) = m \text{ para toda } m \in \mathcal{M}$$

- Nos referimos a este marco de trabajo como criptosistema.
- Los criptosistemas pueden ser construidos e implementados de muchas formas.



- En los sistemas de sustitución mono alfabéticos discutidos anteriormente, las claves correspondientes  $k$  y  $l$  son permutaciones inversas, y por tanto están vinculadas de una forma muy simple. Durante, muchos siglos se ha asumido que cualquier criptosistema práctico debe tener una propiedad similar.
- Ahora está claro que esa presunción es injustificada, pero sistemas con esta propiedad son todavía ampliamente utilizados, y le daremos a ellos un nombre.

### Definición 3 (Criptosistema Clave Simétrica)

*Supongamos que tenemos un criptosistema en el cual  $D_l$  es la inversa de  $E_k$ .*

*El sistema es denominado criptosistema de clave simétrica si, cualquiera de las dos  $k$  o  $l$  es conocida, entonces es fácil de conocer la otra.*

- Durante muchos años los criptógrafos trabajaron para hacer que los sistemas sean seguros ante un ataque de análisis.
- En un sistema mono alfabético, con una clave dada, la letra E (por ejemplo) siempre será reemplazada por la misma letra en el texto cifrado, por ejemplo X.
- El principal problema que presenta es que X será el carácter que más aparece debido a que ha reemplazado a una de las letras que más aparecía en el texto en claro.
- Esta debilidad del sistema puede eliminarse si utilizamos un sistema en el cual  $E$  no siempre reemplazado por la misma letra.



#### Definición 4 (Cifrado Poli alfabético)

*En un sistema poli alfabético la regla para el cifrado es que cada letra en el texto en claro es reemplazado por una letra que depende, no solamente por la letra en sí mismo, también por su posición en el texto.*

- Un método simple de cifrado poli alfabético fue desarrollado en el siglo 16, y es conocido como el *Sistema Vigenere*.
- Básicamente utiliza las permutaciones realizadas en el Sistema Cesar, que son desplazamiento cíclicos de tamaño  $(1 \leq k \leq 28)$ , denotado en el capítulo anterior como  $E_k(x) = [x + k]$ .
- Como en los casos anteriores, eliminaremos los espacios, resultando un texto en claro con todas las letras juntas.



- En el *Sistema Vigenere* la clave es una secuencia  $K = (k_1, k_2, \dots, k_m)$  de números. Si el texto en claro es  $x_1x_2x_3 \dots x_i \dots$ , la regla para el cifrado es

$$E_K(x_i) = [x_i + k_j] \text{ si } i \equiv j \pmod{m}.$$

- Por ejemplo, si la clave tiene longitud  $m=4$ , entonces  $x_1, x_5, x_9 \dots$  son cifrados con un desplazamiento de  $k_1$ ;  $x_2, x_6, x_{10} \dots$  son cifrados con un desplazamiento de  $k_2$  y así sucesivamente.
- Esta descripción representa un sistema de clave simétrica, ya que el descifrado es realizado mediante la inversa de los desplazamientos.
- En la práctica la clave  $K$  es usualmente expresado como una palabra, donde las letras de la palabra representan los desplazamientos relevantes.



## Ejemplo 2

Tenemos la palabra clave *CLAVE*, que representa la clave

$$K = \{3, 12, 1, 22, 5\}$$

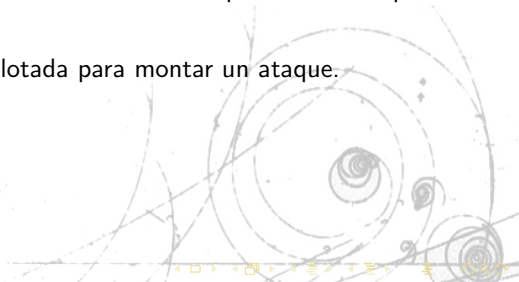
¿Cual será el cifrado del siguiente texto?

*VOTARCAMBIAALGO*

Solución:

V	O	T	A	R	C	A	M	B	I	A	A	L	G	O
3	12	1	22	5	3	12	1	22	5	3	12	1	22	5
Y	A	U	W	W	F	M	N	X	N	D	M	M	C	T

- Durante varios siglos el sistema Vigenere fue concebido como irrompible, y fue denominado como el *cifrado indescifrable*.
- Sin embargo, el sistema posee una vulnerabilidad inherente, debido a que la elección de los desplazamientos deben de estar basados en una regla definitiva, ya que de otra forma el receptor no será capaz de descifrarlo.
- Esta debilidad puede ser explotada para montar un ataque.



- Aunque se pensaba que era irrompible, el Sistema Vigenere era difícil de utilizar. Por ese motivo existía un considerable interés en métodos alternativos.
- Uno de esos métodos era el sistema Playfair, en el cual los diagramas son permutados, en vez de las letras individuales.
- Una clave para el sistema es derivada de una matriz 5x5 que contiene 25 letras (Para tener 25 letras, podemos identificar la Y como W).
- Si no permitimos diagramas con letras dobles de la forma  $LL$ , el número de diagramas posibles es  $25 \times 24 = 600$ , dando la posibilidad de unas 600! claves, por lo cual una búsqueda exhaustiva de la clave es una tarea complicada.



- El orden de las letras en la matriz puede ser aleatorio, o puede ser definida por una palabra clave.
- Figura 1 muestra la matriz basada en la palabra clave *PERSONAL*.
- De esta forma tanto receptor como emisor pueden recordar la clave, y siguiendo unas simples reglas, los procedimientos de cifrado y descifrado son directos.

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

Figura 1: La matriz Playfair con la palabra *PERSONAL*

En el proceso de cifrado los espacios son ignorados, y las letras dobles son separadas mediante una letra comodín tal como  $X$  o  $Z$ . El texto en claro es dividido en diagramas, cada diagrama  $xy$  es cifrado acorde con las siguientes reglas.

- Caso 1** Suponemos que  $x$  e  $y$  están en distintas filas y distintas columnas. Entonces  $xy \mapsto ab$  donde  $x,y,a,b$  son las esquinas de un rectángulo, y  $x,a$  están en la misma fila.
- Caso 2** Suponemos que  $x,y$  están en la misma fila. Entonces  $xy \mapsto uv$ , donde  $u$  y  $v$  son las letras que están a la derecha de  $x$  e  $y$  respectivamente. (Si alguno de  $x$  o  $y$  están al final de la fila, la letra al inicio de la fila es la elegida).
- Caso 3** Suponemos  $x, y$  están en la misma columna. Entonces  $xy \mapsto uv$ , donde  $u$  y  $v$  son las letras inmediatamente debajo de  $x$  e  $y$  respectivamente. (Si alguno de los dos ( $x$  o  $y$ ) están al final de la columna, la letra al inicio de la columna es utilizada).



- Por ejemplo,

$$QA \mapsto KB, FW \mapsto GV, DI \mapsto FD, SH \mapsto BQ$$

- El sistema Playfair es un sistema de clave simétrica.
- La clave de cifrado y descifrado son permutaciones del conjunto de 600 diagramas posibles, Alice utiliza las reglas dadas anteriormente para construir la clave para cifrado y la matriz dada, y Bob utiliza la misma matriz, pero con reglas ligeramente diferentes para construir las claves para el descifrado.
- Las dos claves son permutaciones inversas.



### Ejemplo 3

*Supongamos que recibimos el siguiente texto cifrado, conociendo que el sistema Playfair ha sido utilizado y la palabra clave es PERSONAL. ¿Cual era el texto en claro?*

*DCKOSREBJCARLNB LORISXZ*

Solución:

El texto cifrado es dividido en una secuencia de diagramas como:

*DC KO SR EB JC AR LN BL OR IS XZ*

Utilizando las reglas de forma inversa, el correspondiente secuencia de diagramas es

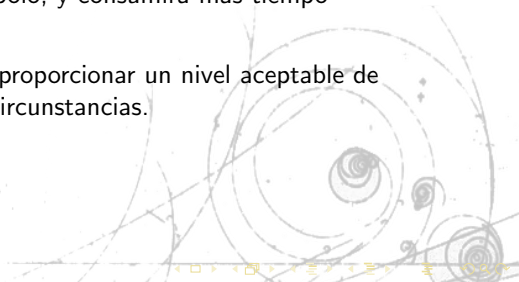
*IN TE RE SA NT EL AC LA SE HO YX*

Donde podemos intuir que el texto es:

*INTERESE NTE LA CLASE HOY*

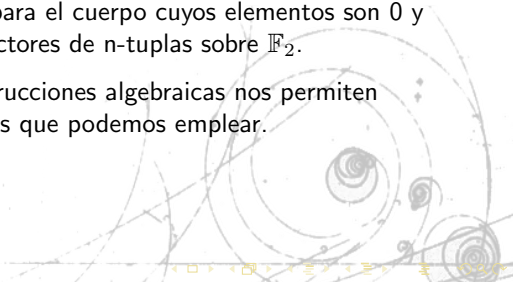


- Como podemos intuir, es posible atacar y romper el sistema *Playfair* mediante análisis de frecuencias, utilizando las tablas de frecuencias para diagramas y siguiendo el mismo procedimiento que detallamos anteriormente.
- Sin embargo, este procedimiento requiere más datos que en el caso de análisis de un simple símbolo, y consumirá más tiempo completarlo.
- Por tanto, el sistema puede proporcionar un nivel aceptable de seguridad en determinadas circunstancias.

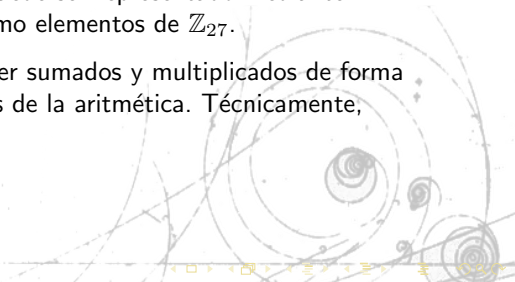


## Algoritmos matemáticos en criptografía

- Hasta ahora hemos usado las matemáticas principalmente con el propósito de explicar como funciona ciertos criptosistemas. En la criptografía actual, las matemáticas tiene un uso más fundamental.
- Los símbolos 0 y 1 en el alfabeto binario  $\mathbb{B}$  tienen propiedades algebraicas. Denotamos  $\mathbb{F}_2$  para el cuerpo cuyos elementos son 0 y 1 y  $\mathbb{F}_2^n$  para el espacio de vectores de n-tuplas sobre  $\mathbb{F}_2$ .
- La utilización de estas construcciones algebraicas nos permiten extender el rango de métodos que podemos emplear.



- En general, los símbolos en un mensaje pueden ser representados por los elementos de una estructura algebraica de muchas formas.
- La forma más simple de representar un conjunto de  $n$  objetos mediante una estructura algebraica es utilizar enteros modulo  $n$ , denotado como  $\mathbb{Z}_n$ .
- Por ejemplo, el alfabeto  $\mathbb{A}$  puede ser representado mediante  $0, 1, 2, \dots, 26$ , denotados como elementos de  $\mathbb{Z}_{27}$ .
- Los enteros *mod*  $n$  pueden ser sumados y multiplicados de forma que ellos satisfacen las reglas de la aritmética. Técnicamente, decimos que  $\mathbb{Z}_n$  es un *anillo*.



- En realidad, a menudo es conveniente utilizar un alfabeto con un *número primo de símbolos*, ya que cuando  $p$  es un número primo, los enteros *mod*  $p$  forman un cuerpo.
- Esto significa cada elemento que no es cero tiene un inverso en la operación de multiplicación. Para enfatizar su naturaleza especial denotaremos este cuerpo como  $\mathbb{F}_p$ .
- Por tanto si extendemos el alfabeto  $\mathbb{A}$ , permitiendo mensajes que contienen comas y puntos, además de los usuales 27 símbolos, entonces tendremos 29 símbolos, un número primo.
- Podemos representar el espacio  $\square$  por el 0, las letras  $A, B, \dots, Z$  por  $1, 2, \dots, 26$ , y la coma y el punto por 27 y 28, todos considerados como elementos del cuerpo  $\mathbb{F}_{29}$ .



- Por ejemplo, si el mensaje es

$I \sqcup CAME, \sqcup I \sqcup SAW, \sqcup I \sqcup CONQUERED.$

lo podemos reemplazar por:

$9\ 0\ 3\ 1\ 13\ 5\ 27\ 0\ 9\ 0\ 13\ 1\ 23\ 27\ 0\ 9\ 0\ 3\ 15\ 14\ 17\ 21\ 5\ 18\ 5\ 4\ 28$

- El punto es que, aunque la aritmética no tiene significado cuando los símbolos son “únicamente” símbolos, podemos realizar todas las operaciones aritméticas, incluyendo una división por un número distinto de cero. Esta situación incrementa enormemente el rango de métodos de cifrado que están disponibles.
- Otra posibilidad es cuando utilizamos la técnica estándar de dividir el flujo de símbolos en bloques con un tamaño apropiado  $m$ , por lo que cada bloque será un  $m$ -vector sobre  $\mathbb{F}_{29}$ . Ahora, no solo es posible aplicar la aritmética en símbolos individuales, podemos utilizar algebra lineal para manipular los vectores.



## Ejemplo 4

*Representamos el mensaje*

*$I \sqcup CAME, \sqcup I \sqcup SAW, \sqcup I \sqcup CONQUERED$ , como una cadena de 2-vectores en  $\mathbb{F}_{29}$ , y establecemos las cadenas de 2-vectores que resultan cuando aplicamos la matriz  $K$ . ¿Como podemos transformar al mensaje original el vector resultante?*

$$K = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}$$

Solución:

El mensaje original es representado por:

$$[9\ 0]'[3\ 1]'[13\ 5]' \dots$$

Y mediante las operaciones

$$K = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 27 \\ 18 \end{pmatrix}$$

son convertidos en

$$[27\ 18]'[13\ 9]'[1\ 12]' \dots$$

Dado la anterior cadena, podemos recuperar la primera aplicando la matriz inversa, la cual en este caso es

$$K^{-1} = \begin{pmatrix} 3 & 25 \\ 27 & 3 \end{pmatrix}$$

- El ejemplo ilustra la básica idea que hay detrás de un acercamiento inicial para utilizar técnicas de matemáticas abstractas en criptografía.
- En dos trabajos publicados sobre 1930, *L.S. Hill* sugería que los mensajes podrían ser cifrados mediante la aplicación de una transformación lineal, es decir, multiplicando los vectores por un matriz.
- Generalmente, en el sistema de Hill la clave es una matriz  $K$  invertible  $m \times n$  y la función de cifrado es

$$E_K(x) = Kx,$$

donde  $x$  es un  $m$ -vector.

- Si  $y = Kx$ , entonces  $K^{-1}y = x$ . Por tanto, la función de descifrado correspondiente a  $E_K$  es dada por  $D_L(y) = Ly$ , donde  $L = K^{-1}$ .
- El sistema de Hill es otro ejemplo de sistema de clave simétrica, ya que  $K^{-1}$  puede ser calculado desde  $K$  utilizando los métodos estándar de la algebra matricial.



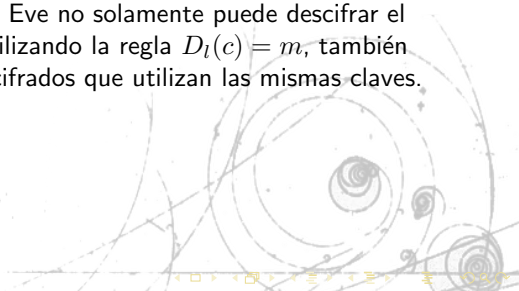
## Métodos de ataque

¿Como puede ser atacado el sistema Hill?

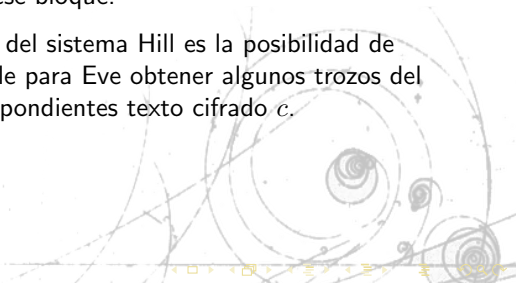
- Un ataque por búsqueda exhaustiva, requiere verificar todas las matrices invertibles  $m \times m$  sobre  $\mathbb{F}_{29}$ .
- Una matriz  $m \times m$  tiene  $m^2$  componentes, y si cada componente es un elemento de  $\mathbb{F}_{29}$ , el número de posibilidades es  $29^{m^2}$ .
- La mayoría de todas estas matrices son invertibles, y consecuentemente  $m$  puede ser escogido de forma que un ataque por búsqueda exhaustiva no sea posible.



- Aunque la búsqueda exhaustiva no sea posible, es posible otros ataques como por ejemplo el ataque de texto cifrado.
- Eve obtiene una pieza de texto cifrado  $c$  e intenta utilizar esta información para encontrar las claves de cifrado  $k$  y descifrado  $l$ .
- Si encontramos estas claves, Eve no solamente puede descifrar el texto cifrado que conoce, utilizando la regla  $D_l(c) = m$ , también puede obtener otros textos cifrados que utilizan las mismas claves.

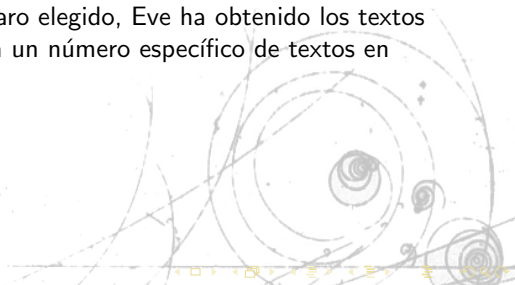


- El ataque mediante análisis de frecuencias es conocido como ataque de texto cifrado conocido. Pero en el sistema Hill este ataque no es posible ya que las letras han sido mezcladas.
- Dada cualquier secuencia de letras puede ser cifrado de muchas formas distintas, dependiendo de la posición en un bloque y las otras letras que suceden en ese bloque.
- Sin embargo, un punto débil del sistema Hill es la posibilidad de otro ataque: puede ser posible para Eve obtener algunos trozos del texto en claro  $m$  y los correspondientes texto cifrado  $c$ .



Los criptógrafos reconocen dos tipos de ataques basados en esta información:

- 1 En un ataque por texto en claro conocido, Eve ha obtenido algún par  $(m_i, c_i)$ , donde cada  $m_i$  es un texto en claro y  $c_i$  es el correspondiente texto cifrado.
- 2 En un ataque de texto en claro elegido, Eve ha obtenido los textos cifrados  $c_i$  correspondiente a un número específico de textos en claro  $m_i$  que ella ha elegido.



## Ejemplo 5

*Supongase que Eve intercepta algún texto cifrado el cual contiene un informe sobre una emboscada, utilizando el sistema Hill con  $m=2$ . Eve sospecha que el texto cifrado*

$[1515]'$  $[2515]'$  $[173]'$

*representa la palabra **AMBUSH** (emboscada). ¿Como puede probar su hipótesis y encontrar la clave?*



Solución:

En 2-vectores sobre  $\mathbb{F}_{29}$  la palabra *AMBUSH* es representada mediante

$$[113]'[221]'[198]'$$

Sabemos que la clave es

$$K = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

La hipótesis de Eve es que la secuencia  $[113]'[221]'[198]'$  es cifrada como  $[1515]'[2515]'[173]'$ . Si la hipótesis es verdadera, los bloques *AM* y *BU* son cifradas como:

$$\begin{pmatrix} 15 \\ 15 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 \\ 13 \end{pmatrix}, \quad \begin{pmatrix} 25 \\ 15 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2 \\ 21 \end{pmatrix}$$

cont. Solución:

Consecuentemente

$$\begin{pmatrix} 15 & 25 \\ 15 & 15 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 13 & 21 \end{pmatrix},$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 15 & 25 \\ 15 & 15 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 13 & 21 \end{pmatrix}^{-1}.$$

Eve haciendo cálculos de algebra elemental en  $\mathbb{F}_{29}$ , puede calcular

$$K = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}.$$



cont. Solución:

Eve puede verificar su hipótesis verificando que, con esta  $K$ , el tercer bloque  $SH$  es cifrado como esperaba:

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 19 \\ 8 \end{pmatrix} = \begin{pmatrix} 17 \\ 3 \end{pmatrix}$$

Ya que ha funcionado entonces es posible aplicar,

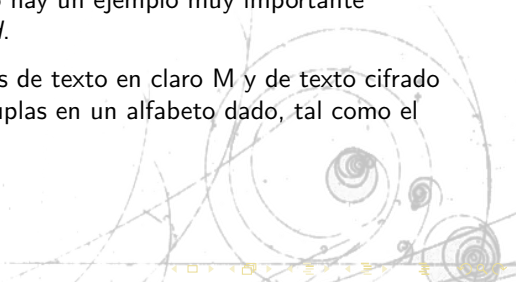
$$K^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$$

al resto del texto cifrado. Si obtenemos un texto en claro con significado, el problema está resuelto. En caso contrario, Eve puede intentar con un trozo de texto distinto.



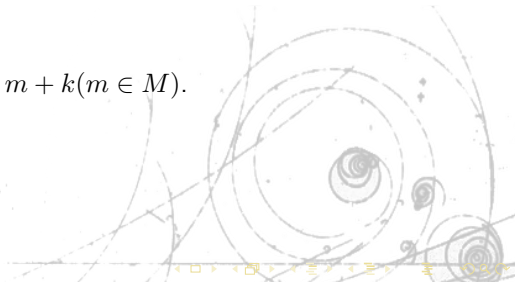
## Secreto Perfecto y One time pad

- Para que un criptosistema sea considerado como un secreto perfecto entonces el número de claves tiene que ser al menos igual al número de mensajes:  $|K| \geq |M|$ .
- El secreto perfecto es una condición muy restrictiva, y los ejemplos son difíciles de obtener. Pero hay un ejemplo muy importante conocido como *one-time pad*.
- En este sistema los conjuntos de texto en claro  $M$  y de texto cifrado son ambos conjuntos de  $n$ -tuplas en un alfabeto dado, tal como el alfabeto español  $\mathbb{A}$ .



- Ya que la teoría es aplicable para cualquier alfabeto, para los propósitos de exposición utilizaremos el alfabeto binario  $\mathbb{F}_2$ .
- Tomaremos  $M = C = \mathbb{F}_2^n$ , el vector de cadenas de longitud  $n$  en  $\mathbb{F}_2$ .
- El conjunto de todas las claves  $K$  es también  $\mathbb{F}_2^n$ , por lo que cumplimos la condición anterior. Para cada  $k \in K$  la función de cifrado  $E_k$  es

$$m \mapsto m + k (m \in M).$$



- La correspondiente función de descifrado es la misma función,  $c \mapsto c + k$ , ya que

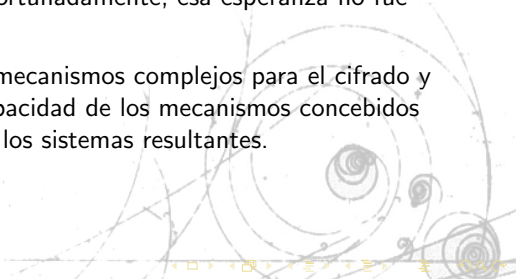
$$(m + k) + k = m.$$

- Claramente tenemos un sistema de clave simétrica.
- La idea detrás del sistema es que el texto en claro es expresado como una cadena  $m$  de  $n$  bits, y es cifrado añadiéndole una cadena de clave arbitraria  $k$  de la misma longitud.
- El sistema cumple con la condición de secreto perfecto si la cadena de la clave es escogida uniformemente aleatoria.
- En otra palabras, cada nuevo mensaje es cifrado utilizando una nueva clave, donde todas las claves son igualmente probables, de hay la razón de su nombre.

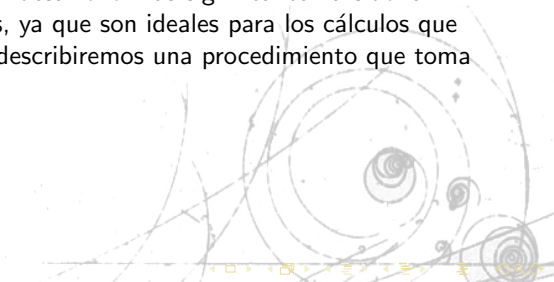


# Métodos Iterativos

- Durante el siglo veinte los avances en los fundamentos matemáticos de la criptografía fueron de la mano con los desarrollos tecnológicos.
- Utilizando máquinas para mezclar el texto en claro de forma eficiente, se esperaba que era posible alcanzar cualquier nivel deseado de seguridad. Desafortunadamente, esa esperanza no fue alcanzada.
- La capacidad para concebir mecanismos complejos para el cifrado y descifrado se adecua a la capacidad de los mecanismos concebidos que pueden atacar y romper los sistemas resultantes.



- La conclusión obtenida fue que los criptosistemas de clave simétrica solamente pueden proporcionar seguridad de forma relativa.
- Los sistemas pueden solamente garantizar la seguridad contra los métodos de ataque conocidos.
- En la parte tecnológica, el desarrollo más significativo ha sido la aplicación de ordenadores, ya que son ideales para los cálculos que involucran iteraciones, y describiremos un procedimiento que toma ventaja de este hecho.



- Supongase que dado un trozo de texto  $X$ , expresado como una cadena de  $n$  bits, y una clave  $k$ , como una cadena de  $s$  bits.
- Sea  $F$  una función que asigna a cada  $X$  y  $k$  otra cadena de  $n$  bits,  $Y = F(k, X)$ . En otras palabras tenemos una función

$$F : \mathbb{F}_2^s \times \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$$

- Por razones de seguridad,  $F$  debe tener unas características especiales, tal como no ser una función lineal.
- Cuando  $s=2$  y  $n=3$  podríamos utilizar la función definida por la clave  $k = \alpha\beta$  y  $X = x_1x_2x_3$  por la regla

$$y_1 = \alpha x_1 x_2 + \beta, y_2 = x_2 + \beta x_3, y_3 = (\alpha + \beta) x_1 x_3.$$

- Para una clave dada, esta función no es lineal, ni es una biyección.

### Definición 5 (Iteración Feistel)

La iteración Feistel asociada con  $F$  y los valores iniciales  $X_0, X_1 \in \mathbb{F}_2^n$ , es definida como sigue.

Sea  $k = (k_1, k_2, \dots, k_r)$  una secuencia de claves, cada una de las cuales es un elemento de  $\mathbb{F}_2^s$ , y  $X_{i+1} = X_{i-1} + F(k_i, X_i)$  ( $i = 1, 2, \dots, r$ ), donde  $+$  representa la operación de suma bit a bit en el espacio de vectores  $\mathbb{F}_2^n$ .



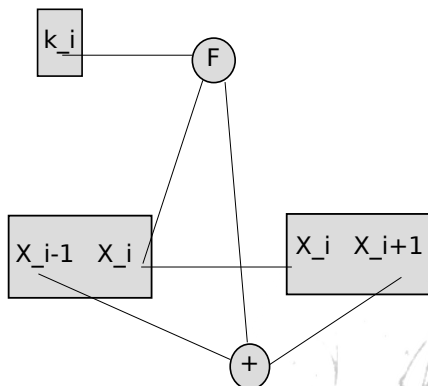


Figura 2: La  $i^{\text{th}}$  ronda de una iteración Feistel



- En criptografía cada paso de la iteración es denominado una ronda. En la  $i$ th ronda la clave  $k_i$  es utilizada para transformar el par  $X_{i-1}X_i$  en  $X_iX_{i+1}$  (Figura 2).
- El cifrado es realizado expresando todos los mensajes como secuencias de cadenas de bits de longitud  $2n$ .
- Si el texto en claro contiene  $m$   $2n$ -cadenas, es dividida en dos partes iguales,  $m = X_0X_1$ , y entonces las  $n$ -cadenas  $X_0$  y  $X_1$  son tomadas como los valores iniciales para la iteración de Feistel.
- El correspondiente texto cifrado es la salida después de  $r$  rondas:  
 $c = X_rX_{r+1}$ .
- Por tanto la función de cifrado para una secuencia de claves  $k$  es dada por

$$E_k(m) = c, \text{ donde } m = X_0X_1, c = X_rX_{r+1}.$$

## Ejemplo 6

Tomaremos  $s=n=3$  y definimos  $Y=F(k,X)$  para una clave  $k = \alpha\beta\gamma$  por la regla

$$y_1y_2y_3 = F(\alpha\beta\gamma, x_1x_2x_3), \text{ donde}$$
$$y_1 = \alpha x_1 + x_2x_3, y_2 = \beta x_2 + x_1x_3, y_3 = \gamma x_3 + x_1x_2.$$

Si la clave es  $k = (100,101,001,111)$  y  $m = 101\ 110$ , calcula  $c = E_k(m)$ .

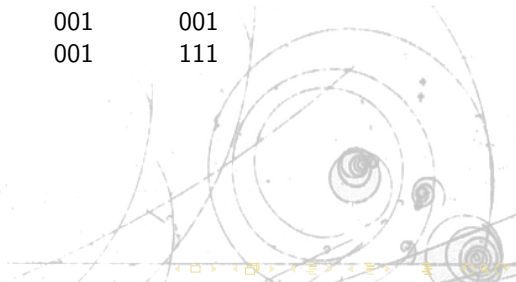


*Solución:*

Los cálculos pueden ser tabulados como sigue:

$i$	$k_i$	$X_i$	$F(k_i, X_i)$	$X_{i+1}$
0	-	101	-	110
1	100	110	101	000
2	101	000	000	110
3	001	110	001	001
4	111	001	001	111

Entonces  $c = 001\ 111$ .



La iteración de Feistel proporciona un mecanismo para mezclar los datos, ya que los parámetros  $n, s, r$  pueden ser tan largos como deseemos, sin embargo es un sistema de clave simétrica trivial.

### Teorema 1

*Supongase que el mensaje  $m = X_0X_1$  es cifrado utilizando el sistema Feistel con una función  $F$  y la secuencia de claves  $k = (k_1, k_2, \dots, k_r)$ , por lo que  $c = E_k(m) = X_rX_{r+1}$ .*

*Entonces el sistema Feistel con la misma función  $F$  y la secuencia de claves  $k^* = (k_r, k_{r-1}, \dots, k_1)$  cuando lo aplicamos a  $c' = X_{r+1}X_r$  obtiene  $m' = X_1X_0$ .*

- Hemos notado que la función  $F$  no puede tener ciertas propiedades. Por ejemplo, si es una función lineal, el sistema Feistel será vulnerable a un ataque de texto en claro conocido.
- Como siempre, la seguridad del sistema también dependerá del tamaño de las claves usadas para evitar ataques de búsqueda exhaustiva.
- Todas las claves  $k_1, k_2, \dots, k_r$  pueden ser escogidas independientemente, pero en práctica es usual extraerlas eligiendo unos conjuntos específicos de  $s$  bits de una clave maestra  $K$ .
- Por ejemplo, si necesitamos 12 claves de longitud 32, Alice y Bob pueden ponerse de acuerdo en una clave maestra de longitud 120, y tomar  $k_1$  como los bits desde 1 a 32,  $k_2$  como los bits 9 a 40,  $k_3$  como los bits 17 al 48, y así continuamente.



## Estándar de Cifrado

- En los 70s la popularidad y el incremento en el uso de procedimientos criptográficos en comercio y la industria dieron lugar a necesidad de establecer un estándar de cifrado.
- El diseño de que cualquier criptosistema tiene que tener en cuenta, que su seguridad solo debe depender de mantener la clave en secreto.
- Un sistema que pretenda ser un estándar de seguridad debe ser capaz de responder estas dos cuestiones básicas.
  - ¿Es posible el ataque por búsqueda exhaustiva?
  - ¿Existe un ataque que es mejor que la búsqueda exhaustiva?



- En 1977 se estableció el primer Estándar de Cifrado de Datos conocidos como DES.
- El componente central de la propuesta es conocido como S-box.
- En términos generales, es una matriz  $S$  donde las filas corresponden a elementos de  $\mathbb{F}_2^M$ , las columnas corresponden a los elementos de  $\mathbb{F}_2^N$  y las entradas a  $\mathbb{F}_2^R$ .
- En DES hay 8 S-boxes, con los parámetros  $M=2$ ,  $N=4$  y  $R=4$ . Un ejemplo es mostrado en la Figura 5.
- Para clarificar, las filas son etiquetadas 0-3 como números binarios,  $0 = 00$ ,  $1 = 01$ ,  $2 = 10$ ,  $3 = 11$ . Las etiquetas de las columnas y las entradas de  $S$  son denotadas por 0-15 de la misma forma: por tanto  $5 = 0101$ ,  $12 = 1100$ , y así con todas.

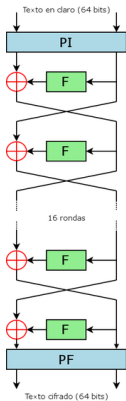


Figura 3: Estructura General DES

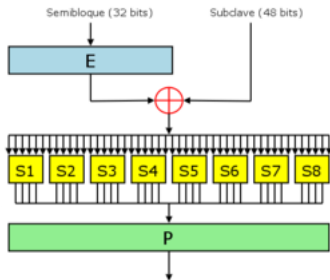
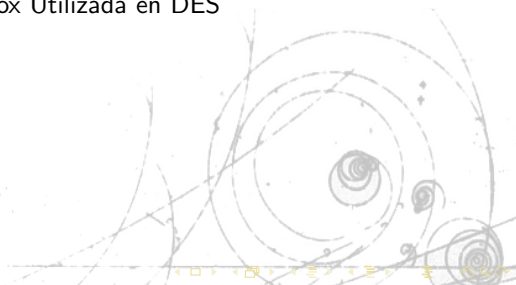


Figura 4: Función Feistel



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
1	10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
2	9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
3	4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13

Figura 5: S-box Utilizada en DES



- Cada S-box en DES determina una función  $f_S : \mathbb{F}_2^6 \rightarrow \mathbb{F}_2^4$  definida por la regla

$$f_S(x_1x_2x_3x_4x_5x_6) = S(x_1x_6, x_2x_3x_4x_5).$$

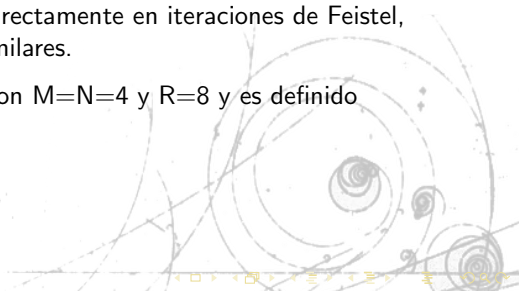
- Por ejemplo, para la S-box dada arriba,

$$f_S(1111001) = S(11, 1100) = S(3, 12) = 6 = 0110.$$

- La función de cifrado DES opera con bloques de 64 bits, utilizando una clave con (esencialmente) 56 bits.
- Es obtenida mediante permutación de los bits y combinando las 8 funciones  $f_S$  definidas por las S-boxes, utilizando una iteración de Feistel con 16 rondas.

- Las S-boxes son construidas de forma que satisfacen un número de criterios, diseñadas para hacerlas más segura frente métodos conocidos de ataques. Dos de esos criterios son:
  - C1 Cada fila de S es una permutación  $\mathbb{F}_2^4$
  - C2 Si  $x, y \in \mathbb{F}_2^6$  cumple que la distancia Hamming  $d(x, y) = 1$ ,  $yx' = f_S(x)$ ,  $y' = f_S(y)$ , entonces  $d(x', y') \geq 2$ .
- La propuesta del DES estableció mucha controversia.
- El método en que se basa la construcción de las S-boxes no fue inicialmente revelado, un hecho que algunas personas tildan de sospechoso.
- Los parámetros escogidos son muy pequeños, principalmente porque la clave maestra solo tiene 56 bits efectivos, esta situación no libera de un ataque por búsqueda exhaustiva.

- Debida a esta situación, en los 90s, estaba claro que DES no podía contestar estas cuestiones satisfactoriamente, por tanto fue reemplazado por el estándar de cifrado avanzado (AES).
- En este sistema la clave maestra tiene 128 bits, y actualmente no es posible realizar un ataque por búsqueda exhaustiva.
- El sistema no está basado directamente en iteraciones de Feistel, pero se aplican principios similares.
- Hay solamente una S-box, con  $M=N=4$  y  $R=8$  y es definido mediante una fórmula.



# El problema de distribución de claves

- En un criptosistema de clave simétrica, Alice y Bob utilizan claves similares.
- Metafóricamente hablando, Alice pone el mensaje en una caja la cual ella cierra utilizando la clave  $k$ .
- Cuando Bob recibe la caja la abre utilizando una clave  $l$  que está muy relacionada con  $k$ , podemos establecer esta relación como  $l = k'$ .
- El principal problema es que Alice y Bob deben comunicarse para ponerse de acuerdo en la clave, y esta comunicación no puede ser por tanto protegida por el proceso de cifrado.
- Esta situación se conoce como el problema de distribución de claves.

- Hay varias formas en las cuales el problema de distribución de claves puede ser solucionado.
- Si tenemos recursos disponibles, podríamos utilizar el one-time pad, de esta manera antes de transmitir cada mensaje Alice envía una nueva clave a Bob de forma segura.
- Claramente, este caso no es práctico para la mayoría de los propósitos. Una medida más práctica es considerar alternativas a los procedimientos de clave simétrica.
- ¿Es posible designar un procedimiento en el cual la clave de Alice y Bob sean independientes, de forma que no necesiten conocer las claves de uno del otro?



Un posible procedimiento sería el descrito a continuación:

- 1 Alice envía a Bob un mensaje cifrado utilizando su clave  $a$ .  
 $A \rightarrow B : E_a(m)$
- 2 Bob cifra el mensaje utilizando su clave  $b$  y lo devuelve a Alice.  
 $B \rightarrow A : E_b(E_a(m))$
- 3 Alice descifra utilizando  $a'$  y lo envía a Bob.  
 $A \rightarrow B : D_{a'}(E_b(E_a(m))) = E_b(m)$
- 4 Bob descifra utilizando  $b'$ .  
 $B : D_{b'}(D_{a'}(E_b(E_a(m)))) = D_{b'}(E_b(m)) = m$

El procedimiento funciona si las operaciones de Alice conmutan con las de Bob.

# El Criptosistema RSA

- En la sección anterior hemos tratado el problema intrínseco que presentan los criptosistemas de clave simétrica debido al problema de la distribución de las claves.
- En 1970s, el desarrollo de la criptografía de clave pública vino a solventar la cita situación.
- La idea fundamental es que un usuario (Bob) tiene dos claves, una pública y otra privada.
- La clave pública es utilizada por Alice y otros para cifrar mensajes que ellos quieren enviar a Bob, y la clave privada es utilizada por Bob para descifrar estos mensajes.
- La seguridad del sistema recae en el aserto que la clave privada de Bob no puede ser calcula, incluso si todo el mundo conoce su clave pública.





- Un método práctico de implementar esta idea fue descubierta por Rivest, Shamir, y Adleman en 1977 y es conocida como criptosistema RSA.
- En el sistema RSA los textos en claro son secuencias de enteros  $\text{mod } n$ , donde es razonable determinar  $n$  como un número de al menos 300 dígitos decimales.

### Definición 6 (Funcion $\phi$ )

*Para cualquier entero positivo  $n$ , el número de enteros  $x$  en el rango  $1 \leq x \leq n$  tal que  $\text{mcd}(x,n)=1$  es denotado como  $\phi(n)$ .  $\phi(n)$  representan también el número de elementos invertibles de  $\mathbb{Z}_n$ .*

- Por ejemplo tomando  $n=14$  los valores de  $x$  tal que  $\text{mcd}(x,14) = 1$  son 1,3,5,9,11,13 por lo que  $\phi(14) = 6$ .
- En este caso las inversas son fácilmente calculables, por ejemplo  $3^{-1} = 5$  ya que  $3 \times 5 = 1 \pmod{14}$ .
- Si  $n$  es un número primo,  $n = p$ , entonces los  $p-1$  números  $x = 1, 2, \dots, p - 1$  satisfacen  $\text{mcd}(x,p) = 1$  y por tanto  $\phi(p) = p - 1$ .

### Lema 1

*Si  $n = pq$ , donde  $p$  y  $q$  son primos, entonces  $\phi(n) = (p - 1)(q - 1)$ .*

En el sistema RSA, los usuarios deben de realizar los siguientes pasos previos a la realización de las funciones de cifrado y descifrado:

- Escoge dos números primos  $p, q$  y calcula

$$n = pq, \phi = (p - 1)(q - 1)$$

- Escoge  $e$  tal que  $\text{mcd}(e, \phi) = 1$ , y calcula

$$d = e^{-1}(\text{mod } \phi)$$

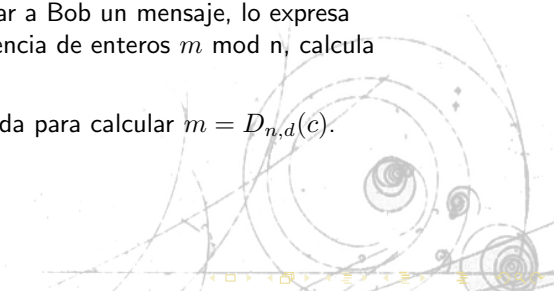
Las funciones de cifrado y descifrado son definidas como sigue:

$$E_{n,e}(m) = m^e \quad (m \in \mathbb{Z}_n),$$
$$D_{n,d}(c) = c^d \quad (c \in \mathbb{Z}_n)$$



El sistema funciona de la siguiente forma:

- Bob utiliza las reglas anteriores para construir, los números  $n$ ,  $\phi$ ,  $e$ , y  $d$ .
- Una vez calculados los elementos, pone a disposición de todos la *clave pública*  $(n,e)$ , pero mantiene la *clave privada*  $d$  en secreto.
- Cuando Alice quiere enviar a Bob un mensaje, lo expresa inicialmente en una secuencia de enteros  $m \bmod n$ , calcula  $c = E_{n,e}(m)$  y envía  $c$ .
- Bob utiliza su clave privada para calcular  $m = D_{n,d}(c)$ .



## Ejemplo 7

Supongamos Bob ha elegido  $p = 47$ ,  $q = 59$ .

(i) Encontrar  $n$  y  $\phi$ , y muestra que  $e=157$  es una valor válido para la clave pública.

(ii) Verifica que su clave privada es  $d=17$ .

(iii) Explicar los pasos que debe realizar Alice para enviar un mensaje cifrado a Bob representado por una secuencia de enteros, por ejemplo  $m=5$ . ¿Como Bob lo podría descifrar?



### Solución:

(i) Tenemos

$$n = 47 \times 59 = 2773, \phi = 46 \times 58 = 2668.$$

La elección  $e = 157$  es válida ya que  $\text{mcd}(2668, 157) = 1$ . Esta condición se cumple ya que 157 es primo y no divide 2668.

(ii) Podemos probar la condición calculando:

$$d \times e = 17 \times 157 = 2669 = 1 \pmod{2668}.$$

continua ...

(iii) Si Alice quiere enviar a Bob un mensaje, buscará su clave pública  $(n,e)=(2773, 157)$ , y realizará un proceso de conversión del mensaje en una secuencia de enteros mod  $n$ . Después cifrará el mensaje utilizando la función de cifrado  $E_{n,e}(m) = m^e$ , y envía el mensaje cifrado a Bob. En este caso

$$c = 5^{157} = 1044 \pmod{2773}.$$

Cuando Bob recibe el texto cifrado  $c = 1044$ , aplicará su función de descifrado  $D_{n,d}(c) = c^d$ , obteniendo

$$m' = 1044^{17} = 5 \pmod{2773}.$$

Podemos probar que  $m' = m$ , por lo que el mensaje original es obtenido.

## Viabilidad de RSA

Daremos una breve explicación de como realizar los cálculos en RSA fácilmente.

Básicamente depende de la ejecución de dos algoritmos, los cuales son usados para reducir problemas complicados a una aritmética básica.

- El algoritmo de Euclides es esencialmente un método para calcular el  $\text{mcd}(a, b)$  de dos enteros  $a$  y  $b$  (asumimos que  $a < b$ ). Depende básicamente del hecho que si

$$b = qa + r \text{ entonces } \text{mcd}(a, b) = \text{mcd}(r, a).$$

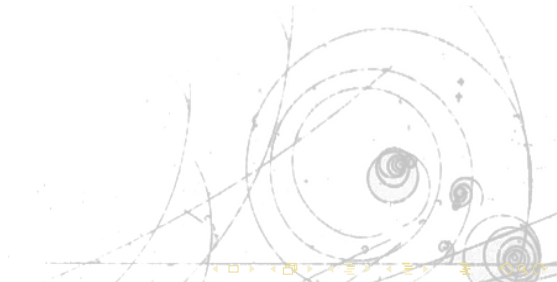
- Podemos reemplazar  $(a, b)$  por  $(a', b') = (r, a)$ , y repetir el proceso.





## Ejemplo 8

*Encuentra el  $\text{mcd}(654, 2406)$ .*



## Ejemplo 8

Encuentra el  $\text{mcd}(654, 2406)$ .

Solución:

El proceso se realiza como sigue:

$$2406 = 3 \times 654 + 444$$

$$654 = 1 \times 444 + 210$$

$$444 = 2 \times 210 + 24$$

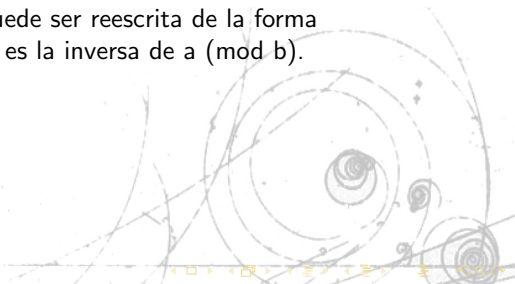
$$210 = 8 \times 24 + 18$$

$$24 = 1 \times 18 + 6$$

$$18 = 3 \times 6.$$

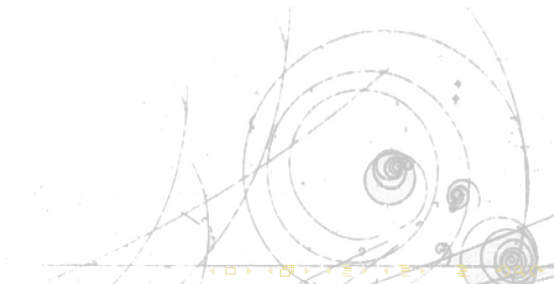
Por tanto 6 es el máximo común divisor que divide 18, 24, 210, 444, 654, 2406, y en general  $\text{mcd}(654, 2406) = 6$ .

- Si  $\text{mcd}(a, b) = 1$ , entonces  $a$  tiene una inversa multiplicativa mod  $b$ , y el algoritmo de Euclides puede ser utilizado para calcularlo.
- El algoritmo es aplicado de forma inversa de la forma  $\lambda a + \mu b$ , donde  $\lambda$  y  $\mu$  son enteros.
- La ecuación  $1 = \lambda a + \mu b$  puede ser reescrita de la forma  $\lambda a = 1(\text{mod } b)$ . Por tanto  $\lambda$  es la inversa de  $a \pmod{b}$ .



## Ejemplo 9

*Encontrar la inversa de 24 mod 31.*



## Ejemplo 9

Encontrar la inversa de 24 mod 31.

Solución:

Primero verificamos que  $\text{mcd}(24, 31) = 1$ :

$$31 = 1 \times 24 + 7$$

$$24 = 3 \times 7 + 3$$

$$7 = 2 \times 3 + 1$$

$$3 = 3 \times 1.$$

Ahora, comenzando con la última ecuación y volviendo hacia atrás tenemos,

$$1 = 7 - 2 \times 3$$

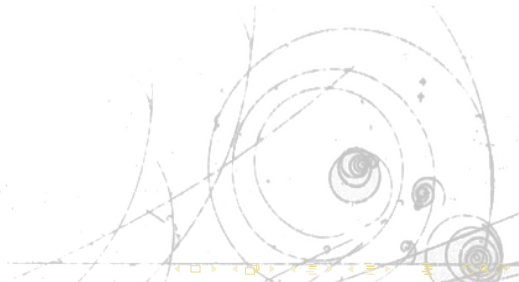
$$= 7 - 2 \times (24 - (3 \times 7)) = 7 \times 7 - 2 \times 24$$

$$= 7 \times (31 - 1 \times 24) - 2 \times 24 = -9 \times 24 + 7 \times 31.$$

Por tanto  $(-9) \times 24 = 1 \pmod{31}$ , por lo que  $24^{-1} = -9 = 22$ .

## Ejemplo 10

*En el ejemplo anterior Bob establece una clave RSA eligiendo  $p = 47$ ,  $q = 59$ , por lo que  $n = 2773$ ,  $\phi = 2668$ . Utiliza el algoritmo de euclides para demostrar que  $e = 157$  es valor válido para su clave pública, y explica como puede calcular su clave privada  $d = e^{-1}(\text{mod } \phi)$ .*



## Ejemplo 10

*En el ejemplo anterior Bob establece una clave RSA eligiendo  $p = 47$ ,  $q = 59$ , por lo que  $n = 2773$ ,  $\phi = 2668$ . Utiliza el algoritmo de euclides para demostrar que  $e = 157$  es valor válido para su clave pública, y explica como puede calcular su clave privada  $d = e^{-1}(\text{mod } \phi)$ .*

### Solución:

El algoritmo de Euclides muestra que  $\text{mcd}(e, \phi) = 1$ :

$$2668 = 16 \times 157 + 156$$

$$157 = 1 \times 156 + 1$$

$$156 = 156 \times 1.$$

$d = e^{-1}$  es calculada:

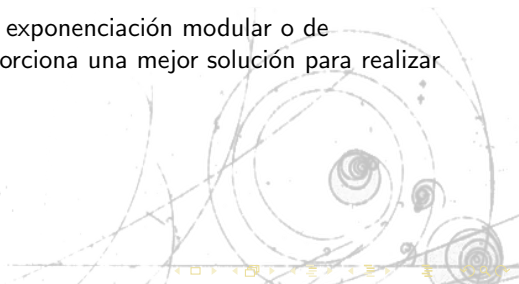
$$1 = 157 - 1 \times 156$$

$$= 157 - 1 \times (2668 - (16 \times 157))$$

$$= (-1) \times 2668 + 17 \times 157.$$

Por tanto  $17 \times 157 = 1(\text{mod } 2668)$ , so por lo que 17 es la clave privada de Bob.

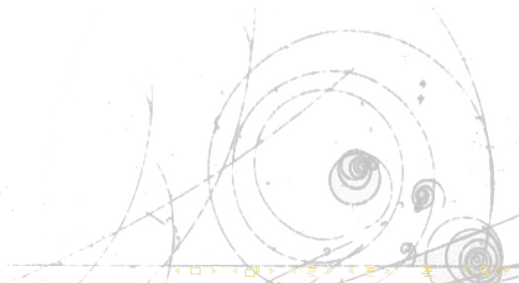
- El otro algoritmo utilizado tiene que ver con las funciones exponenciales  $a^b \pmod{c}$  utilizadas en las funciones de cifrado ( $c = m^e$ ) y descifrado ( $m = c^d$ ).
- Básicamente para calcular la potencia  $b^k$  debemos de utilizar  $k-1$  multiplicaciones, ya que  $b^k = b \times b \times \dots \times b$ .
- Sin embargo el algoritmo de exponenciación modular o de cuadrados repetido nos proporciona una mejor solución para realizar los cálculos.





## Ejemplo 11

*Supongamos que dado el número  $b$ . ¿Cuántas multiplicaciones necesitamos para calcular  $b^{16}$  y cuántas para  $b^{23}$ ?*



## Ejemplo 11

*Supongamos que dado el número  $b$ . ¿Cuántas multiplicaciones necesitamos para calcular  $b^{16}$  y cuántas para  $b^{23}$ ?*

Solución:

El cálculo de  $b^{16}$  requiere solamente cuatro multiplicaciones:

$$b^2 = b \times b, b^4 = b^2 \times b^2, b^8 = b^4 \times b^4, b^{16} = b^8 \times b^8.$$

Sabiendo que  $23 = 16 + 4 + 2 + 1$ , solamente necesitamos tres multiplicaciones más para  $b^{23}$ :

$$b^{23} = b^{16} \times b^4 \times b^2 \times b.$$

Por tanto  $b^{23}$  puede ser realizadas con  $4+3=7$  multiplicaciones.

## Fiabilidad de RSA

Para explicar como funciona RSA, nos basamos es explicar porqué la función  $D_{n,d}$  es la inversa de  $E_{n,e}$ .  
Acorde con la definición, la condición

$$D_{n,d}(E_{n,e}(m)) = m$$

reduce a  $(m^e)^d = m \pmod{n}$ , por lo que debemos probar que esta situación se da cuando  $d$  es la inversa de  $e \pmod{\phi(n)}$ .

### Lema 2

*Si  $\text{mcd}(x,n) = 1$ , por lo que  $x$  tiene una inversa en  $\mathbb{Z}_n$ , entonces  $x^{\phi(n)} = 1$  en  $\mathbb{Z}_n$ .*

## Teorema 2

Si las funciones de cifrado y descifrado RSA son  $E_{n,e}$  y  $D_{n,d}$ , y  $\text{mcd}(m, n) = 1$ , entonces

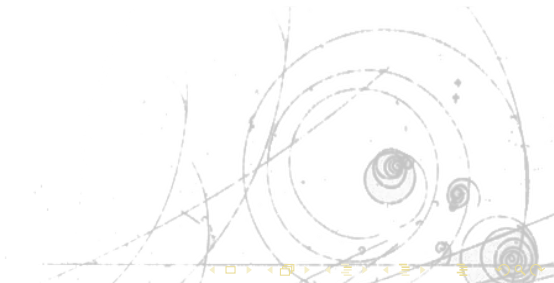
$$D_{n,d}(E_{n,e}(m)) = m.$$

O lo que es lo mismo,  $m^e = m \pmod{n}$ .



## Ejercicio 1

*Realiza una lista de aquellos elementos de  $\mathbb{Z}_{36}$  que tiene inversa multiplicativas, y encuentra la inversa de 5 mod 36.*

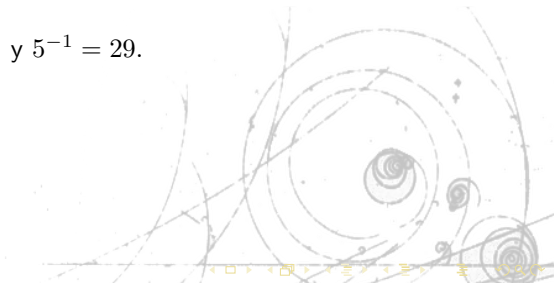


## Ejercicio 1

*Realiza una lista de aquellos elementos de  $\mathbb{Z}_{36}$  que tiene inversa multiplicativas, y encuentra la inversa de 5 mod 36.*

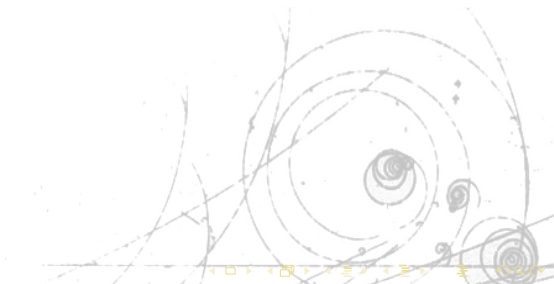
Solución:

El número de elementos es 12, y  $5^{-1} = 29$ .



## Ejercicio 2

Evalúa  $\phi(257)$  y  $\phi(253)$

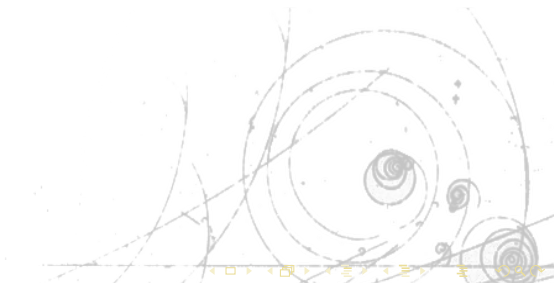


## Ejercicio 2

Evalúa  $\phi(257)$  y  $\phi(253)$

Solución:

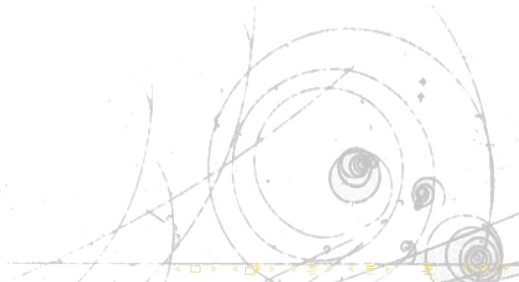
$$\phi(257) = 256 \text{ y } \phi(253) = 220$$





### Ejercicio 3

*Jorge y Antonio utilizan el criptosistema RSA. Jorge publica su clave pública  $n = 77$ ,  $e = 43$ . ¿Cuáles son los primos  $p$  y  $q$ , y cuál es su clave privada?. Antonio envía un mensaje  $m \in \mathbb{Z}_n$ , y Pepe intercepta el mensaje cifrado  $c=5$ . ¿Cuál era  $m$ ?*



### Ejercicio 3

*Jorge y Antonio utilizan el criptosistema RSA. Jorge publica su clave pública  $n = 77$ ,  $e = 43$ . ¿Cuáles son los primos  $p$  y  $q$ , y cuál es su clave privada?. Antonio envía un mensaje  $m \in \mathbb{Z}_n$ , y Pepe intercepta el mensaje cifrado  $c=5$ . ¿Cuál era  $m$ ?*

#### Solución:

En este caso  $n=77$  por lo que es muy fácil factorizar en  $77 = 11 \times 7$ , eso nos da que por ejemplo  $p = 7$  y  $q = 11$ .

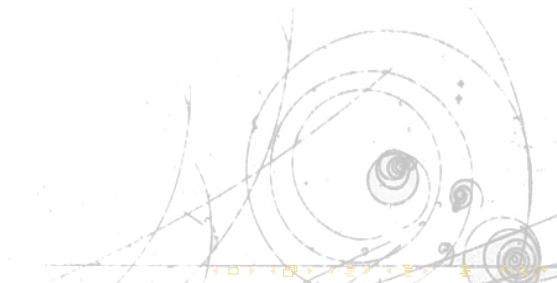
De esta información sabemos el valor de  $\phi$  ya que

$\phi = (p - 1) \times (q - 1) = 6 \times 10 = 60$ , también sabemos  $d = e^{-1} \bmod \phi$ , por lo que  $d = 43^{-1} \bmod 60 = 7$ .

Pepe sabe que  $m = c^d \bmod n$  por lo que  $m = 5^7 \bmod 77 = 47$ .

## Ejercicio 4

*Un usuario de RSA ha anunciado que la clave pública  $n = 2903239$ ,  $e = 5$ . María ha conseguido averiguar que  $n = 1237 \times 2347$ . Verifica que la clave pública es válida y explica por qué la clave privada es  $d = 2319725$ .*



## Ejercicio 4

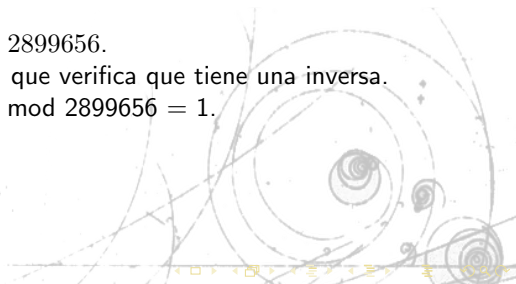
*Un usuario de RSA ha anunciado que la clave pública  $n = 2903239$ ,  $e = 5$ . María ha conseguido averiguar que  $n = 1237 \times 2347$ . Verifica que la clave pública es válida y explica por qué la clave privada es  $d = 2319725$ .*

### Solución:

Sabemos que  $\phi = 1236 \times 2346 = 2899656$ .

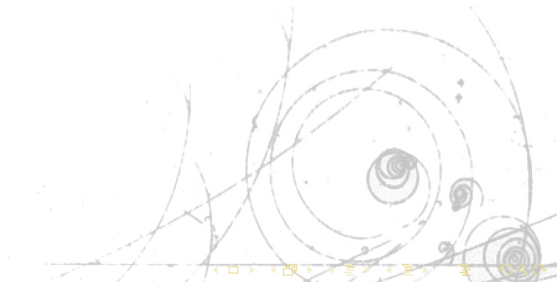
Calculamos  $\text{mcd}(2899656, 5) = 1$  que verifica que tiene una inversa.

$d \times e = 2319725 \times 5 = 11598625 \text{ mod } 2899656 = 1$ .



## Ejercicio 5

*Utiliza el algoritmo de Euclides para mostrar que  $\text{mcd}(15,68)=1$ . Además encuentra la inversa de  $15 \bmod 68$ .*

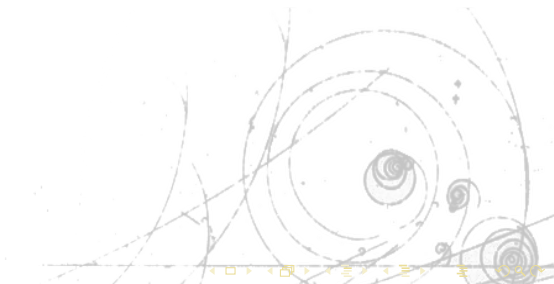


## Ejercicio 5

*Utiliza el algoritmo de Euclides para mostrar que  $\text{mcd}(15,68)=1$ . Además encuentra la inversa de  $15 \bmod 68$ .*

Solución:

$$15^{-1} \bmod 68 = 59.$$



# Certificados de Identidad

La criptografía de clave pública requiere de la utilización de certificados de identidad.

Un certificado de clave pública o certificado de identidad es un documento firmado electrónicamente que vincula una clave pública a una identidad. Por tanto el certificado es utilizado para verificar que una clave pública pertenece a un individuo.

El formato de certificado más usado actualmente es el denominado X.509.<sup>1</sup>

Los certificados X.509 contienen una serie de campos, entre los que podemos encontrar:

- Versión Certificado, Número de Serie, Emisor, Validez, Identificación del Usuario, Información de la clave pública, Algoritmo de la clave pública, Clave pública.

---

<sup>1</sup>Más información en IETF

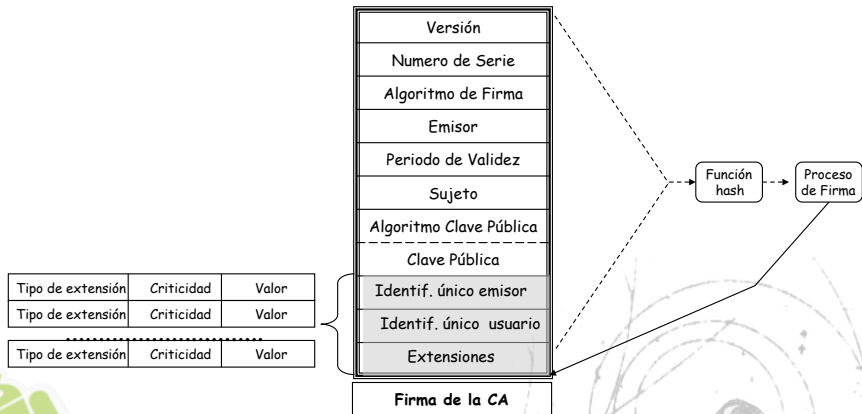


Figura 6: Representación Visual X509v3



Usualmente las conexiones seguras, p.ej. SSL, transmiten cadenas de confianza (o caminos de certificación).

Estas cadenas de confianza son una serie de certificados, que permiten validar los certificados emitidos hasta llegar a un certificado autofirmado.

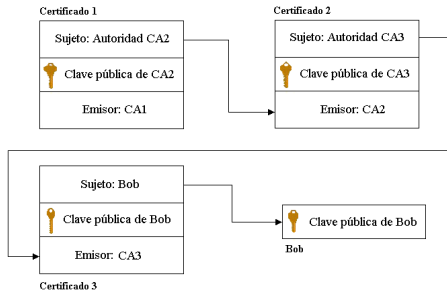
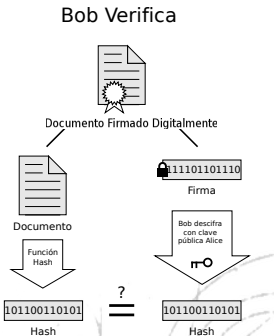
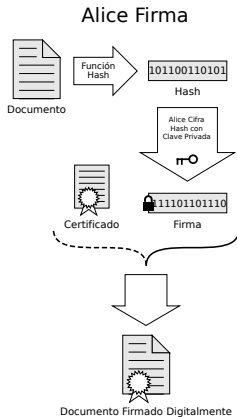


Figura 7: Cadena Confianza Certificados

# Esquema de Firma Digital y Verificación



En caso que los valores hash sean los mismo la firma es correcta

Figura 8: Esquema de Firma Digital

**José A. Montenegro Montes**

*Dpto. Lenguajes y Ciencias de la Computación  
ETSI Informática. Universidad de Málaga*

**monte@lcc.uma.es**

