JAVIER
ECHAIZ

Criptografía

Criptosistemas Clásicos

Clase 6

Prof. Javier Echaiz
D.C.I.C. – U.N.S.
http://cs.uns.edu.ar/~jechaiz
je@cs.uns.edu.ar

Esteganografía

• El mensaje plano puede “ocultarse” de dos formas:

Criptografía vs. Esteganografía

• No es criptografía. La esteganografía “oculta” la existencia del mensaje mientras que la criptografía lo vuelve ininteligible a ojos intrusos.

Algunas técnicas históricas

- Marcado de caracteres.
- Tinta invisible.
- Cinta de maq. de escribir.
- Pinturas.
- Fotos digitales.
- etc.

Ventajas/Desventajas de la Esteganografía

Desventajas

- Mucho overhead para ocultar pocos bits de info.
- Una vez que se “descubre” el sistema es totalmente inútil (igual que si obtienen mi clave secreta bajo criptografía simétrica).

Soluciones:

- Agregar algún tipo de claves.
- **Se puede primero encriptar y luego “ocultar”.**

Ventajas

- El mensaje oculto pasa desapercibido.

Clasificación Histórica de Criptosistemas

Los criptosistemas pueden clasificarse según:

a) Su relación con la Historia en:

- Sistemas Clásicos y Sistemas Modernos

No es ésta la mejor clasificación desde el punto de vista de la computación...

No obstante, permitirá comprobar el desarrollo de estas técnicas de cifrado hoy en día rudimentarias y en algunos casos simples, desde una perspectiva histórica, interesante también como cultura general.

Clasificación actual de los Criptosistemas

o bien según:

b) El tratamiento de la información a cifrar en:

- ***Cifrado en Bloque*** y ***Cifra en Flujo***

c) El tipo de clave utilizada en la cifra en:

- ***Clave Secreta*** y ***Clave Pública***

Cifrado en flujo	→	Próxima clase
Cifrado en bloque	→	Próxima clase
Cifra con clave secreta	→	Clase siguiente a Prox. clase
Cifra con clave pública	→	Clase siguiente a Prox. clase

Primera Aproximación Histórica

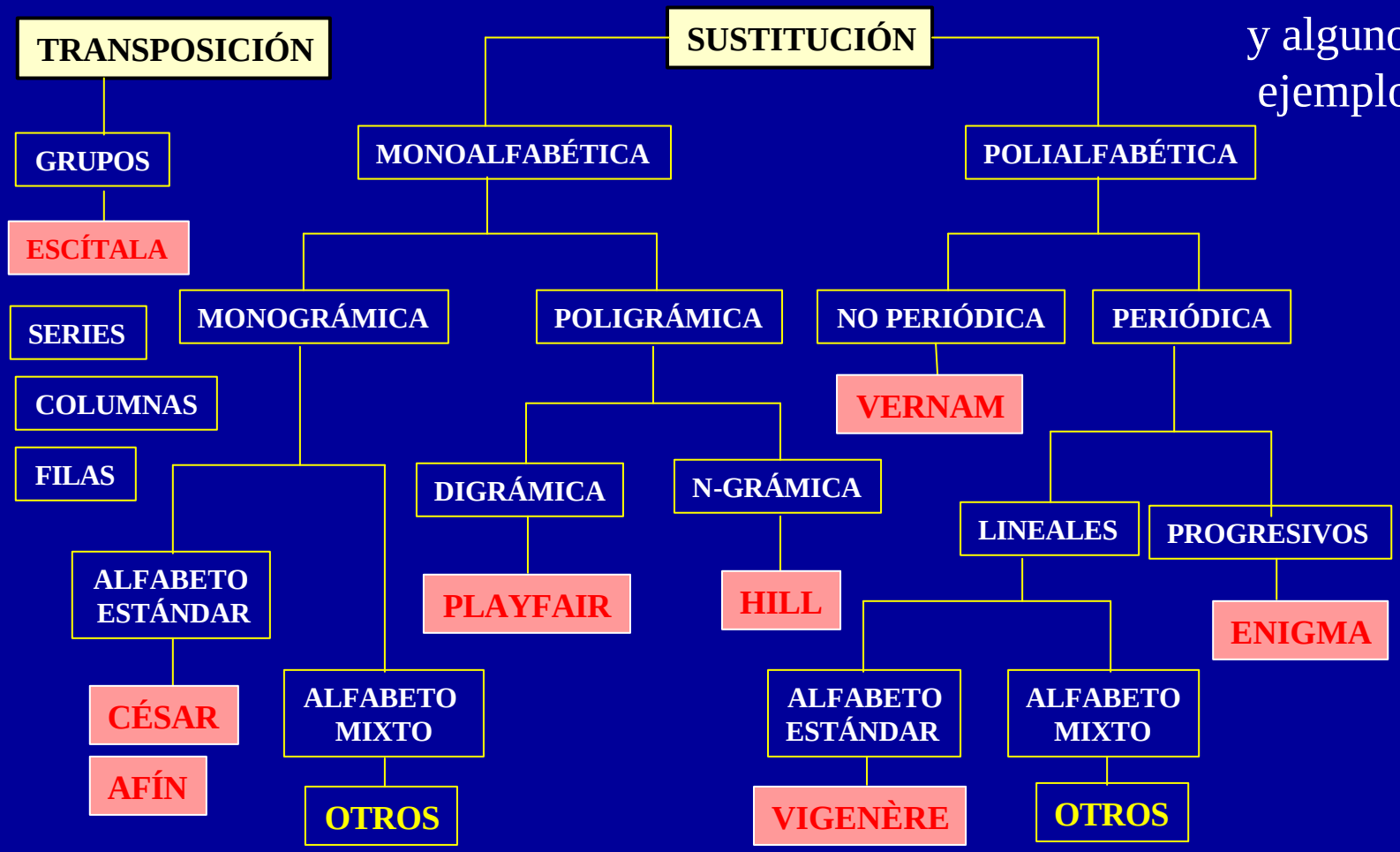
- La criptografía es casi tan antigua como las primeras civilizaciones de nuestro planeta.
- Ya en el siglo V a.C. se usaban técnicas de cifrado para proteger la información.
- Se pretendía garantizar sólo la ***confidencialidad*** y la ***autenticidad*** de los mensajes.
- Los mayores avances se lograron en la Segunda Guerra Mundial: los países en conflicto tenían un gran número de técnicos encargados de romper los mensajes cifrados de los teletipos.

Herramientas de la Criptografía Clásica

Tanto máquinas (artilugios de cifrado) como los algoritmos que trabajaban matemáticamente dentro de un cuerpo finito n, hacen uso de dos técnicas básicas orientadas a caracteres y que, muchos siglos después, propone Shannon:

- **Sustitución**: un carácter o letra se modifica o sustituye por otro elemento en el cifrado.
- **Transposición**: los caracteres o letras del mensaje se redistribuyen sin modificarlos y según ciertas reglas, dentro del criptograma.

Clasificación de los Criptosistemas Clásicos



Hitos históricos en la Criptografía

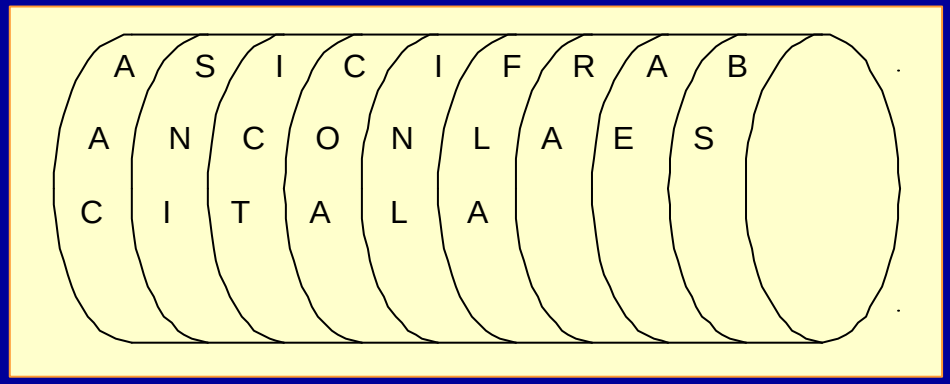
- La criptografía clásica abarca desde tiempos inmemoriales hasta la mitad del siglo XX.
- El punto de inflexión en esta clasificación la marcan tres hechos relevantes:
  - En el año 1948 se publica el estudio de Claude Shannon sobre la Teoría de la Información.
  - En 1974 aparece el estándar de cifrado DES.
  - En el año 1976 se publica el estudio realizado por W. Diffie y M. Hellman sobre la aplicación de funciones matemáticas de un solo sentido a un modelo de cifrado, denominado cifrado con clave pública.

CIFRADO DIGITAL

El cifrador de la Escítala

- La escítala era usada en el siglo V a.C. por el pueblo griego. Consistía en un bastón en el que se enrollaba una cinta y luego se escribía en ella el mensaje de forma longitudinal.
- Al desenrollar la cinta, las letras aparecían sin orden alguno.
- La única posibilidad de “deshacer” este cifrado pasaba por enrollar dicha cinta en un bastón con el mismo diámetro que el usado en el extremo emisor y leer el mensaje de forma longitudinal.

Método de Cifrado de la Escítala



Se trata de un sistema de cifrado por transposición

El texto plano es:

M = ASI CIFRABAN CON LA ESCITALA

El texto cifrado o criptograma será:

C = AAC SNI ICT COA INL FLA RA AE BS

Primer cifrador por Sustitución: Polybios

Es el cifrador por sustitución de caracteres más antiguo que se conoce (siglo II a.C.) pero duplica el tamaño de M (propiedad no muy interesante).

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

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

M_1 = QUÉ BUENA IDEA

C_1 = DA DE AE AB DE AE

CC AA BD AD AE EA

M_2 = LA DEL GRIEGO

C_2 = 31 11 14 15 31 22

42 24 15 22 34

El cifrador del César

En el siglo I a.d.C., Julio César presenta este cifrador cuyo algoritmo consiste en el desplazamiento de tres lugares hacia la derecha de los caracteres del texto plano. Es un cifrador por sustitución monoalfabético en el que las operaciones se realizan módulo n, siendo n igual al número de elementos del alfabeto (latín).

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
M_i	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
C_i	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Alfabeto de cifrado del César para castellano mod 27

Ejemplo de cifrado con cifrador del César

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

$$C_i = M_i + 3 \bmod 27$$

M = EL PATIO DE MI CASA ES PARTICULAR

C = HÑ SDWLR GH OL FDVD HV SDUWLFXÑDU

Cada letra se cifrará siempre igual. Es una gran debilidad y hace que este sistema sea muy vulnerable y fácil de atacar simplemente usando las estadísticas del lenguaje.

Criptoanálisis del Cifrador por Sustitución

A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

Cifrado: $C_i = (M_i + b) \bmod 27$ Descifrado: $M_i = (C_i - b) \bmod 27$

La letra más frecuente del criptograma la hacemos coincidir con la más frecuente del lenguaje, la letra E, y encontramos así b.

C = LZAHL ZBTHW YBLIH XBLKL ILYOH ZLYCH ROKH

Frecuencias observadas en el criptograma: L (7); H (6); Z (3); B (3); Y (3); I (2); K (2); O (2); A (1); T (1); W (1); X (1); C (1); R (1).

Luego, es posible que la letra E del lenguaje (la más frecuente) se cifre como L en el criptograma y que la letra A se cifre como H:

E + b mod 27 = L $\Rightarrow$ b = L - E mod 27 = 11 – 4 mod 27 = 7

A + b mod 27 = H $\Rightarrow$ b = H - A mod 27 = 7 – 0 mod 27 = 7

M = ESTA ES UNA PRUEBA QUE DEBERIA SER VALIDA

Cifrador por sustitución afín mod 27

Cifrado: $C_i = a*M_i + b \bmod 27$

Descifrado: $M_i = (C_i - b) * a^{-1} \bmod 27$ donde $a^{-1} = \text{inv}(a, 27)$

El factor de decimación a deberá ser primo relativo con el cuerpo n (en este caso 27) para que exista el inverso.

El factor de desplazamiento puede ser cualquiera $0 \leq b \leq 27$.

El ataque a este sistema es también muy elemental. Se relaciona el elemento más frecuente del criptograma a la letra E y el segundo a la letra A, planteando un sistema de 2 ecuaciones. Si el texto tiene varias decenas de caracteres este ataque prospera; caso contrario, puede haber ligeros cambios en esta distribución de frecuencias.

El cifrador de Vigenère

Este cifrador polialfabético soluciona la debilidad del cifrado del César de que una letra se cifre siempre igual. Usa una clave K de longitud L y cifra carácter a carácter sumando módulo n el texto en claro con los elementos de esta clave.

$$C_i = M_i + K_i \bmod 27$$

Sea K = CIFRA y el mensaje M = HOLA AMIGOS

M =	H	O	L	A	A	M	I	G	O	S
K =	C	I	F	R	A	C	I	F	R	A
C =	J	W	P	R	A	Ñ	P	L	G	S

sumando mod 27...

Más de un alfabeto: la letra O se cifra de forma distinta.

Observe que el criptograma P se obtiene de un texto L y de un texto I.

¿Es Vigenère un algoritmo seguro?

Si la clave de Vigenère tiene más de 6 caracteres distintos, se logra una distribución de frecuencias en el criptograma del tipo normal, es decir más o menos plana, por lo que se difumina la redundancia del lenguaje.

Aunque pudiera parecer que usando una clave larga y de muchos caracteres distintos y por tanto varios alfabetos de cifrado, Vigenère es un sistema de cifra seguro, esto es falso.

La redundancia del lenguaje unido a técnicas de criptoanálisis muy sencillas, como los métodos de Kasiski y del Índice de Coincidencia, permiten romper el cifrado y la clave de una manera muy fácil y con mínimos recursos.

Veamos un ataque por el método de Kasiski.

Ataque por el método de Kasiski

- El método de Kasiski consiste en buscar repeticiones de cadenas de caracteres en el criptograma. Si estas cadenas son mayores o iguales a tres caracteres y se repiten más de una vez, lo más probable es que esto se deba a cadenas típicas del texto en claro (trigramas, tetragramas, etc., muy comunes) que se han cifrado con una misma porción de la clave.
- Si se detectan estas cadenas, la distancia entre las mismas será múltiplo de la longitud de la clave. Luego, el máximo común divisor entre esas cadenas es un candidato a ser la longitud de la clave, digamos L.
- Dividimos el criptograma en L subcriptogramas que entonces han sido cifrados por una misma letra de la clave y en cada subcriptograma hacemos un ataque simple ahora de tipo estadístico monoalfabético.
- La idea es buscar ahora a través de los tres caracteres más frecuentes en cada subcriptograma las posiciones relativas de las letras A, E y O que en castellano están separadas por 4 y 11 lugares. La letra de la posición que ocupe la letra A (A = 0) será entonces la letra clave correspondiente.

Cadenas repetidas en ataque de Kasiski

Sea el criptograma C de 404 caracteres que vamos a criptoanalizar el siguiente:

```
PBVRQ VICAD SKAÑS DETSJ PSIED BGGMP SLRPW RÑPWY EDSDE ÑDRDP CRCPQ MNPWK
UBZVS FNVRD MTIPW UEQVV CBOVN UEDIF QLONM WNUVR SEIKA ZYEAC EYEDS ETFPH
LBHGU ÑESOM EHLBX VAEEP UÑELI SEVEF WHUNM CLPQP MBRRN BPVIÑ MTIBV VEÑID
ANSJA MTJOK MDODS ELPWI UFOZM QMVNF OHASE SRJWR SFQCO TWVMB JGRPW VSUEX
INQRS JEUEM GGRBD GNNIL AGSJI DSVSU EEINT GRUEE TFGGM PORDF OGTSS TOSEQ
OÑTGR RYVLP WJIFW XOTGG RPQRR JSKET XRNBL ZETGG NEMUO TXJAT ORVJH RSFHV
NUEJI BCHAS EHEUE UOTIE FFGYA TGGMP IKTBW UEÑEN IEEU.
```

Entre otras, se observan las siguientes cadenas (subrayadas) en el criptograma:

- 3 cadenas GGMP, separadas por 256 y 104 posiciones.
- 2 cadenas YEDS, separadas por 72 espacios.
- 2 cadenas HASE, separadas por 156 espacios.
- 2 cadenas VSUE, separadas por 32 espacios.

Luego el período de la clave puede ser mcd (256, 104, 72, 156, 32) = 4. La clave tendrá cuatro caracteres, por lo tanto tomaremos del criptograma el carácter 1º, el 5º, el 9º, etc. para formar el primer subcriptograma C_A; luego el 2º, el 6º, el 10º, etc. para formar el subcriptograma C_B, y así hasta el subcriptograma C_D.

Paso a cifrado monoalfabético en Kasiski

Tenemos 4 subcriptogramas que han sido cifrados con la misma letra de la clave:

C_A = PQAAEPDMRÑEEDCNUSRIECNIONSAAETLUOLAUIEULMNIIEAAOOLU
MNARSOMRSISERNAISIRTMDTOORLIORRENENOAVSNIAEOFAMTEI

C_B = BVDÑTSBPPPDÑPPPBFDPQBUFNUEZCDFBÑMBEÑSFNPBBÑBÑNMKDPF
QFSJFTBPUNJMBNGDUNUFPFSSÑRPFTPJTBTETTJFUBSUTFTPBÑE

C_C = VISSSIGSWWSDCQWZNMWVOEQMVIYESPHEEXEEEWMQRPMVISTMSWO
MOEWQWJWEQEGDISSETEGOOSETYWWGQSXLGMXOHHECEEIGGIWEE

C_D = RCKDJEGLRYDRRMKVVTUVVDLWRKEYEHGSHVPLVHCPRVTVDJJDEIZ
VHSRCVGVXRUGGLJVEGEGRGTQGVJXGRKRZGUJRRVJHHUEYGKUNU

La frecuencia relativa observada en cada uno de los subcriptogramas es:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
C_A	11	0	2	3	12	1	0	0	11	0	0	5	6	9	1	10	2	1	9	7	4	5	1	0	0	0	0
C_B	0	14	1	6	4	12	1	0	0	4	1	0	3	6	8	6	14	2	1	6	9	7	1	0	0	0	1
C_C	0	0	1	2	18	0	7	3	7	1	0	1	7	1	0	0	2	6	1	12	3	0	3	12	3	2	1
C_D	0	0	3	5	7	0	12	6	1	7	5	4	1	1	0	6	2	1	13	2	3	7	14	0	2	3	2

La letra más frecuente del subcriptograma debería corresponder a la letra E del texto en claro, la segunda a la letra A y la tercera a la letra O.

La regla AEO en el ataque de Kasiski

- Si la posición relativa de la letra A es el valor 0, entonces la letra E está cuatro espacios a la derecha de la A (m+4 mod 27) y la letra O está 15 espacios a la derecha de la letra A (m+15 mod 27).
- Buscaremos en cada subcriptograma C_i las tres letras más frecuentes y que cumplan además con esta distribución.
- Es suficiente contar con estas tres letras para que el ataque prospere. No obstante, podemos afinar más el ataque si tomamos en cuenta la siguiente letra frecuente en castellano (S) en posición (m+19) mod 27.

En el ejemplo para C_A se observa que la única solución que cumple con esto es la que coincide la AEO (11, 12, 10) luego la letra clave sería la A. Para C_B elegimos BFP (14, 12, 14) por lo que la letra clave sería B. Para C_C elegimos EIS (18, 7, 12) por lo que la letra clave sería E. Para C_D elegimos RVG (13, 14, 12) por lo que la letra clave sería R.

La clave será K = ABER y M = "Para que la cosa no me sorprenda...". ✌

El índice de coincidencia IC

Cuando encontramos una longitud L de la clave por el método de Kasiski y rompemos el criptogama en L subcriptogamas, podemos comprobar que cada uno de ellos se trata efectivamente de un cifrado monoalfabético aplicando el concepto del índice de coincidencia IC.

$$IC = \sum_{i=0}^{26} p_i^{\,2}$$ para castellano mod 27: $IC = p_A^2 + p_B^2 + ... + p_Z^2 = 0{,}072$

Aunque el estudio de este índice IC queda fuera del contexto de este curso, como para el castellano mod 27 el IC = 0,072, en el ataque de Kasiski se comprueba que para cada subcriptograma su IC esté cercano a este valor. Si el IC es menor que 0,5 es muy probable que no estemos ante un cifrador monoalfabético sino uno polialfabético de periodo 2 o mayor.

En el ejemplo anterior, una vez roto el criptograma en cuatro tenemos: $IC_{CA} = 0{,}070$; $IC_{CB} = 0{,}073$; $IC_{CC} = 0{,}075$; $IC_{CD} = 0{,}065$.

Cifrador poligrámico de Playfair

Los cifrados anteriores trabajan carácter a carácter, es decir son monográmicos. Para aumentar la seguridad del cifrado podemos cifrar por poligramas, bloques de caracteres.

Un cifrador inventado a finales del siglo XIX es el de Playfair que trabaja con una matriz de 5x5 letras, cifrando por digramas. Si el texto plano tiene un número impar de elementos, se rellena con una letra predeterminada, e.g. x.

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

- Si M_1M_2 están en la misma fila, C_1C_2 son los dos caracteres de la derecha.
- Si M_1M_2 están en la misma columna, C_1C_2 son los dos caracteres de abajo.
- Si M_1M_2 están en filas y columnas distintas, C_1C_2 son los dos caracteres de la diagonal, desde la fila de M_1.

Ejemplo de cifrado con Playfair

Si la clave K = BEATLES, eliminando la letra Ñ, se pide cifrar el mensaje M = With a little help from my friends.

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

Se rompe la doble MM agregando una X y se rellena al final con X

M = WI TH AL IT TL EH EL PF RO MX MY FR IE ND SX
C = EP BM TB ME LB BI AB RC UP KY RT MY PC KG DV

Estos sistemas también son criptoanalizables pues en el criptograma C persisten algunas propiedades del lenguaje, en este caso la distribución de digramas típicos del castellano como por ejemplo en, de, mb, nv, tr, mp, etc.

El cifrador de matrices de Hill

En 1929 Lester Hill propone un sistema de cifrado usando una matriz como clave, cifrando Ngramas de forma que:

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ .. \\ C_N \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} & \dots & k_{1N} \\ k_{21} & k_{22} & k_{23} & \dots & k_{2N} \\ k_{31} & k_{32} & k_{33} & \dots & k_{3N} \\ .. & .. & .. & & .. \\ k_{N1} & k_{N2} & k_{N3} & \dots & k_{NN} \end{pmatrix} X \begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ .. \\ M_N \end{pmatrix} \quad \text{mod n}$$

La matriz clave K debe tener inversa K^{-1} en el cuerpo de cifra n. Luego, como $K^{-1} = T_{ADJ(K)}/|K|$ mod n, en donde ADJ(K) es la matriz adjunta, T es la traspuesta y |K| el determinante, este último valor |K| no podrá ser cero ni tener factores en común con n puesto que está en el denominador (concepto de inverso).

Si el texto plano no es múltiplo del bloque N, se rellena con caracteres predeterminados, por ejemplo la letra X o la Z.

Ejemplo de cifrado de Hill

Sea M = AMIGO CONDUCTOR y la clave K la que se muestra:

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 16 & 4 & 11 \\ 8 & 6 & 18 \\ 15 & 19 & 15 \end{pmatrix} \times \begin{pmatrix} 0 \\ 12 \\ 8 \end{pmatrix} \bmod 27$$

K = PELIGROSO será la clave simbólica. Se cifrará el primer trigrama: AMI = 0, 12, 8.

M = AMI GOC OND UCT ORZ

$C_1 = (16*0 + 4*12 + 11*8) \bmod 27 = 136 \bmod 27 = 1 = B$

$C_2 = (8*0 + 6*12 + 18*8) \bmod 27 = 216 \bmod 27 = 0 = A$

$C_3 = (15*0 + 19*12 + 15*8) \bmod 27 = 348 \bmod 27 = 24 = X$

C = BAX PMA BJE XAF EUM (compruebe Ud. los otros trigramas)

Para descifrar encontramos $K^{-1} = \text{inv}(K, 27) = K^{-1} = T_{ADJ(K)}/|K| \bmod n$

$|K| = 16(6*15 - 19*18) - 4(8*15 - 15*18) + 11(8*19 - 15*6) \bmod 27 = 4$

Encontramos luego la matriz adjunta de K, la trasponemos cambiando filas por columnas y la multiplicamos por inv (|K|, 27) = inv (4, 27) = 7 con lo que se obtiene la matriz que se indica (hágalo Ud.)

Ejemplo de descifrado de Hill

$$\begin{pmatrix} M \end{pmatrix} = \begin{pmatrix} K^{-1} \end{pmatrix} \text{ x } \begin{pmatrix} C \end{pmatrix} \bmod n \quad \text{y} \quad K^{-1} = \begin{pmatrix} 18 & 26 & 15 \\ 24 & 6 & 13 \\ 11 & 24 & 10 \end{pmatrix}$$

C = BAXPMABJEXAFEUM y la clave K^{-1} es la que se muestra:

$$\begin{pmatrix} M_1 \\ M_2 \\ M_3 \end{pmatrix} = \begin{pmatrix} 18 & 26 & 15 \\ 24 & 6 & 13 \\ 11 & 24 & 10 \end{pmatrix} \text{ x } \begin{pmatrix} 1 \\ 0 \\ 24 \end{pmatrix} \bmod 27$$

Descifrado del primer trigrama del criptograma: BAX = 1, 0, 24.

C = BAX PMA BJE XAF EUM

$M_1 = (18*1 + 26*0 + 15*24) \bmod 27 = 378 \bmod 27 = 0 = A$

$M_2 = (24*1 + 6*0 + 13*24) \bmod 27 = 336 \bmod 27 = 12 = M$

$M_3 = (11*1 + 24*0 + 10*24) \bmod 27 = 251 \bmod 27 = 8 = I$

M = AMI GOC OND UCT ORZ (compruebe Ud. los otros trigramas)

¿Es seguro el cifrador de Hill?

Si con el sistema de Hill se cifran bloques de 8 caracteres, incluso en un cuerpo tan pequeño como n = 27 el espacio de claves aumenta de forma espectacular, comparable con DES.

Si el módulo de cifra es un primo p, entonces el número de claves válidas es cercano al máximo posible: p^x donde $x = d^2$, con d el tamaño de N-grama o de la matriz clave.

No obstante, el sistema no es seguro. Debido a su linealidad será muy fácil hacer un ataque con texto plano conocido según el método de Gauss Jordan y encontrar así la matriz clave K. Esto es debido a que aparecen los llamados vectores unitarios en el criptograma o en el texto plano, o bien los obtenemos aplicando este método.

Ataque al cifrado de Hill por Gauss Jordan

El método consiste en escribir una matriz 2N-grámica con los elementos del texto en plano y los elementos del criptograma. En esta matriz realizamos operaciones lineales (multiplicar filas por un número y restar filas entre sí) con el objeto de obtener los vectores unitarios.

Por ejemplo podemos romper la matriz clave K teniendo:

```
M = ENU NLU GAR DEL AMA NCH ADE CUY ONO ...
C = WVX IDQ DDO ITQ JGO GJI YMG FVC UÑT ...
```

$$\left(\begin{array}{ccc|ccc} E & N & U & W & V & X \\ N & L & U & I & D & Q \\ G & A & R & D & D & O \\ D & E & L & I & T & Q \\ A & M & A & J & G & O \\ N & C & H & G & J & I \\ A & D & E & Y & M & G \\ C & U & Y & F & V & C \\ O & N & O & U & Ñ & T \end{array}\right) = \left(\begin{array}{ccc|ccc} 4 & 13 & 21 & 23 & 22 & 24 \\ 13 & 11 & 21 & 8 & 3 & 17 \\ 6 & 0 & 18 & 3 & 3 & 15 \\ 3 & 4 & 11 & 8 & 20 & 17 \\ 0 & 12 & 0 & 9 & 6 & 15 \\ 13 & 2 & 7 & 6 & 9 & 8 \\ 0 & 3 & 4 & 25 & 12 & 6 \\ 2 & 21 & 25 & 5 & 22 & 2 \\ 15 & 13 & 15 & 21 & 14 & 20 \end{array}\right)$$

Operaciones en la matriz de Gauss Jordan

Vamos a dejar en la primera columna un número uno en la fila primera y todas las demás filas un cero. Luego multiplicamos el vector (4 13 21 | 23 22 24) por el inv (4, 27) = 7. Así obtenemos 7(4 13 21 | 23 22 24) mod 27 = (1 10 12 | 26 19 6). Si esto no se puede hacer con la primera fila movemos los vectores. Hecho esto vamos restando las filas respecto de esta primera como se indica:

4	13	21	23	22	24
13	11	21	8	3	17
6	0	18	3	3	15
3	4	11	8	20	17
0	12	0	9	6	15
13	2	7	6	9	8
0	3	4	25	12	6
2	21	25	5	22	2
15	13	15	21	14	20

a) 2ª fila = 2ª fila – 13∗1ª fila mod 27
b) 3ª fila = 3ª fila – 6∗1ª fila mod 27
c) 4ª fila = 4ª fila – 3∗1ª fila mod 27
d) 5ª fila ya tiene un 0
e) 6ª fila = 6ª fila – 13∗1ª fila mod 27
f) 7ª fila ya tiene un 0
g) 8ª fila = 8ª fila – 2∗1ª fila mod 27
h) 9ª fila = 9ª fila – 15∗1ª fila mod 27

Matriz clave de Hill criptoanalizada

Repetimos este procedimiento ahora para algún vector en cuya segunda columna tenga un número con inverso en 27 y lo mismo para la tercera columna, moviendo si es preciso los vectores.

Como la mitad izquierda de la matriz 2N era el texto el plano, la parte derecha de la matriz con vectores unitarios corresponderá a la traspuesta de la clave.

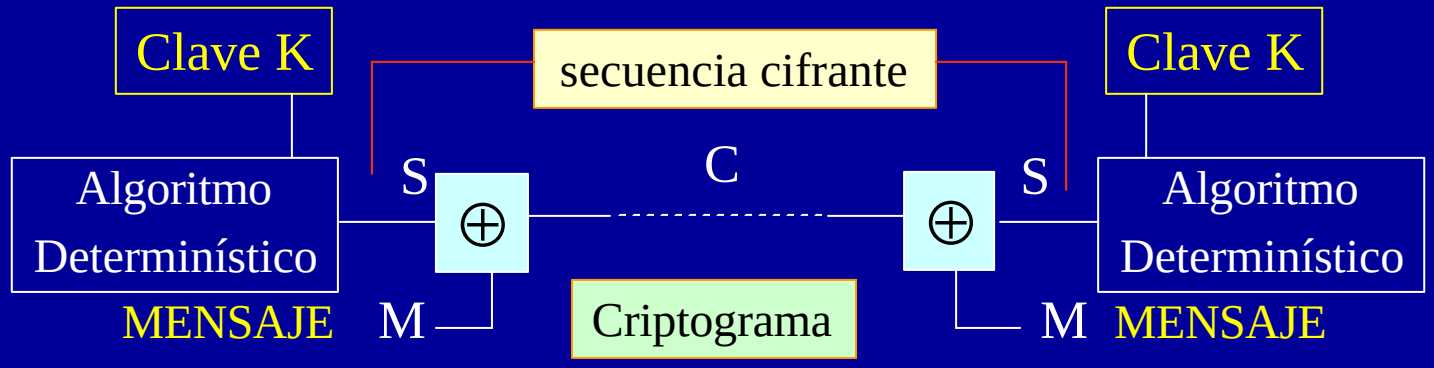
$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 5 & 7 \\ 0 & 1 & 0 & 3 & 5 & 8 \\ 0 & 0 & 1 & 4 & 6 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \Rightarrow \quad K = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Compruebe que la clave es la utilizada en este cifrado.

El cifrador de Vernam

En 1917 Gilbert Vernam (MIT) propone un cifrador por sustitución binaria con clave de un solo uso, basado en el código Baudot de 5 bits:

- La operación de cifrado es la función XOR.
- Usa una secuencia cifrante binaria y aleatoria S que se obtiene de una clave secreta K compartida por emisor y receptor.
- El algoritmo de descifrado es igual al de cifrado por la involución de la función XOR.
- La clave será tan larga o más que el mensaje y se usará una sola vez.



Ejemplo de cifrado de Vernam

Usando el código Baudot (véase próxima diapositiva) se pide cifrar el mensaje M = BYTES con la clave K = VERNAM.

Solución:

B⊕V = 11001⊕11110 = 00111 = U

Y⊕E = 10101⊕00001 = 10100 = H

T⊕R = 10000⊕01010 = 11010 = G

E⊕N = 00001⊕01100 = 01101 = F

S⊕A = 00101⊕00011 = 00110 = I

C = UHGFI

El sistema de Vernam es el único que es matemáticamente seguro e imposible de criptoanalizar ya que la clave se usa una sola vez (*one time pad*), es aleatoria y tanto o más larga que el propio mensaje.

Código Baudot (cifrador de Vernam)

Código Binario	Carácter	Código Binario	Carácter
00000	Blanco	10000	T
00001	E	10001	Z
00010	=	10010	L
00011	A	10011	W
00100	Espacio	10100	H
00101	S	10101	Y
00110	I	10110	P
00111	U	10111	Q
01000	<	11000	O
01001	D	11001	B
01010	R	11010	G
01011	J	11011	↑
01100	N	11100	M
01101	F	11101	X
01110	C	11110	V
01111	K	11111	↓

Técnicas de Transposición (1)

- Hasta este punto analizamos sustitución.
- En esta técnica se permutan los caracteres de M.

Técnica más simple: Rail Fence (prof. 2)

M = LAS VENTANAS DEL SR PUERTAS SON FEAS

L S E T N S E S P E T S O F A

A V N A A D L R U R A S N E S

C = LSETNSESPETSOFAAVNAADLRURASNES

Trivial!

Técnicas de Transposición (2)

- Mejora utilizando una matriz: aparece clave.
- M = LAS VENTANAS DEL SR PUERTAS SON FEAS

K= 4 2 3 1 6 5

M=

L	A	S	V	E	N
T	A	N	A	S	D
E	L	S	R	P	U
E	R	T	A	S	S
O	N	F	E	A	S

¿Sigue siendo Simple?

Si, pues se conserva la frecuencia de cada letra de M

C = VARAEAALRNSNSTFLTEEONDUSSESPSA

Técnicas de Transposición (3)

C = VARAEAALRNSNSTFLTEEONDUSSESPSA

El C anterior es el nuevo M y se repite el algoritmo anterior.

K= 4 2 3 1 6 5

M=
V	A	R	A	E	A
A	L	R	N	S	N
S	T	F	L	T	E
E	O	N	D	U	S
S	E	S	P	S	A

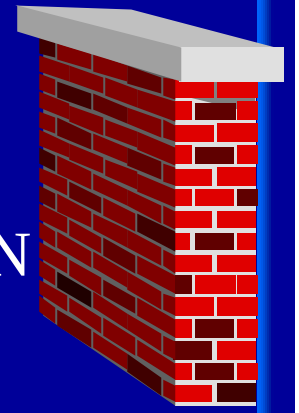
Es claro que la frecuencia de cada letra de M **no** puede variar, pero se trata de una permutación mucho más compleja (menos desestructurada y más dificil para criptoanalizar).

C= ANLDPALTOERRFNSVASESANESAESTUS

En Construcción...

ESTIMADO/A ALUMNO/A:

ESTA PARTE DE LAS TRANSPARENCIAS ESTARÁ EN CONSTRUCCIÓN DURANTE ALGÚN TIEMPO... ☹ ... mis disculpas!



NO OBSTANTE, SEGUIREMOS CON TÉCNICAS DE CIFRADO MODERNAS, DE MAYOR INTERÉS ACADÉMICO. ☺

En el CeCom pueden encontrarse detalles de otros Métodos Clásicos.
Véanse papers (de mi web) + Stallings (Ch. 2) + Pfleeger (Ch.2)

JAVIER
ECHAIZ

Criptografía

Criptosistemas Modernos

Clase 6 (sigue)

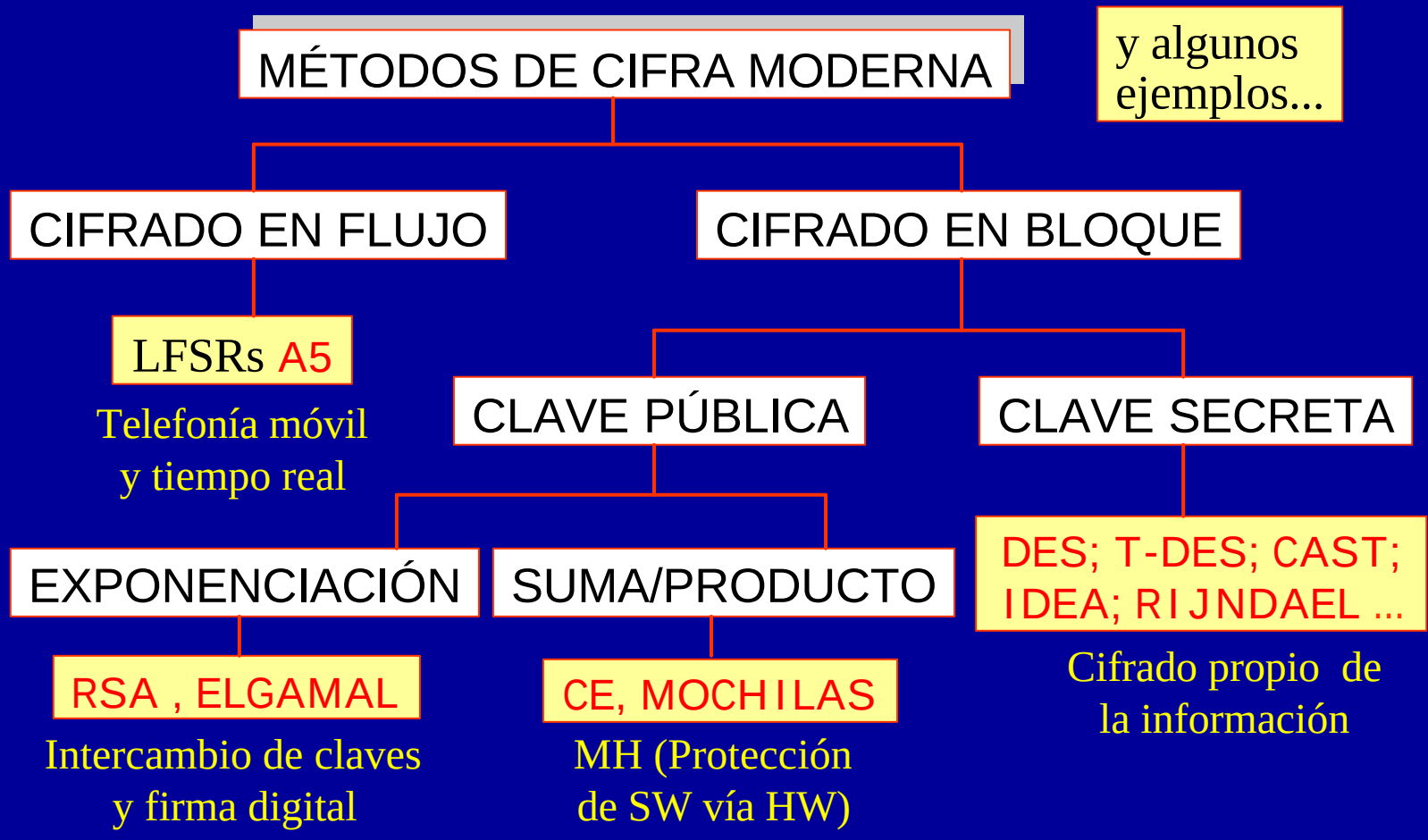
Prof. Javier Echaiz
D.C.I.C. – U.N.S.
http://cs.uns.edu.ar/~jechaiz
je@cs.uns.edu.ar

Conceptos Elementales

Un par de ideas básicas

- Los criptosistemas modernos, cuyo cifrado en bits está orientado a (todos) los caracteres ASCII o ANSI usan por lo general una operación algebraica en Z_n, un cuerpo finito, sin que necesariamente este módulo deba corresponder con el número de elementos del alfabeto o código utilizado. Es más, nunca coinciden; siempre será mucho mayor el cuerpo de trabajo que el alfabeto.
- Su fortaleza está en la imposibilidad computacional de descubrir una clave secreta única, en tanto que el algoritmo de cifrado es (o debería ser) público.

Clasificación de los Criptosistemas



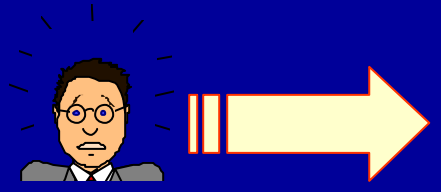
Introducción al Cifrado de Flujo

Usa el concepto de cifrado propuesto por Vernam, que cumple con las ideas de Shannon sobre sistemas de cifrado secreto perfecto, esto es:

a) El espacio de las claves es igual o mayor que el espacio de los mensajes.

b) Las claves deben ser equiprobables.

c) La secuencia de clave se usa una sola vez y luego se destruye (sistema *one-time pad*).



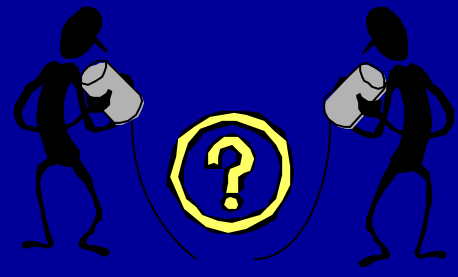
DUDA: ¿Es posible satisfacer la condición **a)**?

Espacio de Claves y del Mensaje

¿Espacio de Claves ≥ Espacio de Mensajes?

1) La secuencia de bits de la clave deberá enviarse al destinatario a través de un canal que sabemos es inseguro (aún no conocemos el protocolo de intercambio de clave de Diffie y Hellman).
2) Si la secuencia es “infinita”, desbordaríamos la capacidad del canal de comunicaciones.

¿Qué solución podemos dar a este problema?



Concepto de Semilla (*seed*)

Si por el canal supuestamente seguro enviamos esa clave tan larga ... ¿por qué entonces no enviamos directamente el mensaje plano y *nos dejamos de historias*? ☺

La solución está en generar una secuencia de tipo pseudoaleatoria con un algoritmo determinístico a partir de una semilla de sólo unos centenares de bits. Podremos generar así secuencias con períodos del orden de 2^n, un valor ciertamente muy alto. Esta semilla es la que se envía al receptor mediante un sistema de cifrado de clave pública y un algoritmo de intercambio de clave y no sobrecargamos el canal.

Técnica de Cifrado en Flujo *(stream)*

- ✓ El mensaje plano se leerá bit a bit.
- ✓ Se realizará una operación de cifrado, normalmente la función XOR, con una secuencia cifrante de bits S_i que debe cumplir ciertas condiciones:
  - Un período muy alto.
  - Aleatoriedad en sus propiedades.

Lo veremos en la próxima clase...

Introducción al Cifrado en Bloque

El mensaje se agrupa en bloques, normalmente de 8 bytes, antes de aplicar el algoritmo de cifrado a cada bloque de forma independiente con la misma clave.

Cifrado con Clave Secreta

Hay algunos algoritmos muy conocidos por su uso en aplicaciones bancarias (DES), correo electrónico (IDEA, CAST) y en comercio electrónico (T-DES).

No obstante, tienen tres puntos débiles.

Debilidades del Cifrado con Clave Secreta

a) Mala gestión de claves. Crece el número de claves secretas en un orden igual a n^2. Muy grave para un valor n grande de usuarios 👎.

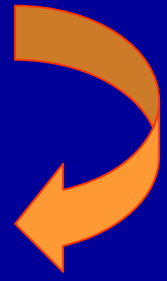
b) Mala distribución de claves. No existe posibilidad de enviar, de forma segura, una clave a través de un medio inseguro 👎.

c) No tiene firma digital. Aunque sí será posible autentificar el mensaje mediante una marca, no es posible firmar digitalmente el mensaje 👎.

¿Por qué Usamos Clave Secreta?

a) Mala gestión de claves 👎.
b) Mala distribución de claves 👎.
c) No tiene firma digital 👎.

¿Tiene algo de bueno el cifrado en bloque con clave secreta?



Sí: la velocidad de cifrado es muy alta 👍

Cifrado en Bloque con Clave Pública

Cifrado con Clave Pública

- Comienza a ser ampliamente conocido a través de su aplicación en los sistemas de correo electrónico seguro (PGP y PEM) al permitir incluir una firma digital adjunta al documento o e-mail enviado.
- Cada usuario tiene dos claves, una secreta o privada y otra pública, inversas dentro de un cuerpo.
- Usan las funciones unidireccionales con trampa.

Funciones Unidireccionales con Trampa

Son funciones matemáticas de un solo sentido (*one-way functions*) y que nos permiten usar la función en sentido directo o de cálculo fácil para cifrar y descifrar (usuarios legítimos) y fuerza el sentido inverso o de cálculo difícil para aquellos (impostores, hackers, etc.) si lo que se desea es atacar o criptoanalizar la cifra.

$f(M) = C$ *siempre es fácil.*

$f^{-1}(C) = M$ *es difícil salvo si se tiene la trampa.*

Funciones con Trampa Típicas (1)

Problema de la factorización

Cálculo Directo: Producto de dos primos grandes **p*q = n**
Cálculo Inverso: Factorización de número grande **n = p*q**

Problema del logaritmo discreto

Cálculo Directo: Exponenciación discreta **$b = a^x \bmod n$**
Cálculo Inverso: Logaritmo discreto **$x = \log_a b \bmod n$**

Funciones con Trampa Típicas (2)

Problema de la mochila

Cálculo Directo: Sumar elementos de mochila con trampa
Cálculo Inverso: Sumar elementos de mochila sin trampa

Problema de la raíz discreta

Cálculo Directo: Cuadrado discreto **x = a*a mod n**
Cálculo Inverso: Raíz cuadrada discreta **n =** Ö**a mod n**

Cifrado con Clave Pública de Destino

NUESTROS PROTAGONISTAS

origen

Mateico

Claves: e_B, n_B, d_B

e_B, n_B: públicas

d_B: privada

e_B y d_B son inversas dentro de un cuerpo Z

Si Mateico realiza la operación con las claves públicas de LaSole (e_A, n_A), la información que se transmite mantiene la confidencialidad: sólo ella puede verla.

$C = E_{eA}(M) \bmod n_A$

destino

LaSole

Claves: e_A, n_A, d_A

e_A, n_A: públicas

d_A: privada

e_A y d_A son inversas dentro de un cuerpo Z

Cifrado con Clave de Destino

Cifrado:

Mateico envía un mensaje M a LaSole

Mateico

Claves: e_B, n_B, d_B

$C = E_{eA}(M) \bmod n_A$

LaSole

Claves: e_A, n_A, d_A

Claves públicas

Clave privada

Descifrado:

$M = E_{dA}[E_{eA}(M)] \bmod n_A$

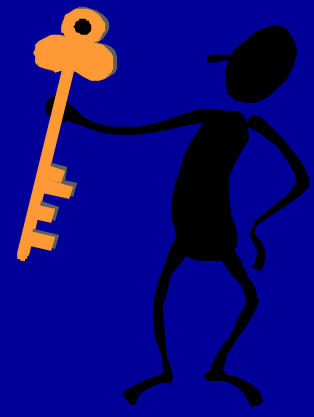
E_{dA} y E_{eA} son inversos

Recupera el texto plano: ***confidencialidad***

¿Y si usamos la Clave Pública de Origen?

Si en vez de utilizar la clave pública de destino, el emisor usa su propia clave pública, el cifrado no tiene sentido bajo el punto de vista de sistemas de clave pública ya que sólo él o ella sería capaz de descifrar el criptograma (deshacer la operación de cifrado) con su propia clave privada.

Esto podría usarse para cifrar de forma local uno o varios ficheros, por ejemplo, pero para ello ya están los sistemas de clave secreta, mucho más rápidos y, por tanto, más eficientes.



¿Y si usamos la Clave Privada de Origen?

Si ahora el emisor usa su clave privada en el cifrado sobre el mensaje, se obtiene una firma digital que lo autentifica como emisor ante el destinatario y, además, a este último le permitirá comprobar la integridad del mensaje.

Veamos antes un ejemplo de algoritmo que usa dos claves

Obviamente, el emisor nunca podrá realizar el cifrado del mensaje M con la clave privada del receptor. ☺

El algoritmo del Mensaje en la Caja

PROTOCOLO: **A** envía a **B** un mensaje **M**

1 A pone el mensaje M en la caja, la cierra con su llave y la envía a B.

2 B recibe la caja, la cierra con su llave y envía a A la caja con las dos cerraduras .

3 A recibe la caja, quita su llave y devuelve a B la caja sólo con la cerradura de B .

4 B recibe la caja, quita su cerradura y puede ver el mensaje M que A puso en el interior de la caja.

¿Todo OK en el Algoritmo de la Caja?

Durante la transmisión, el mensaje está protegido de cualquier intruso por lo que existe integridad del mensaje y hay protección contra una ataque pasivo.

El usuario B no puede estar seguro si quien le ha enviado el mensaje M es el usuario A o un impostor. El algoritmo no permite comprobar la autenticidad del emisor pues no detecta "sustitución" de identidad.

Una ligera modificación del algoritmo anterior nos permitirá cumplir estos dos aspectos de la seguridad informática

Cifrado con Clave Privada del Origen

origen

Mateico

Claves: e_B, n_B, d_B

e_B, n_B: públicas

d_B: privada

e_B y d_B son inversas dentro de un cuerpo Z

Si ahora Mateico realiza la operación de cifrado con su clave privada d_B en el cuerpo n_B LaSole será capaz de comprobar este cifrado ya que posee (entre otras) la clave pública de Mateico. Comprueba así tanto la autenticidad del mensaje como del autor.

$C = E_{dB}(M) \bmod n_B$

destino

LaSole

Claves: e_A, n_A, d_A

e_A, n_A: públicas

d_A: privada

e_A y d_A son inversas dentro de un cuerpo Z

Operación de Cifrado con Clave de Origen

Cifrado:

Mateico firma un mensaje M a LaSole

Mateico

LaSole

Claves: e_B, n_B, d_B

$C = E_{dB}(M) \bmod n_B$

Claves: e_A, n_A, d_A

Clave privada

Claves públicas

Descifrado:

$M = E_{eB}[E_{dB}(M)] \bmod n_B$

E_{dB} y E_{eB} son inversos

Se comprueba la *integridad* del origen

Uso de la Criptografía Asimétrica

¿Qué aplicación tendrán entonces los sistemas de criptografía de clave pública o asimétrica?

- Usando la clave pública del destino se hará el intercambio de claves de sesión de un cifrado con sistemas simétricos (decenas a centenas de bits).
- Usando la clave privada de origen, se firmará digitalmente un resumen (decenas a centenas de bits) del mensaje obtenido con una función hash.

Gestión de Claves (Comparación)

Gestión de claves

Clave Secreta	**Clave Pública**
Hay que memorizar un número muy alto de claves: $\rightarrow n^2$.	Sólo es necesario memorizar la clave privada del emisor.

En cuanto a la gestión de claves, serán más eficientes los sistemas de cifra asimétricos.

Espacio de Claves (Comparación)

Longitud y espacio de claves

Clave Secreta

Debido al tipo de cifrador usado, la clave será del orden de la ***centena*** de bits.

Clave Pública

Por el algoritmo usado en el cifrado, la clave será del orden de los ***miles*** de bits.

3^{128} 3^{1024}

En cuanto al espacio de claves, no son comparables los sistemas simétricos con los asimétricos.

Tiempo de Vida de las Claves

Tiempo de vida de una clave

Clave Secreta

La duración es muy corta. Normalmente se usa como una clave de sesión.

Clave Pública

La duración de la clave pública, que la entrega y gestiona un tercero, suele ser larga.

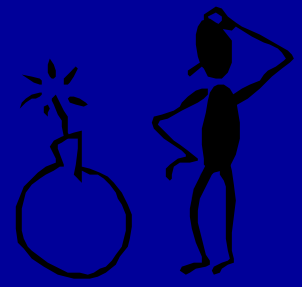
Segundos, minutos

En cuanto al tiempo de vida de una clave, en los sistemas simétricos éste es muchísimo menor que el de los usados en los asimétricos.

Meses, años

Vida de la Clave y Principio de Caducidad

Si en un sistema de clave secreta, ésta se usa como clave de una sesión que dura muy poco tiempo... y en este tiempo es imposible romperla...



¿para qué preocuparse entonces?

La confidencialidad de la información tiene una *caducidad*. Si durante este tiempo alguien puede tener el criptograma e intentar un ataque por fuerza bruta, obtendrá la clave (*que es lo menos importante*) ...

¡pero también el mensaje secreto!

El problema de la Autentificación

Condiciones de la autenticidad:

a) El usuario **A** deberá protegerse ante mensajes dirigidos a **B** que un tercer usuario desconocido **C** introduce por éste. Es la "sustitución" de identidad o problema de la **autentificación del emisor**.

b) El usuario **A** deberá protegerse ante mensajes falsificados por **B** que asegura haberlos recibido firmados por **A**. Es la falsificación del documento o problema de la **autentificación del mensaje**.

Autentificación de Origen y Destino

Autentificación

Clave Secreta

Sólo será posible autentificar el mensaje con una marca pero no al emisor.

Clave Pública

Al haber una clave pública y otra privada, se podrá autentificar el mensaje y al emisor.

En cuanto a la autentificación, los sistemas asimétricos -a diferencia de los simétricos- permiten una **firma digital**.

Velocidad de Cifrado

Velocidad de cifrado

Clave Secreta

La velocidad del cifrado es muy alta. Es el algoritmo de cifra del mensaje.

Clave Pública

La velocidad de cifrado es muy baja. Se usa para el intercambio de clave y la firma digital.

Cientos de M Bytes/seg

Cientos de K Bytes/seg

En cuanto a la velocidad de cifra, los sistemas simétricos son de 100 a 1.000 veces más rápidos que los asimétricos.

Resumen cifrado Simétrico vs. Asimétrico

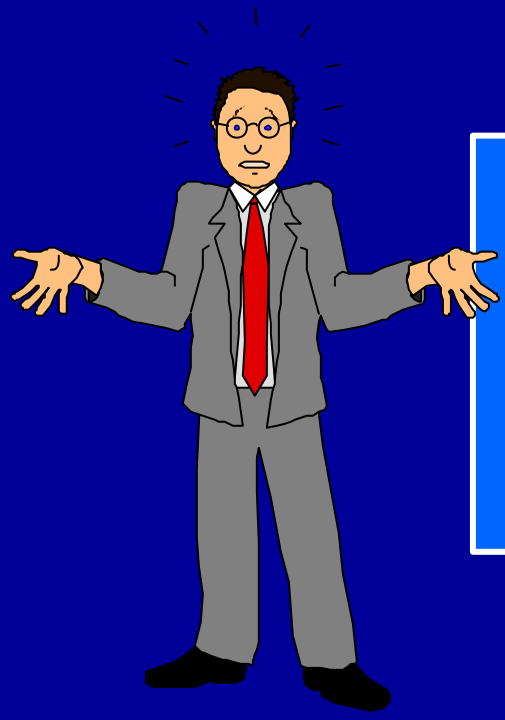
Cifrado Simétrico

- Confidencialidad
- Autentificación parcial
- Sin firma digital
- Claves:
  - Longitud pequeña
  - Tiempo de vida corto
  - Número elevado
- Velocidad alta

Cifrado Asimétrico

- Confidencialidad
- Autentificación total
- Con firma digital
- Claves:
  - Longitud grande
  - Tiempo de vida largo
  - Número reducido
- Velocidad baja

Coming Next!



Cifrado en Flujo y en Bloque con Clave Secreta