CRIPTOANÁLISIS CLÁSICO

DESARROLLO

Módulo I – Criptografía clásica

Definición, objetivos y fundamentos de criptología. Introducción a los criptosistemas. Definiciones. Necesidades. Seguridad teórica y práctica. Criptografía por software y por hardware. Ataques criptográficos. Clasificación general de sistemas criptográficos. Principios de Kerckhoff. Procedimientos clásicos de cifrado: métodos históricos (escítala espartana, etc.). Primalidad, aritmética modular (Z_n, Z^*_n) y funciones numéricas elementales (MCD, mcm). Clave de Julio César. Métodos generales de transposición y sustitución. Cifrador afín. Cifrador de Hill. Claves polialfabéticas (Vigenère). Otros métodos especiales (Playfair, Autokey, etc.). Criptoanálisis elemental: ataque estadístico. Método de Kasiski. Índice de coincidencia e Índice de coincidencia mutuo.

ENCRIPTOR DE HILL

Es otro algoritmo polialfabético inventado en 1919 por Lester S.Hill.

Sea m un entero positivo y $P = C = (Z_{26})^m$ La idea básica es tomar m combinaciones lineales de los m carácteres de un elemento del texto plano, generando así un elemento cifrado.

Por ejemplo, sea $m = 2$ o sea que un elemnto de texto plano es de la forma $x = (x_1, x_2)$ y un elemento cifrado es $y = (y_1, y_2)$ Entonces y_1 podría ser una combinación lineal de x_1, x_2 como

$$y_1 = 11x_1 + 3x_2 \quad , \quad y_2 = 8x_1 + 7x_2 \text{ que en notación matricial es}$$

$$(y_1, y_2) = (x_1, x_2)\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}$$

ENCRIPTOR DE HILL

En general la clave será una matriz K de mxm. Si $x = (x_1,..., x_m) \in P$ y $k \in K$ calculamos

$$y = e_k(x) = (y_1,..., y_m = (x_1,..., x_m) \begin{pmatrix} k_{1,1}\, k_{1,2} \ldots\ k_{1,m} \\ k_{2,1}\, k_{2,2} \ldots k_{2,m} \\ \ldots\ldots\ldots\ldots\ldots \\ k_{m,1}\, k_{m,2} \ldots k_{m,m} \end{pmatrix} \quad \text{o sea } y = xK$$

si la matriz es inversible, entonces $x = yK^{-1}$

Por ejemplo, $\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}^{-1} = \begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix}$ pues $\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}\begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} = \begin{pmatrix} 261 & 286 \\ 182 & 131 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(siempre módulo 26)

ENCRIPTOR DE HILL

Encriptemos la palabra *july*

$ju = (9,20) \quad ly = (11,24)$

$$(9,20)\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} = (99+60, 72+140) = (3,4)$$

$$(11,24)\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} = (121+72, 88+168) = (11,22)$$

Entonces $july \Rightarrow DELW$. Para desencriptar :

$$(3,4)\begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} = (9,20)$$

$$(11,22)\begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} = (11,24)$$

ENCRIPTOR DE PERMUTACION

Los algoritmos que vimos hasta ahora involucran substituciones, o sea que carácteres del texto plano de reemplazan por otros correspondientes al texto cifrado.
La idea de un cifrador de permutaciones es mantener a los carácteres originales, pero alterar su posición. Esta idea se ha usado por siglos, y de hecho la diferencia entre el cifrador de permutaciones y el de subtitución fué señalada por Giovanni Porta en 1563.

Formalmente :

Sea m un entero positivo fijo, $P = C = (Z_{26})^m$ y sea K el conjunto de todas las permutaciones de $\{1,...,m\}$. Para una clave dada k (o sea una permutación π definimos $e_\pi(x_1,...,x_m) = (x_{\pi(1)},...,,x_{\pi(m)}) = (y_1,...,y_m)$ y

$$d_\pi(y_1,...,y_m) = (y_{\pi^{-1}(1)},...,y_{\pi^{-1}(m)})$$

ENCRIPTOR DE PERMUTACION

Como en el cifrador de substituciones es más conveniente usar carácteres alfabéticos en vez de residuos módulo q puesto que no hay operaciones modulares.

Ejemplo :

Sea $m = 6$ y usamos como clave a la permutación π dada por

1	2	3	4	5	6
3	5	1	6	4	2

y su inversa π^{-1}

1	2	3	4	5	6
3	6	1	5	2	4

Si el texto es mañanairemosalcinesolos

lo dividimos en grupos de 6 letras

mañana : iremos : alcine : solos

ñnmaaa : eoismr : cnael : lss?o $\Rightarrow$ necesita "padding"

El método de permutaciones es un caso especial del de Hill. Dada una permutación π de los enteros $\{1,...,m\}$ podemos definir una matriz de permutaciones asociada K_π de mxm según la fórmula

$$k_{i.j} = \begin{cases} 1 \text{ si } i = \pi(j) \\ 0 \text{ de lo contrario} \end{cases}$$

Es fácil ver entonces que el método de Hill usando la matriz K_π es equivalente a la encripción de permutaciones cuando se usa la permutación π

Para el caso anterior la matriz es :

$$K_\pi = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \text{ siendo } K_\pi^{-1} = K_\pi^t$$

CRIPTOGRAFIA CLASICA

se han visto…

- Clave de Corrimiento
- Clave de Permutación
- Clave de Substitución Monoalfabética
- Clave de Sustitución Homofónica
- Clave de Sustitución Polialfabética (Vigenére)
- Clave Afín
- Clave de Hill
- Clave de Vernam
- LFSR (nociones)

CRIPTOGRAFIA CLASICA

comentaremos…

- Clave Autokey
- Claves Beaufort (Vigenére modificado)
- Clave PlayFair
- Nomencladores
- Rotor Jefferson
- Rotores Hebern-Koch
- Rotor Scherbius (Enigma)
- Rotor Hagelin M-209

AUTOKEY

Es un Vigenére polialfabético de s-caracteres, de long variable, en la cual el propio texto cifrado sirve como clave

$k = k_1 k_2 k_3 \ldots k_t$ (secuencia concatenada)

para $i \leq t$; $C_i = (P_i + k_i) \bmod s$

para $i > t$; $C_i = (P_i + C_{i-t}) \bmod s$

(una variante usa P_{i-t} en vez de C_{i-t})

BEAUFORT

Es un Vigenére modificado que emplea la resta modular en vez de la suma. La clave de cada carácter es la propia inversa modular aditiva

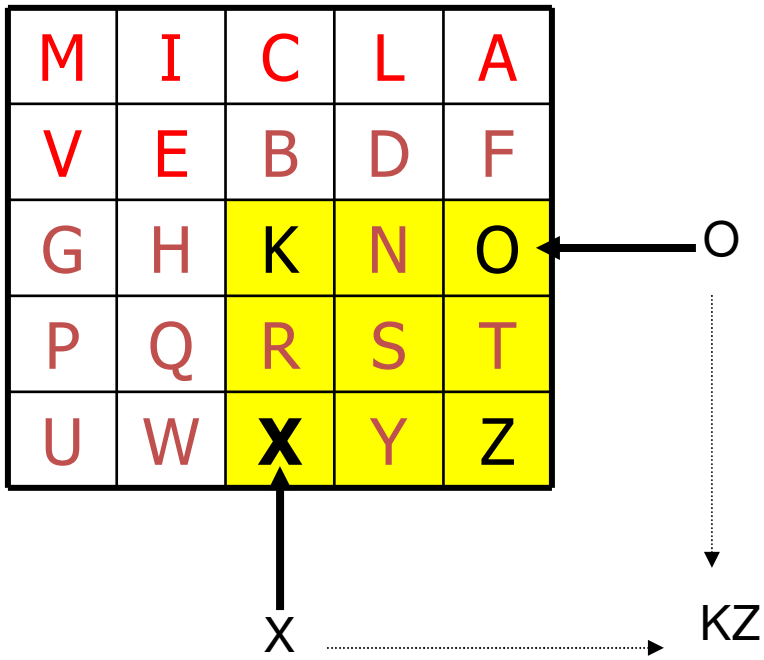
$k = k_1 k_2 k_3 \ldots k_t$ (secuencia concatenada)

Beaufort directo ; $C_i = (k_i - P_i) \bmod s$

Beaufort inverso ; $C_i = (P_i - k_i) \bmod s$

PLAYFAIR

es una sustitucion por digramas; EJ: "XO" → "KZ"



- misma fila: elementos de la derecha circular
- misma Col: elementos inferiores circulares
- idénticos: interponer carácter raro (Z)

Se ataca estadísticamente por digramas

NOMENCLADORES

Libros de Claves

encripción

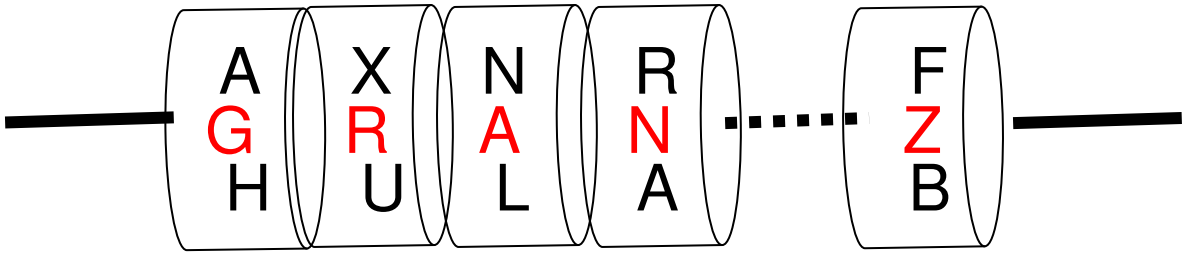
A VX
GAK
HZAM
LPWKM
AL
AMANECER
MG
FOQ
NUYY
SWAAT
...

desencripción

A BATALLA
AA CAMINO
AAA E
AAAA DIRECTO
AAAB OBUS
AB RUTA
ABIX 45
ABUZ DE
BCCAI UNA SOLA
...

ROTOR JEFFERSON

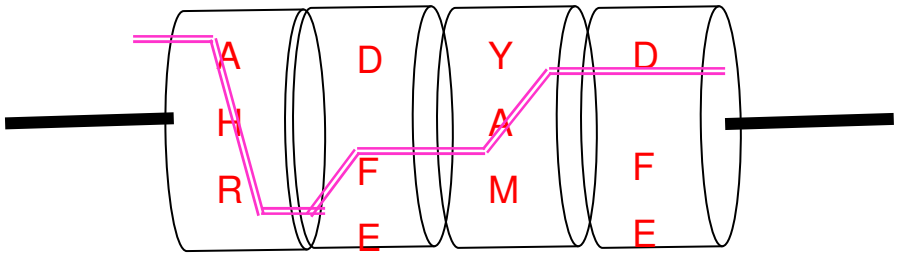
Fines del Siglo XVIII



36 cilindros apareados, cada uno con las 26 letras permutadas en forma diferente, se alinean según el texto plano y luego se sustituye por cualquier otra línea de las 25 restantes

ROTORES HEBERN-KOCH

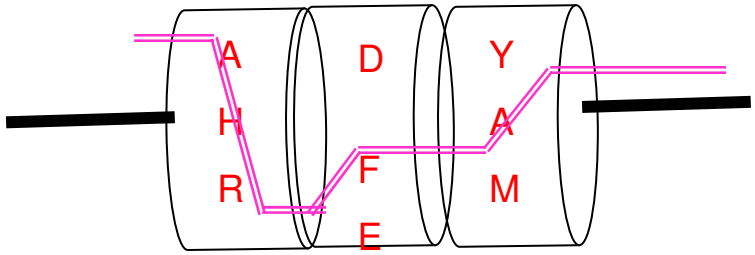
Siglo XX : 1918-1921



máquinas de escribir / calcular electromecánicas modificadas para sustituciones polialfabéticas hardcodeadas por la permutación y posición inicial de cilindros y por el cableado interno entre cilindros y con output luminoso o impreso. Los cilindros se desplazan por engranajes tipo contador de Km

ENIGMA (rotor Scherbius)

1923-1934

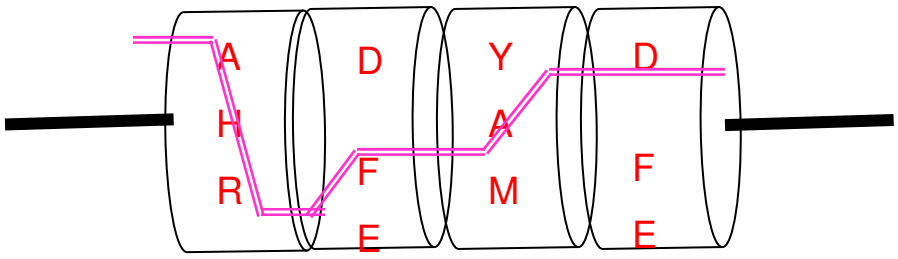


Modelo C (Alemania II GM):

Tres rotores R_1, R_2, R_3. R_1 cicla a R_2 que cicla a R_3, ademas R_2 se cicla a sí mismo. Período: 26.25.26≈17000 La clave era la posición inicial de los rotores (17000), su orden (3!=6) y el estado de un tablero externo que implementaba una sustitución polialfabética fácilmente programable (26!) (manualmente cada hora) mas las sustituciones internas (fijas) de los rotores. Semi-quebrado en 1940 por el equipo de Alan Turing auxiliado por una computadora digital rudimentaria de Ira generación. En 1942 se capturó una Enigma intacta en el U-110 cerca de Islandia. Enigma fue el secreto mejor guardado por los aliados durante 30 años (decenas de miles guardaron silencio, ejemplo único en la historia)

HAGELIN M-209

1934 USA



La máquina del ejército EEUU durante la II GM. Tenía 6 rotores con sustitución polialfabética Beaufort de período 101405850= 26.25.23.21.19.17 letras. El cifrador de rotores mas evolucionado. Hoy día pieza de museo.

CRIPTOANALISIS CLÁSICO

- GENERALMENTE SE ASUME KERCKHOFF
- SELECCIONAR EL ATAQUE ESPECIFICO DEL CRIPTOANALISIS ELEMENTAL PARA LOS CASOS DE CRIPTOGRAFIA CLASICA
- EL CRIPTOANALISIS DEBE ATACAR TANTO A LA PRIMITIVA COMO AL PROTOCOLO, Y EN ESTO PUEDEN FALLAR LAS IMPLEMENTACIONES DE LAS MAS POTENTES PRIMITIVAS ACTUALES
- EN CRIPTOGRAFIA ACTUAL, EL CRIPTOANALISIS ESTARA ORIENTADO A RESOLVER PROBLEMAS COMPLEJOS DE LA TEORIA DE NUMEROS (factorización, logaritmo discreto, raiz modular, generadores, residuosidad cuadrática, etc.)

TERMINOLOGIA BASICA - 1

ATAQUES CRIPTOANALITICOS

Son sistemas que permiten a un ADVERSARIO quebrar un criptosistema o reducir su complejidad interna de manera de facilitar estadísticamente su quiebre respecto al "ataque de fuerza bruta"

ATAQUE DE FUERZA BRUTA

Consiste en explorar sistemáticamente el espacio de claves $\{d_k\}$. Esto implica "tantear" todas las combinaciones posibles para claves legales.

TERMINOLOGIA BASICA - 2

CRIPTOSISTEMA SEGURO

Es aquel para el cual no se conoce mejor sistema de ataque que el de "Fuerza Bruta". Es el ideal perseguido por la criptografía aplicada.
Si el espacio de claves fuese (por ejemplo) del orden de 2^{128} ($\approx 10^{38}$), y se hiciesen tanteos de "fuerza bruta" a razón de 10^{20} claves por segundo (imposible!!!), no alcanzaría la edad del universo desde el big-bang ($\approx 10^{17}$ seg) para quebrar ese sistema. La esperanza de acertar en un tanteo de un espacio |N| es |N|/2.

CRIPTOANALISIS ELEMENTAL: ataque a la sustitución monoalfabética

- El ataque es estadístico. La sustitución simple conserva las frecuencias de letras, digramas, trigramas, etc.
- Frecuencias decrecientes de letras (inglés): @=sp

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

- Frecuencias decrecientes de digramas:

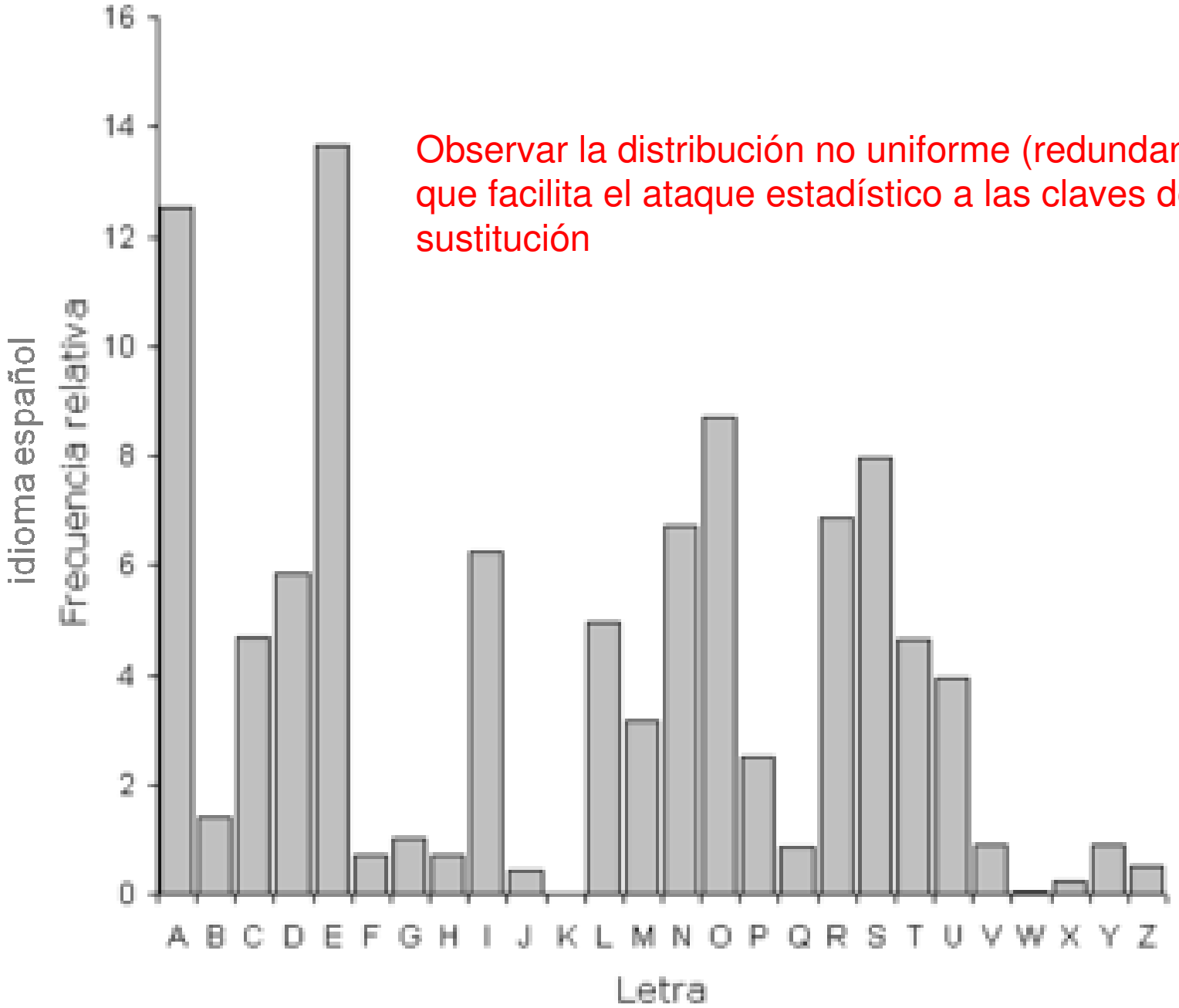
TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI, OF

- Frecuencias decrecientes de trigramas:

THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH

- Tener en cuenta digramas y trigramas prohibidos:

ZZ, TXZ, AAA, KPL, etc



Observar la distribución no uniforme (redundancia) que facilita el ataque estadístico a las claves de sustitución

Ejemplo concreto: el Quijote El texto del Quijote contiene 1.640.502 letras:

LETRA	CANTIDAD	PORCENTAJE
e	229188	14,0%
a	200492	12,2%
o	162512	9,9%
s	125726	7,7%
n	108440	6,6%
r	100953	6,2%
i	90070	5,5%
l	89141	5,4%
d	87237	5,3%
u	79471	4,8%
t	61749	3,8%
c	59435	3,6%
m	44658	2,7%
p	35464	2,2%

LA REDUNDANCIA DEL LENGUAJE LE CONFIERE ROBUSTEZ A LAS COMUNICACIONES

C13R70 D14 D3 V3R4N0 3574B4 3N L4 PL4Y4 0853RV4ND0 D05 CH1C45 8R1NC4ND0 3N 14 4R3N4, 357484N 7R484J484N MUCH0 C0N57RUY3ND0 UN C4571LL0 D3 4R3N4 C0N 70RR35, P454D1Z05, 0CUL705 Y PU3N735. CU4ND0 357484N 4C484ND0 V1N0 UN4 0L4 D357RUY3ND0 70D0 R3DUC13ND0 3L C4571LL0 4 UN M0N70N D3 4R3N4 Y 35PUM4... P3N53 9U3 D35PU35 DE 74N70 35FU3RZ0 L45 CH1C45 C0M3NZ4R14N 4 L10R4R, P3R0 3N V3Z D3 350, C0RR13R0N P0R L4 P14Y4 R13ND0 Y JU64ND0 Y C0M3NZ4R0N 4 C0N57RU1R 07R0 C4571LL0

C0MPR3ND1 9U3 H4814 4PR3ND1D0 UN4 6R4N L3CC10N; 64574M05 MUCH0 713MP0 D3 NU357R4 V1D4 C0N57RUY3ND0 4L6UN4 C054 P3R0 CU4ND0 M45 74RD3 UN4 0L4 L1364 4 D357RU1R 70D0, S010 P3RM4N3C3 L4 4M1574D, 3L 4M0R Y 3L C4R1Ñ0, Y L45 M4N05 D3 49U3LL05 9U3 50N C4P4C35 D3 H4C3RN05 50NRR31R.

CRIPTOANALISIS ELEMENTAL: ataque a la sustitución monoalfabética

- Por ejemplo, en el texto cifrado:
  YIFQFMZRWQFYVECFMDZLCVMRZWNMDZVEJBTXCDDUMJN
  DIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZNZU
  CDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJXZWG
- Se hallaron las siguientes frecuencias decrecientes de letras:

Z, M, C, D, J, F, Y, R, N

- Se les asignan inicialmente (por tanteo):

E, T, A, O, I, N, S, H, R

Usando las frecuencias decrecientes de digramas:

TH, HE, IN, ER, AN, RE, ED, y los digramas prohibidos, se resuelven loas casos dudosos. Por computadora en minutos (o segundos) se resuelve cualquier clave de este tipo

CRIPTOANALISIS ELEMENTAL: ataque al cifrador afín

- Supongamos que el cifrado sea

FMXVEDKAPHFERBNDKRXRSREFMORUDSDKDVSHVUFEDKAPRKD
LYEVLRHHRH

- Las letras mas frecuentes son R y D. Si se le asignan (para inglés) las letras E y T, se obtiene un sistema lineal:

$4a+b=17$ (E→R)

$19a+b=3$ (T→D)

- Resolviendo se obtienen a=6 y b=19 (en Z_{26}), lo que es ilegal (el **MCD(6,26)>1**). Pero tanteando con otras posibilidades, resulta:

$4a+b=17$ (E→R)

$19a+b=7$ (T→H)

Se obtiene a=3 y b=5, lo que permite desencriptar sin problemas: $x=a^{-1}(y-b) \bmod 26$

ALGORITHMSAREQUITEGENERALDEFINITIONSOFARITHMETICPR
OCESSES

CRIPTOANALISIS ELEMENTAL:
ataque a la sustitución polialfabética (tipo Vigenère)

- TEST DE KASISKI

(Friedrich Kasiski – 1863)

Se trata de descubrir la longitud de clave -m- empleada como primer paso para su quiebre

1. BUSCAR DISTANCIAS ENTRE n-GRAMAS (n>=3) IDENTICOS (d1, d2, d3, ….)
2. LA LONGITUD DE LA CLAVE (m) ES MUY PROBABLEMENTE EL MCD(d1, d2, d3, …)

CRIPTOANALISIS ELEMENTAL:
ataque a la sustitución polialfabética (tipo Vigenère)

- INDICE DE COINCIDENCIA

  Wolfe Friedman (1920)

  Probabilidad de que dos elementos x de un string $x_1..x_n$ sean idénticos

$$I_c(x)= (\Sigma f_i\ (f_i-1)\ )/(n(n-1))$$

  f_i:frecuencia observada de x_i, (x_i= A...Z)
  suma desde cero a 25 para letras sin eñe.

- Para 26 letras en strings aleatorios $I_c(x)=0.038$, para cualquier idioma y por la redundancia informática es superior, para inglés es $I_c(x)=0.065$
- Esta distancia estadística permite reconocer longitud (m) de claves, al correlacionar series desplazadas de strings.

CRIPTOANALISIS ELEMENTAL:
ataque a la sustitución polialfabética (tipo Vigenère)

- INDICE MUTUO DE COINCIDENCIAS (método del corrimiento)

  Probabilidad de que un elemento x de un string $x_1..x_n$ sean idéntico a un elemento y de otro string $y_1..y_{n'}$ independiente

$$MI_c(x,y)= (\Sigma f_i\ f_i')/(nn')$$

  f_i: frecuencia observada de x_i, (x_i= A…Z)

  f_i': frecuencia observada de y_i, (y_i= A…Z)

  suma desde cero a 25 para letras sin eñe.

- Presenta máximos cuando hay correlación significativa (debida a la redundancia del lenguaje) y por eso permite inferir claves.

CRIPTOANALISIS ELEMENTAL: ataque al método de Hill

- Es difícil de atacar por ataque CIFRADO-SOLO pero sucumbe muy fácil ante un ataque PLANO-CONOCIDO
- Supongamos que Oscar sabe que se trata de un Hill m=2 posee friday←→PQCFKU
- Entonces, de las tres parejas tipo (fr)->(PQ) sabrá que: ek(5,17)=(15,16) , ek(8,3)=(2,5) y ek(0,24)=(10,20)
- Se plantea y resuelve un sistema lineal y se obtiene la matriz K usada como clave.