GUÍA 1: CRIPTOGRAFÍA CLÁSICA - SOLUCIONES

Ejercicio 1:

- **cifrado de rotación.** *[Capítulo 1 de Katz]*

Gen: elige aleatoriamente $k \in \{0,1,...25\}$

Enc: Si el mensaje m es $m = m_1 m_2 ....m_t$, donde $m_i \in \{0,1,...25\}$ y cada m_i se corresponde con una letra en $\{a, b,...z\}$

entonces el cifrado c es $c = c_1 c_2 ....c_t$, donde $c_i = (m_i + k) (\text{mod } 26)$ $\forall i$. Cada c_i también se corresponde con una letra en $\{a, b,...z\}$

Dec: Si el cifrado es $c = c_1 c_2 ....c_t$, entonces el texto plano obtenido es $m = m_1 m_2 ....m_t$, donde $m_i = (c_i - k) (\text{mod } 26)$ donde $e^{-1}{}_k$ es la función inversa de e_k, que existe por ser biyectiva.

- **cifrado de sustitución monoalfabética**

[Menezes Cap. 1]

Cifrado de Sustitución Simple:
Sea A es un alfabeto de q símbolos y M el conjunto de todas las cadenas de longitud t sobre A.
Sea K el conjunto de todas las permutaciones en el conjunto A.
Para cada $e \in K$, se define una transformación de encripción E_e tal que:
$E_e(m) = (e(m_1) e(m_2) \dots e(m_t)) = (c_1 c_2 \dots c_t) = c$,
donde $m = (m_1 m_2 m_3 \dots m_t) \in M$
Cada símbolo de la tupla, se reemplaza por otro símbolo de A de acuerdo a una permutación fija e.

Es decir, un sistema de sustitución simple considera una permutación del alfabeto (función biyectiva del alfabeto en sí mismo) y lo aplica a cada símbolo que forma el mensaje.

Teniendo en cuenta lo anterior, se definen:

Gen: elige aleatoriamente una función de todas las que hay en K, es decir:

$e_k \leftarrow K = \{e_1, e_2, ..., e_{q!}\}$ (q es la cantidad de símbolos del alfabeto A, y $e_i : A \rightarrow A$ *biyectiva* $\forall i$)

Enc: Si el mensaje m es $m = m_1 m_2 ....m_t$,

entonces el cifrado c es $c = c_1 c_2 ....c_t$, donde $c_i = e_k(m_i)$ $\forall i$

Dec: Si el cifrado es $c = c_1 c_2 ....c_t$, entonces el texto plano obtenido es $m = m_1 m_2 ....m_t$, donde $m_i = e^{-1}{}_k(c_i), \forall i$ donde $e^{-1}{}_k$ es la función inversa de e_k, que existe por ser biyectiva.

- **cifrado de Vigenère Considerado para el alfabeto inglés, de 26 letras.**

Gen: elige aleatoriamente una palabra del alfabeto: k$\leftarrow A^+$

Es decir, $k = k_0 k_1 ....k_{t-1}$ donde $\forall i : k_i \in A$. Se puede considerar que cada letra se corresponde con un número entre 0 y 25 lo cual facilita la definición de operaciones.

Enc: Si el mensaje m es $m = m_0 m_1 ....m_{n-1}$,

entonces el cifrado c es $c = c_0 c_1 ....c_{n-1}$, donde $\forall i : c_i = (m_i + k_j)$ mod (26) con $i \equiv j$ mod (t) $1 \le j \le t$

Dec: Si el cifrado es $c = c_0 c_1 ....c_{n-1}$, entonces el texto plano obtenido es $m = m_0 m_1 ....m_{n-1}$, donde $\forall i : m_i = (c_i - k_j)(26)$ mod (26) con $i \equiv j$ mod (t) $1 \le j \le t$

Ejercicio 2:

GUÍA 1: CRIPTOGRAFÍA CLÁSICA - SOLUCIONES

La composición de funciones biyectivas es biyectiva. Como una sustitución es una función biyectiva, la composición de sustituciones (composición de funciones biyectivas) es también una sustitución. Por lo tanto, la composición de dos sistemas de sustitución simple puede reemplazarse por un solo sistema de sustitución simple (no provee más seguridad que el uso de uno solo)

Demostración formal:

$A=\{m_1, m_2, \ldots, m_q\}$ es el alfabeto de q símbolos.

$K=\{e_1, e_2, \ldots, e_{q!}\}$ es el conjunto de todas las permutaciones posibles de elementos de A, lo que equivale a decir que $K=\{e_i : A \rightarrow A \,/\, e_i \, es biyectiva\}$

$\forall m \in A: \; e_j(e_i(m))=(e_j \circ e_i)(m)$ donde $e_j \circ e_i$ es una función biyectiva porque la composición de funciones biyectivas también lo es. Entonces, como en K están todas las permutaciones de elementos de A, es decir están todas las funciones biyectivas de A en A, también estará $e_k=e_j \circ e_i$. Por lo tanto existe una sustitución e_k tal que $\forall m: e_j(e_i(m))=(e_j \circ e_i)(m)=e_k(m)$

Entonces la composición de sustituciones no provee más seguridad que el uso de una sola sustitución.

Ejercicio 3:

[con criptoclásicos: http://www.criptored.upm.es/software/sw_m001c.htm***]***

PERSEVERA EN AQUELLAS COSAS QUE REALMENTE DESEES CONSEGUIR PORQUE SI ES ASI LO CONSEGUIRAS

La idea es, por un lado, tener en cuenta que se usó una permutación del alfabeto:

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
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

Y por otro lado, que la frecuencia de aparición de cada letra en el texto cifrado se corresponde con la frecuencia de aparición de cada letra en el texto original.

En un texto en castellano, la letra 'E' es de las más frecuentes. En el texto cifrado, la K aparece 13/51 = 25%. Si suponemos que la 'K' reemplaza a la 'E', quiere decir que se usó un cifrado de César de clave 6 ('K'-'E' = 6). Se prueba con esa clave descifrar el resto del texto: si es coherente, ya está. Si no, buscar qué otra letra se puede repetir con un 25% de frecuencia y repetir el proceso.

Ejercicio 4:

a) El resultado sería (sin considerar espacios, tomando el mismo alfabeto del ejercicio 2)

U	N	V	I	N	O	D	E	M	E	S	A
B	A	C	O	B	A	C	O	B	A	C	O
V	**N**	**X**	**W**	**Ñ**	**O**	**F**	**S**	**N**	**E**	**U**	**O**

b) Conviene usar la clave COMPADRE, porque la clave CERO es demasiado corta. La longitud de la clave se conoce como el período del cifrado. Lo ideal es que el período sea lo más grande posible, si pudiera ser del mismo tamaño del mensaje, ¡mejor! Pero eso no sería práctico.

c) Se distinguen dos casos: primero, cuando la clave de ambos cifrados Vigenere tiene igual longitud y segundo, cuando la clave de ambos cifrados es de distinta longitud.

La composición de dos vigenere de claves de igual longitud es un vigenere cuya clave tiene la misma longitud que las primeras, y resulta de combinar las mismas (k' = k1 + k2)

Ejemplo: Encriptar con K1 = BARBA y luego con K2 = JAMON equivale a encriptar de una sola vez con K' = KADPN

Si las claves son de distinta longitud, la nueva clave tiene longitud igual al mínimo común múltiplo entre la longitud de la clave 1 y la longitud de la clave 2 es decir: mcm (long clave1, long clave2). Luego se combinan repitiendo las mismas las veces que sea necesario, obteniendo una nueva clave k'.

Es decir, si $K_1 = k_{11} k_{12} \dots k_{1p}$ y $K_2 = k_{21} k_{22} \dots k_{2q}$, entonces $K' = k'_1 k'_2 k'_3 \dots k'_m$, tal que $k'_1 \equiv (k_{11} + k_{21}) (\mathrm{mod}\ 27)$, $k'_2 \equiv (k_{12} + k_{22}) (\mathrm{mod}\ 27)$, y donde $m = mcm(p, q)$ siendo $p =$ longitud de K_1 y $q =$ longitud de K_2

Ejemplo:

K1 = JAMON →p = 5

K2 = BAR → q = 3

K' debe tener longitud 15:

K1 = JAMONJAMONJAMON

K2 = BARBARBARBARBAR

K' = KADPNABMGÑJRNOE

Es decir, encriptar con K1 y luego con K2 equivale a encriptar una sola vez con K'.

Ejercicio 5:

Criptograma 1

Sustitución monoalfabética. Conserva frecuencias en las nuevas letras, pero usa una permutación del alfabeto original.

Criptograma 2

Sustitución polialfabética

Criptograma 3

Trasposición conserva el alfabeto y la frecuencia de las letras originales.

Ejercicio 6:

[Bishop Chap. 9]

El índice de coincidencia mide las diferencias en las frecuencias de las letras en el texto cifrado. Es definido como la probabilidad de que dos letras elegidas al azar en el texto cifrado sean la misma. Si F_c es la frecuencia de aparición de un carácter cifrado c y N es la longitud del texto cifrado, puede mostrarse que el índice de coincidencia es $IC = \frac{1}{N(N-1)} \sum_{i=0}^{25} F_i (F_i - 1)$.

En idioma ingles, los índices de coincidencia para los diferentes períodos son:

Period	1	2	3	4	5	10	Large
Expected IC	0.066	0.052	0.047	0.045	0.044	0.041	0.038

(más bajo es el IC cuando el período es más largo)

El prusiano Kasiski observó que las repeticiones ocurren cuando los caracteres de la clave aparecen sobre los mismos caracteres en el texto cifrado. El número de caracteres entre repeticiones es un múltiplo del período.

EJEMPLO: THE BOY HAS THE BAG, con clave VIG.

V	I	G	V	I	G	V	I	G	V	I	G	V	I	G
T	H	E	B	O	Y	H	A	S	T	H	E	B	A	G
O	P	K	W	W	E	C	I	Y	O	P	K	W	I	M
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

En el texto cifrado la cadena OPK aparece dos veces.

Esas dos veces están originadas en que la clave VIG se combina con el mismo texto THE.

Las repeticiones en el texto cifrado se dan a **9** caracteres de distancia. Esto es, un múltiplo del período.

El análisis para atacar consistiría en:

1. examinar el texto cifrado para múltiples repeticiones
2. tabular su longitud y el número de caracteres entre sucesivas repeticiones
3. El período debe ser un factor del número de caracteres entre sucesivas repeticiones. Entonces, a partir de las repeticiones, se establecen probables períodos, usando el índice de coincidencia para chequear las deducciones realizadas.
4. Tabular los caracteres para cada letra de la clave por separado y resolver cada una como un cifrado César.

En este ejercicio, el análisis es:

Grupo	Lugares donde aparece	Número de caracteres entre apariciones
JGAZ	0,128	128
NMON	75,256	184
PNFA	87,283	196
AZMJ	130,98	32
NMJ	263,375,475	112,100
AZM	98,130,138	32, 8

El período debe ser un factor de los números que aparecen en la última columna. Como todos son múltiplos de 4, el período debe ser 4 o 2 o 1.

Se parte el texto cifrado en grupos de 4.

J	G	A	Z
N	W	I	N
H	Y	L	Z
D	Y	V	B
B	J	L	C
Q	H	T	N
K	U	D	Q
X	M	O	X
J	N	O	Z

Y se analiza como para cifrado César: en cada columna cuál es la letra con más frecuencia, se estima cuál puede ser, se analizan grupos más frecuentes, etc.

En este caso resulta:

Clave: JUAN

AMANECIA Y EL NUEVO SOL PINTABA DE ORO LAS ONDAS DE UN MAR.

Ejercicio 7:

a) Primero, obtener k a partir de frecuencias del idioma. Luego, reordenar texto.

Ejemplo: HHAVWWWHR

Como la letra más frecuente en español es la E, podemos estimar que H reemplaza la E. (o W)

En el caso de que la H reemplace la E, significaría que k = 3.

Entonces queda: EEXSTTTEO

Luego se reordenan las letras:

N = 2 Falla

EE

XS

TT

TE

O

N = 3 CORRECTO!

EEX

STT

TEO

→ESTETEXTO

b) Para obtener la clave de Cesar, el máximo de pruebas es q (siendo q el número de símbolos del alfabeto). Por cada una de esas pruebas, hay que transponer m veces, entonces requeriría m*q pruebas en total.

Ejercicio 8:

Chosen-plaintext attack:

El adversario tiene la posibilidad de obtener el cifrado de textos planos de su elección.
Intenta determinar el texto plano correspondiente a otro texto cifrado.

- **cifrado de sustitución monoalfabética**

El adversario debería hacer cifrar un texto que contenga todas menos una de las letras del alfabeto.

Ej: Si hace cifrar el texto elegido "abcdefghijklmnopqrstuvwxy", obtendrá $c = c_1 c_2 ... c_{25}$, y conocerá entonces $e_k(x) \forall x$, ya que el cifrado de z, queda unívocamente determinado por ser e_k biyectiva.
Podrían ser menos, si de acuerdo al significado del texto se pueden "deducir" las restantes correspondecias.

- **cifrado de Vigenère**

El adversario debería hacer cifrar un texto de igual longitud del cifrado que quiere determinar, y que contenga la misma letra.

Es decir, si el cifrado es $c = c_0 c_1 ... c_{n-1}$, se puede elegir $m = a^n$. Entonces, al encriptar $m = a^n$ con la clave $k = k_0 k_1 .... k_{t-1}$, el resultado será $k = k_0 k_1 .... k_{t-1} ...$