CENTRO DE INVESTIGACIONES AVANZADAS DEL I.P.N.

CINVESTAV

Códigos y Criptografía

Profesor: Dr. Francisco Rodríguez Henríquez

Proyecto Final

Tema: Criptoanálisis para el cifrado de Vigenére

Alumno:

- Miranda Gómez Oscar

5 de Agosto del 2002

Criptoanálisis para el cifrado de Vigenére

Resumen

El proyecto consiste de un programa hecho en C para la plataforma Windows o MS-Dos el cual se encarga de realizar un criptoanálisis en aquellos textos que han sido cifrados por el método de Vigenére. El programa es capaz de descifrar los archivos sin tener que preguntar ningún tipo de información al usuario, obteniendo muy buenos resultados con textos mayores a 230 caracteres.

Adicionalmente se agrega un programa que sirve para cifrar por el método de vigenere.

Introducción

La criptografía surge de la necesidad de poder mantener fuera del alcance de las miradas ajenas, aquellos documentos de cierta importancia y delicadeza.

Dentro de la historia de la criptografía, se puede hacer una división entre los algoritmos criptográficos que existen:

- Algoritmos clásicos
- Algoritmos modernos

Los mecanismos criptográficos considerados como clásicos, son todos aquellos sistemas de cifrado que surgieron antes de la II Guerra Mundial, o lo que es lo mismo, antes del nacimiento de las computadoras. Estas técnicas tienen en común que pueden ser empleadas usando simplemente papel y lápiz, y que pueden ser criptoanalizadas casi de la misma forma.

Los algoritmos criptográficos modernos, surgen después de la Segunda Guerra Mundial y en su mayoría son de clave asimétrica con un alto grado de seguridad.[2]

Entre los algoritmos clásicos mas representativos se encuentran los siguientes:

- Cifrado por el algoritmo de Julius Caesar (sustitución).
- Sustitución Afín
- Cifrado de Vigenére
- Cifrado por substitución
- Cifrado por bloques, etc.

Para este proyecto, el algoritmo de cifrado que nos interesa es el de Vigenére, el cual es una variante del cifrado de Julius Caesar conocido como sustitución.
Este método criptográfico fue creado por el criptólogo francés Blaise de Vigenére en el siglo XVI, que consiste en un cifrado polialfabético que cuenta con un periodo y con corrimientos, en donde los corrimientos son dados basándose en una llave secreta.

El algoritmo de cifrado de Vigenére se creyó que era muy seguro, hasta que Babage y Kasisky mostraron como atacarlo utilizando el análisis de frecuencia combinado con la distancia entre letras.[1]

Actualmente, todos los métodos criptográficos clásicos han dejado de usarse, sin embargo, para fines educativos se siguen estudiando, ya que permiten comprender sus propiedades básicas, las cuales son útiles para entender mejor a los algoritmos modernos.[2]

Análisis del problema

El problema consiste en poder descifrar cualquier texto que haya sido encriptado con el método de Vigenére, sin tener que contar con la llave con la que fue cifrado el mensaje o texto original. Para lograrlo, es necesario conocer los métodos criptoanalíticos que existen y elegir el que resuelva el problema de manera automática, esto es, que el usuario no tenga que intervenir para ayudar al sistema a encontrar la llave secreta.

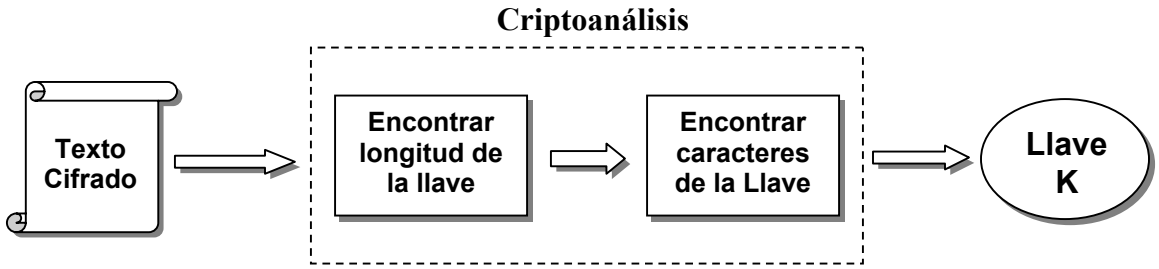

Figura 1: Proceso para obtener la llave secreta

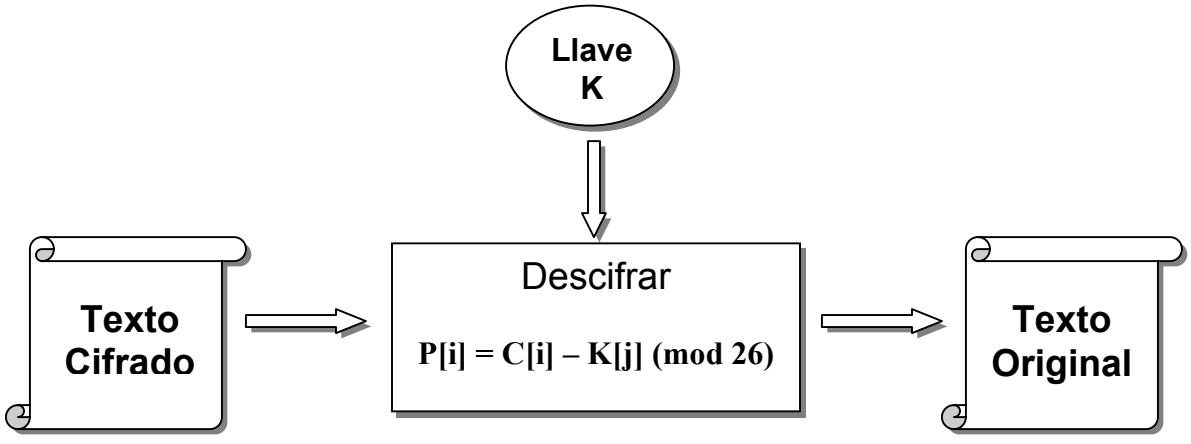

Figura 2: Proceso para descifrar por el método de Vigenére

Para poder desarrollar un método que rompa con el cifrado de Vigenére, es necesario comprender antes la manera en que funciona dicha técnica criptográfica, por lo que a continuación se explica como cifrar y descifrar con el método de Vigenére para después seguir con la explicación paso a paso de cómo realizar el criptoanálisis.

Cifrado

Debido a que el cifrado de Vigenére es una variación del cifrado de Caesar (mejor conocido como cifrado por substitución), al igual que este, toma las letras del alfabeto y las numera desde 0 a 25 para el idioma en inglés y desde 0 a 26 para el español.[1]

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

Tabla 1: Tabla del alfabeto con su numeración correspondiente.

Para empezar a cifrar un texto con este método, lo primero que se hace es elegir una llave, la cual puede ser una o varias palabras.

Por ejemplo, la llave “OSCAR” representada con números sería:

k = (14,18,2,0,17)

Utilizando la llave, lo siguiente es cifrar el mensaje de la siguiente manera:

Texto Claro	e	s	t	o		e	s		u	n	a		p	r	u	e	b	a
Llave	o	s	c	a		r	o		s	c	a		r	o	s	c	a	r
Texto Cifrado	s	k	v	o		v	g		m	p	a		g	f	m	g	b	r

Por *computadora*, el cifrado se realiza mediante la siguiente formula: C[i] = P[i] + K[j] (mod 26), donde C es la letra cifrada, P la letra del texto claro y K la letra de la llave que le corresponde, siendo “i” = 1 hasta el número de letras que tiene el texto a cifrar y para “j” = 1 hasta la longitud de la llave (repitiéndose cada vez que se llega a la última letra de la llave mientras “i” no sea el final del texto).

Ejemplo:

C1 = (P1 + K1)%26 s = e + o (mod 26)
C2 = (P2 + K2)%26 k = s + s (mod 26)
C3 = (P3 + K3)%26 v = t + c (mod 26)
C4 = (P4 + K4)%26 o = o + a (mod 26)

[1] Para la explicación del método tanto en el cifrado, descifrado y criptoanálisis se utilizará el alfabeto en inglés.

C5 = (P5 + K5)%26 v = e + r (mod 26)
C6 = (P6 + K1)%26 g = s+ o (mod 26)
C7 = (P7 + K2)%26 m = u + s (mod 26), etc.

El cifrado puede *hacerse a mano*, y para evitar trabajar con números, debido a que es mas confuso para las personas, se utiliza la tabla de Vigenére, la cual consiste en una matriz de caracteres cuadrada con dimensión de 26x26.

Llave \ Texto Claro	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
A	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
B	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	a
C	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b
D	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c
E	e	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
F	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
G	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	f
H	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g
I	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h
J	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i
K	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j
L	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k
M	m	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l
N	n	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m
O	o	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n
P	p	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
Q	q	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
R	r	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q
S	s	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
T	t	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s
U	u	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t
V	v	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
W	w	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v
X	x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w
Y	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x
Z	z	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

Tabla 2: Tabla de Vigenére.

La forma es que se utiliza la tabla es la siguiente:

Para el mensaje “esto es una prueba” con llave = “oscar”, primero se toma la columna “e” del mensaje a cifrar y luego se toma la fila “o” de la llave, obteniendo la letra cifrada en la celda donde se intercepta la columna y la fila, para este caso es la letra “s”; después se continúa con la columna “s”, fila “s”, donde la letra que se intercepta es la “k” y así sucesivamente hasta cifrar todo el texto.

Descifrado

Para descifrar un texto cifrado con él método de Vigenére, lo único que se tiene que hacer, es restar la llave a dicho texto.

$$P[i] = C[i] – K[j] \pmod{26}$$

Ejemplo:

P1 = (C1 + K1)%26 e = s - o (mod 26)
P2 = (C2 + K2)%26 s = k - s (mod 26)
P3 = (C3 + K3)%26 t = v - c (mod 26), etc.

Criptoanálisis

La seguridad en el método de cifrado de Vigenére depende de dos razones:

- La primera razón, se basa en el hecho de que nadie más conoce la longitud de la llave.
- La segunda razón, consiste en que el número posible de soluciones en comparación con el método de substitución simple, fue incrementado. Por ejemplo, si la longitud de la llave es igual a 5, el número de combinaciones que podrían ser necesarias para descubrir el texto por fuerza bruta sería de 26^5 = 11,881,376.

Actualmente existen técnicas en las que no es necesario realizar fuerza bruta para romper el Cifrado de Vigenére; mediante las siguientes técnicas se verá que es posible encontrar la longitud de la llave, la llave y por lo consiguiente el texto original.

Los pasos que se deben seguir para poder romper el cifrado de Vigenére son los siguientes:

- Encontrar la longitud de la llave
- Encontrar la llave

Una vez que se conoce la llave, lo único que quedaría por hacer, es utilizarla para descifrar el mensaje cifrado.

Encontrar la longitud de la llave

Para encontrar la longitud de la llave, lo que se tiene que hacer es:

1. Comparar cada una de las letras del texto cifrado, con las letras del mismo texto cifrado pero con un desplazamiento que puede ser desde 1 hasta la posible longitud de la llave.
2. Contar el número de coincidencias que se encontraron por cada desplazamiento.

3. Deducir la longitud de la llave basándose en el desplazamiento que tuvo el mayor número de coincidencias.

Ejemplo; se tiene el siguiente texto cifrado:

“skvuuwstacumpojrwnojdjkntwhclvgegtfrguacugtikagutvqfkcrgqjeifsoivblcsesugsrfacs
gojclrwerlvawptrqaqnussrlzqseifbwuciwhvoxfshitokadvgwiuiwvcdusvctfgwncfblgnzrgfe
cqmtsfskvaffagnkovqacojgsfzmeifbwhitwwptvrwnpictnedovgcfaggskotnetsjwnrqgoue
wucczcfuexijceehjgdfggoajsfviuovgsusecnvfsvacemgsvustaehaeelbsntfujcdfrweoetaf
eeqaclzrsfiehwirzrsfyrifveehaeiuovgncckfakckadfqmoeehguiehwtcratkauckgltcfveewv
qdvzuwrjcapcciqgecskvuuwgfeuwngrjckcsgsuvojhwqrzqgurvzseifbsfojqgplfgsngffavmf
gqgshiwoajqjkpkcytawwuqsrgaeodcdcagzaeatwgpdvraehrglgcewucsgojcrvggnvvfhtos
zwoajdjcckwuqsvbwneejaqdvrsvojmuqmlbaeatwgpdvaspeiokgglfs”

Para encontrar la longitud de la llave se empieza con un desplazamiento de 1 y se marcan las coincidencias de la siguiente manera:

s	k	v	u	u	w	s	t	a	c	u	m	p	o	j	r	w	n	o	j	d	j	k	n	t	w	h	c	l	v	g	e
	s	k	v	u	u	w	s	t	a	c	u	m	p	o	j	r	w	n	o	j	d	j	k	n	t	w	h	c	l	v	g

g	t	f	r	g	u	a	c	u	g	t	i	k	a	g	u	t	v	g	f	k	c	r	g	g	j	e	i	f	s	o	i	v	…
e	g	t	f	r	g	u	a	c	u	g	t	i	k	a	g	u	t	v	g	f	k	c	r	g	g	j	e	i	f	s	o	i	…

Lo anterior se tiene que repetir incrementando el desplazamiento de uno en uno hasta la posible longitud de la llave.

Los resultados obtenidos para un desplazamiento desde 1 hasta 7 fueron:

Desplazamiento	Numero de Coincidencias
1	23
2	22
3	23
4	27
5	42
6	32
7	33

Como se puede apreciar, el mayor número de coincidencias se obtuvo en el desplazamiento 5, por lo que se deduce que la longitud de la llave es de 5 caracteres.

Encontrar la llave

Ahora que ya se conoce la longitud de la llave (5 para el ejemplo), se puede dividir el texto cifrado en bloques para obtener todas aquellas letras que han sido

cifradas con el mismo desplazamiento (para este caso los caracteres 1,6,11, 16, etc han sido cifrados con el mismo desplazamiento).

1	2	3	4	5
s	k	v	u	u
w	s	t	a	c
u	m	p	o	j
r	w	n	o	j
d	j	k	n	t
w	h	c	l	v
g	e	g	t	f
r	g	u	a	c

Una vez hecha la separación, cada columna es como si hubiera sido cifrada mediante la técnica de substitución, ya que todas las letras que están en esa columna fueron substituidas con el mismo desplazamiento y por lo tanto es fácil atacar mediante el análisis de frecuencia.

El siguiente paso sería realizar un conteo para saber cuantas veces se repite la misma letra y poder conocer la de mayor frecuencia.

El análisis de frecuencia por columna arrojó los siguiente resultados:

	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	Letras contadas
1	5	8	10	2	1	8	10	7	4	1	0	0	1	0	10	0	10	9	10	1	4	0	15	0	0	5	121
2	14	0	0	1	3	5	14	3	0	10	9	3	5	1	0	0	3	0	16	3	8	7	15	0	1	0	121
3	2	0	16	0	10	7	17	2	2	1	4	0	0	9	6	8	8	2	0	7	6	11	2	0	0	0	120
4	15	0	9	9	15	0	2	1	11	0	0	7	2	7	9	2	0	6	13	6	4	1	0	0	1	0	120
5	0	0	8	2	13	17	4	1	5	10	6	3	0	0	0	0	0	9	1	7	9	16	1	2	0	6	120

A partir de aquí existen dos métodos para encontrar la letra que corresponde a cada columna de la llave.

Primer método:

Este método se basa en decidir cual de todas las letras podría ser la letra “e”, ya que esta letra, tanto en el idioma Ingles como en el Español es la que mayor frecuencia de aparición presenta. Continuando con el ejemplo, para la columna 1 se puede observar que la de mayor frecuencia fue la “w”, por lo que se supone que esta letra tiene mayor porcentaje para ser la letra “e” en el texto original. Para comprobarlo, se tiene que ver si las otras letras que aparecen con una frecuencia alta, concuerdan con la tabla de frecuencias del idioma.

a	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**
0.082	0.015	0.028	0.043	0.127	0.022	0.02	0.061	0.07
j	**k**	**l**	**m**	**n**	**o**	**p**	**q**	**r**
0.002	0.008	0.04	0.024	0.067	0.075	0.019	0.001	0.06
s	**t**	**u**	**v**	**w**	**x**	**y**	**z**	
0.063	0.091	0.028	0.01	0.023	0.001	0.02	0.001	

Tabla 3: Frecuencia de la letras en Ingles

Ejemplo: Si “w” = “e”, entonces la “q”=”i”, “o”=”g”, “g”=”y”, y “c”=“u”

Al comprobar estos resultados con la tabla de frecuencias podemos observar que las letras “y”,”g” y “u” tienen una frecuencia de 0.02 por lo que no es posible que tengan un alto grado de aparición en el texto. Esto nos lleva a deducir que la “w” no es la letra “e” y por lo tanto tenemos que elegir otra letra con frecuencia alta.

Si la letra “o” fuera la “e”, entonces “q”=”a”, “w”=g”, “g”=”q” y “c”=”m”; volviendo a comparar con la tabla de frecuencias se observó que la letra “a” tiene una frecuencia muy alta y a pesar de que la frecuencia de las demás letras esta entre lo normal, al compararlo con los resultados anteriores, se deduce que es mas probable que la letra “o” sea la “e” y por lo tanto nuestra primera letra de llave sería la “o”. Para descubrir las demás letras de la llave se seguiría el mismo procedimiento.

Como se puede observar, el método anterior requiere analizar las letras que tengan una frecuencia alta y no siempre la letra que tiene la mayor frecuencia es la que corresponde a la llave por lo que no es muy útil para ser aplicado en un algoritmo, sin embargo, se utiliza mucho cuando el procedimiento se hace a mano ya que no requiere de muchos cálculos.

Segundo método:

Para encontrar las letras que corresponden a la llave, este método realiza los siguientes cálculos:

1) Obtiene la frecuencia de las letras en el texto cifrado por columnas, para esto sólo divide la frecuencia de cada letra entre el total de letras contadas.

	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	Letras contadas
1	5	8	10	2	1	8	10	7	4	1	0	0	1	0	10	0	10	9	10	1	4	0	15	0	0	5	121

Ejemplo: 5 / 121 = 0.04; La frecuencia de todas las letras para la columna 1 sería:

a	b	c	d	e	f	g	h	i
0.041	0.066	0.082	0.016	0.008	0.066	0.082	0.057	0.033
j	**k**	**l**	**m**	**n**	**o**	**p**	**q**	**r**
0.008	0	0	0.008	0	0.082	0	0.082	0.074
s	**t**	**u**	**v**	**w**	**x**	**y**	**z**	
0.082	0.008	0.033	0	0.123	0	0	0.041	

2) La frecuencia obtenida de todas la letras en la columna se almacena en un vector (***W***), el cual, se multiplica punto a punto por el vector que contiene la tabla de frecuencias de caracteres del idioma (**A**). Este paso se repite 26 veces desde **A[0] hasta A[25]** con la variante de que el vector de la tabla de frecuencias, es rotado circularmente una posición a la derecha por cada ciclo.

El objetivo de esto es obtener todos los posibles valores por cada rotación y así conocer de manera precisa, cual es la letra que corresponde a la columna de la llave.

W1 = (0.041,0.066,0.082,0.016,0.008,0.066,0.082,...,0.123,0,0,0.041)
A[0] = (0.082,0.015,0.028,0.043,0.127,0.022,0.020,...,0.023,0.001,0.020,0.001)

A[1] = (0.001,0.082,0.015,0.028,0.043,0.127,0.022,0.020,...,0.023,0.001,0.020)
A[2] = (0.020,0.001,0.082,0.015,0.028,0.043,0.127,0.022,0.020,...,0.023,0.001)
A[3] = (0.001,0.020,0.001,0.082,0.015,0.028,0.043,0.127,0.022,0.020,...,0.023) , etc.

El resultado de las multiplicaciones punto a punto para la columna 1 sería:

W1.A[0]	**W1.A[1]**	**W1.A[2]**	**W1.A[3]**	**W1.A[4]**	**W1.A[5]**	**W1.A[6]**	**W1.A[7]**	**W1.A[8]**
0.0381	0.0395	0.0397	0.0468	0.0363	0.0342	0.0313	0.0303	0.0380
W1.A[9]	**W1.A[10]**	**W1.A[11]**	**W1.A[12]**	**W1.A[13]**	**W1.A[14]**	**W1.A[15]**	**W1.A[16]**	**W1.A[17]**
0.0405	0.0409	0.0325	0.0389	0.0431	0.0615	0.0387	0.0334	0.0288
W1.A[18]	**W1.A[19]**	**W1.A[20]**	**W1.A[21]**	**W1.A[22]**	**W1.A[23]**	**W1.A[24]**	**W1.A[25]**	
0.0467	0.0304	0.0354	0.0344	0.0371	0.0366	0.0444	0.0422	

Como se puede observar, la multiplicación de las frecuencias en el desplazamiento 14 dio el valor mas grande y sin dudar se puede decir que esta es la letra que corresponde a la llave en la primer columna.

La letra 14 según la tabla 1, es la letra “o”, por lo que, el primer carácter de la llave es esta letra. Para encontrar los demás caracteres de la llave se tienen que volver a realizar los cálculos anteriores, pero ahora enfocados a la columna siguiente, esto se repite hasta encontrar todos los caracteres que forman la llave.

Implementación y/o desarrollo

La herramienta esta diseñada para permitir descifrar cualquier texto que haya sido encriptado con el método de Vigenére, y fue elaborada en ANSI C para MS-DOS, por lo que el tipo de programación es secuencial. La organización del código se encuentra estructurada, esto es, que la herramienta implementa funciones que permiten al programa resolver pequeñas partes del problema, para poder después, enfocarse al problema en general.

Se eligió ANSI C, debido a que este lenguaje de programación es de los mas veloces a la hora de la ejecución, lo cual es necesario para poder descifrar textos muy grandes en periodos de tiempo pequeños, y además cuenta con la características de poder permitir que el código fuente sea compilado en cualquier plataforma ya que la mayoría lo soportan.

Como ya se mostró en la sección anterior, la eficiencia de esta herramienta, se encuentra en la velocidad y precisión con que encuentra la longitud de la llave y las letras que la conforman; Por lo tanto, a continuación se describe la manera en que operan estos algoritmos:

Algoritmo para encontrar la longitud de la llave

Input: Una cadena TextoC con todas la letras del texto cifrado

1. Para x desde 0 hasta longitud(TextoC)-1 hacer
2. Contador = 0
    3. Para y desde 0 hasta longitud(TextoC)-1 hacer
    4. Si TextoC[y] = TextoC[x+y] hacer
        5. Contador = Contador + 1
6. Frecuencia[x] = Contador

Esta función recibe como parámetro el texto cifrado y se encarga de compararlo con el mismo texto pero con un desplazamiento “x”, su objetivo principal, es guardar todas las frecuencias en un vector. La posición en el vector que tenga el mayor número de frecuencias, es tomada como la longitud de la llave.

Algoritmo para encontrar la frecuencia de las letras por columna

Para encontrar los caracteres que forman la llave, es necesario obtener un vector con la frecuencia por columnas (n columnas, donde n es la longitud de la llave). El siguiente algoritmo indica como obtenerlo:

Input: Una cadena TextoC con todas la letras del texto cifrado
El numero de la columna donde se va a hacer el análisis de frecuencia
Un vector Vfrec donde se va a almacenar la frecuencia obtenida de cada letra.

1. Para x desde columna hasta longitud(TextoC) hacer

    2. posición = TextoC[x] – 97
    3. Vfrec[posición] = Vfrec[posición] + 1
    4. LetrasContadas = LetrasContadas
5. Para x desde 0 hasta 25 hacer
    6. Vfrec[x] = Vfrec[x] / LetrasContadas

En el vector Vfrec quedan almacenadas las frecuencias de cada una de las letras del alfabeto que se encontraron en esa columna, en relación a todas las letras del texto cifrado.
Nota: En el paso 2, al carácter obtenido del texto cifrado se le resta 97 para que coincida con la numeración de las letras del alfabeto de la *tabla 1.*

Algoritmo que descubre un carácter de la llave

Input: El vector con las frecuencias obtenidas en el algoritmo anterior
Output: La posición del carácter de la llave correspondiente a la *tabla 1*.

1. Para desplazamiento desde 0 hasta 25 hacer
    2. Resultado = 0
    3. Para x desde 0 hasta 25 hacer
        4. posición = (x+desplazamiento) mod 26
        5. Resultado = Resultado + (Vfrec[posición] * InglesFrecuencia[x])
    6. Vresultado[desplazamiento] = Resultado
7. Mayor = Vresultado[0]
8. posición = 0
9. Para x desde 1 hasta 25 hacer
    10. Si Mayor < Vresultado[x]
        11. Mayor = Vresultado[x]
        12. posición = x
13. Return posición

Este algoritmo multiplica el vector de las frecuencias de las letras de una columna con cada uno de los vectores que contienen las frecuencias del alfabeto en ingles, calculando esta multiplicación punto a punto 26 veces (de 0 a 25), en donde por cada multiplicación, el vector de las frecuencias del alfabeto es recorrido circularmente hacia la derecha un carácter a la vez.

Los resultados de las multiplicaciones por cada desplazamiento son almacenados en otro vector, en el que para finalizar, se tiene que buscar el numero con mayor frecuencia, en donde la posición en la que se encuentre pertenece a la posición del carácter que corresponde a la llave.

NOTA: Para obtener todos los caracteres de la llave, es necesario colocar los algoritmos anteriores en un ciclo que va desde 0 hasta la longitud de la llave –1.

Resultados y Conclusiones

Como resultado de este proyecto se obtuvieron dos programas:

- El primero se llama CIFRAR.EXE el cual es una herramienta que permite cifrar cadenas de caracteres escritas por un usuario mediante el método de Vigenére.
- El segundo se llama VIGENERE.EXE y es el programa más importante ya que tiene como objetivo descifrar textos cifrados por el método de Vigenére,

Los dos programas están diseñados para ser ejecutados en plataformas con sistema operativo Ms-Dos o Windows, sin embargo debido a que el código esta hecho en C, se puede aprovechar las características que este lenguaje ofrece y por lo tanto se puede trasladar a plataformas Linux de manera sencilla.

En cuanto a la medición de velocidad del programa para descifrar texto, no es necesario realizarla debido a que esta es despreciable, ya que el programa no realiza cálculos complejos como otros sistemas criptográficos modernos y su tiempo de descifrado depende sólo de la longitud del archivo.

La eficiencia del programa para descifrar textos, depende mucho del tamaño del archivo y la frecuencia con la que aparece la letra “e” dentro del texto, esto es, si en un archivo con 150 caracteres, es muy frecuente la letra “e”, la probabilidad de encontrar la llave aumenta considerablemente.

Los resultados en cuanto a la eficiencia en el descifrado fueron los siguientes:

1. Es 100% eficiente el programa a partir de los 300 caracteres.
2. Cuando no hay muchas letras “e” el programa necesita hasta de 400 caracteres para obtener un buen resultado.
3. La cantidad mínima de caracteres en el texto para poder aplicarle criptoanálisis es de 253.
4. Se trato de representar lo más fiel posible al método de Vigenére por lo que no existen caracteres fuera del alfabeto ingles como la “ñ”, espacios y símbolos.

Actualmente la aplicación cuenta con las siguientes ventajas y desventajas:

Ventajas

- Obtiene la llave de un texto cifrado de manera automática y sin necesitar ayuda del usuario.
- Obtiene un 100% de éxito al atacar textos cifrados grandes
- El código esta escrito en ANSI C, por lo que es portable para la mayoría de las plataformas y veloz en su ejecución.

Desventajas

- Debido a que el código representa fielmente la técnica de cifrado por Vigenére, este no es capaz de reconocer caracteres que no pertenezcan al alfabeto, tales como, el espacio en blanco, los acentos, números, la letra “ñ”, etc.

Referencias

[1] Introduction to Cryptography with Coding Theory
Lawrence C. Washingtong

[2] Criptografía y Seguridad en Computadores
Manuel José Lucena López, Tercera Edición, Marzo 2002

[3] Handbook of Applied Cryptography
A. Menezes, P. Van Oorschot, and S. Vanstone, CRC Press, 1996.

Direcciones electrónicas

[4] http://islab.oregonstate.edu/koc/ece575/02Project/Mun+Lee/VigenereCipher.html
[5] http://www.geocities.com/Eureka/3999/e_vigenere.htm
[6] http://raphael.math.uic.edu/~jeremy/crypt/vignere.html
[7] http://nob.cs.ucdavis.edu/~bishop/classes/ecs153-1997-winter/vigenere.html
[8] http://astro.ocis.temple.edu/~dhill001/vigenere/vigenere.html