“Cifrado Clásico”

Trabajo Preparatorio N°11

Laboratorio de Seguridad en Redes

Melanny Dávila
Ingeniería en Telecomunicaciones
Facultad de Eléctrica y Eléctronica
Quito, Ecuador
melanny.davila@epn.edu.ec

Alejandra Silva
Ingeniería en Telecomunicaciones
Facultad de Eléctrica y Eléctronica
Quito, Ecuador
alejandra.silva@epn.edu.ec

***Abstract*—En el presente documento se presenta el fundamento teórico acerca de los algoritmos de cifrado, con el fin de poder probar varios algoritmos de cifrado clásico y poder implementar uno de estos de manera sencilla.**

***Index Terms*—Algoritmo, Cifrado, Clásico, Implementar, Sencilla**

I. Introducción

Entre los diferente tipos de algoritmos de cifrado, se tiene uno que es el mas básico y sencillo que es por sustitución, se basa en que las unidades de texto son sustituidas con texto cifrado, existen dos alfabetos en esta sustitución monoalfabético que usa una sustitución fija y polialfabético que usa diferentes sustituciones en diferentes mecanismo del mensaje.

Existen otro dos tipos de cifrado,el primero es cifrado Vigenere es un criptosistema simétrico, es decir, utiliza la misma clave para cifrar y descifrar. Se trata de un algoritmo de cifrado simétrico polialfabético, es decir, es basado en diferentes series de caractéres o letras del cifrado Cesar formando estos caracteres una tabla y finalmente el cifrado Hill este sistema es polialfabético pues puede darse que un mismo carácter en un mensaje a enviar se encripte en dos caracteres distintos en el mensaje encriptado.

II. Objetivos

- Probar varios algoritmos de cifrado clásico
- Implementar un algoritmo sencillo de cifrado clásico.

III. Cuestionario

A. Descargar el software de prueba de algoritmos de cifrado clásico del siguiente enlace: http://www.criptored.upm.es/software/sw_m001c.htm

B. Consultar los mecanismos de criptoanálisis de los algoritmos de cifrado por sustitución (máximo 1 hoja).

Es un método de cifrado por el que unidades de texto son sustituidas con texto cifrado siguiendo un sistema regular; las ”unidades” pueden ser una sola letra o pares de letras, tríos, mezcla de lo anterior, entre otros. El receptor descifra el texto realizado la sustitución inversa.

Las unidades del texto plano mantienen el mismo orden, lo que se cambia son las propias unidades del texto plano, existen diversos tipos de cifrado, el mas común es el cifrado por sustitución simple, que opera sobre un grupo de letras denominado polígrafo mientras que un cifrado monoalfabetico usa una sustitución fija para todo el mensaje y el polialfabeto usa diferentes sustituciones en diferentes momentos del mensaje. [1]

Las ventajas de usar un cifrado César incluyen: [2]

- Uno de los métodos más fáciles de usar en criptografía y puede proporcionar una seguridad mínima a la información.
- Uso de solo una tecla breve en todo el proceso.
- Uno de los mejores métodos para usar si el sistema no puede usar ninguna técnica de codificación complicada.
- Requiere pocos recursos informáticos.

Matemáticamente hablando se define las siguientes operaciones: [2]

- Ordinal de una letra: Es la posición que ocupa la letra en el alfabeto que usamos. Se denota como: ORD (Letra) [ORD (A) = 0].
- Carácter de un número: Es la letra que ocupa la posición del número que tenemos. Se denota como: CAR (x) [donde “x” tiene que ser mayor o igual a cero, y menor o igual al número de letras del alfabeto usado].

Ahora las formulas que se necesitan para el cifrado y descifrado son las sigueintes

- Cifrado: C(x) = x + k (mod T)
- Descifrado: D(x) = x – k (mod T)

(Siendo “x” el ORD (Letra a cifrar), “k” es el número de posiciones a desplazar y “T” es el total de letras en el alfabeto a usar).

Se puede observar que la clave de este cifrado es el número de posiciones a desplazar las letras (k).

C. Consultar el cifrado, descifrado, los mecanismos de criptoanálisis del algoritmo de cifrado Vigenère (máximo 1 hoja).

Es un criptosistema simétrico, es decir, utiliza la misma clave para cifrar y descifrar. Se trata de un algoritmo de cifrado simétrico polialfabético, es decir, es basado en diferentes series de caractéres o letras del cifrado César formando estos caracteres una tabla [3].

En primer lugar, se asocia cada letra con una cifra correspondiente.

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
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

Fig. 1. Tabla de asociación

Se codifica un texto con una palabra agregándole las letras de otra palabra (llamada palabra clave) a cada una de sus letras. La palabra clave se agrega indefinidamente en el texto que se va a cifrar, y después se agrega el código ASCII de cada una de las letras de la palabra clave al texto a cifrar [3]. Por ejemplo, el texto ”rendezvousamidi” con su palabra clave ”bonjour” será codificado de la siguiente manera:

r	e	n	d	e	z	v	o	u	s	a	m	i	d	i
114	101	110	100	101	122	118	111	117	115	97	109	105	100	105

Fig. 2. Texto original

b	o	n	j	o	u	r
98	111	110	106	111	117	114

Fig. 3. Palabra cable

\r+B	E+O	\n+\n	-d+J	E+O	u	+\r	O+B	u+O	s	-a+J	M+O	i	-d+\r	i
114+98	101+111	110+110	100+106	101+111	122+117	118+114	111+98	117+111	115+110	97+106	109+111	105+117	100+114	105+98

Fig. 4. Encriptación final

Para descifrar este mensaje, se necesita la clave secreta y se realiza el descifrado inverso utilizando la sustracción.

A pesar de que el cifrado es mucho más sólido que el cifrado César, aún así se puede romper fácilmente. Cuando los mensajes son mucho más largos que la palabra clave, es posible identificar el largo de la palabra clave y utilizar, para cada secuencia de palabra clave, el método de cálculo de la frecuencia con que aparecen las letras, y determinar así los caracteres de las palabras claves una a la vez. Para evitar este problema, una solución es utilizar una palabra clave que sea casi igual de larga como el texto, a fin de evitar un estudio estadístico del texto cifrado. Este tipo de sistema de cifrado se llama sistema one-time pad [4]. El problema con este tipo de método es la longitud de la clave de cifrado (cuanto más largo el texto a ser cifrado, más grande deberá ser la clave) que impide su memorización e implica una probabilidad mucho más grande de errores en la clave (un solo error hace que el texto sea imposible de leer).

- Método Kasiski: es un método de criptoanálisis aplicado a los cifrados de sustitución polialfabéticos. El mismo que es aplicable si el mensaje es lo suficientemente largo. La existencia de secuencias de caracteres repetidos en el texto cifrado, lo cual significa casi con toda probabilidad que dichas secuencias no sólo eran la misma antes del cifrado sino que además la clave debía coincidir en la misma posición, por lo que lo primero que se debe hacer es detectar secuencias de letras cifradas repetidas. Posterior a eso se puede pensar que el número de caracteres de la clave puede ser el mcd(separación entre posiciones) $= a$.Es decir, la longitud más probable de la clave es a que es el máximo común divisor o mayor número entero que divide a todos estos números (posiciones) sin dejar resto. Una vez que se ha averiguado la longitud de la clave (L), sin equivocación, lo siguiente que se debe hacer es dividir el criptograma en L subcriptogramas (en nuestro caso a), ya que estos han sido cifrados por una misma letra de la clave, a partir de lo que estaría en disposición de realizar un ataque simple de tipo estadístico monoalfabético. De donde se obtendrá la palabra clave, utilizando la tabla y siguiendo los pasos para descifrar este tipo de mensajes.
- Prueba por índice de coincidencia: consiste en extraer una letra de n en el mensaje y medir el índice de coincidencia. Cuanto mayor es, mayor es la probabilidad de que n sea la longitud de la clave. De hecho, tomar una letra en n cuando n es la longitud de la clave equivale a tomar una serie de letras cifradas siempre cifradas con el mismo desplazamiento, por lo tanto, el índice de coincidencia es igual al del texto sin formato [4].

D. Consultar el cifrado, descifrado, los mecanismos de criptoanálisis del algoritmo de cifrado Hill (máximo 1 hoja).

Este sistema esta basado en el álgebra lineal y ha sido importante en la historia de la criptografía. Fue Inventado por Lester S. Hill en 1929, y fue el primer sistema criptográfico polialfabético que era práctico para trabajar con mas de tres símbolos simultáneamente. Este sistema es polialfabético pues puede darse que un mismo carácter en un mensaje a enviar se encripte en dos caracteres distintos en el mensaje encriptado.Las letras se numeran en orden alfabético de forma tal que $A = 0$, $B = 1$, ... ,$Z = 25$, se elije un entero d que determina bloques de d elementos que son tratados como un vector de d dimensiones [5]. También, se elije de forma aleatoria una matriz de d $\times$ d elementos los cuales serán la clave a utilizar. Los elementos de la matriz de d $\times$ d serán enteros entre 0 y 25, además la matriz M debe ser inversible. Para la encriptación, el texto es dividido en bloques de d elementos los cuales se multiplican por la matriz d $\times$ d; todas las operaciones aritméticas se realizan en la forma modulo 26, es decir que $mod26 = 0$, $mod27 = 1$, $mod28 = 2$ etc. Dado un mensaje a encriptar se debe tomar bloques del mensaje de ”d” caracteres y aplicar: $M \times Pi = C$, donde C es el código cifrado para el mensaje Pi [5].

La gran vulnerabilidad de este criptosistema radica en que es muy débil ante un ataque con texto claro conocido, es decir, si el criptoanalista conoce parte del texto en claro correspondiente al texto cifrado del que dispone no tendrá mayor problema para obtener la matriz K con la que se cifró esa parte del texto en claro y, por tanto, estaría en disposición de descifrar todos los mensajes cifrados con dicha clave (K). Esta vulnerabilidad se debe a la linealidad de este criptosistema, por lo que con texto claro conocido y empleando el método de Gauss Jordan no es muy difícil obtener la matriz clave (K) [6].

Sin embargo, Hill plantea a los criptoanalistas problemas mucho mayores a los que planteaba Cesar o Vigenère. Para empezar el espacio de claves es mucho mayor, en este caso es de aC25, es decir las permutaciones de a elementos tomados de entre 25 posibles y usando una matriz mas grande la cantidad de posibles claves se puede hacer tan grande como sea necesario para hacer que sea imposible un ataque por fuerza bruta. Lo mejor que puede hacer un criptoanalista es tratar de conseguir un código para el cual se conozca una parte del mensaje. Y ver si con ambos datos es capaz de encontrar cual fue la matriz utilizada para encriptar el mensaje [6].

REFERENCES

[1] ”Cifrado por sustitucion” (2012).https://es.slideshare.net/GAlbertoHoyos/cifrado-por-sustitucion. accedido (24,ago,2021).

[2] ”¿Qué es el cifrado César y cómo funciona?”.https://ayudaleyprotecciondatos.es/2020/06/10/cifrado-cesar/. accedido (24,ago,2021).

[3] NIHIL, “Criptografía en Python: Cifrado cesar y vigenère”, La cripta del hacker, oct. 25, 2020. https://lacriptadelhacker.wordpress.com/2020/10/25/criptografia-en-python-cifrado-cesar-y-vigenere/ (accedido ago. 25, 2021).

[4] “El cifrado de Vigenère”. https://www.ugr.es/ anillos/textos/pdf/2011/EXPO-1.Criptografia/02a11.htm (accedido ago. 25, 2021).

[5] “Criptosistema Hill”, jun. 26, 2005. https://www.textoscientificos.com/criptografia/hill (accedido ago. 25, 2021).

[6] “Criptografía (XXIII): cifrado de Hill (I), Criptografía (XLIX): el algoritmo DES (I)”, Criptografía (I). https://mikelgarcialarragan.blogspot.com/2015/03/criptografia-i.html (accedido ago. 25, 2021).