Nombre del alumno:

Ignacio Ivan Sanchez Pantoja

Número de matrícula:

18108365

Nombre del profesor:

Israel Alejandro Herrera Araiza

Nombre del curso:

Controles Criptograficos De Seguridad

Actividad:

Fase 1 Actividad de clase Cifrados clásicos y RSA.

Fecha:

24/05/2024

Introducción a los Controles Criptográficos de Seguridad

La criptografía ha sido una herramienta esencial para la protección de la información desde tiempos inmemoriales. Desde los antiguos jeroglíficos egipcios hasta los complejos algoritmos modernos, la criptografía ha evolucionado significativamente, adaptándose a las necesidades de cada era. En este contexto, los controles criptográficos de seguridad juegan un papel crucial en la protección de datos sensibles, asegurando la confidencialidad, integridad y disponibilidad de la información. La criptografía clásica se refiere a los métodos tradicionales de cifrado utilizados desde la antigüedad hasta la era moderna.

A lo largo de este documentos se explora el funcionamiento así como el uso de tanto el cifrado cesar como Vigenère así como el cifrado Hill.

Podemos declarar que la criptografía se utiliza para proteger datos valiosos, asegurando que solo las personas autorizadas puedan acceder a la información. Esto es especialmente crucial en tiempos de conflicto, donde la transmisión segura de mensajes sensibles puede ser la diferencia entre el éxito y el fracaso, como sabemos las técnicas criptográficas previenen que la información caiga en manos enemigas, protegiendo así la seguridad nacional y la privacidad personal. Sin embargo, los métodos de cifrado no son infalibles y pueden ser vulnerables al criptoanálisis, donde los atacantes intentan descifrar la información sin conocer la clave.

En este sentido los criptosistemas clásicos son los mas vulnerables a las técnicas de criptoanálisis, como se exploro en el primer entregable, existen múltiples técnicas que pospone Shannon para poder realizar el análisis de los criptosistemas.

Esquemas Básicos de Comunicación

En el proceso de comunicación en criptografía, se incluyen tres actores principales: el emisor, el receptor y el atacante. Se utiliza una clave para cifrar y descifrar el mensaje, con distintos esquemas según se trate de criptografía simétrica o asimétrica. La criptografía no solo ayuda a mantener la confidencialidad, sino también la integridad y la disponibilidad de la información.

Cifrados Monoalfabeticos empleados en la criptografía clásica:

Uno de los cifrados monoalfabeticos empleados en la criptografía clásica y que posteriormente se vieron en clase es el cifrado cesar, que anteriormente se discutió en el entregable previo, el procedimiento para realizar el cifrado de datos consiste del siguiente:

1) Es necesario definir el alfabeto a utilizar:

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
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

2) Después es necesario codificar el texto en base al alfabeto definido

MENSAJE:	I	V	A	N	S	A	N	C	H	E	Z
	8	21	0	13	18	0	13	2	7	4	25

3) Posteriormente es necesario cifrar el texto realizando un desplazamiento de + 3 caracteres y si el resultado es mayor al alfabeto definido es necesario realizar el operador de modular con el largo del alfabeto para adaptar el resultado a un rango admitido de valores

CIFRADO	0	13	18	5	10	18	5	20	25	22	17

4) El proceso de descifrado es similar es decir es necesario codificar el criptograma, utilizando el alfabeto definido y después restar en este caso 3 caracteres, al tener una cantidad de desplazamiento fija es por ello que se considera un cifrado de tipo monoalfabetico, posterior a realizar la resta de los datos se desprenderán los valores equivalentes al mensaje original, sin embargo es necesario aplicar el operador modular para asegurarnos de no manejar valores negativos:

DESCIFRAD	8	21	0	13	18	0	13	2	7	4	25

Cifrados Polialfabéticos empleados en la criptografía clásica:

un ejemplo de un cifrado que hace uso la sustitución poligráfica es decir que dependiendo de la posición de un carácter en el texto este variara, se desplazara o se utilizara otro alfabeto para cambar la relación del texto plano con el mensaje, tanto el cifrado Hill como Vigenere

Explicación del modo de funcionamiento cifrdo Vigenere

lo primero que hay que tener en cuenta para explicar el cifrado vigenere es el definir el alfabeto que utilizara el cifrado, en este caso se optara por hacer uso del alfabeto ingles como se muestra en la figura:

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
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

Posterior a la definición de nuestro alfabeto es necesario tokenizar o transformar el conjunto de caracteres que conforman el mensaje destinado a cifrar, este mismo proceso es necesario realizarlo también con la clave que será utilizada para cifrar los datos:

TEXTO PLANO																		
I	V	A	N	S	A	N	C	H	E	Z		P	A	N	T	O	J	A
8	21	0	13	18	0	13	2	7	4	25		15	0	13	19	14	9	0

En este caso es mi nombre y mi apellido paterno, mi apellido materno será la clave

p	a	n	t	o	j	a	p	a	n	t
15	0	13	19	14	9	0	15	0	13	19
										TEXTO PLA
I	V	A	N	S	A	N	C	H	E	Z
8	21	0	13	18	0	13	2	7	4	25

Si nos damos cuenta hay un problema y es que la clave debe tener la misma extensión o longitud que el texto a cifrar, para solucionar este inconveniente se puede hacer uso del denominado **periodo criptográfico** esto como es definido por Schneier no es más que repetir los caracteres pertenecientes de la clave una n cantidad determinada de veces para alcanzar la misma longitud que el texto a cifrar o mensaje, desgraciadamente el uso de un periodo criptografico puede involucrar serios inconvenientes, la discusión de dichos inconvenientes queda fuera del alcance de este documento.

Posterior a la generación de una clave con un largo equivalente a la cardinalidad del conjunto de caracteres del mensaje

Se realizara la operación de cifrado de la siguiente forma:

A cada valor del numero que representa tanto el carácter se sumara con cada valor del número que representa el mensaje para después realizar la operación de modular para evitar que si existen valores que sobrepasen el alfabeto empleado, no puedan ser interpretados al momento de codificar el mensaje en el alfabeto dado:

p	a	n	t	o	j	a	p	a	n	t
15	0	13	19	14	9	0	15	0	13	19
I	V	A	N	S	A	N	C	H	E	Z
8	21	0	13	18	0	13	2	7	4	25
					=					
23	21	13	32	32	9	13	17	7	17	44
					mod 26					
23	21	13	6	6	9	13	17	7	17	18
					resultado					
X	V	N	G	G	J	N	R	H	R	S

Como se puede observar en la figura primero se realiza la adición de los valores por ejemplo 19 y 25, esto da como resultado 44 sin embargo al nuestro alfabeto solo poseer 27 caracteres es necesario realizar la operación de modular n -1 es decir 26 para adaptar el resultado

Proceso de descifrado:

las operaciones de descifrado siempre serán la operación análoga a el procedimiento de cifrado utilizado, en este sentido basta con restar en el carácter cifrado previamente tokenizado con nuestra clave utilizada:

p	a	n	t	o	j	a	p	a	n	t
15	0	13	19	14	9	0	15	0	13	19
X	V	N	G	G	J	N	R	H	R	S
23	21	13	6	6	9	13	17	7	17	18
					=					
8	21	0	-13	-8	0	13	2	7	4	-1
					mod 26					
8	21	0	13	18	0	13	2	7	4	25
					resultado					
I	V	A	N	S	A	N	C	H	E	Z

Podemos realzar un ejemplo con los caracteres P de la clave y R de la posición 8:

17 - 15= 2%26 = 2

Esto corresponde al carácter C

como se denoto en el primer entregable previo se considera que este cifrado solo es la aplicación de la propiedad aditiva en la aritmética modular, de hecho es prácticamente la aplicación de 8 cifrados cesar.

Explicación del modo de funcionamiento cifrado Hill:

de la misma forma que con el cifrado Vigenere es necesario definir un 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

Posterior a ello es necesario codificar cada carácter en base al alfabeto:

I	V	A	N	S	A	N	C	E	Z
8	22	0	13	19	0	13	2	4	26

Sin embargo esta cadena de caracteres debe generar una matriz simétrica es decir que sus filas y columnas (mxn) tengan la misma cantidad o por lo menos que su producto punto sea posible realizarlo con la matriz de claves, por lo tanto, es necesario agregar caracteres nulos o una secuencia de caracteres pad para poder completar el tamaño de una matriz simétrica, otra alternativa puede ser utilizar el periodo criptográfico para repetir una contraseña con el fin de generar una simetría entre una cantidad del texto en caso de que la contraseña tenga un tamaño menor que el texto a cifrar.

Una regla del algebra lineal es que para realizar una multiplicación producto punto es necesario que n de la matriz A sea igual a m de la matriz B es decir que el numero de filas de la matriz A sea igual al número de columnas de la matriz B, esto producirá una matriz compuesta del numero de filas de la matriz A x el numero de columnas de la matriz B.

En este caso dado que la matriz de claves es de 3x3

	16	20	0
CLAVE	0	15	16
	13	9	0

Debemos de hacer que el vector mensaje tenga una cantidad de filas similar a la cantidad de columnas de la matriz clave:

I	V	A	N	S	A	N	C	E	Z
8	22	0	13	19	0	13	2	4	26

En ese sentido cada imagen diferente representara una columna, faltando 2 caracteres que se rellenaran con la letra 0 (carácter A), también es posible introducir un nuevo signo para este cometido.

Comprobemos que es posible realizar la operación al ver como se constituye la matriz mensaje:

	8	19	4
A=	22	0	26
	0	13	0
	13	2	0

Si la matriz de claves tiene un tamaño de 3x3(maxna) y esta matriz un tamaño de 4x3(mbxnb)

Entonces (na=mb) que es correcto.

Al realizar la multiplicación de matrices Matriz Clave x Matriz Mensaje se cifrarán los datos, esto hará que el resultado sea:

		180	481	304
B*A	=	690	674	0
		0	195	208
		208	290	32

Sin embargo aun es necesario adaptar este resultado a una forma en la cual pueda ser codificado en base a nuestro alfabeto seleccionado, con base a esta necesidad, se utilizara el operador modular en cada celda del criptograma y después se buscara una equivalencia para codificarlo:

18	22	7
15	26	0
0	6	19
19	20	5

Codificado en el alfabeto seria:

R	V	H
O	Z	A
A	G	S
S	T	F

Lo cual se puede traducir en el string o cadena de caracteres como:

R	O	A	S	V	Z	G	T	H	A	S	F

Ahora bien como se había comentado anteriormente la operación de descifrado no es mas que la realización del inverso de la operación que se realizó para cifrar, esto involucra tener que realizar los pasos del tokenizado o codificado del texto en el alfabeto definido y convertir tanto el criptograma como el conjunto de caracteres de contraseñas en una matriz que sean compatibles, enfatizando en el proceso de descifrado que cambia en relación al proceso de cifrado en este caso se debería aplicar el inverso de una multiplicación es decir pues la división matricial, pero las propiedades matriciales imperan que una división de matrices se debe realizar de la siguiente forma:

$$\frac{C}{B} = C * (B^{-1})$$

Donde C es el criptograma y B el inverso de la matriz clave, para obtener el inverso de una matriz existen un centenar de métodos entre ellos el método de la matriz adjunta, el cual es definido como

$$B^{-1} = \frac{1}{|B|} * adj(B')$$

Es decir la división del valor determinante de la matriz de claves por el producto punto de la matriz transpuesta de b es equivalente a el inverso de la matriz de claves, si bien es importante realizar estas definiciones matemáticas en la práctica gracias a Excel basta con utilizar una función que calculara automáticamente la matriz inversa:

MATRIZ INVERSA		
-0,07759	0	0,17
0,11207	0	-0,1
-0,10506	0,06	0,13

Esto se realizo con la función MINVERSA(C5:E7), posterior a esto ahora si podemos realizar la definición de la división de 2 valores que pospone el algebra lineal, en este caso de la matriz criptograma y la matriz de claves, la cual desprenderá el resultado:

8	19	4
22	0	26
-0	13	0
13	2	0

El cual es equivalente a nuestro mensaje en texto plano una vez decodificado en base al alfabeto definido:

Cifrado RSA

El cifrado RSA es un cifrado que si bien ya no puede ser considerado parte de la criptografía clásica debido al hecho de que este cifrado es clasificado como un cifrado asimétrico (es decir que utiliza 2 claves diferentes una publica y una privada) una publica que puede ser compartida en la red y una privada que se utiliza para cifrar y firmar los mensajes, a grandes rasgos el cifrado RSA a pesar de ser simple ha sobrevivido a lo largo de los años en muchos casos porque "RSA se basa en la dificultad para factorizar grandes números. Las claves pública y privada se calculan a partir de un número que se obtiene como producto de dos primos grandes. El atacante se enfrentara, si quiere recuperar un texto claro a partir del criptograma y la clave pública, a un problema de factorización tendrá que resolver un logaritmo discreto" (LOPEZ, 2011)

Actualmente los algoritmos más eficaces para realizar la factorización de número primos grandes es el algoritmo de **algoritmo de Shor y el algoritmo de criba general de cuerpos numéricos** como se menciona en Lopez.

Para emplear el cifrado RSA es necesario primero determinar 2 números primos, a estos se les denomina como p y q y para realzar este proceso se pueden utilizar cualquiera de las existentes formas para comprobar si un número es primo:

Por ejemplo en este caso se seleccionó el uso de 29 y 31

Posterior a ello debemos calcular n el cual no es más que la multiplicación de ambos números primos:

N = 899

Posterior a ello es necesario determinar el PHI de N esto es conocido como la **Función totiente de Euler** la cual se define como:

$$Phi(n) = (p-1)(q-1)$$

Y nos sirve para poder determinar la inversa de un número dada la condición de que (E será definido más adelante):

$$MCD(E, Phi(n)) = 1$$

En ese sentido es necesario escoger un numero E primo de tal forma que este **número sea la clave publica en conjunción con N,** el funcionamiento debe ser este MCD(E,PHI(N)) = 1, sin embargo en la práctica podemos cumplir esta condición lo cual involucra que E es el inverso de phi(n) es decir si usamos phi(n) mod e nos dará 1 "el elemento simétrico o opuesto para el producto" (LOPEZ, 2011)) lo cual puede ser expresado como:

$$E + PHI(N) \equiv 1\ MOD\ 2\ \#\ esta\ es\ la\ definicion\ de\ simetrico$$

Una pregunta sobre la siguiente expresión podría ser ¿por que usamos 2?, bueno en realidad esto se radica en que en la aritmética modular trabajamos con un arreglo de valores, este arreglo de valores en este caso solo serán 2 elementos E y PHI(N) por lo tanto esto nos debería de dar 1 suponiendo que E y PHI(N) son 3016 y 9 (mas a delante se explicara de donde salen estas cifras):

$$(3016 + 9) mod\ 2 \ = \ 1$$

Ahora bien este digito E es forzosamente necesario que E **multiplicado por otro número denominado d** al momento de realizar la operación modular de PHI(N) sea 1 (matemáticamente conocido como congruencia) y este procedimiento implica que debemos encontrar un numero **d que permita realizar la operación de congruencia**, se realiza este procedimiento para poder decodificar el mensaje codificado, el numero d con corresponderá a nuestra clave privada es decir la que nos sirve para firmar los datos, en este sentido y formalmente Lopez define la siguiente relación matemática para poder encontrar un número que decodifique la clave pública:

$$(d * e) \ \equiv 1\ mod (p-1)(q-1)$$

Desglosado significa:

$$(d * e)\ mod\ phi(n) \ = \ 1$$

Es decir que la multiplicación de d y e debe **resultar en un numero congruente como se había mencionado anteriormente (es decir que la operación de residuo resultante de la función de totient de Euler sea 1)**, Lopez menciona que la calve privada es posible calcularla haciendo uso del algoritmo extendido de Euclides **pero que es necesario conocer el número N es decir el producto de p y q, esto de cierta forma lo realizamos para poder tener una relación entre la clave pública y privada.**

Un ejemplo práctico de esto puede ser con el uso de P y N con valores 53 y 59

$$P \ = \ 53$$

$$Q \ = \ 39$$

De tal forma que:

$$N \ = \ p * q \ = \ 3127$$

Ahora es necesario calcular la función de **totient de Euler**

$$PHI(N) \ = \ (p-1)(q-1) \ = \ 3016$$

Después de esto ahora es necesario encontrar un número e, sin embargo existen unas particularidades extras no mencionadas anteriormente a este punto:

E debe ser mayor a 1 pero menor que phi (N)

Para ello podemos aplicar el siguiente código en python:

```
ev = [e for e,l in zip(range(0,phi),([1]*20)) if gcd(e,phi) == 1]
```

gcd es una función para calcular el mcd es necesario importarla del módulo math

limit en este caso es un límite máximo de ítems a procesar por la lista (se decidió colocar para no obtener todo el rango del número admisible por phi (simplicidad didáctica)), el resultado es el siguiente:

```
>>> ev
[1, 3, 5, 7, 9, 11, 15, 17, 19]
>>>
```

Podemos utilizar en este caso el numero 9

$$E = 9$$

Una vez con este valor es necesario determinar D para que se cumpla la congruencia que se define Lopez, este algoritmo puede ser expresado como:

$$(1 + (k * phi(n)))/e$$

Donde K corresponde a un número real cualquiera:

```
d = [((1+(k*phi))/e) for k in range(0,phi-1)]
```

```
>>> d = [((1+(k*phi))/e) for k in range(0,30)]
>>> print(d)
[0.2, 603.4, 1206.6, 1809.8, 2413.0, 3016.2, 3619.4, 4222.6, 4825.8, 5429.0, 6032.2, 6635.4, 7238.6, 7841.8, 8445.0,
9048.2, 9651.4, 10254.6, 10857.8, 11461.0, 12064.2, 12667.4, 13270.6, 13873.8, 14477.0, 15080.2, 15683.4, 16286.6, 16
889.8, 17493.0]
>>>
```

El numero d debe pertenecer a los números reales es decir debe ser un numero entero

Para este propósito seleccionaremos el número 2413 que se muestra

Entonces:

$$D = 2413$$

Ahora demostraremos la congruencia

$$E * D = 1 \, mod \, phi(n)$$

$$5 * 2413 = 12065 \, \% \, 3016 = 1$$

(en este caso usamos % como símbolo para la operación modular)

Sabemos que la condición se ha cumplido por lo tanto ahora podemos cifrar los datos

Las claves de cifrado públicas y privadas así como sus procedimientos de cifrado/descifrado se ven de finidas como:

Publica (firmado de datos o cifrado):

$$K_{pub}(N, E)$$

$$C = m^e \, mod \, N$$

Privada (descifrado de datos):

$$K_{priv}(N, D)$$

$$M = C^d \ mod \ N$$

Un ejemplo práctico con los datos anteriores y suponiendo que el mensaje son los caracteres que corresponden a “PANTOJA”:

C = [(c**e)%n for c in msj]

[2641, 0, 2307, 2642, 3107, 2763, 0]

RM = [(c**d)%n for c in C]

[15, 0, 13, 19, 14, 9, 0]

```
>>> msj = [15,0,13,19,14,9,0]
>>> C = [ (c**e)%n for c in msj]
>>> RM = [ (c**2413)%3127 for c in C]
```

```
>>> C
[2641, 0, 2307, 2642, 3107, 2763, 0]
>>> RM
[15, 0, 13, 19, 14, 9, 0]
>>>
```

Ahora solo bastaría con decodificar el mensaje según el alfabeto definido que no es mas que el alfabeto que siempre hemos utilizado en los ejemplos anteriores

```
>>> "".join([ascii_uppercase[x%len(ascii_uppercase)] for x in RM])
'PANTOJA'
```

```
from string import ascii_uppercase
C = [ (c**e)%n for c in msj]
RM = [ (c**2413)%3127 for c in C]
"".join([ascii_uppercase[x%len(ascii_uppercase)] for x in RM])
```

Criptoanálisis de Algoritmos Clásicos

El criptoanálisis es la práctica de descifrar códigos sin conocer la clave. En el caso de la sustitución polialfabética, se utilizan métodos como la prueba de Kasiski y la prueba de Friedman para determinar la longitud de la clave y realizar un ataque de fuerza bruta.

La forma para realizar el criptoanálisis al cifrado cesar es utilizando el **índice de coincidencia** que como define schneier no es más que la distancia que existe entre la ocurrencia de 2 caracteres, Friedman describe las tablas de frecuencias de ciertos tipos de cifras tienen características compartidas que se aproximan a las curvas estadísticas que pueden utilizarse para dar solución a una determinada cantidad de cifrados, para poder hacer uso de esta técnica primero debemos determinar el lenguaje en el cual está escrito el criptograma, esto es posible hacerlo utilizando el **índice absoluto de un lenguaje**, el cifrado cesar al ser un cifrado monoalfabetico no oculta la redundancia del idioma del cual está escrito, esto facilita las cosas ya que solo basta con hacer uso de un análisis de variación de frecuencias relativas para encontrar el mensaje.

Sin embargo al ser un cifrado cesar bastaría con determinar el lenguaje en el cual esta escrito el criptograma para efectuar un ataque de fuerza bruta buscando determinar el mensaje más coherente.

En cuestión del cifrado vigenere la forma de efectuar su criptoanálisis se complica ya que se utiliza 8 cifrados cesar, Lopez menciona que los cifrados polialfabeticos no son más que la aplicación de múltiples cifrados monoalfabeticos de una forma cíclica, en base a esto el primer punto de entrada que nos encontraremos como criptoanalistas es el determinar si se está empleando el cifrado vigenere, la forma de realizar esto es utilizando el índice de coincidencia para el criptograma que se tiene, dado que los cifrados polialfabeticos si ocultan la redundancia del texto, para poder realizar el criptoanálisis una vez determinado el texto es necesario aplicar el método de Kasiski para poder determinar la longitud de la clave utilizada, una vez con este hecho es necesario agrupar los caracteres del criptograma en función de la longitud de la posible clave, después es necesario aplicar una análisis de frecuencia para obtener la cantidad de desplazamientos realizados y en función de ello descifrar la clave.

Conclusiones:

La criptografía clásica ha sido una herramienta fundamental para proteger la confidencialidad de la información a lo largo de la historia, especialmente en tiempos de conflicto. Métodos como los cifrados César, Vigenère y Hill establecieron las bases de lo que hoy conocemos como criptografía, permitiendo que solo las personas autorizadas accedan a información sensible. Sin embargo, con el paso del tiempo, estos cifrados demostraron ser vulnerables ante técnicas de criptoanálisis, como el análisis de frecuencias y la prueba de Kasiski, debido a su relativa simplicidad. A pesar de su efectividad en su época, estos métodos son hoy susceptibles de ser descifrados por atacantes con las herramientas adecuadas.

La evolución hacia sistemas más complejos, como el cifrado RSA, que utiliza claves públicas y privadas, ha mejorado la seguridad de los datos, marcando una transición hacia la criptografía moderna. Este tipo de cifrados, basados en la aritmética modular y la dificultad de factorizar grandes números primos, ha mostrado ser más resistente a los ataques, aunque sigue siendo vulnerable a ciertos algoritmos avanzados. En este contexto, las matemáticas juegan un papel crucial, ya que conceptos como el álgebra lineal y la aritmética modular se utilizan tanto en el diseño como en el análisis de los cifrados.

Finalmente, aunque los algoritmos modernos como RSA son mucho más robustos que los sistemas clásicos, el criptoanálisis ha avanzado paralelamente. Esto demuestra que, aunque ningún sistema criptográfico es completamente infalible, el desarrollo de técnicas más sofisticadas ha permitido proteger de manera más efectiva la información en un mundo digital cada vez más amenazado por ataques sofisticados.

Bibliografía

Friedman, W. F. (1987). *THE INDEX OF COINCIDENCE AND ITS APPLICATIONS IN CRYPTANALYSIS.* Laguna Hills, California 92654: AEGEAN PARK PRESS.

LOPEZ, M. J. (2011). *Criptografia y Seguridad en Computadores.* Andalucía: UNIVERSIDAD DE JAEN.

Schneier, B. (1996). Applied cryptography: Protocols, algorithms, and source code in C (2nd ed.).

John Wiley & Sons.