Instituto Politécnico Nacional

Escuela Superior de Ingeniería Mecanica y Electrica

Unidad Culhuacan

Maestría en Ingeniería en Seguridad y Tecnologías de la información

REPORTE TÉCNICO

Implementación del Cifrado de Vigenére en Matlab

Presenta
Jonathan Eduardo García García jgarciag1404@alumno.ipn.mx

Profesor:

Dr. José Portillo Portillo
Introducción a los sistemas de comunicación seguros.

1 de junio de 2022

Índice

1. Objetivo.

1. Implementar el cifrado de Vigenère en matlab
2. Crear una interfaz de usuario para hacer uso del cifrado
3. Generar los histogramas del texto cifrado, texto plano y del idioma

Equipo necesario	Material necesario
Computadora con el Software Matlab.	Apuntes y conocimientos teóricos sobre el cifrado de Vigenère

2. Marco teórico

2.1. Cifrado de Vigenére

El cifrado de Vigenère está basado en el cifrado del Cesar, por lo cual es un cifrado de sustitución. A diferencia del cifrado del cesar, en el cual cada símbolo del texto plano le es sumada una constante k, en el cifrado de Vigenère se tiene un cifrado del Cesar por cada símbolo de una palabra clave. Con lo cual si la palabra clave tiene una longitud m, se tienen m corrimientos diferentes sobre el texto encriptado. De esta forma, no siempre un mismo símbolo en el texto claro se convierte en el mismo símbolo en el texto encriptado.
A cada letra del texto plano se le sumaría una letra de la clave y como la clave suele ser de menor longitud que el texto plano se repetiría para lograr el tamaño del texto plano.
Formalmente el cifrado de Vigenère se puede expresar de la siguiente manera.
$C_i = S_i + K_{imod(m)} mod(n)$

2.2. Entropia

La entropía es un concepto valioso cuando se piensa en hacer criptoanálisis dado que representa la medida promedio de información que tiene un símbolo en algún mensaje, de hecho se puede pensar en calcular la entropía para cierto lenguaje (español, inglés, etc.) y es curioso saber que la entropía de cada lenguaje tiende a cierto valor característico. La cantidad de información de un símbolo B se define como:
$I(B) = \log_2 \frac{1}{P(B)}$

3. Desarrollo

Se tienen los siguientes alfabetos con su respectiva frecuencia:
Español:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	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	27
%12.53	%1.42	%4.68	%5.86	%13.68	%0.69	%1.01	%0.70	%6.25	%0.44	%0.02	%4.97	%3.15	%6.71	%0.31	%8.68	%2.51	%0.88	%6.87	%7.98	%4.63	%3.93	%0.90	%0.01	%0.22	%0.90	%0.52

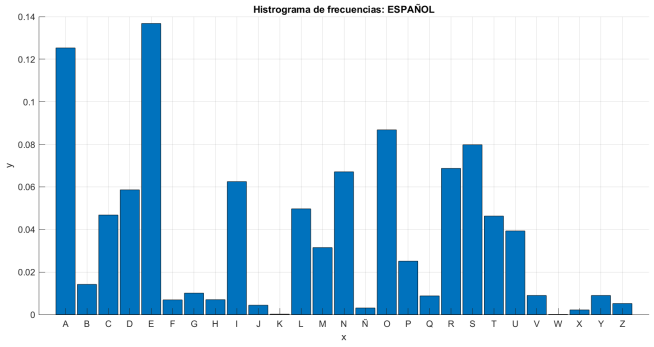

Figura 1: Histograma de frecuencias del idioma Español

Inglés:

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
%8.34	%1.54	%2.73	%4.14	%12.6	%2.03	%1.92	%6.11	%6.71	%0.23	%0.87	%4.24	%2.53	%6.80	%7.70	%1.66	%0.09	%5.68	%6.11	%9.37	%2.85	%1.06	%2.34	%0.20	%2.04	%0.06

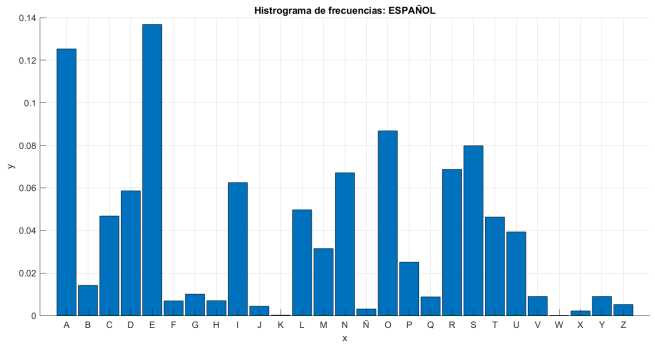

Figura 2: Histograma de frecuencias del idioma Inglés

3.1. Proceso para obtener la llave de cifrado como vector numerico

```
function key=getKeyVector(obj,text)
    if(iscell(text)) %si el texto es una celda
        text=text{1};
    end
    text=convertStringsToChars(text); %convierte la cadena de texto a un arreglo de ...
        caracteres iterable
    key=zeros(1,size(text,2),'single'); %tamano de llave de cifrado
    for i=1:size(text,2) %por cada caracter
        letter=text(i); %obtener el caracter en la posicion i
        letter=obj.LanguageDefinition.normalizeChar(letter);% normalizar el texto (puede que ...
            no este normalizado)
        if(letter<0) %si el caracter en la llave es invalido
            error('La llave de cifrado no es valida caracter no valido:['+text(i)+']');
        end
        key(i)=obj.LanguageDefinition.indexOf(letter)-1; %agregar la representacion numerica ...
            del caracter al vector
    end
end
```

Se toman los caracteres de la llave y se busca su representación numérica en el alfabeto:

Llave	L	L	A	V	E
Valor	12	12	1	22	5

Estos valores servirán como desplazamiento para nuestro texto original repitiendo el patrón hasta terminar la cadena que queremos cifrar.
A cada letra del texto plano se le sumaría una letra de la clave y como la clave suele ser de menor longitud que el texto plano se repetiría para lograr el tamaño del texto plano.

3.2. Implementación del cifrado en matlab

```
function normalizedCipher(obj)
   obj.Key=obj.getKeyVector(obj.Key);
   if(¬obj.Encrypting)
       %la llave de cifrado debe ser negativa para el descifrado
       for i=1:size(obj.Key,2)
           obj.Key(i)=obj.Key(i)*-1;
       end
   end

   key_index=1;
   key_size=size(obj.Key,2);
   ASCII_LENGHT=size(obj.LanguageDefinition.Alphabet,2); % Tamano del alfabeto reducido
   while(obj.Next()) %por cada caracter
       if(obj.Letter<0)
           continue;
       end
       if(key_index>key_size)
           key_index=1;
       end
       key_value=obj.Key(key_index);
       key_index=key_index+1;
       obj.Letter=obj.Letter+key_value; %agregar el desplazamiento de la llave de cifrado
       obj.Letter=mod(obj.Letter,ASCII_LENGHT); %modulo del valor de la letra y el ...
           tamano del diccionario
       obj.Letter=obj.LanguageDefinition.Alphabet(obj.Letter+1).Letter; %obtener la ...
           letra en la nueva posicion del alfabeto cifrado base 1
       obj.ResultText=append(obj.ResultText,obj.Letter); %agregar la letra al ...
           resultado del texto cifrado
   end
end
```

Para cifrar y descifrar el texto plano se siguen los siguientes pasos:

1. Se obtiene la llave de cifrado en su representación numérica

Llave	L	L	A	V	E
Valor	12	12	1	22	5

2. Por cada letra en el texto plano se obtiene su representación numérica

Texto	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
Valor	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

3. Se repite la llave de cifrado las veces que sea necesario para llenar la cadena de texto plano

Texto	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
Valor	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
Llave	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L

4. Se suma el valor de la llave con cada carácter del texto plano

Texto	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
Valor	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
Llave	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L
Valor	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12
Suma	16	14	4	26	10	18	28	9	31	15	23	24	14	36	20	28	29	19	41	25	36	34	24	46	30	38

5. Se aplica la función módulo del tamaño del diccionario a cada carácter

Texto	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
Valor	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
Llave	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L
Valor	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12
Suma	16	14	4	26	10	18	28	9	31	15	23	24	14	36	20	28	29	19	41	25	36	34	24	46	30	38
Módulo	16	14	4	26	10	18	2	9	5	15	23	24	14	10	20	2	3	19	45	25	10	8	24	20	4	12

Finalmente se convierte el resultado del módulo en su representación de carácter y se concatena para obtener el texto cifrado

Texto	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
Valor	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
Llave	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L	L	A	V	E	L
Valor	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12	12	1	22	5	12
Suma	16	14	4	26	10	18	28	9	31	15	23	24	14	36	20	28	29	19	41	25	36	34	24	46	30	38
Módulo	16	14	4	26	10	18	2	9	5	15	23	24	14	10	20	2	3	19	45	25	10	8	24	20	4	12
Cifrado	P	N	D	Z	J	R	B	I	E	O	W	X	N	J	T	B	C	S	O	Y	J	H	X	T	D	L

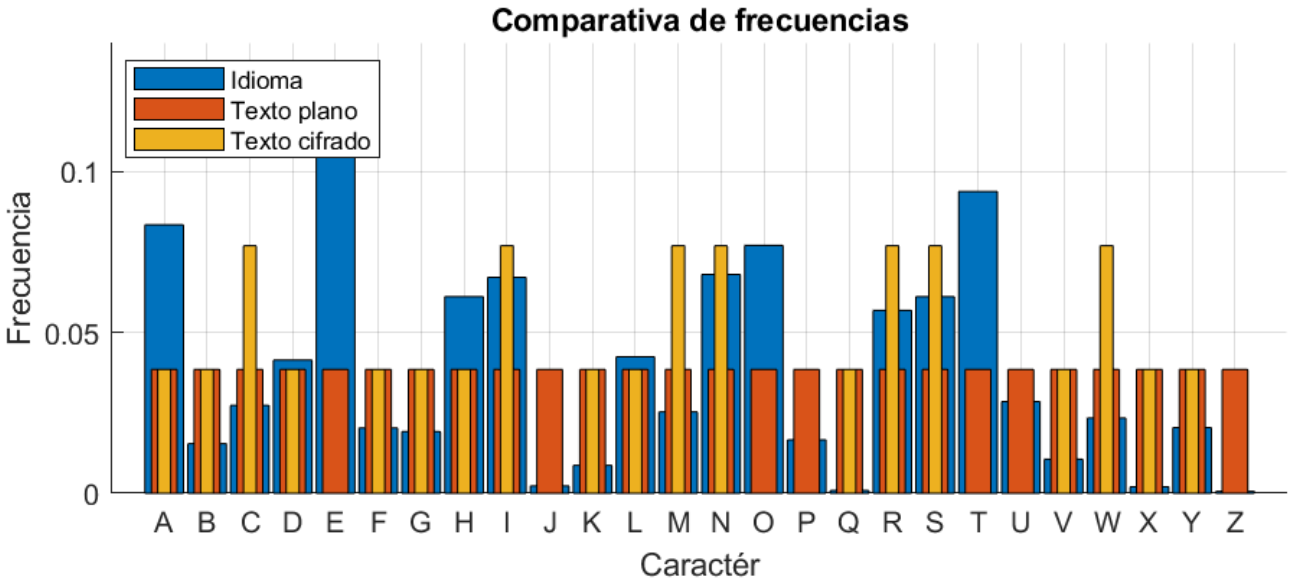

Figura 3: Histograma comparativo de las frecuencias del idioma, el texto cifrado y el texto plano

4. Resultados

Interfaz de usuario mediante la cual el usuario puede introducir texto desde la interfaz o desde un archivo, ingresar una llave numérica o de texto (según sea el caso) y visualizar el resultado del texto cifrado/descifrado. Igualmente se permite seleccionar entre los distintos idiomas que se tienen.

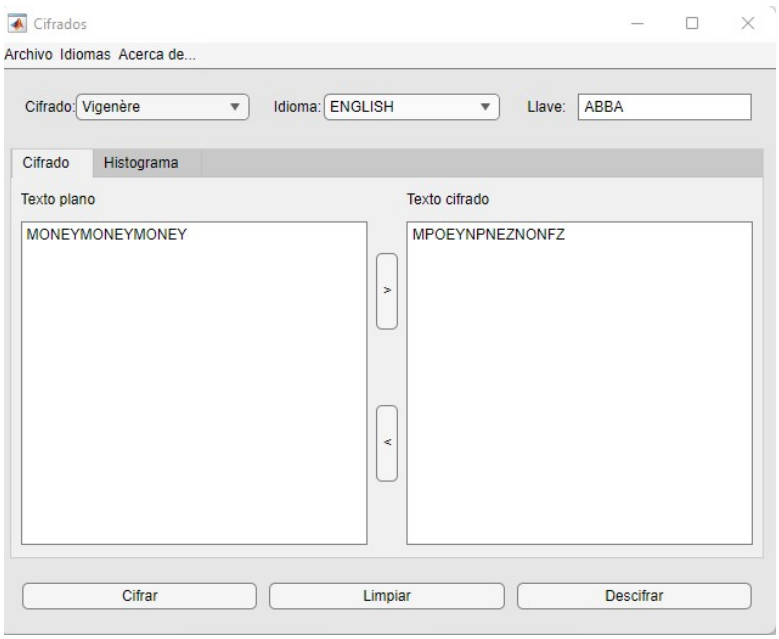

Figura 4: Interfaz de usuario para el cifrado

Interfaz donde el usuario puede comparar los histogramas de las frecuencias generadas para el idioma, texto original y texto cifrado

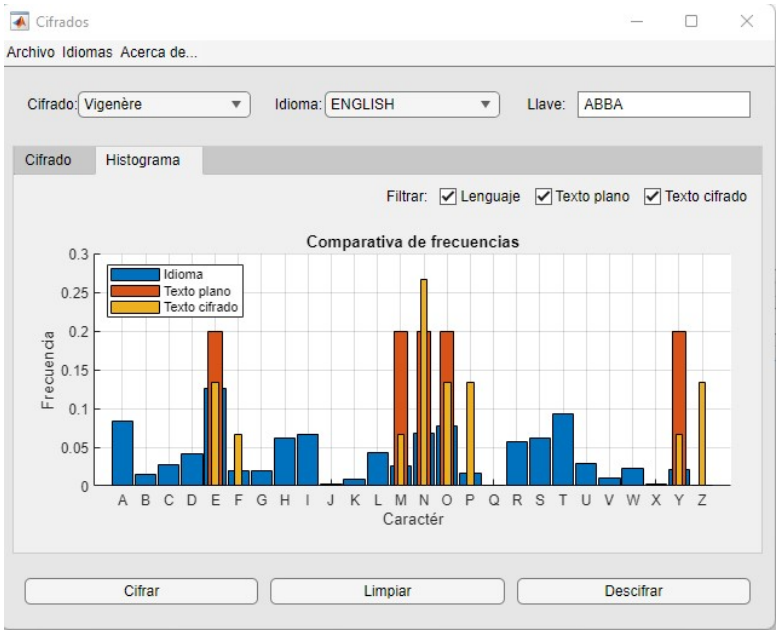

Figura 5: Histograma comparativo de las frecuencias del idioma, el texto cifrado y el texto plano

Interfaz donde se permite al usuario agregar/eliminar y modificar el conjunto de caracteres para cada idioma y la frecuencia de sus caracteres así como agregar/eliminar nuevos idiomas.

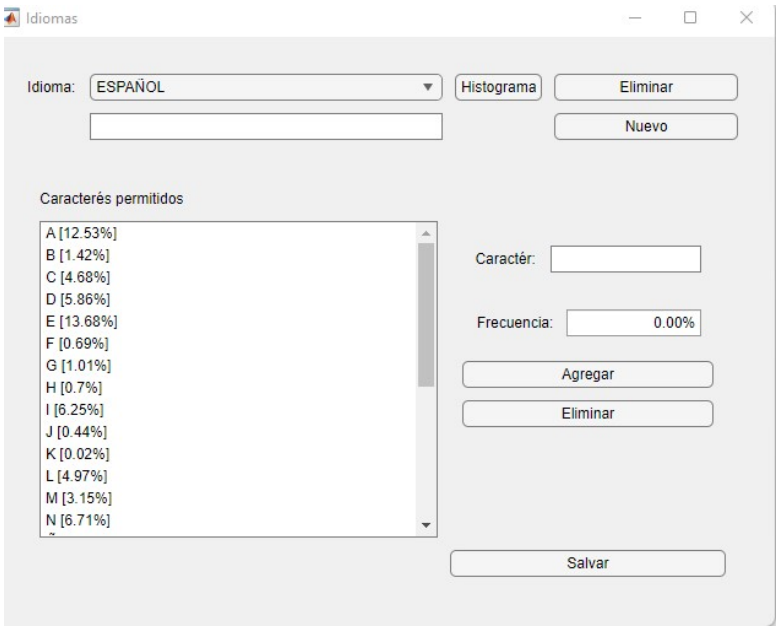

Figura 6: Interfaz para administrar los idiomas

Vista individual del histograma de cada idioma

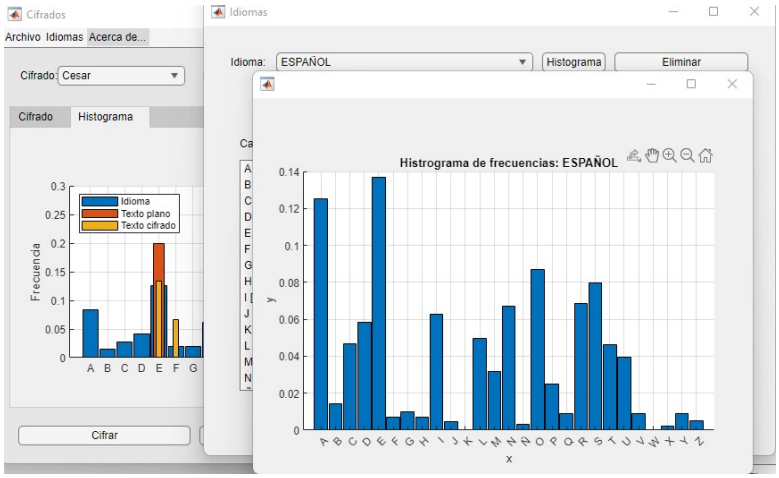

Figura 7: Histograma del idioma

4.1. Comparativa de histogramas con distintas claves

ABCDEFGHIJKLMNÑOPRSTUVWXYZ

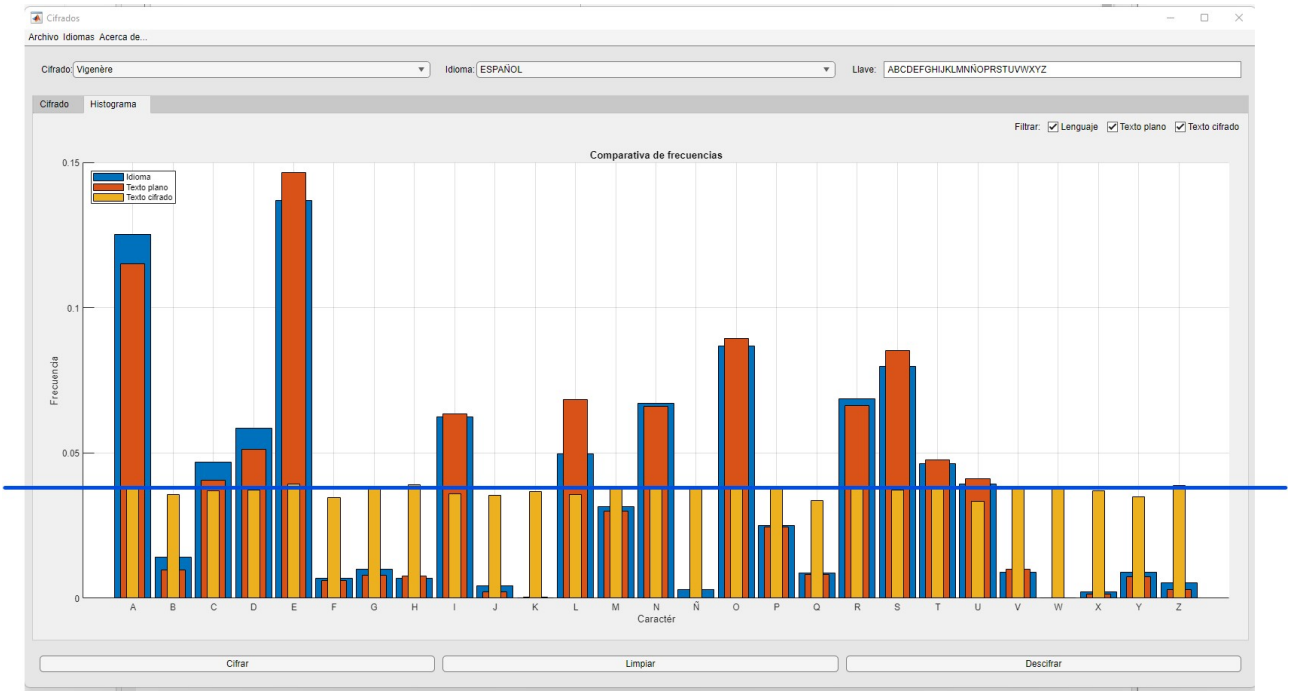

Figura 8: Cifrado con la llave ABCDEFGHIJKLMNÑOPRSTUVWXYZ

THEQUICKBROWNFOXJUMPSOVERTHELAZYDOG

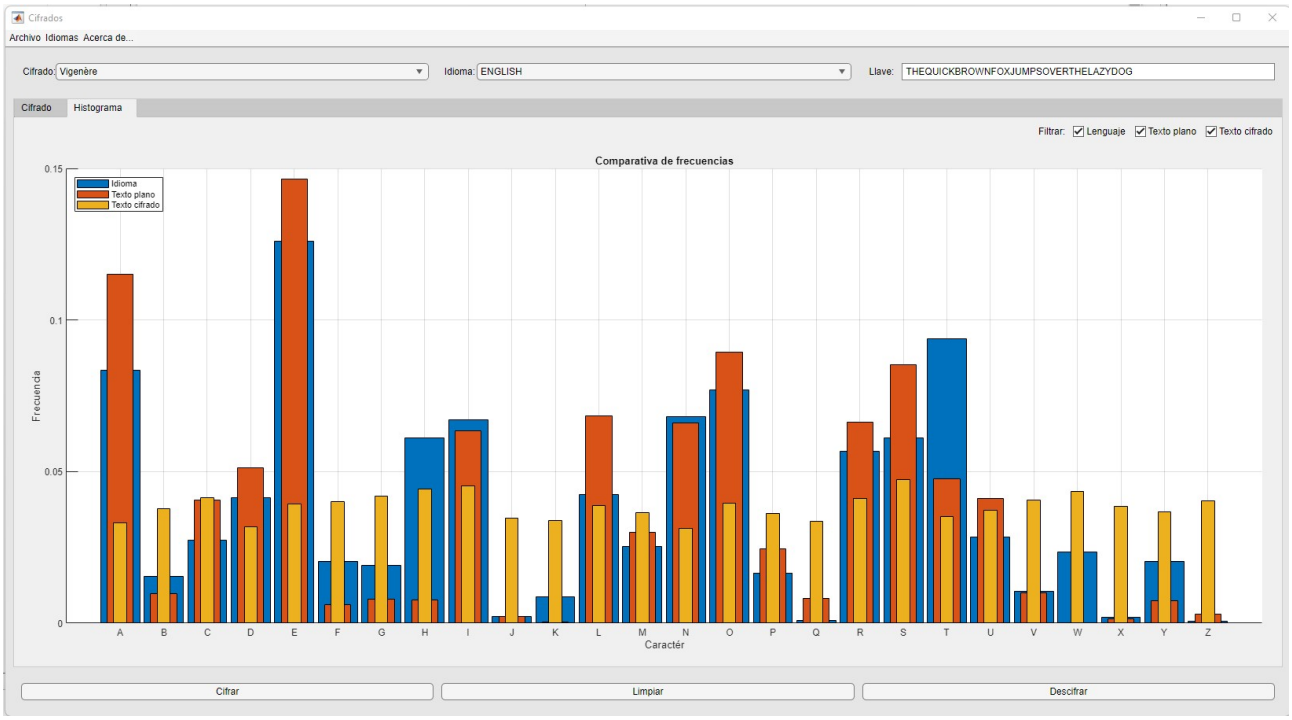

Figura 9: Cifrado con la llave THEQUICKBROWNFOXJUMPSOVERTHELAZYDOG

5. Conclusión

El método de de Vigenère resulta un tanto mas complejo que el cifrado cesar pero en contraste cuando se utiliza con una llave rica en caracteres resulta bastante más efectivo al momento de generar entropía en el texto cifrado y mitigar ataques de análisis de frecuencia. Asimismo, se deberían plantear cuestiones que resulten, al menos, interesantes para realizar como trabajos futuros.

Referencias

[1] S. Gómez, J. D. Arias, and D. Agudelo, "Cripto-análisis sobre métodos clásicos de cifrado," *Scientia et technica*, vol. 2, no. 50, pp. 97–102, 2012.

[2] A. F. Vásquez Henao, "Monografía fundamentos teóricos, matemáticos y estado del arte de las teorías base de la encriptación de datos," 2011.

[3] A. S. González, "La escritura cifrada," *Tabularium*, no. 1, pp. 92–100, 2014.

[4] S. T. Smith, *MATLAB: advanced GUI development*. Dog ear publishing, 2006.

[5] H. Moore and S. Sanadhya, *MATLAB for Engineers*. Pearson Education International New York, 2009.