|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
| ESCUELA PROFESIONAL DE CIENCIA DE LA COMPUTACIÓN | | | |
| ALGORITMOS DE CRIPTOGRAFÍA TRADICIONAL | | | |
| ÁLGEBRA ABSTRACTA (CCOMP 3-1.2) | | | |
|  | | | |
|  | | SEBASTIAN A. PAZ ROCHA | |
|  | | 3er SEMESTRE | |
|  | | 2020 | |
|  | | | |
|  | | | |
| “El alumno declara haber realizado el presente trabajo de acuerdo a las normas de la Universidad Católica San Pablo” | | | |
|  | | | |
|  |  | |  |
|  |  | |

**Criptografía Tradicional**

1. **Cifrados de transposición**
   1. **Cifrado Rail Fence**

El texto claro es escrito hacia abajo diagonalmente en *raíles* sucesivos de una cerca imaginaria, cuando se llega a la base, el texto sube; luego, vuelve a bajar, hasta que el texto sea completado (Hassan & Hizaji, 2017).

La clave del cifrado es el número de raíles, en el caso de la Tabla 1, la clave es cuatro, y el texto claro es “ME GUSTA PROGRAMAR”

Tabla 1

*Cifrado Rail Fence con clave 4*

1

2

3

4

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| M |  |  |  |  |  | A |  |  |  |  |  | A |  |  |  |
|  | E |  |  |  | T |  | P |  |  |  | R |  | M |  |  |
|  |  | G |  | S |  |  |  | R |  | G |  |  |  | A |  |
|  |  |  | U |  |  |  |  |  | O |  |  |  |  |  | R |

*Fuente: Elaboración propia*

Con este cifrado de transposición, el mensaje cifrado es: “MAAETPRMGSRGAUOR”

* 1. **La Escítala**

Fue una herramienta usada por los griegos y sigue un principio parecido al del cifrado Rail Fence, pero la clave es un trozo de madera, la escítala; debían producirse dos idénticas. Se envolvía una tira de pergamino de una letra de longitud alrededor de la escítala, y el mensaje aparecería escrito a lo largo de ella. (Shimeall & Spring, 2014).

La clave del cifrado es el grosor de la Escítala, en el caso de la Tabla 2, la clave es tres, porque el cilindro de madera tiene tal grosor como para mostrar tres letras con cada vuelta de la tira del pergamino. El texto claro es “ME GUSTA PROGRAMAR”

Tabla 2

*Cifrado con la Escítala*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| M |  |  | E |  |  | G |  |  | U |  |  | S |  |  | T |
|  | A |  |  | P |  |  | R |  |  | O |  |  | G |  |  |
|  |  | R |  |  | A |  |  | M |  |  | A |  |  | R |  |

1

2

3

4

*Fuente:Elaboración propia*

Con este cifrado de transposición, el mensaje cifrado es: “MAREPAGRMUOASGRT”

* 1. **Transposición Myszkowski**

El mensaje es escrito horizontalmente a través de las columnas, las cuales son enumeradas basándose en una *palabra clave*. La clave tiene que tener al menos dos letras que se repitan para que el cifrado funcione como es debido.

Las letras en la clave son enumeradas de acuerdo a su orden relativo en el alfabeto; las letras duplicadas tienen el mismo número (Shimeall & Spring, 2014).

Para ilustrar mejor cómo funciona, ver el ejemplo de la Tabla 3, donde el texto claro es “ME GUSTA PROGRAMAR ALGORITMOS DE CIFRADO”, y la clave es “LAPTOP”.

Tabla 3

Cifrado por transposición Myszkowski

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| L | A | P | T | O | P |
| 2 | 1 | 4 | 5 | 3 | 4 |
| M | E | G | U | S | T |
| A | P | R | O | G | R |
| A | M | A | R | A | L |
| G | O | R | I | T | M |
| O | S | D | E | C | I |
| F | R | A | D | O |  |

Texto cifrado:

EPMOSRMAAGOFSGATCOETPRMLOMSIRUORIED

1. **Cifrados de sustitución**

Los cifrados de sustitución se diferencia de los de transposición por lo siguiente: En un cifrado de transposición, el texto claro sólo es reposicionado pero las letras permanecen sin cambios. En contraste, un cifrado de sustitución mantiene la misma secuencia del texto claro, pero sí modifica las entras en sí. (Raggo & Hosmer, 2013)

Como se demostró, los cifrados de transposición están limitados por su principio de reposicionamiento, y pueden ser *crackeados* sin necesidad de una computadora.

Los cifrados de sustitución tienen miles de implementaciones diferentes, lo que los hace mucho más complejos y seguros.

* 1. **Cifrado Polybios**

Hasta donde se sabe, este es el cifrado de sustitución más antiguo, inventado alrededor del año 150 a.C por el historiador griego *Polybios*. (Flores & Reséndiz, 2011)

Este cifrado consiste en la creación de una tabla (matriz) de 5 x 5, en las que se colocan las letras A, B, C, D, E en ambas entradas. Luego, dentro de ella se introducen 25 letras del alfabeto (***i*** y ***j*** van en un mismo espacio); para que se sustituya cada letra por un dígrafo formado por las letras de las entradas, como se observa en la siguiente tabla.

Tabla 4

*Tabla de cifrado Polybios*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E |
| A | A | B | C | D | E |
| B | F | G | H | I/J | K |
| C | L | M | N | O | P |
| D | Q | R | S | T | U |
| E | V | W | X | Y | Z |

Por ejemplo, si se quiere cifrar el mensaje “ES MUCHO TRABAJO PARA UNO SOLO”, se usaría como referencia la Tabla 4, y el texto cifrado sería formado por los dígrafos correspondientes de cada letra, por lo que quedaría de la siguiente manera:

AEDC CBDEACBCCD DDDBAAABAABDCD CEAADBAA DECCCD DCCDCACD

Dado que el cifrado Polybios no posee clave, es muy fácil de descifrar si se sabe cómo armar esa tabla para luego mapear las letras; por lo que es un algoritmo de cifrado obsoleto e inútil, aunque esto no quiere decir que no sea importante para la historia de la criptografía.

* 1. **Cifrado de César**

El cifrado de César es tal vez el algoritmo más representativo de los que están considerados dentro de la criptografía tradicional.

Jason Andress (2014) menciona que se dice que fue inventado por el emperador Julio César; y explica que está basado en la transposición y consiste en cambiar cada letra del *texto claro* por otra letra en otra posición en el alfabeto, en el cifrado original se cambiaban tres posiciones, tal como se muestra en la Figura 1.

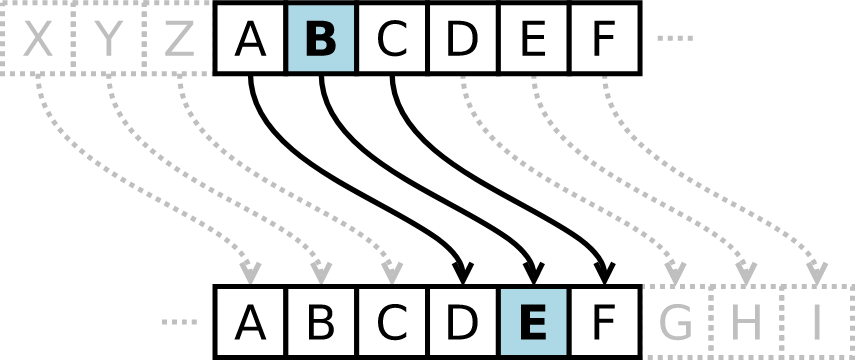


Figura 1: Rotación de clave 3 en el cifrado César

Para descifrar un mensaje en el que se empleó el algoritmo de César, nada más basta restar un número determinado de posiciones a cada letra; este número de posiciones es la *clave*. La clave es necesaria para poder descifrar el mensaje al momento, y debe ser la misma con la que se realizó el cifrado para que se pueda descubrir el mensaje original.

***Criptoanálisis del cifrado César***

Si no se conoce la clave, el cifrado César puede ser quebrantado con métodos estadísticos (análisis de frecuencia).

El uso básico del análisis de frecuencia es, primero, contar la frecuencia de las letras del texto cifrado, y asociarlas con letras posibles del texto claro. Esto basándose en la frecuencia de las letras por idioma, en el caso del español, la Figura 2 indica la frecuencia de algunas letras y palabras.

El primero en estudiar frecuencia de letras fue el árabe matemático Al-Kindi, quien desarrolló este método formalmente (Hassan & Hizaji, 2017).

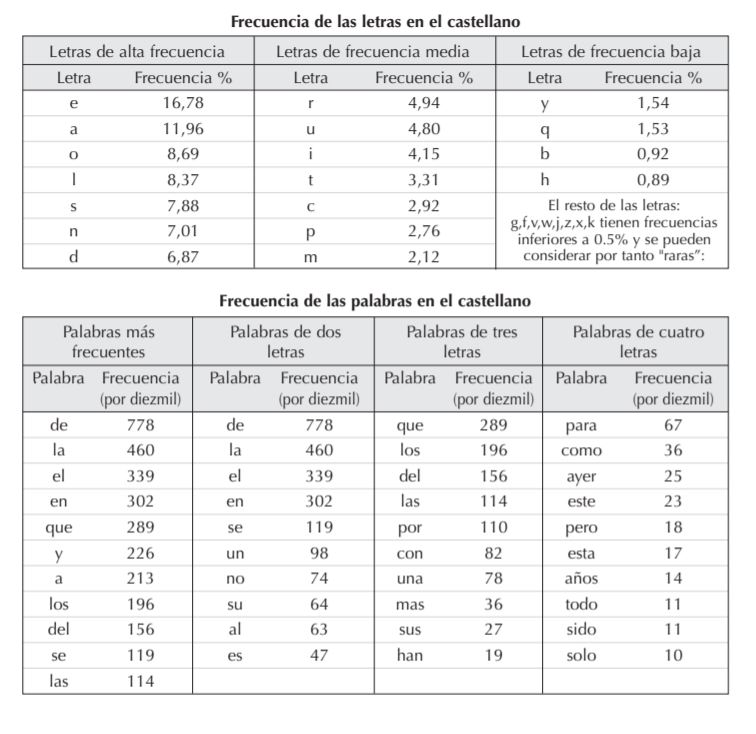


Figura 2: Tabla de análisis de frecuencia en el idioma español

* 1. **Cifrado Afín**

Dado que el cifrado César puede producir a lo más 26 transformaciones distintas del texto (en el alfabeto español), resulta no ser método de encriptación; por eso, el cifrado Afín, ofrece un poco más de seguridad al realizarse ciertas operaciones aritméticas en una ecuación antes de hacer el cambio de las letras. (Lamagna, s.f.)

Dicha ecuación es la siguiente:

Donde ***x*** es el valor numérico de la letra en el alfabeto, ***m*** es el número de letras del alfabeto del texto claro; ***a*** y ***b*** son constantes secretas (claves), y ***y*** es el resultado de la transformación.

Cuando se conoce las claves de un mensaje cifrado por Afín, la ecuación de descifrado es la siguiente:

Para mostrar su funcionamiento, sigue un ejemplo en el cual se cifrará la palabra “PAZROCHA”, haciendo uso del alfabeto inglés (26 letras), con las claves a = 14, b = 11.

**P ->** (14\*15 + 11) mod 26 -> **N**

**A ->** (14\*0 + 11) mod 26 -> **L**

**Z ->** (14\*25 + 11) mod 26 -> **X**

**R ->** (14\*17 + 11) mod 26 -> **P**

**O ->** (14\*14 + 11) mod 26 -> **Z**

**C ->** (14\*2 + 11) mod 26 -> **N**

**H ->** (14\*7 + 11) mod 26 -> **F**

**A ->** (14\*0 + 11) mod 26 -> **L**

Palabra cifrada: NLXPZNFL

***Criptoanálisis del cifrado Afín***

Aunque no se tengan las claves *a* y *b,* el descifrado es bastante sencillo; dado que ya conocemos la ecuación, podemos llegar a los valores de las claves resolviendo dos ecuaciones lineales. Por ejemplo, si asumimos que “SE” es cifrado como “BH”:

**S -> B:** (18\*a + b) = 1 mod 26

**E -> H:** (4\*a + b) = 7 mod 26

Resolviendo estas ecuaciones se obtiene a = 7, b = 5.

* 1. **Cifrado Playfair**

Playfair fue el primer cifrado de sustitución por dígrafos, fue inventado en 1854 por Charles Wheatstone; y fue muy usado durante la Primera Guerra Mundial.

Este algoritmo cifra las letras por pares (dígrafos), en vez de hacerlo individualmente como en un algoritmo de sustitución monoalfabético; y es más difícil de usar el análisis de frecuencia ya que las posibles sustituciones se elevan al cuadrado (26^2 en el caso del alfabeto español). (PracticalCryptography, s.f.)

***Parámetros del algoritmo Playfair a tener en cuenta***

* La clave del algoritmo suele ser una palabra, y se coloca al principio de una matriz de 5 x 5, sin tomar en cuenta las letras que se repiten; cuando la palabra se acaba, se procede a completar la matriz llenando con las letras del alfabeto que faltan (***i*** y ***j*** toman la misma posición). Por ejemplo, si la clave fuese “ABSTRACTA”, la matriz quedaría de la siguiente forma:

Tabla 5

*Matriz del algoritmo Playfair con clave “ABSTRACTA”*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A | B | S | T | R |
| C | D | E | F | G |
| H | I/J | K | L | M |
| N | O | P | Q | U |
| V | W | X | Y | Z |

Fuente: Elaboración propia

* Para realizar este cifrado se deben obviar los signos de puntuación, números y espacios.
* En caso haya dos letras iguales seguidas en el texto claro, como en la palabra “LLAMA”, se reemplaza la segunda ocurrencia por una ‘X’. 🡪 “LXAMA”
* Si el texto claro tiene un número de letras impar, se agrega una ‘X’ al final.
* El algoritmo funciona cuando se separan los caracteres de dos en dos.
* Localización de los dígrafos:

1. Si las letras están en diferentes filas o columnas, se reemplaza el par por las otras letras que están en la misma fila, pero en el otro par de esquinas del rectángulo definido por el par original. Por ejemplo, si se quiere cifrar el dígrafo “FO” (rojo), se tomará el dígrafo “DQ” (verde).
2. Si las letras están en la misma fila, se reemplazará cada una por la que está inmediatamente a su derecha.
3. Si las letras están en la misma columna, se reemplazará cada una por aquella que esté inmediatamente debajo de ella.

Teniendo en cuenta estos parámetros, se procederá a cifrar la frase: “A ESTE PUNTO YA ESTOY CANSADO” con la clave “ABSTRACTA”.

Primero, se debe ignorar todos los espacios, y dividir toda la frase en dígrafos:

“AE ST EP UN TO YA ES TO YC AN SA DO”

Tabla 6

*Tabla de cifrado Playfair*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Texto claro | AE | ST | EP | UN | TO | YA | ES | TO | YC | AN | SA | DO |
| Texto cifrado | SC | TR | KX | NO | BQ | VT | KE | BQ | VF | CV | TB | IW |

Texto cifrado: SCTRKXNOBQVTKEBQVFCVTBIW

***Criptoanálisis del cifrado Playfair***

Hacer el criptoanálisis de Playfair es mucho más difícil que de otros cifrados, ya que sólo existen 27 monografos en el alfabeto español, pero existen más de 600 dígrafos, y es con lo que Playfair funciona.

Aún así, se puede realizar un análisis de frecuencia. Un indicio podría ser los dígrafos inversos, como “SP” y “PS”; seguirían el mismo patrón al descifrarse, por ejemplo “LA” y “AL”. Con eso ya se puede empezar a analizar la clave del cifrado, ya que hay algunas palabras que cuentan con estos dígrafos, como “LAterAL”.

* 1. **Cifrado Vernam**

El cifrado Vernam es un cifrado polialfabético, el cual sigue el mismo principio que el de Vigenère, con la diferencia que su clave es igual de larga que el mensaje, por lo que la clave no se tiene que repetir a lo largo del texto claro.

Por ejemplo, si se quiere cifrar la palabra “ARROZ”, se tiene que usar una clave de la misma longitud, como “HUEVO”. Y se proseguiría según la Tabla 7.

Tabla 7

*Cifrado Vernam*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Texto  Claro | A | R | R | O | Z |
| Posición | 0 | 18 | 18 | 15 | 26 |
| Clave | H | U | E | V | O |
| Posición | 7 | 21 | 4 | 22 | 15 |
| Suma | 7 | 39 | 22 | 37 | 41 |
| Módulo | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 |
| Posición | 7 | 12 | 22 | 10 | 14 |
| Texto  cifrado | H | M | V | K | Ñ |

Pero este cifrado derivó en otro tipo de cifrado Vernam aún más complejo y seguro, el cual es también llamado “Libreta de un solo uso”:

Hay algo muy especial en este cifrado Vernam, y es que es considerado el único con seguridad perfecta (es imposible de descifrar, no hay criptoanálisis); principalmente porque la clave es sólo usada una vez, por eso, debe ser de igual o mayor longitud que el texto claro. Además, la clave tiene que ser completamente aleatoria. (Computer Science Tutor, 2018)

Primero, se debe convertir el texto claro a binario; busca cada letra en su posición de ASCII y hace la conversión de número decimal a binario (ej. a = 97; 97 = 01100001).

Luego, se genera una clave aleatoria que, en binario, tiene que ser de igual o mayor longitud que el texto claro. Finalmente, se produce el texto cifrado al pasar el texto claro y la clave (ambos en binario) a través de una puerta XOR; finalmente se convierte de binario a decimal, y se busca la equivalencia en la tabla de ASCII, para tener el producto cifrado final.

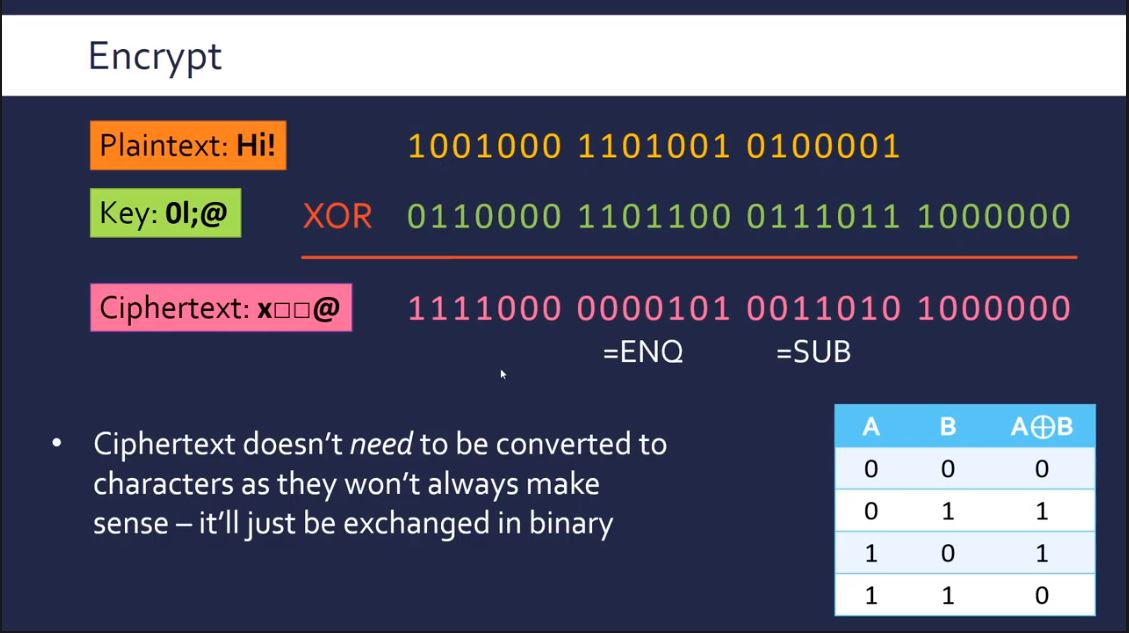


Figura 3: Proceso de cifrado a través de “Libreta de un solo uso” (Vernam)

Como se observa en la figura 3; la clave, en lenguaje binario, es más larga que el texto en claro en sí, lo cual no permite que se repitan patrones para hacer un criptoanálisis. El resultado, luego de pasar ambas cadenas por la puerta XOR, se convierte a decimal y se busca su equivalente símbolo ASCII; en el ejemplo, el mensaje cifrado no cuenta con caracteres, y tampoco necesita hacerlo, lo que vuelve aún más complejo a este tipo de cifrado.

* 1. **Cifrado de Vigenère**

Este cifrado fue inventado por el diplomático francés Blaise de Vigenère en el siglo XVI. Hassan y Hizaji (2017) señalan que este cifrado usa una serie de cifrados César basándose en una *palabra o frase clave.*

Como se explicó antes, el cifrado César cambia las letras usando un solo valor; por ejemplo, en clave 3, una B se convierte en una E; una O, en una R; y así.

Entonces, el cifrado de Vigenère usa varios cifrados César, cambiando el número de posiciones cada vez que pasa a una letra diferente del texto claro (clave de César dinámica).

Para cifrar un mensaje con el algoritmo de Vigenère, se necesita una tabla de Vigenère. Esta tabla consiste en el alfabeto completo escrito 27 veces (en el caso del español) en diferentes filas; cada fila es rotada una posición hasta llegar a la letra final (Figura 4)

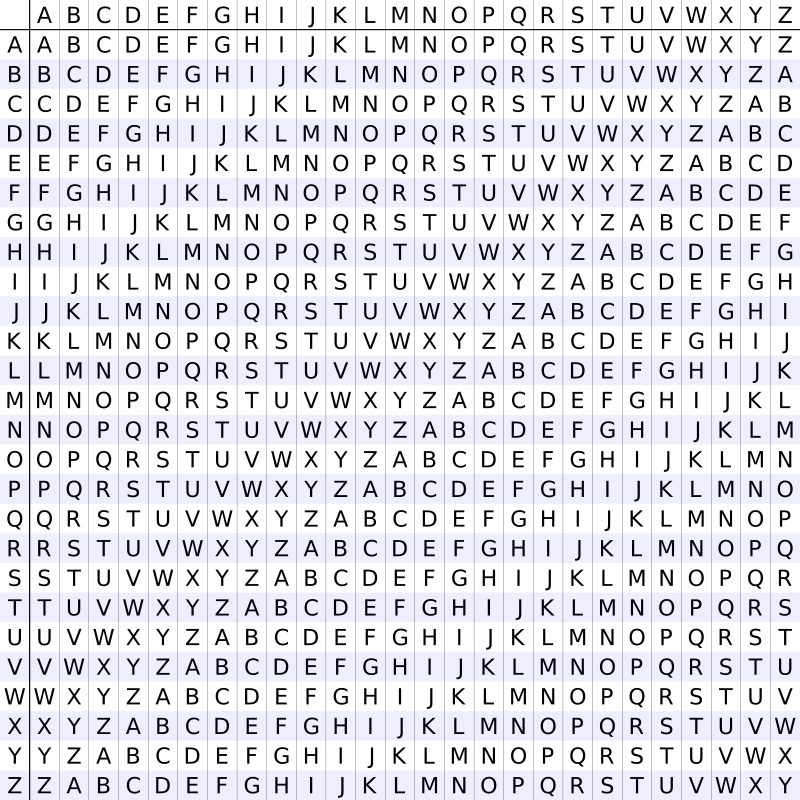


Figura 4: Tabla de Vigenère

Para poder cifrar un mensaje, se necesitará otra tabla para mapear cada letra del texto claro y de la clave. Se sumarán las posiciones de las letras correspondientes y se sacará el módulo 27 para mapear la nueva letra cifrada; en caso las letras de la palabra clave se acaben, se vuelve a repetir hasta que finalice el texto claro; tal como se muestra en la Tabla 8.

Texto claro:

ALGEBRA ABSTRACTA

Clave:

DIFICIL

Tabla 8

*Cifrado de Vigenère*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Texto  Claro | A | L | G | E | B | R | A | A | B | S | T | R | A | C | T | A |
| Posición | 0 | 11 | 6 | 4 | 1 | 18 | 0 | 0 | 1 | 19 | 20 | 18 | 0 | 2 | 20 | 0 |
| Clave | D | I | F | I | C | I | L | D | I | F | I | C | I | L | D | I |
| Posición | 3 | 8 | 5 | 8 | 2 | 8 | 11 | 3 | 8 | 5 | 8 | 2 | 8 | 11 | 3 | 8 |
| Suma | 3 | 19 | 11 | 12 | 3 | 26 | 11 | 3 | 9 | 24 | 28 | 20 | 8 | 13 | 23 | 8 |
| Módulo | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 | mod 27 |
| Posición | 3 | 19 | 11 | 12 | 3 | 26 | 11 | 3 | 9 | 24 | 1 | 20 | 8 | 13 | 23 | 8 |
| Texto  cifrado | D | S | L | M | D | Z | L | D | J | X | B | T | I | N | W | I |

***Criptoanálisis del cifrado de Vigenère***

Mikel García (2015) explica este proceso de criptoanálisis en su blog, en donde menciona las siguientes cosas:

Este cifrado fue considerado invulnerable durante aproximadamente 300 años, hasta que Friedrich Kasiski publicase un método para romper el cifrado de Vigenère en 1863.

La principal fortaleza de este cifrado, es que no se conoce la longitud de la clave; pero si se supiese, se podría fragmentar el texto cifrado teniendo en cuenta dicha longitud, para luego hacer uso del análisis de frecuencia. Esto será posible gracias al método de Kasiski, siempre y cuando la longitud del texto sea lo suficientemente grande para que sea eficaz.

Hay que tener en cuenta que de vez en cuando, en el mensaje cifrado, se repetirán ciertos dígrafos o trígrafos, y la distancia entre ellos **podría** ser un múltiplo de la longitud de la clave. Así que, el primer objetivo es encontrar la ocurrencia de dígrafos o trígrafos en el texto cifrado, como en el ejemplo de la figura 5.

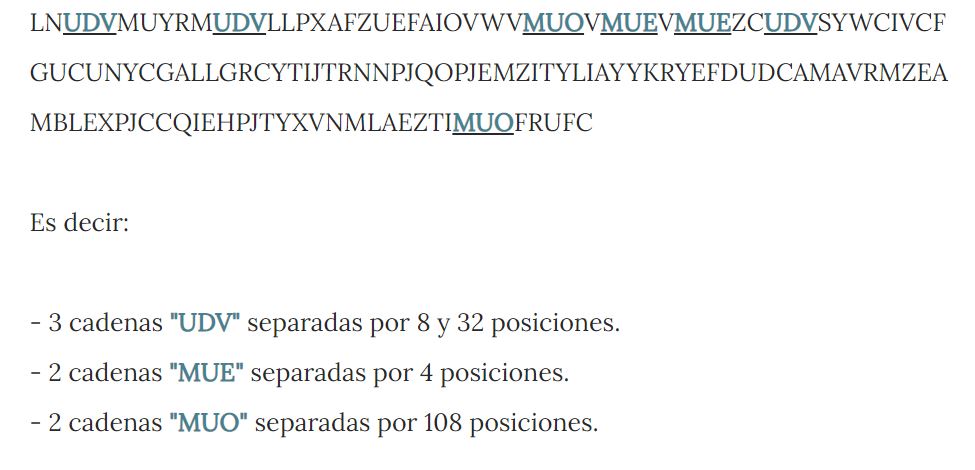


Figura 5: Ejemplo de trígrafos repetidos en un mensaje cifrado

Como se observa, las separaciones de los trígrafos señalados son 4, 8, 32 y 108. Y la longitud de la clave podría ser un MCD de dichos números. A veces pueden variar las separaciones y pueden mostrar números que no tienen divisores en común, con eso no se llagaría a ningún lado en este método, por eso es importante analizar bien los datos que vamos obteniendo. Pero en el conveniente caso de la figura 5, el MCD de dichos números es 4, por lo que podemos suponer que la longitud de la clave es 4.

Después de hallar dicha longitud, se procede a dividir el texto cifrado en bloques de cuatro caracteres cada uno (subcriptogramas).

El primer subcriptograma (CA) contendría los siguientes caracteres del criptograma: 1º, 5º, 9º, etc.

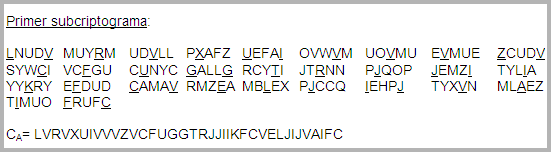


Figura 6: Primer subcriptograma

El segundo subcriptograma (CB) contendría los siguientes caracteres del criptograma: 2º, 6º, 10º, etc.

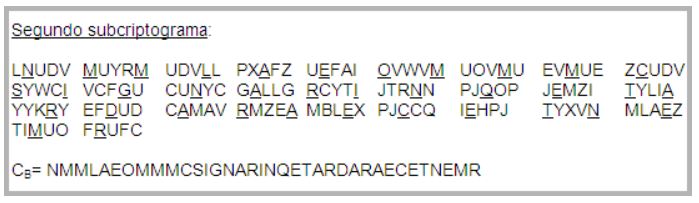


Figura 7: Segundo subcriptograma

El tercer subcriptograma (CC) contendría los siguientes caracteres del criptograma: 3º, 7º, 11º, etc.

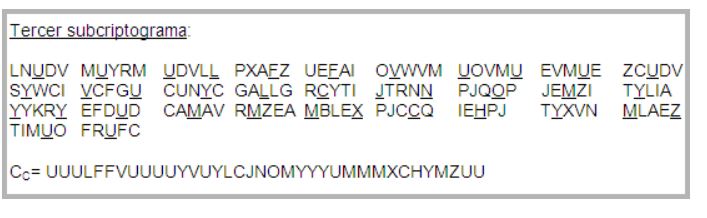


Figura 8: Tercer subcriptograma

Y, finalmente, el cuarto subcriptograma (CD) contendría los siguientes caracteres del criptograma: 4º, 8º, 12º, etc.

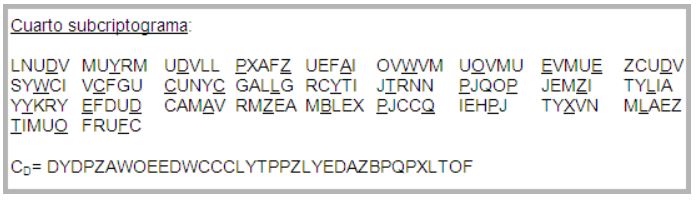


Figura 9: Cuarto subcriptograma

Luego de este exhaustivo análisis, ya se puede realizar un ataque simple de tipo estadístico monoalfabético. Y la frecuencia relativa de los criptogramas sería la siguiente:

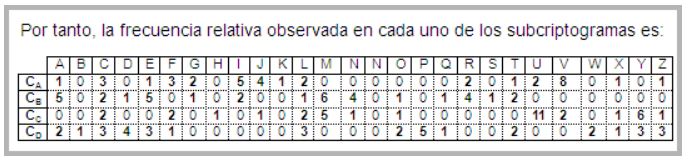


Figura 10: Frecuencia relativa de los criptogramas

A partir de aquí y considerando que la posición relativa de la letra "A" es el valor 0, la letra "E" está 4 espacios a la derecha de la "A" y la letra "O" está 11 de la letra "E", buscaremos en cada subcriptograma (Ci) caracteres frecuentes que cumplan con esa distribución: 0, +4, +11 mod 27:

- Para CA se elige RVG (2, 8, 2), luego la letra clave sería la "R".

- Para CB se elige AEO (5, 5, 1), luego la letra clave sería la "A".

- Para CC se elige UYJ (11, 6, 1), luego la letra clave sería la "U",

- Para CD se elige LOZ (3, 2, 3), luego la letra clave sería la "L".

Con la clave = "RAUL", utilizando la tabla y siguiendo los pasos para descifrar este tipo de mensajes indicados en este post, obtenemos el siguiente texto plano a partir del texto cifrado:

UNASEMANAMASELREGALODELPROBLEMADEMATEMATICASESELLIBROGARDNERPARAPRINCIPIANTESQUESESORTEARAENTRETODASLASPERSONASQUEDESCIFRENESTEMENSAJEFIRMADORAUL

**Referencias bibliográficas**

Andress, J (2014). *The Basics of Information Security (p .71)* (1a ed.). Recuperado de: <https://www.academia.edu/32643426/Andress_Jason_Basics_of_Information_Security_Second_Edition>

Computer Sciece Tutor (2018). *Vernam Cipher (One-Time Pad)*[Video de Youtube]. Recuperado de: <https://www.youtube.com/watch?v=cpqwp2H0SNo>

**García, M (2015). *Criptografía (I): cifrado Vigenère y criptoanálisis Kasiski****.* El blog de García Larragan y Cía. Recuperado de: <http://mikelgarcialarragan.blogspot.com/2015/03/criptografia-i.html>

Hassan, N. A y Hizaji, R (2017). *Data Hiding Techniques in Windows OS*. Syngress publishing (Elsevier).

Lamagna, E. A (s.f.). Classical Cryptography: Affine Cipher. University of Rhode Island, Kingston, Estados Unidos. Recuperado de: <https://www.cs.uri.edu/cryptography/classicalaffine.htm>

PracticalCryptography (s.f.). *Playfair Cipher*. Recuperado de: <http://practicalcryptography.com/ciphers/playfair-cipher/>

Raggo, M y Hosmer, C (2013). *Data Hiding* (Capítulo 1)*.* Syngress publishing (Elsevier).

Shimeall, T. J y Spring, J. M (2014). *Introduction to Information Security* (Capítulo 8). Syngress publishing (Elsevier).