Práctica 2 de Criptografía y Computación

Carlos Núñez Molina
Alessandro Zito
Gabriela Antolinez

1 Cifrado de Vigenére

Primero desciframos el texto cifrado mediante el **Cifrado de Vigenére**. Para ello usamos el método del **Índice de Coincidencia**.

Empezamos calculando el índice de coincidencia de todos los textos cifrados, del 1 al 5. Los índices obtenidos, en orden, son: 0.091, 0.064, 0.063, 0.032, 0.066. Sabemos que el índice de coincidencia medio de los textos en español es de 0.07. Como de todos los cifrados usados solo el de Vigenére cambia el índice de coincidencia del texto original, reduciéndolo al uniformizar las frecuencias de aparición de las distintas letras, sabemos que de todos los textos solo el número 4 ha sido cifrado con Vigenére, al tener un índice de coincidencia de 0.032.

Acto seguido, calculamos la longitud de la clave usada. Para ello, vamos iterando desde $n = 2$ en adelante, dividiendo el texto en sucesivas partes, pegando "saltos" de n letras y calculando el índice de coincidencia de todas estas partes, hasta que todos valgan alrededor de 0.07. Cuando $n = 12$, todas las subcadenas obtenidas tienen un índice de coincidencia alrededor de 0.07, por lo que sabemos que la clave debe ser de longitud 12.

Así, procedo a dividir el texto en las subcadenas correspondientes a dar saltos para $n = 12$ y les aplico un análisis de frecuencias como si cada subcadena hubiera sido cifrada con el Cifrado de Cesar. Las claves obtenidas (correspondientes al número de desplazamientos) son: 18, 4, 13, 0, 2, 8, 12, 8, 4, 13, 20, 15. Al concatenarlas y asociarles su letra correspondiente, obtenemos la clave ***RENACIMIENTO***.

Esta es la clave usada para cifrar el texto 4 con el Cifrado de Vigenére, y con la que podemos descifrar el texto. Para ello, obtenemos la *clave inversa*, es decir, la clave con la que, al cifrar el texto cifrado de nuevo con Vigénere, obtenemos el texto original. Esta clave se obtiene aplicando, para cada letra, los desplazamientos de Cesar en sentido contrario y es: $JWÑAYSOSWÑHM$.

2 Cifrado Escítalo

Por segundo, intentaremos a descifrar los textos mediante el **Cifrado de Escítalo**. Utilizamos la función **Análisis de Frecuencia**. Hacemos el análisis de frecuencia de todos los textos. Podemos ver que en el *texto 3* el análisis de frecuencia es parecida a la de un texto bastante largo escrito en español. Entonces, seguramente el texto será cifrado con Escítalo, porqué este tipo de encriptación no cambia el análisis de frecuencia de un texto. Después de hacer esto, ententamos iterando desde 2 hasta usando la fuerza bruta hasta que hemos descifrado el mensaje, y hacemos esto con la función **cifra_transposicion(texto3,n)**. El numero que sale es $n = 23$.

3 Cifrado de Cesar

Luego, se descifraron los textos que se encontran cifrados con **Cifrado de Cesar** para lo cual se evaluaron las **frecuencias** de los textos. Así, evidenciamos que en el *texto 1* la letra más frecuente es la V con 25.636 por cien, seguida de la letra Z con 7.872 por cien y la letra Ñ con 6.848 por cien; en el *texto 2* las letras más frecuentes son la L con 12.59 por cien y la O con 11.836 por cien. Teniendo en cuenta que las letras más frecuentes en el español son la A y la E, identificamos que éstas se encuentran en la misma distancia entre sí que las letras *V y Z* presentes en el texto 1 y las letras *L y O* del texto 2 dado que en los tres casos existen 4 saltos de separación. Por tanto se evalua sí los textos se encuentran cifrados con Cesar donde el número de saltos corresponde a 21 en el texto 1 y a 7 en el texto 2. Así, se realiza la respectiva sustitución de cada letra teniendo en cuenta los saltos presentes en cada texto. Ejemplo de ello para el texto 1 es la sustitución de la letra V por A, la W por B, la X por C y así de forma sucesiva con las letras restantes, de igual forma para el texto 2 comenzando con la sustitución de la letra L por A, la M por B, la N por C y así de forma progresiva con las letras faltantes. Se comprueba de esta forma que los textos cobran sentido y coherencia conforme se realizan las sustituciones, por lo que se concluye que ellos se encontraban cifrados con Cesar y fueron descifrados de forma adecuada.

4 Cifrado por Permutación

Por último, desciframos el texto cifrado por **permutación**, que es un método de sustitución monoalfabética en el que cada letra se sustituye por cualquier otra. Sabíamos que el texto **5** está cifrado de esta forma porque su índice de coincidencia es de 0.066, muy parecido al índice de coincidencia medio de 0.07 de los textos en español, lo que indica que no se ha cifrado con *Vigénere*. Además, como su tabla de frecuencias difiere de la del español, sabemos que no se ha cifrado con el método del *Escítalo*. Por último, como las dos letras más frecuentes en el texto cifrado son la *B* y la *Z*, sabemos que muy probablemente se corresponda una con la *A* y otra con la *E* (al ser estas las dos letras más

frecuentes del español). Como la B y la Z no están separadas por 4 *saltos*, a diferencia de la A y la E, sabemos que este texto no puede haber sido cifrado por *Cesar*.

Para descifrar este texto, usamos varios métodos. Por una parte, comparamos las tablas de frecuencias del texto cifrado y del español. Por otra parte, comparamos los *n-gramas* más frecuentes del español (obtuvimos esta información de la web *http://corpus.rae.es/frec/1000_formas.TXT*) con los *n-gramas* más frecuentes que aparecían en el texto.

Empezamos encontrando el *3-grama* más frecuente en el texto cifrado, *FJZ*, que se tenía que corresponder con *que*. Así, obtuvimos la sustitución $Z - e$, con lo que la letra B (la más frecuente del texto cifrado) se tenía que corresponder con *a*. La tercera letra más frecuente del texto cifrado, L, la asociamos a *o*. Analizando *2-gramas* y *3-gramas* más frecuentes, sacamos las correspondencias *HBDB - para* y *ÑZ - de*. A partir del *2-grama GB (la)* obtuvimos la correspondencia G - l y la correspondencia V - c la obtuvimos a partir de los *n-gramas VLX (con)* y *VSLX (cion)*. Al sustituir todas estas correspondencias en el texto cifrado, obtuvimos la correspondencia S - i a partir de la palabra *dia*. La correspondencia I - s se obtuvo a partir de los *2-gramas ZG (el), ZI (es), LI (os)* y X - n a partir de *BXÑL (ando), JXÑL (undo)*. Al realizar todas estas sustituciones, pude sacar el resto de correspondencias a partir del texto parcialmente descifrado. De la frase *aPosdespuesensu* obtuve la correspondencia P - $ñ$, de *CasRasonaQanlasAisAascosasenlacasadondesecreiaquecoordina* obtuve C - h, R - t, Q - b, A - m, de *lalluYiosatardedeUunio* obtuve Y - v, U - j, de *anparecidosEtraviesos* obtuve E - y, de *arcadioseNundoEaurelianoseNundo* obtuve N - g, de *durantelainOancia* obtuve O - f, de *sentimientodeKoKobra* obtuve K - z, de *actitudeseMtravagantes* obtuve M - x y, por último, de *Tilogramosdeoro* obtuve T - k.

Así, al realizar estas últimas sustituciones, obtuve el texto 5 descifrado. La clabe k viene dada por todas las sustituciones explicadas, que se enumeran a continuación: A - m, B - a, C - h, D - r, E - y, F - q, G - l, H - p, I - s, J - u, K - z, L - o, M - x, N - g, $Ñ$ - d, O - f, P - $ñ$, Q - b, R - t, S - i, T - k, U - j, V - c, W - w, X - n, Y - v, Z - e.