# Criptografía y Cifrado de Información 2021

Lab 01

15.julio.2021

En este laboratorio implementaremos un análisis de fuerza bruta por frecuencias, para varios cifrados sencillos.

El esquema para construir un análisis de frecuencias por fuerza bruta es el siguiente:

a) Implementar funciones para encriptar y decriptar un texto mediante un cifrado, *e.g.* cifrado *Caesar*, utilizando el alfabeto castellano (27 letras).
(Sugerencia: Probablemente antes se deba construir una función que limpie el texto plano, removiendo caracteres no alfabéticos.)

b) Construir una función que calcule las distribución de los caracteres que aparecen en el texto cifrado. Aquí, por ejemplo, se pueden usar funciones ya implementadas en NLTK. Sin embargo, se espera que su función calcule las probabilidades (las frecuencias divido el total de caracteres). (Sugerencia: Es recomendable completar las letras que no aparezcan en su texto, con probabilidad 0.)

c) Diseñar una métrica, que compare el parecido de la distribución de caracteres en su texto (cifrado o descrifrado), contra la distribución teórica de las letras del castellano. Para obtener la distribución teórica del español, basta conseguir alguna tabla de probabilidades en la web (si no está completa se debe completar las letras que falten).

d) Finalmente, construir una función que haga el proceso de Fuerza Bruta: un ciclo que barre todas las llaves posibles en el espacio de claves. Por cada clave, se almacena en un arreglo la métrica obtenida y se ordena de mejor a peor.
La función debe devolver las $k$ mejores claves. Se entiende que la clave más probable que descifra el texto es la que tiene mejor métrica.

1. Diseñar funciones de encriptado y decriptado, para un texto plano en castellano (27 letras), para los siguientes métodos:

   - Cifrado *Caesar*.
   - Cifrado afín.
   - Cifrado Vigenère.

   Para cada método, muestre ejemplos sencillos de encriptado y decriptado (para verificar que funcionan correctamente).

2. Implementar funciones para calcular la distribución de probabilidades de frecuencias en un texto cifrado. Como se indica en el ítem (b).
   Implementar una función de métrica para comparar dicha distribución contra la distribución teórica de las letras del castellano, como en (c).

3. **Decriptado por Fuerza Bruta.**
   En los archivos de texto `cipher1.txt`, `cipher2.txt` y `cipher3.txt` se encuentran tres textos cifrados, en los que se usaron diferentes métodos, como sigue:

   - `cipher1.txt`, cifrado *Caesar*.
   - `cipher2.txt`, cifrado afín.
   - `cipher3.txt`, cifrado Vigenère.

   En todos los casos, se ha usado el alfabeto castellano de 27 letras.

   Para cada uno de los textos cifrados arriba, implementar un análisis de fuerza bruta por frecuencias. Para cada uno, determinar cuál fue la clave utilizada en cada caso, y decriptar el mensaje.