Ex1: Quebrando Shift Cipher

O Que é a Cifra de César?

A Cifra de César é um tipo de cifra de substituição no qual cada letra no texto simples é substituída por uma letra com um número fixo de posições mais longe no alfabeto. Por exemplo, com uma mudança de 3 posições, a letra A seria substituída por D, B se tornaria E, e assim por diante.

Funcionamento da Cifra de César

1. **Escolha de uma Chave**:
    - A chave na Cifra de César é o número de posições que cada letra no texto é deslocada no alfabeto.

1. **Processo de Codificação**:
    - Cada letra do texto simples é deslocada para frente no alfabeto pelo número de posições especificadas pela chave, resultando na criação do texto cifrado.

1. **Exemplo Simples**:
    - Texto simples: "CRYPTO"
    - Chave: 3
    - Texto cifrado: "FUBSWR"

Vantagens e Limitações

- **Vantagens**:
    - Fácil de implementar e entender.
    - Pode ser útil para cifrar mensagens de baixa segurança.

- **Limitações**:
    - É criptograficamente fraca e facilmente decifrada através de métodos como análise de frequência ou tentativa de todas as possíveis chaves visto que são apenas 26 possíveis chaves.

Implementação

Criptografando a mensagem (cifraCeaser.py)

Para a criptografar a mensagem primeiramente o é perguntado ao usuário o tamanho do deslocamento. Que então pega a mensagem escrita no arquivo mensagem.txt e para cada caracter do alfabeto realiza um deslocamento da (posição do caracter no alfabeto + o deslocamento) mod 26. Posteriormente ele salva essa mensagem criptografada em um arquivo chamado cypher.txt.

Decriptografando

- **Força Bruta (forceDescripCeaser.py)**:
  Devido ao baixo número de combinações possíveis, é possível realizar a descodificação utilizando força bruta. para isso o pograma pega o texto cifrado em cypher.txt e converte para uma lista de caracteres onde para todos os 26 possíveis deslocamentos, é calculado a mensagem gerada por cada chave. Em seguida, fazendo o uso de dígrafos, o programa calcula qual é a resposta mais provavel. Assim retornando ao usuário a mensagem original.
- **Análise de frequência (descripCeaser.py)**:
  - Para esse método, diferente da força bruta, utilizaremos apenas 9 tentativas para descodificar a mensagem original. Esse método utiliza a frequência de cada letra em português para determinar quais seria os deslocamentos mais prováveis. Para isso fazemos um ranqueamento das letras em português para descobrimos as mais recorrentes, em seguida, pegamos a frequência de cada letra dentro do código cifrado e supondo que o texto original era em português, assumimos que a frequência de letras da cifra deve ter um alto grau de correlação com a a frequência de letras em português. Com isso já que cada letra do código cifrado deve corresponder a uma letra do alfabeto deslocado por um número, é escolhido os 3 caracteres mais frequêntes da cifra e é testado o deslocamento de cada um desses caracteres com os 3 caracters mais comuns do português, assim resultando em apenas 9 tentativas. Em seguida, similarmente ao método anterior para se testar cada um dessas tentativas é atribuido um valor numérico da probabilidade da combinação de todos os dígrafos dentro da tentativa, e aquela que possuir uma pontuação maior, é provavelmente a mensagem texto original.