Cifra de Vigenère - Trabalho 1

Pedro Trajano - 190055251

Outubro 2023

1 Introdução

A cifra de Vigenère é um método de criptografia no qual, dada uma mensagem M e uma chave K, uma mensagem criptografada $C = Enc_K(M)$ é gerada. O método é descrito da seguinte forma: mapeamos a mensagem e a chave para suas versões em números inteiros, indicando suas posições no alfabeto, por exemplo, "a = 0", "b = 1" e assim por diante. Para cada letra da mensagem, selecionamos na chave a letra que está na mesma posição. Se a chave for menor que a mensagem, aplicamos o resto da posição da letra na mensagem na chave. Em seguida, somamos as duas letras e colocamos o resultado na mesma posição na mensagem criptografada. Isso pode ser expresso pela seguinte equação:

$$C_i = (M_i + K_i)26 \tag{1}$$

Para descriptografar a cifra, basta subtrair a cifra pela chave:

$$M_i = (C_i - K_i)26 \tag{2}$$

2 Análise da Cifra de Vigenère

Para tentar quebrar a cifra de Vigenère, podemos seguir o seguinte raciocínio. Primeiro, precisamos determinar possíveis tamanhos de chaves. Em seguida, separamos a cifra em diferentes substrings e calculamos a frequência das letras em cada uma delas.

2.1 Determinar tamanho da chave

Para encontrar um possível tamanho de chave, escolhemos um valor N. Calculamos todas as substrings presentes na cifra com tamanho N, mantendo a ordem, e então calculamos a distância entre substrings idênticas. Essa distância revela possíveis padrões na cifra. Como a chave geralmente é menor que a mensagem, certas partes da mensagem estarão cifradas pelo mesmo K_i. Isso nos fornece um ponto de partida para o tamanho da chave.

2.2 Agrupamento de letras

A partir do passo anterior, podemos identificar um possível tamanho da chave N. Por exemplo, se a chave tiver tamanho 5, todos os termos $5i$ serão cifrados pela mesma letra. Isso significa que podemos reduzir o problema para uma série de cifras de César. Criamos NN vetores representando as substrings que têm a mesma letra em suas criptografias.

2.3 Análise de Frequências de Letras

O último passo para encontrar possíveis mensagens é utilizar dados externos, como a frequência com que as letras são utilizadas na língua que está sendo cifrada. Isso nos proporciona uma estimativa estatística de quão próximas as letras da possível mensagem M'

estão da frequência base. O tamanho da mensagem influencia na qualidade do ataque, pois mensagens menores podem se desviar da frequência padrão.

Para cada agrupamento de letras, calculamos a frequência das letras na cifra. Em seguida, deslocamos os termos dessa frequência ao longo das letras e comparamos com a frequência externa. Podemos utilizar métodos como o qui-quadrado ou o produto escalar entre as frequências. Optamos pelo produto escalar, pois para dois vetores vv e v_i, sendo v_i o vetor vv deslocado i posições, o valor máximo desse produto escalar será $v \cdot v_0$. Procuramos maximizar o valor do produto escalar das duas frequências. Encontramos assim uma possível chave e, subsequentemente, uma possível mensagem.

3 Conclusão

O método apresentado é eficaz para mensagens longas e chaves curtas, onde podemos definir frequências com uma certa confiança estatística. No entanto, apresenta limitações em casos em que essas condições não são satisfeitas.