Cifra de Vigenère: Fundamentos, Implementação e Criptoanálise

A Cifra de Vigenère, concebida no século XVI, é um método de criptografia por substituição polialfabética que historicamente representou um avanço em relação às cifras monoalfabéticas. Sua robustez advém da utilização de múltiplos alfabetos de cifragem, selecionados ciclicamente por uma palavra-chave.

Mecanismo da Cifra de Vigenère

A operação da cifra baseia-se na *tabula recta*, uma matriz 26x26 onde cada linha representa um alfabeto de César, deslocado em relação ao anterior. A palavra-chave determina qual linha (alfabeto de cifragem) é utilizada para cada caractere do texto plano.

Processo de Cifragem:

1. **Palavra-Chave:** Uma palavra-chave é definida e repetida, se necessário, para corresponder ao comprimento do texto plano.
2. **Mapeamento Criptográfico:** Para um caractere P_i do texto plano e o caractere correspondente K_i da palavra-chave (alinhada), o caractere cifrado C_i é obtido pela interseção da linha K_i e coluna P_i na *tabula recta*.
    - **Aritmética Modular:** Numericamente, convertendo letras para inteiros (A=0, ..., Z=25), a cifragem é: C_i = (P_i + K_i) mod 26.

Processo de Decifragem:

1. A decifragem é a operação inversa: P_i = (C_i - K_i + 26) mod 26. O termo +26 assegura um resultado positivo antes da operação modular.

Implementação Algorítmica em Python

No código vigenere_attack_multilang (referenciado implicitamente), as funções criptografar_vigenere e descriptografar_vigenere implementam este processo:

- **normalizar_texto(texto, codigo_idioma):** Precede as operações criptográficas. Realiza a conversão para maiúsculas, remoção de caracteres não alfabéticos e, para o português (PT), a transliteração de caracteres acentuados para seus equivalentes não acentuados. Esta etapa garante a conformidade com o alfabeto de 26 caracteres utilizado.
- **criptografar_vigenere(texto_plano_original, chave, codigo_idioma):**
    1. Normaliza o texto_plano_original.
    2. A chave é convertida para maiúsculas.
    3. Itera sobre os caracteres do texto plano normalizado, aplicando a aritmética modular (valor_numerico_P + valor_numerico_K) % 26 para cada par de

caracteres (texto plano, chave). O operador módulo (%) sobre o tamanho da chave gerencia sua repetição cíclica.

- **descriptografar_vigenere(texto_cifrado_normalizado, chave):**
  1. Opera sobre um texto_cifrado_normalizado.
  2. Aplica a aritmética modular inversa (valor_numerico_C - valor_numerico_K + 26) % 26.

Criptoanálise da Cifra de Vigenère por Análise de Frequência

A principal vulnerabilidade da Cifra de Vigenère reside na periodicidade da palavra-chave. Se a chave possui comprimento L, o i-ésimo caractere, o (i+L)-ésimo, o (i+2L)-ésimo, e assim por diante, são todos cifrados com a mesma letra da chave, ou seja, com o mesmo deslocamento de César. Isso transforma o problema polialfabético em L problemas monoalfabéticos distintos.

A criptoanálise efetiva compreende duas fases:

1. **Determinação do Período da Chave (L):**
   - **Teste de Kasiski:** Identifica sequências de caracteres repetidas no criptograma. As distâncias entre ocorrências dessas sequências são frequentemente múltiplos do comprimento da chave L. Fatorando essas distâncias, pode-se inferir L.
   - **Índice de Coincidência (IC):** Este método estatístico, implementado no código, avalia a probabilidade de dois caracteres aleatoriamente selecionados de um texto serem idênticos.
     - Idiomas naturais possuem um IC característico (e.g., Inglês ≈ 0.067; Português ≈ 0.078).
     - Um texto com distribuição uniforme de caracteres (aleatório) possui IC ≈ 1/26 ≈ 0.0385.
     - Ao segmentar o criptograma em L colunas para um L candidato, se L for o período correto, cada coluna (sendo uma cifra de César) exibirá um IC próximo ao do idioma subjacente. Para um L incorreto, as colunas terão um IC médio mais próximo ao de um texto aleatório.
2. **Determinação dos Caracteres da Chave:**
   - Com L conhecido, o criptograma é decomposto em L textos intercalados, cada um cifrado monoalfabeticamente.
   - Para cada uma dessas L colunas, aplica-se a análise de frequência. Testa-se cada um dos 26 possíveis deslocamentos de César.
   - A **estatística Qui-Quadrado (χ^2)** é utilizada para quantificar a semelhança entre a distribuição de frequência de uma coluna decifrada (com um deslocamento candidato) e a distribuição de frequência esperada para o

idioma. O deslocamento que minimiza o valor de χ^2 é o mais provável para aquela coluna, revelando a letra correspondente da chave.

Implementação da Criptoanálise no Código Python

A função realizar_ataque_frequencia_vigenere no código executa a criptoanálise:

1. **Configuração e Normalização:**
   - Seleciona os parâmetros estatísticos (frequencias e ic_esperado) do idioma com base no codigo_idioma.
   - O texto_cifrado_original é processado por normalizar_texto.
2. **Estimativa do Período da Chave (estimar_tamanho_chave):**
   - Itera sobre comprimentos de chave candidatos (de 1 a max_tamanho_chave).
   - Para cada tamanho_chave_teste, o criptograma é dividido em colunas.
   - O IC médio das colunas é calculado (calcular_indice_coincidencia).
   - O tamanho_chave_teste que produz um IC médio mais próximo do ic_esperado_idioma_ref é selecionado como tamanho_chave_estimado.
3. **Decomposição em Colunas (dividir_em_colunas):**
   - O criptograma é segmentado nas colunas_texto_cifrado com base no tamanho_chave_estimado.
4. **Identificação dos Caracteres da Chave (encontrar_letra_chave_da_coluna):**
   - Para cada texto_coluna_atual:
     - São testados os 26 deslocamento_teste possíveis.
     - A coluna é decifrada com o deslocamento corrente.
     - A estatística χ^2 é calculada (calcular_estatistica_qui_quadrado) comparando a distribuição da coluna decifrada com frequencias_idioma_ref.
     - O deslocamento que minimiza χ^2 determina a letra da chave para a coluna.
   - As letras da chave identificadas são concatenadas, formando a chave_estimada_final.
5. **Decifragem com a Chave Estimada:**
   - O texto_cifrado_normalizado é decifrado com a chave_estimada_final através de descriptografar_vigenere.