UnB Instituto de Ciências Exatas Departamento de Ciência da Computação

Segurança Computacional - 2025/1

Professora Priscila Solis

João Victor Prata (202028857), Nikolas Negrao (202024722)

Relatório Trabalho de Implementação 1 - Cifra de Vigenère

Este relatório tem como objetivo descrever sobre a aplicação e funcionamento da cifra de Vigenère, apresentando duas partes: o cifrado/decifrador e o ataque de recuperação de senha por análise de frequência.

Descrição dos Arquivos:

- **vigenere_cypher.py**: arquivo Python contendo o código fonte da cifra com sua implementação e do ataque e sua implementação.

Funcionamento da cifra de Vigenère:

A cifra de Vigenère é uma técnica do tipo substituição polialfabética. Utiliza uma palavra-chave para determinar diferentes deslocamentos para cada letra da mensagem.

Cada letra da chave define um deslocamento com base em sua posição no alfabeto. Por exemplo, a letra 'B' representa um deslocamento de 1, 'C' representa 2, e assim por diante. A chave é repetida ciclicamente sobre o texto, e cada letra da mensagem é cifrada usando o deslocamento correspondente.

A cifra de Vigenère pode ser quebrada usando métodos estatísticos, usando o cálculo do índice de coincidência e a análise de frequência. A decifragem é mais

fácil quando o texto cifrado é longo e a chave é curta. Uma demonstração desse processo foi elaborada no desenvolvimento deste projeto.

Funcionamento do código:

- **Parte I: cifrador / decifrador:**

A função *vigenere_cifrar* recebe como argumentos o texto a ser cifrado e a chave. São removidos os espaços do texto, e todas as letras são convertidas para letras maiúsculas. Em seguida, para cada letra no texto, é feito um deslocamento no alfabeto com base na letra correspondente da chave (também convertida para maiúscula). Esse deslocamento é feito somando a posição da letra da chave à posição da letra do texto (ambas relativas à letra 'A'), e aplicando módulo 26 para manter o resultado dentro dos limites do alfabeto. O resultado é o texto cifrado.

A função *vigenere_decifrar* faz processo reverso da função anterior. Ela também remove espaços e converte o texto para maiúsculo. Em seguida, para cada letra do texto cifrado, realiza-se o deslocamento inverso subtraindo a posição da letra da chave da posição da letra do texto cifrado (novamente com módulo 26). Isso recupera o texto original.

No final do código, na seção de testes e demonstração, a mensagem de exemplo *“ATTACKATDAWN”* é acompanhada por uma senha *“LEMON”*, gerando o texto cifrado *“LXFOPVEFRNHR”*. Em seguida, a mesma mensagem cifrada é usada com a senha para ser decifrada, conseguindo assim recuperar a mensagem decifrada inicial *“ATTACKATDAWN”*.

- **Parte II: ataque de recuperação de senha por análise de frequência:**

A função *analise_de_freq* calcula a frequência relativa de cada letra (A-Z) em um texto cifrado. Ela ignora qualquer caractere que não seja letra e retorna um dicionário com as porcentagens de ocorrência de cada letra, usada para

comparar com as distribuições de frequência esperadas da língua portuguesa ou inglesa.

A função *qui_quadrado* implementa o cálculo do teste do qui-quadrado, utilizado para medir a similaridade entre duas distribuições de frequência: a observada (da função *analise_de_freq*) e a esperada (definida nos dicionários *freq_portugues* e *freq_ingles*). Para cada possível deslocamento de uma letra do alfabeto, calcula-se o valor do qui-quadrado entre a distribuição de frequência resultante e a esperada. O deslocamento que resulta no menor valor do qui-quadrado é considerado o mais provável.

A função *calcular_tamanho_chave* tenta estimar o comprimento da chave utilizada na cifra de Vigenère. Isso é feito usando o Índice de Coincidência (IC), uma medida da probabilidade de duas letras aleatórias em um texto serem iguais. Para cada possível comprimento de chave (de 1 até 20 por padrão), o texto cifrado é dividido em subsequências e é calculado o IC médio entre elas. O tamanho da chave com o maior IC médio é escolhido como o mais provável.

Essas funções são combinadas dentro da função *descobrir_chave*, que primeiro estima o tamanho da chave e, em seguida, tenta determinar cada letra da chave individualmente, com base na análise de frequência e no teste do qui-quadrado.

Dois dicionários foram elaborados contendo frequências de letras para a língua portuguesa e inglesa de acordo com a frequência apresentada no roteiro (citado também nas referências bibliográficas no final deste relatório).

Para as duas línguas portuguesa e inglesa, uma mensagem para cada língua é cifrada com uma chave específica. Tais mensagens são cifradas usando a função *vigenere_cifrar*.

Em seguida, o processo reverso é feito para descobrir a chave a partir da mensagem cifrada (usando a função *descobrir_chave*), e depois decifrar a mensagem com a senha descoberta (com a função *vigenere_decifrar*).

Os algoritmos do processo reverso de descoberta de senha e decifragem geram resultados um pouco destoantes da mensagem original e senha original, mas isso é compreensível, considerando que o deciframento de Vigenère é feito por estimativa com base na frequência de letras. Mesmo assim, o resultado obtido ainda é legível, podendo ser compreendida a mensagem mesmo com algumas letras incorretas. O fato de as mensagens terem mais de 200 caracteres facilita o processo de decifragem.

Por fim, são printados os resultados obtidos. A função *print_block* é para formatação, deixando o texto organizado em blocos de forma a facilitar a leitura do resultado.

Repositório com código fonte:

- **NikolasNP/SC-Projeto1:** https://github.com/NikolasNP/SC-Projeto1/tree/main

---

Referências bibliográficas:

Este projeto foi realizado de acordo com o roteiro do Projeto 1 da disciplina Segurança Computacional, ministrada pela professora Priscila Solis.

- **Frequência de letras**: https://pt.wikipedia.org/wiki/Frequ%C3%AAncia_de_letras
- **Funcionamento da cifra de Vigenere:** https://www.youtube.com/watch?v=SkJcmCaHqS0
- **Quebrando a cifra de Vigenere:** https://www.youtube.com/watch?v=P4z3jAOzT9I&fbclid=IwAR0Aj7nSiyUmdf1XaRI7