Projeto 1 da Disciplina CIC0201 (Segurança Computacional) - Cifra de Vigenère

Maria Eduarda Santos - 190092556[1],
Kléber Rodrigues da Costa Júnior - 200053680[1],

[1]Departamento de Ciência da Computação (CIC) – Universidade de Brasília (UnB)
Brasília, DF – Brasil

{190092556, 200053680}@aluno.unb.br

1. Introdução

A cifra de Vigenère é um tipo de cifra de substituição polialfabética, que foi inventada por Blaise de Vigenère no século XVI. A cifra de Vigenère é um avanço em relação à cifra de César, que é uma cifra de substituição simples, porque usa várias tabelas de substituição, cada uma correspondendo a uma letra da chave [Bruen and Forcinito 2011].

Figura 1. Blaise de Vigenère

A chave é uma palavra ou frase, que é repetida várias vezes para formar uma chave longa o suficiente para criptografar todo o texto. Cada letra da chave é usada para deslocar a letra correspondente do texto original em um determinado número de posições no alfabeto. A cifra de Vigenère foi usada por séculos como um método de criptografia, mas eventualmente foi quebrada por Charles Babbage e Friedrich Kasiski no século XIX. Eles descobriram que, se o comprimento da chave fosse conhecido, a cifra de Vigenère poderia ser facilmente quebrada usando uma técnica chamada análise de frequência. Essa técnica envolve contar com que frequência cada letra aparece no texto cifrado e usar isso para inferir informações sobre a chave.

Embora a cifra de Vigenère não seja mais usada para criptografia, ela ainda é interessante do ponto de vista histórico e é um exemplo de como a criptografia evoluiu ao longo do tempo. Além disso, a cifra de Vigenère ainda é usada em alguns jogos e quebra-cabeças criptográficos [Anderson 2010].

2. Análise de Frequência - Quebrando a Cifra de Vigenère

O processo de análise de frequência é uma técnica criptográfica para quebrar cifras de substituição, incluindo a cifra de Vigenère [fre]. A técnica se baseia no fato de que algumas letras são mais comuns do que outras em um idioma, e essas frequências podem ser usadas para quebrar a criptografia.

O primeiro passo na análise de frequência é contar o número de ocorrências de cada letra no texto cifrado. Isso geralmente é feito usando um histograma, que mostra a

frequência de cada letra. As letras mais comuns no texto cifrado provavelmente correspondem às letras mais comuns no idioma usado para escrever o texto original.

O próximo passo é tentar descobrir o comprimento da chave. Isso pode ser feito usando o índice de coincidência, que é uma medida da probabilidade de que duas letras aleatórias no texto cifrado sejam iguais. Se o índice de coincidência for muito alto, isso sugere que o comprimento da chave é curto. Se for muito baixo, isso sugere que o comprimento da chave é longo.

Uma vez que o comprimento da chave foi estimado, a cifra de Vigenère pode ser quebrada usando tabelas de frequência. Para cada letra na chave, uma tabela de frequência é criada, mostrando a frequência de cada letra no texto cifrado quando aquela letra é usada para criptografar o texto original. Isso pode ser feito calculando a frequência de cada letra em cada deslocamento de posição no alfabeto correspondente à letra da chave. Com as tabelas de frequência para cada letra da chave, é possível criar uma tabela de frequência geral, que mostra a frequência de cada letra no texto cifrado quando a chave inteira é usada para criptografar o texto original. Usando essa tabela de frequência geral, é possível tentar deduzir as letras da chave e, em seguida, descriptografar o texto original usando as tabelas de frequência correspondentes.

3. Implementação

A linguagem de programação Python foi utilizada na construção do algoritmo, a fim de tornar simples a manipulação de textos que serão dados como entrada.

A ideia é utilizar uma tabela (também chamada de tabula recta) que consiste em uma matriz de letras do alfabeto. A primeira linha contém as letras do alfabeto na ordem normal, enquanto as demais linhas são geradas a partir de um deslocamento circular em relação à primeira linha. Cada letra da palavra-chave corresponde a uma linha da tabela, e a cifra é obtida cifrando cada letra do texto original com a letra da tabela que se encontra na mesma linha e coluna da letra da palavra-chave correspondente.

O código implementa as seguintes funcionalidades:

- `key_pattern(key, text)`: essa função recebe a palavra-chave `key` e o texto a ser cifrado/decifrado text, e retorna a palavra-chave repetida até que ela tenha o mesmo tamanho do texto. Caso a palavra-chave seja menor que o texto, ela é repetida várias vezes até que tenha o mesmo tamanho do texto. Caso a palavra-chave contenha caracteres que não são letras do alfabeto, é lançada uma exceção indicando que a chave é inválida.
- `crypt_decrypt(key, text, option)`: essa função recebe a palavra-chave key, o texto a ser cifrado/descifrado `text` e a opção option (que deve ser 'C' para cifrar ou 'D' para decifrar). A função retorna o texto cifrado/decifrado de acordo com a palavra-chave e a opção escolhida. Caso o texto seja vazio ou a palavra-chave tenha tamanho menor que 2 (no caso de cifrar) ou 1 (no caso de descifrar), é lançada uma exceção indicando que o tamanho do texto ou da chave é inválido. A função implementa a lógica da cifra de Vigenère, cifrando cada letra do texto original com a letra da tabela que se encontra na mesma linha e coluna da letra da palavra-chave correspondente.
- `clean_text(text)`: essa função recebe um texto e retorna uma versão do

texto contendo apenas letras maiúsculas do alfabeto. Isso é útil para realizar a análise de frequência dos caracteres.

- `key_size(text)`: essa função recebe um texto cifrado e implementa uma técnica de criptoanálise para estimar o tamanho da palavra-chave utilizada para cifrar o texto. A técnica consiste em comparar triplas de caracteres do texto cifrado para encontrar quais triplas possuem a mesma sequência de caracteres. A ideia é que, se duas triplas possuem a mesma sequência de caracteres, elas foram cifradas utilizando a mesma letra da palavra-chave. A distância entre as triplas no texto cifrado é então utilizada para tentar inferir o tamanho da palavra-chave. A função retorna o tamanho da palavra-chave estimado e pergunta ao usuário se ele deseja continuar com esse tamanho.

4. Arquivos Fonte e Estrutura de Arquivos

O projeto se encontra no GitHub de Kléber Rodrigues, o qula dispõe de um arquivo `README.md`, que detalha o processo de execução do projeto para testes diversos. Os arquivos de teste se encontram na pasta `Testes`, os quais podem ser, ou não, solicitados na execução do arquivo `main.py`, que faz uso do arquivo `vigenere.py`, o qual realiza a quebra de um texto cifrado por meio de análise de frequência. Todos os arquivos se encontram bem documentos, dispondo de comentários e explicações do funcionamento do algoritmo.

Referências

The vigenère cipher: Frequency analysis.

Anderson, R. (2010). *Security Engineering: A Guide to Building Dependable Distributed Systems*. Wiley.

Bruen, A. and Forcinito, M. (2011). *Cryptography, Information Theory, and Error-Correction: A Handbook for the 21st Century*. Wiley Series in Discrete Mathematics and Optimization. Wiley.