Implementação e Ataque da Cifra de Vigenère

Thales Lima Menezes
Dep. Ciência da Computação
Universidade de Brasilia
Brasília, Brasil
170045919@aluno.unb.br

***Resumo*—Este trabalho detalha uma implementação da cifra de Vigenère. Além disso, é explicado um ataque à cifra que possibilita a extração automática da chave utilizada para cifrar uma mensagem, desde que a mensagem seja um texto escrito em inglês ou português e possua tamanho suficientemente grande. É também explicada a implementação dos algoritmos propostos, escrita em Python.**

***Index Terms*—cifra, Vigenère, criptografia**

I. Introdução

A cifra de Vigenère é um método de criptografia que usa uma série de diferentes cifras de César baseadas em letras de uma senha. Trata-se de uma versão simplificada de uma mais geral cifra de substituição polialfabética, inventada por Leon Battista Alberti cerca de 1465.

Neste trabalho, foi desenvolvido um programa que possibilita a cifração e decifração de uma mensagem com uma cifra de Vigenère modificada, que é descrito na seção III. Além disso, é descrito um método de ataque à cifra, que tenta automaticamente descobrir a chave com a qual um texto foi cifrado, na seção IV.

II. Estruturação do Projeto

Neste trabalho, foi implementado um codificador e decodificador da cifra de Vigenère em Python. O projeto foi construído utilizando a versão 3.10.4 e testado em Linux.

O projeto foi dividido em três partes:

1) **cifra.py:** onde estão implementadas as funcionalidades de cifração e decifração da cifra de Vigenère, dada uma chave. Esse módulo é explicado na seção III.
2) **ataque.py:** onde está implementada a função de ataque da cifra de Vigenère, que encontra uma chave automaticamente, dada uma mensagem cifrada. Esse módulo é explicado na seção IV.
3) **test:** possui alguns, poucos, testes, que podem ser executados com o comando `make test`.

III. Implementação

A cifra de Vigenère cada letra da mensagem é "combinada"com uma da chave, por meio da soma módulo 26 dos valores das letras (geralmente, o valor de A é 0, o de B é 1, e assim por diante).

Note que, ao analisarmos a função matemática que descreve a cifra, é possível descrever a função de decifração utilizando à de cifração.

A. Execução do Programa

É necessário especificar uma senha de cifração, por meio da variável de ambiente **PASSWORD**. Fora isso a execução do arquivo cifra.py segue o seguinte formato: **./cifra.py message [-h—–help] [-d]** sendo **-h** para imprimir texto de ajuda, **-d** para decifrar ao invés de cifrar.

IV. Ataque

A. Ideia

É impossível projetar um ataque "universal"à cifra de Vigenère, visto que uma mensagem aleatória, quando cifrada com uma chave qualquer, gera texto cifrado aleatório.

Apesar disso, se a mensagem tiver alguma propriedade que conhecemos a priori, a cifra de Vigenère pode ser quebrada. Assim, nessa seção consideraremos que a mensagem cifrada é um texto relativamente longo, escrito em português ou inglês. O ataque descrito também se aplica a outras línguas, mas focamos nessas duas. Duas propriedades importantes de textos em português e inglês é esperado que as frequências dos caracteres sigam uma distribuição similar a outros textos escritos na mesma língua.

Sabendo disso, podemos elaborar uma estratégia para encontrar uma chave provável, dado um texto cifrado e a informação da língua em que a mensagem foi escrita. A estratégia consiste no seguinte algoritmo:

1) Definir o intervalo de tamanhos possíveis para a chave entre 1 e 20.
2) Para cada tamanho K:
   a) Filtramos as letras do criptograma que foram cifradas pelo i caractere da chave, sendo $0 < i < K$
   b) Testamos todas as 26 letras do alfabeto como chaves de decifração
   c) Comparamos a frequência da mensagem decifrada com a frequência esperada de caracteres para a língua portuguesa ou inglesa.
   d) Selecionamos a letra que possuir a maior compatibilidade com a frequência esperada para a língua
   e) Repetimos o procedimento acima para selecionar letras até completar a chave de tamanho K
3) Após o loop, teremos 20-1=19 chaves com tamanhos variados

4) Serão ordenadas as chaves de acordo com a frequência de letras na mensagem decifrada em comparação com a frequência esperada para a lingua
5) A saída será a chave com maior compatibilidade

B. Execução do Programa de Ataque

É possível especificar a língua que será utilizada, padrão inglês, por meio da variável de ambiente **LANG**. Também é possível personalizar o intervalo de tamanhos possíveis para a chave utilizando as variáveis **MAX_KEY_LENGTH MIN_KEY_LENGTH**. Fora isso a execução do arquivo ataque.py segue o seguinte formato: **./ataque.py message [-h—–help]** sendo **-h** para imprimir texto de ajuda.

V. Conclusão

Neste trabalho, foi introduzida uma cifra de Vigenère e foi feito uma apresentação geral da implementação em Python, podendo adquirir mais informações específicas ao executar os testes e leitura da documentação das funções e código como um todo.

Além disso, foi detalhada uma maneira de quebrar a cifra citada anteriormente, de modo que é possível descobrir automaticamente qual chave foi utilizada para cifrar um texto, desde que a mensagem original seja escrita em inglês ou português, e tenha tamanho consideravelmente maior que o da chave.

Referências