Criptoanálise - Cifra de Vigènere

Igor Pereira Dourado

[1] Escola Politécnica, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Porto Alegre – RS – Brasil

igor.d@edu.pucrs.br

Abstract. *This article aims to describe an implementation made to find the text clear from a cipher text with the Vigenère Cipher, it is important to emphasize that the key used in the encryption is unknown at first, before deciphering. Using ideas and methods such as techniques such as Friedman's coincidence index and alphabetical substitution algorithms to arrive at a coherent and logical end result, with the previously unknown key in hand.*

Resumo. *O presente artigo tem por objetivo a descrição de uma implementação feita para achar o texto claro a partir de um texto cifrado com a Cifra de Vigenère, é importante ressaltar que a chava usada na cifragem é desconhecida em um primeiro momento, antes de decifrarmos. Utilizando ideias e métodos como a técnicas como o índice de coincidência de Friedman e algoritmos de substituições alfabéticas para chegar a um resultado final coerente e lógico, com a chave, anteriormente desconhecida, em mãos.*

1. Introdução

A cifra de Vigenère é uma cifra de substituição de várias tabelas que consiste em evolução cifra de César. A cifra de César é criptografada deslocando cada letra no texto simples por um um certo número de vezes. Na cifra de Vigènere, em vez de executar um único o deslocamento em todo o texto simples, se uma chave for usada para determinar o valor O deslocamento é diferente em toda a mensagem, portanto, não criptografa um caractere de cada vez Composto pelos mesmos caracteres, sendo essa a força da cifra (MIT, 2019).

Na prática, uma cifra consiste em um conjunto de chaves César distintas. entretanto, a chave também tem a capacidade de alterar a forma do caractere original Baseado em mover a posição de caractere para caractere em cada caractere da tecla Emparelhe-o consecutivamente em texto simples. Para encontrar esta chave, você pode usar Kasiski ou método de Friedman (Que foi utilizado para resolver esta implementação junto com a linguagem python e algumas bibliotecas).

Neste artigo será apresentada uma implementação em Python, inicialmente pensada em Java, para a decifragem da cifra de Vigenère com chave desconhecida. Será mostrada uma descrição de cada etapa realizada, detalhando o funcionamento do código. A solução implementada espera receber um texto cifrado desconhecido, e, ao final, mostra ao usuário o texto claro em português ou inglês decifrado [**?**]

2. Base teórica

O problema envolve quebrar o texto cifrado para revelar o texto simples criptografado com Vigênia. Isso não é possível até que o tamanho da chave seja determinado, que por sua vez

explora o estudo de frequências de caracteres entre diferentes idiomas, o que é uma falha fundamental nessa técnica de criptografia. A ação é calculada por Coincidentemente, na permutação dos caracteres cifrados, atinge-se o índice da língua Em português, isso está relacionado ao valor estatístico de repetição de caracteres, não necessariamente ao valor dos próprios caracteres, mascarando-se uns aos outros.

3. Algoritmo, descobrindo o tamanho da chave e solução

A primeira etapa da decifragem consiste na descoberta da chave utilizada para cifrar o texto, visto que apenas o texto cifrado foi fornecido. Para descobrir a chave, é necessário primeiramente descobrir seu tamanho, ou seja, quantos caracteres possui. que pode ser revelado conforme proximidade com o de índice de coincidência, que nos é dado pela seguinte fórmula:

.

.

.

.

$$\frac{\sum_{i=0}^{25} fi(fi-1)}{n(n-1)}$$

Para aproximar os tamanhos de chave possíveis, usa-se uma tabela para armazenar diferentes índices de correspondência entre tamanhos de chave maiores que 1 Verificado. tamanho 1 não é considerado porque torna a cifra simples e trivial como a de César. O tamanho da chave é então aplicado à cifra segmentando-a. Fragmentos na mesma proporção são entendidos aqui como conjuntos de caracteres. Todo para cada um desses conjuntos e calcular o índice médio de coincidência, e quando esses O cálculo muitas vezes corresponde ao próprio idioma, neste caso português e inglês, é Candidatos encontrados para possíveis tamanhos de chave.

Com a espessura da chave já de conhecimento, o texto será dividido para coincidir com esse tamanho em questão, e então, para cada um desses blocos tomamos que o n-ésimo caractere foi cifrado com o n-ésimo caractere da chave. O teste de hipóteses vai de encontro nesse sentido, ele é realizado para procurar convergências entre os resultados gerados das duas distribuições, buscando os valores dos caracteres da chave. Em primeiro lugar, cada caractere de um bloco se desloca até a posição específica 25, da última letra do alfabeto, e utiliza no processo a letra mais frequente do alfabeto (letra "E"para o inglês e letra "A"para o Português). Realizadas as flutuações é conferido a ocorrência de cada caractere sobre o respectivo bloco em questão, e para cada uma dessas ocorrências, ocorre o cálculo de distribuição sobre o numero total de ocorrências. Por fim, a distribuição incrementa ao somatório e o resultado fica armazenado em lista, e o ındice com menor valor na distribuição se torna candidato a deslocamento. Após essas etapas, a letra mais frequente de cada alfabeto da respectiva linguagem selecionada deve revelar o caractere da chave via soma dos índices.

4. Decifrando o texto

Agora com a chave conhecida, a próxima etapa consiste em decifrar todo o texto cifrado pelas Cifras de César simples agora conhecidas. A descriptografia é o mesmo processo da cifragem só que invertido, subtraindo a chave em vez de adicionar, para retornar ao valor

original em texto. Então, basta utilizar o caractere da chave para descobrir o quanto o caractere cifrado foi deslocado, retornando o caractere para sua posição original. Repetindo esse processo por todo texto cifrado irá resultar no texto claro até então desconhecido. Exemplo de parte do texto cifrado fornecido e parte do texto claro encontrado:

Arquivo “plaintextpt.txt”:

Texto cifrado: cyyzvmgurwbszxmehacexuzyfgqeoslnuqqpijhpxmoelhaxmhvi...

Texto claro: quemhacincoentaannostivesseacoragemdepublicarumlivrocomo...

Conclusão

O desenvolvimento desse trabalho me permitiu um maior entendimento do funcionamento da cifra de Vigenère e de qual abordagem deve ser feita no planejamento de algoritmos para decifragem de um texto em que se desconhece informações da chave utilizada na cifragem. Conclui-se que a implementação atingiu o objetivo e conseguiu corretamente determinar a chave e descobrir o texto claro.

5. Referências

Data Science Central (2022) “Chi-Square Statistic”, https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/chi-square

Learn Cryptography (2022) “Vigenère Cipher”, https://learncryptography.com/classical-encryption/vigenere-cipher.