Relatório Segurança Computacional - Projeto 1

Ândrey Galvão Mendes - 18/0097911
Alexandre Augusto de Sá dos Santos - 15/0056940

1 de agosto de 2022

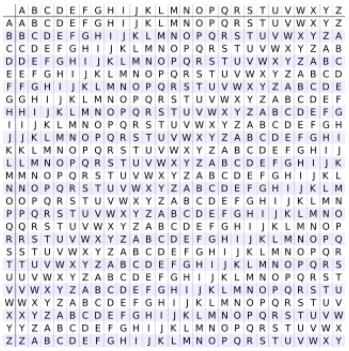
1 Introdução

A criptografia consiste de um conjunto de técnicas e princípios para tornar ilegível alguma mensagem, de forma que apenas as pessoas que conhecem o método utilizado consigam obter a mensagem original. Será explorado neste trabalho a cifra de Vigenère. Para tanto, três funcionalidades serão implementadas, a cifragem de uma mensagem utilizando uma chave, a decifragem de uma mensagem também utilizando chave, e por fim a recuperação da chave a partir da analise de frequência.

2 Implementação

2.1 Cifrador

A cifra de Vigenère utiliza uma chave, onde cada letra da mensagem é mapeada a uma outra letra utilizando a chave.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Figura 1: Mapeamento da cifra de Vigenère

Como pode ser visto na tabela acima, cada linha da tabela é deslocada para a esquerda um número diferente de vezes. Podemos relacionar o número de vezes que cada linha é deslocada associando cada uma das letras do alfabeto a um índice, começando de 0 e indo até 25. Neste caso, o deslocamento de cada linha é exatamente igual ao índice da letra na primeira coluna, correspondente aquela linha.

Sendo assim, usando este calculo para cifrar a mensagem, só será necessário que sejam somados os índices da letra da mensagem e da letra correspondente da chave. Ou seja

$$cifra = (mensagem + chave) mod 26 \tag{1}$$

2.2 Decifrador

Para decifrar a mensagem, deve-se observar na tabela a linha da letra da chave e encontrar naquela linha a letra da mensagem cifrada, podendo então encontrar a letra da mensagem original na primeira linha da tabela. Para isso podemos fazer o processo inverso ao que foi feito durante a cifragem. Ou seja, basta que saibamos a diferença entre cada letra da mensagem cifrada e a letra correspondente da chave. De forma que cada letra da mensagem será dada por

$$mensagem = (cifra - chave) mod 26 \tag{2}$$

2.3 Recuperação da chave

Nos procedimentos para descobrir o tamanho de chave e encontrá-la são utilizadas fórmulas estatísticas, que serão explicadas ao decorrer dessa seção. O primeiro processo é descobrir o tamanho da chave, para isso o texto cifrado foi dividido em conjuntos de letras separadas por um intervalo, sendo esse intervalo o suposto tamanho da chave. Para descobrir se esse intervalo é o correto, precisamos fazer um cálculo nesse conjunto de letras no intuito de descobrir se esse conjunto de letras é o mesmo caractere. Essa técnica é chamada de Índice de Coincidência. Sua fórmula é a seguinte:

$$IC = \frac{F_i * (F_i - 1)}{N * (N - 1)} \tag{3}$$

Sendo F_i a frequência de uma determinada letra do alfabeto no conjunto escolhido e N o tamanho do conjunto.Após ter calculado em todas as letras do conjunto, é tirado a média do valor em relação ao intervalo. E por fim é feito o mesmo processo em todos os intervalos possíveis. O maior valor encontrado é o provável tamanho da chave.

Uma vez descoberto o tamanho da chave, é realizado um procedimento para descobrir os caracteres que a compõem. Considerando que o valor da chave é N, analisamos todas as fatias do texto de tamanho N que foram geradas no processo de descobrir o tamanho. Uma série de deslocamentos é realizado com cada uma das fatias e a cada deslocamento, é verificada a frequência de cada letra do alfabeto na nova fatia. Para descobrir cada caractere da chave, é aplicado o método estatístico Qui-Quadrado.

$$x^2 = \frac{\sum (f_i - F_i)^2}{F_i} \tag{4}$$

Onde f_i é a média da frequência de uma determinada letra na fatia e F_i é a frequência da letra no alfabeto. O programa apresenta um vetor de frequências das letras em inglês e um vetor para português. Ao final temos o Qui-Quadrado para cada letra, e a letra cujo resultado possui o menor valor será considerada a letra da determinada posição da chave. Esse processo é realizado N vezes.

A chave então, é usada para decifrar a mensagem, mostrando-a na tela.

3 Resultados

A partir da análise feita usando testes automatizados com a ferramenta de testes Catch (`https://github.com/catchorg/Catch2`) foi possível garantir o comportamento esperado para mensagens médias e grandes e chaves médias a curtas. Contudo, averiguou-se que ocorreu queda no acerto da análise da mensagem encriptada conforme a mensagem diminuiu e a chave aumentou de tamanho. Além disso, foram observadas repetições no

padrão de palavras durante a análise. Tal repetição pode ser removida ao analisar a senha gerada ao final do processo, caso haja subconjuntos, é possível simplificar a palavra.

4 Conclusão

A partir dos conhecimentos adquiridos e discutidos sobre a técnica de cifra de Vigenère, tornou-se possível fazer uma implementação adequada do algoritmo, isto é, foi garantido um comportamento dentro do padrão. O projeto foi dividido em duas partes, a primeira para cifrar e decifrar uma mensagem usando uma senha, e a segunda parte consistiu na análise da mensagem secreta para atacar a cifra e descobrir a senha e, com isso, decifrar a mensage. Desta forma, pode-se afirmar que a implementação satisfez o objetivo do projeto.

5 Referências

1. https://pages.mtu.edu/ shene/NSF-4/Tutorial/VIG/Vig-Recover.html
2. https://pages.mtu.edu/ shene/NSF-4/Tutorial/VIG/Vig-IOC-Len.html
3. https://pt.wikipedia.org/wiki/Cifra_de_Vigen%C3%A8re