UNIVERSIDADE DO ESTADO DE MATO GROSSO
CAMPUS DE SINOP
FACULDADE DE CIÊNCIAS EXATAS E TECNOLÓGICAS
MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL PROFMAT

APOSTILA DE PROPOSTAS DE ATIVIDADES PARA AULAS DE MATEMÁTICA NO ENSINO FUNDAMENTAL II E NO ENSINO MÉDIO BASEADAS EM TÉCNICAS DE CRIPTOGRAFIA

ISAC ROSA RODRIGUES

Produto final vinculado à dissertação de mestrado intitulada "**USO DE FERRAMENTAS DE CRIPTOGRAFIA NO ENSINO DE MATEMÁTICA NO ENSINO FUNDAMENTAL II E NO ENSINO MÉDIO: PROPOSTAS DE ATIVIDADES**" apresentada ao Programa de Mestrado profissional em Matemática em Rede Nacional – PROFMAT, da Universidade do Estado de Mato Grosso – UNEMAT, como requisito parcial para obtenção do grau de Mestre em Matemática, orientada pelo Prof. Dr. Raul Abreu de Assis e co-orientada pela profa. Dra. Luciana Mafalda Elias de Assis.

UNEMAT
Sinop-MT – 2024

Proposta de Atividades

Título	Apostila de Propostas de Atividades para Aulas de Matemática no Ensino Fundamental II e no Ensino Médio Baseadas em Técnicas de Criptografia
Nível de Ensino	Fundamental II e Médio.
Tipo de atividade	Expositiva e prática podendo ser em grupo ou individual dependendo dos materiais disponíveis e número de participantes.
Duração	5 aulas de 50 minutos para cada atividade proposta.
Objetivos	● Tornar a prática docente mais motivadora e rica de elementos desafiadores; ● Tornar a aula mais dinâmica e participativa, fazendo com que o aluno seja protagonista de seu aprendizado; ● Oportunizar ao estudante expor suas ideias e opiniões sobre a forma como o conteúdo de matrizes foi apresentado; ● Despertar o senso investigativo, bem como a curiosidade naquilo que se está aprendendo.
Conteúdos abordados	● Divisão euclidiana; aritmética modular; progressão aritmética. ● Matrizes e operações entre matrizes; ● Análise combinatória ● Estatística
Material utilizado	● Notebook para o professor; ● Chromebook para o aluno (verificar a disponibilidade na escola) e/ou outro computador que o aluno possa desenvolver as atividades ou celulares de uso pessoal dos alunos. ● Papel cartão para elaboração dos Citales.

SUMÁRIO

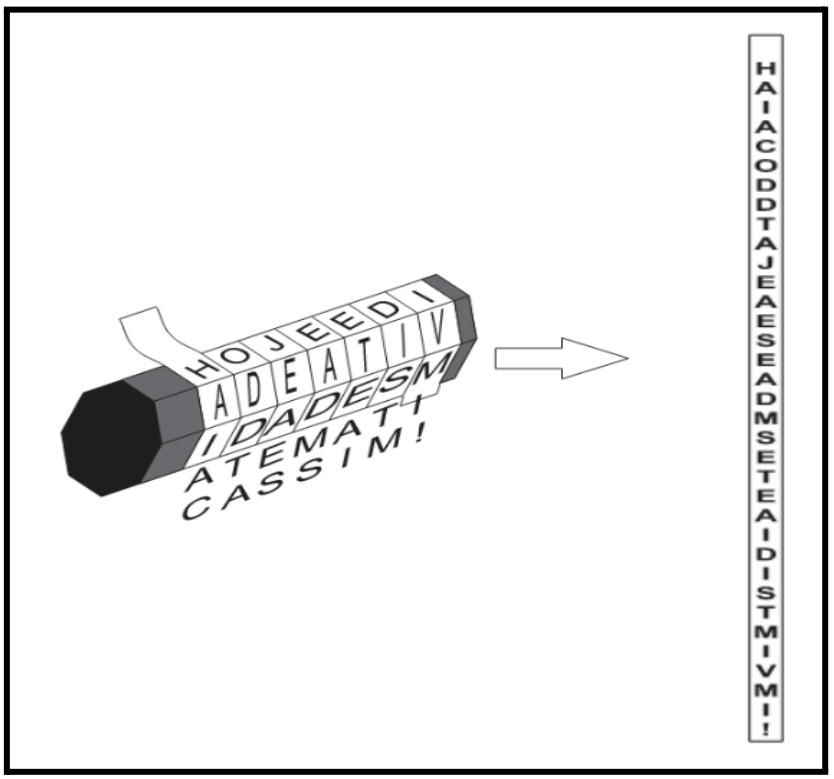
1. A CIFRAGEM POR CITALE (OU CÍTALA ESPARTANA)

Conteúdos relacionados: Divisão euclidiana; aritmética modular; progressão aritmética.

O citale é um aparelho simples para cifragem de mensagens que constitui em um prisma comprido ao redor do qual uma tira de couro, papel, pergaminho ou qualquer material base para escrita é enrolada e o texto a ser cifrado é escrito na tira no sentido longitudinal do citale, em seguida a tira é desenrolada e a sequência de caracteres obtidos ao longo da tira é o texto cifrado que pode ser enviado através do envio da própria tira ou ser transcrito em outro local. Normalmente, omitimos os espaços, acentuação e pontuação para que haja mais eficiência na cifragem. Observe o exemplo:

TEXTO ORIGINAL: Hoje é dia de atividades matemáticas? Sim!

Figura 1: Cifra do citale



Fonte: Elaborado pelo autor

RETIRANDO-SE ESPAÇOS, ACENTUAÇÃO E PONTUAÇÃO:
HOJEEDIADEATIVIDADESMATEMATICASSIM

Para simplificar o processo e a realização das atividades, os textos usados devem ser escritos de modo que ocupem por inteiro o último lado (chamaremos simplesmente de lado as faces retangulares do citale) utilizado do citale, ou seja, ou o número de caracteres do texto deve ser múltiplo do número de voltas dadas com a tira ao redor do citale, ou, ao final do texto, caso o último lado ocupado não tenha sido preenchido por completo, devemos inserir caracteres aleatórios até que todo espaço do último lado seja ocupado. As atividades propostas foram elaboradas considerando que este critério de simplificação tenha sido seguido.

TEXTO CIFRADO: HAIACODDTAJEAESEADMSETEAIDISTMIVMI

Outro detalhe é que não é obrigatório usar todos os lados do citale, no exemplo usamos um citale de 8 lados mas o texto só ocupou 5 desses lados, isso faz com que hajam espaços em branco na tira onde escrevemos o texto, o que dá indício do número de lados do citale utilizado, facilitando uma decifragem, para evitar tal fraqueza na cifragem é possível transcrever o texto cifrado para outro lugar eliminando-se os espaços em branco, uma outra abordagem nos permite manter o texto na tira, para isso basta preencher os lados não ocupados pelo texto com caracteres aleatórios que serão facilmente reconhecidas como um artifício de simples preenchimento de espaço por quem decifrar o texto.

O método básico de decifragem do citale é manter o texto na tira e enrolar esta em um citale idêntico ao que foi usado para cifrar o texto. Porém existem outros métodos, usados quando o texto é transcrito para outro lugar ou quando não se conhece as características do citale usado para cifrar o texto.

Se observarmos com atenção a estrutura do citale, podemos notar que as letras consecutivas no texto original estão sempre a uma mesma distância umas das outras no texto cifrado, dessa forma, se enumerarmos a posição de cada letra no texto cifrado, a divisão do número relativo a posição de letras consecutivas no texto original pelo número de lados do citale (seja tanto o número total de lados como o número de lados efetivamente utilizados), deixará um mesmo resto e assim, o texto poderá ser decifrado desde que identifiquemos o número de lados em questão.

Isso pode ser feito dividindo o número relativo à posição das letras no texto cifrado por algum número inteiro, agrupando aquelas cujos restos forem iguais e analisando se a sequência de letras obtidas faz algum sentido, caso não faça, repetimos o processo com outro número como divisor e vamos repetindo até encontrarmos o divisor adequado. Seguindo com nosso exemplo, enumeramos os caracteres do texto cifrado:

Tabela 1: Posição das letras no texto cifrado

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

H	A	I	A	C	O	D	D	T	A	J	E	A	E	S	E	A	D

19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	
M	S	E	T	E	A	I	D	I	S	T	M	I	V	M	I	

Fonte: Elaborado pelo autor

O menor número possível de lados para um prisma é 3, portanto esse será o primeiro número pelo qual dividiremos os números relativos às posições das letras no texto cifrado.

Nas tabelas a seguir escrevemos cada número relativo a posição de uma letra no texto cifrado na forma do teorema da divisão euclidiana, “r” indica o resto da divisão e “q” o quociente, iniciamos na tabela onde usamos o número 3 como divisor:

Tabela 2: Caracteres do texto e a divisão euclidiana da posição relativa pelo divisor 3

resto 1	resto 2	resto 0
H:1=0*3+1; r=1, q=0	A:1=0*3+2; r=2, q=0	I:1=1*3+0; r=0, q=1
A:4=1*3+1 r=1, q=1	C:5=1*3+2 r=2, q=1	O:6=2*3+0 r=0, q=2
D:7=2*3+1 r=1, q=2	D:8=2*3+2 r=2, q=2	T:9=3*3+0 r=0, q=3
A:10=3*3+1 r=1, q=3	J:11=3*3+2 r=2, q=3	E:12=4*3+0 r=0, q=4
A:13=4*3+1 r=1, q=4	E:14=4*3+2 r=2, q=4	S:15=5*3+0 r=0, q=5
E:16=5*3+1 r=1, q=5	A:17=5*3+2 r=2, q=5	D:18=6*3+0 r=0, q=6
M:19=6*3+1 r=1, q=6	S:20=6*3+2 r=2, q=6	E:21=7*3+0 r=0, q=7
T:22=7*3+1 r=1, q=7	E:23=7*3+2 r=2, q=7	A:24=8*3+0 r=0, q=8
I:25=8*3+1 r=1, q=8	D:26=8*3+2 r=2, q=8	I:27=9*3+0 r=0, q=9
S:28=9*3+1 r=1, q=9	T:29=9*3+2 r=2, q=9	M:30=10*3+0 r=0, q=10
I:31=10*3+1 r=1, q=10	V:32=10*3+2 r=2, q=10	M:33=11*3+0 r=0, q=11
I:34=11*3+1 r=1, q=11		

Fonte: elaborado pelo autor

Como teste, vamos agrupar algumas letras correspondentes ao resto 1:

1,4,7,10,13,16 => HADAAE

A sequência obtida não parece fazer muito sentido, portanto vamos testar usando o número 4 como divisor:

Tabela 3: Caracteres do texto e a divisão euclidiana da posição relativa pelo divisor 4

resto 1	resto 2	resto 3	resto 0
H:1=0*4+1 r=1, q=0	A:2=0*4+2 r=2, q=0	I:3=0*4+3 r=3, q=0	A:4=1*4+0 r=0, q=1
C:5=1*4+1 r=1, q=1	O:6=1*4+2 r=2, q=1	D:7=1*4+3 r=3, q=1	D:8=2*4+0 r=0, q=2
T:9=2*4+1 r=1, q=2	A:10=2*4+2 r=2, q=2	J:11=2*4+3 r=3, q=2	E:12=3*4+0 r=0, q=3
A:13=3*4+1 r=1, q=3	E:14=3*4+2 r=2, q=3	S:15=3*4+3 r=3, q=3	E:16=4*4+0 r=0, q=4
A:17=4*4+1 r=1, q=4	D:18=4*4+2 r=2, q=4	M:19=4*4+3 r=3, q=4	S:20=5*4+0 r=0, q=5
E:21=5*4+1 r=1, q=5	T:22=5*4+2 r=2, q=5	E:23=5*4+3 r=3, q=5	A:24=6*4+0 r=0, q=6
I:25=6*4+1 r=1, q=6	D:26=6*4+2 r=2, q=6	I:27=6*4+3 r=3, q=6	S:28=7*4+0 r=0, q=7
T:29=7*4+1 r=1, q=7	M:30=7*4+2 r=2, q=7	I:31=7*4+3 r=3, q=7	V:32=8*4+0 r=0, q=8
M:33=8*4+1 r=1, q=8	I:34=8*4+2 r=2, q=8		

Fonte: Elaborado pelo autor

Agrupando algumas letras correspondentes ao resto 1:

1,5,9,13,17,21 => HCTAAE

Novamente não há sentido aparente, vamos usar o 5 como divisor em um novo teste:

Tabela 4: Caracteres do texto e a divisão euclidiana da posição relativa pelo divisor 5

resto 1	resto 2	resto 3	resto 4	resto 0
H:1=0*5+1 r=1, q=0	A:2=0*5+2 r=2, q=0	I:3=0*5+3 r=3, q=0	A:4=0*5+4 r=4, q=0	C:5=1*5+0 r=0, q=1
O:6=1*5+1 r=1, q=1	D:7=1*5+2 r=2, q=1	D:8=1*5+3 r=3, q=1	T:9=1*5+4 r=4, q=1	A:10=2*5+0 r=0, q=2
J:11=2*5+1 r=1, q=2	E:12=2*5+2 r=2, q=2	A:13=2*5+3 r=3, q=2	E:14=2*5+4 r=4, q=2	S:15=3*5+0 r=0, q=3
E:16=3*5+1 r=1, q=3	A:17=3*5+2 r=2, q=3	D:18=3*5+3 r=3, q=3	M:19=3*5+4 r=4, q=3	S:20=4*5+0 r=0, q=4
E:21=4*5+1 r=1, q=4	T:22=4*5+2 r=2, q=4	E:23=4*5+3 r=3, q=4	A:24=4*5+4 r=4, q=4	I:25=5*5+0 r=0, q=5
D:26=5*5+1 r=1, q=5	I:27=5*5+2 r=2, q=5	S:28=5*5+3 r=3, q=5	T:29=5*5+4 r=4, q=5	M:30=6*5+0 r=0, q=6
I:31=6*5+1 r=1, q=6	V:32=6*5+2 r=2, q=6	M:33=6*5+3 r=3, q=6	I:34=6*5+4 r=4, q=6	

Fonte: Elaborado pelo autor

Agrupando algumas letras correspondentes ao resto 1:

1,6,11,16,21,26 => HOJEED

Agora, aparentemente temos uma palavra do texto original, para verificar se estamos no caminho correto ou se a palavra apareceu ao acaso, vamos agrupar em sequência as letras correspondentes aos restos 1,2,3,4 e 0:

Resto 1 => 1,6,11,16,21,26,31 => HOJEEDI
Resto 2 => 2,7,12,17,22,27,32 => ADEATIV
Resto 3 => 3,8,13,18,23,28,33 => IDADESM
Resto 4 => 4,9,14,19,24,29,34 => ATEMATI
Resto 0 => 5,10,15,20,25,30,35 => CASSIM

HOJEEDIADEATIVIDADESMATEMATICASSIM => Hoje é dia de atividades matemáticas? Sim!

É interessante notar que, se ao realizarmos as divisões, formos agrupando os resultados em células de tabelas cujo número de colunas for igual ao divisor usado,

quando usarmos o divisor correto o texto já estará decifrado se lermos as letras coluna por coluna como ocorreu na tabela 3, essa é uma característica que pode reduzir o trabalho da decifragem mas que pode não ser mencionada a fim de que os alunos tenham a oportunidade de notar tal detalhe. Seguindo essa abordagem, no momento em que os alunos realizarem as atividades propostas é aconselhável que os mesmos realizem cada divisão presente nas tabelas anteriores anotando os restos para que tenham a oportunidade de reconhecer os padrões que surgem. Após isso, é conveniente anotar em cada teste com determinado divisor apenas os números que já sabemos que deixarão o mesmo resto sem fazer as divisões uma a uma. Por exemplo: ao dividir os números naturais por 5 como foi feito na última tabela, já sabemos (e espera-se que os alunos percebam) que os números que deixarão resto 1 são os termos de uma *PA* de termo inicial 1 e razão 5: 1,6,11,16,21,26,31… os que deixarão resto 2 são os termos de uma PA de termo inicial 2 e razão 5: 2,7,12,17,22,27,32… e assim por diante. Isso se justifica justamente pelo caráter de repetição periódica da posição de letras consecutivas do texto original no texto cifrado. Se uma letra qualquer tem a posição x_1 no texto original e posição a_1 no texto cifrado e o citale utilizado para cifragem tem r lados, a próxima letra do texto original x_2 estará r posições à frente de x_1 no texto cifrado, enquanto x_3 estará r posições à frente de x_2, ou seja, se chamarmos de $P(x)$ a posição no texto cifrado de um caractere de posição genérica x no texto original, temos:

$$P(x_1) = a_1$$
$$P(x_2) = P(x_1) + r = a_1 + r$$
$$P(x_3) = P(x_2) + r = a_1 + 2r$$
$$\vdots$$
$$\vdots$$
$$P(x_n) = a_1 + (n-1)r$$

1. 1 ATIVIDADES PROPOSTAS:

01) A figura “citale de 7 lados” presente ao final desta apostila é a planificação de um citale de 7 lados, recorte-a nas linhas contínuas, faça as dobras nas linhas pontilhadas e cole as abas para montar um citale (aconselhamos colar a folha em uma cartolina ou similar antes de recortar para dar maior resistência). Use-o para decifrar o texto a seguir:

O N E P E S R S E E O U E A O T D D V M F L A E E A A A C S P A S I Z O G E M
C S E M I N A O N R T R D D M A - O A E U O D G D N N R S A A O D T E E N L S
O E C E O I A E S E L U L Q M D R E N E U T E U N I U E O L M A V G L R E C O
E A E N A A T R L S O I C I S I P D N H V O L L E D O E P E A L A D S A I

02) As sequências de letras a seguir são o resultado de uma cifragem por citale de uma frase. Em cada caso determine o número de lados do citale utilizado e a frase original:

a) B D A D V E O I T O O S M A O S C .

b) M A U L A T M G S O E U Q D C E U O O M E O N L M O V E M U E V E R
N A D O C S E Q E S R U R E A E Q E A E U M O N E B M T E O E R R R N
E A A O G P A S U R T U E O E M I V O A A A Q V Q D U E U E E Z E A E T
M M U E C I N R O Z A D N A O E S D T V E E I O G S N L U E H T I A A

03) Usando o citale construído na atividade (01) cifre algum texto não muito curto (em torno de 400 caracteres sem contar os espaços) e entregue a um colega enquanto ele lhe entrega um texto dele também cifrado. Determine um método de decifrar o texto rapidamente sem usar o citale e sem realizar todas as divisões realizadas no exemplo.

04) No texto foi recomendado que o último lado do citale ocupado pelo texto a ser cifrado fosse ocupado por completo, inclusive apesar da prática de se retirar a pontuação, um ponto de exclamação foi mantido no nosso exemplo devido a essa recomendação. Pense na estrutura do citale e do texto cifrado na tira que foi utilizada e indique um método de se contornar esse problema de modo que tal recomendação não precise ser seguida.

2. CIFRAGEM POR OPERAÇÕES COM MATRIZES

Conteúdos relacionados: Matrizes, operações com matrizes, conversão de dados alfabéticos para dados numéricos.

Nesta seção, apresentaremos como realizar uma cifragem utilizando operações entre matrizes. Iniciamos, verificando que uma mensagem qualquer pode ser organizada em tabelas. As letras e outros caracteres que formam um texto qualquer normalmente são dispostos em linhas, com o início do texto na esquerda da primeira linha superior e sua continuidade se dando para a direita e para as linhas que se seguem abaixo sendo este o sentido convencional de escrita e leitura (da esquerda para a direita e de cima para baixo) no Português e em diversos outros idiomas. Dessa forma é possível alinhar verticalmente os caracteres que compõem o texto em colunas de modo que cada linha tenha a mesma quantidade de caracteres (com a possível exceção da última). Observe um exemplo:

Frase na forma original:
O FUTURO NÃO É MAIS COMO ERA ANTIGAMENTE

Frase após o alinhamento (aqui consideramos o espaço entre palavras como um caractere representado por "_"):

Tabela 5: Caracteres da frase distribuídos em uma tabela 2x20

O	_	F	U	T	U	R	O	_	N	Ã	O	_	É	_	M	A	I	S	_
C	O	M	O	_	E	R	A	_	A	N	T	I	G	A	M	E	N	T	E

Fonte: Elaborado pelo autor

Podemos dizer que os elementos da frase formam uma tabela de 2 linhas e 20 colunas, se não nos importarmos em deixar partes de uma mesma palavra em linhas distintas podemos montar tabelas de diferentes números de linhas e colunas:

Tabela 6: Caracteres da frase distribuídos em uma tabela 4x10

O	_	F	U	T	U	R	O	_	N
Ã	O	_	É	_	M	A	I	S	_
C	O	M	O	_	E	R	A	_	A
N	T	I	G	A	M	E	N	T	E

Fonte: Elaborado pelo autor

Se associarmos um número distinto a cada caractere distinto usado na escrita, podemos converter um texto em uma sequência de números e, em seguida, a sequência de números em uma tabela, obtendo assim uma matriz de entradas numéricas. Se, por exemplo, definirmos as seguintes correspondências entre letras e números:

Tabela 7: Correspondência entre caracteres textuais e numéricos

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T

21	22	23	24	25	26	27
U	V	W	X	Y	Z	_

Fonte: Elaborado pelo autor

A Tabela 6, pode ser convertida na seguinte matriz:

$$M = \begin{bmatrix} 15 & 27 & 6 & 21 & 20 & 21 & 18 & 15 & 27 & 14 \\ 1 & 15 & 27 & 5 & 27 & 13 & 1 & 9 & 19 & 27 \\ 3 & 15 & 13 & 15 & 27 & 5 & 18 & 1 & 27 & 1 \\ 14 & 20 & 9 & 7 & 1 & 13 & 5 & 14 & 20 & 5 \end{bmatrix}$$

Tendo convertido nosso texto em uma matriz, é possível cifrá-lo usando alguma operação que receba cada número da matriz e retorne um novo número. Evidentemente isso só fará sentido se usarmos alguma operação que possa ser "desfeita" para que o texto possa ser decifrado posteriormente. Multiplicar a matriz por algum escalar seria uma possibilidade, porém resultaria em uma cifragem fraca, assim, vamos optar pela multiplicação da nossa matriz por uma outra. A operação inversa (que desfaz) a multiplicação de uma matriz qualquer A por uma outra matriz B é a multiplicação da matriz resultante pela inversa de *B*, denotada como B^{-1} (BOLDRINI, 1986). Portanto, se encontrarmos uma matriz que possa ser multiplicada pela matriz do nosso exemplo (a multiplicação de matrizes somente está definida para casos específicos) e que seja invertível (nem toda matriz possui matriz inversa) podemos usar a multiplicação de matrizes para cifrar nosso texto.

Tendo definido nossa técnica de cifragem, as propriedades das matrizes nos impõe duas limitações: primeiramente, como a multiplicação de matrizes não é comutativa a ordem em que realizamos essa operação é fundamental para o sucesso da técnica, devemos portanto multiplicar nossa matriz por uma matriz inversível (que daqui em diante será chamada de chave) para cifrar o texto, e para decifrar devemos multiplicar a matriz obtida na cifragem pela inversa da chave exatamente nessa ordem para cifragem e decifragem. A segunda limitação resulta do fato de que só possível multiplicar matrizes quando uma delas (dita a primeira) tem o número de colunas igual ao número de linhas da outra (dita a segunda), portanto devido à ordem da multiplicação na cifragem nossa matriz deve ter o mesmo número de colunas que o número de linhas da chave e portanto a conversão do texto em matriz e a criação da chave devem resultar em matrizes que obedeçam esse quesito.

Ao exemplificar a conversão de texto em números e a criação de matrizes com esses dados obtivemos a seguinte matriz 4x10. Para cifrar o texto nesse caso, seria necessário uma chave com 10 linhas, como apenas matrizes quadradas possuem inversas (mas não necessariamente todas) teríamos portanto uma chave de ordem 10. Como o processo de multiplicação de matrizes aumenta consideravelmente o número de passos conforme o número de colunas da primeira matriz (e consequentemente o número de linhas da segunda) aumenta, é conveniente distribuir nossos dados em uma matriz com menor número de colunas, o que também nos facilitará a criação da chave. Para nosso exemplo vamos distribuir nossos dados em uma matriz 10x4 (para aplicar aos alunos é aconselhável usar uma matriz com um número ainda menor de colunas, para o caso do nosso texto poderia ser utilizada uma matriz 20x2 por exemplo):

$$\begin{bmatrix} 15 & 27 & 6 & 21 \\ 20 & 21 & 18 & 15 \\ 27 & 14 & 1 & 15 \\ 27 & 5 & 27 & 13 \\ 1 & 9 & 19 & 27 \\ 3 & 15 & 13 & 15 \\ 27 & 5 & 18 & 1 \\ 27 & 1 & 14 & 20 \\ 9 & 7 & 1 & 13 \\ 5 & 14 & 20 & 5 \end{bmatrix}$$

Agora precisamos obter como chave uma matriz de ordem 4 invertível, com algumas contas é possível verificar que a seguinte matriz atende à nossa necessidade:

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 1 \end{bmatrix}$$

A multiplicação de nossa matriz pela chave resulta:

$$\begin{bmatrix} 15 & 27 & 6 & 21 \\ 20 & 21 & 18 & 15 \\ 27 & 14 & 1 & 15 \\ 27 & 5 & 27 & 13 \\ 1 & 9 & 19 & 27 \\ 3 & 15 & 13 & 15 \\ 27 & 5 & 18 & 1 \\ 27 & 1 & 14 & 20 \\ 9 & 7 & 1 & 13 \\ 5 & 14 & 20 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 42 & 60 & 75 & 48 \\ 53 & 72 & 69 & 71 \\ 43 & 31 & 45 & 44 \\ 67 & 72 & 58 & 94 \\ 47 & 74 & 82 & 66 \\ 31 & 56 & 58 & 44 \\ 46 & 42 & 25 & 64 \\ 61 & 49 & 55 & 75 \\ 23 & 22 & 34 & 24 \\ 30 & 59 & 44 & 50 \end{bmatrix}$$

A sequência numérica obtida com a matriz resultante é a mensagem cifrada que pode ser enviada com relativa segurança a um determinado destinatário que conheça a matriz chave e portanto seja capaz de usar sua inversa para decifrar a mensagem através da obtenção da matriz original.

Um detalhe interessante é que muitos dos elementos da matriz obtida na cifragem são números que não aparecem na tabela de correspondência entre caracteres e números. Portanto, alguém que interceptasse a mensagem, mesmo conhecendo a correspondência entre letras e números sequer seria capaz de atribuir letras aos números que compõem a mensagem cifrada, impossibilitando o mesmo de aplicar alguma técnica de decifragem baseada na frequência dos caracteres.

Caso seja de interesse do professor, os cálculos das operações matriciais podem ser realizados pelos alunos com o auxílio de *softwares* específicos para tal. Uma opção indicada é a calculadora online e gratuita de matrizes disponível em: https://matrixcalc.org/, outra ferramenta que pode ser utilizada é o editor de planilhas Microsoft Excel.

2. 1 ATIVIDADES PROPOSTAS:

01) Durante a explicação do método foi afirmado, sem demonstração, que a chave proposta atenderia às nossas necessidades (ser invertível) e que era fácil verificar isso com algumas contas. Faça tal verificação.

02) Encontre a matriz inversa da matriz chave utilizada no exemplo do texto e multiplique a matriz resultante pela matriz que você encontrou. Confira se o resultado é a matriz original que havíamos multiplicado pela matriz chave.

03) Considerando que, após a definição de uma matriz chave, haja a intenção de utilizá-la para cifrar outros textos, a matriz formada pelos dados numéricos do novo texto a ser cifrado deve ter o número de colunas igual à ordem da matriz chave. Entretanto, é possível que o número de caracteres do novo texto não seja múltiplo da ordem da matriz chave. Como isso interfere no processo de cifragem?

04) Encontre métodos para contornar a situação descrita acima.

05) A sequência numérica a seguir resultou de um texto cifrado pelo método que vimos nos exemplos, utilizando a mesma tabela de associação entre caracteres e números e a mesma chave. Use a matriz inversa encontrada na atividade (02) e decifre o texto:

56	58	70	71	42	41	35	56	22	64	53	38	45
59	36	72	55	86	72	82	50	73	68	70	28	44
37	36	70	95	97	95	42	41	35	56	42	41	35
56	42	41	35	56	22	64	53	38	45	59	36	72
64	95	90	91	43	71	75	58	55	60	34	82	17
36	24	30	59	41	63	64	39	44	55	43	45	23
39	46	38	34	27	50	57	82	86	74	36	80	67
58	62	58	72	75	59	58	72	72	34	61	64	49

06) Encontre uma matriz de ordem 3 que atenda às propriedades necessárias para que possa ser usada como chave para cifragem. Forneça essa matriz a algum colega para que ele a utilize para cifrar um texto qualquer que você não conheça, receba a sequência numérica obtida pelo seu colega com o uso da sua chave e decifre o texto:

07) Analise com atenção todos os passos da cifragem por multiplicação de matrizes. Procure determinar ao menos um meio de tornar a cifragem mais forte (mais difícil de ser decifrada):

3. CIFRAGEM POR TRANSPOSIÇÃO E CIFRAGEM POR SUBSTITUIÇÃO

Conteúdos relacionados: Princípio fundamental da contagem; análise combinatória.

A cifragem por citale e alguns outros métodos específicos de cifragem pertencem a um grupo geral de cifras chamadas cifras de transposição. Esse tipo é caracterizado por uma troca da posição dos caracteres do texto original no texto cifrado, ou seja, cada letra de uma frase por exemplo, estará em uma posição diferente no texto cifrado daquela em que se encontrava no texto original porém nessa nova posição ela mantém sua identidade, ou seja é a mesma letra. Cifras desse grupo têm características de tratamento de caracteres opostas às do grupo chamado cifras de substituição, nas quais os caracteres mantém sua posição inalterada durante a cifragem, porém são trocados por outros caracteres de acordo com alguma pré-determinação. Um exemplo desse segundo grupo é cifragem por multiplicação de matrizes onde inicialmente cada letra do texto original é trocada por um número que mantém nesse estágio da cifragem a posição da letra à qual ele foi associado, depois cada número é trocado por outro obtido da multiplicação de matrizes porém a posição desse outro número no texto cifrado é a mesma do número anterior e consequentemente da letra do texto original.

❖ Cifragem por transposição:

Nas cifras de transposição temos efetivamente como resultado um anagrama do texto original, anagrama é um termo normalmente usado para se referir à palavras compostas exatamente pelas mesmas letras cada qual na mesma quantidade mas em ordem diferentes. Por exemplo, a palavra bola é um anagrama da palavra loba e a palavra socar é um anagrama da palavra rocas. Qualquer combinação nova das letras originais é um anagrama sem que seja necessário que a nova combinação tenha algum sentido textual, assim ablo e ocsar também são anagramas de loba e rocas respectivamente. É possível cifrar qualquer texto apenas gerando um anagrama qualquer do mesmo, e isso possibilita um número de diferentes cifragens que aumenta rapidamente conforme se aumenta o número de caracteres do texto.Observe:

- uma palavra de 2 letras só possui 2 anagramas:
  ai/ia;
  ou/uo

- uma palavra de 3 letras possui 6 anagramas:
  foi/fio/ofi/oif/ifo/iof
  olá/oal/loa/lao/alo/aol

- uma palavra de 4 letras possui 24 anagramas:
  beco/beoc/boce/boec/bcoe/bceo/ebco/eboc/ecbo/ecob/eocb/eobc/
  cebo/ceob/cbeo/cboe/cobe/coeb/oceb/ocbe/oecb/oebc/obce/obec

É notável que a cada letra que adicionamos à palavra, o número de anagramas possíveis é o número de anagramas para a quantidade anterior de letras vezes o número de letras atual. Isso se deve ao fato de que o números de anagramas é o número de permutações simples possíveis para dada quantidade de letras, logo para uma palavra de *n* letras temos *n!* anagramas possíveis. Observe, porém, que entre todos os anagramas possíveis podemos ter alguns repetidos caso hajam letras repetidas na palavra, o que certamente acontecerá caso queiramos formar anagramas de uma frase ou um texto, observe um exemplo com a palavra uau, onde enumeramos cada letra com um índice para facilitar a distinção entre o primeiro “u” e o segundo:

$u_1a_2u_3/u_1u_3a_2/a_2u_3u_1/a_2u_1u_3/u_3a_2u_1/u_3u_1a_2$

Temos, como era de se esperar, 6 anagramas para uma palavra de 3 letras porém apenas 3 anagramas distintos. Como, para uma palavra, frase ou texto qualquer, temos um número específico de possíveis anagramas que corresponde ao número de possíveis cifragens por transposição. Entretanto é evidente que a elaboração de dois textos cifrados que sejam iguais, a partir de um mesmo texto original não tem sentido lógico, devemos distinguir o número de cifras de transposição possíveis dos números de cifras de transposição que resultam em textos distintos.

É evidente que, apesar do número de cifragens por transposição ser altíssimo para textos que possuem um número alto de caracteres, na prática, a cifragem não pode ser feita de qualquer modo aleatório pois isso impossibilitaria a decifragem posterior. Assim, devemos cifrar um texto por transposição através de algum método bem definido de cifragem, como o citale. Nesses casos, a cifra definida para transpor os caracteres pode ser representada matematicamente por uma função bijetiva sobre o conjunto das posições do caractere no texto, enquanto que o processo de decifragem é a correspondente função inversa.

Algumas cifras simples de transposição:

- Cifra das colunas:

Nessa cifra os caracteres do texto a ser cifrado são escritos em colunas formando uma grade com um número de linhas pré definido, quando o número de linhas é atingido em uma coluna a escrita continua na coluna seguinte, ao final, caso a última coluna não seja inteiramente preenchida é possível inserir caracteres aleatórios para completar a grade. A mensagem cifrada é formada pelos caracteres na ordem em que ficam dispostos nas linhas da grade. Por exemplo, para uma grade com 4 linhas a frase “Não revele essa mensagem” é escrita assim na grade:

N	E	E	A	S	M
A	V	E	M	A	R
O	E	S	E	G	T
R	L	S	N	E	I

E portanto, cifrada assim:

N E E A S M A V E M A R O E S E G T R L S N E I

Para decifrar a mensagem basta que o destinatário conheça o número fixo de linhas utilizadas na grade e divida o número de caracteres da mensagem por esse número obtendo assim o número de colunas nas quais ele deve distribuir os caracteres da mensagem cifrada escrevendo-os linha por linha e finalmente lendo coluna por coluna. Essa cifra é muito semelhante a do citale, porém sem a necessidade de um objeto físico para realizar a cifragem. Especificamente, se fizermos um processo análogo à cifra das colunas, porém distribuindo os caracteres em linhas numa grade com o número de colunas fixo obteremos exatamente uma cifra por citale.

- Cifra por transposição de colunas

Nesta cifra, uma palavra chave é escrita na primeira linha de uma grade e, em seguida, a mensagem a ser cifrada é escrita abaixo, linha por linha com um total de colunas igual ao número de letras da palavra chave. A mensagem cifrada é obtida através da transcrição das colunas por ordem alfabética das letras da palavra chave. Por exemplo, para cifrar a mensagem “alguns infinitos são maiores que outros” usando a palavra chave “cifra”, a grade é escrita como:

C	I	F	R	A
A	L	G	U	N
S	I	N	F	I
N	I	T	O	S
S	A	O	M	A
I	O	R	E	S
Q	U	E	O	U
T	R	O	S	A

Como o anagrama da palavra chave em que as letras estão em ordem alfabética é ACFIR, escrevendo as letras da coluna de cada uma dessas letras temos:

Coluna do A: N I S A S U A

Coluna do C: A S N S I Q T

Coluna do F: G N T O R E O

Coluna do I: L I I A O U R

Coluna do R: U F O M E O S

E a mensagem cifrada será:

NISASUAASNSIQTGNTO
REOLIIAOURUFOMEOS

Para decifrar a mensagem, basta que o destinatário conheça a palavra chave utilizada na grade e divida o número de caracteres da mensagem pelo número de letras da palavra chave. Dessa forma, ele obtém o número de linhas nas quais ele deve distribuir os caracteres da mensagem cifrada, escrevendo-os então, coluna por coluna abaixo de cada letra da palavra chave na ordem alfabética das mesmas e, finalmente, lendo linha por linha.

- Cifras das espirais

Este é um conjunto de cifras que são na verdade variações de uma mesma ideia: escreve-se os caracteres da mensagem a ser cifrada em uma grade retangular como nas cifras anteriores. Em seguida, obtêm-se o texto cifrado transcrevendo os caracteres na ordem em que aparecem segundo uma “espiral

retangular” construída sobre a grade. Por exemplo, para cifrar a frase “cifras das espirais” podemos escrever os caracteres da mensagem linha por linha numa grade retangular e transcrevê-los para a mensagem cifrada segundo uma espiral retangular iniciada no canto superior direito e direcionada para o centro da grade no sentido anti-horário, assim:

As inúmeras variações dessa ideia podem ser obtidas pelas combinações das diferentes opções da distribuição dos caracteres na grade com as diferentes opções de construção da espiral:

1. a mensagem pode ser escrita na grade linha por linha ou coluna por coluna.
2. a espiral pode ser construída em sentido horário ou anti-horário.
3. a espiral pode ser construída das bordas para o centro ou do centro para as bordas.
4. a espiral pode iniciar (quando construída da borda para o centro) ou terminar (quando construída do centro para a borda) em qualquer um dos quatro vértices da grade.

❖ Cifragem por substituição:

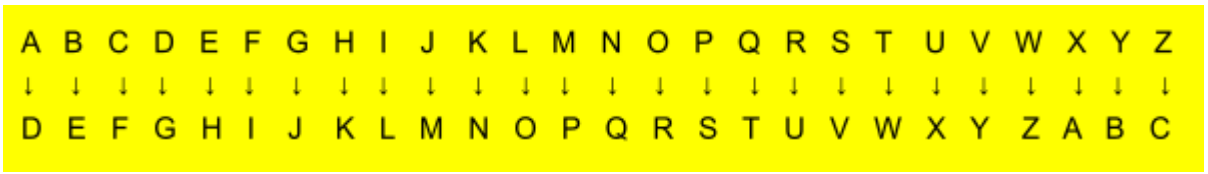
Quando se trata da cifragem por substituição o número possível de cifras deixa de depender do número de caracteres do texto e passa a depender do número de caracteres existentes que podem ser usados para substituir os originais. Em teoria, o número de possíveis cifragens de qualquer texto por substituição é infinito visto que existem infinitos caracteres que podem ser usados como substitutos dos originais (uma prova disso é a troca de letras por números, como existem infinitos números, qualquer sequência textual pode ser convertida em infinitas sequências numéricas). Evidentemente, assim como na cifragem por transposição, na cifragem por substituição é necessário que o destinatário conheça de antemão o método de substituição específico usado em uma mensagem para que possa decifrá-la.

Algumas cifras simples de substituição:

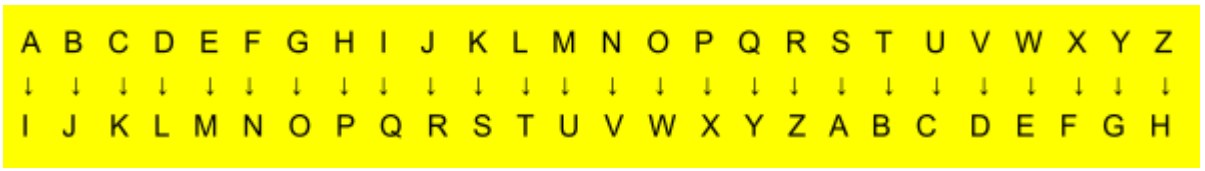
- Cifra de César

Na chamada cifra de César cada letra do texto original é trocada por uma letra que se encontra à uma certa quantidade de posições à frente ou antes dessa letra no alfabeto. Por exemplo, podemos cifrar uma frase trocando cada letra da mesma pela letra que se encontra 3 posições à frente no alfabeto, a letra A por exemplo seria trocada pela letra D, a letra M pela letra P e a letra X pela letra pela

letra A, de um modo geral para essa configuração associamos as letras do alfabeto convencional por letras do alfabeto com deslocamento de 3 posições:



Seguindo esta mesma ideia, podemos definir outras cifras distintas, obtidas dos possíveis deslocamentos do alfabeto. Usando um deslocamento de 8 posições por exemplo, podemos cifrar a frase “existem várias possibilidades” através da seguinte relação:



Obtendo:

M F Q A B M U D I Z Q I A X W A A Q J Q T Q L I L M A

Além dos diferentes deslocamentos do alfabeto da cifra em relação ao original, é possível definir alfabetos “fora de ordem” por meio de permutações das letras do alfabeto original e usar esses “rearranjos” como base para a substituição, isso gera um número altíssimo de cifras distintas.

- Cifra de deslocamento por palavra chave

Uma cifra de substituição bastante interessante por, tal qual a cifra de César, ser de fácil implementação é a cifra de deslocamento por palavra chave. Nela, uma palavra chave (ou frase) é pré definida e o alfabeto cifrado será iniciado pelas letras da palavra chave, desconsiderando letras repetidas (e espaços caso se use uma frase), o restante do alfabeto cifrado é composto simplesmente pelas letras que não constam na palavra chave, as quais podem ser colocadas em ordem, ou assim como no caso da cifra de César, podemos ainda rearranjar as letras restantes. Por exemplo, se usarmos a palavra chave “criptografia”, retirando as letras repetidas obtemos a sequência C R I P T O G A F que iniciará nosso alfabeto para a cifragem, se mantermos as letras restantes em ordem obtemos:

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

Um exemplo usando a mesma palavra chave, porém rearranjando a letras restante:

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

A frase “os resultados podem variar” cifrada por (I), fica:

K Q N T Q U E S C P K Q L K P T H V C N F C N

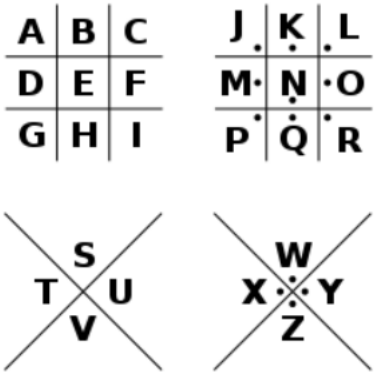
A mesma frase, cifrada por (II), fica:

Q W V T W Y L X C P Q W S Q P T M Z C V F C V

- Cifra do chiqueiro

A cifra do chiqueiro, também conhecida como cifra maçônica, por ter sido muito usada por maçons a partir do século XVII, substitui as letras do alfabeto por símbolos específicos através da associação em quatro grades geométricas como a seguinte:

Figura 02: Grades da cifra do chiqueiro



Fonte:https://pt.planetcalc.com/7842/ (acesso em 10/04/2024)

Assim como no caso da cifra de César, é possível obter um número elevado de distintas cifras do chiqueiro através da distribuição de um rearranjo do alfabeto

original nas grades. Usando a associação do exemplo anterior, a frase “use rearranjos” fica cifrada como:

Figura 03: Frase cifrada pela cifra do chiqueiro

Fonte: Elaborado pelo autor

- Código Morse

Apesar de não se tratar de uma cifra no sentido de um método criado para ocultar o significado de mensagens, o famoso código Morse se baseia na substituição de caracteres por símbolos específicos através de uma correspondência pré definida, o mesmo foi criado para o envio de mensagens por telégrafo e consiste em uma tabela de correspondência entre caracteres usuais e caracteres formados por específicas combinações de pontos e traços, como na figura abaixo:

Figura 04: Tabela de caracteres do código morse

A	•-	N	-•	0	-----
B	-•••	O	---	1	•----
C	-•-•	P	•--•	2	••---
D	-••	Q	--•-	3	•••--
E	•	R	•-•	4	••••-
F	••-•	S	•••	5	•••••
G	--•	T	-	6	-••••
H	••••	U	••-	7	--•••
I	••	V	•••-	8	---••
J	•---	W	•--	9	----•
K	-•-	X	-••-	.	•-•-•-
L	•-••	Y	-•--	,	--••--
M	--	Z	--••	?	••--••

Fonte: https://www.significados.com.br/codigo-morse/ (acesso em 10/04/2024)

Atualmente, com o advento da internet, a comunicação por código morse não apresenta segurança alguma, pois o mesmo é facilmente identificado devido à aparência dos símbolos e a partir da identificação é possível obter a tabela de associação rapidamente em uma pesquisa online. Porém, antes de tal possibilidade se difundir, o código Morse podia ser usado como meio de criptografia (fraca) caso fosse desejável esconder uma mensagem de alguém que sabia-se não ser conhecedor do código.

- Barreira do idioma

O caso do código Morse, em que um meio de comunicação escrito não foi desenvolvido para ocultar mensagens, mas em determinadas circunstâncias acaba por fazê-lo, pode ocorrer com as escritas de algumas civilizações antigas que utilizavam outros conjuntos de símbolos (alfabetos) para representar os fonemas de seus idiomas. É possível (e há casos) que no decorrer da história humana, alguns idiomas sejam extintos e a leitura de sua escrita deixe de ser possível de modo direto devido a inexistência de indivíduos “alfabetizados” em tal idioma e de textos claramente escritos no idioma em questão e em outro conhecido. Nessas situações os textos deixados por tais civilizações estão, de certa forma, criptografados para os leitores da atualidade e os métodos para decifrá-los são os mesmos geralmente usados para textos intencionalmente criptografados. Com isso é possível realizar atividades de treino de cifragem e decifragem, contextualizadas com um componente histórico e linguístico através da associação do nosso alfabeto com os diferentes alfabetos conhecidos atualmente.

3. 1 ATIVIDADES PROPOSTAS:

01) Já comentamos no texto que o número de possíveis cifras por transposição de um texto de n caracteres é n!, porém que o números de possíveis textos cifrados distintos será menor que n! para textos originais que apresentem caracteres repetidos (o que obrigatoriamente acontecerá quando *n* for maior que o número de caracteres existentes em determinado idioma), suponha que certa frase tenha 6 caracteres, porém 1 deles é repetido 3 vezes. Qual o número de possíveis cifragens por transposição para esta frase? Qual o número de cifragens distintas por transposição para esta frase?

02) Qual é o número de distintas cifragens por transposição para a frase: “HOJE VAMOS APRENDER”?

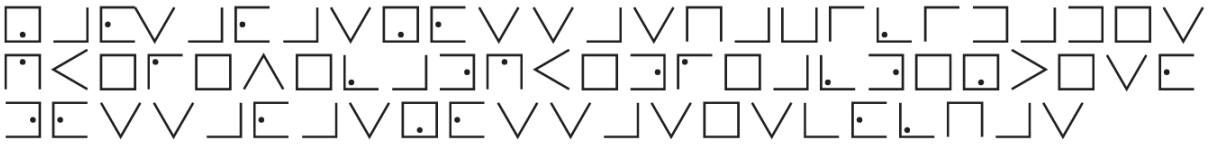
03) Considerando os quatro detalhes a respeito da distribuição de caracteres e construção da espiral na cifra das espirais, qual o total de cifras possíveis desse mesmo tipo para uma mesma frase?

04) Na cifra de César, na cifra de deslocamento por palavra chave e na cifra do chiqueiro (assim como em outras não abordadas aqui), é possível obter uma alto número de cifras utilizando diferentes rearranjos do alfabeto original. Na cifra de César, para o alfabeto latino (26 letras), quantas cifras distintas é possível obter por simples deslocamentos do alfabeto original? Quantas aproximadamente é possível obter através de rearranjos do alfabeto original?

05) Na cifra de deslocamento por palavra chave, se usarmos, como nos exemplos anteriores, a palavra chave CRIPTOGRAFIA, quantas cifras distintas podemos obter a partir de rearranjos do restante das letras do alfabeto?

06) Escolha uma palavra chave (ou frase) e calcule o número de cifras distintas disponíveis a partir de rearranjos do restante das letras do alfabeto:

07) Use as grades de associação da cifra do chiqueiro para decifrar a seguinte mensagem:



4. ANÁLISE DE FREQUÊNCIA

Conteúdos relacionados: Frequência; porcentagem; análise linguística, método hipotético-dedutivo.

A análise de frequência é uma técnica usada em criptoanálise para quebrar cifras de substituição. Em todas as cifras de substituição que vimos até o momento, cada caracter de um texto original é trocado por um único caractere específico no texto cifrado. O conjunto de caracteres usados na cifra constitui o que chamamos de alfabeto cifrado e como, nesses casos, temos um único alfabeto cifrado, tais cifras são chamadas de cifras de substituição monoalfabéticas.

As cifras monoalfabéticas sofrem de uma fraqueza muito grande devido à característica presente em qualquer linguagem escrita de que algumas letras são muito mais frequentes que outras ao longo de textos, isso é verdade em qualquer idioma e não apenas para letras como para determinadas combinações de letras em sequência e para palavras. Por exemplo, em português, a letra “A” é a mais comum, ou seja, a mais abundante em textos em comparação com as outras letras do alfabeto, sendo seguida pelas letras “E”, “O”, “S” e “R”, a abundância ou carência de determinadas letras é determinada por frequências relativamente constantes em textos de um mesmo idioma desde que tais textos sejam longos o suficiente. Dessa forma, como em uma cifra de substituição monoalfabética cada caracter é substituído sempre por um mesmo caracter no texto cifrado, as frequências relativas são preservadas. Assim, desde que se conheça o idioma no qual o texto original foi escrito e as frequências relativas das letras nesse idioma, é razoável supor que o caractere mais frequente do texto cifrado esteja substituindo a letra mais frequente do idioma, o segundo caracter mais frequente esteja substituindo a segunda letra mais frequente e assim por diante.

É evidente que, para letras de baixa frequência, ou letras de frequência parecida as hipóteses sobre quais caracteres as substituem não serão tão certeiras,

mas, desde que se use as informações mais distintas de frequências, e se decifre ao menos algumas partes do texto, o contexto extraído dessas partes já decifradas pode ser usado como auxílio no processo de decifrar as partes restantes. Por exemplo, se um determinado caractere em um longo texto cifrado que sabe-se que foi escrito originalmente em português aparece com maior frequência que todos os outros e supõe-se que o texto tenha sido cifrado por substituição monoalfabética, podemos supor com boa segurança que esse caractere está substituindo a letra “A”. Além disso, se, em algumas partes do texto esse caractere está seguido de um outro cuja frequência seja parecida com a das letras “H” e “Q” (que possuem frequências bem próximas no português) podemos determinar com relativa confiança que esse caracter é um substituto da letra “Q” ao nos atentarmos que em português a letra “A” é bem mais frequentemente seguida pelo “Q” do que pelo “H”. Percebemos que um conhecimento em linguística é um importante aliado na análise de frequência, muitos outros fatores característicos do idioma podem ser levados em consideração na análise, se em um texto cifrado os espaços entre as palavras foram mantidos. É possível, por exemplo, analisar as probabilidades das palavras desse idioma começar e terminar com determinadas letras e usar essas informações para auxiliar a análise de frequência. Abaixo segue uma tabela de frequência das letras do alfabeto em textos em português:

Figura 05: Frequência relativa das letras em textos em português

Letra	Freqüência		Letra	Freqüência
A	14.63%		N	5.05%
B	1.04%		O	10.73%
C	3.88%		P	2.52%
D	4.99%		Q	1.20%
E	12.57%		R	6.53%
F	1.02%		S	7.81%
G	1.30%		T	4.34%
H	1.28%		U	4.63%
I	6.18%		V	1.67%
J	0.40%		W	0.01%
K	0.02%		X	0.21%
L	2.78%		Y	0.01%
M	4.74%		Z	0.47%

Fonte: https://www.gta.ufrj.br/grad/06_2/alexandre/criptoanalise.html#:~:text=As%20vogais%20A%2C%20E%2C%20I,4%20dos%20textos%20em%20Portugu%C3%AAs. (acesso em 17/06/2024)

Vejamos um exemplo de decifragem por análise de frequência onde sabemos que o texto original está em português e foi cifrado por substituição monoalfabética:

Texto cifrado (para simplificação os espaços entre palavras foram mantidos):

xitidzia aiw vi uqvpi vili pcuqtlm wxqvqiw vwaai qvmaowbidmt nwvbm lm uioqi kixihma lm kicaiz ozivlma awnzqumvbwa m biujmu lm zmumlqi twa

Se contarmos os caracteres do texto obtemos os seguintes valores:

i: 20 vezes ≈ 17%
m: 14 vezes ≈ 12%
a: 11 vezes ≈ 9,5%
v: 9 vezes ≈ 7,8%
w: 9 vezes ≈ 7,8%
q: 8 vezes ≈ 6,9%
l: 7 vezes ≈ 6%
u: 7 vezes ≈ 6%
z: 5 vezes ≈ 4,3%
b: 4 vezes ≈ 3,4%
t: 4 vezes ≈ 3,4%
x: 3 vezes ≈ 2,6%
o: 3 vezes ≈ 2,6%
c: 2 vezes ≈ 1,7%
d: 2 vezes ≈ 1,7%
k: 2 vezes ≈ 1,7%
n: 2 vezes ≈ 1,7%
p: 2 vezes ≈ 1,7%
h: 1 vez ≈ 0,9%
j: 1 vez ≈ 0,9%
e: 0 vezes ≈ 0%
f: 0 vezes ≈ 0%
g: 0 vezes ≈ 0%
r: 0 vezes ≈ 0%
s: 0 vezes ≈ 0%
y: 0 vezes ≈ 0%

Comparando as frequências no texto cifrado com as fornecidas pela tabela anterior, podemos supor com relativa segurança que a letra “i” está substituindo a letra “a” e que a letra “m” está substituindo a letra “e”, vamos trocar essas letras no texto cifrado:

xAtAdzAa aAw vA uqvpA vAlA pcuqtlE wxqvqAw vwaaA qvEaowbAdEt nwvbE lE uAoqA kAxAhEa lE kAcaAz ozAvlEa awnzquEvbwa E bAujEu lE zEuElqA twa

A terceira letra mais comum no texto cifrado é o “a” e no português é o “o”, podemos portanto supor que um seja substituto do outro, isso faz todo sentido se considerarmos apenas a frequência das letras, porém teríamos no texto a sequência de letras “ooa” na palavra “vwaaA” o que é muito incomum pois na maioria das

palavras com duplicação da letra “o” a letra seguinte é uma consoante ou já se trata do final da palavra (cooptar; cooperação; voo; coordenador; álcool; zoológico; enjoo; etc), esse detalhe torna pouco provável a hipótese de que o “a” substitui o “o”. Vamos supor, então, que ele seja um substituto para a quarta letra mais comum do português, o “s” nesse caso para a palavra mencionada teremos a tríade “ssa” que é muito comum no português, especificamente no fim de uma palavra como é nesse caso (nossa; vossa; possa; pressa; submissa; remessa; massa; essa; fossa; etc), trocando o “a” pelo “s” temos:

xAtAdzAS SAw vA uqvpA vAlA pcuqtlE wxqvqAw vwSSA qvESowbAdEt nwvbE lE uAoqA kAxAhES lE kAcSAz ozAvlES SwnzquEvbwS E bAujEu lE zEuElqA twS

Essa segunda abordagem parece fazer muito sentido devido a quantidade de vezes em que o “s” apareceu no final de palavras o que é frequente no português devido aos plurais. Vamos procurar agora pela letra que substitui o “o”, as próximas letras mais frequentes do texto cifrado são o “v” e o “w” cujas frequências são idênticas. Se supormos que o “v” esteja substituindo o “o” teremos a palavra “oa” que não faz sentido algum como palavra individual (em português “oa” é um sufixo) e a palavra “OAlA” que também não parece fazer sentido qualquer que seja a letra que introduzirmos no lugar do “l”, vamos então supor que o substituto do “o” seja o “w” o que a princípio não parece gerar contradições ou raridades linguísticas, fazendo a substituição temos:

xAtAdzAS SAO vA uqvpA vAlA pcuqtlE OxqvqAO vOSSA qvESoObAdEt nOvbE lE uAoqA kAxAhES lE kAcSAz ozAvlES SOnzquEvbOS E bAujEu lE zEuElqA tOS

Como já usamos a hipótese de que o “w” substitui o “o” vamos tentar definir que letra está sendo substituída pelo “v” no texto cifrado, já temos hipóteses para as 4 letras mais comuns no português, a quinta letra mais comum é o “r”, se supormos que o “v” seja seu substituto, Neste caso, teremos as palavras “ra” e “RAlA” que faz sentido supondo que o texto fale sobre rãs (ra seria rã sem a acentuação e rala pode ser raça caso o “l” esteja substituindo o “c”) porém teremos também a palavra ROSSA inexistente no português. Apesar da possibilidade de se tratar de um erro de ortografia para a palavra roça, vamos supor que não seja esse o caso. Analisaremos agora a hipótese de que o “v” seja um substituto para a sexta letra mais comum no português o “i”, nesse caso surge a palavra IOSSA, inexistente no português e a palavra IAlA, com o “l” ainda a ser trocado, as palavras obtidas pela substituição do “l” nessa última são consideravelmente raras, considerando essas duas observações vamos descartar essa hipótese. Finalmente podemos supor que o “v” seja um substituto para a sétima letra mais comum no português o “n”, nesse caso teremos algumas palavras bastante comuns como NOSSA, NA e as palavras possíveis de se obter substituindo o “l” em NAlA (nada, nata, naja), vamos seguir por este caminho e realizar a substituição:

xAtAdzAS SAO NA uqNpA NAlA pcuqtlE OxqNqAO NOSSA qNESoObAdEt nONbE lE uAoqA kAxAhES lE kAcSAz ozANlES SOnzquENbOS E bAujEu lE zEuElqA tOS

Agora vamos tentar identificar no texto as letras que substituem o “r” e o “i”. Supondo, inicialmente, que o “q” substitui o “r” teremos a palavra (parcialmente decifrada) RNESoObAdEt, e a dupla “RN” apesar de aparecer em várias palavras (discernimento, cerne, escárnio, adorno, etc) não aparece no início de palavras em português e portanto essa hipótese é bem improvável. Por outro lado partirmos da hipótese de que o “q” substitui o “i” não obtemos situações incomuns ou impossíveis para as sequências de letras que surgem, vamos então realizar tal substituição:

xAtAdzAS SAO NA uINpA NAlA pcuItlE OxINIAO NOSSA INESoObAdEt nONbE lE uAoIA kAxAhES lE kAcSAz ozANlES SOnzIuENbOS E bAujEu lE zEuEIlA tOS

A próxima letra mais comum do texto cifrado é o “l” e a próxima hipótese razoável com base nas frequências é que ele seja um substituto para o “r”, porém isso nos dá a palavra “RE” (que pode ser a palavra ré sem a acentuação), aparecendo três vezes no texto, sendo essa uma palavra que normalmente é comum apenas em contextos musicais (nota ré) ou sobre manobras e movimentos orientados (marcha ré) é mais provável que o “l” substitua a próxima letra mais comum do português, a letra “d”, pois nesse caso a palavra que aparece três vezes será DE, uma preposição extremamente comum mesmo em textos curtos, a partir de tal hipótese temos:

xAtAdzAS SAO NA uINpA NADA pcuItDE OxINIAO NOSSA INESoObAdEt nONbE DE uAoIA kAxAhES DE kAcSAz ozANDES SOnzIuENbOS E bAujEu DE zEuEDIA tOS

Diferindo um pouco da distribuição de frequências de letras no português, em nosso texto o “r” evidentemente tem uma frequência mais baixa que o esperado. Vamos prosseguir procurando determinar qual letra o substitui nessa cifra, vamos supor que seja a próxima letra mais frequente em nosso texto o “u”, seguindo essa hipótese, temos:

xAtAdzAS SAO NA RINpA NADA pcRItDE OxINIAO NOSSA INESoObAdEt nONbE DE RAoIA kAxAhES DE kAcSAz ozANDES SOnzIRENbOS E bARjER DE zEREDIA tOS

Nesse caso o único indício forte de que podemos estar corretos é a palavra RINpA que pode ser RINHA, porém como isso é pouco para tomarmos uma decisão vamos supor que o substituto do “r” seja o “z” (próxima letra mais comum após o “u” em nosso texto cifrado) para essa hipótese temos:

xAtAdRAS SAO NA uINpA NADA pcuItDE OxINIAO NOSSA INESoObAdEt nONbE DE uAoIA kAxAhES DE kAcSAR oRANDES SOnRIuENbOS E bAujEu DE REuEDIA tOS

Nesse caso também temos um único indício forte de estarmos corretos, a palavra oRANDES que pode ser GRANDES, nesse ponto poderíamos escolher uma das hipóteses e continuar nossos testes porém se escolhermos o substituto errado para o “r”. Podemos gastar muito tempo e energia na decifragem antes de ficar evidente nosso erro, portanto vamos tentar melhorar nossas duas hipóteses supondo que entre as duas letras que analisamos “u” e “z” aquela que não for substituta do “r” será substituta do “m” a próxima letra mais frequente do português que ainda não analisamos, para essa abordagem nossas hipóteses nos dão:

HIPÓTESE I (r foi substituído por u e m foi substituído por z):

xAtAdMAS SAO NA RINpA NADA pcRItDE OxINIAO NOSSA INESoObAdEt nONbE DE RAoIA kAxAhES DE kAcSAM oMANDES SOnMIRENbOS E bARjER DE MEREDIA tOS

HIPÓTESE II (r foi substituído por z e m foi substituído por u):

xAtAdRAS SAO NA MINpA NADA pcMItDE OxINIAO NOSSA INESoObAdEt nONbE DE MAoIA kAxAhES DE kAcSAR oRANDES SOnRIMENbOS E bAMjEM DE REMEDIA tOS

Agora fica evidente que, ao menos entre essas duas hipóteses, a segunda é bem mais razoável, pois temos pela segunda hipótese a palavra REMEDIA, muito mais provável de ser parte do texto original que MEREDIA. Além disso, oRANDES obtido pela segunda hipótese pode ser, como já comentamos, a palavra GRANDES. Porém, para a primeira hipótese não temos letras que possam substituir a primeira letra de oMANDES de modo a fazer algum sentido. Pela análise dessa palavra em particular, também vemos que, na segunda hipótese, o “o” deve ser substituto do “g” o que condiz com a palavra MAoIA, que no caso será MAGIA. Portanto, vamos adotar a segunda hipótese e já adicionar a nova hipótese de que “o” substitui “g”:

xAtAdRAS SAO NA MINpA NADA pcMItDE OxINIAO NOSSA INESGObAdEt nONbE DE MAGIA kAxAhES DE kAcSAR GRANDES SOnRIMENbOS E bAMjEM DE REMEDIA tOS

Poderíamos continuar fazendo hipóteses baseadas em comparações das frequências de cada letra no texto cifrado e no português, porém a essa altura, já temos alguns indícios de quais são determinadas palavras (na verdade já tínhamos tais indícios alguns passos atrás, mas dei continuidade a análise de frequência para ilustrar melhor o método). Assim pode ser bem mais vantajoso elaborar as próximas

hipóteses a partir disso, principalmente porque as letras que faltam são aquelas pouco frequentes e algumas delas têm frequências muito parecidas. Por exemplo, a palavra parcialmente decifrada INESGObAdEt, muito provavelmente é INESGOTÁVEL, confiantes de que isso está correto, elaboramos a hipótese de que “b” substitui “t”, “d” substitui “v” e “t” substitui “l”, aplicando essa hipótese temos:

xALAVRAS SAO NA MINpA NADA pcMILDE OxINIAO NOSSA INESGOTAVEL nONTE DE MAGIA kAxAhES DE kAcSAR GRANDES SOnRIMENTOS E TAMjEM DE REMEDIA LOS

Agora, devido a primeira e sétima palavras do texto parece muito evidente que “x” substitui “p”, aplicando essa hipótese obtemos:

PALAVRAS SAO NA MINpA NADA pcMILDE OPINIAO NOSSA INESGOTAVEL nONTE DE MAGIA kAPAhES DE kAcSAR GRANDES SOnRIMENTOS E TAMjEM DE REMEDIA LOS

Pelas palavras nONTE e SOnRIMENTOS, temos que o “n” substitui o “f”, e encontramos:

PALAVRAS SAO NA MINpA NADA pcMILDE OPINIAO NOSSA INESGOTAVEL FONTE DE MAGIA kAPAhES DE kAcSAR GRANDES SOFRIMENTOS E TAMjEM DE REMEDIA LOS

Nesse momento já é bem notável pela palavra MINpA e o contexto em que está inserida que “p” substitui “h” o que nos leva a concluir por HcMILDE que “c” substitui “u” e consequentemente por kAUSAR que “k” substitui “c”:

PALAVRAS SAO NA MINHA NADA HUMILDE OPINIAO NOSSA INESGOTAVEL FONTE DE MAGIA CAPAhES DE CAUSAR GRANDES SOFRIMENTOS E TAMjEM DE REMEDIA LOS

Finalmente, decifrando o “h” como “z”, o “j” como ”b” e inserindo a pontuação e acentuação necessárias, temos o texto decifrado:

PALAVRAS SÃO NA MINHA NADA HUMILDE OPINIÃO, NOSSA INESGOTÁVEL FONTE DE MAGIA, CAPAZES DE CAUSAR GRANDES SOFRIMENTOS E TAMBÉM DE REMEDIÁ-LOS.

Aqui, cabem alguns comentários adicionais sobre o uso da análise de frequência: evidentemente apesar da ideia principal ser bastante simples, a aplicação eficiente envolve muito mais que uma simples comparação de frequências relativas, sendo importante um bom domínio e conhecimento de linguagem, a

habilidade de elaborar boas hipóteses com base em certa quantidade de dados além de testá-las e também, em diversos momentos um pouco de intuição.

Como vimos no exemplo anterior, as frequências de caracteres em um texto curto podem apresentar consideráveis divergências com as frequências do idioma, no nosso exemplo, o “r” era a nona letra mais frequente, enquanto de um modo geral no português ele é quinta. Essas divergências tendem a diminuir em textos mais longos, porém outro detalhe importante é que elas podem se manter ou até serem mais presentes quando analisamos textos relacionados a contextos muito específicos, um documento militar ou um artigo sobre biologia por exemplo, as frequências das letras podem diferir bastante do padrão observado em textos gerais, desse modo, alguém que procura decifrar um texto por análise de frequência, e que sabe que o conteúdo do texto cifrado pertence à um contexto específico, deve preferencialmente usar como base uma tabela de frequências de letras obtidas da análise de textos que também pertençam a esse mesmo contexto.

Outro detalhe é que, conforme avançamos na decifragem de um texto, podemos realizar testes de hipóteses mais gerais sobre o alfabeto cifrado utilizado. No nosso exemplo, depois de decifrar umas quatro ou cinco letras, poderíamos ter comparado a posição de cada letra do texto original no alfabeto com a posição da letra que o substituiu a fim de verificar se a cifra utilizada foi a cifra de César, se o fizéssemos, perceberíamos que, de fato, para cifrar este texto foi usada uma cifra de César com deslocamento de oito posições. Ao contrário de outros métodos de decifragem, percebemos que a análise de frequência pode ser bem flexível, combinando racionalidade e criatividade na hora de elaborar hipóteses. Portanto, ao aplicarmos esse método, caso um palpite lhe pareça uma boa opção de caminho, é interessante analisá-lo, pois, isso não lhe tomará muito tempo e se o palpite se mostrar correto, pode reduzir muito os esforços restantes na tarefa de decifrar o texto.

Para auxiliar nas atividades relacionadas a esse tópico é possível usar algoritmos em alguma linguagem de programação para realizar as tarefas de contagem de caracteres, substituição de um caractere por outro para testar as hipóteses e até mesmo para realizar uma cifragem após um alfabeto cifrado ser definido. No final desse material constam alguns algoritmos em linguagem python que podem ser copiados e colados em um ambiente de programação para a execução. Para a linguagem python existem editores onlines dispensando o download de softwares um exemplo é o encontrado no endereço https://programiz.pro/ide/python. Consta também um algoritmo para implementar a cifra descrita a seguir.

Devido à análise de frequência, qualquer cifra de substituição monoalfabética se torna bastante insegura. Uma das ideias desenvolvidas por criptógrafos para garantir a segurança no envio de mensagens foram as chamadas cifras de substituição polialfabéticas, nas quais, como o nome já indica, mais de um alfabeto cifrado é usado. É evidente que ao utilizar uma cifra de substituição polialfabética deve-se ter bem definido qual dos alfabetos será usado em cada parte do texto, a

fim de que a mensagem possa ser decifrada através de passos bem definidos. Para exemplificar, imagine que dois alfabetos cifrados sejam definidos como segue na tabela abaixo, onde na primeira linha temos o alfabeto padrão e nas seguintes os dois cifrados:

Tabela 08: Dois alfabetos cifrados

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

Fonte: Elaborado pelo autor

De posse de tais ferramentas podemos determinar uma cifra onde a primeira letra de um texto original é cifrada de acordo com o primeiro alfabeto, a segunda de acordo com o segundo, e de um modo geral letras de posição ímpar são cifradas pelo primeiro e as de posição par pelo segundo. O termo “cifra polialfabética” por exemplo, seria cifrado como “ebyxq tgvopsnqltdomq”. Algumas vantagens desse tipo de cifra ficam evidentes já com esse exemplo: a letra “a”, de maior ocorrência na palavra original foi cifrada por duas letras distintas reduzindo sua frequência na palavra cifrada, além de uma mesma letra na palavra original poder ser cifrada por duas letras distintas, uma mesma letra repetida na palavra cifrada pode estar cifrando duas letras distintas, como ocorre por exemplo com a letra “t” que em sua primeira aparição está cifrando a letra “p” e na segunda a letra “e”.

Um exemplo relativamente famoso de cifra polialfabética é a chamada cifra de Vigenère, que usa 26 alfabetos cifrados distintos, cada qual sendo obtido por um deslocamento distinto do alfabeto padrão (ou seja, cada um representa uma cifra de César), incluindo um que curiosamente é o próprio alfabeto padrão (deslocamento de zero posições). A partir de uma palavra, frase, ou sequência qualquer de letras (de preferência fácil de se memorizar) chamada chave, um alfabeto específico dentre os 26 é selecionado para cifrar cada letra do texto original. Vamos observar um exemplo para entender como especificamente a chave atua na escolha dos alfabetos cifrados. Para facilitar o uso de cada alfabeto no momento correto, é construída uma tabela contendo o alfabeto padrão na primeira linha e os alfabetos cifrados nas linhas seguintes, cada qual resultando do deslocamento por uma posição do alfabeto acima, essa tabela é chamada de tábua reta (ou tábula recta, como é mais encontrada na literatura relacionada), ou quadrado de Vigenère:

Figura 05: Quadrado de Vigenère

Alfabeto correto	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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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
26	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

FONTE: https://sca.profmat-sbm.org.br/profmat_tcc.php?id1=1831&id2=286 (acesso em 20/06/2024)

De posse do quadrado de Vigenère, escolhemos uma chave. Vamos usar a palavra “SOMA”, então associamos a primeira letra da chave com a primeira da mensagem, a segunda letra da chave com a segunda da mensagem e assim sucessivamente, quando as letras da chave acabarem (nesse caso na quarta letra) associamos a próxima letra da mensagem com a primeira da chave novamente e continuamos até que todas as letras da mensagem estejam associadas à uma letra da chave, um jeito prático de fazer isso é escrever a chave repetidas vezes acima da mensagem combinando as duas letra por letra, se quisermos cifrar a frase “não temo o que vem depois”, por exemplo, escrevemos:

PALAVRA - CHAVE: S O M A S O M A S O M A S O M A S O M A
MENSAGEM ORIGINAL: n ã o t e m o o q u e v e m d e p o i s

Agora, de acordo com a letra da chave escolhemos uma linha do quadrado, a linha que se inicia por essa letra e que define um dos 26 alfabetos, em seguida procuramos a intersecção dessa linha com a coluna iniciada pela letra correspondente na mensagem e usamos a letra dessa intersecção para cifrá-la, por exemplo, para o nosso caso, a primeira letras da chave é o “S” e a primeira letra da mensagem é o “n”, portanto, esse “n” será cifrado pela intersecção da linha iniciada pelo “S” (linha 18) com a coluna iniciada pelo “n”, nessa intersecção se encontra a letra “f” que portanto será a substituta do “n”, no próximo passo a letra da chave é o “O” que define a linha 14 e está relacionado com o “a” da mensagem na intersecção da linha 14 com a coluna do “a” encontramos a letra “o” que será a substituta do “a”, prosseguindo com o processo obtemos:

PALAVRA - CHAVE:	S	O	M	A	S	O	M	A	S	O	M	A	S	O	M	A	S	O	M	A
MENSAGEM ORIGINAL:	n	ã	o	t	e	m	o	o	q	u	e	v	e	m	d	e	p	o	i	s
MENSAGEM CIFRADA:	f	o	a	t	w	a	a	o	i	i	q	v	w	a	p	e	h	c	u	s

Como era de se esperar em um cifra polialfabética, uma mesma letra original foi cifrada por mais de uma letra e uma mesma letra cifrada representa mais de uma letra original. Como nossa chave tem 4 letras, alternamos no uso de 4 alfabetos cifrados, uma chave maior resultaria no uso de mais alfabetos tornando a cifra ainda mais segura.

Para decifrar, realizamos o processo inverso: associamos as letras da chave com as letras da mensagem cifrada, podemos escrever a chave acima da mensagem cifrada como fizemos para cifrar. A partir de cada letra da chave definimos uma linha do quadro e procuramos nessa linha a letra cifrada relacionada a letra da chave em questão, a letra que inicia a coluna onde a letra cifrada foi encontrada é a letra original. No nosso exemplo, iniciamos com a letra “S” que define a linha 18, nela procuramos a letra “f” e como a mesma se encontra na intersecção da linha 18 com a coluna do “n” então a letra original que foi cifrada como “f” é o “n”.

A abordagem padrão da técnica de análise de frequência não funciona contra a cifra de Vigenère, pois o uso de vários alfabetos disfarça a frequência das letras na mensagem, no nosso exemplo temos as seguintes frequências para as letras na mensagem original e na mensagem cifrada:

MENSAGEM ORIGINAL:

LETRA	o	e	m	n	a	t	q	u	v	d	p	i	s
OCORRÊNCIA	4	4	2	1	1	1	1	1	1	1	1	1	1

MENSAGEM CIFRADA:

LETRA	a	o	w	i	f	t	q	u	v	e	p	h	c	s
OCORRÊNCIA	4	2	2	2	1	1	1	1	1	1	1	1	1	1

Enquanto existem duas letras mais frequentes na mensagem original, na cifrada só há uma, o “a”, que, nesse caso, cifra duas letras distintas da mensagem original.

Apesar da segurança relativa da cifra de Vigenère (que inclusive foi conhecida por muitos anos como a cifra indecifrável), ainda é possível decifrá-la. Em um texto relativamente longo (quanto mais longo maior a chance do padrão mencionado a seguir ocorrer) é bem provável que determinadas palavras de alta ocorrência sejam cifradas em mais de um momento no texto pela mesma parte da chave e portanto resultem em sequências iguais de letras no texto cifrado, o espaço entre essas repetições dá ao criptoanalista uma noção do tamanho (número de letras) da chave utilizada, analisando diferentes repetições que surgem em textos longos é possível elaborar uma boa hipótese sobre esse tamanho da chave quando o espaçamento entre repetições distintas indicam um mesmo tamanho.

A partir dessa hipótese, é possível separar o texto em “blocos” das letras que se supõe terem sido cifradas a partir de uma mesma letra da chave, por exemplo: se as repetições indicam que a chave contém 7 letras, então a primeira, a oitava, a décima quinta letras do texto cifrado e todas as outras que estão deslocadas de um múltiplo de 7 a partir da primeira terão sido cifradas pela primeira letra da chave e portanto formam um bloco, da mesma forma a segunda letra do texto cifrado e todas as outras que estão deslocadas de um múltiplo de 7 a partir dela terão sido cifradas pela segunda letra da chave e formarão outro bloco, em cada bloco a frequência de letras segue o padrão normal e portanto nos blocos é possível aplicar a análise de frequência, como em cada bloco temos efetivamente as letras obtidas por uma cifra de César, a análise de frequência nos permite determinar o deslocamento do alfabeto cifrado usado em cada bloco e consequentemente a letra da palavra chave que determinou esse alfabeto, determinando a palavra chave podemos decifrar o texto facilmente como descrito anteriormente.

4. 1 ATIVIDADES PROPOSTAS:

01) Usando análise de frequência, decifre o texto a seguir, originalmente em português e cifrado por substituição monoalfabética:

QSOQL G GXZKG LQWT Q CTKRQRT RTKKQRTOKQ G GFZTD T IOLZGKOQ G QDQFIQ T XD DOLZTKOG DQL IGPT T XDQ RQROCQ HGK OLLG EIQDQ-LT HKTLTFZT.

02) Usando análise de frequência, decifre o texto a seguir, originalmente em português e cifrado por substituição monoalfabética:

T OVPOCT VLHTC TVUOKIO KL SLCDRLKBO, COMOHTKIL T MTVBT OWBOKVTL IO LUOTKLV TRNDV O VOCOKTV PHLCOVBTV. TV TMOV VLYCOMLTJ TV TCMLCOV, OKXNTKBL LV VLKV IT KTBNCORT COVVLTJ. T UTIT ZTVVL, LV OWZHLCTILCOV OKULKBCTJ T CDXNORT OVULKIDIT KTV TCOTV TDKIT KTL IOVYCTMTITV. L VLH, ULJ VON YCDHSL DKBOKVL, TXNOUO L VLHL TCDIL, ZCOZTCTKIL-L ZTCT TV VOJOKBOV XNO, OJ YCOMO, VOCTL ZHTKBTITV.

03) Usando análise de frequência, decifre o texto a seguir, originalmente em português e cifrado por substituição monoalfabética:

RI UBRI 7RUUZW TRU ZEOUZ HI TZLUHI, RXZUZ7ZELR UZ7PUIRI ZIIZE7BHBI THUH HI 7UBHOPUHI YPZ VHJBOHW ZIOZ WPELR. H IZUZEBLHLZ LR 7ZEHUBR Z YPZJUHLH HTZEHI TZQR 7HEOR LRI THIIHURI Z R WPUWPUBR LHI XRQVHI HR AZEOR. H VHUWREBH ZEOUZ H OZUUH Z RI IZUZI YPZ EZQH ABAZW Z ZABLZEOZ ZW 7HLH LZOHQVZ. ER 7RUHÇHR LHI IZQAHI, UZTQZOHI LZ ABLH Z WBIOZUBRI, RI ABH4HEOZI LZI7RJUZW VBIORUBHI HEOB3HI 3UHAHLHI EHI UR7VHI.

04) Realize uma análise de frequência no texto cifrado pela cifra do chiqueiro na atividade de número 07 do tópico anterior (III - CIFRAGEM POR TRANSPOSIÇÃO E CIFRAGEM POR SUBSTITUIÇÃO).

05) Em grupos, cifrem textos em português com pelo menos 80 palavras, usando substituição monoalfabética, troquem os textos entre os grupos e tentem decifrar o texto recebido através da análise de frequência.

06) Cifre um texto com a cifra de Vigenère, realize uma análise de frequência do texto cifrado e compare o resultado com a frequência das letras no texto original. Argumente sobre a eficiência da análise de frequência contra a cifra de Vigenère:

07) Em grupos, cifrem textos em português com pelo menos 80 palavras, usando a cifra de Vigenère, troquem os textos entre os grupos informando também a quantidade de letras da chave utilizada e tentem decifrar o texto recebido identificando qual é a chave através da análise de frequência nos blocos de letras obtidos do texto cifrado.

5. MATERIAIS DE APOIO:

- CÓDIGO EM PYTHON PARA CONTAR AS LETRAS E ALGARISMOS DE UM TEXTO:

```python
from collections import Counter
import re

def contar_letras(texto):
    # Normaliza o texto removendo acentos e convertendo para minúsculas
    texto_normalizado = re.sub(r'[áàâã]', 'a', texto.lower())
    texto_normalizado = re.sub(r'[éê]', 'e', texto_normalizado)
    texto_normalizado = re.sub(r'[í]', 'i', texto_normalizado)
    texto_normalizado = re.sub(r'[óôõ]', 'o', texto_normalizado)
    texto_normalizado = re.sub(r'[ú]', 'u', texto_normalizado)

    # Remove caracteres que não são letras ou números
    texto_normalizado = re.sub(r'[^a-z0-9]', '', texto_normalizado)

    # Conta a ocorrência de cada letra e número
    contagem_letras = Counter(texto_normalizado)

    return contagem_letras

# Exemplo de uso
texto = "INSIRA SEU TEXTO AQUI"
contagem = contar_letras(texto)

print(contagem)
```

Fonte: Elaborado pelo autor através do chatbot ChatGPT; versão: GPT 4o. Disponível em: https://chatgpt.com/ (acesso em 20/05/2024)

- CÓDIGO EM PYTHON PARA SUBSTITUIR UMA LETRA OU ALGARISMO AO LONGO DE UM TEXTO POR UMA LETRA ESPECIFICADA:

```python
def substituir_letra(texto, letra_antiga, letra_nova):
    # Substitui todas as ocorrências da letra antiga pela letra nova no texto
    texto_substituido = texto.replace(letra_antiga, letra_nova)
    return texto_substituido

# Exemplo de uso
```

```
texto = "INSIRA SEU TEXTO AQUI"
letra_antiga = 'INSIRA AQUI A LETRA OU ALGARISMO A SER SUBSTITUÍDO'
letra_nova = 'INSIRA AQUI A LETRA SUBSTITUTA'

# Normalizar o texto (removendo acentos, mas mantendo a capitalização)
import re

def normalizar(texto):
    texto_normalizado = re.sub(r'[ÁÀÂÃáàâã]', 'a', texto)
    texto_normalizado = re.sub(r'[ÉÊéê]', 'e', texto_normalizado)
    texto_normalizado = re.sub(r'[Íí]', 'i', texto_normalizado)
    texto_normalizado = re.sub(r'[ÓÔÕóôõ]', 'o', texto_normalizado)
    texto_normalizado = re.sub(r'[Úú]', 'u', texto_normalizado)
    return texto_normalizado

texto_normalizado = normalizar(texto)

# Substituir letra especificada
texto_substituido = substituir_letra(texto_normalizado, letra_antiga, letra_nova)

print(texto_substituido)
```

Fonte: Elaborado pelo autor através do chatbot ChatGPT; versão: GPT 4o. Disponível em: https://chatgpt.com/ (acesso em 20/05/2024)

- CÓDIGO EM PYTHON PARA CIFRAR UM TEXTO POR SUBSTITUIÇÃO MONOALFABÉTICA A PARTIR DE UM ALFABETO CIFRADO DADO (QUE PODE CONTER ALGARISMOS):

```
def substituir_letras(texto, alfabeto_original, alfabeto_cifrado):
    # Cria um dicionário de substituição baseado nos alfabetos fornecidos
    substituicao = str.maketrans(alfabeto_original, alfabeto_cifrado)

    # Aplica a substituição no texto
    texto_cifrado = texto.translate(substituicao)

    return texto_cifrado

# Alfabetos fornecidos (os alfabetos devem conter o mesmo número de caracteres que devem ser digitados no campo especificado sem espaços)
alfabeto_original = "abcdefghijklmnopqrstuvwxyz"
alfabeto_cifrado = "INSIRAAQUISEUALFABETOCIFRADO"
```

```
# Exemplo de uso
texto = "INSIRA SEU TEXTO AQUI."

# Normalizar o texto (removendo acentos e convertendo para minúsculas)
import re

texto_normalizado = re.sub(r'[áàâã]', 'a', texto.lower())
texto_normalizado = re.sub(r'[éê]', 'e', texto_normalizado)
texto_normalizado = re.sub(r'[í]', 'i', texto_normalizado)
texto_normalizado = re.sub(r'[óôõ]', 'o', texto_normalizado)
texto_normalizado = re.sub(r'[ú]', 'u', texto_normalizado)

# Substituir letras conforme os alfabetos fornecidos
texto_cifrado = substituir_letras(texto_normalizado, alfabeto_original,
alfabeto_cifrado)

print(texto_cifrado)
```

Fonte: Elaborado pelo autor através do chatbot ChatGPT; versão: GPT 4o. Disponível em: https://chatgpt.com/ (acesso em 20/05/2024)

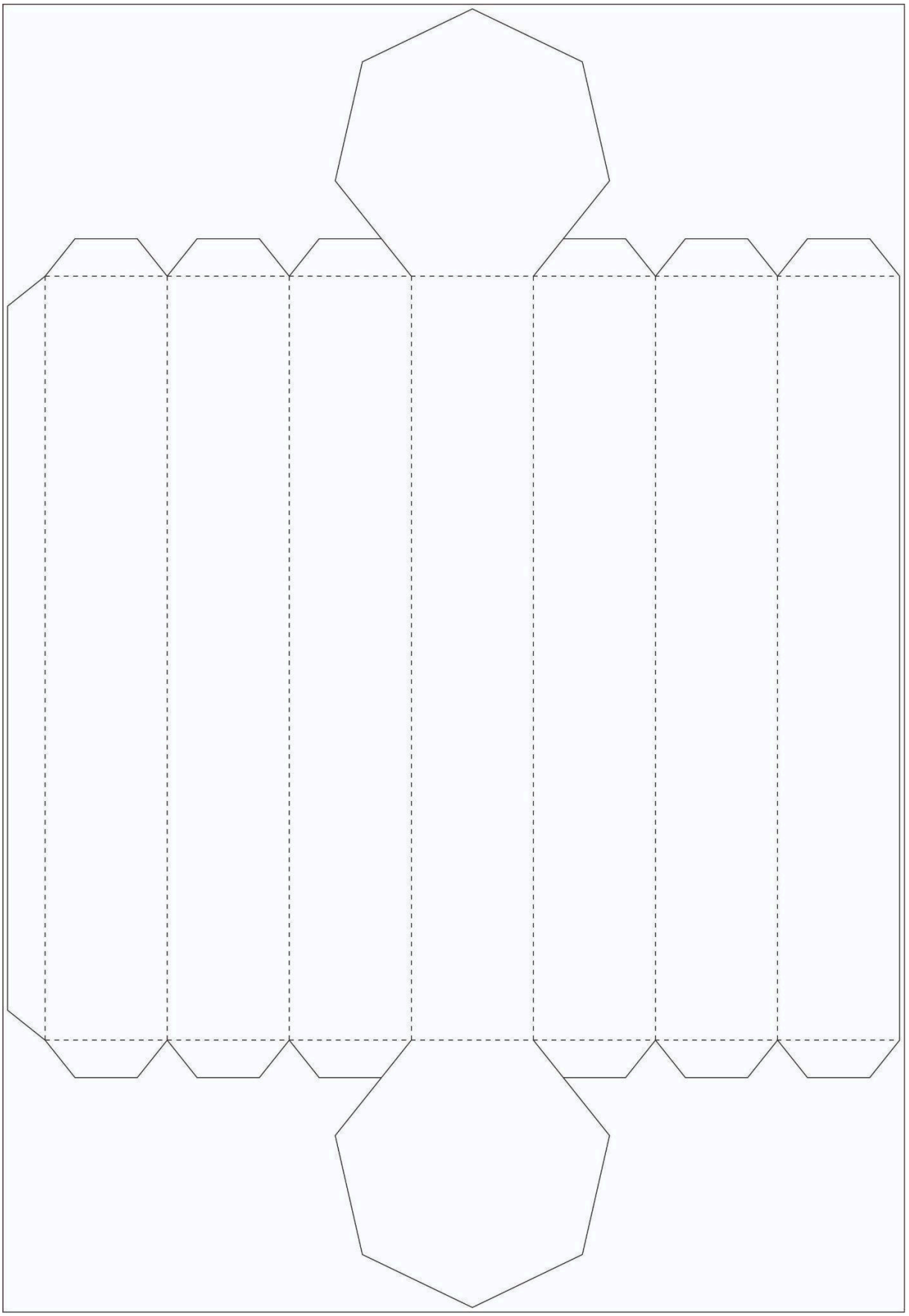
- CÓDIGO EM PYTHON PARA CIFRAR UM TEXTO POR CIFRA DE VIGENÈRE:

```
# criptografa um texto com a cifra de Vigenère

def lettre(c):
    # retorna verdadeiro se for uma letra sem acento
    car = ord(c.upper())
    return car>64 and car<91

def decalage(c,k):
    # altera uma letra para maiúscula. Outras letras não são modificadas
    car = ord(c.upper())
    if lettre(c):
        car += k
        while car>90:
            car -= 26
        while car<65:
            car += 26
        return chr(car)
    else:
        return ""
```

```
def vigenere(message,chave,crypte):
    # executa a mudança com base na chave e nos caracteres da mensagem
    n = 0
    chiffre=''
    for c in message:
        if lettre(c):
            k = ord(chave[n%len(chave)])-65
            if crypte:
                chiffre += decalage(c,k)
            else:
                chiffre += decalage(c,-k)
            n+=1
        else:
            chiffre += c
    return chiffre


# teste
chave = "INSIRA SUA CHAVE AQUI"
texte="INSIRA SEU TEXTO AQUI"
texte_code = vigenere(texte,chave,True)
print(texte_code)
texte_decode = vigenere(texte_code,chave,False)
print(texte_decode)
```

Fonte: https://www.apprendre-en-ligne.net/crypto/python/vigenere/vigenere.py (adaptado pelo autor)

- Planificação de um citale de 7 faces retangulares:



Fonte: Elaborado pelo autor