Criptografia

Cifras simétricas clássicas

Roadmap

- Teoria dos números aplicada às cifras simétricas
- Classificação das cifras
- Cifra clássicas de substituição
- Cifra clássicas de transposição

Universidade Federal do Ceará
Campus de Crateús

Ciência da Computação
Sistemas de Informação

emerson@crateus.ufc.br

Teoria dos números

Euclides — matemático grego que viveu em Alexandria por volta de 300 a.C.

Muitas áreas da matemática são relevantes para criptografia.

No entanto, o ramo mais importante da matemática para a criptografia é a **teoria dos números**.

Alguns dos teoremas e provas na teoria dos número já estavam incluídos na obra clássica de Euclides — *Elementos*.

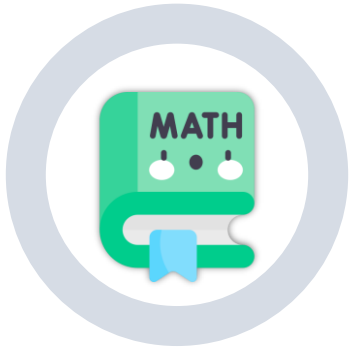
Números primos

- Os **números primos** despertam o interesse dos matemáticos deste a Grécia antiga.
- Na era moderna, os **números primos** são de extrema importância para a **criptografia**.

Definição 1 (número primo). *Um inteiro $p > 1$ é chamado de **primo** se seus únicos fatores positivos são 1 e o próprio p.*

Definição 2 (número composto). *Um inteiro $n > 1$ que não é primo é chamado de **composto**.*

Números primos



- Os números primos são os **elementos de construção** de todos os outros números inteiros positivos.
- Essa propriedade é tão importante que o chamamos de Teorema Fundamental da Aritmética.

Teorema 1 (Teorema fundamental da aritmética). *Todo inteiro positivo* $n > 1$ *pode ser escrito univocamente (exceto quanto à ordem dos fatores) como o produto de números primos, na forma*

$$n = p_1^{e_1} p_2^{e_2} \cdots p_n^{e_n}$$

em que p_i *são primos distintos e cada* $e_i \geq 1$ *é a multiplicidade de* p_i.

Exemplo

$3600 = 2^4 \cdot 3^2 \cdot 5^2$. Note que os fatores são escritos em ordem crescente.

Números primos

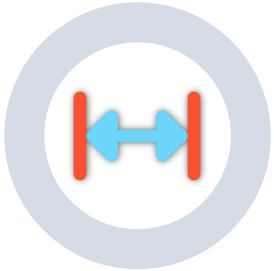
Perguntas e respostas importantes sobre os primos

Quantos números primos existem?

- ***Euclides de Alexandria*** *provou, no século III a.C, que há* ***infinitos*** *números primos.*

É fácil encontrar números primos?

- *No começo, é muito fácil achar números primos****: 2,3, 5, 7, 11, 13, 17, 19, 23, 29...***
- *Entretanto, esse padrão não continua. À medida que se avança nos inteiros positivos, os números primos se* ***distanciam*** *uns dos outros.*



Existe uma fórmula que descreva precisamente qual o próximo primo de uma lista?

- ***Não existe padrão*** *na formação dos números primos.*
- *A sequência de números primos* ***parece totalmente aleatória****.*
- *A falta de padrão é considerado um dos grandes* ***mistérios da matemática****.*

Aritmética modular

- Em vez de fazermos aritmética sobre o conjunto de todos os **inteiros** $\mathbb{Z}$, faremos operações em um outro conjunto, $\mathbb{Z}_n$, em que $\boldsymbol{n}$ é um inteiro positivo, chamado **módulo**.
- $\mathbb{Z}_n$ é chamado de **conjunto de inteiros módulo $\boldsymbol{n}$**, e definido como:

$$\mathbb{Z}_n = \{0,1,2,\dots,n-1\}.$$

> *Note que o conjunto $\mathbb{Z}_n$ é finito, contendo todos os números naturais de 0 a $n-1$.*

- Quase todos os algoritmos de criptografia são baseados em aritmética dentro de um **conjunto finito de elementos**.

> *Essa técnica de executar a aritmética em um **conjunto finito de números inteiros** é conhecida como **aritmética modular**.*

Aritmética modular

Mas é possível realizar operações aritméticas apenas dentro dos ***limites*** *do conjuntos* $\mathbb{Z}_n$*?*

A resposta é ***SIM****. Vejamos!*

- Considere um relógio analógico. Dado uma hora qualquer, se incrementar a hora em 1, terá o seguinte:

  1h, 2h, 3h,..., 11h, 12h, 1h, 2h, 3h,..., 11h, 12h...

- Mesmo que continue adicionando uma hora, nunca sairá do conjunto .
- Para converter entre o relógio de 24 horas e o de 12 horas, basta tomar o resto da divisão do valor no sistema de 24 horas por 12.

Aritmética modular

- A **aritmética modular** é um sistema de **aritmética de congruências** — isto é, que opera sobre os **restos** dos inteiros divididos por um **valor fixo**, o módulo.

> **Definição 3 (Relação de congruência).** *Dado um inteiro positivo n, chamado de **módulo**, diz-se que dois inteiros a e b são **congruentes módulo n**, se n divide a diferença $a - b$. Isso o é escrito como:*
>
> $$a \equiv b \pmod{n}$$

> *Dizer que $a \equiv b \pmod{n}$ é equivalente a dizer que a e b deixam o **mesmo resto** quando divididos por n.*

> *Na aritmética modular, se dois números a e b são congruentes, então consideramos a e b iguais.*

Classes de resíduos

- Observe que o operador $(\text{mod } n)$ mapeia todos os inteiros para o conjunto

$$\mathbb{Z}_n = \{0,1,2,\dots,n-1\}.$$

- Qualquer inteiro é **congruente módulo n** a um único inteiro no conjunto $\mathbb{Z}_n$.
- O conjunto $\mathbb{Z}_n$ é conhecido como o **conjunto de classes de resíduos** módulo n.

Isto significa que cada inteiro em $\mathbb{Z}_n$ representa uma classe de resíduos módulo n, que são denotadas por $[0],[1],[2],\dots,[n-1]$.

Definição 3 (classe de resíduos). *Uma classe de resíduos consiste em todos os inteiros com o mesmo resto, quando divididos por n, tal como*

$$[x] = \{a \in \mathbb{Z} \mid a \equiv x \pmod{n}\}$$

Classes de resíduos

- As classes de resíduos módulo 4 são:

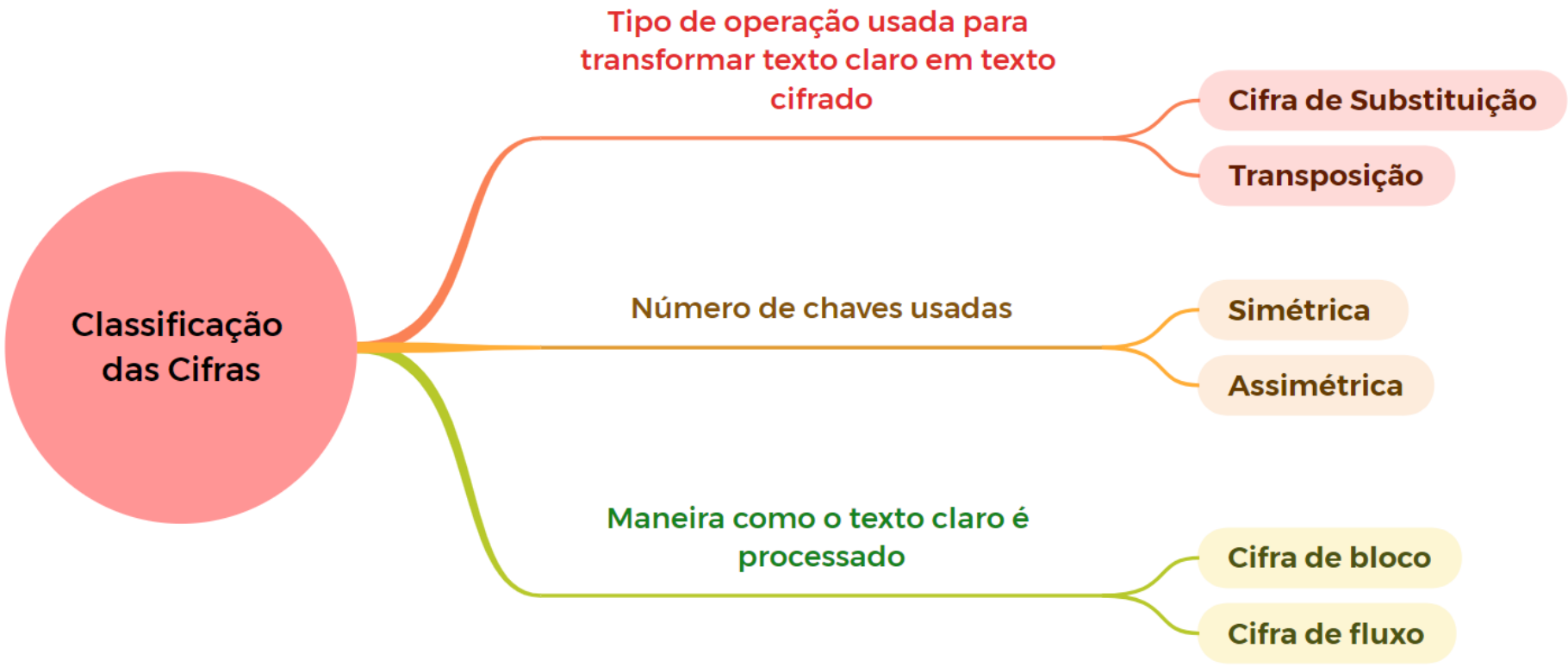
$$[0]=\{\ldots, -16, -12, -8, -4, 0, 4, 8, 12, 16,\ldots\}$$
$$[1]=\{\ldots, -15, -11, -7, -3, 1, 5, 9, 13, 17,\ldots\}$$
$$[2]=\{\ldots, -14, -10, -6, -2, 2, 6, 10, 14, 18,\ldots\}$$
$$[3]=\{\ldots, -13, -9, -5, -1, 3, 7, 11, 15, 19,\ldots\}$$

> *De todos os inteiros em uma classe de resíduos, o **menor não negativo** é aquele normalmente usado para representá-la.*

> *Por exemplo, a classe representada ao número 2 é exatamente o conjunto dos inteiros que são congruentes com 2 módulo 4.*

Classificação das cifras

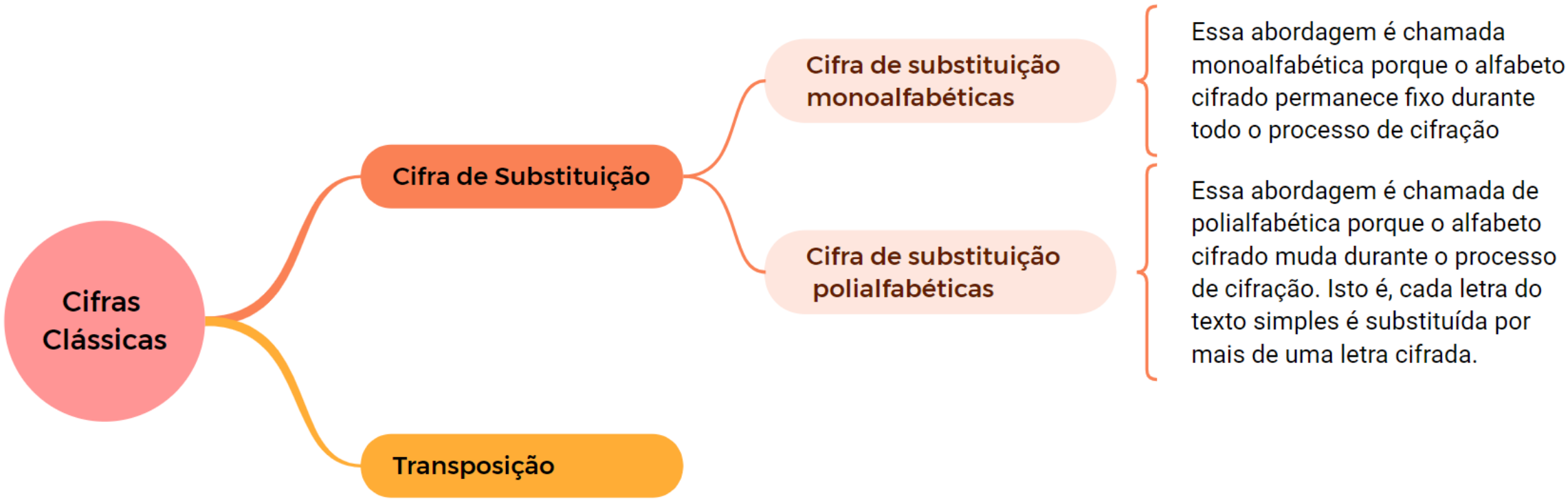
Os sistemas criptográficos são caracterizados em três dimensões **independentes**.



Cifras simétricas

Cifras clássicas

- Os dois blocos básicos de construção de todas as técnicas de criptografia são a **substituição** e a **transposição**.



Cifras de substituição monoalfabética

Cifra de César

- A cifra de substituição mais antiga e mais simples é creditada a **Júlio César**.
- A cifra, que ficou conhecida como **cifra de César**, envolve a substituição de **cada letra do alfabeto** pela letra que está **três** posições adiante.

```
Original: abcdefghijklmnopqrstuvwxyz
Cifrado : DEFGHIJKLMNOPQRSTUVWXYZABC
```

- Ou seja, a letra "**A**" é substituída pela letra "**D**", a letra "**B**" é substituída pela letra "**E**", e assim por diante.

> *Note que o alfabeto recomeça no final, de modo que a letra após dizer é **Z** é **A**.*

Cifras de substituição monoalfabética

Cifra de César

- Suponha que Alice queira enviar para Bob a seguinte mensagem de forma segura:

  aprendendo criptografia na ufc

  DSUHQGHQGR FULSWRJUDILD QD XIF

- Se Bob não souber o que Alice fez, não compreenderá a mensagem, pois a mensagem que ela enviou não faz sentido.

- No entanto, se Bob souber que Alice aplicou a cifra de César, ele pode decifrar a mensagem.

- Basta que Bob substitua cada letra pela que vem **três** posições antes no alfabeto.

Cifras de substituição monoalfabética

Cifra de César

- A cifra de Cesar, apesar de simples, é um **criptossistema**.

> **Relembrando!**
>
> *Um **sistema criptográfico** é um par de algoritmos, um para converter de um texto simples para um texto cifrado e o outro para converter de um texto cifrado para um texto simples.*

- O **protocolo** para transformar a mensagem.
  - *Neste caso, consiste na substituição de uma letra por outra que está à frente.*
- Uma **chave** específica.
  - *No caso da cifra de César, a chave é **3** — i.e., cada letra deve ser substituída pela letra que está **três** posições a frente.*

Cifras de substituição monoalfabética

Cifra de César

Generalizando!

- Vamos atribuir um equivalente numérico a cada letra.

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Para **cifrar**, substitua cada letra p do texto claro pela letra C do alfabeto cifrado deslocando k posições. k é a chave e pode assumir um valor no intervalo de $\mathbf{1}$ a $\mathbf{25}$.

$$C = (p + k) \bmod 26$$

- O protocolo de **decifração** é:

$$p = (C - k) \bmod 26$$

Cifras de substituição monoalfabética

Cifra de César

Quebrando a cifra de Cesar

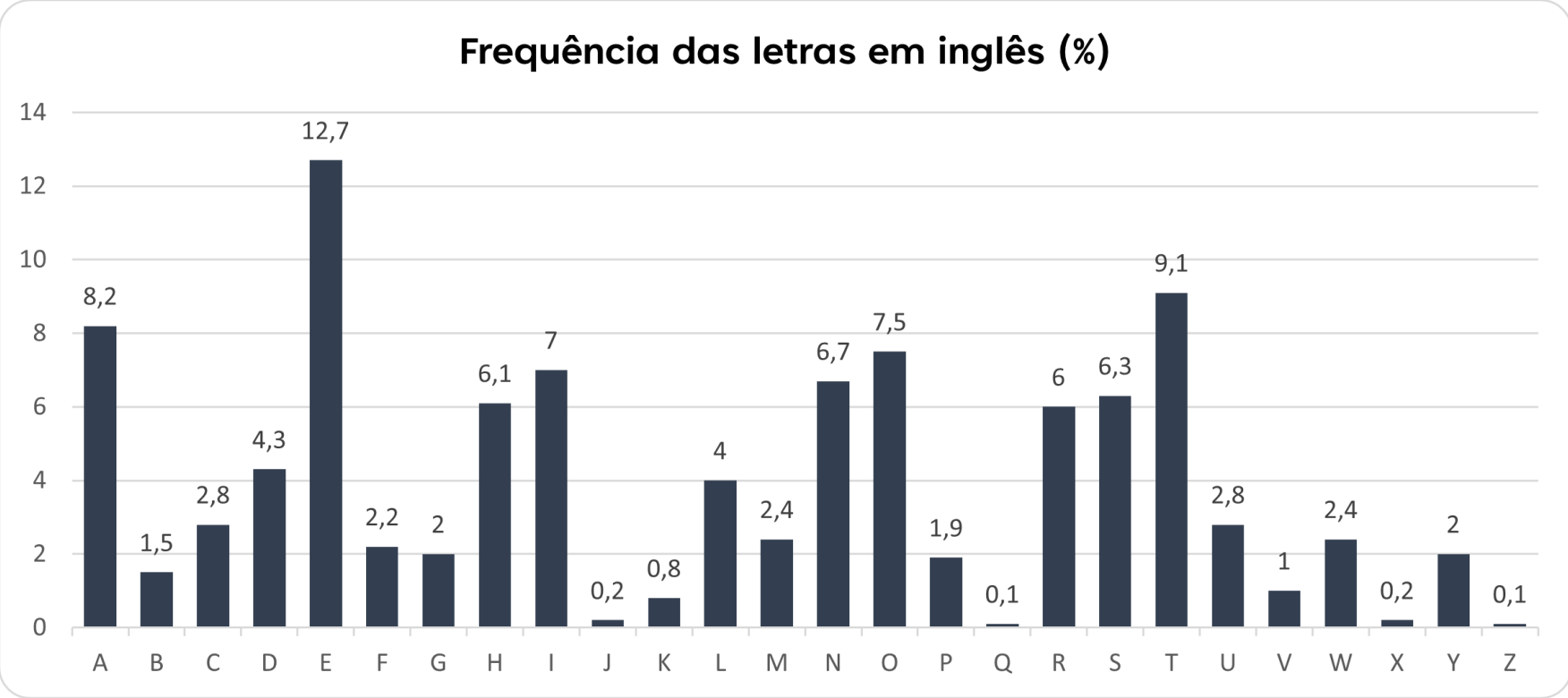
- Uma criptoanálise por **força bruta** é fácil. Basta tentar todas as **25** chaves possíveis.

Chave 1	crtgpfgpfq etkrvqitchkc pc whe
Chave 2	bqsfoefoep dsjquphsbgjb ob vgd
Chave 3	aprendendo criptografia na ufc
Chave 4	zoqdmcdmcn bqhosnfqzehz mz teb
Chave 5	ynpclbclbm apgnrmepydgy ly sda
Chave 6	xmobkabkal zofmqldoxcfx kx rcz
Chave 7	wlnajzajzk ynelpkcnwbew jw qby
Chave 8	vkmziyziyj xmdkojbmvadv iv pax
Chave 9	ujlyhxyhxi wlcjnialuzcu hu ozw
Chave 10	tikxgwxgwh vkbimhzktybt gt nyv
Chave 11	shjwfvwfvg ujahlgyjsxas fs mxu
Chave 12	rgiveuveuf tizgkfxirwzr er lwt
Chave 13	qfhudtudte shyfjewhqvyq dq kvs
Chave 14	pegtcstcsd rgxeidvgpuxp cp jur
Chave 15	odfsbrsbrc qfwdhcufotwo bo itq
Chave 16	nceraqraqb pevcgbtensvn an hsp
Chave 17	mbdqzpqzpa odubfasdmrum zm gro
Chave 18	lacpyopyoz nctaezrclqtl yl fqn
Chave 19	kzboxnoxny mbszdyqbkpsk xk epm
Chave 20	jyanwmnwmx larycxpajorj wj dol
Chave 21	ixzmvlmvlw kzqxbwozinqi vi cnk
Chave 22	hwyluklukv jypwavnyhmph uh bmj
Chave 23	gvxktjktju ixovzumxglog tg ali
Chave 24	fuwjsijsit hwnuytlwfknf sf zkh
Chave 25	etvirhirhs gvmtxskvejme re yjg

Cifras de substituição monoalfabética

Análise de frequência

- As **cifras monoalfabéticas** são fáceis de ser quebradas porque refletem os dados de frequência relativa das letras do alfabeto original.



Cifras de substituição polialfabética

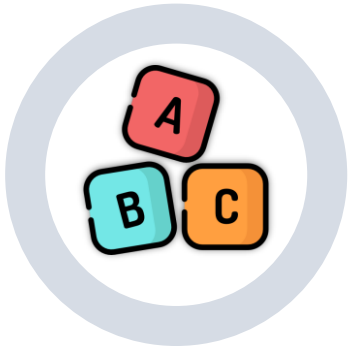
- Para dificultar a **análise de frequência**, pode-se construir uma cifra mais forte, conhecida como **cifra polialfabética.**

> *Em uma **cifra de substituição polialfabética**, cada letra do texto simples é substituída por mais de uma letra cifrada, tornando o trabalho mais difícil para o criptoanalista.*

- Por exemplo, a letra "**e**" pode ser atribuída a vários símbolos cifrados diferentes, conhecidos como **homófonos**.
  - Neste caso, cada homófono seria usado em rodizio ou aleatoriamente.
  - Ou seja, se um "g" for cifrado como um "**X**" em um ponto, ele não será necessariamente cifrado como um "**X**" posteriormente na mensagem.

Cifras de substituição polialfabética

Cifra de Vigenère



- A **cifra de Vigenère**, criada no século XVI, é a cifra polialfabética mais conhecida e uma das mais simples.
- Esta cifra é semelhante **à cifra de César**, exceto que as letras não são deslocadas **por um valor fixo**.

- No caso da cifra de Vigenère, o valor do **deslocamento muda para cada letra** definido por uma **chave**.

> *A **chave** é uma simples coleção de letras que representam números com base em sua posição no alfabeto.*

- Dito de outra forma, o conjunto de **regras de substituição** consiste nos **26 alfabetos** cifrados de César, com deslocamentos de **0** a **25**.

Cifras de substituição polialfabética

Cifra de Vigenère

EXEMPLO

*Para auxiliar no uso deste esquema, uma matriz conhecida como **tabela de Vigenère** é usada.*

- **TEXTO SIMPLES :** prove theth eorem
- **CHAVE :** GRAPE GRAPE GRAPE
- **TEXTO CIFRADO:** VIOKI ZYEIL KFRTQ

	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

Cifras de substituição polialfabética

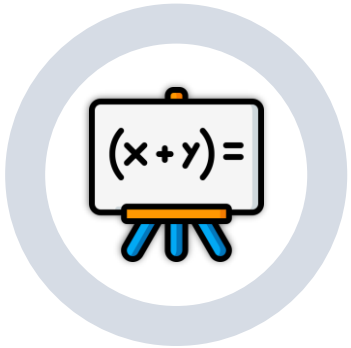
Cifra de Vigenère

Em resumo!

- Para **cifrar**, usa-se “G” como a chave para a primeira letra, “R” como a chave para a segunda e assim por diante.
  - Assim que a chave termina, recomeça do início da palavra.

- Para **decifrar**, se o primeiro caractere do texto cifrado é “V” e a primeira letra da chave é “G”, então percorre-se na coluna “G” até encontrar o “V”.
  - A letra “V” aparece na linha “p”. Portanto, a primeira letra do texto simples é “p”.
  - Continua-se fazendo isso para cada letra no texto cifrado.

Descrição algébrica da cifra de Vigenère



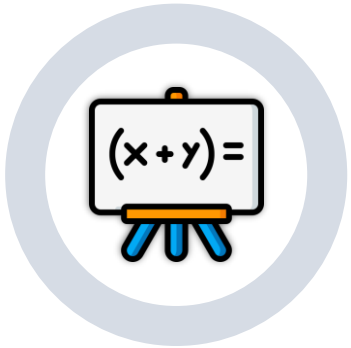
- Em vez de pensar na mensagem como sendo composta por **letras**, vamos pensar nela como sendo composta por **números**, de **0** a **25**.
  - Assim, **A** será **0**, **B** será **1** e assim por diante.

Notação

- Denotamos o **texto simples** por $\boldsymbol{P} = \boldsymbol{p_0, p_1, p_2, \ldots, p_{i-1}}$, em que $\boldsymbol{p_i}$ é o i-ésimo número em $\boldsymbol{P}$.
- Denotamos a **chave** por $\boldsymbol{K} = \boldsymbol{k_0, k_1, k_2, \ldots, k_{m-1}}$, em que $\boldsymbol{k_m}$ é o m-ésimo número em $\boldsymbol{K}$.

> *Normalmente, a chave é menor que a mensagem. Assim, repete-se a chave até o tamanho do texto simples.*

Descrição algébrica da cifra de Vigenère



A equação geral do processo de **cifração** é:

$$C_i = (p_i + k_j) \bmod 26$$

E a equação geral do processo de **decifração** é:

$$p_i = (C_i - k_j) \bmod 26$$

Cifras de substituição polialfabética

Descrição algébrica da cifra de Vigenère

Exemplo!

Chave : deceptivedeceptivedeceptive

Texto simples : wearediscoveredsaveyourself

Texto cifrado : ZICVTWQNGRZGVTWAVZHCQYGLMGJ

- Expresso numericamente, temos:

*	3	4	2	4	15	19	8	21	4	3	4	2	4	15	19	8	21	4	3	4	2	4	15	19	8	21	4
@	22	4	0	17	4	3	8	18	2	14	21	4	17	4	3	18	0	21	4	24	14	20	17	18	4	11	5
#	25	8	2	21	19	22	16	13	6	17	25	6	21	19	22	0	21	25	7	2	16	24	6	11	12	6	9

Legenda: * chave | @ texto claro | # texto cifrado

Cifras de transposição

- Um tipo **diferente** de cifra é obtido realizando algum tipo de **permutação** nas letras do texto simples. Essa técnica é chamada de **cifra de transposição**.

Relembrando!

Uma **permutação** é um conjunto finito de elementos S em uma sequência ordenada de todos os elementos de S, com cada um aparecendo exatamente uma vez.

- Por exemplo, se $S = \{a, b, c\}$, existem seis permutações de S:

$$\{abc\}, \{acb\}, \{bac\}, \{bca\}, \{cab\}, \{cba\}$$

> *Existem* $n!$ *permutações de um conjunto de* n *elementos, porque o primeiro elemento pode ser escolhido de uma das* n *maneiras, o segundo de* $n - 1$ *maneiras, o terceiro de* $n - 2$ *maneiras, e assim por diante.*

Cifras de transposição

- **Note que!** Uma cifra de **transposição** não substitui um símbolo por outro; na verdade, ela modifica a localização dos símbolos.
- Um símbolo na **primeira** posição do texto claro pode aparecer na **décima** posição do texto cifrado.

> *Em outras palavras, uma cifra de transposição reordena (transpõe) os símbolos.*

Cifras de transposição

Cifra de rail fence (cerca de trilhos)

*A cifra de **rail fence** é uma técnica na qual o texto simples é escrito como uma sequência de diagonais e depois lido como uma sequência de linhas —como se estivéssemos descendo e subindo diagonalmente nos trilhos de uma cerca.*

Exemplo!

- Para cifrar a mensagem "`meet me after the toga party`" com uma cerca de trilho **profundidade 2**, faz-se:

m		e		m		a		t		r		h		t		g		p		r		y
	e		t		e		f		e		t		e		o		a		a		t	

- A **chave** é um número de linhas na cerca. A mensagem cifrada é:

`MEMATRHTGPRY ETEFETEOAAT`

Cifras de transposição

Cifra de rail fence (cerca de trilhos)

- Note que no exemplo anterior, temos uma mensagem com 23 caracteres. Portanto, precisamos de uma grade com **23 colunas** e **2 linhas**, que a chave.
- Para **decifrar** a texto cifrado `MEMATRHTGPRY ETEFETEOAAT`, basta preencher os espaços destacados em ordem de cima para baixo.

Preenche a primeira linha completa antes de passar para a segunda linha.

M		E		M		A		T		R		H		T		G		P		R		Y
	E		T		E		F		E		T		E		O		A		A		T	

- Por fim, basta ler o texto na diagonal para recuperar o texto simples.

Cifra de rail fence (cerca de trilhos)

Quebrando a cifra

- A cifra da cerca de trilhos não é muito segura, porque **não há muitas chaves possíveis.**
- Note que não faz muito sentido se o comprimento da chave for maior do que o comprimento do texto simples.
- Dessa forma, o número de chaves possíveis pode ser facilmente verificado por computador, ou mesmo a mão.

Fim!

[Aula 03] Cifras simétricas clássicas