

Cifras de Substituição Simples

Agora que você está familiarizado com a Cifra de César, vamos apresentar uma generalização desse conceito: as **cifras de substituição simples**.

Em uma cifra de substituição simples, cada letra é substituída individualmente de acordo com um **alfabeto de substituição**. Esse alfabeto pode ser uma rotação fixa do alfabeto normal (como a cifra de César) ou algum embaralhamento mais complexo.

Alguns exemplos notáveis de cifra de substituição simples são:

Cifra de Atbash

“ Seu nome tem origem da primeira, última, segunda e penúltima letra Hebraica (Aleph-Taw-Bet-Shin)

Nessa cifra, cada letra é mapeada para o alfabeto invertido, ou seja, a primeira vira a última, a segunda vira a penúltima e assim por diante.

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

Assim, se usarmos essa cifra em `may the force be with you`, obteremos:

original:	M	A	Y	T	H	E	F	O	R	C	E	B	E	W	I	T	H	Y	O	U
cifrado:	N	Z	B	G	S	V	U	L	I	X	V	Y	V	D	R	G	S	B	L	F

A Cifra de Atbash pode ser interpretada como um caso particular da Cifra de Affine, uma cifra que usa aritmética modular para encriptar.

Cifra da Palavra-Chave

A Cifra da Palavra-Chave ou *keyword cipher* consiste em escolher uma **chave** e usá-la para decidir como as letras serão substituídas.

As palavras repetidas dessa chave serão removidas e a própria chave será o começo do alfabeto a ser mapeado. O resto das letras continuarão em ordem alfabética, tirando as letras já usadas.

Por exemplo, escolhendo a chave `Marvin`, o novo alfabeto terá esse formato

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

Assim, ao encriptar a mensagem `Arthur Dent`, obteremos:

```
original: A R T H U R D E N T
cifrado:  M P S C T P V I J S
```

Detectando

Como mencionado na seção de Cifra de César, uma mensagem encriptada por uma cifra de substituição simples terá uma distribuição de frequência das letras semelhante ao da língua usada, mas com as letras trocadas.

Essa distribuição de frequência de um texto pode ser identificada através de uma **análise de frequência**.

Nas línguas naturais, algumas letras aparecem mais frequentemente que outras, como uma espécie de digital do idioma. Por exemplo, a letra mais comum na língua inglesa é o **"e"**, em português é o **"a"**.

Essa análise de frequência pode ser feita simplesmente contando as letras do texto. Existem ferramentas online para isso como o site **dcode** ou pode ser feito rapidamente com um biblioteca em Python, onde `text` é o texto a ser analisado:

```
from collections import Counter
Counter(text.upper()).most_common()
```

Solucionando

O ponto fraco de cifras de substituição simples é que elas são muito suscetíveis à **análises de frequência**.

Assim, se você tiver um texto de tamanho razoável, por volta de 50 caracteres, é possível analisar a frequência com que as letras aparecem e deduzir qual foi o alfabeto de substituição usado.

O site [guaballa](#) é um excelente decodificador de cifras de substituição simples.

Referências

Cifra de Atbash: [Jeremiah's Game](#)

Cifra da Palavra-Chave: [GeeksforGeeks](#)

[Learn Cryptography](#)

Exercícios

[OverTheWire: Krypton 3](#)

Revision #4

Created Sat, Oct 6, 2018 6:48 PM by [Andrew](#)

Updated Mon, Jan 28, 2019 10:01 PM by [Andrew](#)