Universidade Federal do Rio de Janeiro - UFRJ
Gabriel Martins

Cifra de Vigenère

Rio de Janeiro
2017

Números Inteiros e Criptografia

Professor: Luis Menasché
Período: 2017.2
Aluno: Gabriel Martins Machado Christo
DRE: 117217732

SUMÁRIO

Introdução

A cifra de Vigenère é um método de criptografia simétrica, que usa uma série de diferentes cifras de César, baseadas nas letras de uma determinada senha (chave).

Assim como a cifra de César, a cifra de Vigenère é um método de criptografia simétrica, ou seja, utiliza a mesma chave para criptografar e descriptografar. Porém, diferente da cifra de César, a cifra de Vigenère utiliza uma longa senha alfabética como chave, e múltiplas cifras de césar (múltiplos alfabetos) baseadas nesta senha, o que neutraliza o problema da frequência de letras (uma letra ser codificada de apenas uma maneira, que facilita o processo de quebra da cifra). O tipo dessa substituição denomina-se polialfabética monográfica.

A invenção da cifra de Vigenère é erradamente atribuída a Blaise de Vigenère. Ela foi originalmente descrita por Giovan Batista Belaso no seu livro datado de 1553, com o título *“La cifra del. Sig. Giovan Batista Belaso”*.

Encriptação e Decriptação

Para cifrar, é usada uma tabela de alfabetos que consiste no alfabeto escrito 26 vezes em diferentes linhas, cada um deslocado ciclicamente do anterior por uma posição. As 26 linhas correspondem às 26 possíveis cifras de César.

	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

Cada letra da palavra escolhida como chave vai indicar a linha a ser utilizada para cifrar ou decifrar uma letra da mensagem.

- Exemplo:
  Texto: “OHOGAS”
  Chave: “KKEAEMEN”

Escolhendo a chave e repetindo-a até ter o comprimento do texto a cifrar, obtemos “KKEAEM”.

A primeira letra do texto, “O”, é cifrada usando o alfabeto na linha “K”, que é a primeira letra da chave. Olhando na tabela mostrada, vemos que a letra na linha “K” e coluna “O” é “Y”. Para a segunda letra do texto, ver a letra na linha “K” e coluna “H” , que é “R”. Repete-se o processo até a última letra do texto:

- Exemplo:
  Texto: “OHOGAS”
  Chave: “KKEAEMEN”
  Texto cifrado: “YRSGEE”

A decriptação é feita através do processo inverso. Por ser um método simétrico, necessita-se da mesma chave utilizada na encriptação.

- Exemplo:
  Texto cifrado: “YRSGEE”
  Chave: “KKEAEMEN”

A primeira letra do texto cifrado, “Y”, foi cifrada utilizando o alfabeto da linha “K”, que é a letra correspondente na chave. Para obter a letra original basta verificar qual é a coluna onde se encontra a letra “Y”, na linha “K”. Repete-se o processo até a última letra do texto cifrado.

- Exemplo:
  Texto cifrado: “YRSGEE”
  Chave: “KKEAEMEN”
  Texto original: “OHOGAS”

Forma Algébrica

A cifra de Vigenère também pode ser vista de forma algébrica. Primeiro mapeamos as letras A–Z nos números inteiros 0–25:

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
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

Depois, aplicamos a adição do módulo 26, que é a quantidade de letras do nosso alfabeto. Assim obtemos duas equações:

Encriptação: $C_i=P_i+K_i$ (mod 26)
Decriptação: $P_i=C_i-K_i+26$ (mod 26)

*O operador mod é o resto da divisão por 26

Onde:

C = Letra cifrada
P = Letra original
K = Chave de criptografia
i = Índice

- Exemplo de encriptação:
  Texto original: “COLLIER”
  Chave: “REPROVEI”

A primeira letra do texto original é “C”, que vale 2. A letra de índice correspondente na chave é “R”, que vale 17. Assim temos: C=2+17 (mod 26)=19 mod 26=19. Na tabela, 19 corresponde a letra “T”. Assim, obtemos nossa letra cifrada. Repetindo o processo até a última letra do texto original, obtemos:

Texto cifrado: “TSACWZV”

- Exemplo de decriptação:
  Texto cifrado: “TSACWZV”
  Chave: “REPROVEI”

Começando arbitrariamente pela última letra do texto cifrado,achamos “V” que vale 21. A letra com índice correspondente da chave é “E” que vale 4. Assim temos P=21-4+26 (mod 26)=43 mod 26=17. Na tabela, o número 17 corresponde a letra “R”, que é justamente a última letra do texto original. Repetindo o processo para todas as letras do texto cifrado, obtém-se:

Texto original: “COLLIER”

Quebra da Cifra

A grande vantagem da cifra de Vigenère é que as cifras polialfabéticas são mais difíceis de quebrar por análise de frequência, porém, se a chave fosse curta e constantemente repetida, como resultado, palavras comuns provavelmente iriam aparecer criptografadas segundo as mesmas letras da chave.

Avaliando algumas características do texto cifrado era possível definir o tamanho dessa chave e a partir disso, analisar as linhas de letras cifradas pela mesma linha da tabela, de modo que passava a ser possível empregar um método padrão baseado na frequência de certas letras na linguagem. Essa descoberta de padrões repetidos no texto, levaria à descoberta da chave de criptografia.

Dentre as principais técnicas de *data mining (extração não trivial de informação implícita, desconhecida, e potencialmente útil, a partir de um conjunto de dados)*, utilizadas nos processos de criptoanálise, estão:

Sequence Mining: Esta técnica é usada para pesquisar padrões estatisticamente relevantes em sequências de dados. Na criptoanálise, a técnica de sequence mining poderá ser particularmente útil na pesquisa de padrões de letras (que se repetem ao longo da cifra), permitindo determinar o tamanho da chave com mais precisão.

Modelos de Previsão: Constituem o tipo de *data mining* mais conhecido e usado. Estes modelos são usados para prever o valor de um atributo, a partir de um conjunto de outros atributos conhecidos. Assim, podem ser usados modelos de previsão para se prever as palavras a decifrar. A ideia é, considerando que algumas letras da palavra-chave estão erradas, conseguir encontrar as palavras corretas através de métodos de previsão.

Graph Mining: Recorrendo a grafos, é possível representar o relacionamento entre as diversas palavras, e pontuações, de uma qualquer linguagem. Assim, é possível fazer a reconstrução de um texto, cujas pontuações, acentuações e espaços foram omitidos (fato que acontece na cifra de Vigenère), e com eventuais erros.

Top K: Pode ser definido como sendo uma estrutura de dados onde é possível armazenar um máximo de K elementos. Nesta estrutura, o critério de seleção consiste num valor de ranking dos elementos, permanecendo na estrutura um máximo de K elementos, ordenados por ranking. Assim, existem para o efeito dois vetores ordenados, um para os valores de ranking dos elementos, e outro para os próprios elementos. Sempre que um elemento é inserido ou empurrado para a posição K+1, será eliminado.

Por muito tempo a cifra ficou marcada como confiável e chegou a ser classificada como inquebrável pelo matemático Charles Lutwidge Dodgson e elogiada pela revista Scientific American.

Charles Babbage foi o primeiro a descobrir como quebrar a cifra de Vigenère, em 1854, durante a Guerra da Crimeia, em análise matemática de cifras polialfabéticas. Os documentos dessa análise só foram publicados tempos depois, por outro pesquisador, o Prússio Friedrich Kasiski.

Ao contrário do esperado, esse acontecimento ajudou a fortalecer a cifra, como iremos ver adiante, nas melhorias da cifra.

Melhorias

Após a quebra da cifra por Charles Babbage, viu-se que uma fragilidade era o tamanho da chave utilizada. A primeira melhoria da cifra foi a utilização de uma palavra chave com o mesmo tamanho da mensagem, o que prevenia repetições de padrões, que possibilitavam a análise de frequência de letras.

Para evitar vulnerabilidades estatísticas pelo uso de textos sensíveis como frase-chave, surgiu mais uma melhoria para a cifra: a utilização de sequências aleatórias de letras como frase-chave. Esse foi o início do conceito de chaves de uso único.

Além das duas melhorias acima citadas, tem-se inúmeras variantes da cifra de Vigenère surgidas ao longo do tempo. Dentre elas:

- Cifra de Vernam (One-time pad)
- Cifra de Beaufort
- Cifra de Gronsfeld
- Cifra de auto-chave
  E muitas outras…

Iremos falar especificamente sobre três das cifras mencionadas: a cifra de Gronsfeld,a cifra de auto-chave e a cifra de Vernam:

- **Cifra de Gronsfeld** - Criada pelo belga José de Bronckhorst, Conde de Gronsfeld, em 1734. É exatamente igual a cifra de Vigenère, exceto que números são usados como chave. Isso faz com que diferente da cifra de Vigenère, que tem 26 possíveis deslocamentos, a cifra de Gronsfeld tenha apenas 10 (dígitos de 0 a 9). Esses números podem ser retirados de sequências numéricas, como a série de fibonacci, por exemplo, ou qualquer outra sequência pseudo-randômica. Fora a chave de criptografia e o número de deslocamentos não há nenhuma outra diferença entre as cifras.

- **Cifra de auto-chave** - Como já dito anteriormente, a cifra de vigenère é erradamente atribuída a Blaise de Vigenère. Porém, ele de fato criou uma cifra: a cifra de auto-chave, em 1586. Se a chave de criptografia for curta, palavras comuns acabam tendo a mesma substituição alfabética, que facilita a análise de frequência de letras. A cifra de auto-chave resolve esse problema utilizando a própria mensagem como chave. Quando a chave alfabética acaba (que indica a escolha inicial do alfabeto cifrante), a chave se torna a i-ésima (a partir da primeira letra) letra da mensagem original, escolhendo assim os alfabetos subsequentes. Ocorre esse processo até a última letra da mensagem original ser cifrada.

- **Cifra de Vernam (One-time pad)** - Inventada por Gilbert Vernam em 1917. Essa cifra usa como chave um texto que nunca é repetido e que possui o mesmo tamanho da mensagem. É a única cifra verdadeiramente indecifrável conhecida. Existe ainda a One-time pad (OTP), que é baseada na cifra de Vernam. A diferença é que no One-time pad, a cifragem é aplicada bit a bit (através da operação XOR (ou exclusivo)), ao invés de usar caracteres.

BIBLIOGRAFIA

<https://pt.wikipedia.org/wiki/Cifra_de_Vigen%C3%A8re>
<https://en.wikipedia.org/wiki/Vigen%C3%A8re_cipher>
<http://br.ccm.net/contents/145-a-codificacao-com-a-cifra-de-vigenere>
<http://www.numaboa.net.br/criptografia/substituicoes/polialfabeticas/506-vigenere>
<http://mcsesolution.com.br/blog/2010/08/31/a-matematica-da-cifra-de-vigenere/>
<https://www.kaspersky.com.br/blog/vigenere-cipher-history/5688/>
<http://www.wikiwand.com/pt/Hist%C3%B3ria_da_criptografia#/>
<http://practicalcryptography.com/ciphers/vigenere-gronsfeld-and-autokey-cipher/>
<https://joseluistabaracarbajo.gitbooks.io/criptografia-clasica/content/Cripto12.html>
<http://www.numaboa.net.br/criptografia/substituicoes/polialfabeticas/803-gronsfeld>
<http://www.numaboa.net.br/criptografia/substituicoes/polialfabeticas/342-Substituicoes-polialfabeticas>
<http://www.w3ii.com/pt/cryptography/traditional_ciphers.html>
<http://wikipedia.qwika.com/en2pt/Substitution_cipher>
<https://shellterlabs.com/pt/training/get-started/understanding-crypto/>
<http://www.cryptomuseum.com/crypto/vernam.htm>
<https://pt.wikipedia.org/wiki/One-time_pad>
<http://www.pro-technix.com/information/crypto/pages/vernam_base.html>
<https://en.wikipedia.org/wiki/Sequential_pattern_mining>
<http://wiki.di.uminho.pt/twiki/pub/Education/Criptografia/CriptografiaMestrados0607/Relatorio-Joel.pdf>

- Links acessados em 16/09/2017