Criptografia: introdução

André Gustavo dos Santos[1]

[1] Departamento de Informática
Universidade Federal de Viçosa

INF230 - 2021/1

UFV

Conteúdo

1 Criptografia clássica

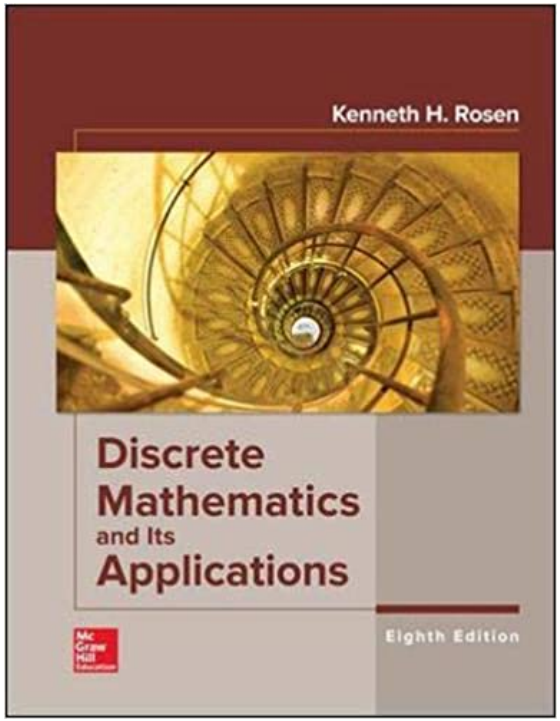
2 Criptoanálise

3 Outras cifras de deslocamento

4 Cifras de bloco

UFV

Os slides seguintes são baseados nas seções 4.6.1 e 4.6.2 do livro texto da disciplina:



ROSEN, Kenneth H.
Discrete mathematics and its applications.
McGraw-Hill Education, 8th edition, 2018

UFV

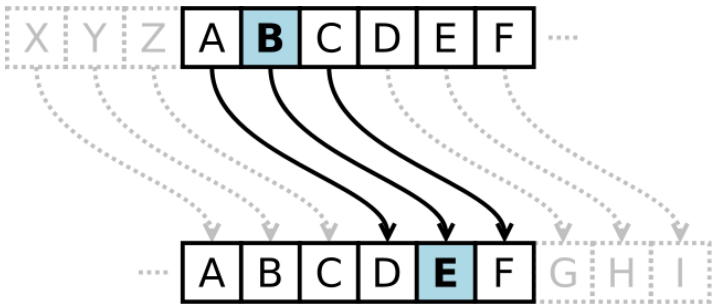
Introdução

- Criptografia é o processo de transformar informação de forma que não possa ser recuperada facilmente sem algum conhecimento especial
- Teoria de números tem um papel fundamental na criptografia
- É a base de muitos sistemas clássicos de criptografia, usados há milhares de anos até o século XX
- Também é base dos sistemas de criptografia moderna
- A seguir são descritos alguns sistemas clássicos de criptografia, como cifra por deslocamento e por transposição

UFV

Criptografia clássica

- Um dos usos mais antigos de criptografia que se tem notícia foi usado por Júlio César
- Ele criava mensagens secretas substituindo cada letra pela letra 3 posições à frente no alfabeto (sendo as 3 últimas substituídas pelas 3 primeiras)
  - Por exemplo, B virava E e X virava A
  - Este é um exemplo de codificação (criptografia), tornar uma mensagem secreta



Fonte: wikipedia

- Matematicamente, podemos representar este processo no conjunto $\mathbb{Z}_{26}$
  - As letras de A a Z são representadas por números de 0 a 25 (ex.: A é 0, J é 9)
  - A codificação é uma função em $\mathbb{Z}_{26}$: $f(p) = (p + 3) \bmod 26$
  - Ou seja, na mensagem criptografada, uma letra representada pelo inteiro p é substituída pela letra representada por $(p + 3) \bmod 26$

Criptografia clássica

Que mensagem secreta é criada pela cifra de César para "MEET YOU IN THE PARK"?

- Substituir cada letra pelo número correspondente
  MEET YOU IN THE PARK $\Rightarrow$ 12 4 4 19 24 14 20 8 13 19 7 4 15 0 17 10
- Substituir cada número p por $f(p) = (p + 3) \bmod 26$
  15 7 7 22 1 17 23 11 16 22 10 7 18 3 20 13
- Substituir cada número pela letra correspondente
  PHHW BRX LQ WKH SDUN

- Para recuperar a mensagem original devemos usar a função f^{-1}, a inversa de f
- Note que a inversa de f é dada por $f^{-1}(p) = (p - 3) \bmod 26$
- Ou seja, cada letra deve ser substituída pela letra 3 posições atrás no alfabeto
- O processo de determinar a mensagem original a partir da mensagem secreta é chamado decodificação (descriptografia)

UFV

Cifra de deslocamento

- Há várias formas de generalizar a cifra de César
- Por exemplo, em vez de deslocar os códigos 3 posições podemos deslocar k
  $f(p) = (p + k) \text{mod } 26$
- Este método é chamado cifra de deslocamento (*shift cipher*)
- A decodificação pode ser feita com a função inversa
  $f(p) = (p - k) \text{mod } 26$
- O valor k é chamado de chave

UFV

Cifra de deslocamento

Codificação de “STOP GLOBAL WARMING” pela cifra de deslocamento com $k = 11$

- Substituir cada letra pelo número correspondente
  STOP GLOBAL WARMING $\Rightarrow$ 18 19 14 15 6 11 14 1 0 11 22 0 17 12 8 13 6
- Substituir cada número p por $f(p) = (p + 11) \bmod 26$
  3 4 25 0 17 22 25 12 11 22 7 11 2 23 19 24 17
- Substituir cada número pela letra correspondente
  DEZA RWZMLW HLCXTYR

Decodificação de “LEWLYPLUJL PZ H NYLHA ALHJOLY”, codificada com $k = 7$

- Substituir cada letra pelo número correspondente
  LEWLYPLUJL PZ H NYLHA ALHJOLY
  $\Rightarrow$ 11 4 22 11 24 15 11 20 9 11 15 25 7 13 24 11 7 0 0 11 7 9 14 11 24
- Substituir cada número p por $f^{-1}(p) = (p - 11) \bmod 26$
  4 23 15 4 17 8 4 13 2 4 8 18 0 6 17 4 0 19 19 4 0 2 7 4 17
- Substituir cada número pela letra correspondente
  EXPERIENCE IS A GREAT TEACHER

UFV

Criptoanálise

- Criptoanálise é a arte de decodificar uma mensagem secreta sem conhecer a chave (e/ou o método) que a gerou
- É um processo bem difícil e trabalhoso, especialmente quando não se conhece o método usado na codificação da mensagem
- A cifra de deslocamento, no entanto, é facilmente quebrada quando se sabe que foi este o método usado
- Note que existem 26 chaves possíveis, se incluirmos $k = 0$ (sem codificação)
- Basta tentar cada uma e ver qual das decodificações produz um texto com sentido
- Mais que isso, é possível usar análise de frequência
  - Letras mais frequentes em textos em inglês:
    E 13%, T 9%, A 8%, O 8%, I 7%, N 7%, S 7%, H 6%, R 6%
  - A letra mais comum no texto cifrado é, provavelmente, a codificação da letra E
  - A diferença dos códigos entre a letra mais frequente e a letra E produz uma chave k
  - Se a mensagem decodificada com $-k$ fizer sentido, assumimos tê-la decifrado
  - Senão, fazemos o mesmo processo supondo que a mais frequente é a codificação de T
  - Se ainda não fizer sentido, repetimos o processo, da mais comum para a menos comum

UFV

Criptoanálise

Interceptamos a mensagem secreta “ZNK KGXRE HOXJ MKZY ZNK CUXS”

Se sabemos ter sido codificada por deslocamento, tentamos análise de frequência

- A letra mais frequente na mensagem é K
- Provavelmente é o código da letra E, a mais frequente em inglês
- Se é esse caso, temos que $10 = 4 + k.$ mod 26, então $k = 6$
- Deslocando as letras com $k = -6$ encontramos a mensagem
  THE EARLY BIRD GETS THE WORM
- Como é um texto com sentido, assumimos que descobrimos a chave correta

UFV

Cifra de Vigenère[1]

- Cifra de deslocamento baseada em múltiplas cifras de César
- No lugar de uma chave k, usa uma chave múltipla, por exemplo uma palavra

Codificar "ATTACK AT DAWN" com a chave "LEMON"

- Repetir a senha tantas vezes quanto necessário até o tamanho da mensagem
- Substituir p_i por $c_i = (p_i + k_i) \bmod 26$
- `ATTACKATDOWN` *(mensagem original)*
  `LEMONLEMONLE` *(chave)*
  `LXFOPVEFRNHR` *(mensagem cifrada)*
- Nesse caso:
  A + L = 0 + 11 = 11 = L
  T + E = 19 + 4 = 23 = X
  ...

- Para fazer criptoanálise por frequência, é necessário saber o tamanho da chave

[1] O método foi erroneamente atribuído a Blaise de Vigenère; foi primeiramente descrito por Giovan Battista Bellaso em 1553 e usado por Leon Battista Alberti bem antes, por volta de 1467.

UFV

Cifra afim

- Outra forma de tentar aumentar a segurança da cifra de deslocamento é usar:

$$f(p) = (ap + b) \bmod 26$$

  em que a e b são inteiros escolhidos de tal forma que f seja bijetora.
- Esta função é uma bijeção se e somente se $\text{mdc}(a, 26) = 1$
- Tal função é chamada de transformação afim

A letra K é representada por que letra na codificação por $f(p) = (7p + 3) \bmod 26$?

- K é representada pelo número 10
- $f(10) = (7 \cdot 10 + 3) \bmod 26 = 21$
- 21 representa a letra V, então K é substituída por V

Cifra afim

- Como decodificar uma mensagem codificada por cifra afim?
- Seja $c = (ap + b) \text{mod } 26$, com $\text{mdc}(a, 26) = 1$, a codificação de p
- Devemos resolver a congruência $c \equiv (ap + b) \text{mod } 26$ para p
- Como $\text{mdc}(a, 26) = 1$, sabemos que a possui um inverso $\bar{a}$ modulo 26
- Então:

$$
\begin{aligned}
c &= (ap + b) \text{mod } 26 \\
c - b &= (ap + b - b) \text{mod } 26 \\
c - b &= ap \text{ mod } 26 \\
\bar{a}(c - b) &= \bar{a}ap \text{ mod } 26 \\
\bar{a}(c - b) &= p \text{ mod } 26 \\
p &= \bar{a}(c - b) \text{ mod } 26
\end{aligned}
$$

UFV

Cifras de bloco

- Cifras de deslocamento e afim são cifras por caractere, pois cada letra é substituída por alguma outra letra
- São cifras monoalfabéticas, enquanto a cifra de Vigenère é polialfabética
- Uma forma mais segura é substituir blocos de caracteres por blocos de caracteres
- Um exemplo é a cifra por transposição, em que a senha é uma permutação σ de $\{1, 2, \ldots, m\}$ para algum inteiro m
- Primeiramente a mensagem é dividida em blocos de tamanho m (se o tamanho da mensagem não for divisível por m, é completada com letras aleatórias no final)
- Um bloco $p_1 p_2 \ldots p_m$ é criptografado como $c_1 c_2 \ldots c_m = p_{\sigma(1)} p_{\sigma(2)} \ldots p_{\sigma(m)}$

UFV

Cifras de bloco

- Os exemplos a seguir como chave a permutação σ de $\{1, 2, 3, 4\}$ em que $\sigma(1) = 3, \sigma(2) = 1, \sigma(3) = 4, \sigma(4) = 2$

Codificar “PIRATE ATTACK” com a permutação σ acima

- A mensagem é dividida em blocos de 4 letras: PIRA TEAT TACK
- Em cada bloco, a 1ª letra vai para a 3ª posição, a 2ª letra para a 1ª posição, etc
- A mensagem codificada é então: IAPR ETTA AKTC

Decodificar “SWUETRAEOEHS” com a permutação σ acima

- A mensagem é dividida em blocos de 4 letras: SWUE TRAE OEHS
- A permutação inversa σ^{-1} envia 1 para 2, 2 para 4, 3 para 1 e 4 para 3
- A mensagem decodificada é então: USEW ATER HOSE, lida USE WATER HOSE

UFV