Criptografia
- Cifras clássicas

Prof. Dr. Avelino F. Zorzo
Escola Politécnica - PUCRS

Propriedades

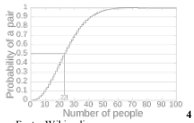
- Duas propriedades de uma cifra segura
  - Confusão
    - » Substituição
    - » Objetivo: fazer a relação entre a chave e o texto cifrado o mais complexa
  - Difusão
    - » Transposição
    - » Objetivo: re-arranjar os bits na mensagem de forma que redundância no texto claro se espalhe no texto cifrado

Revisão: quatro tipos de ataques

1. Ataque **só com texto cifrado** (***Ciphertext-only***)
   - Enigma: elos e correntes
2. Ataque **com texto claro conhecido** (***Known-plaintext***)
   - Enigma: *cribs*
3. Ataque **com texto claro escolhido** (***Chosen-plaintext***)
   - Quebra de código na batalha de Midway (II GM)
4. Ataque **com texto cifrado escolhido** (***Chosen-ciphertext***)
   - O objetivo é deduzir a chave (ataque na hora do almoço)

Outros tipos de ataques

- Encontro no meio do caminho
  - 2^{2k} operações é reduzido para 2^k operações mais 2^k armazenamentos
- Paradoxo do anivesário
  - Probabilidade que 2 pessoas em uma sala com 23 pessoas tenham o mesmo aniversário é de aproximadamente 0.507
  - $P_2(m,n) = 1 - e^{\wedge}(-n^2/2.m)$
  - $P_2(365, 23) = 0.507$
  - $P_2(365, 30) = 0.706$



Fonte: Wikipedia

Outros tipos de ataques

- Ataque por força bruta
- Ataque do homem no meio do caminho
- Criptoanálise diferencial
- Análise de frequência
- …

Aritmética modular

- Definição:
  - Suponha a e b inteiros, e m um inteiro positivo. Então escrevemos $a \equiv b$ se m divide a-b.
  - $a \equiv b \pmod m$ é chamado de congruência
  - m é chamado de módulo.
- Exemplo:

$2 \equiv 11 \pmod 3$ pois $2 \bmod 3 = 11 \bmod 3 = 2$
$12 \equiv 19 \pmod 7$ pois $12 \bmod 7 = 19 \bmod 7 = 5$

Aritmética modular

- Aritmética módulo m é definido como:
  - $Z_n = \{0, \ldots, n-1\}$
  - Duas operações: + e x
    - » Funciona como adição e multiplicação
    - » Resultados são reduzidos ao módulo n
    - » $a,b \in Z_n$, $a+b \in Z_n$
    - » $a,b \in Z_n$, $axb \in Z_n$
- Satisfaz a maioria das regras aritméticas conhecidas, *e.g.* adição é fechada, multiplicação é comutativa, etc.

Criptografia clássica

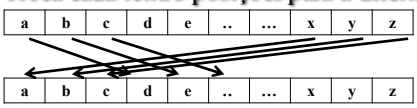
- Criptografia clássica
  - Baseada em caracteres (humanos)
- Criptografia moderna
  - Baseada em entradas binárias (computadores)
- O que mudou?
  - 26 elementos para 2 elementos.
  - Mas, a filosofia permanece basicamente a mesma

Cifras de substituição

1. Cifras monoalfabéticas
   - Exemplo: a → b, b → c
2. Cifra polialfabéticas
   - Composta por diversas cifras monoalfabéticas
3. Cifras de substituição monofônicas
   - Exemplo: a → f ou 5 ou 8 ou 9
4. Cifra de substituição com poligramas
   - Substitui grupos de letras, AB → PQ

Cifra de César – cifra de troca

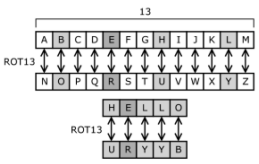
- Nomeada em homenagem a Júlio César
- Usada por centenas de anos
- Troca cada letra 3 posições para a direita



- Exemplo
  - atacaraoamanhecer → dwdfdudrdpdqkhfhu

ROT13 – cifra de troca

- Outro exemplo de cifra monoalfabética
- Comum encontrar em sistemas Unix
- Cada letra é rotacionada 13 posições
- Pergunta
  - Por que não ROT14?



(Fon te: Wik ipedia .org)

Criptoanálise - cifra de troca

- Ataque de força bruta
  - Dado JBCRCLQRWCRVNBJENBWRWN, é possível encontrar o texto claro?

```
jbcrclqrwcrvnbjenbwrwn      (K=0)
iabqbkpqvbqumaidmavqvm      (K=1)
hzapajopuaptlzhclzupul      (K=2)
gyzozinotzoskygbkytotk      (K=3)
fxynyhmnsynrjxfajxsnsj      (K=4)
ewxmxglmrxmqiweziwrmri      (K=5)
dvwlwfklqwlphvdyhvqlqh      (K=6)
cuvkvejkpvkogucxgupkpg      (K=7)
btujudijoujnftbwftojof      (K=8)
astitchintimesavesnine      (K=9)
```

Qual o problema?

- Cifras de troca (módulo 26) não são seguras, pois podem ser quebradas por **pesquisa exaustivas**
- Existem somente 26 chaves possíveis
- Na média, o texto claro pode ser descoberto depois de somente 26/2=13 tentativas.
- ***Lição***: para uma cifra ser segura, a quantia de chaves deve ser grande
- Mas o contrário é verdadeiro?

Cifras de substituição

- Número de símbolos no alfabeto = q
  - q! diferentes chaves
  - Alfabeto português = 26 letras
  - 26! aproximadamente 4×10^{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

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

- Ataque de análise de frequência

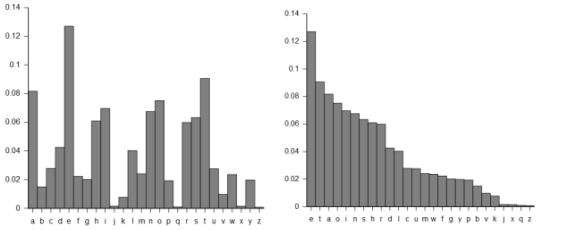
Criptoanálise – cifra de substituição

- Frequência das letras no Inglês:

letra	prob	letra	prob	letra	prob
A	.082	J	.002	S	.063
B	.015	K	.008	T	.091
C	.028	L	.040	U	.028
D	.043	M	.024	V	.010
E	.127	N	.067	W	.023
F	.022	O	.075	X	.001
G	.020	P	.019	Y	.020
H	.061	Q	.001	Z	.001
I	.070	R	.060		

  - Sequência de 2 letras mais comuns: TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI, OF
  - Sequência de 3 letras mais comuns: THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH

Criptoanálise – cifra de substituição



(a) Frequência das letras em Inglês **(b) Frequência das letras em Inglês ordenado**

(Fonte: wikipedia)

Criptoanálise – cifra de substituição

- Ataque só com o texto cifrado
- Texto cifrado interceptado:

BTLDXFETMDGLGMVMYFQEMQAPMVBZQMXZQEGZVXFTL
XGUWFVXBFWDYUXUQFQXUBGQZBMYMBBFHQXFPXGU
VHISUBXZVCMGQVXGUBFAUITUMCUTVXGZVIFFCXTMBUV
BTLDXFETMDGLPTFWZXVZQZXZMYMQAYZWZXUAHVUIL
XGUUELDXZMQVVFWUPFHTXGFHVMQALUMTVMEFXFXGU
XKUQXZUXGBUQXHTLKGUTUZXDYMLUAMBTHBZMYTFYU
ZQXGUFHXBFWUFPIFXGKFTYAKMTVBFWDYUXUAZQ
QZQUXUUQVZJXLXGTUUXGUIFFCBFOUTVXGFVUMVDUBXV
FPXGUGZVXFTLKGZBGKUTUWFVXVZEQZPZBMQXXFXGU
AUOUYFDWUQXFPXGUVHISUBX

Criptoanálise – cifra de substituição

- Frequência das letras no texto cifrado:

letra	prob	letra	prob	letra	prob
A	.023	J	.003	S	.005
B	.054	K	.015	T	.054
C	.010	L	.030	U	.120
D	.026	M	.061	V	.066
E	.018	N	.000	W	.023
F	.090	O	.005	X	.118
G	.066	P	.020	Y	.028
H	.023	Q	.059	Z	.064
I	.018	R	.000		

  - Sequência de 2 letras mais comuns:
    - » XG (16), GU (11), XF (8), QX (7), VX (7), BF (6), UX (6), ZQ (6)
  - Sequência de 3 letras mais comuns:
    - » XGU (10), BFW, FPX, FXG, GZV, LDX, LXG, MQA, PXG, UBX, UQX, UXU, VXG – (all 3 times)

Criptoanálise – cifra de substituição

- Assumindo que:
  - U e X são as letras mais frequentes no texto cifrado, logo assumir que elas sejam E eT texto claro.
- Sequência de 2 letras mais comum: XG.
- Sequência de 3 letras mais comum: XGU
  - Assumir que X=T e G=H.
  - THE é a sequência de 3 letras mais comum no Inglês: U=E.
- XF é uma sequência de 2 letras comum.
  - Anteriormente X = T. XF é TO ou TI. O é um pouco mais frequente que I em Inglês, assim F = O.

Criptoanálise – cifra de substituição

- Texto cifrado atual

BTLDtoETMDhLhMVMYo QEMQAPMVBZQMtZQEhZ VtoTL
theWoVtBoWDYeteQoQteBhQZBMY MBBoHQtoPthe
VHISeBtZVCMhQVtheBo AeITeMCeT VthZVIoCtTMBe V
BTLDtoETMDhLPToWZtVZQZtZMYMQA YZWZteA HVeIL
theeELDtZMQVVoWePoHTthoHV MQALeMT VMEototthe
tKeQtZethBeQtHTLKheTeZtDYMLeAMBTHBZ MYToYe
ZQtheoHtBoWeFPIothKoTYAK MTVBoW DYeteAZ Q
QZQeteeQVZJtLthTeetheIooCBFOeTVthFVeMVDeBtV
oPthehZVtoTLKhZBhKeTeWoVtVZEQZPZB MQttothe
AeOeYoDWeQtoPtheVHISeBt

- X = T, G = H, U = E, F = O
  - Após mais alguma análise:
    - » QX (AT, IT, NT) e UQX → Q = N
    - » MQA = AND

Criptoanálise – cifra de substituição

- Texto claro original (com espaços):

cryptography has a long and fascinating history
the most complete nontechnical account of the
subject is kahns the codebreakers this book traces
cryptography from its initial and limited use by
the egyptians some four thousand years ago to the
twentieth century where it played a crucial role
in the outcome of both world wars completed in
nineteen sixty three the book covers those aspects
of the history which were most significant to the
development of the subject

(Fonte: Handbook of Applied Cryptography, A. Menezes, P. van Oorschot, e S. Vanstone)

Qual o problema?

- Uma grande quantidade de possibilidades de chaves não é suficiente para uma cifra ser segura.
- Substituição só provê confusão.
- ***Lição***: uma cifra segura deve combinar confusão e difusão.

Cifra de Vigenère

- Uma cifra polialfabética baseada na ideia de combinar algumas Cifras de César em uma única cifra.
- Nomeada em homenagem a Blaise De Vigenère, um diplomata francês em 1586

k = A L V A R O A L V A R O A L V A
m = B E R E A D Y A C K A T D A W N

c = B P M E R R Y L X K R H D L R N

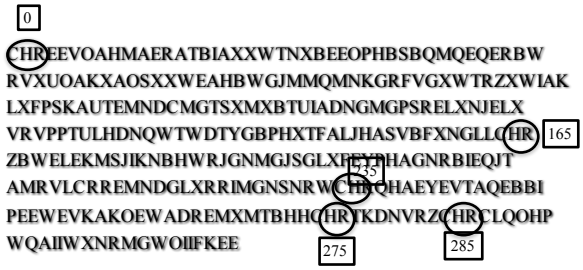
Criptoanálise – Cifra de Vigenère

- Dois passos para a criptoanálise
  1. Encontrar o tamanho m da chave
  2. Encontrar cada letra da chave

Tamanho da chave

- Primeiro método: Teste de Kasiski
  - Descrito por Friedrich Kasiski em1863
  - Procurar por segmentos idênticos e contar quantas posições os separam.

Examplo: Kasiski



Tamanho da chave

- Segundo método: índice de coincidência
  - Descrito por William Friedman
  - Suponha $x = x_1x_2..x_n$ é uma sequência de n caracteres do alfabeto
  - O **índice de coincidência** de **x** é definido como a probabilidade de dois elementos aleatórios de x serem idênticos.

Índice de coincidência

- Suponha uma sequência de n letras em Inglês
  - Frequência de $a = f_0$
  - Frequência de $b = f_1$ ...
  - Frequência de $z = f_{25}$
  - Assim o índice de coincidência é:

$$I_c(x) = \frac{\sum_{i=0}^{25}\binom{f_i}{2}}{\binom{n}{2}} = \frac{\sum_{i=0}^{25} f_i(f_i-1)}{n(n-1)} \approx \sum_{i=0}^{25} p_i^2$$

Índice de coincidência

- Índice de coincidência de uma frase em Inglês:

$$I_c(x) \approx \sum_{i=0}^{25} p_i^2 = 0.065$$

- O mesmo raciocínio pode ser aplicado se x é um texto cifrado obtido através de uma cifra monoalfabética.
- Re-escreva o texto cifrado c como :

$$\begin{array}{lcl} c_1 & = & c_1c_{m+1}c_{2m+1}\ \ldots \\ c_2 & = & c_2c_{m+2}c_{m+2}\ \ldots \\ & \ldots & \\ c_m & = & c_mc_{2m}c_{3m}\ \ldots \end{array}$$

- Se $c_1, c_2, \ldots, c_m$ são construídos de forma que m é o tamanho da chave, então cada $I_c(c_i)$ deveria ser de aproximadamente 0.065

Índice de coincidência

- Se m não for o tamanho da chave, as sequências c_i pareceriam aleatórias. Cadeias aleatórias tem:

$$I_c \approx 26\left(\frac{1}{26}\right)^2 = \frac{1}{26} = 0.038$$

- Exemplo: tabela com I_c para valores de m:

m	I_c
1	0.043
2	0.052; 0.051
3	0.05; 0.059; 0.045
4	0.049; 0.053; 0.052; 0.051
5	0.034; 0.05; 0.048; 0.038; 0.045
6	0.063; 0.07; 0.083; 0.062; 0.071; 0.048
7	0.033; 0.041; 0.038; 0.046; 0.041; 0.04; 0.047

- Este método mostra que $m = 6$.

Próximo passo

- Quebrar cada pedaço como na cifra de troca

```
C H R E E
V O A H M
A E R A T
B I A X X
W T N X B
E E O P H
B S B Q M
. .
```

Fraqueza de Vigenère

- Cifra de Vigenère envolve alguma transposição
- Entretanto, a transposição não distribui as informações de maneira aleatória no texto cifrado.
- ***Lição***: texto cifrado não deve ter qualquer padrão identificável.
- Em outras palavras, uma cifra segura deve produzir um texto cifrado que é indistinguível de um texto aleatório.

Outras formas de transposição

- Texto claro: "this sentence is secret"

```
t h i s s
e n t e n
c e i s s
e c r e t
```

- Se transforma em: tecehnecitirsesesnst

```
3 1 2 5 4
t h i s s
e n t e n
c e i s s
e c r e t
```

- E depois: hnecitirtecesnstsese

Primeiro trabalho

- Dado um texto cifrado encontre o texto claro
- Escreva um relatório em uma página explicando como foi feita a criptoanálise e parte do texto cifrado e do texto claro.