Trabalho Prático 1

Laboratório de Informática (LEI)

2006/2007

Instruções

Neste documento enuncia-se o primeiro trabalho prático, com um peso de 40%. O trabalho deverá ser realizado por grupos formados por dois ou três elementos.

A data de entrega será anunciada na página da disciplina,
`http://wiki.di.uminho.pt/wiki/bin/view/Education/LaboratoriosInformatica1`

Juntamente com o trabalho deverá ser entregue um relatório sucinto em LaTeX onde deve constar:

- identificação dos elementos do grupo
- resumo do trabalho efectuado

1 Introdução

As cifras de substituição são um conjunto de cifras clássicas onde pedaços de texto em claro são substituídos por texto de cifra correspondente, mantendo a sua posição. Para obter a mensagem em claro, basta inverter o processo.

Na cifra de César, por exemplo, escolhe-se um número n entre 1 e 26 e transporta-se cada letra n posições no alfabeto. Para $n = 3$, obtém-se

```
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
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 (3)
```

A mensagem "Ola mundo" passa a ser "RODPXQGR" (note-se que se retirou os espaços e passou-se todas as letras para maiúsculas)

A cifra de Vigenère, "le chiffre indéchiffrable", usa o mesmo princípio mas com vários alfabetos de cifra simultâneos. Qual dos alfabetos deve ser

utilizado para cada letra é determinado por uma palavra-chave previamente escolhida.

Por exemplo, se a palavra-chave escolhida for "GATO", os alfabetos a utilizar são:

```
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
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 (6)
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)
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 (19)
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 (14)
```

e a mensagem "o rato roeu a rolha da garrafa de rum do rei da russia" passa a ser "URTHURHSAAKCRHTRGGTFXAYOJEKISDHFKIWOXULGOA",

```
GATOGATOGATOGATOGATOGATOGATOGATOGATOGATOGATOGA
oratoroeuarolhadagarrafaderumdoreidarussia
URTHURHSAAKCRHTRGGTFXAYOJEKISDHFKIWOXULGOA
```

No fundo, é uma variante da cifra de César em que, em vez de se transportar cada letra um número fixo de posições, esse número é determinado pela palavra-chave (a posição no alfabeto de cada letra).

A segurança da cifra depende essencialmente de que haja uma informação que seja mantida secreta: *a chave*. No caso da cifra de César, a chave é o número de posições a transpor (no exemplo, 3). No caso da cifra de César, a chave é a palavra-chave escolhida (no exemplo, "GATO").

Uma nota final: actualmente estas cifras já não são seriamente usadas dada a facilidade com que normalmente são quebradas com técnicas simples de cripto-análise como a "análise de frequências" e o poder de cálculo dos computadores actuais. As tarefas finais neste trabalho mostram isso. No entanto, as famosas máquinas Enigma usadas na Segunda Guerra Mundial usavam cifras de substituição. E a "one-time-pad", uma cifra de substituição cuja chave (que tem de ser completamente aleatória) tem o tamanho da mensagem, apesar de impraticável, é, de facto, inquebrável.

2 Referências

Ver, por exemplo, os artigos no Wikipedia (http://en.wikipedia.org) sobre a "Caesar cipher" e "Vigenere cipher".

O "The Code Book" do Simon Singh introduz estas e outras cifras clássicas juntamente com uma interessante caminhada pela história da criptografia.

3 Tarefas

1. Implemente uma função que dada um número inteiro devolve uma string com a lista de digitos por extenso e a sua função inversa.

   ```
   *LI0607> ord2ext 1332795503
   UM TRES TRES DOIS SETE NOVE CINCO CINCO ZERO TRES
   *LI0607> ext2ord UM TRES TRES DOIS SETE
   13327
   ```

2. Implemente as funções de cifragem e decifragem correspondentes à cifra de César. A mensagem a cifrar deverá ser primeiro limpa de caracteres que não pertençam ao alfabeto comum e convertida para maiúsculas.

3. Implemente as funções de cifragem e decifragem correspondentes à cifra de Vigenère. Tal como a cifra de César, a mensagem a cifrar deverá ser primeiro limpa de caractéres que não pertençam ao alfabeto comum e convertida para maiúsculas.

4. Uma mensagem m foi cifrada usando a cifra de César dando origem à mensagem cifrada

   "RLWRGRLVGRLVCHURQRYHRLWRTXDWURXPCHURVHWH"

   Sabe-se, no entanto, que a mensagem original consiste numa listagem, por extenso, de dígitos (ver primeiro ponto). Pretende-se saber qual a mensagem original e a chave usada para cifrar.

   Para isso, implemente uma função que faça um ataque à força bruta, experimentando as chaves todas e procurando as palavras "UM", "DOIS", etc., no texto resultante da aplicação da função de decifragem com cada chave.

   Comente a vantagem de conhecer previamente parte do texto em claro.

   Opcional: Adicionalmente, implemente uma função que dada a mensagem decifrada contendo uma listagem de digitos (e que não contêm espaços), extraia o número correspondente.

   ```
   *LI0607> extrair "DOISZERODOISDOISZERO"
   20220
   ```

5. **(Pergunta de Valorização)** As duas mensagens seguintes foram cifradas usando a cifra de Vigenère:

"RJVIJIXJSIOEYHDSDSXMEWNEXZOMOO"
"AIKWSQTVBRQSUWIEUDAEFEFQTANFDE"

Tal como no ponto anterior, sabe-se que as mensagens originais consistem numa listagem, por extenso, de digitos. Implemente um ataque à força bruta e obtenha as mensagens originais (e chaves respectivas).

O ataque, principalmente para a segunda mensagem, deverá durar mais tempo a computar que no ponto anterior. Tente optimizar o tempo de computação, sem comprometer a clareza e generalidade do programa.

Nota: Se a chave tem o mesmo tamanho da mensagem e é verdadeiramente aleatória, tem-se um exemplo de um "one-time pad", que é inquebrável. A cifra de Vigenère usa geralmente uma chave consideravelmente mais curta que a mensagem.