Cifra Homofônica

Mestrado em Ciência da Computação
Estruturas de Dados
Prof. Dr. Paulo Roberto Gomes Luzzardi
Mestranda: Nelsi Warken

Sumário

1. Introdução

- **Criptologia** é a área de conhecimento em que estão reunidos os estudos da criptografia e da criptoanálise.

- **Criptografia** é o conjunto de princípios e técnicas empregadas para cifrar a escrita, de modo que apenas os que têm acesso às convenções combinadas possam lê-la. A **mensagem é modificada** por uma **função de codificação** com o auxílio de um **valor especial ou chave**, e transformada assim no **criptograma.**

  Para que o receptor possa ler a mensagem, ele modifica o criptograma com uma chave e uma função de decodificação, obtendo novamente o texto original.

- **Criptoanálise** ou **Criptanálise** é o conjunto de **técnicas e métodos** para **decifrar** uma escrita de **sistema desconhecido**. Termo 'decifrar' sendo usado com o significado de descobrir a mensagem original de um criptograma sem possuir a chave de decodificação (usuário não legítimo).

2. História da Criptologia

Durante muitos séculos, a criptografia foi tratada como uma arte. Seu nome vem das palavras gregas *kryptós* (oculto, **secreto**, obscuro, ininteligível) e *graphein* (**escrita**). Os romanos já utilizavam tais conhecimentos para **guerras** e **segredos de estado**.

A "arte de cifrar mensagens" tem origens muito mais **antigas**. Qualquer escrita desconhecida pode ser considerada uma cifra.

Os métodos criptográficos desenvolvidos na **antiguidade** eram baseados essencialmente em técnicas de **substituição** e **transposição** simples, já que o uso de **cálculos** matemáticos **complexos** era pouco prático. Uma das mais antigas conhecidas: 400 anos antes de Cristo.

2. História da Criptografia - Aplicações

487 a.C. – **Militares**

- Segredos Nacionais;
- Planos, datas, tropas;
- Estratégias.

1960... – avanço das **telecomunicações** – aplicações civis

- Empresas, informação interna;
- Troca de Informações entre instituições.

1969... – **Internet** – globalização da Informação e da criptografia

- Aplicações pessoais, marketing, vendas;
- Comunicações entre computadores:
- Correio eletrônico, redes sem fios.
- Segurança da informação: **confidenciabilidade**, **integridade** e **autenticidade**.

3. Tipos de Cifras

- Cifras de Transposição
- Cifras de Substituição
  - Substituições Monoalfabéticas
    - Monográfica ou monogrâmica
    - Poligráfica ou Poligrâmica
    - Tomográfica ou Tomogrâmica
      - **Homofônica**
  - Substituições Polialfabéticas

3. Tipos de Cifras

- **Cifras de transposição**: misturam as letras do texto original de acordo com uma regra reversível qualquer. O texto cifrado é obtido através da **permutação** do **texto original**.

- **Cifras de substituição**: produzem criptogramas nos quais **as letras** do texto original, tratadas individualmente ou em grupos de comprimento constante, **são substituídas por outras letras, figuras, símbolos** ou uma combinação destes de acordo com um sistema pré-definido.

  As **tabelas de substituição** contém os caracteres que serão substituídos e os caracteres de substituição. Estas tabelas também são conhecidas como **cifrantes** ou **alfabetos cifrantes**.

3. Tipos de Cifras (Substituição)

Substituição Monoalfabética: quando **um cifrante** ou **alfabeto** é aplicado.

Substituição Polialfabética: quando **mais de um cifrante** é utilizado para cifrar um texto claro.

3. Tipos de Cifras (Monoalfabéticas)

Substituição Monogrâmica ou Monográfica: **um dos caracteres do texto claro** é substituído por um **outro**, cada caracter é tratado individualmente.

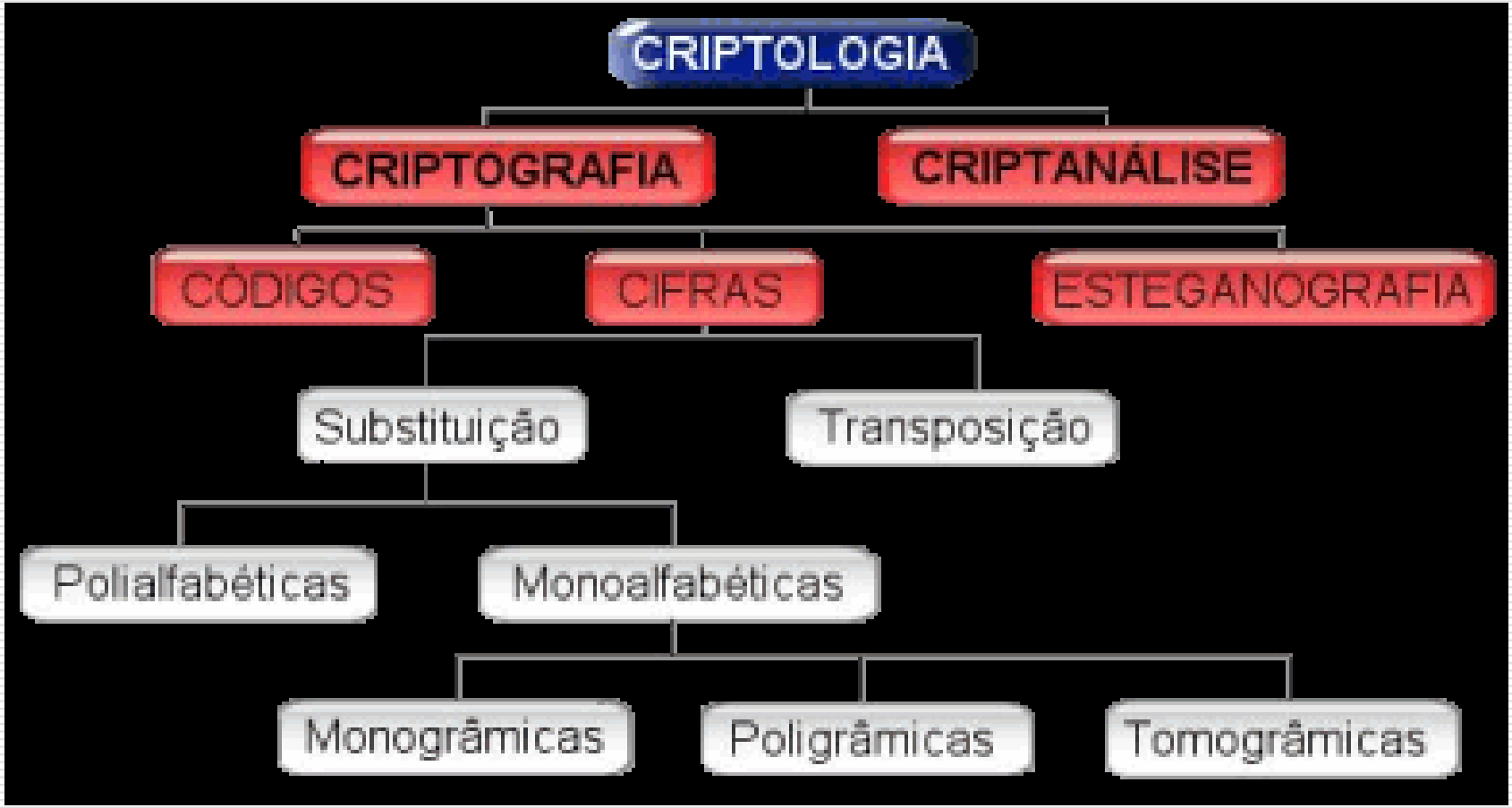
Na criptografia contemporânea, que usa computadores, substitui-se blocos de bits ao invés de caracteres. O princípio, porém, é o mesmo.

Substituição Poligrâmica: substituir **grupos de caracteres do texto original** por **um ou mais caracteres**. O comprimento da mensagem cifrada, geralmente, não é o mesmo da mensagem original.

Substituição Tomogrâmica: cada **um dos caracteres** da mensagem clara é substituído por **dois ou mais símbolos**.

Todas as classificações acima são **Substituições Monoalfabéticas**.

3. Tipos de Cifras



3. Tipos de Cifras (Tomogrâmica)

Substituição Homofônica (Monofônica): substituição monoalfabética tomogrâmica (grego - mesmo som). Sequências diferentes de letras que são pronunciadas de forma semelhante. Na criptologia, é uma cifra que **substitui cada um dos caracteres do texto claro por um de vários símbolos possíveis, todos com o mesmo significado.**

4. Cifra Homofônica

Na substituição homofônica, os símbolos de substituição para a letra A são **homófonos**: todos são pronunciados como A.

Convertem cada **caractere** em outro qualquer de um conjunto específico de caracteres. Assim, a letra '**a**' pode ser codificada nos valores '**29**', '**35**', '**82**' ou '**87**', por exemplo, a critério do **remetente**. No criptograma resultante, qualquer um desses valores será traduzido na letra 'a'. Utilizar sempre o **mesmo numero** de simbolos para a substituição.

Esta técnica foi utilizada por Antoine e Bonaventure Rossignol, no século XVII.

.

4.Cifra Homofônica

- Em 1401, Simeone de Crema usou uma **chave** na qual cada **vogal** do texto original possuía **vários equivalentes**.
- Em 1595, Henrique IV, rei da Inglaterra, utilizava uma cifra **homofônica** particular para os **assuntos sigilosos**. Um pouco mais elaborada, além das vogais havia outras **letras frequentes** com **mais de um substituto**.
- Em 1628, Luis XIII usava uma cifra homofônica **própria**. Na mesma época, a correspondência de **assuntos estrangeiros** entre Constantinopla e a França também era cifrada e possuía um método próprio.

5. Implementação

- **Função Codificar**

  Substituir um **caracter lido**, alfabético (“a” a “z”) por uma **sequencia de três digitos**, chamada “**chave**”. Cada letra pode ter de 1 a 14 chaves.

  As chaves estão colocadas numa **matriz de 26 linhas e 15 colunas**. A primeira coluna de cada letra indica o numero de chaves existentes para esta letra a ser codificada. A escolha da chave é feita **randomicamente**.

5. Implementação - Codificação

- Podemos **substituir** as **letras** do alfabeto por **números**, atribuindo às letras de maior frequência uma quantidade maior de chaves.
- Pode ser feito um **estudo de frequência** de letras em textos.
- Caracter **não alfabético** do texto original fica **igual** na saída criptografada.
- Letras **maiusculas** são transformadas em **minusculas**, antes da substituição.

5. Implementação – Matriz Substituição

```
/* a */ 14, 9, 37, 88, 89, 125, 281, 310, 442, 563, 629, 795, 870, 986, 911,
/* b */ 4, 11, 14, 59, 70, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* c */ 5, 25, 31, 55, 61, 90, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* d */ 4, 26, 40, 69, 95, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* e */ 12, 5, 16, 44, 60, 76, 94, 99, 657, 520, 415, 333, 588, 0, 0,
/* f */ 3, 19, 24, 67, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* g */ 3, 777, 97, 28, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* h */ 4, 6, 15, 39, 68, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* i */ 10, 29, 33, 47, 50, 578, 819, 666, 418, 374, 930, 0, 0, 0, 0,
/* j */ 3, 12, 46, 62, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* k */ 1, 83, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* l */ 5, 2, 8, 23, 58, 526, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* m */ 8, 18, 21, 64, 75, 395, 555, 746, 815, 0, 0, 0, 0, 0, 0,
/* n */ 6, 10, 35, 77, 65, 111, 813, 0, 0, 0, 0, 0, 0, 0, 0,
/* o */ 11, 20, 27, 53, 73, 74, 80, 84, 98, 381, 549, 763, 0, 0, 0,
/* p */ 6, 32, 36, 71, 85, 111, 987, 0, 0, 0, 0, 0, 0, 0, 0,
/* q */ 4, 22, 51, 17, 48, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* r */ 6, 7, 38, 42, 63, 91, 92, 0, 0, 0, 0, 0, 0, 0, 0,
/* s */ 6, 34, 43, 54, 72, 87, 96, 0, 0, 0, 0, 0, 0, 0, 0,
/* t */ 4, 13, 56, 66, 78, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* u */ 7, 4, 45, 49, 82, 288, 435, 613, 0, 0, 0, 0, 0, 0, 0,
/* v */ 3, 1, 30, 79, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* w */ 3, 86, 100, 999, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* x */ 4, 3, 41, 648, 888, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* y */ 2, 93, 772, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* z */ 4, 52, 590, 724, 876, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 };
```

5. Implementação - Decodificação

- **Função Decodificar**

  A cada **3 digitos lidos** procurar na tabela a **chave**, formada por estes tres numeros. A busca em cada linha é limitada pelo numero de chaves da linha na matriz.

  Se algum caracter **não for digito**, ou se a **chave não for encontrada**, será devolvido o que foi digitado.

6. Exemplo

- Texto para criptografar:

  Cifra Homofonica de Rossignol

- Primeira decodificação

090819019042911 039074815084024020111050061629
026016 042549072034418097077098023

- Segunda decodificação

025418019063795 068381395098067027077047031310
095094 092027054087029097111763526

6. Exemplo

```
/* a */ 14, 9, 37, 88, 89, 125, 281, 310, 442, 563, 629, 795, 870, 986, 911,
/* b */ 4, 11, 14, 59, 70, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* c */ 5, 25, 31, 55, 61, 90, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* d */ 4, 26, 40, 69, 95, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* e */ 12, 5, 16, 44, 60, 76, 94, 99, 657, 520, 415, 333, 588, 0, 0,
/* f */ 3, 19, 24, 67, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* g */ 3, 777, 97, 28, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* h */ 4, 6, 15, 39, 68, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* i */ 10, 29, 33, 47, 50, 578, 819, 666, 418, 374, 930, 0, 0, 0, 0,
/* j */ 3, 12, 46, 62, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* k */ 1, 83, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* l */ 5, 2, 8, 23, 58, 526, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* m */ 8, 18, 21, 64, 75, 395, 555, 746, 815, 0, 0, 0, 0, 0, 0,
/* n */ 6, 10, 35, 77, 65, 111, 813, 0, 0, 0, 0, 0, 0, 0, 0,
/* o */ 11, 20, 27, 53, 73, 74, 80, 84, 98, 381, 549, 763, 0, 0, 0,
/* p */ 6, 32, 36, 71, 85, 111, 987, 0, 0, 0, 0, 0, 0, 0, 0,
/* q */ 4, 22, 51, 17, 48, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* r */ 6, 7, 38, 42, 63, 91, 92, 0, 0, 0, 0, 0, 0, 0, 0,
/* s */ 6, 34, 43, 54, 72, 87, 96, 0, 0, 0, 0, 0, 0, 0, 0,
/* t */ 4, 13, 56, 66, 78, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* u */ 7, 4, 45, 49, 82, 288, 435, 613, 0, 0, 0, 0, 0, 0, 0,
/* v */ 3, 1, 30, 79, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* w */ 3, 86, 100, 999, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* x */ 4, 3, 41, 648, 888, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* y */ 2, 93, 772, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
/* z */ 4, 52, 590, 724, 876, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0 };
```

Texto original:

Cifra Homofonica de Rossignol

Primeira decodificação:

090819019042911

039074815084024020111050061629

026016

042549072034418097077098023

Segunda decodicação

025418019063795

068381395098067027077047 03

310 095094

092027054087029097111763526

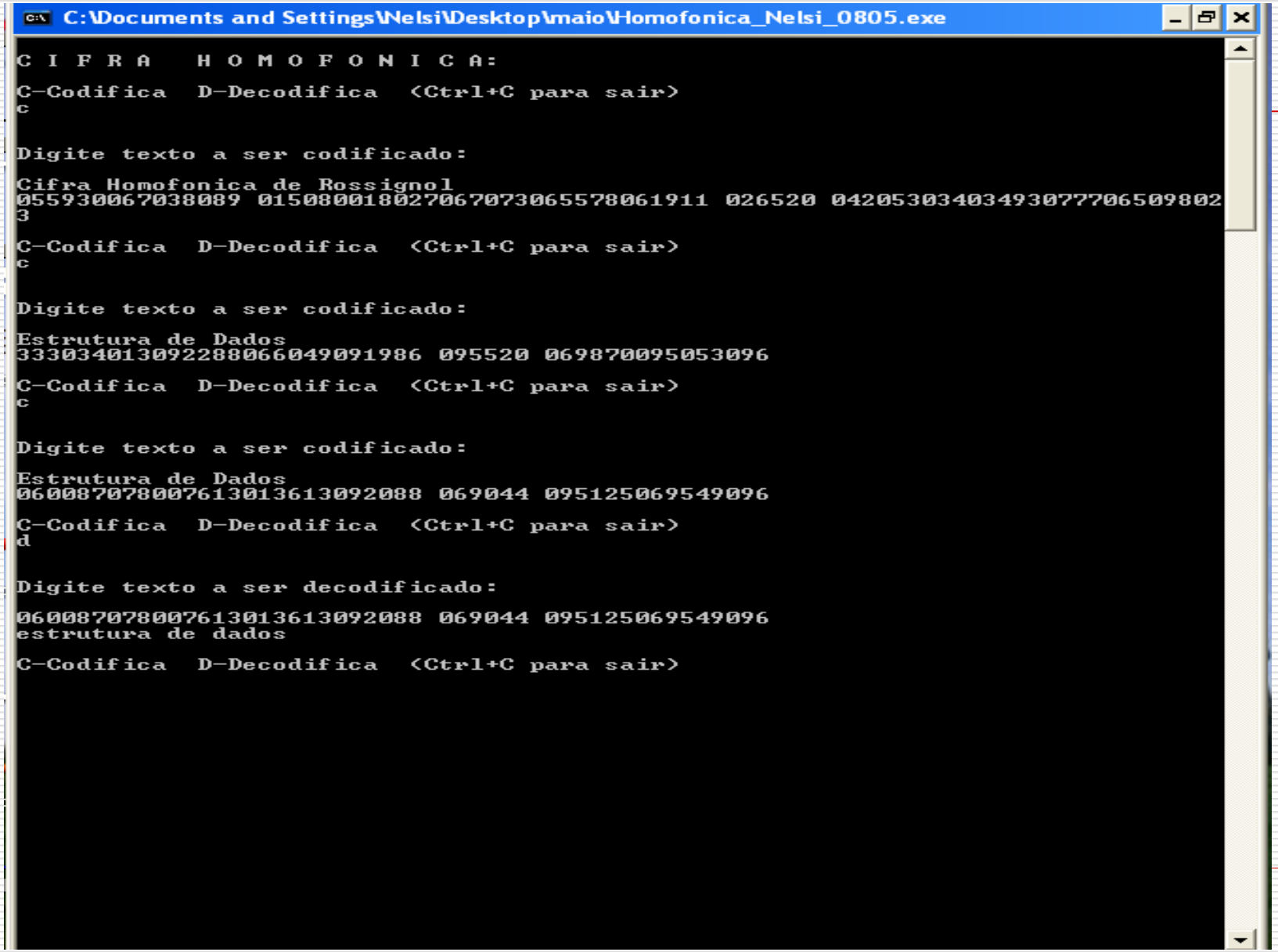
7. Código Fonte

```
printf( "\n\nC-Codifica  D-Decodifica  (Ctrl+C para sair)\n" );
      funcao = getc( stdin );
       fflush( stdin );
       if( funcao == 'C' || funcao == 'c' )    //  Codificar
       {
// O comando abaixo inicializa o gerador de números randômicos para não ser sempre o mesmo.
          srand( ( unsigned )time( &randomico ) );
          puts( "\n\nDigite texto a ser codificado:\n" );
          fflush( stdin );
          while( (entrada = tolower( getchar( ) ) ) != 10 )
          {
          if( isalpha( entrada ) )
          {
//            teste = (entrada - 'a');
// calculo da linha = letra lida - 'a' --->  c - a = 2 (terceira linha)
// calculo da chave = randomico de 1 ate o valor da coluna zero (numero de chaves)
//                  da respectiva linha.
// Exemplo: letra do texto = "i" --> linha 8, coluna 0 = 10. Sera procurada uma chave,
// randomicamente, entre as colunas 1 e 11.
             i = entrada - 'a';
             j = ( 1 + rand() % alpha[i][0]);
             entrada = alpha[i][j];
             printf( "%03d", entrada );
          }
          else
             printf( "%c", entrada );   //  caracter nao existe na tabela
          }
          }
```

7. Código Fonte

```
if ( funcao == 'D' || funcao == 'd')
        {       /* decodificar */
          puts( "\n\nDigite texto a ser decodificado: \n" );
          fflush( stdin );
          while( ( entrada = getchar( ) ) != 10 )
          if( isdigit( entrada ) )
          {
            entrada -= '0';   //
            if( digito == 3 )
            {
              codigo = dig1 * 100 + dig2 * 10 + entrada;
              for( i = 0 ; i < 26 && digito == 3 ; i++ )
                  for( j = 1 ; j <= alpha[i][0] && digito == 3 ; j++ )
                      if( codigo == alpha[ i ][ j ] )
                      {
                         digito = 1;
                         entrada = i + 'a';
                         putchar (entrada);
                      }
              if (digito == 3)
              {
                   printf( "%03d", codigo );
                   digito = 1;
              }
            }
            else
            {
              if (digito == 2)
              {
                  dig2 = entrada;
                  digito++;
              }
              else
              {
                  dig1 = entrada;
                  digito++;
              }
            }
          }
          else
            putchar( entrada );
          }
        }
    }
}
```

8. Execução


C:\Documents and Settings\Nelsi\Desktop\maio\Homofonica_Nelsi_0805.exe
C I F R A H O M O F O N I C A:
C-Codifica D-Decodifica (Ctrl+C para sair)
c
Digite texto a ser codificado:
Cifra Homofonica de Rossignol
055930067038089 015080018027067073065578061911 026520 042053034034930777065098023
C-Codifica D-Decodifica (Ctrl+C para sair)
c
Digite texto a ser codificado:
Estrutura de Dados
333034013092288066049091986 095520 069870095053096
C-Codifica D-Decodifica (Ctrl+C para sair)
c
Digite texto a ser codificado:
Estrutura de Dados
060087078007613013613092088 069044 095125069549096
C-Codifica D-Decodifica (Ctrl+C para sair)
d
Digite texto a ser decodificado:
060087078007613013613092088 069044 095125069549096
estrutura de dados
C-Codifica D-Decodifica (Ctrl+C para sair)

8. Execução

C I F R A H O M O F O N I C A:

C-Codifica D-Decodifica (Ctrl+C para sair)

D

Digite texto a ser decodificado:

025418067042037 039381064763024073065666055281 040588 092381054096666097035053002

cifra homofonica de rossignol

C-Codifica D-Decodifica (Ctrl+C para sair)

D

Digite texto a ser decodificado:

090819019042911 039074815084024020111050061629 026016 042549072034418097077098023

cifra homofonica de rossignol

C-Codifica D-Decodifica (Ctrl+C para sair)

D

Digite texto a ser decodificado:

076072066092613056004091442 095520 040795069074072 - 077588002034578 100795063083016010

estrutura de dados - nelsi warken

C-Codifica D-Decodifica (Ctrl+C para sair)

9. Melhorias

- Inclusão de **codificação** para **espaço em branco**, **números**, letras **acentuadas**, caracteres **especiais**, etc.
- Estudo detalhado de **frequências de letras**, **símbolos**, **acentuações**, para melhorar o número de chaves para cada linha da matriz.
- Inclusão de ***nulos***, **símbolos** ou **códigos** incluídos no **criptograma**, mas que não significam nada e são **ignorados** na hora de decodificar.
- **Grupos de caracteres** que podem ser transformados em **outras chaves**.
- A **chave** pode ser aumentada para **4 ou mais números**.