Exercício : 1 _____ / _____ / _____

Esteganografia

Em países onde a censura é abundante, os dissidentes tendem a usar a tecnologia para burlar sua rigidez. A criptografia permite que mensagens secretas sejam enviadas mas se o governo está de olho sobre os dissidentes (os "de sempre", segundo o chefe de polícia de Getúlio Vargas, ou o ator Claude Rains que faz o chefe de polícia francês em Casablanca), o simples fato de enviar uma mensagem secreta já pode ser visto como delito reprimível.

A idéia é comunicar mensagens que não sejam entendidas como tal. Usa-se a esteganografia, das palavras gregas que significam "escrita cifrada". Heródoto escreveu sobre um general que raspou a cabeça de um escravo, tatuou lá uma mensagem, esperou o cabelo crescer e o enviou atravessando as linhas inimigas, que não perceberam a passagem do mensageiro. As técnicas modernas são conceitualmente as mesmas, só com uma largura de banda mais alta e com latência mais baixa.

A seguir duas mensagens inocentes de um gato:

A segunda imagem contém além da imagem do gato, a íntegra do Hino Nacional Brasileiro com um total de 3060 caracteres.

O truque está em usar uma imagem BMP em true-color, que como vimos admitem 16.777.216 (2^{24}) cores. Se usurparmos um único bit de cada byte, haverá uma diminuição nas cores, mas garantidamente nenhum olho humano será capaz de percebê-la. A idéia é deixar apenas 7 bits para codificar a cor e usar o bit menos significativo de cada byte para codificar a mensagem. A quantidade de cores agora é de apenas 2.097.152 (2^{21}), mas como está se usando o bit menos importante, a diferença é imperceptível.

Não fiz isto neste exercício, mas a boa prática manda que a mensagem seja enviada criptografada, quando então um degrau significativo na segurança da mensagem será adicionado.

Outra observação bem importante é que esta técnica só se aplica a imagens NOVAS. Se o autor pegar uma imagem manjada na Internet e codificar nela uma mensagem, qualquer desconfiança poderá ser confirmada simplesmente rodando um FILE COMPARER (utilitário FC) sobre a imagem enviada e a que existe por aí na Internet. O FC vai apontar diferenças de conteúdo, e vai escancarar o uso da esteganografia.

Finalmente, para encerrar o nosso curso de resistência às ditaduras, segue a recomendação de criar um site dedicado ao ditador cheio de imagens laudatórias dele, de suas pretensas conquistas e essas coisas. Cada imagem dessa pode carregar mensagens secretas e tão virulentas (do ponto de vista do ditador) quanto se queira.

Padrão BMP

O bloco de controle tem a seguinte configuração

Tipo	Tam	Desl	Nome	Descrição	exemplo
char	2	0	type	constante igual a BM	424D
long int	4	2	size	tamanho do arquivo	00 01 45 16
int	2	6	reserved	zeros	00 00
int	2	8	reserved	zeros	00 00
long int	4	A	off	desloc até a imagem	00 00 00 36
long int	4	E	size	resto do bloco controle	28 00 00 00
long int	4	12	width	n. de colunas	00 00 00 96
long int	4	16	height	n. de linhas	00 00 00 B8
int	2	1A	planes	planos	00 01
int	2	1C	bitcount	bits / pixel	00 18
long int	4	1E	compress	compressao	00 00 00 00
long int	4	22	sizeimage	tamanho da imagem	00 01 44 E0
long int	4	26	pix/m H	pixels / m na horizontal	00 00 00 00
long int	4	2A	pix/m V	na vertical	00 00 00 00

Note que

- Neste arquivo não há tabela de cores; ele é true color, já que o campo *bitcount* vale 18 (24 em decimal).
- Portanto, no endereço 36 está a cor azul do primeiro pixel da imagem.
- Cada linha, no seu final pode ter *stuffing bits* de maneira a alinhar a próxima linha em múltiplo de 4 bytes.
- Se um determinado pixel tiver R = G = B, pode-se concluir que este pixel é monocromático, ou como se diz "branco e preto".
- a ordem das cores no arquivo é BGR ao contrário do nome do esquema que é RGB.

Como fazer

A seguir o começo do primeiro arquivo do gato (sem nada)

```
ende   0  1  2  3  4  5  6  7   8  9  A  B  C  D  E  F ascii
----  -- -- -- -- -- -- -- --  -- -- -- -- -- -- -- -- ----------------
0000  42 4D 16 45 01 00 00 00  00 00 36 00 00 00 28 00 BM.E......6...(.
0010  00 00 96 00 00 00 B8 00  00 00 01 00 18 00 00 00 ................
0020  00 00 E0 44 01 00 00 00  00 00 00 00 00 00 00 00 ...D............
0030  00 00 00 00 00 00 B6 CD  C5 B7 CE C6 B8 CF CB B7 ................
0040  D0 CC AD C5 C5 8D A9 AA  63 7E 82 41 5F 64 34 51 ........c~.A_d4Q
0050  5A 26 45 4E 20 3E 49 2B  49 54 3C 5A 65 41 5F 6A Z&EN >I+IT<ZeA_j
0060  31 4F 5A 20 3E 49 2E 4C  57 3D 5B 66 44 60 6B 3B 1OZ >I.LW=[fD`k;
0070  57 62 3B 57 62 4F 6B 76  68 83 91 74 8F 9D 4B 66 Wb;WbOkvh..t..Kf
0080  74 51 6C 7A 4E 6B 79 4C  69 77 4C 69 77 47 64 72 tQlzNkyLiwLiwGdr
0090  49 69 76 57 79 86 94 B6  C3 69 8C 99 84 A9 B7 66 IivWy....i.....f
00A0  8B 99 69 90 9E 87 B1 BE  5B 84 93 6A 95 A4 89 B6 ..i.....[..j....
00B0  C4 85 B4 C2 8A B8 C9 87  B8 C8 70 A3 B3 5B 8E 9E ..........p..[..
00C0  65 99 A9 80 B5 C2 9E D5  E2 8C C4 CF B5 EA F7 9E e...............
00D0  D3 E0 8A BF CC A7 DC E9  86 BD CC 62 9B AA 67 9F ...........b..g.
00E0  B0 65 A0 B0 6C A4 B7 72  A7 BB 67 99 AF 55 84 9A .e..l..r..g..U..
00F0  50 7C 93 57 83 9A 66 92  A9 65 94 AA 72 9C B3 6A P|.W..f..e..r..j
0100  94 AB 57 7E 94 57 7E 94  5E 86 99 56 7E 91 4F 7A ..W~.W~.^..V~.Oz
0110  8B 5E 89 9A 47 72 83 51  7C 8D 58 81 90 44 6D 7C .^..Gr.Q|.X..Dm|
...
```

Agora, o começo do segundo arquivo (com o hino codificado)

```
ende   0  1  2  3  4  5  6  7   8  9  A  B  C  D  E  F ascii
----  -- -- -- -- -- -- -- --  -- -- -- -- -- -- -- -- ----------------
0000  42 4D 16 45 01 00 00 00  00 00 36 00 00 00 28 00 BM.E......6...(.
0010  00 00 96 00 00 00 B8 00  00 00 01 00 18 00 00 00 ................
0020  00 00 E0 44 01 00 00 00  00 00 00 00 00 00 00 00 ...D............
0030  00 00 00 00 00 00 B6 CC  C4 B6 CF C6 B9 CF CB B7 ................
0040  D1 CD AC C5 C4 8C A8 AB  62 7E 83 41 5F 65 34 51 ........b~.A_e4Q
0050  5B 27 44 4F 20 3F 48 2B  49 55 3C 5B 65 40 5E 6B ['DO ?H+IU<[e@^k
0060  31 4E 5B 20 3E 49 2E 4D  57 3D 5A 66 45 60 6A 3B 1N[ >I.MW=ZfE`j;
0070  57 62 3A 56 62 4F 6A 77  69 82 91 75 8E 9D 4A 66 Wb:VbOjwi..u..Jf
0080  75 50 6C 7A 4E 6A 78 4D  69 76 4C 69 76 46 64 73 uPlzNjxMivLivFds
0090  49 68 77 57 79 87 94 B6  C3 68 8C 98 84 A8 B6 67 IhwWy....h.....g
00A0  8A 98 69 90 9E 87 B0 BF  5B 85 92 6A 94 A4 88 B7 ..i.....[..j....
00B0  C5 84 B5 C2 8A B9 C8 87  B9 C9 70 A2 B3 5A 8E 9F ..........p..Z..
00C0  65 98 A8 80 B4 C3 9E D5  E3 8C C5 CF B5 EA F6 9F e...............
00D0  D3 E0 8A BF CD A7 DC E9  87 BC CC 62 9A AB 66 9E ...........b..f.
00E0  B1 64 A0 B0 6C A4 B6 73  A7 BA 66 98 AE 55 84 9B .d..l..s..f..U..
00F0  51 7D 92 56 83 9B 66 92  A9 64 94 AA 72 9C B2 6B Q}.V..f..d..r..k
0100  95 AA 57 7F 94 57 7E 95  5F 86 98 56 7E 91 4E 7B ..W..W~._..V~.N{
0110  8B 5F 88 9A 47 72 82 51  7D 8C 58 81 91 45 6C 7D ._..Gr.Q}.X..El}
...
```

Nele está codificada o Hino Nacional Brasileiro. Nos bytes 55, 56, ..., 70 (em decimal) ou 36, 37, 38, 39, 3A, 3B, 3C, 3D, 3E, 3F, 40, 41, 42, 43, 44 e 45 está codificado um número binário contendo o comprimento da mensagem, neste caso 3060. Que começa no byte 71 (decimal) ou 46 (hexadecimal).

A solução do problema pode ser obtida de duas maneiras:

- Escrevendo um programa que leia a imagem e recupere os bits montando conjuntos de 8 em 8 bits e consultando uma tabela ASCII ou
- Olhando o arquivo através de um utilitário binário (por exemplo wxHexEditor ou então o site http://www.onlinehexeditor.com/) e investigando os bytes em questão. Se o byte tiver conteúdo PAR, é porque foi codificado um ZERO nele. Se o conteúdo for ÍMPAR, codificou-se um UM nele. Deve-se recuperar os bits de 8 em 8 e depois consultar manualmente uma tabela ASCII.

☞ Para você fazer

Você recebeu um arquivo de nome

`XM01.BMP`

Nele está codificada uma pequena mensagem. Os dois primeiros bytes da mensagem codificam seu tamanho, como visto acima.

A mensagem achada é: