Segurança e privacidade na internet por esteganografia em imagens

Anderson Rocha[1] Siome Goldenstein[1] Heitor Costa[2] Lucas Chaves[2]

[1] *Instituto de Computação*
Universidade Estadual de Campinas
{anderson.rocha,siome}@ic.unicamp.br
[2] *Departamento de Ciência da Computação*
Universidade Federal de Lavras
{heitor,lucas}@ufla.br

Abstract

A esteganografia, arte e ciência das comunicações secretas, inclui um vasto conjunto de métodos para comunicações secretas tais como tintas "invisíveis", micro-pontos, arranjo de caracteres (character arrangement) entre outras. Os principais objetivos deste trabalho foram pesquisar as principais técnicas de esteganografia em imagens digitais da atualidade e desenvolver um software capaz de permitir a comunicação segura pela internet.

1. Introdução

Uma das áreas que tem recebido muita atenção recentemente é a *esteganografia*. Esta é a arte de mascarar informações como uma forma de evitar a sua detecção. *Esteganografia* deriva do grego, sendo *estegano = esconder, mascarar* e *grafia = escrita*. Logo, *esteganografia* é a arte da *escrita encoberta* ou, de forma mais abrangente, é a arte das comunicações encobertas [1].

A *esteganografia* inclui um vasto conjunto de métodos para comunicações secretas desenvolvidos ao longo da história. Atualmente, trabalha-se na estruturação e no desenvolvimento da *esteganografia digital*. Esta consiste em um conjunto de técnicas e algoritmos capazes de permitir uma comunicação digital mais segura em um tempo em que até os *e-mails* dos usuários de computadores podem estar sendo lidos e os seus passos em um computador pessoal rastreados.

Este artigo apresenta as principais técnicas de *esteganografia* da atualidade. O principal objetivo do trabalho foi aumentar a robustez das técnicas existentes. Para isso, criou-se uma nova técnica *esteganográfica*, unindo a força da cifragem de blocos do algoritmo criptográfico DES [2] e um conjunto de permutações cíclicas, às técnicas existentes.

2. Terminologia

Segundo o modelo geral de ocultamento de dados (*information hiding*), o dado embutido (*embedded data*) é a mensagem que se deseja enviar de maneira secreta. Freqüentemente, este dado é escondido em uma mensagem inócua (sem maior importância) conhecida como mensagem de cobertura (*cover-message*). Após o processo de inserção dos dados na mensagem de cobertura, obtém-se o chamado estego-objeto (*stego-object*) que é uma mensagem inócua contendo secretamente uma mensagem de maior importância. A Figura 1 apresenta como o processo pode ser interpretado. Um indivíduo escolhe o dado a ser escondido e, a partir de uma senha, mascara estes dados em uma imagem de cobertura previamente selecionada. O resultado é a estego-imagem a ser enviada [1]. Uma estego-chave (*stego-key*) ou *senha* é utilizada para controlar o processo de ocultamento de forma a restringir a detecção e/ou recuperação dos dados do material embutido.

3. Importância do trabalho desenvolvido

Juntamente com a *criptografia*, a *esteganografia* apresenta-se como uma tecnologia apta a auxiliar as pessoas a aumentarem sua privacidade *on-line*. No entanto, este alto grau de sigilo preocupa as autoridades políticas e policiais. Planos terroristas que ameacem a segurança nacional de um país podem ser elaborados

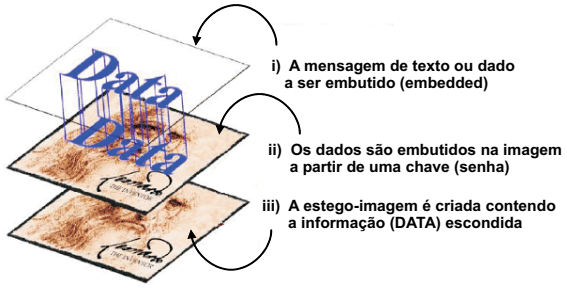

Figura 1: Exemplo de ocultamento de uma mensagem

totalmente em segredo. Segundo a *Eletronic Frontier Foundation* [3], muitas propostas de controle de privacidade já existem ou estão em andamento tais como: PATRIOT [1], Carnivore[2], DMCA[3], CAPPS II[4], entre outros.

Poucos sabem, mas a maioria do conteúdo em circulação pela *Internet* é constantemente vigiado pelo projeto *Echelon*. Este projeto visa filtrar toda a informação em circulação pela rede em busca de terroristas. O projeto existe em uma associação dos países EUA, Reino Unido, Austrália e Canadá. Outro projeto não menos relevante é o *UKUSA*. O termo é uma justaposição das siglas UK (*United Kingdom*) e USA (*United States of America*). Este projeto visa filtrar toda e qualquer informação em nome da segurança nacional dos países envolvidos. O *Echelon* e o *UKUSA* são dois atentados contra a liberdade de expressão dos cidadãos.

O controle de privacidade não pode ser justificado por uma visão deturpada de que segurança e privacidade são termos antagônicos. Não é verdade que, caso as pessoas não possam ser vigiadas, elas representam perigo ou para outras pessoas ou para o país.

4. Estado da arte

A abordagem mais comum de inserção de mensagens em imagens é relativa à técnicas baseadas em inserção LSB (*Least Significant Bit*) e podem ser aplicadas a cada *byte* de uma imagem de 32-*bits*. Estas imagens possuem cada *pixel* codificado em quatro *bytes*. Um para o canal **alfa** (*alpha transparency*), outro para o canal **vermelho** (*red*), outro para o canal **verde** (*green*) e outro para o canal **azul** (*blue*).

[1] *Provide Appropriate Tools Required to Intercept and Obstruct Terrorism.*

[2] Programa do FBI para vigiar o correio eletrônico.

[3] *Digital Milenium Copyright Act*

[4] *Computer Assisted Passenger Pre-Screening System*

Seguramente, pode-se selecionar um *bit* (o menos significativo) em cada *byte* do *pixel* para representar o *bit* a ser escondido sem causar alterações perceptíveis na imagem [1, 4].

Para entender melhor, suponha que se deseja esconder a letra **E** dentro da porção de imagem da Figura 2. Na Figura 2, têm-se três *pixels* da imagem de cobertura. Como a letra **E** pode ser escrita em forma binária segundo seu código ASCII como **10000011**, é suficiente utilizar apenas os dois primeiros *pixels* da imagem. Assim, utilizando-se a técnica LSB, tem-se o resultado mostrado na Figura 3. Na Figura 3, os *bits* destacados pelo quadrado representam as modificações necessárias nos LSBs para esconder a letra **E**.

```
(00100111 11101001 11001000 11101010) [a, R, G, B]
(10100111 11001000 11101001 11101000) [a, R, G, B]
(11001000 00100111 11101001 00100111) [a, R, G, B]
```

Figura 2: Porção de uma imagem de cobertura

```
(00100111 11101000 11001000 11101010) [a, R, G, B]
(10100110 11001000 11101001 11101001) [a, R, G, B]
(11001000 00100111 11101001 00100111) [a, R, G, B]
```

Figura 3: Porção da estego-imagem gerada pela porção de imagem 2

4.1 Softwares de esteganografia disponíveis

Aplicações de *esteganografia* estão disponíveis na *internet* para executar em uma grande variedade de plataformas.

Duas ferramentas capazes de trabalhar em associação com a *criptografia* são *Ray Arachelian's White Noise Storm*[5] e o *S-Tools*[6].

Colin Maroney desenvolveu o *Hide and Seek*[7]. Esta ferramenta é capaz de mascarar uma lista de arquivos em uma imagem, mas não faz uso de *criptografia*. Niels Provos desenvolveu o *Outguess*[8] capaz de esconder mensagens com alto grau de robustez em vários formatos de arquivos de imagens.

Duas outras ferramentas de destaque são *Jpeg-Jsteg*[9] capaz de fazer o mascaramento de informações utilizando os *pixels* mais significativos de uma imagem *jpg*

[5] `ftp.csua.berkeley.edu/pub/cypherpunks/`

[6] `ftp.idea.sec.dsi.unimi.it/pub/security/`

[7] `ftp.csua.berkeley.edu/pub/cypherpunks/`

[8] `www.outguess.org/`

[9] `ftp.funet.fi/pub/crypt/steganography`

e o *Camaleão*[10]. O *Camaleão* é o sistema desenvolvido pelos autores deste trabalho e é apresentado em mais detalhes na Seção 5.

5. Resultados e implementação

Ao longo deste trabalho, foi desenvolvido o *Camaleão: um software para proteção digital utilizando esteganografia* que permite a comunicação segura pela *internet* por fazer uso da *esteganografia*.

5.1 Robustez do software

A robustez da solução implementada se deve a três fatores: as senhas ou chaves de deslocamento, as permutações cíclicas entre os blocos e a cifragem interna dos blocos que podem ocorrer na mensagem antes do mascaramento.

5.1.1 As senhas

A senha, ou chave de deslocamento, é uma seqüência numérica de tamanho n. Os elementos pertencentes à chave estão mapeados no intervalo $\{0, \ldots, m\}$ sendo m um valor máximo. O valor m é dado a partir do módulo k em que se deseja criar a chave. Exemplo: Seja C uma chave de tamanho $n = 5$ e módulo $k = 3$. Cada elemento de C será um número inteiro pertencente ao intervalo $\{0, 1, 2\}$. A chave é criada a partir de um gerador pseudo-aleatório de números.

5.1.2 Cifragem de blocos e permutações cíclicas

Os autores desenvolveram uma abordagem chamada *cifragem por blocos com permutações cíclicas*. O processo é apresentado na Figura 4.

A mensagem a ser escondida é dividida em N blocos de tamanho fixo. A partir desse momento, os N blocos são independentemente criptografados com o algoritmo DES usando como chave simétrica a chave de deslocamento gerada. Isto produz N blocos criptografados C. Os blocos C são então permutados entre si ciclicamente de acordo com a chave de deslocamento. Em seguida, mais um processo de *criptografia* é feito internamente em cada bloco. Desta vez, o processo de *criptografia* é feito através de permutações cíclicas. O conjunto de *bits* resultante de todas estas operações é então escondido na imagem de cobertura a partir da chave de deslocamento gerada. Todo este processo é feito de modo a manter a *estego-imagem* resultante estatisticamente o mais próximo possível da imagem original.

[10]Disponível em `http://andersonrocha.cjb.net`

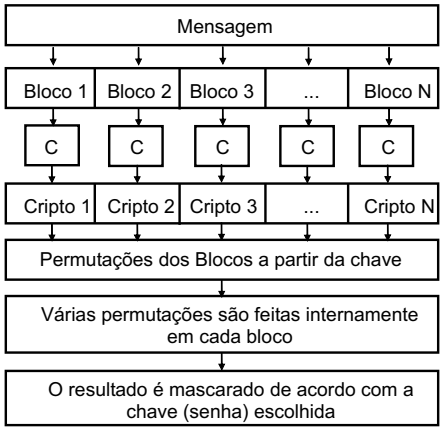

Figura 4: Exemplo de ocultamento de uma mensagem

6. Considerações finais

Este artigo apresentou as principais técnicas de mascaramento, em especial, mascaramento em imagens. Foi mostrado o método de mascaramento de dados em imagens baseado na cifragem de blocos e permutações cíclicas que foi utilizado na implementação feita.

A *esteganografia*, quando bem utilizada, fornece meios eficientes e eficazes na busca por proteção digital. Associando *criptografia* e *esteganografia*, como solução implementada, as pessoas têm o poder de comunicar-se em segredo pela *internet* mantendo sua identidade íntegra e secreta tendo mais uma opção para exercerem seu direito à liberdade e privacidade.

Referências

[1] F. A. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding - a survey," in *Proceedings of IEEE*, July 1999, pp. 1062–1078, special issue on Protection on multimedia content.

[2] B. Schneier, *Applied Cryptography*. New York: John Wiley & Sons, 1995, iSBN 0-47111-709-9.

[3] EFF, "EFF – The Eletronic Frontier Foundation," Disponível em www.eff.org, 2003, Último acesso em 10 de agosto de 2004.

[4] P. Wayner, *Disappearing cryptography*. San Francisco: Morgan Kaufmann Publishers, 2002, ISBN 1-55860-769-2.