Análise de Similaridade Estrutural de Imagens Esteganografadas com Python

Ewerton da Silva Farias, Geoflly de Lima Adonias e Carlos Danilo Miranda Regis
Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa – PB
E-mails: ewerton.farias.ee@gmail.com, gladonias@ieee.org, carlos.regis@ifpb.edu.br

***Resumo*—A segurança da informação e o direito à privacidade são temas que vem sendo amplamente debatidos por governos e pela população mundial. Embora seja possível combinar diversas formas de proteção para que seja atingido um maior nível de confidencialidade da informação, nós também podemos caracterizá-las em basicamente duas formas: criptografia e esteganografia, sendo a primeira uma forma de ocultar o significado e a última uma forma de ocultar a presença da informação. Neste artigo abordaremos a técnica, da esteganografia, de substituição do *bit* menos significativo de uma imagem, descrevendo seu funcionamento e mostrando resultados de avaliação do grau de similaridade estrutural entre as imagens originais e suas respectivas imagens esteganografadas, destacando o índice de eficiência do nosso algoritmo descrito neste trabalho.**

***Palavras-chave*—Esteganografia, python, imagem, lsb, criptografia, ssim, segurança, informação, similaridade estrutural.**

I. Introdução

Com o desenvolvimento de novas tecnologias notou-se um aumento significativo no uso de plataformas digitais com o intuito de melhorar e proporcionar maior agilidade e segurança na troca de mensagens e informações no nosso cotidiano. Pessoas e empresas buscam cada vez mais segurança em suas transações pela rede mundial de computadores, a fim de proteger seus dados digitais que trafegam constantemente em uma rede vulnerável à ataques de pessoas e *softwares* maliciosos.

Um dos ramos da criptografia é a esteganografia, que é uma palavra de origem grega que significa a arte da escrita escondida, onde *estegano* significa *esconder* e *grafia* signfica *escrita*. A estegoanálise por sua vez é a arte de detectar mensagens escondidas nos mais diversos meios de comunicação. A esteganografia inclui um amplo conjunto de métodos e técnicas para prover comunicações secretas desenvolvidos ao longo da história. Dentre as técnicas se destacam: tintas invisíveis, micropontos, arranjo de caracteres (*character arrangement*), assinaturas digitais e canais escondidos (*covert channels*) [1].

Na criptografia, é utilizada uma determinada chave, e a mensagem fica ilegível para qualquer pessoa ou sistema que não possua a chave. Entretanto, qualquer pessoa ou sistema inteligente saberá que naqueles dados há uma informação relevante. Com a esteganografia acontece diferente, uma vez que usa-se uma espécie de portadora para transportar a mensagem a ser transmitida, e essa portadora é um elemento que possui uma informação de sentido próprio e independente, sendo que a probabilidade da transmissão causar qualquer suspeita é mínima. Obviamente, pode-se usar primeiro a criptografia na mensagem a ser transmitida, e em seguida inserí-la na mensagem portadora, tornando o sistema ainda mais seguro, uma vez que alia a segurança contra interceptação (esteganografia) à segurança contra o acesso à informação (criptografia) [2].

Neste artigo, temos por objetivo inserir textos com tamanhos diferentes (informação) dentro de imagens com variações de cores distintas (portadora) utilizando a técnica *Least Significant Bit* (LSB), e avaliar o grau de semelhança estrutural entre as imagens originais e suas respectivas cópias esteganografadas utilizando a técnica *Structural Similarity* (SSIM), a fim de constatarmos o percentual de sucesso do processo de esteganografia.

II. Metodologia

O trabalho consiste no desenvolvimento de um algoritmo utilizando a linguagem de programação python, onde aplica-se a esteganografia fazendo-se uso da técnica LSB.

O código tem como objetivo encapsular uma informação (texto) dentro de uma imagem, ambos inseridos pelo usuário, de modo que as alterações presentes na nova imagem gerada pelo programa, não sejam perceptíveis ao olho humano.

Depois, é possível extrair a informação da imagem esteganografada, e avaliar o histograma de ambas as imagens com o intuito de verificar as diferença no nível de frequência entre as imagens original e esteganografada. Após a finalização do processo, as imagens são submetidas a uma análise de similaridade estrutural utilizando a técnica *Structural Similarity* (SSIM).

A. Imagens

Para análise, tratamos como critério de escolha das imagens, a serem submetidas ao processo de esteganografia, o formato *Joint Photographic Experts Group* (JPEG), resoluções similares e que possuíssem diferentes variações de cores e luminosidade entre si, conforme podemos ver na Figura 1.

B. A Técnica LSB

A técnica LSB consiste em utilizar o *bit* menos significativo de cada *pixel* (ou de cada cor) da imagem, para ocultar a mensagem [3].

Cada *pixel* de uma imagem corresponde a 24 *bits* ou 3 *bytes*, 1 *byte* corresponde a cor vermelha, 1 *byte* a cor azul e o

outro *byte* a cor verde. Assim a técnica LSB altera os *bits* menos significativos de um *byte*, e dessa forma não ocorre nenhuma perturbação na imagem visível ao olho humano, como podemos ver na Figura 2.

C. A Técnica SSIM

O índice SSIM é um método para medir a semelhança entre duas imagens. O índice SSIM pode ser visto como uma medida de qualidade de uma das imagens a ser comparado, desde que a outra imagem seja considerada a de referência [4].

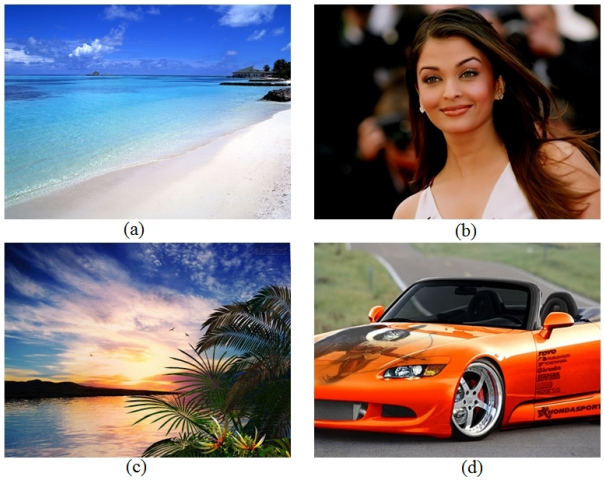

Figura 1: Imagens utilizadas na esteganografia.

Figura 2: (a) Imagem Original e (b) Imagem Esteganografada.

D. O Programa

Para o desenvolvimento do *software* utilizamos as seguintes bibliotecas: *Python Imaging Library (PIL)*, *OpenCV*, *stepic*, *matplotlib* e *PySSIM*.

A biblioteca *stepic* é a responsável pelo processo de esteganografia, seu funcionamento se dá com a leitura dos *pixels* da esquerda para a direita começando do topo, lendo três *pixels* por vez, e cada um deles contém três valores: *Red*, *Green* e *Blue (RGB)*. Os valores devem ser convertidos ímpar para 1 e par para 0 [5].

O primeiro passo se dá ao receber a imagem, indicada pelo usuário, após a entrada da imagem é solicitado ao usuário um texto a ser encapsulado na imagem. Para teste de comportamento das imagens foram inseridos textos de 196, 397 e 649 caracteres.

Após a inserção do texto é gerada a imagem esteganografada e são exibidos os gráficos contendo o histograma da estego-imagem comparado com o da imagem original. Por último, as imagens originais são comparadas com as imagens esteganografadas por meio da técnica SSIM.

III. Resultados

Com o intuito de avaliar a eficácia do processo de esteganografia, utilizou-se a ferramenta SSIM que avalia o grau de similaridade entre duas imagens e retorna um valor em uma escala de 0 a 1, na qual 0 indica que ambas as imagens comparadas não possuem similaridades entre si e 1 indica que elas são idênticas.

Para os testes, utilizamos as imagens da Figura 1 e textos de 196, 397 e 649 caracteres. Cada imagem foi esteganografada com os três textos e os resultados para cada figura estão registrados na Tabela I:

Texto	Fig. (a)	Fig. (b)	Fig. (c)	Fig. (d)
196	0.965462	0.980480	0.953276	0.979711
397	0.965463	0.980446	0.953267	0.979710
649	0.965461	0.980440	0.953278	0.979709

Tabela I: Valores de similaridades entre as imagens.

Pode ser observado que para a imagem (b) foram obtidos os melhores resultados por se tratar de uma imagem com menor variação de cores e luminosidade.

IV. Conclusão

Após realizarmos as avaliações entre as imagens originais e suas respectivas cópias esteganografadas, identificamos que quanto mais cores (e/ou detalhes) numa imagem e quanto maior a quantidade de dados a serem inseridos, existe uma maior probabilidade da imagem adquirir ruídos, e isso torna-se um aspecto chave, no uso da esteganografia, quando o assunto é manter o sigilo dos dados inseridos na portadora.

Os resultados demostraram que é possível aumentar o nível de confidencialidade dos dados a serem transmitidos, apenas levando em consideração aspectos gráficos da imagem e desconsiderando o uso de métodos mais sofisticados de criptografia devido a sua simplicidade e alto nível de eficiência.

Referências

[1] S. Petitcolas, F. A. P.; Katzenbeisser, *Information hiding techniques for steganography and digital watermarking*, 1st ed., 1999.

[2] B. W. Y. K. Huang, I. Glesk, E. E. Narimanov, T. Wang, and P. R. Prucnal, "Combining cryptographic and steganographic security with self-wrapped optical code division multiplexing techniques," Eletronic Letters, v. 43, Tech. Rep., Dezembro 2007.

[3] F. L. T. Jascone, "Protótipo de software para ocultar texto criptografado em imagens digitais," Trabalho de Conclusão de Curso Ciências da Computação, Universidade Regional de Blumenau, 2003.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr 2004.

[5] L. Domnitser, "How Stepic Hides Data in Images," http://domnit.org/blog/2007/02/stepic-explanation.html, 2007, [Acessado em 30 de setembro de 2014].