Instituto Tecnológico de Buenos Aires

StegoBMP

72.44 Criptografía y Seguridad - 2024Q1

Tomás Marengo - 61587

tmarengo@itba.edu.ar

Índice

1. Introducción

La esteganografía es la práctica de ocultar información dentro de otro objeto, de manera que su existencia no sea perceptible a simple vista. A lo largo de la historia, se han utilizado diversas técnicas esteganográficas para proteger la confidencialidad de la información. En el contexto digital, una de las técnicas más comunes consiste en ocultar datos en archivos de imagen modificando sus bits menos significativos (LSB, por sus siglas en inglés).

Este trabajo práctico tiene como objetivo implementar y probar un sistema de esteganografía digital utilizando diferentes métodos de inserción en los bits menos significativos de imágenes BMP. Además, se exploran técnicas de cifrado para proteger la información antes de ocultarla, garantizando así una capa adicional de seguridad. El objetivo es garantizar que la información sea indetectable y, en caso de ser descubierta, inaccesible sin la clave adecuada.

2. Diseño del sistema

El sistema está diseñado en una arquitectura modular que consta de las siguientes partes:

- **Operator:** Clase base que contiene los parámetros comunes a las operaciones de ocultamiento y extracción.
- **Embedder:** Clase derivada de Operator, responsable del ocultamiento de datos en una imagen BMP.
- **Extractor:** Clase derivada de Operator, responsable de la extracción de datos de una imagen BMP.
- **SteganographyUtil:** Contiene métodos estáticos para la encriptación y desencriptación de datos, así como constantes utilizadas en todo el sistema.
- **Steganography:** Punto de entrada del programa que parsea los argumentos de la línea de comandos y coordina las operaciones de ocultamiento y extracción.

3. Stegoanálisis de los archivos provistos

4. Cuestiones a Analizar

- *Discutir los siguientes aspectos relativos al documento.*
  - *a) Organización formal del documento:*
    - **Respuesta:** contiene por lo menos parte de la estructura típica de un paper: encabezamiento del informe, título, autoría, resumen (abstract), instroducción, discusión y conclusión, y referencias; en el medio de esto, todo el detalle del método.
  - *b) La descripción del algoritmo:*
    - **Respuesta:** propone un ejemplo y el pseudocódigo no es complejo, por lo que no es complicado de leer.
  - *c) La notación utilizada, ¿es clara? ¿hay algún error o contradicción?*
    - **Respuesta:** es claro pero yo agregaría casos relacionados a la extracción y no sólo al ocultamiento, ya que quedan preguntas sin responder, por ejemplo cuál es la mejor forma de guardar los patrones.
- *Esteganografiar un mismo archivo en un .bmp con cada uno de los tres algoritmos, y comparar los resultados obtenidos. Hacer un cuadro comparativo de los tres algoritmos estableciendo ventajas y desventajas.*
  - **Respuesta:**

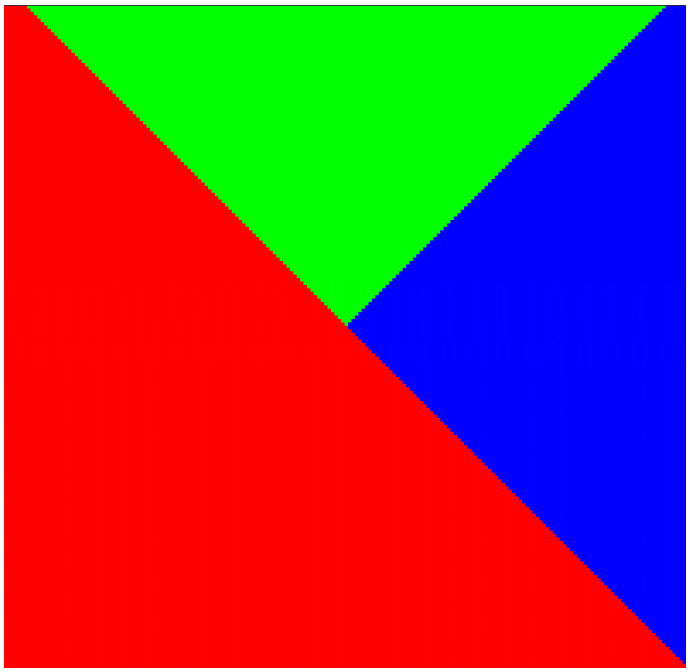

Figura 1: Imagen con un mensaje oculto con LSB1.

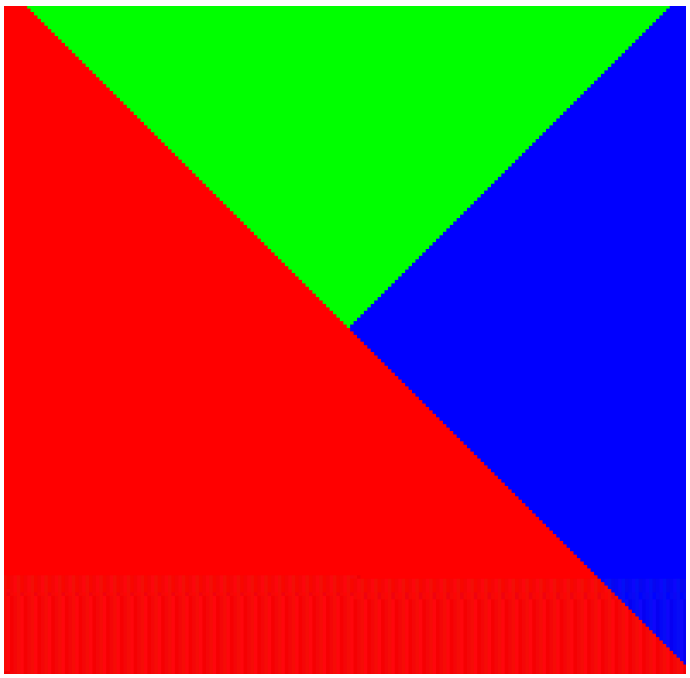

Figura 2: Imagen con un mensaje oculto con LSB4.

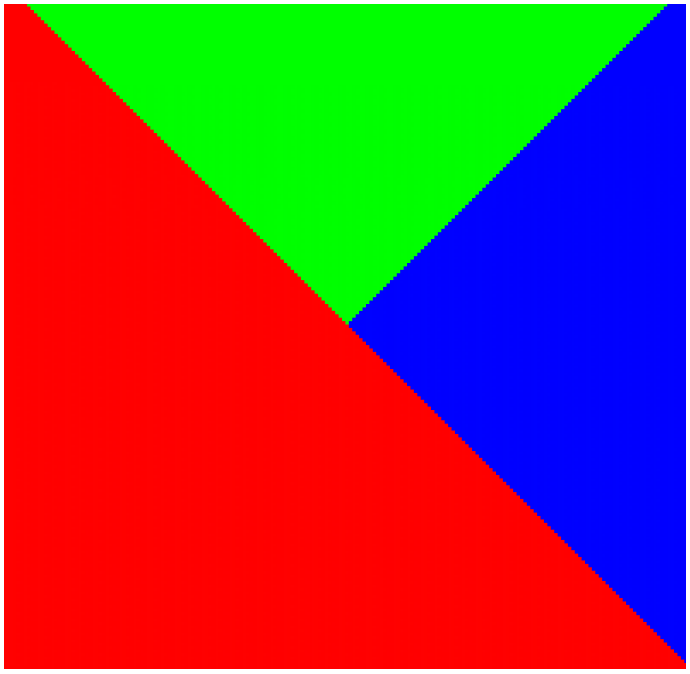

Figura 3: Imagen con un mensaje oculto con LSBI.

En LSB4 se puede notar claramente en la parte de abajo cómo cambió el color rojo. En LSB1 hay que forzar mucho la vista para poder notar un cambio en este color. Finalmente, en LSBI no se nota nada (en este caso particular).

Método	Ventajas	Desventajas
LSB1	◦ Fácil de implementar ◦ Baja distorsión visual ◦ Alta capacidad de almacenamiento	◦ Baja robustez contra ataques de compresión y manipulación de imagen ◦ Fácil de detectar con análisis estadístico
LSB4	◦ Mayor capacidad de almacenamiento que LSB1	◦ Moderada distorsión visual ◦ Baja robustez contra ataques de compresión y manipulación de imagen
LSBI	◦ Mayor robustez contra ataques ◦ Buena calidad de imagen ◦ Balance entre capacidad y seguridad	◦ Más complejo de implementar ◦ Menor capacidad de almacenamiento comparado con LSB4

Cuadro 1: Comparación de métodos de esteganografía: LSB1, LSB4 y LSBI

- *Explicar detalladamente el procedimiento realizado para descubrir qué se había ocultado en cada archivo y de qué modo. Indicar qué se encontró en cada archivo.*
  - **Respuesta: TERMINAR**
- *Algunos mensajes ocultos tenían, a su vez, otros mensajes ocultos. Indica cuál era ese mensaje y cómo se había ocultado.*
  - **Respuesta: TERMINAR**
- *Uno de los archivos ocultos era una porción de un video de una película, donde se ve ejemplificado una manera de ocultar información. ¿Qué se ocultaba y sobre qué portador?*
  - **Respuesta: TERMINAR**
- *¿De qué se trató el método de esteganografiado que no era LSB1 ni LSB4 ni LSBI? ¿Es un método eficaz? ¿Por qué?*
  - **Respuesta: TERMINAR**

- *¿Por qué la propuesta del documento de Majeed y Sulaiman es realmente una mejora respecto de LSB común?*
    - **Respuesta:** porque añade varias capas de seguridad: al invertir los bits de ciertos píxeles en la imagen, después de aplicar LSB1, se dificulta la recuperación de mensajes secretos por parte de un atacante ya que primero debería deducir el patrón utilizado para invertir o saber de dónde recuperarlo. Además al no usar los pixéles rojo, estos actúan como datos de ruido, aumentando la complejidad del proceso de extracción. También busca minimizar la cantidad de píxeles modificados en comparación con los métodos LSB1 y LSB4. Esto conduce a una mejora en la calidad de la imagen esteganográfica resultante
- *En la implementación se optó por guardar los patrones invertidos al final del mensaje. ¿Es esta una opción más segura que guardarlos antes del mensaje? ¿Por qué?*
    - **Respuesta:** Bueno acá tenemos el problema que le planteé a Ana. En principio lo más seguro siempre es al final, porque en principio, valga la redundancia, es más difícil saber dónde termina que dónde empieza (trivial). El problema de guardarlo al final es tener algún dato (que debería no ser tan trivial, sino sería lo mismo que guardarlo al comienzo) que nos diga dónde comienzan los patrones en ese "final". Por ejemplo, nosotros sabíamos que empezaban luego de la extensión, pero nunca podríamos saber dónde terminaba la extensión, ya que el carácter nulo también podría estar invertido.
- *¿De qué otra manera o en qué otro lugar podría guardarse el registro de los patrones invertidos?*
    - **Respuesta:** Como mencioné, se podría agregar al principio o al final. Además, se podría guardar en cualquier otro lado, pero deberíamos tener otro dato para saber dónde (por ejemplo si decidieramos incluirlo en la mitad, tendríamos que saber dónde está la mitad efectivamente). Caemos nuevamente en tener que ocultar otro dato (en nuestro ejemplo, la ubicación de la mitad) para saber donde está otro dato (nuestros patrones). Esto dificultaría un poco el análisis que tenga que hacer el atacante, pero cualquiera que sepa cómo funciona nuestro algoritmo, sabrá donde están esos datos.
- *¿Qué dificultades encontraron en la implementación del algoritmo del paper?*
    - **Respuesta:** Sin duda el manejo de bits y bytes, que siempre es un dolor de cabeza (operaciones de shifteo, and y or a nivel bit, etc.)
- *¿Qué mejoras o futuras extensiones harías al programa stegobmp?*

- **Respuesta:** Un análisis estadístico profundo testeando los diferentes algoritmos (y agregándo otros que encuentre publicados por la comunidad). De esta manera sacar conclusiones matemáticas de cuáles sirven más y en qué casos. Por ejemplo, pueden haber imágenes con colores de píxeles particulares donde un algoritmo sirva más que otro.