

Método Esteganográfico usando Recorrido de Grafos en Imágenes

Nils Murrugarra-Llerena^{1,2} Fredy Carranza-Athó^{1,2} Iván Vaca-Poquioma^{1,2}

¹ Sociedad de Estudiantes de Ciencia de la Computación

²Universidad Nacional de Trujillo

nineil.cs@gmail.com, ed.ededc@gmail.com, coren55@gmail.com

Resumen

El presente trabajo propone una técnica para ocultar datos (del tipo texto) en imágenes, pero con la intención de dispersarlos de un modo no continuado, evitando así que los datos se encuentren juntos y sea fácil de detectarlos. Para lograr nuestro objetivo, nos basamos en recorrido de grafos. Nuestro método resulta ser efectivo ya que al analizar los histogramas de las imágenes resultantes con las originales, la variación es mínima.

1. Introducción

(Fridrich and MGoljan, 2002) Desde tiempos antiguos los seres humanos hemos deseado enviar mensajes ocultos destinados a una persona en especial pero que este mensaje no pueda ser obtenido por terceras personas, algunos casos son (Plata, 2007) (Villagrán, 2002):

En Grecia, se ocultó un mensaje enviado hacia Esparta. En esos tiempos se usaban tabloncillos cubiertos con cera para escribir mensajes, por lo que el mensaje oculto fue escrito directamente en el tablón, luego cubierto con cera y finalmente un nuevo mensaje fue escrito sobre la cera. Otra técnica fue la de ocultar los escritos en papel utilizando tinta invisible, hecha en base al zumo del limón, la que se hacía visible al calentar el papel.

Durante la segunda guerra mundial, se microfilmaban los mensajes hasta reducirlos al tamaño de un punto, con lo que ese mensaje podía ser enviado como el punto de la "i" dentro de otro mensaje, sin levantar sospecha alguna.

Actualmente, con el uso de los computadores y el intercambio de información a través de medios informáticos, se ocultan mensajes en archivos que a simple vista son comunes y corrientes, como fotografías o archivos de música, pero que con un software adecuado podría obtenerse un mensaje que se encuentre oculto dentro de esos archivos.

En el presente trabajo detallaremos la manera de ocultar mensajes en imágenes, ocultando la mayor cantidad posible de información, además distribuyendo de la mejor manera los píxeles a modificar en la imagen.

La estructura del trabajo consta de una sección de Trabajos Previos donde se comenta acerca de trabajos relacionados hechos anteriormente. Una sección denominada Método EStegano Graf, donde se explica el método desarrollado. Las secciones Experimentos y Discusión de resultados donde se presentan los datos obtenidos y se los analiza. Finalmente la sección Conclusiones la cual indica las ventajas y desventajas de nuestro método así como Trabajos Futuros donde se comenta el alcance posterior de lo expuesto.

2. Trabajos Previos

Como uno de los inicios de la esteganografía aplicada a imágenes tenemos un método que consta en poner los datos tales como son en la imagen de salida, lo cual en la mayoría de los casos nos generara ruido.

A causa de este problema nace el método del bit menos significativo, propuesto por Derek Upham's(Provos and Honeyman, 2003), el cual consiste en tomar los bits menos significativos de los píxeles de la imagen, y en estos bits ocultar los datos necesarios, por supuesto esta técnica nos permitirá ocultar menos cantidad de información pero la hará casi imperceptible al ojo humano.

Por ejemplo si deseáramos almacenar la letra *a* (código ASCII 97) tenemos:

97 en binario=01100001

Para el caso de una imagen si deseamos almacenar un byte (8 bits) necesitamos tres píxeles (debido a que cada píxel almacena 3 bytes de datos, y nosotros almacenamos un bit por cada byte de la imagen). Por lo tanto, distribuiríamos el código ASCII del carácter de la siguiente manera(Ver Cuadro 1)(Agreda and Shimokawa, 2006):

Píxel	R	G	B
Píxel 1	0	1	1
Píxel 2	0	0	0
Píxel 3	0	1	-

Cuadro 1: Valores de reemplazo

El método del bit menos significativo tomado en cuenta escoge los píxeles a modificar de una manera secuencial, aglomerando los cambios realizados en la imagen en su parte inicial.

3. Método Estego-graf

Antes de iniciar esta sección, decir que el nombre de Estego-graf, es el nombre que pusimos al método utilizado. Se debe a la combinación de técnicas Esteganográficas y recorrido de grafos.

Como primer paso, se necesita una imagen en RGB, el formato no es de mucha importancia en este caso. Pero si se necesita que la imagen este en un formato de tres capas.

El primer paso, es considerar las entradas. Para ello es necesario obtener un archivo de texto plano y una imagen con las características mencionadas anteriormente. Y como último parámetro, es una matriz (M) de n filas, por dos columnas. Donde n representa la cantidad total de caracteres que se incrustarán en la imagen y las dos columnas que identifican la posición a ocultar de cada carácter en la imagen.

Dicha matriz M , necesita cierta seguridad, dado que la ubicación secuencial no sería tan buena, se busca conseguir una manera indistinta en el orden de los píxeles, pero la cual debe poder ser obtenida más de una vez con los mismos resultados.

Para realizar lo antes mencionado nos basaremos en conocimientos de teoría de grafos (Brassard and Bratley, 1997) (Cairo and Guardati, 2002), ya que los recorridos que se realizarán dentro de la matriz de la imagen, no será otra cosa que un recorrido en profundidad de un árbol, con lo cual lograremos calcular la matriz M . De esta manera el proceso nos proporcionará una serie de posiciones mucho más difícil de poder reconocer al aplicar un análisis sobre la imagen.

Veamos los pasos generales para realizar el recorrido de grafos:

1. Definimos nuestros operadores (ver Figura 1):

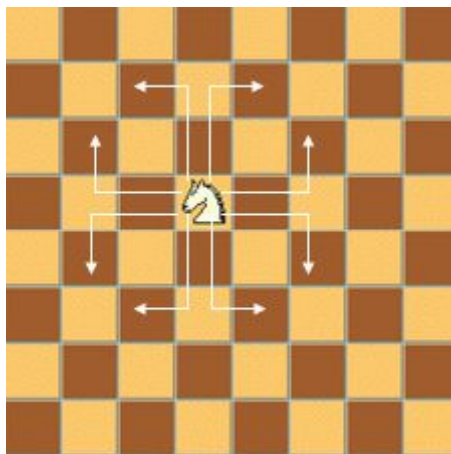


Figura 1: Movimientos del caballo

2. Definimos nuestras restricciones:

Los límites de la matriz (representación de la imagen) no deben de ser excedidos.

3. Elaboración del grafo de acuerdo a los movimientos del caballo (Ver Figura 2):

Los nodos representan una posición dentro de la matriz (fila, columna).

Las aristas representan uno de los movimientos posibles del caballo. Por ejemplo, en el que cada nodo se expande en ocho hijos, si ellos son posiciones válidas.

Por supuesto para la realización del recorrido de grafos debemos de contar con una posición inicial, y el proceso a realizar es un recorrido en profundidad con los posibles movimientos del caballo hasta que el número de posiciones visitadas es igual al número de caracteres del mensaje.

Tengamos en cuenta que el número de caracteres a ocultar sea menor o igual a la mitad del número de píxeles de la imagen ya que a menor cantidad de caracteres el proceso de determinar las posiciones será más rápido, ya que buscare el camino en el grafo con el cual visitara el número de caracteres determinado; debemos tener en cuenta que no será factible usar todos los píxeles de la imagen en el proceso de ocultamiento ya que daría prácticamente el mismo

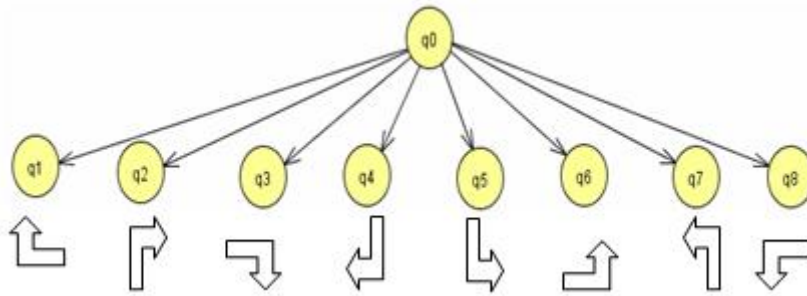


Figura 2: Operadores de movimiento del caballo

resultado con un método de posiciones secuenciales y además demoraría mucho más tiempo de procesamiento.

Al procesar la imagen y esconder el texto en ella, los caracteres quedarán lo mas dispersos posibles, siendo casi imposibles lograr su reunión de manera que el mensaje sea legible sin saber la manera en que se distribuyeron los píxeles. El sistema solo deberá conocer el punto de inicio donde se comenzó a simular el recorrido del caballo de manera que en el proceso de descryptación se produzca una matriz igual a M, que indique de la misma manera donde están ocultos los caracteres.

Para el reemplazo de los bits, se utiliza el Algoritmo 1:

Algoritmo 1 Incrustar los píxeles en la imagen

para todo píxel RGB **hacer**

Byte R ← reemplazar los tres LSB por los tres primeros bits del carácter

Byte G ← reemplazar los tres LSB por los tres bits siguientes del carácter

Byte B ← reemplazar los dos LSB por los dos últimos bits del carácter

fin para

4. Experimentos y Resultados

Los experimentos a continuación presentados fueron realizados en una foto de 361*241 píxeles de resolución. Entre los experimentos realizados iremos variando la cantidad de píxeles a incrustar, entre 3000 caracteres y 1500 caracteres, en cada ítem se compararan las imágenes inicial y final y los histogramas en rojo, verde y azul

1. Para ocultar un texto de: 3000 caracteres aproximadamente.(Ver Figura 3 a 6)

Los valores promedio de color los apreciamos en el cuadro 2.

2. Para ocultar un texto de: 1500 caracteres aproximadamente(Ver Figura 7 a 10).

Histograma	Original	Oculto
Histograma Rojo	134.9803	134.9256
Histograma Verde	104.939	104.8767
Histograma Azul	85.1796	85.1671

Cuadro 2: Promedio color imagen original (izquierda) y promedio color imagen con texto oculto(derecha) en figuras 3 a 6

Los valores promedio de color los apreciamos en el cuadro 3.

Histograma	Original	Oculto
Histograma Rojo	134.9803	134.9541
Histograma Verde	104.939	104.9112
Histograma Azul	85.1796	85.1743

Cuadro 3: Promedio color imagen original (izquierda) y promedio color imagen con texto oculto (derecha) en figuras 7 a 10



Figura 3: Imagen Original (Izquierda) e Imagen con texto oculto (Derecha)

Como otro punto importante compararemos el método usado en este trabajo con el del bit menos significativo usando posiciones secuenciales(Balocco and Lopez, 2005), los experimentos fueron realizados en una imagen de 50*50 pixeles ocultando 490 caracteres. Ver Figura 11.

5. Discusión de los Experimentos

De los experimentos realizados podemos apreciar que la imagen original con la de texto oculto no muestran diferencias notorias a simple vista, además teniendo en cuenta los histogramas estos tampoco varían mucho y por ultimo hemos tomado en cuenta un valor promedio de color de cada histograma (tomando por separado el rojo, verde y azul), siendo la máxima variación entre el original y el de salida de 0.0623.

Comparando el método propuesto con uno de posiciones secuenciales(ver figura 11), podemos decir que el método propuesto dispersa los pixeles a modificar, dando lugar esto a que los

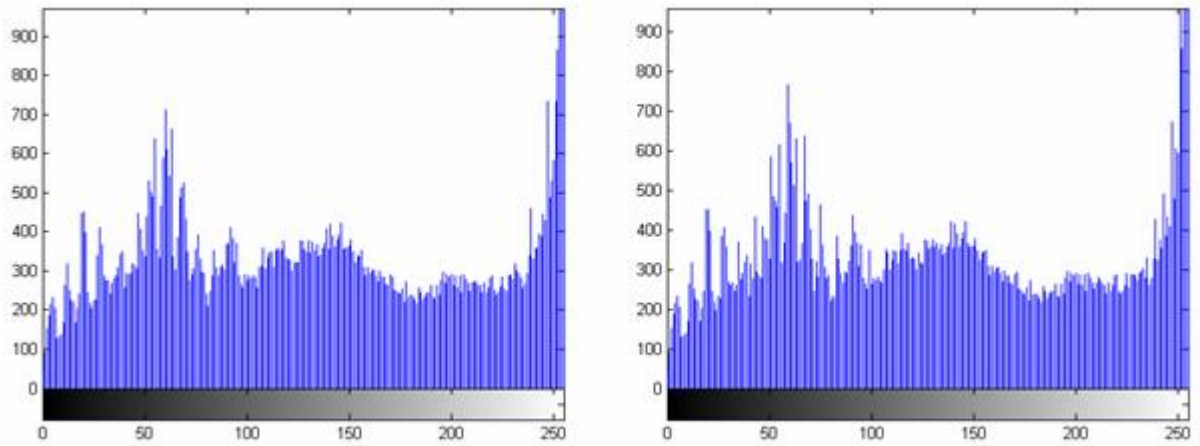


Figura 4: Histograma Rojo Imagen Original (Izquierda) e Histograma Rojo Imagen con texto oculto (Derecha)

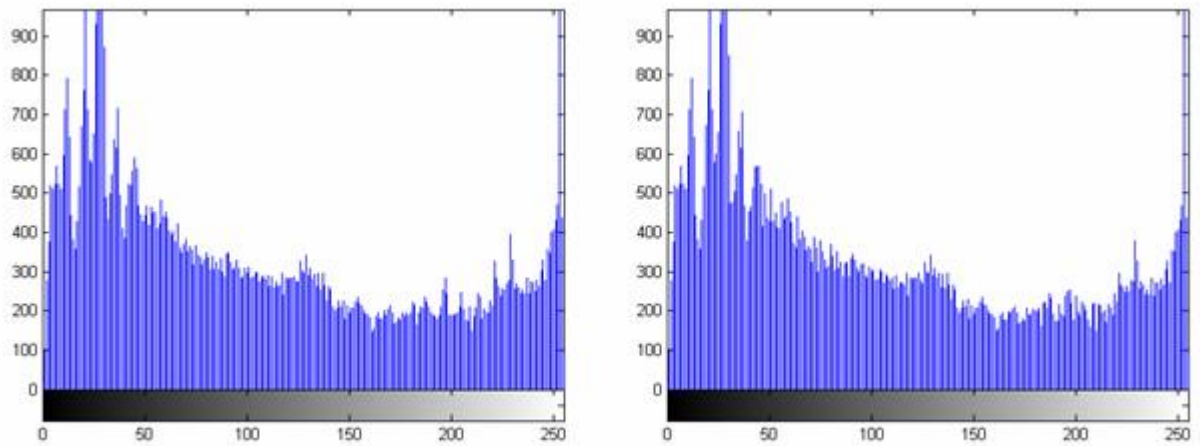


Figura 5: Histograma Verde Imagen Original (Izquierda) e Histograma Verde Imagen con texto oculto (Derecha)

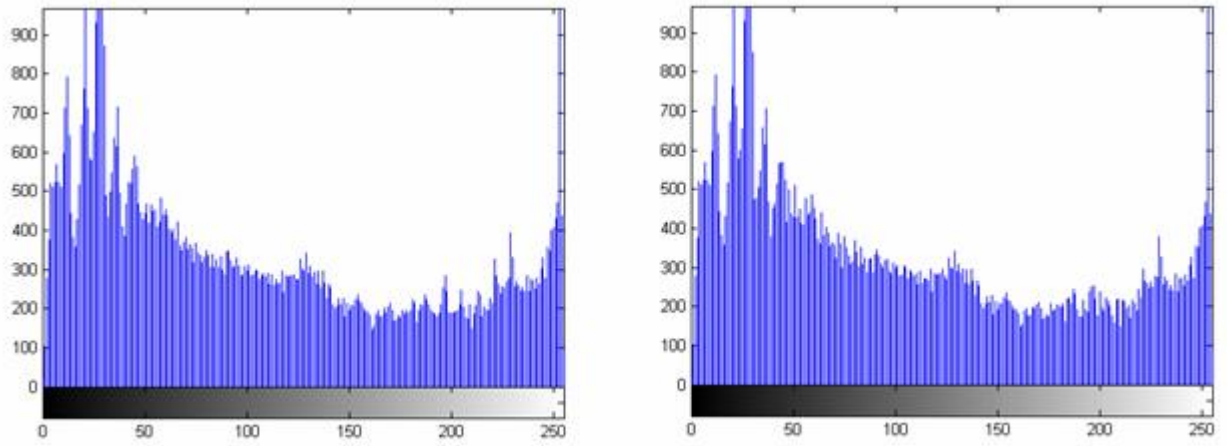


Figura 6: Histograma Azul Imagen Original (Izquierda) e Histograma Azul Imagen con texto oculto (Derecha)



Figura 7: Imagen Original (Izquierda) e Imagen con texto oculto (Derecha)

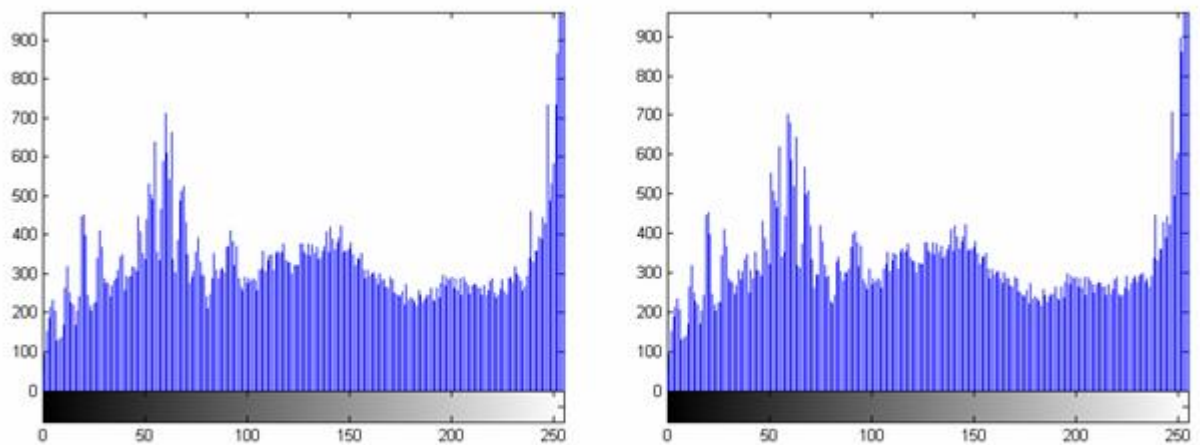


Figura 8: Histograma Rojo Imagen Original (Izquierda) e Histograma Rojo Imagen con texto oculto (Derecha)

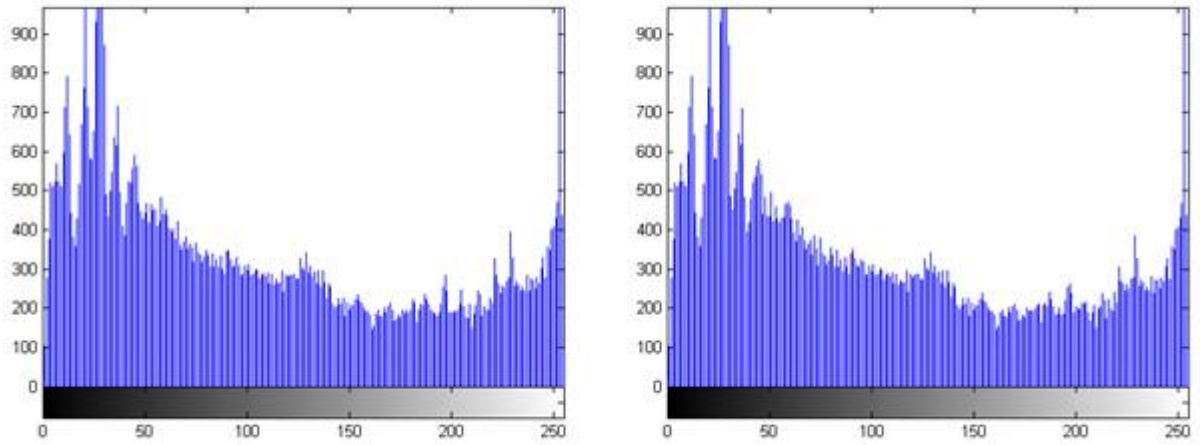


Figura 9: Histograma Verde Imagen Original (Izquierda) e Histograma Verde Imagen con texto oculto (Derecha)

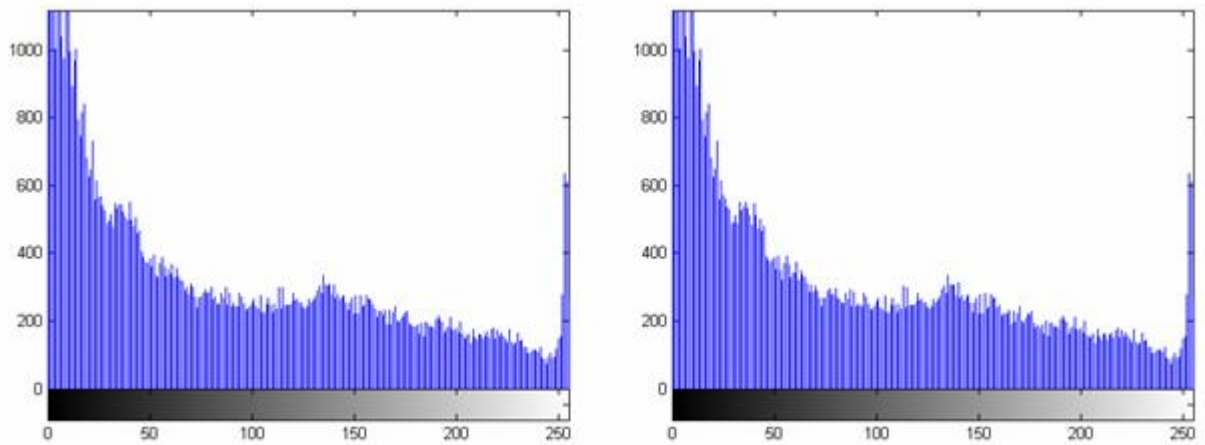


Figura 10: Histograma Azul Imagen Original (Izquierda) e Histograma Azul Imagen con texto oculto (Derecha)

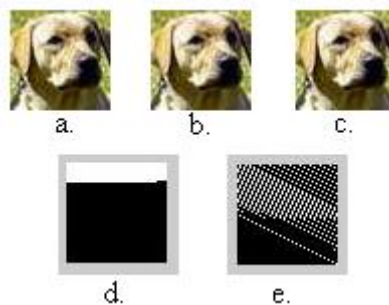


Figura 11: a. Imagen Original, b. Imagen con texto oculto con posiciones secuenciales, c. Imagen con texto oculto con posiciones dispersas (usando recorrido de grafos), d. Máscara de posiciones secuenciales, e. Máscara de posiciones dispersas

cambios en la imagen no se aglomeren en una sola región.

Por las consideraciones antes mencionadas y por la manera de dispersar los píxeles decimos que nuestro algoritmo es bueno.

6. Conclusiones

El método estego-graf produce resultados efectivos, dado que las mediciones realizadas a los histogramas demuestran cambios leves después de la incrustación de los datos.

El uso de recorrido de grafos en el algoritmo nos permite hallar posiciones no secuenciales, las cuales nos ayudan a que no se aglomere en un solo lugar los cambios realizados en la imagen.

Una de las desventajas, es el tiempo que demanda el cálculo de las posiciones, pero el cual es compensado con la efectividad del método. Por ejemplo en el caso de los métodos de posiciones secuenciales, estos no se demoran en el cálculo de las posiciones, pero los cambios realizados a la imagen se aglomeran en el inicio de la imagen. En cambio al usar posiciones dispersas distribuimos de mejor manera los cambios a costo de tiempo de procesamiento.

Otra de las desventajas, es que el método de dispersión no logra dispersar los píxeles a modificar en toda la imagen de una manera uniforme. Por ejemplo en el caso de la Figura 11, podemos ver que en la zona inferior izquierda no se ha realizado ningún cambio.

7. Trabajos Futuros

Dentro de la implementación, se tomo también en cuenta el ocultamiento de una imagen dentro de otra, para lo cual, se consideró reemplazar los bits de una imagen, pero en este caso la imagen a ocultar deberá estar en escala de grises. Esta restricción se hace dado que las imágenes en RGB triplicarían el tamaño y el tiempo de la obtención de las posiciones sería mucho mayor. Como la razón de una imagen RGB a una en escala de grises es de 3:1, obviamente para ocultar una imagen dentro de otra, una debe tener una resolución igual o mayor.

Si bien es cierto hasta al momento existen implementaciones como imágenes y texto plano, esta técnica se puede ampliar hacia archivos de sonido y video, ya que solo se necesitaría su representación binaria.

Referencias

- Agreda, C. and Shimokawa, E. (2006). Desarrollo de técnicas de esteganografía. *Universidad Nacional de Trujillo*.
- Balocco, H. and Lopez, E. (2005). Ocultamiento de información en imágenes. <http://cpdsifich.wikidot.com/local-files/tpsaplicacion/2005-LopezBalocco-EsteganoRGB.pdf>.
- Brassard, G. and Bratley, P. (1997). *Fundamentos de Algoritmia*. Prentice Hall.

- Cairo, O. and Guardati, M. (2002). *Estructuras de Datos*. McGrawHill.
- Fridrich, J. and M Goljan (2002). Practical steganalysis: State of the art. *Security and Watermarking of Multimedia Contents*.
- Plata, J. (2007). Historia de la criptografía y esteganografía. <http://atalap.blogspot.com/2007/09/historia-de-la-criptografa-y.html>.
- Provos, N. and Honeyman, P. (2003). Hide and seek: An introduction to steganography. *IEEE Journal*.
- Villagrán, M. (2002). Orígenes de la esteganografía. <http://www.vsantivirus.com/esteganografia.htm>.