

Selección óptima de métodos de sustitución en aplicaciones esteganográficas

Ing. Gustavo Rodríguez Medina¹, Mg. Ing. G. Sergio Navas², Mg. Ing. Jorge Eterovic³

Gabinete de Computación / Fac. de Ingeniería / Univ. Nacional de San Juan ^{1,2}

Av. Libertador Gral. San Martín 1109 (oeste) – San Juan
0264 – 4211700 (Int. 285¹ / 435²)

Escuela de Posgrado / Univ. Nacional de la Matanza ³

Moreno 1623 - 1^o piso, Ciudad Autónoma de Buenos Aires ³
011 – 4331-9816/17

grodriguez@unsj.edu.ar ¹, snavas@unsj.edu.ar ², eterovic@unlam.edu.ar ³

Resumen

La Esteganografía hace uso de diferentes técnicas para ocultar información sensible en un Portador, y que la misma pase inadvertida para terceros [1]. Generalmente, los objetos portadores utilizados son elementos multimediales, tales como imágenes, audio y video digital. El proceso de ocultación de la información debe procurar muy alta efectividad y razonable eficiencia [2].

Al estudiar y al entrar en detalles en las técnicas esteganográficas, comienzan a surgir algunos parámetros o variables de medida, las cuales permiten comparar entre las técnicas aplicadas. Estas características permiten poner al alcance del usuario la posibilidad de elegir, según los requisitos u objetivos planteados, qué características preservar o minimizar, tales como: *Capacidad* de recepción de información del portador, nivel de *Perceptibilidad* visual y nivel de *Detectabili-*

dad del mensaje oculto en el portador. (contra ataques de esteganálisis).

Entre ellas se presenta una relación de compromiso, es decir, al manipular una variable de interés ésta puede presentar una mejora en detrimento de otra.

Como solución a lo anterior, en este trabajo se presentan tres propuestas para la búsqueda y selección óptima de métodos de Sustitución LSB, aplicados a imágenes Bitmap, basándose en Capacidad de alojamiento, requerimientos de Imperceptibilidad visual y de Indetectabilidad del mensaje a ocultar.

Los algoritmos desarrollados se implementan en Matlab©, conformando un Software esteganográfico prototipo completamente funcional.

Palabras clave:

Esteganografía, Sustitución LSB, métodos óptimos.

Contexto

El presente trabajo expone de manera resumida los principales resultados obtenidos en el marco de la tesis "*Estudio, análisis, desarrollo y propuestas de algoritmos para la selección óptima de métodos de sustitución en aplicaciones esteganográficas*", para la Maestría en Informática de la UNLaM, bajo la dirección de docentes/investigadores de la Univ. Nacional de la Matanza y de la Univ. Nacional de San Juan.

Introducción

Las técnicas esteganográficas intentan ocultar un objeto software de cualquier tipo, el cual representa la información de interés, camuflándolo en otro, denominado portador, anfitrión o cubierta, con el fin de ser enviados desde un emisor hacia un receptor, a efectos de lograr una transmisión encubierta [1].

Si bien podrían ser utilizados portadores de cualquier tipo, pero debido a que los sentidos humanos son limitados, el ojo y oído no pueden detectar cambios sutiles en presentaciones visuales o de audio, por lo cual los objetos software multimediales son más aptos para encubrir información [2]. Estos resultan ser los preferidos en aplicaciones esteganográficas como elemento portador dada la eficacia de las técnicas que aprovechan tales limitaciones. El presente trabajo se orienta a la utilización de portadores imagen Bitmap en formato BMP color de 24 bits.

En síntesis, un portador imagen puede ser utilizado para ocultar, a la vista de intrusos, cualquier mensaje u objeto software (archivo), codificándolo como sutiles cambios en los colores de los píxeles (sus componentes RGB) que no deben ser percibidos por el ojo humano; de tal forma que el Mensaje embebido en el Portador, denominado como "*Estego-*

portador", pueda ser transmitido, sin que sea detectado el hecho; para luego aplicar el proceso inverso (decodificación) de modo que el receptor pueda recuperar y disponer del Mensaje enviado [1].

La figura 1 expone un esquema general del sistema de esteganografiado.

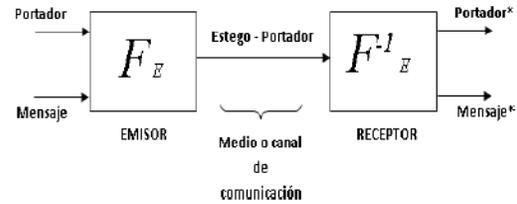


Figura 1. Esquema general de un sistema de Esteganografía Digital.

Otros métodos esteganográficos sacan provecho de la incapacidad de aplicaciones software para detectar la presencia de objetos extraños (archivos) "insertos" en portadores [3][4]. Tales métodos, si bien pueden utilizar cualquier tipo de portador (incluso ejecutables), poseen debilidades que los ponen muy en desventaja frente a las técnicas de sustitución [2].

Esteganografía por Sustitución

En general, cualquier archivo multimedia contiene áreas donde los datos son "poco perceptibles a los sentidos". Esas áreas se pueden sustituir por otros datos, realizando cambios que no sean detectados, en particular, en este caso, por el sentido de la vista [2]. Esto permite encubrir información de interés dentro de un archivo portador, haciendo que el mismo *parezca* igual al original.

Los métodos de sustitución consisten, básicamente, en utilizar los bits menos significativos de los bytes componentes de color del portador (RGB) para sustituirlos por bits correspondientes a un mensaje a encubrir.

La técnica más conocida y utilizada es la llamada LSB¹ de 1 bit, que se basa en usar el bit menos significativo de cada byte del portador, sustituyéndolo por bits correspondientes al mensaje. La imagen resultante del proceso (Estego-portador) sufre cambios respecto a la original que son realmente *imperceptibles* para el ojo humano, si el portador no es una imagen de color plano (o único).

Una de las desventajas del citado método es su limitada capacidad: sólo hasta un 12,5% de la imagen puede dedicarse a “mensaje”. Tal límite comienza a incrementarse si se utilizan variantes de la técnica base LSB, las que hagan uso de más de 1 bit por byte del portador. Así por ejemplo, es posible utilizar los dos bits menos significativos de cada byte (cada canal de color), obteniéndose una capacidad del doble (25%), pero en detrimento de la *Perceptibilidad* [2].

Por lo tanto, en alguna medida, se estará limitado en la cantidad de bits a sustituir del portador, dependiendo del contenido de ruido de la imagen (categoría de la imagen), o dicho de otra manera, de los cambios bruscos de color entre píxeles adyacentes. Resulta intuitivo expresar que portadores ruidosos admitirán mayor cantidad de bits a sustituir insertándose así el mensaje de interés, mientras que aquellos de colores uniformes (menos ruidosos) serán afectados en mayor medida (efectos visuales perceptibles no deseados). Surge la necesidad de contar con alguna *guía* para que la tarea de ocultar información en un portador imagen, de forma eficaz, no se convierta en una serie de intentos por prueba y error, o que la misma solo quede en manos de un experto.

Algoritmos de Sustitución desarrollados

Durante la elaboración del trabajo de investigación y desarrollo, indicado anteriormente, se han elaborado algoritmos para la implementación del método base de *Sustitución LSB de 1 bit*, y variantes a esta, los que permiten aplicar también la sustitución en diversas cantidades de bits en los canales o planos de color RGB de la imagen portadora. De esta manera, se han implementado como funciones de Matlab la variante de *Sustitución LSB Simple* (reemplazando igual cantidad de bit en cada canal de color del portador), la variante de *Sustitución LSB Compuesta* (sustituyendo diferentes cantidades de bits en cada canal de color del portador), la variante de *Sustitución LSB con Alternancia de canal*, en la cual se alternan los canales de color que se reemplazan en el portador por el mensaje, como así también la implementación de técnicas que dispersan los bits del mensaje en el portador a efectos de evitar la detectabilidad. De esta última, se implementó una variante de *Sustitución con Dispersión DELTA Fija* y otra de *Sustitución con Dispersión DELTA Aleatoria*, siendo el DELTA la separación máxima (en cantidad de píxeles) entre los píxeles afectados por la operación de sustitución en el portador.

También se han desarrollado aquellos algoritmos que permiten recuperar un mensaje oculto en un portador BMP, que haya utilizado algunas de las técnicas antes enunciadas, conformando así el esquema completo de un sistema esteganográfico presentado en la figura 1. Además, se desarrollaron otros algoritmos que han servido como herramientas de apoyo y contribución al alcance de los objetivos propuestos, tales como un algoritmo para la *Categorización Automática de imágenes*, un algoritmo de análisis que implementa *Filtros por Capas de Bits y de Canales de Color*, etc.

¹ *Least Significant Bit* (Sustitución del Bit Menos Significativo)

Tablas de métodos de Mejores Relaciones

Tal como se ha expuesto, dependiendo de la cantidad de bits a sustituir del portador, del método utilizado, del contenido de ruido de la imagen portadora (categoría) y del tamaño del mensaje a ocultar, el resultado, el *Estego-portador*, presentará efectos visibles, o no, y será más o menos detectable el mensaje ocultado, cuando terceros intenten descubrirlo mediante filtros gráficos y/o análisis estadístico de la imagen, u otras técnicas (esteganálisis).

A la vez, también se conoce de la dependencia entre las variables indicadas como *Perceptibilidad* visual, *Detectabilidad* del mensaje oculto y la *Capacidad* de recepción de información del portador.

Por lo cual, la tarea de operar (minimizar, mantener o aumentar) las variables mencionadas, al ocultar un mensaje, se transforma más bien en una serie de intentos de pruebas y error, echando mano a la intuición, al no contar con alguna guía o camino a seguir para aplicar de manera efectiva y eficiente la tarea de ocultamiento de información.

Para dar solución a esto, es que se toma como punto de partida la Tabla de "*Mejores Relaciones Perceptibilidad - Capacidad*" [2], la que presenta un conjunto de métodos esteganográficos, que aseguran las mejores relaciones posibles entre la categoría del portador (de 1 a 4), la Capacidad requerida del propio portador para ocultar el mensaje (indicado como filas de la tabla) y el nivel de Perceptibilidad que resulta al aplicar el método (dispuesto como columnas de la tabla). Esto sirve de guía y **asegura resultados óptimos** en cuanto a la *Perceptibilidad* visual a lograr.

En relación a resultados que aseguren cubrir aspectos de *Detectabilidad*, se ha trabajado sobre otra tabla que resume los mejores métodos que aseguran cubrir esta variable, a la cual se le ha denominado

Tabla de "Mejores Relaciones Detectabilidad - Capacidad - Perceptibilidad"

La primer tabla, contempla métodos que permiten ocultar tamaños de mensajes que representan hasta un 62,5% del tamaño del portador, mientras el límite de la segunda es del 33,3%.

Algoritmos de Búsqueda y Selección óptima de Métodos de Sustitución

Basándose en la tabla de *Mejores Relaciones Perceptibilidad - Capacidad* y en la tabla de *Mejores Relaciones Detectabilidad - Capacidad - Perceptibilidad*, se proponen tres algoritmos, considerando los posibles requerimientos de un usuario:

- Aplicaciones que requieren *Cubrir Perceptibilidad* (Perceptibilidad crítica).
- Aplicaciones que requieren *Cubrir Detectabilidad* (Detectabilidad crítica, solución de compromiso).
- Aplicaciones con *Solución por Defecto*.

Algoritmo para Cubrir Perceptibilidad

El algoritmo recorre tabla de "*Mejores relaciones Perceptibilidad - Capacidad*", y selecciona aquel método que permita obtener la menor perceptibilidad posible que cumpla con la Capacidad requerida. Dicho requisito de Capacidad estará en función de los tamaños de los archivos correspondientes al mensaje y al portador (seleccionados por el usuario) y también a la categoría de imagen, calculada por el software.

Si según el método de Sustitución LSB seleccionado, resultasen valores de Perceptibilidad "Nula" o "Muy Baja", el software aplica e informa el método, como así también el valor de *Perceptibilidad* logrado. Si con el método encontrado

solo se cubre *Perceptibilidad* con un valor "Bajo", se advierte al usuario para confirmar su aplicación. Si la *Perceptibilidad* obtenida fuera de valor "Media", no se aplica ningún método y se informa que no se pueden cubrir los requerimientos, sugiriéndose algunas alternativas.

Algoritmo para *Cubrir Detectabilidad*

El algoritmo recorre tabla de *Mejores Relaciones Detectabilidad - Capacidad - Perceptibilidad* y selecciona aquel método que permita obtener la menor *Perceptibilidad* que cumpla con la *Capacidad* requerida, priorizando *Detectabilidad*.

Si con el método de *Sustitución* seleccionado se obtienen valores de *Perceptibilidad* "Nula", "Muy Baja" o "Baja" (la menor que se logre), siempre que resultare posible aplicar algún método con *Alternancia* o *Dispersión*, los que aseguran una *Detectabilidad* "Muy Baja", entonces se aplica dicho método encontrado. Si esto último no fuera posible, se aplica algún método *Combinado* que asegura una cobertura de *Detectabilidad* en forma "Parcial" o "Baja". En todos los casos se informa el método aplicado y los correspondientes valores logrados de *Perceptibilidad* y *Detectabilidad*. En el caso que no se pueda aplicar ningún método que cumpla los requisitos también se informa.

Algoritmo de *Solución por Defecto*:

Con este algoritmo se pretende ofrecer un tipo de solución donde el usuario, para todos en general y en particular para aquellos con escasos conocimientos sobre el tema, donde el algoritmo busca la alternativa óptima para aplicar un método para ocultar información. El usuario permite que el algoritmo de *Solución por Defecto* tome la decisión de aplicar la mejor alternativa posible. Básicamente el algoritmo trata de aplicar algún método que cubra *Perceptibilidad*, y si se puede,

además, trata de cubrir también *Detectabilidad*.

Si es posible, se debe aplicar algún método con *Dispersión*, *Alternancia* o *Combinado*, con el menor valor de *Perceptibilidad* posible. Valores admitidos de *Perceptibilidad* son "Nula", "Muy Baja" y "Baja". En el caso que sea "Baja", se informa que se está en un valor límite. Si el valor fuera "Media", se consulta al usuario si se aplica o no el método encontrado. En todos los casos se informan los valores obtenidos de *Perceptibilidad*, *Detectabilidad* y el método aplicado.

Este algoritmo resulta ser una combinación de los dos anteriores, con algunas salvedades.

Líneas de Investigación, Desarrollo e Innovación

La línea de investigación corresponde a la temática de *Esteganografía*, la que se enmarca en el área de Seguridad Informática. En este sentido, los autores del presente trabajo han elaborado dos tesis de Maestría, como así también la publicación de algunos resultados en diferentes congresos.

Los resultados obtenidos desde el año 2006, aportan estudios y desarrollos innovadores en el área.

Resultados y Objetivos

Se han efectuado diversas pruebas a los 3 algoritmos de *Búsquedas y Selección de Métodos Óptimos: Cubrir Perceptibilidad, Cubrir Detectabilidad y Solución por defecto*.

Para cada una de las pruebas se utilizan Portadores imagen BMP de 24 bits, de categoría 1, 2 y 3, como así también archivos mensajes de diferentes tamaños, con el fin de obtener diversos valores de

Capacidad requerida (relación tamaño del mensaje y el portador). Esto permite que a través del Software Prototipo desarrollado, se seleccionen diversos pares de portadores y mensajes, con el propósito de comprobar que cada uno de los algoritmos recorren las correspondientes tablas de métodos óptimos de manera adecuada, y de acuerdo a la descripción anteriormente efectuada de cada uno de los tres algoritmos.

Por cada prueba se obtienen ventanas con resultados en el Software Prototipo, las que indican: el método aplicado, los niveles cubiertos de *Perceptibilidad* y *Detectabilidad*, relación mensaje / portador, el porcentaje de ocupación del mensaje en el portador, etc. Dichos valores han sido tomados y dispuestos en tablas de resultados, para cada uno de los algoritmos. El Software Prototipo también presenta gráficamente los resultados de cada prueba, las imágenes *Estegoportadores* obtenidos, con el fin de comparar con la imagen portadora original.

Todas las tablas de resultados obtenidas, de cada uno de los algoritmos desarrollados, para portadores de categorías 1, 2 y 3, y para diversos tamaños de archivos mensajes, se han contrastado con las correspondientes tablas de *Mejores Relaciones Perceptibilidad - Capacidad* y de *Mejores Relaciones Detectabilidad - Capacidad - Perceptibilidad*, obteniendo excelentes resultados.

Formación de Recursos Humanos

En la temática de Esteganografía se viene trabajando entre ambas universidades citadas desde el año 2006.

En ese año, el trabajo conjunto dio como resultado una Tesis de Maestría denominada "*Exploración de efectos esteganográficos sobre portadores imagen de mapa de bits utilizando diferentes técnicas y algoritmos*".

En la actualidad se encuentra finalizada y en etapa de presentación otra Tesis de Maestría, sobre la cual se basa la presente publicación, cuyo título es "*Estudio, análisis, desarrollo y propuestas de algoritmos para la selección óptima de métodos de sustitución en aplicaciones esteganográficas*".

Referencias

- [1] C. Gustavo Rodriguez M. Estudio, análisis, desarrollo y propuestas de algoritmos para la selección óptima de métodos de sustitución en aplicaciones esteganográficas. Argentina. Univ. Nacional de la Matanza – Escuela de Posgrado. 2015.
- [2] G. S. Navas. Exploración de efectos Esteganográficos sobre portadores imagen de mapa de bits utilizando diferentes técnicas y algoritmos. Argentina. Univ. Nacional de la Matanza – Escuela de Posgrado. 2006.
- [3] Ross J. Anderson, Fabien A. Petitcolas. On The Limits of Steganography [en línea]. United Kingdom. IEEE Journal of Selected Areas in Communications, 16(4):474-481, Mayo 1998. Special Issue on Copyright & Privacy Protection. ISSN 0733-8716. [Citado 09 Septiembre 2013]. Disponible en World Wide Web: <http://www.cl.cam.ac.uk/~rja14/Papers/jsac98-limsteg.pdf>
- [4] Abbas Cheddad, Joan Condell, Kevin Curran, Paul McKeivitt. Digital image steganography: Survey and analysis of current methods [en línea]. El Sevier (Ed), Journal of Signal Processing, 90(3), 727-752, marzo 2010. [Citado 12 Junio 2014]. Disponible en World Wide Web: <http://www.sciencedirect.com/science/article/pii/S0165168409003648>