SECRETO COMPARTIDO EN IMÁGENES CON ESTEGANOGRAFÍA

Alan Pomerantz
Agustín Marseillan
Juan José Marinelli

23-06-2014 I.T.B.A
Criptografía y Seguridad

Introducción

El presente informe explica el desarrollo del trabajo práctico número dos de la materia Criptografía y Seguridad.
El mismo consiste en desarrollar un esquema de secreto compartido de *Blackely,* de parámetros k y n. El mismo hace uso del concepto de esteganografía, basándose en el ocultamiento de la información de una imagen en imágenes portadoras llamadas "*sombras" (parámetro n)* y se concibe la idea de la existencia de un valor "*k" (*Menor o igual a n*)* , cantidad mínima de sombras necesarias para recuperar la información ocultada (el secreto).
Consigna: Realizar un programa en lenguaje C que implemente el algoritmo de Secreto Compartido en Imágenes descripto en el documento "Improvements in Geometry-Based Secret Image Sharing Approach with Steganography" cuyos autores son Mustafá Ulutas, Vasif V. Nabiyev y Guzin Ulutas de la universidad de Karadeniz, Turquía.
El programa permitirá
- Distribuir una imagen secreta de extensión ".bmp" en otras imágenes también de extensión ".bmp" que serán las sombras en un esquema (k, n) de secreto compartido.
- Recuperar una imagen secreta de extensión ".bmp" a partir de k imágenes, también de extensión ".bmp"

Manejo de Archivos:

Para el desarrollo del trabajo práctico, se desarrollaron funciones de manejo de archivos, tanto para leer, como para escribir los mismos, y para utilizarlos, tanto en la función de

distribución (encripción) como en la de recuperación (desencripción).
Para el manejo de archivos de imagen del tipo bitMap File, se estudió la estructura de los mismos.
El formato BMP es un formato de archivos binario de imagen bastante simple. Consta de dos partes
- Header de 54 bytes
- Cuerpo de tamaño variable.
El encabezado contiene información acerca del archivo: tamaño de archivo, ancho de imagen, alto de imagen, bits por píxel, si está comprimido, etc

En el cuerpo del archivo bmp, están los bits que definen la imagen propiamente dicha. Las imágenes utilizadas son de 8 bits por píxel, lo que concluye en que son imagenes en tonos de grises: el píxel de valor 0x00 es de color negro y el píxel 0xFF es de color blanco.

Función de Encripción (Distribución de secreto)

Tanto las imágenes portadoras como la imagen secreta se parten en bloques de tamaño k. Según ese valor k, se determina cuántos bits en cada píxel serán destinados a ocultar el valor B que resulta del cálculo de la sombra a partir del bloque secreto. Suponiendo que el esquema de reparto sea (3, n). Cada grupo de 3 pixeles contiene los
coeficientes de un plano. Suponiendo que tomamos el primer bloque de 3 pixeles en las 3 imágenes sombras que se poseen y generamos un sistema de ecuaciones, el esquema es el siguiente:

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases}$$

fig0. Sistema de ecuaciones

Es importante determinar que los coeficientes sean linealmente independientes para que el sistema resulte compatibles determinado, y no con infinitas soluciones. Entonces, si hubiera alguna ecuación que, en combinación con alguna otra genera ese problema hay que cambiarla sutilmente. El cambio propuesto es simplemente añadirle al bit una unidad.

Función de Desencripción (Recuperación del secreto)

La función de desencripción se centra básicamente en la resolución del sistema de ecuaciones propuesto en un esquema de distribución. Teniendo en cuenta el valor del parámetro k, las imagenes son separadas en grupos de dicha longitud. Se toman los primeros k píxeles y se los separa según sus bits más significativos y menos significativos. Se toman los menos significativos concatenados para formar el resultado de la fórmula 1 y los más significativos como coeficientes de la ecuación.
Al realizar este procedimiento con todas las sombras necesarias para recuperar el secreto se obtiene un sistema de ecuaciones el cual se resuelve aplicando el complemento modular al mismo y obteniendo así los valores de los píxeles de la imagen real. Repitiendo este procedimiento para todos los grupos de k píxeles formamos nuevamente la imagen oculta.

Análisis del documento de Ulutas

1. Discutir los siguientes aspectos relativos al documento de Ulutas y sus colegas:

a. Organización formal del documento.

El documento está organizado como un paper de investigación científica. Posee lenguaje formal y concreto, una introducción teórica e histórica del problema, un desarrollo de la solución, y referencias a trabajos pasados. Comienza con el planteo de la existencia de dos esquemas distintos para el secreto compartido, y los contrasta. Luego plantea el Esquema de Blakley, enunciando su correspondiente esquema, y los métodos de encripción y desencripción propuestos. Concluye la explicación con una serie de experimentos realizados en el apartado 4. Experimental Results, para luego finalizar en las conclusiones, siempre referenciando, a lo largo de todo el paper de investigación, a los autores y fuentes citadas.

b. La descripción del paso 7 del algoritmo de reconstrucción.

Asume que dicho set o sistema de ecuaciones no poseen dependencia lineal entre sí, y no establece la metodología ad-hoc para solventar dicha situación, dejando librado a quien desee implementar este esquema, la solución a dicho problema.

c. La notación utilizada, ¿es clara? ¿cambia a lo largo del documento?

La notación utilizada es clara, y mantiene consistencia a lo largo de todo el documento. Cabe aclarar que se requiere de un marco teórico; premisa de la cual parte el paper investigativo,

asumiendo que el lector posee conocimientos básicos de la matemática necesaria y requerida para el correcto entendimiento del presente paper. También hace alusión a funciones criptográficas (MD5) que el lector debe tener conocimiento de la existencia y propósito de las mismas.
En la sección resultados, quizá no resulte del todo claro la presencia de los términos USC-SIPI image y PSNR (cuyos valores oscilan entre 38 y 39, en los ejemplos entregados). Concluimos que el PSNR refiere a Peak signal-to-noise ratio, término utilizado en ingeniería para definir la relación entre la máxima energía posible de una señal y el ruido que afecta a su representación fidedigna

2. En el método original de secreto compartido de Blakley se descartan las sombras que tengan ceros. ¿por qué? ¿Por qué crees que Ulutas y sus colegas no tuvieron en cuenta eso?

En el método original de secreto compartido de Blakley se descartan las sombras que tienen cero, ya que a la hora de realizar el esquema de desencripción, el sistema no presente infinitas soluciones.
En el paper, motivo explicado en ***La descripción del paso 7 del algoritmo de reconstrucción,*** se omite la posibilidad de solventar dicha situación, por lo cual Ulutas y sus colegas, no lo tuvieron en cuenta.

3. Una vez recuperada la imagen secreta, ¿es esta imagen exactamente igual a la imagen ocultada?¿Por qué? (Tener en cuenta sólo la matriz de píxeles, no el encabezado).

A la hora de encriptar una imagen, en la necesidad de crear un sistema de ecuaciones, cuyos individuos no sean linealmente dependientes entre sí (con el objetivo de que el mismo tenga una única solución y sea SCD), se genera una perturbación o modificación ligera de los bits menos significativos de cada byte participante del proceso. Segun el enunciado se indica *'Al modificar puede ser conveniente elegir aleatoriamente cuál píxel modificar y cuánto.'* . El criterio de modificación seleccionado por el grupo consiste en alterar (a los 4 o 3 bits menos significativos (según el parámetro k)), sumándole una unidad, siempre operando modularmente en la unidad correspondiente al parámetro k.
Por ejemplo, si un byte resulta ser linealmente dependiente, a su valor menos significativo (Ej: 13) se le añade una unidad (obteniendo 14) y se le aplica la operación modular (14%15 = 14).
Al realizar la desencripción, desde el punto de partida, estaremos en búsqueda de hallar una imagen que ha sido ya alterada inicialmente, por lo cual, se concluye que, una vez recuperada la imagen secreta, esta difiere de la original.

4. Discutir los siguientes aspectos relativos al algoritmo implementado:

a. Facilidad de implementación

Los imprevistos relativos a la implementación se presentaron conforme se avanzó con el desarrollo del mismo. Inicialmente se comenzó utilizando una estructura que albergaba toda la información relativa al header de la imagen BMP. Se desistió del uso de la misma por problemas de compatibilidad (Se trabajo en entorno UNIX como en entorno Windows).

También se detectaron problemas de compatibilidad con Windows, a la hora de la escritura de la imagen, que no existen en entorno UNIX. Se intentó investigar el origen de la modificación de algunos bytes (Utilizando a notepad++ con el plugin HexEditor) pero no se pudo concluir nada.
Luego se avanzó al desarrollo del trabajo en sí. Comenzando por los métodos de encripción y decriptación referidos a k=2.
La extensión a k=3 fue relativamente directa, ya que en su esencia realizan lo mismo.
Hubo algunos problemas a la hora de definir decalajes de bits, y la ubicación del bit de autenticación, pero las mismas, luego de algunas horas de debugging, fueron solucionadas.

b. Posibilidad de extender el algoritmo para que se usen imágenes en color.

Segun el esquema de colores que se utilice, se requerirán una cierta cantidad de bits, que manejen los valores por cada canal. Habría que modificar también el módulo en el que se trabaja (ya no tenemos 8 bits). y se debe tener cierto cuidado a la hora de modificar los bits y trabajar con las ecuaciones. A simple vista, y conceptualmente, creemos posible, con el conocimiento que se tiene, de realizar una extensión a imágenes en color.

c. Ventajas respecto del algoritmo original de Shamir (mencionar por lo menos 2)

Las ventajas que introduce el algoritmo propuesto por el paper de investigación, frente al esquema original de Shamir, y a otros esquemas similares (citados por el paper) es la aparición del bit de chequeo o autenticación. El mismo permite a quien recupera la imagen, garantizar de que la misma consta del correcto chequeo de paridad y que no fue alterada.
El algoritmo propuesto sugiere la modificación de los bits menos significativos de la imagen, provocando en la misma una

modificación ligera impercetible ante posibles atacantes que pudieran darse cuenta de la existencia de un secreto en la misma.
También, al utilizar (el algoritmo) un approach geométrico, y no polinomial, para albergar secretos de tamaño NxN por ejemplo, no requiere utilizar sombras de tamaño 2Nx2N (Como lo requieren otros métodos de secreto compartido), ya que con una sombra del mismo tamaño (NxN) lo puede lograr.
Esto mejora la complejidad espacial del método, y el ancho de banda utilizado para las distintas operaciones de transmisión.

5. ¿Qué dificultades tuvieron en la lectura del documento y/o en la implementación?

Quizás la mayor dificultad, radicada en el error de comenzar a desarrollar e ir leyendo el paper 'on the go', fue a la hora de comprender el funcionamiento, tanto de la encripción como de la desencripción.
Hubieron algunos problemas de compatibilidad entre los entornos Windows y Unix, a la hora del desarrollo, como se indicó anteriormente.
Y quizás hubiese sido deseable contar con un set de imágenes de prueba mayor, para poder operar con las mismas.
Un error conceptual que inocentemente se incurrió, fue la de intentar guardar el secreto, en 3 sombras que eran equivalentes al secreto. Ésto provocó que todos los sistemas de ecuaciones resultaran Linealmente Dependientes. El resultado fue una imagen totalmente difusa, a la hora de desencriptar el secreto.
Se hizo mucho hincapié en las explicaciones brindadas en la clase práctica de la materia, en donde se indicó, entre otras cosas, como se debería haber obtenido el bit de autenticación

(concatenaciones, y operaciones necesarias), y algunos detalles de implementación general.

6. ¿Qué extensiones o modificaciones harían a la implementación o al algoritmo?

Cuestiones relacionadas a la reusabilidad del código para extenderlo, son posibles modificaciones que se hubieran realizado a la implementación y al algoritmo, si se hubiese contado con mayor tiempo para realizarlo.
La posibilidad de incluir k=4 no es una extensión tan directa como se desearía. Lo mismo respecto a la posibilidad de que se opere con imágenes a color.

7. ¿En qué situaciones aplicarían este tipo de algoritmos?

Este tipo de algoritmos, categorizados dentro de lo que se denomina 'esteganografìa', se aplican en situaciones en donde se desea comunicar algo a alguien, a través de un canal encubierto. El algoritmo aplica a imagenes, pero la idea se puede extender a audios, videos, etc, respetando la idea original y los conceptos propios de la esteganografìa.
También, como un indica un ejercicio de la guía de ejercicios prácticos de la materia, el sistema de secreto compartido puede utilizarse para transmitir información entre n individuos, generando la restricción de que k <=n individuos (y no menos) son necesarios para la recuperación del mensaje original.

Resultados

Se realizaron pruebas tanto de encripción como desencripción, utilizando las imágenes provistas por la cátedra.

A continuación se muestran los resultados de la desencripción.

<u>Desencripción k=2</u>

A partir de las siguientes sombras (provistas por la cátedra)

fig1. y fig2. Imagenes Sombra

y se obtuvo la siguiente imagen, que era secreta

fig3. Imagen Secreta (Logo de una reconocida banda)

Desencripción k=3

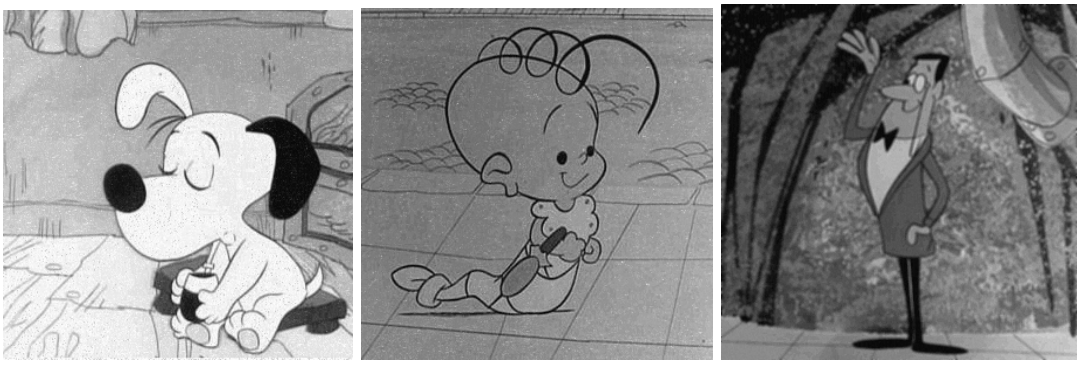

fig 4. Sombras para k=3

y se obtuvo la siguiente imagen, que era secreta

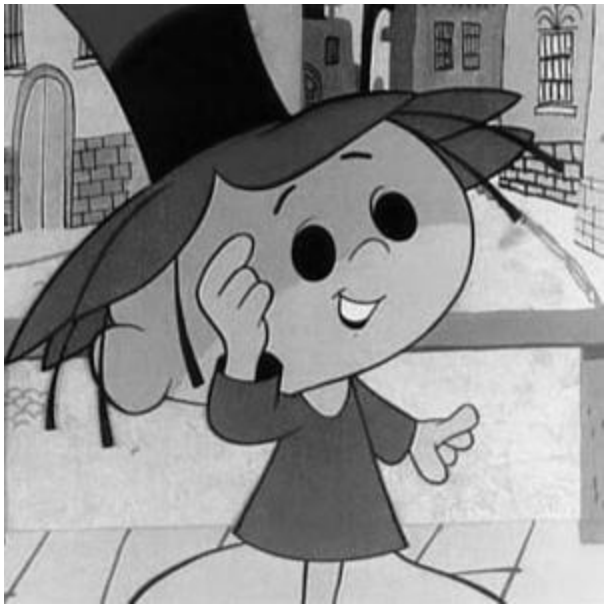

fig5. Imagen Secreta (Hijitus)

Referencias

Capítulo 15 de Computer Security – Art and Science, Matt Bishop, Addison-Wesley, 2004
Capítulo 10 y 12 de Handbook of Applied Cryptography, Alfred J. Menezes, Paul C. Van
Oorschot, Scott A. Vanstone, CRC Press, 1997
"Improvements in Geometry-Based Secret Image Sharing Approach with Steganography", de
Mustafá Ulutas, Vasif V. Nabiyev, y Guzin Ulutas.
"Secreto Compartido", de Ana María Arias Roig

Sobre Criptografía Visual
Página de Criptografía visual de Doug Stinson:
http://cacr.uwaterloo.ca/~dstinson/visual.html
"Visual Cryptography", Moni Naor y Adi Shamir.
http://www.wisdom.weizmann.ac.il/~naor/PUZZLES/visual_pap.ps.gz

Sobre Formato BMP
http://www.fileformat.info/format/bmp/corion.htm