Criptografía y Seguridad (72.04)

"SECRETO COMPARTIDO EN IMÁGENES CON ESTEGANOGRAFÍA"

- Informe -

24 de Junio de 2019

Grupo 12

Katan Johnathan 56653
Heimann Matías 57503
Lo Coco Juan Pablo 57313

Índice

Consideraciones Previas

Cambiamos el parámetro -dir del ejecutable por -l (L minúscula).

Objetivos

- Introducirlos en el campo de la criptografía visual y sus aplicaciones, a través de la implementación de un algoritmo de Secreto Compartido en Imágenes.
- Introducirlos en el campo de la esteganografía y sus aplicaciones.
- Implementar y analizar un algoritmo descripto en un documento científico.

Cuestiones a Analizar

Interpretación de los Documentos.

1. Discutir los siguientes aspectos relativos al documento.
    a. Organización formal del documento.
    b. La descripción del algoritmo de distribución y la del algoritmo de recuperación.
    c. La notación utilizada, ¿es clara? ¿cambia a lo largo del documento? ¿hay algún error?

Previo al realizamiento de la implementación del algoritmo procedimos a la lectura de los documentos donde Li Bai y Azzahra y Sugeng exponen sus propuestas de imagenes secretas usando matrices de proyecciones.

En el documento de Li Bai, encontramos que hay algunas operaciones del ejemplo que están mal hechas (tienen errores), por ende al seguir el ejemplo uno puede perderse o deducir que se encuentran mal las operaciones. Por otro lado no queda claro cómo genera las matrices G en el ejemplo, y la notación en la explicación llega a confundir más. De esta situación Azzahra y Sugeng no se salvan. Si bien explican un poco mejor cómo proceder con la realización de las matrices Gi, algunos índices pueden no llegar a entenderse del todo.

Ambos documentos tienen la desventaja de que solo realizan un único ejemplo, tomando n = 4 y k = 2, lo cual para la implementación se complicaba probar ejemplos de n = 8 y k = 4.

Algo rescatable de ambos documentos es que deciden llamar a las matrices intermedias de la misma manera, lo cual ayuda a la modularización de código para la integración.

2. ¿Por qué la propuesta de Azzahra y Sugeng supone una mejora a la propuesta de Li Bai?

La propuesta de Azzahra y Sugeng supone una mejora a la propuesta de Li Bai dado que puede asegurar que las imágenes recibidas no provengan de fuentes deshonestas y/o falsas. Otra mejora implica que cada proveedor de imágenes puede validar su imagen para autenticar el secreto. Esta mejora lo realiza mediante la imagen de “Marca de Agua” que permite la verificación de una imagen.

Implementación de Algoritmo

3. ¿Qué dificultades se encuentran al elegir pares (k, n) distintos de los establecidos en este TP?

La principal ventaja de usar pares de (k,n)={(2,4), (4, 8)}, es que se puede usar imágenes del mismo tamaño de la imagen a cubrir que de las imágenes portadora y que además uno puede saber cuantos bits tengo que guardar en un byte. Esta última disminuye la complejidad de implementación; en caso de tener valores distintos habría que hacer conjuntos de bytes de diversos tamaños.

4. ¿Por qué es importante controlar el rango de A y el resultado de At .A?

Es importante que el rango de A sea k porque en el caso de que sea menor no se va a poder calcular la proyección. También, el resultado de A transpuesta * A es importante porque lo necesitamos para sacar la inversa de este resultado y así calcular la proyección.

Básicamente si estos resultados no dan como se espera no se va a poder distribuir el secreto ya que nunca se van a poder armar las matrices doble S y R, y así el resto de las que derivan de estas.

5. ¿Por qué es válida la forma de generar los Xi?

Es válida porque utilizando bases de exponentes pertenecientes a aritmética modular 251 al elevarlos desde 0 a k-1 podemos obtener todos valores linealmente independientes. Ya que el primero va a ser 1 para todos y el segundo va a ser el número elegido como base. Entonces todos quedan linealmente independientes y se cumple la condición pedida para el armado de vectores Xi

6. La imagen RW que se obtiene es una imagen “con ruido”. ¿Sería necesario ocultarla mediante esteganografía?¿Cómo podría hacerse?

No hace falta utilizar esteganografía, ya que, para poder darle uso a las matrices Rw y así recuperar la imagen marca se debe realizar previamente un recupero de la matriz R, la cual está puesta con esteganografía en las imágenes portadoras

7. ¿Por qué siempre hay que indicar n, aún al recuperar?

Al recuperar, para saber la cantidad de matrices Sh y matrices Rw que se deben formar y sus filas y columnas, se debe conocer el valor de n. Este valor, solo con k, no es posible calcularlo para un par (k,n) genérico, y por eso es que se debe enviar el valor de n aún al recuperar.

Manejo de archivos BMP

8. ¿En qué otro lugar puede guardarse el número de sombra?

Un lugar posible para guardar el número de sombra es en una posición determinada del arreglo de pixeles de la imagen. Como no voy a usar más de un byte para identificar el número de sombra, es suficiente un solo byte del arreglo de pixeles. También, al ser una imagen de 24 bits por pixel, realizar un cambio en un solo byte es imperceptible al ojo.

Por ejemplo, se podría guardar el número de sombra en el primer byte del arreglo de pixeles de la imagen, teniendo en cuenta que no se esté pisando ningún otro valor guardado previamente por esteganografía. Se podrían utilizar los bits del primer byte que no fueron usados por algún método de esteganografía (LSB o LSB2) previamente.

Aspectos relativos al Algoritmo Implementado

9. Discutir los siguientes aspectos relativos al algoritmo implementado:
    a. Facilidad de implementación
    b. Posibilidad de extender el algoritmo o modificarlo.

Como grupo aceptamos la propuesta de la cátedra de distribuir en tareas los distintos elementos que conforman al problema para que una vez en funcionamiento puedan ser integrados. De paso, distribuimos las tareas entre los integrantes de grupo. Al mismo tiempo observamos que otros grupos tomaron las mismas estrategias para la distribución de tareas y cada uno de los integrantes de nuestro grupo comentaba con sus pares de otros grupos las soluciones implementadas. Esto nos permitió poder comparar y sugerir mejores propuestas que hicieron que la difícil tarea de la implementación se alivianara.

Luego de distribuir las tareas entre los miembros del equipo y realizar los distintos módulos del programa por separado, realizamos una serie de tests antes de integrarlos, para asegurarnos que funcionen correctamente y evitar problemas en la integración de los módulos.

En cuanto a la parte de la resolución de matrices se procedió al armado de una librería de operaciones matriciales. La complejidad vino del lado que al realizar operaciones

en aritmética modular, en especial la resolución de sistemas de ecuaciones, no hay tantos ejemplos como para comprobar los resultados.

Sobre la implementación de la librería de manejo de imágenes BMP, se implementaron funciones básicas de lectura y escritura de imágenes. La implementación, inicialmente resultó sencilla debido a la gran cantidad de documentación y recursos disponibles para aprender sobre este tipo de formato de archivos. Al integrar este módulo con el resto y realizar pruebas, surgieron complicaciones. Por ejemplo, al recuperar la imagen secreta notamos que se veía una silueta, pero con mucho ruido y platinada. Después de investigar durante un tiempo, logramos identificar que entre el header y el bitmap había una serie de bytes que indicaba una paleta de colores. Al incluir esa paleta en las imágenes en blanco y negro, se lograban ver bien.

Respecto a esteganografía, la implementación básica de ocultamiento de mensajes tanto con el método LSB como LSB2 resultó sencilla y fácil de testear. Las complicaciones surgieron cuando se tuvo que integrar este módulo con el módulo de manejo de imágenes BMP y de matrices, al esconder y extraer la información de las imágenes portadoras. Si bien cada módulo se testeó por separado para evitar problemas en la integración, esta integración resultó complicada, surgiendo varios errores en el proceso, hasta que se pudieron solucionar. Creemos que en parte esta complicación surge de la dificultad de debuggear la información que se maneja entre estos módulos, ya que varias veces la imagen recuperada que se obtenía era una imagen solo con ruido, y no se lograba identificar si el error pertenece a la implementación del algoritmo de los papers, o si tenía que ver con el manejo de imágenes bmp, o con esteganografía.

La implementación se redujo al procesamiento de imágenes en blanco y negro donde cada pixel podía tomar un valor entre 0 y 255. En estas imágenes dado una imagen, se calculaban sumas de bits (sombras) que luego se agregan a los bit menos significativos de imágenes portadoras, provocando en las imágenes portadoras lo que conocemos como ruido. Las sumas de bits se encuentran en valores en módulo 251, dado que 251 es el numero primo mas grande, menor a 255.

Este valor base de 251 podría ser extendido para cualquier otro primo y el número de 255 también podría ser extendido para cualquier otro número. Esta posibilidad de extensión nos permite pensar que tranquilamente podría ser extensivo a imágenes en formato RGB, donde tendremos mayor cantidad de valores. Así como usamos imagenes a color, tambien podriamos realizar estenografía mediante audio o video, u otro formato de archivo. Lo unico que cambiaria en estos elementos sería la forma de representar los problemas, como se agregaría el ruido organizado (las sombras en nuestro problema) a las otras representaciones y la base de la aritmética modular para resolver el problema.

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

Investigando encontramos que el concepto de secreto compartido tiene muchas aplicaciones en la Informática. Uno de las principales aplicaciones que tiene la tecnica de secreto compartido en la actualidad es en dejar marcas de agua digitales a los productos,

tal como los derechos de autor, datos de proveedor, entre otros. Esto lo haríamos distribuyendo una cierta información(que puede ser los datos de fabricante) entre N productos, que pueden ser idénticos o no, de tal manera que pueda distribuirlos con su “ruido específico” y para identificar que pertenecen a mi gama de producto verificó el producto con ruido con la marca de agua. De esta manera, si alguien quisiera sacarle la información que tiene el producto para poner sus datos al producto, debería juntar todos los productos de la gama, conseguir la información guardada con estenografía y extraer la información original para modificarla.

Conclusiones

En primer lugar, este trabajo práctico resultó ser una excelente oportunidad para aprender más en detalle el funcionamiento de un esquema de secreto compartido, y cómo implementarlo. También resultó interesante aprender cómo se manipulan y representan las imágenes en formato BMP, y como encriptarlas a través del esquema propuesto en los papers.

Así como resultó interesante los aprendizajes obtenidos durante el desarrollo del trabajo práctico, el mismo también nos demostró las dificultades que conllevan la implementación de un algoritmo de secreto compartido aplicado a imágenes. Entre estas dificultades, algunas mencionadas previamente en el informe, cabe destacar el proceso de integración de los distintos módulos que componen la implementación del algoritmo. Durante este proceso es que se presentaron la mayor cantidad de problemas, aun habiendo realizado tests en cada uno de los módulos.

A su vez, el trabajo práctico presentó una oportunidad para reflexionar acerca de la esteganografía como método de ocultación de la información. Si bien es un método de protección de información no llega a considerarse como una metodología de criptografía. A pesar de esto se utilizan técnicas esteganográficas para aumentar la seguridad de sistemas que son encriptados.

Por último, las cuestiones a analizar presentadas por la cátedra nos permitieron reflexionar mejor sobre el esquema propuesto de secreto compartido, sus características, y posibles alternativas en la implementación.