



Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Eléctrica

“Esteganografía”

Curso:	EL65C
Profesores:	Alberto Castro R. Rodrigo Werlinger C.
Integrantes:	Paula Uribe J. Jorge Wuth S.
Fecha:	21/09/05

Introducción

En el marco del curso de Seguridad en Telecomunicaciones se realiza el presente trabajo de investigación que apunta a estudiar un sistema de seguridad particularmente desconocido para nosotros.

La idea es poder obtener un rango amplio de sistemas de seguridad para poder compararlos entre ellos y evaluar ventajas y desventajas de cada uno, para poder tener una opinión fundada a la hora de elegir un sistema a utilizar.

El presente trabajo está estructurado de forma de tener tanto un marco teórico que respalde esta técnica, así como ejemplos de implementación. Adicionalmente se incluyen aspectos teóricos de seguridad y un estudio de ésta

Objetivos

- Conocer cómo opera este sistema de seguridad en la teoría
- Lograr implementar y entender implementaciones prácticas de la esteganografía
- Discutir ventajas y desventajas
- Estudio de la seguridad de la esteganografía

¿Qué es la Esteganografía?

La esteganografía es el arte y la ciencia de ocultar información cuya presencia no puede ser detectada. Motivado por la creciente preocupación sobre la protección de la propiedad intelectual en Internet, el interés en técnicas para ocultar información ha ido creciendo durante los últimos años.

Esta técnica consiste precisamente en ocultar un texto o cualquier otra información, dentro de un archivo de gráfico o de audio, asegurado con una clave de acceso sólo conocida por la persona que creó dicho archivo, quien también será el encargado de hacerla saber a quien tenga que descubrir el contenido de dicha foto o audio, sin la cual sería casi imposible de obtener. Al ver la fotografía o escuchar el archivo de audio, no será notado nada extraño, pero si contamos con un programa designado a tal fin (la esteganografía) y la clave de acceso obtenida de quien creó el archivo, podremos descubrir lo que lleva dentro.

La esteganografía con un adversario pasivo está muy bien ilustrado en el “*problema de los prisioneros*” de Simmons. Alice y Bob están en la cárcel y quieren crear un plan de escape. Toda su comunicación es observada por el adversario (la guardia, Eva), quien impedirá su plan transfiriéndolos a una cárcel de alta seguridad tan pronto como detecten alguna señal de mensaje oculto. Alice y Bob triunfarán si Alice puede enviar información a Bob sin que Eva sospeche.

Ocultar información de adversarios activos es un problema diferente, porque la existencia de un mensaje oculto es públicamente sabida. La esteganografía contra adversarios activos puede ser dividida en sello de agua y huella digital. El sello de agua reemplaza objetos digitales con una identificación de origen, todos los objetos se marcan de la misma manera. La huella digital, por el contrario, apunta a identificar copias individuales de un objeto, agregándole una marca única en cada copia que es distribuida. Si luego, es descubierta una copia ilegal, el dueño de los derechos puede identificar al comprador decodificando la información oculta.

Un modelo común y terminología para ocultar información es el que se describe a continuación: un mensaje original, inalterado se llama “*texto cubierto*”. El emisor trata de ocultar un mensaje agregado, transformando el texto cubierto utilizando una clave secreta. El mensaje resultante se llama “*estegotexto*” y es enviado al receptor. Similar a la criptografía, se asume que el adversario tiene completo conocimiento sobre el sistema, excepto por la clave secreta que comparten receptor y emisor, que garantiza seguridad.

Se da una noción de seguridad que está basada en test de hipótesis: mediante observación de un mensaje enviado por Alice, el adversario tiene que decidir si corresponde a un texto cubierto “C” o contiene un mensaje agregado y corresponde a un estegotexto “S”. Este es el problema de distinguir dos explicaciones diferentes para los datos observados que se investiga en estadística y en teoría de la información como “*test de hipótesis*”. Se usa la función de entropía como la medida básica de la información contenida en una observación.

Modelo y Definición de Seguridad

La figura 1 muestra nuestro modelo de estegosistema. Eva observa un mensaje que es enviado desde Alice hacia Bob. Ella no sabe si Alice envía el legítimo texto cubierto C o el estegotexto S conteniendo información oculta para Bob. Modelamos esto dejando a Alice operar estrictamente en uno de los dos modos: ella está activa (y su salida es S) o inactiva (enviando texto cubierto C).

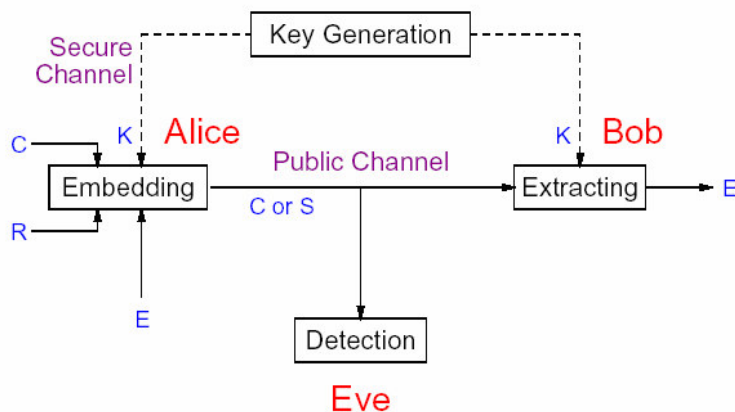


Figura 1: el modelo de un estegosistema con clave secreta con adversario pasivo. Muestra el mensaje agregado E , el texto cubierto C , el estegotexto S , La fuente aleatoria privada de Alice R , y la clave secreta K compartida por Alice y Bob. Alice está enviando ya sea texto cubierto C o estegotexto S .

Si Alice está activa, ella transforma C para que contenga un mensaje agregado E, usando una clave secreta, K. Alice puede usar una fuente privada aleatoria R para agregar. La salida del proceso de ocultar la información es el estegotexto S. Bob debe ser capaz de recuperar E a partir del estegotexto S, y de la clave K. Expresado en términos de entropía, el sistema satisface:

1. $H(S|CEKR)=0$. El estegotexto es determinado únicamente por las entradas de Alice.
2. $H(E)>0$. Hay incerteza sobre el mensaje agregado.
3. $H(E|SK)=0$. Bob debe ser capaz de decodificar el mensaje agregado únicamente.

Si Alice está inactiva, ella envía texto cubierto C y no hay mensaje agregado.

Eva, en base a observación del mensaje enviado debe decidir si fue generado de acuerdo a la distribución del texto cubierto C o de acuerdo a la distribución modificada del estegotexto S , es decir, si Alice está activa o no. Como esta tarea es un problema de test de hipótesis, cuantificamos la seguridad de un estegosistema en términos de la entropía relativa (distancia) entre P_C y P_S .

Teorema 1: Un estegosistema como el definido antes con texto cubierto C y estegotexto S se dice ϵ -seguro contra adversarios pasivos si

$$D(P_C || P_S) \leq \varepsilon$$

Donde D es la entropía relativa entre dos distribuciones (P_C y P_S).

Si $\varepsilon=0$, el estegosistema se dice perfectamente seguro.

Consideremos el proceso de decisión de Eva para una regla de decisión particular, dada por una partición binaria (C_0, C_1) de un conjunto C de posibles textos cubiertos. Ella decide que Alice está activa si y sólo si el mensaje observado c está contenido en C_1 . Idealmente, ella siempre detectaría un mensaje oculto (pero esto ocurre sólo si Alice elige una codificación tal que el texto cubierto y el estegotexto son disjuntos). Si Eva falla al detectar que está observando un estegotexto S , ella comete un error de tipo II, con probabilidad β . El error opuesto es el error de tipo I: Eva decide que Alice envía un estegotexto cuando realmente se envió un texto cubierto C . Como un caso especial, se puede asumir que Eva nunca comete un error de tipo I. El error de tipo I tiene probabilidad α .

Así, de la desigualdad

$$d(\alpha, \beta) \leq D(P_{Q_0} \| P_{Q_1})$$

con $d(\alpha, \beta) = \alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta}$: entropía binaria relativa, se obtiene:

Teorema 2: En un estegosistema que es ϵ -seguro contra adversarios pasivos, la probabilidad β de que el adversario no detecte un mensaje oculto y la probabilidad α de que el adversario falsamente detecte un mensaje oculto satisface

$$d(\alpha, \beta) \leq \epsilon$$

En particular, si $\alpha=0$, entonces

$$\beta > 2^{-\epsilon}$$

donde δ es el límite superior de $D(P_C \| P_S)$

En un sistema perfectamente seguro tenemos que $D(P_C \| P_S)=0$ y por lo tanto $P_C=P_S$. Entonces, no se puede obtener información de si Alice está activa sólo observando el mensaje.

Estegosistemas incondicionalmente seguros

Asumimos primero que la distribución de texto cubierto es conocida y se diseñan funciones para ocultar mensajes correspondientes.

Asumiendo que el texto cubierto es un string uniformemente distribuido de n bits, para un n positivo. El generador de claves elige la clave K de n bits con distribución uniforme y la envía a Alice y Bob. La función para ocultar información (si Alice está activa) consiste en realizar un XOR entre este mensaje particular de n bits, e , y K , entonces $S=e \oplus K$, y Bob puede decodificarlo haciendo la operación $e=S \oplus K$. El estegotexto resultante, S , está uniformemente distribuido en el conjunto de strings de n bits, y por lo tanto, $D(P_C \| P_S)=0$. Entonces, la conversión de un paso provee perfecta seguridad esteganográfica si el texto cubierto es aleatorio uniformemente.

Esta técnica oculta una imagen monocromática, partiéndola en dos grupos aleatorios de puntos. Cuando estos son superpuestos, la imagen aparece. También es posible producir dos imágenes, tal que ambas juntas revelen el mensaje oculto.

Para distribuciones generales de texto cubierto, se describe ahora una sistema que oculta un mensaje de 1 bit en estegotexto. La extensión para mensajes de mayor largo es directa. Sea el texto cubierto C con tamaño $|C|$ con una distribución arbitraria P_C . Alice construye la función de agregado de una partición de C en dos partes, tales que a ambas partes se le asigna aproximadamente la misma probabilidad. En otras palabras, sea

$$C_0 = \min_{C' \subseteq C} \left| \sum_{c \in C'} P_C(c) - \sum_{c \notin C'} P_C(c) \right| \quad \text{y} \quad C_1 = C \setminus C_0$$

Alice y Bob comparten una clave de 1 bit, $K \in \{0,1\}$. Definamos C_0 como una variable aleatoria designada por C_0 y distribución P_{C_0} igual a la distribución condicional $P_{C|C \in C_0}$ y definamos C_1 similarmente sobre C_1 . Entonces Alice opera el estegotexto para agregar un mensaje $e \in \{0,1\}$ como $S = C_e \oplus K$. Bob puede decodificar el mensaje porque conoce que $e=0$ si y sólo si $S \in C_k$.

Teorema 3: El estegosistema de mensaje de un bit descrito antes es seguro contra adversarios pasivos.

En efecto, sea $\delta = P[C \in C_0] - P[C \in C_1]$. Mostramos sólo el caso $\delta > 0$. Es directo verificar que

$$P_S(c) = \begin{cases} P_C(c)/(1+\delta) & \text{Si } c \in C_0 \\ P_C(c)/(1-\delta) & \text{Si } c \in C_1 \end{cases}$$

Sigue que:

$$\begin{aligned} D(P_C \| P_S) &= \sum_{c \in C} P_C(c) \log \frac{P_C(c)}{P_S(c)} \\ &= \sum_{c \in C_0} P_C(c) \log(1+\delta) + \sum_{c \in C_1} P_C(c) \log(1-\delta) \\ &= \frac{1+\delta}{2} \cdot \log(1+\delta) + \frac{1-\delta}{2} \cdot \log(1-\delta) \\ &\leq \frac{1+\delta}{2} \cdot \frac{\delta}{\ln 2} + \frac{1-\delta}{2} \cdot \frac{-\delta}{\ln 2} \\ &= \frac{\delta^2}{\ln 2} \end{aligned}$$

usando que $\log(1+x) \leq x/\ln 2$.

Supongamos que el mensaje oculto fue agregado antes de que se aplicara compresión de datos a S . La compresión de datos es un proceso determinístico, por lo tanto, el lema 1 es válido y muestra que si comenzamos con un estegosistema ϵ -seguro, la seguridad del sistema comprimido es también a lo más ϵ . Para ponerlo de otro modo, la compresión de datos no puede dañar la seguridad del estegosistema.

Implementaciones

A continuación se describirá como ocultar ficheros o información de manera manual o con la ayuda de algún software esteganográfico, además se analizará la seguridad que ofrecen las distintas técnicas.

Esteganografía manual

Es posible ocultar información de manera manual para lo cual se necesita un editor hexadecimal, un archivo que será el portador de la información y, por su puesto, la información que se desee ocultar. Existe un método muy simple para ocultar información y consiste en agregarla después del **EOF** (End Of File) del archivo portador de manera que sea ignorada por el programa que interprete dicho archivo. Esto se ilustra con el siguiente ejemplo.

La información a ocultar será el texto *"Este texto está oculto"*, el editor que se utilizará será WinHex y el archivo portador será la imagen JPG que se muestra a continuación



Si abrimos el archivo imagen con el editor veremos en su parte final las siguientes líneas.

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
00013120	FA	85	14	51	48	41	45	14	50	01	45	14	50	01	45	14	ú . QHAE . P . E . P . E .
00013136	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013152	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013168	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013184	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013200	50	01	45	14	50	01	45	14	50	07	FF	D9					P . E . P . E . P . E . P . E .

La primera columna indica el offset, la columna central es la vista hexadecimal de la imagen y la tercera columna es la interpretación en texto de los valores hexadecimales correspondiente al código ASCII. Se puede observar la cadena **FF D9** al final del archivo, esta cadena le indica al programa que muestra la imagen (programa interprete, por ejemplo Paint) el fin de la imagen, corresponde al **EOF** para el formato **JPG** por lo que todo lo que se encuentre después de dicha cadena no se visualizará ni interferirá con la imagen original. De esta manera podemos ocultar nuestra información simplemente añadiéndola después del **EOF** quedando como se muestra a continuación:

Offset	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
00013120	FA	85	14	51	48	41	45	14	50	01	45	14	50	01	45	14	ú . QHAE . P . E . P . E .
00013136	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013152	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013168	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013184	50	01	45	14	50	01	45	14	50	01	45	14	50	01	45	14	P . E . P . E . P . E . P . E .
00013200	50	01	45	14	50	01	45	14	50	07	FF	D9	45	73	74	65	P . E . P . E . P . y Û Este
00013216	20	74	65	78	74	6F	20	65	73	74	E1	20	6F	63	75	6C	texto está ocul
00013232	74	65															to

Esta operación se simplifica aún más ingresando la información directamente en la tercera columna que es la de modo texto y el editor se encarga de su correspondiente hexadecimal. Al guardar el archivo y visualizarlo se obtiene la siguiente imagen



Al comparar la imagen original y la modificada no se observan diferencias ya que la información agregada es ignorada por el programa visualizador.

Este método, si bien es simple, es muy inseguro. Cualquier persona que sospeche que se ha ocultado información podrá acceder a ella abriendo la imagen con algún editor hexadecimal, o incluso en un editor corriente donde se mostrará la representación en texto de la imagen observándose claramente el texto "oculto" al final del archivo. A pesar de esto existen softwares de pago que ocultan información de esta manera, incluso hay algunos que aseguran un nivel de cifrado seguro, lo cual es falso ya que no encriptan la información, y

algoritmos de compresión avanzada, lo cual de hecho es cierto pues comprimen con WinZip antes de "ocultar".

Métodos informáticos

Existen dos métodos ampliamente usados para ocultar información y que son seguros y robustos. El primero y más robusto utiliza algoritmos como la transformada de coseno discreta, que está basada en la transformada discreta de Fourier pero utilizando únicamente números reales, es uno de los métodos más seguros y resistentes a ataques que existen.

El segundo método se basa en la manipulación de los bits menos significativos del archivo (**LSB**, Least Significant Bit), es ampliamente usado por los softwares de esteganografía y consiste en alterar los archivos aumentando o disminuyendo en un bit cada grupo de bytes que conforman el archivo, por ejemplo, en un archivo de imagen **BMP** de 24 bits de profundidad cada color se especifica con 24 bits, es decir grupos de 3 bytes, obteniéndose una variedad de 16.777.216 colores, si de cada grupo alteramos el bit menos significativo del último byte se estará cambiando el color de un píxel por el color inmediatamente siguiente, lo cual será imperceptible para el ojo humano dada la gran variedad de colores disponibles y la poca diferencia entre colores adyacentes.

Para poder aplicar este método primero se debe conocer la estructura del formato del archivo, para así saber los grupos de bytes que lo conforman y obtener el bit menos significativo del último byte de cada grupo, ya que de otra manera se podría estar alterando el bit menos significativo pero del primer byte de cada grupo, lo cual provocaría un cambio no despreciable en el archivo portador, o peor aún, se podría alterar el encabezado del archivo dejándolo inutilizable. Otra desventaja de este método es que no es aplicable directamente a archivos que posean algún nivel de compresión como es el caso de archivos **JPG** o **Mp3** entre otros, ya que se debe respetar la estructura del algoritmo de compresión, lo cual es bastante complicado de hacer en forma manual, sin embargo, debido a la gran disminución de tamaño con respecto al archivo sin comprimir estos formatos son muy usados, pero el proceso se realiza con la ayuda de algún software especializado.

A continuación se ocultará un archivo de texto de 1Kb dentro de la imagen portadora anterior, el proceso se realizará con el programa STEGANOS File manager incluido en el Steganos Security Suite 7.

Se escoge la opción Steganos File Manager, en *File* se elige *New encrypted file*, con *Add* le indicamos la información a ocultar, en este caso un archivo de texto, luego al cerrar con *Close encrypted file* pregunta si se desea sólo encriptar o encriptar y esconder, al seleccionar esta última pedirá el archivo portador que puede ser de formato JPG, JPEG, BMP o WAV. Usamos como archivo portador la imagen original del ejemplo anterior. Al finalizar se asigna una clave para poder recuperar la información oculta.

Con el fin de comparar los cambios producidos en la imagen se abren la original y la modificada con el editor hexadecimal, a continuación se muestran ambas imágenes y un extracto de su código hexadecimal.



0576	C5	C6	C7	C8	C9	CA	D2	D3	D4	D5	D6
0592	E3	E4	E5	E6	E7	E8	E9	EA	F2	F3	F4
0608	FA	FF	DA	00	0C	03	01	00	02	11	03
0624	9A	28	A2	AC	E6	0A	28	A2	80	0A	28
0640	80	0A	28	A2	80	0A	28	A2	80	0A	28
0656	77	7A	A5	AD	A1	28	CE	5E	5F	F9	E6
0672	A6	A2	DB	B2	13	69	6A	CB	94	55	1B
0688	F6	34	72	28	CE	D6	C7	23	D4	62	AF
0704	4D	35	74	14	51	45	21	85	14	51	40
0720	14	51	40	05	14	51	40	05	14	51	40
0736	14	51	40	05	14	51	40	05	14	51	40
0752	14	51	40	05	14	51	40	05	14	51	40
0768	35	DC	46	8C	ED	F7	54	12	69	D4	11
0784	FA	AD	D5	E2	FC	AD	E4	44	47	0A	87

0576	C5	C6	C7	C8	C9	CA	D2	D3	D4	D5	D6
0592	E3	E4	E5	E6	E7	E8	E9	EA	F2	F3	F4
0608	FA	FF	DA	00	0C	03	01	00	02	11	03
0624	9A	28	A2	AC	E6	0A	28	A2	80	0A	28
0640	80	0A	28	A2	80	0A	28	A2	80	0A	28
0656	77	5A	A5	AD	A9	28	CE	5E	5F	F9	E6
0672	77	B2	15	FA	97	28	AA	36	3A	9C	77
0688	03	3B	5B	1C	8F	5E	2A	F5	12	8B	8B
0704	A2	8A	43	0A	28	A2	80	0A	28	A2	80
0720	28	A2	80	0A	28	A2	80	0A	28	A2	80
0736	28	A2	80	0A	28	A2	80	0A	28	A2	80
0752	28	A2	80	0A	28	A2	80	0A	28	A2	80
0768	CF	DD	50	4D	3A	82	33	C1	A0	0E	66
0784	53	E4	46	47	0A	87	E6	23	DD	BF	C2

Se puede observar que a partir de la línea 656 los códigos varían, sin embargo las imágenes correspondientes parecen idénticas.

En este método, la integridad de la información oculta dependerá de la integridad del archivo en el que se ocultó, ya que cualquier variación en los bits del archivo portador podría modificar la información oculta. Por esto, si se ha ocultado información en algún archivo, ésta se perdería con algún eventual cambio de formato del archivo portador, los cuales generalmente se hacen para disminuir su tamaño, como ocurre al pasar de WAV a Mp3 de BMP a JPG.

Conclusiones

En la teoría mostrada, se cuenta con 3 teoremas que avalan el buen comportamiento en cuanto a seguridad de este sistema. Haciendo uso de la estadística, es posible construir una base teórica que permite comparar distintos sistemas de seguridad en forma objetiva.

Podemos observar de las imágenes anteriores que es realmente imposible para el ojo humano percibir los pequeños cambios producidos en el archivo portador y que los métodos expuestos ocultan eficazmente cualquier información que pueda representarse por bits, hoy en día todo.

Estos métodos se pueden aplicar a prácticamente todo tipo de archivos aunque los más usados son los archivos de imágenes y sonido que se valen de imperfecciones en nuestros sentidos para así ocultar información.

Se implementaron distintos métodos comprobándose la facilidad e inseguridad de algunos, pero que sin embargo pueden ser útiles para transmitir información siempre que no se levanten sospechas, ya que como se vio un análisis muy simple puede revelar el mensaje oculto.

Es bueno notar que este buen comportamiento en cuanto a seguridad puede en algunos casos representar una característica no deseable. Nos referimos a situaciones en las que se le da un mal uso a las ventajas de la esteganografía, como por ejemplo a ataque terroristas en la red, o a robos de información cuyos resultados son camuflados con esta técnica. Es un tema que queda para la discusión, pero que en ningún caso opaca las ventajas de esta herramienta, por lo que se siguen desarrollando técnicas cada vez más sofisticadas y utilizando en forma crecientemente masiva.

Bibliografía

- Trabajo Práctico “Esteganografía”, Universidad John F. Kennedy, Argentina, 1998
- Christian Cachin, “An Information-Theoretic Model for Steganography”, MIT Laboratory for Computer Science, Cambridge, USA, 1998
- Jessica Fridrich, Miroslav Goljan, Rui Du, “Detecting LSB Steganography in Color and Gray-Scale Images”, State University of New York, Binghamton
- “Introducción a la Esteganografía”, www.death-master.tk
- Niels Provos, Peter Honeyman, “Hide and Seek: An Introduction to Steganography”, IEEE Security & Privacy, Mayo/Junio 2003