

FACULTAD DE CIENCIAS, UNAM



Modelado y programación

Proyecto 2: Esteganógrafo

27 de abril de 2022

Garzón Castro Oscar René

Profesor: José de Jesús Galaviz Casas

Ayudante: Luis Soto Martínez

Ayud. Lab.: Miren Jessamyn Hernández Leyva

1. Definición del problema

Se tiene que realizar un esteganógrafo. El cual va tomar un texto y de cierta forma lo oculta en una imagen. El punto clave es que el esteganógrafo puede develar el texto oculto en la imagen.

2. Análisis del problema.

Para realizar el esteganógrafo se va a utilizar el último bit significativo (LSB) de cada valor rojo, verde y azul en un pixel. De tal forma que para guardar un caracter de 8-bits vamos a necesitar 3 pixeles.

En total vamos a tener 9 bits disponibles. Notemos que el segundo bit de izquierda a derecha siempre va a ser cero, ya que los caracteres usados tienen un valor unicode entre 32 y 127 por lo que el binario mayor va a ser 01111111. Es decir ese número siempre sería par lo cual levantaría sospechas si alguien analiza la imagen. Para evitar esto el segundo bit no se va a modificar y será par o impar de acuerdo al color original. El primer bit es análogo al segundo por lo que tampoco se va a modificar.

Se va utilizar un formato PNG para que la foto no pierda información al guardarse en el disco duro.

Para hacer un uso mas eficiente de los caracteres disponibles a las palabras con acento se les va remover este, las mayúsculas se convertiran a minúsculas y la letra ñ se pasará a la n. Si el texto tiene algún símbolo que no este en 32-127 de los caracteres en ascii el programa lo va remplazar con un espacio en blanco.

3. Selección de la mejor alternativa.

Para este programa se consideró usar Python o Java. La biblioteca PIL en Python está muy bien documentada, además de tener funciones de alto nivel que te ayudan a trabajar con los pixeles sin tener que preocuparte por detalles demasiados técnicos. Las bibliotecas para procesar imágenes en Java no están bien documentadas y tienes que manejar los pixeles a nivel de bits, lo que lo hace demasiado tedioso. Por lo tanto, Python es el lenguaje utilizado para realizar el esteganógrafo.

Además, Python tiene la biblioteca NumPy, con la cual podemos modificar la dimensión del arreglo de manera sencilla. También NumPy y PIL interactúan de manera armónica al pasar de arreglo a imagen y viceversa

4. Pseudocódigo.

Ocultar

1. Obtener el texto a ocultar
2. Pasar el texto a su valor decimal en Unicode y de ahí a binario
3. Obtener el arreglo de pixeles de la imagen donde se va ocultar el texto
4. Crear un nuevo arreglo tal que los LSB de los colores rojo, verde y azul estén modificados para representar los valores en binario del texto.

5. Agregó para indicar que ahí termina el texto oculto
6. Convertir este arreglo a imagen y guardarla en la computadora del usuario.

El arreglo original de la foto tiene tres dimensiones lo que complica moverse por grupos de tres píxeles a través del arreglo. Por lo que se decidió reducir una dimensión del arreglo. Así solo se tiene que avanzar de tres en tres a lo largo del arreglo para avanzar por la imagen. Este enfoque se utiliza tanto para ocultar como para develar.

Develar

1. Obtengo el arreglo de pixeles de la imagen
2. Reduzco una dimensión del arreglo para que sea mas fácil iterar por los pixeles
3. Obtengo el caracter representado en cada tercia de pixeles
4. Añado cada caracter a una string
5. Cuando encuentre paro de buscar en el arreglo
6. Imprimo la string en un archivo
7. Guardo el archivo

El programa verifica que sus entradas sean fotos en formato PNG y archivos de texto que terminen .txt. Si es un formato diferente la entrada será rechazada. También te avisa si el texto es muy largo para guardar en la imagen. Además, alerta si el texto o la imagen no han sido encontradas.

Para que el programa funcione correctamente las imágenes a utilizar tienen que estar en la carpeta esteganografo/imagenes y los textos en esteganografo/textos.

5. Proyección a futuro.

Se podrían añadir más formatos de audio, video o gifs para ocultar textos. Además, se podría encriptar el texto inicial para aumentar la seguridad del mensaje.

Hay muchos estenógrafos en la web que ocultan información en imágenes o archivos de audio. Todos estos sitios son de uso gratuito. Además, este método estenográfico ya es muy conocido por lo que no es viable ocultar mensajes con él. En conclusión, es un proyecto de formación más que algo con lo que se podría generar un ingreso.

6. Referencias

<https://docs.python.org/3/tutorial/inputoutput.html#reading-and-writing-files>
<https://docs.python.org/3/library/index.html>
<https://towardsdatascience.com/image-processing-with-python-5b35320a4f3c>
<https://pillow.readthedocs.io/en/stable/reference/Image.html#functions>
<https://numpy.org/doc/stable/reference/generated/numpy.asarray.htm>