Grupo 2:

stenography-crypto

Informe Trabajo Práctico Especial

72.44 - Criptografía y Seguridad - Primer Cuatrimestre 2022.

Catalán, Roberto José - 59174

Dell'Isola, Lucas - 58025

Pecile, Gian Luca - 59235

Índice

Cuestiones a Analizar

Pregunta 1

Discutir los siguientes aspectos relativos al documento.

a. Organización formal del documento.

b. La descripción del algoritmo.

c. La notación utilizada, ¿es clara? ¿Hay algún error o contradicción?

En cuanto al documento del enunciado, se le podría agregar un índice para mejor organización y a su vez, reducir el largo de algunas secciones como la introducción de modo que no parezca tan abrumador al empezar a leerlo.

La descripción del algoritmo por otro lado, es clara y concisa donde el uso de ejemplos e imágenes lo hace más simple de comprender. El único problema es en la sección del algoritmo del paper que la explicación se siente escueta y al momento de implementar el algoritmo aparecen problemas que se debieron preguntar en clase como por ejemplo el guardado de los patrones en LSBI. Esto se terminó concordando como:

```
LSB1(patron) || LSBInverted(Tamaño real || datos archivo || extensión)
```

La notación utilizada es clara y se asemeja a la vista en clase, la única posible contradicción se da con la definición del empaquetado para LSBI debido principalmente a la falta de claridad en el *paper*.

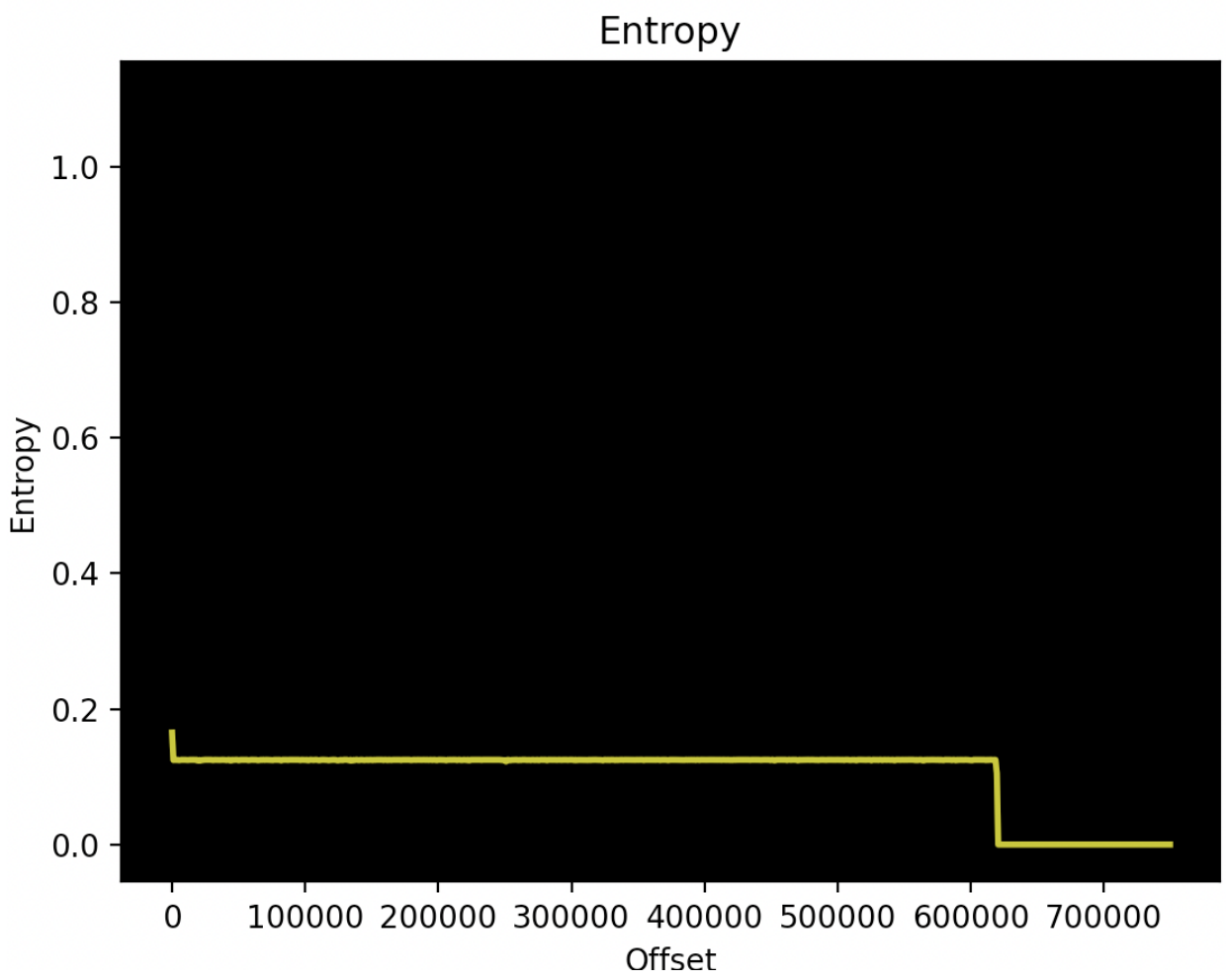
El uso de ejemplos de ejecución y su descripción es de gran utilidad para el entendimiento del programa a realizar.

Pregunta 2

Esteganografiar un mismo archivo en un .bmp con cada uno de los tres algoritmos, y comparar los resultados obtenidos. Hacer un cuadro comparativo de los tres algoritmos estableciendo ventajas y desventajas.

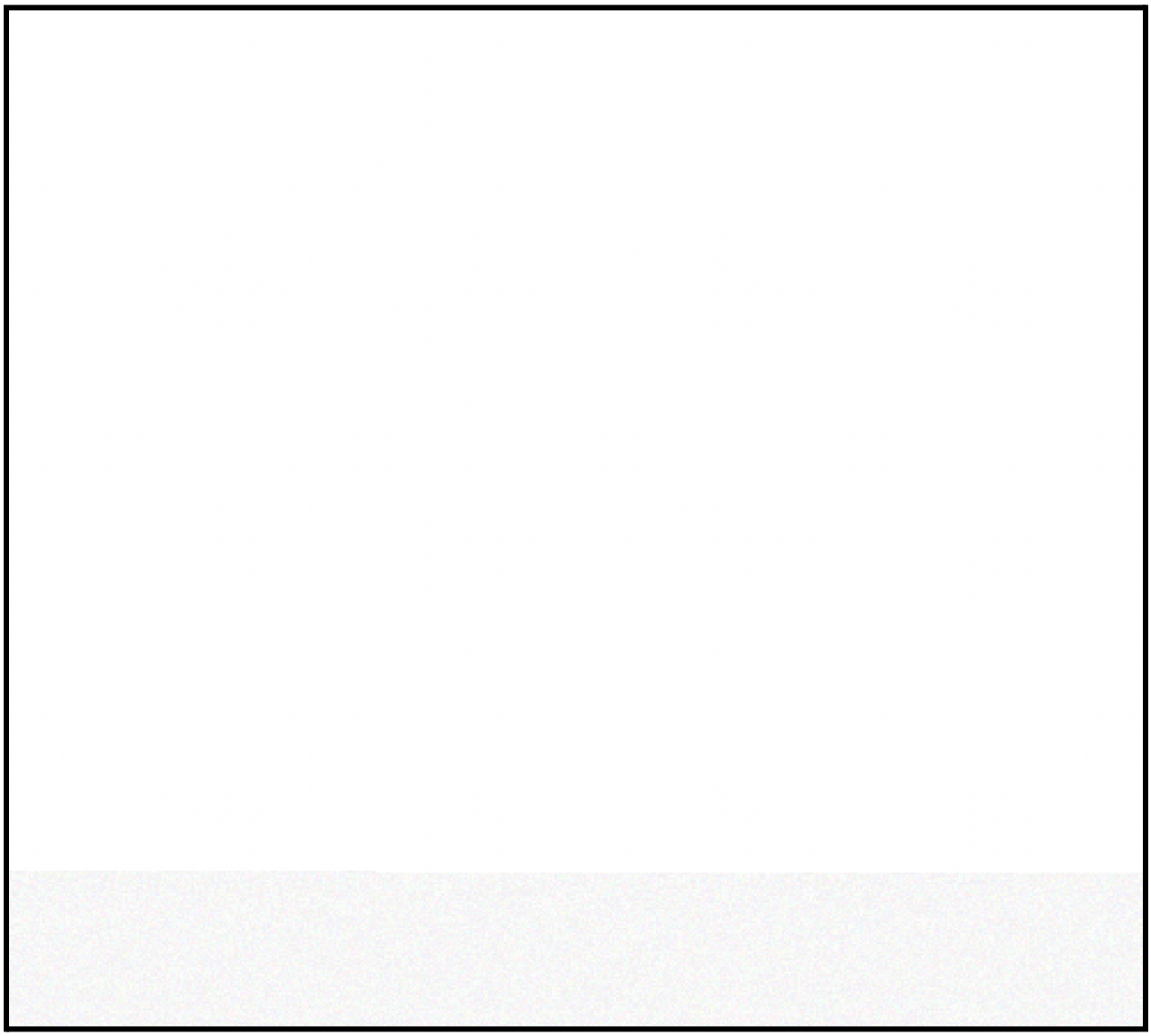
Se esconde el archivo pepe.png -con un tamaño de 77 Kb- en el archivo blanco.bmp -con un tamaño de 770 Kb- (el cual es una imagen blanca para observar el ruido). Luego se procede a mostrar un gráfico de la entropía para el archivo obtenido:

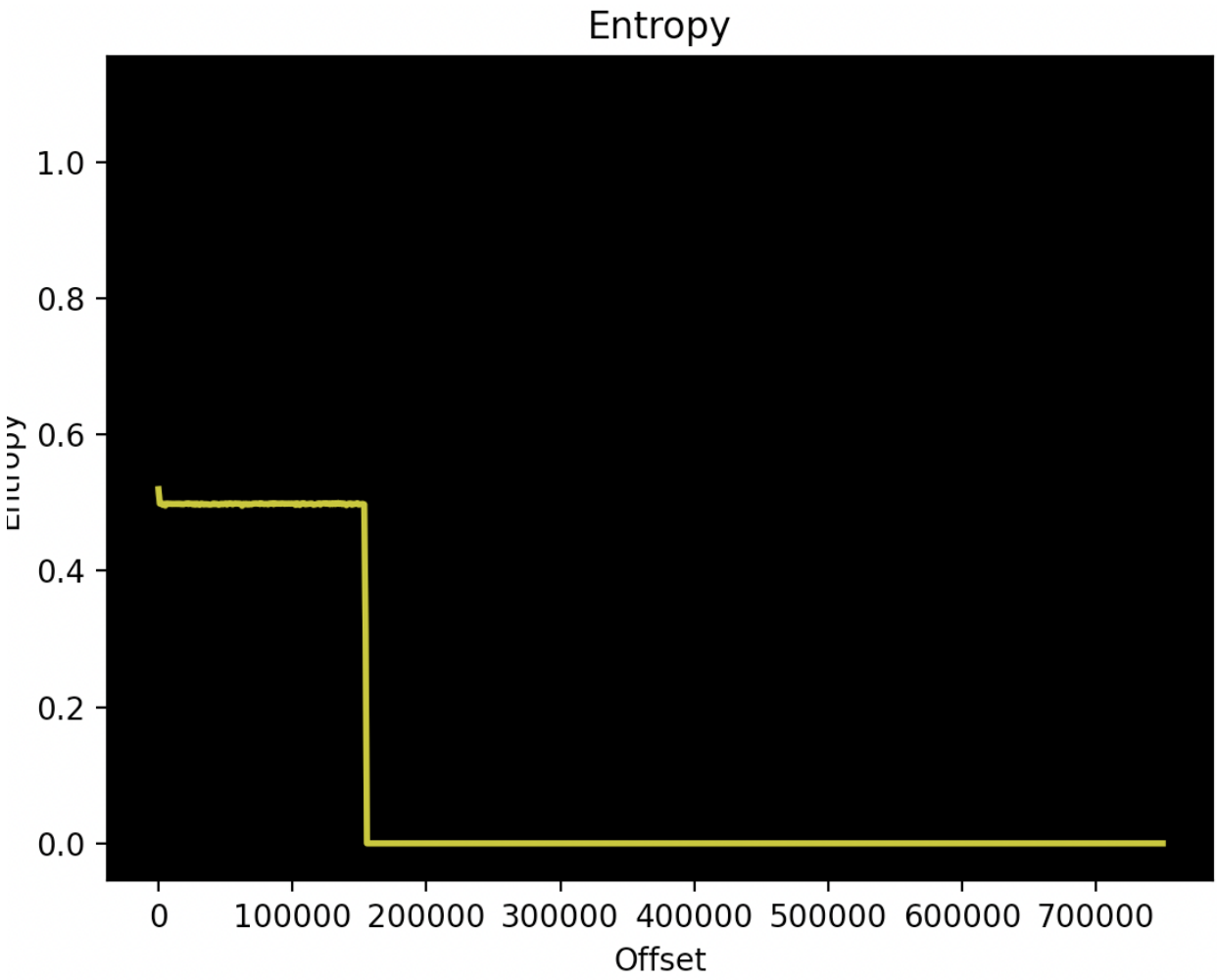
LSB1



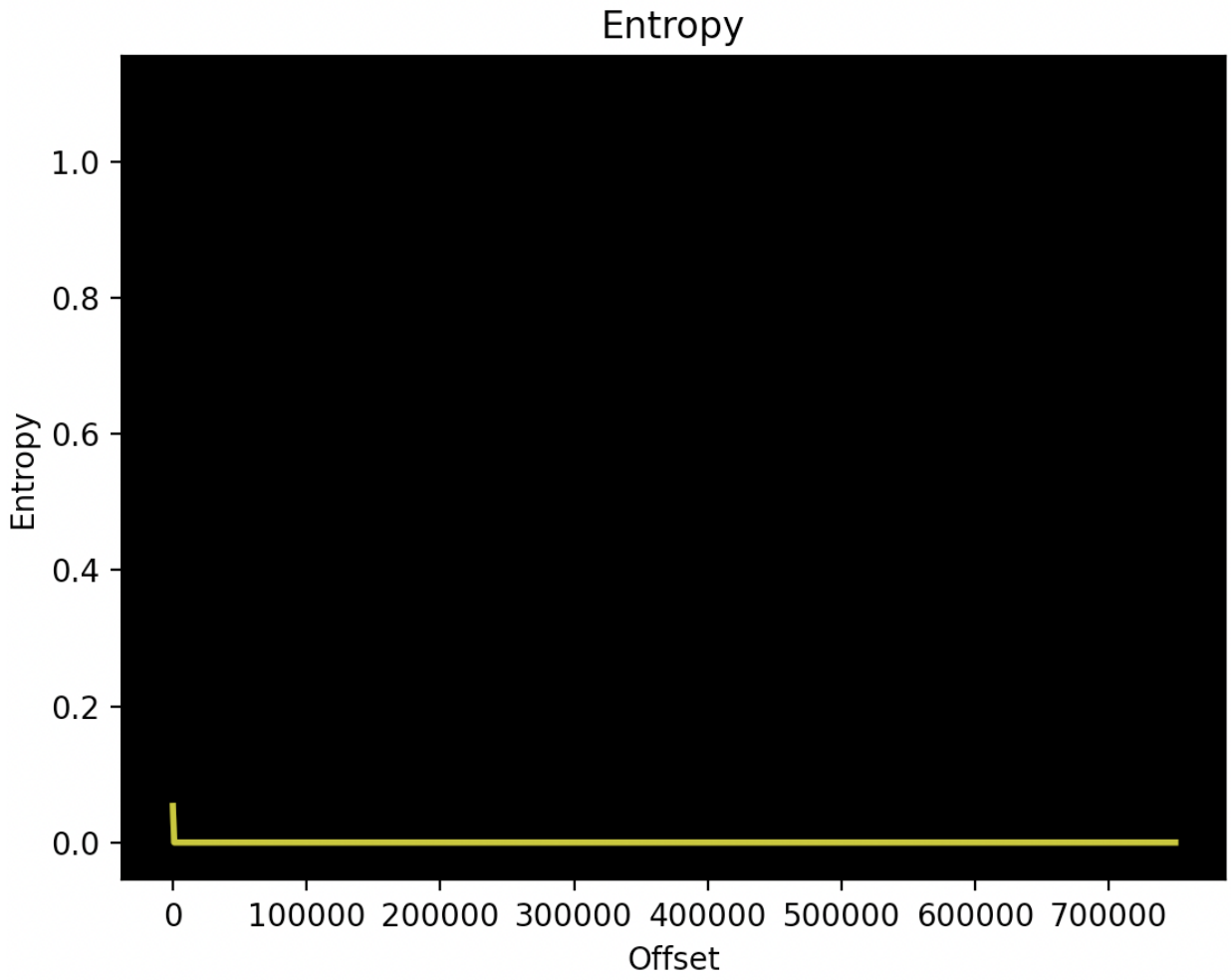
LSB4

Se observa ruido en la imagen blanca (con un recuadro negro) a continuación:




Entropy
1.0
0.8
0.6
0.4
0.2
0.0
Entropy
0
100000
200000
300000
400000
500000
600000
700000
Offset

LSBI


Entropy
1.0
0.8
0.6
0.4
0.2
0.0
Entropy
0
100000
200000
300000
400000
500000
600000
700000
Offset

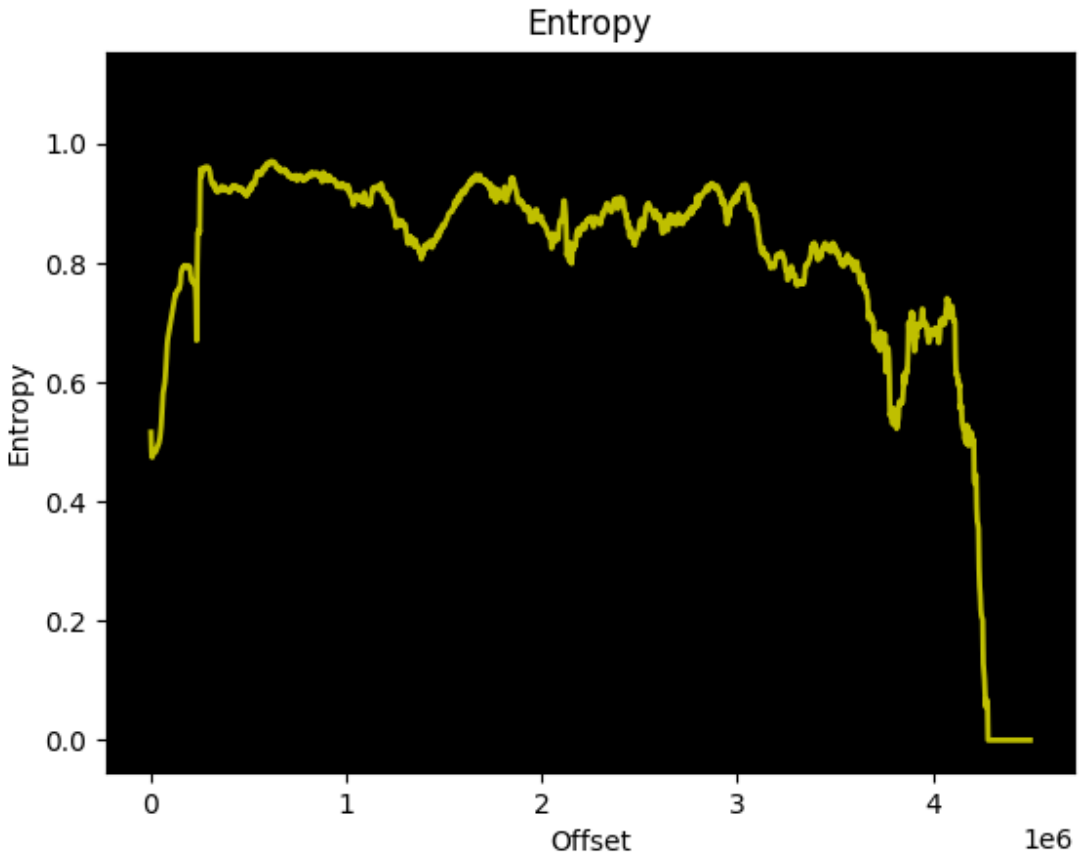
Algoritmo	Ventajas	Desventajas
LSB1	Utiliza el bit menos significativo para ocultar el mensaje, haciendo que las modificaciones a los píxeles sean mínimas o incluso nulas.	El tamaño de los mensajes que puede esconder es muy chico.
LSB4	Permite esconder mensajes de mayor tamaño que otros métodos ya que usa los 4 bits menos significativos.	Hay mucha mayor presencia de ruido que los otros métodos.
LSB Improved	No resulta tan evidente el hecho de que se esconde un mensaje al analizar la entropía.	Se pierde eficiencia al embeber. Requiere 4 bytes más que los métodos anteriores.

Pregunta 3

Explicar detalladamente el procedimiento realizado para descubrir qué se había ocultado en cada archivo y de qué modo. Indicar qué se encontró en cada archivo.

Usando la librería vista en clase, BinWalk, se obtuvieron los gráficos con la entropía para cada imagen a analizar sus contenidos. A su vez se hizo uso de la herramienta HexEdit recomendada por la cátedra para el análisis hexadecimal de los archivos. A continuación se detalla para cada archivo los descubrimientos:

Back



Se puede ver que el archivo no está encriptado por el gráfico de la entropía. Se realizan pruebas con los algoritmos LSB1, LSB4 y LSBI.

Cuando se realizó la prueba con LSB4 se encontró la siguiente imagen:

Analizando el contenido del archivo con el editor hexadecimal se puede ver:

```
0000A790   00 00 00 08  00 00 00 73  6F 6C 32 2E  74 78 74 45  .......sol2.txtE
0000A7A0   8E 41 0E C2  30 0C 04 EF  48 FC 61 1F  80 2A 78 02  .A..0...H.a..*x.
0000A7B0   E2 CA 27 56  49 A0 96 92  B8 4A D2 56  E5 F5 38 E1  ..'VI....J.V..8.
0000A7C0   C0 C5 B2 D7  B3 6B 3B 7A  22 49 26 42  C5 9A 71 9B  .....k;z"I&B..q.
0000A7D0   CE A7 A1 BD  24 5A D1 92  68 32 11 43  2B B4 DD 53  ....$Z..h2.C+..S
0000A7E0   2B 58 9D 08  7C 40 64 FD  2D 2A 42 5A  24 7C 98 D1  +X..|@d.-*BZ$|..
0000A7F0   D4 1B 12 32  AE 3D EA 5E  C5 7A A7 D9  A0 81 45 30  ...2.=.^.z....E0
0000A800   BE B5 48 4B  8A A3 8F C9  78 8B 25 16  D6 BA 6B F1  ..HK....x.%...k.
0000A810   F6 48 63 0F  D0 56 14 2C  6E 96 CD 88  87 E6 BF F5  .Hc..V.,n.......
0000A820   32 6C 38 FE  A6 99 47 7F  7F DA D3 D6  0F 16 59 1A  2l8...G.......Y.
0000A830   07 A1 6E 8D  4D A7 2F 50  4B 01 02 1F  00 14 00 00  ..n.M./PK.......
0000A840   00 08 00 02  8A BC 54 83  81 95 36 98  00 00 00 E8  ......T...6.....
0000A850   00 00 00 08  00 24 00 00  00 00 00 00  00 20 00 00  .....$....... ..
0000A860   00 00 00 00  00 73 6F 6C  32 2E 74 78  74 0A 00 20  .....sol2.txt..
0000A870   00 00 00 00  00 01 00 18  00 76 F2 D9  C1 CF 72 D8  .........v....r.
0000A880   01 76 F2 D9  C1 CF 72 D8  01 66 8C 13  F4 CE 72 D8  .v....r..f....r.
0000A890   01 50 4B 05  06 00 00 00  00 01 00 01  00 5A 00 00  .PK..........Z..
0000A8A0   00 BE 00 00  00 00 00 _                            .......
---  back-LSB4-extract.png       --0xA8A7/0xA8A7--100%-------------------------
```

Buscando la firma del archivo, se pueden ver los bytes `50 4B 03 04` que corresponde a la firma de un archivo de compresión de tipo zip.

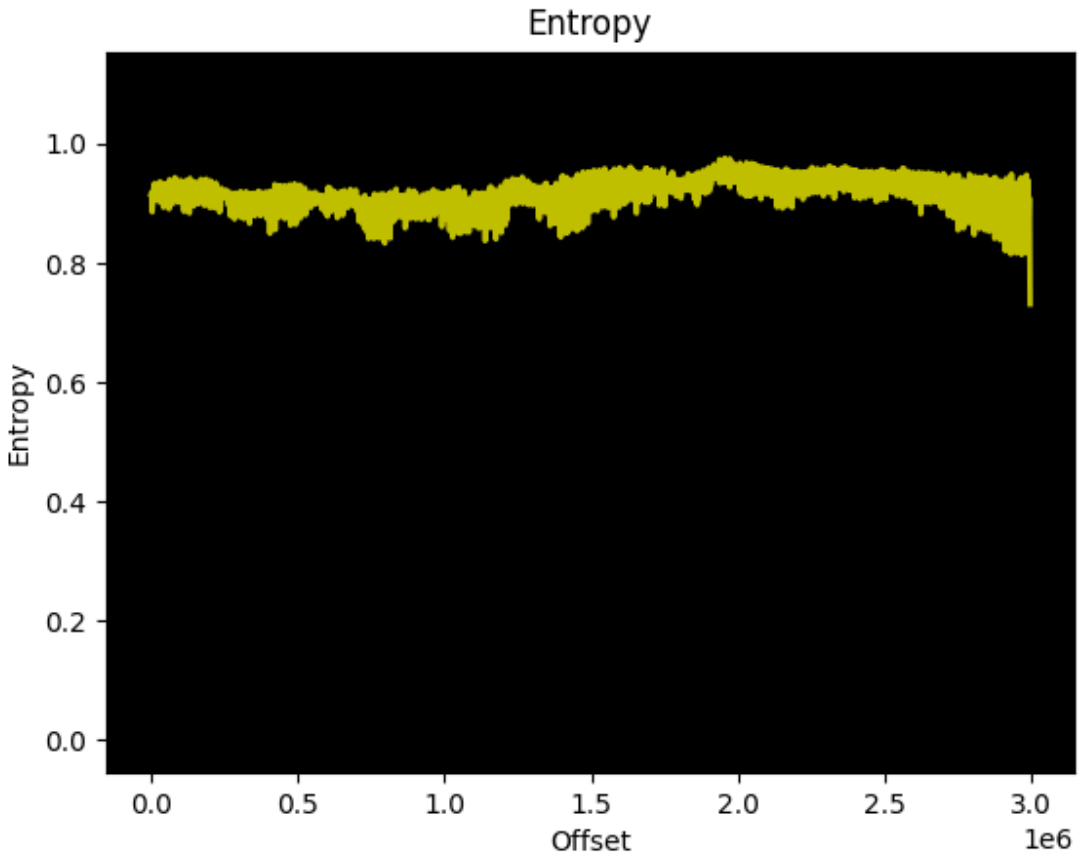
Cambiando el archivo a .zip y extrayendo su contenido el .txt obtenido tiene el siguiente contenido:

```
cada mina es un 1.
cada fila forma una letra.
Los ascii de las letras empiezan todos en 01.
Asi encontraras el algoritmo y el modo
La password esta en otro archivo
Con algoritmo, modo y password hay un .wmv encriptado y oculto.
```

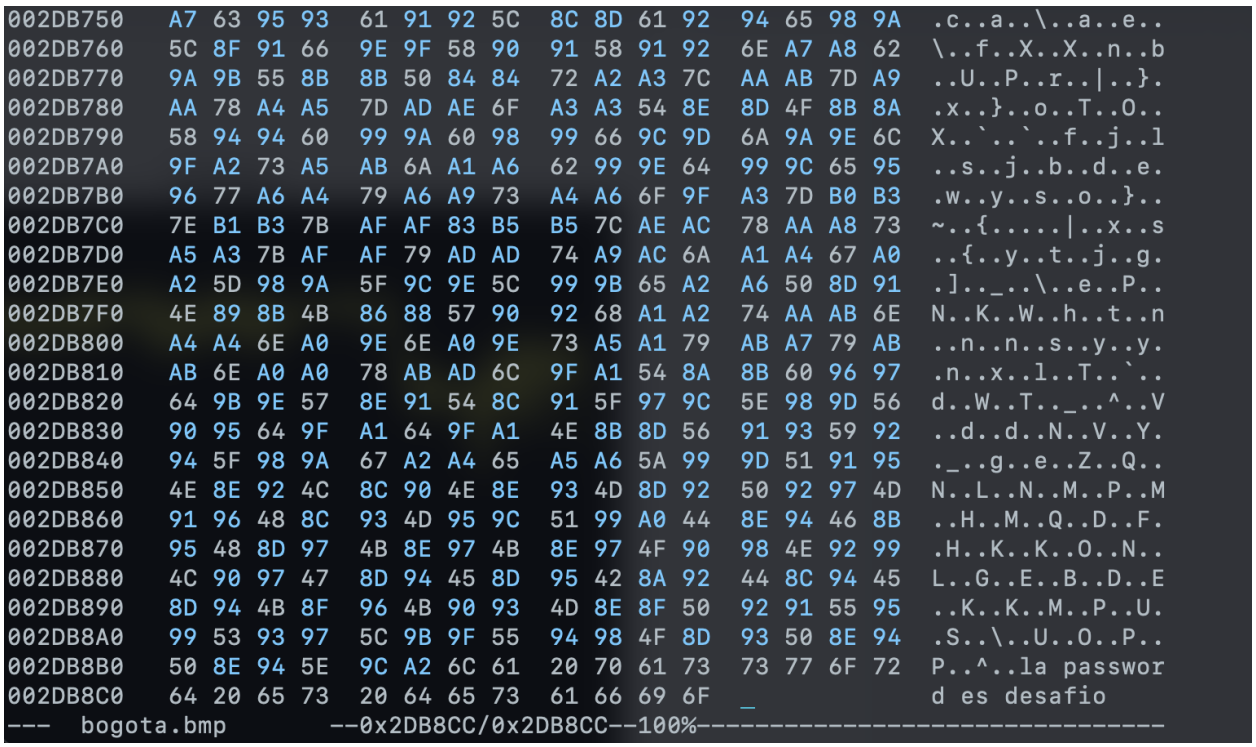
Con esto interpretamos nuevamente la imagen:

Por lo tanto, sabemos que uno de los archivos se encuentra encriptado con DES y en modo CBC.

Bogota

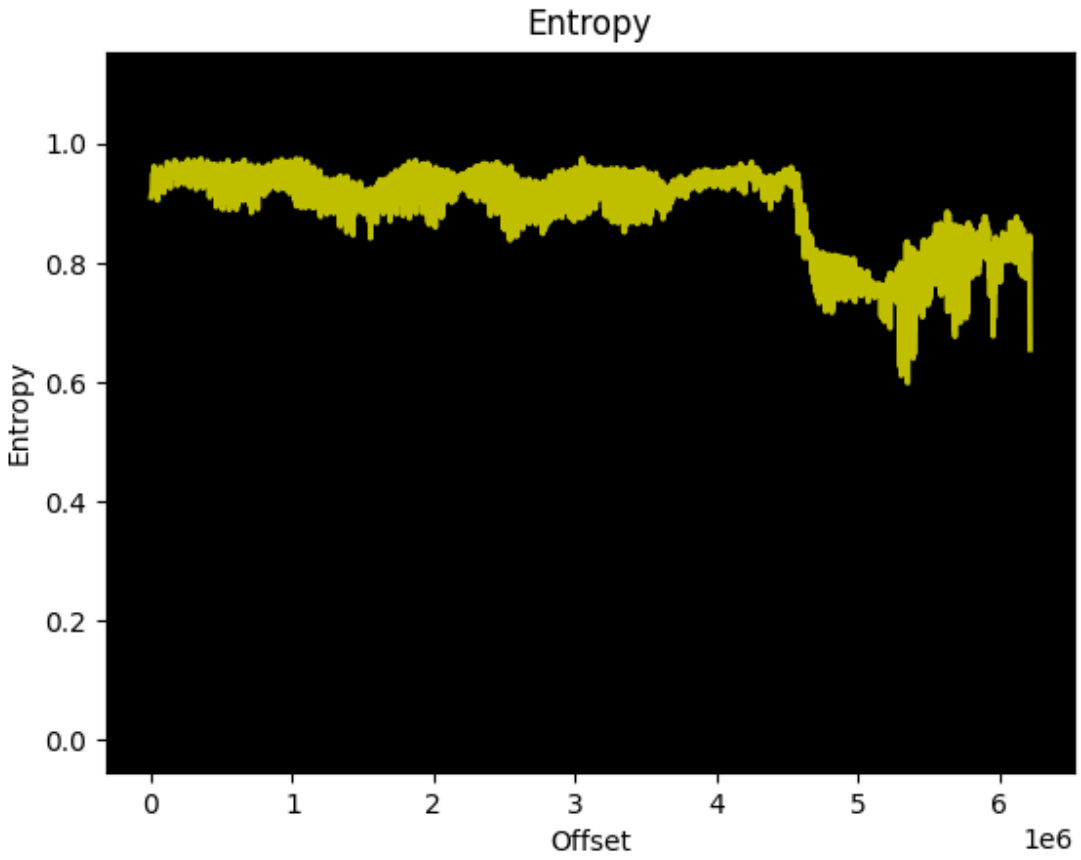


Viendo la irregularidad presente en el final del archivo se decide analizar con el editor hexadecimal la posición, donde se encuentra lo siguiente:



Por lo tanto, de la metadata y las pistas anteriores se sabe que la clave de la encriptación asociada al algoritmo de encriptación DES con modo CBC es “desafio”.

Budapest

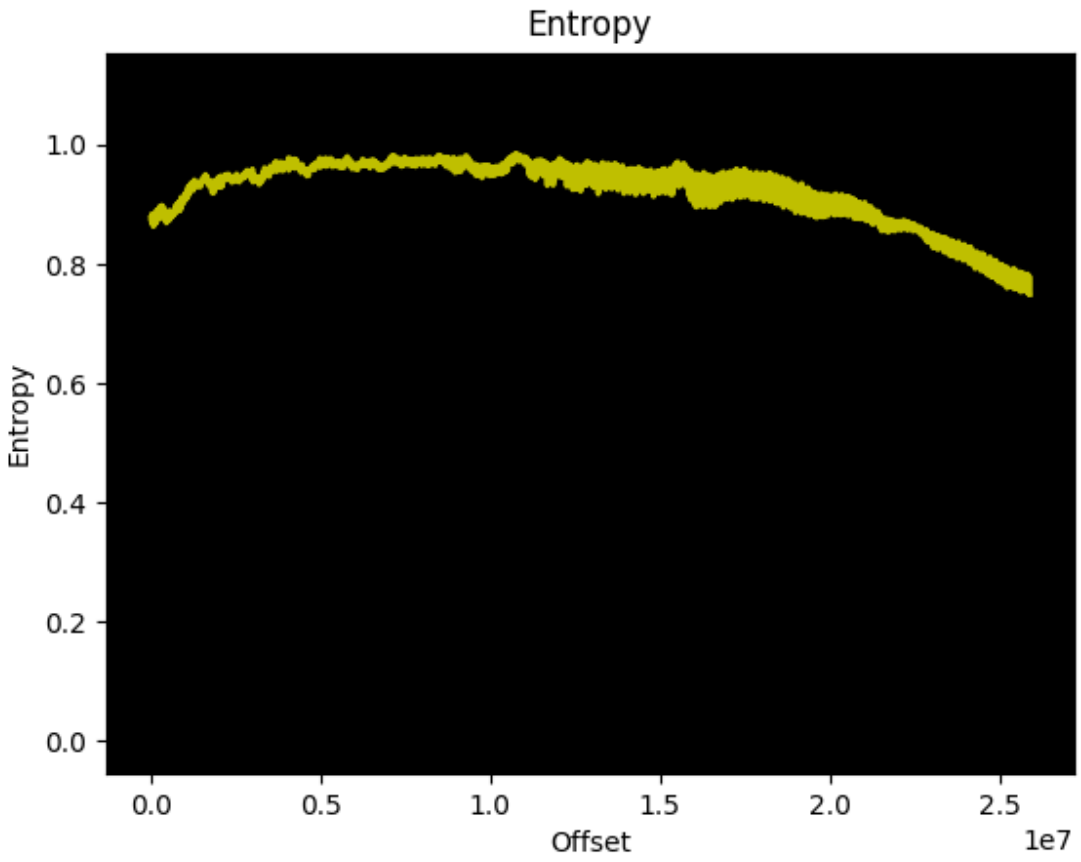


Se puede deducir que algo se encuentra escondido en el archivo por las irregularidades presentes.

En primera instancia, al usar LSB1 se obtiene como resultado un archivo el cual no tiene extensión esto se debe a que es el método equivocado y se tiene que usar LSBI con el cual se obtiene un PDF con la pista “al .png cambiarle la extensión por .zip y descomprimir”.

Madridoso

Dicho archivo es el más pesado de los previamente analizados (25.9 Mb). Posiblemente sea el que contenga el video mencionado en la pregunta 5. A continuación se analiza la entropía del mismo:



Se puede observar que posiblemente el archivo se encuentra encriptado, por lo tanto se prueba con los algoritmos LSB1, LSB4 y LSBI usando el algoritmo de encriptación DES con modo CBC y clave “desafio” previamente encontrados.

Se sospecha que con LSB1 se obtiene el resultado pero posiblemente haya un error con la clave debido a la siguiente salida obtenida:

Luego de una actualización del archivo, se encontró como salida un video de extensión .wmv que se encuentran detallados en la pregunta 5 sus contenidos.

Pregunta 4

Algunos mensajes ocultos tenían, a su vez, otros mensajes ocultos. Indica cuál era ese mensaje y cómo se había ocultado.

La pregunta describe el caso de Back que se encuentra detallado en su sección del archivo. De manera resumida, el resultado que se obtuvo se interpretó por su firma como png pero el misma contenía otra firma de tipo zip. Al cambiar el tipo de archivo y descomprimir el mismo se pudo hallar un archivo txt contenido.

Pregunta 5

Uno de los archivos ocultos era una porción de un video, donde se ve ejemplificado una manera de ocultar información ¿qué se ocultaba según el video y sobre qué portador?

En el video se puede ver que se encuentran observando una imagen que contiene información oculta ya que es más grande de lo que debería.

Pregunta 6

¿De qué se trató el método de estenografiado que no era LSB1 ni LSB4 ni LSBI? ¿Es un método eficaz? ¿Por qué?

El método que no era LSB1, ni LSB4 o LSBI se trató sobre esconder metadata dentro del archivo.

Este no es tan eficaz como el resto de los métodos ya que es fácil de detectar realizando un análisis de la entropía y viendo en detalle esa zona con un editor

hexadecimal. Incluso si no se fuera a realizar un análisis de la entropía, se pueden ver diferencias de tamaño entre lo que muestran los headers y el tamaño actual del archivo. Si fuera de mayor tamaño el secreto se volvería más evidente.

Pregunta 7

¿Por qué la propuesta del documento de Akhtar, Khan y Johri es realmente una mejora respecto de LSB común?

El algoritmo LSBI propuesto por Akhtar, Khan y Johri propone mejorar la calidad de la imagen resultante luego de su uso, a su vez haciendo el secreto más seguro a través de la inversión de bits ya que puede “evadir” ciertos métodos de análisis.

Pregunta 8

¿De qué otra manera o en qué otro lugar podría guardarse el registro de los patrones invertidos?

Una posibilidad de guardar los patrones invertidos es al principio de la información del archivo, empaquetandolo en los primeros 4 bytes. En caso de no necesitar invertir nada se dejan en 0.

Pregunta 9

Leer el Segundo esquema y analizar (sin implementar) cuáles serían las ventajas que pueden verse y sus desventajas o inconvenientes podría tener su implementación.

Ventajas	Desventajas
Permite invertir menos píxeles, por lo tanto permitiendo que llame menos la atención en ciertos análisis de stenagroafia.	Supone que el receptor tiene la imagen “cover” original, ya que la necesita para poder analizar si debe invertir o no el último bit.

Pregunta 10

¿Qué dificultades encontraron en la implementación del algoritmo del paper?

La principal dificultad encontrada en el diseño del algoritmo ocurrió al momento de decidir cuándo es conveniente leer o escribir en cuanto a las modificaciones que se hicieron. Esto puede ser desarrollado de una manera más eficiente ya que se repite código.

Además, hubo dificultades cuando se necesitan guardar los 4 bits que son necesarios para comprender qué invertir y qué no.

Por último, en el *paper* no especifica cómo conviene analizar si hubo un cambio o no y se puede perder eficiencia revisando constantemente los cambios.

Pregunta 11

¿Qué mejoras o futuras extensiones harías al programa stegobmp?

Como primera extensión se plantearía soporte a más variedades de archivos que no se limiten únicamente a archivos de tipo bmp.

A su vez, otra posible extensión para el programa stegobmp es la posibilidad de calcular la entropía como parámetro similar a -embed o -extract y se obtiene como resultado una representación gráfica de la entropía del archivo ingresado por parámetros similar a cómo funciona en la librería binwalk.

Otra posible extensión es la generalización del algoritmo LSB para soportar más variaciones que no se limiten a LSB1 y LSB4.

Bibliografía

- El documento "An Improved Inverted LSB Image Steganography" de autores Nadeem Akhtar, Shahbaaz Khan y Pragati Johri
- Sobre archivos bmp.
- Cummings, Jonathan y otros: Steganography and Digital Watermarking.
- Gómez Cárdenas, Roberto: Esteganografía.
- Johnson, Neil F. y Jajodia, Sushil: Exploring Steganography. Seeing the Unseen. Disponible.
- Binwalk.
- GitHub - livz/cloacked-pixel: LSB steganography and detection.
- HEXEDIT.
- 504B0304 File Signatures.