

Grupo 2:

stenography-crypto

Informe Trabajo Práctico Especial

72.44 - Criptografía y Seguridad - Primer Cuatrimestre 2022.

Catalán, Roberto Jose - 59174

Dell'Isola, Lucas - 58025

Pecile, Gian Luca - 59235

Índice

Cuestiones a Analizar	2
Pregunta 1	2
Pregunta 2	2
LSB1	3
LSB4	3
LSBI	6
Pregunta 3	7
Back	8
Bogota	11
Budapest	13
Madridoso	14
Pregunta 4	15
Pregunta 5	15
Pregunta 6	15
Pregunta 7	16
Pregunta 8	16
Pregunta 9	16
Pregunta 10	17
Pregunta 11	17
Bibliografía	17

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*.

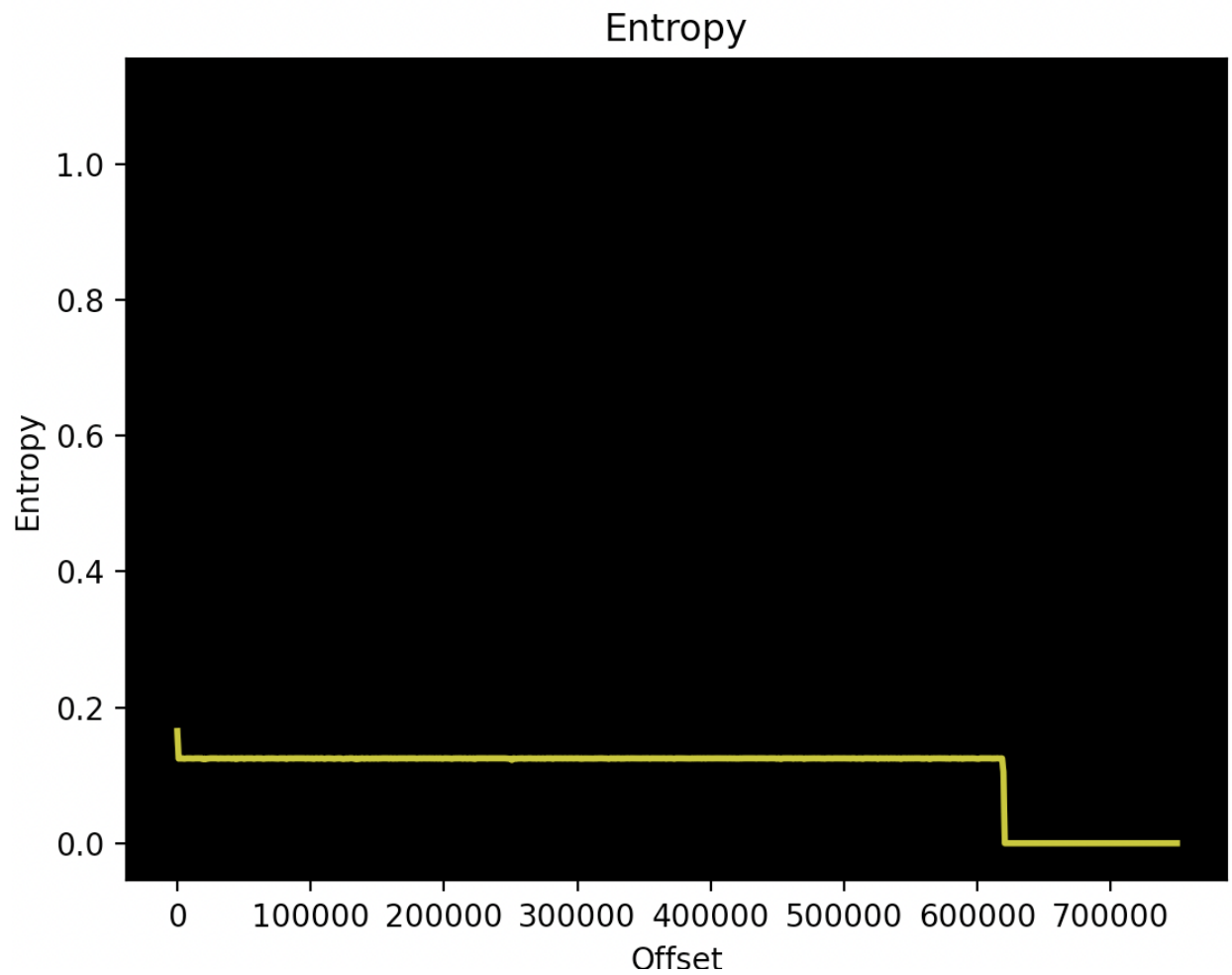
El uso de ejemplos de ejecución y su descripción es de gran utilidad para el entendimiento del programa a realizar.

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 en el archivo blanco.bmp (el cual es una imagen blanca para observar el ruido).

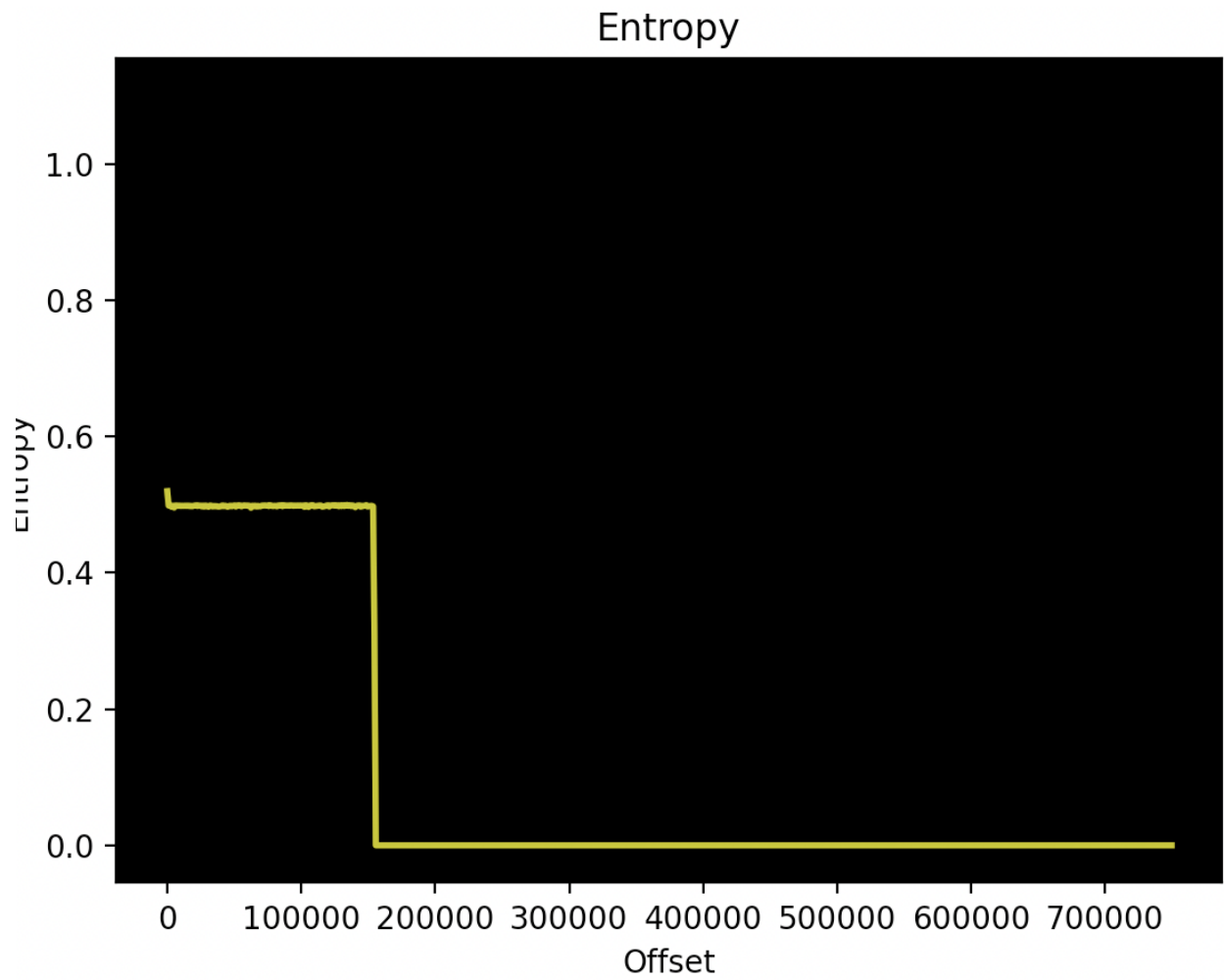
LSB1



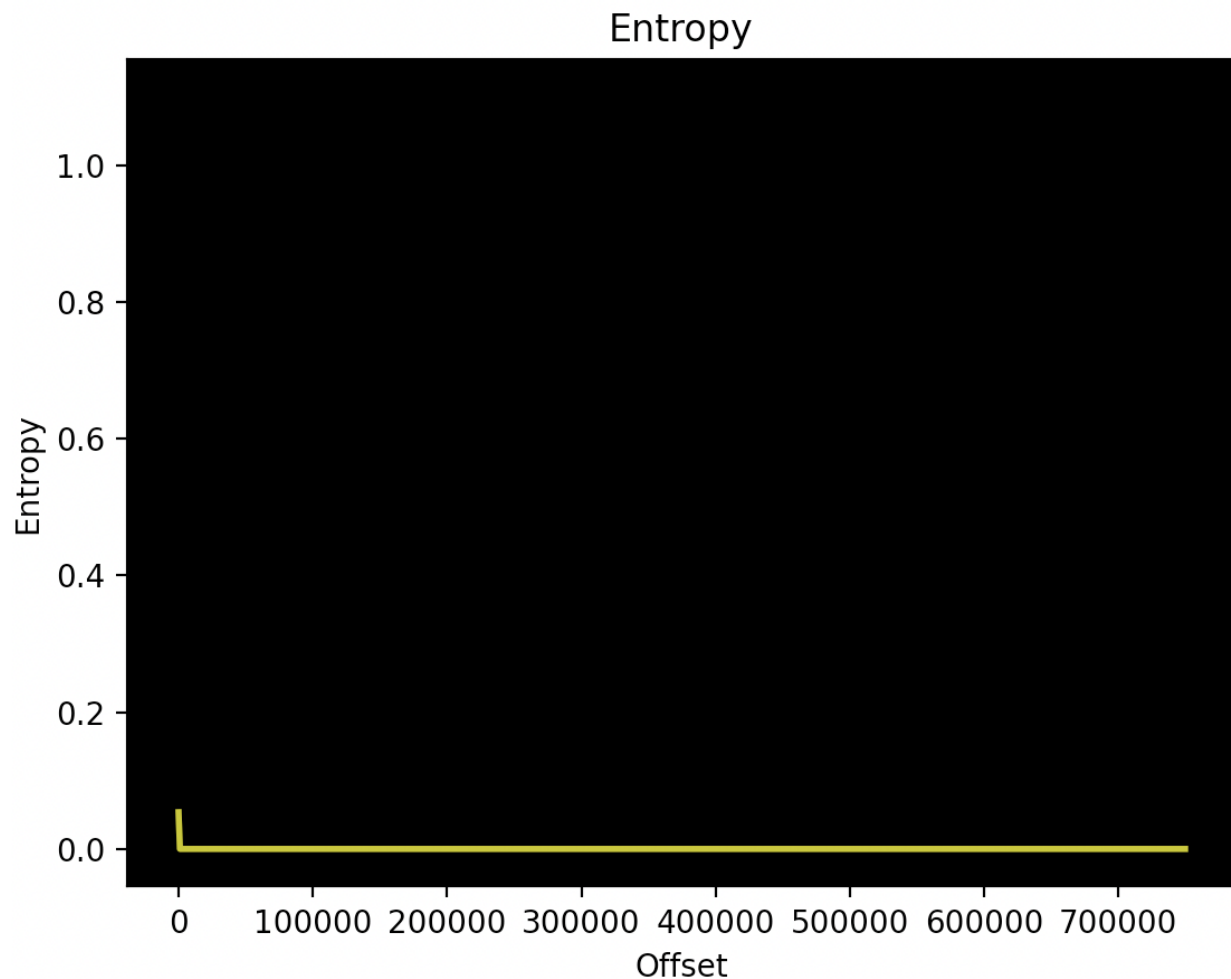
LSB4

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





LSBI



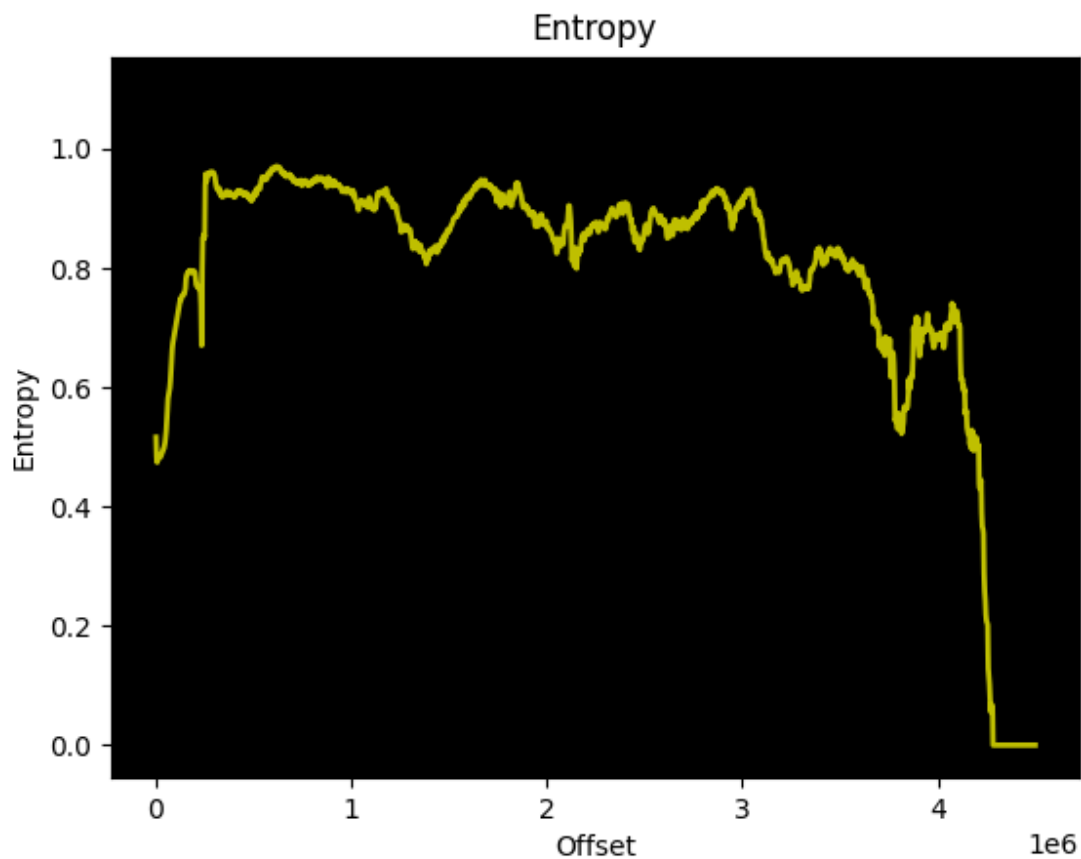
Algoritmo	Ventajas	Desventajas
LSB1	Utiliza el bit menos significativo para ocultar el mensaje, haciendo que las modificaciones a los píxeles sean mínimas o incluso nulas.	El tamaño de los mensajes que puede esconder es muy chico.
LSB4	Permite esconder mensajes de mayor tamaño que otros métodos ya que usa los 4 bits menos significativos.	Hay mayor presencia de ruido.
LSB Improved	Permite mayor complejidad	Se pierde eficiencia a la

	al momento de descifrar el mensaje oculto.	hora de manejar el secreto. Requiere más espacio 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 218...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 00 00 00 00 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.

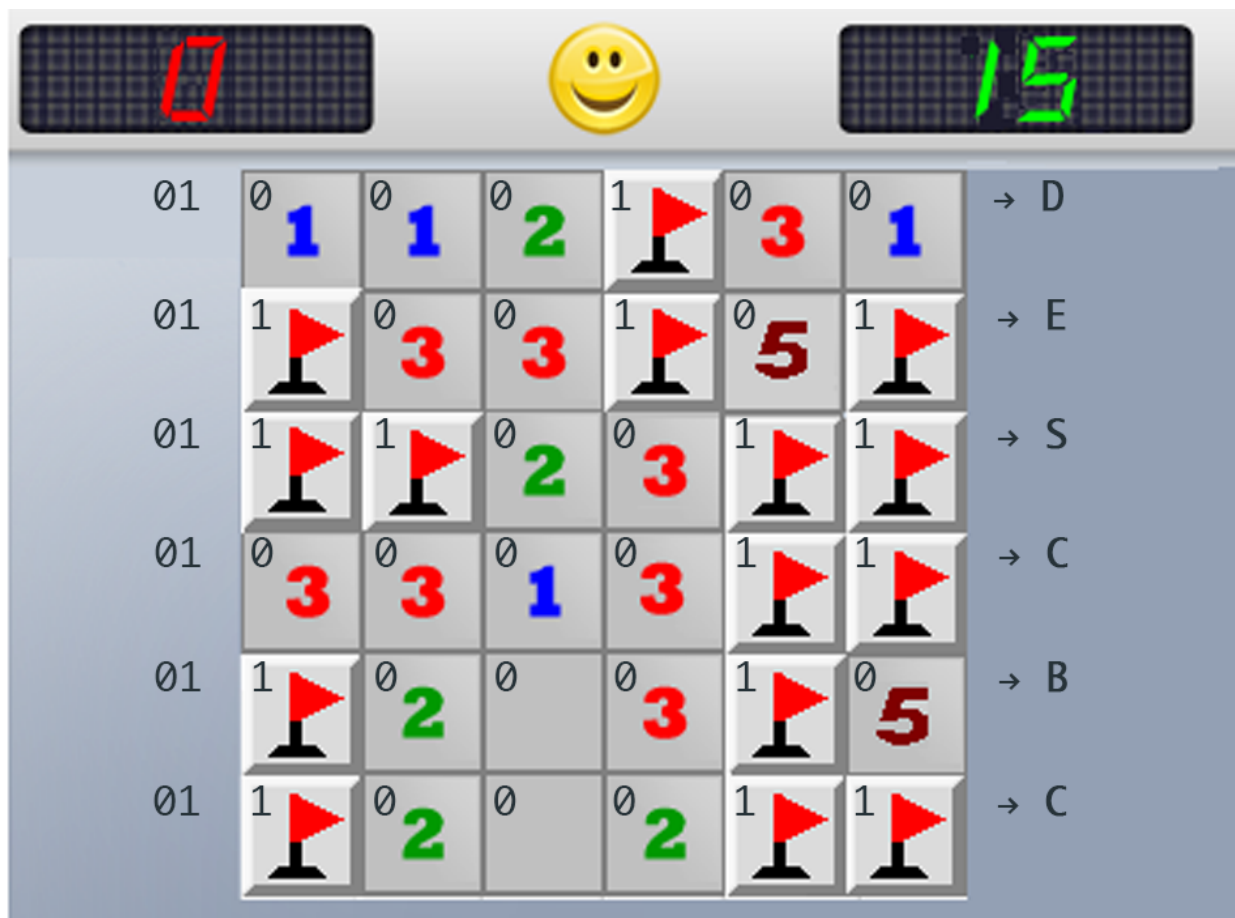
Los ascii de las letras empiezan todos en 01.

Asi encontraras el algoritmo y el modo

La password esta en otro archivo

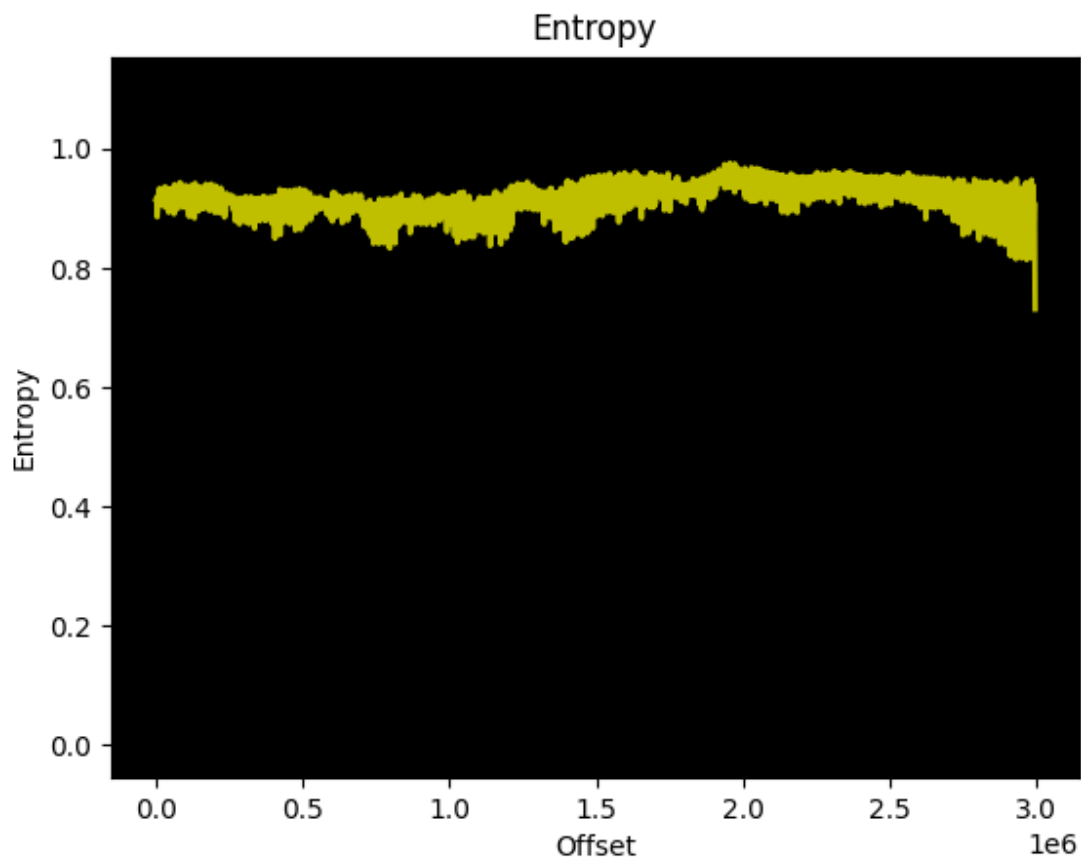
Con algoritmo, modo y password hay un .wmv encriptado y oculto.

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:

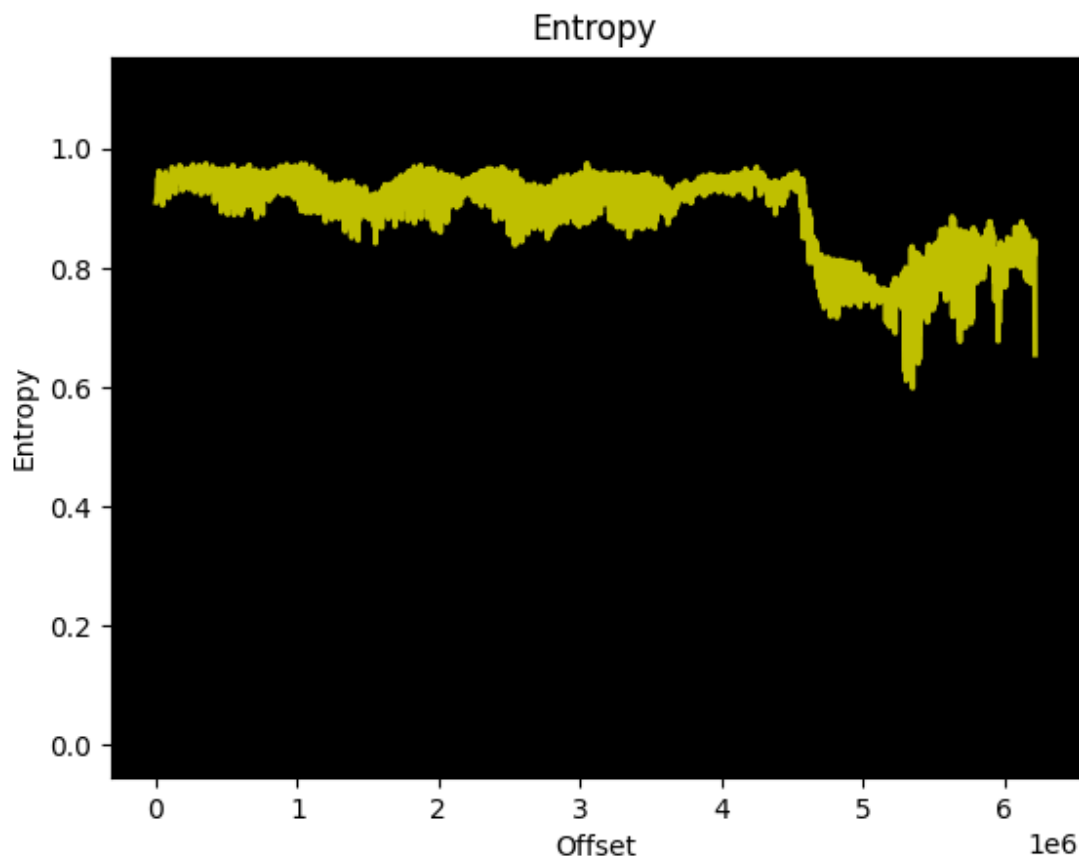
```

002DB750  A7 63 95 93 61 91 92 5C 8C 8D 61 92 94 65 98 9A .c..a..\..a.e..
002DB760  5C 8F 91 66 9E 9F 58 90 91 58 91 92 6E A7 A8 62 \..f..X..X..n..b
002DB770  9A 9B 55 8B 8B 50 84 84 72 A2 A3 7C AA AB 7D A9 ..U..P..r..|..}.
002DB780  AA 78 A4 A5 7D AD AE 6F A3 A3 54 8E 8D 4F 8B 8A .x..}..o..T..O..
002DB790  58 94 94 60 99 9A 60 98 99 66 9C 9D 6A 9A 9E 6C X..`..`..f..j..l
002DB7A0  9F A2 73 A5 AB 6A A1 A6 62 99 9E 64 99 9C 65 95 ..s..j..b..d..e.
002DB7B0  96 77 A6 A4 79 A6 A9 73 A4 A6 6F 9F A3 7D B0 B3 .w..y..s..o..}..
002DB7C0  7E B1 B3 7B AF AF 83 B5 B5 7C AE AC 78 AA A8 73 ~..{.....|..x..s
002DB7D0  A5 A3 7B AF AF 79 AD AD 74 A9 AC 6A A1 A4 67 A0 ..{..y..t..j..g.
002DB7E0  A2 5D 98 9A 5F 9C 9E 5C 99 9B 65 A2 A6 50 8D 91 .].._..\..e..P..
002DB7F0  4E 89 8B 4B 86 88 57 90 92 68 A1 A2 74 AA AB 6E N..K..W..h..t..n
002DB800  A4 A4 6E A0 9E 6E A0 9E 73 A5 A1 79 AB A7 79 AB ..n..n..s..y..y.
002DB810  AB 6E A0 A0 78 AB AD 6C 9F A1 54 8A 8B 60 96 97 .n..x..l..T..`..
002DB820  64 9B 9E 57 8E 91 54 8C 91 5F 97 9C 5E 98 9D 56 d..W..T.._..^..V
002DB830  90 95 64 9F A1 64 9F A1 4E 8B 8D 56 91 93 59 92 ..d..d..N..V..Y.
002DB840  94 5F 98 9A 67 A2 A4 65 A5 A6 5A 99 9D 51 91 95 ._.g..e..Z..Q..
002DB850  4E 8E 92 4C 8C 90 4E 8E 93 4D 8D 92 50 92 97 4D N..L..N..M..P..M
002DB860  91 96 48 8C 93 4D 95 9C 51 99 A0 44 8E 94 46 8B ..H..M..Q..D..F.
002DB870  95 48 8D 97 4B 8E 97 4B 8E 97 4F 90 98 4E 92 99 .H..K..K..O..N..
002DB880  4C 90 97 47 8D 94 45 8D 95 42 8A 92 44 8C 94 45 L..G..E..B..D..E
002DB890  8D 94 4B 8F 96 4B 90 93 4D 8E 8F 50 92 91 55 95 ..K..K..M..P..U.
002DB8A0  99 53 93 97 5C 9B 9F 55 94 98 4F 8D 93 50 8E 94 .S..\..U..O..P..
002DB8B0  50 8E 94 5E 9C A2 6C 61 20 70 61 73 73 77 6F 72 P..^..la passwor
002DB8C0  64 20 65 73 20 64 65 73 61 66 69 6F _ d es desafio
--- bogota.bmp --0x2DB8CC/0x2DB8CC--100%-----

```

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 “desafío”.

Budapest

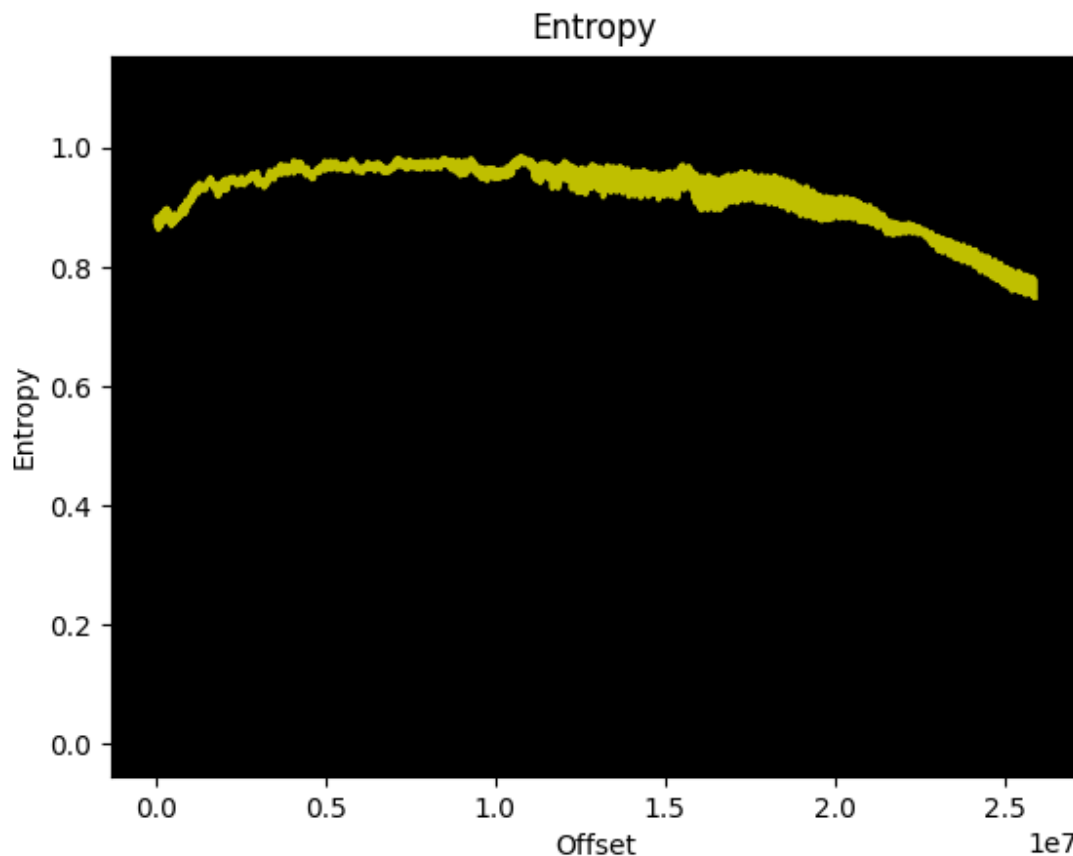


Se puede deducir que algo se encuentra escondido en el archivo por las irregularidades presentes.

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:



```
Run: extract-madridoso-LSB1-...
/Users/gipefile/Library/Java/JavaVirtualMachines/openjdk-18.0.1.1/Contents/Home/bin/java ...
Exception in thread "main" java.security.InvalidKeyException: Given final block not properly padded. Such issues can arise if a bad key is used during decryption.
    at java.base/com.sun.crypto.provider.CipherCore.unpad(CipherCore.java:858)
    at java.base/com.sun.crypto.provider.CipherCore.fillOutputBuffer(CipherCore.java:938)
    at java.base/com.sun.crypto.provider.CipherCore.doFinal(CipherCore.java:736)
    at java.base/com.sun.crypto.provider.DESCipher.engineDoFinal(DESCipher.java:315)
    at java.base/java.security.Cipher.doFinal(Cipher.java:2286)
    at Encryption.CipherAlgorithm.transform(CipherAlgorithm.java:58)
```

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 mismo 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 stenografía.	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.](#)