Abertay University

How and Why Digital Steganography is Used Today

Mairi McQueer – 1700231@uad.ac.uk

Introduction to Security – CMP110

BSc Ethical Hacking Year 1

2017/18

Abstract

This report is on the basics of Digital Watermarking (DW) and Digital Steganography (DS). It also looks into the reasons why people use DS today in order to bypass government restrictions on what data can be shared, for example pro-Shia Islam media in Saudi Arabia. It also looks into how watermarks are used to protect intellectual property online from people who wish to copy or distribute it without prior permission from the content creator.

DW, and by definition DS, is demonstrated using the least significant bit substitution method of watermarking. This was done using the software S-Tools, where a text file was encoded into an image file then a separate text file was retrieved from another image, demonstrating how DW can be achieved relatively simply. The program StegSpy was then used to show how DS is not invisible and how easy it is to discover if an image has had information encoded into it.

Contents

1 Introduction

1.a Background

Watermarks originated in Medieval Europe as a way of identifying the manufacturer or the specific type of paper. These early watermarks were created from linen or other materials being cut up, mixed with water and placed into a mould with a wire mesh on top. As the water drained through the mesh it created a lined pattern on the paper. A thicker wire image could be placed on top of the mould to create the watermarked image, which could then be seen when the paper was held up to some light.

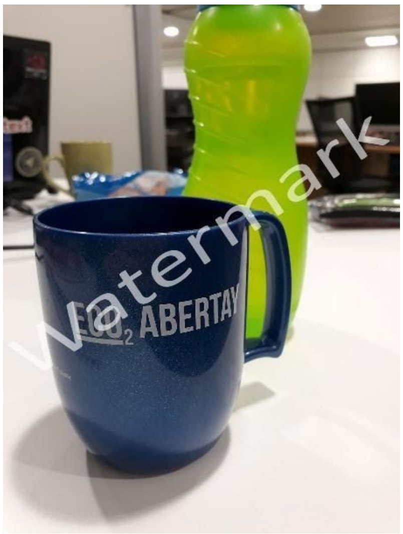


Figure 1: image with paper watermark

Digital Watermarking (DW) is often confused with paper watermarks on digital images, (See Figure 1) where a logo or text is placed over a picture in order to try and stop another person replicating of copying the image without prior consent from the content creator. DW is far more complex, it involves encoding data about the image in such a way that has minimal effect on the image quality and cannot be easily perceptible to anyone looking at it. Data usually contained in a watermark includes; information on the copyright owner, image creator and authorised consumer, as well as the requirements for handling the property rights.

The reason DW is sometimes preferred over simply encrypting the image file is that once the file is decrypted it can then be copied and distributed relatively easily but with DW the circulation of the data can be tracked and the original owner easily identified. (Barni et al, 1998)

According to Barni et al's 1997 report titled 'A DCT-Domain System for Robust Image Watermarking', a good watermark should be:

- Unobtrusive, little to no quality is lost so that it appears invisible to the observer and meaning those with malicious intent cannot easily find it to remove it.
- Readily extractable, meaning that the owner of the image and any quality control authority can extract the watermark to view relatively easily.
- Robust, making it difficult for malicious users to remove and any attempts to do so significantly reduces the image quality.
- Unambiguous, in identifying the owner and the content creator.
- Innumerable, meaning that many DWs can be produced and still be told apart from each other.

Steganography's etymology is ancient Greek from 'steganos' and 'graphein', meaning covered and to write. One of the earliest recorded uses of steganography comes from ancient Greeks tattooing messages on slaves heads, waiting for their hair to grow back, then shaving their hair again to reveal the message underneath. From there on the ancient Chinese military officers and diplomats hid messages on thin silk or paper sheets.

Modern day steganography can be sorted into three categories:

- Technical, could also be considered physical steganography. Uses techniques such as invisible ink or microdotting, where secret information in photographs are shrunk down to the size of a full stop and placed in an inconspicuous magazines or newspapers. Microdots were used predominantly in the Second World War and invisible ink is still very much used today, in the form of ink that can be seen under ultraviolet light, for detecting counterfeits.
- Linguistic, simply put this is written steganography. Linguistic steganography be broken down further into two subcategories:
    - Semagrams involve hiding information using techniques such as, adding extra spaces or changing the font of certain words in a text document such as a newspaper. Semagrams can also be hidden in images or clothing.
    - Open codes are less obvious and hide messages in plain sight. For example a block of unassuming text in a newspaper becomes a secret message when the first letter of each word is removed. E.g. "Humiliated engineer leads protest" hides the word "help".
- Digital Steganography (DS) is hiding information in images, videos or music files digitally, this tends to be done using least significant bit substitution but DS can also be, creating secret hard drive partitions and storing data there or redundant pattern encoding. Redundant pattern encoding is where the secret information is dispersed all over the image protecting it more from resizing and cropping.

DW in itself is a form of DS. However it differs from generic DS as all the data encoded into the image is entirely related to said image, whereas DS tends to encode data that is completely separate to the original image. This is done to try and make it inconspicuous as the information is meant to remain hidden to all apart from the authorised users.

In modern society DS is used for sending messages and data covertly. This can be employed by those wishing to bypass government censorship of media, for example sending messages about the Tiananmen square protests in China. It has also been suggested that DS was used to send messages between terrorists in Al Qaeda in preparation for the 9/11 attacks on the World Trade Centre in America. (Shih, 2008. p.140) Although both these examples are quite extreme in nature there are plenty of more mundane uses for DS, such as for storing data that one does not wish to be readily available to others or for simply communicating with another person more securely.

1.b Aims

The main aim of this report is to explain the of Digital Watermarking, its uses in modern society and its limitations through achieving the following goals:

- Encoding an encrypted message onto an image
- Revealing the message on said image with the password
- Revealing if there is any steganography on the image without knowing the password

2 Procedure

2.a Overview

Using the software S-Tools data can be encoded into an image using Least Significant Bit (LSB) substitution and can only be revealed if the user has the correct passphrase.
The author has decided to use S-Tools software to encode and encrypt a text file to an image, then place a password on it and finally reveal the encoded data. The watermarked image will also be put through the software StegSpy to test if the steganography can be found without the password.

2.b Creating the watermark

1. Start the program by running the S-Tools.exe file
2. Drag the image file into the blank box (Appendix A)
    1. If the image file is not in the correct format, place it into an application such as paint and save it as .bmp
3. Create your watermark by creating a .txt file on an application such as notepad and putting your message on there (Appendix B)
4. Drag the .txt file on top of the image file
5. Create a password and choose an encryption type (IDEA or MDC is recommended) (Appendix C)
6. Save the newly watermarked image as a .bmp file

2.c Revealing the watermark

1. Start the program by running the S-Tools.exe file
2. Drag the watermarked .bmp file into the blank box
3. Right-click on the image and select ‘reveal’
4. Enter the password
5. Right-click the file name and save the file onto your computer (Appendix D)
6. Open the file to view the hidden data (Appendix E)

2.d Uncovering steganography

1. Start the program StegSpy
2. Press the run button
3. Select the watermarked .bmp file
4. Wait for the program to finish checking the image (Appendix F)

3 Results

All of the aims stated above were met successfully, the DS worked as expected. The text file was both encrypted and encoded successfully making use of IDEA. The passphrase for this was "password123", onto the image file without any noticeable loss of quality. (Appendix G) A problem encountered during this process was that the original image was in JPEG format so would not work with the S-Tools program, this was resolved by converting the image file into a bitmap. Converting the file resulted in a slight increase of image quality. (Appendix H)

Although not in the original aims, the author decided to test the maximum amount of data the software would allow to be encoded. At this time a text file was created and the Bee Movie script was added, then placed into the image. The script was then added again and the process repeated until an error message appeared explaining that the file was too large to be hidden in the photo. This revealed that S-Tools is capable of inserting around 3MB, 57 Bee Movies (a singular Bee Movie equalling 52.84 KB), onto any given image file. (Appendix I)

The revealing of the text watermarked onto the image also worked as expected (Appendix E) as did the StegSpy software, which uncovered that the image had hidden data on it. Due to the time taken for the process both the user and windows were of the belief that the program had crashed but eventually StegSpy returned to a responsive state and gave the desired result. The results proved that DS was contained within this image. (Appendix F)

Comparing the original image to the first watermarked image, with the two lines of text, there is no noticeable difference between them. (Appendix G) However when comparing the original file to the second watermarked image the difference is slightly more noticeable especially at the edge of shadows, which appear blurrier in the watermarked photo. (Appendix J)

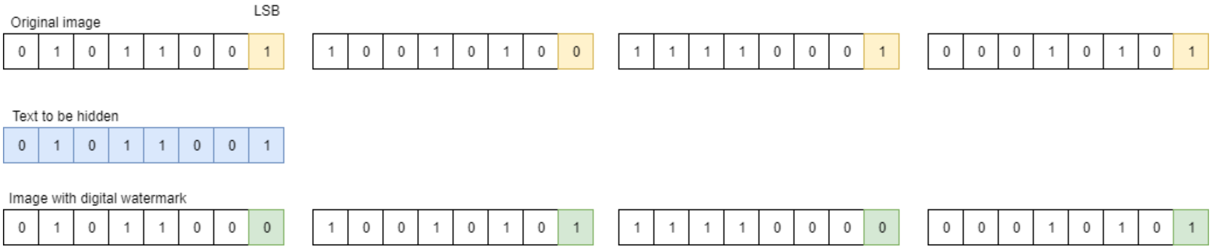
4 Discussion

4.a General Discussion

Although the images are referred to as watermarked the author is fully aware that due to the irrelevance of some of the text placed into these photos it is technically not a watermark, but instead simply DS. This was done intentionally as the text is simply to demonstrate how watermarks are applied.

These results demonstrate how S-Tools is a convenient and relatively easy way to hide data onto digital media. This means that the content creator can place their information onto an image or video in a way that is relatively imperceptible to the human eye and allows them to add a password onto the file. Requiring the password makes it much more difficult for a person without the rights to the image to uncover and potentially remove the watermark, but not impossible as the password may be revealed after a brute-force attack. The encryption options also allow for further security as someone trying to remove or read the information hidden on the bmp file would find it very difficult to encrypt, especially if IDEA or MDC is used.

As stated above the software used for this report makes use of LSB substitution. Which works by replacing the LSB of each pixel with a bit from the binary form of the text to be hidden, as seen in the figure below. By using the LSB the change is not enough to be noticeable by anyone comparing the original and watermarked image. It is possible to increase the amount of bits used, although this does increase the difference between the original and new image as more quality is lost. Most people cannot tell the difference up to three least significant bits but after four the quality really begins to decline to the point where it's noticeable.



0 1 0 1 1 0 0 1 | 1 0 0 1 0 1 0 0 | 1 1 1 1 0 0 0 1 | 0 0 0 1 0 1 0 1

0 1 0 1 1 0 0 1

0 1 0 1 1 0 0 0 | 1 0 0 1 0 1 0 1 | 1 1 1 1 0 0 0 0 | 0 0 0 1 0 1 0 1

Figure 2: LSB

As seen in appendix C, there are four options for encrypting the data to be hidden in the image file:

- IDEA, or International Data Encryption Algorithm is widely considered the best publicly known algorithm, as no successful attacks have been published so far. It uses an 128 bit key and the patent for this is held by Ascom tech, although it is free to use for non-commercial purposes.
- DES, stands for Data Encryption Standard and is easily the least secure out of the four options. Published in 1975 by IBM it is also a block cipher, however it's key is only 56 bits and it uses symmetric key encryption. Symmetric cryptography means that the same key is used for encryption and decryption. Was used by the government of the United States but they formally withdrew from using it mid 2005 and now favour 3DES.
- Triple DES, also known as 3DES, simply put performs DES three times. It is now considered obsolete as it has been replaced by the Advanced Encryption Standard. It encrypts the data three times and creates three keys. (See figure 3)
- MDC, Message Digest Cipher, not to be confused with Modern Detection Code encryption. Originally developed by Ronald Rivest, there are currently 6 versions, the most recent one was published in 2008. The first five versions are very susceptible to brute force attacks and

therefore are not considered to be very secure, hence why only MD6 is really used for encryption today.

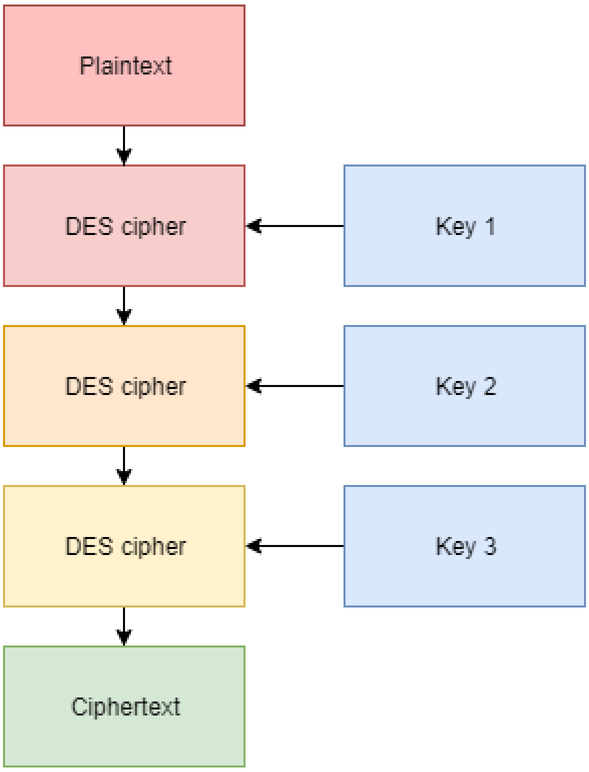


Figure 3: Triple DES

All of these options use a block cipher, meaning that it encrypts information as a whole block rather than one bit at a time. Block ciphers are believed to be stronger than stream ciphers however they can be slower due to the large amount of data being passed to the cipher.

IDEA was choses to encrypt the practice watermark as it is very secure so would make it more difficult, if not impossible, for someone without the password to view or remove the watermark.

StegSpy finding the DS reveals how although watermarks and hidden data may be invisible to the human eye, software can uncover it quickly. This does provide a flaw with DS as a technique to send data discreetly as if the steganography can be discovered quickly the actual data can be unearthed using a brute-force attack on the password.

This being said it is very unlikely that all images a person is sending or receiving will be checked for DS, still making a valuable tool for any person wishing to send information on a smaller scale without it being intercepted. An example would be if a person wishes to send media that is illegal in their country, such as any movie involving time travel in China, to another person then they are very unlikely to have their messages looked at any further than face value. Thus making DS a very useful tool to those persons. This example also shows how DW would also be relevant, since the film is likely to have been obtained illegally the DW will allow the movie to be tracked as to who the original owner is and who has the actual rights to own it.

Another issue with using DS tools such as S-Tools to send data discretely is the sheer quantity of media files required in order to send large amounts of data, as demonstrated above S-Tools has a limit of just over 3MB of text that can be placed onto an image. This means that it is only really practical for small scale communications, as sending large amounts of seemingly 'random' images could start to appear suspicious increasing the chance that they will be checked for hidden data.

4.c Conclusions

DS is seemingly obsolete for the communication of large amounts of sensitive data. However is still useful for sending messages to another party discreetly, as it is relatively easy to use and can hide a fairly good amount of text. For someone who isn't having their communications looked at by someone who is motivated enough to spend the time trying to get the passphrase he encryption makes it very secure.

4.d Future Work

Digital Steganography, specifically DW is a vast topic and this report barely covers the basics of it. Future work would definitely include looking more in-depth at the types of DW, such as; robust, semifragile and fragile watermarks. Specifically the differences between them, the benefits and the limitations of using each one. As well as doing further research into the algorithms involved in:

- Blind and non-blind watermarking like those designed by Cox et al or Swanson et al.
- Encryption techniques, for example IDEA or MD6.

Understanding the algorithms and the proper workings of these techniques would take a considerable amount of time that was not afforded to the author for this report so must be left for future research into the topic.

References

Alembic rare books(2015)*Watermarks & Foolscaps: Exploring the History of Paper Production.* Available at: https://alembicrarebooks.com/blogs/alembic-rare-books-blog/40160515-watermarks-foolscaps-exploring-the-history-of-paper-production (Accessed: 30 April 2018).

Barni, M. et al (1998) *'A DCT-Domain System for Robust Image Watermarking'.* Italy: Elsevier Inc. (Accessed: 28 April 2018).

Cox, I.et al (2008) *Digital Watermarking and Steganography.* Massachusetts: Elsevier Inc.

INFOSEC Institute(2011) CISSP- *Steganography, An Introduction Using S-Tools.* Available at: http://resources.infosecinstitute.com/cissp-steganography-an-introduction-using-s-tools/ (Accessed: 29 April 2018).

Jyothi, B. et al(2010) *'Implementation and Analysis of Email Messages Encryption and Image Steganography Schemes for Image Authentication and Verification'.* India: International Journal of Computer Applications. Available at: https://pdfs.semanticscholar.org/344c/815b8fbeaa716449e4057e3315aa03ad52d0.pdf (Accessed: 30 April 2018).

Shih, F. (2008) *Digital Watermarking and Steganography: Fundamentals and Techniques.* Florida: Taylor & Francis Group.

Siper, A. et al (2005) *'The Rise of Steganography'.* New York: Pace University. Available at: http://csis.pace.edu/~ctappert/srd2005/d1.pdf (Accessed: 29 April 2018).

https://searchsecurity.techtarget.com/definition/International-Data-Encryption-Algorithm

https://searchsecurity.techtarget.com/definition/Data-Encryption-Standard

https://www.techopedia.com/definition/4144/triple-des

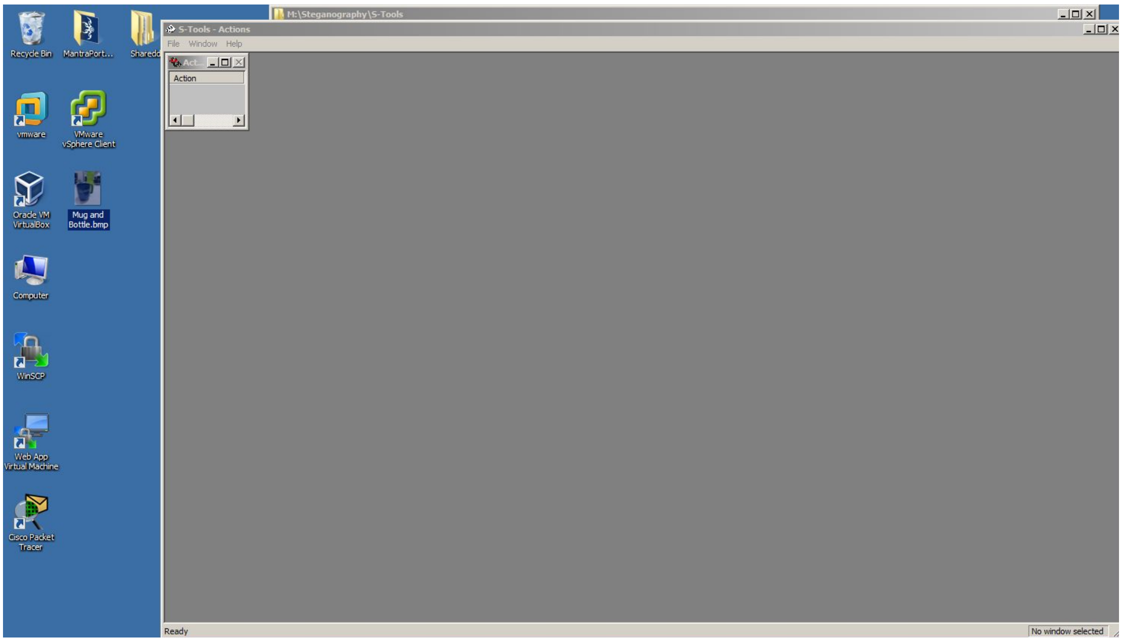
https://www.ibm.com/support/knowledgecenter/en/SSB23S_1.1.0.14/gtps7/s7symm.html

Bee Movie 2007 Hickner. S, Smith. S DreamWorks Animation

Appendices

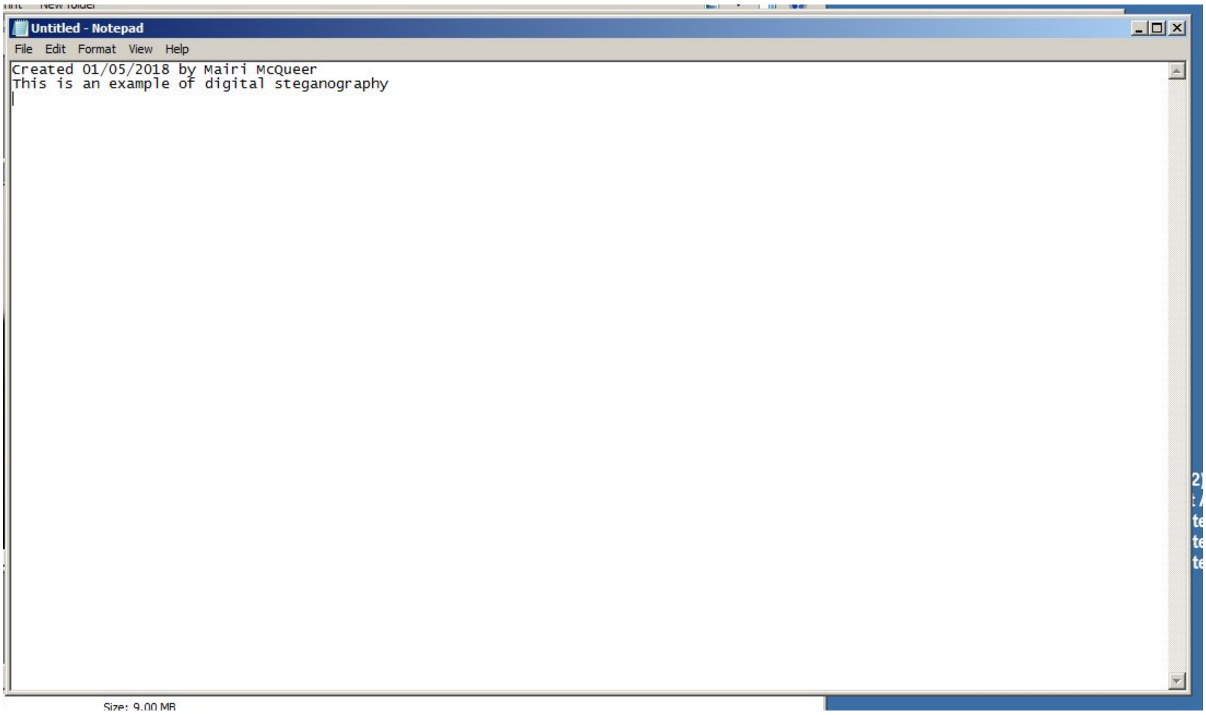
Appendix A

Step 2 of creating the watermark: Drag the image file into the blank box.



Appendix B

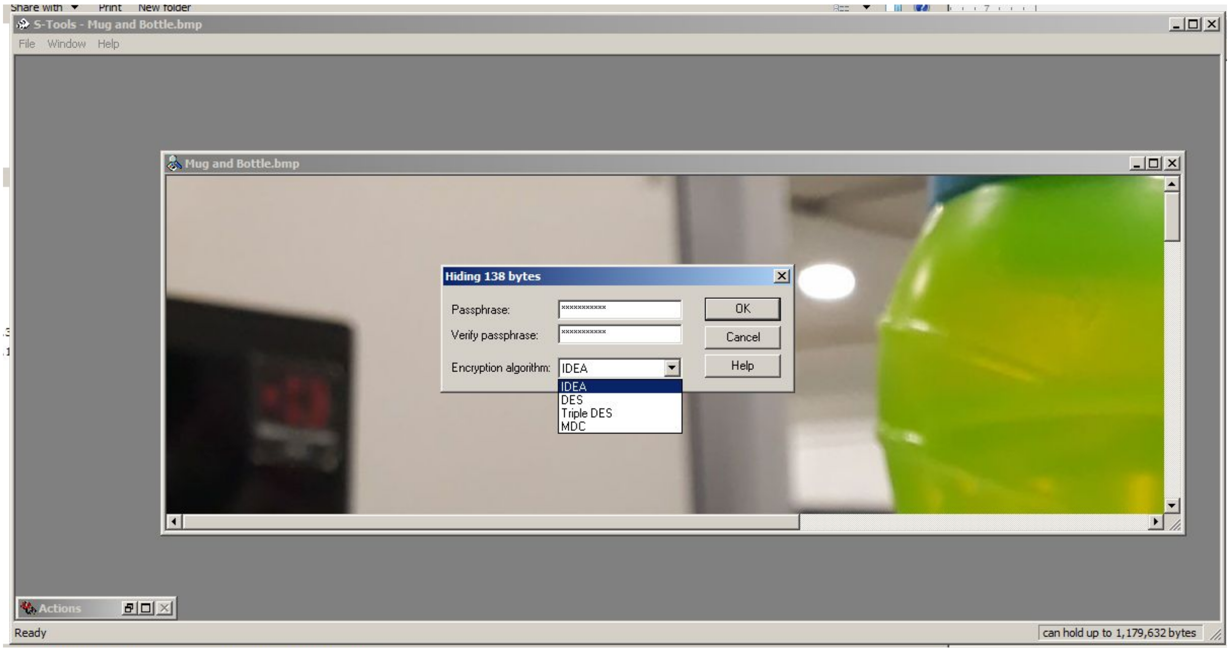
Step 3 of creating the watermark: Create a text file on notepad



```
Created 01/05/2018 by Mairi McQueer
This is an example of digital steganography
```

Appendix C

Step 5 of creating the watermark. This image also shows the four encryption options; IDEA, DES, 3DES and MDC.



Hiding 138 bytes

Passphrase: ***********

Verify passphrase: ***********

Encryption algorithm: IDEA

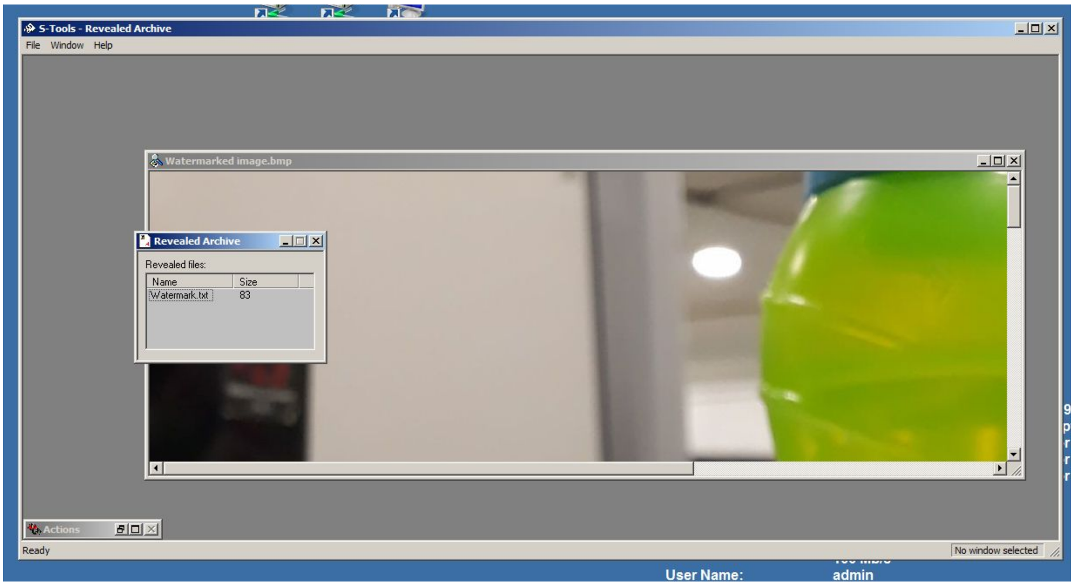
IDEA
DES
Triple DES
MDC

Ready

can hold up to 1,179,632 bytes

Appendix D

Step 5 of Revealing the watermark: Save the revealed file onto your computer.



Revealed files:

Name	Size
Watermark.txt	83

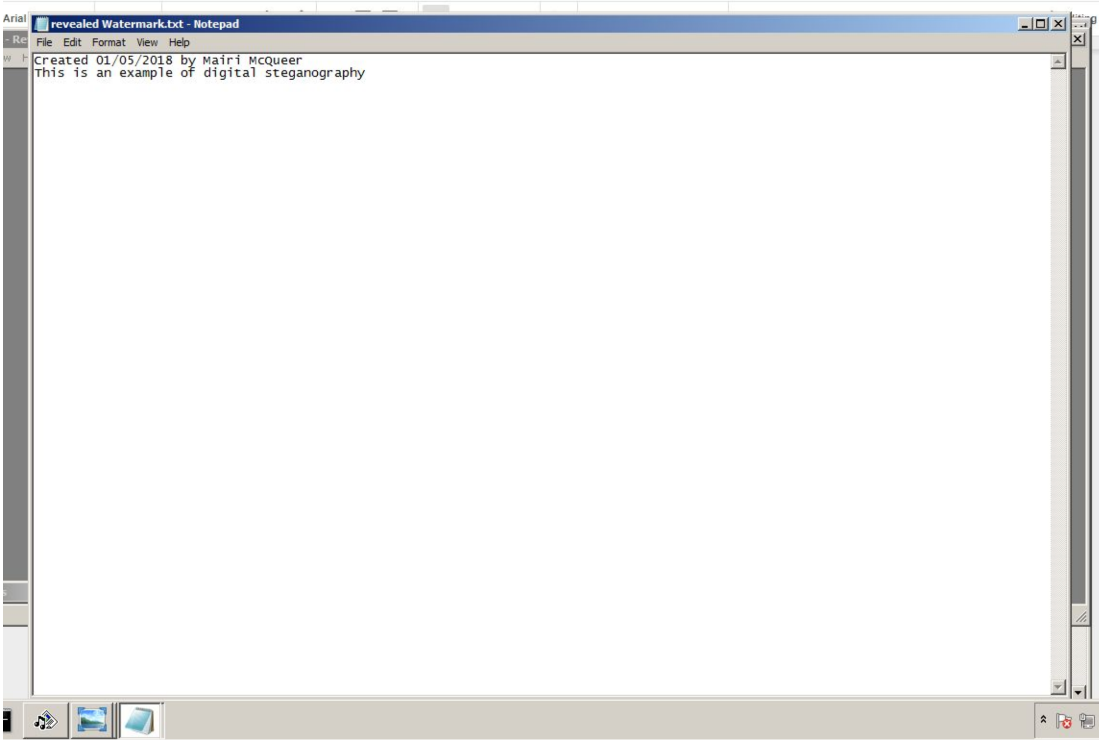
Ready

No window selected

User Name: admin

Appendix E

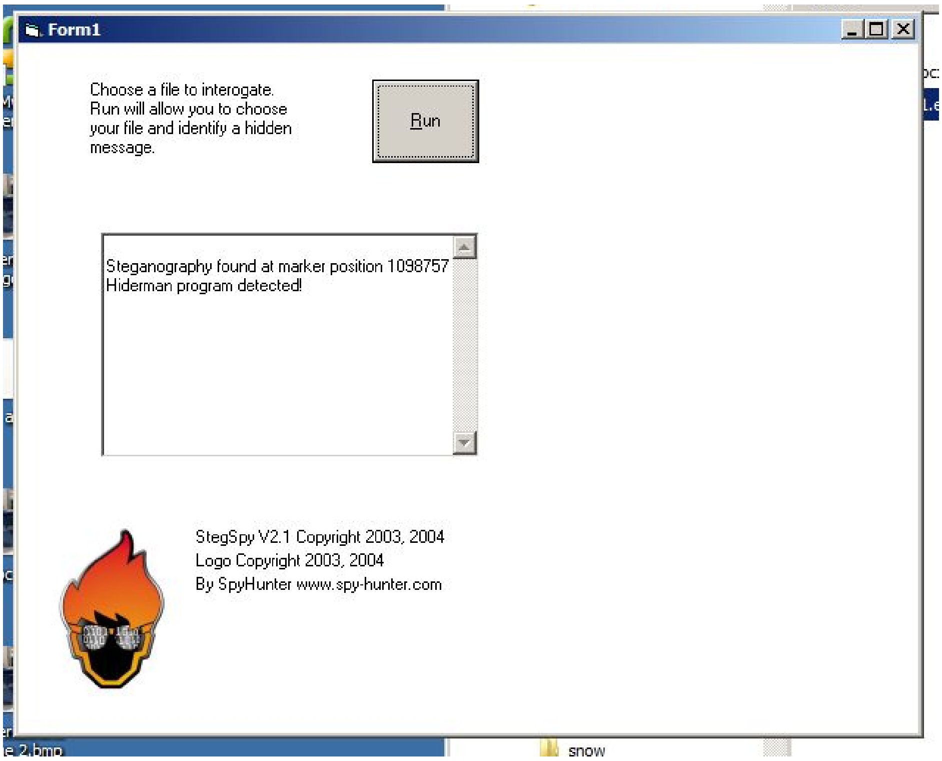
Step 6 of revealing the watermark: Open the uncovered file from the watermarked image.



```
Created 01/05/2018 by Mairi McQueer
This is an example of digital steganography
```

Appendix F

Step 4 of uncovering steganography, the image shows how StegSpy has successfully found the location of the hidden data on the image file.



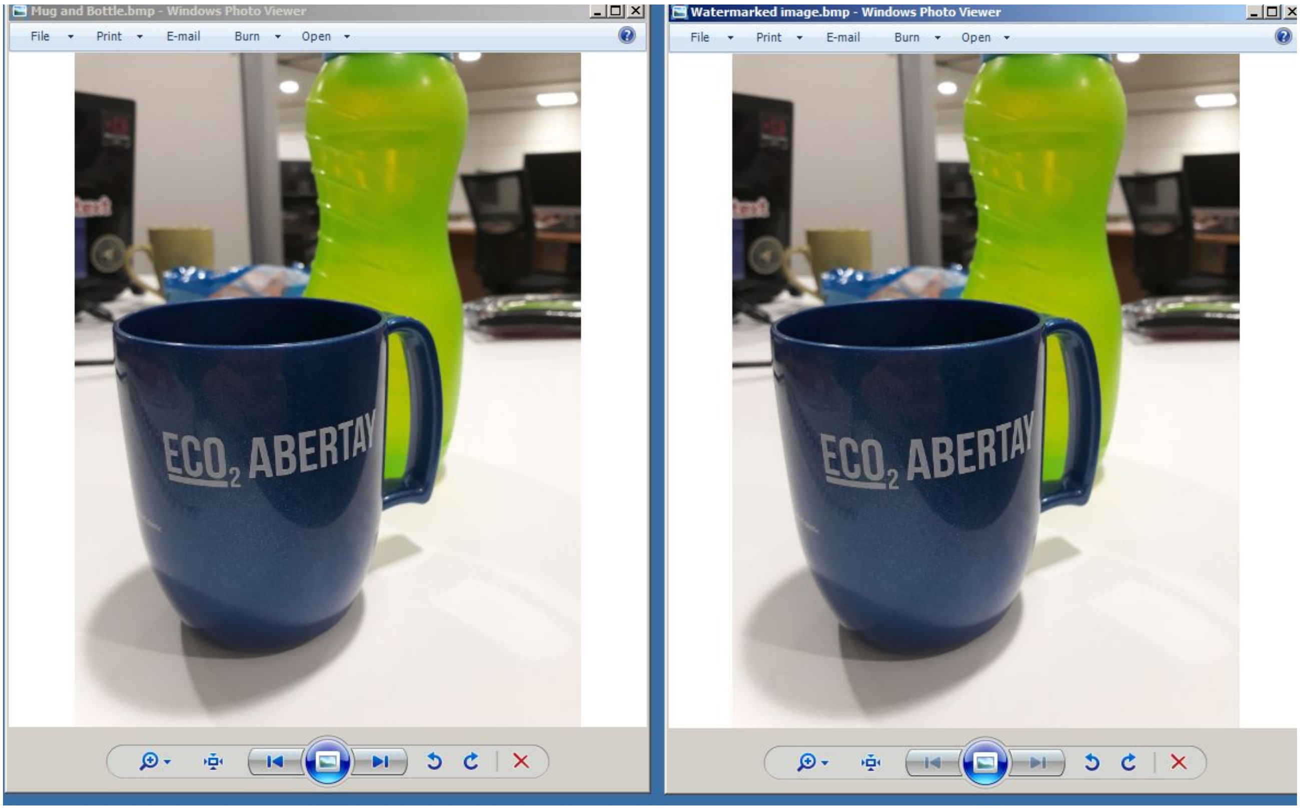
Choose a file to interogate. Run will allow you to choose your file and identify a hidden message.

```
Steganography found at marker position 1098757
Hiderman program detected!
```

StegSpy V2.1 Copyright 2003, 2004
Logo Copyright 2003, 2004
By SpyHunter www.spy-hunter.com

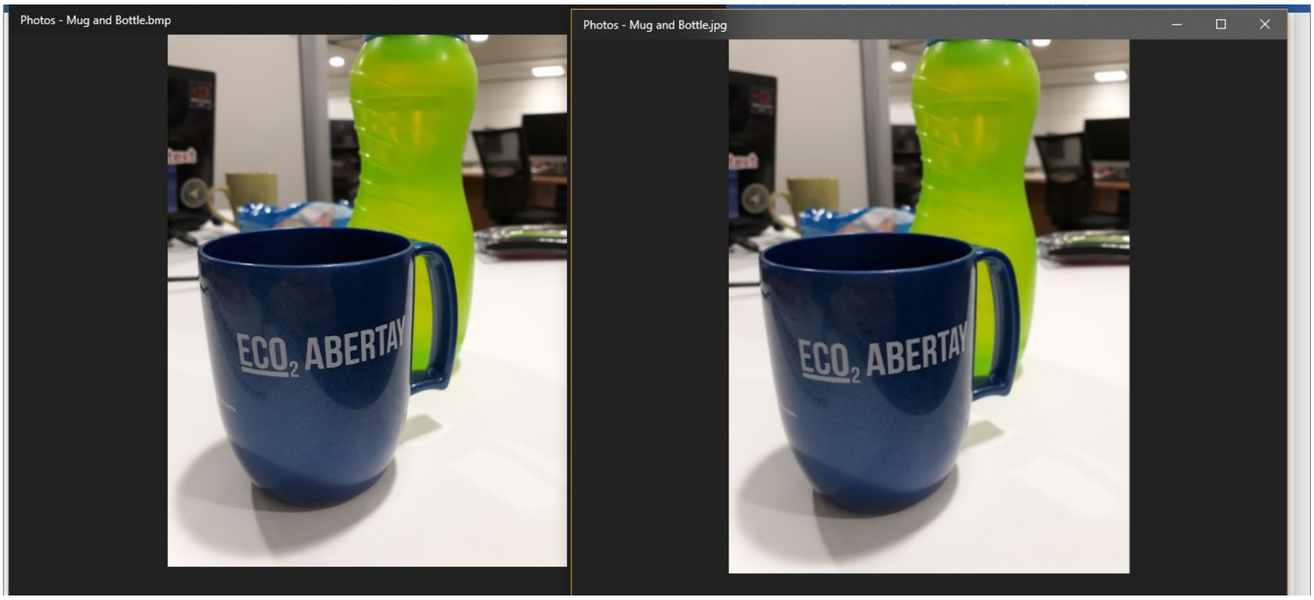
Appendix G

Comparison between original file and first watermarked file. (This watermark was the two lines of text)


Mug and Bottle.bmp - Windows Photo Viewer
File
Print
E-mail
Burn
Open
Watermarked image.bmp - Windows Photo Viewer
File
Print
E-mail
Burn
Open

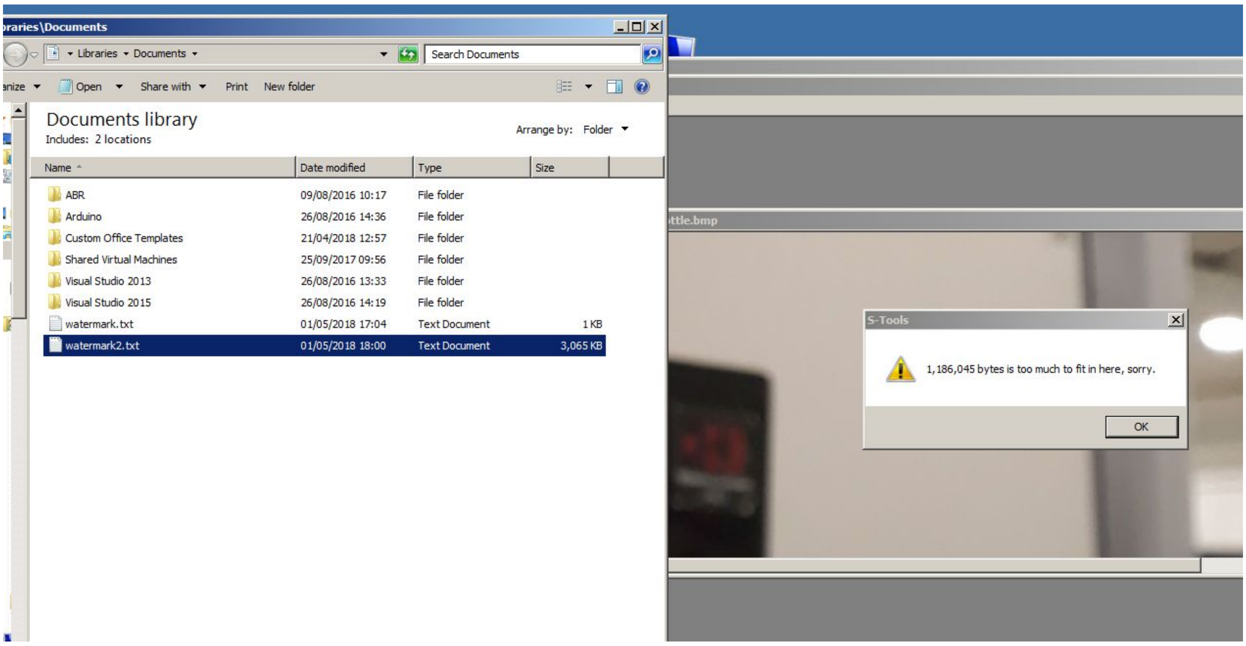
Appendix H

Comparison between the original JPEG image and the converted bitmap image.



Appendix I

Image shows how the text file of over 3MB was too large to be converted. On the left the large text file is highlighted in blue and on the right is the error message from S-Tools.

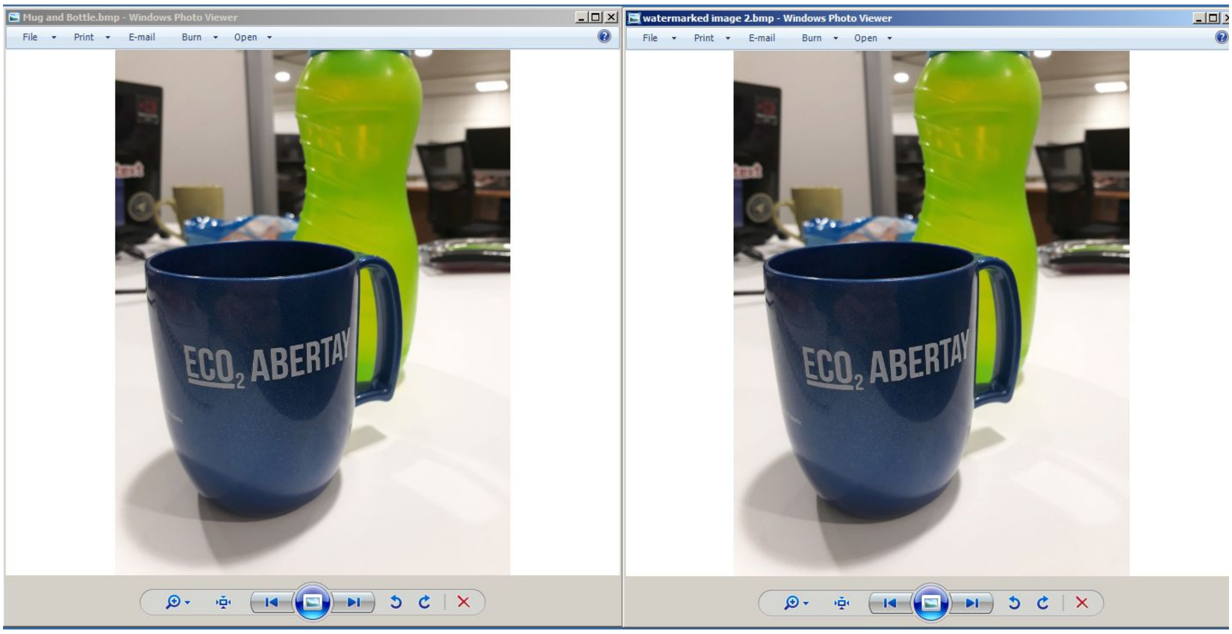


Name	Date modified	Type	Size
ABR	09/08/2016 10:17	File folder	
Arduino	26/08/2016 14:36	File folder	
Custom Office Templates	21/04/2018 12:57	File folder	
Shared Virtual Machines	25/09/2017 09:56	File folder	
Visual Studio 2013	26/08/2016 13:33	File folder	
Visual Studio 2015	26/08/2016 14:19	File folder	
watermark.txt	01/05/2018 17:04	Text Document	1 KB
watermark2.txt	01/05/2018 18:00	Text Document	3,065 KB

1,186,045 bytes is too much to fit in here, sorry.

Appendix J

Comparison between the original image file and the second watermarked file, containing 3MB of Bee Movie.


Mug and Bottle.bmp - Windows Photo Viewer
File ▾ Print ▾ E-mail Burn ▾ Open ▾
ECO2 ABERTAY
watermarked image 2.bmp - Windows Photo Viewer
File ▾ Print ▾ E-mail Burn ▾ Open ▾
ECO2 ABERTAY