What a Digital Forensics Investigator should know about Steganalysis of Digital Content

Abel Melinte
B00137882

Department of Informatics, School of Informatics and Engineering, Technological University Dublin, Dublin 15

[Word Limit for Assignment: 3500 Actual Word Count:]

Digital Forensics and Cyber Security
Computer and Network Forensics
12/12/2022

Declaration

I hereby certify that this material, which I now submit to Technological University Dublin in partial fulfilment of the requirements for the degree of Bachelor of Science (Honours) in Computing in Digital Forensics and Cyber Security, is entirely my own work except where otherwise stated, and has not been submitted for assessment for an academic purpose at this or any other academic institution other than in partial fulfilment of the requirements of that stated above.

Signed: Abel Melinte Dated: 02/12/2022

Plagiarism

I declare that the work I am submitting for assessment by the Institute examiner(s) is entirely my own work, except where the author or source has been duly referenced and attributed.

I confirm that this material has not been previously submitted for a degree or any other qualification at TUD or any at other institution.

I further confirm that I have read and understood the Institute policy on plagiarism in assignments and examinations (3AS08.doc) and that I am not, so far as I aware, in breach of any of these regulations.

Name: Abel Melinte Student ID: B00137882

Course: Digital Forensics and Cyber Security Module: Computer & Network Forensics

Signed: Abel Melinte Date: 02/12/22

Table Of Contents

Table Of Figures

1. Introduction

Steganography is an extremely powerful method that has been used for many years and it is also still being used until this day. Steganography is the creativity and science of unseen communication, in other words it means that someone can embed an important message within another message, which will lead the reader in thinking that the message is completely harmless. When a person comes across something suspicious within a file for example an encrypted message, this encourages the person do some investigation and have the curiosity to decrypt the encrypted message. (J.R Krenn 2004)

When the internet was created and released to the public the most highlighted section of communication and information technology has been security of information. The term "Cryptography" was fabricated as method for having the power to be able to secure the information using multiple methods to be able to decrypt data and be able to encrypt the data. The method that has been used to administer this is labelled as "Steganography". (J.R Krenn 2004)

Steganography in today's world has its own personal goal to keep its data undetectable. If you notice that the hidden content is not revealed to anyone it doesn't mean it's not there, it means that it's hidden from the interacted user. (J.R Krenn 2004)

2. Literature Review

2.1 What is Steganography

The word Steganography comes from Greek word “Stegos”, Stegos in the Greek language means “to cover over with silence” which in Steganography terms it means “to keep secret” and in Greek “Graphia” means “Writing” so when you collide the two together you get “Secret Writing” which Steganography is all about. Steganography is used in many ways but the main way it is being used is to embed secret details or secret messages into normal texts and forward it to the set destination without being known that is contains anything secret even it’s being looked at in detail (J.R. Krenn, 2004).

2.2 The History of Steganography

Steganography was first discovered and used by Herodotus in 500 BC. The story behind it was that Histiaeus (the ruler of Miletus) was departed from his home city, he planned to return to Miletus which was difficult as it was in control of his son in law, Aristagoras, so what Histiaeus planned to do was to stage a revolt in Ionia. This is where Steganography got introduced, he began to shave one of the slaves’ head that he had, and he tattooed a secret message that he didn’t want anyone to see on the slave’s head. He then held the slave captive until the slave began to grow his hair back, once the slaves’ hair grew back and prevent that tattoo from being visible, Histiaeus sent the slave to Aristagoras with the manual to cut the slaves head and view the secret message that has been tattooed to his head. This was the first-time steganography has been used and because of Histiaeus, it’s still used to this day of course with different methods, more advanced (Josh Lake, 2019)

2.3 Difference between Steganography and Cryptography

Steganography is more related on hiding the secret information/messages, whereas Cryptography is more related to being certain that the secret information cannot be access by anyone who have their hands on the file/image/video etc. Whenever someone has the intention to use Steganography, the person should be extremely certain that nobody will be able to see any sign of hidden communication occurring within that specific file/image/video. Whoever plans to use Cryptography they make sure that nobody has (Josh Lake, 2019)

2.4 Different methods of Steganography

In the Steganography space, mostly all digital file formats can be used for steganography, yet formats with a lot of redundancy are better since they are more efficient. There are multiple types of steganography which will take time to explain each into deep detail so here are the most common types of steganography that are being used (Josh Lake, 2019).

- Invisible Ink
- Null Ciphers

- Micro-dots
- Video Steganography
- Audio Steganography
- Picture Steganography

(Josh Lake, 2019).

2.4.1 Invisible Ink

The invisible ink was when someone used a specific ink so whenever someone looked at that letter or message, they wouldn't be able to see what he has written with that ink only the message with the normal ink would have been shown to the human eye. Invisible Ink was mainly seen when World War 2 was taken place, they used it for sending letters that did not seem odd whatsoever, but it had invisible ink inserted between the normal ink that was being used (Neil F. Johnson, Sushil Jajodia, 1998)

2.4.2 Null Ciphers

When it comes to Null Ciphers it's simple to use, all what it takes is to hide the actual hidden message among the normal messages. The most common use of Null Ciphers is making every nth letter, word as part of a hidden message. (Josh Luke, 2019)

Example

I dislike **Steganography** as it **is** not as **active** as other methods.

In this example every 3rd word is a hidden text, so the final hidden text will be

"Steganography is active"

2.4.3 Micro-dots

The micro-dots where first developed in half the years of the 19th century; they haven't been used to often for steganography until it got to World War 1. Micro-dots involved in minimizing an image or even a message to the size of a full stop, this gave the advantage of allowing people to communicate and being able to forward hidden information without being suspicious about any hidden information (Josh Lake, 2019).

2.4.4 Image Steganography

Images are one of the most known cover objects that are used in steganography. There are numerous different picture file formats in the world of digital images. the majority of them for certain applications. There are multiple steganographic algorithms for these various image file types. (T. Morkel, J.H.P. Eloff, M.S. Olivier, 2014)

2.4.5 Video Steganography

Video steganography is an addon to Image Steganography, a video can also be several multiple images allocated together. There are multiple different views that highlight the difference between video and image steganography. Whenever it comes to video steganography, whoever feels suspicious that the video file could possibly contain some hidden information, the investigator could have the intention to lower the video compression, change the framerate of the suspicious video file, adding and deleting frames of the video. In image steganography an investigator cannot do this as an image is static and if it contained any hidden information, it would be on that single image or frame. (Ritesh Upadhyay, Prof. Y. S. Thakur, Dir. D. K. Sakravdia, 2015)

2.4.6 Audio Steganography

Audio Steganography is the method of having the ability to hide secret audio in a specific audio file. The easiest and most effective method for audio steganography is the Least Significant Bit (LSB) alteration approach. The first method involves randomizing the host message's bit number, which is used to incorporate a secret message; the second method involves randomizing a sample digit carrying the next part of the encrypted/hidden message. (Muhammad Asad, Junaid Gilani, Adnan Khalid, 2011)

2.4.7 Bacon's Cipher

The Bacon Cipher was found and created by a politician, Francis Bacon in the 17th century. This method was highly investigated as it was more interesting than any other methods that there was back then, this was because it was hiding any implemented secret message in the format of the real text instead of its content. The Bacon Cipher was pointing more towards the steganography side instead of the cryptography side this was because the message was just normal like any other text rather than being encrypted and looking like random characters. The Bacon Cipher method has been successful and used numerous times in the recent years. An example of someone who used the Bacon Cipher were the "Notorious White Supremacist Gang", the Aryan Brotherhood, they used this method to have the ability to encode all the gang members and orders to go forward and attack their opponents/rival gangs. (Josh Lake, 2019)

2.5 Steganography Tools to Encrypt

There are multiple tools that can be used to encrypt/decrypt files, the top tools that are open source and free to use are these.

- Xiao Steganography
  Xiao Steganography is a fantastic, free Windows program that belongs to the category Security software and the subsection Encryption. Nakasoft.

- Steghide
  Steghide is a program for steganography that can conceal data in different types of images and audio files. The embedding is robust to first-order statistical tests since the colour-respectively sample-frequencies are left alone.

- Crypture
  software for steganography. Files are encrypted (1024 bit key) and kept in Windows bitmap files by Crypture. Only 6 KB, no installation necessary. fills all bits with noise so that it can evade detection by common steganalysis techniques. Additionally encrypted and dispersed is the data header.

- SteganographX Plus
  With the help of the SteganographX application, you may conceal text inside of a bitmap image. In BMP files with 16, 24, or 32 bits of color, the text will not be seen and the changes made to the bitmap image won't be obvious.

- rSteg
  A simple software program called rSteg was created expressly to make it easier for you to encrypt and conceal text messages within photographs. Similar functionality is displayed in the final images as it was in the originals. You can still open them in the viewer of your choice.

- OpenPuff
  OpenPuff is a specialized steganography program with distinct features that is appropriate for covert delivery of extremely sensitive data.

  And much more…

  (Pavitra Shankdhar, 2020)

Here is a quick guide on how to use one of the tools above, how to encrypt hidden text to an image while using OpenPuff.

After launching OpenPuff you're present with this software.

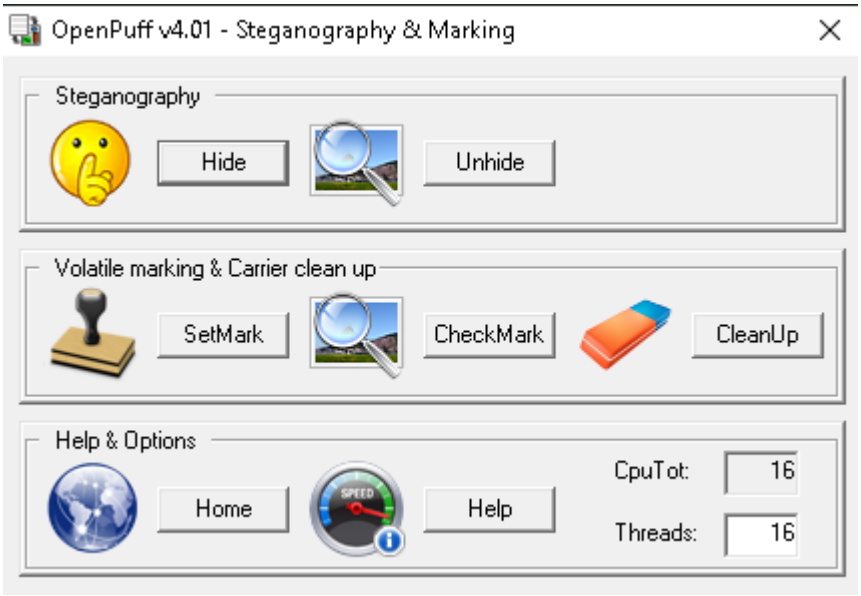

Figure 1: OpenPuff

To encrypt a text, you click on "Hide" and then the that will redirect you to the next section where it will show the full panel of OpenPuff.

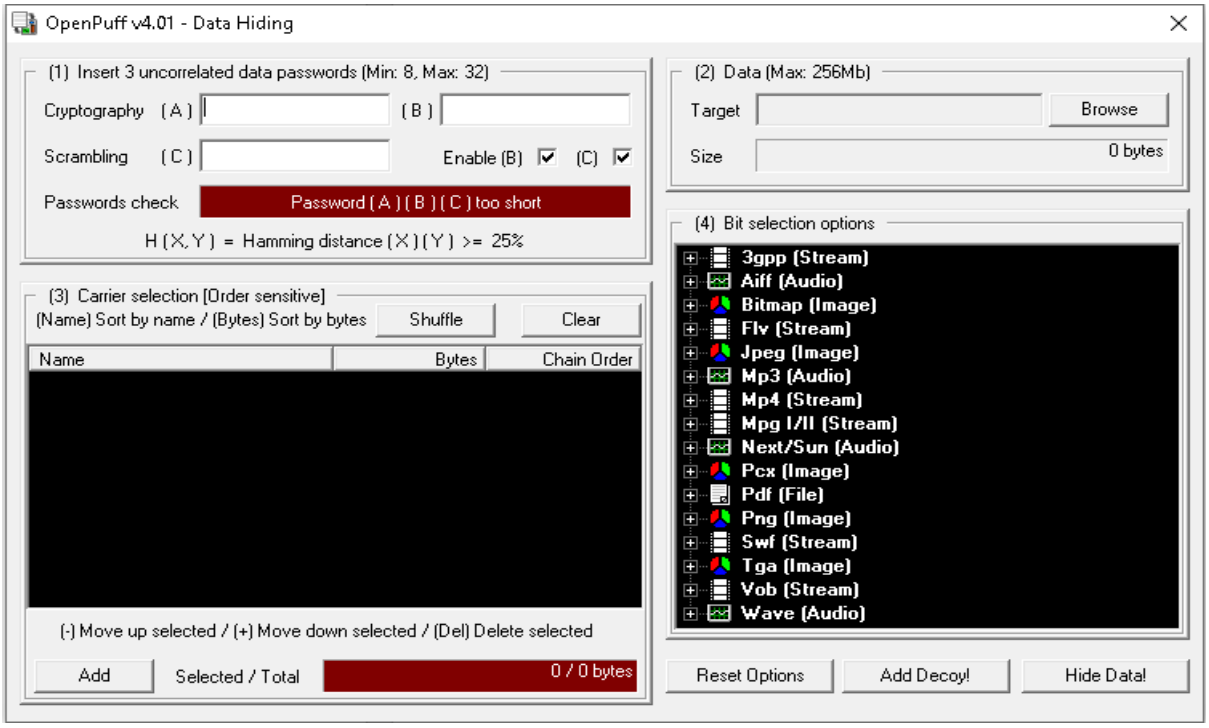

Figure 2: OpenPuff Panel

Then you're presented with a panel that requires you to add a password to access the file (1). You then need to add the txt file that will hold the hidden text within the image. (2)

Figure 3: Notepad

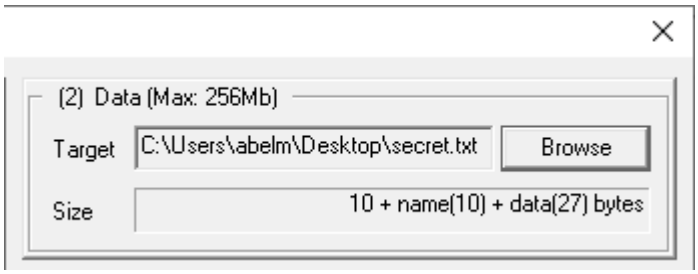

Figure 4: OpenPuff Selecting Notepad

You then search for an image that will be encrypted by the text you have created (3).

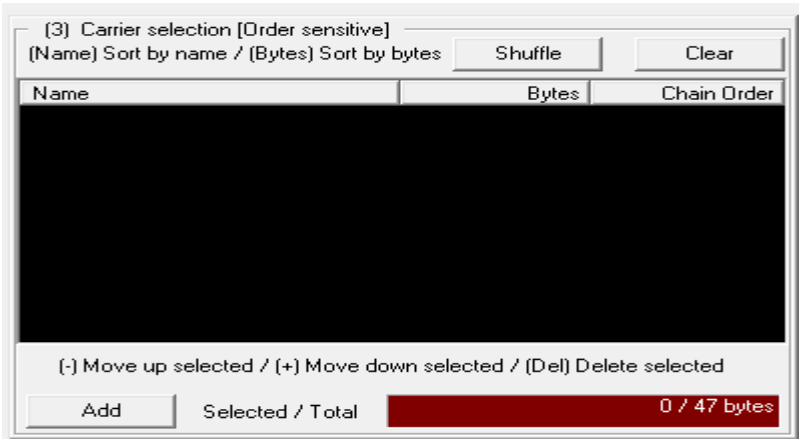

Figure 5: OpenPuff Selecting Image

Result :

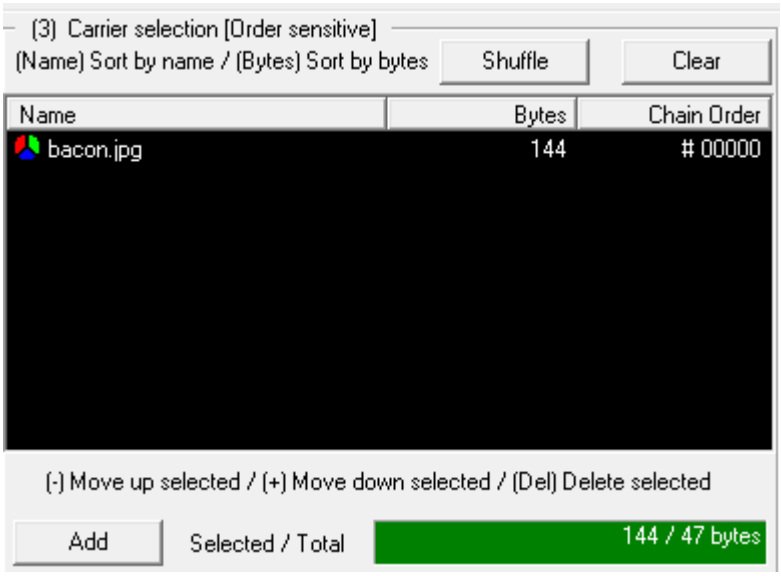

Figure 6: OpenPuff Selected Image

Then finally you add the bit selection output, which in this case will be JPEG (Maximum)(4)

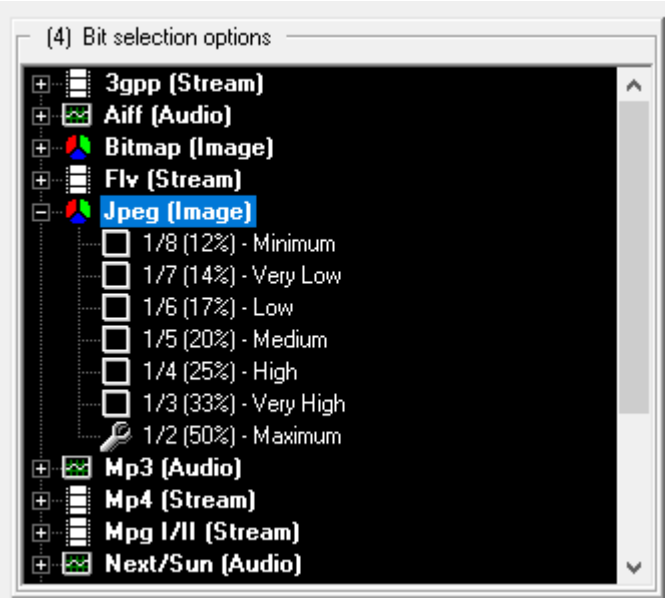

Figure 7: OpenPuff Bit Output

After the user has done these steps, you click on “Hide Data” and select a directory where you want your encrypted image to be hidden.

Figure 8: Hide Data Button

That was the last step, you now have an encrypted image that contains a secret text that no user will be able to view whenever checking the image out.

2.6 What is Steganalysis

Steganalysis is the area of study of having the ability to detect the presence of any encrypted or hidden information that may be embedded into media files such as images, videos, text. Steganalysis has received much audience and attention in the security and forensics science as it's so strong that it can prevent any embedded hidden messages from getting forward and being able to reach its destination. Steganalysis overall is quite challenging because of the lack of knowledge about any specific characteristics of the media, for example an image, video or audio which can be exploited to be able to hide any information and detect the same. Whenever approaching steganalysis it could also depend on the steganography algorithms that will be used. In steganalysis there are three main steganalysis algorithms, these are Video, Image and Audio. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.7 Image Steganalysis

The two main categories of image steganalysis algorithms are Specific and Generic algorithms. The Specific approach is one of a class of image steganalysis methods with a high success rate for spotting the underlying steganographic algorithm that was utilized. The "Generic" approaches the image steganalysis that it is individualistic of the partial section of the steganography algorithm that uses to hide messages and produce high standard results. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.7.1 Image Steganalysis Algorithms

Image Steganography algorithms are more focused within the embedding method which is named as the Least Significant Bit (LSB) embedding. Each individual pixel that is contained in an image is mainly known as a 24-bitmap value, a mixture of three bytes representing the R, G and B value for the popular colours Red, Green, and finally Blue. Whenever there is a higher RGB value that means that it implies a higher intensity. LSB embedding utilizes the fact that whenever changing the least crucial section of each of the three bytes of the pixel, it would only produce a slight change to the intensity of the colour the is being represented by the pixel and the change is not noticeable to the human eye. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.7.2 Palette Image Steganalysis

Palette Image Steganalysis is mainly focused on the GIF images. The GIF extension to images manages to 8 bits p/pixel and when it comes to the colours of a pixel, the colours are referenced from palette table that have a sum of 256 distinct colours to the 24-bit R, G, B colour space. The LSB embedding towards the GIF image switches the 24-bit R, G, B value of a specific pixel which this could modify the hue of the palette (among the 256 the pixel's distinct colour(s). The steganographic algorithm's strength is in lowering the likelihood that a pixel's palette colour will change and in limiting any noticeable distortion that incorporating the secret image can bring. A statistical study of the palette table is used to do the steganalysis of the GIF stego picture "vis-à-vis" when there is a noticeable entropy rise, the image and detection are then made.

2.7.3 Generic Image Steganalysis

When it comes to the Generic Image algorithm, it is referred to as the Universal or Blind Steganalysis algorithms, mainly work extremely good on all unknown and known steganography algorithms. When a message is inserted, these steganalysis techniques take use of how certain inherent attribute of the cover images alter. The goal is to pinpoint the distinctive aspects of an image that are monotonous and undergo statistically significant changes because of message embedding. To precisely and separate these changes, the general steganalysis algorithms are created. The selection of the appropriate attributes, which shouldn't change across photos of various types, has a significant impact on the prediction's accuracy.

2.8 Audio Steganalysis

The rapid advancement of VoIP and other P2P audio services aims to provide a variety of alternatives for communication conversion. With the aid of steganography tools, a minor adjustment to the binary chain of audio tests can quickly convert the transmission. Audio signals contain a typical excess and an unexpected nature that makes them the best feasible candidates to be utilized as a cover to convert communications so it can subsequently be able to disguise the secret messages. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.8.1 Audio Steganalysis Algorithms

There are four main audio steganography algorithms which are, Phase encoding, Low-bit encoding, Echo data hiding. The disadvantages to these four main algorithms are that it can be utilized as steganalysis. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.8.2 Phase Encoding

Phase Encoding is mainly focused on that fact which the phase components of sound are not noticeable to the human ear as normal noise is. The message bits are embedded as the shift phases in the phase range of the digital signal. This usually heads towards to the unheard embedding in terms of Signal-to-Perceived Noise Ratio (SPNR) and any hidden messages or information gets camouflaged inside the audio signal. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

The fact that the concealed message is only contained in the early part of the audio signal is demonstrated by phase coding's usual characteristic, which is primarily the lowest data transfer rate. By altering the phase link between the recurrence components of that particular segment, an increase in the portion's extent would have the reverse effect and cause a ripple. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.8.3 Low-bit Encoding

Low-bit Encoding is the binary type of the hidden data message which is exchanged along with the Least Significant Bit that representative of the audio file. Overall, even if this method is quite simple

and can also be used to be embedding larger messages, this method is not capable of protecting the secret messages from small mitigations that have the ability to reveal because of the format changing or lossy conversion. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.8.4 Echo Data Hiding

Echo Hiding is the information that is embedded by presenting an echo into a discrete audio signal. For example, SS coding, "Echo Hiding" gives access for a better standard data movement rate, and it also presents superior hale whenever it is set side by side to the noise-inducing strategies. In order to strongly hide the specific data, there are three parameters that need to be involved which are, decay rate, amplitude and finally offset from the real signal. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

2.8.5 Spread Spectrum Coding

Spread Spectrum also known as "SS", is the coding method which its function is to spread the bits of any secret/hidden data message alongside the frequency of a specific audio signal. The one thing is, unlike the LSB coding, the SS coding method sticks to spreading the hidden information while using a code which is liberated to the cover signal. Somehow the SS Coding method is actually able to perform quite better than the LSB coding and also even the other phase coding techniques. However, when it comes to the LSB coding method, the SS method is able to register noise to an audio file. Unfortunately, this vulnerability can easily be label for steganalysis.

2.9 Video Steganalysis

Whenever it comes to Video Steganalysis, it is extremely similar to Image Steganalysis but the only different is that the video has a frame-by-frame whereas the image is just a static frame without any other frames linked to it. (Natarajan Meghanathan, Lopamudra Nayak, 2010)

3.0 Steganalysis Tools to Decrypt

There are multiple tools that can actually be used to decrypt encrypted files, here is a concise list of the main tools that are used whenever some wants to discover steganography taken place within a file.

- StegSecret Tool
- StegoHunt Tool
- StegExpose Tool

- Virtual Steganographic Laboratory Tool
- Stegdetect Tool

(Hegarty, 2018).

3.0.1 StegSecret Tool

A steganalysis program called StegSecret, a Java-based open project, can find hidden data in various forms of digital material, including video, photos, and audio. The program is intended to find hidden data in various steganography. (Hegarty, 2018).

3.0.2 StegoHunt Tool

A software steganalysis program called StegoHunt is used to find the involvement of hidden data. It is commercially available. Payloads can be examined, feature cracked, and extracted using the StegoHunt tool. To brute force attack the encrypted data that was initially hidden, the application also provides password dictionaries. (Hegarty, 2018).

3.0.3 StegExpose Tool

A command-line interface steganalysis tool called StegExpose specializes in identifying the Least Significant Bit (LSB) approach, which is used to obfuscate data with an image. (Hegarty, 2018).

3.0.4 Virtual Steganographic Laboratory Tool

The Virtual Steganographic Laboratory (VSL) is a steganalysis tool that is available for free and is employed to find hidden data inside the Least Significant Bit. (Hegarty, 2018).

3.0.5 Stegdetect Tool

An automated open source steganalysis program called Stegdetect was created to find hidden data in digital photographs like the JPEG format. Using steganographic tools like jphide, jsteg, outguess, and others, Stegdetect has to ability decipher hidden data. (Hegarty, 2018).

3. Discussion

The outcomes of this research paper have highlighted the main topics of what a Digital Forensic Investigator should know about Steganalysis whenever he must work in something related to Steganography. The paper has been completed and has covered over strong information such as,

- What Steganography is overall.
- The History about it
- The Differences between Steganography and Cryptography
- The Different Methods that can be used such as
    - Invisible Ink
    - Null Ciphers
    - Micro-Dots
    - Image Steg.
    - Video Steg.
    - Audio Steg.
    - Bacon Cipher
- Steganography tools to encrypt
- What Steganalysis is overall
- Image Steganalysis
- Image Steganalysis Algorithms
- Audio Steganalysis
- Audio Steganalysis Algorithms
    - Phase encoding
    - Low-bit Encoding
    - Echo-data Hiding
- Video Steganalysis
- Steganalysis Tools that can be used to decrypt

4. Conclusion

In conclusion, overall, a digital forensics investigator will have a strong understanding of steganalysis to effectively investigate digital content and uncover any hidden or obscured information. This knowledge may include understanding common steganography methods, using appropriate tools and software for steganalysis, and following best practices for conducting a thorough examination. Additionally, an investigator should be conscious of potential challenges of steganalysis, as well as any legal considerations related to its use in an investigation. With a solid foundation in steganalysis, a digital forensics investigator can confidently and successfully uncover hidden information to support their investigations. When this paper will be fully read and used it as a learning paper, they will learn anything that is related Steganography and not just a summary but in strong deep detail.

5. Bibliography

Khare, P. and Singh, J. (2011) *Digital Image Steganography - ResearchGate*. Available at: https://www.researchgate.net/profile/Dr-Jaikaran-Singh/publication/216052617_DIGITAL_IMAGE_STEGANOGRAPHY/links/56f08f4808ae70bdd6c94df9/DIGITAL-IMAGE-STEGANOGRAPHY.pdf?origin=publication_detail (Accessed: December 10, 2022).

Morkel, T., Eloff, J.H.P. and Oliver, M.S. (2005) *An overview of image steganography - martinolivier.com*. Available at: http://martinolivier.com/open/stegoverview.pdf (Accessed: December 10, 2022).

Provos, N. and Honeyman, P. (2003) *Hide and seek: An introduction to steganography | IEEE journals ...* Available at: https://ieeexplore.ieee.org/document/1203220 (Accessed: December 12, 2022).

Krenn, J.R. (2004) *Steganography and Steganalysis*. Available at: https://www.krenn.nl/univ/cry/steg/article.pdf (Accessed: December 10, 2022).

Lake, J. (2021) *What is steganography and how does it differ from cryptography?, Comparitech*. Available at: https://www.comparitech.com/blog/information-security/what-is-steganography/ (Accessed: December 10, 2022).

Johnson, N.F. and Jajodia, S. (1998) *Exploring steganography: Seeing the unseen*. Available at: https://ieeexplore.ieee.org/abstract/document/4655281 (Accessed: December 10, 2022).

Upadhyay, R., Thakur, Y.S. and Sakravdia , D.K. (2015) *A Comparative Study of Un-optimized and Optimized Video Steganography* . Available at: https://www.ijcsit.com/docs/Volume%206/vol6issue04/ijcsit20150604102.pdf (Accessed: December 12, 2022).

Asad, M., Gilani, J. and Khalid, A. (2011) *An enhanced least significant bit modification technique for audio steganography*. Available at: https://ieeexplore.ieee.org/document/6020921 (Accessed: December 11, 2022).

Shankdhar, P. (2020) *Best tools to perform steganography [updated 2020], Infosec Resources*. Infosec Resources. Available at: https://resources.infosecinstitute.com/topic/steganography-and-tools-to-perform-steganography/#gref (Accessed: December 11, 2022).

Meghanathan, N. and Nayak, L. (2010) *STEGANALYSIS ALGORITHMS FOR DETECTING THE HIDDEN INFORMATION IN IMAGE, AUDIO AND VIDEO COVER MEDIA*. Available at: https://www.researchgate.net/profile/Mdjulkar-Nayeen-Mahi/post/steganography_using_DWT_quantitative_embedding_process/attachment/59d635cd79197b8077993492/AS%3A385703046336515%401468970180555/download/1010s4.pdf (Accessed: December 11, 2022).

Reilly, M. (2019) *Deep Learning Steganalysis*. Available at: https://vle-bn.tudublin.ie/pluginfile.php/298415/mod_resource/content/1/11711_Matthew_Reilly_B00092951_MatthewReilly_Thesis_232438_1722120062%20%281%29.pdf (Accessed: December 11, 2022).

Hetzl, S. (2003) *Steghide*. Available at: https://steghide.sourceforge.net/ (Accessed: December 12, 2022).

Nakasoft (2007) *Xiao Steganography*, *Softonic*. Available at: https://xiao-steganography.en.softonic.com/ (Accessed: December 12, 2022).

LeeLu (2010) *Download STEGANOGRAPHX plus 2.0*, *softpedia*. LeeLu Soft. Available at: https://www.softpedia.com/get/Security/Encrypting/SteganographX.shtml (Accessed: December 12, 2022).

Marculescu, A. (2010) *Download rSteg 0404*, *softpedia*. Abhinav Kumar Kushwaha, Alok Ranjan, S... Available at: https://www.softpedia.com/get/Security/Security-Related/rSteg.shtml (Accessed: December 12, 2022).

Oliboni, C. (2004) *OpenPuff - yet not another steganography SW*, *OpenPuff - Steganography & Watermarking*. Available at: https://embeddedsw.net/OpenPuff_Steganography_Home.html (Accessed: December 12, 2022).