TJHSST Computer Systems Lab Senior Research Project
Implementation of Least Significant Bit Steganography and its Steganalysis
2009-2010

Deniz Oran

January 27, 2010

Abstract

Although steganography is an ancient form of unobtrusive and covert communications, its implementation has become significantly more efficient due to the advent of image storage technology and the associated amounts of space available to encode a message. An implementation of Least Significant Bit (LSB) steganographical techniques with image files to covertly encode both text and image communications in the JAVA programming language is achieved through the inherant "noise" in most image formats. Thus far, the .PNG image format has been successfully used to encode a text message. After enabling the encoding of the images, a Graphical User Interface will display the status of the operation and will enable the detection of both steganographically compromised host or carrier images and will attempt to display the communication hidden within.

Keywords: Steganography, Steganalysis, Least Significant Bit encoding, image carrier

1 Introduction - Elaboration on the problem statement, purpose, and project scope

1.1 Scope of Study

Steganography has evolved historically with the digital age. More and more data is being sent to other computers through the internet and an inexpensive method of covertly sending that data was discovered. Encryption is typically expensive in terms of purchasing software and the computing power needed to encrypt and decrypt at the other end. Steganography

is rapidly evolving as an alternative to encrpytion and is considered a possible method of espionage and cyber-terrorism. Research in the field has rapidly accelerated to deter the misuse of steganographical encoding techniques. This program will be an introduction to the field of steganography because it uses one of the fundamental encoding methods.

1.2 Expected results

A comprehensive GUI that enables the user to encode text messages or images into a carrier image. The program will also enable a user to upload a suspected compromised image to be tested for its legitimacy. Because Least Significant Bit steganography is the method of choice for professional software, it is also the easiest to detect because most software is made to be able to detect its own products. As a result, further work with mixing cryptography and steganography will have to be done to improve the program's resiliency. The advantage of using LSB encoding is its high storage capacity. A file must have eight bytes for each encoded character, which is relatively low compared to other alternatives.

The initial research consisted of learning the binary structure of the various image formats and how to manipulate binary code. In addition, pixel structures and their manipulation was also researched in order to hide communications in their unused space. There are also numerous available steganography techniques, but few can encode as much information as LSB encoding. Some techniques involve digital watermarks, random insertion of fake messages, and password protection.

1.3 Type of research

This project would be user-inspired research because the ultimate goal is to effectively use the software in real-world communications. The program is not going to utilize something such as a command prompt because a GUI would better facilitate usability. Integrating the encoding and extraction code with the GUIis necessary for this program to be successful.

2 Background and review of current literature and research

The art and science of steganography has existed since the time of the Ancient Greeks and continues to be used for secure and covert communication. The initial method of unobtrusively concealing a sent message was to shave the hair off of a servant, inscribe the desired message, wait for the hair to grow back, and then send the servant to the recipient. A method that involved writing a on a wax tablet, covering it with fresh wax, and sending it by public means to therecipient was an even more progressive and economical method of secret communication. Steganography is therefore the opposite of encryption, in which a message is made unintelligible, but openly transmitted through public means. The concealment of the fact that a message is even being transferred is the true essence of steganography. The classic struggle of covertly communicating is best illustrated through the prisoners' dilemma,

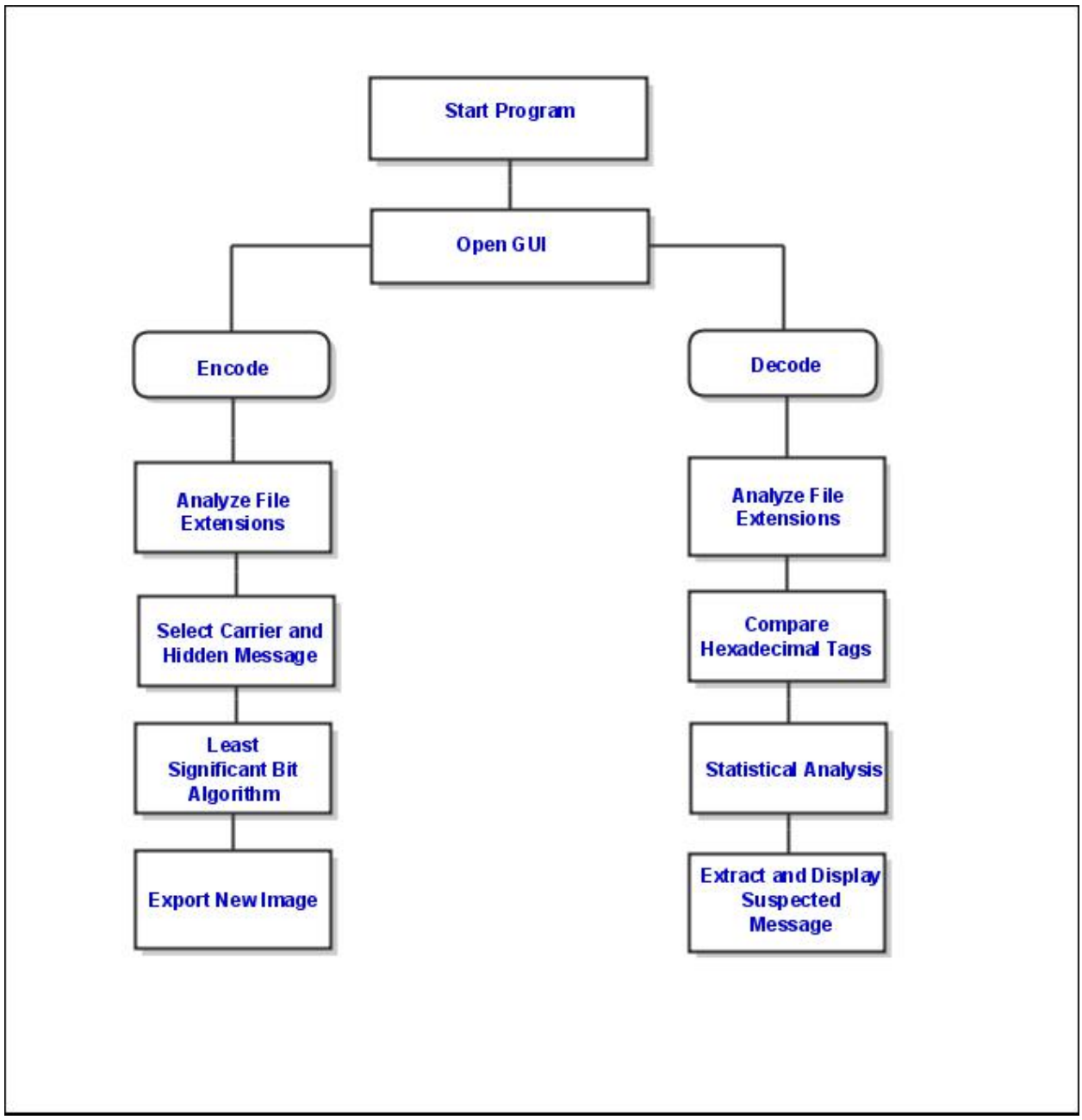
in which two prisoners must exchange plans to escape without provoking the suspicion of the warden. Encryption would be an ineffective method in this situation because as soon as the prison warden doesn't comprehend the communication he immediately shuts down all forms of communication between the prisoners, thus foiling their attempt at escaping.

3 Procedures and Methodology

The JAVA programming language will be used because of it's secure and stable structure along with its popularity. The most effective and most common way of covertly encoding a message or image is by hiding it in unused portions of commonly used lossless image formats such as .GIF or .PNG. To do this, a technique called Least Significant Bit encoding is employed to edit numerous picayune parts of the image and placing parts of binary code that can be compiled by thereader to form an image or a test message. The program therefore reads the entire binary composition of the image into a cubic data structure or three-dimensional array storing the pixel row, column, and color value. The least significant bits are then extracted and compiled together into a new matrix that is saved as either a new image or text message.

The second portion of the project is detecting the efficacy of the steganography. Clearly the first test would simply be visually ensuring that the original image and the carrier appear identical. The second, and more sophisticated, method, is through statistical analysis. A program will convert the suspected image into Hexadecimal code and will analyze the tag that is associated with the image format. Another method involves analyzing the consistency of the color transitions among the pixels of the image. The most likely method used to extract the information hidden by the encoding is simply reversing the encoding process. This method is usually guess by commercial software because of the prevalence of LSB encoding. To deter this, though, techniques such as inserting the information at varying places and "padding" some bits with '0's can be implemented. In addition, a commercial steganalysis program may be used to either confirm or deny the professional quality of the program.

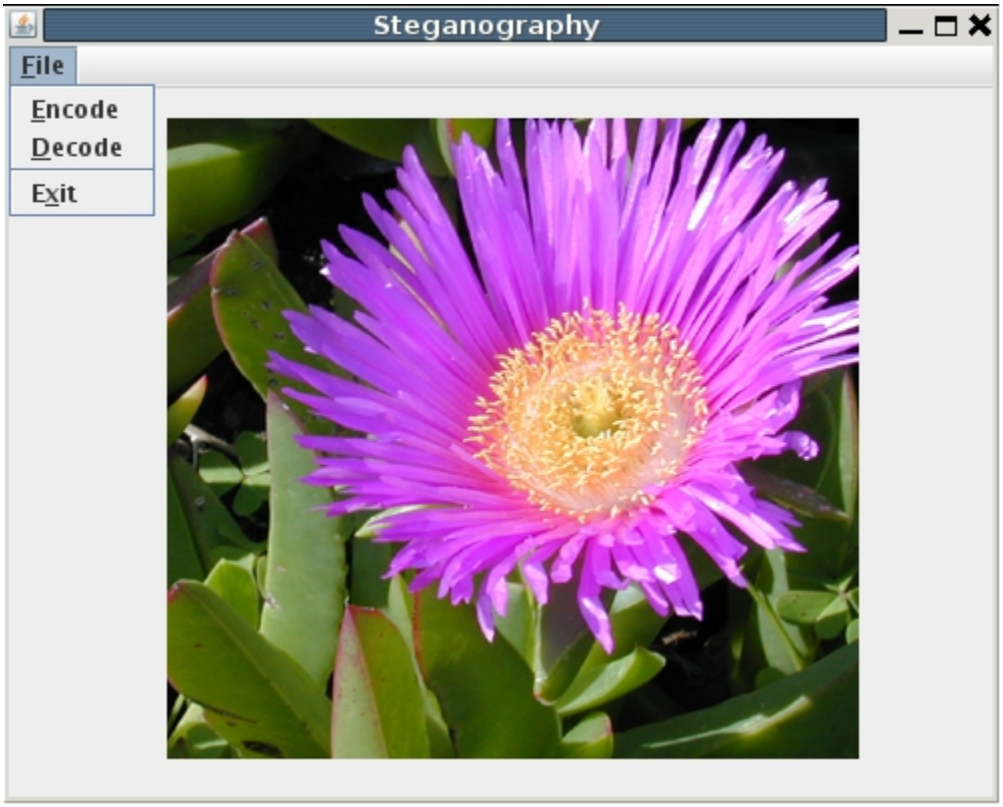
In order to compare the altered image with the original, both will be placed side by side for visual inspection because the essence of steganography is not for the message to be electronically tested. Regardless, if the image is suspected, the uncompromised carrier will be required to compare the images bit by bit. Otherwise, if the program knows that Least Significant Bit encoding was used, it can simply look for irregularities within the least significant bits. After the detection program is coded and tested, commercial programs will used to detect messages encoded by my program. This will ensure the validity of the program and will see how viable the program can be if implemented in a real-world situation or within the intelligence communitty.



4 Expected Results

An advanced implementation of steganography that passes at least a visual inspection will be able to be used and will not incite the suspicions of someone intercepting the message. Because of the GUI, the program will be incredibly easy to use and will display the image and what is being encoded. Then, if the decoding feature is used, it will display what is purportedly being hidden. In order to demonstrate that the program worked, the compromised image will be compared to the original. The the binary code for each will be compared and

will yield the discrepancy due to the usage of steganography. So far the GUI looks like the following:



A project utilizing steganography was previously done in this Senior Technology Lab in 2007 with .wav audio files. In contrast to the image steganography I am implementing, a future researcher can further investigate detection methods or steganography in .exe or other formats with white space. Similar to image steganography, far more information can be hidden in videos, which are just a compination of a series of images and sound. Videos can be effective because they vary in file size so a large file size would not be suspected. The only disadvantage is in fact the video size, which prevents the easy transmission of the message.

References

[1] Anderson "On the Limits of Steganography", 1998

[2] Bahar, H.B. "Image Steganography, a New Approach for Transferring Security Information", August 2008

[3] Dabeer, Onkar "Detection of Hiding in the Least Significant Bit", 2004

[4] Thampi, Sabu "Information Hiding Techniques: A Tutorial Review", 2004