REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
28-08-2012	Related Material	-

4. TITLE AND SUBTITLE
Steganography: LSB Methodology (Progress Report)

5a. CONTRACT NUMBER
W911NF-11-1-0174

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
206022

6. AUTHORS
Angel Sierra, Dr. Alfredo Cruz (Advisor)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Polytechnic University of Puerto Rico
377 Ponce De Leon
Hato Rey
San Juan, PR 00918 -

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
58924-CS-REP.12

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
In computer science, steganography is the science of concealing information within a computer file. Electronic communications may include the transport layer, document file, image file, program or protocol.

When working we steganography two major work areas can be identified, steganography and

15. SUBJECT TERMS
steganography; JPEG images; LSB Embedding

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Alfredo Cruz
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 787-622-8000

Report Title

Steganography: LSB Methodology (Progress Report)

ABSTRACT

In computer science, steganography is the science of concealing information within a computer file. Electronic communications may include the transport layer, document file, image file, program or protocol.

When working we steganography two major work areas can be identified, steganography and steganalysis. The goal of steganography is to hide the fact that a covert communication is present within an innocuous communication, in this work area you have access to the original work. The goal of steganalysis is to detect when a covert communication is occurring, in this work area you don't have access to the original work.

Aside of the background research performed to the concept of steganography we have done the following:

- Prepared the environment to run and test the Digital Invisible Toolkit in a windows 7 laptop
- Tested encode, decode, stegAnalysis and benchmark functions by providing a color image as the cover work and a notepad document as the message input.
- Reviewed the battlesteg algorithm java code.

POLYTECHNIC UNIVERSITY OF PUERTO RICO

Steganography: LSB Methodology

Progress Report

Angel N. Sierra

8/2/2012

Steganography is a method that conceals information within computer files. This concept is known to provide “security by obscurity”. In this report you will find some background information in regards to the basic concepts of steganography along with the research work performed until August 2, 2012.

Steganography Background

In computer science, steganography is the science of concealing information within a computer file. Electronic communications may include the transport layer, document file, image file, program or protocol.

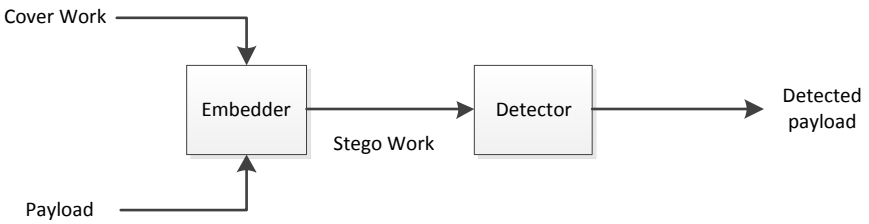

Figure 1: Generic steganography systems

When working we steganography two major work areas can be identified, steganography and steganalysis. The goal of steganography is to hide the fact that a covert communication is present within an innocuous communication, in this work area you have access to the original work. The goal of steganalysis is to detect when a covert communication is occurring, in this work area you don't have access to the original work.

While studying steganography concept, four major steganalytic methods were found.

Least Significant Bit Embedding

> [142] J. Fridrich, M. Gojan, and R. Du. Reliable detection of LSB steganography in grayscale and color images. In J. Dittmann, K. Nahrstedt, and P. Wohlmacher, editors, Proceedings of the ACM, Special Session on Multimedia Security and Watermarking, Ottawa, Canada, October 5, 2001, pages 27 – 30, 2001.
>
> [444] A. Westfeld and A. Pfitzmann. Attacks on steganographic systems. In A. Pfitzmann, editor, Information Hiding, 3rd International Workshop, IH'99, Dresden, Germany, September 29 – October 1, 1999, volume 1768 of LNCS, pages 61 – 75. Springer-Verlag, New York, 2000.

JPEG Images

> [132] H. Farid and L. Siwei. Detecting hidden messages using higher-order statistics and support vector machines. In F. A. P. Petitcolas, editor, Information Hiding, 5th International Workshop, IW 2002, Noordwijkerbout, The Netherlands, October 7 – 9, 2002, volume 2578 of LNCS, pages 340 – 354. Springer-Verlag, New York, 2002.
>
> [144] J. Fridrich, M. Goljan, and R. Du. Steganalysis based on JPEG compatibility. In A. G. Tescher, editor, Special Session on Theoretical and Practical Issues in Digital Watermarking and Data

Hiding, SPIE multimedia systems and applications IV, Denver, CO, August 20 – 24, 2001, volume 4518, pages 275 – 280, 2001.

Audio Files

[443] A. Westfeld. Detecting low embedding rates. In F. A. P. Petitcolas, editor, Information Hiding, 5th International Workshop, IH 2002, Noordwijkerbout, The Netberlands, October 7 – 9, 2002, volume 2578 of LNCS, pages 324 – 339. Springer-Verlag, Berlin, 2002.

Multimedia Files

[147] J. Fridrich, M. Goljan, and D. Soukal. Perturbed quantization steganography using wet paper codes. In J. Dittman and J. Fridrich, editors, Proceeding ACM Multimedia and Security Workshop , Magdeburg, Germany, Septeber 20 – 21, 2004, pages 4 – 15. ACM Press, New York, 2004.

[442] A. Westfeld. High Capacity despite better steganalysis (F5 – a steganographic algorithm. In I. S. Moskowitz, editor, Information Hiding, 4th International Workshop, volume 2137 of LNCS, pages 289 – 302. Springer-Verlag, New York, 2001.

As it can be seen Steganography could be very abroad in terms of techniques and algorithms. For this reason we have decided to start studying one of the most basic techniques for steganography known as Least Significant Bit (LSB).

Also, detecting modifications in a cover image using LSB embedding method is more difficult than using JPEG embedding method as expressed in J. Fridrich, M. Gojan and R. Du paper titled “Reliable detection of LSB steganography in grayscale and color images”.

From a general perspective Figure 2 presents the scheme used for steganography, in this scheme we need to generate an embedding function and extracting function. In the embedding function we use the cover work (i.e. source), the message that wants to be included and the stego key as inputs. In the extraction function we only need the stego key as input in order to obtain the output message embedded before.

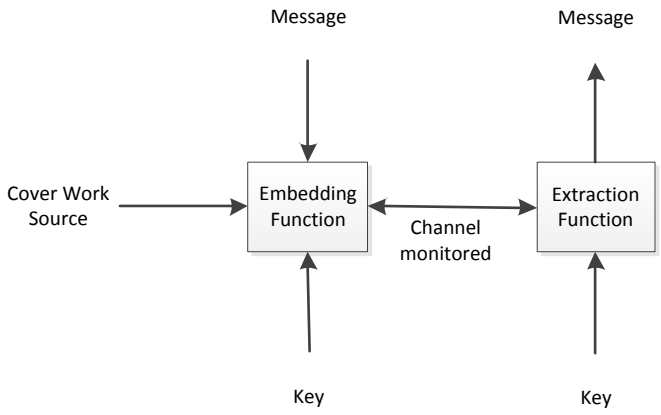

Figure 2: Steganographic Embedding Scheme

In the embedding function for this scheme we need to provide a cover work, a message and a stego key. The cover work is nothing more than the media where the message is to be embedded, in the LSB case it would be an image with GIF or BMP file format. The second required input will be the message, which in the java program we are working is a document with .txt file format (e.g. notepad document). The third and last input is a stego Key which is used as a security measure to ensure that the message is retrieved by the person holding the right key. Once the embedding function gets the inputs requested it will generate a new image called "the stego image" which will be used as an input by the extraction function.

In the extraction function the stego image and the stego key is used as input, this function will generate a .txt file with the embedded message.

Steganography Tool

For our research purposes we have used an open source program in java called Digital Invisible Toolkit written by Kathryn Hempstalk found in a web page known as SourceForge (http://diit.sourceforge.net/). This open source program includes several algorithms for encoding information and several algorithms for analyzing the results to look for steganography. The hiding algorithms all work on the least significant bits. The more sophisticated algorithms use Sobel filters or Laplace filters to flag the best pixels for hiding information. These tend to flag the edges between objects where the intensity of the image changes quickly.

This tool offers four distinctive functions related to steganography, Encode (i.e. embedding function), Decode (i.e. extraction function), StegAnalysis and Benchmark. The encode and decode functions are the typical primary functions of a steganography program and the authors offers six different algorithms to perform the task as shown in Table 1.

Name	Description
BattleSteg	Hides by filtering the image to obtain a list of ships (best places to hide), then randomly 'shoots' at the image until a 'ship' is 'hit'. For a short while the shots are clustered around the ship, and then it moves away and begins randomly shooting again.
BlindHide	Starts writing at (0, 0) and moves along each pixel, colour and bit in scan lines across the image. Uses pure steganography.
DynamicBattleSteg	Works like BattleSteg only uses dynamic programming to prevent lots of memory from being used. This is NOT compatible with normal BattleSteg.
DynamicFilterFirst	Works like FilterFirst only uses dynamic programming to prevent lots of memory from being used. This is NOT compatible with normal FilterFist
FilterFirst	A pure steganography method, FilterFirst uses edge detecting filters to obtain an ordered list of the best places to hide. It then writes to this list in the order given.
HideSeek	HideSeek randomly picks a pixel / colour / bit and hides there. If it picks a bit it's written to before, it will skip over it and go onto the next randomly selected bit

Table 1: Encode and Decode algorithms

The stegAnalysis function offers three different types of analysis to determine the presence of a message in the stego-work (i.e. RS Analysis, Sample Pairs, and Laplace Graph).

The benchmark function was created to compare the cover-work with the stego-work and compare the difference found after the modification of each algorithm. This function uses eight different analysis to observe the differences (i.e. average absolute difference, correlation quality, Laplacian mean squared error, lp norm, mean squared error, normalized cross-correlation, peak signal to noise ratio and signal to noise ratio).

Work Summary

Current Work

Aside of the background research performed to the concept of steganography we have done the following:

- Prepared the environment to run and test the Digital Invisible Toolkit in a windows 7 laptop
- Tested encode, decode, stegAnalysis and benchmark functions by providing a color image as the cover work and a notepad document as the message input.
- Reviewed the battlesteg algorithm java code.

Future Work

In the near future we expected to continue our research as follows:

- Attempt to gather more specific information about the six different algorithms used for the encode and decode functions

- Study and Research the StegAnalysis functions and determine how they are used to determine if an image contains any hidden message.
- Provide a more deep study to the statistical analysis performed for the benchmark.
- Provide a more deep review to each one of the algorithm code from a programming perspective
- Study the possibility of moving the program from a desktop application to a web application

Project Opportunities

Based on the information gathered we could start different projects in parallel to have a more practical and complete understanding of the steganography concept. Some of the projects ideas are as follow:

- JPEG Embedding – The purpose of this project is to investigate and find a way to embed a hidden message in a JPEG image and compare its efficiency with LSB embedding method.
- Audio File Embedding – The purpose of this project is to investigate and find a way to embed a hidden message in an audio file to obtain a more practical experience of this methodology
- Digital Invisible Toolkit benchmark – The purpose of this project is to see the effects of grayscale, image quality, image size and message size in a cover work. We will be preparing a wide sample of pictures encoding each one of them with different size message and prepare a database where the different statistical analysis results will be stored for comparison and study purposes
- Digital Invisible Toolkit as a Web application for steganography – The idea of this project is to provide users a program that could encode and decode messages in cover-work without the need of having an application to be installed in each machine used.
- Digital Invisible Toolkit as a Mobile application for steganography – The idea of this project is to provide users the option of embedding hidden messages to pictures they take with their own phone and allow them to send the stego-work as a SMS or e-mail message.

References

Cox, I. J., Miller, M. L., Bloom, J. A., Fridrich, J., & Kalker, T. (2008). *Digital Watermarking and Steganography, Second Edition.* Burlington, MA: Morgan Kaufmann.

Wayner, P. (2009). *Disappearing Cryptography Information Hiding: Steganography & Watermarking, Third Edition.* Burlington, MA: Morgan Kaufmann.