STEGANOGRAPHY IN GIF ANIMATED GRAYSCALE IMAGES

MASSALIN YERLAN

A project report submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Computer Science (Information Security)

Faculty of Computing
Universiti Teknologi Malaysia

MAY 2013

This thesis is dedicated to my family for their endless support and encouragement.

ACKNOWLEDGEMENT

First and foremost, I would like to express heartfelt gratitude to Allah, most gracious, most compassionate, who allowed me to join this remarkable University, then to my supervisor **Dr. Bakhtiari Majid** for his constant support during my study at UTM. He inspired me greatly to work in this project. His willingness to motivate me contributed tremendously to our project. I have learned a lot from him and I am fortunate to have him as my mentor and supervisor.

Besides that, I would like to thank the authority of Universiti Teknologi Malaysia (UTM) for providing me with a good environment and facilities to complete this project with software which I needed during process.

ABSTRACT

Computer steganography is the one of the most rapidly developing areas, because for a long time it was underestimated and neglected. The computer steganography gives an opportunity to embed information within a digital object, at the same time, striving to minimize the impact on the carrier object. One of the most popular carrier media are still images due to their immense popularity in the Internet compared with other formats of digital objects such as video or audio formats. And one of the most suitable formats from among still images is BMP format, due to the absence of compression loss and larger capacity, that allows to hide larger amounts of data, than for example, JPEG images. Herewith, the author proposes an survey in the area of Bitmap images. So there is a certain logic in our work, when we start from BMP survey and proceed to an animated GIF format, since the frames inside GIF format are of BMP format as well. And we propose an algorithm based on LSB method, that selectively changes the LSB according to the properties in the images.

ABSTRAK

Steganografi Komputer adalah salah satu kawasan yang paling pesat membangun, kerana untuk masa yang lama ia dipandang ringan dan diabaikan. Steganografi komputer memberi peluang untuk menerapkan maklumat dalam objek digital, pada masa yang sama, berusaha untuk meminimumkan kesan pada objek pengangkut. Salah satu media pembawa yang paling popular adalah imej-imej pegun kerana populariti yang besar di Internet berbanding dengan format lain objek digital seperti video atau format audio. Dan salah satu daripada format yang paling sesuai dari kalangan imej masih format BMP, kerana ketiadaan kehilangan mampatan dan kapasiti yang lebih besar, yang membolehkan untuk menyembunyikan lebih besar jumlah data, daripada contohnya, JPEG imej. Bersama-sama ini, penulis mencadangkan satu kajian dalam bidang imej Bitmap. Jadi ada logik tertentu dalam kerja kita, apabila kita mula daripada kajian BMP dan teruskan ke format GIF animasi, kerana bingkai dalam format GIF adalah format BMP juga. Dan kita mencadangkan satu algoritma berdasarkan kaedah LSB, yang terpilih berubah LSB mengikut sifat-sifat dalam imej.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	x
	LIST OF FIGURES	xi
1	**INTRODUCTION**	
	1.1 Initial Overview	1
	1.2 Problem Background	2
	1.3 Problem Statement	4
	1.4 Project Objectives	4
	1.5 Scope of the Study	5
	1.7 Thesis Contributions	6
	1.8 Summary	7
2	**LITERATURE REVIEW**	
	2.1 Introduction	8
	2.1.1 Practical Aspects of Steganography	11
	2.2.1 Definition of a Stegosystem	12
	2.2.2 General Requirements for a Stegosystem	14
	2.2.3 Key Components of a Stegosystem	14
	2.1.1 Some Common Issues in Hiding Information	16

2.1.1.1 Payload Issue 16

2.1.1.2 Integrity and Robustness Issue 17

2.2 ±1 Embedding Steganalysis Classification 18

2.2.1 Training Phase 19

2.2.1 Testing Phase 22

2.2.1.1 Cross Validation 22

2.2.1.2 Performance Measures 23

2.3 ±1 Embedding Benchmark Review 26

2.3.1 ±1 Embedding Steganography 26

2.3.2 ±1 Embedding Stegoanalyzers 28

2.3.2.1 Stego Noise (WAM) Statistics 28

2.5 Bitmap Domain 30

2.5.1 A Bitmap Header 31

2.5.2 A Bitmap Data Block 32

2.6 Grayscale Image Domain 33

2.7.1 GIF Format Structure 38

3 **METHODOLOGY**

3.1 Introduction 41

3.2 Research Framework 41

3.3. Phase One: Domain Analysis 42

3.4.1 Design Sub-Phase 44

3.4.2 Implementation Sub-Phase 45

3.4.2.1 Algorithm Implementation General Schedule 45

3.4.2.3 Implementation System 47

3.5 Phase Three: Evaluation and Benchmarking Phase 47

3.5.1 Evaluation Criteria 48

3.5.1.1 Algorithm Implementation General Schedule 48

3.5.1.2 Histogram Method 50

3.6 Summary 50

4 **STEGOSYSTEM DESIGN**

4.2 Choosing a Stegosystem Type 51

4.3 Cover Image Choice Rule 52

4.4 Algorithm General Outline 52

4.4.1 Pre-embedding Procedures 53

4.4.1.1 Laplace Transformation 53

4.4.2 Embedding Method 58

4.5 Stegosystem Special Features 59

4.5.1 Undetectability Issue 59

4.5.2 Keep the Payload Ratio 60

5 IMPLEMENTATION AND TESTING

5.1 The Algorithm Refinement 61

5.2 Survey on Image Modification by Proposed LSB Method 64

5.2.1 General Guideline 64

5.2.2 Experimental Basic Data 64

5.2.3 Simple Comparison Method 66

5.2.4 Histogram 69

5.2.5 PSNR 71

6 CONCLUSION

6.1 General Summary 73

6.2 Complex Efficiency Evaluation 74

6.3 Recommendation for Future Works 75

REFERENCES 76

APPENDIX A 78

APPENDIX B 82

APPENDIX C 83

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Binary classification rates	24
2.2	Bitmap Header Structure	31
2.3	Bitmap information block structure	32
4.1	Dispersion parameters	57
5.1	Statistics of the input Test Images	65
5.2	Results of the Embedding Process	66
5.3	Stego bits extraction from the container	69
5.4	PSNR results.	72

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
2.1	Steganography and its related fields	9
2.2	Stegosystem structure	15
2.3	Bitmap block structure	33
2.4	A 4-bit color image	34
2.5	A 4-bit Gradient	34
2.6	A 8-bit color image	35
2.7	A 8-bit Gradient	35
3.1	Research framework structure	42
3.2	Algorithm building roadmap	46
3.3	A cover image – the Engine	48
3.4	A stego image – Vinnie-Pooh	48
4.1	An original image	53
4.2	An image with Laplace filter applied	53
5.1	A Stego-algorithm pseudocode	64
5.2	A cover image – the Engine	65
5.3	A stego image – Vinnie-Pooh	65
5.4	Comparison of changes in hex mode	67
5.5	Change in color depth caused by stego tool	68
5.6	Original image histogram (2nd frame)	70
5.7	A stego image histogram (2nd frame)	70
5.8	Original image histogram (3rd frame)	70
5.9	A stego image histogram (3rd frame)	70
5.10	Original image histogram (4th frame)	70
5.11	A stego image histogram (4th frame)	70
5.12	Original image histogram (5th frame)	71

5.13 A stego image histogram (5th frame) 71

CHAPTER 1

INTRODUCTION

1.1 Initial Overview

Steganography word is derived from the two Greek words *steganos* meaning "covered or protected" and *graphei* meaning "writing", and thus means "concealed writing". It is the art of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message, a form of security achieved through obscurity. The first use of the term was recorded in 1499 by Johannes Trithemius in his Steganographia. Although cryptography and steganography could be considered is relatives to each other in sense of achieving of confidentiality in communication of information assets, nevertheless, goals of cryptography and steganography are different.

Steganography refers to hiding fact of delivery information asset rather than to securing information asset itself from unauthorized access as cryptography does. Steganography provides concealing the very fact of information existence, whereas the goal of cryptography is to secure information, existence of which is well-known. Apart from this, cryptography performs secure communication from eavesdropping as well. In pre-digital world the hidden message could be in invisible ink between the visible lines of a private letter, like sympathetic ink. As to digital world, hiding messages in digital objects, concerns a wide range of information assets that surpass the scope of any subject in field of information security, whereas cryptography for instance is restricted just on securing communication between sender and recipient as well as information itself. In such stego messages can hide in something common: audio, video files, articles, and even nowadays executables, or they supposed to assume any other possible form.

For this reason sometimes it is quite hard to track down all possible ways of information leakage, the fact that places steganography on distinguished position in information security field. Why steganography sometimes is preferred to cryptography. No everybody knows, that in England if one refuses to share one's password of one's encrypted hard disk to police or other control service on request, for example customs, laws of this country allow to jail such a person up to 2 years. In other words what is good for one person or community is not the same for another person or community.

In our thesis we refer to steganography in pixel domain, techniques that allow to hide message in still images, after the image is nothing more than a grid or matrix that stores the numerical values of the elements called pixels.

1.2 Problem Background

Over the last decade, steganography, in particular image steganography, has become a matter of great interest. Considerable variety of techniques has been proposed especially for wide known pixel domain formats like jpeg, gif, etc. Generally adopted idea about undetectability runs as follows: less embedding causes less detection. The point of view as arguable and is not completely correct (Cancelli, G. and Barni, M., 2007), but it is good start point of research for improving initial steganographic techniques. Furthermore, a new approach has been appeared whose principle lies in channel coding techniques oriented to reduce embedded message of a stego work, e.g. Wet Paper Coding work of Fridrich, J., *et al.* (2005). Another approach, especially in JPEG domain is to adjust in some way image statistics containing embedded message, regards image cover using subset of support function. Some authors (Fridrich, J., *et al.* 2007) recently even went on further, trying to calculate upper limit of stego message payload to embed before exposing work onto detection threat by using common JPEG stegoanalyzers.

A stegoanalyzer algorithm should decide whether is given image is a cover or stego object. But it runs its function in such way that it designed only for specific steganography hiding method. In other words conventional steganography methods suffer lack of heuristics methods. However, current steganalysis trends are shifting towards so called *blind* steganalysis, which is oriented to detect algorithms without knowledge of their structure and design, and consequently designed to detect widest possible range of stego Works, even those whose implementation previously has not been known.

A classics of steganalysis is a steganographic algorithm which known as ±1 embedding, or as LSB matching, and hides messages within pixel images. Due to its effectiveness based on simple concept and resistance to detectability, ±1 embedding has become as an industrial standard in steganography science.

The technique sometimes is also called LSB replacement. The asymmetry produced by LSB embedding effect, assumes a form of a statistical anomaly which is reflected in the histogram depicting intensities in the form of intensity values pairs; and it approximately matches the frequency in case when the cover Work contains stego message. The above mentioned circumstance is a subject for steganalysis purpose.

±1 embedding or the LSB matching, is a more refined variant of simple LSB replacement method. Unlike simple replacing the LSB with the message bit, the corresponding pixel value incremented or decremented in random way, whenever the LSB value is changed. Hence, the asymmetry which is peculiar to LSB, is almost erased after flipping manipulations with the bits. But although statistical anomalies more subtle and discrimination accuracy is significantly lower than for LSB embedding, they still appear after applying flipping effects which, in turn, permits discrimination between cover and stego Works. The latter circumstance is the subject of interest for stego analysts.

1.3 Problem Statement

Conventional steganograhy tools rely on JPEG, BMP, TIFF, and other formats with large capacity, but not so many surveys were dedicated to a GIF format, due to belief that embedding bits affect cover GIF more drastically then normal images due to poor pallette comprising only 256 colours for RGB scheme. The LSB method is a simple technique example for embedding hidden messages onto still images. But for handling such a still image type with insufficient depth as 8-bit of GIF, there is a need to for proposing a technique based on LSB, that embeds a concealed message inside animated GIF as a cover for embedding message, but at the same time deliberately chooses areas within GIF's frames to embed message, avoiding thus distortion of the plain areas with less pixels density.

1.4 Project Objectives

Due to restricted number of its colors in range of 256 colors, the GIF format is rather vulnerable to possible stegoanalysis attacks in case if it is used for stego container.

In the domain of GIF animated image steganography we define our objective as building a key-based stegosystem with an algorithm that would be able to embed bitmap stego-image into bitmap frames located inside GIF animated format in grayscale palette with less affect on cover image and consequently less detectablility together with absence of possibility to retrieve it for unauthorized person without a secret key.

1.5 Scope of the Study

We claim to make research into “passive” steganography field in pixel domain and its state-of-art techniques, and how to improve and strengthen those techniques against common steganalysis methods. These methods are meant to detect and retrieve LSB embedding steganography and are called LSB steganalysis methods. So, we construct our review on LSB steganography from mutual adversaries sides: hiding and detection. From the side of hiding we introduce such effective solution for resistance to ±1 steganalysis as an animated GIF cover image and improvements related to methods of hiding hidden message in GIF image area.

The main function of the steganalysis algorithm is to make decision whether an image is a normal or contains hidden message. Some stegoanalyzers go even beyond “binary” detection manner, trying to estimate the embedded message size and sometimes the essence of the message. In our research, we don't focus on the first stage, viewing analysis as a “binary” point of view, i.e. whether a given image is stego or non-stego Work, due to the definition that has been made clear in Chapter 2.2.1, we consider robustness and payload as well.

While analyzing adversaries' tools are represented with stegoanalyzer techniques, we drive our attention to analysis algorithm features, bypassing such a bulky concept as classifier. Classification has a long history and we do not go into details on the subject of classification. Instead of this we propose reader to proceed to Pattern classification of Duda, R. *et al.*, 2000.

Concerning reviewing steganalysis techniques, we do not concentrate on relative advantages of one or another classification algorithms, like Support Vector Machine or Fisher Linear Discriminant (FLD). But we mainly focus on generic properties and issues that are proper and can be applied to all classification algorithms. In particular, we consider common to all algorithms two phases which are general components of a classification system: training and test phases.

While common stegoanalyzers base on dictionary developed with training phase to differ cover work from stego work, *blind* steganalysis are similar in their functionality to heuristics algorithms. In other words they are not aware of the underlying steganographic algorithm structure. Consequently, blind method is expected to detect the a message embedded with different algorithms, even including unknown algorithms. On the other side, targeted steganalysis knows about given underlying steganographic algorithm, for which it was specially designed. In our reveiw, we concern about targeted steganalysis, and in particular, the ±1 embedding detection.

1.7 Thesis Contributions

The Thesis contribution has two type of outputs: BMP survey and GIF animated object research. From BMP format, first, we review the structure and related features of this format and describe its difference to other formats like JPEG. In particular, we describe difference in hiding messages in BMP format comparing to other known formats.

In the part, dedicated to the GIF image area, we also reprsent typical features, and advantages and drawback in the sense of using this format as a means for establishing covert channel, or in other words using this format as a cover image for embedding messages.

1.8 Summary

This thesis is about research of steganography issues in the pixel domain.

In Section 1, we have made the reader well informed about background steganography field, we outlined the problem, and objectives of research.

In Section 2, information is expanded into details of steganography of BMP images and GIF images.

Regarding BMP steganography, we have made a review of stganography field in the Section 2.3.1, by making review of LSB matching and an improved LSB replacement algorithm, also called ±1 embedding. Also, in Section 2.5 we propose a new algorithm

In Section 3, we assume a new approach, based on GIF grayscale domain, by improving such parameters as detectability and payload. Moreover we compose a set of parameters to evaluate effectiveness of a new method.

And in Chaper 4 we get initial results based on approaches highlighted and proposed in Chapters 2 and 3.

In Chaper 5 we get final results based on approach by embedding message into a GIF cover using proposed method in Chapter 4.

Eventually, in Chapter 6, we summarize our achieved work results, and provide recommendations of future works.

REFERENCES

1. Fridrich, J., Goljan, M., Lisonek, P. and Soukal, D., “Writing on wet paper,” *IEEE Transactions on Signal Processing, vol. 53, no. 10 Part 2,* 2005. 3923–3935.
2. Fridrich, J., Pevnẏ, T. and Kodovskẏ, J., “Statistically undetectable jpeg steganography: dead ends challenges, and opportunities,” in *Proceedings of the 9th workshop on Multimedia & security*. ACM New York, NY, USA, 2007. 3–14.
3. Zhang, J., Cox, I. J. and Doërr, G., “Steganalysis for LSB Matching in images with high-frequency noise,” *IEEE 9th Workshop on Multimedia Signal Processing (MMSP),* 2007. 385–388.
4. R. Duda, P. Hart, and D. Stork, Pattern classification. *Wiley-Interscience*, 2000.
5. Cox, I. J., Miller, M., Bloom, J., Fridrich, J. and Kalker, T., Digital watermarking and steganography. *Morgan Kaufmann*, 2007.
6. Shannon, C., “Communication Theory of Secrecy Systems,” *Bell System technical Journal*, *vol. 28.* 1954, 656–715.
7. Simmons, G. J., “The prisoners’ problem and the subliminal channel,” i*n Advances in Cryptology: Proceedings of CRYPTO’83. Plenum Pub Corp, 51–67*, 1984.
8. Dabeer, O., Sullivan, K., Madhow, U., Chandrasekaran, S. and Manjunath, B. S., “Detection of hiding in the least significant bit,” *IEEE Transactions on Signal Processing*, vol. 52, no. 10, October 2004, 3046–3058.
9. Ker, A. D., “A general framework for structural analysis of LSB replacement,” *Proceedings of the 7th Information Hiding Workshop, LNCS*, vol. 3727, June 2005, 296–311.
10. Currie, C.E.I., “Surmounting the Effects of Lossy Compression on Steganography” *19th National Information Systems Security Conference,*

Baltimore, Maryland, 22—25 Oct. 1996, 194-201.

11. Anonymous, “Algorithm hides data inside unaltered images”, *Information Security Monitor*, 1999. -1999. -Vol. 14, 8, 10.
12. Kahn, D., “The History of Steganography”, *Proceedings of the first Workshop of Information Hiding*. 1996, 1-5.
13. Jost, P., Vandergheynst, P., and Frossard, P., “Redundant image representations in security applications,” *International Conference on Image Processing (ICIP),* vol. 4, 2004.
14. Genne, O.V., “The main provisions of steganography” *Journal "Information Security. Confident"*, 2000. - № 3.
15. Gribunin, V.G., Okov I.N., Turintsev I.V., “Digital steganography”, *M. Solon Press*, 2002. - 272.