JPEG Steganography Application for Mobile Devices

By

Degrano, Karla Joahnna Shayne A.

Mahilum, Jedd Benedict Kris T.

Sua, Susan S.

SCHOOL OF ARTS AND SCIENCES

ATENEO DE DAVAO UNIVERSITY

MARCH 2009

JPEG Steganography Application for Mobile Devices

An Independent Study

Presented to

The Faculty of the Computer Studies Division

Ateneo de Davao University

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Science in Computer Science

By

Degrano, Karla Joahnna Shayne A.

Mahilum, Jedd Benedict Kris T.

Sua, Susan S.

SCHOOL OF ARTS AND SCIENCES

ATENEO DE DAVAO UNIVERSITY

MARCH 2009

ACKNOWLEDGMENTS

The proponents have undergone and survived a lot of hurdles as they accomplished this study. And so, they would like to extend their deepest gratitude to a number of people.

First of all, the proponents would like to thank the CS Division faculty and staff, most importantly Ma'am Ma. Teresa T. Quindoy – Subject Coordinator, for inspiring and driving the proponents to work on the fulfillment of their study.

Second, they would like to express their appreciation to their fellow batch mates Tetle, Ruki, Paul, Kcy, Nicoy, Janx, Sonito, Ming, Nanay, Will, Hao, MarVince, Kakai, Prenzy, and Arian. Without the unceasing support that these people have given, this study would not have reached its completion.

Third, the proponents would like to show their gratitude to their adviser, Mr. Antonio "Tony" Bulao III, who has been with the proponents in their ups and downs. Without his guidance and advises, the proponents would never have been able to continue the study.

Fourth, to the proponents' parents and family, who have contributed to the accomplishment of the study by aiding the proponents in terms of the expenses, the proponents would forever be indebted to you.

Last, but most definitely not the least, the proponents would like to extend their praise and gratitude to God, Almighty Father, for without His never-ending support, providence, and intervention, the proponents would have long been lost in the path of downfall.

Once again, the proponents express their heart-felt deepest gratitude to all

TABLE OF CONTENTS

CHAPTER 4 THEORETICAL BACKGROUND

CHAPTER 5 RESULTS AND DISCUSSION

ABSTRACT

Nowadays, mobile devices, such as the common cellular phone, have become more than just simple telecommunications devices. Instead, they have become an extension of the owner's personal life. These mobile devices now hold sensitive and very personal data owned by the user. Hence, a need for a better security mechanism on these mobile devices arises from this situation. This can be done by applying the concept of Steganography, particularly Image Steganography, on these mobile devices. By hiding the sensitive data inside a JPEG image, these data are now free from unauthorized viewing.

Keywords:

Steganography, Image Steganography, Mobile Devices, Security

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

It truly is remarkable how technology now drives the world. From small household machines to supercomputers that hold tremendous data, indeed, we are living in the technological era. However, living in this kind of world also has its downfalls. One might be surprised to know that the September 11 attack of terrorists against the United States of America was aided by a technology to have effective covert communications. Many say that the terrorists used Steganography to communicate with one another while effectively hiding their plans from other people. Although there has never been any proof that the terrorists did indeed use Steganography for their attack, the assumption still points out the effectiveness of Steganography as a means of obscuring data. Indeed, Steganography is one of the fundamental ways by which data can be kept confidential.

But what really is Steganography. Is it a new kind of security measure developed recently? The answer is no. It has been around since the times of ancient Rome. At that time, text was traditionally written on wax that was poured on top of stone tablets. When the sender of the information wanted to obscure the message – for purpose of military intelligence, for instance – they would use Steganography. The wax would be scraped off and the message would be inscribed or written directly on the tablet. They would then pour wax on top of the message, thereby obscuring not just its meaning but its very existence.

So how can the concept of Steganography be of use to us today since using stone tablets would be very impractical? Like many security tools nowadays, Steganography can be used for a variety of reasons. For one, it can be used for legitimate purposes such as watermarking images for reasons such as copyright protection. In addition, it can be used as a way to make a substitute for a one-way hash value where you take a variable length input and create a static length out put string to verify that no changes have been made to the original variable length input (*Applied Cryptography, Bruce Schneier, John Wiley and Sons Inc., 1996).* Furthermore, Steganography can be used to tag notes to online images and other images. Finally, it can be used to maintain the confidentiality of valuable information, to protect the data from possible sabotage, theft, or unauthorized viewing (*Steganography: Hidden Data, by Deborah Radcliff, June 10, 2002).*

By now, Steganography sounds a lot like Cryptography. So what is their difference? First, let us define Cryptography. Webopedia.com defines Cryptography as the art of protecting information by transforming it (encrypting it) into an unreadable format, called cipher text. Only those who possess a secret key can decipher (or decrypt) the message into plain text. As clearly stated, Cryptography only alters the message by converting it into gibberish text. It does not hide it from plain sight. Therefore, the difference between Steganography and Cryptography is that in the latter, one can tell that a message has been encrypted, but he/she cannot decode the message without knowing the proper key whereas in the former, the message itself may not be difficult to decode, but most people would not be able to detect the presence of that message. For this reason, the combined power of the Steganography and Cryptography can provide two levels of security.

As aforementioned, we now live in a technological world. The world which once was so vast and big is now getting smaller. With the current models of mobile phones, most, if not all, people can access the "world" by means of the Internet which they can also access from their mobile phones. More and more people are thus become highly dependent on their mobile phones. Some would even consider these gadgets as an extension of their personal lives. As a result, these mobile phones now hold private and very sensitive data of the users. Thus, a need for better security arises from this fact. This gives the proponents their main drive for this study. That is, to develop a JPEG Steganography application for mobile phones. So far, there has been no Steganography application that is implemented on mobile phones. The proponents decided to choose to work with JPEG images since this is the most common file type on modern mobile phones. With this JPEG Steganography application, mobile phone users can have the guarantee that their private and very sensitive data stored in their mobile phones are safe from any form of unauthorized access or viewing.

1.2 Statement of the Problem

The research study sought to answer the following problem: How would the use of JPEG Steganography improve file secrecy on mobile devices?

In particular, it sought to answer the following questions:

- How will the Steganography application hide files on a JPEG image (JPEG is the common image file type on mobile devices)
- What are the different approaches for Image Steganography? Is JPEG Steganography the most appropriate technique to use?
- What will happen to the image after the files are hidden in it? What will happen to the files when the image is altered? How will the files be retrieved after the Steganography process?
- How will related and existing PC-based Steganography applications aid the proponents in their study

1.3 Objectives of the Study

The main objective of the proponents' study was the development of a Steganography application that would have improved file secrecy on mobile devices. The following are the proponents' sub-objectives on their study:

- To identify the process of hiding information on JPEG images;
- To be able to name the different (common) approaches to Image Steganography and compare these approaches to JPEG Steganography.
- To determine whether the JPEG image retains its image properties after the files are hidden in it and whether future alterations of the JPEG image affect the hidden files;
- To find out the process of retrieving the hidden files and verifying that these hidden files could, indeed, still be retrieved; and,
- To have a feel on the concept of Steganography by trying the different PC-based Steganography applications present.

1.4 Significance of the Study

The study intends to enhance file security of present mobile devices by developing a JPEG Steganography application, which can easily be installed on these devices.

The application will provide a better means of protection from unauthorized viewing and transferring of sensitive data and the possibility of data sabotage by hiding these sensitive data on a JPEG image, which is commonly found on mobile devices. This will give an impression to the public that there seems to be no such data attached on the JPEG image; thus ensuring that these files will be secured.

Moreover, the study will also talk about the different approaches to Image Steganography, which might be useful for future related studies on this concept.

Lastly, the study and the application may serve as a form of related works/literature for further investigations and developments on the use of JPEG Steganography not only on mobile phones but also, possibly, on other forms of communications media.

1.5 Scope and Limitations of the Study

The study will focus on the task of enhancing file secrecy on mobile devices. This will be achieved by incorporating the concept of JPEG Steganography as a mode of protecting sensitive data by hiding them on a JPEG image. Thus, the study is geared towards the development of a JPEG Steganography application for MMS-capable mobile devices.

As mentioned, the mobile devices must be MMS-capable since these types of devices support JPEG images, which are the main medium for the application of the JPEG Steganography concept.

The JPEG Steganography application to be developed must not only hide files; it must also allow file retrieval. The proponents will not assure that the files stored or hidden on the image will be retained once the image is altered, however, the proponents will try to develop ways that will allow file retention even if the image is altered.

The proponents will develop the JPEG Steganography application using Java and they are looking into either NetBeans IDE 6.1 or EclipseME for the development environment.

1.6 Definition of Terms

In this study, certain terms have been mentioned that seem to be too technical to be easily understood. For the convenience of the readers, these terms shall be clarified and clearly defined as follows:

Binary Representation: A binary representation is a number written using Base 2. Each digit may have only one of two possible values (0 or 1) hence the name, binary. Binary representations have many applications in computer science because they map well onto the "off" and "on" states of electronic transistors.

Bit: A bit is a binary digit, taking a logical value of either "1" or "0" (also referred to as "true" or "false" respectively). Binary digits are a basic unit of information storage and communication in digital computing and digital information theory.

Bitmap: In computer graphics, a bitmap is a type of memory organization or image file format used to store digital images. The term bitmap comes from the computer programming terminology, meaning just a map of bits, a spatially mapped array of bits.

Byte Array: Simply an array of bytes.

Byte Code: Byte code is a term which has been used to denote various forms of instruction sets designed for efficient execution by a software interpreter as well as being suitable for further compilation into machine code.

Cover Image: Image used to hide a file in Image Steganography.

Discrete Cosine Transform: A discrete cosine transform (DCT) expresses a sequence of finitely many data points in terms of a sum of cosine functions oscillating at different frequencies. DCTs are important to numerous applications in science and engineering, from lossy compression of audio and images (where small high-frequency components can be discarded), to spectral methods for the numerical solution of partial differential equations.

File: A computer file is a block of arbitrary information, or resource for storing information, which is available to a computer program and is usually based on some kind of durable storage. A file is durable in the sense that it remains available for programs to use after the current program has finished.

File Hiding: The process of hiding files by means of Steganography.

File Retention: The ability of a file to retain any hidden files within it after changes are made to that file.

File Retrieval: The process of retrieving the hidden file inside another file and reconstructing it to its original format.

Filename Extension: A filename extension is a suffix to the name of a computer file applied to indicate the encoding convention (file format) of its contents.

Hexadecimal: In mathematics and computer science, hexadecimal (also base-16, hexa, or hex) is a numeral system with a radix, or base, of 16. It uses sixteen distinct symbols, most often the symbols 0–9 to represent values zero to nine, and A, B, C, D, E, F (or a through f) to represent values ten to fifteen.

IDE: An integrated development environment (IDE) also known as integrated design environment or integrated debugging environment is a software application that provides comprehensive facilities to computer programmers for software development.

Image Steganography: The application of Steganography process on Images.

J2ME: In computing, the Java Platform, Micro Edition or Java ME (still commonly referred to by its previous name: Java 2 Platform, Micro Edition or J2ME) is a specification of a subset of the Java platform aimed at providing a certified collection of Java APIs for the development of software for tiny, small and resource-constrained devices. Target devices range from industrial control and automotive devices to cell phones and set-top boxes.

J2SE: Java 2 Platform, Standard Edition or Java SE is a widely used platform for programming in the Java language. It is the Java Platform used to deploy portable applications for general use. In practical terms, Java SE consists of a virtual machine, which must be used to run Java programs, together with a set of libraries (or "packages") needed to allow the use of file systems, networks, graphical interfaces, and so on, from within those programs.

JPEG: In computing, JPEG is a commonly used method of compression for photographic images. The degree of compression can be adjusted, allowing a selectable tradeoff between storage size and image quality. JPEG typically achieves 10:1 compression with little perceptible loss in image quality.

JPEG Steganography: Image Steganography specifically applied on JPEG.

Least Significant Bit: In computing, the least significant bit (lsb) is the bit position in a binary integer giving the units value, that is, determining whether the number is even or odd. The lsb is sometimes referred to as the right-most bit, due to the convention in positional notation of writing less significant digits further to the right. It is analogous to the least significant digit of a decimal integer, which is the digit in the ones (right-most) position

Lossless: Lossless data compression is a class of data compression algorithms allowing the exact original data to be reconstructed from the compressed data.

Lossy: A lossy compression method is one where compressing data and then decompressing it retrieves data that may well be different from the original, but is close enough to be useful in some way.

MIDP: Mobile Information Device Profile (MIDP) is a specification published for the use of Java on embedded devices such as mobile phones and PDAs. MIDP is part of the Java Platform, Micro Edition (Java ME) framework and sits on top of Connected Limited Device Configuration (CLDC), a set of lower level programming interfaces.

MMS: Multimedia Messaging Service, or MMS, is a telecommunications standard for sending messages that include multimedia objects (images, audio, video, rich text). MMS is an extension of the SMS standard, allowing longer message lengths and using WAP to display the content.

Most Significant Bit: In computing, the most significant bit (msb) is the bit position in a binary number having the greatest value. The msb is sometimes referred to as the left-most bit on big-endian architectures, due to the convention in positional notation of writing more significant digits further to the left.

Pixel: In digital imaging, a pixel (or picture element) is the smallest item of information in an image. Pixels are normally arranged in a 2-dimensional grid, and are often represented using dots, squares, or rectangles. Each pixel is a sample of an original image, where more samples typically provide more-accurate representations of the original.

Processing Power: Also referred as the number of instructions per second that the processor for a device can execute.

RGB: The RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green, and blue.

RGB Component: RGB color model of a particular pixel in an image.

Simulation: Simulation is the imitation of some real thing, state of affairs, or process. The act of simulating something generally entails representing certain key characteristics or behaviours of a selected physical or abstract system.

Steg File: The retrieved file from a Steg Image.

Steg Image: The resulting image after Image Steganography process is done.

Steganography: Steganography is the art and science 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 through obscurity.