

POLYTECHNIC UNIVERSITY
Department of Computer and Information Science

Introduction to Steganography

K. Ming Leung

Abstract: This material deals with the history and an introduction to the art and science of steganography. A simple example illustrating how a large number of hidden images can be embedded within a given image.

Directory

- [Table of Contents](#)
- [Begin Article](#)

Copyright © 2000 mleung@poly.edu
Last Revision Date: November 11, 2004

Table of Contents

1. Introduction
2. History and Steganography
3. Embedding Images in the LSB of an Image

1. Introduction

Steganography simply takes one piece of information and hides it within another.[1, 2] Computer files (images, sounds recordings, even disks) contain unused or insignificant areas of data. Steganography takes advantage of these areas, replacing them with information (encrypted mail, for instance). The files can then be exchanged without anyone knowing what really lies inside of them. An image of the space shuttle landing might contain a private letter to a friend. A recording of a short sentence might contain your company's plans for a secret new product. Steganography can also be used to place a hidden "trademark" in images, music, and software, a technique referred to as watermarking.

Steganography sometimes is used when encryption is not permitted. Or, more commonly, steganography is used to supplement encryption. An encrypted file may still hide information using steganography, so even if the encrypted file is deciphered, the hidden message is not seen.

Steganography (literally meaning covered writing) dates back to

ancient Greece, where common practices consisted of etching messages in wooden tablets and covering them with wax, and tattooing a shaved messenger's head, letting his hair grow back, then shaving it again when he arrived at his contact point.

Until recently, information hiding techniques received very much less attention from the research community and from industry than cryptography. This situation is, however, changing rapidly and the first academic conference on this topic was organized in 1996. The main driving force is concern over protecting copyright; as audio, video and other works become available in digital form, the ease with which perfect copies can be made may lead to large-scale unauthorized copying, and this is of great concern to the music, film, book and software publishing industries. At the same time, moves by various governments to restrict the availability of encryption services have motivated people to study methods by which private messages can be embedded in seemingly innocuous cover messages.

Steganography has gained much more popularity since the September 11 attacks on the U.S. Some people suspected that terrorists connected with the incidents might have used it for covert communica-

tions. While no such connection has been proven, the concern points out the effectiveness of steganography as a means of obscuring data. Indeed, along with encryption, steganography is one of the fundamental ways by which data can be kept confidential.

Steganalysis is the detection and the task of the recovery of hidden data.

2. History and Steganography

Throughout history, a multitude of methods and variations have been used to hide information. David Kahn's "The Codebreakers" [3] provides an excellent accounting of this history. Bruce Norman recounts numerous tales of cryptography and steganography during times of war in "Secret Warfare: The Battle of Codes and Ciphers" [4].

One of the first documents describing steganography is from the Histories of Herodotus. In ancient Greece, text was written on wax covered tablets. In one story Demeratus wanted to notify Sparta that Xerxes intended to invade Greece. To avoid capture, he scraped the wax off of the tablets and wrote a message on the underlying wood.

He then covered the tablets with wax again. The tablets appeared to be blank and unused so they passed inspection by sentries without question.

Another ingenious method was to shave the head of a messenger and tattoo a message or image on the messengers head. After allowing his hair to grow, the message would be undetected until the head was shaved again.

Another common form of invisible writing is through the use of invisible inks. Such inks were used with much success as recently as World War II. An innocent letter may contain a very different message written between the lines. Early in World War II steganographic technology consisted almost exclusively of invisible inks. Common sources for invisible inks are milk, vinegar, fruit juices and urine. All of these darken when heated.

With the improvement of technology and the ease as to the decoding of these invisible inks, more sophisticated inks were developed which react to various chemicals. Some messages had to be "developed" much as photographs are developed with a number of chemicals in processing labs.

Null ciphers (unencrypted messages) were also used. The real message is "camouflaged" in an innocent sounding message. Due to the "sound" of many open coded messages, the suspect communications were detected by mail filters. However "innocent" messages were allowed to flow through. An example of a message containing such a null cipher from is:

Fishing freshwater bends and saltwater
coasts rewards anyone feeling stressed.
Resourceful anglers usually find masterful
leapers fun and admit swordfish rank
overwhelming anyday.

By taking the third letter in each word, the following message emerges:

Send Lawyers, Guns, and Money.

The following message was actually sent by a German Spy in World War II:

Apparently neutral's protest is thoroughly
discounted and ignored. Isman hard hit.

Blockade issue affects pretext for embargo
on by products, ejecting suets and
vegetable oils.

Taking the second letter in each word the following message emerges:

Pershing sails from NY June 1.

Note that i represents 1 in Roman numerals. As message detection improved, new technologies were developed which could pass more information and be even less conspicuous. The Germans developed microdot technology which FBI Director J. Edgar Hoover referred to as "the enemy's masterpiece of espionage." Microdots are photographs the size of a printed period having the clarity of standard-sized type-written pages. The first microdots were discovered masquerading as a period on a typed envelope carried by a German agent in 1941. The message was not hidden, nor encrypted. It was just so small as to not draw attention to itself (for a while). The information can be read by the intended recipient using a microscope.

With many methods being discovered and intercepted, the Office of Censorship took extreme actions such as banning flower deliveries

which contained delivery dates, crossword puzzles and even report cards as they can all contain secret messages. Censors even went as far as rewording letters and replacing stamps on envelopes.

With every discovery of a message hidden using an existing application, a new steganographic application is being devised. There are even new twists to old methods. Drawings have often been used to conceal or reveal information. It is simple to encode a message by varying lines, colors or other elements in pictures. Computers take such a method to new dimensions as we will see later.

Even the layout of a document can provide information about that document. There is a series of publications dealing with document identification and marking by modulating the position of lines and words. Similar techniques can also be used to provide some other "covert" information just as 0 and 1 are informational bits for a computer. As in one of their examples, word-shifting can be used to help identify an original document. A similar method can be applied to display an entirely different message. Take the following sentence (S0), produced for example in HTML:

<PRE>

<TT>

We explore new steganographic and cryptographic algorithms and techniques throughout the world to produce wide variety and security in the electronic web called the Internet.

</TT></PRE>

and apply some word shifting algorithm (S1) to produce the sentence in HTML format:

<PRE>

<TT>

We
explore
<fontsize="-1">
new steganographic and cryptographic algorithms
and techniques throughout

the world

```
<font size="-1">&nbsp;</font>
to produce
<font size="+1">&nbsp;</font>
wide
<font size="-1">&nbsp;</font>
variety and security in the electronic
<font size="+1">&nbsp;</font>
web
<font size="-1">&nbsp;</font>
called the Internet.
</TT></PRE>
```

This is achieved by expanding the space before **explore**, **the**, **wide**, and **web** by one point and condensing the space after **explore**, **world**, **wide** and **web** by one point in sentence S1. Independently, the sentences containing the shifted words appear harmless, but combining this with the original sentence produces a different message:

explore the world wide web.

3. Embedding Images in the LSB of an Image

We illustrate here how secret images can be embedded in the least significant bits of a seemingly harmless image.

In the following Matlab program that was adopted from Moler,[5] **A** is a 64×64 matrix containing double precision real numbers for an image. The numbers are all scaled to lie between 0 and 1. We can use the Matlab function `imagesc` to display it as a 32-bit gray scale image.

It turns out that hidden in the less significant bits of the original data are numerous other images. They can be extracted in Matlab in the following way. Consider a particular element, say a , of matrix **A**. Since a lies between 0 and 1, it has a sign of +1 and an exponent $E = 0$, and can therefore be represented as

$$a = d_0 + \frac{d_1}{2} + \frac{d_2}{2^2} + \cdots + \frac{d_{52}}{2^{52}}.$$

Note that for arbitrary integers n and m , where $n \geq m$ and $m \geq 0$,

$$\begin{aligned}\lfloor 2^n a \rfloor &= \lfloor d_0 2^n + d_1 2^{n-1} + \cdots + d_n 2^0 + d_{n+1} 2^{-1} + \cdots d_{52} 2^{n-52} \rfloor \\ &= d_0 2^n + d_1 2^{n-1} + \cdots + d_n 2^0,\end{aligned}$$

and so

$$\begin{aligned}u &= \text{mod} (\lfloor 2^n a \rfloor, 2^{n-m+1}) \\ &= 2^{n-m+1} (d_m 2^{-1} + \cdots d_n 2^{-n+m+1}) \\ &= d_m 2^{n-m} + \cdots d_n 2^0.\end{aligned}$$

The maximum value of u is $2^{n-m+1} - 1$, and its minimum is 0. Thus $v = u + 1$ has maximum value 2^{n-m+1} and minimum value 1. The resulting matrix \mathbf{U} can therefore be displayed as a $(n - m + 1)$ -bit gray scale image.

Here we can obtain a total of 16 images by selecting m from the array

$$p = [1 \ 6 \ 11 \ 16 \ 17 \ 18 \ 19 \ 20 \ 25 \ 30 \ 35 \ 36 \ 40 \ 44 \ 48 \ 52],$$

and n from the array

$$q = [5 \ 10 \ 15 \ 16 \ 17 \ 18 \ 19 \ 24 \ 29 \ 34 \ 35 \ 39 \ 43 \ 47 \ 51 \ 52].$$

The complete Matlab program is show below.

```
% steganall2.m
% modified from steganall.m written by C. Moler.
% Steganography: Here are all images hidden in the
% default CDATA for the IMAGE command.

p = [1  6 11 16 17 18 19 20 25 30 35 36 40 44 48 52];
q = [5 10 15 16 17 18 19 24 29 34 35 39 43 47 51 52];

clf      % clear current figure.
image    % returns a handle to an IMAGE object.
imageh = get(gca,'child');
% A is a 64x64 matrix with values in [0 1].
A = get(imageh,'cdata')/32;
clf
shg
```

```
colormap(gray(32));  
bigscreen;  
  
% Image obtained from the entire original data:  
subplot(1,2,1);  
imagesc(A);  
title('Image from the full data');  
axis image; axis ij; axis off;  
  
% This shows the image obtained using only  
% the top most significant bit of the data:  
subplot(1,2,2);  
m = 1; n = 5;  
U = mod(floor(2^n*A),2^(n-m+1));  
V = U+1;  
imagesc(V)    %Scale data and display as image  
title([int2str(m) ':' int2str(n)])  
axis image; axis ij; axis off;
```

pause;

% The following shows that 15 other images are
% secretly embedded in the original one.

```
for k = 1:16
    subplot(4,4,k)
    m = p(k);    n = q(k);
    U = mod(floor(2^n*A),2^(n-m+1));
    V = U+1;
    imagesc(V)    %Scale data and display as image
    title([int2str(m) ':' int2str(n)])
    axis image; axis ij; axis off;
end
```


References

- [1] Neil F. Johnson, Zoran Duric, and Sushil G. Jajodia, *Information Hiding: Steganography and Watermarking - Attacks and Counter-measures (Advances in Information Security, Volume 1)*, Kluwer Academic Publishers, 2001.
- [2] Eric Cole, *Hiding in Plain Sight: Steganography and the Art of Covert Communication*, Wiley, 2003.
- [3] David Kahn, *The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*, An authoritative history of cryptography in general.
- [4] Bruce Norman, *Secret Warfare: The Battle of Codes and Ciphers*, Borgo Press, 1990.
- [5] The Matlab program was adopted from C. Moler, *Numerical Computing with Matlab* at the [*Mathworks site*](#).

