



Global Information Assurance Certification Paper

Copyright SANS Institute
Author Retains Full Rights

This paper is taken from the GIAC directory of certified professionals. Reposting is not permitted without express written permission.

Interested in learning more?

Check out the list of upcoming events offering
"Security Essentials: Network, Endpoint, and Cloud (Security 401)"
at <http://www.giac.org/registration/gsec>

Table of Contents	1
Trang_Nguyen_GSEC.doc.....	2

© SANS Institute 2005, Author retains full rights.

The Secret Hidden in Steganography, Cryptography, HTML and the Web

Trang D. Nguyen
January 10, 2005
GSEC Practical Assignment
Version 1.4c

Abstract

Steganography and cryptography are technologies that can be used for secret and secured communications. Hypertext markup language (HTML) and the World Wide Web are technologies that can be used for open communications. When combined, steganography, cryptography, HTML and the Web can be effective tools for covert communications among terrorists, drug traffickers and their accomplices.

The focus in this paper is to demonstrate the core proof of concept in the combined use of steganography, cryptography, HTML and the Web for secret communication. This is not a step-for-step on how to use the software nor is it an attempt to show all possible combinations that make use of the same technology to produce the results. This is about how steganography, cryptography, HTML and the Web can be used for secret communication and that security technology need to be develop to focus on securing the Web from misguided use.

Let's Talk the Talk

One to one communication exists in many forms. We talk to one another face to face. We use wired telephones to communicate over land lines and cellular telephones to talk over the air wave for mobility. We digitally communicate using instant messaging and e-mail. E-mail has become the most popular communication medium for non-face-to-face communication. Each day, computer users send each other billions of e-mail messages. The problem with using these common means of communication for covert communication is security. They do not provide the confidentiality of secrecy and they do not meet the anonymity goal of covert operation.

Face to face discussion in a room can be intercepted by using covert listening devices and audio surveillance spy equipment. Native language conversation is easy to understand and foreign language communication can be translated.

Wired telephone and cellular telephone communications are susceptible to interception by telephone tapping devices¹. Wired telephone conversation can be recorded using wire tapping devices and transmitters. Analog cell phone scanners can intercept cellular telephone communication². With the appropriate hardware and software, digital cellular communication can also be intercepted and decrypted.

*"One (method) is using a personal computer, a cell phone, and the software. The software catches the "waves" over the phone. As you know, today's cell phones transmissions are encrypted. The software decrypts that because the code is a joke. The only problem is to get the software for "normal" users. The trouble for finding that is very high."*¹

Obtaining software to decrypt digital cellular communication is not a problem for the intelligent communities. They have the resources and access to high-tech devices and software that would easily decrypt cellular communication.

Although encrypted voice communication equipment such as the STU-III exists for secured communication, they are not available commercially for purchase by terrorists, drug traffickers and their underworld accomplices. Secure Telephone Units, Third Generation, (STU-III) are encrypted telephones in use by the intelligent communities and the Department of Defense.

Instant messages and **standard** e-mail are in plain text. The message can be captured during transmission over a network. They can be intercepted before they are delivered to the recipient as they traverse the network.

To secure the privacy of e-mail so that only the intended recipient can read the message, e-mail encryption protocols have been developed to ensure privacy and provide authentication of the sender for non-repudiation. The two current adopted protocols are S/MIME (RSA Data Security, Inc.) and PGP (PGP, Inc.). They are different in many ways and are not designed to be interoperable. In addition to S/MIME and PGP there are many proprietary encrypted e-mail solutions available from vendors.

Although secured email provides confidentiality through encryption, the problem of anonymity for covert communication still exists. The sender and the recipients are easily identified. Over a period of monitoring communication activities, the organization structure can be mapped out. Monitoring communication activities can start on either the sender or the receiver end of the communication.

Terrorists use the Web in their propaganda war. Their websites are published and information is disseminated in anonymity, could they be using the Web for covert communication? Using the vast cyberspace of the World Wide Web as channels for dissemination of information would make it more difficult to discover and intercept covert communication.

Walk the Walk - Terminology

In depth analysis and explanation of steganography, cryptographic, hypertext markup language and the World Wide Web are beyond the scope of this paper. Introductory to each technology is necessary to make it easier for readers to understand the content and its overall concept.

Steganography

Steganography is a Greek word which means “covered writing”. The goal of steganography is to hide the existence of a secret message in a carrier.³ The “writing” is the hidden message. The hidden message can be any plain or encrypted digital data stream. The “covered” is the courier or carrier of the message. “Stego-carrier” is a stegos file that has a hidden message. It is possible to embed the message in image, audio, video, text, document file such as HTML and PDF (portable document file), and even executable files using steganography software.

Early generation of steganography software uses a password as protection scheme to hide a plain text message in the carrier. Today, steganography is a form of encryption that goes beyond the basic encryption of hiding plain text file in a carrier. Modern steganography software uses the password or pass-phrase to encrypt the hidden message through symmetric cryptography. The encryption and password protection provided by the steganography software are layers of security as additions to the password protection and encryption that may have already been performed on the message.

Cryptography

There are two types of cryptosystems: symmetric and asymmetric.

Symmetric cryptography is also known as secret key cryptography. It uses the same key to encrypt and decrypt a message. To ensure confidentiality, the one key must be kept secret between the sending and receiving end-entities. Symmetric cryptography is generally faster than asymmetric cryptography and is used for encrypting data prior to sending through the unsecured networks. Some well known symmetric cryptographic algorithms are DES (Data Encryption Standard), 3DES (Triple-DES – based on using DES three rounds encryption), IDEA (International Data Encryption Algorithm), AES (Advanced Encryption Standard), RC2, RC4, RC5, and RC6, Blowfish, CAST5 and CAST6.

Asymmetric cryptography is also known as public key cryptography. It uses two keys known as the private key and the public key pair. The key pair is generated together and works together in the encryption and decryption scheme. One key is used to encrypt the message. The other key is used to decrypt the message. The one key cannot encrypt and then decrypt the encrypted message. The private key is tied to its owner, must be kept private, secured and accessible only by its owner. The public key is available to the public at large. The most commonly used public key algorithm is RSA named after its inventors Ron Rivest, Adi Shamir and Leonard Adelman. Other well known asymmetric cryptographic algorithms are Diffie-Hellman (Whitfield Diffie and Martin Hellman), ElGamal (Taher ElGamal), DSS (Digital Standard Signature), and ECC (Elliptic curve cryptosystems).

Hypertext Markup Language (HTML)

HTML is a markup language. HTML's primary use is to create Web pages. Tim Berners-Lee, a programmer at the European Center for Particle Physics (CERN), is credited with the creation of HTML and its framework.⁴ The World Wide Web Consortium (W3C) oversees the standardization and development of HTML. The language is based on the Standard Generalized Markup Language (SGML). Berners-Lee developed Hypertext markup language for the scientific community to publish and share text document, a language that is platform, network and terminal independent.⁴

HTML is an interpreted language of the Web that is based on tags, sometimes referred to as elements. Most tags have an opening and a closing tag pair. The opening tag acts as an "on" switch and the closing tag acts as an "off" switch. Standalone or singleton tags do not have a closing tag.

Tags have names. Opening tag is enclosed within a less-than and greater-than symbol. Closing tag adds a forward slash before the tag name, before enclosing with the less-than and greater-than symbol. In the most basic form [table-1], HTML is a string of character without embedded spaces.

Basic HTML Tag Format

	Opening Tag	Closing Tag
Tag pair	<HTML>	</HTML>
Singleton		

Table-1

The World Wide Web (WWW)

The Web began as a text based file environment viewable only by text browsers such as Lynx. In 1993, Marc Andreessen and Eric Bina developed a graphical user interface (GUI) at the National Center for Supercomputing Application (NCSA) at the University of Illinois in Urgana-Champaign.⁵ Mosaic is the foundation for the development of GUI web browsers such as Netscape, Internet Explorer and Opera.

Today as part of the Internet, the World Wide Web becomes a global communication medium for businesses, individuals and possibly terrorists, drug traffickers and their underworld accomplices.

"In 1988, around half of the thirty organizations designated as "Foreign Terrorist Organizations" under the U.S. Antiterrorism and Effective Death Penalty Act of 1996 maintained websites; by 2000, virtually all terrorist groups had established their presence on the Internet."⁶

"Recently, there have been rumors about terrorist using steganography to hide

their communication and secret plans.”⁷

Neils Provos and Peter Honey Research created a detection framework to capture and “*analyzed two million images downloaded from eBay auctions and one million images downloaded from a USENET archive but have not been able to find a single hidden image.*”^{8,9}

Does this mean that steganography is not being used by terrorists, drug traffickers and their underworld accomplices? We don’t know for sure because the terrorists, drug traffickers and their underworld accomplices didn’t make public their methods of covert communication. What we know for sure is that the scope of the research was narrow and the capability of the steganalysis software, “Stegdetect” was limited.

eBay requires valid verifiable identity of the sellers and buyers such as credit card and/or bank account information. Surely, the terrorists would not want to expose their anonymity by communicating with accomplices on eBay.

“Stegdetect” detection framework was build around detecting stego-carrier created by the JSteg (JSteg-Shell), JPHide, Invisible secrets, OutGuess version 01.3b, F5 (header analysis), AppendX and Camouflage.¹⁰ It was not designed to detect steganography other than those and it only works with Joint Photographic Expert Group image (JPEG) files. Stego-carrier can be created using image, audio, video, text, document file such as HTML and PDF (portable document file), and even executable files as carrier.

The Proof before the Concept

To test the effectiveness of “Stegdetect” and to determine if there is steganography software that do not use the popular hiding technique of JSteg, JPHide, OutGuess, F5, AppendX and Camouflage, I created a simple message file with “notepad” name “Secret Plan.txt”. The message file contains the following text:

*This message is a plain text message detailing the covert operation of a terrorist group or drug trafficking group. The content is sensitive, so it will be encrypted and delivered in a covert manner such as using steganography, cryptography, HTML and the web. After you (my accomplice) extract this hidden plan, message, the password, the encryption key and all the secrets from the stego-carrier file, replace the stego-carrier file from the drop-site with the sanitized carrier file. The password to drop-site where you downloaded this stego-carrier is “dup*load1t”. We registered our next drop site at geocities as w3cz34.*

I used “Steganography” version 1.61 from Secure Kit Incorporated (www.securekit.com) to create the stego-carrier file. “Steganography” encrypts the hidden content using 256bit symmetric cryptography and is easy to use.

Figure-1 shows the ease of using “Steganography”. In a few easy steps, computer

users can hide and unhide a secret message in a carrier.

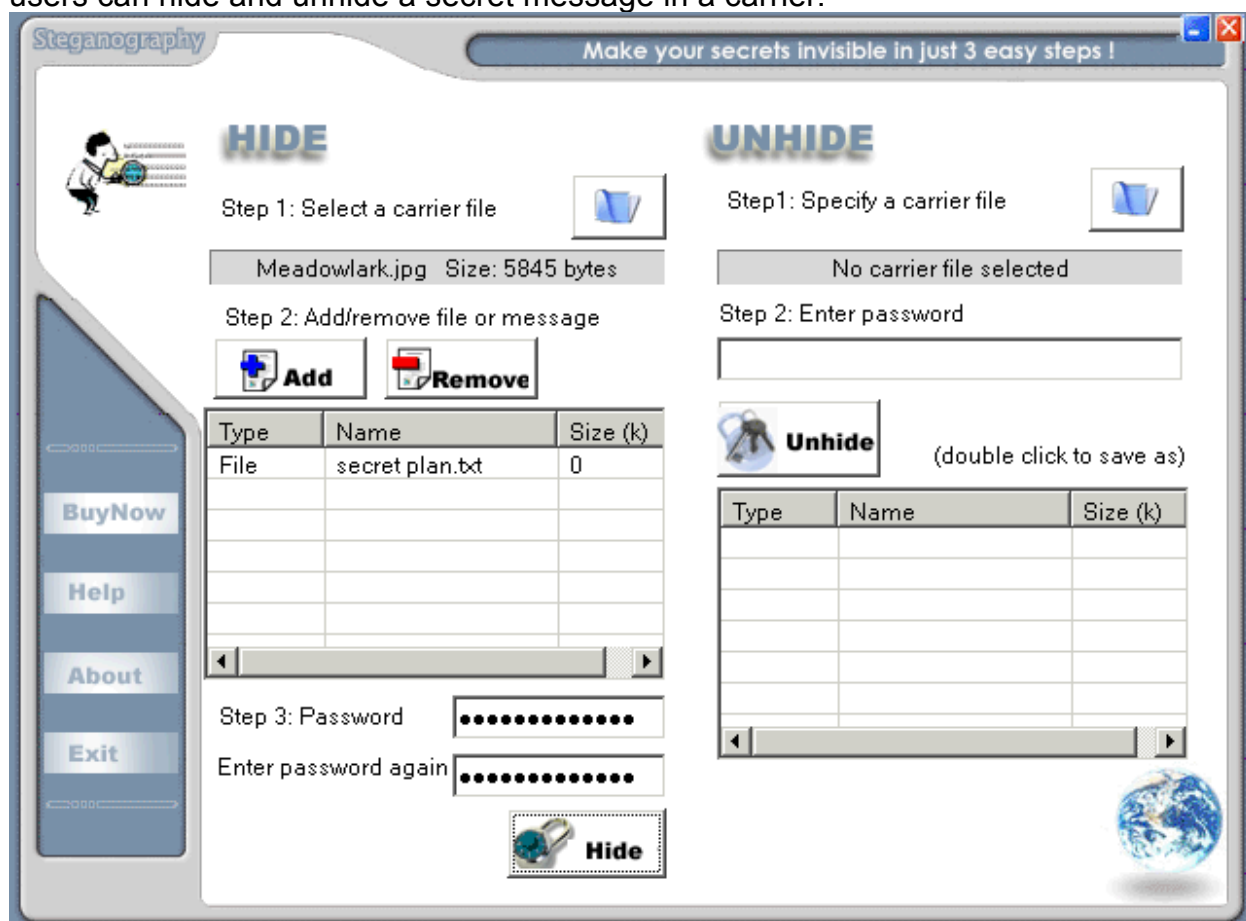


Figure-1

- Step 1: Select the carrier file which is the image file "Meadowlark.jpg"
- Step 2: Select the message file which is the text file "Secret Plan.txt"
- Step 3: Enter a password to secure and encrypt the stego-carrier
- Step 4: Click [Hide] and name the file in the "Save As" dialog.

I saved the file as "Stego_x.jpg" and used "Stegdetect" to detect steganographic content in "Stego_x.jpg". The result was negative. [figure-2]



Figure-2

The negative result from "Stegdetect" shows that there is one or more steganography method that is beyond the popular steganalysis capability of "Stegdetect". It also means that there are steganographic software can be used by terrorists and drug

traffickers to evade steganalysis. Steganalysis is the investigation of steganography or hidden information. When the detection confirms the existence of the hidden message, attempts can be made to extract the hidden content in order to compromise its confidentiality and integrity. There are additional steps can be taken to reduce the chance of being discovered by steganalysis.

Proof of Concept

Study has shown that the larger the hidden message the easier it is to detect through steganalysis. To reduce the likelihood of detection the hidden message file should be as small as possible. The message file “Secret Plan.txt” used in the “Proof before the Concept” was small. It was a simple text file that was created using “notepad”.

Since the message file can be any digital data stream, such as a Microsoft Word file. The “Secret Plan.txt” file was converted to a Microsoft Word document file and named “Secret Plan.doc”. The origin text file was 1K in size. The Word document file was 24K in size.

Let the process begins... Squeeze...

PKWare, Zip Commander, Winrar and Winzip are some popular file compression software that can be used to compress or reduce the size of the file before using steganography. Winzip also have security featured such as password protection and encryption.

Figure-3 shows that Winzip version 9.0 (www.winzip.com) achieved a compression ratio of 84% on the “Secret Plan.doc”.

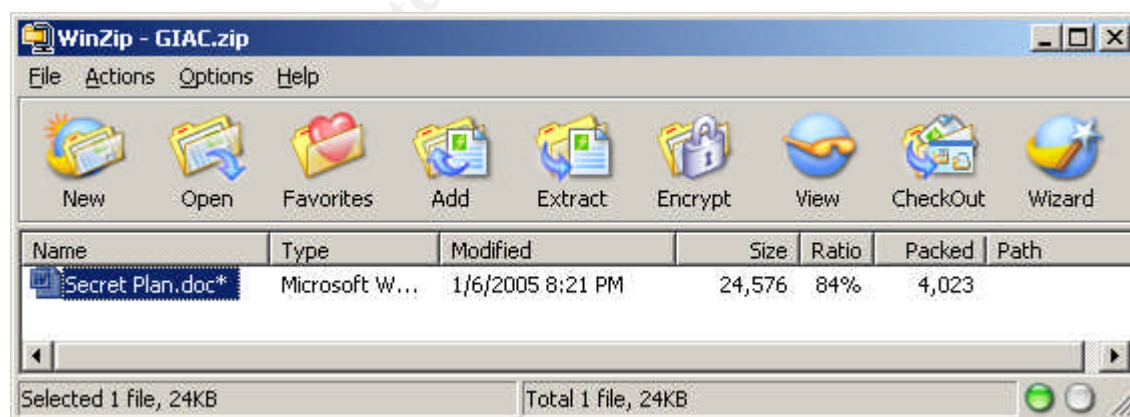


Figure-3

Applying Symmetric Cryptography...

Winzip encrypted the “Secret Plan.doc” using 256-bit AES (Advanced Encryption Standard) symmetric cryptography in creating the “GIAC.zip” file.

Asymmetric for Non-repudiation...

To secure the covert message and ensure non-repudiation, a characteristic of public key encryption, PGPmail [figure-4] from PGP Corporation (<http://www.pgp.com/>) can be used to encrypt and optionally digitally signed the message. PGP and S/MIME are two adopted standards for securing e-mail using asymmetric cryptography.



Figure-4

The process of encrypting a file is simple using PGPmail graphical user interface. Click the [Encrypt] button (second button from the left) to start the “Select File(s) to Encrypt” dialog. After a file or files have been selected for encryption, PGPmail displays the “Key Selection Dialog” which is a list of recipient’s public keys [figure-5].

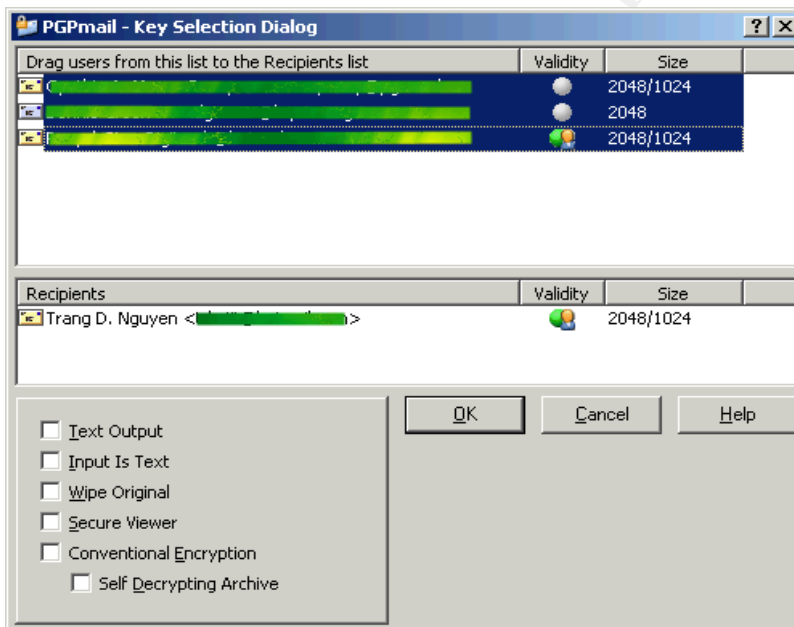


Figure-5

The next step is to drag users from the list (above) to the recipients list (below) and then Click [OK]. PGPmail encrypts the files using the recipient’s public key and optionally signed the message with the sender’s private key. In this case, the recipient’s (Trang D. Nguyen – Figure-5) encryption algorithm was a Diffie-Hellman.

The file “GIAC.zip” [figure-6] created previously by Winzip was encrypted by PGPmail. The output file “GIAC.zip.pgp” [figure-6] has the concatenated file extension of .pgp to indicate that it’s a PGP encrypted file.

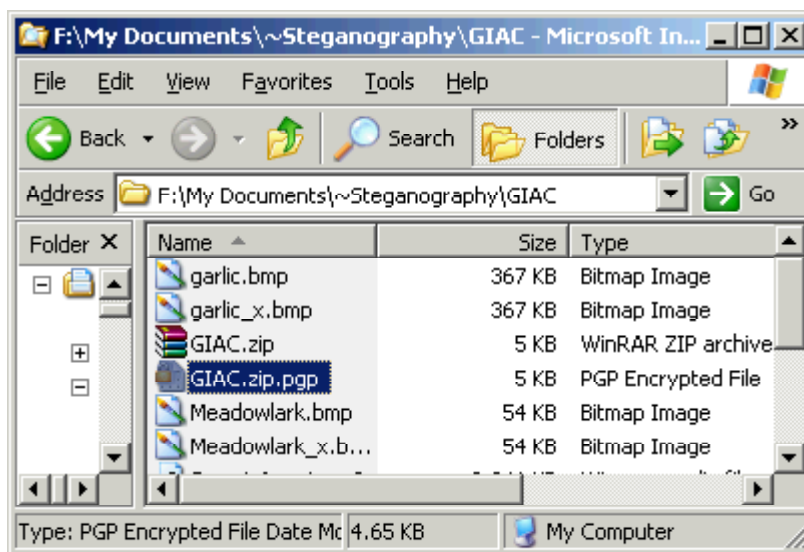


Figure-6

Although the covert message is ready for posting on the Web using invisible HTML links, steganography will be applied so that visible and invisible HTML links can be demonstrated.

Steganography

Software specializing in steganography still exists today; however, steganography is becoming an added feature of security software such as “SecurEngine Professional” by Adrien Pinet (<http://secureengine.isecurelabs.com>) and “Computer-Security” from Adolix (<http://www.adolix.com/computer-security/>).

In the “Proof before the Concept”, I used “Steganography” version 1.61 from Secure Kit, Inc. to create the stego-carrier. In this process, I used “Computer-Security” from Adolix to create the stego-carrier.

“Computer-Security” supports an extensive list encryption such as 3Way, Blowfish, Gost, Q128, Safer-SK128, SCOP, SHARK, Square, TEA, TEA Extended, Twofish, Cast 128, DES Triple 24byte, Diamond II, FROG, NewDES, RC2, Rijndael, Sapphire II, and SkipJack.

Figure-7 shows Adolix’s “Computer-Security” steganography hide process screen. The procedure is similar to other steganographic software. The basic steps are:

1. Select file or files to hide
2. Type a password or password phase to secure the hidden file
3. Choose an encryption algorithm
4. Select a carrier file
5. Name and save the stego-carrier file

The World Wide Web

The Web is dynamic in nature. It is impossible to track changes on the ever changing Web. The dynamic of the Web and the inheritance of unsecured web servers in the demilitarized zone (DMZ) can be very effective communication channels for covert operations. The cyber-war has been focused mainly on defending attacks on computer networks through the Internet. We cannot forget the potential covert use of the Web in the war on terrorism and the war on drug.

Corporate public web servers in the quasi-public screened network are not as secured as the corporate internal network servers. The hypertext transfer protocol (HTTP) TCP/IP port 80 is opened to allow HTTP traffic. Microsoft Internet Information Services (IIS) requires constant security vulnerabilities patches. A compromised web server can become a drop site for terrorists and drug traffickers. Network security on a compromised system thus becomes an agent and support of terrorists and drug traffickers.

Terrorists, drug traffickers and accomplices do not need to compromise corporate web servers in the DMZ to achieve their missions. There are many free web hosting sites on the Internet that can be used. A simple Internet search using the keyword "free web hosting" yields several pages of free web hosting service providers and software for publishing web pages.

Free web hosting sites such as "geocities.yahoo.com", "www.tripod.lycos.com", "angelfire.lycos.com" and www.blackplanet.com do not require identification. Anyone can use false identity to register for a free website. The process is simple and relatively easy. Free online websites are excellent for covert operation; they can be registered and abandoned after use. This makes it very difficult for intelligent communities to monitor covert communications.

Hypertext Markup Language and the Web

HTML is the interpreted language of the Web. It is based on tags, easy to learn and simple to code. A few simple HTML statements to create a hypertext link are all that is needed to post the stego-carrier online for download by accomplice. The link can be posted visible or invisible on the web page.

Visible Posting

The image on the web page image in figure-9 is a stego-carrier file. The image is clearly visible however one must know that it is a stego-carrier and the password or pass phrase before the hidden content can be extracted.

To download the stego-carrier image, the recipient moves the mouse over the image, clicks the right mouse button and "Save Picture As..." a file on his computer. Once it is

downloaded the hidden content can be extracted using the software that created the file in reverse order. Where is the password? It's right on the page. The password is hidden but in plain sight.

HTML source code:

```
<html>
<body>
<center><b>Stego-carrier</b></center>
<image src="Meadowlark_x.bmp">
</body>
</html>
```

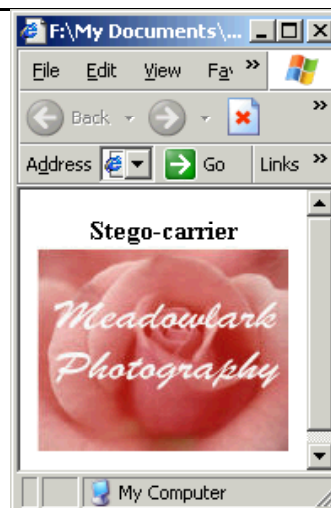
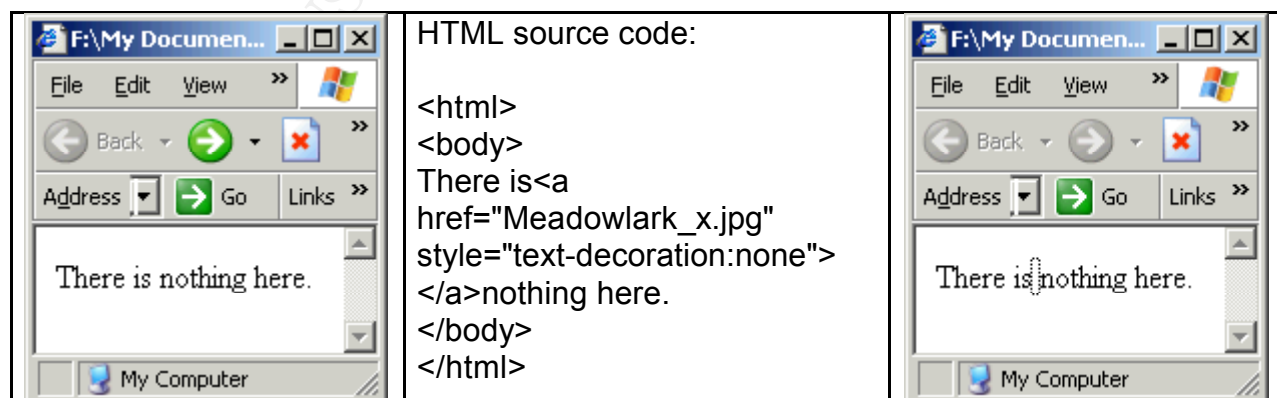


Figure-9

Invisible Posting

Steganography is covered or hidden writing. Invisibility is an attribute of steganography. It can be achieved in HTML by posting the stego-carrier as a link around a blank space within the text.

In figure-10 example, the innocent text on the web page contains a hypertext link to the stego-carrier file. The image on the left shows the normal display. The image on the right shows the existence of a hypertext link on the blank space between the word "is" and the word "nothing". The recipient can download the stego-carrier to his computer by placing the mouse cursor over the space until the hypertext link cursor the finger is visible. The recipient clicks the right mouse button and "Save Target As..." to download the stego-carrier file.



Invisible HTML link on a site is difficult to locate. First you need to know the correct uniform resource locator (URL) address page which contains the invisible link. Then by moving the mouse over the entire web page with time and some luck, the link “finger” mouse pointer will show up. That’s if the programmer had not changed the link “finger” mouse pointer to an arrow mouse pointer.

It can be challenging looking for the invisible link. A quick change in the code, it can be almost impossible to locate the stego-carrier on the page. Look very carefully at the graphic in figure-11; you will notice there is a very small (red) dot between the letter “e” and the period. I chose an image that would make the red dot so that it would show up on the page. If I had made that a white dot it would be impossible to detect its presence visually. The graphic dot is only 1 pixel in size. As it exists, it would be impossible to position the mouse cursor over the dot to extract the file. To perform the “Save Picture As...”, the recipient must first increase the text size the browser display to make the large enough to position the mouse pointer over it.

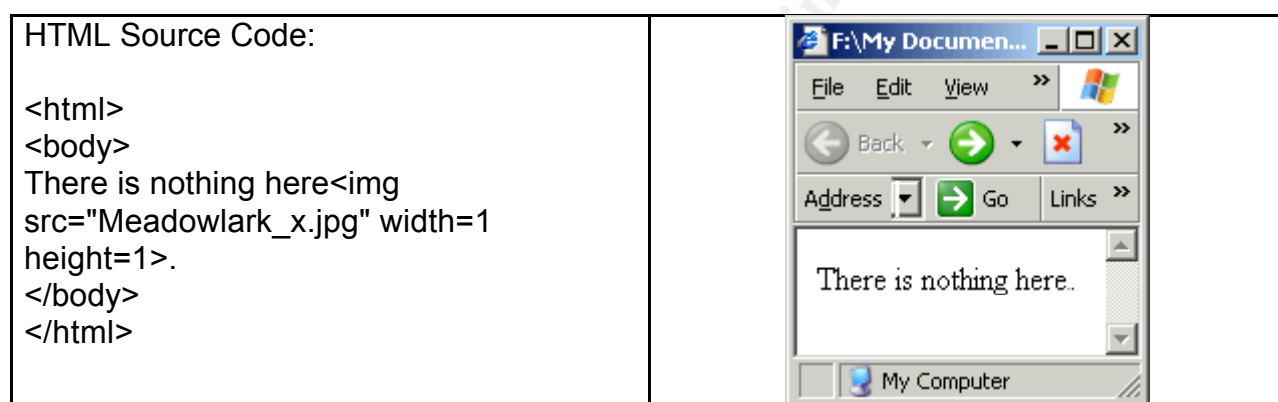


Figure-11

Steganographic HTML

Text file, PDF and HTML files can also be used as stego-carrier. HTML is a text file and can be viewed by any text editor. The encrypted hidden message is converted to white spaces and appended to the end of the HTML.

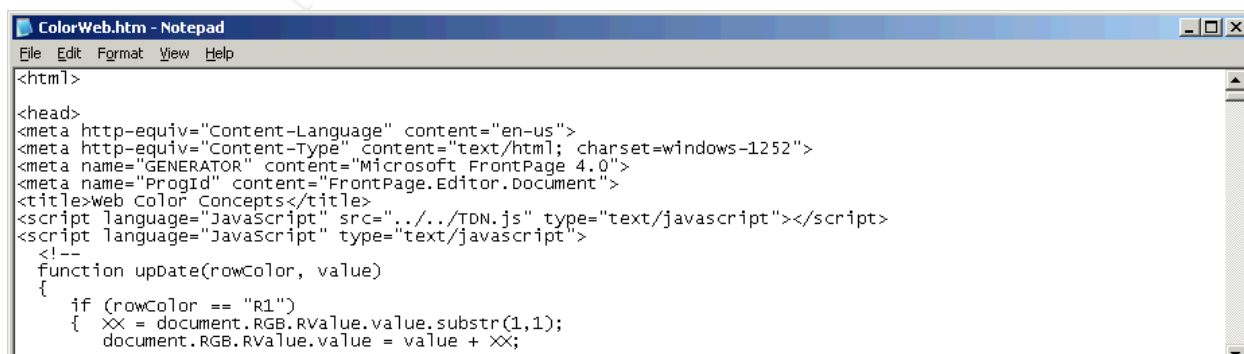
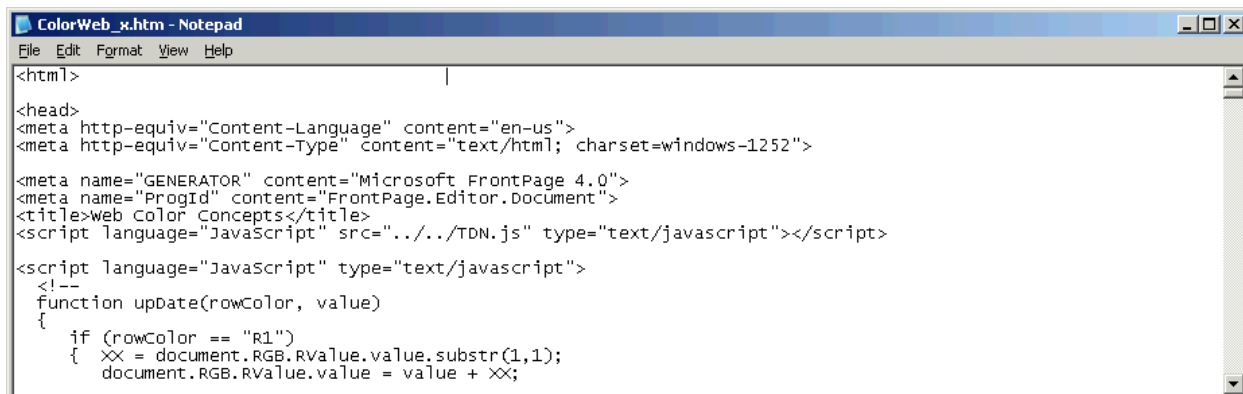


Figure-12a



```
<html>

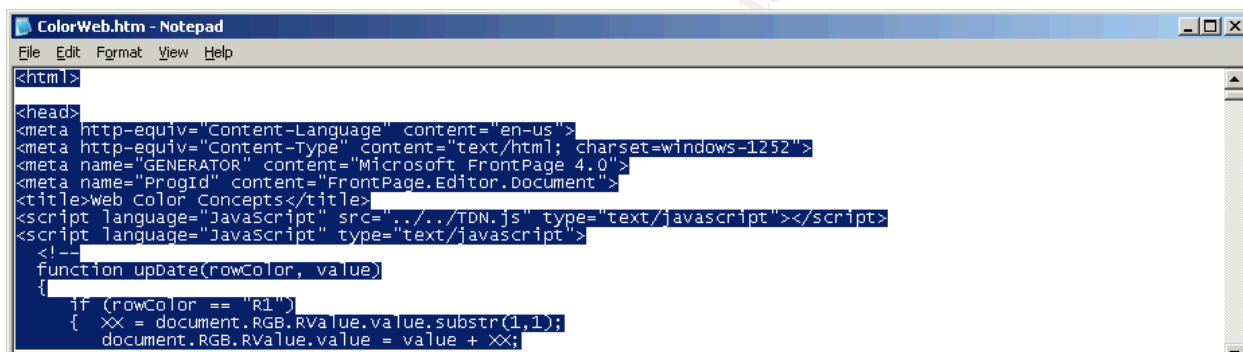
<head>
<meta http-equiv="Content-Language" content="en-us">
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">

<meta name="GENERATOR" content="Microsoft FrontPage 4.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<title>web Color Concepts</title>
<script language="JavaScript" src="../../TDN.js" type="text/javascript"></script>

<script language="JavaScript" type="text/javascript">
<!--
function update(rowColor, value)
{
    if (rowColor == "R1")
    { XX = document.RGB.RValue.value.substr(1,1);
      document.RGB.RValue.value = value + XX;
    }
}
```

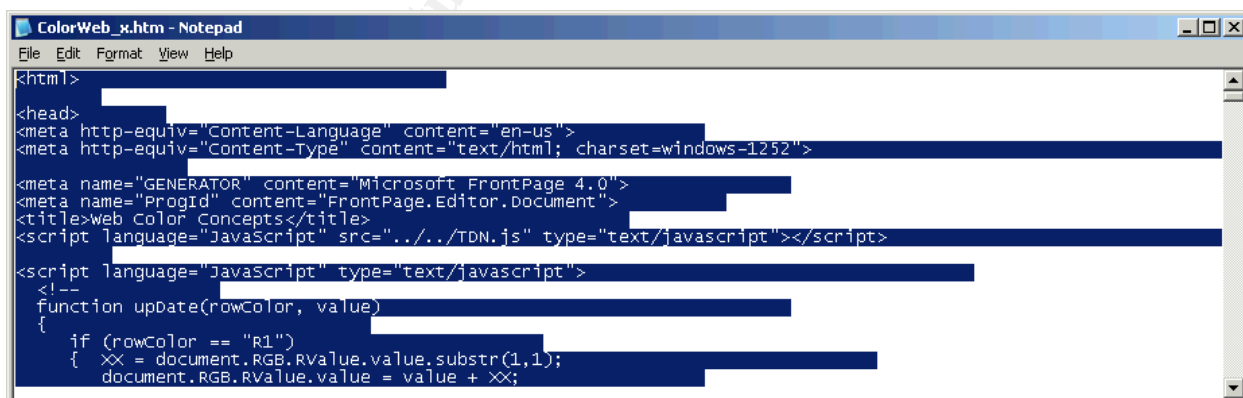
Figure-12b

View with the naked eye, the text in figure-12a and figure-12b windows look at exactly alike. However if the text are selected as in figure-13a and figure-13b, white spaces appended at the end of each HTML line of code are clearly visible. The white spaces are characters the encrypted hidden file.



```
<html>
<head>
<meta http-equiv="Content-Language" content="en-us">
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">
<meta name="GENERATOR" content="Microsoft FrontPage 4.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<title>web Color Concepts</title>
<script language="JavaScript" src="../../TDN.js" type="text/javascript"></script>
<script language="JavaScript" type="text/javascript">
<!--
function update(rowColor, value)
{
    if (rowColor == "R1")
    { XX = document.RGB.RValue.value.substr(1,1);
      document.RGB.RValue.value = value + XX;
    }
}
```

Figure-13a



```
<html>
<head>
<meta http-equiv="Content-Language" content="en-us">
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">
<meta name="GENERATOR" content="Microsoft FrontPage 4.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<title>web Color Concepts</title>
<script language="JavaScript" src="../../TDN.js" type="text/javascript"></script>
<script language="JavaScript" type="text/javascript">
<!--
function update(rowColor, value)
{
    if (rowColor == "R1")
    { XX = document.RGB.RValue.value.substr(1,1);
      document.RGB.RValue.value = value + XX;
    }
}
```

Solutions, The Big Question Mark?

Are there any solutions? The solution to the problem depends on the security zone. Defending cyberspace is no easy task. Cyberspace must be defined as battle zones such as corporate war zone and the web or free web hosting war zone.

Corporate web security administrators and web developers must work together to defend the public web servers that they are responsible. The servers must be defended against intrusion and secured against unauthorized access. Security patches must be applied in a timely manner. Security best practices need to apply to prevent web servers from becoming drop points for intruders.

The integrity of data files must be carefully monitored to ensure that they have not been modified at run time. Change auditing solution tools such as Tripwire for Servers, a commercial product from Tripwire.com can be implemented to detect and alert administrators when file or directory changes have occurred. Change management methodology can be implemented to compare the original files stored as read only files with copy of the files published on the public web directory at time intervals for changes. Changes from external intrusion, malicious tampering, or accident should be monitored.

The web or free web hosting war zone must be defended by the industry at large. To prevent the Web from being used maliciously, concerted effort by the free web hosting and ISP sponsored web hosting communities are needed. The free web hosting sites need to monitor their servers for malicious use. Effective steganalysis software needs to be developed to detect stego-carriers. Browsers must have capability to detect and remove invisible links source code.

The industry of “free web hosting sites” need to come to term that their services may be used by illicitly by terrorists, drug traffickers and the underworld.

Conclusion:

Steganography, cryptography, HTML and the Web can be an effective tools for covert communication among terrorists. The Web is a big frontier of the cyber war and is a major challenge for the intelligent communities in the war on terrorism and the war on drug. Current steganalysis technology is inadequate to detect hidden payload in stega-carrier. Corporate web administrator and the industry at large especially web hosting sites and technology developers need to focus on developing tools to counter the potential misguided use of the Web. Using the World Wide Web for covert operation is not limited to terrorists; it extends to the drug traffickers and other social disorder disobedient groups.

References:

- [1] Phone Losers of America,
URL: http://www.phonelosers.org/article_recording_telephone_calls.html,
(11/26/04 17:00)
- [2] Cell phone scanners,
URL: <http://www.spyequipmentguide.com/cell-phone-scanners.html>,

(11/26/04 17:30)

- [3] Whitman, M.E., Mattord, H.J., Principles of Information Security, Boston: Thomson Course Technology, 2003, Appendix, pg 326
- [4] Holzschlag, M.E., Using HTML 4 – Special Edition, Principles of Information, Indianapolis: Que-Macmillan, 2000, pg 11
- [5] Home of Mosaic, URL: <http://archive.ncsa.uiuc.edu/SDG/Software/Mosaic/NCSAMosaicHome.html>, (12/20/04 19:00)
- [6] Gabriel Weimann, “How Modern Terrorism Uses The Internet”, Special Report 116, United States Institute of Peace
URL: <http://www.usip.org/pubs/specialreports/sr116.html>; (01/05/05 18:45)
- [7] Niels Provos, Defeating Steganalysis, URL: <http://niels.xtdnet.nl/stego/>, (01/05/05 17:45)
- [8] Niels Provos and Peter Honeyman, “Detecting Steganographic Content on The Internet”, ISOC NDSS'02, San Diego, CA, February 2002. [August 2001, CITI Techreport], URL: <http://niels.xtdnet.nl/papers/detecting.pdf> (01/05/05 20:00)
- [9] Niels Provos, Scanning USENET for Steganography, 2001,
URL: <http://niels.xtdnet.nl/stego/usetnet.php>, (01/05/05 20:00)
- [10] Niels Provos, Steganography Detection with StegDetect, 1999-2004,
URL: <http://www.outguess.org/detection.php>