Global Information Assurance Certification Paper

Steganography

The New Terrorist Tool?

Danley Harrison

Summary

Steganography is the art of hiding communications without the unintended party's knowledge. Steganography has again come into focus by the crypto analytical community and the media because of the ease that information can be hidden in today's many forms of digital communication. Steganography techniques can be applied in many ways including images, invisible inks, microdot printing, spread spectrum, burst communication, messages, viruses inside of a seemingly benign communication and now the Internet. [1]

Steganography is achieved by using stegno systems software. These software programs can hide information not only in images but can hide messages inside of audio, video and spam[2]. According to the media, it is rumored that this was a tool used to conceal information and to assist terrorist cells in their campaigns.

***New-York*. Copyright photo courtesy of Patrick Loo, University of Cambridge (pl201@cam.ac.uk)**

Digital Watermarking, another form of information hiding, is gaining interest with the current industry demand for providing a way to digitally fingerprint audio and video. This is being used to detect copyright violators and provide an aid to help in prosecuting those that violate copyright laws.

This paper will provide a brief history of Steganography, an explanation of a few of the terms used, review of the method of hiding information using text, audio, and images. It will give a brief discussion on the differences between digital watermarking and steganography, and the current trends in the industry with the increasing demands for digital watermarking and fingerprinting. A look at how to identify hidden messages using Steganalysis[3] is reviewed. It has become a relevant topic since the terrorist attacks on the United States. Finally, several Internet links to popular Steganographic systems software are given.

History of Steganography

An important sub-discipline of information hiding is steganography. While the art of cryptography is about protecting the content of messages, steganography is about concealing their very existence. Steganography was first used by the Greeks to send secret messages. One of the first documents mentioning Steganography is from **Histories of Herodotus**[4]. The Greek historian, Herodotus wrote how documents with strategic information were transferred during a battle. Text was written on tablets that were then covered with wax to hide the original message. The messenger could transport the undetectable information hidden on the tablets. Upon delivery the wax would be melted and the message would appear. Other stories documented by Herodotus indicated it was a common practice to shave the heads of slaves and tattoo messages. After the hair had grown back, the slave would be sent on their way with the message that was undetectable until their hair was shaved off. This worked as long as the non-intended recipient did not have the key; which was the clue to shave the heads of message carrier.

In fact the word Steganography literally means covered writing and was derived from Greek[5]. Since the security of steganograpic system rely on only the intended recipient knowing the code it is very difficult to detect the embedded message without the application of a correct encoding system.

> Steganography is the art and science of communicating in a way which hides the existence of the communication. In contrast to cryptography, where the "enemy" is allowed to detect, intercept and modify messages without being able to violate certain security premises guaranteed by a cryptosystem, the goal of steganography is to hide messages inside other "harmless" messages in a way that does not allow any "enemy" to even detect that there is a second secret message present. [6]

Other methods to circumvent detection of hidden information include invisible inks made from various substances like fruit juices and even urine which when heated would appear. The carrier who seems to possess blank paper when searched, but when their final destination was reached the message would be recovered by adding heat or immersing in a solution not unlike developing a photographic picture. Unsophisticated, but effective, these techniques allowed information to be transmitted in secrecy.

Steganography in text relies on an encoding system where it might use the last word in every sentence of a letter. These are the hardest to crack unless the secret key is known. This holds true to "Kerckhoff"'s Principle"[7] in cryptography, which states that the security of a cryptographic system should rely only on the key material. Johannes Trithemius[8] (1462-1516), a German monk wrote a volume called *Steganographia*, which was a body of research in Steganography and cryptography. Because he feared the reaction of the authorities during those times, his work was not officially published. Gaspari Schotti[9] (1608-1666) produced the earliest book on Steganography in 1665. It was called *Steganographica.*

Today, the ***Voynich Manuscript***[10] is one of the most well know Steganographic text that has defied decryption. This late 15th century manuscript is written in an unknown cipher code that has resisted all efforts at reading for the last 80 years. This manuscript is a medieval book that was written in an unidentified cipher. This manuscript is still being analyzed to unlock its mysteries by crypto-researchers Jim Reed from AT&T Bell Labs[11]. No one had been able to crack this encoding key so it remains locked. Again holding true to "Kerckhoff's Principle".

During the World War II, "microdot", another form of Stegano became very popular. This method was used to send information around with spies avoiding detection. Pages of information was reduced to "microdot" by photographic reduction and then placed on top of printed periods or commas in innocent cover material[12]. Hiding these dots as periods at the end of a sentences permitted data to be hidden directly in printed documents. Because the technology used to create these "microdots" was very expensive, it was not available to the average person. The government was not very concerned with private sector using such form of information hiding. This soon changed with the introduction of microelectronics and computers.

With the advent of spread spectrum, frequency hopping and burst communications, the National Security Agency and other security agencies were getting concerned with the entrance of the private sector into their exclusive domain of hidden communication.[13] TRANSEC or Transmission Security is the term used for the Government's monitoring this branch that includes Steganography. Given that nearly all communication channels like radio broadcast and landline communications carries some kind of noise, this noise could be replaced with secret signals that has been encoded into a form indistinguishable from the noise, and without knowledge of the cipher, the encoded signal could be transmitted concealed. Several attempts were made by Government Agencies to prevent patent applications for inventions that circumvented the analysis of these types of communication[14].

Our government did research on embedding hidden signals to aid in their communications in restricted areas:

> In 1973 TRW began designing a satellite system for use by the CIA in communicating with agents in "denied areas" Code-named Pyramider, the system employed frequency-hopping. This provided the agent with large "safe areas" in cities, where the signals could be hidden among random urban radio transmissions. The system was also capable of reducing aircraft interception in remote areas to a radius of twenty nautical miles.[15]

Information Hiding refers to both digital watermarking and Steganography. Digital audio, video, and pictures are increasingly embedded with distinguishing but imperceptible marks, which may contain a hidden copyright notice or serial number or even help to prevent unauthorized copying directly. Inserting a digital watermark into an image, software or music is a method that is currently being suggested and investigated as a way

to protect and stop the illegal copying of digital media[16]. To circumvent or remove digital watermarking would render the original information degraded. However, modification to the original ("cover") data medium containing steganographic information would destroy the hidden message or "embedded data". One distinction between Steganography and Watermarking is stego information is not obvious to the unintended addressee, while watermarks can possibly reveal themselves as an option.

Steganography tools hide large blocks of information; whereas watermarking tools embed less information. In watermarking the embedded data is distributed redundantly throughout the entire cover image[17]. Both of these methods insert embedded data and manipulate the images, but it also causes degradation and distortion to the original properties of the image.

There is a current industry demand for providing a way to digitally watermark and fingerprint audio, video and program files to detect copyright violators and provide an aid to help in prosecuting those that violate copyright laws. Record companies shut down Napster because they provide music that was copywrited but not digitally fingerprinted.

Before the tragic event of September 11, 2001 USA Today reported that terrorists were using the Internet to transmit hidden communications. During the recent U.S. Embassy bombing case in Africa, several documents came to light in the media that suggest Osama bin Laden [18]and his associates have been using steganography to hide terrorist target plans inside pornography and MP3 and other files that are freely distributed over the Internet. These claims proved to be false according to research done by Niels Provos[19] Using a web crawler that downloaded over two million images from EBay's auction site, not a single message was acquired. Even then there has been a major shift of focus toward identifying the use of this technology because of the ability to conceal strategic and damaging information on the Internet.

Since September 11, the emphasis on detecting hidden communications has again become a hot interest area to law enforcement and counter intelligence agencies. They are interested in understanding these technologies and their weaknesses, so as to detect and trace allegedly hidden messages in communications that various terrorist cells are concealing using various methods. Earlier last year the intelligence agencies investigated a web site thinking that it held stego information that the al-Qaeda terrorist cell could use to continue their terrorist attack on the United States. No evidence was found so far but acknowledgment was given in the public media that this kind of information hiding could be used. There are still frequent reports in the news media that the same tactics are being used.

Because of the ease by which any unsophisticated individual can now conceal information using stego-software, there is a general concern that since the World Wide Web is so vast, many messages could be hidden on the Internet making it virtually impossible to find. But finding the message does not always guarantee that the information could be recovered. There are hundreds of Steganography software programs that are on available the Internet as freeware or shareware. Anyone with access

to a computer can download stego-systems software and transport their hidden communications using the Internet.

Terminology and Application of Steganography

Before going further, some explanations of the terms used in the field of Steganography is necessary. The popular model of hiding data in other forms of data can be described as follows:

> The *embedded* data is the message that one wishes to send secretly. It is usually hidden in an innocent message referred to as a *cover-text*, or *cover-image* or *cover-audio* as appropriate, producing the *stego- text* or other *stego-object*. A *stego-key* is used to control the hiding process so as to restrict detection and/or recovery of the embedded data to parties who know it (or who know some derived key value).

When discussing stego the terms *cover* and *container* are used to describe the original data. This cover data can consist of audio, video, pictures such as BMP, GIF, JEPG and still images. This is the medium that is used to hold the embedded secret data. The embedded secret data is called the "*message*". The *container* can be any type of digitial information that's transmitted through digital or analog transmission system. Since most communication channels transmit signal that also include some form of noise. The noise can now be replaced with this hidden message that's placed in the container. The "container" information can be any of the different media and graphic files such as JEPG, BMP, GIF, MP3, Wav, or AVIs. The terminology to describe this sub-set of Information Hiding was defined at the **Information Hiding Workshop**[20] that took place in Cambridge, England in April 1996.

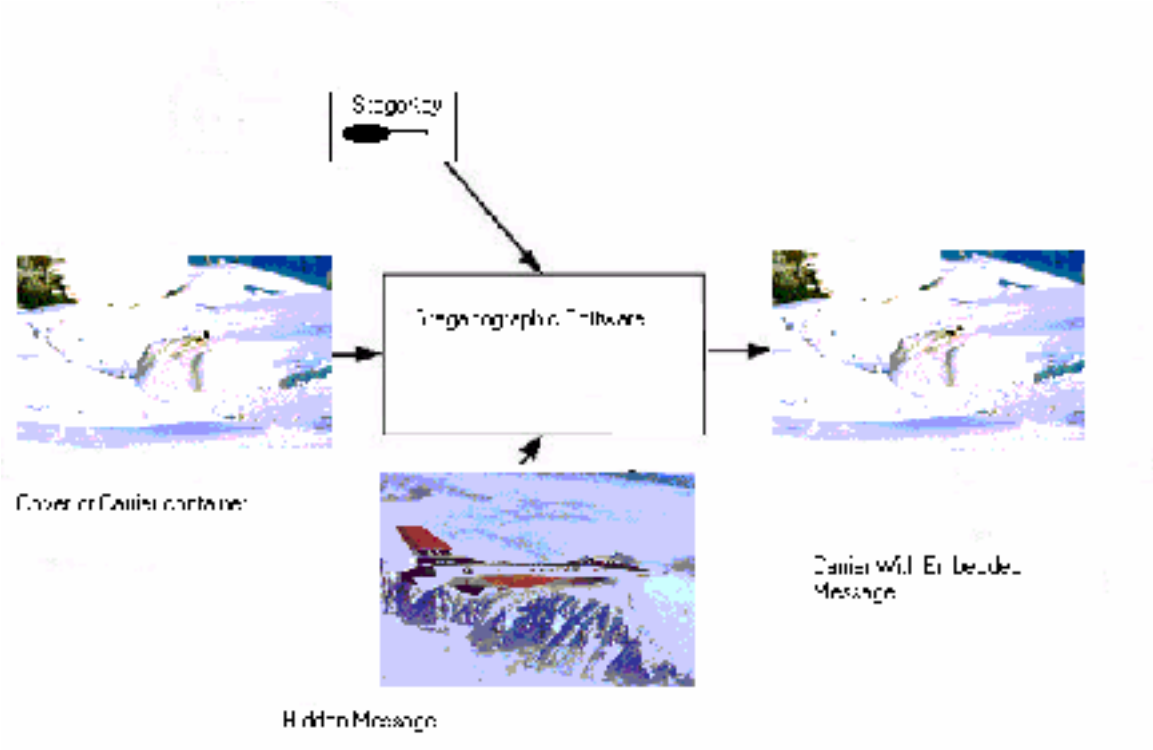

Figure 1: Steganography Using Software To Embed Message

Hiding messages inside of a container is done by usually using the least significant bit information from the original "container". By encoding information into the least significant byte, information can be embedded into the "container" with little degradation of that sound or image file. Because these least significant bits are used to represent other attribute in the original file, the container file must be large enough to hold its information and the embedded message. As shown in Figure 1, the cover image of the bunny rabbit is used to embed the photograph of the F14 jet fighter and the Stego software makes use of the least significant bit of the container file to now hold the *secret* message. The only way to view the embedded *secret* file is to know the stego-key.

Detecting Hidden Information

This section will give short description on research detecting a file that has hidden communications. This science is call steganalysis. Today, steganographic software enables communications to be hidden in sound, graphic, and even blank media. This allows communications to remain virtually impossible to detect and even decoded once it is detected. This basic design principle of steganographic systems, i.e. replacing high entropy noise with a high entropy secret transmission, is quite obvious. There are a number of simple software tools been published for e.g. hiding files in the least significant bits of digital images or for transforming PGP messages into files resembling pure random byte sequences.[21]

> Common communication systems have a huge number of characteristics and only a small fraction of what looks like noise can actually be replaced by the statistically very clean noise of a cryptographic cipher text. Noise in communication systems is often created by modulation, quantization and signal crossover and is heavily influenced by these mechanisms and in addition by all kinds of filters, echo cancelation units, data format converters, etc. Many steganographic systems have to work in noisy environments and consequently require synchronization and forward error correction mechanisms that also have to be undetectable as long as the secret key is unknown.[22]

Because the purpose of *steganography* is having to covert communication between two parties whose existence is unknown to a possible attacker, a successful attack consists of detecting the existence of this communication. Current concern in the private sector is using steganalysis to identify threats to companies internal data systems and keep proprietary information from being exchanged outside the company.

Most of the tools differ in their approaches for hiding information. Cryptologist must be perceptive of which tool is being used and which, if any, stegokey is used; detecting the hidden embedded information may become quite difficult. However, some of the tools construct stego-images with attributes that act as signatures for the steganography tool or method used.

The goal of steganography is to avoid drawing attention to the transmission of a hidden embedded message. If attention is raised, then this goal is defeated. Ascertaining and rendering useless, such covert messages is a new art form known as steganalysis[23].

> To begin evaluate images for additional, hidden information, the concept of defining a "normal" or average image was deemed desirable. Defining a normal image is somewhat difficult when considering the possibilities of digital photographs, paintings, drawings, and graphics. Only after evaluating many original images and stego-images as to color composition, luminance, and pixel relationship do anomalies point to characteristics that are not "normal" in other images. Several patterns became visible when evaluating many images used for applying steganography. The chosen message and known cover attacks were quite useful in detecting these patterns. In images that have color palettes or indexes, colors are typically ordered from the most used colors to the least used colors to reduce table lookup time. The changes between color values may change gradually but rarely, if ever, in one bit shifts. Gray-scale image color indexes do shift in 1-bit increments, but all the RGB values are the same. Applying a similar approach to monochromatic images other than gray-scale, normally two of the RGB values are the same with the third generally being a much stronger saturation of color. Some images such as hand drawings, fractals and clip art may shift greatly in the color values of adjacent pixels. However, having occurrences of single pixels outstanding may point to the existence of hidden information[24].

Recognizing whether or not a file contains hidden embedded data requires evaluation of the compromised file to the real thing—this is not always possible. The eye cannot always categorize photographic loss because most steganography programs use slight algorithmic change of the color palette table (that's why black and white photos are usually the best). And, even if you did suspect that a secret message possibly hidden inside one of your files, you need to know which software program was used, and then identify the password to open the file (if encrypted, which it probably is).

Two aspects of attacks on steganography are uncovering and damaging of the embedded message. Any image can be manipulated with the intent of destroying some hidden information whether an embedded message exists or not. Uncovering the existence of a hidden message will save time in the message removal phase by processing only those images that contain hidden information. Uncovering an embedded message also defeats the primary goal of steganography, that of cloaking the very existence of a hidden message.

Digital watermarking is used to copyright and identify digital media. This useful aspect of Steganography provides a mean to identify ownership. This method also discourages illegal copying of copyrighted material. Several companies embed licensing information into the logo file of their software programs that will not allow installation and configuration until a serial number is embedded into the program.

Until recently, information hiding techniques received much less attention from the research community and from industry than cryptography, but this is changing rapidly since the first academic conference on the subject was organized in 1996. The main motivating force is concern over copyright, as audio, video and other multimedia works become available in digital form[25], the effortlessness 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.

Since the emergence of MP3 digital audio files, the Music industry has been struggling for a way to identify and stop illicit duplication of digital media. MP3Stego[26] is a stenography tool for MP3 audio streams designed by Fabien A. P. Petitcolas at the University of Cambridge, UK. MP3Stego will conceal information in MP3 files during the compression process. The data is first compressed, encrypted and then secreted in the MP3 bit stream. Although MP3Stego has been written with steganographic applications in mind it might be used as a copyright marking system for MP3 files As we can see, this technology has the potential to extend beyond the world of espionage and military applications to the consumer market.

There are other applications in the real world that has increase the interest in information hiding:

- Terrorist cells can use inconspicuous communications to send plans.

- Intelligence and Military organizations have need for inconspicuous communications signal detection techniques that help to identify the enemy and results in speedy counter attacks.
- Law enforcement agencies and counter intelligence are interested in learning how stego-systems (software) can be compromised, so they can detect hidden communications[27].

So are terrorist cells using high-tech stego tools to insert hidden messages into ordinary communication? So far, the experts do not think this is being done and there is no way to tell. But, it is obvious, that our military and intelligence services seem to feel this is happening. Any news videos of Bin Laden will be scrutinized and examined looking for any signatures of stego software signatures in order to counter-attack such attempts. However, users of the Internet need anonymous communications, industry needs new tools to embed copyright mark invisibly in digital media, but spies can abuse these tools to pass on hidden communications in ordinary data over the public networks. For these reasons it can expected to hear a lot more about Steganography.

Steganography Software References

Many software applications are available that provide steganographic results. This following list gives a sample of the software available for the PC platform. All effort is being made to credit the authors of the software listed in this paper. However, some authors wish to remain anonymous.

1.Anonymous, Author alias: Black Wolf. *StegoDos - Black Wolf's Picture Encoder v0.90B*, Public Domain. ftp://ftp.csua.berkeley.edu/pub/cypherpunks/steganography/stegodos.zip.
2. Arachelian, R.: *White Noise Storm™ (WNS)*, Shareware (1994) ftp://ftp.csua.berkeley.edu/pub/cypherpunks/steganography/wns210.zip.
3. Brown, A.: *S-Tools for Windows*, Shareware 1994. ftp://idea.sec.dsi.unimi.it/pub/security/crypt/code/s-tools3.zip (version 3), ftp://idea.sec.dsi.unimi.it/pub/security/crypt/code/s-tools4.zip (version 4)
4. Hastur, H: *Mandelsteg,* ftp://idea.sec.dsi.unimi.it/pub/security/crypt/code/
5. Machado, R.: *EzStego, Stego Online, Stego*, http://www.stego.com
6. Maroney, C.: *Hide and Seek*, Shareware. ftp://ftp.csua.berkeley.edu/pub/cypherpunks/steganography/hdsk41b.zip (version 4.1), http://www.rugeley.demon.co.uk/security/hdsk50.zip (version 5.0), http://www.cypher.net/products/ (version 1.0 for Windows 95)
7. Repp, H.: *Hide4PGP,* http://www.rugeley.demon.co.uk/security/hide4pgp.zip
8. Hansmann F.: *Steganos.* Deus Ex Machina Communications. http://www.steganography.com.
9. Digimarc Corporation: *PictureMarc™*, *MarcSpider™,* http://www.digimarc.com
10. Kutter, M., Jordan, F.: *JK-PGS (Pretty Good Signature)*. Signal Processing Laboratory at Swiss Federal Institute of Technology (EPFL). http://ltswww.epfl.ch/~kutter/watermarking/JK_PGS.html
11. MediaSec Technologies LLC.: *SysCop™*, http://www.mediasec.com/
10. Signum Technologies, *SureSign*, http://www.signumtech.com
12. Upham, D.: *Jpeg-Jsteg*. Modification of the Independent JPEG Group's JPEG software (release 4) for 1-bit steganography in JFIF output files. ftp://ftp.funet.fi/pub/crypt/steganography
13. http://members.tripod.com/steganography/stego/software.html
14. http://www.jjtc.com/ihws98/jjgmu.html
15. Watermark and Steganography Analysis and Testing Tools Kuhn, M.: StirMark. URL: http://www.cl.cam.ac.uk/~fapp2/watermarking/image_watermarking/stirmark (1997)
16. Sanders, D.: Stegodetect. Steganography detection tool. (1997)
17. Anonymous: unZign. Watermarking testing tool available at http://altern.org/watermark/ - the author may be contracted through unzign@hotmail.com (1997)
18. URL: http://members.tripod.com/steganography/stego/software.html

Images

1. F-16. Courtesy of the Signal and Image Processing Institute at the University of Southern California.
2. Arctic Hare. Copyright photos courtesy of Robert E. Barber, Barber Nature Photography (BarberR@aol.com)
3. New-York. Copyright photo courtesy of Patrick Loo, University of Cambridge (pl201@cam.ac.uk
4. The Information Hiding Homepage Digital Watermarking & Steganography http://www.cl.cam.ac.uk/~fapp2/steganography/index.html
5. Digital Water Mark example. URL: http://www.research.ibm.com/image_apps/wmsample.html
6. Photo Database http://www.cl.cam.ac.uk/%7Efapp2/watermarking/benchmark/image_database.html - Lea

[1]

[2] http://www.spammimic.com/.

[3] Johnson, N.F. and S. Jajodia. "Steganalysis of Images Created Using Current Steganography Software." Center for Secure Information Systems George Mason University Fairfax, Virginia 22030-4444. http://www.ise.gmu.edu/~njohnson/ihws98/jjgmu.html.

[4] http://www.geocities.com/Athens/8744/herhist.htm.

[5] The Oxford English Dictionary. Clarendon Press, Oxford, 1933.

[6] Kuhn, Markus. Steganography Mailing List. URL: http://www.cl.cam.ac.uk/~mgk25/. URL: http://www.iks-jena.de/mitarb/lutz/security/stegano.html. 1995-07-03

[7] Johnson, N.F. and S. Jajodia. "Steganalysis of Images Created Using Current Steganography Software." Center for Secure Information Systems George Mason University Fairfax, Virginia 22030-4444. http://www.ise.gmu.edu/~njohnson/ihws98/jjgmu.html

[8] Peterson, Joseph H. URL:http://www.esotericarchives.com/tritheim/stegano.htm (03/21/2002).

[9] http://es.rice.edu/ES/humsoc/Galileo/Catalog/Files/schott.html

[10] The Voynich Manuscript. URL: http://www.crystalinks.com/voynich.html

[11] http://www.research.att.com/~reeds/voynich.html (Bell Labs is now part of Avaya and is known as Avaya Labs).

[12] Newman, B. Secrets of German Espionage. London: Robert Hale Ltd, 1940; J. E. Hoover, "The enemy's masterpiece of Espionage." The Reader's Digest , vol. 48, pp. 49{53, May 1946, London edition. Inc., 3rd edition. 1994, ISBN 0-471-59342-7.

[13] Kirovski and Malvar, Henrique. "Robust Covert Communication over a Public Audio Channel Using Spread Spectrum." URL: http://www.cs.ucla.edu/~darko/papers/SSW2.pdf http://citeseer.nj.nec.com/484260.html

[14] Bamford, James. The Puzzle Palace. New York: Penguin Group, 1983, page 447- 450. ISBN 0140067485.

[15] Lindsey, Robert. The Falcon and the Snowman .New York: Simon & Schuster, 1979, page 218

[16] http://www.research.ibm.com/image_apps/wmsample.html; THE NEW YORK TIMES. "Digital Copyright Agreement for Video." February 17, 1999 http://www.nytimes.com/library/tech/99/02/biztech/articles/17blue.html.

[17] Johnson, N.F. and S. Jajodia. "Steganalysis: The Investigation of Hidden Information." IEEE Information Technology Conference , Syracuse, New York, USA, September 1st - 3rd, 1998: 113-116. URL: http://www.jjtc.com/ihws98/jjgmu.html.

[18] Kelley, Jack. Terror groups hide behind Web encryption. USA TODAY. 02/05/2001 http://www.usatoday.com/life/cyber/tech/2001-02-05-binladen.htm.

[19] Provos, Niels. "Scanning USENET for SteganographyContent on the Internet." Fri Jan 4 2002. URL: http://www.citi.umich.edu//u/provos/stego/usenet.php

[20] First Information Hiding Workshop held in Cambridge, UK in April 1996; Anderson, R., (ed.): Information Hiding: First International Workshop, Cambridge, UK. Lecture Notes in Computer Science, Vol. 1174. Springer-Verlag, Berlin Heidelberg, New York (1996).

[21] Johnson, N.F. and S. Jajodia . "Steganalysis: The Investigation of Hidden Information." IEEE Information Technology Conference , Syracuse, New York, USA, September 1st - 3rd, 1998: 113-116. URL: http://www.jjtc.com/ihws98/jjgmu.html

[22] Kuhn, Markus. Steganography Mailing List. URL: http://www.cl.cam.ac.uk/~mgk25/. URL: http://www.iks-jena.de/mitarb/lutz/security/stegano.html. 1995-07-03
URL: http://www.cl.cam.ac.uk/~fapp2/steganography/mailing_list.html.
Markus Kuhn was a Computer Science student -- University of Erlangen,

[23] Johnson, N.F. and S. Jajodia. "Steganalysis of Images Created Using Current Steganography Software." Center for Secure Information Systems George Mason University Fairfax, Virginia 22030-4444. http://www.ise.gmu.edu/~njohnson/ihws98/jjgmu.html.

[24] Kuhn, Markus. Steganography Mailing List. URL: http://www.cl.cam.ac.uk/~mgk25/. URL: http://www.iks-jena.de/mitarb/lutz/security/stegano.html. 1995-07-03
URL: http://www.cl.cam.ac.uk/~fapp2/steganography/mailing_list.html.
Markus Kuhn was a Computer Science student -- University of Erlangen,

[25] Digital Copyright Agreement for Video By THE NEW YORK TIMES
http://www.nytimes.com/library/tech/99/02/biztech/articles/17blue.html
[2626] [24] Petitcolas, F. MP3Stego. URL: http://www.cl.cam.ac.uk/~fapp2/steganography/mp3stego/

[27] [25] Petitcolas, F. Anderson, J. and Kuhn, M. G. "Information Hiding: A Survey", IEEE Special Issue on Protection of Multimedia Content, 1999. URL: http://citeseer.nj.nec.com/petitcolas99information.html