Journal of Advances in Science and Technology

Vol. IV, No. VII, November-2012, ISSN 2230-9659

A SURVEY ON DIFFERENT APPLICATION OF DATA HIDING

www.ignited.in

Journal of Advances in Science and Technology
Vol. IV, No. VII, November-2012, ISSN 2230-9659

A Survey on Different Application of Data Hiding

Amit Upadhyaya[1] Dr. Rajesh Kumar Pathak[2] Dimple Jayaswal[3]

[1]Ph.D. (Pursuing) CMJ University, Shillong, Meghalaya

[2]Ph.D. GNIET Engineering College Greater Noida, Utter Pradesh

[3]M.Tech. (Pursuing) RTU, Kota, Rajasthan

Abstract-Steganography is the art of hiding the fact that communication is taking place. The purpose of steganography is covert communication to hide the existence of a message from a third party. Techniques to hide valuable information within seemingly harmless messages have been widely used for centuries. Typically, their use is appropriate when encryption is not available or not adequate (e.g. when available cryptography is too weak), or simply when it is convenient that no external observer can infer that some information is being exchanged.

In the digital era, new cover mediums for hiding data in communication are constantly being proposed, from the classical image files (such as bmp, gif, and jpg formats) to audio files(i.e. wav and mp3), text and html documents, emails disguised as spam, TCP/IP packets, executables programs, DNA strands, etc. In this work, we present and analyze on different applications of data hiding. These application are used continuously in digital world i.e. Forensic Science, Games, intelligence bureau and anti counterfeiting and authentication.

KeyWords: Application, Data Hiding, Stegnography, Forensic Science, Anticounterfieting

----------♦----------

1. INTRODUCTION

With the development of computer and expanding its use in different areas of life and work, the issue of information security has become increasingly important. One of the grounds discussed in information security is the exchange of information through the cover media. To this end, different methods such as cryptography, steganography, coding, etc have been used. The method of steganography is among the methods that have received attention in recent years[1]. The main goal of steganography is to hide information in the other cover media so that other person will not notice the presence of the information. This is a major distinction between this method and the other methods of covert exchange of information because, for example, in cryptography, the individuals notice the information by seeing the coded information but they will not be able to comprehend the information. However, in steganography, the existence of the information in the sources will not be noticed at all. Most steganography jobs have been carried out on images, video clips, texts, music and sounds .Nowadays, using a combination of steganography and the other methods, information security has improved considerably. In addition to being used in the covert exchange of information, steganography is used in other grounds such as copyright, preventing e-document forging. Steganography is derived from the Greek for covered writing and essentially means “to hide in plain sight”. Steganography is the art of inconspicuously hiding data within data. The main goal of steganography is to hide information well enough such that the unintended recipients do not suspect the steganographic medium of containing hidden data Simple steganographic techniques have been in use for hundreds of years, but with the increasing use of files in an electronic format new techniques for information hiding have become possible. Most steganography jobs have been carried out on different storage cover media like text, image, audio or video. Steganography [2] & encryption are both used to ensure data confidentiality however the main difference between them is that with encryption anybody can see that both parties are communicating in secret. Steganography hides the existence of a secret message and in the best case nobody can see that both parties are communicating in secret. This makes steganography suitable for some tasks for which encryption aren’t, such as copyright marking. Table 1 shows a comparison of different techniques for communicating in secret [4]. Encryption allows secure communication requiring a key to read the information. An attacker cannot remove the encryption but it is relatively easy to modify the file, making it unreadable for the intended recipient.

TABLE 1
COMPARISON OF SECRET COMMUNICATION TECHNIQUES.

Secret Communication Techniques	Confidentiality	Integrity	Un removability
Encryption	Yes	No	Yes
Digital Signatures	No	Yes	No
Steganography	Yes/No	Yes/No	Yes

Steganography hides the covert message but not the fact that two parties are communicating with each other. The stego process generally involves placing a hidden message within some transport medium, called the carrier . The secret message is embedded within the carrier to form the stego medium. The use of a stego key may be employed for encryption of the hidden message and/or for randomization within the stego scheme. In summary:

stego_medium = hidden_message + carrier + stego_key

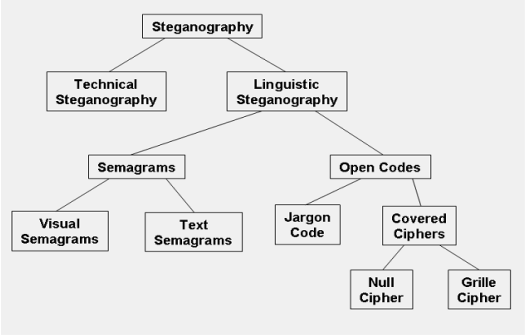

Figure 1 shows a common taxonomy of stenographic techniques [6,7]:

Technical steganography uses scientific methods to hide a message, such as the use of invisible ink or microdots and other size reduction methods.

• Linguistic steganography hides the message within the carrier in some non-obvious ways and is further categorized as semagrams or open codes.

• Semagrams hide information by the use symbols or signs. A visual semagram uses innocent-looking or everyday physical objects to convey a message, such as doodles or the positioning of items on a desk (or Web site). A text semagram hides a message by modifying the appearance of the carrier text, such as subtle changes in font size or type, adding extra spaces, or different flourishes in letters or handwritten text.

• Open codes hide a message within a legitimate carrier message in ways that are not obvious to an unsuspecting observer. The carrier message is sometimes called the overt communication while the hidden message is the covert communication. This category is subdivided into jargon codes and covered ciphers.

• Jargon code, as the name suggests, uses language that is understood by a group of people but is meaningless to others. Jargon codes in clude war chalking (symbols used to indicate the presence and type of wireless network signal), underground terminology, or an innocent conversation that conveys special meaning because of facts known only to the speakers. A subset of jargon codes are cue codes, where certain pre-arranged phrases convey meaning.

• Covered, or concealment, ciphers hide a message openly in the carrier medium so that it can be recovered by anyone who knows the secret for how it was concealed. A grille cipher employs a template that is used to cove r the carrier message; the words that appear in the openings of the template are the hidden message. A null cipher hides the message according to some prearranged set of rules, such as "read every fifth word" or "look at the third character in every word."

2. FEATURES

Data-hiding techniques should be capable of embed-ding data in a host signal with the following restrictions and features:

a. The host signal should be non objection ally degraded and the embedded data should be minimally perceptible. (The goal is for the data to remain hidden. As any magician will tell you, it is possible for something to be hidden while it remains in plain sight; you merely keep the person from looking at it. We will use the words hidden, inaudible, unperceivable, and invisible to mean that an observer does not notice the presence of the data, even if they are perceptible.)

b. The embedded data should be directly encoded into the media, rather than into a header or wrap-per, so that the data remain intact across varying data file formats.

c. The embedded data should be immune to modifications ranging from intentional and intelligent attempts at removal to anticipated manipulations, e.g., channel noise, filtering, resampling, cropping,encoding, lossy compressing, printing and scanning, digital-to-analog (D/A) conversion, and analog to digital (A/D) conversion, etc.

d. Asymmetrical coding of the embedded data is desirable, since the purpose of data hiding

is to keep the data in the host signal, but not necessarily to make the data difficult to access.

e. Error correction coding should be used to ensure data integrity. It is inevitable that there will be some degradation to the embedded data when the host signal is modified. [5]

f. The embedded data should be self-clocking or arbitrarily re-entrant. This ensures that the embedded data can be recovered when only fragments of the host signal are available, e.g., if a sound bite is extracted from an interview, data embedded in the audio segment can be recovered. This feature also facilitates automatic decoding of the hidden data, since there is no need to refer to the original host signal.

3. APPLICATION

Steganography is employed in various useful applications, e.g., copyright control of materials, enhancing robustness of image search engines and smart IDs (identity cards) where individuals' details are embedded in their photographs. Other applications are video-audio synchronization, companies' safe circulation of secret data, TV broadcasting, TCP/IP packets (for instance a unique ID can be embedded into an image to analyze the network traffic of particular users) [8], and also checksum embedding [9]. Petitcolas [10] demonstrated some contemporary applications, one of which was in Medical Imaging Systems where a separation is considered necessary for confidentiality between patients' image data or DNA sequences and their captions, e.g., physician, patient's name, address and other particulars. A link however, must be maintained between the two. Thus, embedding the patient's information in the image could be a useful safety measure and helps in solving such problems. Steganography would provide an ultimate guarantee of authentication that no other security tool may ensure.

Many departments or organizations use steganography to hide the data. Some of them use stegnography for illegal purposes that should be detected. To detect the hiding data, there are lots of tools available. In this paper we discuss about some applications of hiding data.

3.1 ANTICOUNTERFEITING

Recent advances in the technology for ink-jet printers and scanners have created a new way for computer users to inexpensively create high-quality color reproductions of any original document. One obvious problem that arises from these advances is the ability to casually create counterfeit currency (or Some other valuable document) that looks real enough to fool anyone who does not inspect it care-fully. The embedding of a digital signature into a secure document that can be recognized by a printer as it prints the document is a potential deterrent to casual counterfeiters. If a secure document is recognized, the printer can refuse to complete the print job and issue an appropriate warning. For this deterrent to serve as a viable solution, however, it must provide a high level of security while requiring only a minimal addition to a printer. Potentially, such anticounterfeiting methods could be used to protect currency, stock certificates, bank checks, or airline tickets from illicit reproduction. These methods are of particular applicability when unsuspecting individuals are exchanging documents for something of value.

To serve as an example, the implementation of an encoding and detecting algorithm, Tartan Threads,[11,12] based upon multiple-layer encoding, is detailed and evaluated regarding its ability to address the casual counterfeiting problem. (The name Tar-tan Threads was chosen because the method uses striped patterns that are reminiscent of the security threads found in U.S. paper currency.) Tartan Threads is designed to hold information in a fixed-size linearly contiguous space to allow for time-efficient decoding with a high degree of certainty. When used in conjunction with a system that marks continuous-tone printouts with a unique digital sig-nature, it provides ink-jet printers copy protection commensurate with color copiers. Marvel et al.[13] have implemented a blind digital steganography system called Spread Spectrum Image Steganography (SSIS) built upon a two-dimensional spread-spectrum method. Through the use of image restoration techniques, an estimate of the original image is recreated and then subtracted from the en-coded document to reveal the encoded information. As compared to Tartan Threads, SSISyields a higher encoding bandwidth and a lower perceptibility. The encoding, however, is not intended to survive a printing and scanning image cycle, nor is it amenable to quick decoding.

Herrigel et al. [14] and Fridrich et al. [15] both describe image watermarking methods built on spread-spectrum techniques combined with a public-key encryption system for authentication of both the author and purchaser of digital images. Herrigel's technique, like Tartan Threads, redundantly encodes several small areas of the image with identical watermarks, although Herrigel et al.'s watermark is in the Fourier transform domain rather than the spatial domain.

The technique is intended to survive cropping, as the areas are tiled and encoded in the single orientation. Rotation and scaling transformations are handled through the analysis of these encoded blocks in a

polar coordinate space. By calculating the Fourier transform of the entire image, it is possible to characterize any rotation or scaling that had been applied to the image. Fridrich et al. combines global and local encoding schemes. As an additional security mea-sure, encoding patterns are generated using a secret key. Since these techniques involve two-dimensional encoding methods, decoding requires extensive processing times for larger images.

These methods focus primarily on marking images that are intended to be distributed in digital form. They may tolerate some loss but they are not in-tended to survive the combination of sampling and quantization errors introduced by a printing and scanning cycle, the nonlinear effects produced in various printing processes, etc. Tartan Threads are in-tended for images that will be distributed in printed form. The encoding survives even at low scanning resolutions (100 dpi [dots per inch]) and can be de-coded without complicated analysis of the encoded image— only a small contiguous area of the image needs to be processed. Any encoding on an actively circulated document, such as currency, must be detectable even after wear from usage. (An interesting area for further work would be to create a model for how such documents typically wear over time—it is unknown if such a model publicly exists.) Since different users—with different equipment—will potentially be scanning the protected document, the encoding must also survive any no geometric transformations and lossy image transformations that may result from the use of different image file formats, compression methods, slight rotations, imperfect color sampling, and re-sampling at varying resolutions and offsets. Ideally, the encoding would also survive any transformation a user is likely to apply to the digitized image data that does not call attention to the human eye. Because an ink-jet printer renders images line-by-line and often has only enough memory to buffer a few image lines (typically between 16 kilobytes to 1 megabyte of buffer memory for even the more sophisticated models), all decoding must rely on only a small portion of the document data. Ideally, the decoder would be created with inexpensive hardware capable of affordable searching for encoded information in every print line. Any solution must also have extremely low probabilities of false triggering to prevent disrupting consumers who are using their printers for legitimate purposes.

3.2 GAME APPLICATIONS

The idea of hiding information into game strategies can be applied in two different scenarios. The first and more straight-forward one consists in playing a new game from scratch. However, a similar approach can be put in practice by adding extra information to an already played game. In this case, two parties can append comments and variations at different stages of the game.

A) Play a new game:

Both parts of the communication channel should have access to exactly the same software, and share a common secret key (used for encrypting the hidden contents) and some other parameters (i.e. the board size, the software version, etc.) to assure that the software's internal states are reproducible by the two parties. The main idea is to compute for each position played in the

game all the movements which are over a certain threshold value T, and then codify in the selection of the actual played move some bits of the message we want to hide. Say that, at a given position there are gm (good moves) moves not worst than the given threshold T (analogously, we can fix a certain number n and pick the move within the list of best n moves), then sort them by their value (according to the evaluation function) and select the ith move to codify the binary representation of the number i. In this way, in each position we will be able to hide around log2 (gm) bits of information. The choice of threshold T allows us to adjust some aspects of the scheme. By increasing T, the channel capacity will grow, since we will have more gm to embed hidden bits. Alternatively, decreasing T will result in a higher invisibility of the hidden contents, for chosen strategies do not deviate significantly from optimal.

B) Inserting data in to a game:

Inserting data presents the benefits of simplicity, high embedding capability, and security. It simply consists in including variations and comments in already played games.

The general idea is basically to follow the overall afore-mentioned procedure for hiding some bits in every move, but in this case moves correspond to variants of the main line actually played on the game. At certain positions in the game we would introduce variants of the main line played. Each of these variants could embed bits by various means: choosing at what movement does the variant begin, the length of the variant, and obviously, in every move of the variant, just by using the same algorithm described above for codifying hidden data in moves.

Generally, this embedding technique will allow hiding much more information than the first one, and will be in most cases harder to detect. However, it also has the important drawback of a much lower robustness: it suffices to delete all comments to erase the hidden data, but at the cost of severely decreasing the attractiveness and usefulness of the game file. [16]

3.3 FORENSIC SCIENCE

Some people use stegnagraphy for illegal purposes. To detect them forensic department use tools for detecting the steganography or hiding data. In market currently many softwares are available that can detect the presence of stego programs, detect

suspect carrier files, and disrupt steganographically hidden messages.

The detection of stego software on a suspect computer is important to the subsequent forensic analysis. As the research shows, many stego detection programs work best when there are clues as to the type of stego that was employed in th e first place. Finding stego software on a computer would give rise to the suspicion that there are actually stego files with hidden messages on the suspect computer. Furthermore, the type of stego software found will directly impact any subsequent steganalysis; e.g., S-Tools might direct one's attention to GIF, BMP, and WAV files while JP Hide-&-Seek might direct the analyst to look more closely at JPEG files. [17]

WetStone Technologies' Gargoyle (formerly StegoDetect) software [18] can be used to detect the presence of stego software. Gargoyle employs a proprietary data set (or hash set) of all of the files in the known stego software distributions, comparing them to the hashes of the files subject to search. Gargoyle data sets can also be used to detect the presence of cryptography, in stant messaging, key logging, Trojan horse, password cracking, and other nefarious software.

AccessData's Forensic Toolkit (FTK) [19] and Guidance Software's EnCase [20] can use the HashKeeper [21], Maresware [22], and National Software Reference Library (NSRL) [23] hashsets to look for a large variety of software. In general, these data sets are designed to exclude hashes of known "good" files from search indexes during the computer forensic analysis. Gargoyle can also import these hash sets.

The detection of steganography software continues to become harder for another reason -- the small size of the software coupled with the increasing storage capacity of removable media. S-Tools, for example, requires less than 600 KB of disk space and can be executed directly, without additional installation, from a floppy or USB memory key; under those circumstances, no remnants of the program would be found on the hard drive.

The second important function of steganography detection software is to find possible carrier files. Ideally, the detection software would also provide some clues as to the stego algorithm used to hide information in the suspect file so that the analyst might be able to attempt recovery of the hidden information.

4. CONCLUSION

As today communication in digital form is necessary for people and organizations. This communication should be safe and secure. To do so there are lots of techniques and tools available in the market. Stegnography is also a technique to hide data or to kept communication secured. Steganography can be used in different manner as in games (to hide user's detail and some content of stages), anticounterfieting (to print secure document) and in forensic department (to hide the illegal contents). Steganoraphy is the wide area in which data hiding is done through various techniques. This paper concludes that data hiding can be used in legal and illegal manner.

5. REFERENCES

5.1 Mohammad Shirali-Shahreza , "A ne w method for real time steganography", ICSP 2006 Proceedings of IEEE .

5.2 Yuk Ying Chung, fang FeiXu , "Development of video watermarking for MPEG2 video" City university of Hong Kong ,IEEE 2006.

5.3 C. Lu, J. Chen and K. Fan, "Real-time Frame-Dependent Video Watermarking in VLC Domain", Signal Processing : Image Communication 20, 2005, pp. 624–642.

5.4 Jonathan Cummins, Patrick Diskin, Samuel Lau and Robert Parlett, " Steganography and digital watermarking " School of Computer Science, The University of Birmingham. 2003. www.cs.unibo.it/people/phdstudents/scacciag/home_files/teach/datahide.pdf.

5.5 P. Sweene, Error Control Coding (An Introduction), Prentice-Hall International Ltd., Englewood Cliffs, NJ (1991).

5.6 Arnold, M., Schmucker, M., and Wolthusen, S.D. Techniques and Applications of Digital Watermarking and Content Protection. Norwood (MA): Artech House, 2003.

5.7 Bauer, F.L. Decrypted Secrets: Methods and Maxims of Cryptology , 3rd ed. New York: Springer-Verlag, 2002.

5.8 N.F. Johnson and S. Jajodia, Exploring steganography: Seeing the unseen, IEEE Computer, 31(2)(1998) 26-34.

5.9 W. Bender, W. Butera, D. Gruhl, R. Hwang, F.J. Paiz and S. Pogreb, Applications for data hiding, IBM Systems Journal, 39 (3&4)(2000) 547-568.

5.10 F.A.P. Petitcolas, "Introduction to information hiding", in: S. Katzenbeisser and F.A.P. Petitcolas, (ed.) (2000) Information hiding

techniques for steganography and digital watermarking, Norwood: Artech House, INC.

5.11 R. Hwang,A Robust Algorithm for Information Hiding in Dig-ital Pictures , M.Eng. thesis, MIT, Cambridge, MA (May 1999).

5.12 F. Paiz, Tartan Threads: A Method for the Real-Time Digital Recognition of Secure Documents in Ink Jet Printers, M.Eng. thesis, MIT, Cambridge, MA (May 1999).

5.13 L. M. Marvel, C. G. Boncelet, and C. T. Retter, “Reliable Blind Information Hiding for Images,” Information Hiding: Second International Workshop , D. Aucsmith, Editor, Lecture Notes in Computer Science1525 , Springer-Verlag, Portland, OR (April 15–17, 1998), pp. 48 – 62.

5.14 A. Herrigel, J. J. K. O Ruanaidh, H. Petersen, S. Pereira, and T. Pun, “Secure Copyright Protection Techniques for Dig-ital Images,” Information Hiding: Second International Work-shop , D. Aucsmith, Editor, Lecture Notes in Computer Sci-ence 1525, Springer-Verlag, Portland, OR (April 15–17, 1998), pp. 169 –190.

5.15 J. Fridrich, “Robust Digital Watermarking Based on Key-Dependent Basis Functions,”Information Hiding: Second In-ternational Workshop , D. Aucsmith, Editor, Lecture Notes in Computer Science1525, Springer-Verlag, Portland, OR (April 15–17, 1998), pp. 143–157.

5.16 Steganography in games: A general methodology and its application to the game of Go Julio C. Hernand ez-Castro, Ignacio Blasco-Lopez, Juan M. Estevez- Tapiador, Arturo Ribagorda-Garnach o Computer Science Department, Carlos III University of Madrid, Avda. Universidad 30, 28911 Legane´ s, Madrid, Spain.

5.17 StegoArchive.com Web Site. URL: http://www.stegoarchive.com/. Last accessed: 2003-12-30.

5.18 WetStone Technologies Web Site. "Gar goyle." URL: http://www.wetstonetech.com/ gargoyle_ns.html. Last accessed: 2003-12-29.

5.19 AccessData Web site. "Forensi c Toolkit product page." URL: http://www.accessdata.com/Product04_Overvi ew.htm. Last accessed: 2003-12-29.

5.20 Guidance Software Web Site. URL: http://www.guidancesoftware.com/. Last accessed: 2003-12-29.

5.21 Hashkeeper Files Web Site. URL: http://www.hashkeeper.org/files/. Last accessed: 2003-12-29.

5.22 Maresware Web site. "Hash Set CD." URL: http://www.dmares.com/maresware/ hash_cd.htm. Last accessed: 2003-12-29.

5.23 National Software Reference Libr ary (NSRL) Project Web Site. URL: http://www.nsrl.nist.gov/. Last accessed: 2003-12-29.