An Inductive Approach to the Knack of Steganology

Remya A R
Research Scholar
Dept. of Computer Applications
Cochin University of Science
and Technology, Kochi-22

A Sreekumar,Ph.D
Associate Professor
Dept. of Computer Applications
Cochin University of Science
and Technology, Kochi-22

ABSTRACT
Steganology, the knack of unseen communication has a significant pose in the modern digital era. Steganography, the practice of concealing a message in any digital medium to the extent that it is not even revealed that a hidden communication is happening through. Steganalysis, the practice to listen the communication medium that results in detection of the presence of any hidden communication. This paper presents a brief overview in the field of Steganology, especially steganography.

General Terms
Security, Communication systems.

Keywords
Steganography, Steganalysis, Cover Medium, Stego Object, Stego Medium

1. INTRODUCTION
Any message or data passed through digital devices is termed as DIGITAL COMMUNICATION. Protecting or securing digital communication is a significant challenge and can be achieved by cryptography, steganography or digital signature techniques.

Cryptography ensures the communication to be private by transforming the information into an unreadable format such as scrambling. The technique used to accomplish this is termed as encryption and the reverse is termed as decryption. Encryption mechanisms uses encryption key to encrypt the messages at the sender side and the decryption mechanisms which uses decryption key to decrypt the messages at the receiver side. That is both scrambling and unscrambling can be achieved with the aid of keys.

Steganography [1] skins the existence of stealthy data in the communication. Using this method the secret data is embedded into a media which works as the cover and cater the communication obscure to human eye. Steganography systems should be non-vulnerable to the systems which check the statistical patterns for identifying the presence of any hidden data in communication.

Digital signature helps to retain the authenticity of an electronic communication and thereby maintains the trust between the sender and the receiver. A digital signature guarantees that the electronic document to be authentic the authorship be affirmed. Any change to the document will invalidate the signature or integrity of the document. Digital signature rely on encode and decode techniques which work hand in hand.

2. THE KNACK OF STEGANOGRAPHY
The word Steganography is derived from the Greek word Steganographia which means covered writing. Steganography is inimitable from other data hiding techniques because of the use of a covered medium which passes across the intervening channel without any evidence of the secret data transmission.

i. Steganography in Early Days:

A. Practical Steganography

Steganography, the most famous sub-discipline of Information Hiding has got a very long historical background can be traced backed to 440 BC. Herodotus the ancient Greek historian narrates the story of secret communication happened during the war between the Persian Empire and Greek city states where in the messages where wrote on shaved the head of Greek messenger and once his hair grown back, he could travel and reveal the message to the desired recipients without making any suspicion. Another way of sending secret information was by removing the wax of writing tablet with a wood which has some messages written on it and then coated with wax, to get an appearance of a blank one. Chinese history reveals the usage of small piece of silk for writing the secret messages and foils these into small balls, to make it swallowed by the messenger. In 1860s, a French Photographer, Dragon makes tiny images which result in the development of microscopic images. Another form is the architecture: the work of paintings, sculptures etc. appear different from certain angles which can be used to hide some secret information.

B. Text Steganography

Text steganography[2][3] involves hidden messages embedded in character based text or natural language and can be anything from changing words with in text to generating random character sequence generation or acrostic. Johannes Trithemius (1462-1516), the author of the first printed book on cryptology invented a steganographic cipher in which each letter was represented as a word taken from a succession of columns to form a legitimate prayer *Ave-Maria-Cipher*. Ancient Chinese used data hiding technique where in the sender wrote messages on the paper through the holes of the mask and embeds the cover message once the mask is removed and the receiver, who had the same mask, would retrieve the hidden message. In the early 16th century Cardan, an Italian Mathematician, reinvented the method of using Paper Mask which is now known as *Cardan Grille*. Semagrams [4] that hide information by the use of symbols or

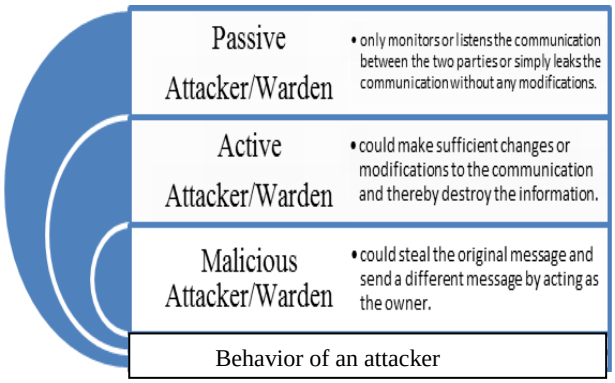
signs and Open Codes [3] which employ legitimate carrier messages to hide a covert message are key Linguistic Steganography techniques.

ii. Steganography in Digital Age

A. Modern Steganography

In 1983, Gustavus . J. Simmons laid the foundations of modern steganography and introduced the concept of subliminal channel; authentication without secrecy. He elucidated it with the prisoner's problem [5].Alice and Bob are jailbirds who communicate to develop escape plans and their communication monitored by warden named Wendy. If Alice and Bob use cryptography for communication of escape plan Wendy will recognize it compromise their plan. The only way for them to communicate is to send an innocuous message and embed the conceding information in it. The transmission channel is no more visible by Wendy; it is a subliminal channel [6] [7].

Behavior of an attacker, in this case the warden Wendy can be illustrated as:

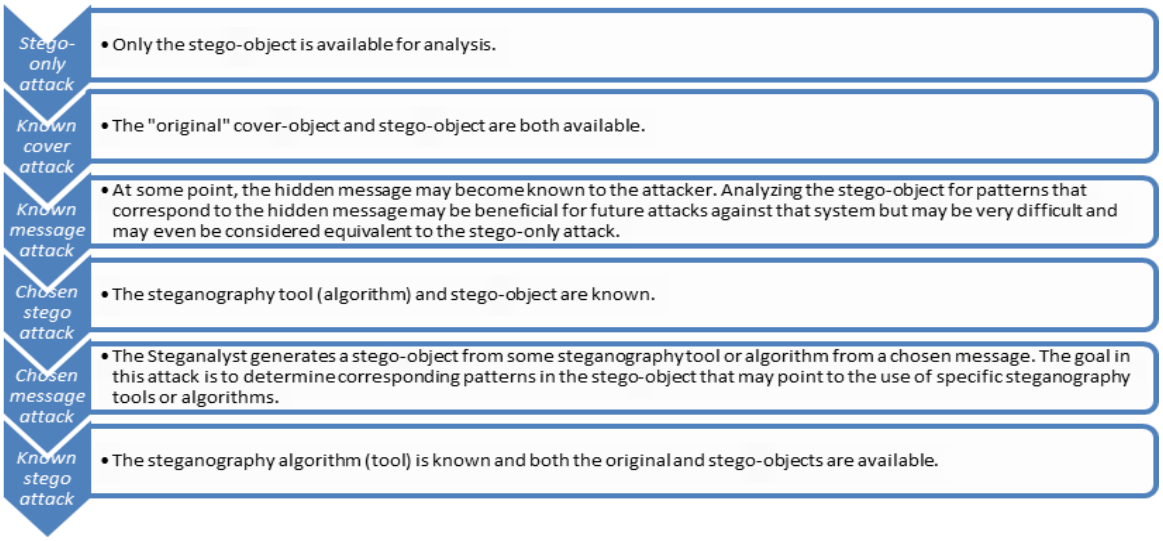


Modern-day Steganography generally refers to hiding information in electronic/digital media such as images, audio, video or even text files. With digital technique it is easy to hide and extract data and as the size of the information to be hidden is small compared to the data in which it is hidden (cover) the electronic media is best suited. Steganography technique in the digital era may be concealing messages within the lowest bits of noisy images or sound files or Chaffing [8] [9] and windowing techniques [8] [9] or Mimic functions [10] convert one file to have the statistical profile of another. Cryptography provides confidentiality to the message transferred whereas steganography brings in the feature of invisibility to the transferred message.

B. Steganalysis

A threat on steganographic system is obvious and in any such circumstances the robustness of the system depends on its ability to resist against the following characteristics: detecting, extracting and disabling the embedded information. The technique used to achieve uncovering of hidden messages generated by steganographic system is termed as Steganalysis [12]. A steg-analyser is the people involved in the detection and try to manipulate the message to retrieve or destroy the hidden information in it. The complexity of steganalysis will increase as complexity of steganographic technique increase.

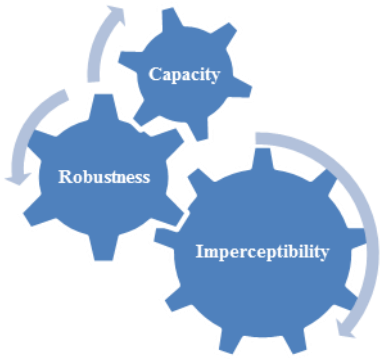
The attacks[1] which are available to the Steganalyst can be described as: Stego-only attack, known-cover attack, known-message attack, chosen-stego attack, chosen-message attack, known-stego attack etc.



3. STEGANOGRAPHIC SYSTEM

Steganography in digital communication guarantees two main directions in information hiding: protection against the detection of the existence of information and hiding the data.

3.1 Steganographic Requirements



A. Invisibility or Imperceptibility

The first and foremost requirement for any steganography algorithm is the invisibility. The strength of steganography lies in its ability to be overlooked by the human eye or systems which check the presence of any hidden data in communication. At any instance if one can comprehend that an image has been tampered with, the algorithm is compromised.

B. Robustness

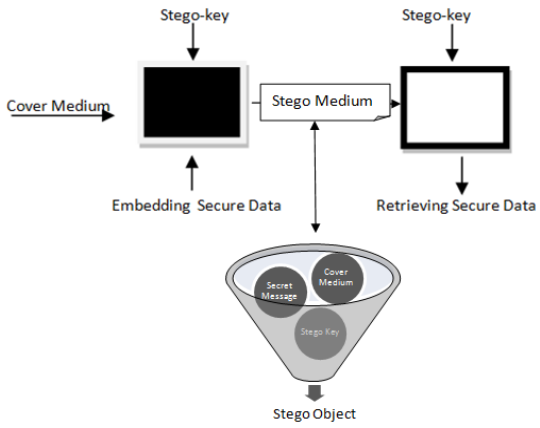
At any point of communication the information may undergo changes by an active warden in an attempt to remove or tamper the hidden information and it is desirable that the steganography algorithms be robust against these malicious or unintentional changes.

C. Payload Capacity

Steganography aims hidden communication and therefore requires ample embedding capacity. Capacity refers to the maximum volume of information that can be entrenched into a cover-object and then can be reliably recovered from the Stego-object (distorted version), without being discovered.

3.2 Steganographic Process

The steganography process involving the secure data to be send across the communication channel should be embedded in the pre-determined cover medium using the Stego-key. The cover medium with the Stego-key, termed as the Stego-medium is communicated over the channel. In order to retrieve the embedded information, the receiver should know the Stego-key.



The Steganographic Process

3.3 Steganography Protocols

In literature there are basically three types of steganographic protocols [10] [11] [14] [15]: *pure steganography*, *secret key steganography* and *public key steganography*.

A. Pure steganography doesn't require the exchange of the cipher such as a Stego key and is the least secure mean to communicate secretly. The sender and receiver rely on the presumption that no third party is overhearing the communication. The quadruple = <C, M, D, E>, where C is the set of possible covers, M the set of secret messages with $|C| \geq |M|$, E: $C \times M \rightarrow C$ the embedding function and D : $C \rightarrow M$, the extraction function, with the property that $D(E(c,m)) = m$ for all $m \in M$ and $c \in C$ is called a pure steganographic system.

B. Secret key steganography system is analogous to a symmetric cipher, where the sender elects the cover and entrenches the stealthy message by means of a secret key which is acknowledged by the receiver who can reverse the process and extract the stealth message. Any attacker will be incapable to acquire evidence of the encoded information. The quintuple = <C, M, K, DK, EK>, where C is the set of possible cover, M the set of secret messages with |C| ≥ |M|, K the set of secret keys, EK : C×M×K → C and DK : C×K → M with the property that DK(EK(c, m, k), k) = m for all m ∈ M, c ∈ C and k ∈ K, is called a secret key steganographic system.

C. Public key steganography system is analogous to a Public key Cryptography where both parties to communicate steganographically with no prior exchange of secrets. Public key steganography does not rely on the exchange of a secret key. These systems require the use of two keys: one private key and one public key. The public key which is stored in the public database is used in the embedding process, while the private key is used to extract the information.

3.4 Steganographic Methods

Steganographic methods can be classified in two different ways:

A. According to the type of cover medium used for the communication

a. Digitized photographs [16], with different image formats and videos are used as the cover medium. Photographs should be represented by a matrix of numbers. Image formats include JPEG, BMP, GIF etc.

b. Digitized sound [17] as the cover medium employs the property of representing pressure variations as a sequence of numbers in different time slices. The representations are imprecise and thereby providing enough space for hiding the information.

c. Text [18] as cover is achieved either by line-space encoding or word-space encoding by varying the positions of lines or words in the document. Or can be achieved by changing the heights of letter strokes.

d. Executable [19] are dealing with Mimic Functions proposed by Wayner [61] or CFGs. Making use of the changes in statistical profile of Mimic Functions or the different statistical properties of English language such as the non-uniformity in the distribution of characters.

B. According to the cover modification applied in the secure embedding process

a. Substitution Systems

i) Least Significant Bit Substitution [20] [21] [22]: In this technique LSBs of the selected cover medium are replaced with the secure data and it works on the notion that the LSB of each cover-element by the secret message is exchanged. Secret message can either be a 0 or 1. Substitution can also be done like replacing two LSBs of one cover element with two message bits.

ii) Pseudorandom Permutations [23] [24]: As an alternative of embedding the message bits chronologically in the increasing order of the index values for the cover, respective element in the pseudorandom number sequence could be generated anywhere from 1 to length (n), where n is the cover and the secret message bits could be entrenched at the cover-bits corresponding to the index values generated. Spread the message in the cover in a rather random manner. Using pseudorandom number generator for LSB substitution may result in collision. An indexing scheme based on pseudorandom permutation is used to overcome this problem.

iii) Image Downgrading and Covert Channels [25]: Another approach in substitution systems is where the secret message and covert channels are both images. Consider the case in which both the cover as well as secret images has equal dimensions: At the sender end, four LSBs of cover's gray scale or color values are exchanged with the four MSBs of the secret image. At the receiver end, extracting the four LSBs out of the Stego-image will permit access to the MSBs of the secret message.

iv) Cover-Regions and Parity Bits [24]: A cover-region is represented by any non-empty subset of {c1,…..,cl(c)}. Dividing the cover in several disjoint regions helps to store one bit of information in a whole cover-region rather than in a single element.

A parity bit of a region I can be calculated by:
In the embedding step, l(m) disjoint cover-regions Ii ($1 \leq i \leq$ l(m)) are selected, each encodes one secret bit mi in the parity bit p(Ii). If the parity bit of one cover-region Ii does not match with the secret bit mi to encode, one LSB of the values in Ii is flipped. This will result in p(Ii) = mi. In the decoding process, the parity bits of all selected regions are calculated and lined up to reconstruct the message. Again, the cover-regions can be constructed pseudo randomly using the Stego-key as a seed.

v) Quantization and Dithering [26] [27]: Quantization involves entrenching secure information in quantized images by utilizing the difference in the intensity values of the adjacent pixels. Dithering is basically the process of creating an indexed image approximation of the RGB image in the array RGB by dithering the colors in color map, i.e., it changes the colors of pixels in a neighborhood so that the average color in each neighborhood approximates the original RGB color.

vi) Palette-based and Binary Images [28] [29]: In a palette-based image only a subset of colors from a specific color space can be used to colorize the image. Every palette based image format consists of two parts – a palette and the actual image data – both of these can be manipulated for embedding secret messages making use of their LSB values. Binary images contain redundancies in the way black and white pixels are distributed.

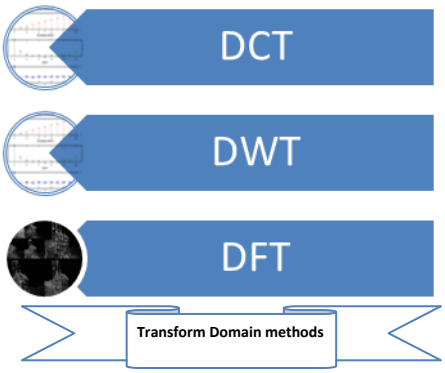
vii) Unused or Reserved space in Computer System [30]: Using these method helps in hiding the secret message without perceptually degrading the carrier. Another method incorporates the creation of hidden partition in

file systems. Unused space in file headers of image and audio can also be used to hold "extra" information. Another method is using unused space in packet header of TCP/IP packets

b. Transform Domain Techniques

Advancement over the above method is the Transform domain techniques where the messages are hidden in the significant areas of the cover image. These types of systems are more susceptible to attacks such as compression, cropping and some image processing techniques. This technique embeds a low amplitude signal with low bandwidth in a medium that presents a higher bandwidth.

Important transform domain techniques include: *Discrete Cosine Transform, Fourier Transform and Wavelet Transform* – these techniques are applied because the Linear Complexity plays a very important role in the case of cryptography as well as steganography.



i) DCT [31] [32] based Steganography entrench the text message in least significant bits of the Discrete Cosine (DC) coefficient of digital picture. When information is veiled inside image or video, the program hiding the information usually performs the DCT and it works by slightly changing each of the images in the video, only to the level that is not noticeable by the human eye.

ii) DFT [33] [34] methods are analogous to the DCT in terms of modifying the frequency coefficients in the mid-band region. [35] Nyquist rate is used in image processing by padding an image with zeroes before the DFT is taken. The DFT is taken on each color plane separately, so the focus will be on one arbitrary plane.

iii) DWT [36] is used to transform a spatial domain into frequency domain. The use of wavelet in image stenographic model lies in the fact that the wavelet transform clearly separates the high frequency and low frequency information on a pixel by pixel basis

c. Spread Spectrum Techniques and Information Hiding

In Spread Spectrum Image Steganography (SSIS) [37] [38] the secret message is held as a Gaussian noise in a cover image and at low noise power levels the cover image degradation is imperceptible by the human judgment and at higher levels the noise appears as speckles or image distortion.

d. Statistical Steganography
Statistical technique [39] otherwise termed model based technique tends to amend the statistical

properties of an image in addition to preserving them in the embedding process. The modification is trivial that it takes advantage of human weakness in detecting luminance variation. This technique modifies the cover image to make a sort of significant change in the statistical characteristics if a "1" is transmitted or it is left unchanged. To send multiple bits, the image is fragmented into sub-images, each corresponding to a single bit of the message.

e. Distortion Techniques:

In distortion technique, for a chosen cover-element, in order to encode a 0, sender leaves the pixel unchanged, to encode a 1, a random value Δx is added to the pixel's color. Receiver performs a comparison of all selected pixels of the Stego-image and cover-image. Distortion techniques involve sender and receiver of message to know the original cover image so that the receiver can make use of decoding functions to check for the differences between the original cover image and the distorted cover image received and restore the secret message. If the Stego-image is different from the cover image at the given message pixel, then the message bit is a "1" else "0." The sender can alter the "1" value pixels in such mode that there is no change in statistical properties of the image.

f. Cover Generation Techniques:

The Cover generation technique creates a digital object only for the purpose of being a cover for secret communication. This is achieved through Mimic Functions or Automated Generation of English Texts (CFG) [40].

4. CONCLUSION

The art of unseen communication has been discussed in brief which facilitate the study on steganography especially image steganography in an effectual manner. An organized survey has been done to identify the types of steganography: steganography in early days and steganography in digital era; Steganographic system with its requirements, the process, and the protocols used in communication. This paper also presents the various types of methods employed in the unseen communication.

5. REFERENCES

[1] Stefan Katzenbeisser and Fabien A. P. Petitcolas , "Information Hiding Techniques for Steganography and Digital Watermarking", Computer Security Series, Artech House, Boston, 2000

[2] K. Bennett, "Linguistic Steganography: Survey, Analysis, and Robustness Concerns for Hiding Information in Text", Purdue University, CERIAS Tech. Report, 2004

[3] M. K. Kaleem, "An Overview of various forms of Linguistic Steganography and their Applications in protecting data", Journal of Global Research in Computer Science, Volume 3, No. 5, May 2012

[4] Cerkez Paul S , “Automated Detection of Semagram-Laden Images”, Graduate School of Computer and Information Sciences, Nova Southeastern University, 2012

[5] Gustavus J. Simmons , “The prisoner's problem and the subliminal channel”, Sandia National Laboratories, Albuquerque, NM 87185

[6] B. Schneier, “Secrets & Lies; Digital Security in a Networked World”, John Wiley & Sons, 2000

[7] “Eliminating Steganography in Internet Traffic with Active Wardens” Gina Fisk, Mike Fisk, Christos Papadopoulos and Josh Neil, Los Alamos National Laboratory and University of Southern California

[8] Ronald L. Rivest , “Chaffing and Winnowing: Confidentiality without Encryption”, CryptoBytes Summer 1998 – The Technical News Letter of RSA Laboratories

[9] Amitabh Maurya, Pankaj Kumar Saini and Navnish Goel, “Chaffing and winnowing without using steganography and encryption technique”, International Journal of Information Technology and Knowledge Management, July-December 2011, Volume 4, No. 2, pp. 515-517

[10] Peter Wayner, "Mimic Function and Tractability”, Presented By: Quanzhong Li, November 8, 1999

[11] Scott Craver , "On Public-key Steganography in the Presence of an Active Warden", Intel Corporation, Microcomputer Research Labs, 2200 Mission College Blvd., Santa Clara, CA 95052-8119; Department of Mathematical Sciences, Northern Illinois University, DeKalb, IL 60115

[12] Neil F. Johnson and Sushil Jajodia , "Steganalysis: The Investigation of Hidden Information", 1998 IEEE Information Technology Conference Information Environment for the Future Cat No98EX228 (1998)

[13] Jonathan Cummins, Patrick Diskin, Samuel Lau and Robert Parlett , “Steganography And Digital Watermarking”, School of Computer Science, The University of Birmingham

[14] Zaidoon Kh. AL-Ani, A.A.Zaidan, B.B.Zaidan and Hamdan.O.Alanazi , "Overview: Main Fundamentals for Steganography", Journal Of Computing, Volume 2, Issue 3, March 2010, Issn 2151-9617

[15] Luis von Ahn and Nicholas J. Hopper, "Public-Key Steganography", Computer Science Dept, Carnegie Mellon University, Pittsburgh PA 15213 USA

[16] Mehdi Kharrazi1, Husrev T. Sencar, and Nasir Memon , "Image Steganography: Concepts and Practice", WSPC/Lecture Notes Series: 9in x 6in, April 22, 2004

[17] Kaliappan Gopalan and Stanley Wenndt , “Audio steganography for covert data transmission by imperceptible tone insertion”, Department of Engineering, Purdue University Calumet, Hammond, IN 46323 ; Multi-Sensor Exploitation Branch, AFRL/IFEC, Rome, NY 13441

[18] Debnath Bhattacharyya, Poulami Das, Samir Kumar Bandyopadhyay and Tai-hoon Kim , "Text Steganography: A Novel Approach", International Journal of Advanced Science and Technology, Vol. 3, February, 2009

[19] Bertrand Anckaert, Bjorn De Sutter and Koen De Bosschere , "Covert Communication through Executables", Electronics and Information Systems Department, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium

[20] Ravindra Gupta, Akanksha Jain, Gajendra Singh, "Combine use of Steganography and Visual Cryptography for Secured Data hiding in Computer Forensics", International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012,4366 - 4370

[21] Jasvinder Kaur,Manoj Duhan,Ashok Kumar,Raj Kumar Yadav , "Matrix Matching Method For Secret Communication Using Image Steganography", ANNALS OF FACULTY ENGINEERINGHUNEDOARA– InternationalJournalOf Engineering, 2012

[22] Jasvinder Kaur, Manoj Duhan,Ashok Kumar , "Digital Logic Embedding Using Single Row", International Journal on Computer Science and Engineering (IJCSE), Vol. 3 No. 12 December 2011

[23] Jessica Fridrich and Miroslav Goljan , "Digital image steganography using stochastic modulation", Department of Electrical and Computer Engineering, SUNY Binghamton, Binghamton, NY, 13902-6000, USA

[24] Dr. Natarajan Meghanathan, "Least Significant Bit (LSB)-based Steganography"

[25] Charles Kurak And John Mc Huges , “A Cautionary Note On Image Downgrading", IEEE 1992

[26] Jiri Fridrich & Du Rui , "Secure Steganographic Methods for Palette Images", Information Hiding Lecture Notes in Computer Science Volume 1768, 2000, pp 47-60

[27] J.M. Buhmann, D.W. Fellner, M.Held,J.Kettererand J. Puzicha, "Dithered Color Quantization", Wiley Online Library, Computer Graphics Forum, Volume 17, Issue 3, Article first published online: 25 DEC 2001

[28] Jiri Fridrich, "A New Steganographic Method for Palette-Based Images", in Proceedings of the IS&T PICS conference, 1999

[29] Xuefeng Wang, Zhen Yao & Chang-Tsun Li , "A Palette-Based Image Steganographic Method Using Colour Quantisation", IEEE 2005

[30] Abdelrahman Desok , "Noiseless Steganography: The Key to Covert Communications", CRC Press

[31] Ekta Walia,Payal Jain and Navdeep, “ An Analysis of LSB & DCT based Steganography”, Global Journal of Computer Science and Technology, P a g e | 4 Vol. 10 Issue 1 (Ver 1.0), April 2010

[32] Ken Cabeen and Peter Gent , "Image Compression and Discrete Cosine Transform"

[33] Matthew Warren Dodd , "Applications of the Discrete Fourier Transform in Information Theory and Cryptology", PhD Thesis, Royal Holloway and Bedford New College, University of London

[34] James L. Massey, “The Discrete Fourier Transform in Coding and Cryptography”, ITW 1998, San Diego, CA, February 8 – 11

[35] Robert T. McKeon, Member, Steganography using the Fourier Transform and Zero-Padding Aliasing Properties (May 2006), IEEE

[36] Po-Yueh Chen and Hung-Ju Lin , "A DWT Based Approach for Image Steganography", International Journal of Applied Science and Engineering, 2006. 4, 3: 275-290

[37] Frederick S. Brundick and Lisa M. Marvel, "Implementation of Spread Spectrum Image Steganography", ARMY Research Laboratory, March 2001

[38] Luis Perez-Freire, Pierre Moulin and Fernando Perez-Gonzalez, "Security of spread-spectrum-based data hiding", SPIE Proceedings Vol. 6505, February 2007

[39] Sharon Rose Govada, Bonu Satish Kumar, "Text Steganography with Multi level Shielding", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012

[40] Dr. Natarajan Meghanathan, "Information Hiding through Cover Generation: Using Context-Free Grammar"