ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 4, April 2015**

# Extracting Spread-Spectrum Hidden Data from Digital Media

Jyothi B[1]. Rameshkumar H K[2].

M.Tech Student, Dept. of Computer Science and Engineering, STJIT, Ranebennur , India

Assistant Professor, Dept. of Computer Science and Engineering, STJIT, Ranebennur , India

**ABSTRACT**: We consider the problem of extracting blindly data embedded over a wide band in a spectrum (transform) domain of a digital medium (image, audio, video). We develop a novel multi-carrier/signature iterative generalized least-squares (M-IGLS) core procedure to seek unknown data hidden in hosts via multi-carrier spread-spectrum embedding. Neither the original host nor the embedding carriers are assumed available. Experimental studies on images show that the developed algorithm can achieve recovery probability of error close to what may be attained with known embedding carriers and host autocorrelation matrix.

**KEYWORDS**: Authentication, annotation, blind detection, covert communications, data hiding, information hiding, spread-spectrum embedding, steganalysis, steganography, watermarking.

## I. INTRUDUCTION

DIGITAL data embedding in digital media is an information technology field of rapidly growing commercial as well as national security interest. Applications may vary from annotation, copyright-marking, and watermarking, to single stream media merging (text, audio, image) and covert communication. In annotation, secondary data are embedded into digital multimedia to provide a way to deliver side information for various purposes; copyright-marking may act as permanent "iron branding" to show ownership; fragile watermarking may be intended to detect future tampering; hidden low-probability to-detect (LPD) watermarking may serve as identification for confidential data validation or digital fingerprinting for tracing purposes. Covert communication or steganography, which literally means "covered writing" in Greek, is the process of hiding data under a cover medium (also referred to as host), such as image, video, or audio, to establish secret communication between trusting parties and conceal the existence of embedded data. As a general encompassing comment, different applications of information hiding, such as the ones identified above, require different satisfactory trade-offs between the following four basic attributes of data hiding:

(i) Payload - information delivery rate; (ii) robustness - hidden data resistance to noise/disturbance; (iii) transparency - low host distortion for concealment purposes; and (iv) security - inability by unauthorized users to detect/access the communication channel.

Recently, developing data embedding technologies are being seen to pose a threat to personal privacy, commercial, and national security interests. In this work, we focus our attention on the blind recovery of secret data hidden in medium hosts via multi-carrier/signature direct-sequence spread-spectrum (DS-SS) transform domain embedding. Neither the original host nor the embedding carriers (signatures or spreading sequences) are assumed known (fully blind data extraction). This blind hidden data extraction problem has also been referred to as "Watermarked content Only Attack" (WOA) in the watermarking security context.

While passive detection-only of the presence of embedded data is being intensively investigated in the past few years, active hidden data extraction is a relatively new branch of research. In blind extraction of SS embedded data, the unknown host acts as a source of interference/disturbance to the data to be recovered and, in a way, the problem parallels blind signal separation (BSS) applications as they arise in the fields of array processing, biomedical signal processing, and code-division multiple-access (CDMA) communication systems. Under the assumption that the embedded secret messages are independent identically distributed (i.i.d.) random sequences and independent to the cover host, independent component analysis (ICA) may be utilized to pursue hidden data

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 4, April 2015**

extraction. However, ICA-based BSS algorithms are not effective in the presence of correlated signal interference as is the case in SS multimedia embedding and degrade rapidly as the dimension of the carrier (signature) decreases relative to the message size. In, an iterative generalized least squares (IGLS) procedure was developed to blindly recover unknown messages hidden in image hosts via SS embedding. The algorithm has low complexity and strong recovery performance. However, the scheme is designed solely for single-carrier SS embedding where messages are hidden with one signature only and is not generalizable to the multicarrier case. Realistically, an embedded would favour multicarrier SS transform-domain embedding to increase security and/or payload rate.

## II.RELATED WORK

**1. M-IGLS Based Extracting Hidden Data from Digital Media**

Authors:Y.Singston Albert Dhas, D.Abisha

Data hiding and extraction schemes are increasing in today's communication world due to rapid increment of data tracking and tampering attacks. So we need an efficient and robust data hiding schemes to protect from these attacks. In this project the blindly extraction technique is considered. Blindly extraction means the original host and the embedding carriers are not need to be known. Here, the hidden data embedded to the host signal, via multicarrier SS embedding. The hidden data is extracted from the digital media like audio, video or image. The extraction algorithm used to extract the hidden data from digital media is Multicarrier Iterative Generalized Least Squares (M-IGLS).It is a low complexity algorithm and it attains the probability of error recovery equals to known host and embedding carriers. It's a peak signal to noise ratio value obtained is high.

**2.A Lossless Watermarking Based Authentication System For Medical Images**

Authors:Samia Boucherkha and Mohamed Benmohamed

In this paper we investigate the watermarking authentication when applied to medical imagery field. We first give an overview of watermarking technology by paying attention to fragile watermarking since it is the usual scheme for authentication.

We then analyse the requirements for image authentication and integrity in medical imagery, and we show finally that invertible schemes are the best suited for this particular field. A well-known authentication method is studied. This technique is then adapted here for interleaving patient information and message authentication code with medical images in a reversible manner, that is using lossless compression. The resulting scheme enables on a side the exact recovery of the original image that can be unambiguously authenticated, and on the other side, the patient information to be saved or transmitted in a confidential way. To ensure greater security the patient information is encrypted before being embedded into images.

**3.Modern Steganographic technique: A survey**

Authors:Pratap Chandra Mandal

Steganography is one of the methods of secret communication that hides the existence of hidden message. It can be defined as the study of invisible communication that usually deals with the ways of hiding the existence of the communicated message. The hidden message may be text, image, audio, video, etc. The files can be a cover image after inserting the message into the cover image using stego-key. It is referred to as stego-image. Steganography is now more important due to the exponential growth and secret communication of potential computer users on the internet. In this paper I have analysed various steganographic techniques. It also given an overview of steganography, different methods of steganography, its applications, how it is different from cryptography.

## III. EXISTING SYSTEM

Recently, developing data embedding technologies are being seen to pose a threat to personal privacy, commercial, and national security interests. In this work, we focus our attention on the blind recovery of secret data hidden in medium hosts via multi-carrier/signature direct-sequence spread-spectrum (DS-SS) transform domain embedding. Neither the original host nor the embedding carriers (signatures or spreading sequences) are assumed known (fully blind data extraction). This blind hidden data extraction problem has also been referred to as "Watermarked content Only Attack" (WOA) in the watermarking security context.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 4, April 2015**

In the existing system reversible data hiding technique the image is compressed and encrypted by using the encryption key and the data to hide is embedded in to the image by using the same encryption key. The user who knows the secret encryption key used can access the image and decrypt it after extracting or removing the data hidden in the image. After extracting the data hidden in the image then only can be the original image is retrieved.

Disadvantages:

- Less security.
- Provides only Image and File Steganography.
- Fewer users friendly.
- Does not Provides Compression and File Security.

## IV. PROPOSED SYSTEM

We propose the information hiding concept to reduce the risk of using cryptographic algorithms alone. Data hiding techniques embed information into another medium making it imperceptible to others, except for those that are meant to receive the hidden information and are aware of it presence. It focuses on methods of hidden data in which cryptographic algorithms are combined with the information hiding techniques to increase the security of transmitted data. We focus our attention on the blind recovery of secret data hidden in medium hosts via multi-carrier/signature direct-sequence spread-spectrum transform domain embedding.

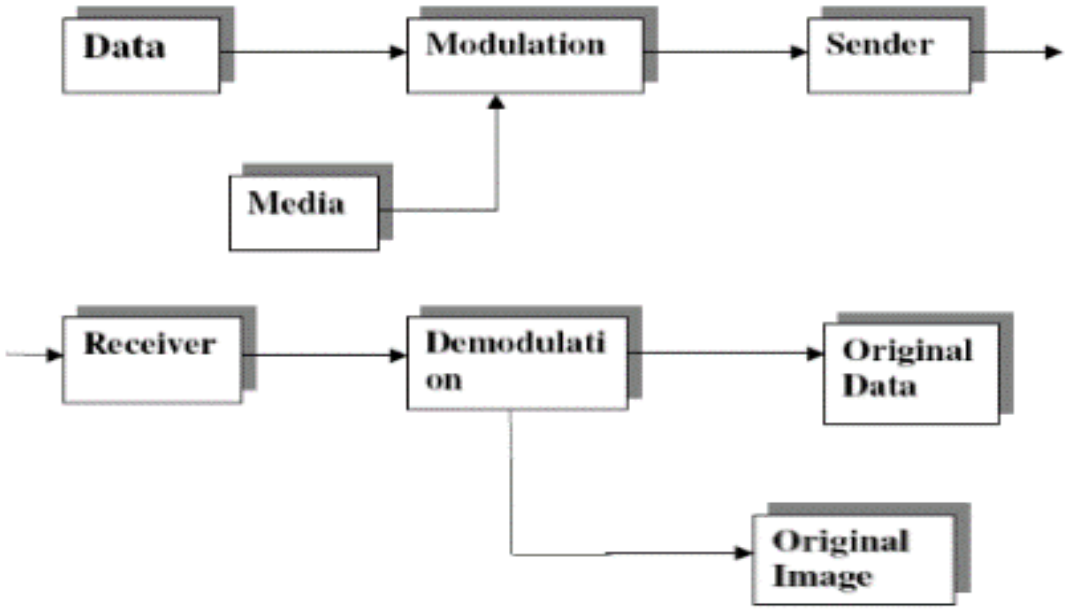


**Fig. : System architecture**

Advantages:

- Many options for users friendly with application.
- Provides Compression and File Security.
- Current Steganography Provides all Types of File Embed Process
- Steganography provides Encryption ,Compression Activities for file to increase security
- Current Steganography is High Security
- Steganography provides all types of Files
- A relatively good text file size.

## V. MODULE DESCRIPTION

Number of Modules:

After careful analysis the system has been identified to have the following modules:

1. Steganography
2. Multi-Carrier Spread Spectrum Embedding

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

3. Image encryption and watermarking
4. Image decryption and extraction

**1. Steganography**

Steganography includes the concealment of information within computer files. In digital steganography, electronic communications may include steganographic coding inside of a transport layer, such as a document file, image file, program or protocol Digital steganography can hide confidential data (i.e. secret files) very securely by embedding them into some media data called "vessel data." The vessel data is also referred to as "carrier, cover, or dummy data". In Steganography images used for vessel data. The embedding operation in practice is to replace the "complex areas" on the bit planes of the vessel image with the confidential data. The most important aspect of Steganography is that the embedding capacity is very large. For a 'normal' image, roughly 50% of the data might be replaceable with secret data before image degradation becomes apparent.

**2. Multi-Carrier Spread Spectrum Embedding:**

The technique of spread spectrum may allow partly to fulfill the above requirements. Advantages of spread spectrum techniques are widely known: Immunity against multi-path distortion, no need for frequency planning, high flexibility and variable data rate transmission. The capability of minimizing multiple access interference in direct-sequence code-division-multiple-access system is given by the cross-correlation properties of spreading codes. In the case of multi-path propagation the capability of distinguishing one component from others in the composite received signal is offered by the auto-correlation properties of the spreading codes.

**3. Image encryption and watermarking:**

The host image is an 8-bit or higher grey level image which must ideally be the same size as the plaintext image or else resized accordingly using the same proportions. Pre-conditioning the cipher and the convolution processes are undertaken using a Discrete Fourier Transform (DFT).

The output will include negative floating point numbers upon taking the real component of a complex array. The array must be rectified by adding the largest negative value in the output array to the same array before normalization.For color host images, the binary cipher text can be inserted into one or all of the RGB components.The binary plaintext image should have homogeneous margins to minimize the effects of ringing due to ‘edge effects’ when processing the data using Fourier transform.

Image decryption and extraction:
(i) The correlation operation should be undertaken using a DFT.
(ii) For color images, the data is decomposed into each RGB component and each 1-bit layer is extracted and correlated with the appropriate cipher.
(iii) The output obtained in Step 3 has a low dynamic range and therefore requires to be quantized into an 8-bit image based on floating point numbers within the range max (array)-min (array).

**4. Image decryption and extraction:**

(i) The correlation operation should be undertaken using a DFT.
(ii) For color images, the data is decomposed into each RGB component and each 1-bit layer is extracted and correlated with the appropriate cipher.
(iii) The output obtained in Step 3 has a low dynamic range and therefore requires to be quantized into an 8-bit image based on floating point numbers within the range max (array)-min (array).

## VI. EXPERIMENTAL RESULTS

The proposed method is to extract the hidden data from the digital media. Here blindly recovery of data is considered. That is the original host end embedding carrier is not need to be known. This method uses multicarrier embedding and DCT transformation for the embedding the data into the host image. The M-IGLS algorithm is used for the extraction purpose. This algorithm is a low complexity algorithm and it attains the probability of error

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 4, April 2015**

recovery equals to known host and embedding carriers. It is used as a tool to analyse the performance of the data hiding schemes.

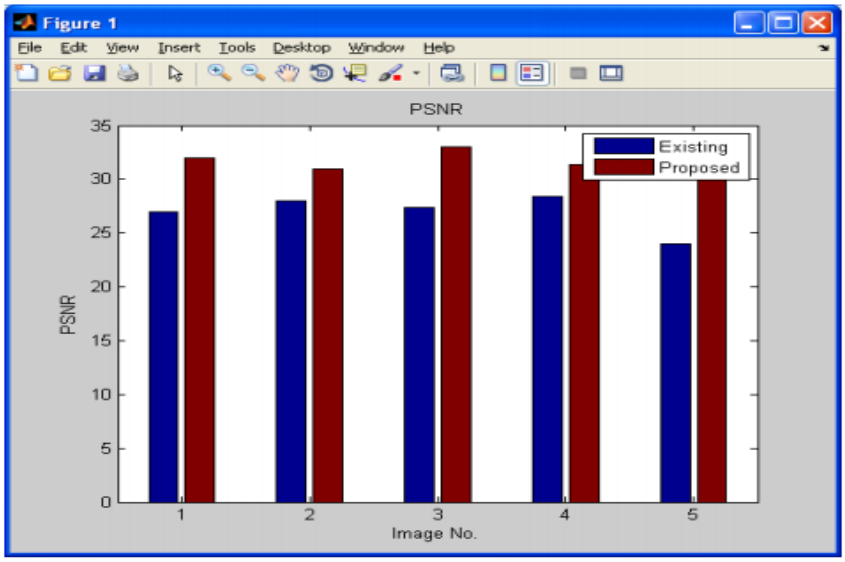


**Fig: Graph for PSNR verses image number**

In this performance graph peak signal-to-noise ratio,extract the hidden data from the digital media.PSNR is most commonly used to measure the quality of reconstruction of image compression. PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. For color images the image is converted to a different color space and PSNR is reported against each channel of that color space. Typical values for the PSNR in lossy image and video compression are between 30 and 50dB.This was obtained in our proposed system. For higher value the bit rate is to kept better.

## VII. CONCLUSION AND FUTURE WORK

We considered the problem of blindly extracting unknown messages hidden in image hosts via multi carrier/signature spread-spectrum embedding. Neither the original host nor the embedding carriers are assumed available. We developed a low complexity multi-carrier iterative generalized least-squares(M-IGLS) core algorithm. Experimental studies showed that M-IGLS can achieve probability of error rather close to what may be attained with known embedding signatures and known original host autocorrelation matrix and presents itself as an effective countermeasure to conventional SS data embedding/ hiding, This technique is enhanced by using harmony search algorithm where it provides low time consumption and high attack resistance.

## ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my Institution Sri Taralabalu Jagadguru Institute of Technology, Guide and Staff members for their continuous support for this survey and finally my friends for their coordination in this work

## REFERENCES

[1] A. Singh and L. Liu, "Sharoes: A data sharing platform for outsourced enterprise storage environments," in Proceedings of the 24th International Conference on Data Engineering, ICDE. IEEE, 2008, pp. 993–1002.
[2] 104th United States Congress, "Health Insurance Portability and Accountability Act of 1996 (HIPAA)," Online at http://aspe.hhs.gov/admnsimp/pl104191.htm, 1996.
[3] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in CCS '07: Proceedings of the 14th ACM Conference on Computer and Communications Security, New York, NY, USA, 2007, pp. 598–609.
[4] Y. Deswarte, J.-J. Quisquater, and A. Sa¨ıdane, "Remote integrity checking," in 6th Working Conference on Integrity and Internal Control in Information Systems (IICIS), S. J. L. Strous, Ed., 2003, pp. 1–11.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 4, April 2015**

[5] D. L. G. Filho and P. S. L. M. Barreto, "Demonstrating data possession and uncheatable data transfer," Cryptology ePrint Archive, Report 2006/150, 2006.
[6] E. Mykletun, M. Narasimha, and G. Tsudik, "Authentication and integrity in outsourced databases," Trans. Storage, vol. 2,no. 2, 2006.
[7] F. Seb´e, J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J.-J. Quisquater, "Efficient remote data possession checking in critical information infrastructures," IEEE Trans. on Knowl. and Data Eng., vol. 20, no. 8, 2008.
[8] M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage services honest," in HOTOS'07:Proceedings of the 11th USENIX workshop on Hot topics in operating systems, Berkeley, CA, USA, 2007, pp. 1–6.
[9] M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," Cryptology ePrint Archive, Report 2008/186, 2008.
[10] Y. Wang and P. Moulin, "Perfectly secure steganography: Capacity, error exponents, and code constructions," IEEE Trans. Inf. Theory, vol. 54, no. 6, pp. 2706–2722, Jun. 2008.
[11] F. Cayre, C. Fontaine, and T. Furon, "Watermarking security: Theory and practice," IEEE Trans. Signal Process., vol. 53, no. 10, pt. 2, pp. 3976–3987, Oct. 2005.
[12] L. Pérez-Freire, P. Comesana, J. R. Troncoso-Pastoriza, and F. Pérez-González, "Watermarking security: A survey," LNCS Trans. Data Hiding Multimedia Security, vol. 4300, pp. 41–72, Oct. 2006.
[13] M. Barni, F. Bartolini, and T. Furon, "A general framework for robust watermarking security," ACM J. Signal Process., Special Section: Security of Data Hiding Technologies, vol. 83, pp. 2069–2084, Oct. 2003.
[14] L. Pérez-Freire and F. Pérez-González, "Spread-spectrum watermarking security," IEEE Trans. Inf. Forensics Security, vol. 4, no. 1, pp. 2–24, Mar. 2009.
[15] S. Lyu and H. Farid, "Steganalysis using higher-order image statistics," IEEE Trans. Inf. Forensics Security, vol. 1, no. 1, pp. 111–119, Mar.2006.

## BIOGRAPHY

**Jyothi B** received Bachelor degree B.E from G.M Institute of Technology, Davangere Vishweshwaraya Technological University, Belgaum, Karnataka, India . She is now pursuing Masters Degree M.Tech Computer science and engineering department at Sri Taralabalu Jagadguru Institute of Technology, Karnataka, India.

**Rameshkumar H K Assistant Professor** Dept. of Computer science and Engineering,Sri Taralabalu Jagadguru Institute of Technology, Karnataka, India