International Journal on Advances in Engineering Technology and Science
Volume: 2, Issue: 4, July-August 2016, e-ISSN: 2455-3131

Hide Secret Massages in Audio File in Secure Wireless Communication using Embedded Algorithm

Akhilesh A. Waoo[1], Sanjay Sharma[2], Sumit Thakur[3]

[1]Research Scholar, Department of Computer Application, MANIT
MANIT, Bhopal, India
[1]akhileshwaoo@gmail.com

[2]Professor, Department of Computer Application, MANIT
MANIT, Bhopal, India
ssharma66@rediffmail.com

[3]AP, Bansal College of Engineering
Bhopal, India
sumit.thk@gmail.com

Abstract— **Now a day, rapid growth of Internet and its users increases range of data types which emphasize the security problem and also flood of multimedia contents in the structure of transmission of such data. The Encryption system for data is used for hiding data and messages by supervised trained dictionary. It involves hiding the secret text inside the cheating text. This cheating text is then hiding in Audio file. If this cheating text is intercepted by unauthorized person, the secret text may still be undetected by them. There are many reasons to hide data but most important is to put off unauthorized persons from becoming conscious of the existence of such a message in the network of communication. This hidden message is information which is not immediately noticeable which must be discovered or uncovered and interpreted before it can be known to anybody. This paper proposes an approach for generating cheating text and hiding cheating text in audio/sound file this high capacity audio algorithm based on the wavelet packet transform with adaptive hiding in LSB (Least Significant Bits). The adaptive hiding is determined depend on the cover signal strength and bits block matching between message and cover signals.**

Keywords—**Audio hiding wavelet packet transform, Data Hiding data, least significant bit. Adaptive hiding, Steganalysis**

I. INTRODUCTION

The Internet has revolutionized the modern world and the numerous Internet based applications that get introduced these days add to the high levels of comfort and connectivity in every aspects of human life. The expansion of data transmission through Internet made the process of improving data protection inevitable. The broadband Internet access has enabled a bigger flow of audio content through this network. When the data protection is based only on our system, it's often related to LSB technique. The Difference between the original images and the one with data embedded in still evident. This difference tends to be decreased, because of the signal for cryptanalyst's and hackers. Audio hiding file system is a useful means for transmitting convert battlefield information via and innocuous cover audio signal. Audio is an important communication way for people, and therefore is a convenient medium secure communications. In order to discriminate stegno audios from clear normal ones, that embed random data into a (possibly) stegno audio file by using a certain steganography tool. It was found that the variation in some statistical features of audio file is significantly different between clear audio files and stegno ones which already contain hidden messages embedded by the same tool. In this paper, that can detect the existence of hidden messages, and also identify the tools used to hide them.

The other Peer-to-Peer (P2P) audio services provide vast opportunities for covert communications by slightly altering the binary sequence of the audio signal with existing system tools, after converting communication channels may be relatively easy to establish. Moreover, the inherent redundancy in the audio signal and its transient the unpredictable characteristics imply a high hidden capacity [3]. This is further aided by the fact that the human ear is insensitive to small distortions in the audio signal. Information hiding in digital documents provides a means for overcoming those problems. Depending on what information in which form is hidden in the audio, one can distinguish at least two types of data hiding schemes: non-robust, undetectable data hiding, and robust audio watermarking. In the first case, a digital audio serves as a container for a secret message. In the second application, robust audio watermarking, a short message (a watermark) is embedded in the audio in a robust manner. By robustness we mean the ability to survive common audio

processing operations, such as lossy compression, filtering, noise adding, geometrical transformations, etc. Such robust watermark can be obviously used for copyright protection, fraud detection (verification of audio integrity), authentication, etc. At this point we emphasize that cryptographic authentication protocols cannot solve all the issues related to authentication.

II. LITERATURE SURVEY

Saswati Ghosh et. al. [8] uses double layered secure data transfer technique for cryptography and audio steganography for mobile network. This paper shows the characters of chipper text data are converted to bit values and are encrypted by using XOR (Ex-OR) operation using a symmetric key technique by using a secret key-box; it is again scrambled and then divided into 2-bit blocks. These blocks from MSB are replaced by the 2 LSB of each byte of cover audio bit stream. Saswati Ghosh et al used trick for this steganography method hides behind logic of selecting the next byte of the cover audio and additive technique with a key constraint is used as the proposed algorithm. Upon replacing by the secret bit blocks, the byte in the cover audio is break into two parts i.e. nibbles and added overlooking the carry and is converted to decimal values. Second and third bits are taken together with the decimal value which is considered as key constraint for this. These 2 values are added and next byte is chose after counting that numbers of byte positions. This paper presented technique which overcome many of the limitations like selection of chipper text size and cover audio format, noise removal etc. which supports the user to transfer data from mobile network in a more secured and efficient way.

Generally the wav files were used in the previous methods, and also the file size of the chipper text was limited. But, in our proposed algorithm, any kind of audio file can be used not only wav or mp3 as in the previous method. Audio steganography technique has the main advantage that it is very hard to detect the interferences at the time of hearing the stego audio file. In this, along with a strong cryptography technique, steganography technique is combined, to make this double-layered method a hard breakable one. According to the previous studies and proposed method provides an efficient and helpful technique for the users who want to send secret data over wireless network hiding from the eavesdroppers and man in middle.

In paper [17], Hardik K. Molia and Hardik A Gohel discussed about various attacks in computer networks for social sites. Also in paper [18], Hardik Gohel discussed about improving the same concept with keys and authentication tools to secure data. The unauthorized persons can use it for mal-activities, it is better to hide data to keep their communications secret and to coordinate attacks. In paper [9], Confused Document Encrypting (CDE) Scheme is a method used for data hiding. This involves concealing the secret text inside the cheating text. If the cheating text is captured, the secret text may still be undetected which focuses on reducing the amount of data transmission in delivering confused documents. They can use any piece of information on the internet as a cheating text. In this method, the sender only needs to transmit the encrypted URL (Uniform Resource Locator) to the receiver. Now receiver can follow the URL and download the cheating text. In this way, we avoids transmitting large amounts of cheating text, which is a major drawback of traditional confused document encrypting schemes. They proposed a new approach which can improve traditional confused document encrypting schemes by improvement using reduction of their transmission or broadcasting overhead, and thus make it suitable for wireless environment communication with low data rate.

To reduce the overhead of transmission of cheating text, the sender only need to send an URL in encrypted format, and the receiver can follow the URL to get the cheating text. No doubt, this method significantly reduces transmission requirements. Just for the sake of example, the sender can now only send a 60-byte URL, and the receiver can obtain a cheating text maximum of 30,000-40,000 bytes from that URL. This is very useful in wireless communication environment with low data rate transmission. Even a mobile phone using GPRS can easily use this approach to send out a secret and important message.

This study certainly decreases the transmission overhead, and reduces bandwidth consumption for mobile environment due to. This method demonstrates good performance with 50% overhead reduction for Chinese text in Big5 code. One of the interesting study can be based on performance by applying this method with difference character sets on various languages.

Haider Ismael Shahadi et. al. [10] in 2011, proposes a new high capacity audio/sound steganography algorithm based on the wavelet packet transform with adaptive hiding in LSB. They shows that message can be embedded up to 42 % of the total size of the cover audio signal with at least of 50 dB signal-to-noise ratio (SNR).

The following important stages are repeated to hide each secret message segment in one cover segment:

a. Cover Signal Decomposition and Preparing Stage - Every segment of the input audio cover signal is decomposed using L-levels of Haar DWPT to obtain 2L signals one represents the approximation coefficients signal and the others represent details coefficients signals. They separately produced signal with length of Z/2Lsamples. They select 2L-2 from the details signal starting from the highest frequency component for embedding of the secret message.
b. Key Generating and Secret Message Embedding Stage – Pre-processed Message segment (MP) that has size of MxN and the matrix of embedding positions contents that has size WxM are fed to the bits block matching process. In the process of bits block matching, the bits blocks of MP and EPC,

matrix are compared to compute matching between each bits block (row) of MP and whole blocks (rows) of EPC to obtain the blocks matching matrix BM having th size of MxW.

c. Stego-Key Embedding Stage- Because of the arbitrary distribution of the message blocks in embedding process in the previous section, the recovery algorithm of the proposed scheme will need stego key and message size to extract the message blocks from the stego-signal. Therefore, the stego key will embed with the message size in the lowest frequency details signal (D2 L -1). They choose this signal to embed the stego-key as it is having maximum power between all other details signals to make the stego-key more resistance against distortion or lost.

III. PROPOSED TECHNIQUE

In this paper text documents are considering, where each document consists of an ordered list of sentences [11], and each sentence consists of an ordered list of words. Each word treat as contained in the text as a Scrt_text associated with its tag Document, including noun (n.), verb (v.). For example, the word "are" contained in the text is depicted by (v.), where (v) is the Scrt_text of "are". Without loss of generality, here we create a dictionary and simplified the tag document for denoting a generalized word. The generalization relation between two words having the same tag document, which is a partial relation such that:

Let w1 = (Scrt_text 1|doc1) and w2 = (Scrt_text 2|doc2),

We have that w1 _ w2 implies Scrt_text1 = Scrt_text2 A vocabulary, denoted as V = {w1,w2, . . . ,wn}, is a collection of a limited number of distinct words.

A phrase is an ordered list of words, denoted as

s = w1w2 ……..wk.

A phrase can also contain generalized words. Within context of mining [12, 13], sequence patterns a word is an item and a phrase is a sequence.

A sentence is a grammatical complete phrase, denoted as s#. A document is a set of sentences. Paper not concentrates in the order of the context of sequence data mining though the document which is logically an ordered list of sentences. Moreover, in the same context, a document can be generalized to be a set of phrases.

Data-1"The actors in this film are all also very good. This is a good film without big budget sets. Very good sound, picture, and seats."

Example-1 Data-1 contains 3 sentences. If paper considers only the nouns, verb, contained in the data, Data-1 corresponds to a document Doc1

After apply cheating text algorithm:-

"The actresses in this film are all also very bad. This is a bad film without small budget sets. Very bad sound, picture, and seats."

On the other hand, importing the tags document into the data model it makes possible to focus on specific part of text [14], such as for building text class descriptors by nouns.

A. Encryption

Step 1: Use the cheating text to generate the Index of Words (IOW). In this type dictionary, each row stores a noun word and verb word, all word index of its occurrence in the cheating text [11].

Step 2: Using the IOW and the plaintext, paper can obtain a plaintext file. For each Word in plaintext, we look up the corresponding row in dictionary and choose a word from the row. The randomness here is significant, because it can improve the degree of safety.

Step 3: Encrypt the plaintext file with the public-key cryptography with any ciphering algorithm can be used, such as the International Data Encryption.

After generation of cheating text hiding data inside Audio files the technique usually used is low bit encoding which is somewhat similar to LSB that is generally used in Images. Spread Spectrum is another method used to conceal information inside of an audio file. This method works by adding random data to the signal the information is conceal inside a carrier and spread across the frequency spectrum. The thing that makes this method of concealing information inside of audio files better than other methods is that it can actually improve the sound of the audio inside an audio file.

There are some major audio hide algorithms are available like Low-bit encoding, Phase encoding, Spread spectrum coding and Echo data hiding.

a) Low-bit Encoding

Low-bit encoding [16], the binary version of the secret data message is substituted with the least significant bit (LSB) of each sample of the audio cover file. Though this method is simple and can be used to embed larger messages, the method cannot protect the hidden message from small modifications that can arise as a result of format conversion or lossy compression.

b) Phase Coding

Phase coding is based on the fact that the phase components of sound are not as perceptible to the human ear as noise [15]. Message bits are encoded as phase shifts in the phase spectrum of a digital signal.

This leads to inaudible encoding in terms of the Signal-to-Perceived Noise Ratio (SPNR) and the secret message gets camouflaged in the audio signal, not detectable by the Steganalysis methods based on SPNR.

B. Proposed Embedding algorithm

Get Plain Text Data (Tdx)

Generate Index of Words Ix

x= 1 2 3 …xn-1

Create grammatical phrase G(Ix (Tdx))

Generalized data in a set of phrases Gzn(Ix (Tdx))

Covert in Secret Message Sdx = Gzn(Ix (Tdx))

Segmentation Adx

Scaling and Converting

DWPT L-Level

Audio_bin=de2bi(Adx)

identity=[1 0 1 0 1 0 1 0]

Extract Secret Data

Embedded Secret Data in Audio File

Set in Bits Blocks using Positions Contents

Sdx (1:8)=bitset(Sdx (1:8+ Adx), identity(1:8))

Segment Collecting

Data Hidden in Audio Signal Adx = Sdx (1:8)

C. Proposed Extracting algorithm

Taken Data Hidden Audio Adx

Segmentation

Scaling and Converting

Audio_bin=de2bi(Adx)

identity=[1 0 1 0 1 0 1 0]

Bits Blocks base Matching

DWPT L-Level

Extract Secret Data

Sdx (1:8)=bitset(Sdx (1:8- Adx), identity(1:8))

Generalized data in a set of phrases

Sdx = Gzn(Ix (Tdx))

Generate the Index of Words

Create grammatical complete phrase

Tdx = G(Ix (Tdx))

Covert in Plain Data Tdx

The receiver is mandated to know the message length in order to use and extract the embedded message from the cover signal. A characteristic feature of phase coding is the low data transmission rate owing to the fact that the secret message is encoded only in the first segment of the audio signal. On the contrary, an increase in the length of the segment would have a ripple effect by altering the phase relations between the frequency components of the segment; thereby making detection easier.

IV. TEST RESULTS AND ANALYSIS

The proposed in algorithm uses a dictionary to store the word, so in worst case, the size of dictionary may be large as the original data. We tested ten (10) files in our program, and the results are shown in Table 1.

In our experiment here we embedding data in audio file only and after send embed file through email or any electronic media to destination and after receiving same data file they extract data from audio file and use the information.

TABLE I
RESPECT TO RESPONSE TIME AND USER LOAD

S. No	Size Byte	Time in Second	
		Embedding data	**Extracting data**
1	36	1.01	0.98
2	67	1.23	1.02
3	90	1.45	1.12
4	160	1.9	1.32
5	234	2.32	1.45
6	453	2.54	1.65
7	512	2.67	1.69
8	589	2.72	1.71
9	687	2.89	1.89
10	784	3.1	1.92

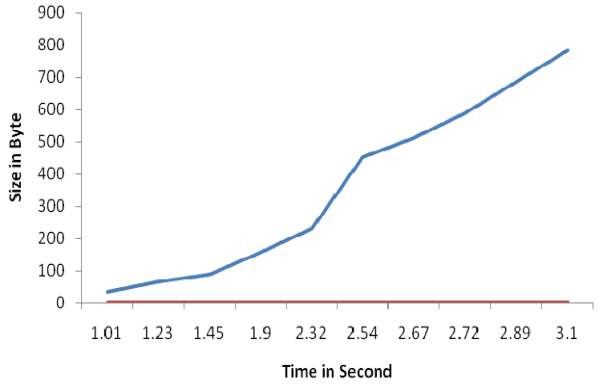

Fig 1. Shows graphical results of embedding data with respect to response time and user load.

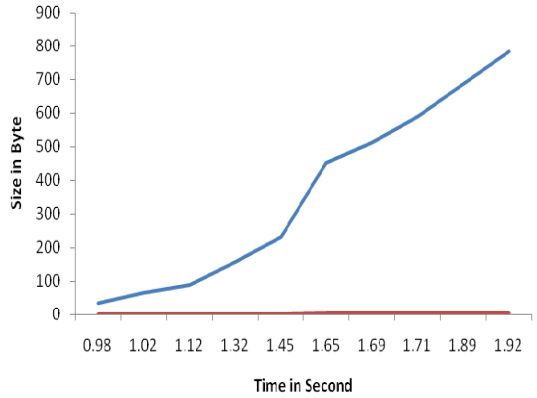

Fig 2. Shows graphical results of extracting data with respect to response time and user load.

V. DISCUSSION

Here we also encrypt textual data in cipher text because of many researchers have made a great effort on developing cryptosystems, which is difficult and complicated to reverse from the cipher text into the original data. However, cipher text usually seems meaningless as it is encrypted. When a hacker intercepts this meaningless data [4], he knows that it has been encrypted. In other words, as soon as he gets the encrypted data, he knows that he had obtained an important message. There is an old saying "Trying to cover up a misdeed, one only makes it more conspicuous". Therefore, a better approach to protect the message may suggest us to choose a different technique for concealing Data Hiding and Encrypting Scheme.

Any attempt to authenticate the digital audio file by appending information will fail. The embedded information will be transparent to the human ear, but it should be detectable using a sophisticated algorithm provided a secret key is available [5] and [6] and [7]. There have been some improvements by altering the modified (embedded) picture with a substitution matrix [1, 2]. The unwanted effect is increased the complexity of computational system.

Strengthening the proposed solutions with an existing cryptographic algorithm Advanced Encryption Standard can partially improve this unwanted effect. Because AES has a K level security from the aspect of cryptography they offered increase in security of data transfer with the combination of the AES algorithm and steganography [1, 2], The process of receiving the message and the removal of the protection is, however, equally important. Looking at the complexity and implementation, these processes are considered symmetric.

Cryptographic authentication deals with authenticating the sender of the message over insecure channels. However, once the message (audio) is decrypted, the audio is unprotected and can be copied and further distributed. Unlike classical paintings that can be studied for authenticity using sophisticated experimental techniques, a digital artwork is just a collection of bits. A visible signature in the corner of the audio can be easily replaced or removed with advanced audio processing software packages, such as Photo-Shop. Additional information in the audio file header can be erased or changed as well.

VI. CONCLUSION

In our research paper we proposed algorithm for secure transferring of data. Here we took plain data than convert it in cheating data after hiding data in audio signal to hide data. There are many reasons to hide data in audio signal but important is to avoid access of unauthorized persons from becoming aware of the existence of a original message. The results produces arbitrary, of the block matching generate an random key for embedding process, and cheating data increasing the security of massage information in the proposed algorithm. Audio data hiding can also be used in corporate world.

In our research we test in 10 different size of file and results showed that the algorithm has a better performance. As these tests demonstrate, authentication schemes and our algorithms carry varying amounts of overhead, and therefore have vastly different performance characteristics. Hiding data can also be used in the non-commercial field to hide information that keeps data private. Terrorists can also use data hiding to keep their communications secret and to coordinate attacks. Another advantage for the proposed algorithm is the reconstruction of the actual secret messages does not require the original cover audio signal and that's why the cover signal can be any recorded audio by the hiding side. We implemented our algorithm using a Mat lab tools.

REFERENCES

[1] A. Brazil, Path Relinking and Aes, "Cryptography in Color Image Steganography," M. Sc. Dissertation, Computer Institute, UFF (available at http://www.ic.uff.br/PosGraduacao/lista_dissertacao.php ?ano=2008, 2008.

[2] A. L. Brazil, A Sanchez, A. Conci, N. Behlilovic, "An Hybrid of Genetic and Path Relinking Algorithms for Steganography," Proceedings of 53st International Symposium ELMAR, pp. 285, Zadar, Croatia, 2011.

[3] Er. Niranjan Singh and Dr. Bhupendra Verma, "Quality and Distortion Evaluation of Audio Signal by Spectrum," *International Journal of Computer Science and Security (IJCSS)*, Volume (6): Issue (1), PP 629-636, 2012.

[4] F. Cayre, O. Devillers, F. Schmitt, and H. Maıˆtre, "Watermarking Triangle Meshes for Authentication and Integrity," INRIA Research Report RR-5223, June 2004.

[5] M. Nosrati, R. Karimi and M. Hariri, "An introduction to steganography methods", World Applied Programming, Vol (1), No (3), pp.191-195, August, 2011.

[6] M. Salem Atoum, M. Suleiman A. Rababaa, Dr. S. Ibrahim, O. A. Ahmed, "A Steganography Method Based on Hiding secrete data in MPEG/Audio Layer III", IJCSNS *International Journal of Computer Science and Network Security*, VOL.11 No.5, May, pp 184-188, 2011.

[7] M. L. Mat Kiah, B. B. Zaidan, A. A. Zaidan, A. M. Ahmed1 and S. Hasan A. bakri, "A review of audio based steganography and digital watermarking", *International Journal of the Physical Sciences* Vol. 6(16), pp. 3837-3850, 18 August, 2011.

[8] S. Ghosh, D. De and D. Kandar, "A Double Layered Additive Space Sequenced Audio Steganography Technique for Mobile Network", International Conference on Radar, Communication and Computing (ICRCC), SKP Engineering College, Tiruvannamalai, TN., India. 21 - 22 December, 2012. Pp .29-33. 978-1-4673-2758-9/12/$31.00 ©2012 IEEE, 2012.

[9] T. Yao and Q. Wu, "On the Study of Overhead Reduction for Confused Document Encrypting Schemes," MelT 2010, 201 0 IEEE.

[10] Haider Ismael Shahadi, Razali Jidin, "High Capacity and Inaudibility Audio Steganography Scheme", 7th International Conference on Information Assurance and Security (IAS), IEEE 2011.

[11] H. Schmid. "Probabilistic Part-of-Speech tagging using decision trees," In NeMLaP, 1994.

[12] R. Agrawal and R. Srikant, "Mining sequential patterns. In ICDE," pages 3–14, 1995.

[13] S. Jaillet, A. Laurent, and M. Teisseire, "Sequential patterns for text categorization," *Intelligent Data Analysis Journal*, 10(3):199–214, 2006.

[14] R Weis and S Lucks, "Cryptographic Hash Functions-Recent Results on Cryptanalysis and their Implications on System Security," 5th System Administration and Network Engineering Conference, pp 15-19, 2006.

[15] W. Bender, D. Gruhl and N. Morimoto, "Techniques for Data Hiding," *IBM Systems Journal*, vol. 35, no. 3, pp. 313 – 336, 1996.

[16] R. Sridevi, A. Damodaram and S.V.L. Narasimham, "Efficient Method of Audio Steganography by Modified LSB Algorithm and Strong Encryption Key with Enhanced Security," *Journal of Theoretical and Applied Information Technology*, vol. 5, no. 6, pp. 768 – 771, June 2009.

[17] H. K. Molia, H. A Gohe, "Protection of Computer Networks from the Social Engineering Attacks", *International Journal on Advances in Engineering, Technology and Science (IJAETS),* Volume: 1, Issue: 1, pp. 1-5, October 2015.

[18] H. Gohel, "Intelligent Web based Secure Browsing Implementation," *International Journal on Advances in Engineering, Technology and Science (IJAETS),* Volume: 1, Issue: 1, pp. 14-16, October 2015.