Research Journal of Applied Sciences, Engineering and Technology 4(14): 2255-2258, 2012
ISSN: 2040-7467

Submitted: March 15, 2012 Accepted: April 10, 2012 Published: July 15, 2012

Information Security using Audio Steganography -A Survey

B. Santhi, G. Radhika and S. Ruthra Reka
Department of Information and Communication Technology, SASTRA University, India

Abstract: The most important application of internet is data transmission. Unfortunately this is less secured because of advanced hacking technologies. So, for secured data transmission we make use of steganography. This is the art of hiding information where the existence of data is unknown. Any medium like music, video, text, speech, etc can be used. In this study, the selected medium is audio. This study discusses about the existing audio steganographic techniques along with their advantages and limitations. Also an algorithm implementing parity and LSB methods is proposed. This mitigates the limitations of the existing methods discussed, thus increasing security and reducing computational load and code complexity.

Key words: Cryptography, human auditory system, LSB, parity coding, phase shifting, temporal domain

INTRODUCTION

Audio steganography is the technique of hiding information inside an audio signal. As data is embedded in the signal, it gets modified. This modification should be made imperceptible to the human ear. Image can also be taken as a medium but audio steganography is more challenging because of the characteristics of Human Auditory System (HAS) like large power, dynamic range of hearing and large range of audible frequency.

Cryptography involves the encryption of message. It makes no attempt to hide the encrypted message. In steganography the original message is not altered but the very existence is hidden from the intruder by embedding the message in the selected medium.

Audio steganography:

Cover signal + Target data = Stego signal→ (Transmitted)

Cryptography in steganography:

Cover signal + Encrypted data = Stego signal

where,

Message data + Encryption Key = Encrypted data

So, audio steganography is preferable than cryptography because it is more confidential as it hides the existence of message. For providing more security we can incorporate cryptography along with steganography. There are several methods for audio steganography in literature. This study was developed with a motive of analyzing many steganographic techniques in use and determining their strengths and limitations. A new technique is proposed considering such existing methods and providing a better algorithm in terms of simplicity, computational load and security.

EXISTING METHODOLOGY

A very popular audio steganography method is the LSB (Least Significant Bit) algorithm. In this, the least significant bit of the cover signal is used to hide the message. This is a very simple method. The changes made to the LSB bit will not be reflected in the final stego signal. Being a simple technique, a very high level of security is not achieved. Modifications are done to the existing LSB method to improve security. Few of such techniques are explained in this study. Apart from security, certain other parameters like time complexity, computational load, SNR (Signal to Noise Ratio), BER (Bit Error Rate), efficiency, etc are to be considered for Audio Steganography.

LSB possesses advantages like low computational load and simplicity yet lacks in security. Hence Enhanced Audio Steganography (EAS) with additional layers of encryption and decryption can be implemented. Here before encoding the data, it is encrypted. The main features of EAS are the size of the file is not changed after encoding and there is no software available to determine the sound variation caused by bit level manipulations. By Sridevi *et al*. (2005), a solution is suggested to both of these issues thus supporting different audio formats and reducing the time for encoding and decoding are discussed. Data is embedded in such a way that each character requires eight 254/255 bytes. Frequency chart analysis shows that the efficiency of the algorithm in terms of security is the advantage of this method and the

Corresponding Author: B. Santhi, Department of Information and Communication Technology, SASTRA University, India

constraint is that sound quality depends upon the size of the audio selected and length of message embedded. This also supports any audio format.

Kekre *et al*. (2010), proposes two methods that are used with LSB. First method is parity coding and the other is XORing of LSB. Initially the cover audio signal is read and one should ensure that the size of the message to be embedded is less than the cover audio signal. In parity method, the parity bit is considered before directly replacing the LSB. Depending upon the message bit to be embedded, the LSB is either flipped or retained. If the message bit is 0, LSB has to be modified in such a way that parity of the sample is even. Else if the message bit is 1, LSB is modified in such a way that parity of the sample is odd. In second method, XORed operation between the LSB and the next bit has to be equivalent to the message bit to be embedded. If equal, the LSB is retained, else flipped. Also they reduce the computational load and the capacity of the cover audio is increased. From experimental results it is found that the encryption with steganography provides better security.

Embedding data in the higher LSB layers is prone to less attack than those embedded in the lower layers. But embedding in higher LSB will result in distortion. Therefore further steps have to be included to reduce these distortions. Following few papers handles these concepts.

Nedeljko and Tapio (2005) proposes a method that will enable to embed data extending it from the fourth to sixth LSB layers with minimum distortions in two steps. Watermarked bit is embedded in any higher LSB layer using novel LSB method in the first step and second step changes white noise properties by shaping the impulse noise which is caused by the embedding bit. The stated algorithm by Nedeljko and Tapio (2005) involves various cases. Depending upon the message bit to embedded, the particular case is chosen. Experimental results shows that SNR value is around 8QS (Quantization Step) in the standard algorithm which uses the 4^{th} LSB and in the proposed algorithm from 1-4QS. Average power of the introduced noise is 9.31 dB lesser than the conventional method. When the data to be embedded is watermarked, the error rate is much reduced. The BER is also very less compared to the standard method. The adversary cannot exactly detect the bit layer where data is embedded. But this in turn makes the algorithm much more complex.

The algorithm explained by Mazdak *et al*. (2009), is a robust steganographic method that embeds the data in the multiple, vague and higher LSB layers. Generally there are two types of attacks namely unintentional attacks like re-sampling, re-quantization, lossy compression, etc and intentional attacks like cropping, recycling, re scaling, white colored noise, etc. By Mazdak *et al*. (2009), solutions are suggested for both these type of attacks. The intentional attack like revealing the hidden message and unintentional attack like distortions with high average power are dealt. Generally messages are embedded in the first LSB bit. In order to provide a solution for the intentional attack, the data bits are embedded in the bit other than 1^{st} LSB bit. As a solution to the second problem, the bits other than the selected bits for embedding are altered reducing the distortion. These solutions are provided in the following approaches: Alteration, Modification, Verification and Reconstruction. Intelligent algorithm that either uses Genetic algorithm or symbolic AI system is essential for bit alteration and to embed the message in the deeper layers, making the code very complex. The pros of this study are robustness and high capacity.

Even though we perform embedding in different deep layers, to confine ourselves to fixed deep layers we go for 4^{th} and 1^{st} layer. The results by Samir and Biswajita (2011) experimentally proved to yield a stego signal that does not differ much from cover audio signal. To ensure no information loss we check with a string length embedded in a first 8 samples. Message is embedded in the samples selected by the small logic as follows: consider the prime number next to the string length and determine all the prime numbers following it upto the number of characters to be hidden. Each character requires three samples.

To embed the message, we convert it to 7 bits binary equivalent and remove the MSB making it 6 bits thus paring 2 bits per sample. In each pair MSB is selected and embedded in the 4^{th} LSB of the cover signal and certain changes are made to the other bits to reduce distortion, which follows the same logic specified by Nedeljko and Tapio (2005). The other bit of the pair is embedded in the LSB after this. At the receiver end, MSB (7^{th} LSB layer) is chosen as 1 for upper case letters and 0 for numbers and special characters. Distortion is experimentally proved to be very less. Capacity and robustness is found to increase when 4^{th} and 1^{st} LSB layers are used. The extraction algorithm is also very simple but this study is defined on .wav file format and supports mainly upper case characters.

LSB is combined with other techniques in different domains like temporal, cepstral and transform domain. Ahmad and Mohammad (2009) selected temporal domain. This domain reduces computational load and supports fast implementation. We calculate the hearing threshold in the temporal domain which is exploited as the embedding threshold, yielding more capacity compared to uniform embedding pattern. Proposed method uses compression of information using lossless compressor, thus increasing total bit rate. With the key to pseudo random number generator, the compressed data is encrypted. Embedding threshold is calculated based on the cover signal.

Thresholds are assigned according to the samples magnitude. Lower magnitudes have less embedding capacity than the higher ones. SNR comparison of audio and stego through experimental and theoretical approaches are equal. BER calculated for pop, classic, country and speech resulted zero, thus efficient. In presence of slight distortions, this method cannot achieve full recovery.

Gopalan (2004) proved a capacity of 62 bps for embedding. BER is less than 2 to 2.5 for clean cover and noisy speech. As data can be successfully embedded in the spectral points, a technique was used in which first step is done by extending it to the log spectral domain in perceptually masked regions is proposed. In this study, every frame of speech sample is represented as convolution between Excitation Source Signal and Vocal Tract System (VTS). This convolution is converted to addition and Inverse Fourier Transform is applied to this complex log spectrum. This converts vocal tract model to lower indices and excitation to higher indices or frequencies in cepstral domain.

In order to produce distortion free stego signal, data is embedded in the lower range of frequencies by modifying mean removed cepstrum. Frame synchronization and robustness can be achieved by modifying the mean cepstrum values near the rising energy points. Certain pre-processing steps are required for extension of spectral to log spectral domain. Cepstrum modified frame is quantized to 16 bits for transmission after its conversion to time domain. Modification with the VTS affects only the lower or fundamental frequencies. Frame with voiced/unvoiced transitions causes errors when tried to modify. Even in noise audio signal with the frequency samples of 8 KHz and frequency range of 51 to 150, stego is imperceptible to host. Same samples with higher frequencies slight distortions were observed. Considering cepstral domain, real cepstrum has application like speaker identification and speech analysis, complex cepstrum helps in obtaining cepstrum-modified speech by embedding and watermarking.

The already discussed domains were temporal and cepstral. Ahmad and Mohammad (2008) selected Integer Wavelet Transform domain. Among different types of wavelet transform, Haar Discrete Wavelet Transform is used. This result in wavelet co-efficient which converted to integer, are used to calculate the embedding threshold. Based on the threshold, the encrypted data is embedded. Now the signal is converted to wavelet domain and Inverse Wavelet Discrete Transform is applied before transmission. At the receiver end the reverse process followed by data extraction using the secret key. The secret key is given as input to the PN (Pseudo Random Number) Generator which arranges the covert data and according to the threshold, the data packet is formed. SNR is one of the important parameters determining the efficiency of the algorithm. It is also proved that the SNR values for theoretical and experimental results are equal. Hence without any actual embedding effort, SNR is calculated. By this method, SNR value is about 40 dB which is higher than other methods. They also has increased payload, full recovery with better robustness.

Gurvinder *et al*. (2009), proposed algorithm for both image and audio steganography. Regarding audio steganography, it states the technique of producing an echo signal from the original signal. Data is embedded in the echo signal varying its parameters like decay rate, offset and amplitude. These parameters should be set below human audible range for imperceptibility. The original signal is segmented into blocks and each block is given the value 0 or 1 depending upon the secret message. The original signal is echoed and the message is embedded into it. This is then given to the mixer, the output of which is fed to the encoder. At the receiver end, auto correlation and decoding is done to separate the secret signal and the original signal. Further development of this echo hiding is Time Spread Echo Hiding. Here several echoes are produced and are stretched in time from which only certain echoes are selected for embedding data based on the PN generator.

The main idea proposed by Samir *et al*. (2011), is based on psycho acoustic theory of persistence and phase shifting. Persistence of hearing is based on the fact that two sounds successively with a difference of less than one-tenth of a second hit our ears, then the difference between the sounds is imperceptible. $\varnothing$ is called the phase shift, the change of which is same as the shift in time. Samir *et al*. (2011), used uncompressed audio format (WAV format). Consider S1, T1 and S2, T2 as the sampling rate and total duration of the cover and target file. The target file is split after every $\varnothing$ seconds, $\varnothing$ varying between 0 and 1/10. Hence the total samples is equal to S2 * $\varnothing$ in target audio file. Phase shift and cover audio will be ($\varnothing$ * S2)/S1. In cover audio file, data is embedded at the interval of (T1*$\varnothing$)/T2 seconds. Regarding the computational complexity of time, it is O(max(m, n)) for both sender and receiver where m and n are number of samples in the cover and target audio file respectively. Steganalysis results prove that the insertion is quite undetectable and capacity is large.

Few limitations that are generally observed in this study are: Certain algorithms are very time consuming and have complex implementation, they don't support all audio formats. Even though full recovery and security are very essential and possible, there are some algorithms that cannot achieve it. As each and every algorithm has its own advantages and May 28, 2012limitations, they are

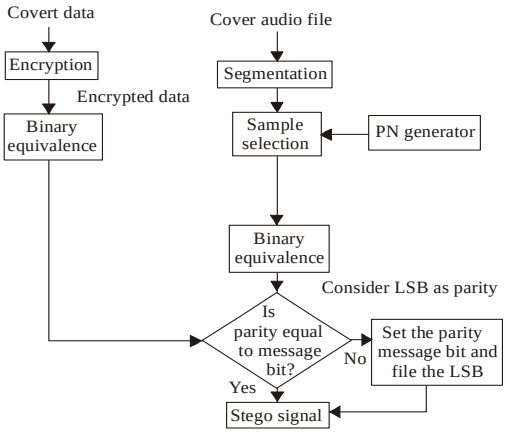

Fig. 1: Work flow diagram of proposed algorithm

chosen depending upon the parameters required. This study focuses the limitations in security and proposes the algorithm in the following section.

PROPOSED ALGORITHM

Based on the above study, the simple method for audio steganography is selected. To improve the efficiency and robustness this study proposes the novel technique that slightly modify the LSB technique. Parity coding is included. Figure 1 depicts the workflow of this proposed algorithm.

Proposed Algorithm uses even parity for detection. Data bit is embedded in LSB or higher layers based on the sample position. The function selected for determining the sample position has to be shared between the sender and the receiver. This method may expect a better level of security using these modifications. In future, the practical implementation of this algorithm will be done and performance analysis may be discussed in detail.

CONCLUSION

Thus this study analyses various audio steganographic techniques and determine their advantages and limitations. Also this study concludes that each technique is efficient depending upon the required parameter. Finally a solution is proposed for the security problem, the main parameter of steganography.

REFERENCES

Ahmad, D. and P. Mohammad, 2008. Adaptive digital audio steganography based on integer wavelet transform. Circ. Syst. Signal Process., 27: 247-2590.

Ahmad, D. and P. Mohammad 2009. Adaptive and Efficient Audio DataHiding Method in Temporal Domain. International Conference on Information and Communication Systems.

Gopalan K., 2004. Cepstral Domain Modification of Audio Signals for Data Embedding-Preliminary Results. Proceeding of 16th Annual Symposium on Electronic Imaging--Security, Steganography and Watermarking of Multimedia Contents VI, San Jose, CA, January.

Gurvinder, S., S.K. Dey, S. Dubey and S. Katiyal, 2009. Increasing the efficiency of Echo Hiding Digital Audio Steganography. 4th National Conference, INDIACom-2010 Computing for National Development, Bharati Vidya peeth's Institute of Computer Applications and Management, New Delhi.

Kekre, H.B., A. Archana R. Swarnalata and A. Uttara, 2010. Information hiding in audio signals. Int. J. Comput. Appl., 7(9).

Mazdak, Z., A.M. Azizah, B.A. Rabiah, M.Z. Akram and A.., Shahidan, 2009. A genetic-algorithm-based approach for audio steganography. World Acad. Sci. Eng., Technol., 52: 360-363.

Nedeljko, C. and S.A. Tapio, 2005. Increasing robustness of lsb audio steganography by reduced distortion lsb coding. J. Universal Comput. Sci. 11(1): 56-65.

Samir, K.B. Tuhin, U.P. and Ra., Avishek, 2011. A robust audio steganographic technique based on phase shifting and psycho-acoustic persistence of human hearing ability. Int. J. Comput. Corporate Res., 1(1).

Sami, K.B. and D. Biswajita, 2011. Higher LSB layer based audio steganography technique, IJECT, 2(4).

Sridevi, R., A. Damodaram and S.V.L. Narasimham, 2005. Efficient method of audio steganography by modified LSB algorithm and strong encryption key with enhanced security. J. Theor. Appl. Inf. Technol., 769-771.