A view on latest audio steganography techniques

Article · April 2011

DOI: 10.1109/INNOVATIONS.2011.5893859

CITATIONS
47

READS
1,118

4 authors, including:

Fatiha Djebbar
United Arab Emirates University
17 PUBLICATIONS **203** CITATIONS
SEE PROFILE

Habib Hamam
Université de Moncton
199 PUBLICATIONS **1,269** CITATIONS
SEE PROFILE

Karim abed-meraim
Université d'Orléans
413 PUBLICATIONS **8,871** CITATIONS
SEE PROFILE

A view on latest audio steganography techniques

Fatiha Djebbar*, Beghdad Ayad†, Habib Hamam‡and Karim Abed-Meraim§
*Université de Bretagne Occidentale, Brest, France and UAE University, UAE
Email: fdjebbar@uaeu.ac.ae
†Al Ain University, Al Ain, UAE
‡Faculty of Engineering Université de Moncton, Moncton, NB, Canada
§Telecom ParisTech, Paris, France

***Abstract*—Steganography has been proposed as a new alternative technique to enforce data security. Lately, novel and versatile audio steganographic methods have been proposed. A perfect audio Steganographic technique aim at embedding data in an imperceptible, robust and secure way and then extracting it by authorized people. Hence, up to date the main challenge in digital audio steganography is to obtain robust high capacity steganographic systems. Leaning towards designing a system that ensures high capacity or robustness and security of embedded data has led to great diversity in the existing steganographic techniques. In this paper, we present a current state of art literature in digital audio steganographic techniques. We explore their potentials and limitations to ensure secure communication. A comparison and an evaluation for the reviewed techniques is also presented in this paper.**

***Index Terms*—Digital data security, audio steganography, information hiding.**

I. Introduction

The proliferation of digital data in their various formats has attracted a special interest from researchers to ensure their security. Techniques such as encryption and watermarking are already used in this regard. However, the need for new techniques and new algorithms to counter constantly-changing malicious attempts to the integrity of digital data has become a necessity in today's digital era. Steganography, which literary means "covered writing" has drawn more attention in the last few years. Its primary goal is to hide the fact that a communication is taking place between two parties. The sender embeds secret data of any type using a key in a digital cover file to produce a stego file, in such a way that an observer cannot detect the existence of the hidden message [1]. At the other end, the receiver processes the received stego-file to extract the hidden message. An example of audio steganography is depicted in Fig. 1 where the cover file being used is a digital audio signal. An obvious application is a covert communication using innocuous cover audio signal, such as telephone or video conference conversations.

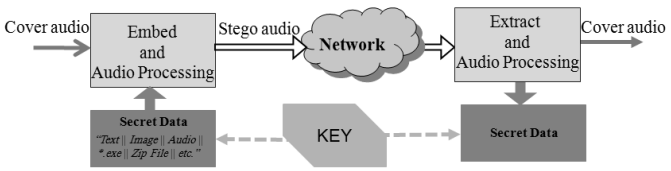

Fig. 1: Blocks diagram for audio steganography.

To minimize the difference between the original medium and the one obtained after embedding the hidden data, recent steganography techniques benefit from the natural limitations in the auditory and visual perceptions of human in one hand, and on the other hand from the properties of digital media through utilizing them as a cover to vehicle secret communications. Image and video based steganography relies on the limited human visual system in remarking luminance variation at levels greater than 1 in 240 in uniform gray levels or 1 in 30 of random patterns [1]. However, audio-based steganography exploits the masking effect property of Human Auditory System (HAS) [2] as explained later in this paper.

Various features influence the quality of audio steganographic methods. The importance and the impact of each feature depends on the application and the transmission environment. The most important properties include robustness to noise and to signal manipulation, security and hiding-capacity of embedded data. Robustness requirement is tightly related to the application and is the most challenging to satisfy in a steganographic system. In addition, there is a tradeoff between robustness and hiding-capacity. Generally, they hardly coexist in the same steganographic system.

In this review, the use of audio files as a cover medium to vehicle secret communications is thoroughly investigated. Several works in audio steganography are discussed in this paper. The reminder of this paper is organized as follows: Latest audio steganography techniques and their evaluation are presented in Sections II, III, IV, and V respectively. Finally, conclusions and future work are presented in Section VI.

II. Temporal domain

A. Low-bit encoding

Also known as LSB (Least Significant Bit), it is one of the earliest methods used for information hiding [1]. It consists in embedding each bit from the message in the least significant bit of the cover audio in a deterministic way Fig. 2. Thus, for a 16 kHz sampled audio, 16 kbps of data is embedded. The LSB method allows high embedding capacity for data and is relatively easy to implement or to combine with other hiding techniques. However, this technique is characterized by low robustness to noise addition and thus by low security as well since it is very vulnerable even to simple attacks. Filtering, amplifying, noise addition or lossy compression of the stego-audio will very likely destroy the data. Furthermore,

an attacker can easily uncover the message by just removing the entire LSB plane. In [3], a simple LSB strategy has been applied to embed a voice message in wireless communication. However, in an attempt to increase the hiding capacity while minimizing the error on the stego audio, [4] adopted minimum error replacement method while embedding four bits per sample. The embedding error is then diffused on the next four samples. Hidden data channel's capacity with the latter method has reached 176.3 kbps in 44.1 kHz signal.

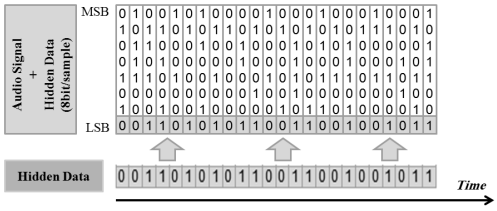

Fig. 2: LSB in 8b/sample signal is overwrote by one bit of the embedded data.

To improve robustness against distortion and noise of LSB method, [5], [6] and [7] have increased the depth of the embedding layer from 4th to 6th and 8th LSB layer without affecting the perceptual transparency of the stego audio signal. In [5] and [6], only bits at the sixth position of each 16 bits sample of the original host signal are replaced with bits from the message. To minimize the embedding error, the other bits can be flipped in order to have a new sample that is closer to the original one. On the other hand, [7] has shifted the the LSB embedding to the eighth bit and has avoided hiding in silent periods or near silent points in the host signal. The fact that the embedding occurs in the eighth bit will slightly increase the robustness of this method compared to the conventional LSB methods. However, the hiding capacity will decrease since some of the samples have to be left unchangeable to preserve the audio perceptual quality of the audio signal. In addition, the easiness of message retrieval is still one of the major drawback of the LSB and its variant, knowing by fact that embedded bits are at sixth or eighth position from the stego audio signal.

B. Echo hiding

Echo hiding method embeds data into audio signal by introducing short echo to the host signal. The nature of the echo is a resonance added to the host audio. Therefore, the problem of sensitivity of the HAS to the additive noise is avoided. After the echo has been added, the stego signal retains the same statistical and perceptual characteristics. Data is hidden by manipulating three parameters of the echo signal: initial amplitude, offset (delay) and decay rate so that the echo is not audible [8]. For a delay up to 1 ms between the original signal and the echo, the effect is indistinguishable. In addition to that, the amplitude and the decay rate could be set to values under the audible threshold of human ear. Considering the latter, data could be hidden without being perceivable. However, the drawback is lenient detection and low detection ratio that restrict application. Due to low embedding rate and low security, no audio steganography system based on echo hiding has been presented in recent researches, to the best of our knowledge. Moreover, only few techniques have been proposed for audio watermarking. To improve the watermark system robustness against common signal processing, an interesting echo hiding-time spread technique has been proposed in [9]. Compared to the conventional echo-hiding system, the proposed method detects the watermark bit based on the correlation amount at the receiver not on the delay. In addition to that, the watermark is spreaded into the whole signal. The presented system is cepstral content based in which the original signal cepstral portion of error is removed at the decoder. This led to a better detection rate.

III. Transform domain

A. Frequency domain

1) Tone insertion: Tone insertion techniques rely on the inaudibility of lower power tones in the presence of significantly higher ones. A phenomenon of the auditory masking of HAS in the spectral domain. The "masking" effect is a property of HAS which make any weak speech component imperceptible by listeners in presence of a much louder one [10]. Embedding data by inserting inaudible tones in cover audio signal is presented in [11] and [12]. More precisely, to embed one bit in a speech frame, a pair of tones is generated at two chosen frequencies f_0 and f_1. The power level of the two masked frequencies is set to a known ratio of the general power of each audio frame. By inserting tones at known frequencies and at low power level, concealed embedding and correct data extraction are achieved. The author of this method acknowledge a hiding capacity of 250 bps when four tones are inserted in each speech spectrum. Any attempt to further increase the capacity must use more than four tones. Yet, the stego-speech quality can be negatively affected. Tone insertion method can resist some of the unintentional attacks such as low-pass filtering and bit truncation. Besides the low embedding capacity, embedded data can be retrieved since inserted tones are easy to detect.

2) Phase coding: Phase coding exploits the human audio system insensitivity to relative phase of different spectral components. It is based on replacing selected phase components from the original speech spectrum with hidden data. However, to insure inaudibility, phase components modification should be kept small [13]- [16]. It is also noted that among data hiding techniques, phase coding tolerates better signal distortion [1]. Authors in [13] have inserted data in phase components using an independent multi-band phase modulation. In this approach, imperceptible phase modifications are achieved using controlled phase alteration of the host audio. Quantization Index Modulation (QIM) method is applied on phase components, where phase value of a frequency bin is replaced by the nearest o point to hide '0' or x point to hide '1' as shown in Fig. 3. The method has an embedding capability of 20 to 60 bps in 44.1 kHz, is resistant to compression but do not survive low-pass filtering.

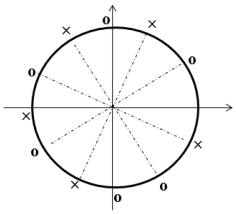

Fig. 3: Phase quantization for a frequency component [13].

For greater embedding capacity, [14] has applied QIM on the phase of the strongest harmonic with a step size of $\pi/2^n$. At three different quantization levels (n=1, 2 and 3) about 20 kbits per minute were reliably embedded and robust to MP3 encoder with a Bit Error Rate (BER) near zero. Despite the fact that phase quantization is robust to perceptual audio compression, HAS is not very sensitive to phase distortion [1]. Consequently, an intruder can also introduce imperceptible frequency modulation and eventually destroy the used phase quantization scheme.

Using allpass digital filters (APFs), authors in [15] embed data in selected subbands using distinct patterns of APF. For hiding capacity of 1.2Kbps, the proposed scheme is robust against noise addition, compression with an error less than 2%, random chopping, re-quantization and re-sampling. To further increase the robustness of phase hiding scheme, a set of nth order APFs were used in [16]. The value of n is an even positive integer and pole locations my be chosen in a variety of ways. Data is embedded in selected APF parameters and retrieved using the power spectrum to estimate APF pole locations. This method resists standard data manipulation attacks and realizes an embedding capacity of 243 bps while providing comparable perceptual performance and better robustness.

3) Amplitude modification: HAS characteristics depend more on frequency as it is more sensitive to amplitude components. Thus, frequency domain is better than time domain for data hiding. In addition, HAS has certain peculiarities that must be exploited for hiding data effectively. The "masking effect" phenomenon masks weaker frequency near strong resonant frequency [34]. An original method has been proposed in [17] where the original odd magnitude frequency components are interpolated to generate the even samples that are used for embedding data bits. In the receiver, the original odd samples and the interpolated even samples are the same as in the coder. The method has a capacity of 3 kbps and provides robustness against common audio signal processing such as echo, added noise, filtering, resampling and MPEG compression. Sine frequencies within the range of 7 kHz to 8 kHz contribute minimally to wide-band speech intelligibility. Authors in [18] presented a method for hiding text in wide-band speech signal without degrading the speech quality. Locations from high frequency components were partially replaced by ASCII representation of the text message. Embedding capacity of 8 Kbps in 16 kHz signal has been achieved in noise free environment. To further increase the embedding capacity and to ensure the security of hidden data, [19] and [20] presented a method that limits the impact of high data capacity embedding on the quality of stego speech. It consists in finding potential embedding areas under the speech spectrum and uses the energy of each frequency bin component to determine the maximum number of bits that can confine without distorting the cover speech. The embedding in the selected frequencies occurs below a distortion level to limit the impact of the hiding on the stego-speech. To ensure security to embedded data, the algorithm uses multiple parameters that can be adjusted as shown in Fig. 4. The proposed steganographic system achieves large hiding capacity of 20Kbps in wave files of 16kHz and resists noise addition at lower hiding capacity.

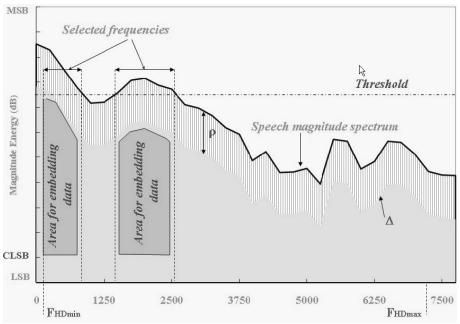

Fig. 4: Set of parameters to define potential area for data embedding [20]

4) Cepstral domain: Known also as log spectral domain, data in this method is embedded in cepstrum coefficients which endure with high fidelity most common signal processing attacks than the original samples in time domain. In addition, cepstrum alteration at frequencies that are in the perceptually masked regions of the majority of cover speech frames ensures inaudibility of the resulting stego. Employing cepstral domain modification has been proposed by [21]. The cover signal is transformed into cepstral domain and data is embedded in selected cepstrum coefficient by applying statistical mean manipulation. In this method, an embedding rate of 20 to 40 bps is achieved while guarantying robustness to common signal attacks. In [22], cepstrums of two selected frequencies f_1 and f_2. in each energetic frame are modified slightly to embed bit '1' or '0'. This method ensured a reliable embedding rate of approximately 54 bits/s. For more security of the embedded data, the latter author suggests in [23] to use the latter algorithm and embed data with different arbitrary frequency components at each frame.

5) Spread spectrum: Spread spectrum technique spreads hidden signal data through the frequency spectrum. Spread Spectrum (SS) is a concept developed in communications to ensure a proper recovery of a signal sent over a noisy channel by producing redundant copies of the data signal. Basically, Data is multiplied by an M-sequence code known to sender and receiver [24], then embedded in the cover audio. Thus, If noise corrupts some values, there will still be copies of each value left to recover the embedded message. In [25] conventional direct sequence spread spectrum (DSSS) technique was applied to hide confidential information in

MP3 and WAV audio digital signals. To control stego speech distortion, [26] and [27] have proposed an embedding method where spreaded data is hidden under a frequency mask. In [26] spread spectrum is combined to phase shifting to increase the robustness of transmitted data against additive noise and allows easy detection of the embedded data. In this method, a reliable hiding capacity of 3 bps was attained. For a better hiding rate of 20 bps, [27] used SS technique in sub-band domain. Appropriately chosen subband coefficients were selected to address robustness problem and resolve synchronization uncertainty at the decoder.

B. Wavelet domain

Audio steganography based on Discrete Wavelet Transform (DWT) is described in [28]. Data is embedded in the LSBs of the wavelet coefficients achieving high capacity of 200 kbps in 44.1 kHz audio signal. To improve embedded data imperceptibility, [29] employed a hearing threshold when embedding data in the integer wavelet coefficients, while [30] avoided data hiding in silent parts of the audio signal. Even though data hiding in wavelet domain procures high embedding rate, data extraction at the receiver side might have some errors.

IV. Encoder domain

When considering data hiding for real time communications, speech codecs such as: AMR, ACELP, SILK at their respective coding rate are employed. Passing through one of the codecs, the transmitted signal is coded and compressed according to the codec rate then decompressed at the decoder end. Thus, the signal data at the receiver is not exactly the same as it was at the sender part, which affects the hidden data retrieval correctness and what makes these techniques very challenging.

To survive speech codecs, authors in [31] have embedded data in the bitstream of an ACELP speech codec. The technique consists in hiding data jointly with the analysis-by-synthesis codebook search. The authors applied the concept on the AMR speech codec (12.2 kbit/s mode) which hides 2 kbit/s of data in the bitstream or, alternatively, reduces the codec bit rate by 1 kbit/s. Surviving speech codec, robustness, reverberations and background noises were realized in [32]. The technique hides data into speech and music signals of various types using subband amplitude modulation. Nishimura's results [32] showed that reverberant speech signals with different background noises of 10 dB SNR can transmit more than 90% of embedded data at a rate of 48 bps, with minor degradation in recognizing syllables. Embedding data in LPC vocoder was proposed by [33]. The authors used an auto-correlation based pitch tracking algorithm to perform a voiced/unvoiced segmentation. They replaced the linear prediction residual in the unvoiced segments by a data sequence. Once the residual's power is matched, this substitution does not lead to perceptual degradation. The signal is conceived using the unmodified LPC filter coefficients. Linear prediction analysis of the received signal is used to decode hidden data. The technique offers a reliable hiding rate of 2kbps.

Exploiting the LSB technique to hide data in speech codec is described in [34]. The technique consists in embedding data in the LSB of the Fourier transform of the prediction residual of the host speech. An LPC filter is used to automatically shape the spectrum of flickering LSB noise. It yields that embedded data noise is substantially less audible in this system. A hiding rate of 80 bps is attained for a continuous AM broadcast encoded at 2.4 Kbps. The idea of reference [35] is the fact that by dividing a speech signal into segments, it is permitted to reasonably modify some acoustic parameters below the audible threshold of the human ear. The modifications are then not perceivable if the original signal is not available for comparison. Peak, phase, frequency and time transformation methods may be applied on the segments. Moreover, fundamental frequency and segment duration modifications may also be applied.

Authors in [36] have presented a lossless steganography technique for G.711-PCMU telephony speech coder. Data in this case is represented by folded binary code which codes each speech sample with a value between -127 and 127 including -0 and +0. One bit is embedded in 8 bits speech data whose absolute amplitude is zero. Depending on the number of samples whose absolute amplitudes are 0, a potential hiding rate ranging from 24 to 400 bps is obtained. To increase the hiding capacity the same authors [37] have introduced a semi-lossless technique for G.711-PCMU, where speech sample amplitudes are amplified with a defined level 'i'. Samples with absolute amplitudes vary from 0 to i are utilized in the hiding process.

V. Audio Steganography Evaluation

To evaluate the performance of the reviewed techniques, signal-to-noise ratio SNR is utilized [38]. SNR's value indicates the distortion amount induced by embedded data in the cover audio signal $s_c(m,n)$. SNR value is given by the following equation:

$$SNR_{dB} = 10\log_{10}\left(\frac{\sum_{n=1}^{N}|s_c(m,n)|^2}{\sum_{n=1}^{N}|s_c(m,n)-s_s(m,n)|^2}\right) \quad (1)$$

Where $s_s(m,n)$ is the stego-audio signal such as: $m = 1,...M$ and $n = 1,...N$, where M is the number of frames in milliseconds (ms) and N is the number of samples in each frame. Table I shows SNR values for maximum allowed hiding rate in selected temporal domain steganography softwares available online [39] and [40]. The cover audio files are WAV format speeches of 4 to 10 ms length sampled at 16 kHz with 16 bits per sample. Hidden message in both s-tools and Steghide softwares is compressed and encrypted before the embedding process. The noise level induced by the embedding in the cover speech signal is depicted in Fig. 5.

To control the distortion induced by the embedding process, most audio steganography methods based on frequency domain use a perceptual model to determine the permissible amount of data embedding without distorting the audio signal. Many audio steganography algorithms use most often frequency masking [19], [20], [26] and [27] and auditory masking

TABLE I: Objective performance of audio steganography softwares

Software	Hiding rate(Kbps)	SNR
H4PGP	32	53.5
s-tools	18	68.44
Steghide	18	57.8

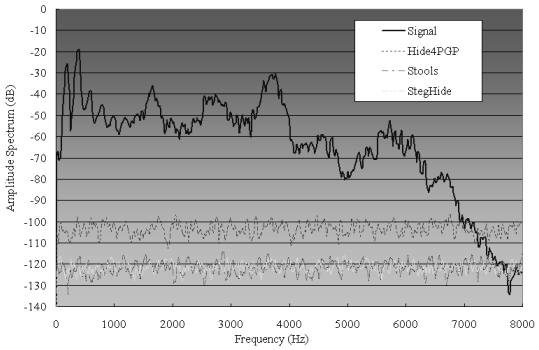

Fig. 5: Noise level in a speech frame induced by each software tool appearing in Table I.

[29], as the perceptual model for steganography embedding. In addition, some frequency domain approaches, i.e, phase embedding implicitly inherit the phase properties which include robustness to common linear signal manipulations such as: amplification, attenuation, filtering, resampling, etc.

In a more challenging environment, such as real time applications, encoder domain methods enssure robustness to embedded data against compression and noise addition. These two properties are explicitly sought when designing a robust steganographic system. However, embedded data integrity in encoder domain could be compromised if a voice encoder/decoder exists in the network. Furthermore, hidden data could be subject to transformation if a voice enhancement algorithm such as echo or noise reduction is placed in the network.

In order to prevent easy detection of embedded data, below are few points to be noted:

1) Prudent steganography practice assumes that the cover utilized to hide messages should not raise any suspicion to opponents. Known-carrier attack can be simply avoided by creating new cover audio prior to data embedding. Publicly available cover audio signals may compromise the steganography system reliability by a simple comparison between the original and the stego-audio file.
2) Human Auditory System (HAS) toleration to common sounds alterations (e.g. loud sounds tend to mask out quiet sounds) [1] have a dual effect on audio steganography system: (1) steganographer take to their advantage the fact that noise created by data embedding using frequency masking tend to be unnoticeable and (2) lossy audio data compression such as MPEG/Audio encoding will remove masked frequencies and consequently the embedded data [41].
3) Hiding in silent intervals of audio signal or in high LSB layers of audio samples can compromise the steganographic system by a simple listening test.
4) Adding an encryption layer prior to the embedding process reinforce the security of steganographic systems and makes hidden message detection harder. Encrypted data have generally high degree of randomness which adapts very well with the nature of audio files [42].
5) Embedding data in an indeterministic way makes its extraction challenging [20]. Hence, if the existence of steganographic message is discovered, it would not necessarily lead to its extracting.

To sum up, strengths and weaknesses of the reviewed techniques are shown in Table II. Focusing on factors such as security considerations concerning hostile channel attacks, robustness or larger hiding capacity depend on the application and used channel transmission conditions.

TABLE II: Summary and evaluation of Audio steganography methods

Hiding domain	Methods	Strengthes	Weaknesses	Hiding rate
Temporal domain	Low bit encoding	Simple and easy way of hiding Information with high bit rate	Easy to extract and to destroy	16 kbps
	Echo hiding	Resilient to lossy data compression algorithms	Low security and capacity	40-50 bps
Frequency Domain	Magnitude spectrum	Longer message to hide and less likely to be affected by errors during transmission	Low recovery quality	3 Kbps
	Tone insertion	Imperceptibility and concealment of embedded data	Lack of transparency and security	250 bps
	Phase spectrum	Robust against signal processing manipulation and data retrieval needs the original signal	Low capacity	333 bps
	Spread spectrum	Provide better robustness	Vulnerable to time scale modification	20 bps
	Cepstral domain	Robust against signal processing operations	Perceptible signal distortions and low robustness	54 bps
Wavelet Domain	Wavelet coefficients	Provide high embedding capacity	lossy data retrieval	200 kbps
Codecs domain	Codebook modification	High robustness	Low embedding rate	2 kbps
	Bitstream hiding	High robustness	Low embedding rate	400bps

VI. Conclusion

To ensure digital information security, various techniques have been presented in recent researchers work. Audio steganography, in particular, addresses issues related to the need to secure and preserve the integrity of data hidden in voice communications, even when the latter passes through insecure channels. This paper presents a review of the current state of art literature in digital audio steganography techniques and approaches. We presented some of the most interesting audio steganography techniques. We discussed their potentials

and limitations in ensuring secure communication. From our point of view, a comparison and an evaluation for the reviewed techniques has been also given. The advantage on using one technique over another one depends strongly on the type of the application and its exigencies such as hiding capacity or the type of attacks that might encounter the transmitted signal.

REFERENCES

[1] Walter Bender, Daniel Gruhl, Norishige Morimoto, Anthony Lu, "Techniques for Data Hiding", IBM Systems Journal, vol. 35, no. 3 and 4, pp. 313-336, 1996.

[2] E. Zwicker and H. Fastl, Psychoacoustics, Springer Verlag, Berlin, 1990.

[3] K. Gopalan, "Audio steganography using bit modification", Proceedings of International Conference on Multimedia and Expo, Vol. 1, pp.629-632, 6-9 July 2003.

[4] N. Cvejic, T. Seppiinen, "Increasing the capacity of LSB-based audio steganography", IEEE Workshop on Multimedia Signal processing, pp. 336 -338, 2002.

[5] N. Cvejic, T. Seppanen, "Increasing Robustness of LSB Audio Steganography Using a Novel Embedding Method", Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC04), vol. 2, pp. 533, 2004.

[6] N. Cvejic, and T. Seppnen, "Reduced distortion bit-modification for LSB audio steganography", Journal of Universal Computer Science, vol. 11, no.1, pp. 56-65, January 2005.

[7] Mohamed A. Ahmed, Miss Laiha Mat Kiah, B.B. Zaidan and A.A. Zaidan, "A Novel Embedding Method to Increase Capacity and Robustness of Low-bit Encoding Audio Steganography Technique Using Noise Gate Software Logic Algorithm", Journal of Applied Sciences, vol. 10, pp. 59-64, 2010.

[8] D. Gruhl, W. Bender, "Echo hiding", Proceeding of Inforomation Hiding Workshop, pp. 295315, 1996.

[9] Erfani, Y. and Siahpoush, S, "Robust audio watermarking using improved TS echo hiding", Digital Signal Processing, vol. 19, pp.809-814, September 2009.

[10] B. Paillard, P. Mabilleau, S. Morissette, J. Soumagne, "PERCEVAL: Perceptual Evaluation of the Quality of Audio Signals", journal of Audio Engeneering Society, vol. 40, pp 21-31, February 1992.

[11] K. Gopalan, et al, "Covert Speech Communication Via Cover Speech By Tone Insertion", Proceeding of IEEE Aerospace Conference, Big Sky, MT, March 2003.

[12] K. Gopalan and S. Wenndt, "Audio Steganography for Covert Data Transmission by Imperceptible Tone Insertion", WOC 2004, Banff, Canada July 8 10, 2004.

[13] Gang. L, A.N. Akansu, M. Ramkumar, "MP3 resistant oblivious steganography", Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, Vol. 3, pp.1365-1368, 7-11 May 2001.

[14] X. Dong, M. Bocko, Z. Ignjatovic, "Data hiding via phase manipulation of audio signals", IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 5, pp. 377-380, 17-21 May 2004.

[15] R. Ansari, H. Malik, and A. Khokhar, "Data-hiding in audio using frequency-selective phase alteration", IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '04), pp. 389-392, Montreal, Quebec, Canada, May 2004.

[16] H. M. A. Malik, R. Ansari, and A. A. Khokhar, "Robust Data Hiding in Audio Using Allpass Filters", IEEE Transactions on Audio, Speech and Language Processing, vol. 15, no. 4, pp. 1296 - 1304, May 2007.

[17] Mehdi Fallahpour and David Megias, "High capacity audio watermarking using FFT amplitude interpolation", IEICE Electron. Express, Vol. 6, No. 14, pp.1057-1063, 2009.

[18] F. Djebbar, D. Guerchi, K. Abed-Maraim and H. Hamam, "Text-in speech spectrum steganography", ISSPA Mai 2010, Malaysia, 2010.

[19] F. Djebbar, K. Abed-Maraim, D. Guerchi, and H. Hamam, "Energy based text-in speech spectrum hiding using speech mask properties", ICSRA Mai 2010, China, 2010.

[20] F. Djebbar, H. Hamam, K. Abed-Maraim, D. Guerchi, "Controlled Distortion for High Capacity Data-in-speech Spectrum Steganography", 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06), Germany, Oct 2010.

[21] X. Li and H.H. Yu, "Transparent and robust audio data hiding in cepstrum domain", Proc. IEEE International Conference on Multimedia and Expo, (ICME 2000), New York, NY, 2000.

[22] K. Gopalan, "Audio Steganography by Cepstrum Modification", Proc. of the IEEE 2005 International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05), Philadelphia, March 2005.

[23] K. Gopalan, "A unified audio and image steganography by spectrum modification", IEEE International Conference on Industrial Technology (ICIT), pp.1-5, 10-13 Feb. 2009.

[24] Khan, K. "Cryptology and the origins of spread spectrum", IEEE Spectrum 21, pp. 70-80, 1984.

[25] S. Hernandez-Garay, R. Vazquez-Medina, L. Nino de Rivera, V. Ponomaryov, "Steganographic communication channel using audio signals", 12th International Conference on Mathematical Methods in Electromagnetic Theory, (MMET), pp. 427 - 429, 2 July 2008.

[26] H. Matsuka, "Spread spectrum audio steganography using sub-band phase shifting", In IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP06), pp. 36, Pasadena, CA, USA, December 2006.

[27] X. Li, H.H. Yu, "Transparent and robust audio data hiding in subband domain", Proceedings of the Fourth IEEE International Conference on Multimedia and Expo, (ICME 2000), New York, NY, pp. 397400, 2000.

[28] N. Cvejic, T. Seppanen, "A wavelet domain LSB insertion algorithm for high capacity audio steganography", Proc. 10th IEEE Digital Signal Processing Workshop and 2nd Signal Processing Education Workshop, pp. 5355, 1316 October 2002.

[29] Mohammad Pooyan, Ahmed Delforouzi, "Adaptive Digital Audio Steganography Based on Integer Wavelet Transform", Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2007), vol. 2 pp. 283 - 286, 2007.

[30] S. Shirali-Shahreza and M. Shirali-Shahreza, "High capacity error free wavelet domain speech steganography", Proc. 33rd Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2008), pp. 17291732, 30 March 2008.

[31] B. Geiser, P. Vary, "High rate data hiding in ACELP speech codecs", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008), pp. 4005 - 4008, 4 April 2008.

[32] A. Nishimura, "Data hiding for audio signals that are robust with respect to air transmission and a speech codec", IIH-MSP'08, pp. 601-604, 15-17 Aug 2008.

[33] K. Hofbauer and G. Kubin, "High-rate data embedding in unvoiced speech," in Proc. Int. Conf. Spoken Language Processing (INTERSPEECH), Pittsburgh, PY, USA, pp. 241-244, September 2006.

[34] G.S.Kang, T.M.Moran, D.A.Heide, "Hiding Information Under Speech", Naval Research Laboratory, Washington, DC 20375-5320, NRL/FR/5550–05-10,126, 2005.

[35] Ponomar, Marina, "Data hiding in speech signals on the basis of the modification of segment pitch and duration", 19th International Congress on Acoustics ICA2007MADRID, 2-7 Sept. 2007, Madrid, Spain, 2007, CAS-03-023, p.46.

[36] Naofumi Aoki, "A Technique of Lossless Steganography for G.711 Telephony Speech", International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2008), pp. 608-611, 2008.

[37] Naofumi Aoki, "A Semi-Lossless Steganography Technique for G.711 Telephony Speech", International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2010), pp. 534-537, 2010.

[38] Y. Hu, P. Loizou, "Evaluation of objective quality measures for speech enhancement", IEEE Transactions on Speech and Audio Processing, 16(1), 229-238, 2008.

[39] http://www.jjtc.com/Security/stegtools.htm

[40] Steghidehttp://steghide.sourceforge.net/

[41] Robert Krenn, "Steganography and steganalysis", http://www.krenn.nl/univ/cry/steg/article.pdf, January 2004.

[42] H. Farid, "Detecting Steganographic Messages in Digital Images", Technical Report TR2001-412, Dartmouth College, Computer Science Department, http://www.cs.dartmouth.edu/~farid/publications/tr01.pdf, 2001.