Hindawi
Mathematical Problems in Engineering
Volume 2017, Article ID 3597695, 9 pages
https://doi.org/10.1155/2017/3597695

Research Article

Semifragile Speech Watermarking Based on Least Significant Bit Replacement of Line Spectral Frequencies

Mohammad Ali Nematollahi,[1] Chalee Vorakulpipat,[2] and Hamurabi Gamboa Rosales[3]

[1] *Department of Computer Engineering, Islamic Azad University, Safadasht Branch, Tehran, Iran*
[2] *Cybersecurity Laboratory, Wireless Innovation and Security Research Unit, National Electronics and Computer Technology Center (NECTEC), 112 Phahonyothin Road, Khlong Nueng, Khlong Luang District, Pathumthani 12120, Thailand*
[3] *Department of Electronics Engineering, Universidad Autónoma de Zacatecas, 98000 Zacatecas, ZAC, Mexico*

Correspondence should be addressed to Mohammad Ali Nematollahi; greencomputinguae@gmail.com

Received 19 July 2016; Revised 24 November 2016; Accepted 19 December 2016; Published 10 January 2017

Academic Editor: Xinkai Chen

Copyright © 2017 Mohammad Ali Nematollahi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are various techniques for speech watermarking based on modifying the linear prediction coefficients (LPCs); however, the estimated and modified LPCs vary from each other even without attacks. Because line spectral frequency (LSF) has less sensitivity to watermarking than LPC, watermark bits are embedded into the maximum number of LSFs by applying the least significant bit replacement (LSBR) method. To reduce the differences between estimated and modified LPCs, a checking loop is added to minimize the watermark extraction error. Experimental results show that the proposed semifragile speech watermarking method can provide high imperceptibility and that any manipulation of the watermark signal destroys the watermark bits since manipulation changes it to a random stream of bits.

1. Introduction

Digital watermarking has various applications in media security, copyright protection, and authentication as a complementary technique. Currently, many studies [1, 2] are exploring speech authentication systems. Semifragile speech watermarking, as the most important part of digital watermarking technology, verifies that a speech is genuine. Semifragile speech watermarking is a general authentication method [3, 4] applied for verification. It is required in several applications, including forensics, telephone banking, VoIP, police security, air traffic control for VHF communication, and general biometrics.

The majority of speech watermarking algorithms apply LPCs to embed the watermark [3, 5–7]. However, even without attacks, the estimated and modified LPCs vary from each other because they have a multivariate Gaussian distribution [8]. Researchers have tried to combat or bypass the variations by analysis by synthesis (AbS) techniques. Although some recent works in [3, 5–7] have applied LPCs for speech watermarking, this paper attempts to increase the imperceptibility and fragility of the watermarking. In order to do so, the disadvantages of the previous works are rectified. In [3], the LPCs are quantized through converting LPCs to Reflection coefficients (RCs) and applying inverse sine (IS). However, the stability of all-pole filter cannot be guaranteed. Therefore, the LPCs are converted to the LSFs and then quantized to guarantee the stability and reduce the spectral sensitivity [5]. However, the quantization strategy suffers from some limitations (e.g., amplitude scaling), furthermore, it degraded the imperceptibility of the watermarked speech signals. In [6, 7], the formants of the speech signal have modified to carry the watermark bits. However, this approach reduces the spectral sensitivity and it cannot be applied to nonvoice segments properly. Furthermore, formants carry most of the specific information of the speaker's voice which can be degraded by this approach. In this paper, the proposed approach converts the LPCs to LSFs and applies LSBR method to provide trade-off among capacity, imperceptibility, and fragility. Furthermore, this approach is completely blind

that neither quantization step nor other parameters (for controlling the degree of shift in each formants enhancement) are required. Moreover, a checking loop is considered to reduce the watermark extraction error and decrease the watermark distortion.

The remainder of the paper is organized as follows. The first part is an overview of speech watermarking techniques and linear predictive analysis (LPA) theory. The second part proposes a new semifragile speech watermarking technique. The final part is a discussion of experimental results, conclusions, and future works and research opportunities in this area.

2. Digital Speech Watermarking

A watermark is a method to protect digital channel communication. A watermark can be applied as a header in a digital telephone recording like an analog header to demonstrate that the signal has not been tampered with. Speech watermarks can be embedded as a mark or time-stamp inside a speech signal to prevent any channel modification including intentional or unintentional manipulation. Speech features should not be affected by the watermark, particularly when speaker identity is crucial, such as in forensic applications.

Before moving to the next section, an overview of speech watermarking technology, limitations, and problems follows.

2.1. Terminology. Watermarking is the technique and the art of hiding additional data (such as watermarked bits, logo, and text messages) in the host signal that may be an image, video, audio, speech, or text without the existence of additional information being perceived. The additional information that is embedded into the host signal should be extractable and must resist against various intentional and unintentional attacks.

2.2. Challenges in Digital Speech Watermarking. To develop a speech watermark, various factors must be considered. First, the embedded watermark must react to intentional changes such as replacement or modification but withstand unintentional attacks such as quantization and amplitude modification to provide authentication. Second, it must permit trade-off of authentication lengths. Although a long authentication is preferred for the extraction process, a short length can be used to precisely detect frames that are under attack. Another main challenge is trading off capacity, imperceptibility, and robustness. All of those criteria oppose each other, and meeting them is impossible or very difficult.

2.3. Review of Related Work on Speech Watermarking. Special characteristics such as production and perception of a speech signal distinguished it from other types of signal such as audio. Although, many techniques are available for embedding a mark into a speech signal, they can be roughly classified into the following categories.

2.3.1. Transform Domain. Several techniques have been proposed. The first method is based on auditory masking [13, 14] that uses unimportant perceptual components of speech segments to insert watermark bits. The majority of these techniques rely on features of the human auditory system (HAS) to ensure that the watermark cannot be heard. The core of this technique is that, in the presence of a louder sound (masker), a lower sound (masked) is not heard. This method is dependent on the temporal and spectral features of the masker and the masked and is divided into frequency and temporal masking. The second method is called spread spectrum (SS) [15, 16]. This method spreads hidden pseudorandom data throughout the frequency spectrum and extracts watermarks by calculating the correlation between pseudorandom noise data and the watermarked speech signal. Linear SS applies DS/BPSK (direct sequence spread spectrum/binary phase shift keying) to embed confidential data into the host speech signal.

Phase modulation is the third method; this technique embeds the watermark bits by modifying the phase of speech to preserve the power spectrum without any changes. Instead of MSE distortion as in other techniques, in phase modulation techniques, watermarked and original speech have the same power spectrum. The basis for this method is that HAS is less sensitive to absolute phase compared to relative phase or amplitude. There are two well-known methods of phase modulation; the first is called phase modification, which embeds the watermark into different bands. The second one is called phase coding, which uses one frame for all of the watermark data.

Although there is no standard definition available for this phase, three definitions have been used frequently in various references [17]. These definitions include the autoregressive (AR) phase, DFT phase, and lapped orthogonal transforms (which has no direct phase manipulation). In the fourth method, the speech signal is transformed to cepstrum by the log spectral domain, and then the watermark bits are embedded in the cepstrum coefficients [18]. This method can provide the proper robustness, inaudibility, and capacity. The fifth method, which is named amplitude coding [19], is based on applying frequency masking transformation, the wideband magnitude speech spectrum is calculated, then a secure embedding area is found that is usually located between 7 kHz to 8 kHz, and watermark bits are embedded in this area. This method has been shown to provide better intelligibility, inaudibility, and capacity.

2.3.2. Parametric Modeling. In contrast to other signal types such as audio, image, and video, a speech signal can be modeled by an all-pole filter (autoregressive (AR)). The first technique indirectly modifies or quantizes (AR) parameters such as LPC and line spectral pair (LSP) to embed the watermark [3, 5–7, 20]. The second technique embeds the watermark in the bit stream of codec such as G.729 [21], ACELP [22], G.711-PCMU [23], and G.723.1 [24] to bypass speech compression attacks. This technique might embed the watermark during or after speech compression.

2.3.3. Patchwork Method. The fundamental concept behind this method is manipulation of two sets of speech signals to determine the difference between them. This can be done by statistical methods to change the variance, energy, or mean of

two sets [25]. Performance of this method is directly related to the distance (mean, variance, and energy) between the two sets. If this distance is large, the watermark extraction is easier, but the speech is imperceptibly degraded.

2.4. Linear Predictive Analysis. For speech production, glottal excitation excites the vocal tract and is then filtered by lip radiation. Equation (1) shows speech signals in the frequency domain:

$$S(z) = G(z) \cdot H(z) \cdot R(z), \tag{1}$$

where $G(z)$ may be an impulse train generator for voiced speech or white noise for fricative and unvoiced speech signals. $H(z)$ denotes the vocal tract and $R(z)$ corresponds to lip radiation.

Linear predictive analysis can model quasistationary (between 20 and 30 ms) parts of a speech signal as a linear combination of past samples (LPCs) and errors (LPC residual). While LPCs model the vocal tract system with Pth-order real coefficients $[a1, a2, \ldots, ap]$ as in (2), the LPC residual provides information about excitation sources as in (3):

$$\hat{s}(n) = \sum_{i=1}^{P} a_k s(n-k), \tag{2}$$

$$\text{Residual error} = s(n) - \hat{s}(n). \tag{3}$$

The source-filter that models the envelope spectrum corresponding to the resonance of the vocal tract is a pth-order complex polynomial as in

$$|H(z)| \equiv \left|\frac{1}{A(z)}\right| = \left|\frac{1}{1-\sum_{k=1}^{p} a_k Z^{-k}}\right|, \tag{4}$$

where $z = r\exp(i2\pi f/F_s)$ corresponds to a polar number, r is its magnitude, and $i2\pi f/F_s$ is its phase (angle) and F_s is sampling frequency. Root-solving method [26] is applied to the LP polynomial to estimate the maxima location of the vocal tract resonances (formats). Therefore, $R1, \ldots, Rp$ are the P roots of $A(z)$ as in

$$A(z) = 1 - \sum_{k=1}^{p} a_k Z^{-k} = Z^{-p} \prod_{k=1}^{P} (z - R_k). \tag{5}$$

3. Proposed Semifragile Speech Watermarking Algorithm

This paper proposes a new blind semifragile speech watermarking technique by applying statistical methods in contrast to previous works [3, 4] that embed the watermark in LPCs by quantization. The semifragile speech watermarking technique proposed is based on LSBR of LSFs that is very sensitive against any manipulation. This speech watermarking technique can provide authentication over an unknown channel. Not only can the proposed method provide imperceptibility, but any manipulation on the watermark signal destroys the watermark bits and changes it to a random bits stream. Any small manipulation of the speech signal can change the LSFs; therefore, LSF is a good candidate for semifragile speech watermarking. However, quantization of LPCs can seriously degrade quality of the speech signal. For this reason, another LPC representation known as LPS or LSF has been applied. The LPS is less sensitive to the watermarking than LPC. Details of the embedding and extraction process are presented in the following algorithm.

3.1. Embedding Process

(a) Segment the original speech signal into frame F_i with lengths of 20 ms to 30 ms.

(b) Apply LPA on each frame to compute the LPCs (LP_i) as in (2).

(c) Convert the LPCs to LSFs based on

$$\begin{aligned} P(z) &= A(z) + z^{-(p+1)} A\left(z^{-1}\right), \\ Q(z) &= A(z) - z^{-(p+1)} A\left(z^{-1}\right), \end{aligned} \tag{6}$$

where $A(z)$ is an LPC polynomial as described in (5), P is the order of LPA, and $P(z)$ and $Q(z)$ are two decompositions symmetrical and antisymmetrical polynomials, respectively. LSP or LSF coefficients are the roots of $Q(z)$ and $P(z)$.

(d) Find the maximum LSF coefficient for embedding the watermark bit. Embedding the watermark in the maximum LSF can improve imperceptibility because when the LSF is larger, the ratio between the watermark LSF and the original LSF ((LSF + wm)/LSF) decreases.

(e) Apply least significant bit replacement (LSBR) to embed the watermark bits in the maximum LSF coefficients. LSBR is selected due to good fragility properties and simplicity of the embedding technique.

(f) Convert LSFs to LPC coefficients ($\widehat{\mathrm{LP}}_i$) based on (7) where $A(z)$ is made up from $P(z)$ and $Q(z)$.

$$A(z) = \frac{P(z) + Q(z)}{2}. \tag{7}$$

(g) Synthesize the speech frame ($\widehat{F}_i$) by using ($\widehat{\mathrm{LP}}_i$).

(h) To overcome a statistical feature of LPCs (multivariate Gaussian distribution) that can occur differences between embedding and extraction of LPCs even in the absence of any attack, a checking loop is added to reduce the watermark extraction error. If the embedding and extraction error is less than a threshold amount, reconstruct the watermarked speech signal based on modified frames $\widehat{F}_i$. Otherwise, process it through the embedding loop again. The threshold is application-dependent which is trading off among time, imperceptibility, and robustness. If the usage is for a real-time system that needs to watermark the speech signal quickly, the suitable threshold should

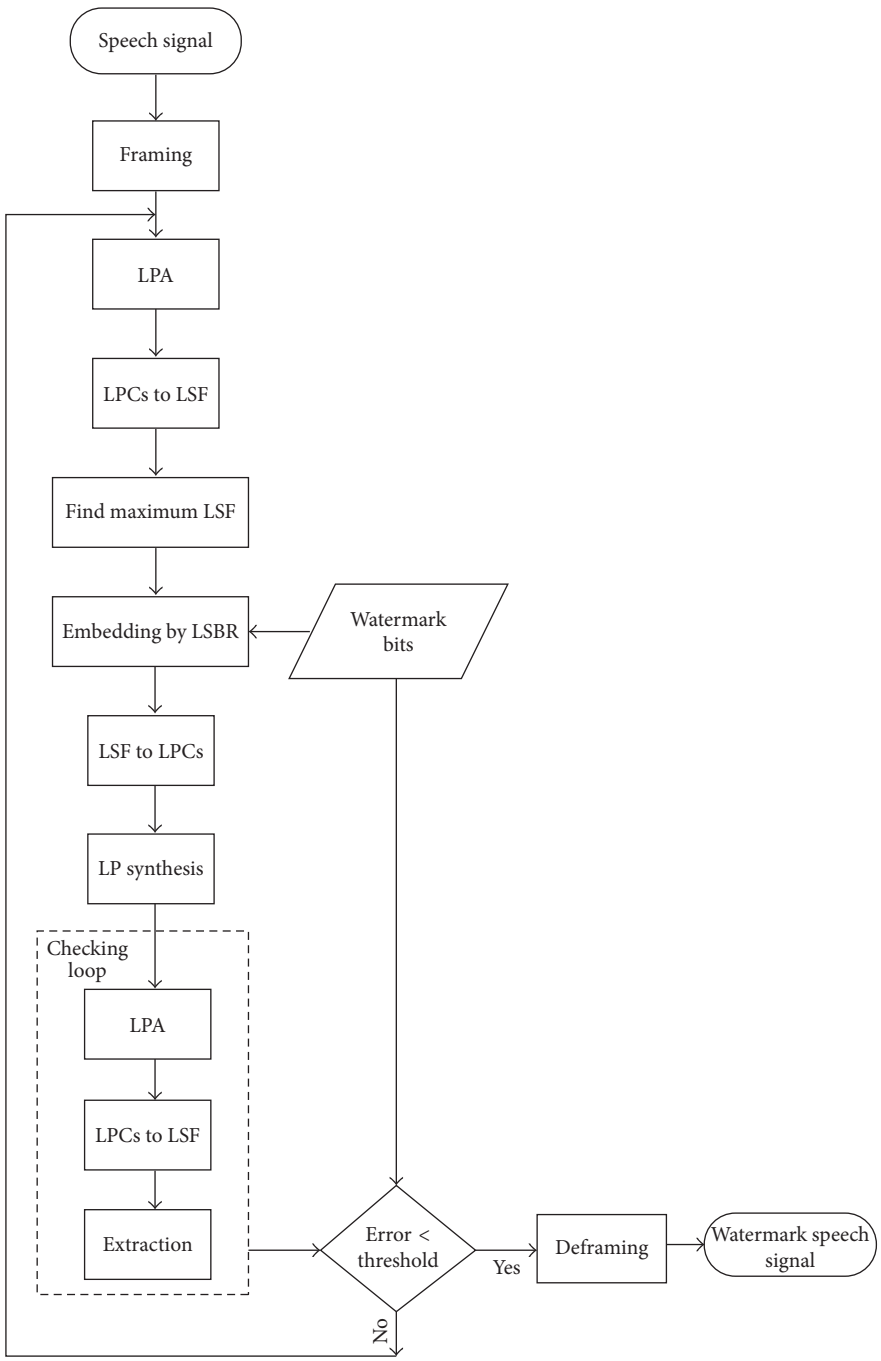


Figure 1: Block diagram of embedding process in the proposed fragile digital speech watermarking technique.

be small enough to have less number of iteration in checking loop. If the usage is for a highly imperceptible or a highly robust system, the number of iteration in checking loop should be large enough to minimize the BER. In this work, the amount of threshold is assumed to be 0.02.

Figure 1 shows the block diagram of the embedding process in the proposed semifragile speech watermarking technique.

Due to the selection of a simple technique for the embedding process, extraction of the watermark is the reverse of the embedding process that is described as follows.

3.2. Extraction Process

(a) Segment the watermarked speech signal into frames ($\widehat{F_i^I}$) with the same length as when embedded.

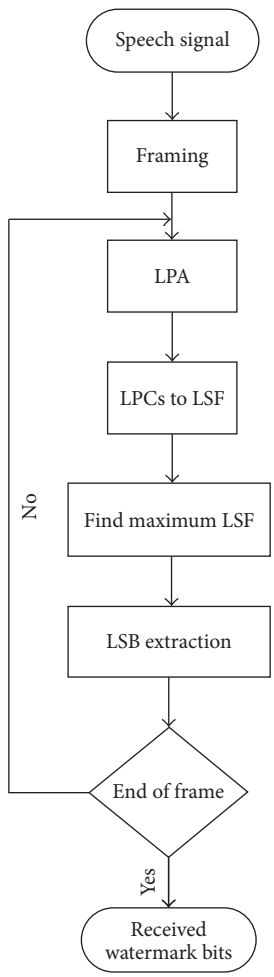


Figure 2: Block diagram of the extraction process in the proposed fragile speech watermarking technique.

(b) Apply LPA on frames ($\widehat{F_i'}$) to compute LPCs (LP_i) as in (2).

(c) Convert LPCs (LP_i) to LSFs based on (6).

(d) Find the maximum LSF and read the Least Significant Bit (LSB).

(e) Construct the extracted LSBs to a watermark bit stream.

Figure 2 shows the block diagram of the extraction process in the proposed semifragile speech watermarking technique.

4. Experimental Results

To evaluate performance of the proposed watermarking scheme in a real situation, a series of simulations were implemented and tested in MATLAB (R2010b). The speech signals used in the experiments were captured from 50 speakers (25 men and 25 women). The speech is mono in wave format at $F_s = 8$ kHz, 16 bit/sample, the bandwidth is 4 kHz, and a H-360 Logitex microphone was used for speech recording. Simulation was performed to evaluate the fragility performance of the proposed fragile speech watermarking technique. As discussed, LPCs were computed from each frame. Then, a watermark was embedded in the LSB of the maximum value of LPCs. However, quantization of LPCs may significantly degrade the quality of the speech signal. Figure 3 presents a comparison between quantization of LSF and LPC for various quantization bits. As seen, for almost all quantization bits, quantization of LSF was approximately 15 dB SNR higher than quantization of LPC. Therefore, LSF has been selected for the developed fragile digital speech watermarking technique. Figure 3 also shows that when the number of quantization bits was increased, the quality of the speech signal was increased.

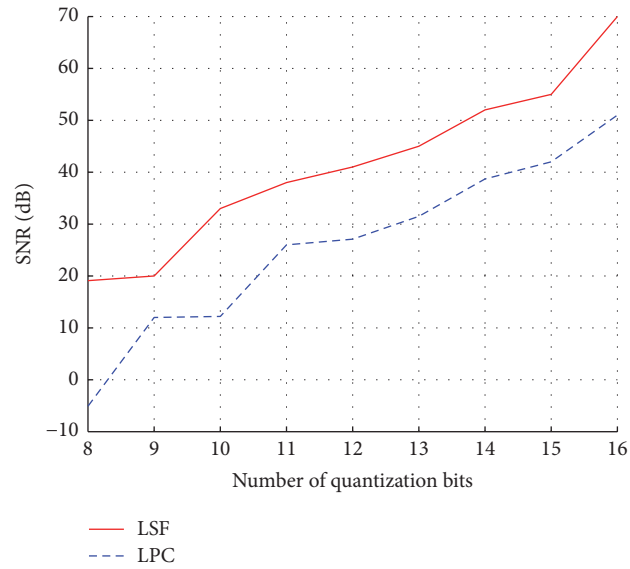


Figure 3: Comparison between quantization of LPC and LSF for various quantization bits.

As discussed in Section 1, due to the statistical nature of LPCs, when the watermark embeds into the LPCs, identical LPCs are not extracted due to residual error in the LP before LPCs are watermarked. Therefore, iterations must be performed using analysis by synthesis (AbS) to reduce the error between LPCs embedding and extraction. Figure 4 shows the number of required iterations in the AbS loop for different quantization bits. Although more quantization bits improve the imperceptibility of the watermark speech signal (see Figure 3) additional iteration should be applied to improve the probability of correct detection of the watermark. For example, for 12 quantization bits, only 6 iterations were performed to reach zero error. However, perfect watermark detection ($P_c = 1$) never will be achieved for 16 quantization bits, even if an infinite number of iterations are performed.

For evaluating the fragility property of the developed semifragile digital speech watermarking technique, some attacks including AWGN, low pass filter (LPF), band pass filter (BPF), high pass filter (HPF), median filter, and resampling were designed. Without any attack, the probability of correct detection of a watermark is one.

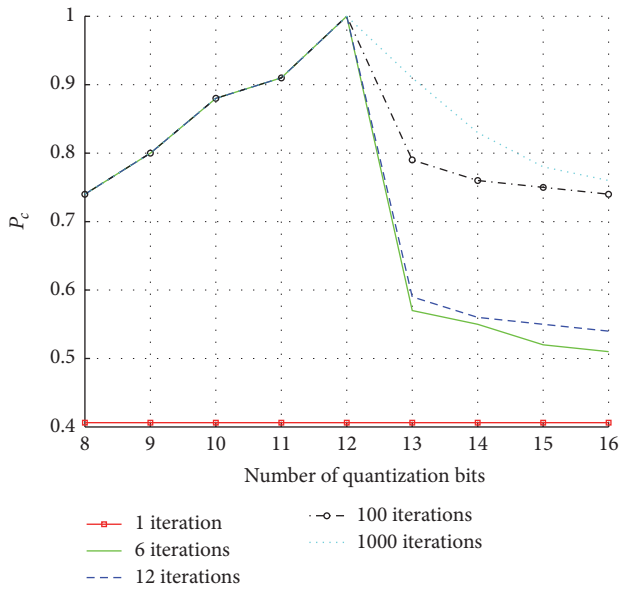


Figure 4: Comparison of the number of required iterations to reach a high probability of correct detection of watermark for various quantization bits.

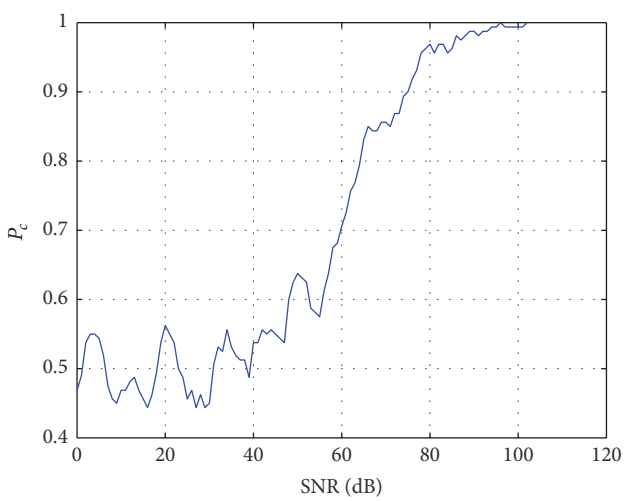


Figure 5: Probability of correct detection of a watermark for different SNR for an AWGN attack.

4.1. AWGN Attack. In this attack, the watermarked speech signals were passed through an AWGN channel with different SNRs. Figure 5 shows the probability of correct detection of a watermark in the range of 0 dB to 120 dB. As seen, the probability of correct detection was less than 90% for SNR = 75 dB. The watermark was extracted without error for SNRs higher than 104 dB.

4.2. LPF Attack. In this attack, the watermarked speech signals were passed through LPF with different passbands within the range of 100 Hz to 7500 Hz. Figure 6 shows the probability of correct detection of a watermark for various passbands. For all passbands, the probability of correct detection was less than 50%. Therefore, any manipulation by LPF can be detected.

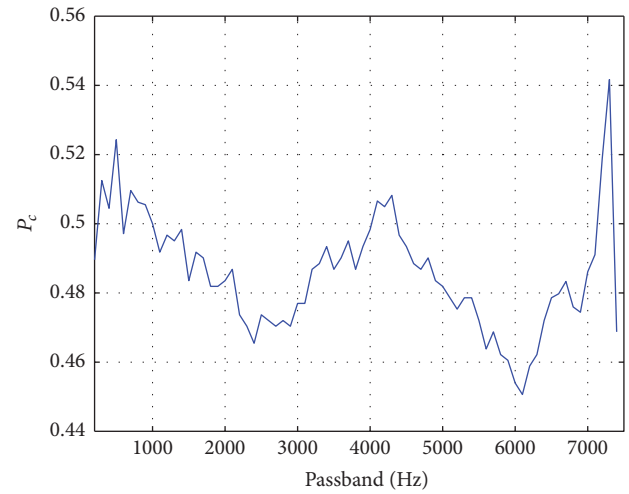


Figure 6: Probability of correct detection of a watermark for different passbands for an LPF attack.

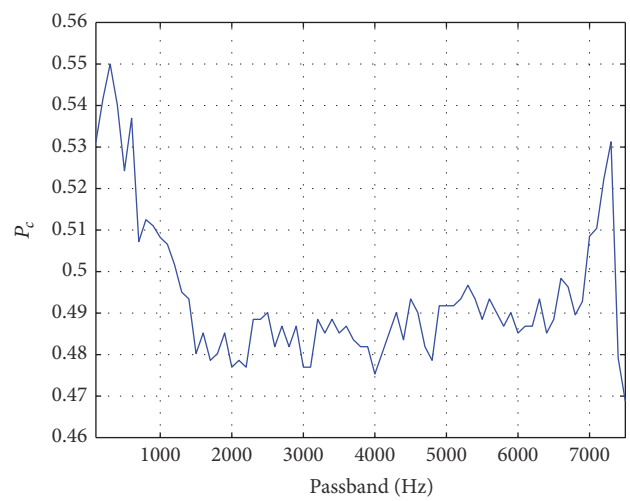


Figure 7: Probability of correct detection of a watermark for different cutoff frequencies approximately 4 KHz for a BPF attack.

4.3. BPF Attack. In this attack, the watermarked speech signals were passed through a BPF with bandwidth between 100 Hz to 7500 Hz and central frequency of 4 KHz. By changing the bandwidth of the BPF, the watermarked speech signal was filtered. Then, the fragile watermark was extracted as in Figure 7. The random nature of the extracted watermark shows that any BPF can be detected.

4.4. HPF Attack. In this attack, the watermarked speech signals were passed through an HPF with bandwidth of 200 Hz to 7500 Hz by selecting various bandwidths. Figure 8 presents the probability of correct detection of the watermark for all of the bandwidths that was approximately 50%.

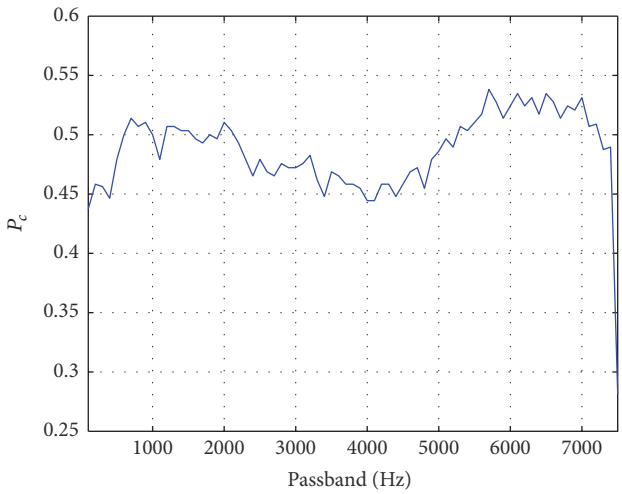


Figure 8: Probability of correct detection of a watermark for different cutoff frequencies for an HPF attack.

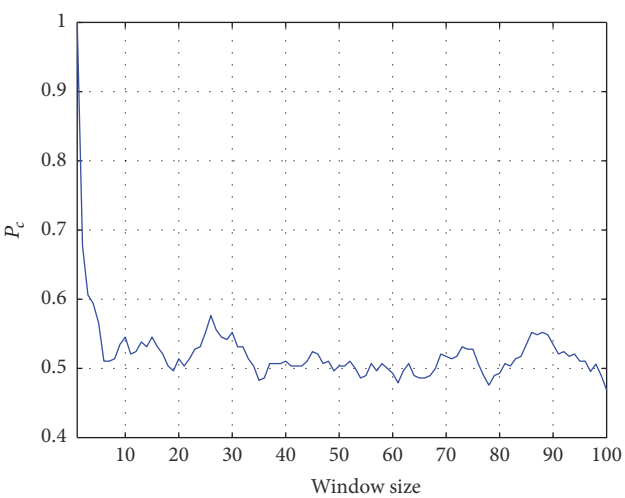


Figure 9: Probability of correct detection of a watermark for different window sizes for median filter attack.

4.5. Median Filter Attack. In this attack, the speech watermarked signals were passed through a median filter with window sizes from 1 to 100. Figure 9 shows the probability of correct detection of the watermark for all window sizes. Apart from window size of 1, the watermark bits were extracted randomly for the rest of the window sizes.

4.6. Resampling Attack. In this attack, the watermarked speech signals were first downsampled with a factor, and then they were upsampled with the previous factor. Figure 10 presents the probability of correct detection of a watermark for the range of 1 to 1/20. Except for 1, the rest of the sampling factors randomly changed the extracted watermark bits.

As seen from Figures 5 and 10, the random nature of the extracted watermark bits demonstrates the fragile property of the developed semifragile digital speech watermarking technique. Therefore, any manipulation (here, only conventional signal processing operation) of the watermarked speech signal was detected by the developed semifragile digital speech watermarking technique. Imperceptibility was high after embedding the fragile digital speech watermark and will not degrade speaker recognition performance.

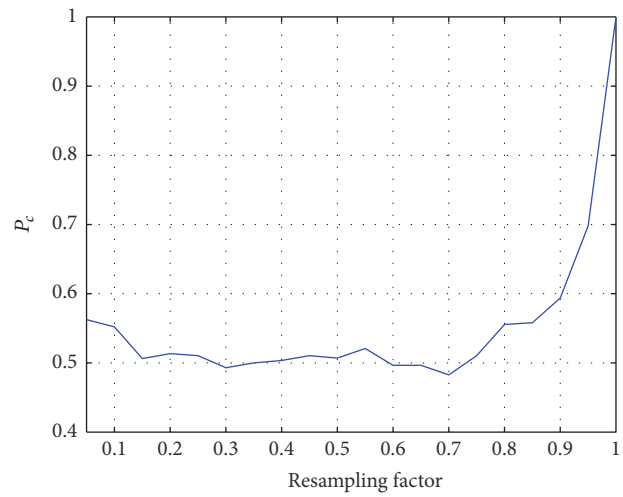


Figure 10: Probability of correct detection of a watermark for different sampling factors for resampling attack.

5. Discussion

The factors in the watermarking problem compete with each other. Figure 11 shows triangles for different systems where each of the systems only focuses on a watermarking criterion. As seen, more concentration in a watermarking criterion degrades the other watermarking criteria. Only systems that provided reasonable and acceptable performance by trading off among different watermark criteria including capacity, robustness, and recognition performance have been selected. Each criterion in each axis was normalized into the range 0 to 1 with respect to the maximum amount at that axis due to provide better visualization. For example, the amount for the capacity axis was divided by 32 bps. Each axis was organized in ascending order for better consistency. Whenever a criterion is increased, this criterion is farther from the axis origin.

Table 1 illustrates a comparison of state-of-the-art semifragile watermarking techniques in terms of their average time, bit error rate (BER), signal-to-noise ratio (SNR), and capacity. As seen, the developed semifragile speech watermarking technique can provide high imperceptibility compared to other techniques. The lower band of the capacity for the proposed technique is equal or higher than other techniques. Since any manipulation of the watermark signal must change the watermark bits to a certain random stream of bits, the BER of the proposed technique can provide more fragility than the other techniques.

In real-time applications, it is significant to evaluate the checking loop impact on the embedding process. Table 1 presents the required time for embedding process for all

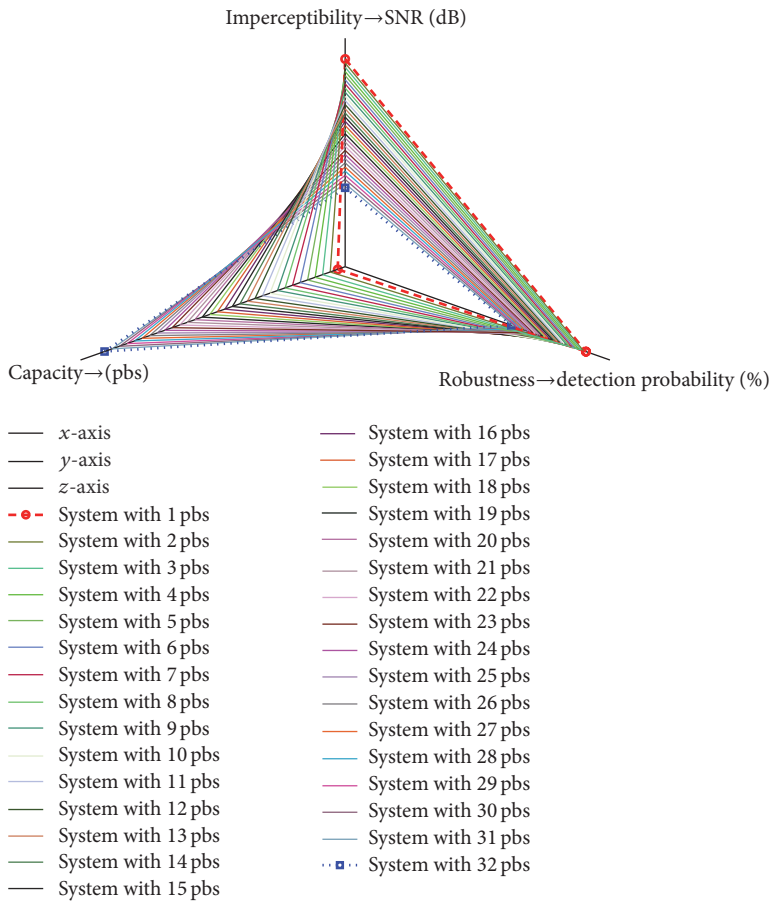


Figure 11: Triangles for the effect of capacity on other watermarking criteria.

Table 1: Comparison of different semifragile watermarking techniques in terms of their time, BER, SNR, and capacity.

Watermark techniques	BER (%)	Speech SNR (dB)	Capacity (bps)	Total time (s)
Semifragile LSF-LSBR	0.5032	47.32	33.33–50	321
DWPT-QIM [9]	0.4367	43.39	31.25–1000	232
AbS [3]	0.4780	28.08	33.33–50	309
LSF [7, 10]	0.5127	30.32	33.33–50	298
DT-CWT [11]	0.1367	31.36	15.66–976.56	351
Genetic algorithm [12]	0.4513	29.30	N/A	411

watermarking techniques. Apparently, the proposed technique was induced more delay to embedding process due to the checking loop. It must be noted that the "profile ('memory', 'on')" MATLAB function was used to compute CPU time for each watermarking technique in this paper.

6. Conclusion and Future Work

This paper presents a new blind semifragile speech watermarking algorithm to handle the limitations and problems of recent approaches. The embedding process inserts the watermark inside the maximum LSF which is more robust against various unintentional attacks. A checking loop has been added to reduce the watermark extraction error. Experimental results show that this algorithm is fragile and imperceptible under various intentional attacks. It can be concluded that LSF outperforms LPC for watermarking.

In the future, a research study and design for a speech watermarking and encryption system that can improve the security of the system with multilevel security should be conducted. Another direction might be synchronization for this watermarking scheme.

Competing Interests

The authors declare that they have no conflict of interests.

References

[1] M. Faundez-Zanuy, J. J. Lucena-Molina, and M. Hagmüller, "Speech watermarking: an approach for the forensic analysis of digital telephonic recordings," *Journal of Forensic Sciences*, vol. 55, no. 4, pp. 1080–1087, 2010.

[2] S. Saraswathi, "Speech authentication based on audio watermarking," *International Journal of Information Technology*, vol. 16, no. 1, 2010.

[3] B. Yan and Y.-J. Guo, "Speech authentication by semi-fragile speech watermarking utilizing analysis by synthesis and spectral distortion optimization," *Multimedia Tools and Applications*, vol. 67, no. 2, pp. 383–405, 2013.

[4] B. Yan, Z. Lu, S. Sun, and J. Pan, "Speech authentication by semi-fragile watermarking," in *Knowledge-Based Intelligent Information and Engineering Systems*, vol. 3683 of *Lecture Notes in Computer Science*, pp. 497–504, Springer, Berlin, Germany, 2005.

[5] S. Wang and M. Unoki, "Watermarking method for speech signals based on modifications to LSFS," in *Proceedings of the 9th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP '13)*, pp. 283–286, IEEE, Beijing, China, October 2013.

[6] S. Wang and M. Unoki, "Watermarking of speech signals based on formant enhancement," in *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO '14)*, IEEE, September 2014.

[7] S. Wang, M. Unoki, and N. S. Kim, "Formant enhancement based speech watermarking for tampering detection," in *Proceedings of the 15th Annual Conference of the International Speech Communication Association*, Singapore, 2014.

[8] J. R. Deller, J. G. Proakis, and J. H. Hansen, *Discrete-Time Processing of Speech Signals*, IEEE, New York, NY, USA, 2000.

[9] M. A. Nematollahi, M. A. Akhaee, S. A. R. Al-Haddad, and H. Gamboa-Rosales, "Semi-fragile digital speech watermarking for online speaker recognition," *Eurasip Journal on Audio, Speech, and Music Processing*, vol. 2015, no. 1, article 31, 2015.

[10] S. Wang and M. Unoki, "Speech watermarking method based on formant tuning," *IEICE Transactions on Information and Systems*, vol. 98, no. 1, pp. 29–37, 2015.

[11] M.-Q. Fan, P.-P. Liu, H.-X. Wang, and H.-J. Li, "A semi-fragile watermarking scheme for authenticating audio signal based on dual-tree complex wavelet transform and discrete cosine transform," *International Journal of Computer Mathematics*, vol. 90, no. 12, pp. 2588–2602, 2013.

[12] M. Zamani and A. B. A. Manaf, "Genetic algorithm for fragile audio watermarking," *Telecommunication Systems*, vol. 59, no. 3, pp. 291–304, 2015.

[13] F. Djebbar, K. Abed-Meraim, D. Guerchi, and H. Hamam, "Energy based text-in speech spectrum hiding using speech mask properties," in *Proceedings of the 2nd International Conference on Signal Processing, Robotics and Automation (ICSRA '10)*, Wuhan, China, May 2010.

[14] S. Van De Par, A. Kohlrausch, R. Heusdens, J. Jensen, and S. H. Jensen, "A perceptual model for sinusoidal audio coding based on spectral integration," *Eurasip Journal on Applied Signal Processing*, vol. 2005, no. 9, pp. 1292–1304, 2005.

[15] S. Arora and S. Emmanuel, "Adaptive spread spectrum based watermarking of speech," 2013.

[16] Q. Cheng and J. S. Sorensen, "Spread spectrum signaling for speech watermarking," Google Patents, 2005.

[17] M. A. Nematollahi and S. A. R. Al-Haddad, "An overview of digital speech watermarking," *International Journal of Speech Technology*, vol. 16, no. 4, pp. 471–488, 2013.

[18] K. Gopalan, "A unified audio and image steganography by spectrum modification," in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT '09)*, February 2009.

[19] F. Djebbar, B. Ayad, K. Abed-Meraim, and H. Hamam, "Unified phase and magnitude speech spectra data hiding algorithm," *Security and Communication Networks*, vol. 6, no. 8, pp. 961–971, 2013.

[20] A. Gurijala, "Speech watermarking through parametric modeling," ProQuest, 2007.

[21] J. Singh, P. Garg, and A. Nath De, "A combined watermarking and encryption algorithm for secure VoIP," *Information Security Journal: A Global Perspective*, vol. 18, no. 2, pp. 99–105, 2009.

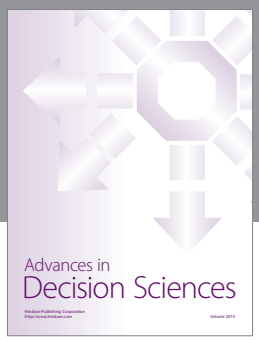
[22] B. Geiser and P. Vary, "High rate data hiding in ACELP speech codecs," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 4005–4008, Las Vegas, Nev, USA, April 2008.

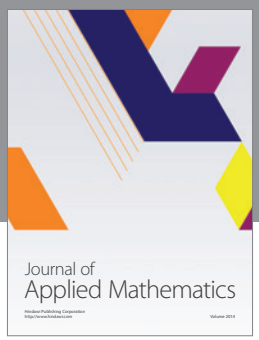
[23] N. Aoki, "A technique of lossless steganography for G.711 telephony speech," in *Proceedings of the 4th International Conference on Intelligent Information Hiding and Multiedia Signal Processing (IIH-MSP '08)*, August 2008.

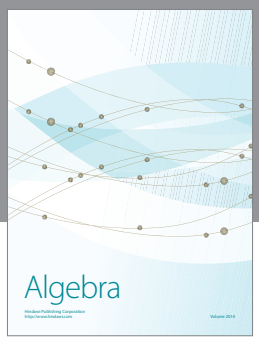
[24] Y. F. Huang, S. Tang, and J. Yuan, "Steganography in inactive frames of VoIP streams encoded by source codec," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 2, pp. 296–306, 2011.

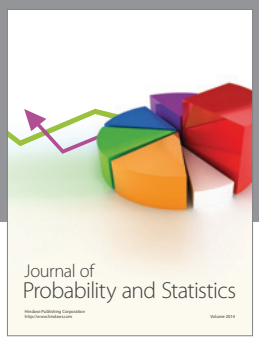
[25] I.-K. Yeo and H. J. Kim, "Modified patchwork algorithm: a novel audio watermarking scheme," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 4, pp. 381–386, 2003.

[26] G. K. Vallabha and B. Tuller, "Systematic errors in the formant analysis of steady-state vowels," *Speech Communication*, vol. 38, no. 1-2, pp. 141–160, 2002.









Hindawi

Submit your manuscripts at
https://www.hindawi.com