Steganography Audio Based on Zero-Tree Wavelet Transform Algorithm

R. Esfahani*, A. R. Matinfar

Abstract

Audio steganography is very important, equally steganography in other media (image, video, etc.). In this paper is presented steganography of audio based on embedded zero-tree waveleten transform algorithm. Improving the resistance against white noise and additive noise with the lowest SNR is one of the important topics in steganography. The proposed algorithm is more robust against normal white noise than uniform noise and has a bit error rate of less than 1 bit against SNRs higher than 10db. According to the obtained BER, if the proposed algorithm is attacked, the hidden signal is lost completel. Also, the proposed method is resistant against additive noise. The proposed algorithm has the least changes in the sound smoothness criterion in frequency domain with Capstrom distance scale and audio files in the form of music with soft tone (loudness), and the increase of secret message does not have much effect on creating disturbances in the frequency domain. The proposed algorithm of the frequency spectrum does not change the audio signal much, and it also follows the property of the hearing threshold level, and high-pitched music with male speech has the best results, so it is favorable to the spectrum structure of Bark. Also, the proposed algorithm has favorable results in the time domain. The lowest SNR is related to high-pitched music with female speech, which has an SNR of about 13db. According to the obtained results, we will have the worst case of embedding a secret message by choosing the audio signal with female speech. Because there is a certain smoothness in the fmale speech signal. Therefore, this uniform state will be lost to some extent by embedding a secret message in this type of audio signal, and the CZD criterion will increase according to the component-by-component comparison of the two main signals and the signal containing the secret message.

Key Words: *Embedded Zero Tree Wavelet Transform (EZW), White Noise, Additive Error, BSD and MBSD Criteria, CDM Kapstrom Distance Scale, ISD and COSH Distance, Short Time Radon Transform STFRT, SNR and CZD*

*Assistant Professor- Imam Hossein Comprehensive University - (resfahani@ihu.ac.ir) - writer - in - change

نشریه علمی پدافند غیرعامل

سال چهاردهم، شماره ۴، پاییز ۱۴۰۲، (پیاپی ۵۵): صص ۱۳۰-۱۱۵

علمی - ترویجی

نهان‌نگاری در صوت مبتنی بر الگوریتم تبدیل موجک درخت صفر درج‌شده

رضا اصفهانی*[۱]، احمدرضا متین‌فر[۲]

تاریخ دریافت: ۱۴۰۲/۰۱/۰۸

تاریخ پذیرش: ۱۴۰۲/۰۵/۱۶

چکیده

نهان‌نگاری در صوت همانند نهان‌نگاری در رسانه‌های دیگر (تصویر، ویدئو و ...)، از اهمیت ویژه‌ای برخوردار است. در این مقاله نهان‌نگاری در صوت مبتنی بر الگوریتم تبدیل موجک درخت صفر درج‌شده ارائه شده است. ارتقاء مقاومت در برابر نویز سفید و نویز افزایشی با کم‌ترین SNR یکی از موضوعات مهم در نهان‌نگاری است. الگوریتم پیشنهادی، در برابر نویز سفید نرمال بیشتر از نویز یکنواخت مقاوم بوده ودر برابر SNRهای بالاتر از 10db نرخ بیت خطای کم‌تر از ۱ بیت دارد. در صورت حمله به الگوریتم پیشنهادی، با توجه به BER به‌دست آمده، سیگنال مخفی کاملاً تخریب می‌شود. همچنین روش پیشنهادی در برابر اضافه‌کردن نویز افزایشی مقاوم است. الگوریتم پیشنهادی در حوزه فرکانس با مقیاس فاصله کپستروم و فایل‌های صوتی به فرم موزیک با تُن (بلندی) ملایم، کم‌ترین تغییرات را در معیار همواری صوت دارد و افزایش پیام مخفی تاثیر چندانی در ایجاد اغتشاش در حوزه فرکانس ندارد. الگوریتم پیشنهادشده طیف فرکانسی، سیگنال صوتی را چندان تغییر نمی‌دهد و همچنین از خاصیت سطح آستانه شنوایی پیروی می‌کند و موزیک تن بالا با گفتار مرد دارای بهترین نتایج می‌باشند بنابراین از هر جهت مساعد با ساختار طیف بارک می‌باشند. همچنین الگوریتم پیشنهادی نتایج مطلوبی در حوزه زمان دارد. کم‌ترین SNR مربوط به موزیک تن بالا با گفتار زن می‌باشد که دارای SNRی در حدود 13 db می‌باشد. با توجه به نتایج حاصل‌شده، با انتخاب سیگنال صوتی با گفتار زن بدترین حالت درج پیام مخفی را خواهیم داشت. زیرا در سیگنال گفتار زن همواری خاصی محسوس است. بنابراین با درج پیام مخفی در این نوع سیگنال صوتی، تا حدودی این فرم یکنواختی از بین خواهد رفت و با توجه به مقایسه‌شدن مولفه به مولفه دو سیگنال اصلی و سیگنال حاوی پیام مخفی، میزان معیار CZD افزایش می‌یابد.

کلیدواژه‌ها: تبدیل موجک درخت صفر درج‌شده (EZW)، نویز سفید، خطای افزایشی، معیار BSD و MBSD، مقیاس فاصله کپستروم CDM، فاصله ایتاکورا - سایتو ISD و فاصله COSH، تبدیل رادون زمان کوتاه STFRT، SNR و فاصله سزنا-کوشی CZD

[۱] استادیار دانشگاه جامع امام حسین (ع)، تهران، ایران- (resfahani@ihu.ac.ir)- نویسنده مسئول

[۲] استادیار دانشگاه جامع امام حسین (ع)، تهران، ایران

۱- مقدمه

با توجه به وجود فرمت‌های مختلف صوتی مانند Wma، Ra، mp3، wav، Mid، Cdda و Aiff، می‌توان از این رسانه (صوت) برای هنر پنهان‌کردن اطلاعات در آن بهره جست. لذا بررسی ویژگی‌های صوت، فرمت‌ها و کُدرهای صوتی، برای استفاده از روش‌های نهان‌نگاری در آن‌ها لازم و ضروری است [۳،۲،۱]. به‌طور کلی دو زیر فضای آنالوگ و زیر فضای دیجیتال برای محیط‌های صوت وجود دارد. برای ذخیره فایل‌های صوتی دیجیتال، مستقیماً از سیگنال صوتی آنالوگ نمونه‌برداری می‌شود. به‌عبارت دیگر فایل‌های صوتی دیجیتال از تبدیل صداهای آنالوگ به حوزه دیجیتال به‌دست آمده‌اند. فرآیند تبدیل اصوات آنالوگ به دیجیتال شامل دو مرحله چندی‌سازی‌کردن و نمونه‌برداری زمانی می‌باشد. الگوی چندی‌سازی[1]: چندی‌سازی هر مقدار ورودی را به مقادیر گسسته تبدیل می‌کند. معمول‌ترین فرمت‌های ذخیره صدا با کیفیت بالا استفاده از فرمت‌های ۱۶ بیتی خطی چندی‌ساز شده است. همانند آن چیزی که در WAV[2] یا AIFF[3] استفاده می‌شود. مقداری از دامنه سیگنال‌های صوتی به‌وسیله این فرمت دچار اعوجاج می‌شود.

نرخ نمونه‌گیری زمانی: فرایندی است که مقادیر آنالوگ در تقسیمات زمانی منظم نمونه‌برداری می‌شوند. پُر استفاده‌ترین نرخ نمونه‌گیری زمانی[4] برای صوت شامل 8KHz , 9.6KHz , 10KHz , 12KHz , 16KHz , 22.05KHz و 44.1KHz هستند. نرخ نمونه‌گیری تعداد نمونه‌های سیگنال صوتی است که در یک ثانیه از فریم صوت وجود دارد.

وسیله پخش یا محیط پخش در یک سیگنال صدا به محیطی گفته می‌شود که سیگنال در آن از کُدکننده به کُدگشا می‌رود. چهار محیط پخش شامل «محیط دیجیتالی نظیر به نظیر»، «محیطی با استفاده از کم یا زیادکردن و نمونه‌گیری مجدد»، «پخش آنالوگ و نمونه‌گیری مجدد» و «محیط هوا» است. محیط پخش و همچنین وسیله ذخیره سیگنال، در نوع الگوریتم‌هایی که برای پنهان‌کردن اطلاعات به‌کار می‌رود، اهمیت دارند [۴, ۵].

در بخش دوم موضوعات اولیه و اساسی نهان‌نگاری در فایل صوتی و تبدیل موجک درخت صفر درجشده (EZW) شرح داده شده است. ضرورت این یادآوری‌ها، به کاربرد و استفاده آن در الگوریتم پیشنهادی ارتباط دارد. در بخش سوم، روش پیشنهادی نهان‌نگاری در صوت مبتنی بر EZW ارائه شده است که شامل نحوه عملکرد سیستم، طراحی و نتایج پیاده‌سازی و ارزیابی الگوریتم پیشنهادی است.

[1] Quantisation
[2] Windows Audio Visual
[3] Audio Interchange File Format
[4] Temporal

۲- نهان‌نگاری در فایل صوتی

در حالت کلی برای نهان‌نگاری در فایل‌های صوتی مراحل زیر انجام می‌شود:

- نحوه عملکرد کد کننده MPEG: در کد کننده MPEG، مدل فیزیوآکوستیگ سیگنال‌های صوتی مورد آنالیز قرار می‌گیرد و میزان نویز ماسک‌شده موجود به عنوان تابعی از فرکانس محاسبه می‌شود.
- تطبیق (هم تراز کردن) زمانی داده‌های صوتی: در مدل فیزیوآکوستیکی که کم‌ترین پیچیدگی را دارد، محدوده فرکانسی به ۳۲ زیرباند فرکانسی تقسیم می‌شود[5]. در فشرده‌سازی داده با استفاده از الگوریتم MPEG برای هر باند بحرانی، از تکنیک بانک فیلتر استفاده می‌شود [۶]. که به طریقه رابطه (۱) محاسبه می‌شود.

$$S_t[i] = \sum_{k=0}^{63}\sum_{j=0}^{7} M[i][k] \times (C[k+64j] \times x[k+64j]) \quad (۱)$$

و

$$M[i][k] = \cos\left[\frac{(2i+1)\times(k-16)\times\pi}{64}\right]$$

که C[k] همان پنجره ماسک‌کننده در مدل مورد استفاده‌شده می‌باشد. مدل فیزیوآکوستیک از یک نگاشت زمان به فرکانس مستقل و جدا به‌جای بانک فیلتر چند فازه استفاده می‌کند. زیرا به دقت وضوح فرکانس بیشتری در محاسبات تعیین سطح آستانه ماسک‌کننده نیاز دارد. طبق پیاده‌سازی نرم‌افزاری، یک وزن‌دِه هانینگ[6] بر روی داده‌ها قبل از تبدیل فوریه اِعمال می‌شود تا اثرات لبه‌ها را در فریم صوتی کاهش دهد. تبدیل فوریه با توجه به طول پنجره آنالیز، ۵۱۲ در نظر گرفته می‌شود.

- فرایند محاسبات طیفی: در این مرحله، سطح بلندی هر طیف صوتی متناسب با زیر باند فرکانسی مورد نظر (Lsb[7])، محاسبه می‌شود.
- مجزا کردن مقادیر طیفی به مولفه‌های ضربه و غیر ضربه‌ای: هر دو مدل ضربه و مولفه‌های شبه نویزی را از سیگنال صوتی تعیین و جدا می کنند. این امر به این دلیل می باشد که قابلیت ماسک گذاری این دو نوع سیگنال متفاوت است.
- اعمال یک تابع پخش کننده: توانایی ماسک‌کنندگی یک سیگنال داده شده در محدوده ی باند بحرانی سیگنال انتشار (پخش) می‌شود.
- قراردادن باند پایین برای مقادیر سطوح آستانه: هر دو مدل شامل یک سطح آستانه ماسک‌کننده مطلق (سطح آستانه سکوت) می باشند که از روی آزمایشات به‌دست آمده است.

[5] فیلترشدن به ۳۲ زیر فضای یکسان وابسته به نرخ نایکوئیست سیگنال
[6] Hanning
[7] Level subband

سطح آستانه سکوت، در باند پایین شنوایی صوت وجود دارد.

- یافتن سطح آستانه ماسک‌کننده برای هر زیر باند: هر دو مدل فیزیوآکوستیکی، سطح آستانه ماسک‌کننده‌ای با دقت تفکیک فرکانسی بالا برای هر زیر باند بطور مجزا محاسبه می‌کنند که توسط بانک فیلتر چند فازه به دست می آید.
- محاسبه نسبت سیگنال به ماسک: این نسبت به عنوان نسبتی از انرژی سیگنال در طول زیر باند به حداقل سطح آستانه ماسک‌کنندگی در آن زیر باند محاسبه می‌شود. در ادامه مولفه‌های ضربه و غیر ضربه یک سیگنال به همراه فضای ماسک‌کنندگی این مولفه‌ها در مقابل سطح آستانه مطلق ماسک‌کنندگی نمایش داده شده است.

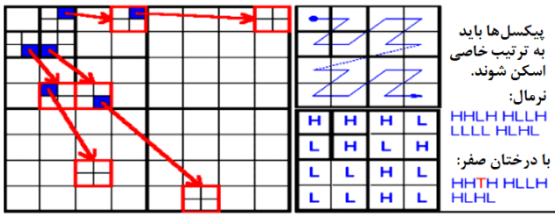
۲-۱- تبدیل موجک درخت صفر درج‌شده (EZW[1])

تبدیل موجک گسسته یک سیگنال دلخواه x به‌صورت رابطه (۲) بیان می‌شود:

$$\mathrm{DWT}(\mathrm{Scale},\mathrm{Position}) = \sum_{n=1}^{N} f(n)\Psi(\mathrm{Scale},\mathrm{Position},n) \quad (۲)$$

که $\Psi(.)$ توابع متنوعی می‌باشد که می‌توان از آن‌ها بسته به کاربرد مورد نظر استفاده نمود. بهترین فیلترهای مورد استفاده در تجزیه سیگنال صوتی، استفاده از فیلترهای 'db' می‌باشد. در این مقاله از تبدیل موجک با فیلتر 'db' در چهار سطح استفاده شده است. الگوریتم تبدیل موجک درخت صفر درج‌شده در واقع الگوریتمی است که برای فشرده‌سازی تصویر و انتقال آن ارائه شده است. EZW کدکننده‌ای می‌باشد که براساس تبدیل موجک بنا شده است. کد کننده EZW بر پایه کدکردن پیشرو پایه‌گذاری شده است و یک تصویر را به دنباله‌ای از رشته‌ها با دقت افزایشی (پیشرو) فشرده می‌سازد. نکته مورد توجه، "استفاده از وابستگی بین ضرایب تبدیل موجک در طول کل سیگنال و بخش بزرگی از سیگنال کدشده که در زیر سطح آستانه قرار دارند" می‌باشد. **در این حالت است که بحث درخت صفر مطرح می‌شود.** الگوریتم EZW شرح داده شده توسط شپیرو از روش کدکردن در سطح آستانه استفاده می‌کند [۷] . مساله مهم در این‌جا از دست دادن موقعیت ضرایب می‌باشد. بدون اطلاعات موقعیت، حتی اگر تمامی ضرایب هم کد شوند، ضرایب کدگشا، قادر به بازسازی سیگنال کدشده نمی‌باشد. این موضوع در نحوه انتخاب موقعیت، توسط کدکننده نهفته است. EZW از یک مرتبه پیمایش تعیین شده‌ای از موقعیت‌های ضرایب تبدیل موجک استفاده می‌کند. در شکل (۱) نمونه ای از ترتیب پیمایش ضرایب نمایش داده شده است.

[1] Embedded Zerotrees of Wavelet Transforms



شکل (۱): نمایش ترتیب پیمایش ضرایب تبدیل موجک [۷، ۸].

هر چند که استفاده از درخت صفر موقعیت‌های بسیاری را برای کدکردن ایجاد می‌کند، اما چندین حالت متفاوت از پیمایش به شرطی که زیرباندهای کوچک‌تر تماماً قبل از رفتن به سطح بالاتر پیمایش شوند نیز امکان‌پذیر است [۹، ۱۰].

۳- روش پیشنهادی

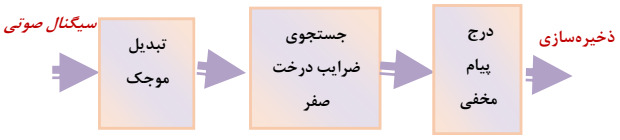
۳-۱- خلاصه عملکرد سیستم در روش پیشنهادی

در این بخش هدف ارائه الگوریتم نهان‌نگاری جهت ذخیره داده مخفی در فایل‌های صوتی می‌باشد. اساس طراحی این الگوریتم برپایه " الگوریتم فشرده‌سازی درخت صفر در تصاویر" می‌باشد. این الگوریتم از ویژگی طبیعی که در تصاویر وجود دارد، استفاده می‌کند. ویژگی که در تصویر طبیعی وجود دارد، تمایل تصاویر به توزیع نرمال می‌باشد. بنابراین از میان ضرایب موجود در ترکیب تصویر، ضرایبی که بزرگتر باشند مولفه‌های اصلی سازنده تصویر به شمار می‌آیند. بنابراین در بیان تصویر می‌توان از ضرایب کوچکتر صرف نظر کرد. الگوریتم درخت صفر بیانگر این موضوع است که ضرایب بزرگ در تشکیل تصویر نقش اصلی و سازنده آن را دارند و ضرایب کوچکتر تنها سازنده جزئیات تصویر می‌باشند. در الگوریتم درخت صفر ضرایب تصویر بر اساس تبدیل موجک در ۳ یا ۴ مرحله بیان می‌شوند و در ادامه ضرایبی را که در تخمین هر مرحله از مقدار قبل کوچکتر باشند را به عنوان ضرایب کم اهمیت در نظر می‌گیرد. زیرا این ضرایب در هر مرحله تخمین تبدیل موجک تنها سازنده جزئیات در بخش قبل می‌باشند و ضریب اصلی در مرحله اول تمامی این ضرایب را می‌توانند تحت پوشش قرار دهد زیرا ضریب مرحله بالاتر، بزرگتر از ضرایب مراحل پایین‌تر می‌باشد.

از این ایده که در " الگوریتم درخت صفر ضرایبی را می‌توان یافت که در تشکیل سیگنال کم‌ترین تاثیر را دارند در نهان‌نگاری صوتی استفاده شده است".

زیرا "همانند تصاویر، سیگنال‌های صوتی طبیعی نیز تمایل به توزیع نرمال دارند. علاوه بر آن سیگنال‌های صوتی واقعی، ترکیبی از چندین منبع صوتی مجزا می‌باشند که بر طبق اصل نرمال جمع بی نهایت سیگنال طبیعی با توزیع متفاوت، توزیع نرمال خواهد بود می‌توان این نتیجه را گرفت که سیگنال‌های طبیعی صوتی توزیع نرمال خواهند داشت".

پس با پیداکردن ضرایب کم اهمیت در سیگنال طبیعی صوتی می‌توان این ضرایب را به‌نحوی دلخواه تغییر داد که علاوه بر دارا بودن خاصیت درخت صفر حاوی اطلاعات داده مخفی نیز باشد.
بلوک دیاگرام مراحل الگوریتم پیشنهادی ذخیره داده مخفی درون سیگنال صوتی در شکل (۲) آورده شده است.



شکل (۲): بلوک دیاگرام مراحل الگوریتم پیشنهادی ذخیره داده مخفی درون سیگنال صوتی

در مرحله اول اعمال تبدیل موجک و به دست آوردن ضرایب آن می‌باشد. تعداد مراتب تبدیل موجک و نوع فیلتر مورد استفاده در تبدیل موجک بسیار متنوع می‌باشد. فیلترهای تبدیل موجک برای کاربردهای متنوع و سیگنال‌های متفاوت تنوع زیادی یافته اند به طوری که با دانستن ویژگی ساختاری سیگنال می‌توان ضرایب بهتری از تبدیل موجک به دست آورد. البته ویژگی که در تمامی فیلترهای تبدیل موجک رعایت می‌شود استفاده از بردارهای پایه متعامد می‌باشد که با تکیه بر این موضوع می‌توان به این نکته دست یافت که ضرایب به دست آمده از هر مرحله برای تخمین سیگنال اصلی یکتا خواهد بود. زیرا همبستگی بردارهای متعامد صفر می‌باشد. فشردگی و کشیدگی فیلترهای مورد استفاده وابسته به کاربرد مورد نظر متفاوت می‌تواند باشد. به عنوان مثال در تخمین سیگنال‌های صوتی بیشتر از فیلترهای db استفاده می‌شود زیرا فیلترهای db دارای ویژگی اغتشاش سریع زمانی (چولگی بالا) می‌باشد. همچنین در کاربردهایی که سیگنال دارای خاصیت متقارن می‌باشد می‌توان از فیلترهای Sym استفاده نمود. مرحله دوم استخراج ضرایب درخت صفر می‌باشد. در این قسمت ضرایب سیگنال صوتی باید به نحوی تغییر یابد که متناسب با فرمت ضرایب تصویر باشد. زیرا ضرایب تصویر دو بعدی می‌باشند و ضرایب سیگنال صوتی یک بعدی هستند. پس از پیاده‌سازی سیگنال صوتی در قالب تصویر الگوریتم درخت صفر جهت یافتن ضرایب درخت صفر اعمال می‌گردد. موضوع آخر مربوط به درج پیام مخفی درون سیگنال صوتی می‌باشد. در این قسمت به تعریف روش و الگویی جهت درج پیام مخفی پرداخته می‌شود. الگوی ذخیره‌سازی نقش مهمی در بازیابی داده مخفی و همچنین پایداری الگوریتم در برابر حملات و همچنین میزان ظرفیت نهان‌نگاری دارد. و در نهایت سیگنال حامل داده‌های مخفی ذخیره می‌شود. در بازیابی داده‌های نهان‌نگاری‌شده از درون سیگنال حامل برعکس مراحل بالا دنبال می‌شود. در ابتدا سیگنال صوتی حامل داده مخفی به فرمت داده‌های صوتی شکل داده می‌شوند و در ادامه الگوریتم درخت صفر برای یافتن مکان‌های ضرایب درخت صفر اعمال می‌شود و سپس بر طبق الگوریتم ذخیره‌سازی داده مخفی به استخراج پیام پرداخته می‌شود.

۳-۲- طراحی سیستم روش پیشنهادی

۳-۲-۱- طراحی الگوریتم ذخیره‌سازی پیام مخفی

در این بخش به پیشنهاد الگوریتم مربوطه پرداخته می‌شود و نتایج کلی به همراه آزمون آن آورده شده است. همان‌طور که پیشتر ذکر گردید این پیشنهاد الگوریتم بر پایه تکنیک‌های تبدیل موجک و درخت صفر بنا شده است. در این قسمت روند طراحی الگوریتم بر اساس تئوری‌های موجود بیان‌شده، و با استفاده از الگوریتم گام به گام آزمون حدسیات ذکر می‌گردد. همان‌طور که بیان شد، الگوریتم درخت صفر بر اساس یافتن ضرایبی از تبدیل موجک می‌باشد که کم‌ترین تاثیر را در ساختار والد و اولاد تبدیل موجک دارد. بنابر تعریف درخت صفر، اگر تمامی اولاد یک والد از مقدار والد خود کم‌تر باشد آنگاه والد و اولاد آن را درخت صفر می‌نامند که با توجه به تعریف، تنها مقدار والد در ساختار سیگنال موثر بوده و می‌توان از ضرایب اولاد صرف نظر کرد. با توجه به ساختار دو بعدی الگوریتم درخت صفر، می‌بایستی این الگوریتم را برای تبدیل موجک سیگنال صوتی مناسب نمود. با توجه به الگوریتم درخت صفر که بیشترین عملکرد آن معطوف به جزئیات و مولفه‌های بالاگذر سیگنال می‌شود، بهترین مکان جهت اضافه‌کردن سیگنال اصلی و ضرایب تبدیل موجک مرتبه اول آن، ابتدای دنباله به دست آمده Xw می‌باشد. در ادامه نیز نحوه چیدمان ضرایب درخت صفر نمایش داده شده است. فرض شود سیگنال دلخواه یک سیگنال تصادفی S به طول ۱۶ نمونه باشد. از این سیگنال در سه سطح و با استفاده از فیلتر 'haar' تبدیل موجک گرفته می‌شود و در ادامه ضرایب درخت صفر آن محاسبه می‌شود. به کمک MATLAB موضوعات کلی زیر پیاده‌سازی شد.:

- محاسبه تبدیل موجک از سیگنال تصادفی S در سه سطح با استفاده از فیلتر Haar
- چیدمان سیگنال X و بازسازی آن در دو بعد مطابق الگوریتم EZW
- محاسبه مکان ضرایب درخت صفر
- جایگذاری پیام (در این مثال آرایه ای از ۱) به جای ضرایب درخت صفر
- تبدیل معکوس گرفتن از ضرایب تبدیل موجک و بازیابی سیگنال جایگذاری‌شده با داده مخفی
- آزمون درستی بازیابی داده: اعمال الگوریتم درخت صفر بر روی سیگنال بازیابی‌شده و مقایسه مکان ضرایب درخت صفر به‌دست آمده با مکان ضرایب درخت صفر سیگنال اصلی

مراحل جایگذاری داده مخفی به ۵ بخش متفاوت تقسیم می‌شود:

۱- محاسبه تبدیل موجک در سه سطح از سیگنال مورد نظر

۲- چیدن ضرایب سیگنال مورد نظر به فرمت مناسب در دوبعد

۳- محاسبه درخت صفر ضرایب سیگنال

۴- جایگذاری داده مخفی در مکان‌های مورد نظر

۵- بازیابی سیگنال با استفاده از تبدیل معکوس تبدیل موجک

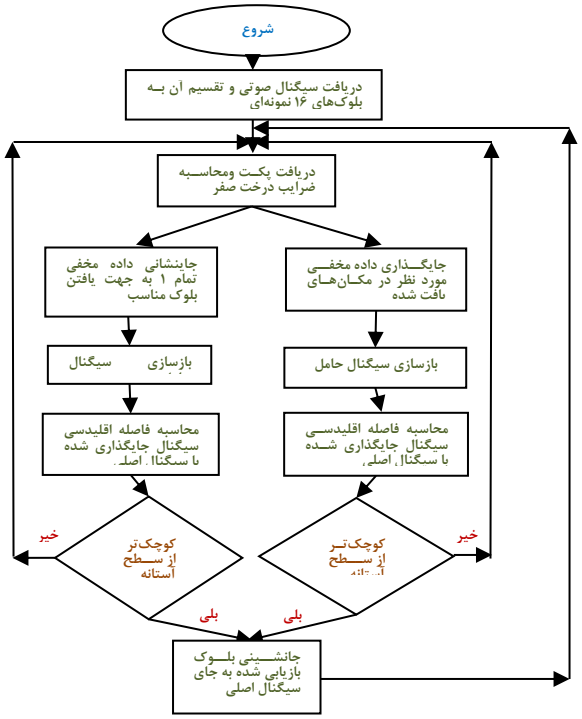
پس از جایگذاری داده مخفی آزمون درستی بازیابی داده پنهان‌شده یا همان Verification باید صورت پذیرد. یعنی آیا در بازیابی داده مخفی از ضرایب درخت صفر، مکان‌های جدید به دست آمده از اعمال تبدیل EZW بر روی سیگنال حامل اطلاعات مخفی با مکان‌های ضرایب درخت صفر سیگنال اصلی تطابق دارد یا خیر؟ برای این منظور ۳ مرحله بالا بر روی سیگنال بازیابی‌شده (سیگنال حامل اطلاعات مخفی) انجام می‌گیرد. در صورت درستی عملکرد بازیابی مکان‌های درخت صفر، میزان اغتشاش واردشده در سیگنال حامل و سیگنال اصلی با استفاده از محاسبه فاصله اقلیدسی میان این دو سیگنال محاسبه می‌شود. نتایج عددی الگوریتم سیگنال تصادفی S در فضای دو بعدی و همچنین ضرایب آن در مقیاس ۲۵۶ تا ۲۵۶ را بازسازی شد. ماتریس خروجی الگوریتم درخت صفر پس از حذف‌کردن ضرایب کم ارزش به‌دست آمد. در ادامه با استفاده از آن، و تبدیل معکوس تبدیل موجک، سیگنال اصلی بازسازی شد. در این حالت سطح آستانه را برای یافتن درخت صفر ضرایب سیگنال برابر یک شانزدهم بزرگترین مقدار سیگنال قرار دادیم. بدیهی است که برای بدست آوردن قدرت مقاومت بیشتر سطح آستانه را می‌توان بزرگتر انتخاب نمود. هرچه سطح آستانه بزرگتر باشد، مقاومت الگوریتم در برابر حملات بیشتر خواهد شد اما میزان خرابی سیگنال نسبت به قبل آن بیشتر خواهد شد.

همان‌طور که اشاره شد، پس از ترکیب سیگنال حامل با پیام مخفی باید صحت عملکرد الگوریتم در بازیابی رشته داده پیام مخفی سنجیده شود. بنابراین یک بار دیگر الگوریتم را بر روی سیگنال بازسازی‌شده اعمال شد. نتیجه حاصله از خروجی درخت صفر این فرایند باید با نتیجه درخت صفر سیگنال اصلی تطابق داشته باشد. یعنی هر دو خروجی درخت‌های صفر حاصل از الگوریتم بر روی دو سیگنال اصلی و بازسازی‌شده نتایج یکسانی را داشته است. حال که طرح اولیه طراحی الگوریتم خروجی مثبتی را در بر داشت، در ادامه، نحوه ذخیره‌سازی داده‌های پیام مخفی درون مکان‌های درخت صفر آورده شده است.

۳-۲-۲- پیاده‌سازی الگوریتم ذخیره‌سازی پیام مخفی

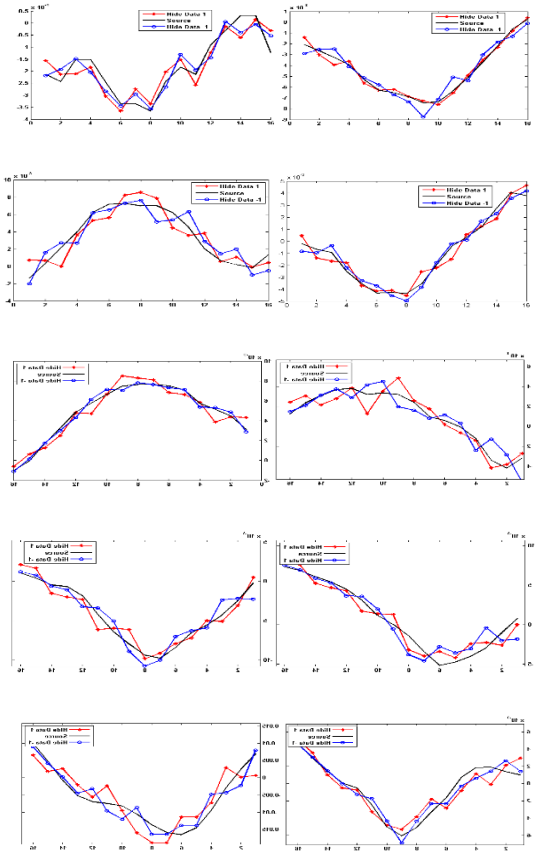
الف) الگوریتم پیشنهادی اول: در این مقاله از تبدیل موجک مرتبه سوم با فیلترهای db استفاده شده است. بدیهی است که برای به دست آوردن ضرایب مناسب می‌توان از فیلترهای دیگر تبدیل موجک استفاده نمود. اما به دلیل این که فیلترهای db سازگاری بیشتری با سیگنال‌های طبیعی دارند، این نوع فیلترها انتخاب شده‌اند. علاوه بر این، در بسیاری از مقاله‌های پژوهشی در رابطه با نهان‌نگاری و نشانه گذاری با استفاده از تبدیل موجک، فیلتر مورد استفاده، فیلترهای db بوده است. انتخاب درجه تبدیل موجک وابستگی زیادی به طول سیگنال در نظر گرفته شده دارد. معمولاً در پژوهش‌های انجام‌شده از مراتب ۳، ۴ و ۵ تبدیل موجک استفاده می‌شود [۱۲،۱۱]. در این‌جا به‌دلیل این که طول سیگنال اولیه را ۱۶ در نظر گرفتیم، طول ضرایب تبدیل موجک مرتبه ۳ آن با استفاده از فیلتر Haar برابر با ۴ خواهد شد که در استفاده و تخمین سیگنال صوتی اولیه از لحاظ محاسبات و پیچیدگی بهترین انتخاب می‌باشد. زیرا انتخاب ضریبی با طول ۴ به مفهوم تخمین یک نقطه از سیگنال با همسایگی ۴ تایی می‌باشد. یعنی هر ضریب از سیگنال اصلی با پنجره‌ای به طولی ۴ مولفه از فیلتر انتخاب‌شده تخمین زده می‌شود. در ذخیره‌سازی داده باید توجه زیادی به تاثیر پیام مخفی در شکل سیگنال بازسازی‌شده داشت. بنابراین برای آزمون کارایی الگوریتم بدترین شرایط را برای داده مخفی در نظر می‌گیرند، حالت‌هایی که معمولا در شرایط آزمون برای سیگنال مخفی در نظر می‌گیرند عبارتند از سیگنال مخفی با مولفه‌های تماماً یک، تماماً صفر یا یک در میان صفر و یک [۱۷،۱۶،۱۵،۱۴،۱۳]. برای درج بیت‌های پیام مخفی مکان‌هایی از درخت صفر در نظر گرفته می‌شوند که با مکان‌های قرار گیری ضرایب مولفه‌های تبدیل موجک مرتبه سوم سازگاری داشته باشند. همان‌طور که قبلا نیز ذکر شد، این امر به دلیل این است که بازسازی سیگنال اصلی از روی مولفه‌های مرتبه سوم صورت می‌پذیرد. در این قسمت پس از تولید سیگنال تصادفی دلخواه، تبدیل موجک مرتبه سوم آن با استفاده از فیلتر db1 محاسبه می‌شود و سپس از ضرایب آن برای انتخاب درخت صفر از الگوریتم EZW استفاده می‌شود. مکان‌هایی که از الگوریتم EZW حاصل می‌شوند با مکان‌های ضرایب تبدیل موجک مرتبه سوم تطبیق داده می‌شوند و ضرایبی را که متناظر با آن‌ها در ماتریس درخت صفر، مقدار صفر دارند، به عنوان مکان‌های درج پیام مخفی انتخاب می‌شوند.

در ادامه تمامی سیگنال درج‌شده را یکسان قرار داده می‌شود. در مرحله اول سیگنال درج‌شده را تماما یک در نظر گرفته و سپس سیگنال اصلی از روی این ضرایب تغییر یافته بازسازی می‌شود. برای آزمون نتیجه عملکرد الگوریتم مجددا از این سیگنال بازسازی‌شده تبدیل موجک در سه سطح گرفته‌شده و درخت صفر ریشه‌های آن بازیابی می‌گردد. در شکل (۳) روند شرح داده‌شده به همراه فلوچارت برنامه نمایش داده شده است.



شکل (۳): فلوچارت پیشنهادی اول

در شکل (۴) چند نمونه از عملکرد الگوریتم ذکرشده که با دنباله تماماً یک و تماماً صفر نهان‌نگاری‌شده، نمایش داده شده است.



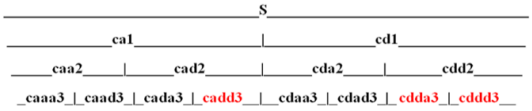
شکل (۴): چند نمونه از عملکرد الگوریتم پیشنهادی اول، با دنباله تماماً یک و تماماً صفر نهان‌نگاری‌شده

همان‌طور که مشاهده می‌شود، میزان تفاوت میان سیگنال‌های حامل با سیگنال اصلی ناچیز (در حدود 3-10) می‌باشد. در شکل ۴، منحنی بدون علامت و ساده مربوط به سیگنال اصلی است. در این آزمایش، سیگنال صوتی اصلی به بلوک‌های ۱۶ نمونه‌ای تقسیم‌بندی شده‌اند و در تمامی بلوک‌ها پیام مخفی به دو صورت دنباله ای از ۱ (منحنی با علامت *) و ۱- (منحنی با علامت دایره‌های کوچک) نهان‌نگاری شده است. دامنه ۱ به منظور درج کد ۱ و ۱- به منظور درج کد صفر است که در هنگام نهان‌نگاری دنباله پیام مخفی دامنه‌ها در نظر گرفته‌شده، در حد آستانه‌ای ضرب می‌شوند. اما نکته‌ای که در این‌جا باید مَد نظر گرفته شود این است که درج بیت پیام مخفی باعث خرابی در ساختار سیگنال صوتی نشود. زیرا خرابی و شکست در سیگنال زمانی همانند اضافه‌شدن تُن فرکانسی در طیف فرکانسی سیگنال و ایجاد اغتشاش و خرابی در سیگنال می‌شود. بنابراین بعد از انجام الگوریتم، فاصله اقلیدسی سیگنال بازسازی‌شده و سیگنال اصلی محاسبه می‌شود و در صورتی که مقدار این فاصله از یک میزان معین (حد آستانه) بیشتر نباشد، با توجه به الگوریتم درخت صفر بلوک درج‌شده را می‌توان به‌عنوان بلوک مساعد جهت نهان‌نگاری در نظر گرفت و می‌توان داده‌های بازیابی‌شده با استفاده از پیام مخفی را به‌جای داده‌های اصلی سیگنال قرار داد. در صورتی که میزان فاصله اقلیدسی بیشتر از حد استانه مورد نظر باشد سیگنال اصلی بدون تغییر باقی می‌ماند. در اینجا دلیل استفاده از فاصله اقلیدسی بدین جهت می‌باشد که تغییرات صورت پذیرفته بر روی ضرایب تبدیل موجک با استفاده از فیلتر Haar صورت پذیرفته است. بنابراین بازسازی دوباره سیگنال اصلی در حوزه زمان می‌باشد. بنابراین با توجه به تغییرات زمانی صورت پذیرفته‌شده بر روی سیگنال ساده ترین معیار جهت اندازه‌گیری تغییرات دو سیگنال در حوزه زمان استفاده از معیار فاصله اقلیدسی می‌باشد. البته لازم به ذکر است که استفاده از چندین معیار به عنوان استخراج کننده ویژگی دارای مزیت‌های بیشتری می‌باشد اما با اضافه‌کردن یک معیار جدید میزان پیچیدگی محاسبات و پیاده‌سازی آن چندین برابر می‌شود. با افزایش حد آستانه حجم ذخیره داده‌های پیام مخفی افزایش می‌یابد اما در عوض سیگنال بازسازی‌شده بیشتر دچار خرابی و اغتشاش می‌شود. بنابراین بین میزان حد آستانه کیفیت و ظرفیت نهان‌نگاری رابطه معکوس وجود دارد یعنی با افزایش ظرفیت نهان‌نگاری، کیفیت سیگنال صوتی کاهش پیدا می‌کند. در نتایج به‌دست آمده حد آستانه $0.1<Threshold<0.8$ عملکرد خوبی در نهان‌نگاری داده‌ها در هر دو وضعیت داده صفر و یک دارد. بنابراین در این مقاله بیشتر از این محدوده آستانه جهت آزمون الگوریتم استفاده شده است. بازیابی اطلاعات ذخیره‌شده عکس مراحل بالا را طی می‌کند.

بدین صورت که مکان‌های ضرایب درخت صفر سیگنال حامل، با استفاده از الگوریتم درخت صفر جستجو می‌شوند. در ادامه ضرایب درخت صفر به عنوان داده‌های نهان‌نگاری‌شده استخراج می‌شوند. در صورتی که داده‌ها بزرگتر از صفر باشند به عنوان داده نهان‌نگاری‌شده ۱ و در صورتی که داده‌ها کوچکتر از صفر باشند داده‌های به عنوان داده نهان‌نگاری‌شده صفر در نظر گرفته می‌شوند.

ب) الگوریتم پیشنهادی دوم: علاوه بر الگوریتم پیشنهادشده الگوریتم دیگری نیز قابل دستیابی می‌باشد. همان‌طور که قبلا ذکر شد داده‌های پیام به جای ضرایب درخت صفر در سیگنال صوتی جایگذاری می‌شوند. در الگوریتم پیشنهادشده به ازای هر ضریب درخت صفر یافته‌شده درون هر بلوک یک عدد از دنباله پیام مخفی قرار می‌گیرد. بنابراین احتمال خراب‌شدن داده مورد نظر در اثر حملات افزایش می‌یابد. در الگوریتم پیشنهادی دوم، به‌جای تمامی ضرایب درخت صفر که از یک بلوک به دست می‌آید تنها یک عدد از دنباله پیام مخفی ذخیره‌سازی شود در این صورت با تغییر یکی از ضرایب درخت صفر پیام مخفی دچار خرابی نخواهد شد.

ج) الگوریتم پیشنهادی سوم: مطابق شکل (۵) همچنین در آزمون‌های انجام‌شده مشاهده گردید که مکان‌هایی از درخت صفر نسبت به دیگر مکان‌ها بسیار تکرار می‌شوند. یعنی ضرایب این مکانها صرف نظر از مقادیر آن‌ها دارای ارزش ساختاری چندانی در سیگنال صوتی نمی‌باشند. این مکانها عبارتند از Cadd3، Cdda3 و Cddd3. این مکان‌ها در شکل زیر نمایش داده شده‌اند.



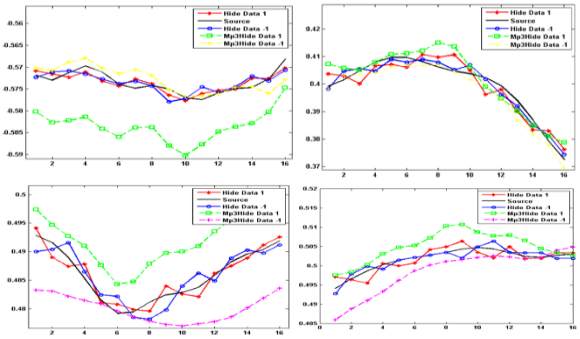
شکل (۵): مکان‌های درخت صفر در الگوریتم پیشنهادی سوم

همان‌طور که قبلاً نیز ذکر شد، مولفه‌های فرکانس پایین (DC) ساختار کلی سیگنال را تشکیل می‌دهند و مولفه‌های فرکانس بالا جزئیات سیگنال را تشکیل می‌دهد. ضرایب تبدیل موجک Cadd3 مربوط به تشکیل جزئیات فرکانس‌های پایین (DC) سیگنال صوتی می‌باشد پس بنابراین منطقی می‌باشد که در تشکیل ساختار اصلی سیگنال صوتی تاثیر چندانی نداشته باشد. همچنین Cdda3 و Cddd3 به ترتیب مربوط به تشکیل ساختار فرکانس پایین و فرکانس بالای جزئیات سیگنال صوتی (مولفه‌های فرکانس بالای سیگنال) می‌باشند. به دلیل این که باز جزئیات (مولفه‌های فرکانس بالا) تاثیر چندانی در تشکیل ساختار شکل موج صوتی ندارند بنابراین تجزیه این عوامل به زیر مجموعه فرکانس پایین و فرکانس بالا نیز اهمیت چندانی ندارند. با توجه به این مطلب در می‌یابیم تجزیه درخت صفر نتایج یکسانی را با مفاهیم واقعی دارد. بنابراین با صرف نظر از این ضرایب می‌توان داده مخفی را در این مکان‌ها نهان‌نگاری نمود. یکی از مزیت‌های این کار افزایش سرعت پردازش می‌باشد.

د) الگوریتم سرفایل گذاری پیام مخفی: در این قسمت به شرح نحوه سرفایل گذاری بر روی داده‌های پرداخته می‌شود. سرفایل گذاری داده‌ها از این لحاظ دارای اهمیت می‌باشد که ممکن است طول پیام مخفی از ظرفیت نهان‌نگاری کوچک‌تر باشد در این صورت برای دیکد پیام درج‌شده با مشکل روبه رو خواهیم شد. زیرا بازیابی اطلات به‌صورت کاملا کور صورت می‌پذیرد و در دیکدکردن داده مخفی به سیگنال اصلی اولیه که بر روی آن نهان‌نگاری اعمال شده است دسترسی نخواهیم داشت. بنابراین الگوریتم رمزگشایی داده مخفی تمامی طول ظرفیت نهان‌نگاری یک سیگنال را به عنوان داده مخفی بازیابی خواهد نمود. بنابراین باید الگوریتمی جهت مشخص نمودن ابتدا و انتهای داده مخفی مشخص نمود. همچنین از دیگر کاربردهای سرفایل گذاری بر روی پکت‌های داده مخفی می‌توان به نهان‌نگاری داده در چندین فایل اشاره نمود. به عبارت دیگر درصورتی که حجم یا ظرفیت نهان‌نگاری سیگنال صوتی کوچکتر از حجم پیام مخفی باشد، می‌توان پیام مخفی را به چندین پکت تقسیم نمود و هر پکت را به‌طور جداگانه درون فایل‌های متفاوت و مجزا ذخیره نمود. البته سرفایل در نظر گرفته‌شده باید دارای سه خصوصیت یکتایی، طول کوچک و افزونگی کم باشد. سرفایل در نظر گرفته‌شده نباید در طول داده‌های پیام مخفی وجود داشته باشد زیرا در غیر این صورت دنباله استخراجی اشتباه خواهد بود. علاوه براین طول سر پکت نباید از یک حدی بیشتر باشد. زیرا طول زیاد سرفایل از میزان ظرفیت نهان‌نگاری کاسته خواهد کرد. الگوریتم استانداردی که برای سرفایل گذاری داده‌های مخفی در نظر گرفته شده است، الگوریتم HDLC می‌باشد که یک الگوریتم استاندارد برای ارسال داده بر روی هر نوع خط ارتباطی می‌باشد. در این الگوریتم دنباله "01111110" به عنوان سر فایل هر نوع پکت داده اعم از داده‌های مخابراتی یا غیره در نظر گرفته می‌شود. زیرا دو خصوصیت طول کم و افزونگی کم را دارد. به دلیل این که این دنباله ممکن است درون داده‌های یک پکت ظاهر شود، قبل از اضافه‌کردن این سرفایل به داده‌ها عمل Bit Stuffing بر روی داده‌ها صورت می‌پذیرد. بدین صورت که ابتدا درون دنباله رشته داده‌های پیام به جستجوی دنباله "11111" پرداخته می‌شود و صرف نظر از این که مقدار داده چه عددی باشد یک عدد "0" به انتهای این دنباله اضافه می‌شود. در این صورت درون دنباله داده پیام مخفی دیگر ۶ عدد ۱ پشت سر هم نخواهیم داشت و تنها با اضافه‌کردن سرفایل به دنباله ۶ عدد ۱ پشت سر هم خواهیم داشت که نشانگر ابتدا و انتهای فایل خواهد بود. در این مقاله پس از دریافت پکت داده مخفی و

تبدیل آن به دنبـــاله‌ای از "0" و "1" عملیـات Bit Stuffing بـر روی دنباله اعداد انجام می‌شود و در نهایـت دنبالـه "01111110" به انتهای داده‌های پیام مخفی اضافه می‌شود.

در پیاده‌سازی الگوریتم پیشنهادی، در صورتی که مقدار خروجـی خطای حاصل از الگوریتم پیشنهادی صفر باشـد بـه مفهـوم ایـن است که بلوک دریافتی حاوی اطلاعـات پیـام مخفـی مـی‌باشـد. درصورتی که مقدار خطا مخـالف صـفر باشـد بلـوک مـورد نظـر مناسب نهان‌نگاری نمی‌باشد. در این مقاله با استفاده از این توابـع فایل MP3 به فایل صوتی تبدیل می‌شود و سپس عملیـات مـورد نظر بر روی فایل صوتی انجام می‌پذیرد. خـوبی ایـن کـار در ایـن است که بـا یکبـار MP3 کـردن فایـل صـوتی مقـادیر زیـادی از نمونه‌هایی که در فایل صوتی شنیده نمی‌شـوند و در طـی پخـش فایل ماسک می‌شوند، حذف خواهند شد. بنابراین این اطمینان را خواهیم داشت که سیگنال نهان‌نگـاری‌شـده در اثـر MP3 شـدن تغییرات چندانی نخواهد کـرد. در شـکل (۶) چنـد نمونـه بلـوک صوتی نهان‌نگاری‌شده که یک بار MP3 شـده اسـت و دوبـاره بـه فایل صوتی دیکد شـده اسـت را در مقایسـه بـا سـیگنال صـوتی نمایش داده شده است.



شکل (۶): چند نمونه بلوک صوتی نهان‌نگاری‌شده یک بار MP3 شده و دوباره به فایل صوتی دیکُدشده، در مقایسه با سیگنال صوتی

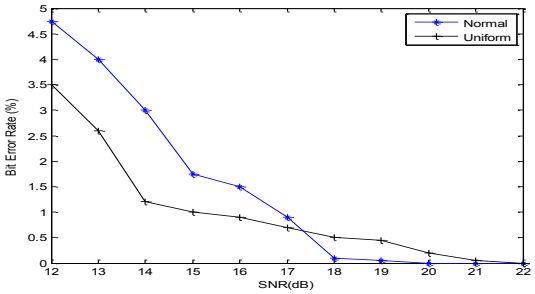
همان‌طور که در شکل (۶) مشاهده می‌شود، Mp3 نمـودن فـایلی صوتی که یک بـار Mp3 شـده اسـت تغییـرات چنـدانی نخواهـد داشت. در این شکل، نمودارها به‌صورت منحنی با علامت * بـرای پیام مخفی با دنباله‌ای از ۱، منحنی بدون علامـت و سـاده بـرای سیگنال اصلی، منحنی با علامت دایـره‌هـای کوچـک بـرای پیـام مخفی با دنباله‌ای از ۱-، منحنی با علامت مستطیل‌های کوچـک برای MP3 با دنباله‌ای از ۱ و منحنی با علامت + برایMP3 پیـام مخفی با دنباله‌ای از ۱- مشخص شده است.

۳-۳- ارزیابی الگوریتم پیشنهادی

۳-۳-۱- بررسی عملکرد الگوریتم پیشنهادی و پایـداری در برابر حملات

در این بخش به کارایی الگوریتم پیشـنهادی در برابـر حمـلات و نویز کانال پرداخته شد. همان‌گونه که قبلا هم اشاره شد، حملات در نهان‌نگاری به دو مفهوم و نوع کلی تقسیم‌بندی می‌شوند. نـوع اول، حملاتـی مـی‌باشـند کـه بـه منظـور تخریـب سـیگنال نهان‌نگـاری‌شـده بـر روی سـیگنال حامـل پیـام مخفـی صـورت می‌پذیرد. هدف از این نـوع حمـلات از بـین بـردن پیـام مخفـی درج‌شده درون سیگنال حامل و جلوگیری از ارسال پیام مخفی به مقصد می‌باشد. نوع دوم، حملاتی هستند که به منظـور کشـف و دیکدکردن پیام مخفی بر روی سیگنال حامل پیام مخفی صـورت می‌پذیرد. هدف از حملات نوع دوم به دست آوردن محتوی پیـام مخفی و آگاهی از آن می‌باشد. به دلیل این که بستر کاری در این مقاله سیگنال صوتی دیجیتال مـی‌باشـد، مبـاحثی از قبـل نـویز محیط، نویز کانال و نویز گیرنده بی مفهوم خواهد بود. به عبـارت دیگر، در برخورد با سیگنال‌های دیجیتال ایـن فرضـیات در نظـر گرفته می‌شود که سیگنال دیجیتال از لحاظ نـویز ایمـن خواهـد بود. زیرا در صورتی که چنین فرضیه ای در نظر گرفته نشـود بـه هیچ یک از فایل‌های دیجیتـال کـه بـر روی یـک بسـتر منطقـی ذخیره می‌شوند نمی‌توان اعتماد کرد. زیرا با جابجایی و تغییـر در یک بیت از فایل دیجیتال، محتوای کل فایل از بین خواهد رفـت. از دیدگاهی دیگر، بررسی نویز میزان قابلیت اعتماد بـه الگـوریتم ذخیره‌سازی و نهان‌نگاری را افـزایش مـی‌دهـد. بنـابراین از ایـن دیدگاه، بررسی نویز و تاثیر آن بر سیگنال نهان‌نگاری‌شده خالی از لطف نخواهد بـود. از ایـن رو در ادامـه بـه بررسـی ایـن موضـوع خواهیم پرداخت. بررسی و آزمون حملات بر روی سیگنال حامل پیام مخفی در این مقاله به دو صورت اضافه‌کردن نویز سـفید بـه سیگنال و اندازه‌گیری میزان SNR سیگنال و همچنین حملات از نوع اضافه‌کردن مولفه به نمونه‌های سیگنال حامل پیـام مخفـی و اندازه گیری میزان خطـای بازیـابی داده مخفـی صـورت خواهـد پذیرفت.

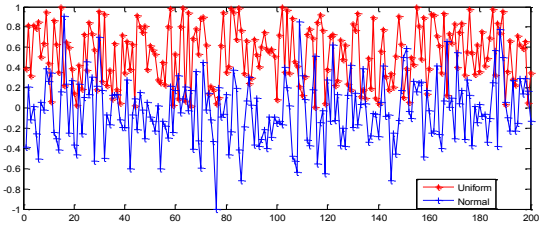
- **افزودن نویز سفید و بررسی میزان سیگنال به نـویز:** در این آزمون که با نرم افزار Matlab صـورت انجـام مـی‌شـود یـک سیگنال صوتی با فرکانس ۴۴۱۰۰ هرتز و کوانتیزاسیون ۱۶ بیت در نظر گرفته شده است. نویز اعمالی دو نوع نویز سفید نرمال، و نویز سفید یکنواخت انتخاب شده است. نتایج این تحقیق مطابق شکل (۷) می‌باشد.



شکل (۷): افزودن نویز سفید و بررسی میزان سیگنال به نویز

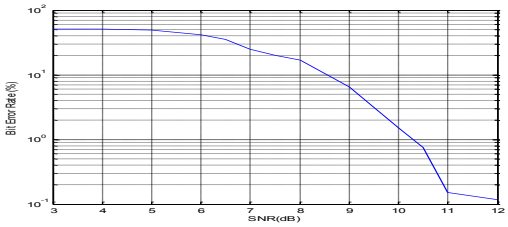
همان‌طور که از شکل (۷) مشخص است، پایداری الگوریتم پیشنهادی در برابر نویز سفید نرمال (منحنی با علامت +) بیشتر از نویز یکنواخت (منحنی با علامت *) می‌باشد.

در شکل (۸) دو نمونه از توزیع‌های به‌کار برده‌شده نمایش داده شده است. حالت نرمال با علامت + و حالت یکنواخت با علامت * مشخص شده است.



شکل (۸): توزیع به‌کار برده‌شده نویز سفید نرمال، و نویز سفید یکنواخت

در ادامه به دلیل کاربردی‌تر بودن روش دوم از الگوریتم پیشنهادی، به بررسی بیشتر میزان تاثیر نویز بر روی این الگوریتم و آزمون‌های آن پرداخته شده است. نمونه‌های مورد استفاده برای آزمون آماری ۱۰ نمونه صوتی متفاوت می‌باشد. نمونه‌های صوتی شامل ۲ نمونه‌ی صوتی با گفتار مرد، ۲ نمونه ی صوتی با گفتار زن، ۲ نمونه ی صوتی با گفتار کودک، ۲ نمونه ی صوتی موسیقی و ۲ نمونه ی صوتی مختلط می‌باشد. نرخ خطای بیت (BER) برای ۱۰ بار تکرار به‌صورت متوسط برای هر کدام از نمونه‌ها بر حسب لگاریتم محاسبه شده است. نتایج محاسبات خطا در شکل (۹) آورده شده است.



شکل (۹): نتایج محاسبات خطا

همان‌طور که مشاهده می‌شود، الگوریتم پیشنهادی در برابر SNR های بالاتر از 10db نرخ بیت خطای کم‌تر از ۱ بیت دارد.

○ **بررسی عملکرد الگوریتم پیشنهادی در برابر خطای افزایشی**: در این قسمت مقاومت الگوریتم پیشنهادی را در برابر حملات افزایشی بررسی می‌کنیم. نحوه عملکرد حملات افزایشی بدین صورت می‌باشد که در بین داده‌های سیگنال صوتی بلوک‌هایی به‌صورت تصادفی اضافه می‌گردد. بدین شکل، در صورتی که الگوریتم به کار رفته نسبت به مکان چه در حوزه زمان و چه در حوزه فرکانس ایمن نباشد، بازیابی اطلاعات مخفی به‌طور صحیح صورت نخواهد پذیرفت. به عنوان مثال در صورتی که از روش LSB برای نهان‌نگاری استفاده شود، با اضافه‌کردن یک نمونه داده در میان نمونه‌های سیگنال صوتی اصلی، تمامی اطلاعات مخفی‌شده درون سیگنال حامل در صورتی از الگوریتم‌های تصحیح خطا استفاده نشده باشد از بین خواهد رفت. این در حالی است که رخداد خطا به‌صورت انتشاری در میان تمامی داده‌ها پخش خواهد شد. بنابراین با توجه به اهمیت پایداری الگوریتم در برابر حملات تخریبی، در این بخش به بررسی پایداری الگوریتم پیشنهادی در برابر نویز افزایشی خواهیم پرداخت. نحوه ایجاد نویز افزایشی به چندین صورت مختلف امکان‌پذیر می‌باشد. بنابراین از میان روش‌های مختلف، چند روش زیر برای آزمون این قسمت در نظر گرفته شده است:

الف) اضافه نمودن یک نمونه تصادفی در میان داده‌های سیگنال صوتی: در این قسمت بدون درنظر گرفتن محتوای سیگنال، نمونه‌هایی در میان نمونه‌های سیگنال اصلی درج می‌شود. با توجه به این که تعداد حالت‌های درج نمونه در این حالت بسیار زیاد می‌باشد، سه روش زیر برای آزمون این بخش انتخاب شد:

الف-۱) اضافه‌کردن یک نمونه معین به‌صورت یک در میان در بین نمونه‌های سیگنال صوتی: با توجه به این که حملات صورت پذیرفته‌شده نباید کیفیت سیگنال صوتی را مخدوش کند، بنابراین در این حالت به‌صورت یک در میان متوسط هر دو نمونه مجاور محاسبه می‌گردد و سپس این مقدار به عنوان نمونه اضافه‌شده به سیگنال اصلی در نظر گرفته می‌شود. با این روش، کیفیت سیگنال صوتی تغییری نخواهد داشت و تنها حجم فایل صوتی دو برابر می‌شود. این روش بر روی ۱۵ نمونه سیگنال صوتی متفاوت به‌صورت تصادفی انجام شده است. رابطه (۳) نحوه به دست آوردن نمونه اضافه‌شده را نمایش می‌دهد.

$$YT(N)=(YS(N-1)+ YS(N+1))/2; \qquad (۳)$$

در این حالت درصورتی که نرخ بیت نمونه برداری سیگنال اصلی به ۲ برابر حالت قبل تغییر کند، سیگنال صوتی از لحاظ کیفیت تغییر چندانی نخواهد کرد.

الف-۲) اضافه‌کردن یک نمونه معین به‌صورت تصادفی در بین نمونه‌های سیگنال صوتی: تفاوت این روش با روش الف در انتخاب مکان‌های جایگذاری نمونه مورد نظر می‌باشد. در این روش به تعداد ۱۰ درصد از کل نمونه‌های سیگنال صوتی، نمونه مورد نظر در بین نمونه‌های سیگنال صوتی جایگذاری می‌شود.

الف-۳) اضافه‌کردن یک نمونه تصادفی به‌صورت تصادفی در بین نمونه‌های سیگنال صوتی: در این روش، علاوه براین که دامنه‌های نمونه درج‌شده به‌طور تصادفی انتخاب می‌گردند، مکان‌های درج نیز به‌طور تصادفی به تعداد ۱۰ درصد از کل نمونه‌های سیگنال صوتی انتخاب می‌شوند.

درصد نرخ خطای بیت برای سه روش اضافه نمودن یک نمونه تصادفی در میان داده‌های سیگنال صوتی به ترتیب الف-۱

برابر با ۵۷، الف-۲ برابر با ۲,۵ و الف-۳ برابر با ۱۱ حاصل شد. همانطور که مشاهده می‌شود، در صورتی که به طور یک در میان به نمونه‌های سیگنال اصلی متوسط دو نمونه مجاور اضافه گردد، درصد خطا به نحو چشمگیری افزایش می‌یابد. این روش، حمله موفقی خواهد بود در صورتی که نرخ بیت نمونه برداری سیگنال صوتی به دو برابر افزایش یابد که با این وجود کیفیت سیگنال اصلی تغییر چندانی نخواهد کرد. اما به دلیل این که دیکد داده‌های پیام مخفی با نیمی از داده‌های اصلی درج‌شده صورت می‌پذیرد، درصد خطا به طور قابل ملاحظه ای افزایش می‌یابد. در صورتی که نرخ بیت سیگنال صوتی بدون تغییر در نظر گرفته شود با توجه به یک درمیان اضافه‌کردن نمونه‌ها می‌توان این نتیجه را گرفت که با این کار سیگنال صوتی را به طور کل تخریب نموده ایم. در این دو حالت به میزان صد درصد به سیگنال، داده افزوده می‌شود. و همانطور که BER نشان می‌دهد، آشکارسازی پیام مخفی کاملا تصادفی می‌شود. به عبارت دیگر سیگنال مخفی کاملا تخریب می‌شود.

ب) اضافه نمودن یک بلوک داده تصادفی در میان بلوک‌های سیگنال صوتی: در این روش یک بلوک با داده‌های تصادفی در میان داده‌های سیگنال صوتی اضافه می‌گردد. با توجه به این نکته که الگوریتم پشنهادی بر پایه بلوک‌های ۱۶ نمونه ای از سیگنال صوتی استوار است، بنابراین بلوک‌های نمونه درج‌شده را نیز با همین طول در نظر می‌گیریم. تعداد بلوک‌های در نظر گرفته‌شده در این قسمت برابر ۱۰ درصد تعداد کل بلوک‌های ۱۶ تایی موجو در سیگنال می‌باشد. همچنین دامنه نمونه‌های موجود درون بلوک‌های تصادفی به‌صورت نویز سفید گوسی انتخاب شده است. درصد نرخ خطای بیتِ اضافه نمودن یک بلوک داده تصادفی در میان بلوک‌های سیگنال صوتی در روش بلوک تصادفی نویز گوسی با طول ۱۶ نمونه، برابر با ۲ به‌دست آمد. همان‌طور که از نتایج مشهود است، روش پیشنهادی در برابر اضافه‌کردن نویز افزایشی به‌صورت بلوک مقاوم به‌نظر می‌رسد. علت این امر بدین خاطر می‌باشد که بلوک‌های اضافه‌شده به‌صورت نویز گوسی فاقد ویژگی بلوک‌هایی است که الگوریتم پیشنهادی بر روی این بلوک‌ها عملیات نهان‌نگاری را پیاده‌سازی می‌کند. به‌عبارت دیگر نویزهای گوسی شکل قله مانند خود را تاحدی حفظ می‌کنند.

با توجه به الگوریتم درخت صفر و نتایج آزمایشات، چنین به‌نظر می‌رسد که بلوک‌هایی مساعد نهان‌نگاری می‌باشند که دارای شیب تقریبا ثابتی باشند و در ابتدا یا انتهای بلوک دارای پیک کوچکی باشند. از اینرو هرچند بلوک نویز گوسی که سیگنال صوتی اضافه شود، تاثیر چندانی بر الگوریتم درج پیام مخفی پیشنهادی نخواهد داشت.

ج) اضافه نمودن بلوک‌های سکوت در میان پیچ‌های صوتی سیگنال صوتی: در این روش بلوک‌هایی با توزیع یکنواخت در میان بلوک‌های سیگنال صوتی حامل پیام مخفی اضافه می‌شود. به دلیل این که بلوک‌های سکوت تقریبا دارای شکل یکنواخت و مقدار دامنه کوچک هستند، در اینجا برای بررسی بیشتر الگوریتم پیشنهادی، از توزیع یکنواخت در انتخاب نمونه‌های بلوک درج‌شده استفاده شده است. با توجه به این که معمولا بلوک‌های سکوت پیش و یا پس از یک گام صوتی در سیگنال صوتی ظاهر می‌شوند، بنابراین بهترین مکان جهت اضافه‌کردن بلوک درج‌شده پیش و یا پس از بلوک‌های سکوت به‌نظر می‌رسد. در ادامه برای یافتن بلوک‌های سکوت از دو تکنیک ساده آماری میانگین و واریانس نمونه‌ها استفاده شده است. در اینجا، بلوک‌هایی که میانگین آماری آن‌ها از یک سطح آستانه کوچکتر باشند و همچنین واریانس نمونه‌های آن بزرگتر از ۰,۷۳ باشد به عنوان بلوک سکوت در نظر گرفته می‌شود. درصد نرخ خطای بیت برای حالت اضافه نمودن بلوک‌های سکوت در میان پیچ‌های صوتی سیگنال صوتی در روش بلوک سکوت نویز یکنواخت با طول ۱۶ نمونه برابر با ۱,۲۵ به‌دست آمد.

نتیجه این بخش نیز مانند بخش قبلی مشابه می‌باشد. زیرا بلوک‌هایی که دارای توزیع یکنواخت هستند، مستعد درج پیام مخفی با استفاده از روش الگوریتم پیشنهادی نخواهند بود.

۳-۳-۲- بررسی عملکرد الگوریتم پیشنهادی در برابر فشرده‌سازی و MP3

در این قسمت به پایداری الگوریتم پیشنهادی در برابر فشرده‌سازی و MP3 پرداخته شده است. همان‌طور که در قبل شرح داده شد، پیام مخفی بر روی نمونه‌های PCM صوت درج می‌شوند. نحوه ذخیره‌سازی اطلاعات بدین صورت می‌باشد که در ابتدا فایل MP3 انتخاب‌شده را به فرمت WAV تبدیل کرده و سپس الگوریتم پیشنهادی بر روی آن اعمال می‌شود.

در ادامه خروجی الگوریتم که به فرمت WAV می‌باشد به MP3 تبدیل می‌شود. در انتها برای اندازه گیری و محاسبه میزان خطای حاصله، فایل MP3 به دست آمده، دوباره به WAV تبدیل می‌شود و با استفاده از این فایل الگوریتم دیکدکننده پیشنهادی بر روی آن اعمال می‌شود. در نهایت میزان خطای حاصله بر حسب بیت محاسبه می‌شود. برای انجام این آزمون از فایل‌های صوتی با نرخ بیت ۱۲۸ کیلو بیت بر ثانیه استفاده شده است.

نمونه‌های صوتی شامل ۲ نمونه ی صوتی با گفتار مرد، ۲ نمونه ی صوتی با گفتار زن، ۲ نمونه‌ی صوتی با گفتار بچه، ۲ نمونه‌ی صوتی بی کلام و ۲ نمونه‌ی صوتی مختلط می‌باشد. در جدول (۱) نتایج آزمون فشرده‌سازی MP3 برای مقادیر متفاوت سطح آستانه آورده شده است.

جدول (۱): نتایج آزمون فشرده‌سازی MP3 برای مقادیر متفاوت سطح آستانه

سیگنال	نرخ خطای بیت (٪)		
نرخ نمونه برداری / سطح آستانه	۱۲۸kbps	۹۶kbps	۶۴kbps
۰,۵	۱۵,۳	۲۲,۴	۲۴,۱
۰,۷	۹,۸	۲۱,۸	۲۰,۳
۰,۹	۶,۶	۱۸,۹	۲۶,۶
۱,۱	۷,۲	۱۷,۱	۱۹,۷

۳-۳-۳- بررسی الگوریتم پیشنهادی از لحاظ آماری

از آن‌جایی که صوت به دسته‌های گوناگونی طبقه‌بندی می‌شود، بنابراین برای بررسی جامع تر الگوریتم پیشنهادی ۲۰ قطعه صوتی در نظر گرفته شد. این مجموعه شامل ۵ قطعه صوتی آهنگ (بی کلام)، ۵ قطعه صوتی گفتار زن، ۵ قطعه صوتی گفتار مرد و ۵ قطعه موسیقی آمیخته گفتار و آهنگ می‌باشد. در ادامه تمامی روش‌های آزمون بر روی این مجموعه اعمال می‌گردد.

○ بررسی آماری الگوریتم پیشنهادی در حوزه فرکانس

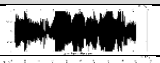
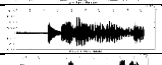
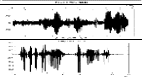
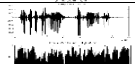
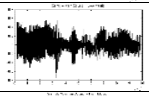
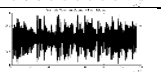
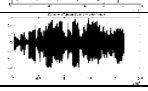
مقیاس فاصله کپستروم CDM: همانطور که می‌دانیم این معیار بیانگر میزان هموای طیف سیگنال می‌باشد. در رابطه (۴) ضرایب کپستروم سیگنال اصلی را $C_x(k)$ و ضرایب کپستروم سیگنال نهان‌نگاری‌شده را $C_y(k)$ مینامیم. $d(c_x, c_y, m)$ فاصله بین ضرایب کپستروم فریم m نامیده می‌شود.

$$CD = \frac{\sum_{m=1}^{M} w(m)\, d(c_x, c_y, m)}{\sum_{m=1}^{M} (m)} \qquad (۴)$$

$$d(c_x, c_y, m) = \left[[c_x(0) - c_y(0)]^2 + 2 \sum_{k=1}^{L} [c_x(k) - c_y(k)]^2 \right]^{\frac{1}{2}}$$

همان‌طور که در رابطه (۴) مشاهده می‌شود، آنالیز کیفیت صوت به روش CDM، برای هر قطعه از صوت (فریم) به‌طور جداگانه در نظر گرفته می‌شود. یعنی صوت را به فریم‌های با طول دلخواه تقسیم می‌کنند و فاصله اقلیدسی بین فریم نهان‌نگاری‌شده و اصلی را محاسبه می‌کنند همچنین به دلیل این که انرژی هر فریم (بلندی) متفاوت می‌باشد، هر فریم را در متوسط انرژی آن که به‌صورت w در فرمول فوق نمایش داده شده است، ضرب می‌کنند. در اینجا به دلیل آن که فایل‌های صوتی از قبل دسته‌بندی شده‌اند و طول هر فایل صوتی کوتاه در نظر گرفته شده است، نیازی به فریم‌بندی فایل صوتی نداریم. علاوه بر این برای درک محسوس تر و همچنین مقایسه بین فایل‌های صوتی متفاوت باید معیارهای به دست آمده را نرمالیزه کرد. به‌منظور این هدف معیار CDM به دست آمده را بر طول هر فایل صوتی تقسیم کرده و میانگین تعداد فایل‌های مشابه از یک نوع را محاسبه می‌کنیم تا در مقایسه بین فایل‌های صوتی مختلف، دید بهتری به ما دهد. نتایج بررسی این روش در جدول (۲) آورده شده است.

جدول (۲): نتایج بررسی آماری الگوریتم پیشنهادی در حوزه فرکانس با مقیاس فاصله کپستروم

	معیار CDM			
دنباله پیام مخفی / نوع صوت	شکل موج	تماماً یک	تماماً صفر	یکی در میان یک و صفر
موزیک با تن بالا		۴,۵۴۲*۱۰$^{-۴}$	۶,۴۳۱*۱۰$^{-۴}$	۷,۳۲۱*۱۰$^{-۴}$
موزیک با تن ملایم		۴,۵۵۴*۱۰$^{-۵}$	۷,۳۳۸*۱۰$^{-۵}$	۲,۴۵۶*۱۰$^{-۵}$
گفتار زن		۷,۳۴۲*۱۰$^{-۴}$	۵,۳۳۱*۱۰$^{-۴}$	۷,۲۲*۱۰$^{-۴}$
گفتار مرد		۵,۳۳۱*۱۰$^{-۳}$	۶,۳۲۴*۱۰$^{-۳}$	۷,۳۴۲*۱۰$^{-۳}$
موزیک تن بالا با گفتار مرد		۱,۴۴۲*۱۰$^{-۴}$	۱,۳۲۴*۱۰$^{-۳}$	۴,۳۴۲*۱۰$^{-۴}$
موزیک تن پایین با گفتار مرد		۱,۳۴۹*۱۰$^{-۴}$	۱,۰۰۸*۱۰$^{-۳}$	۴,۲۳۱*۱۰$^{-۴}$
موزیک تن بالا با گفتار زن		۵,۲۲۱*۱۰$^{-۳}$	۱,۵۴۱*۱۰$^{-۴}$	۳,۸۳۷*۱۰$^{-۳}$
موزیک تن پایین با گفتار زن		۸,۳۲۴*۱۰$^{-۳}$	۹,۵۳۴*۱۰$^{-۳}$	۸,۸۷۴*۱۰$^{-۳}$
مختلط		۲,۶۳۵*۱۰$^{-۳}$	۸,۳۳۳*۱۰$^{-۳}$	۷,۶۱۲*۱۰$^{-۴}$

همان‌طور که از متوسط معیار محاسبه‌شده در جدول (۲) به‌نظر می‌رسد، فایل‌های صوتی به فرم موزیک با تن (بلندی) ملایم کم‌ترین تغییرات را در معیار همواری صوت دارد. این امر بدین خاطر است که موزیک با تن ملایم همواری ملایمی دارد. فاصله ایتاکورا- سایتو ISD و فاصله COSH: معیار ISD فاصله بین طیف توان سیگنال نهان‌نگاری‌شده و سیگنال اصلی را محاسبه می‌کند و COSH نسخه ی متقارنی از فاصله ISD می‌باشد. فرمول آن‌ها به‌شکل رابطه (۵) بوده که نتایج بررسی طبق جدول (۳) به‌دست آمده است:

$$IS = \int_{-\pi}^{\pi} \left(\log\frac{Y(w)}{X(w)} + \frac{X(w)}{Y(w)} - 1\right)\frac{dw}{2\pi} \quad (۵)$$

و

$$COSH = \int_{-\pi}^{\pi} \left[\frac{1}{2}\left(\frac{Y(w)}{X(w)} + \frac{X(w)}{Y(w)}\right) - 1\right]\frac{dw}{2\pi}$$

جدول (۳): نتایج بررسی فاصله ایتاکورا - سایتو

معیار	ISD / COSH		
دنباله پیام مخفی / نوع صوت	تماماً یک	تماماً صفر	یکی در میان یک و صفر
موزیک با تن بالا	-۲,۳۶۱ / -۳,۵۶۸	-۲,۵۳۱ / -۲,۹۸۸	-۳,۰۱۴ / -۱,۲۳۳
موزیک با تن ملایم	۶,۲۴۷ / ۷,۴۳۲	۶,۲۳۱ / -۳,۲۲۵	-۱۰,۳۴۴ / -۸,۳۴۲
گفتار زن	-۱,۲۴۵ / ۰,۱۱۱	-۱,۲۲۵ / -۱,۸۸۶	۱۵,۵۶۷ / ۸,۳۳۲
گفتار مرد	-۲۳,۲۲۷ / ۱۲,۶۳۵	-۲۱,۷۸۸ / -۲۳,۸۴۴	۲۰,۱۳۴ / ۷,۰۰۱
موزیک تن بالا با گفتار مرد	-۱,۴۶۳ / -۱,۱۶۵	-۱,۳۴۹ / -۱,۲۳۱	-۴,۵۲۱ / -۱,۳۲۴
موزیک تن پایین با گفتار مرد	۱۳,۲۷۴ / ۷,۲۴۱	۱۲,۱۴۵ / ۷,۳۲۱	-۱۷,۰۲۳ / -۳,۲۲۳
موزیک تن بالا با گفتار زن	-۴,۵۴۹ / -۵,۳۲۸	-۳,۱۹۵ / -۱,۴۵۹	-۱۵,۳۳۶ / -۶,۲۴۲
موزیک تن پایین با گفتار زن	-۱,۵۱۹ / -۱,۵۵۸	-۱,۴۳۴ / -۱,۶۵۶	-۱,۴۳۸ / -۴,۵۴۸
مختلط	-۵,۳۲۶ / -۳,۶۵۲	-۷,۲۲ / -۲,۹۸۲	۷,۷۵۲ / ۵,۹۵۳

تبدیل رادون یک سیگنال در واقع تصویرکردن سیگنال مورد نظر در زاویه مورد نظر می‌باشد. بنابراین این معیار فاصله میزان تشابه تبدیل فوریه سیگنال را با چرخش‌های متفاوت آن در زوایای مختلف محاسبه می‌کند. جدول (۴) نتایج بررسی این معیار را نشان می‌دهد.

جدول (۴): نتایج بررسی معیار تبدیل فوریه- رادون زمان کوتاه

معیار / ۴۵ درجه / ۹۰ درجه	تماما یک	تماما صفر	یکی در میان یک و صفر
دنباله پیام مخفی / نوع صوت			
موزیک با تن بالا	۰,۰۱۰۱ / $۳,۹۴۰*۱۰^{-۳}$	۰,۰۱۰۸ / $۱,۹۷۱*۱۰^{-۳}$	۰,۰۱۱۳ / $۲,۹۷۳*۱۰^{-۳}$
موزیک با تن ملایم	۰,۰۱۵۴ / ۰,۰۱۱۳	۰,۰۱۲۲ / ۰,۰۱۱۷	$۹,۹۲۸۴*۱۰^{-۳}$ / ۰,۰۰۲۹
گفتار زن	۰,۰۱۸ / $۷,۴۸۲۲*۱۰^{-۳}$	$۹,۵۹۹۹*۱۰^{-۴}$ / $۲,۰۹۷۳*۱۰^{-۴}$	$۸,۰۰۷۵*۱۰^{-۴}$ / $۷,۸۲۲۷*۱۰^{-۴}$
گفتار مرد	۰,۰۰۳۲ / ۰,۰۲۳۴	$۸,۱۷۸۸*۱۰^{-۳}$ / $۵,۰۰۴۸*۱۰^{-۳}$	$۱,۷۷۰۲*۱۰^{-۳}$ / $۶,۹۹۲۱*۱۰^{-۳}$
موزیک تن بالا با گفتار مرد	$۳,۸۳۸۶*۱۰^{-۳}$ / $۷,۰۷*۱۰^{-۳}$	$۸,۲۲۳۸*۱۰^{-۳}$ / $۳,۳۲۲۶*۱۰^{-۳}$	$۴,۹۹۲*۱۰^{-۳}$ / $۷,۳۴۹*۱۰^{-۳}$
موزیک تن پایین با گفتار مرد	۰,۰۱۲۳ / ۰,۱۰۰۲	$۹,۲۳۱۴*۱۰^{-۳}$ / $۱,۶۰۱۶*۱۰^{-۳}$	$۱,۴۳۵۰*۱۰^{-۳}$ / $۳,۱۹۹۷*۱۰^{-۳}$
موزیک تن بالا با گفتار زن	$۰,۸۶۴*۱۰^{-۳}$ / $۳,۲۱۹*۱۰^{-۳}$	$۶,۱۹۲*۱۰^{-۳}$ / $۳,۲۹۱*۱۰^{-۳}$	۰,۰۰۱۲ / $۹,۰۳۴*۱۰^{-۳}$
موزیک تن پایین با گفتار زن	۰,۰۰۲۴ / ۰,۰۱۰۳	$۶,۳۲۹*۱۰^{-۳}$ / $۵,۸۶*۱۰^{-۳}$	$۲,۳۴۱*۱۰^{-۳}$ / $۴,۴۶۹*۱۰^{-۳}$
مختلط	$۹,۵۸۳۰*۱۰^{-۴}$ / $۹,۳۳۲*۱۰^{-۳}$	$۷,۵۴۲۶*۱۰^{-۴}$ / $۳,۳۳۱*۱۰^{-۳}$	$۶,۲۵۵۳*۱۰^{-۴}$ / $۹,۳۴۲*۱۰^{-۳}$

با توجه به خصوصیت الگوریتم تبدیل رادون و این که تبدیل رادون سیگنال داده شده را به فضایی با زاویه داده شده تبدیل می‌کند، نمونه‌های جدید به فضایی با بردار واحد وابسته نگاشته می‌شوند. به عبارت دیگر در این فضای جدید می‌توان میزان پراکندگی نمونه‌ها را صرف نظر از وابستگی بردارهای پایه نسبت به یکدیگر به دست آورد. با توجه به در نظر گرفتن دو زاویه ۴۵ درجه و ۹۰ درجه برای تبدیل رادون، بدین صورت می‌توان در مورد سیگنال‌های مورد آزمایش بحث نمود که با این آزمون می‌توان میزان پراکندگی سیگنال را حول دامنه صفر به دست آورد. در ادامه با توجه به این که سیگنال استفاده‌شده در تبدیل رادون، ضرایب تبدیل فوریه سیگنال صوتی می‌باشد، مفهوم این آزمون به معنی اندازه گیری میزان فرکانس‌های سیگنال با دامنه کوچک می‌باشد. این معیار تا حدی مشابه معیار اندازه گیری تن‌های زیر سطح آستانه شنوایی می‌باشد با این تفاوت که در نمودار آستانه شنوایی سطح آستانه به ازای فرکانس‌های مختلف میزانی متفاوت دارد. بنابراین در صورتی که میانگین تبدیل رادون یک سیگنال به صفر نزدیک باشد به مفهوم این است که سیگنال مورد آزمون از همواری نسبتا یکنواختی برخوردار است. در مثال جدول (۷) با توجه با این که تمامی ضرایب در یک رنج و نزدیک به مقدار صفر می‌باشند، می‌توان این نتیجه را گرفت که افزایش پیام مخفی تاثیر چندانی در ایجاد اغتشاش در حوزه فرکانس ندارد.

- **بررسی آماری الگوریتم پیشنهادی در حوزه ادراک**

معیار BSD و MBSD: طبق بحث‌های قبل، معیار BSD بر پایه محاسبه اختلاف دامنه بین دو سیگنال در حوزه طیف بارک[1] می‌باشد. بر اساس مدل فیزیوآکوستیک[2]، طیف فرکانسی صوت به

[1] طیف بارک یک تقسیم‌بندی فرکانسی است که با توجه به مدل مطلق شنوایی محدوده فرکانسی صوت را به نواحی بحرانی مختلفی تقسیم می‌نماید.

[2] Physioacoustic

نواحی یا زیر باندهایی تقسیم می‌شود که سطح بلندی شنوایی در آن نواحی متفاوت می‌باشد. برای محاسبه معیار BSD، طیف فرکانسی صوت مورد نظر را متناظر با طیف بارک، ناحیه‌بندی کرده و سپس فاصله اقلیدسی میان مجموعه مولفه‌های فرکانسی سیگنال اصلی با سیگنال نهان‌نگاری‌شده در آن نواحی محاسبه می‌شود. فرمول محاسبه آن طبق رابطه (۶) است:

$$D_{xy} = \left|S_x(i) - S_y(i)\right| \quad (۶)$$

که S طیف توان زیر باند مورد نظر می‌باشد.

تنها تفاوت معیار MBSD با BSD در نسبت دادن وزن‌های متفاوت به هر زیر باند می‌باشد. با توجه به این موضوع که هر تن (بلندی) با توجه به باند فرکانسی متناظر با آن دارای سطح تاثیر شنوایی متفاوت می‌باشند. بنابراین برای هر زیر باند وزن خاصی با توجه به طیف فرکانسی متناظر با آن نسبت می‌دهند. وزن هر زیر باند با توجه به سطح آستانه بارک و سطح دامنه سیگنال تعیین می‌گردد. رابطه محاسبه معیار MBSD طبق رابطه (۷) است:

$$MBSD = \sum_{k=1}^{C} M(i) D_{xy}(i) \quad (۷)$$

که M(i) وزن‌های اختصاص داده شده به هر زیر باند است. نتایج این دو روش در جدول (۵) آورده شده است.

جدول (۵): نتایج بررسی معیار BSD و MBSD

معیار	BSD / MBSD	تماماً یک	تماماً صفر	یکی در میان یک و صفر
دنباله پیام مخفی / نوع صوت				
موزیک با تن بالا	BSD	۰,۲۵۲۱	۰,۱۰۳۵	۰,۰۹۵۲
	MBSD	۰,۲۰۵۷	۰,۰۵۰۶	۰,۰۲۷۱
موزیک با تن ملایم	BSD	۰,۱۲۲۹	۰,۲۹۰۱	۰,۱۰۲۱
	MBSD	۰,۱۰۹۰	۰,۰۱۷۷	۰,۰۲۴۳
گفتار زن	BSD	۰,۱۳۵۹	۰,۱۱۱۳	۰,۱۱۹۸
	MBSD	۰,۰۹۳۲	۰,۰۹۵۰	۰,۰۰۳۲
گفتار مرد	BSD	۰,۴۶۱۸	۰,۴۳۲۲	۰,۳۷۲۱
	MBSD	۰,۰۲۷	۰,۰۲۳۹	۰,۰۹۹
موزیک تن بالا با گفتار مرد	BSD	۰,۰۲۴	۰,۰۲۸	۰,۰۲۴
	MBSD	۰,۰۱۵	۰,۰۸۷	۰,۰۰۲
موزیک تن پایین با گفتار مرد	BSD	۰,۰۵۱	۰,۰۶۱	۰,۰۱۳
	MBSD	۰,۰۱۱	۰,۰۱۴	۰,۰۱۳
موزیک تن بالا با گفتار زن	BSD	۰,۳۴۶	۰,۱۷۶	۰,۰۹۸
	MBSD	۰,۰۱۶	۰,۰۱۱	۰,۰۱۰
موزیک تن پایین با گفتار زن	BSD	۰,۱۶۳	۰,۲۲۸	۰,۱۱۹
	MBSD	۰۰۸۲.	۰,۰۱۳	۰,۰۶۴
مختلط	BSD	۰,۰۶۷	۰,۱۱۴۳	۰,۰۲۲
	MBSD	۰,۲۹۰	۰,۱۹۲	۰,۰۲۹

آزمون BSD و MBSD مشابه آزمون اندازه‌گیری STFRT ، اندازه‌گیری میزان اغتشاش سیگنال در حوزه فرکانس می‌باشد. با این تفاوت که در معیار BSD یا MBSD طیف فرکانسی از فیلتر سطح آستانه شنوایی عبور داده می‌شود. در صورتی که معیار BSD یا MBSD به مقدار صفر نزدیک باشد می‌توان این نتیجه را گرفت که درج پیام مخفی درون سیگنال حامل علاوه بر این که در حوزه فرکانس اغتشاش وارد نمی‌کند، بلکه خاصیت سطح آستانه شنوایی صوت را نیز بهم نمی‌زند. همانطور که از جدول بالا پیداست الگوریتم پیشنهادشده طیف فرکانسی سیگنال صوتی را چندان تغییر نمی‌دهد و همچنین از خاصیت سطح آستانه

شنوایی پیروی می‌کند. علت این امر نیز واضح می‌باشد زیرا الگوریتم پیشنهادی بر روی سیگنال صوتی اعمال شده است که حاصل خروجی تبدیل MP3 می‌باشد. بنابراین سیگنال صوتی مبرا از داده‌هایی است که درون سطح آستانه شنوایی قرار می‌گیرند. همچنین با توجه به جدول بالا موزیک تن بالا با گفتار مرد دارای بهترین نتایج می‌باشند. زیرا با توجه به ساختار فرکانسی این نوع سیگنال‌ها، تن‌هایی با دامنه کوچک و یا غیر تن کم‌تر در طیف سیگنالی آن‌ها موجود می‌باشد. بنابراین از هر جهت مساعد با ساختار طیف بارک می‌باشند.

- **معیارهای حوزه زمان**

نسبت سیگنال به نویز SNR: معیار SNR معروفترین مقیاس اندازه‌گیری اغتشاش در حوزه ی زمان می باشد که سیگنال اصلی و سیگنال نهان‌نگاری‌شده را نمونه به نمونه با یکدیگر مقایسه می کند. این معیار حاوی اطلاعاتی در مورد اغتشاشات جمع شونده برروی سیگنال ایستان می‌باشد. رابطه محاسبه SNR طبق رابطه (۸) است.

$$SNR = 10\log_{10}\frac{\sum_{i=1}^{N} x^2(i)}{\sum_{i=1}^{N}(x(i)-y(i))^2} \quad (۸)$$

نتایج این روش در جدول (۶) آورده شده است.

جدول (۶): نتایج بررسی نسبت سیگنال به نویز

معیار SNR			
دنباله پیام مخفی / نوع صوت	تماما یک	تماما صفر	یکی در میان یک و صفر
موزیک با تن بالا	۲۲,۲۰۱۰	۲۰,۱۹۰۰	۲۰,۰۰۱۰
موزیک با تن ملایم	۱۸,۷۲۰۵	۱۹,۳۱۴۳	۱۹,۹۶۴۲
گفتار زن	۲۰۳۲۹۱	۲۰,۶۵۰۱	۲۰,۱۷۸۲
گفتار مرد	۱۵,۸۱۲۰	۱۱,۹۳۸۸	۱۳,۰۲۲۲
موزیک تن بالا با گفتار مرد	۱۸,۱۲۱۱	۱۴,۹۵۲۶	۱۷,۳۴۲۱
موزیک تن پایین با گفتار مرد	۱۷,۲۳۲۰	۱۶,۹۲۳۱	۱۵,۳۰۷۵
موزیک تن بالا با گفتار زن	۱۳,۳۸۳۹	۱۴,۳۴۷۲	۱۴,۹۲۷۳
موزیک تن پایین با گفتار زن	۱۹,۶۲۰۵	۲۲,۸۸۲۴	۲۱,۳۲۶۵
مختلط	۱۵,۲۳۷۷	۱۶,۸۸۷۲	۱۴,۹۹۳۸

با توجه به رابطه SNR که در بالا ذکر شده است، مفهوم SNR بدین معنی است در صورتی که سیگنال نهان‌نگاری‌شده شدیداً در حوزه زمان مورد اغتشاش واقع شود فاصله آن تا سیگنال اصلی بسیار زیاد خواهد شد. بنابراین مخرج کسر رابطه SNR دچار رشد فزاینده ای می‌شود. دراین صورت مقدار SNR به سمت صفر و حتی منفی بی نهایت میل می‌کند. بنابراین هرچه در این معیار میزان SNR بیشتر باشد، دلیل بر تطابق بیشتر سیگنال اصلی با سیگنال نهان‌نگاری‌شده دارد. با توجه به نتایج جدول بالا این نتیجه را می‌توان گرفت که الگوریتم پیشنهادی نتایج مطلوبی در حوزه زمان نیز دارد. کم‌ترین SNR مربوط به موزیک تن بالا با گفتار زن می‌باشد که دارای SNRی در حدود 13 db می‌باشد. دلیل این نتیجه می‌تواند به علت اغتشاش شدید زمانی سیگنال مربوطه باشد.

فاصله سزنا-کوشی CZD: این معیار که فرمول آن در رابطه (۹) آورده شده است، یک همبستگی بر پایه استاندارد متریک است که به‌طور مستقیم بردار نمونه‌ها را در حوزه‌ی زمان مقایسه می‌کند. این معیار، میزان شباهت بین دو سیگنال را اندازه گیری می‌کند.

$$C = \frac{1}{K}\sum_{i=0}^{K-1}\left(1-\frac{2\min(x(i),y(i))}{x(i)+y(i)}\right) \quad (۹)$$

نتایج این آزمون در جدول (۷) آورده شده است.

جدول (۷): نتایج بررسی فاصله سزنا-کوشی

معیار CZD			
دنباله پیام مخفی / نوع صوت	تماما یک	تماما صفر	یکی در میان یک و صفر
موزیک با تن بالا	$-2,32*10^{-3}$	$-1,14*10^{-3}$	-۰,۰۰۲۴
موزیک با تن ملایم	۰,۶۱۹۲	-۰,۰۵۲۸	۰,۱۶۳۴
گفتار زن	۰,۷۵۴۲	۰,۰۳۴۳	-۰,۴۶۲۱
گفتار مرد	-۰,۰۰۹۵	۰,۰۱۵۴	۰,۱۰۳۲
موزیک تن بالا با گفتار مرد	$-8,34*10^{-3}$	۰,۰۹۷۳	-۰,۰۱۸۳۰
موزیک تن پایین با گفتار مرد	۰,۱۰۲۲	-۰,۱۲۴۸	۰,۱۹۳۹
موزیک تن بالا با گفتار زن	$9,2731*10^{-4}$	۰,۰۲۱۴	۰,۰۲۷۷
موزیک تن پایین با گفتار زن	-۰,۰۶۵۷	-۰,۰۱۳۶	-۰,۰۹۸۵
مختلط	۰,۰۲۲۳	$-8,377*10^{-4}$	$1,43*10^{-4}$

در این معیار در صورتی که دو سیگنال مورد آزمون یکسان باشند، مقدار این معیار برابر با صفر خواهد شد. این معیار به بررسی تک تک عناصر دو سیگنال مورد آزمون نسبت به یکدیگر می‌پردازد. همانطور که از رابطه بالا پیداست در صورتی که مولفه‌های سیگنال نهان‌نگاری‌شده اختلاف کمی با مولفه‌های سیگنال اصلی داشته باشند، میزان این معیار به سمت صفر میل می‌کند. با توجه به جدول بالا می‌توان دریافت که الگوریتم پیشنهادی در حوزه زمان اغتشاش چندانی را نیز بر سیگنال صوتی متحمل نمی‌کند. از نتایج حاصله در جدول بالا می‌توان دریافت که با انتخاب سیگنال صوتی با گفتار زن بدترین حالت درج پیام مخفی را خواهیم داشت. زیرا در سیگنال گفتار زن همواری خواصی محسوس است. بنابراین با درج پیام مخفی در این نوع سیگنال صوتی، تا حدودی این فرم یکنواختی از بین خواهد رفت و با توجه به مقایسه‌شدن مولفه به مولفه دو سیگنال اصلی و سیگنال حاوی پیام مخفی، میزان معیار CZD افزایش می‌یابد.

۴- نتیجه‌گیری

الگوریتم نهان‌نگاری بایستی مناسب طراحی شود و ضمن هنر حفظ سیگنال پوشانه، در صورت حمله، سیگنال مخفی آن نیز آشکار نشود. در الگوریتم پیشنهادی این موضوعات وجود دارد و به‌طور کلی نتایج و ویژگی‌های زیر در آن به‌دست آمده است:

- پایداری الگوریتم پیشنهادی، در برابر نویز سفید نرمال بیشتر از نویز یکنواخت است ودر برابر SNR های بالاتر از 10db نرخ بیت خطای کم‌تر از ۱ بیت دارد.
- در صورت حمله به الگوریتم پیشنهادی، با توجه به BER به‌دست آمده، آشکارسازی پیام مخفی کاملا تصادفی‌شده و سیگنال مخفی کاملا تخریب می‌شود.
- همچنین روش پیشنهادی در برابر اضافه‌کردن نویز افزایشی به‌صورت بلوک مقاوم به‌نظر می‌رسد و مستعد درج پیام مخفی با استفاده از روش الگوریتم پیشنهادی نخواهند بود.
- نتایج بررسی آماری الگوریتم پیشنهادی در حوزه فرکانس با مقیاس فاصله کپستروم همان‌طور که از متوسط معیار محاسبه‌شده در نتایج به‌دست آمده، فایل‌های صوتی به فرم موزیک با تن (بلندی) ملایم کم‌ترین تغییرات را در معیار همواری صوت دارد.
- در الگوریتم پیشنهادی، افزایش پیام مخفی تاثیر چندانی در ایجاد اغتشاش در حوزه فرکانس ندارد.
- الگوریتم پیشنهادشده طیف فرکانسی سیگنال صوتی را چندان تغییر نمی‌دهد و همچنین از خاصیت سطح آستانه شنوایی پیروی می‌کند.
- همچنین با توجه به نتایج حاصل‌شده، موزیک تن بالا با گفتار مرد دارای بهترین نتایج می‌باشند بنابراین از هر جهت مساعد با ساختار طیف بارک می‌باشند.
- الگوریتم پیشنهادی با توجه به جداول (۶) و (۷)، نتایج مطلوبی در حوزه زمان دارد.
- کم‌ترین SNR مربوط به موزیک تن بالا با گفتار زن می‌باشد که دارای SNRی در حدود 13db می‌باشد. دلیل این نتیجه می‌تواند به علت اغتشاش شدید زمانی سیگنال مربوطه باشد. با توجه به نتایج حاصل شده، با انتخاب سیگنال صوتی با گفتار زن بدترین حالت درج پیام مخفی را خواهیم داشت.

۵- مراجع

[1] C. Kratzer, J. Dittmann, T. Vogel and R. Hillert, "Design and evaluation of steganography for voice-over-IP," 2006 IEEE International Symposium on Circuits and Systems, Kos, Greece, p. 4, 2006. doi: 10.1109/ISCAS.2006.1693105.

[2] L. Boney, A. H. Tewfik, and K. N. Hamdy, "Digital watermarks for audio signals," Proceedings of the Third IEEE International Conference on Multimedia Computing and Systems, Hiroshima, Japan, pp. 473-480,1996. doi: 10.1109/MMCS.1996.535015.

[3] Cvejic, Nedeljko, "Algorithms For Audio Watermarking and Steganography, 2004.

[4] S. S. Agaian, D. Akopian, O. Caglayan, and S. A. D'Souza, "Lossless Adaptive Digital Audio Steganography," Conference Record of the Thirty-Ninth Asilomar Conference onSignals, Systems and Computers, 2005., Pacific Grove, CA, USA, pp. 903-906, 2005. doi: 10.1109/ACSSC.2005.1599886.

[5] W. C. CHU, "Speech Coding Algorithms-Foundation and Evaluation of Standatadized Coder, 2003.

[6] J. Herre and S. Dick, "Psychoacoustic Models for Perceptual Audio Coding—A Tutorial Review," Applied Sciences, vol. 9, no. 14, p. 2854, Jul. 2019, doi: 10.3390/app9142854.

[7] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," in IEEE Transactions on Signal Processing, vol. 41, no. 12, pp. 3445-3462, Dec. 1993, doi: 10.1109/78.258085.

[8] V. Ralph Algazi and Robert R. Estes Jr. "Analysis-based coding of image transform and subband coefficients", Proc. SPIE 2564, Applications of Digital Image Processing XVIII, 22 August 1995.

[9] K. L. Narasihimhaprasad, M. V. Nagabhushanam, V. V. Satyanarayana Tallapragada, and J. Krishna Sunkara, "Embedded Zero-Tree Wavelet Coding with Selective Decomposition Bands," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, pp. 0445-0449, 2019. doi: 10.1109/ICCSP.2019.8698062.

[10] J. Antonio Alvarez-Cedillo, T. Alvarez-Sanchez, M. Aguilar-Fernandez, and J. Sandoval-Gutierrez, 'Many-Core Algorithm of the Embedded Zerotree Wavelet Encoder', Coding Theory. IntechOpen, Mar. 11, 2020. doi: 10.5772/intechopen.89300.

[11] Rekik, Siwar et al. "Speech steganography using wavelet and Fourier transforms." EURASIP Journal on Audio, Speech, and Music Processing, pp. 1-14, 2012.

[12] H. Inoue, A. Miyazaki, A. Yamamoto, and T. Katsura, "A digital watermark based on the wavelet transform and its robustness on image compression," Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269), Chicago, IL, USA, vol. 2, pp. 391-395, 1998. doi: 10.1109/ICIP.1998.723388.

[13] Z. Xu, K. Wang, and X. -h. Qiao, "Digital Audio Watermarking Algorithm Based on Quantizing Coefficients," 2006 International Conference on Intelligent Information Hiding and Multimedia, Pasadena, CA, USA, 2006, pp. 41-46, doi: 10.1109/IIH-MSP.2006.265115.

[14] Xiang, Shijun, Hyoung Joong Kim, and Jiwu Huang. "Audio watermarking robust against time-scale modification and MP3 compression." Signal Processing 88.10, pp. 2372-2387, 2008.

[15] M. Baziyad, et al., "Maximizing embedding capacity for speech steganography: a segment-growing approach," Multimedia Tools and Applications, vol. 80, pp. 24469-24490, 2021.

[16] F. J. Farsana and K. Gopakumar, "Speech encryption algorithm based on nonorthogonal quantum state with hyperchaotic keystreams." Advances in Mathematical Physics 2020, pp. 1-12, 2020.

[17] A. Kuznetsov, et al., "Direct Spread Spectrum Technology for Data Hiding in Audio," Sensors 22.9, 3115, 2022.