

روشی جدید در واترمارکینگ صوت با استفاده از طیف گسترده تعمیم یافته و بالاترین ماسک آستانه شنیداری

فهیمه فتاح پور
دانشگاه سمنان

خشایار یغمایی
دانشگاه سمنان

Fattahpour14@gmail.com

Khyaghmaie@semnan.ac.ir

چکیده: در این مقاله، روشی جدید برای واترمارکینگ صوت با استفاده از کدهای شبه نویز تعمیم یافته در بسط سیگنال واترمارک ارائه می شود. در این روش که بر پایه کدگذاری کانولوشنی و کدگشایی ویتربی طراحی شده است، حضور واترمارک به لحاظ ادراکی قابل تشخیص نیست. به منظور افزایش ظرفیت جاسازی واترمارک، واترمارک در مناطق خاصی از سیگنال میزبان جاسازی می شود که نتیجه آن افزایش آستانه ماسک گذاری می باشد. با استفاده از این آستانه ماسک-گذاری، تزریق توافقی واترمارک، آشکارسازی کور، مقاومت و ظرفیت جاسازی واترمارک نسبت به روش های مشابه افزایش یافته است. همچنین نتایج شبیه سازی در مقایسه با روش های مشابه، بر مقاومت سیگنال واترمارک در برابر حملاتی از قبیل اضافه کردن نویز، فشرده سازی، فیلتر کردن پایین گذر و بازنمونه برداری دلالت دارند.

با تزریق سیگنال واترمارک در سیگنال اصلی، از شفافیت سیگنال اصلی کاسته نشود. SNR محاسبه شده برای سیگنال واترمارک شده نسبت به سیگنال اصلی باید بیشتر از 20 dB و بار مفید برای واترمارک 20bps باشد. واترمارک باید در برابر پردازش های متفاوت سیگنال و حملاتی همچون تبدیلات A/D و D/A، اضافه شدن یا ضرب شدن با نویز و فشرده سازی MP3 از خود مقاومت نشان دهد. همچنین واترمارک باید در برابر آشکارسازی ناخواسته، حذف و جاسازی ناخواسته از خود مقاومت نشان دهد. این نیازمندی ها، چالش های بزرگی را برای یک واترمارکینگ صوت مقاوم بوجود می آورد. الگوریتم های موجود در زمینه واترمارکینگ صوت به دو دسته تقسیم می شوند. الگوریتم های حوزه زمان [4,5] و الگوریتم های حوزه فرکانس [6-8]. مهمترین ضعف یک الگوریتم واترمارکینگ صوت، کم بودن بار مفید واترمارک و پایین بودن مقاومت الگوریتم می باشد [9,10].

واژه های کلیدی: آنتروپی ادراکی، طیف گسترده، کدگذاری کانولوشنی، کدگشایی ویتربی، مدل روان شنیداری.

۱- مقدمه

واترمارکینگ نقش مهمی در امنیت سیستم های چند رسانه ای بازی می کند. مبنای واترمارکینگ بر اساس خصوصیات سیستم شنوایی و بینایی انسان است بطوریکه اطلاعات جاسازی شده (واترمارک) در سیگنال اصلی غیر قابل درک باقی می ماند و این تکنیک در برابر حملات و اعوجاجات از خود مقاومت نشان می دهد. به دلیل حساسیت بیشتر سیستم شنوایی انسان، واترمارکینگ صوت بیشتر از واترمارکینگ تصویر مورد چالش قرار می گیرد [1]. در دهه اخیر، تکنیک های متنوعی از واترمارکینگ صوت پیشنهاد شده اند که شامل LSB، کدینگ فاز، مدولاسیون فاز، کدینگ پژواک، ویولت و طیف گسترده می باشند [2]. در مقایسه با تکنولوژی واترمارکینگ تصویر، واترمارکینگ صوت کمتر مورد توجه قرار گرفته است. لازمه اجرای صحیح الگوریتم واترمارکینگ صوت بر طبق مقررات IFPI برقراری سه شرط زیر است [3]:

یکی از روش های مهم در زمینه واترمارکینگ صوت تکنیک طیف گسترده است. در ارتباطات طیف گسترده، پهنای باند سیگنال در فرستنده افزایش می یابد تا سیگنال در برابر تداخلات از خود مقاومت نشان دهد. واترمارکینگ طیف گسترده به دلیل استفاده از جملات شبه نویزی که به عنوان واترمارک مورد استفاده قرار می گیرند، اخیراً هم در حوزه واترمارکینگ صوت و هم در حوزه واترمارکینگ تصویر مورد استفاده قرار گرفته است. تکنیک طیف گسترده یکی از مقاوم ترین تکنیک های الگوریتم واترمارکینگ صوت می باشد [11]. هدف طیف گسترده بسط اطلاعات جاسازی شده در طیف فرکانسی سیگنال اصلی است. در واترمارکینگ صوت، به منظور غیر قابل درک بودن واترمارک جاسازی شده در سیگنال میزبان اغلب از مدلی که بر اساس خصوصیات سیستم شنوایی انسان کار می کند و مدل روان شنیداری نامیده می شود، استفاده می کنند. درجه موفقیت سیستم واترمارکینگ، به مشخصات مدل روان شنیداری مورد استفاده، بستگی دارد. یکی از ملزومات سیستم واترمارکینگ حفظ کیفیت ادراکی سیگنال میزبان است. این

موضوع به عنوان شفافیت سیستم شناخته می‌شود. آستانه ماسک-گذاری، یا کمترین آستانه شنوایی، نقش مهمی در مدل روان‌شنیداری بازی می‌کند و این موضوع، کلید موفقیت سیستم‌های واترمارکینگ گفتار بر پایه مدل روان‌شنیداری است. در این مقاله، ابتدا اشاره‌ای به تکنیک طیف گسترده با استفاده از کد شبه‌نوین (PN) معمولی، که طیف گسترده معمولی نیز خوانده می‌شود، خواهیم داشت. در ادامه به منظور بهبود عملکرد سیستم بازیابی سیگنال واترمارک، تکنیک طیف گسترده با استفاده از کد شبه‌نوین تعمیم‌یافته و مراحل تولید و بازیابی سیگنال واترمارک در این روش پیشنهاد می‌شوند. نرخ خطای بیت سیگنال واترمارک بازیابی‌شده نیز در هر دو مورد طیف گسترده معمولی و تعمیم‌یافته مورد بررسی قرار می‌گیرد. در ادامه براساس مدل آنالیز ادراکی [12] به منظور افزایش ظرفیت پنهان‌سازی داده واترمارک، ماسک مناسب که در مکان بالاتری قرار دارد، پیشنهاد می‌شود. از جمله مزایای بالاترین ماسک آستانه شنیداری پیشنهادی، نرخ فشرده‌سازی بیشتر در کاربردهای فشرده‌سازی صوت و بار مفید یا ظرفیت پنهان‌سازی اطلاعاتی بیشتر در سیستم واترمارکینگ، انرژی بیشتر برای سیگنال واترمارک و در نتیجه افزایش سرعت آشکارسازی و بازیابی واترمارک و افزایش مقاومت واترمارک در برابر فشرده‌سازی سیگنال یا حملات حذفی می‌باشد. به‌کارگیری ماسک آستانه شنیداری بالا، احتمال اعوجاج قابل شنیدن را افزایش می‌دهد. برای مقابله با این اثر، در الگوریتم پیشنهادی، واترمارک قبل از تزریق به سیگنال اصلی به نحو مناسب میزان می‌شود. در روش ارائه شده، با محاسبه ماسک شنیداری مناسب، شفافیت ادراکی سیستم مطابق با استاندارد ISO/MPEG در نظر گرفته شده است [12]. سیستم ارائه شده بلایند است؛ بدین معنی که گیرنده در فرآیند بازیابی به سیگنال صوتی اصلی نیازی ندارد.

۲- تکنیک طیف گسترده

قاعده کلی برای تکنیک طیف گسترده این است که پهنای باندی بیشتر از حد معمول اشغال شود [13]. بخاطر افزایش پهنای باند، چگالی طیف توان سیگنال در کانال، دقیقاً شبیه به سیگنال نوین، کم خواهد شد. بسط سیگنال با استفاده از ترکیب آن با یک کد انجام می‌شود که وابسته به نوع داده فرستاده شده است. از جمله مزایای تکنیک طیف گسترده می‌توان به موارد زیر اشاره کرد [13]:

هنگامی که سیگنال حول یک باند فرکانسی بزرگ بسط داده شود، چگالی طیف توان سیگنال به طور چشمگیری کاهش خواهد یافت.

امکان دستیابی تصادفی افزایش داده می‌شود، یعنی هنگامی که تعداد کدهای بیشتری تولید شود، به تعداد کاربر بیشتری اجازه استفاده از باند فرکانسی داده خواهد شد.

هر چه تعداد کاربر بیشتر باشد، تداخلات محدودتر می‌شوند.

امنیت: بدون دانستن کد بسط، بازیابی داده فرستاده‌شده غیر ممکن است.

هر چه پهنای باند بیشتری برای سیستم مورد استفاده قرار گیرد اعوجاج نیز کمتر خواهد بود.

در فرستنده به منظور بسط سیگنال واترمارک از یک دنباله شبه نوین استفاده می‌شود. در گیرنده نیز برای رفع بسط از همان دنباله شبه نوینی که در فرستنده استفاده شده، استفاده می‌شود.

دنباله‌های (کدهای) شبه نوین نقش کلیدی در سیستم‌های طیف گسترده جمله مستقیم دارند؛ زیرا برای بسط و رفع بسط سیگنال باند پایه مورد استفاده قرار می‌گیرند. به طور کلی دو نوع جمله شبه‌نوین وجود دارد: جمله شبه نوین معمولی و تعمیم‌یافته [14].

اجزای دنباله شبه نوین معمولی $\{+1, -1\}$ و اجزای دنباله شبه‌نوین تعمیم‌یافته $\{+1, 0, -1\}$ هستند. در تکنیک طیف گسترده با ضرب سیگنال واترمارک در جمله شبه‌نوین (در حوزه زمان)، سیگنال واترمارک، بسط داده می‌شود. در [15] بسط سیگنال واترمارک با استفاده از جمله شبه نوین معمولی انجام داده شده است. در این مقاله به منظور کاهش نرخ خطای بیت واترمارک بازیابی‌شده، از جمله شبه نوین تعمیم‌یافته استفاده می‌کنیم.

۳- تولید واترمارک در تکنیک طیف گسترده تعمیم‌یافته

به منظور بسط سیگنال واترمارک با استفاده از کد شبه‌نوین معمولی، از کدگذارهای ساده مثل کد تکرار استفاده می‌شود [15]؛ اما در بسط سیگنال واترمارک با استفاده از کد شبه‌نوین تعمیم‌یافته، از کدگذار کانولوشنی و کدگشای متناظر با آن، یعنی کدگشای ویتربی استفاده می‌کنیم [16]. طول کدگذار کانولوشنی را برابر با K در نظر می‌گیریم. کدگذار کانولوشنی بیت ها را به L سمبل نگاشت می‌دهد. کلمه کد مربوط به سیگنال واترمارک w برای طیف گسترده جمله مستقیم تعمیم‌یافته که به صورت کانولوشنی کدگذاری شده است، به صورت معادله (1) است:

$$w(t) = \sum_{k=0}^{L-1} w_k(t - kT_s) \quad (1)$$

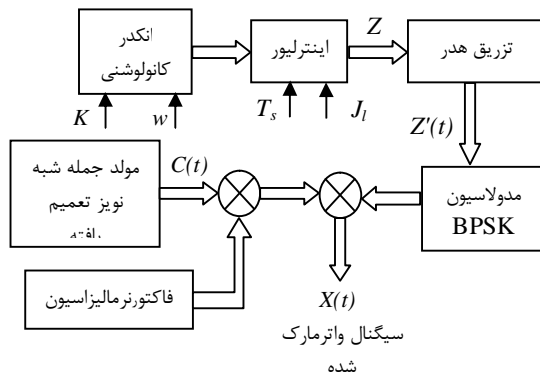
سمبلها به صورت کانولوشنی اینترلیو می‌شوند تا به شکل کد اینترلیو شده به صورت معادله (2) در آیند:

$$x(t) = \sum_{l=0}^{L-1} w_l(t - lT_s) \quad (2)$$

$$L = \frac{5 \times K}{r} \quad 0 \leq t < T$$

که در این معادله $u(t)$ فقط در بازه $(0, T_s]$ برابر با 1 است و $L/T = T_s$

شنوایی انسان، محل جاسازی واترمارک را تعیین خواهد کرد. در حقیقت واترمارک در فضاهای توخالی از سیگنال میزبان، که با استفاده از مدل روان شنیداری تعیین شده، جاسازی می شود. با حذف اجزای فرکانسی پایین تر از ماسک آستانه و تزریق سیگنال واترمارک بسط یافته در این ناحیه، عمل جاسازی سیگنال واترمارک در سیگنال اصلی انجام خواهد شد.



شکل (1): تولید واترمارک در تکنیک طیف گسترده تعمیم یافته

در برخی از الگوریتم های طیف گسترده، به منظور رسیدن به مقاومت بالاتری از سیگنال واترمارک، آن را در کل سیگنال میزبان، گسترده می سازند. اگر چه این روش از مقاومت بالایی برخوردار است اما محاسبات پیچیده و زمانبری داشته و برای کاربردهایی که نیاز به کدگذاری و کدگشایی سریع دارند مناسب نمی باشد. بدین منظور با کم کردن بعضی از سطوح مقاومت، واترمارک را در فضای خاصی از سیگنال میزبان جاسازی می کنیم. در روش پیشنهادی، مناطقی با طولی بیشتر از طول هر فریم واترمارک و انرژی ای بیشتر از متوسط انرژی واترمارک، در نظر می گیریم و واترمارک را فقط در آن مناطق تزریق می کنیم. همان گونه که بیان شد، الگوریتم هایی که از تمامی سیگنال میزبان به منظور جاسازی داده واترمارک استفاده می کنند، دلیل وجود پردازش هایی همچون کدگذاری و کدگشایی که شامل چندین تبدیل فوریه و عکس تبدیل فوریه است، پردازش هایی زمانبر و از لحاظ محاسباتی سنگین هستند. با جاسازی واترمارک در مناطق محدودی از سیگنال میزبان، بجای جاسازی در کل سیگنال میزبان، عملیات جاسازی ساده تر شده و سرعت آشکارسازی نیز افزایش می یابد. در این صورت مشاهده می شود که آستانه ماسک گذاری نسبت به آنچه که قبلاً محاسبه شده است در مکان بالاتری قرار می گیرد که دلیل آن، جاسازی واترمارک در مناطق انرژی بالای سیگنال اصلی است. با بالا رفتن مکان ماسک آستانه، فضای خالی بیشتری در زیر این آستانه بوجود خواهد آمد؛ در این صورت واترمارک بیشتری از لحاظ ظرفیت در این فضای بوجود آمده در سیگنال اصلی، تزریق خواهد شد. در سیستم واترمارکینگ، بین ماسک آستانه ای که در مکان بالایی قرار دارد و قابلیت شنیده شدن

مدت زمان سمبل است. نرخ کد کدگذار کانولوشنی است و برابر با نسبت تعداد سمبل های ورودی (m) به تعداد سمبل های خروجی (n) می باشد [8]. در معادله (2) جمله $\frac{Z'(t)}{Z}$ اشاره به ضریب اینترلیوینگ سمبل ها دارد. در مرحله بعد سیگنال $Z'(t)$ با هدف تعیین مکان دقیق سیگنال واترمارک در سیگنال واترمارک شده در گیرنده، به کلمه کد اینترلیوشده اضافه می شود. سیگنال حاصل توسط مدولاسیون BPSK مدوله می شود و با استفاده از جمله شبه نویز $C(t)$ به صورت معادله (3) بسط داده می شود:

$$X(t) = \frac{Z'(t)}{Z} C(t) \quad (3)$$

که در این معادله E انرژی هر کلمه کد است. T طول هر چیپ است. N ، تعداد اجزای دنباله شبه نویز و فرکانس حامل به صورت ω_c است. جمله شبه نویز $C(t)$ به صورت معادله (4) تعریف می شود:

$$C(t) = \sum_{i=1}^N a_i \cos(\omega_c t + \theta_i) \quad (4)$$

که در این معادله $v(t)$ فقط در بازه $[0, T_c]$ برابر با 1 است. ضریب جمله شبه نویز (a_i) برای DSSS تعمیم یافته به صورت معادله (5):

$$a_i = \frac{1}{\sqrt{N}} \quad (5)$$

و برای DSSS معمولی به صورت معادله (6) است:

$$a_i = \frac{1}{\sqrt{L}} \quad (6)$$

که در این دو معادله P_r اشاره به احتمال رخداد بیت در جمله PN دارد و مستقل از $\frac{Z'(t)}{Z}$ است. به منظور جبران انرژی از دست رفته ناشی از وجود صفر در جمله شبه نویز تعمیم یافته، از فاکتور نرمالیزاسیون $\frac{1}{\sqrt{L}}$ استفاده می کنیم [16]؛ که u اشاره به کل تعداد چیپ های غیر صفر به ازای L سمبل کد شده سیگنال واترمارک دارد و N ، تعداد اجزای جمله شبه نویز است. خروجی این قسمت سیگنال صوتی واترمارک خواهد بود. در شکل (1) دیاگرام مربوط به تولید سیگنال واترمارک در تکنیک طیف گسترده تعمیم یافته نشان داده شده است.

۴ - جاسازی واترمارک در تکنیک طیف گسترده تعمیم یافته

یافته با استفاده از بالاترین ماسک آستانه شنیداری

پس از بسط سیگنال واترمارک، به منظور جاسازی واترمارک در سیگنال صوتی اصلی از یک مدل روان شنیداری مطابق با استاندارد ISO/MPEG استفاده می شود [15]. مدل روان شنیداری آستانه شنیداری سیگنال صوتی اصلی را مشخص می کند. کمترین آستانه

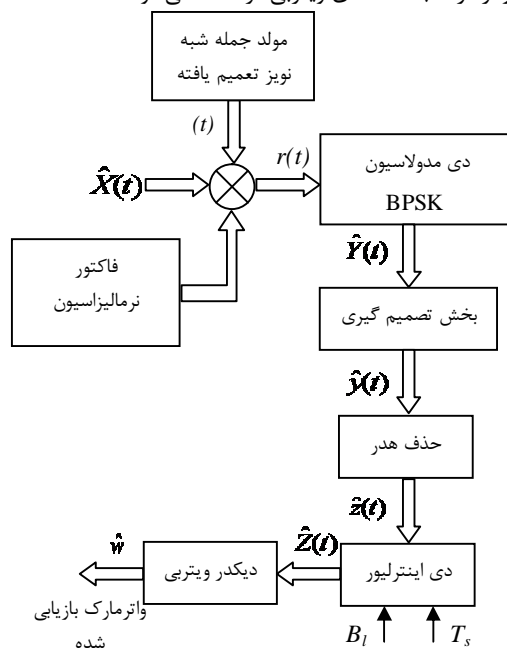
سیگنال ورودی $\hat{Y}(t)$ با آستانه صفر مقایسه می‌شود. همان‌طور که در معادله (8) نشان داده شده، اگر سیگنال ورودی از آستانه صفر بزرگتر باشد خروجی +1 و اگر سیگنال ورودی از آستانه صفر کوچکتر باشد، خروجی -1 خواهد بود:

$$\hat{Y}(t) = \begin{cases} 1 & \text{if } Y(t) > 0 \\ -1 & \text{if } Y(t) < 0 \end{cases} \quad (8)$$

پس از حذف هدر سمبل‌ها به صورت معادله (9) تحت دی‌اینترلیونگ قرار می‌گیرند:

$$\hat{z}(t) = \hat{Y}(t) \cdot T_s \quad (9)$$

در این معادله $u(t)$ فقط در فاصله $[0, T_s]$ برابر با 1 است و $T/L = T_s$ مدت زمان سمبل است و جمله $\hat{z}(t)$ اشاره به ضریب دی‌اینترلیونگ سمبل‌ها دارد. سمبل‌ها به منظور بازیابی بیت‌های اصلی پیام واترمارک به کدگشای ویتربی فرستاده می‌شوند



شکل 3: سیستم بازیابی سیگنال واترمارک در تکنیک طیف گسترده تعمیم‌یافته

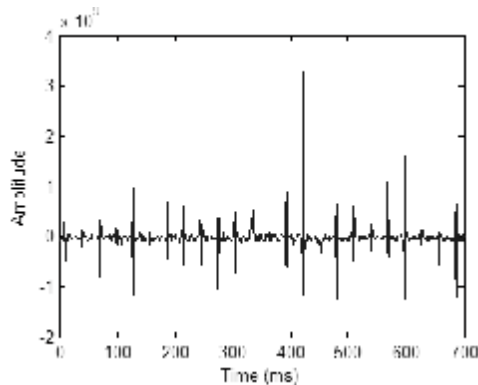
6 - کدگشای ویتربی

کدگشای ویتربی یکی از انواع الگوریتم‌های کدگشایی است که به منظور کدگشایی کدهای کانولوشنی مورد استفاده قرار می‌گیرد. با استفاده از کدگشای ویتربی بار محاسباتی که در ساختار شبکه کدگذار وجود دارد کاهش می‌یابد [8]. در این مقاله طول کدگشا را $5 \times K$ در نظر می‌گیریم، زیرا بر طبق آزمایشات انجام شده، با این مقدار از طول

سیگنال واترمارک، مصالح‌های وجود دارد. داشتن بالاترین آستانه ماسک‌گذاری تا حد ممکن و جلوگیری از ایجاد اعوجاج قابل شنیدن، از موضوعات کلیدی به هنگام استفاده از مدل روان‌شنیداری است. بدین‌منظور در مدل پیشنهادی از میزان‌کردن واترمارک‌ها قبل از تزریق در سیگنال صوتی اصلی استفاده می‌کنیم. در این سیستم، به منظور میزان‌کردن تزریق سیگنال صوتی واترمارک، فریم‌های واترمارک 75% همپوشان هستند. با میزان‌کردن مناسب واترمارک، انرژی سیگنال میزبان و سیگنال واترمارک‌شده تقریباً با هم برابر خواهد بود.

5 - بازیابی واترمارک در تکنیک طیف گسترده تعمیم‌یافته

در گیرنده، ابتدا از یک فیلتر وقتی با دقت بالا به منظور آشکارسازی سیگنال هدر در فریم $r(t)$ استفاده می‌کنیم؛ (سیگنال $r(t)$ سیگنال رفع بسط شده در پروسه بازیابی واترمارک می‌باشد) تا مکان حضور سیگنال $header(t)$ در سیگنال صوتی واترمارک‌شده آشکار شود [15]. یک نمونه از خروجی فیلتر در شکل (2) نشان داده شده است. پیک، مکان نمونه‌ای که $header(t)$ در فریم $r(t)$ از آنجا شروع می‌شود را نشان می‌دهد. این ردیابی برای همه فریم‌های درون سیگنال واترمارک‌شده انجام می‌شود و همه مکان‌هایی که پیک در آن‌ها رخ می‌دهد برای آنالیز بعدی ذخیره می‌شوند. برای هر مکان پیک که در سیگنال واترمارک‌شده پیدا می‌شود، یک فریم انتخاب شده $\hat{X}(t)$ با طولی مشابه سیگنال واترمارک‌شده، پردازش می‌شود.



شکل (2): آشکارسازی پیک در فیلتر وقتی

این پردازش در شکل (3) نشان داده شده است. سیگنال ورودی در همان جمله شبه‌نویزی که در فرستنده استفاده شده است، ضرب می‌شود و همزمان تحت دی‌مدولاسیون BPSK، به صورت معادله (7) قرار می‌گیرد.

$$\hat{z}(t) = \hat{Y}(t) \cdot T_s \quad (7)$$

بعد از عملیات ضرب و دی‌مدولاسیون، چپ‌ها به سمبل تبدیل می‌شوند و به بخش تصمیم‌گیری فرستاده می‌شوند. در بخش تصمیم‌گیری،

کدگشا، خطای بوجود آمده توسط طول کدگشا قابل چشم‌پوشی خواهد بود [8]. مزیت کدگشای ویتربی در این است که زمان کدگشایی در آن ثابت است. همچنین پیچیدگی کدگشا تابعی از تعداد سنبل‌ها در کلمه کد نمی‌باشد. معیار مورد استفاده در کدگشایی ویتربی، معیار فاصله همینگ است [8].

۷- نتایج شبیه‌سازی

در این سیستم، سه نمونه سیگنال صوتی (موسیقی Light (A)، موسیقی پاپ (B) و سیگنال صحبت (C)) به عنوان سیگنال میزبان، به منظور مقایسه بین دو حالت تکنیک طیف گسترده معمولی و تکنیک طیف گسترده تعمیم‌یافته پیشنهادی، انتخاب شده‌اند. طول هر سه نمونه سیگنال، 90 ثانیه است و با نرخ نمونه برداری 1/44 کیلوهرتز و کوانتیزاسیون 16 PCM بیتی برای پردازش انتخاب شده‌اند. واترمارک مورد استفاده در این سیستم، یک متن 40 کاراکتری است. طول کدگذار کانولوشنی را $K=7$ در نظر می‌گیریم (طول یک کدگذار کانولوشنی مناسب K می‌باشد [16]). مقاومت روش پیشنهادی بر روی سیگنال‌های واترمارک‌شده مورد آزمایش قرار گرفت. در این آزمایش، سیگنال واترمارک‌شده مورد حملاتی از قبیل اضافه کردن نویز، فشرده‌سازی، فیلتر کردن پایین‌گذر، باز نمونه‌برداری و حملات برشی قرار می‌گیرد. عملکرد صحیح سیستم بازبازی سیگنال واترمارک، به نرخ بازبازی واترمارک (WDR) و نرخ دقت بیت (BAR) بستگی دارد. نرخ بازبازی واترمارک را به صورت نسبت تعداد فریم‌های صحیح بازبازی شده واترمارک به تعداد فریم‌های جاسازی شده واترمارک تعریف می‌کنیم [17]. نرخ دقت بیت را نیز به صورت نسبت تعداد بیت‌های صحیح بازبازی شده در هر فریم به تعداد بیت‌های جاسازی شده تعریف می‌کنیم [17]. عملکرد کلی سیستم به صورت معادله (10) تعریف می‌شود:

(10)

در این معادله، DPM میزان عملکرد بازبازی صحیح سیگنال واترمارک در روش پیشنهادی است [17]. نتیجه بازبازی صحیح سیگنال واترمارک به مقدار DPM در شرایطی که سیگنال واترمارک‌شده در معرض حملات زیر قرار می‌گیرد، بستگی دارد:

بازنمونه‌برداری: سرعت نمونه برداری سیگنال صوتی واترمارک‌شده تا فرکانس 05/22 kHz کاهش یافت و پس از آن، این سرعت تا فرکانس 1/44kHz افزایش یافت.

فشرده‌سازی: سیگنال واترمارک‌شده دستخوش کدگذاری و کدگشایی ISO/MPEG¹ LayerIII با نرخ بیت 128 Kbs قرار گرفت.

فیلتر پایین‌گذر: سیگنال واترمارک‌شده تحت فیلتر پایین‌گذر با فرکانس

قطعی برابر با 4 KHz قرار گرفت.

اضافه شدن نویز: سیگنال واترمارک‌شده، در معرض حمله نویز سفید گوسی قرار گرفت.

حمله برشی: به منظور ایجاد حملات غیرهمزمان‌شده، از هر 100 نمونه سیگنال صوتی واترمارک‌شده، 1 نمونه به صورت تصادفی حذف شد.

بعد از جاسازی واترمارک و رسم ماسک آستانه شنیداری بر حسب باندهای بحرانی، بر طبق شکل (4)، مشاهده می‌شود که نمودار ممتد بالاترین ماسک آستانه پیشنهادی و مربوط به حالتی است که واترمارک در مناطق انرژی بالای سیگنال میزبان جاسازی شده است و نمودار نقطه‌چین، کمترین ماسک آستانه و مربوط به زمانی است که واترمارک در کل سیگنال میزبان جاسازی می‌شود. منظور از مناطق انرژی بالای سیگنال میزبان، مناطقی است که انرژی آن مناطق از متوسط انرژی مناطق انرژی بالای سیگنال صوتی واترمارک (که به طور عمده در فرکانس‌های پایین قرار دارد) بیشتر باشد.

جدول (1): DPM برای سه نمونه سیگنال صوتی واترمارک‌شده

سیگنال صوتی	A		B		C	
	معمولاً	پیشنهادی	معمولاً	پیشنهادی	معمولاً	پیشنهادی
نوع کدگذاری						
بازنمونه برداری	۰.۸۴	۱	۰.۷	۱	۰.۸۷	۱
فشرده سازی	۰.۶	۱	۰.۴۳	۱	۰.۴	۱
فیلترپایین	۰.۴۲	۱	۰.۴۴	۱	۰.۴۶	۰.۹۹۵
نویزسفید گوسی	۰.۴۱	۱	۰.۵	۱	۰.۲	۰.۹۹۸
حمله برشی	۰.۶	۱	۰.۸۹	۱	۰.۶۵	۰.۹۹۹

همان‌گونه که مشاهده می‌شود خط ممتد به اندازه 35 dB بالاتر از نمودار مربوط به کمترین ماسک آستانه قرار می‌گیرد. نتیجه این افزایش، افزایش بار مفید یا افزایش مقاومت واترمارک در برابر آشکارسازی و حملات مخرب خواهد بود. به بیان دیگر، فضای بیشتری برای پنهان‌سازی داده بوجود می‌آید و ظرفیت پنهان‌سازی به طور ملموسی بهبود خواهد یافت.

در سیستم واترمارک‌کینگ، انرژی واترمارک، بیشتر در فرکانس‌های پایین قرار می‌گیرد. فشرده‌سازی استاندارد صوت MPEG، معمولاً اجزای فرکانس بالای سیگنال را از بین می‌برد. در نتیجه در روش بالاترین ماسک آستانه شنیداری، با انتخاب سیگنال در فرکانس‌های پایین، واترمارک تزریق‌شده، بهتر می‌تواند فشرده‌سازی صوت MPEG را تحمل کند.

EH: مدت زمان صرف شده (ثانیه) برای کدگذاری (E) در صورتیکه واترمارک در مناطق انرژی بالای سیگنال میزبان جاسازی شود.

DH: مدت زمان صرف شده (ثانیه) برای کدگشایی (D) در صورتیکه واترمارک در مناطق انرژی بالای سیگنال میزبان جاسازی شود.

EW: مدت زمان صرف شده (ثانیه) برای کدگذاری (E) در صورتیکه واترمارک در کل سیگنال میزبان جاسازی شود.

DW: مدت زمان صرف شده (ثانیه) برای کدگشایی (D) در صورتیکه واترمارک در کل سیگنال میزبان جاسازی شود.

ES: مدت زمان ذخیره شده (%) برای کدگذاری (E) سیگنال،
 $(EW/(EH-EW)=ES)$.

DS: مدت زمان ذخیره شده (%) برای کدگشایی (D) سیگنال،
 $(DW/(DH-DW)=DS)$.

می‌باشند.

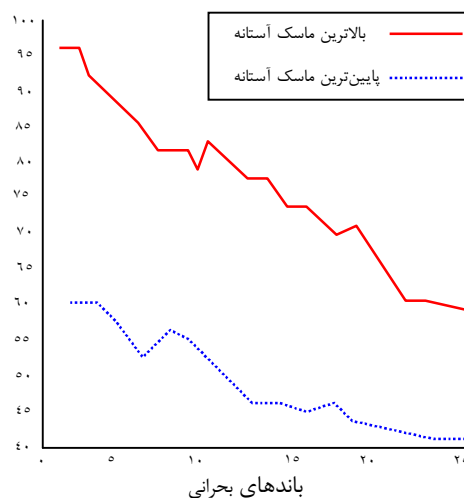
جدول (3): مقایسه سرعت اجرای الگوریتم

	EH	DH	EW	DW	ES	DS
A	5/10	859/1	931/7	85/10	99/85	69/88
B	570/1	766/2	054/7	108	90/76	77/80
C	391/5	930/5	718/6	24/13	16/92	71/95

مشاهده می‌شود که از نظر زمان مورد نیاز برای محاسبه، در حالتی که واترمارک در مناطق انرژی بالای سیگنال میزبان جاسازی شود، نسبت به حالتی که واترمارک در کل سیگنال میزبان جاسازی شود، به طور متوسط 85% بهبودی حاصل شده است. در این روش واترمارک جاسازی شده که حداکثر ظرفیت آن 40 کاراکتر بود، به راحتی در برابر فشرده‌سازی صوت MPEG از خود مقاومت نشان داد. سیگنال واترمارک شده نیز مورد آزمایش شنیداری قرار گرفت و از لحاظ ادراکی هیچ تفاوتی بین سیگنال واترمارک شده و سیگنال اصلی احساس نشد. در ادامه به منظور مقایسه تکنیک طیف گسترده معمولی و طیف گسترده تعمیم یافته در فرآیند بازیابی سیگنال واترمارک جاسازی شده در سیگنال اصلی، سه نمونه سیگنال صوتی را مورد بررسی قرار می‌دهیم. نتایج ارائه شده در جدول 4. BER محاسبه شده (%) برای واترمارک بازیابی شده از سیگنال‌های مورد حمله قرار گرفته را نشان می‌دهد.

با توجه به جدول ذیل مشاهده می‌شود که تکنیک طیف گسترده تعمیم یافته در فرآیند بازیابی سیگنال صوتی واترمارک، بهتر از تکنیک طیف گسترده معمولی عمل می‌کند.

مکان ماسک آستانه



شکل (4): ماسک‌های فرکانسی متفاوت

در این آزمایش برای هر فریم، 4196 نمونه در نظر گرفته شد؛ که بعد از انتقال به حوزه فرکانس، بخش‌های فرکانسی که طیف توان آنها پایین تر از ماسک آستانه بود انتخاب شده تا واترمارک بجای آنها تزریق شود. هر چه تعداد نمونه بیشتری پایین تر از ماسک آستانه باشد، ظرفیت پنهان سازی داده نیز بیشتر خواهد بود. نتیجه مقایسه ظرفیت بر حسب بیت برای هر سه مدل سیگنال صوتی، در جدول 2 آورده شده است. همانطور که در این جدول نیز مشخص شده است در هر دو حالت ماسک پایین و ماسک بالا، سیگنال گفتار بیشترین میزان جاسازی در ظرفیت واترمارک را داراست. با مقایسه بین کمترین و بالاترین مدل ماسک آستانه، مشاهده می‌شود که در روش بالاترین مدل ماسک آستانه، به طور متوسط در ظرفیت جاسازی واترمارک 58% بهبود ایجاد شده است.

جدول (2): مقایسه ظرفیت بین حالت‌های مختلف ماسک آستانه

نوع صوت	کمترین ماسک	بالاترین ماسک	بهبود در ظرفیت
A	960	1520	58%
B	1397	2212	3/58%
C	1476	2338	4/58%

در بعضی کاربردها سرعت پردازش سیستم نسبت به مقاومت آن از اولویت برخوردار است. با توجه به آنچه که در بخش 4 توضیح داده شد، برای رسیدن به سرعت بالا، واترمارک باید در مناطق انرژی بالای سیگنال به جای کل سیگنال میزبان جاسازی شود. جدول 3 نتایج مقایسه شده بین دو حالت (واترمارک در مناطق انرژی بالای سیگنال میزبان و واترمارک در کل سیگنال میزبان) را نشان می‌دهد. در جدول

جدول (4): BER بر حسب درصد برای واترمارک بازیابی شده از سه نمونه سیگنال صوتی

نوع صوت	A		B		C	
	معمولی	تعمیم یافته	معمولی	تعمیم یافته	معمولی	تعمیم یافته
بازنمونه برداری	4/1	0	9/2	0	3/1	4/0
فشرده سازی	9	2/0	3/25	4/0	3/28	6/0
فیلتر پایین‌گذر	23	1	23	2/1	5/20	5/1
نویز سفید گوسی	4/27	2/0	3/12	2/0	55	5/7
حمله برشی	9	0	2	0	3/5	2/1

۸ - نتیجه‌گیری

در این مقاله، تکنیک طیف گسترده با استفاده از کد شبه‌نویز تعمیم‌یافته به منظور بهبود عملکرد سیستم بازیابی سیگنال واترمارک، بر پایه کدگذاری کانولوشنی و کدگشایی ویتربی در سیستم واترمارکینگ پیشنهاد شد. به منظور جاسازی سیگنال واترمارک در سیگنال اصلی از یک مدل روان‌شنیداری مطابق با استاندارد ISO/MPEG استفاده شد. به منظور سنجش مقاومت سیگنال واترمارک جاسازی‌شده در سیگنال اصلی، سیگنال واترمارک‌شده در معرض حملاتی از قبیل اضافه کردن نویز، فشرده‌سازی، فیلتر کردن پایین‌گذر، بازنمونه‌برداری و حملات برشی قرار گرفت. نتایج تجربی نشان‌دهنده افزایش مقاومت سیگنال واترمارک نسبت به روش پیشین (تکنیک طیف گسترده معمولی) و کاهش نرخ خطای بیت در فرآیند بازیابی سیگنال واترمارک هستند. در ادامه، بالاترین ماسک آستانه شنیداری به منظور دارا بودن بیشترین میزان ظرفیت داده جاسازی‌شده در سیگنال اصلی، بر پایه مدل آنتروپی ادراکی و تکنیک طیف گسترده، در سیستم واترمارکینگ مورد بررسی قرار گرفته و روشی برای بهره‌گیری مناسب از ویژگی‌های این ماسک ارائه شد. به منظور افزایش در ظرفیت پنهان‌سازی داده واترمارک، تزریق واترمارک در محل‌های خاصی از سیگنال صوتی میزبان انجام شد. همچنین به منظور حفظ شفافیت سیگنال اصلی، سیگنال واترمارک قبل از تزریق در سیگنال اصلی، میزان شدت تجربی نشان‌دهنده شفافیت و حفظ کیفیت ادراکی سیگنال واترمارک‌شده هستند. علاوه بر آن بدلیل جاسازی واترمارک در مناطق انرژی بالای سیگنال میزبان، الگوریتم پیشنهادی از مقاومت

خوبی در برابر الگوریتم‌های فشرده‌سازی همچون فشرده‌سازی صوت MPEG برخوردار است.

مراجع

- [۱] I. J. Cox, M. L. Miller, and J. A. Bloom, in *Digital Watermarking*, Academic Press, San Diego, Calif, USA, 2002.
- [۲] W. Bender, D. Gruhl, N. Morimoto, and A. Liu, "Techniques for data hiding," *IBM Systems Journal*, Vol. 35, no. 3-4, pp. 313-336, 1996.
- [۳] S. Katzenbeisser and F. A. P. Petitcolas, Eds., *Information Hiding Techniques for Steganography and Digital Watermarking*: Artech House, Inc., 2000.
- [۴] M. A. Gerzon and P. G. Graven, "A high-rate buried-data channel for audio CD," *Journal of the Audio Engineering Society*, Vol. 43, no. 1/2, pp. 3-22, 1995.
- [۵] D. Gruhl et al., "Echo hiding," in *Proc. 1st Information Hiding Workshop*, LNCS, Vol. 1174, Berlin, Germany, pp. 295-315, 1996.
- [۶] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3/4, pp. 131-336, 1996.
- [۷] M. Cooperman and S. Moskowitz, "Steganographic Method and Device," *U.S. Patent* 5 613 004, Mar 1997.
- [۸] J. Huang, Y. Wang, and Y. Q. Shi, "A blind audio watermarking algorithm with self-synchronization," in *Proc. IEEE Int. Symp. Circuits and Systems*, Vol. 3, pp. 627-630, 2002.
- [۹] P. Bassia et al., "Robust audio watermarking in the time domain," *IEEE Trans. Multimedia*, Vol. 3, no. 2, pp. 232-241, 2001.
- [۱۰] B.-S. Ko et al., "Time-spread echo method for digital audio watermarking using PN sequences," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Vol. 2, pp. 2001-2004, 2002.
- [۱۱] W. Cheng, F. Han, M. Tung, K. Xu, "Robust audio steganography using direct-sequence spread spectrum technology", 2007.
- [۱۲] J. Johnson, "Estimation of perceptual entropy using noise masking criteria," in *Proc. ICASSP-88*, pp. 2524-2527, May 1988.
- [۱۳] Kirovski, D. Malvar, H. S., "Spread spectrum watermarking of audio signals," *Signal processing, IEEE Transaction on [see also Acoustics, Speech, and Signal processing, IEEE Transaction on]*, Vol. 51 pp. 1020-1033, no. 4, April 2003.
- [۱۴] H. Ramaswamy, "Generation of generalized signature sequences," *MSc thesis*, Cleveland state university, 2004.
- [۱۵] A. R. Garcia, "Digital watermarking of audio signals using a psychoacoustic auditory model and spread

spectrum theory, "*MSc thesis, University of Miami*, April 1999.

[16] M .Vellala, "Coded generalized of direct-sequence spread-spectrum with specific codes, "*MSc thesis, Cleveland state university*, 2004.

[17] H .Malik, A .Khokhar, R .Ansari, "Robust audio watermarking frequency selective spread spectrum theory, "*Department of Electrical and Engineering University of Illinois at Chicago*, 2004.