مجموعه مقالات پانزدهمین کنفرانس بین‌المللی انجمن رمز ایران
دانشکده مهندسی برق، دانشگاه تربیت دبیر شهید رجایی، تهران، ۶-۷ شهریور ۱۳۹۷

بهبود امنیت پنهان‌نگاری سیگنال گفتار براساس تصادفی‌سازی در حوزه موجک

مهدی صالحی فر[۱]
[۱]کارشناسی ارشد مهندسی برق مخابرات، دانشگاه صنعتی امیرکبیر، تهران
salehifar@aut.ac.ir

چکیده

این مقاله پیرامون پنهان‌نگاری سیگنال گفتار به عنوان داده محرمانه در سیگنال پوششی، با توجه به سه اصل امنیت، دقت و ظرفیت و با تمرکز بیشتر روی دو مقوله اول صورت پذیرفته است. برای تحقق این هدف، از پنهان‌نگاری در ضرایب موجک از مرتبه ی تصادفی به ازای هر فریم، ا ستفاده شده ا ست. روش مذکور، به همراه دو تکنیک کاهش همب ستگی بین داده‌های محرمانه و جانمایی داده‌های محرمانه، برا ساس اولویت آن‌ها (اولویت بندی داده‌های محرمانه و نیز داده‌های سیگنال پو ش شی)، سبب امنیت و دقت بالا، در کنار ظرفیت قابل قبول (بطور متوسط بیست و هشت درصد طول زمانی سیگنال گفتار پوششی) خواهد شد.

کلمات کلیدی

پنهان‌نگاری گفتار، پنهان‌سازی داده، تبدیل موجک گسسته، پنهان‌نگاری تصادفی

۱- مقدمه

تاکنون تکنیک‌های زیادی برای مخفی‌سازی سیگنال پنهان[۱] (سیگنالی که قصد پنهان‌سازی آن را داریم) در سیگنال پوششی ایجاد شده‌اند. تکنیک‌هایی مانند LSB[۲] حوزه‌ی زمان که ظرفیت پنهان‌سازی بالایی دارند ولی از امنیت پائینی برخوردار هستند [12]و یا تکنیک‌های مبتنی بر کدینگ فاز که امنیت خوبی دارند ولی ظرفیت پنهان‌نگاری‌شان بسیار محدود می‌باشد [2]و نیز تکنیک‌های طیف گسترده به سبب نویزی که در سیگنال پوششی ایجاد می‌کنند، سبب شک در وجود سیگنال پنهان می‌شود [8]و برخی تکنیک‌های دیگر که هر کدام معایب و محاسن خود را دارند [14].

همواره، سه مقوله‌ی امنیت، ظرفیت و دقت از مهمترین موارد مطرح در حوزه‌ی پنهان‌نگاری هستند. تکنیک‌های مبتنی بر پنهان‌نگاری داده در حوزه‌ی تبدیل، به طور نسبی، هر سه پارامتر مهم مطرح شده را برآورده می‌سازند [9]. گسترش و تداوم استفاده از روش‌های پنهان‌نگاری گفتار مبتنی بر تبدیلاتی چون موجک، خود دلیلی بر اثر بخشی روش‌های مبتنی بر تبدیلات می‌باشد [7,13].

ایده اصلی موجک، تحلیل بر اساس مقیاس است. هر سیگنال را می‌توان با مجموعه‌ای از نسخه‌های مقیاس شده و شیفت یافته، بیان کرد. تبدیل موجک، به سبب قابلیت چند رزولوشنی خود، جنبه‌هایی از سیگنال را روشن می‌کند که سایر تبدیل‌ها از آن عاجزند و به این دلیل امکان خوبی را در اختیار قرار می‌دهد (رزولوشن فرکانسی-زمانی).

در روش ارائه شده در این مقاله، از ویژگی چند رزولوشنی تبدیل موجک بهره می‌بریم و از هر فریم سیگنال پوششی، تبدیل موجکی با مرتبه تصادفی (رزولوشن تصادفی) گرفته می‌شود و از این رو است که محل‌های پنهان‌سازی در هر فریم بصورت تصادفی تعیین می‌گردند که این موضوع سبب ایجاد امنیت خواهد شد. علاوه بر آن، از داده‌های محرمانه نیز تبدیل موجک با بیشترین مرتبه گرفته خواهد شد. سپس به منظور فشرده‌سازی داده‌های محرمانه، گزینشی از ضرایب موجک حاصل صورت می‌پذیرد و بعد از آن، ضرایب منتخب، اولویت‌بندی می‌شوند. ضرایب با اولویت بالاتر در محل‌های مطمئن‌تر و ضرایب با اولویت پائین‌تر در محل‌های با ریسک بیشتر از محل‌های تصادفی در نظر گرفته شده، جانمایی می‌شوند. ضمناً برای بالا بردن امنیت، داده‌های محرمانه‌ی منتخب، لابه‌لا قرارداده[۳] می‌شوند تا داده‌های ذخیره شده در هر فریم سیگنال

پوششی با یکدیگر همبستگی کمتری داشته باشند. متناسب با روش تصادفی معرفی شده، کلیدی جهت کشف سیگنال محرمانه در گیرنده، تولید می‌شود.
از آنجا که در روش ارائه شده، بجای رمزنگاری برای کاهش همبستگی بین داده‌ها از لابه‌لاقراردادن داده‌ها استفاده شده است، بنابراین در صورت بروز نویز تنها بخش‌هایی از داده‌ها دچار تغییر می‌شوند و تمام داده‌ها غیر قابل بازیابی نخواهد شد و با توجه به نوع داده‌های پنهان شده که ضرایب تبدیل موجک سیگنال گفتار می‌باشند، حساسیت زیادی در برابر نویز ندارند که می‌توان در مقاله‌ای دیگر به آن پرداخت.

۲- پنهان نگاری مبتنی بر تبدیل موجک

۲-۱- موجک هار

موجک هار، توسط دو موجک پدر و مادر ساخته می‌شود. این موجک‌ها به ترتیب با Φ و Ψ نشان داده می‌شوند. موجک پدر را "تابع مقیاس" نیز می‌نامند. موجک پدر هار بصورت زیر تعریف می‌شود:

$$\Phi(t) = \begin{cases} 1 & 0 \le t \le 1 \\ 0 & Otherwise \end{cases} \quad (۱)$$

همچنین موجک پدر به عنوان "تابع مشخصه" بازه‌ی واحد نیز شناخته می‌شود.
موجک مادر هار بصورت زیر تعریف می‌شود:

$$\Psi(t) = \begin{cases} 1 & 0 \le t < \frac{1}{2} \\ -1 & \frac{1}{2} \le t < 1 \\ 0 & Otherwise \end{cases} \quad (۲)$$

موجک‌های پدر و مادر به صورت زیر با هم ارتباط دارند:

$$\Psi(t) = \Phi(2t) - \Phi(2t-1) \quad (۳)$$

مطابق با این نوع نام‌گذاری، نسل اول دختران تعریف می‌شوند:

$$\Psi_{1,0}(t) = \Psi(2t) \quad و \quad \Psi_{1,1}(t) = \Psi(2t-1) \quad (۴)$$

اگرچه این طور به نظر می‌رسد که دختران فقط از مادر مشتق شده‌اند ولی می‌توان آن‌ها را از طریق پدر نیز تعریف کرد:

$$\Psi_{1,0} = \Phi(4t) - \Phi(4t-1) \quad و \quad \Psi_{1,1} = \Phi(4t-2) - \Phi(4t-3) \quad (۵)$$

در حالت کلی می‌توان گفت نسل nام دختران، 2^n موجک دارند:

$$\Psi_{n,k}(t) = \Psi(2^n t - k) \quad 0 \le k \le 2^n - 1 \quad (۶)$$

اعضای این نسل، در بازه‌هایی به طول $2^{-(n+1)}$ ثابت خواهند بود.
موجک پسران:
برای هر مقدار صحیح مثبت n داریم:

$$\Phi_{n,k}(t) = \Phi(2^n t - k) \quad (۷)$$

که در آن $0 \le k \le 2^n - 1$.

۲-۲- موجک هار و تجزیه متعامد

قبل از بیان مفهوم تقریب یک تابع توسط تابعی دیگر، لازم است که مفهوم فاصله‌ی بین توابع بیان گردد. این کار در فضاهای برداری با استفاده از مفهوم ضرب داخلی انجام می‌شود. برای هر دو تابع f و g در متعلق به یک فضای برداری تعریف شده در ($+\infty$ و $-\infty$)، ضرب داخلی بصورت زیر تعریف می شود:

$$< f, g > = \int_{-\infty}^{+\infty} f(t)g(t)dt \quad (۸)$$

با استفاده از ضرب داخلی، طول و فاصله نیز می‌توانند تعریف شوند. طول یا نرم تابع f بصورت زیر تعریف می‌گردد:

$$\|f\| = < f, f >^{\frac{1}{2}} = \left(\int_{-\infty}^{+\infty} f^2(t)dt\right)^{\frac{1}{2}} < \infty \quad (۹)$$

و فاصله‌ی بین دو تابع f و g نیز به قرار زیر است:

$$\|f - g\| = \sqrt{\int_{-\infty}^{+\infty} (f(t) - g(t))^2 dt} \quad (۱۰)$$

با معرفی مفهوم ضرب داخلی، می‌توان یک فضای برداری را فضای ضرب داخلی نامید.
تئوری تجزیه‌ی متعامد: اگر W زیرفضایی با بعد متناهی از فضای ضرب داخلی V باشد، آنگاه هر $v \in V$ را می‌توان به صورت یکتایی به شکل $v = w + w^{\perp}$ نوشت که $w \in W$ و $w^{\perp} \in W^{\perp}$. (اساس این تئوری $V = W \oplus W^{\perp}$ است).
این تئوری به صورت زیر به کار می‌رود: فرض کنید که فضای ضرب داخلی V را داریم و W زیرفضایی از آن با پایه‌های متعامد $\{w_1, w_2, \dots, w_k\}$ است و نیز فرض کنید که $v \in V$. آنگاه w، همان طور که در تئوری تجزیه‌ی متعامد توصیف شد، بردار زیر خواهد بود:

$$w = \frac{<v,w_1>}{<w_1,w_1>} w_1 + \frac{<v,w_2>}{<w_2,w_2>} w_2 + \cdots + \frac{<v,w_k>}{<w_k,w_k>} w_k = \sum_{i=1}^{k} \frac{<v,w_i>}{<w_i,w_i>} w_i \quad (۱۱)$$

بردار w، تصویر متعامد v روی W نامیده می‌شود. بردار $w^{\perp}$ ، باقی‌مانده نام دارد ($w^{\perp} = w - v$) [1].
تحلیل چند رزولوشنی، دنباله ای تودرتو

$$\cdots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \quad (۱۲)$$

از زیرفضاهای $L^2(R)$ (توابع مربع انتگرال‌پذیر روی R) با تابع مقیاس φ می‌باشد به قسمی که:

- $\cup_{n \in Z} V_n$ در $L^2(R)$ چگال است.
- $\cap_{n \in Z} V_n = \{0\}$
- $f(t) \in V_n$ اگر و تنها اگر $f(2^{-n}t) \in V_0$.

$\{\varphi(t-k)\}_{k \in Z}$ پایه‌ای متعامد یکه برای V_0 است.
برای اطلاع بیشتر به [1] رجوع کنید.

۲-۳- آماده‌سازی سیگنال گفتار محرمانه با کمک تبدیل موجک هار

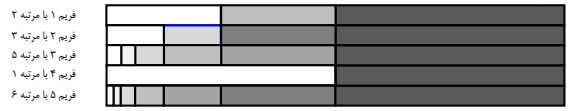
از آنجایی که داده‌ی محرمانه نیز از جنس سیگنال گفتار است و هر فریم آن، حجمی معادل یک فریم سیگنال پوششی دارد، نمی‌توان تمام نمونه‌ها را عیناً، در سیگنال پوششی جا داد. از این رو، می‌بایست شاخص‌هایی از سیگنال محرمانه را برای پنهان‌نگاری استخراج نمائیم که به کمک آن‌ها بتوانیم سیگنال گفتاری قابل فهم و شبیه به سیگنال گفتار اولیه را در گیرنده، تولید نمائیم. روش‌های مختلفی برای این کار وجود دارند، مانند استفاده از "وکودر پیشگویی خطی تحریک شده با مانده[۴]"، "وکودر پیشگویی با تحریک کد[۵]" و "کد کردن زیر باند[۶]" [3,10]. روش بکار رفته برای این منظور، روش کد کردن زیر باند

می‌باشد [3]. برای این منظور، ابتدا سیگنال گفتار محرمانه را فریم‌بندی می‌کنیم. سپس، متناسب با تعداد سمبل‌های موجود در فریم‌ها، بالاترین مرتبه‌ی ممکن برای تجزیه‌ی متعامد داده‌های هر فریم را مشخص می‌کنیم (مرتبه همان رزولوشن است). حال، با استفاده از فرمول (۱۱)، ضرایب موجک هار را برای هر فریم، به ازای بالاترین مرتبه، حساب می‌کنیم. حال می‌توان کلیات را به همراه جزئیات با ارزش بالاتر را درون آرایه‌ای ذخیره کرد و از باقی مقادیر که جزئیاتی با ارزش کمتر می‌باشند، چشم‌پوشی کرد. با این کار حجم داده‌ها به شدت کاهش خواهد یافت. هرچه از جزئیات بیشتری چشم پوشی کنیم، حجم داده‌های محرمانه کمتر خواهد شد و در عوض آن، کیفت سیگنال گفتار محرمانه‌ی بازسازی شده نیز کاهش خواهد یافت.

۲-۴- آماده سازی سیگنال گفتار پوششی با کمک تبدیل موجک هار

ابتدا می‌بایست، سیگنال گفتار پوششی، تحت یک پیش پردازش قرار گیرد تا ظرفیت‌های آن برای پنهان‌نگاری بدست آید، سپس فرآیند پنهان‌نگاری آغاز گردد.

برای آماده‌سازی سیگنال گفتار پوششی، ابتدا این سیگنال را فریم‌بندی می‌کنیم. سپس با توجه به تعداد نمونه‌ی موجود در هر فریم، بیشترین رزولوشن ممکن برای تبدیل موجک را محاسبه می‌نماییم. حال دنباله‌ای از اعداد تصادفی کوچکتر مساوی این مقدار (بیشترین رزولوشن ممکن برای تبدیل موجک) را تولید می‌کنیم. این دنباله، کلید پنهان‌نگاری است و لازم است که در زمان کشف نیز در اختیار باشد. بدین طریق تعیین می‌کنیم که در هر فریم، می‌بایست ضرایب موجک چه مرتبه‌ای محاسبه گردند و در کجاها پنهان‌سازی صورت پذیرد. مثلاً فرض کنید که بالاترین مرتبه ۶ باشد و دنباله، برابر $\{2,3,5,1,6\}$ باشد، آنگاه خواهیم داشت:



شکل (۱) : محاسبه ضرایب موجک هر فریم سیگنال پوششی با مرتبه تصادفی

داده‌های محرمانه، در قسمت‌های تیره رنگ که معرف جزئیات سیگنال صوت پوششی هستند پنهان خواهند شد. حجم کلید نیز ناچیز خواهد بود. مثلاً برای این حالت، تنها سه بیت جهت بیان مرتبه‌ی تصادفی موجک هر فریم کافی خواهد بود. این کلید از طریق کانالی دیگر، در اختیار گیرنده‌ی پیام قرار خواهد گرفت تا به واسطه‌ی آن بتواند عمل کشف پیام محرمانه را انجام دهد. با توجه به تصادفی بودن این فرآیند، ظریفت ایجاد شده برای پنهان‌نگاری، در هر بار اجرای این پروسه، متفاوت خواهد بود. از این رو، می‌بایست ظرفیت پنهان‌نگاری در هر مرتبه، محاسبه گردد تا از کافی بودن آن، قبل از ادامه‌ی فرآیند، اطمینان حاصل کنیم.

عمل پنهان‌نگاری در حوزه‌ی موجک صورت می‌پذیرد، حال از آنجا که ضرایب استفاده شده برای انجام پنهان‌نگاری در هر فریم متفاوت خواهد بود و نیز اینکه برای تعیین این محل‌ها از الگوریتم خاصی پیروی نمی‌کنیم، بلکه کاملاً تصادفی انتخاب می‌شوند، همبستگی بین محل‌های پنهان‌نگاری در هر فریم، نسبت به سایر فریم‌ها و یا بین سیگنال‌های مختلف پنهان‌نگاری شده توسط این الگوریتم، وجود ندارد و با بازگشت به حوزه‌ی زمان، این داده‌ها که به طور تصادفی پنهان شده‌اند، در سیگنال صوت پوششی پخش می‌شوند؛ از این رو است که این روش سبب بالا رفتن امنیت می‌گردد.

۲-۵- پنهان‌نگاری داده‌های محرمانه فشرده شده در ظرفیت‌های کشف شده‌ی حوزه‌ی موجک سیگنال گفتار پوششی

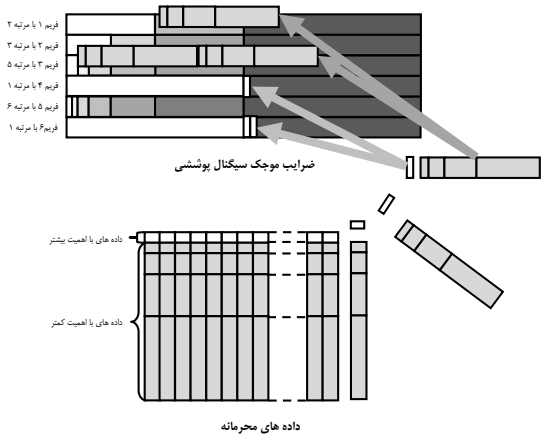
همان‌گونه که ذکر شد، در فاز پنهان‌نگاری، ابتدا داده‌های محرمانه‌ی ذخیره شده در ماتریس را به دو دسته‌ی داده‌های با اهمیت بیشتر و داده‌های با اهمیت کمتر دسته‌بندی می‌کنیم. داده‌های با اهمیت بیشتر، داده‌های بیانگر کلیات هر فریم گفتار محرمانه می‌باشند و داده‌های با اهمیت کمتر، بیانگر جزئیات هر فریم می‌باشند.

با توجه به محل تعیین شده برای جانمایی داده‌های محرمانه در هر فریم، داده‌های با اهمیت کمتر را در کم ارزش‌ترین بیت فریم‌های بامرتبه‌ی بالاتر، جاگذاری می‌کنیم. از آنجا که این جاگذاری تنها در کم‌ارزش‌ترین بیت حوزه‌ی موجک صورت می‌پذیرد و نمونه‌های انتخاب شده برای این جاگذاری، از نمونه‌های معرف جزئیات می‌باشند، لذا تأثیر قابل لمسی (توسط سیستم ادراکی انسان) در کیفیت سیگنال پوششی نخواهد گذاشت.

داده‌های محرمانه‌ی با اهمیت بالاتر را در بیت‌های با ارزش بیشتری از ضرایب موجک مرتبه‌ی پائین‌تر پنهان می‌کنیم.

باید توجه داشت که اگر در بیت‌های با ارزش بالاتر فرکانس‌های پایین‌تر معرف جزئیات، جاگذاری انجام می‌دادیم (اگر خلاف آنچه ذکر شد، عمل می‌کردیم)، ضرایب با اهمیت بالاتر دستخوش تغییرات بیشتری می‌شد که این مطلوب نیست. از طرفی ظرفیت مناسب پنهان‌نگاری جزو شاخص‌های کار بوده است پس تزریق در بیت‌های با ارزش بیشتر یک راهکار برای این مساله است. البته بیت‌های با ارزش بالاتر، کمتر نیز دستخوش تغییر می‌شوند.

این روش جاگذاری، دو حسن دارد: اول اینکه، در فریم‌هایی از سیگنال پوششی که ضرایب موجک با مرتبه‌ی پائین‌تر را حساب می‌کنیم، جزئیات حاصل، از ارزش کمتری برخوردار هستند، از این رو در صورت تزریق داده‌ی محرمانه در بیت‌های با ارزش بالاتر در آن‌ها، اطلاعات چندان مهمی، دستخوش تغییر نخواهند شد. حسن دوم این است که به علت تصادفی انتخاب شدن مرتبه‌ی موجک برای هر فریم، از آنجا که این احتمال دارای توزیع یکنواخت است بنابراین با احتمال بالائی فریم‌های با مرتبه‌ی پائین پشت سر هم قرار نمی‌گیرند، از این رو، این تزریق در بیت های با ارزش بالاتر با احتمال خوبی در فریم‌های متوالی رخ نمی‌دهد، بدین سبب تأثیر آن پخش و غیر قابل تشخیص خواهد شد. به منظور درک بیشتر از روند کار به شکل (۲) رجوع کنید.



شکل (۲) : پنهان‌نگاری بر اساس اولویت داده‌های محرمانه و داده‌های پوششی

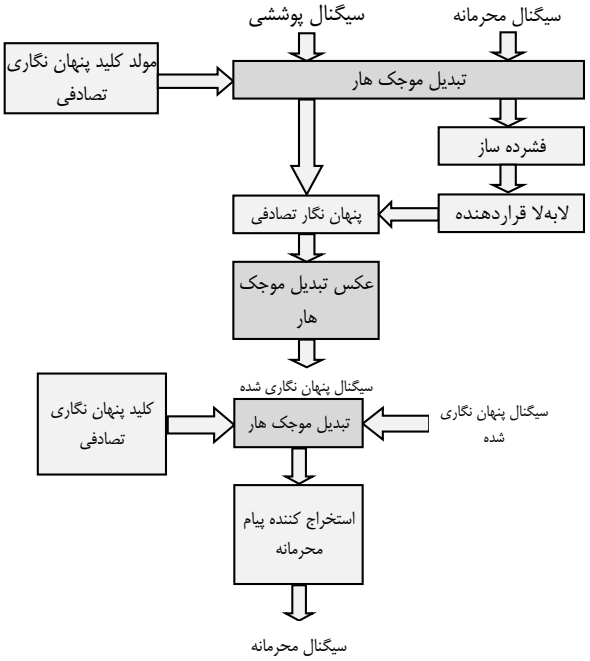
۲-۶- کشف سیگنال گفتار محرمانه با استفاده از کلید پنهان‌نگاری

در سمت دریافت کننده‌ی پیام، سیگنال گفتار دریافتی، فریم بندی می‌شود. سپس با توجه به کلید، از هر فریم، تبدیل موجک با مرتبه‌ی مناسب گرفته می‌شود و داده‌های محرمانه از بیت‌هایی که در آن‌ها جایگزین شده بودند، بازخوانی خواهند شد. برای کشف پیام در گیرنده، لازم است که کلید صحیح در اختیار باشد. در روش ارائه شده، تعیین محل داده‌های محرمانه تنها از طریق کلید پنهان‌نگاری ممکن خواهد بود و با توجه به اینکه در تولید این کلید از هیچ الگوریتمی پیروی نمی‌کنیم و تنها بر حسب تصادف آن را تولید می‌نمائیم، این کلید قابل کشف نخواهد بود. ضمناً به دلیل آنکه داده‌های محرمانه نیز در فرآیند پیش‌پردازش، لابه‌لا قرارداده شده‌اند، همبستگی بین آن‌ها کاهش یافته و با ادغام این دو روش، امنیت خوبی برای داده‌های محرمانه ایجاد شده است که کشف آن‌ها را مشروط به در اختیار داشتن کلید پنهان‌نگاری ساخته است و در صورت در اختیار بودن حجم بالایی از سیگنال‌های پنهان‌نگاری شده، نمی‌توان به ارتباطی بین آن‌ها پی برد.

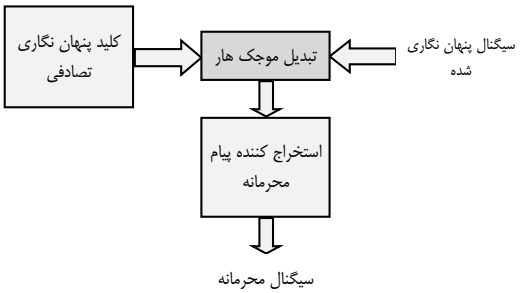
۲-۷- بازسازی سیگنال گفتار محرمانه

پس از استخراج داده‌های محرمانه از سیگنال پوششی، مقادیر در یک آرایه mxn ذخیره می‌شوند که m معرف تعداد سمبل‌ها در هر فریم و n معرف تعداد فریم‌ها است. با توجه به آنکه داده‌های اولیه، برای امنیت بیشتر لابه‌لا قرارداده شده بودند، می‌بایست عکس آن، به ماتریس حاصل اعمال شود تا چیدمان درست، ایجاد گردد.

در شکل‌های (۳) و (۴) شماتیک روش ارائه شده، آورده شده است.



شکل (۳) : فرآیند پنهان‌نگاری پیام محرمانه



شکل (۴) : فرایند استخراج پیام محرمانه

۳- نتیجه گیری

۳-۱- دادگان مورد استفاده

جهت سنجش عملکرد روش ارائه شده در این مقاله، از ۲۰ سیگنال گفتار محرمانه (با طول زمانی بین ۱/۹۱۷ تا ۵۸/۱ ثانیه) و نیز ۵۰ سیگنال گفتار پوششی (با طول زمانی بین ۶۱/۳۱ تا ۱۹۱/۱۴ ثانیه) با مشخصات زیر استفاده شده است:

جدول (۱) : مشخصات سیگنال‌های گفتار بکار رفته جهت ارزیابی الگوریتم

سیگنال	نوع کدینگ	نرخ نمونه برداری	کانال ها	رزولوشن
پوششی	Windows PCM(Wav)	16 KHz	mono	16 bits
محرمانه	Windows PCM(Wav)	8 KHz	mono	8 bits

جهت سنجش دقیق عملکرد الگوریتم پنهان‌نگاری ارائه شده، سیگنال‌های پوششی، سیگنال‌هایی بسیار تمیز و باکیفیت (با نسبت سیگنال به نویز ۶۰ دسیبل) می‌باشند.

۳-۲- ارزیابی عملکرد الگوریتم پنهان‌نگاری ارائه شده

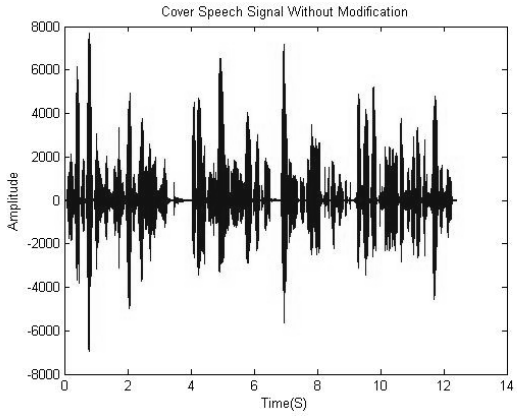
هرچه سیگنال گفتار پوششی اولیه تفاوت کمتری نسبت به سیگنال گفتار پنهان-نگاری شده داشته باشد، روش پنهان‌نگاری از امنیت بالاتری برخوردار خواهد بود. برای این منظور از معیار PESQ[۷] جهت سنجش بهره می‌بریم. محاسبه‌ی این معیار در کنار مقایسه‌ی مقادیر SDR[۸]، می‌تواند، سنجش خوبی را سبب گردد. البته مقایسه‌ی ظاهری سیگنال‌های حوزه‌ی زمان و فرکانس پوششی و پنهان‌نگاری شده نیز خالی از لطف نخواهد بود. بنابر این جهت ارزیابی کارایی روش پنهان‌نگاری، از چند روش استفاده شده است:

- مقایسه‌ی شکل حوزه‌ی زمان سیگنال گفتار پوششی و پنهان‌نگاری شده
- مقایسه ی طیف سیگنال گفتار پوششی و پنهان‌نگاری شده
- استفاده از معیار PESQ جهت قیاس سیگنال‌های گفتار پوششی و پنهان-نگاری شده
- محاسبه ی SDR قطعه‌ای (زمانی)
- میزان خطا در بازیابی اطلاعات محرمانه
- مقایسه‌ی هیستوگرام سیگنال گفتار پوششی و پنهان‌نگاری شده
- بررسی ظرفیت پنهان‌نگاری

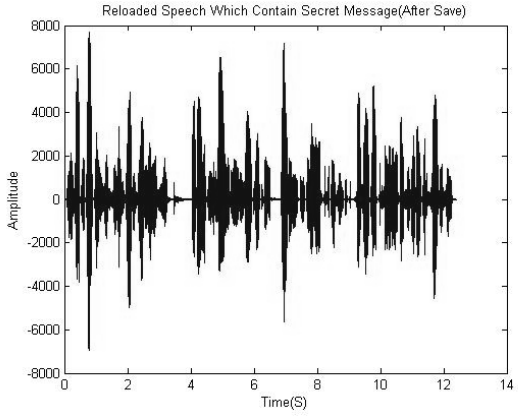
که در ادامه، نتایج حاصل بیان خواهد شد.

۳-۳- مقایسه شکل حوزه زمان سیگنال‌های گفتار پوششی و پنهان نگاری شده

چنانچه به منظور ارزیابی روش ارائه شده در این مقاله، از سیگنال گفتار پوششی شکل (۵) جهت پنهان‌نگاری داده‌های محرمانه مورد نظر، استفاده نمائیم، سیگنال گفتار پنهان‌نگاری شده نمایش داده شده در شکل (۶) حاصل خواهد شد. همان‌طور که مشاهده می‌شود، سیگنال گفتار پوششی و پنهان‌نگاری شده در حوزه زمان، کاملاً شبیه یکدیگرند و این در کنار معیارهای دیگری که در ادامه ذکر خواهد شد، دلیلی بر عملکرد خوب روش ارائه شده در این مقاله است. با تکرار سنجش فوق بر روی سایر سیگنال‌های معرفی شده در زیربخش ۳-۱، نتایجی مشابه حاصل شد که در اینجا به منظور جلوگیری از تکرار از ذکر آن‌ها اجتناب می‌نمائیم.

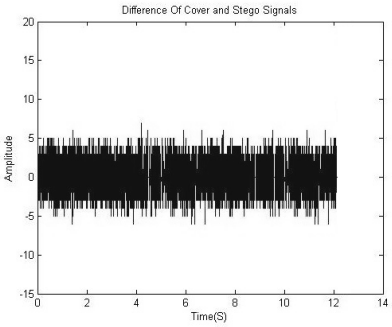


شکل (۵) : سیگنال گفتار پوششی در حوزه‌ی زمان



شکل (۶) : سیگنال گفتار پنهان‌نگاری شده در حوزه‌ی زمان

حال چنانچه تفاضل دو سیگنال گفتار پوششی و پنهان‌نگاری شده را رسم نمائیم (شکل (۷))، مشاهده خواهیم کرد که سیگنال باقیمانده بسیار کم دامنه و نویز مانند است (متوسط دامنه سیگنال باقیمانده حدود $\frac{1}{400}$ متوسط دامنه سیگنال-های گفتار پوششی و پنهان‌نگاری شده می‌باشد) و هیچ شباهتی به سیگنال گفتار ندارد. این نیز دلیلی دیگر بر کارایی روش ارائه شده در این مقاله می‌باشد.

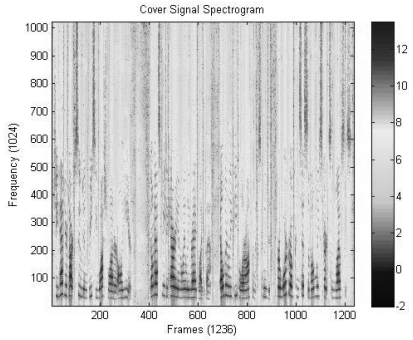


شکل (۷) : تفاضل سیگنال‌های گفتار پوششی و پنهان‌نگاری شده در حوزه‌ی زمان

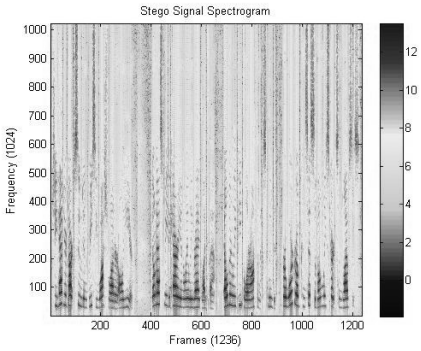
۳-۴- مقایسه طیف سیگنال گفتار پوششی و پنهان-نگاری شده

چنانچه طیف دو سیگنال معرفی شده در بخش قبلی را طی شکل‌های (۸) و (۹) با یکدیگر مقایسه نمائیم، خواهیم دید که طیف سیگنال‌های گفتار پوششی

و پنهان‌نگاری شده، بسیار شبیه یکدیگر است و این نیز مبین عملکرد خوب الگوریتم پنهان‌نگاری ارائه شده طی این مقاله می‌باشد.
توجه شود که در شکل‌های (۸)، (۹) و (۱۰)، مقدار ۱۰۲۳ در محور عمودی، معرف ۱۰۲۴ امین ضریب فوریه است که معادل فرکانس ۱۶ کیلوهرتز می‌باشد.

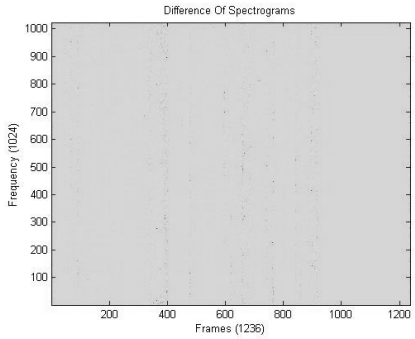


شکل (۸) : طیف سیگنال گفتار پوششی



شکل (۹) : طیف سیگنال گفتار پنهان‌نگاری شده

چنانچه تفاضل دو طیف سیگنال گفتار پوششی و پنهان‌نگاری شده را رسم نمائیم (شکل (۱۰))، مشاهده خواهیم کرد که تفاوت این دو طیف بسیار ناچیز است و این خود دلیلی بر کارایی روش ارائه شده در این مقاله می‌باشد.



شکل (۱۰) : تفاضل طیف سیگنال‌های گفتار پوششی و پنهان‌نگاری شده

۳-۵- استفاده از معیار PESQ جهت قیاس سیگنال-های گفتار پوششی و پنهان‌نگاری شده

چنانچه سیگنال‌های گفتار پوششی معرفی شده در زیربخش ۳-۱ را جهت سنجش کارایی روش ارائه شده بکار ببریم و سیگنال‌های گفتار پنهان‌نگاری شده حاصل را با سیگنال‌های گفتار پوششی متناظر، طبق معیار PESQ مقایسه نمائیم، جدول (۲) حاصل می‌شود. به منظور حفظ اختصار، از بیان تمام موارد اجتناب شده است ولی نتایج حاصل مشابه موارد ذکر شده در جدول (۲) می‌باشد. نزدیکی مقادیر حاصل، به مقدار بیشینه (عدد پنج) بیانگر کارایی روش ارائه شده در این مقاله می‌باشد.
در جدول (۲)، NB مخفف Narrow Band، WB مخفف Wide Band، MOS مخفف Mean Opinion Score و نیز LQO مخفف Listening Quality Objective می‌باشد [4,5,6,9].
NB MOS طبق استاندارد P.862.1 معیاری برای تخمین MOS در حالت باند باریک است، WB MOS طبق استاندارد P.862.2 معیاری برای تخمین MOS در حالت باند پهن است که معرف مقادیر نگاشت شده به LQO می-باشند، زیرا MOS معیاری ادراکی است در حالی که ما در اینجا از سنجش PESQ که معیاری عقلی است، جهت تخمین آن بهره برده‌ایم از این رو است که آن را طبق استاندارد P.800.1، MOS LQO نامیده‌ایم.

جدول (۲) : معیار PESQ به ازای سنجش‌های مختلف

سیگنال	NB MOS LQO	WB MOS LQO
۱	4.449	4.496
۲	4.449	4.496
۳	4.449	4.496
۴	4.539	4.518
۵	4.540	4.518
۶	4.540	4.518
۷	4.536	4.603
۸	4.533	4.594
۹	4.536	4.601
۱۰	4.531	4.597
۱۱	4.533	4.602
۱۲	4.533	4.592
۱۳	4.541	4.520
۱۴	4.540	4.519

۳-۶- محاسبه SDR قطعه‌ای (زمانی)

ابتدا سیگنال گفتار پوششی و پنهان‌نگاری شده، از نظر زمانی همردیف می‌شوند و سپس فرمول زیر اعمال می‌گردد. M تعداد فریم‌های سیگنال و N طول هر فریم می‌باشد.

$$SDR_{seg} = \frac{10}{M}\sum_{m=0}^{M-1}\log_{10}\frac{\left(\sum_{n=N_m}^{N_m+N-1}(S_n)^2\right)}{\sum_{n=N_m}^{N_m+N-1}(S_n-\tilde{S}_n)^2} \qquad (۱۳)$$

در فرمول بالا، S_n معرف سیگنال گفتار پوششی و $\tilde{S}_n$ معرف سیگنال پنهان-نگاری شده می‌باشد.
این معیار، برای تعیین کیفیت سیگنال مناسب است و در کنار معیار PESQ می‌تواند روش خوبی برای ارزیابی باشد.
چنانچه سیگنال‌های گفتار پوششی معرفی شده در زیربخش ۳-۱ را جهت سنجش کارایی روش ارائه شده بکار ببریم و از سیگنال‌های گفتار پنهان‌نگاری شده حاصل به همراه سیگنال‌های گفتار پوششی متناظر، طبق فرمول (۱۳) استفاده نمائیم، جدول (۳) حاصل می‌شود. به منظور حفظ اختصار، از بیان تمام موارد اجتناب شده است ولی نتایج حاصل مشابه موارد ذکر شده در جدول (۳)

می‌باشد. مقادیر بالای SDRقطعه‌ای، بیان‌گر کارایی روش ارائه شده در این مقاله می‌باشد.

جدول (۳) : مقدار SDR قطعه‌ای (زمانی)

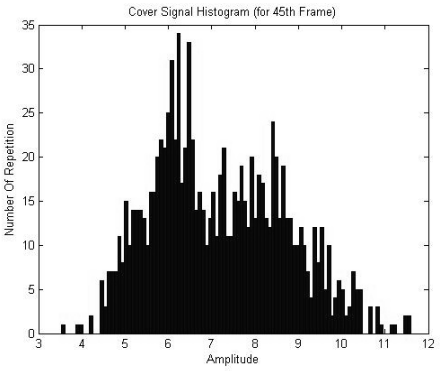
سیگنال	SegSDR
۱	62.3905
۲	61.2508
۳	63.8002
۴	66.4733
۵	61.0961
۶	64.2434
۷	56.6359
۸	54.5281
۹	61.8165
۱۰	53.1149
۱۱	54.8312
۱۲	53.2545
۱۳	61.0241
۱۴	64.0803

۳-۷- میزان خطا در بازیابی اطلاعات محرمانه

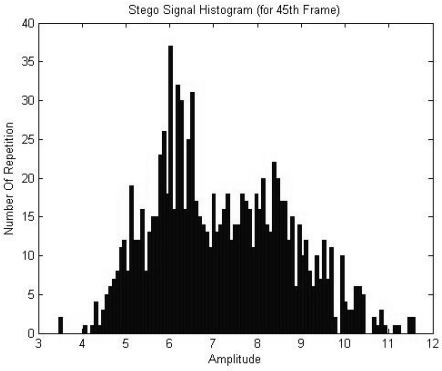
در ادامه چنانچه سیگنال‌های گفتار پوششی ذکر شده را جهت سنجش کارایی روش ارائه شده بکار ببریم و داده‌های محرمانه اولیه را با داده‌های محرمانه بازیابی شده قیاس نمائیم، متوجه خواهیم شد که همان‌گونه که انتظار آن را داشتیم، میزان خطا در بازیابی اطلاعات محرمانه در روش ارائه شده در این مقاله، صفر خواهد بود.

۳-۸- مقایسه هیستوگرام سیگنال گفتار پوششی و پنهان‌نگاری شده

از آنجا که اثر هر تغییر حوزه تبدیلی در حوزه زمان پخش می‌شود، پی‌گیری اثر جاگذاری‌های صورت گرفته در حوزه فرکانس، مشهودتر خواهد بود در حالی که این آثار در حوزه زمان چندان قابل رؤیت نخواهند بود. لذا در شکل‌های (۱۱) و (۱۲)، هیستوگرام سیگنال‌های گفتار پوششی و پنهان‌نگاری شده در حوزه فرکانس برای فریم چهل و پنجم (فریمی که بیشترین تزریق داده محرمانه را در آن صورت گرفته است)، مورد قیاس قرار گرفته‌اند. شباهت زیاد هیستوگرام این دو سیگنال، دلیل دیگری بر کارایی روش ارائه شده در این مقاله می‌باشد.



شکل (۱۱) : هیستوگرام سیگنال گفتار پوششی



شکل (۱۲) : هیستوگرام سیگنال گفتار پنهان‌نگاری شده

۳-۹- ظرفیت پنهان‌نگاری

با توجه به تصادفی بودن روش پنهان‌نگاری و نیز قابل تغییر بودن میزان فشرده-سازی سیگنال محرمانه (سیگنالی که قرار است پنهان شود)، نمی‌توان عدد دقیقی برای معرفی ظرفیت پنهان‌سازی این روش بیان کرد ولی به صورت تجربی در روش ارائه شده طی این پژوهش به طور متوسط ۲۸٪ طول زمانی سیگنال گفتار پوششی، ظرفیت پنهان‌نگاری برای سیگنال گفتار محرمانه وجود دارد. در صورت تزریق در تعداد بیت‌های بیشتر در عوض تزریق در کم ارزش-ترین بیت و یا فشرده‌سازی بیشتر سیگنال محرمانه، این مقدار قابل افزایش خواهد بود.

۳-۱۰- مقایسه روش ارائه شده با روشی مشابه

آقای رکیک و همکارانش در سال ۲۰۱۲ روشی مشابه را ارائه دادند که در آن ضرایب فرکانس‌های پائین سیگنال گفتار محرمانه در ضرایب کم دامنه و فرکانس بالای تبدیل فوریه دامنه حاصل از تبدیل موجک سیگنال گفتار پوششی پنهان می‌شد [11]. در جدول (۴) نتایج حاصل از دو روش جهت مقایسه، ذکر شده است.

جدول (۴) : مقایسه مقدار متوسط PESQ و SDR

مقدار متوسط	روش آقای رکیک	روش ارائه شده در این مقاله
PESQ	3.68	4.497
SDR	32.11	59.895

دو روش مقایسه شده، دارای شرایط نزدیکی هستند:

- در هر دو روش، سیگنال محرمانه و پوششی از جنس صوت می‌باشد.
- پنهان‌نگاری در هر دو روش در حوزه تبدیل انجام می‌شود.

– از تبدیل موجک در هر دو روش استفاده شده است.

بنابراین، جدول (۴) به خوبی نشان دهنده برتری و کارآمدی روش ارائه شده در این مقاله می‌باشد.

۳-۱۱- جمع بندی

شباهت کامل سیگنال‌های گفتار پوششی و پنهان‌نگاری شده در دو حوزه زمان و فرکانس، تطابق ادراکی بسیار بالای این دو سیگنال (طبق معیارهای PESQ و SDRقطعه‌ای)، عدم وجود خطا طی بازیابی اطلاعات محرمانه در روش ارائه شده، ظرفیت پنهان‌نگاری مناسب، شباهت زیاد هیستوگرام دو سیگنال و نیز مقایسه انجام شده با روشی مشابه؛ مبین کارایی و امنیت بالای این روش از پنهان‌نگاری می‌باشد.

روش ارائه شده در این مقاله از پنج رویکرد زیر جهت ایجاد امنیت بهره می‌برد:

– کاهش همبستگی داده‌های محرمانه

– اولویت بندی داده‌های محرمانه

– از بین بردن همبستگی بین محل‌های پنهان‌سازی (تعیین تصادفی محل‌های پنهان‌سازی از میان محل‌های مناسب برای این منظور)

– اولویت بندی محل‌های پنهان‌سازی انتخاب شده به صورت تصادفی

– پنهان‌سازی براساس اولویت داده‌های محرمانه و نیز اولویت محل‌های پنهان-سازی

تعیین تصادفی محل‌های پنهان‌سازی در حوزه موجک، مزیت اصلی این روش از پنهان‌نگاری نسبت به روش‌های مشابه است که با از میان بردن همبستگی بین محل‌های پنهان‌سازی سبب امنیت بیشتر آن خواهد شد. بدین صورت که در صورت موجود بودن الگوریتم پنهان‌نگاری، کشف اطلاعات محرمانه تنها با در اختیار داشتن کلید تصادفی بکار رفته در آن پنهان‌نگاری، ممکن خواهد بود و حتی با تحلیل حجم بسیار بالایی از داده‌های پنهان‌نگاری شده، نمی‌توان به محل‌های پنهان‌سازی پی‌برد.

مراجع

[1] Aboufadel, E., Schlicker, S., *Discovering Wavelets*, Wiley-Interscience Publication, 1999.

[2] Bender, W., Gruhl, D., Morimoto, N., "*Techniques for Data Hiding"*, IBM Syst, Vol.35, No. 3, pp. 313–336, 1996.

[3] Huang, X. D., Acero, A., Hon, H. W., "*Spoken Language Processing"*, Prentice-Hall, 2000.

[4] ITU-T P.862.1, *Mapping function for transforming,* 2003.

[5] ITU-T P.862, *Raw Result Scores to MOS-LQO*, 2003.

[6] ITU-T P.800.1, *Mean Opinion Score (MOS) Terminology* , 2006.

[7] Indra, M., Reddy, S., Kumar, *A. P. S., "Secured Data Transmission Using Wavelet Based Steganography and Cryptography by Using AES Algorithm",* International Conference on Computational Modeling and Security, 2016.

[8] Kirovski, D., Malvar, H., "*Spread Spectrum Watermarking of Audio Signals*", IEEE Trans. Signal Process, Vol. 51, No. 4, pp. 1020-1033, 2003.

[9] Pooyan, M., Delforouzi, A., "*LSB-based Audio Steganography Method Based on Lifting Wavelet Transform"*, IEEE International Symposium on Signal Procssing and Information Technology, pp. 600-603, 2005.

[10] Rabiner, L., Juang, B. H., "*Fundamentals of Speech Recognition",* Prentice-Hall, 1993.

[11] *Rekik,S., Guerchi,D., Selouani,S. A., Hamam,H., "Speech Steganography Using Wavelet and Fourier Transforms"*, EURASIP Journal on Audio, Speech and Music Processing, 2012.

[12] Sridevi, R., Damodaram, A., Narasimham, S. V. L., "*Efficient Method of Audio Steganography by Modified LSB Algorithm and Strong Encryption Key with Enhanced Security",* J. Theor. Appl. Inf. Technol, Vol. 5, No. 6, pp. 768-771, 2009.

[13] *Seyyedi,S. A., Sadau, V., Ivanov, N., "A Secure Steganography Method Based on Integer Lifting Wavelet Transform"*, International Journal of Network Security, Vol. 18, No. 1, PP.124-132, 2016.

[14] Yan, D., Wang, R., Yu, X., Zhu, J., "*Steganography for MP3 Audio by Exploiting the Rule of Window Switching"*, ELSEVIERE, pp. 704-716, 2012.

زیرنویس‌ها

[1] Stego signal

[2] Least Significant Bit

[3] Interleave

[4] Residual Excited Linear Prediction

[5] Code Excited Linear Prediction

[6] Sub-band Coding

[7] Perceptual Evaluation of Speech Quality

[8] Signal to Distortion Ratio