طرح نهان نگاری در سیگنال صحبت با شفافیت بالا با استفاده از الگوریتم ژنتیک پیوسته

حجت اله مقدسی [۱]، محمد فخر دانش [۲]، کیان کیقباد [۳]

۱ کارشناس ارشد مخابرات رمز، دانشگاه صنعتی مالک اشتر، h_moghadasi@mut.ac.ir

۲ عضو هیات علمی مجتمع فناوری اطلاعات ارتباطات و امنیت (ICT)، دانشگاه صنعتی مالک اشتر

۳ استادیار مجتمع فناوری اطلاعات ارتباطات و امنیت (ICT)، دانشگاه صنعتی مالک اشتر

تاریخ دریافت: ۹۳/۲/۱۰ تاریخ پذیرش: ۹۴/۱۰/۲

چکیده

در این مقاله یک روش نهان‌نگاری در سیگنال صحبت در حوزه‌ی موجک و با جانشینی غیرمستقیم بیتی ارائه می‌شود. در این روش ابتدا از نمونه‌های گسسته سیگنال میزبان و سیگنال صحبت مخفی، تبدیل موجک گرفته می‌شود. سپس با به‌کارگیری الگوریتم وراثتی پیوسته، ضرایب موجک سیگنال صحبت مخفی، به‌طور مؤثر و کارآمد در ضرایب موجک سیگنال میزبان ادغام می‌شوند. در این الگوریتم برای بهبود عملکرد انتخاب، از روش‌های مقیاس کردن شایستگی برای تنظیم فشار انتخاب استفاده شده است. همچنین بر حفظ تنوع جمعیت از یک‌طرف و تلاش برای جلوگیری از ایجاد گونه‌های غالب از طرف دیگر، ترکیبی از جایگزینی نخبه‌گرا و جایگزینی ازدحامی به همراه روش جلوگیری از ازدواج با محارم استفاده می‌شود. به دلیل خطای کوانتیزه کردن، در گیرنده بین سیگنال مخفی قبل و بعد از نهان نگاری اختلاف وجود دارد، اما این اختلاف دارای الگوی نویز گوسی مناسب است، به همین دلیل، فشرده‌سازی آن قضیه اول شانون را ارضا کرده و به حد شانون که همان آنتروپی است، بسیار نزدیک است. بنابراین نویز را نیز به روش فشرده‌سازی بدون اتلاف هافمن، همراه سیگنال نهان نگاشته ارسال می‌کنیم. نتایج تجربی و بررسی آزمایش‌ها کیفیت مطلوب سیگنال نهان‌نگاشته و ظرفیت بالای نهان‌نگاری را نشان می‌دهد. از طرفی تحلیل آماری نتایج بهبود مقادیر میانگین، واریانس، چولگی و کشیدگی را در مقایسه با سه روش پوشش فرکانسی، جانشینی کم‌ارزش‌ترین بیت و پوشش مؤثر موجک در حوزه‌ی زمان و فرکانس نشان می‌دهد. همچنین با به‌کارگیری الگوریتم وراثتی پیوسته به‌طور متوسط ۱۰٪ به ظرفیت نقاط مناسب برای ادغام افزوده شده است.

کلیدواژه

جانشینی غیرمستقیم بیتی ، الگوریتم ژنتیک پیوسته، قضیه اول شانون، آنتروپی، جایگزینی ازدحامی، نخبه‌گرایی

مقدمه

امروزه با توسعه‌ی سریع فن‌آوری‌های نوین، دسترسی غیرمجاز افراد و سازمان‌های دولتی و خصوصی به اطلاعات در حال گسترش است. امنیت اطلاعات یک حوزه مهم است که در سالیان اخیر پیشرفت‌های زیادی کرده است و با تکنیک‌هایی از قبیل رمزنگاری[۱]، نهان نگاری[۲]، نشانه‌گذاری[۳] آمیخته شده است. دانش نهان نگاری ابزاری مهم و نیرومند برای تبادل امن اطلاعات هست. نهان نگاری علم و هنر انتقال پیام‌های مخفی به طریقی است که به جز گیرنده‌ی مورد نظر، هیچ شخص دیگری از وجود پیام مطلع نشود. در نهان نگاری، اغتشاش و تغییری که از لحاظ ادراکی یا تحلیلی قابل کشف باشد، نباید در درون سیگنال میزبان رخ دهد. منظور از کلمه‌ی ادراک، قوه‌ی ادراک آدمی (مثل بینایی یا شنوایی) است که نبایستی قادر به تشخیص حضور وجود پیام مخفی در سیگنال میزبان شود و منظور از کلمه‌ی تحلیل این است که با استفاده از تحلیل‌های آماری نبایستی وجود پیام در سیگنال میزبان تشخیص داده شود. در یک سیستم نهان نگاری سیگنالی که ارسال مخفی آن هدف اصلی سیستم است، سیگنال پیام یا داده‌ی جاسازی شده، سیگنالی که به عنوان بستری برای مخفی کردن سیگنال پیام بکار می‌رود، سیگنال پوشش یا میزبان و سیگنالی که در آن سیگنال پیام جاسازی شده است را سیگنال

1 Cryptography
2 Steganography
3 Watermarking

نهان نگاری‌شده یا نهان نگاشته می‌نامند. سه پارامتر وجود دارد که کیفیت کار را در بحث مخفی سازی اطلاعات نشان می‌دهد که عبارت است از: شفافیت که توانایی برای اجتناب از سوء ظن درباره وجود یک پیام مخفی است، ظرفیت نهان نگاری که نشان‌دهنده میزان اطلاعات پنهانی در سیگنال میزبان است و مقاومت که قابلیت عدم آسیب‌پذیری در برابر حملات عمدی یا غیر عمدی است. تفاوت نهان نگاری و رمزنگاری در این است که در نهان نگاری اصل وجود پیام، مخفی و نامفهوم است ولی در رمزنگاری محتوای پیام، نامفهوم و مخفی است. مزیت نهان نگاری در مقابل رمزنگاری این است که در نهان نگاری پیام، توجهی را به خود جلب نمی‌کند، درحالی‌که یک پیام رمزنگاری شده، صرف‌نظر از اینکه به چه میزان غیرقابل شکست باشد، هر ناظری را نسبت به خود بدبین کرده و می‌تواند به تنهایی حاکی از یک توطئه باشد. تفاوت نهان نگاری و نشانه‌گذاری در این است که برخلاف نهان نگاری که هدف آن قرار دادن یک پیام به گونه‌ای است که وجود پیام، سری و مخفی باشد، هدف از نشانه‌گذاری، قرار دادن علائم و اطلاعاتی در محصولات دیجیتال است که امکان حذف یا تغییر این علائم وجود نداشته باشد. از این طریق حق مالکیت یک محصول دیجیتال حفظ می‌شود. البته باید توجه کرد که در نشانه‌گذاری نیز مانند نهان نگاری مخفی بودن پیام می‌تواند مطرح باشد ولی تأکید اصلی بر عدم امکان تغییر یا حذف علائم است، نه بر عدم امکان تشخیص وجود پیام. در مقابل، در نهان نگاری نیز عدم امکان تغییر یا حذف پیام می‌تواند مطرح باشد ولی اولویت بر عدم امکان کشف وجود پیام است؛ به عبارت دیگر، در نشانه‌گذاری اهمیت سیگنال میزبان بیشتر از سیگنال پیام است، ولی در نهان نگاری، اهمیت سیگنال پیام بیشتر از سیگنال میزبان است.

به طور کلی یک سامانه‌ی مخفی سازی اطلاعات در رسانه‌های مختلف نظیر صوت، تصویر و ویدئو بایستی ظرفیت بالای نهان نگاری، شفافیت آماری (تحلیلی) و ادراکی مطلوب قبل و بعد از نهان نگاری را تضمین کند [۱۰]. شفافیت آماری به مراتب از شفافیت ادراکی سخت تر هست زیرا سیستم شنیداری انسان مشمول فریب خوردن هست ولی آمار و ارقام مشمول فریب خوردن نیست. به عنوان نمونه امکان پوشش یک سطح پایین صدا در یک سطح بالای صدا بدون تشخیص سیستم ادراکی وجود دارد [۱۶]. الگوریتم‌های نهان‌نگاری صوتی عموماً به دو گروه حوزه‌ی زمان و حوزه‌ی تبدیل تقسیم‌بندی می‌شوند. الگوریتم‌های حوزه‌ی تبدیل نسبت به الگوریتم‌های حوزه‌ی زمان در مقابل حملات مقاوم‌تر می‌باشند، اما مقاومت آن‌ها در سیگنال‌هایی که تعداد مؤلفه‌های حوزه‌ی تبدیل‌شان بسیار کم می‌باشد، رضایت‌بخش نیست [۱۳]. در ضمن این الگوریتم‌ها نسبت به الگوریتم‌های حوزه‌ی زمان پیچیده‌تر و از زمان بیش‌تری استفاده می‌کنند [۱۲]. حوزه‌ی تبدیل، حوزه‌ی پرکاربردتری است و یکی از معروف‌ترین حوزه‌ها، حوزه‌ی تبدیل موجک است که در سال‌های اخیر به دلیل ظرفیت بالای ارسال تا ۲۰۰ kbps توجه بیشتری را به خود جلب کرده است [۱۸]. در این مقاله نیز نهان نگاری در حوزه‌ی تبدیل موجک و به صورت جانشینی غیرمستقیم صورت می‌گیرد. قبل پرداختن به روش پیشنهادی به چند روش به کار گرفته‌شده در مقالات اشاره می‌شود و سپس الگوریتم پیشنهادی را با دو رویکرد تشریح می‌کنیم.

روش افزایش ظرفیت نهان‌نگاری در فایل‌های صوتی با استفاده از روش ذخیره‌ی غیریکنواخت در ضرایب موجک مترقی (LWT)[۴] توسط امید اسلامی و همکارانش ارائه شد. این روش یک رویکرد نهان‌نگاری مبتنی بر تغییر بیت‌های کم‌ارزش ارائه می‌دهد که در آن ضرایب کم‌ارزش موجک مترقی سیگنال صوتی بیت‌ها را متناسب با بیت‌های داده تغییر می‌دهد. تعداد بیت‌های قابل‌تغییر در هر زیرباند با آستانه‌ی شنوایی که برای آن زیرباند تعریف می‌شود، متناسب است. تأکید اصلی در این روش، افزایش ظرفیت نهان‌نگاری بدون کاهش کیفیت شنیداری صوت نهان نگاشته و بازیابی کامل اطلاعات مخفی‌شده در صوت است. در این روش ابتدا فایل صوتی موردنظر به اندازه‌های ۶۴ تایی فریم‌بندی می‌شود. سپس به نمونه‌های گسسته در زمان سیگنال پوشش (میزبان)، تبدیل موجک مترقی با فیلتر هار اعمال می‌شود تا ضرایب صحیح به دست آیند. با استفاده از این ضرایب، آستانه‌ی شنوایی را محاسبه می‌کنیم. آستانه‌ی شنوایی (آستانه‌ی مطلق) در حقیقت مرز بین شنیدن و نشنیدن گوش است به‌شرط این‌که هیچ سیگنال قابل شنیده شدن دیگری در محیط اطراف گوش در حال پخش نباشد. در حوزه‌ی زمان این آستانه به‌صورت یکنواخت توزیع می‌شود ولی در حوزه‌ی فرکانس حد این آستانه در فرکانس‌های مختلف، متفاوت است به‌طوری‌که آستانه‌ی شنوایی در فرکانس‌های بالا و خیلی پایین، بالا است ولی در فرکانس‌های میانی میزان آستانه‌ی سکوت بسیار پایین است. اگر از سیگنال صوت میزبان تبدیل موجک بگیریم و این‌کار را تا آن‌جا انجام دهیم که تمام زیر باندها متعلق به سطح پنجم باشند نتیجه‌ی آن، رشته‌هایی خواهد بود که به ۳۲ زیرباند قابل‌تفکیک است. این نوع اعمال تبدیل موجک اصطلاحاً موجک پاکت[۵] گفته می‌شود. در این روش، اطلاعات در ضرایب تبدیل موجک صوت مخفی می‌شود. میزان اطلاعات واردشده به هر ضریب، به حساسیت شنوایی گوش انسان در آن محدوده بستگی دارد. هرچه محدوده‌ی فرکانس زیرباند موردنظر به فرکانس‌های پایین و یا بالا نزدیک‌تر باشد، حساسیت کمتری نسبت به تغییرات اعمال‌شده در آن از خود نشان می‌دهد، درنتیجه حجم بیشتری از اطلاعات را در این دسته از ضرایب می‌توان مخفی کرد. این روش به‌طور میانگین قابلیت دارا بودن

4 Lifting Wavelet Transform
5 Packet Wavelet

ظرفیتی معادل ۳۶۷ بیت در هر فریم ۶۴ تایی را دارا است. در این روش سیگنال به نویز (SNR)[6] عددی معادل ۵۰/۳۷ dB است، که نشان از کارایی بالای ادغام است. از طرفی میانگین آرا امتیازات (MOS)[7] ، به‌طور متوسط دارای مقدار ۴/۸ است که نشان‌دهنده‌ی کیفیت شنیداری بالای فایل‌های صوتی است [۲].

یکی دیگر از روش‌ها، ادغام ظرفیت بالا در کم‌ارزش‌ترین بیت ضرایب در حوزه‌ی موجک می‌باشد که توسط سیوجیک[8] و همکارانش ارائه شد. در این روش ابتدا سیگنال میزبان در نمونه‌های ۵۱۲ تایی فریم‌بندی می‌شود. سپس از سیگنال فوق، تبدیل موجک گسسته در سطح پنجم گرفته می‌شود و ۳۲ زیرباند ساخته می‌شود. سپس با توجه به بیش‌ترین مقدار در هر زیرباند، ضرایب موجک سیگنال میزبان مقیاس‌دهی می‌شوند و یک آرایه از اعداد باینری حاصل‌شده و سپس مکمل ۲ آن‌ها محاسبه می‌شود. در این‌کار برای جانشینی مناسب سیگنال پیام مخفی در سیگنال میزبان یک تضعیف اولیه انجام می‌گیرد که به‌عنوان نمونه برای جانشینی در ۸ بیت کم‌ارزش، تضعیف ۴۸ dB- بر روی سیگنال پیام مخفی صورت می‌گیرد. مقایسه‌ی روش نهان‌نگاری فوق در ضرایب کم‌ارزش در حوزه‌ی موجک نسبت به نهان‌نگاری در ضرایب کم‌ارزش در حوزه‌ی زمان، بهبود SNR را به میزان ۲۰ dB نشان می‌دهد. هم‌چنین به ازای نهان‌نگاری در ۸ بیت کم‌ارزش ضرایب موجک سیگنال میزبان، مقدار میانگین آراء امتیازات (MOS) برای سبک‌های مختلف موسیقی محاسبه‌شده و به‌طور متوسط مقدار ۴/۸ به‌دست‌آمده است. این روش بر روی نمونه‌های ۱۰ ثانیه‌ای اعمال‌شده و به‌طور متوسط نرخ پنهان‌سازی اطلاعات تا ۲۲۰/۵kbps را فراهم می‌آورد [۱۴]. روش دیگر استفاده از مکانیزم پوشش فرکانسی برای پنهان کردن متن در سیگنال صحبت است. در این روش هر نمونه از پیام مخفی در یک بخش (مؤلفه) فرکانسی سیگنال میزبان به شرط ارضا شدن یک سطح آستانه مشخصی، مخفی می‌شود. با این کار کیفیت خوب سیگنال نهان نگاشته تضمین می‌شود زیرا تعداد بیت‌های جایگزین شده در سیگنال میزبان محدود به سطح اعوجاج می‌شود. در این روش یک تضعیف اولیه ضروری است. این تضعیف باعث می‌شود که بیشینه دامنه سیگنال مخفی از بیشینه سیگنال میزبان کم‌تر شود. بیشینه محدوده دینامیکی سیگنال پیام مخفی ۲ بیت کمتر از سیگنال میزبان است [۱۶]. روش دیگر نهان‌نگاری صوتی مبتنی بر جابجایی طیف[9] توسط اسکوپین[10] و همکارانش ارائه شد که ادغام سیگنال پیام مخفی در سیگنال پیام میزبان در حوزه فرکانس است. در این کار، طیف سیگنال مخفی در بالاترین فرکانس طیف سیگنال میزبان قرار می‌گیرد. لازم به توضیح است که شفافیت ادراکی بر این حقیقت استوار است که سیستم شنیداری انسان در فرکانس‌های بالا حساسیت کمتری دارد. بنابراین پیام مخفی می‌تواند بدون مظنون بودن به وجودش در فرکانس‌های بالا پنهان شود. در این طرح محدوده سیگنال میزبان ۱۸ KHz تا ۲۲ KHz و محدوده سیگنال صحبت مخفی ۴ KHz در نظر گرفته شده است [۲۳]. روش دیگر نهان‌نگاری صوتی مبتنی بر طیف گسترده[11] نیز توسط اسکوپین و همکارانش ارائه شد. این روش نهان‌نگاری در حوزه‌ی فرکانس، فرآیند ادغام را برای سیگنال میزبان و پیام مخفی انجام می‌دهد. ایده‌ی اصلی طرح روش طیف گسترده بر مبنای توزیع طیف پیام مخفی در طیف سیگنال میزبان است. با توجه به این‌که طیف سیگنال میزبان در محدوده‌ی ۲۰ KHz تا ۲۲ kHz است، پهنای باند آن دارای تعداد مؤلفه‌های بسیار زیادتری نسبت به طیف سیگنال پیام مخفی (سیگنال صحبت در محدوده‌ی ۴ kHz) می‌باشد، لذا تشخیص توزیع و پراکندگی طیف سیگنال صحبت در طیف سیگنال میزبان توسط شخص مهاجم بسیار دشوار است. هم‌چنین برای اطمینان از مخفی بودن اطلاعات و جلوگیری از سوءظن به وجود پیام مخفی، یک تضعیف اولیه نیز بر روی سیگنال پیام مخفی صورت می‌گیرد که در روش گفته شده این تضعیف برابر ۴۰ dB- است. با این‌کار طبق خاصیت پوشش سیستم شنیداری انسان که در قسمت قبل توضیح داده شد، مخفی ماندن اطلاعات در سیگنال میزبان تضمین می‌شود [۲۳]. روش دیگر نهان نگاری صوتی با استفاده از الگوریتم ژنتیک و تجزیه ماتریسی است. در این روش با استفاده از روش تجزیه ماتریس‌ها هر فریم سیگنال صدا تجزیه‌شده، سپس با استفاده از الگوریتم ژنتیک بهترین مکان برای جاسازی بیت نهان نگاشته که مقاومت بالایی در برابر حملات دارد جستجو می‌شود و بیت مورد نظر در محل بهینه جاسازی می‌گردد. به منظور بالا بردن مقاومت در برابر حملات مختلف، در هر فریم فقط یک بیت جاسازی می‌شود. استفاده از روش تجزیه ماتریس‌ها و الگوریتم ژنتیک باعث بالا رفتن مقاومت سیگنال نهان نگاشته و همچنین بالا رفتن سرعت اجرای الگوریتم نسبت به روش‌های معمولی می‌گردد [۱۱]. روش دیگر با استفاده از پوشش موجک موثر است. ابتدا از هر دو سیگنال میزبان و سیگنال پیام مخفی تبدیل موجک گرفته می‌شود. سپس پیشنهاد می‌شود که ضرایب موجک بر اساس یک آستانه به دو دسته با معنی (باارزش) و بدون معنی (بدون ارزش) دسته‌بندی می‌شوند زیرا ضرایب با دامنه پایین نمی‌توانند انرژی با معنی برای سیگنال فراهم بیاورند و آن‌ها را صفر در نظر می‌گیریم که این فرآیند پردازش آستانه نامیده می‌شود. ضرایب موجک پیام مخفی به گونه‌ای مرتب می‌شوند که شبیه نماینده موجک سیگنال میزبان به نظر آیند. در عمل اگر زمان غیر خاموشی سیگنال پیام مخفی کمتر یا مساوی زمان غیر خاموشی سیگنال میزبان باشد پیام

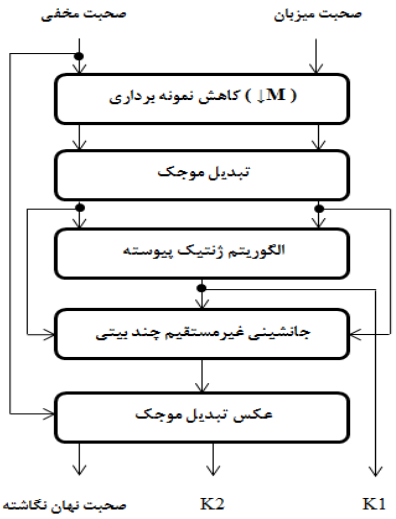
6 Signal To Noise Ratio
7 Mean Opinion Score
8 Cvejic
9 Shift Spectrum
10 Skopin

11 Spread Spectrum

مخفی مشابه پیام سیگنال میزبان اصلی به نظر می‌رسد. در این کار سیگنال پیام مخفی dB ۱۲ کمتر از سیگنال پیام میزبان است. یک کلید مخفی که تعیین‌کننده موقعیت‌های اولیه هر داده است نیز در نظر گرفته می‌شود. این روش یک نهان نگاری مناسب سیگنال صحبت ارائه می‌دهد [۱۵]. در ادامه الگوریتم پیشنهادی در دو قسمت معرفی می‌گردد که قسمت اول فرایند نهان نگاری در بخش فرستنده و قسمت دوم فرایند استخراج در بخش گیرنده را مشخص می‌کند.

فرایند نهان نگاری اطلاعات

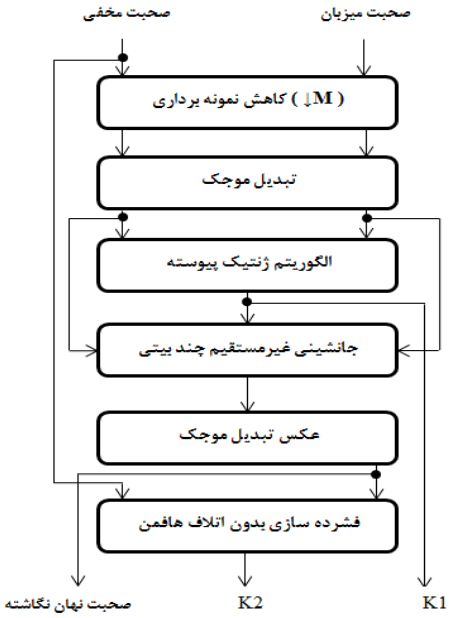
در این قسمت، دو بلوک دیاگرام پیشنهاد می‌شود که هر کدام ویژگی‌ها و رویکرد خاص خود را دارند. در شکل۱ و ۲ دو بلوک دیاگرام نهان نگاری را مشاهده می‌کنیم که به ترتیب به صورت زیر قابل توضیح هستند:



شکل۱. بلوک دیاگرام - ۱ فرآیند نهان نگاری در فرستنده

کاهش نمونه‌برداری

به دلیل آنکه نمونه‌های صحبت میزبان و صحبت مخفی دارای حجم بالایی است، از این رو برای پردازش بهتر و سریع تر، فرآیند کاهش نمونه‌برداری انجام می‌شود. در این روش برای ۲ ثانیه از نمونه‌ی سیگنال صحبت، پس از آزمون‌های مختلف، کاهش نمونه‌برداری با مقیاس ۶ پیشنهاد شده است (↓M=۶). این بخش برای هر دو بلوک دیاگرام ۱ و ۲ مشترک است. با این کار بدون آنکه کیفیت سیگنال صحبت تغییری محسوسی کند، حجم محاسبات پایین، و جستجوی فضای مورد بررسی الگوریتم ژنتیک که در ادامه توضیح داده می‌شود بهتر انجام می‌گیرد.



شکل۲. بلوک دیاگرام -۲ فرآیند نهان نگاری در فرستنده

تبدیل موجک

در این مرحله از تبدیل موجک گسسته دودویی با فیلتر هار[۱۲] استفاده شده است. دو ویژگی مهم تبدیل موجک، یکسان نبودن پهنای باند فیلترها و استفاده از کاهش نمونه‌برداری است. از این دو ویژگی در حذف نویز و همچنین فشرده‌سازی استفاده می‌شود. در تبدیل موجک دودویی در هر مرحله از انتقال اطلاعات از حوزه-ی زمان به حوزه‌ی موجک، سیگنال به دو قسمت ضرایب تقریب[۱۳] و ضرایب تفصیل[۱۴] تجزیه می‌شود. ضرایب تقریب دربردارنده‌ی مؤلفه‌های پایین سیگنال (L) و ضرایب تفصیل دربردارنده‌ی مؤلفه‌های بالای سیگنال (H) است. خصوصیت فشرده‌سازی بیانگر این است که تبدیل موجک سیگنال‌های واقعی در چند ضریب عمده پراکنده است. ما اطلاعات سیگنال صحبت مخفی را در ضرایب تقریب سیگنال میزبان قرار می‌دهیم. استفاده از پنجره‌های با طول کوچک تر در فرکانس‌های بالاتر باعث افزایش دقت زمانی در این فرکانس‌ها شده و از سوی دیگر وجود چنین ساختار پنجره‌ای با حلزونی گوش انسان نیز مطابقت بیشتری دارد. مزیت دیگر تبدیل موجک کمتر بودن حجم محاسبات آن در مقایسه با تبدیل فوریه است. در تبدیل موجک گسسته حجم محاسبات از O(n) و در تبدیل فوریه گسسته از O(nlogn) است که n تعداد نمونه‌های سیگنال است [۱]. این بخش نیز برای دو بلوک دیاگرام یکسان است.

12 Haar
13 Approximation
14 Detail

الگوریتم ژنتیک

الگوریتم ژنتیک یک روش جستجوی موثر در فضاهای بسیار بزرگ است که در نهایت منجر به جهت‌گیری به سمت پیدا کردن یک جواب بهینه می‌گردد که با استفاده از روش‌های معمول نمی‌توان به آن پاسخ مشخصی داد. الگوریتم ژنتیک یکی از انشعاب‌های الگوریتم‌های تکاملی است که در حقیقت روش جستجوی کامپیوتری بر پایه‌ی الگوریتم‌های بهینه‌سازی و بر اساس ساختار ژن‌ها و کروموزوم‌هاست [۷و۸]. اصول الگوریتم ژنتیک توسط جان هالند[15] و همکارانش در دانشگاه میشیگان در سال ۱۹۶۲ ارائه شد. در سال ۱۹۷۵ تحقیقات در این زمینه منجر به انتشار کتاب " تطابق در سیستم‌های طبیعی و مصنوعی" گردید [۲۱]. گلدبرگ[16]، این الگوریتم را به عنوان الگوریتم ژنتیک ساده معرفی کرد. این الگوریتم از لحاظ محاسباتی ساده و درعین‌حال در جستجویش قدرتمند است [۹].

استراتژی حل مسئله برای بلوک دیاگرام – ۱ بر اساس الگوریتم ژنتیک (استراتژی اول)

در این بخش استراتژی حل مسئله برای بلوک دیاگرام -۱ بر اساس تعریف الگوریتم ژنتیک روی مسئله فوق (نهان نگاری) باهدف یافتن بهترین مکان‌ها برای ادغام سیگنال صحبت مخفی در سیگنال صحبت میزبان با شبه کد زیر را توضیح می‌دهیم:

```
Begin
Chromosome = random permutation of size of cover
Set counter_i
for GA=1 to Generation do
   for i=1 to number of population do
      for j=1 to size of cover do
 Z_i(j) ← ceil (cte*cover(Chromosome_i(j))*work(j)^-1)
      end //
    if size of Z_i(j) ≤ 5bit then
    counter_i ← counter_i +1
    end if
  end //
Fitness Evaluation {
Calculate counter_i
}
Calculate bestchromosome
Selection for mating pool
Cycle crossover
 Static & Swap mutation
Elitism selection
Crowding replacement
end_GA //
K1= bestchromosome
```

منظور از cover و work به ترتیب ضرایب موجک سیگنال صحبت میزبان و مخفی است. در این مقاله، cte برابر ۲، پیشنهاد شده است. (در این حالت تضعیف dB ۶ را لحاظ کرده‌ایم.) منظور از generation، نسل (تعداد تکرار الگوریتم) است. اگر R فضای جستجوی کلی باشد که برابر با تعداد نمونه‌های سیگنال صحبت میزبان یا مخفی است (درتمامی قسمت-های این تحقیق تعداد نمونه‌های یکسانی از هر دو سیگنال انتخاب شده است.) در این صورت کروموزوم‌ها و جمعیت در الگوریتم بالا به صورت زیر تعریف می‌شوند:

$$\text{Chromosome}_j = [a_1,a_2,\ldots,a_{\text{length(R)}}] \quad , \quad j=1,2,\ldots,10 \qquad (۱)$$

$$a_k \in \{1,2,\ldots,\text{length(R)}\} \quad , \quad 1 \le k \le \text{length(R)} \qquad (۲)$$

$$\text{population} = \{\text{Chromosome}_1,\ldots,\text{Chromosome}_{10}\} \qquad (۳)$$

برای هر جمعیت j تایی، k بار مرحله‌ی زیر انجام می‌شود:

$$Z_j(k) = ceil\left(\frac{cte*cover(chromosomej(k))}{work(k)}\right) \qquad (۴)$$

منظور از $ceil$ در اینجا کوچک‌ترین عدد صحیح بزرگ‌تر یا مساوی یک عدد اعشاری است. برای جلوگیری از افت کیفیت سیگنال نهان نگاشته در این قسمت یک کران پایین d_{min} و یک کران بالای d_{max} در نظر گرفته می‌شود. برای هر j، یک شمارنده با مقداردهی اولیه صفر تخصیص داده و به ازای k هایی که Z های به دست آمده در این محدوده قرار بگیرند، شمارنده‌ی متناظرش یک واحد به صورت زیر افزایش می‌یابد:

$$\text{If } Z_j(k) \in \{d_{min},\ldots,d_{max}\} \rightarrow \text{counter}_j = \text{counter}_j+1 \qquad (۵)$$

در این روش d_{min} برابر با ۲ و d_{max} برابر با ۳۱ پیشنهاد شده است. حال تابع شایستگی (برازندگی) به صورت زیر تعریف می‌شود:

$$\text{Fitness function} = \text{calculate } \{ \text{counter}_j \} \text{ , } j=1,2,\ldots,10 \qquad (۶)$$

در این روش بهترین کروموزوم از یک نسل به نسل بعدی به طور مستقیم انتقال داده شده است (نخبه‌گرایی[17]). بر اساس تحقیقات انجام‌گرفته توسط هاپت[18] بهتر است که اندازه‌ی جمعیت برای الگوریتم‌های ژنتیک پیوسته کمتر، از ۱۶ باشد [۲۰]. از این رو در هر دو استراتژی اندازه‌ی جمعیت ۱۰ در نظر گرفته شده است. از آنجایی که روش‌های انتخاب متناسب با برازندگی و از جمله چرخ رولت دارای فشار انتخاب[19] بالا است، کاربرد مستقیم آن‌ها باعث فقدان تنوع ژنی و در نتیجه همگرایی زودرس می‌شود، به همین دلیل ما از روش‌های مقیاس کردن شایستگی[20] برای تنظیم فشار انتخاب استفاده می‌کنیم. در این مقاله از انتخاب بولتزمن به صورت زیر استفاده شده است:

$$P_i = \frac{Fitness\ value}{\sum_{i=1}^{10}(Fitness\ value_i)} \qquad (۷)$$

15 John Holland
16 Goldberg

17 Elitism
18 Haupt
19 Selection Pressure
20 Fitness Scaling

$$Q_i = exp\left(\frac{(P_i-1)*generation}{2*max\{generation\}}\right) \quad (۸)$$

پس از محاسبه‌ی مقادیر Q_i، از روش چرخ رولت، کروموزوم‌های والد را انتخاب می‌کنیم. با ترکیب روش بولتزمن و چرخ رولت در حقیقت در ابتدای اجرای الگوریتم ژنتیک، شایستگی مقیاس شده-ی افراد ضعیف و قوی به یکدیگر نزدیک بوده و بدین وسیله فشار انتخاب پایین نگه داشته می‌شود. در ادامه‌ی الگوریتم، فاصله‌ی شایستگی افراد قوی و ضعیف زیادتر شده و فشار انتخاب رفته‌رفته افزایش می‌یابد. در این بلوک دیاگرام جهش به صورت استاتیکی و از نوع جابه‌جایی[۲۱] تعریف می‌شود. در این نوع جهش، دو نقطه از یک کروموزوم به طور اتفاقی انتخاب شده و موقعیت آن دو تغییر می‌کند. برای ادغام کروموزوم‌ها از روش ادغام چرخشی گلدبرگ (cx)[۲۲] استفاده شده و به صورت تک نقطه‌ای است [۵]. این نوع ادغام با این محدودیت انجام می‌شود که مقدار هر آلل[۲۳] (مقدار یک ژن) از یکی از دو والدین می‌آید. این ادغام به محل برش وابسته نیست و یک چرخه مشابه از جایگشت‌های جبری را برای فرزندان طی می‌کند.

استراتژی حل مسئله برای بلوک دیاگرام – ۲ بر اساس الگوریتم ژنتیک (استراتژی دوم)

این بخش شباهت زیادی با بخش قبلی دارد با این تفاوت که قسمت‌هایی مانند بخش‌بندی فضای جستجو، نوع جهش و قسمت ممانعت ازدواج با محارم در آن اعمال شده است. استراتژی حل مسئله بر اساس تعریف الگوریتم ژنتیک روی مسئله فوق برای بلوک دیاگرام -۲ را با شبه کد زیر توضیح می‌دهیم:

```
Begin
Chromosome = random permutation of size of cover
Set  counter_i
R=total search space
R_1, R_1, ..., R_n ←R  // divide into subspace
for  GA=1 to  Generation  do
   for  i=1 to  number of population  do
      for  j=1 to  size of  cover  do
 Z_i(j) ← ceil (cte*cover(Chromosome_i(j))*work(j)^-1)
      end //
    if size of Z_i(j) ≤ 5bit  then
    counter_i ← counter_i +1
    end if
  end //
Fitness Evaluation  {
Calculate  counter_i
}
Calculate  bestchromosome
Selection for mating pool
Cycle crossover
Dynamic & sawp  mutation
Elitism selection
Crowding replacement
Incest prevention                                    end_GA
//
K1= bestchromosome
```

ابتدا فضای جستجو به چند زیر فضا تقسیم می‌شود (در این تحقیق پس از بررسی‌های مختلف، فضای جستجو به ۸ زیر فضا تقسیم شده است.) با فرض اینکه فضای کل R باشد، داریم:

$$R=R_1 U R_2 U \ldots U R_8 \quad (۹)$$

برای هر R_i که $i=1,2,\ldots,8$ کروموزوم‌ها و جمعیت به صورت زیر تعریف می‌شوند:

$$Chromosome_j = [a_1,a_2,\ldots,a_{length(Ri)}] \quad , \quad j=1,2,\ldots,10 \quad (۱۰)$$

$$a_k \in \{1,2,\ldots,length(R_i)\} \quad , \quad 1\leq k \leq length(Ri) \quad (۱۱)$$

تشکیل جمعیت، نحوه‌ی به دست آمدن Z ها و محاسبه شایستگی کروموزوم‌ها مشابه استراتژی بکار رفته در بلوک دیاگرام -۱ است. تفاوت این بخش در این است که ما در حقیقت i تا جمعیت j تایی که طولشان length (R_i) است، را در اختیار داریم. بهترین - کروموزوم‌ها از جمعیت فوق را از نسل‌فعلی به نسل بعدی به طور مستقیم انتقال می‌دهیم. البته نبایستی کروموزوم‌های نخبه زیادی را انتقال مستقیم داد، زیرا باعث ایجاد یک گونه[۲۴] غالب و در نتیجه به همگرایی زودرس[۲۵] منجر می‌شود. به همین دلیل در روش پیشنهادی الگوریتم مربوط به بلوک دیاگرام -۲ ما فقط بهترین کروموزوم را انتقال مستقیم می‌دهیم. کروموزوم‌های ضعیف به طور قطعی کنار گذاشته نمی‌شوند زیرا تمام خصوصیات کروموزوم‌های ضعیف، ضعیف نیست و حذف آن‌ها می‌تواند باعث کاهش غنای ژنی در نسل‌های آینده بشود [۶]. به همین دلیل ضعیف‌ترین کروموزوم نیز در هر نسل حفظ شده است. در اینجا نیز از روش‌های مقیاس کردن شایستگی برای تنظیم فشار انتخاب استفاده می‌کنیم. روش مورد استفاده در بلوک دیگرام -۲ نیز انتخاب بولتزمن و مشابه توضیحات گفته‌شده در استراتژی قبل است. برای ادغام کروموزوم‌ها نیز از روش ادغام چرخشی گلدبرگ و به صورت تک نقطه‌ای استفاده شده است. از آنجایی که در مسائل جایگشتی نرخ جهش‌های بزرگ مناسب نیست [۲۵]، به همین دلیل دراستراتژی دوم نرخ جهش کم و به صورت دینامیکی با رابطه‌ی زیر در نظر گرفته می‌شود:

$$P_m(t) = \left(2 + \frac{(length(Ri)-2)*t}{(generation-1)}\right)^{-1} \quad (۱۲)$$

که در رابطه‌ی بالا t شمارنده نسل فعلی است. این جهش نیز از نوع جابه‌جایی تعریف می‌شود.

21 Swap
22 Cycle Crossover
23 Allele

24 Specie
25 Premature Convergence

جانشینی غیرمستقیم چند بیتی

بعد از انتخاب بهترین کروموزوم که همان کلید اصلی k_1 (کلید اول) است برای هر دو بلوک دیاگرام، Z_{final} به صورت زیر محاسبه می‌شود:

$$Z_{final}(k) = ceil\left(\frac{cte*cover(K_1(k))}{work(k)}\right) \qquad (۱۳)$$

Z_{final} به دست آمده در حقیقت نماینده‌ی غیرمستقیم ضرایب موجک سیگنال صحبت مخفی است. این ضرایب می‌تواند حداکثر در ۵ بیت کم ارزش سیگنال صحبت میزبان جاسازی شود.

عکس تبدیل موجک

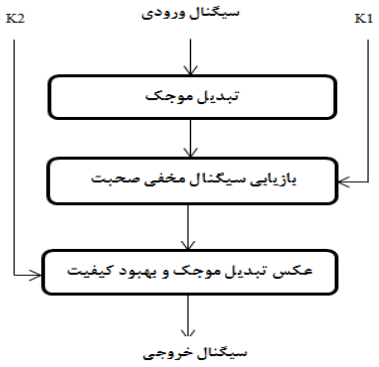
در این بخش از سیگنال نهان نگاشته، عکس تبدیل موجک گرفته و سیگنال را به حوزه‌ی زمان برمی‌گردانیم. در بلوک دیاگرام -۱ برای افزایش کیفیت سیگنال صحبت مخفی در قسمت گیرنده، اختلاف بین سیگنال صحبت مخفی قبل و بعد از نهان نگاری به عنوان کلید فرعی (k_2) همراه با سیگنال صوتی نهان نگاشته به گیرنده ارسال می‌شود. اما در بلوک دیاگرام -۲ این اختلاف به بخش بعد که فشرده‌سازی بدون اتلاف است، منتقل می‌شود.

فشرده‌سازی بدون اتلاف هافمن

به دلیل کوانتیزه کردن، بین سیگنال اصلی مخفی و سیگنال استخراج‌شده از سیگنال نهان نگاشته، اختلاف وجود دارد که سبب می‌شود کیفیت سیگنال مخفی در گیرنده پس از استخراج پایین بیاید. بررسی‌ها نشان داده این اختلاف را می‌توان با یک نویز گوسی کانال مدل کرد. اما این نویز را با فشرده‌سازی بدون اتلاف هافمن می‌توان فشرده کرد. کلید k_2 همان اختلاف فشرده شده است. این بخش را فقط در بلوک دیاگرام-۲ داریم.

فرایند استخراج اطلاعات

این بخش برای هر دو بلوک دیاگرام یکسان است و شامل سه مرحله‌ی تبدیل موجک، بازیابی سیگنال مخفی صحبت و عکس تبدیل موجک و بهبود کیفیت صحبت مخفی با اضافه کردن کلید دوم است. شکل۳ مراحل مختلف استخراج اطلاعات را نشان داده است که به صورت زیر قابل توضیح است (منظور از سیگنال ورودی همان سیگنال صحبت نهان نگاشته که بسیار شبیه سیگنال صحبت میزبان است و منظور از سیگنال خروجی، سیگنال صحبت مخفی است که هدف بازیابی بدون اتلاف یا کم اتلاف آن است) :



شکل۳. بلوک دیاگرام فرآیند استخراج در گیرنده

تبدیل موجک

همانند قسمت فرستنده در اینجا نیز از سیگنال ورودی تبدیل موجک با فیلتر هار می‌گیریم و سیگنال را به حوزه‌ی موجک انتقال می‌دهیم. بقیه‌ی توضیحات مشابه بخش متناظر در فرستنده است.

بازیابی سیگنال مخفی صحبت

در این بخش کلید k_1 را روی ضرایب موجک سیگنال به دست آمده از قسمت قبل اعمال کرده و سیگنال صحبت مخفی $G_2(w)$ را که در حوزه‌ی موجک است، به صورت زیر به دست می‌آوریم:

$$G_2(w) = ceil\left(\frac{cte*G_1((K_1(w))}{pe(w)}\right) \qquad (۱۴)$$

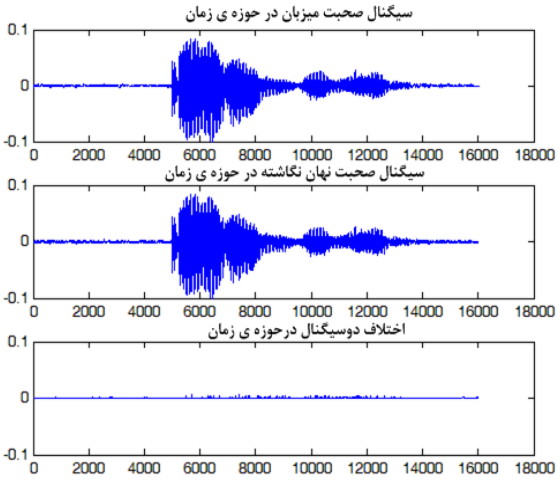
که در رابطه‌ی فوق $G_1(w)$ سیگنال نهان نگاشته در حوزه‌ی موجک و pe کم ارزش‌ترین بیت ضرایب در حوزه‌ی موجک است که در کار پیشنهادی حداکثر ۵ بیت کم ارزش است.

عکس تبدیل موجک و بهبود کیفیت

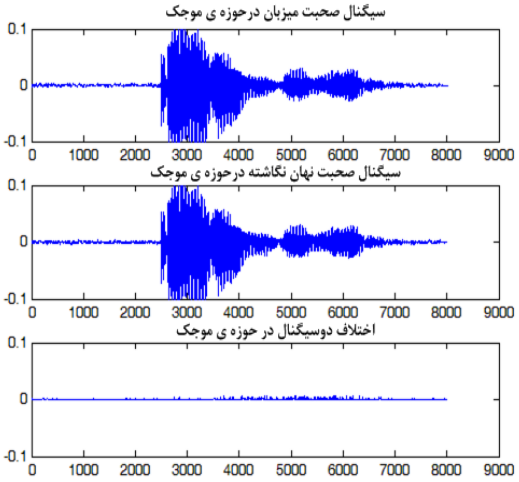
در این بخش از سیگنال $G_2(w)$، عکس تبدیل موجک گرفته‌شده و سیگنال به حوزه‌ی زمان برگردانده می‌شود. سپس کلید k_2 برای افزایش کیفیت سیگنال صحبت مخفی به آن اضافه می‌شود. این قسمت یک مرحله تکمیلی و فقط برای بهبود هرچه بیشتر صدا در بخش گیرنده انجام می‌شود.

آزمایش‌های تجربی

در این مقاله نمونه‌های سیگنال صحبت میزبان و مخفی هر کدام ۲ ثانیه، فرکانس نمونه‌برداری ۴۴/۱ KHz و برای هر نمونه ۱۶ بیت اختصاص داده شده است. در شکل ۴ و شکل ۵ سیگنال صحبت میزبان و نهان نگاشته و اختلاف آن‌ها در دو حوزه‌ی زمان و موجک نشان داده شده است. (به دلیل شباهت بسیار زیاد، شکل‌ها فقط برای بلوک دیاگرام -۲ رسم شده است.)

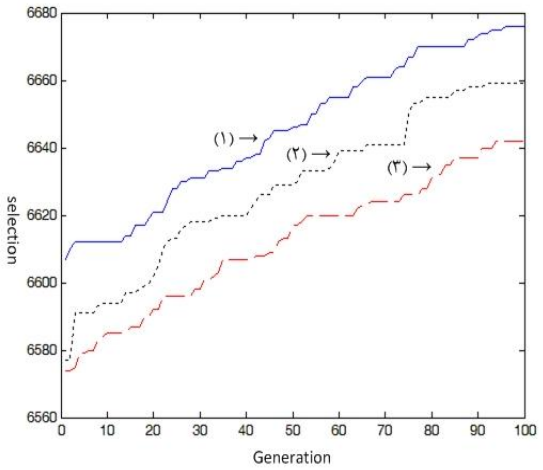


شکل۴. نمایش سیگنال میزبان و نهان نگاشته و اختلاف آن‌ها در حوزه‌ی زمان



شکل۵. نمایش سیگنال میزبان و نهان نگاشته و اختلاف آن‌ها در حوزه‌ی موجک

همان طور که در قسمت استراتژی‌ها بحث شد برای تنظیم شدت انتخاب[۲۶] روش چرخ رولت و انتخاب بولتزمن را به صورت ترکیبی استفاده می‌کنیم. شکل ۶ میزان تأثیر جایگزینی ازدحامی، جایگزینی نسلی[۲۷] و ترکیب چرخ رولت و انتخاب بولتزمن را نشان می‌دهد.



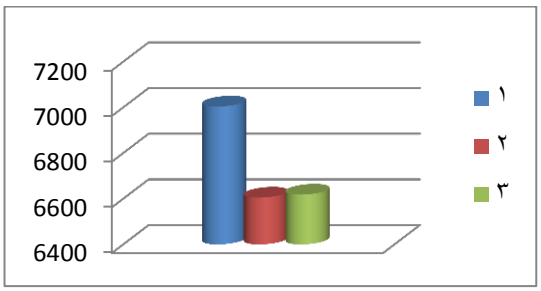
شکل۶. ۱- با استفاده از جایگزینی ازدحامی و با انتخاب ترکیبی چرخ رولت و بولتزمن ۲- بدون جایگزینی ازدحامی و با انتخاب ترکیبی چرخ رولت و بولتزمن ۳- با جایگزینی نسلی و بدون انتخاب ترکیبی چرخ رولت و بولتزمن

همان طور که از شکل ۶ پیداست، در حالت -۱ به دلیل اینکه فشار انتخاب کنترل شده است، افراد با شایستگی پایین در نسل‌های اولیه از شانس متعادلی برای تولیدمثل و مشارکت در نسل بعد برخوردارند و همچنین به دلیل استفاده از جایگزینی ازدحامی، علاوه بر شایستگی افراد، ملاحظات تنوع جمعیت نیز مدنظر قرار می‌گیرد. در حالت-۲ تنوع ژنی پایین است و امکان همگرایی زودرس وجود دارد. در حالت-۳ به دلیل عدم استفاده از انتخاب بولتزمن، فشار انتخاب کنترل شده نیست و از طرفی به دلیل استفاده جایگزینی نسلی و از بین رفتن همه کروموزوم‌های نسل والد، تنوع ژنی کاملاً تصادفی است. در این تحقیق ترکیبی از جایگزینی نخبه گرا و جایگزینی ازدحامی[۲۸] به همراه روش جلوگیری از ازدواج با محارم[۲۹] استفاده شده است. مادامی که یک جمعیت تکامل می‌یابد، اعضای آن به یکدیگر شبیه‌تر می‌شوند. در بحث جلوگیری از ازدواج با محارم سعی می‌شود والدهایی که با یکدیگر تفاوت (فاصله) زیادی دارند با یکدیگر جفت‌گیری انجام داده تا تکامل بیشتری حاصل شود. در این روش فاصله‌ی اقلیدسی افراد با بهترین فرد محاسبه شده و چنانچه از سطح آستانه‌ی تفاوت مشخصی بیشتر باشد آن را در فهرست والدین قرار می‌دهیم [۱۷]. در روش پیشنهادی سطح آستانه ۲۰۰۰۰ در نظر گرفته شده است. چنانچه در پایان هر ارزیابی تغییری در جمعیت والدین ایجاد نشود، سطح آستانه‌ی تفاوت ۱۰۰۰ واحد کم می‌شود. همچنین برای حفظ تنوع جمعیت، از روش جایگزینی ازدحامی استفاده می‌شود. در این کار ضریب ازدحام (CF)[۳۰] ۳ در نظر گرفته‌شده و به این تعداد از اعضای جمعیت به صورت تصادفی، انتخاب و جایگزین عضو با بالاترین میزان شباهت می‌شود. نتایج به

26 Selection Intensity
27 Generational Replacement

28 Crowding Replacement
29 Incest Prevention
30 Crowding Factor

دست آمده از شکل ۷ نشان می‌دهد که جلوگیری از ازدواج با محارم به تنهایی نمی‌تواند شایستگی فرزندان به دست آمده را ارتقاء دهد ولی ترکیب آن با جایگزینی ازدحامی نتیجه‌ی مطلوب تری می‌دهد.



شکل۷. ۱- با استفاده از جلوگیری از ازدواج با محارم و جایگزینی ازدحامی ۲- با استفاده از جلوگیری از ازدواج با محارم و بدون جایگزینی ازدحامی ۳- بدون استفاده از جلوگیری از ازدواج با محارم و بدون جایگزینی ازدحامی (اعداد به دست آمده در جدول نشان‌دهنده‌ی تعداد مکان‌های مناسب برای ادغام است)

در این مقاله دو تحلیل نهان کاوی که یکی تحلیل زمانی و دیگری تحلیل فرکانسی است، به کار می‌رود. و در تمامی قسمت‌های این نهان کاوی، ۴ رابطه زیر محاسبه می‌شود:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \quad (۱۵)$$

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{(\mu-1)} \quad (۱۶)$$

$$sk = \frac{\sum_{i=1}^{N}(x_i-\mu)^3}{(N-1)\sigma^3} \quad (۱۷)$$

$$k = \frac{\sum_{i=1}^{N}(x_i-\mu)^4}{(N-1)\sigma^4} \quad (۱۸)$$

که در روابط بالا x_i سیگنال ورودی، N تعداد نمونه‌ها، μ میانگین[31]، σ^2 واریانس[32]، sk چولگی[33] و k کشیدگی[34] است. ما تحلیل را در دو حوزه‌ی زمان و فرکانس و با توجه به روابط زیر ارائه می‌دهیم [۱۹و۲۶]:

$$X[n]=\log_{10}(|v[n]|+1) \quad (۱۹)$$
$$X_a[n]=[v[2]-v[1] \dots v[n]-v[n-1]]$$

$$X_b[n]=[X_a[2]-X_a[1] \dots X[n]_a-X_a[n-1]]$$

$$X(f)=abs(fft\{ X_b[n]\}) \quad (۲۰)$$

که در روابط بالا $v[n]$ سیگنال صوت ورودی است. نتایج روش پیشنهادی با روش کم ارزش‌ترین بیت [۱۴]، پوشش فرکانسی [۲۳] و پوشش موجک موثر [۱۵] مقایسه شده و در جدول‌های ۱ و ۲ آمده است. اعداد به دست آمده به صورت درصد اختلاف بین سیگنال میزبان و سیگنال نهان نگاشته برای چهار پارامتر ذکر شده است. در جدول زیر اعداد به دست آمده هر چقدر کوچک تر باشند، نشان‌دهنده‌ی آن است که سیگنال نهان نگاشته شباهت بیشتری به سیگنال میزبان دارد و فرایند ادغام در قسمت فرستنده بهتر صورت گرفته است.

جدول۱. مقادیر به دست آمده در حوزه‌ی زمان

روش‌ها	میانگین (%)	واریانس (%)	چولگی (%)	کشیدگی (%)
کم ارزش‌ترین بیت	۱/۱۱۰	۰/۵۴۳	۰/۳۷۷	۰/۳۵۲
پوشش فرکانسی	۴/۲۰۰	۲/۳۷۰	۶/۵۹۰	۲/۷۹۰
پوشش موجک موثر	۰/۱۰۲	۰/۲۲۶	۰/۱۹۲	۰/۰۸۶
روش پیشنهادی ۱	۰/۰۲۷	۰/۰۱۳۶	۰/۱۳۰۰	۰/۱۹۳۳
روش پیشنهادی ۲	۰/۰۱۸	۰/۰۹۴	۰/۰۱۶	۰/۰۰۸

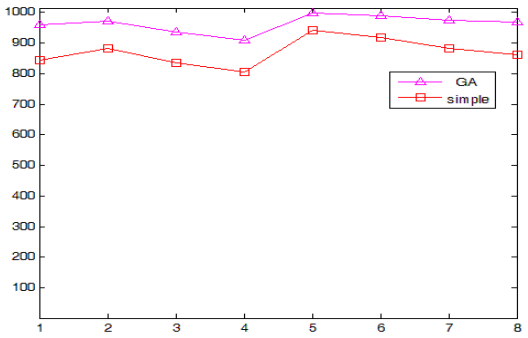
جدول۲. مقادیر به دست آمده در حوزه‌ی فرکانس

روش‌ها	میانگین (%)	واریانس (%)	چولگی (%)	کشیدگی (%)
کم ارزش‌ترین بیت	۲/۵۸۰	۲/۸۷۰	۹/۴۴۰	۵/۶۹۰
پوشش فرکانسی	۱۲/۶۴	۳/۳۴۰	۱/۵۱۰	۲/۶۰۰
پوشش موجک موثر	۰/۱۵۰۰	۱/۰۴۰۰	۴/۳۵۴۹	۰/۹۳۶
روش پیشنهادی ۱	۰/۶۸۶۲	۰/۶۲۶۵	۳/۳۹۳۱	۳/۹۷۹۹
روش پیشنهادی ۲	۰/۰۱۷	۰/۰۱۸	۲/۰۶۰	۲/۹۹۰

مقایسه‌ی روش پیشنهادی برای دو بلوک دیاگرام گفته شده نشان می‌دهد که اعداد به دست آمده از بلوک دیاگرام- ۲ بهتر (کوچک-تر) است. در حقیقت بلوک دیاگرام -۱ دارای سادگی بهتر و پیچیدگی کمتر و زمان پردازش کمتر و بلوک دیاگرام- ۲ دارای دقت بهتر و پیچیدگی بیشتر است. اولویت‌بخشی به سرعت یا دقت عامل اصلی در انتخاب یکی از این دو بلوک دیاگرام است. همچنین تحلیل مقادیر به دست آمده روی بلوک دیاگرام -۲ نشان می‌دهد که اعداد به دست آمده در روش پیشنهادی به دلیل استفاده از الگوریتم ژنتیک جهت یافتن بهترین مکان‌ها برای نهان نگاری، آماره‌های مرتبه‌ی اول تا چهارم آن در حوزه‌ی زمان بسیار پایین تر از سه روش بالا است. در حوزه‌ی فرکانس نیز آماره‌های مرتبه‌ی اول و دوم بسیار پایین تر از سه روش گفته شده است و تنها مقادیر چولگی و کشیدگی با سه روش بالا قابل‌مقایسه و تنها مقدار کمی بیشتر است. شکل ۸ تأثیر الگوریتم ژنتیک را در انتخاب تعداد مکان‌های مناسب نشان می‌دهد. در این شکل تعداد انتخاب‌ها برای ادغام قبل از استفاده از الگوریتم ژنتیک و بعد از استفاده از الگوریتم ژنتیک را در هر ۸ زیر بخش برای اجرای الگوریتم پس از طی ۵۰۰ نسل نشان می‌دهد. همان طور که از

31 Mean
32 Variance
33 Skewness
34 Kurtosis

شکل پیداست بیش‌ترین افزایش ظرفیت مربوط به زیر بخش اول با ۱۱/۵٪ و کمترین افزایش ظرفیت مربوط به زیر بخش پنجم با ۵/۵٪ است. به طور متوسط می‌توان گفت با بکارگیری الگوریتم ژنتیک حدود ۱۰ ٪ ظرفیت انتخاب نقاط ادغام مناسب برای نهان نگاری افزایش می‌یابد.



شکل۸. تعداد انتخاب‌ها قبل و بعد از الگوریتم ژنتیک

همچنین مقادیر سیگنال به نویز (SNR_{dB}) و معیار شباهت ساختاری (SSIM)[35] برای دو بلوک دیاگرام با رابطه‌های زیر محاسبه شده اند:

$$SNR_{dB} = 10 \log \frac{\sum_n x^2(n)}{\sum_n (x(n)-y(n))^2} \qquad (۲۱)$$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \qquad (۲۲)$$

که در روابط بالا x سیگنال میزبان و y سیگنال نهان نگاشته است. مقادیر μ_x، μ_y، σ^2_x، σ^2_y و σ_{xy} به ترتیب میانگین سیگنال میزبان، میانگین سیگنال نهان نگاشته، واریانس سیگنال میزبان، واریانس سیگنال نهان نگاشته و کوواریانس بین سیگنال میزبان و نهان نگاشته است. ضرایب C_1 و C_2 جهت پایدارسازی روش در هنگام کوچک بودن میانگین و واریانس استفاده می‌شود. در این مقاله C_1=0.001 و C_2=0.003 در نظر گرفته شده است. از نظر معیار SSIM هر چه دو سیگنال x و y شباهت بیشتری به داشته باشند، مقدار SSIM به یک نزدیک‌تر است. مقادیر محاسبه شده از رابطه-های بالا در جدول۳ نشان داده شده است.

جدول۳. مقادیر سیگنال به نویز و معیار شباهت ساختاری

	بلوک دیاگرام-۱	بلوک دیاگرام-۲
سیگنال به نویز(dB)	۴۹/۳۴۲۸	۴۹/۶۷۶۷
شباهت ساختاری(x,y)	۰/۸۴۷۷	۰/۸۶۰۱

نتایج جدول۳ نشان می‌دهد که نسبت سیگنال به نویز هر دو روش پیشنهادی قابل‌قبول و نزدیک به هم است. ولی معیار شباهت ساختاری برای بلوک دیاگرام-۲ به یک نزدیک تر است که نشان‌دهنده‌ی شباهت بیشتر بین سیگنال میزبان و سیگنال نهان نگاشته است.

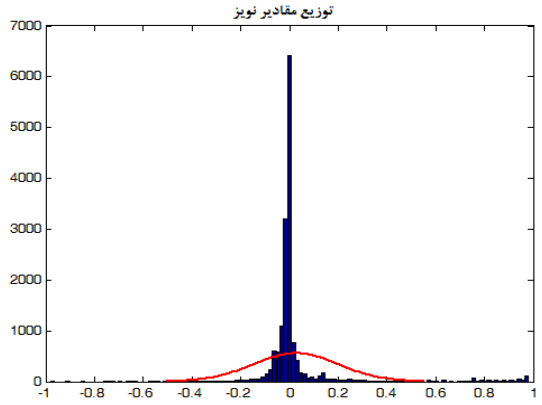
در ادامه در جدول ۴ مقادیر ظرفیت روش پیشنهادی با روش‌های گفته شده مورد مقایسه قرار گرفته است.

جدول ۴. مقایسه مقادیر ظرفیت

ظرفیت ادغام (bit)	روش‌های به‌کاررفته
۵	روش پیشنهادی
۵	روش پوشش موجک مؤثر [۱۵]
۵	روش کم‌ارزش‌ترین‌بیت [۱۴]
۴	روش موجک مترقی [۲]
۳	روش پوشش فرکانسی [۲۳]

همانطور که از مقادیر جدول ۴ مشاهده می شود ظرفیت اغام در مقایسه با دو روش اول (روش موجک موثر و روش کم ارزش‌ترین بیت) افزایش نکرده ولی مقادیر ممان‌های اول تا چهارم در حوزه زمان و فرکانس بهبود یافته است. همچنین در مقایسه با دو روش بعدی (روش موجک مترقی و روش پوشش فرکانسی) ظرفیت ادغام بهبود یافته است. نتیجه آنکه بکار بردن روش پیشنهادی یا باعث افزایش ظرفیت شده و یا اینکه با بهبود مقادیر ممان‌ها باعث افزایش شفافیت می‌شود.

همچنین بررسی‌ها نشان می‌دهد که اختلاف بین سیگنال صحبت مخفی قبل و بعد از استخراج را می‌توان با یک نویز کانال مدل کرد. چنانچه توزیع این مقادیر را در بازه‌ی [-1,1] رسم شود، متوجه می‌شویم که از توزیع گوسی تبعیت می‌کند. در شکل ۹ این توزیع رسم شده است.



شکل۹. توزیع گوسی مقادیر نویز مدل شده‌ی کانال

35 Structural SIMilarity Index

طبق قضیه‌ی اول شانون (قضیه کد گذاری بدون اتلاف منبع)، آنتروپی منبع، حد ضروری بر روی تعداد بیت‌های مورد نیاز برای نمایش منبع باهدف بازیابی است [۳]، ما از کد گذاری هافمن برای فشرده‌سازی بدون اتلاف مقادیر نویز مدل شده‌ی کانال استفاده می‌کنیم. این نوع کد گذاری، برای هر سمبل، تعداد بیتی متناسب با عکس احتمال رخداد آن در نظر می‌گیرد. محاسبات نشان می‌دهد که این نرخ فشرده‌سازی به حد شانون از قضیه‌ی - اول که همان آنتروپی منبع است، بسیار نزدیک است [۴]. نتایج در جدول ۵ آمده است. که در روابط زیر آنتروپی منبع کد نشده (H)، طول متوسط کد هافمن ($\bar{n}$)، اضافات منبع بدون کد بندی (ρ)، اضافات منبع با کد بندی هافمن (ρ')، بازده‌ی منبع بدون کد بندی (e)، بازده‌ی منبع با کد بندی هافمن (e') از روابط زیر محاسبه می‌شوند (s تعداد سمبل‌های منبع و p_i احتمال رخداد آن‌هاست):

$$H=-\sum_{i=1}^{s} p_i log_2 p_i \quad (۲۳)$$

$$\bar{n}=\sum_{i=1}^{s} n_i p_i \quad (۲۴)$$

$$e=\frac{H}{log_2^s} \quad , \quad \rho=1-e \quad (۲۵)$$

$$e'=\frac{H}{\bar{n}} \quad , \quad \rho'=1-e' \quad (۲۶)$$

جدول ۵. نتایج به دست آمده از کد بندی هافمن

مقادیر به دست آمده	مشخصات کمی ارزیابی
۹/۳۱۳۱	آنتروپی منبع کد نشده
۹/۳۳۸۱	طول متوسط کد هافمن
۱۵٪	اضافات منبع بدون کد بندی
۰/۲۷٪	اضافات منبع با کد بندی هافمن
۸۵٪	بازده‌ی منبع بدون کد بندی
۹۹/۷۳٪	بازده‌ی منبع با کد بندی هافمن

بررسی نتایج جدول، عملکرد موفق کد بندی هافمن در فشرده‌سازی داده با راندمان بسیار بالا را نشان می‌دهد. در حقیقت با این نوع فشرده‌سازی به جای نرخ ارسال ۱۶ بیت بر سمبل، از نرخ ارسال ۹/۳ بیت بر سمبل برای ارسال سیگنال نویز استفاده می‌شود.

نتیجه‌گیری

در این مقاله یک روش نهان نگاری صوتی در حوزه‌ی موجک با دو رویکرد ارائه شد. به دلیل استفاده از الگوریتم ژنتیک و طراحی مناسب پارامترهای آن، مکان‌های مناسبی برای نهان نگاری انتخاب شد. آمارگان مرتبه‌های بالاتر در مقایسه با سه روش جانشینی مستقیم LSB، پوشش فرکانسی و پوشش موجک موثر در حد قابل قبولی، در حوزه‌ی زمان و فرکانس بهبود یافت. مزیت روش فوق عدم وابستگی به سیگنال میزبان است و به دلیل بکار بردن روش جانشینی غیرمستقیم LSB، شباهت بین سیگنال صحبت نهان نگاشته و سیگنال صحبت میزبان بالا می‌رود. ویژگی بلوک دیاگرام -۱ سادگی بهتر و سرعت بیشتر و بلوک دیاگرام -۲ پیچیدگی بیشتر ولی دقت بالاتر است. به دلیل خطای کوانتیزه کردن، بین سیگنال صحبت مخفی قبل و بعد از استخراج اختلافی وجود دارد ولی این اختلاف از توزیع گوسی پیروی می‌کند. فشرده‌سازی بدون اتلاف هافمن این مقدار، به حد شانون از قضیه- ی اول آن که همان آنتروپی است بسیار نزدیک استب. این مقدار فشرده شده را همراه سیگنال اصلی ارسال می‌کنیم. مزیت دیگر این روش این است که با وجود اینکه ما حداکثر در ۵ بیت نهان نگاری را انجام داده‌ایم ولی الگوی مشخصی در تعداد بیت انتخابی نداریم که این می‌تواند یک الگوی تصادفی برای نهان نگاری باشد. این روش از روش نهان نگاری است که برخط نمی‌باشد و امنیت آن فقط به امنیت کلید وابسته است که با اصول پیشنهادی کرکهف سازگار است. ما در این کار از الگوریتم ژنتیک استفاده کردیم ولی به عنوان یک پیشنهاد می‌توان دیگر الگوریتم‌های فرامکاشفه‌ای نظیر شبیه‌سازی حرارتی یا الگوریتم کلونی مورچه‌ها را مورد بررسی و تحلیل قرارداد. روش پیشنهادی ارائه شده می‌تواند در انواع روش‌های نهان نگاری سیگنال صحبت در سیگنال صحبت (speech in speech) مورد استفاده قرار گیرد.

مرجع‌ها

[۱] آیت، سعید، مبانی پردازش سیگنال گفتار، انتشارات دانشگاه پیام نور، ۱۳۸۷

[۲] اسلامی، امید، "افزایش ظرفیت استگانوگرافی در فایل‌های صوتی با استفاده از روش ذخیره‌ی غیریکنواخت در ضرایب موجک مترقی"، هفتمین کنفرانس بین‌المللی انجمن رمز، دانشگاه صنعتی خواجه‌نصیرالدین طوسی، شهریور ۱۳۸۹.

[۳] پروکیس، جان‌جی، مخابرات دیجیتال، ترجمه‌ی کلانتری، محمد اسماعیل، انتشارات فدک ایستاتیس، ۱۳۸۹

[۴] تیموری، مهدی، شبیه‌سازی و تحلیل سیستم‌های مخابراتی دیجیتال با استفاده از Matlab، انتشارات کیان رایانه سبز، ۱۳۹۰

[۵] شمس جاوی، محمد، پیاده‌سازی و حل مسائل کاربردی با الگوریتم ژنتیک، انتشارات فرا هوش، ۱۳۹۰

[۶] نظام‌آبادی پور، حسین، الگوریتم وراثتی مفاهیم پایه و مباحث پیشرفته، انتشارات دانشگاه شهید باهنر کرمان، ۱۳۸۹

[۷] شهریار شاه حسینی، هادی، الگوریتم‌های تکاملی، مبانی، کاربردها، پیاده‌سازی، انتشارات دانشگاه علم و صنعت ایران، ۱۳۹۱

[۸] عالم تبریز، اکبر، الگوریتم‌های فرا ابتکاری در بهینه‌سازی ترکیبی، انتشارات صفار، ۱۳۹۰
[۹] فروزان، محمدرضا، روش‌های نوین بهینه‌سازی، ۱۳۹۰
[۱۰] دلفروزی، احمد، "ارائه یک روش پنهان نگاری اطلاعات در صوت دیجیتالی مبتنی بر تبدیل موجک مترقی"، سیزدهمین کنفرانس ملی انجمن کامپیوتر ایران، دانشگاه صنعتی شریف، ۱۳۸۶
[۱۱] محسن فر، سید محمدرضا، "نهان نگاری صوتی مقاوم با استفاده از الگوریتم ژنتیک و تجزیهQR"، یازدهمین کنفرانس سیستم‌های هوشمند، دانشگاه خوارزمی، ۱۳۹۱
[۱۲] مهدوی جعفری، سمیه ، "یک الگوریتم نهان‌نگاری صوتی مبتنی براساس خوشه‌بندی نمونه‌ها"، هفتمین کنفرانس بین‌المللی انجمن رمز، دانشگاه صنعتی خواجه نصیرالدین طوسی، شهریور ۱۳۸۹.
[13] A.N., Lemma, et. al., "A Temporal Domain Audio Watermarking Technique", IEEE Trans. Signal Proc.,Vol. 51, No. 4, April 2003.
[14] Cvejic, Nedeljko, and Tapio Seppanen. "A wavelet domain LSB insertion algorithm for high capacity audio steganography." In Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop. Proceedings of 2002 IEEE 10th, pp. 53-55. IEEE, 2002.
[15] Ballesteros L, Dora M., and Juan M. Moreno A. "Highly transparent steganography model of speech signals using Efficient Wavelet Masking." Expert Systems with Applications 39, no. 10 (2012): 9141-9149.
[16] Djebbar, Fatiha, Beghdad Ayad, Habib Hamam, and Karim Abed-Meraim. "A view on latest audio steganography techniques." In Innovations in Information Technology (IIT), 2011 International Conference on, pp. 409-414. IEEE, 2011.
[17] Eshelman, Larry J., Richard A. Caruana, and J. David Schaffer. "Biases in the crossover landscape." In Proceedings of the third international conference on Genetic algorithms, pp. 10-19. Morgan Kaufmann Publishers Inc., 1989.
[18] Djebbar, Fatiha, Beghdad Ayad, Habib Hamam, and Karim Abed-Meraim. "A view on latest audio steganography techniques." In Innovations in Information Technology (IIT), 2011 International Conference on, pp. 409-414. IEEE, 2011.
[19] Liu, Qingzhong, Andrew H. Sung, and Mengyu Qiao. "Temporal derivative-based spectrum and mel-cepstrum audio steganalysis." Information Forensics and Security, IEEE Transactions on 4, no. 3 (2009): 359-368.
[20] Haupt, Randy L. "Optimum population size and mutation rate for a simple real genetic algorithm that optimizes array factors." In Antennas and Propagation Society International Symposium, 2000. IEEE, vol. 2, pp. 1034-1037. IEEE, 2000.
[21] Holland, John H. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. U Michigan Press, 1975.
[22] Delforouzi, Ahmad, and Mohammad Pooyan. "Adaptive digital audio steganography based on integer wavelet transform." Circuits, Systems & Signal Processing 27, no. 2 (2008): 247-259.
[23] Skopin, Dmitriy E., Ibrahim MM El-Emary, Rashad J. Rasras, and Ruba S. Diab. "Advanced algorithms in audio steganography for hiding human speech signal." In Advanced Computer Control (ICACC), 2010 2nd International Conference on, vol. 3, pp. 29-32. IEEE, 2010.
[24] Shahreza, S. S., and Mohammad T. Manzuri Shalmani. "High capacity error free wavelet domain speech steganography." In Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, pp. 1729-1732. IEEE, 2008.
[25] Whitley, Darrell, Timothy Starkweather, and Dan Shaner. The traveling salesman and sequence scheduling: Quality solutions using genetic edge recombination. Colorado State University, Department of Computer Science, 1991.
[26] Qi, Yin-Cheng, Liang Ye, and Chong Liu. "Wavelet domain audio steganalysis for multiplicative embedding model." In Wavelet Analysis and Pattern Recognition, 2009. ICWAPR 2009. International Conference on, pp. 429-432. IEEE, 2009.