دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

گزارش پروژه پژوهشی درس پردازش گفتار

موضوع پروژه:

Speech Steganography

پنهان سازی گفتار

نگارش:

تینا خواجه 93210761

شیما شرافتی 93207981

صبا اثنی عشری92204958

استاد درس:

دکتر حسین صامتی

مقدمه

یکی از مسائل مورد توجه در زمینه‌ی ارتباطات دیجیتال از راه دور که به صورت سیگنال های صوتی، تصویری و... صورت می‌گیرد، مسئله‌ی امنیت اطلاعات منتقل شده می‌باشد. این نوع ارتباط که از راه دور صورت می گیرد در معرض خطراتی همچون دسترسی افراد غیر مجاز و تغییر اطلاعات می باشد. تحقیقات بسیاری به منظور فراهم شدن روشی مناسب برای انتقال اطلاعات به منظور جلوگیری از دسترسی افراد غیرمجاز انجام شده و راه‌حل هایی برای این مسائل ارائه شده‌است که از این میان می‌توان به دو روش رمزنگاری[1] و پنهان‌کاری[2] اشاره کرد.

رمز نگاری از طریق درهم‌کردن سیگنال ارتباطی، در فرستنده و ارسال سیگنال به هم ریخته به جای سیگنال اولیه در کانال ارتباطی، امنیت لازم را فراهم می آورد. روش رمزنگاری که الگوریتم‌های مختلفی برای آن ارائه شده‌است، طوری سیگنال را تغییر می‌دهد که در صورت دسترسی فرد غیر مجاز به سیگنال تغییر یافته، اطلاعات غیر قابل فهم بوده و به سختی بتوان به سیگنال اصلی رسید. نکته قابل توجه در مورد روش‌های مختلف استفاده شده با عنوان رمزنگاری این است که فرد غیر مجاز بعد از مشاهده سیگنالی که غیرقابل فهم بوده به رمزی بودن اطلاعات مظنون شده و ممکن است این موضوع باعث تحریک وی به تلاش برای پی بردن به اطلاعات و یافتن رمز شود که این موضوع یکی از نقاط ضعف روش رمزنگاری محسوب می‌گردد.

یکی دیگر از روش های موجود با هدف فراهم آوردن امنیت اطلاعات روش **پنهان سازی** یا Steganography است. این کلمه از ترکیب دو کلمه یونانی stego به معنای پنهان کردن و graphy به معنای نوشتن گرفته شده است. استفاده از روش پنهان‌سازی به زمان‌های گذشته بر می‌گردد و تکنیک‌های مختلفی در طول زمان برای انتقال پیام‌های محرمانه از طریق پنهان‌سازی به کار گرفته شده‌است.

استفاده از این روش برای اطلاعات دیجیتال عبارت است از پنهان کردن اطلاعات در فایل‌های کامپیوتری از جمله فایل های صوتی، عکس، تصویر و... . در این روش سیگنال محرمانه مورد نظر در یک سیگنال ارتباطی دیگر قرار گرفته و پنهان می‌شود. عملیات پنهان‌سازی سیگنال محرمانه با بهره گیری از خطای موجود در حس‌های پنج‌گانه انسانی از قبیل شنوایی و بینایی صورت می‌گیرد و به گونه‌ای انجام می‌شود که توسط این حس‌ها قابل تشخیص نباشد و در صورت

[1]-Cryptography
[2] -Steganography

دسترسی فرد غیر مجاز به آن تنها اطلاعات سیگنال دوم که از نظر امنیتی اهمیت ندارد توسط فرد قابل مشاهده باشد و اثری از عملیات پنهان سازی مشهود نباشد.

پنهان سازی داده‌ها کاربردهای فراوانی در زمینه‌های مختلف دارد که از آن جمله می‌توان به سازمان‌های اطلاعاتی، سازمان‌های نظامی، تصاویر پزشکی، پخش‌های تلویزیونی و رادیویی، سیستم های رادار و سنجش از راه دور و... اشاره کرد.

با توجه به کاربرد بسیار پنهان‌سازی، روش های مختلفی با این هدف ارائه شده اند. روش های مطرح شده برای پنهان سازی خود با چالش‌های مختلفی از جمله میزان کیفیت سیگنال حامل و سیگنال بازسازی شده در گیرنده و نرخ ارسال اطلاعات سیگنال محرمانه روبه‌رو می‌باشد. تحقیقات بسیاری در این زمینه برای بهبود معیارهای اشاره شده انجام شده است.

در گزارش پیش رو ابتدا چند روش موجود برای پنهان سازی معرفی خواهد شد و سپس روش پیاده‌سازی شده در پروژه با جزئیات ارائه می‌شود.

روش های مختلفی برای پنهان سازی یک سیگنال در سیگنال دیگر با عنوان سیگنال حامل ارائه شده است. این روش ها با معیار های مختلفی ارزیابی می شوند. این معیار ها عبارتند از:

۱. میزان مقاومت روش در مقابل افزودن نویز به داده مخفی شده در هنگام فشرده‌سازی و تبدیل فرمت‌ها.

۲. امنیت داده محرمانه

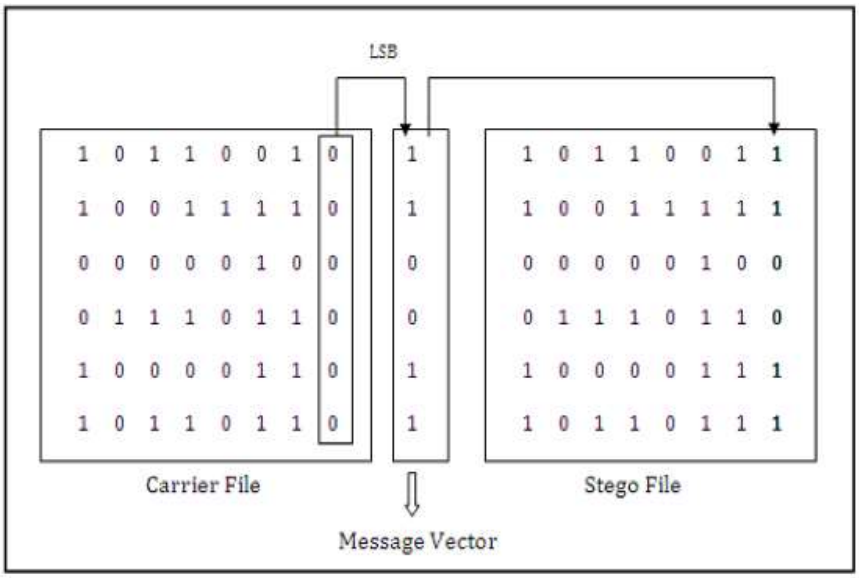
۳. توانایی ذخیره داده پنهان شده با نرخ بالا در سیگنال حامل

در ادامه به معرفی چند روش و بررسی آن‌ها از منظر معیارهای معرفی شده می‌پردازیم.

LSB (Least Significant Bit)

ساده‌ترین روشی که برای پنهان سازی اطلاعات استفاده می‌شود روش LSB می باشد. این روش بر این اصل استوار است که قسمت پر اهمیت تر اطلاعات در بیت‌های پر ارزش تر و قسمت کم ارزش اطلاعات در بیت کم‌ارزش قرار دارد. ابتدا نمایش باینری سیگنال محرمانه را به‌دست آورده و سپس هر یک از بیت های موجود را به ترتیب در یک بایت از

سیگنال حامل قرار می‌دهیم. این روش با توجه به اینکه بیت های داده را در بیت‌های کم ارزش ذخیره می‌نماید تغییرات اندکی در سیگنال پوشش ایجاد می نماید که این تغییرات توسط گوش قابل تشخیص نمی‌باشد. در این روش احتمال تخریب و از بین رفتن اطلاعات کمتر می‌باشد و همچنین از دیگر مزیت‌های این روش می توان به نرخ بالای انتقال اطلاعات اشاره نمود. مشکل موجود در این روش این است که داده ذخیره شده در سیگنال حامل با دستکاری، تغییر فرمت و فشرده‌سازی به راحتی می‌تواند تحت تاثیر قرار گرفته و از بین برود. پنهان کردن داده ها در سیگنال صوتی با روش LSB در سیگنال حوزه زمان یکی از ساده‌ترین الگوریتم های موجود است که نرخ داده پنهان سازی بالایی را نیز فراهم می آورد.



شکل ۱ – مثالی از الگوریتم LSB

Phase Coding

یکی از ویژگی‌های دستگاه شنوایی انسان این است که نسبت به فاز سیگنال صوتی حساسیت کمتری دارد. روش phase coding با بهره گیری از این ویژگی، پیغام محرمانه را در قسمت فاز سیگنال حامل ذخیره می‌نماید. به صورت خلاصه می توان گفت روند پنهان‌سازی از طریق گام های زیر انجام می‌شود:

۱. سیگنال حامل به قسمت‌هایی با اندازه سیگنال محرمانه تقسیم می شود.

۲. با استفاده از تبدیل فوریه، سیگنال به دو بخش اندازه و فاز تقسیم می‌شود.

۳. تفاوت فاز بین قسمت‌های مختلف سیگنال ذخیره می گردد (هر چند که گوش انسان به فاز وابسته نیست اما به اختلاف فاز حساس بوده و این اختلاف باید حفظ گردد).

۴. با استفاده از فاز قسمت اول که در آن اطلاعات ذخیر شده‌اند و ماتریس اختلاف فاز به تهیه ماتریس فاز جدید می پردازیم.

بعد از تکمیل مراحل بالا و به‌دست آوردن ماتریس فاز جدید، فرستنده با استفاده از ماتریس فاز جدید و ماتریس اندازه نگه‌داری شده و با اعمال عکس تبدیل فوریه دوباره سیگنال را بازسازی کرده و ارسال می‌کند.

به منظور بازسازی سیگنال پنهان شده نیز همگام سازی بین گیرنده و فرستنده باید صورت گیرد و و با دانستن طول سیگنال پنهان شده عملیات بازیابی در گیرنده صورت می‌گیرد.

مشکل موجود در این روش کم بودن نرخ ارسال اطلاعات می باشد که باعث شده این روش برای داده‌های با طول کم مناسب باشد.

Spread Spectrum

در این روش اطلاعات به جای تمرکز در یک طیف فرکانسی خاص در پهنای طیف فرکانسی گسترده و پخش می شود. این روش اجازه می دهد که در صورت از بین رفتن بعضی طیف‌های فرکانسی اطلاعات همچنان قابل بازیابی باشند. این روش پنهان سازی را از طریق فراهم آوردن نرخ انتقال اطلاعات متوسط و قابل قبول و در عین حال مقاوم بودن در مقابل تغییرات دو روش قبلی را بهبود بخشیده است.

Echo Hiding

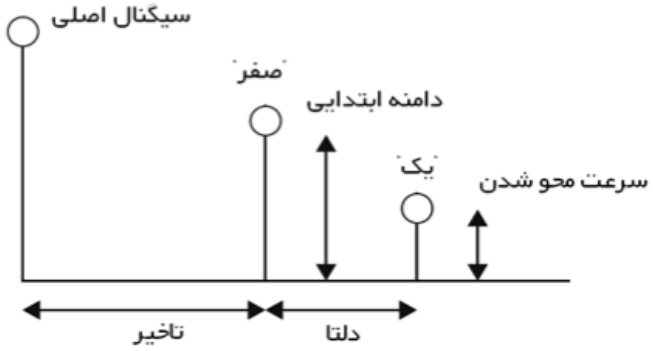
در این روش تلاش بر آن است تا سیگنال f(t) را از طریق معرفی سیگنال $f(t + \Delta t)$ پنهان نماییم.

$$cf(t + \Delta t) = f(t) + \propto f(t - \Delta t)$$

اطلاعات در سیگنال از طریق معرفی و افزوردن پژواک مصنوعی ذخیره می شود. در این روش سیگنال حامل به قسمت هایی به اندازه طول سیگنال پیغام تقسیم می شود و در هر بلاک یک بیت از پیام محرمانه ذخیره می‌گردد. به منظور

ذخیره بیت‌های صفر و یا یک دو نوع مختلف تاخیر Δt و $\Delta t'$ را معرفی می کنیم. در واقع این روش داده را در سیگنال حامل از تغییر سه پارامتر : ۱ـ دامنه سیگنال ۲ـنرخ میرا شوندگی ۳ـمیزان تاخیر پنهان می سازد. مقدار تاخیر پژواک در فریم‌ها به گونه‌ای انتخاب می شود که اثر تخریبی از نظر شنونده نداشته باشد. در نهایت سیگنال حامل و سیگنال تاخیر یافته و تضعیف شده با هم جمع شده و ارسال می‌شود.

به منظور بازیابی پیام در گیرنده اطلاعاتی همچون طول پیام باید بین فرستنده و گیرنده به اشتراک گذاشته شود تا از هر قسمت یک بیت استخراج گردد. اطلاعات ذخیره شده در سیگنال ارسال شده از طریق محاسبه میزان شباهت یک سیگنال به خود قابل استخراج می باشد. این روش در مقابل تغییر و از بین رفتن اطلاعات موجود در سیگنال تا حد قابل قبولی مقاوم می‌باشد.



شکل ۲ - پارامترهای قابل تنظیم در روش Echo

در این روش‌ها سیگنال حامل به حوزه دیگری منتقل شده و سپس با بهره‌گیری از ویژگی های سیگنال در آن حوزه سیگنال پیغام را پنهان کرده و سپس سیگنال جدید بدست آمده را به حوزه زمان برگردانده و ارسال می‌کنیم.

حوزه‌های مرسوم برای پنهان سازی معمولا عبارتند از : تبدیل کسینوس گسسته، تبدیل فوریه گسسته و تبدیل ویولت گسسته. چنین روش‌هایی که از ویژگی سیگنال در حوزه‌های دیگر از جمله حوزه فرکانس و... بهره می‌برد معمولا نسبت به روش هایی که بر پایه حوزه زمان در مقابل نویز و از بین رفتن اطلاعات مقاوم تر می‌باشد.

سیگنال گفتار دارای خواص بسیاری می باشد که در پروژه پیش رو سعی برآن شده تا با بهره گیری از این ویژگی‌ها پنهان سازی سیگنال گفتار در سیگنال گفتار صورت گیرد. در روش ارائه شده تلاش بر آن است که ضمن پنهان سازی کیفیت سیگنال حامل تا حد ممکن تحت تاثیر واقع نشده و همچنان قابل فهم باشد.

کاربرد های مختلفی برای روش ارائه شده می توان در نظر داشت که از آن میان می توان به کاهش فضای ذخیره سازی گفتار، کاهش هزینه ارسال اطلاعات ، امنیت انتقال سیگنال گفتار و... اشاره نمود.

الگوریتم پنهان سازی و مراحل پیاده سازی

همان طور که گفته شد، در مبحث پنهان سازی سیگنال صوت، هدف پنهان کردن یک سیگنال گفتار حاوی پیغامی محرمانه در سیگنال دیگر و متفاوت از سیگنال محرمانه به نام حامل است. بنابراین در این پروژه هدف پنهان کردن یک پیغام صوتی محرمانه در سیگنال صوتی دیگری است به طوری که :

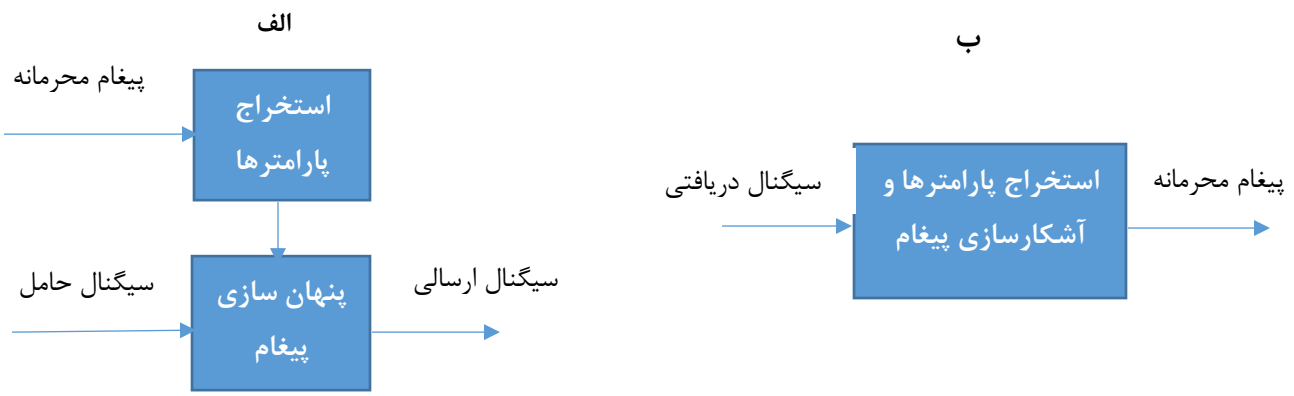
۱. سیگنال صوتی پس از ذخیره پیغام(stego) بیشترین شباهت را به صوت حامل داشته باشد، به گونهای که مشخص نشود این سیگنال حاوی پیغامی محرمانه است.

۲. پیغام محرمانه طوری در صوت حامل ذخیره شود که سیگنال بازسازی شده در گیرنده قابل فهم باشد.

شکل ۱، نمای کلی مراحل پنهان‌سازی و آشکارسازی پیام محرمانه را در قسمت گیرنده و فرستنده نشان می‌دهد. با نگاهی به شکل در می‌یابیم که فرستنده دارای دو بخش زیر است:

۱. دریافت پیام محرمانه و استخراج ویژگی‌هایی از آن برای نمایش این پیام

۲. دریافت سیگنال حامل و پارامترهای مربوط به سیگنال محرمانه از بخش ۱ و سپس ذخیره‌ی این پارامترها در سیگنال حامل و ارسال سیگنال به‌دست‌آمده بر روی کانال



شکل ۱ – نمای کلی الگوریتم پنهان سازی پیغام

الف: فرستنده – ب: گیرنده

در قسمت گیرنده نیز سیگنال حاوی پیام محرمانه دریافت می‌شود و طبق الگوریتمی که در قسمت دوم فرستنده ذکر شد، پارامترهای مربوط به پیام محرمانه از سیگنال دریافتی استخراج می‌شود.
در ادامه دو بخش اصلی فرستنده به تفصیل معرفی می شود و جزئیات پیاده‌سازی انجام شده در این پروژه برای هر بخش شرح داده خواهد شد. سپس توضیح مختصری نیز در مورد پیاده سازی گیرنده ارایه می شود.

فرستنده

۱. استخراج پارامترها

برای ارسال پیغام محرمانه به جای پنهان سازی کل پیغام در سیگنال حامل، پارامترهایی از آن استخراج میشود (طوری که این پارامترها با دقت خوبی سیگنال محرمانه را بازسازی نماید) و این پارامترها در سیگنال حامل پنهان می‌شود. دلایل بسیاری برای استفاده از پارامترها به جای استفاده از کل پیغام وجود دارد. از جمله‌ی این دلایل می‌توان به موارد زیر اشاره نمود:

۱. در بحث پنهان‌سازی سیگنال گفتار، هدف پنهان سازی پیغام محرمانه است طوری که سیگنال حاوی پیغام بیشترین شباهت را با سیگنال اولیه قبل از پنهان سازی داشته باشد (تا حد ممکن مشخص نشود که این سیگنال حاوی پیغام است) و همچنین سیگنال محرمانه‌ی بازسازی شده در قسمت گیرنده قابل درک باشد (لزومی به کیفیت بسیار بالا برای این سیگنال وجود ندارد). با توجه به توضیحات فوق، در صورت استفاده از پارامترهایی از سیگنال به جای کل سیگنال می‌توان به هر دو هدف نایل شد.

۲. تعداد مکان‌های مناسب در سیگنال حامل برای پنهان سازی سیگنال محرمانه محدود است. (به دلیل محدود بودن پهنای باند سیگنال گفتار) و بنابراین نمیتوان از تعداد زیادی پارامتر برای ارسال پیغام محرمانه استفاده کرد.

۳. در صورت انتخاب درست پارامترها و تعداد آن، می‌توان به نرخ بیشتر ارسال (تعداد اطلاعات محرمانه ارسالی به ازای سیگنال حامل) دست یافت.

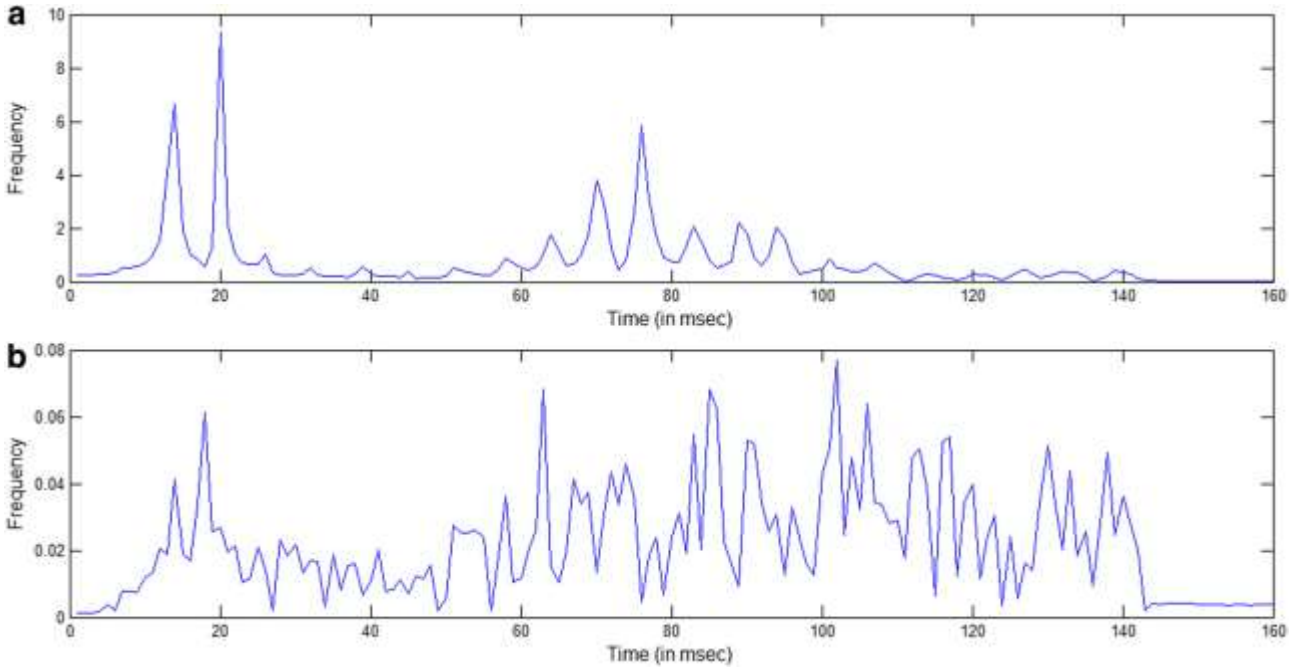
در این پروژه از ضرایب LPC برای بازسازی سیگنال محرمانه استفاده می‌کنیم. برای توضیح بیشتر ضرائب LPC به ۱ مراجعه شود.

۲. پنهان سازی

پس از استخراج پارامترها از سیگنال محرمانه باید این پارامترها را در سیگنال حامل پنهان کنیم. الگوریتم‌های بسیاری در این زمینه موجود است که اکثر آن‌ها از ویژگی‌های سیگنال گفتار برای پنهان‌سازی پیغام استفاده می‌کنند. در ادامه به شرح این ویژگی‌ها می‌پردازیم.

۲,۱ ویژگی‌های طیفی سیگنال گفتار

سیگنال گفتار یک سیگنال با پهنای باند محدود است که معمولا پهنایی بین ۴ تا ۷ کیلوهرتز (برای حالت باند باریک و باند پهن) دارد. شکل ۲، طیف سیگنال گفتار باند پهن را برای یک واج voiced و یک واجunvoiced را نشان می‌دهد.



شکل ۲ – طیف سیگنال گفتار برای یک frame

a: واج voiced – b: واج unvoiced

با توجه به شکل، در می‌یابیم که سیگنال گفتار عموما سیگنالی با مؤلفه‌های فرکانس پایین است و مؤلفه‌های بالای ۴ تا ۷ کیلوهرتز به تعداد محدودی در طیف فرکانسی آن وجود دارد. بنابراین می‌توان دریافت که در صورتی که مؤلفه‌های فرکانس بالا را از سیگنال گفتار حذف کنیم، با وضوح قابل قبولی همچنان قابل فهم است و به سیگنال قبلی شباهت دارد.

بنابراین برای پنهان سازی پیغام در سیگنال گفتار در صورتی که از مؤلفه‌های فرکانس بالای سیگنال حامل برای این کار استفاده نماییم می‌توانیم انتظار داشته باشیم که سیگنال به‌دست‌آمده (stego) تا حد زیادی به سیگنال اولیه (حامل) شباهت داشته باشد. بنابراین باید ابتدا به نحوی فرکانس‌های بالا و پایین سیگنال حامل را جدا کرده و پارامترهای مربوط به پیغام محرمانه را در فرکانس‌های بالای سیگنال حامل ذخیره نماییم.
برای جداسازی مؤلفه‌های فرکانس پایین و بالای سیگنال صوتی از تبدیل موجک استفاده می‌کنیم. در ادامه به توضیح مختصری از تبدیل موجک و سپس دلیل انتخاب این تبدیل می‌پردازیم.

۲,۱,۱ تبدیل موجک [3]

هرچند تبدیل فوریه یک ابزار مفید برای تحلیل مؤلفه های فرکانسی یک سیگنال است اما با محاسبه تبدیل فوریه‌ی یک سیگنال، نمیتوان به طور دقیق مشخص کرد که در چه زمانی یک فرکانس خاص در سیگنال ظاهر شده‌است. تبدیل فوریه زمان کوتاه [4]از یک پنجره لغزان برای به دست آوردن اسپکتروگرام سیگنال استفاده میکند تااطلاعات زمان و فرکانس هر دو حفظ شود. اما هنوز یک مشکل باقی می‌ماند: طول پنجره رزولوشن فرکانسی را کنترل می‌کند. به نظر میرسد تبدیل موجک راه حل این مشکل است. تبدیل‌های موجک بر پایه موجک‌های[5] کوچک با طول محدود هستند. نسخه انتقال یافته از موجک‌ها دقیقا بر روی زمانی که مورد نظر ماست قرار می‌گیرند در حالی که نسخه تغییر مقیاس یافته ازموجک‌ها به ما اجازه می‌دهد که سیگنال را در مقیاس های مختلف بررسی کنیم.

ایده‌ی اصلی در استفاده از تبدیل موجک تحلیل براساس مقیاس است. هر سیگنال می‌تواند براساس نسخه‌های انتقال یافته و مقیاس شده‌ی موجک مادرنشان داده شود. تحلیل موجک می‌تواند جنبه‌ای از داده را روشن کند که تحلیل‌های دیگر سیگنال قادر به انجام آن نیستند. جنبه‌هایی همچون روند تغییر سیگنال با زمان، ناپیوستگی‌های موجود در مشتق‌های بالاتر سیگنال، نقاط شکست و خود شباهتی.
تحلیل موجک اجازه‌ی تجزیه‌ی سیگنال به دو بخش فرکانس بالا و فرکانس پایین را به ما می‌دهد. نقاط لبه‌ای سیگنال محدود به فرکانس‌های بالای سیگنال هستند. برای تحلیل فرکانس‌های بالای سیگنال از مجموعه‌ای از فیلترهای بالاگذر و برای تحلیل فرکانس‌های پایین از مجموعه‌ای از فیلترهای پایین گذر استفاده می‌شود. در واقع فیلترهای با فرکانس قطع مختلف برای تحلیل سیگنال در رزولوشن‌های مختلف استفاده می‌شود.

---
[3] Wavelet Transform
[4] Short Fourier Transform
[5] Mother Wavelets

تبدیل موجک براساس مقیاس‌ها و مکان‌های توان دوانجام می‌شود. موجک مادر با توان دو مقیاس بندی می‌شود و با اعداد طبیعی انتقال داده می‌شود تا موجک‌های دختر تولید شود. از این موجک‌ها برای ساخت سیگنال اصلی استفاده می‌شود.به‌صورت دقیق‌تر هر تابع$f(t) \in L^2(R)$رامی‌توان به صورت زیر نشان داد:

$$f(t) = \sum_{j=1}^{L} \sum_{k=-\infty}^{\infty} d(j,k)\varphi(2^{-j}t - k) + \sum_{k=-\infty}^{\infty} a(L,k)\emptyset(2^{-L}t - k)$$

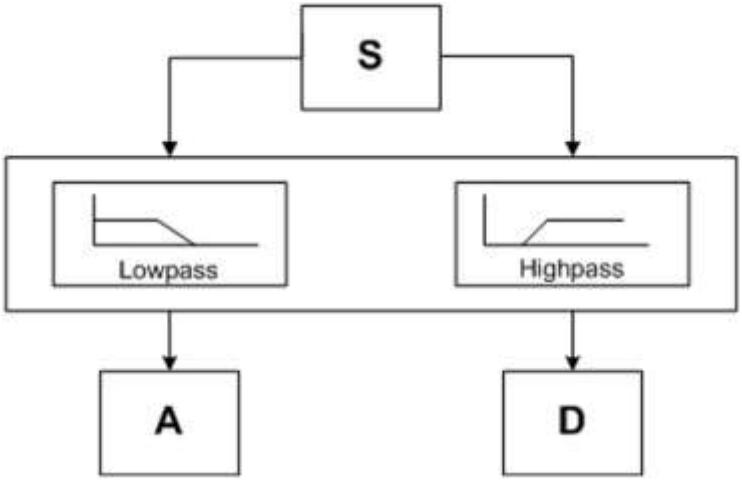
تابع$\varphi(t)$ به عنوان موجک مادر و تابع$\emptyset(t)$ به عنوان تابع مقیاسی [6]در نظر گرفته می‌شود.ضرائب$a(L,k)$ به عنوان ضرائب تقریبی [7]در مقیاس L وضرائب$d(j,k)$ به عنوان ضرائب جزئی [8]در مقیاس j شناخته می‌شود. ضرائب تقریبی و جزئی به صورت زیر به‌دست می‌آیند.

$$a(L,k) = \frac{1}{\sqrt{2^L}} \int_{-\infty}^{\infty} f(t)\emptyset(2^{-L}t - k)dt$$

$$a(L,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(t)\varphi(2^{-j}t - k)dt$$

با توجه به توضیحات داده شده، می‌توان دریافت که تبدیل موجک همچون دو فیلتر پایین گذر و بالاگذر عمل کرده و سیگنال را به دو بخش فرکانس پایین و فرکانس بالا یا تقریب(بخش‌های مقیاس بالا و فرکانس پایین سیگنال) و جزئیات(بخش‌های مقیاس پایین و فرکانس بالای سیگنال) تبدیل می‌کند(شکل ۳).

[6] Scaling Function
[7] Approximation Coefficients
[8] Detail Coefficients

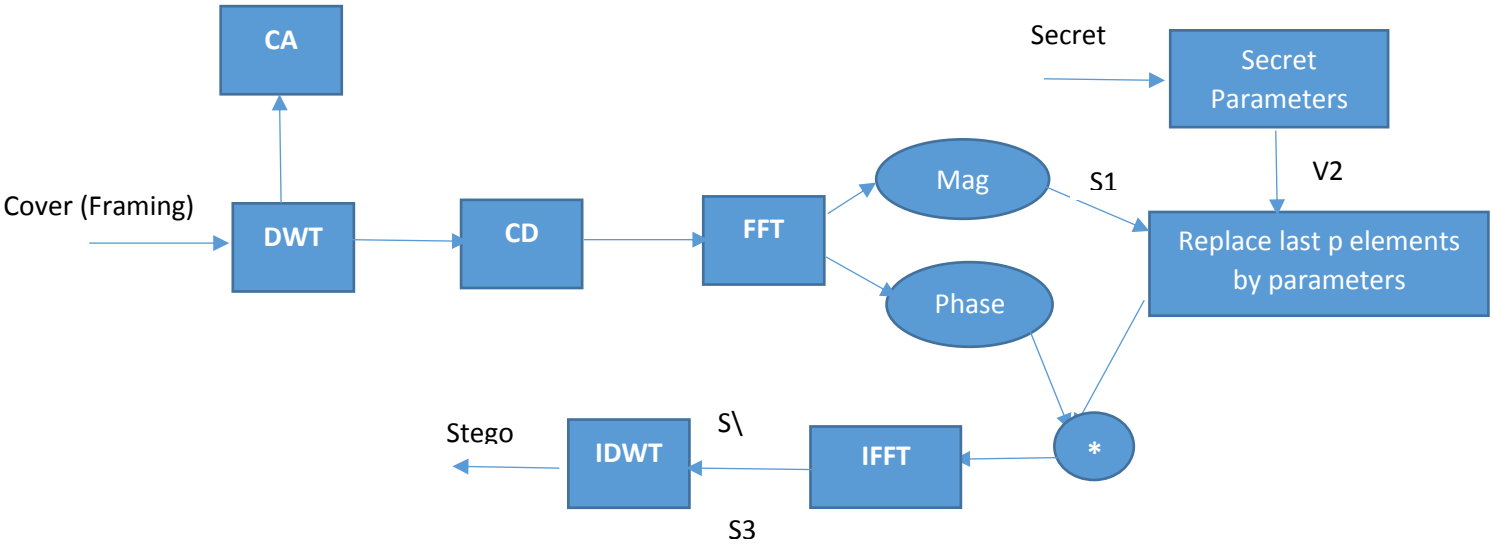


شکل ۳ – شبیه سازی تبدیل

۲,۲ الگوریتم پنهان سازی

شکل ۴، مراحل اصلی الگوریتم به کار رفته در این پروژه را نشان می‌دهد. با توجه به این شکل، ابتدا سیگنال حامل (cover) وارد شده و فریم بندی می‌شود. در این پروژه فریم ها بازه‌های ۲۰ میلی ثانیه ای یا ۱۶۰ نقطه‌ای با میزان همپوشانی ۱۰ میلی ثانیه یا ۸۰ نقطه انتخاب شده‌اند. سپس از هر فریم تبدیل موجک گرفته می‌شود. همان‌طور که در بخش قبل گفته شد پارامترهای مربوط به سیگنال محرمانه در قسمت‌های فرکانس بالای سیگنال حامل ذخیره می‌شود. بنابراین پس از اعمال تبدیل موجک، از بخش جزئیات (CD) تبدیل فوریه گرفته و پارامترهای به دست‌آمده از سیگنال محرمانه در دامنه‌ی این تبدیل ذخیره می‌شود (استفاده از تبدیل فوریه برای پیچیدگی بیشتر الگوریتم پنهان سازی است). در آخر نیز با استفاده از مؤلفه‌های فرکانس بالای سیگنال حامل (CD) که اکنون پارامترهای مربوط به پیغام را در خود دارد (s3) و مؤلفه‌های فرکانس پایین آن (CA) معکوس تبدیل موجک(IDWT) گرفته می‌شود تا به سیگنال stego برسیم.

ذکر این نکته نیز ضروری است که برای استفاده از ضرائب بردار s3 در اعمال تبدیل IDWT این ضرائب بر مقداری تقسیم می‌شود زیرا همان طور که گفتیم CA مربوط به بخش مقیاس پایین و فرکانس بالای سیگنال است و در صورتی که مقادیر مربوط به این ضرائب زیاد باشد سیگنال stego حالتی غیر طبیعی به خود میگیرد و پنهان بودن پیغامی محرمانه در آن مشهود است.



شکل ۴ – مراحل الگوریتم پنهان سازی

در ادامه نحوه ذخیره سازی پارامترهای LPC و مشکلات مواجه شده در این زمینه معرفی خواهد شد.

۲,۲,۳ ذخیره سازی پارامترهای LPC

با توجه به شکل ۴، پارامترهای مربوط به ضرائب LPC (بردار v2) در p عنصر نهایی بردار s1 ذخیره می‌شود. یکی از مشکلاتی که در این زمینه پیش می‌آید این است که مقادیر بردار v2 اعداد صحیح هستند در حالی که این مقادیر باید در دامنه تبدیل FFT به دست آمده (بردار s1) ذخیره شوند و همان طور که می‌دانیم s1 دارای مقادیر مثبت است.

برای رفع این مشکل دو راه وجود دارد.

۱. استفاده از ضرائب LSF

۲. ذخیره علامت های مربوط به ضرائب LPC

در این پروژه از راه‌کار دوم استفاده شده‌است. یعنی پس از قرار دادن هر ضریب در بردار s1 علامت مربوط به آن نیز در در عنصر بعدی در s1 ذخیره می‌شود.

مشکل دیگری که در مرحله ذخیره سازی پارامترها به وجود می‌آید این است که اندازه بردار v2 ممکن است بیش از نصف اندازه بردار s1 باشد (از آن‌جایی که بردار s1 مربوط به دامنه FFT است باید از نصف قرینه باشد – شکل ۵). برای رفع این مشکل نیز چند راه‌حل به کار گرفته شده‌است:

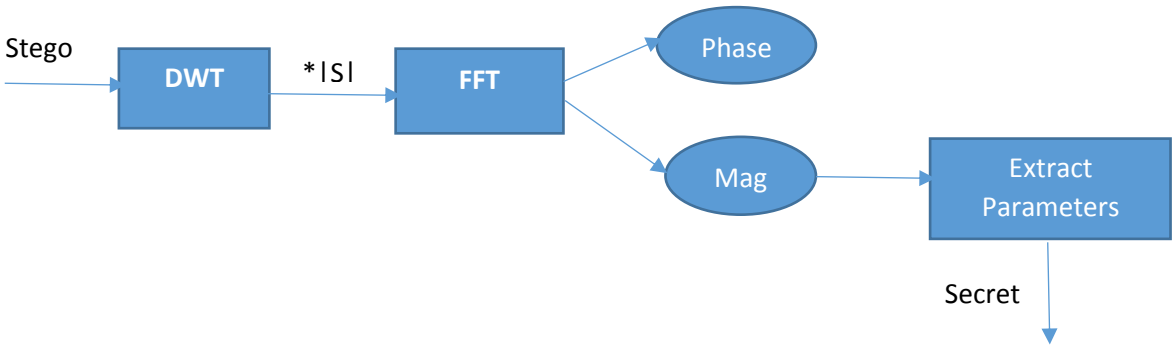
۱. به جای ذخیره‌ی همه‌ی پارامترها، پارامترها به صورت k تا در میان ذخیره شوند: این کار باعث افت کیفیت سیگنال محرمانه بازسازی شده می‌شود ولی میتوان k را طوری تعیین نمود که سیگنال بازسازی شده از کیفیت قابل قبولی برخوردار باشد.

۲ اندازه فریم های سیگنال حامل بیش از اندازه‌ی فریم‌های سیگنال محرمانه باشد: هرچند بزرگ‌تر بودن طول فریم‌های s1 باعث می‌شود نرخ ارسال پیغام محرمانه کاهش یابد ولی میتوان نسبت طول دو فریم را طوری تنظیم کرد که نسبت مناسبی بین طول بردارهای v2 و s1 برقرار شود و همچنین نرخ ارسال قابل قبولی داشته باشیم.

شکل

گیرنده

با توجه به شکل ۶، در این قسمت مراحل طی شده در فرستنده را به صورت معکوس طی می‌نماییم تا به پارامترهای مربوط به سیگنال محرمانه برسیم. سپس بر اساس این پارامترها (در این‌جا پارامترهای LPC) سیگنال را بازسازی می‌کنیم.



شکل ۶ – مراحل الگوریتم آشکارسازی پیغام محرمانه در گیرنده

ارزیابی

مقدمات ارزیابی

برای ارزیابی تکنیکِ ارائه شده شبیه سازی های متعددی با استفاده از بانک اطلاعاتی NOIZEUS صورت گرفته است. این بانک اطلاعاتی شامل 30 جمله موجود در پایگاه دادهِ جمله ای IEEE می باشد و صدا ها توسط 3 گوینده‌ی زن و مرد بیان شده است.

این 30 جمله که 15 مورد آن توسط مرد و 15 مورد دیگر ان توسط زن گفته شده است شامل تمام فونیم‌های انگلیسی می‌باشد. فرکانس اصلی نمونه‌ها25KHZ می باشد که به 8 KHZ تعدیل یافته‌اند.

برای ارزیابی 4 مجموعه تست در نظر گرفته شده است.در مجموعه‌ی اول هر یک از 15 فایل سخنرانی مرد را در هر 15 فایل سخنرانی زن پنهان سازی می شود. مجموعه‌ی دوم هر یک از 15 فایل سخنرانی زن را در هر 15 فایل سخنرانی مرد پنهان سازی می شود. در مجموعه‌ی سوم هر یک از 15 فایل سخنرانی مرد در 14 فایل سخنرانی باقی مانده دیگر پنهان سازی می‌شود و در مجموعه‌ی آخر هر یک از 15 فایل سخنرانی زن در 14 فایل سخنرانی باقی مانده پنهان سازی می شود.

هر مجموعه تست برای 5 نوع مختلف از خانواده‌ی wavelet ها (Haar, Daubechies, Symlets, Coiflets, and BiorSplines) تکرار شده است. و در نهایت 4210 تست صورت گرفته است.
برای ارزیابی تاثیر تکنیک DWT-FFT دو آزمایش مختلف با استفاده از متد DWT-FTT و با استفاده از متد FFT صورت گرفته است.

نتیجه ارزیابی

یکی از معیارهای ارزیابی هر سیستم استگانوگرافی مقایسه بین سیگنال حامل و سیگنال حاملِ حاوی پیام می باشد.که این مقایسه را می‌توان با تست‌های objective و subjective صورت داد.
در تست subjective چند تست مقایسه ای غیر رسمی صورت گرفته که در این تست‌ها سیگنال حامل و سیگنال حاملِ حاوی پیام با ترتیبی تصادفی برای تعدادی شنونده پخش می‌شوند و شنونده باید بین سیگنال حامل و سیگنال حاملِ حاوی پیام پخش شده سیگنال با کیفیت بالاتر را مشخص کند. در این تست اکثر شنوندها نتوانستند متوجهِ تفاوتی بین این دو سیگنال شوند

در تست های objective از معیار SegSNR (نسبت سیگنال به نویز) استفاده شده است که به صورت زیر تعریف می شود.

$$\mathrm{SegSNR(dB)} = 10\log_{10}\left(\frac{\sum_{m=0}^{159}[s_1(m)]^2}{\sum_{m=0}^{159}[s_1(m)-s_3(m)]^2}\right)$$

معیار دیگری که برای این کار وجود دارد PESQ می‌باشد که ارزیابی درکی از سیگنال گفتار می‌باشد که یک تست objective برای ارزیابی کیفیت گفتار می‌باشد.

که s_1 و s_3 به ترتیب سیگنال حامل و سیگنال حاملِ حاوی پیام می باشد.در این کار فریم‌های گفتار بازه های 20 ms (160 نمونه در هر فریم) در نظر گرفته شده است.

در جدول 1 میانگین مقدارSegSNR را برای 4 مجموعه متفاوت از تست ها با استفاده از الگوریتم DWT-FFT آورده شده است. در جدول 2 میانگین مقدارSegSNR را برای همان مجموعه از تست ها با فقط استفاده از الگوریتمFFT آورده شده است.

با توجه به نتایج حاصل شده از این جداول در می یابیم که بالاترین SegSNR زمانی حاصل می شود که سیگنال cover و سیگنال secret هر دو از فایل سخنرانی مرد باشند . البته این نتایج قابل پیش‌بینی بودند به این دلیل که صدای مرد زیر می باشد و دارای فرکانس‌های پایین بیشتری نسبت به صدای زن می‌باشد و چون در این روش ما اطلاعات را در فرکانس‌های بالا ذخیره می‌نماییم اطلاعات کمتری را از دست می‌دهیم.

بعد از این بیشترین SegSNR مربوط به زمانی می‌شود که سیگنال cover از فایل سخنرانی مرد و سیگنال secret از فایل سخنرانی زن باشد.

و بین دو حالت بعدی تفاوت چندانی دیده نمی‌شود

Table 1 SNR of the DWT-FFT-based hiding approach (paper result)

Cover signals	**Secret signals**	**SegSNR(dB)**
Female	*Male*	31.86
Male	Female	32.70
Male	Male	34.45
Female	Female	31.13
Average		**32.54**

Table 1 SNR of the DWT-FFT-based hiding approach (our result)

Cover signals	**Secret signals**	**SegSNR(dB)**
Female	*Male*	22.87
Male	Female	27.02
Male	Male	30.34
Female	Female	34.49
Average		**25.93**

Table 2 SNR of the FFT-based hiding approach

Cover signals	Secret signals	SegSNR(dB)
Female	*Male*	51.46
Male	Female	52.62
Male	Male	54.37
Female	Female	51.09
Average		**52.39**

با مقایسه این دو جدول متوجه می شویم که کیفیت سیگنال پنهان سازی شده در حالتی که از الگوریتم FFT استفاده شده است نسبت به حالتی که از DWT-FFT استفاده شده است بهتر می‌باشد هر چند در الگوریتم DWT-FTT مقاومت سیگنال پنهان سازی شده در مقابله با تکنیک های شناسایی افزایش می‌دهد.

در مرحله بعد از wavelet های مختلف برای مقایسه تاثیرشان در کیفیت سیگنال پنهان سازی شده استفاده شده است. جدول 3 نشان دهنده‌ی نتایج این تست‌ها می‌باشد.

Table 3 Different wavelets results of DWT-FFT-based steganography systems(paper result)

Wavelet name		Haar	Daubechies(db1)	Symlets(sym1)	Coiflets(coif1)	biorSplines (bior)
Cover signals	Secret signals	SegSNR(dB)	SegSNR(dB)	SegSNR(dB)	SegSNR(dB)	SegSNR(dB)
Female	*Male*	31.53	31.86	31.48	31.41	31.39
Male	Female	31.98	32.70	31.86	31.96	31.91
Male	Male	34.12	34.35	34.08	34.08	34.04
Female	Female	30.79	31.13	30.68	30.76	30.71
Average		32.11	32.51	32.03	32.05	32.01

Wavelet name		Haar	Daubechies(db1)
Cover signals	Secret signals	SegSNR(dB)	SegSNR(dB)
Female	*Male*	28.17	28.25
Male	Female	28.95	29.06
Male	Male	31.54	31.87
Female	Female	27.07	27.35
Average		28.93	29.13

Table 3 Different wavelets results of DWT-FFT-based steganography systems(our result)

باتوجه به جدول مشخص می‌شود که انواع مختلف wavelet ها نتایج تقریبا یکسانی دارند بنا براین این روش وابسته به نوع خاصی از wavelet نیست و SegSNR تفاوت چندانی برای wavelet های مختلف قائل نمی‌شود

معیار مقایسه SegSNR فقط مشخص کننده‌ی کارایی می‌باشد در حالی که PESQ معیار معتبرتری برای ارزیابی کارایی تکنیک مخفی سازی می‌باشد و تکنیکی برای ارزیابی کیفیت سیگنال فراهم می‌کند.

با استفاده از این تکنیک می توان میزان تخریب سیگنال حامل در حالتی که شامل سیگنال پیام و در حالتی که شامل سیگنال پیام نیست را بدست اورد.

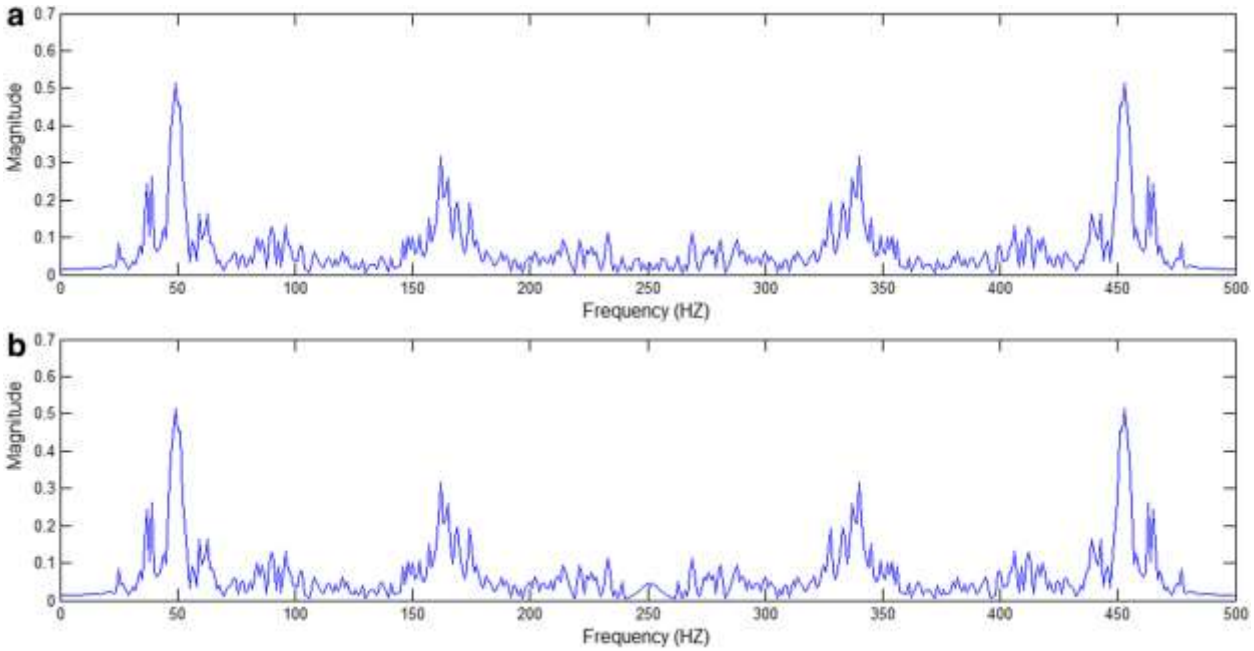
در جدول 4 میانگین مقدار PESQ با استفاده از دو تکنیکDWT-FFT و FFT آورده شده است.

Table 4 PESQ of DWT-FFT and FFT-based hiding approach

Speaker	**PESQ**	
	DWT-FFT	FFT
Female	3.58	4.12
Male	3.78	4.16
Average	**3.68**	**4.14**

شکل 1 دامنه تغییراتِ PESQ را برای 20 سیگنال صوتی و برای هر دو تکنیک DWT-FFT و FFT نشان می دهد . تکنیک مخفی سازی دارای میانگینِ PESQ 3.68 و 4.14 به ترتیب برای DWT-FFT و FFT می‌باشد .

شکل نشان دهنده‌ی مقدار سیگنال حامل و سیگنال حامل معادل پس از پنهان کردن ضرایب LPC پیام در آن می باشد.



تحلیل PESQ نشان می‌دهد که سیگنال حامل و سیگنال حاملِ حاوی پیام دارای کیفیت‌های مشابه ای می‌باشند.

تست های subjective و objective برای ارزیابی کارایی نشان می‌دهند که روش مخفی سازی ارئه شده هیچ گونه شکی در رابطه با وجود یک سیگنال پیام در سیگنال حامل ایجاد نمی‌کند در حالی که قادر است سیگنال پیام را به طور کامل بازسازی کند .

سیگنال پیام ساخته شده از هر دو روش DWT-FFT و FTT کاملا قابل درک است اما یک سری اعوجاجات در ان قابل شنیدن می‌باشداما آنچه هدف این روش می باشد قابل فهم بودن پیام و رساندن آن به گیرنده می‌باشد

جدول 5 نشان دهنده ی تاثیر الگوریتم‌های پیاده سازی شده بر اساس معیار SegSNR می‌باشد

Table 5 Impact of the hiding process on the secret speech in terms of SegSNR

Speaker	**SegSNR**	
	DWT-FFT	FFT
Female	21.76	24.64
Male	23.89	26.28
Average	**22.83**	**25.46**

Table 5 Impact of the hiding process on the secret speech in terms of SegSNR(our result)

Speaker	**SegSNR**	
	DWT-FFT	FFT

Female	13.43	24.64
Male	15.29	26.28
Average	**14.36**	**25.46**

نتیجه گیری

در این پروژه یک سیستم استگانوگرافی امنیتی جدید ارائه شده است که هدف از آن پنهان سازی پیام صوتی می‌باشد. در این روش پیام صوتی به گونه ای پنهان می شود که در گفتار حامل قابل شنیدن نیست و همچنین به دلیل استفاده از روش‌های پیچیده ای برای پنهان سازی به راحتی توسط افراد استراق سمع کننده قابل بازیابی نیست که هدف اصلی در الگوریتم های استگانوگرافی نیز همین می‌باشد.

REFERENCES

1) Rekik et al.: Speech steganography using wavelet and Fourier transforms. EURASIP Journal on Audio, Speech, and Music Processing 2012 2012:20.
2) s.Rekik , D. Guerchi, H.Hamam, Audio Steganography Coding Using the Discrete Wavelet Transforms
3) M.Nosrati, R.Karimi, Audio Steganography: A Survey on Recent Approaches World Applied Programming, Vol (2), No (3), March 2012. 202-205