تشخیص پنهان نگاری در استانداردهای کدگذار صوتی LPC، CELP و MELP با استفاده از شبکه عصبی LVQ

تاریخ دریافت:
۲۲ اردیبهشت ماه ۱۴۰۳

تاریخ پذیرش:
۲۱ خرداد ماه ۱۴۰۳

پوریا اعتضادی فر*[۱]، سعید طلعتی[۲]، محمدرضا حسنی آهنگر[۳]، مهدی ملازاده[۴]

۱. استادیار، دانشکده مهندسی برق و کامپیوتر، دانشگاه جامع امام حسین (ع)، تهران، ایران.
۲. دانشجوی دکتری، دانشگاه جامع امام حسین (ع)، تهران، ایران.
۳. استاد، دانشگاه جامع امام حسین (ع)، تهران، ایران.
۴. استادیار، دانشگاه جامع امام حسین (ع)، تهران، ایران.

چکیده

امنیت اطلاعات یکی از مسائل بسیارمهم حوزه جنگ الکترونیک است که همواره موردتوجه محققان زیادی قرار می گیرد. پنهان نگاری یکی از روش های ارتباطات امن بوده که هدف آن مخفی کردن اطلاعات در بستری از داده و محتوا است؛ داده های صوتی دارای پیچیدگی فشرده سازی بسیار بالایی هستند که رفتار آنها را شبیه به آنتروپی می‌کند و در نتیجه تشخیص پنهان نگاری در این سیگنال ها بسیار مشکل است. در این مقاله روشی جدید در تشخیص پنهان نگاری در سه استاندارد کدگذار صوتی LPC، CELP و MELPکه جزو قدرتمندترین روش های کدگذاری صوتی هستند ارائه می شود. داده های مخفی در بیت کم ارزش رسانه حامل پنهان شده اند، تحلیل صوت با استفاده از ویژگی آزمون اجرا ارائه می شود. پس از بررسی صوت اصلی و رسانه حامل، ویژگی‌های متفاوت بین این اصوات استخراج شده و برای آموزش شبکه عصبی هوشمند (LVQ) استفاده می‌شود. مرحله طبقه‌بندی داده ها با استفاده از لایه‌های این شبکه عصبی انجام‌شده و الگوریتم پیشنهادی برای این صداها تست‌شده است. میزان میانگین دقت آشکارسازی پنهان نگاری روش پیشنهادی ۹۳.۵۹٪ است که برتری این روش را در مقایسه با سایر روش ها نشان می دهد.

واژه‌های کلیدی: پنهان نگاری، نهان کاوی، آشکارسازی، تخمین خطی، شبکه عصبی LVQ.

Detection of Steganography in LPC, CELP and MELP Audio Standards Encoder Using LVQ Neural Network

Pouriya Etezadifar*[1], Saeed Talati[2], Mohammadreza Hassani Ahangar[3], Mahdi Mollazade[4]

1. Assistant professor, Department of Electrical Engineering, Imam Hossein University, Tehran, Iran.
2. Ph.d Student, Imam Hossein University, Tehran, Iran.
3. Professor, Imam Hossein University, Tehran, Iran.
4. Assistant professor, Imam Hossein University, Tehran, Iran.

Abstract

Information security is a critical aspect of Electronic Warfare today and is a major focus for many researchers. Steganography, a secure communication method, involves concealing information within other data or content, with audio data offering a higher capacity for hiding information. This article introduces a novel approach for detecting steganography in three audio encoders: LPC, CELP, and MELP, known for their effectiveness in audio encoders. The Steganography technique discussed here involves embedding data in the least significant bit of the audio media. Audio analysis is carried out using the RUNS test features. Variations are identified and utilized to train a neural network (LVQ) by comparing these features in the cover and stego audio files. The neural network is then employed for classification, and the proposed algorithm is tested on the audio samples. The proposed classification method demonstrates an average accuracy of 93.59% in detecting steganography, showcasing its effectiveness compared to other existing methods.

Key words: Steganography, Steganalysis, Detection, Linear Predictive, LVQ Neural Network.

ایمیل نویسنده مسئول: petezadifar@ihu.ac.ir

۱-مقدمه

پنهان‌نگاری علم و هنر مخفی کردن اطلاعات است که به‌وسیله آن می‌توان اطلاعات محرمانه را درون یک پوشانه مخفی کرد[۱]. در حال حاضر کاربردهای اصلی پنهان‌نگاری پوشش‌های دیجیتالی همچون تصویر، صوت، ویدئو و متن هستند[۲]. با رواج شبکه‌های اجتماعی که از فایل‌های صوتی پشتیبانی می‌کنند، انتقال صوت به یک روش محبوب تبدیل‌شده است. با توجه به عدم حساسیت بالای سیستم شنوایی انسان و ظرفیت بالای فایل‌های صوتی، پنهان‌نگاری در صوت دیجیتال به‌عنوان پوشش بسیار مورداستفاده قرار می‌گیرد[۳]. سیگنال‌های صوتی دارای پیچیدگی فشرده‌سازی بسیار بالایی هستند که رفتار آن‌ها را شبیه به آنتروپی می‌کند و درنتیجه تشخیص پنهان‌نگاری در این سیگنال‌ها بسیار مشکل است[۴]. در پنهان‌نگاری هدف افزودن اطلاعات اضافی در رسانه حامل است به‌طوری‌که تغییری در رسانه حامل به وجود نیاید؛ برای تشخیص پنهان‌نگاری می‌بایست برخی ویژگی‌های محیط پوشش و پوشانه استخراج‌شده و از آن‌ها برای تشخیص و تجزیه تحلیل استفاده کرد[۵]. در این مقاله، در ابتدا به بررسی استانداردهای کدگذار صوتی بر پایه پیش‌بینی‌خطی[۱] می‌پردازیم و روش‌های الگوریتم صوتی کدگذار صوتی پیش‌بینی خطی[۲] (LPC) [۶]، پیش‌بینی خطی برانگیخته از کد (CELP[۳])[۷]، پیش‌بینی خطی تحریک مختلط (MELP[۴])[۸] بررسی خواهند شد و پس‌ازآن در این سه استاندارد کدگذار صوتی پنهان‌نگاری صورت خواهد پذیرفت و در ادامه روش پیشنهادی تشخیص پنهان‌نگاری این مقاله شرح داده می‌شود و الگوریتم تشخیص پنهان‌نگاری به روش پیشنهادی بررسی‌شده و ویژگی آماری آزمون اجرا موردبررسی قرار خواهند گرفت، سپس شبکه عصبی هوشمند LVQ موردمحاسبه قرار می‌گیرد. در انتها دقت آشکارسازی روش تشخیص پنهان‌نگاری پیشنهادی با سایر روش‌ها مقایسه شده و مزیت روش پیشنهادی عنوان خواهد شد. در ادامه به تشریح الگوریتم‌های کدگذار صوتی بر پایه پیش‌بینی خطی می‌پردازیم.

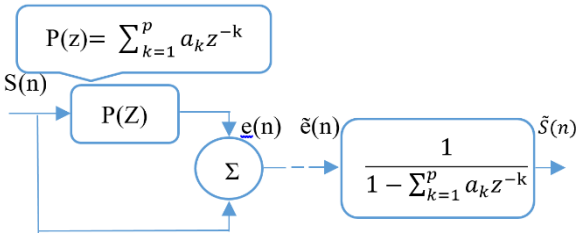
۲- کدگذار صوتی بر پایه پیش‌بینی خطی

کدگذارهای صوت عموماً در نرخ بیت‌های کمتر از ۴/۸ کیلوبیت بر ثانیه استفاده می‌شوند. هدف اصلی کدگذارهای صوت[۵] (برخلاف کدگذارهای موجک[۶] که هدفشان به‌دست آوردن سیگنال گفتار با شبیه‌ترین حالت ممکن به سیگنال اصلی با حداکثر میزان سیگنال به نویز است) ساخت گفتار مصنوعی است که دارای کیفیت مشابه سیگنال اصلی باشد. در حالت کلی سیگنال گفتار دارای دو نوع همبستگی کوتاه‌مدت (همبستگی میان نمونه‌های مجاور یک سیگنال) و همبستگی بلندمدت (همبستگی میان نمونه‌هایی که به فاصله یک Pitch از یکدیگر قرار دارند) است که فیلتر صوتی بیانگر همبستگی کوتاه‌مدت سیگنال گفتار بوده و به‌عبارت‌دیگر این فیلتر خطای تخمین میان نمونه‌های مجاور را کمینه می‌کند. اگر تخمین اُام را با $\tilde{S}(n)$ نشان دهیم آنگاه خواهیم داشت[۹]:

$$\tilde{S}(n) = \sum_{k=1}^{p} a_k - s(n-k) \qquad (۱)$$

$$e(n) - \tilde{S}(n) = s(n) - \sum_{k=1}^{p} a_k - s(n-k) \qquad (۲)$$

شکل ۱ بلوک دیاگرام تخمین‌گر سیگنال صوت را ترسیم می‌کند.



شکل ۱- ساختار کلی بلوک دیاگرام تخمین‌گر سیگنال صوت[۹]

باید توجه داشت که:

- واریانس e(n) خیلی کمتر از تغییرات s(n) است.
- در گیرنده فیلتر موردنظر از جنس IIR است.

پس بایستی ضرایب a_k به نحوی انتخاب شوند که فیلتر پایدار باشد. پس خواهیم داشت[۹]:

$$S(w) = E(w) * \frac{1}{1-\sum_{k=1}^{p} a_k e^{-jkw}} \qquad (۳)$$

که مطابق رابطه‌ی ۳، مقدار E(w) بیانگر فیلتر جریان هوای عبوری از حنجره بوده و $1 - \sum_{k=1}^{p} a_k e^{-jkw}$ بیانگر فیلتر صوت است. می‌دانیم پوش سیگنال گفتار برابر است با[۹]:

$$H(z) = \frac{1}{1-\sum_{k=1}^{p} a_k z^{-k}} \qquad (۴)$$

$$e_n(m) = S_n(m) - \sum_{k=1}^{p} a_k S_n(m-K) \qquad (۵)$$

که در این رابطه $e_{(n)}$ خطای تخمین سمبل mاُام برای فریم nاُام است و داریم[۹]:

$$E_n(m) = \sum e^2(n) \qquad (۶)$$

برای محاسبه a_k باید مقدار E_n حداقل گردد. پس خواهیم داشت[۹]:

1 Linear Predictive
2 Linear Predictive Coding
3 Coded Exited Linear Predictive
4 Mixed-Excitation Linear Predicttive
5 Voice Coders
6 Wavelet Coders

$$E_n = \sum_{m=n-M}^{n+M}[S_n(k) - \sum_{k=1}^{p} a_k S_n(m-K)]^2 \quad (۷)$$

چون[۹]:

$$\rightarrow \sum 2(S_n(m) - \sum_{k=1}^{p} a_k S_n(m-K) * (S_n(m-i)) = 0$$

$$\rightarrow f(x) = \begin{cases} \sum(S_n(m)S_n(m-i) = \sum_k a_k \sum_m S_n(m-K)\, S_n(m-i) \\ 1 \le i \le p \end{cases} \begin{matrix} x < i \\ , \ 1 \le k \le p \end{matrix} \quad (۸)$$

با تعریف دو پارامتر زیر داریم[۹]:

$$\varphi_n(\mathrm{i},\mathrm{k}) \triangleq \sum S_n(m-K)S_n(m-i) \quad (۹)$$

که از روابط ۸ و ۹ نتیجه می‌گیریم[۹]:

$$\rightarrow \sum_{k=1}^{p} a_k \varphi_n(\mathrm{i},\mathrm{k}) = \varphi_n(\mathrm{i},0) 1 \le i \le p$$

$$a_1\varphi_n(1,1) + a_n\varphi_n(1,2) + \cdots a_p\varphi_n(1,\mathrm{p}) = \varphi_n(1,0) \quad (۱۰)$$

$$\underbrace{\begin{bmatrix} \varphi_n(1,1) & \cdots & \varphi_n(1,\mathrm{p}) \\ \vdots & \ddots & \vdots \\ \varphi_n(\mathrm{p},1) & \cdots & \varphi_n(\mathrm{p},\mathrm{p}) \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} a_1 \\ \dots \\ a_p \end{bmatrix}}_{\mathbf{B}} = \underbrace{\begin{bmatrix} \varphi_n(1,0) \\ \dots \\ \varphi_n(\mathrm{p},0) \end{bmatrix}}_{\mathbf{C}}$$

درنتیجه خواهیم داشت[۹]:

$$A = B^{-1}C \quad (۱۱)$$

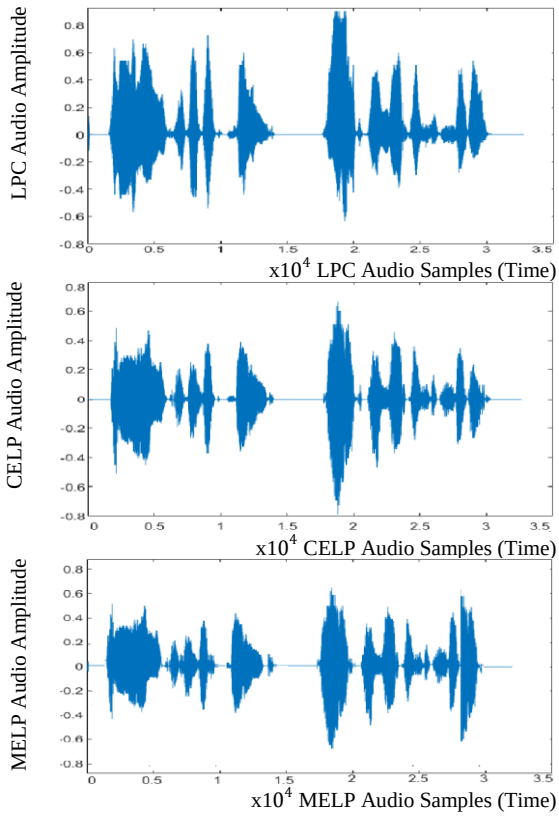
با توجه به اینکه در این مقاله تشخیص پنهان‌نگاری در سه کدگذار صوتی LPC ، CELP و MELP بررسی خواهند شد، این ۳ الگوریتم کدگذار صوتی در ادامه بررسی می‌شوند. الگوریتم کدگذار صوتی پیش‌بینی خطی (LPC)، یک روش فشرده‌سازی صوت است که برای کاهش حجم فایل‌های صوتی با حفظ کیفیت استفاده می‌شود. استفاده از این کدگذار امکان فشرده‌سازی مؤثر در صوت با استفاده از روش‌های آماری و پیش‌بینی را فراهم می‌کند. این روش از تحلیل زمانی سیگنال صوتی استفاده می‌کند و سعی می‌کند با استفاده از تجزیه سیگنال به قطعه‌های کوچک‌تر، ویژگی‌های آماری و طیفی سیگنال را استخراج کند. در فرآیند کدگذاری، سیگنال صوتی به قطعات کوچک‌تر تقسیم می‌شود و سپس برای هر قطعه از سیگنال، یک مدل آماری پیش‌بینی خطی ایجاد می‌شود. این مدل توانایی پیش‌بینی نمونه‌های بعدی سیگنال صوتی را بر اساس نمونه‌های قبلی دارد. با استفاده از ضرایب پیش‌بینی خطی، سیگنال صوتی با دقت بالا بازسازی می‌شود. مزیت اصلی این کدگذار در فشرده‌سازی صوتی، کاهش حجم فایل‌های صوتی است که با استفاده از آن می‌توان اطلاعات آماری و چگالی طیفی سیگنال صوتی را با دقت بالا بازسازی کرد و درنتیجه، حجم فایل‌های صوتی را به‌طور قابل‌توجهی کاهش داد. از این کدگذاری برای استفاده در سیستم‌های ارتباطی، ارسال صوت در شبکه‌ها و فشرده‌سازی صوت در ذخیره‌سازی‌های مختلف، از جمله فرمت‌های فایل صوتی، استفاده می‌شود[۷]. پیش‌بینی خطی برانگیخته از کد (CELP)، یک کدگذار صوتی است که هدف اصلی آن کاهش حجم فایل‌های صوتی با حفظ کیفیت صوت است. این کدگذار بر اساس الگوریتم‌های فشرده‌سازی، اطلاعات غیرضروری و تکراری در سیگنال صوتی را حذف می‌کند. این الگوریتم معمولاً در تبدیل موجک، تحلیل زمان–فرکانس و فشرده‌سازی بر پایه مدولاسیون بهره می‌برد. مزایای استفاده از این کدگذار شامل کاهش فضای ذخیره‌سازی موردنیاز برای فایل‌های صوتی، امکان انتقال سریع‌تر فایل‌های صوتی از طریق شبکه و امکان استفاده راحت‌تر از فایل‌های صوتی در دستگاه‌های قابل‌حمل است[۸]. پیش‌بینی خطی تحریک مختلط (MELP) یک استاندارد کدگذاری صوتی است که برای فشرده‌سازی صوت استفاده می‌شود. این استاندارد برای کاهش حجم داده‌های صوتی طراحی‌شده است. از این کدگذار عموماً برای انتقال صوت در شبکه‌های تلفن همراه، نظامی و کاربردهایی که نیاز به فشرده‌سازی صوت دارند، استفاده می‌شود. از ویژگی‌های MELPمی‌توان به کیفیت بالا در فشرده‌سازی صوت، نیاز به پردازش کم، پهنای باند کم برای انتقال صوت اشاره کرد.

شکل۲- الگوریتم تشخیص تصویر

۳-پنهان‌نگاری در LPC، CELP و MELP

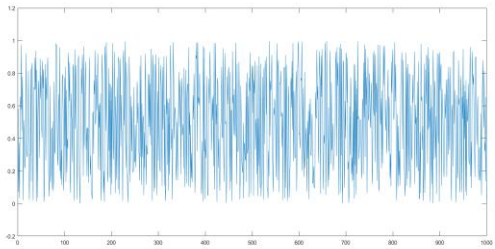
در این مقاله در ابتدا ۱۰۰ صوت انتخاب شد، این صوت توسط نرم‌افزار متلب در قالب کدگذارهای صوتی LPC، CELP و MELP قرار گرفت که درنتیجه ۳۰۰ صوت تولید شد و در ادامه با استفاده از روش بیت کم‌ارزش مطابق شکل ۳ از یک دنباله تصادفی برای پنهان–نگاری در صوت موردنظر استفاده‌شده است. شکل ۳ هیستوگرام خروجی طیف گفتار شبیه‌سازی‌شده با استفاده از نرم‌افزار متلب در سه استاندارد

کدگذار صوتی LPC، CELP و MELP است.



شکل۳- خروجی‌های هیستوگرام طیف گفتار (به ترتیب) LPC، CELP و MELP

در این مقاله مطابق شکل ۴ از یک دنباله تصادفی برای پنهان کردن در اصوات موردنظر استفاده‌شده است.



شکل٤- رسانه مورداستفاده پنهان‌نگاری

الگوریتم اولیه پنهان‌نگاری به روش بیت کم‌ارزش بدین شرح است:

- یک دنباله تصـــادفی برای پنهان کردن اطلاعات درون گفتار انتخاب می‌شود.
- دنباله تصـــادفی درون بیت کم‌ارزش صـــوت موردنظر جاســـازی می‌شود.
- رسانه‌ی حامل ایجاد می‌شود.

با توجه به شباهت بالای هیستوگرام خروجی صوت رسانه حامل و صوت بدون پنهان‌نگاری امنیت بالای این روش مشهود است.

۳-۱-استخراج ویژگی‌ها

با توجه به اینکه در هدف پنهان‌نگاری افزودن پیام مخفی درون رسانه حامل به طوریست که تغییری در این رسانه به وجود نیاید ازاین‌رو برای تشخیص پنهان‌نگاری می‌بایست ویژگی‌های رسانه حامل استخراج‌شده و از آن‌ها برای تشخیص و تجزیه تحلیل اطلاعات استفاده کرد. در روش پیشنهادی تشخیص پنهان‌نگاری این مقاله از آزمون آماری اجرا برای استخراج ویژگی‌های داده‌های پنهان استفاده خواهد شد.

۳-۲-ویژگی آزمون اجرا[1]

آزمون اجرای تصادفی یک آزمون آماری برای تشخیص تصادفی بودن داده‌هاست که این روش جایگزینی برای آزمایش همبستگی خودکار داده‌هاست (برای تایید اینکه آیا داده‌ها با مقدار تاخیر همبستگی دارند یا خیر) در صورتی‌که که همبستگی بین نمونه‌ها وجود نداشته باشد می‌توان نتیجه گرفت داده‌ها تصادفی هستند. با توجه به اینکه این سه کدگذار رفتاری تصادفی دارند (همبستگی بین نمونه‌ها وجود ندارد) برای بررسی تصادفی بودن (غیر همبستگی) یک رشته بیت از این آزمون استفاده می‌شود. تمرکز این آزمون روی تعداد کل اجراها در دنباله است با این تعریف که یک اجرا، دنباله‌ای از بیت‌های برابر و بی‌وقفه (بدون تغییر حالت از یک به صفر یا برعکس) باشد[۱۱]. یک آزمون اجرا از طول k شامل دقیقا k بیت برابر است و با بیتی با مقدار مخالف، قبل و بعد آن، مرزبندی‌شده است. هدف آزمون اجرا این است که تعیین کند آیا تعداد اجراهای یک‌ها و صفرهای طول‌های مختلف با آنچه که از یک دنباله تصادفی انتظار هست برابر است یا خیر. این آزمون تعیین می‌کند که آیا نوسان بین صفر و یک بسیار سریع یا بسیار کند است، در ادامه به بررسی تابع فراخوانی برای این آزمون خواهیم پرداخت. اگر n طول رشته بیتی باشد که قصد داریم میزان تصادفی بودن آن را بررسی کنیم، تابع فراخوانی که در این آزمون برای محاسبه‌ی میزان تصادفی بودن پیشنهاد شده به صورت رابطه ۱۲ تعریف می‌شود [۱۱]:

$$A = \mathrm{RUNS}(n) \qquad (۱۲)$$

که در این رابطه n طول رشـته بیت اسـت، و ورودی دیگری که توسط این تابع مورد استفاده قرار می‌گیرد نیز ε است که بیانگر دنباله ای از بیت‌هاسـت که توسـط یک الگوریتم RNG یا PRNG [2] مطابق رابطه ۱۴تولید می‌شود [۱۱]:

[1] Runs Test
[2] PRNG (pseudorandom number generator), RNG(random number generator)

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_2}, \dots, \boldsymbol{\varepsilon_n} \quad (۱۴)$$

که مطابق رابطه‌ی ۱۴، $\boldsymbol{\varepsilon_1}$ معرف اولین بیت و $\boldsymbol{\varepsilon_n}$ معرف nامین بیتِ مورد ارزیابی است. همچنین، این رشته بیت‌ها، برای نمونه‌های تصادفی از رشته‌های تصادفیِ تولید شده‌اند و برای بررسی رشته‌های غیرتصادفی از رشته‌های بیت یک فایل استفاده می‌شود. برای انجام هر یک از این آزمون‌ها یک آماره آزمون خواهیم داشت که مقدار $P-value$ برای آن محاسبه می‌شود؛ آماره این توزیع به صورت زیر تعریف می‌شود [۱۱]:

$$B = V_n(obs) \quad (۱۵)$$

که در رابطه ۱۵، $V_n(obs)$ آماره آزمون است که بیانگر تعداد کل اجراها (یعنی تعداد کل اجراهای صفر و تعداد کل اجراهای یک) را در طول کل n بیت نشان می‌دهد. توزیع مرجع برای آماره این آزمون نیز (χ^2) است.

٤-الگوریتم محاسبه ویژگی در آزمون آزمون اجرا

در ادامه گام‌های مورد نیاز برای محاسبه‌ی ویژگی مورد نظر برای این آزمون بررسی خواهد شد. (همچنین باید به این نکته اشاره شود که آزمون اجراها یک آزمون فرکانس را به عنوان پیش نیاز انجام می‌دهد):

- گام اول: محاسبه نسبت آزمون π در دنباله ورودی برای "یک" ها با محاسبه‌ی رابطه ۱۶ [۱۱]:

$$\pi = \frac{\sum_j \varepsilon_j}{n} \quad (۱۶)$$

- گام دوم: آماره‌ی آزمون را با استفاده از رابطه ۱۷ محاسبه می‌شود [۱۱]:

$$\begin{cases} \varepsilon_k = \varepsilon_{k+1} & \rightarrow r(k) = 0 \\ in\ other\ wise \rightarrow & r(k) = 1 \end{cases} \quad (۱۷)$$

- گام سوم: محاسبه رابطه ۱۹ [۲۰]:

٤-۱-تشخیص پنهان‌نگاری

در کنار گسترش روش‌های مختلف پنهان‌نگاری در صوت، روش‌های متنوع تشخیص پهان‌نگاری ایجادشده است که هدف آن استفاده از ویژگی‌های صوت و الگوریتم‌های پردازشی برای کشف اطلاعات پنهان‌شده است. توانایی تشخیص دیتا در صوت پنهان‌شده تحت تأثیر فاکتورهای مختلفی قرار دارد که ازجمله می‌توان به دیتای جاسازی‌شده، فرمت پوشانه انتخابی و روش جاسازی دیتای موردنظر در پوشانه اشاره کرد[۱۰].

٤-۲-دقت آشکارسازی

برای ارزیابی عملکرد از معیار دقت آشکارسازی[1] استفاده خواهد اما برای دانستن این معیار لازم است معیارهای TPR و TNR مطابق روابط (۲۹ و ۲۸) تعریف شوند که TPR به عنوان نسبت، تعداد موارد مثبت که به درستی به عنوان مثبت صحیح (TP) طبقه بندی شده اند وتعداد موارد مثبت که به اشتباه به عنوان منفی کاذب طبقه بندی شده اند (FN) هستند، تعداد موارد مثبت واقعی مطابق رابطه ۲۰ محاسبه می‌شود[۲۲]:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{(\mathrm{TP}+\mathrm{FN})} \quad (۲۰)$$

TNR به عنوان نسبت تعداد موارد منفی، که به درستی به عنوان منفی طبقه بندی شده اند (TN) و تعداد موارد منفی که به اشتباه به عنوان مثبت کاذب طبقه بندی شده اند (FP) است، تعداد موارد منفی صحیح مطابق رابطه ۲۱ محاسبه می‌شود[۲۲]:

$$\mathrm{TNR} = \frac{\mathrm{TN}}{(\mathrm{TN}+\mathrm{FP})} \quad (۲۱)$$

برای ارزیابی عملکرد از معیار دقت آشکارسازی استفاده خواهد شد که در بازشناسی الگو، بازیابی اطلاعات و طبقه‌بندی آماری کاربرد فراوانی دارد؛این معیار از رابطه ۲۲ محاسبه می‌شود[۲۲].

$$\mathrm{TRP} = \frac{(\mathrm{TP}+\mathrm{TN})}{(\mathrm{TP}+\mathrm{TN}+\mathrm{FN}+\mathrm{FP})} \times 100\% \quad (۲۲)$$

که صورت کسر رابطهِی ۲۸ موارد صحیح دسته‌بندی شده و مخرج، تمام موارد دسته‌بندی شده است.

٤-۳-نرخ خطای بیت

نرخ خطای بیت به عنوان درصد مواردی که به اشتباه نسبت به

$$P - value = \mathrm{erfc}\left(\frac{|V_n(obs) - 2n\pi(1-\pi)|}{2\sqrt{2n}\,\pi(1-\pi)}\right) \quad (۱۹)$$

تعداد کل موارد در مجموعه داده طبقه بندی شده اند تعریف شده و از رابطه‌ی ۲۳ محاسبه می‌گردد.

$$\mathrm{TRP} = \frac{(\mathrm{FN}+\mathrm{FP})}{(\mathrm{TP}+\mathrm{TN}+\mathrm{FN}+\mathrm{FP})} \times 100\% \quad (۲۳)$$

که صورت کسر این رابطه، موارد اشتباه دسته‌بندی شده است.

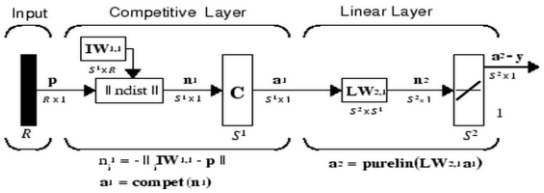
٤-٤-شبکه عصبی برای طبقه بندی داده‌ها

در این مقاله از شبکه عصبی رقمی‌ساز بردار یادگیری[2] که یک روش دسته‌بندی الگو است، که در آن هر واحد خروجی بیانگر دسته یا گروه خاصی از الگوهای ورودی است. در این شبکه‌ها موقعیت واحدهای خروجی با تنظیم وزن‌های آنها از طریق آموزش با نظارت در حین آموزش مشخص می‌شود. در حقیقت هرکدام از خروجی‌ها نمایش‌دهنده

[1] Detection Accuracy

[2] Learning Vector Quntization

یک کلاس بوده و هرکدام توسط بردار وزن آن کلاس مشخص می‌شود. بردار وزن هرکدام از کلاس‌ها توسط یکی از مجموعه‌های آموزشی مقداردهی اولیه‌شده و سپس توسط الگوریتم‌های یادگیری با نظارت بهینه می‌شود. پس از آموزش، شبکه بردارهای ورودی را با اختصاص دادن هرکدام از آنها به یک واحد خروجی که بردار وزن آن به بردار ورودی نزدیک‌تر است، دسته‌بندی می‌کند[۱۳]. مطابق شکل ۵ این شبکه از دو لایه تشکیل شده که نرون‌های لایه دوم به تعداد کلاس‌ها هستند. ابتدا وزن‌های اولیه به هر خروجی داده می‌شود که معرف بردار وزن آن کلاس است. سپس داده‌های آموزشی با کلاس مربوطه به شبکه داده می‌شود. سپس وزن‌های هر مرکز با داده‌های آموزشی برچسب دار تحت شرایط هم‌کلاس بودن یا نبودن بهینه می‌شود. به‌این‌ترتیب در صورت هم‌کلاس بودن به آن ورودی نزدیک و در صورت غیر هم‌کلاس بودن از آن ورودی دور می‌شود.



شکل ۵- ساختار شبکه عصبی LVQ پیشنهادی تحت آموزش

با توجه به روابط زیر الگوریتم LVQ داده‌ها را از یک فضای نامتناهی به یک فضای محدود نگاشت می‌کند. طبق رابطه ۲۴ نرون برنده نرونی خواهد بود که کمترین فاصله را با داده دارد. در این صورت وزن آن نرون تصحیح شده و به‌اصطلاح به سمت داده حرکت می‌کند [۱۳]:

$$\left.\begin{matrix}\forall x \in X \\ X => \infty\end{matrix}\right\} \; i \in I, m_i \in X, I < \infty \qquad (۲۴)$$

اگر مطابق رابطه ۲۵ داشته باشیم [۱۳]:

$$\begin{cases} M = (m1, m2, m3 \dots mk) \\ X: \; x(0), x(1), x(2), x(3), x(4), \dots \end{cases} \qquad (۲۵)$$

که M بردار مراکز خوشه‌ها و X بردار داده آموزش است.

در آن صورت تابع هزینه به‌صورت رابطه ۲۶ تعریف می‌شود [۲۳]:

$$E = \int || x - m_C ||^2 \, p(x) \, dx \qquad (۲۶)$$

در این رابطه m_C مقدار برنده شده و $p(x)$ توزیع احتمالی است که در رابطه ۲۷ داریم [۱۳]:

$$c = arg_i \min (||x - m_i||)^2 \qquad (۲۷)$$

با حل این معادله رابطه ۲۸ را خواهیم داشت [۱۳]:

$$m_i(t+1) = m_i(t) + \Delta m_i(t) \qquad i = 1,2,3 \dots k \qquad (۲۸)$$

که در رابطه ۲۹ داریم [۱۳]:

$$\Delta m_i(t) = \begin{cases} \alpha(t)[x(t) - m_i(t)] & i = c \\ 0 & i \neq c \end{cases} \qquad (۲۹)$$

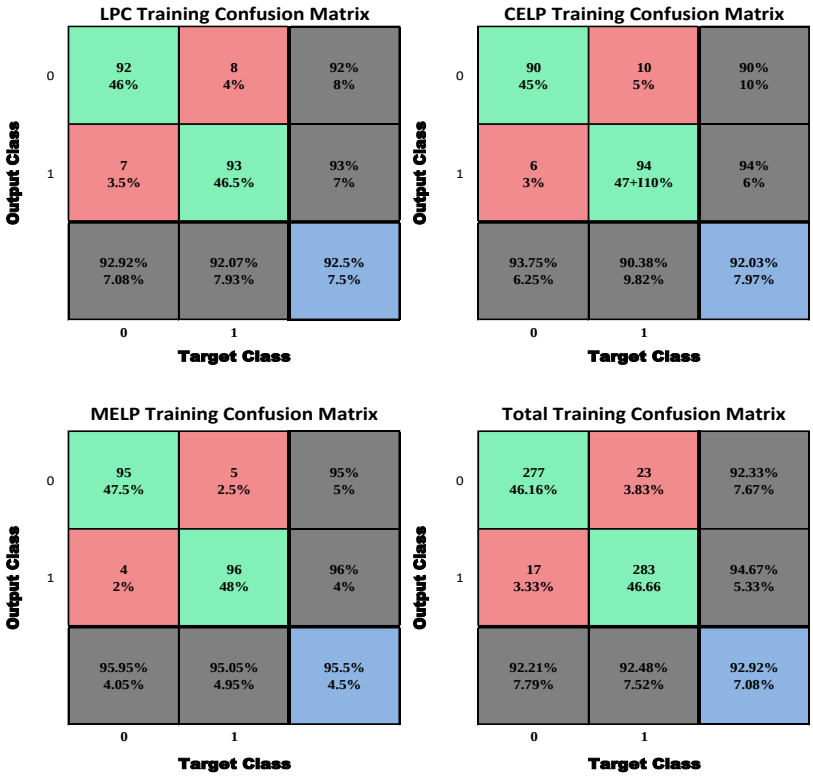
این معادله برای یادگیری نظارت‌شده به‌صورت رابطه ۳۰ تعریف می‌شود [۲۳]:

$$m_i(t+1) = m_i(t) + \Delta m_i(t) \qquad (۳۰)$$

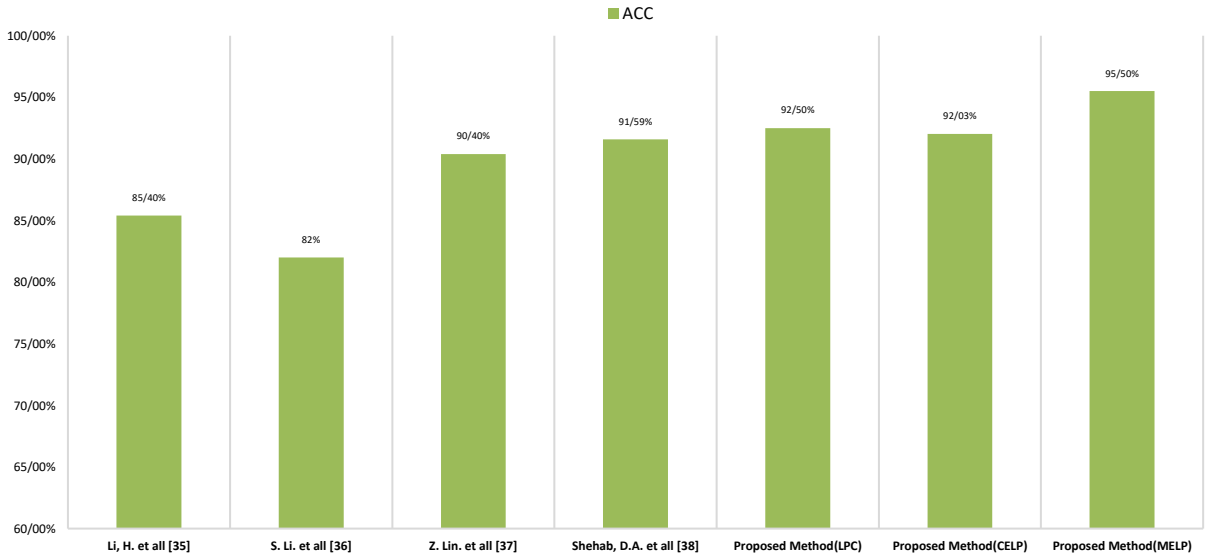
به‌نحوی‌که در رابطه‌ی ۳۱ داریم [۱۳]:

$$\Delta m_i(t) = \begin{cases} \alpha(t)[x(t) - m_i(t)] & i = c \\ -\alpha(t)[x(t) - m_i(t)] & i = c \\ 0 \;\; i \neq c & \end{cases} \qquad (۳۱)$$

در رابطه ۳۷ و در حالت اول m_i و $x(t)$در یک کلاس هستند ولی در حالت دوم m_i و $x(t)$ در یک کلاس نیستند. بعد از یادگیری، شبکه LVQورودی را به کلاسی که برداری با نزدیک‌ترین فاصله به آن باشد، نسبت می‌دهد. در این مقاله از شبکه عصبی LVQبه‌عنوان تحلیل‌کننده استفاده‌شده است. شبکه عصبی چندلایه یک مولد شبکه عصبی مصنوعی با تغذیه مستقیم است که دسته‌های داده‌های ورودی را روی دسته خروجی مناسب ترسیم می‌کند. شبکه مورد آزمایش دارای ۴ لایه است که یک‌لایه ورودی، دولایه پنهان و یک‌لایه خروجی دارد. تعداد نرون‌های لایه ورودی دو (صدای با پنهان‌نگاری و بدون پنهان‌نگاری)، تعداد ویژگی‌ها برابر ۱(ویژگی اجرا)، و تعداد نرون‌های لایه خروجی ۲(تشخیص صوت با پنهان‌نگاری، یا بدون پنهان‌نگاری) است که مشخص می‌کند که آیا صوت گفتار با پنهان‌نگاری است یا خیر. شبکه عصبی (LVQ) طبقه‌بندی کننده داده در این مقاله در ابتدا، ویژگی‌های مبتنی بر پیش‌بینی خطی را استخراج کرده و هر پوشش را یک و استگو را صفر برچسب‌گذاری می‌کند. سپس از ابزار تشخیص الگو به عنوان مدل تشخیص پنهان‌نگاری استفاده می‌کند، که در این مقاله ۳۰۰ پوشش و ۳۰۰ استگو تولید شده و توسط LSB آزمایش شده است. شکل ۶ ماتریس ابهام شبکه عصبی مصنوعی(که معیاری برای طبقه‌بندی‌کننده داده‌هاست) را نشان می‌دهد. همانطور که در ماتریس ابهام شکل ۶ دیده می‌شود بهترین دقت آشکارسازی تشخیص پنهان‌نگاری روش پیشنهادی این مقاله به ترتیب مربوط به MELP، LPC و سپس CELP است، در ادامه و در شکل ۷ به مقایسه طبقه‌بندی TP، TN و دقت آشکارسازی روش پیشنهادی با سایر مقالات می‌پردازیم. در شکل ۷ میزان طبقه بندی مثبت صحیح، طبقه بندی منفی صحیح و دقت آشکارسازی روش پیشنهادی در کنار سایر روش ها آورده شده است. در ادامه و در جدول ۱ بررسی جامعی از میزان FN، FP، TN،TP و دقت آشکارسازی روش پیشنهادی و سایر روش ها آورده شده است.



شکل٦- دقت طبقه بندی شبکه عصبی LVQ با استفاده از معیار SNR برای ویژگی آزمون اجرا



شکل۷- مقایسه دقت آشکارسازی روش پیشنهادی در برابر سایر روش‌ها

جدول۱- مقایسه ارزیابی روش پیشنهادی تشخیص پنهان‌نگاری با سایر روش‌ها

روش تشخیص پنهان نگاری		FP	FN	TP	TN	ACC
Li, H. et all [35]		٪۱۴.۶	٪۱۴.۶	٪۸۵.۴	٪۸۵.۴	٪۸۵.۴
S. Li. et all [36]		٪۸	٪۲۸	٪۷۲	٪۹۲	٪۸۲
Z. Lin. et all [37]		٪۱۰.۲	٪۸.۱۶	٪۹۱.۸۴	٪۸۹.۸	٪۹۰.۴
Shehab, D.A. et all [38]		٪۲.۰۶	٪۱۴.۷۶	٪۸۵.۲۴	٪۹۷.۹۴	٪۹۱.۵۹
Proposed Method: (Detection of Steganography Using LVQ Neural Network)	LPC	٪۷.۰۸	٪۷.۹۳	٪۹۲.۰۷	٪۹۲.۹۲	٪۹۲.۵
	CELP	٪۴.۲۶	٪۹.۴۴	٪۹۰.۵۶	٪۹۵.۷۴	٪۹۲.۰۳
	MELP	٪۴.۰۵	٪۴.۹۵	٪۹۵.۰۵	٪۹۵.۹۵	٪۹۵.۵
Averraged Proposed Method		٪۵.۱۴	٪۷.۵	٪۹۴.۸۶	٪۹۲.۵	٪۹۳.۵۹

۵-نتیجه‌گیری

در این مقاله روشی جدید در تشخیص پنهان‌نگاری در سه استاندارد کدگذار صوتی LPC، CELP و MELP ارائه شد. برای جاسازی داده‌های مخفی از بیت کم‌ارزش استفاده شد که نتایج نشان‌دهنده مقاومت بالای این روش بود. همچنین برای تحلیل صوت از ویژگی آزمون اجرا استفاده‌شد و پس از بررسی این مقادیر در صوت پوشش و استگو، ویژگی‌های متفاوت بین این صوت استخراج و برای آموزش شبکه عصبی هوشمند بدون نظارت رقمی‌ساز بردار یادگیر (LVQ) استفاده شد. مرحله طبقه‌بندی با استفاده از لایه‌های این شبکه عصبی انجام و الگوریتم پیشنهادی برای این اصوات تست‌شد. همانطور که در جدول ۱ دیده شد میزان میانگین مربوط به FN، FP، TP، TN و دقت آشکارسازی پنهان‌نگاری روش پیشنهادی به ترتیب برابر با٪۵.۱۴، ٪۷.۵۰، ٪۹۴.۸۶، ٪۹۲.۵۰ و ٪۹۳.۵۹ است؛ که نشان‌دهنده برتری محسوس روش پیشنهادی در برابر سایر روش‌ها است.

٦-مراجع

[1] J. Jezdimirović, N. Pekez and J. Kovačević, "Security enhancement of LSB-based audio steganography method," 2023 Zooming Innovation in Consumer Technologies Conference (ZINC), Novi Sad, Serbia, 2023, pp. 77-82, doi: 10.1109/ZINC58345.2023.10174020.

[2] M. H. Noor Azam, F. H. Mohd Ridzuan, and M. N. S. Mohd Sayuti, "Optimized Cover Selection for Audio Steganography Using Multi-Objective Evolutionary Algorithm", JICT, vol. 22, no. 2, pp. 255–282, Apr. 2023.

[3] Talati, S., EtezadiFar, P., Hassani Ahangar, M. R., Molazade, M., Investigation of Steganography Methods in Audio Standard Coders: LPC, CELP, MELP. Majlesi Journal of Telecommunication Devices, 2023; 12(1): 7-15.doi: 10.30486/mjtd.2022.695928.

[4] Etezadifar, P., Talati, S., Hassani Ahangar, M. R., Molazade, M. Comparison of Standards Digital Audio Encoders LPC, CELP, and MELP based on the Quality and Complexity of the Content in the Transmitted. Majlesi Journal of Telecommunication Devices, 2023; doi: 10.30486/mjtd.2023.1988165.1035.

[5] Talati, S. and Reza, R., 2023. Presenting a New Method of Image Steganalysis Based on MLP Neural Network. Passive Defense Quarterly, 14(4), pp.21-32

[6] Standard, Federal. "1015, Telecommunications: Analog to Digital Conversion of Radio Voice By 2400 Bit/Second Linear Predictive Coding." National Communication System Office of Ttchnology and Standards (1984).

[7] M. Schroeder and B. Atal, "Code-excited linear prediction (CELP): High-quality speech at very low bit rates," ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1985, pp. 937-940, doi: 10.1109/ICASSP.1985.1168147.

[8] A. V. McCree and T. P. Barnwell, "A mixed excitation LPC vocoder model for low bit rate speech coding". IEEE Transactions on Speech and Audio Processing, vol. 3, no. 4, pp. 242-250, July 1995.

[9] Bishnu Atal "The History of Linear Prediction ". ICASSP '78. IEEE Signal Processing Magazine, vol.23, no2, march 2006. 154-161.

[10] Y. Ren, D. Liu, C. Liu, Q. Xiong, J. Fu and L. Wang, "A Universal Audio Steganalysis Scheme Based on Multiscale Spectrograms and DeepResNet," in IEEE Transactions on Dependable and Secure Computing, vol. 20, no. 1, pp. 665-679, 1 Jan.-Feb. 2023, doi: 10.1109/TDSC.2022.3141121.

[11] Bujang MA, Sapri FE. An Application of the Runs Test to Test for Randomness of Observations Obtained from a Clinical Survey in an Ordered Population. Malays J Med Sci. 2018 Jul; 25(4):146-151.

[12] Good IJ. The serial test for sampling numbers and other tests for randomness. Mathematical Proceedings of the Cambridge Philosophical Society. 1953; 49(2):276-284. doi:10.1017/S030500410002836X

[13] Mohammed J. Alhaddad, Monagi H. Alkinani, Mohammed Salem Atoum and Alaa Abdulsalm Alarood. "Evolutionary Detection Accuracy of Secret Data in Audio Steganography for Securing 5G-Enabled Internet of Things," Symmetry 2020, 12, 2071; doi: 10.3390/ sym12122071.

[14] Talati, Saeed. Mohamadreza Hasani Ahangar, ".Analysis, Simulation and Optimization of LVQ Neural Network Algorithm and Comparison with SOM", MJTD, 2020, 10.1.

[15] A. V. McCree and T. P. Barnwell, "A mixed excitation LPC vocoder model for low bit rate speech coding," in IEEE Transactions on Speech and Audio Processing, vol. 3, no. 4, pp. 242-250, July 1995.

[16] J. J. D. van Schalkwyk, D. J. Joubert and J. G. van der Linde, "Linear predictive speech coding at 2400 b/s," in Transactions of the South African Institute of Electrical Engineers, vol. 84, no. 3, pp. 146-152, June 1993.

[17] Weiran Lin, Soo Ngee Koh and Xiao Lin, "Mixed excitation linear prediction coding of wideband speech at 8 kbps," 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100), 2000, pp. II1137-II1140 vol.2, doi: 10.1109/ICASSP.2000.859165.

[18] Bujang MA, Sapri FE. An Application of the Runs Test to Test for Randomness of Observations Obtained from a Clinical Survey in an Ordered Population. Malays J Med Sci. 2018 Jul; 25(4):146-151. doi: 10.21315/mjms2018.25.4.15. Epub 2018 Aug 30. PMID: 30914857; PMCID: PMC6422539.

[19] Muhammad Asad; Junaid Gilani; Adnan Khalid, An enhanced least significant bit modification technique for audio steganography, International Conference on Computer Networks and Information Technology, July 2011.

[20] L. Gang, A. N. Akansu and M. Ramkumar, "Mp3 resistant oblivious steganography," in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Salt Lake City, USA, 2001, pp. 1365–8.

[21] H. Liu, J. Liu, R. Hu, X. Yan and S. Wan, "Adaptive audio steganography scheme based on wavelet packet energy," in IEEE International Conference on High Performance and Smart Computing (HPSC), 2017.

[22] Talati, Saeed, Hassani Ahangar, Mohammad Reza. (2021). "Radar Data Processing Using a Combination of Principal Component Analysis Methods and SelfOrganized and Digitizing Learning Vector Neural Networks", Electronic and Cyber Defense, 9 (2), pp. 1- 7.

[23] Talati, S., Etezadifar, P. (2020). 'Providing an Optimal Way to Increase the Security of Data Transfer using Watermarking in Digital Audio Signals', Majlesi Journal of Telecommunication Devices, 9(1), pp. 35-46.

[24] Etezadifar, P., Talati, S., Hassani Ahangar, M. R., Molazade, M. (2023). 'Investigation of Steganography Methods in Audio Standard Coders: LPC, CELP, MELP', Majlesi Journal of Telecommunication Devices, doi: 10.30486/mjtd.2022.695928.

[25] Hashemi, Seyed & Barati, Shahrokh & Talati, S. & Noori, H. (2016). A genetic algorithm approach to optimal placement of switching and protective equipment on a distribution network. Journal of Engineering and Applied Sciences. 11. 1395-1400.

[26] Hashemi, Seyed & Abyari, M. & Barati, Shahrokh & Tahmasebi, Sanaz & Talati, S. (2016). A proposed method to controller parameter soft tuning as accommodation FTC after unknown input observer FDI. Journal of Engineering and Applied Sciences. 11. 2818-2829.

[27] O. Sharifi-Tehrani and S. Talati. (2017) "PPU Adaptive LMS Algorithm, a Hardware-Efficient Approach; a Review on", Majlesi Journal of Mechatronic Systems, vol. 6, no. 1.

[28] S. Talati, A. Rahmati, and H. Heidari. (2019) "Investigating the Effect of Voltage Controlled Oscillator Delay on the Stability of Phase Lock Loops", MJTD, vol. 8, no. 2, pp. 57-61.

[34] Talati, Saeed, Hassani Ahangar, Mohammad Reza. (2021). “Radar Data Processing Using a Combination of Principal Component Analysis Methods and Self-Organized and Digitizing Learning Vector Neural Networks”, Electronic and Cyber Defense, 9 (2), pp. 1-7.

[35] S.-b. Li, H.-z. Tao, and Y.-f. Huang, “Detection of quantization index modulation steganography in g. 723.1 bit stream based on quantization index sequence analysis,” Journal of Zhejiang University SCIENCE C , vol. 13, no. 8, pp. 624–634, 2012.

[36] S. Li, Y. Jia, and C.-C. J. Kuo, “Steganalysis of qim steganography in low-bit-rate speech signals,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 5, pp. 1011–1022, 2017.

[37] Z. Lin, Y. Huang, and J. Wang, “Rnn-sm: Fast steganalysis of voip streams using recurrent neural network,” IEEE Transactions on Information Forensics & Security, vol. PP, no. 99, pp. 1–1, 2018.

[38] Shehab, D.A. and Alhaddad, M.J., 2022. Comprehensive survey of multimedia steganalysis: Techniques, evaluations, and trends in future research. Symmetry, 14(1), p.117.

[29] Saeed Talati, mohamadreza Hasani Ahangar (2020) “Analysis, Simulation and Optimization of LVQ Neural Network Algorithm and Comparison with SOM”, MJTD, vol. 10, no. 1.

[30] Talati, S., & Hassani Ahangar. M. R. (2020) “Combining Principal Component Analysis Methods and Self-Organized and Vector Learning Neural Networks for Radar Data”, Majlesi Journal of Telecommunication Devices, 9(2), 65-69.

[31] Hassani Ahangar, M. R., Talati, S., Rahmati, A., & Heidari, H. (2020). The Use of Electronic Warfare and Information Signaling in Network-based Warfare. Majlesi Journal of Telecommunication Devices, 9(2), 93-97.

[32] Aslinezhad, M., Mahmoudi, O., & Talati, S. (2020). Blind Detection of Channel Parameters Using Combination of the Gaussian Elimination and Interleaving. Majlesi Journal of Mechatronic Systems, 9(4), 59-67.

[33] Talati, S., & Amjadi, A. (2020). Design and Simulation of a Novel Photonic Crystal Fiber with a Low Dispersion Coefficient in the Terahertz Band. Majlesi Journal of Mechatronic Systems, 9(2), 23-28.