

Speech Steganalysis of Least Significant Bits Based on the Percentage of Equal Adjacent Samples

S. Yazdanpanah, M. Kheyrandish*, M. Mosleh

*Department of Computer Engineering, Dezful Branch, Islamic Azad University, Dezful, Iran

(Received: 06/04/2020, Accepted: 26/10/2020)

ABSTRACT

The popularity of audio formats usually attracts the attention of intruders and criminals to use this medium as a cover for establishing their secret communications. The extensive use of this formats, along with various modern techniques, designed for audio steganography, can cause the cyber spaces to be insecure environments. In order to deal with threats, some audio steganalysis techniques have been presented that statistically analyze various audio formats, such as music, MP3, and VoIP, efficiently. Among the presented approaches, combining the techniques of signal processing and machine learning has made possible the creation of steganalyzers that are highly accurate. However, since the statistical properties of audio files differ from purely speech ones, the current steganalysis methods cannot detect speech stego files, accurately. Another issue is the large number of analysis dimensions which increase the implementation cost, significantly. As response to these issues, this paper proposes the percentage of equal adjacent samples (PEAS) feature, as a one-dimensional feature for speech steganalysis. Using a classifier, based on the Gaussian membership function, on stego instances with 50% embedding ratio, the evaluation results for the designed steganalyzer, show a sensitivity of 99.82%. Additionally, it can efficiently estimate the length of a hidden message with the desirable accuracy. Also, the PEAS steganalysis was evaluated on a database, containing classic music instances, and the results show an 81.2% efficient performance.

Keywords: Speech steganalysis, audio steganalysis, digital signal processing, LSB, Steganography

* Corresponding Author Email: kheyrandish@iaud.ac.ir

علمی - پژوهشی

نهان‌کاوی گفتار در بیت‌های کم ارزش بر مبنای درصد نمونه‌های مجاور یکسان

سعید یزدان‌پناه^۱، محمد خیراندیش^{۲*}، محمد مصلح^۳

۱- دانشجوی دکتری، ۲- استادیار، ۳- دانشیار، گروه مهندسی کامپیوتر، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران.

(دریافت: ۱۳۹۹/۰۱/۱۸، پذیرش: ۱۳۹۹/۰۸/۰۵)

چکیده

عمومیت فایل‌های صوتی، اغلب توجه مهاجمین و عناصر مخرب را برای استفاده از این حامل، جهت پوشش‌دهی ارتباطات محرمانه خود جلب می‌نماید. گستردگی استفاده از این قالب‌ها، به‌همراه رویکردهای متعدد و مدرنی که برای نهان‌نگاری در فایل‌های صوتی طراحی شده‌اند، می‌توانند فضای سایبری را به محیطی نا امن بدل نمایند. در راستای مقابله با این تهدیدات، امروزه روش‌های متعدد نهان‌کاوی ابداع شده‌اند که با دقت بالایی قادر به تحلیل آماری قالب‌های مختلف صوتی، مانند MP3 و VoIP هستند. در میان راه‌حل‌های ارائه‌شده، ترکیب روش‌های پردازش سیگنال و یادگیری ماشین، امکان ایجاد نهان‌کاو‌هایی با دقت بسیار بالا را فراهم نموده است. با این وجود، از آنجا که ویژگی‌های آماری فایل‌های صوتی گفتاری متفاوت از نمونه‌های دیگر صوتی است، روش‌های جاری نهان‌کاوی قادر نیستند به شکل مؤثری فایل‌های حامل گفتاری را تشخیص دهند. مشکل دیگر، ابعاد بالای تحلیلی است که به شکل چشمگیری هزینه پیاده‌سازی را افزایش می‌دهد. در پاسخ به مشکلات ذکرشده، این مقاله ویژگی یک‌بعدی "درصد نمونه‌های مجاور یکسان" را به‌عنوان فاکتور جداسازی نمونه‌های نهان‌نگاری شده از پاک مطرح می‌کند. نتایج نشانگر حساسیت ۹۹/۸۲٪ نهان‌کاو طراحی شده با استفاده از دسته‌بند تابع عضویت گاوسی، در نرخ نهان‌نگاری ۵۰٪ است. علاوه بر این، این نهان‌کاو قادر است با دقت مطلوبی حجم پیام مخفی‌شده را تخمین بزند. عملکرد الگوریتم طراحی‌شده بر روی یک پایگاه داده متشکل از نمونه‌های موسیقی کلاسیک نیز ارزیابی شده و نتایج حاکی از کارایی ۸۱/۲٪ آن هستند.

کلیدواژه‌ها: نهان‌کاوی گفتار، نهان‌کاوی صوتی، پردازش سیگنال‌های صوتی، LSB، نهان‌نگاری

۱- مقدمه

استفاده قرار گیرند [۲]. از میان این رسانه‌ها، فایل‌های صوتی بخش قابل توجهی از ترافیک چندرسانه‌ای حال حاضر را تشکیل می‌دهند و در نتیجه، فرصت ویژه‌ای را برای نهان‌نگاری فراهم می‌آورند که در آن، انتقال اطلاعات محرمانه و غیرمجاز توجه کمتری را به خود جلب می‌کند [۴].

با وجود پیامدهای امنیتی، در مقایسه با تصویر، توجه کافی به امنیت رسانه صوتی گفتاری نشده است. برخی ایده‌های استفاده شده در نهان‌کاوی تصویری، پس از اصلاح ساختار الگوریتم، برای پشتیبانی از نهان‌کاوی قالب‌های صوتی نیز استفاده شده‌اند. با این وجود، از آنجا که این نهان‌کاوها از ابتدا بر اساس ویژگی‌های سیگنال‌ها و فایل‌های صوتی طراحی نشده‌اند، قادر به ارائه حساسیت‌های بالا نیستند [۵].

از یک دیدگاه، روش‌های نهان‌کاوی را می‌توان به دو نوع فعال و غیرفعال تقسیم کرد. در روش فعال، هرگاه وجود پیام در حامل شناسایی می‌شود، بدنه فایل به‌گونه‌ای دست‌کاری می‌گردد که اطلاعات گنجانده‌شده در آن، تخریب شده و عملاً از ارسال اطلاعات پیشگیری می‌شود. روش غیر فعال، در چنین مواردی، از

اگرچه رمزنگاری به‌شکل ایده‌آلی مانع دسترسی غیرمجاز به داده‌ها می‌شود، اما قادر به اختفای الگوهای ارتباطی نیست و متخصصین قادر خواهند بود اطلاعات حساسی مانند طول پیام، دفعات و مدت ارتباطات و هویت فرستنده/گیرنده را استخراج کنند [۱]. بعد دیگری از فن‌آوری‌های امنیتی، نهان‌نگاری است که علم اختفای داده‌های محرمانه در بدنه فایل‌های دیجیتال، به شکل غیرمحسوس است. مهم‌ترین اهداف نهان‌نگاری، پیشگیری از تحلیل ترافیک و حفظ حریم خصوصی است. در صورت به‌کارگیری نهان‌نگاری، در سوی دیگر، نیاز به روش‌هایی برای کشف ارتباطات مخفی است که نهان‌کاوی نامیده می‌شود. در حال حاضر، نهان‌کاوی به مهم‌ترین ابزار جرم‌شناسی برای نظارت بر ارتباطات داده‌ای مشکوک و کنترل تروریست‌ها، هکرها و در کل، قانون‌شکنان تبدیل شده است [۲ و ۳].

ویدئو، تصویر و صوت مهم‌ترین رسانه‌های دیجیتالی هستند که می‌توانند به‌عنوان حامل داده‌های محرمانه نهان شده مورد

* رایانامه نویسنده پاسخگو: Kheyrandish@iaud.ac.ir



نیستند به شکل مؤثری اثرات نهان‌نگاری ضربی را تشخیص دهند. بنابراین، در این زیرگروه، مقادیر قدرمطلق لگاریتمی نمونه‌های داده شده صوتی به‌عنوان عامل تشخیص استفاده می‌شوند [۱].

$$s[n] = c[n](1 + m[n]) \quad (۱)$$

$$e[n] = c[n].m[n] \quad (۲)$$

گروه الگوریتم‌های جهانی شامل دو دسته از روش‌هاست: کالیبره و غیر کالیبره. اولین چالش روش اول، تشخیص ویژگی‌هایی است که به‌جای وابستگی به سیگنال‌های اصلی، وابسته به پیام نهان شده هستند [۱]. تخمین میزبان خود-مولد اولین زیرشاخه روش‌های کالیبره شده است. مهم‌ترین روش‌ها در این زیرشاخه، از شاخص‌های کیفیت صوتی^۳ (AQM)، فاصله Hausdorff، انتشار گاوسی عمومی و نویز مدل‌سازی شده بر اساس مدل ترکیبی گاوسی استفاده کرده‌اند [۱]. دومین زیرگروه، شامل الگوریتم‌های منبع ثابت است که در آن، x و y به‌ترتیب نشانگر سیگنال‌های میزبان و حامل هستند [۱۲]. تابعی $M(x, y)$ تابعی است که اختلافات بین دو سیگنال ذکر شده را اندازه می‌گیرد و قابل تعمیم به تمامی سیگنال‌ها می‌باشد. سومین زیرگروه، خطی، مدل فضای میزبان را تخمین می‌زند [۱۳] و در صورتی که مدل چندبعدی فضای میزبان شناخته شده باشد، قابل استفاده است. آخرین زیرگروه از روش‌های کالیبره، روش نهان‌نگاری مجدد است که سیگنال‌های داده شده را مجدداً نهان‌نگاری کرده و از آنها برای تشخیص نهان‌نگاری بهره می‌برد [۱۴]. ویژگی مشترک تمامی روش‌های کالیبره‌سازی ذکر شده، نیاز به سیگنال‌های میزبان برای فرآیند کالیبره‌سازی است.

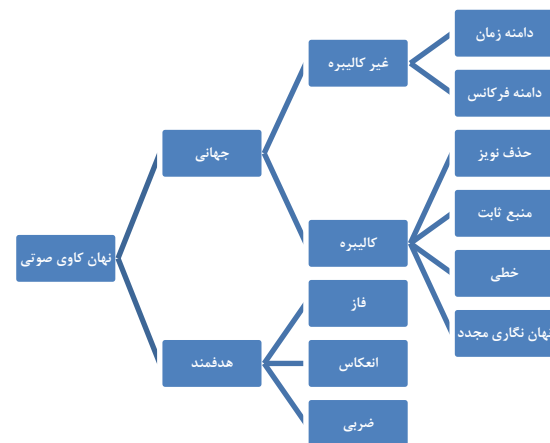
در روش‌های غیر کالیبره، ویژگی‌های نهان‌کاوی مستقیماً از سیگنال‌ها، در دامنه زمانی و یا فرکانسی، استخراج می‌شوند. مهم‌ترین ویژگی‌های دامنه زمانی که در نهان‌کاوی استفاده می‌شوند، از معادلات پیش‌بینی خطی [۱۵] و مدل‌های آماری هیستوگرام [۱۶] به‌دست می‌آیند و علاوه بر آن، شامل ویژگی‌های مبتنی بر آشوب [۱۷] و مارکوف [۱۸] نیز می‌باشند. در حوزه فرکانس، مهم‌ترین منابع برای نهان‌کاوی، شامل ضرایب کپسترال فرکانس مل^۴ (MFCC) [۱۹]، مشتقات MFCC سیگنال‌های صوتی [۱]، ترکیب ویژگی‌ها [۱] و مدل روان‌شناختی معکوس سیستم شنوایی انسان^۵ [۲۰] هستند.

اگرچه تاکنون تحقیقات متعددی در زمینه نهان‌کاوی صوتی

تحویل اطلاعات به گیرنده ممانعت به عمل می‌آورد [۱۶].

از آنجا که نهان‌نگاری اطلاعات در سیگنال‌های حامل، ویژگی‌های آن را تغییر می‌دهد، در صورت داشتن ویژگی‌های اطلاعات سیگنال اصلی، می‌توان یک نهان‌کاوی ایده‌آل را طراحی کرد. عدم وجود این اطلاعات، محققان را به سوی استفاده از روش‌های آماری در الگوریتم‌های یادگیری ماشین سوق داده است [۱ و ۷].

از دیدگاه دیگر، با تکیه بر ویژگی‌های سیگنال اصلی، الگوریتم‌های نهان‌کاوی به دو دسته هدفمند و جهانی تقسیم می‌شوند. شکل (۱) این روش تقسیم‌بندی الگوریتم‌های نهان‌کاوی را نمایش می‌دهد.



شکل (۱): تقسیم‌بندی الگوریتم‌های نهان‌کاوی صوتی بر مبنای ویژگی‌های سیگنال

کدگذاری فاز، نهان‌کاوی انعکاس و روش ضربی، سه زیرگروه روش هدفمند می‌باشند [۸]. کدگذاری فاز بر این مبنا طراحی شده است که نسبت فازها درون تمام بلوک‌ها حفظ می‌شود؛ اما این نسبت در بین بلوک‌ها متفاوت است [۹]. در روش‌های نهان‌نگاری انعکاس، مجموعه‌ای از هسته‌ها وجود دارند که یکی از آنها، بر مبنای بیت‌های پیام انتخاب شده و با قطعات فایل حامل ترکیب می‌شود [۱۰]. روش‌های جلو-عقب^۱ و مثبت-منفی^۲ دو نمونه از هسته‌هایی هستند که به این منظور طراحی شده‌اند. فرآیند نهان‌نگاری در زیرگروه روش‌های ضربی (که در اینجا از آن با عنوان نهان‌نگاری ضربی یاد می‌شود) را می‌توان به شکل معادله (۱) شرح داد که در آن، $s[n]$ ، $c[n]$ و $m[n]$ به‌ترتیب فایل حامل، میزبان و پیام هستند [۱۱]. برای نهان‌کاوی این روش، نویز بر مبنای معادله (۲) محاسبه می‌شود. نهان‌کاوی دیگر قادر

³ Audio Quality Metrics

⁴ Mel-frequency cepstral coefficients

⁵ Reversed Psychoacoustic Model of Human Hearing

¹ Forward-Backward

² Positive-Negative

[۲۲]، به‌منظور شناسایی فایل‌های حامل در قالب Wav که برای نهان‌نگاری آنها از نرم‌افزارهای Steghide، S-tools4 و Hide4PGP استفاده شده بود، از معیار نرخ فشردگی بدون اتلاف بهره بردند. در تحقیقی دیگر، [۲۳]، تینت و میا نهان‌کاو خود را بر مبنای ویژگی‌های نرخ عبور از صفر، ضرایب کپسترال فرکانس مل، ویژگی‌های کوتاه مدت انرژی و شار طیفی^۲ طراحی نمودند. ارزیابی عملکرد نهان‌کاو آنها حاکی از حساسیت مناسب آن در تشخیص فایل‌های نهان‌نگاری شده با استفاده از S-Tools4 و Hide4PGP است.

گوش مجازی بر مبنای مقیاس معکوس مل، یک روش بسیار حساس به مناطق با حساسیت بالا است [۲۰] که در مقایسه با روش‌های پیشین مبتنی بر MFCC، توانسته است به دقت بالاتری دست یابد. قاسم‌زاده و ارجمند، در نسخه توسعه‌یافته این روش، مدل مرتبه دوم مارکوف^۳ و ضرایب کپسترال مل^۴ (2D-MM) خود را بهبود دادند و یک مدل بر مبنای خصوصیات روان‌شناختی سیستم شنوایی انسان را طراحی کردند [۲۴]. HITIT نهان‌کاو است که توسط [۲۵]، بر مبنای ویژگی‌های تبدیل سریع فوریه در شبکه‌های عصبی مصنوعی طراحی شد و قادر است با دقت ۷۵٪، نمونه‌های حامل را از پاک جداسازی نماید.

در حیطه نهان‌کاو MP3، یو و همکاران [۲۶] از کالیبره کردن فشردسازی برای کشف فایل‌های حامل ایجادشده توسط MP3Stego استفاده کردند. روش آنها در نرخ‌های بالای فشردسازی، عملکرد مناسبی را از خود نشان داد. مدتی بعد یان و همکاران [۲۷] نهان‌کاو جدیدی را برای کشف فایل‌های حامل ایجادشده توسط MP3Stego معرفی کردند. این نهان‌کاو، از ویژگی انحراف از معیار دنباله دیفرانسیل مرتبه دوم در مرحله کوانتیزه کردن، برای جداسازی نمونه‌ها استفاده می‌کند. طبق ادعای محققین، نهان‌کاو آنها عملکرد بهتری را در مقایسه با دیگر نهان‌کاوهای MP3Stego از خود نشان می‌دهد.

با توجه به حجم گسترده استفاده از فناوری VoIP در اینترنت و به‌ویژه در پیام‌رسان‌ها و شبکه‌های اجتماعی، در اختیار داشتن نهان‌کاوهایی که برای تحلیل داده‌های VoIP طراحی شده‌اند، بسیار حیاتی است. به‌همین منظور، رن و همکاران [۲۸] یک روش نهان‌کاو چند نرخ تطابقی^۵ (AMR) را پیشنهاد دادند که برای حمله به حامل‌های نهان‌نگاری شده

صورت گرفته، اما روش‌های بسیار محدودی در حیطه نهان‌کاو صوتی گفتاری ارائه شده‌اند. علاوه بر این، تلاش‌های صورت گرفته توانایی و حساسیت لازم را برای حفظ امنیت فضای سایبری در اختیار ندارند. این تحقیق یک روش نوین نهان‌کاو صوتی گفتاری، در حیطه روش‌های جهانی غیرکالیبره را ارائه می‌کند. مهم‌ترین مزیت روش ارائه‌شده، تحلیل تک‌بعدی نمونه‌های داده‌شده در دامنه زمانی است. ویژگی‌ای که به این منظور طراحی شده است، درصد نمونه‌های مجاور یکسان، قادر به جداسازی نمونه‌های حامل از پاک، با یک بار پردازشی بسیار محدود است.

ساختار مقاله در بخش‌های بعدی به این صورت سازماندهی شده است: بخش دوم ادبیات تحقیق در حیطه‌های نهان‌کاو Wav، MP3، گفتار و صدا روی پروتکل اینترنت^۱ (VoIP) را مورد بررسی قرار می‌دهد. مفاهیم، ویژگی طراحی شده و جزئیات روش طراحی شده در بخش سوم شرح داده شده‌اند. بخش چهارم، با معرفی معیارهای ارزیابی آغاز شده و با ارزیابی نتایج و مقایسه نتایج با روش‌های مشابه، ادامه می‌یابد. در نهایت، بخش پنجم نتایج حاصل را ارائه می‌کند.

۲- ادبیات تحقیق

استفاده از ارتباطات صوتی در اینترنت، به‌شکل روزافزونی در حال گسترش است و اغلب شبکه‌های اجتماعی و پیام‌رسان‌ها مانند WhatsApp، WeChat، LINE و Facebook این ویژگی را به خدمات خود افزوده‌اند [۲۱]. در نتیجه، این رسانه ارتباطی عمومی، به یک کانال ارتباطی محبوب برای مهاجمین و قانون شکنان بدل شده است. علاوه بر این، طیف گسترده‌ای از ابزارها و الگوریتم‌ها، مانند MP3Stego، S-Tools و Hide4PGP ابداع شده‌اند که اختفای داده در این حامل‌ها را ساده می‌سازند. در پی توسعه روش‌های پیشرفته نهان‌نگاری، نهان‌کاو صوتی به یک ابزار حیاتی برای تأمین امنیت فضای سایبری تبدیل شده است.

لیو و همکاران [۱۸]، در سال ۲۰۰۹، یک روش بر مبنای کپسترال مل و مشتقات کپسترال فوریه، برای نهان‌کاو فایل‌های Wav را ارائه کردند. روش آنها عملکرد قابل اطمینانی را در نرخ‌های نهان‌نگاری بالا از خود نشان داد. برای توسعه این روش، محققین مشتقات درجه دوم ویژگی‌های انتقال مارکوف سیگنال را به‌همراه ضرایب کپسترال مل به‌کار گرفتند تا حساسیت نهان‌کاو خود را بالاتر ببرند. پس از آن، جبار و آیاد در

^۲ Spectral flux

^۳ Second order Markov Model

^۴ Mel cepstrum coefficients

^۵ Adaptive Multi-Rate

^۱ Voice over Internet Protocol

نهان‌کاو، برای جداسازی نمونه‌ها، از ترکیب ماشین بردار پشتیبان و روش جستجوی شناور روبه‌جلو ترتیبی^۴ (SFFS) استفاده می‌کند. نتایج ارزیابی، دقت عمل ۷۷/۳٪ تا ۱۰۰٪ را برای آن نشان می‌دهد. در تحقیقی دیگر، قاسم‌زاده و ارجمندی [۲۰]، یک روش نهان‌کاو صوتی مبتنی بر ماشین بردار پشتیبان را با استفاده از کپسترال مل معکوس طراحی کردند. ارزیابی آن در برابر دو روش واترمارکینگ و دو روش نهان‌نگاری، نشان‌دهنده دقت عمل ۹۱٪ تا ۹۹٪ است.

کوکال و همکاران یک روش نهان‌کاو گفتار را بر مبنای ویژگی‌های غیرخطی واریانس بردار تأخیر طراحی کردند [۳۴]. عملکرد این روش بر مبنای پایگاه داده صوتی TIMIT ارزیابی گردید و نتایج مطلوبی، به‌ویژه در کشف فایل‌های نهان‌نگاری شده با استفاده از Steghide، Hide4PGP، STOMOD، FHSS و DSSS حاصل شد. در گام بعدی، ویژگی‌های مبتنی بر واریانس بردار تأخیر، با ویژگی‌های مبتنی بر آشوب ترکیب شدند تا حساسیت نهان‌کاو بهبود یابد. نتایج ارزیابی این نهان‌کاو در جدول (۵) ارائه شده است.

روش نهان‌کاو گفتار فشرده میانو و همکاران در [۳۵]، سیگنال‌های فشرده‌شده گفتاری را برای کشف پیام‌های گنجانده‌شده مخفی، با استفاده از روش‌های آنتروپی^۵ و مارکوف^۶ مورد تحلیل قرار می‌دهد. در روش مذکور، از احتمالات گذر مارکوف برای ارزیابی وابستگی‌های بین پالس‌های مجاور و از آنتروپی، برای ارزیابی اختلالات انتشارهای پالس ترکیبی استفاده شده است. برای نهان‌کاو نمونه‌های داده شده، احتمالات گذر مارکوف و آنتروپی مشترک و مشروط پالس‌ها استخراج شد. سپس با استفاده از داده‌های جمع‌آوری شده، یک ماشین بردار پشتیبان، فایل داده‌شده را به‌عنوان پاک یا حامل دسته‌بندی می‌کند.

نهان‌کاو با استفاده از شبکه‌های باور عمیق که توسط پاولین و همکاران [۶] پیشنهاد شد، روشی است که در آن، نهان‌کاو طراحی شده قادر است نمونه فایل‌های نهان‌نگاری شده با استفاده از FreqSteg، Hide4PGP و StegHide را شناسایی کند. علاوه بر این، نهان‌کاو قادر است نام ابزار استفاده‌شده در نهان‌نگاری را نیز تعیین نماید.

به‌صورت سنتی، روش‌های نهان‌کاو VoIP بخش Payload بسته‌های داده‌ای را برای کشف داده‌های مخفی مورد بررسی قرار

توسط AMR [۲۹ و ۳۰]، بر مبنای احتمال وضعیت تکراری پالس عمل می‌کند. در یک تحقیق مشابه، تیان و همکاران [۲] از ویژگی‌های آماری زوج پالس‌ها برای کشف حامل‌های AMR که توسط روش‌های Geiser و Miao نهان‌نگاری شده‌اند، استفاده نمودند. عملکرد این نهان‌کاو زمانی قابل اطمینان است که نرخ نهان‌نگاری بیش از ۸۰٪ باشد. برای حمله به کتابچه کد‌ها فم‌ن که به‌منظور طراحی نمونه‌های حامل بر مبنای کدگذاری صوتی پیشرفته^۱ (AAC) طراحی شده است، رن و همکاران [۳۱] مدل مارکوف کتابچه کد مجاور را به‌کار گرفتند و تحت شرایط خاص، موفق به دستیابی به دقت ۱۰۰٪ شدند.

اولین نهان‌کاو بر مبنای یادگیری ماشین، توسط اوزر و همکاران [۳۲] و با ترکیب ماشین‌های بردار پشتیبان و معیارهای سنجش کیفیت صوت طراحی شده است. این نهان‌کاو، در برابر دو الگوریتم نهان‌نگاری و چهار الگوریتم واترمارکینگ ارزیابی گردیده و نتایج، حاکی از دقت عملکرد ۸۷٪ تا ۱۰۰٪ آن است. روش‌های کشف واترمارک و کشف نهان‌نگاری، به‌ترتیب به حساسیت ۶۹٪ و ۷۳٪ دست یافتند. در یک روش مدرن‌تر، جانسون و همکاران [۱۳] از تحلیل مؤلفه اصلی^۲ (PCA) برای استخراج مدل آماری نمونه‌های صوتی داده شده استفاده کردند. سیگنال‌ها ابتدا توسط تبدیل فوریه زمان کوتاه تجزیه شدند و سپس داده‌های آماری استخراج شده از PCA، برای طبقه‌بندی، به ماشین بردار پشتیبان داده شدند. تحلیل عملکرد این نهان‌کاو در برابر نهان‌نگار Hide4PGP و روش کم‌ارزش‌ترین بیت^۳ در نرخ نهان‌نگاری ۱۰۰٪، نشانگر حساسیت ۸۳/۱٪ تا ۱۰۰٪ آن است.

روش کاهش اعوجاجات مرزی توسط آلتون و همکاران [۵]، با تحلیل نمونه فایل‌های نهان‌نگاری شده، توسط دو روش واترمارکینگ، نمونه اعوجاجات مورفولوژیکی را برای استفاده در یک شبکه عصبی مورد تحلیل قرار می‌دهد. بسته به روش واترمارکینگ استفاده شده، دقت این روش بین ۸۰٪ تا ۱۰۰٪ نوسان دارد. در یک روش دیگر بر مبنای یادگیری ماشین، کرتزر و دیتمن [۱۹] از ضرایب کپسترال فرکانس مل (MFCC) به‌عنوان ویژگی مورد استفاده دسته‌بند استفاده کردند. برای دسته‌بندی فایل‌ها، ضرایب استخراج شده به یک تابع شعاعی در ماشین بردار پشتیبان داده شدند.

یوروکلا و همکاران [۳۳] در نهان‌کاو خود از ویژگی واریانس بردار تأخیر برای کشف وجود داده‌های مخفی استفاده کردند. این

^۴ Sequentialforward Floating Search Method

^۵ Entropy

^۶ Markov

^۱ Advanced Audio Coding

^۲ Principal Component Analysis

^۳ Least Significant Bit-LSB

جدول (۱): علائم و اختصارات

علامت	شرح
D	آرایه داده‌ای نمونه‌های وارد شده از فایل Wav
D_i	آمین نمونه در آرایه داده‌ای وارد شده از فایل Wav
Sp	آرایه نمونه‌های گفتار
Si	آرایه نمونه‌های سکوت
$Cnt(Sp)$	تعداد نمونه‌های از نوع گفتار در آرایه Sp
$Cnt(Si)$	تعداد نمونه‌های از نوع سکوت در آرایه Si
Sp_i	آمین نمونه در آرایه نمونه‌های گفتار
Si_i	آمین نمونه در آرایه نمونه‌های سکوت
$PEAS_{Sp}$	درصد نمونه‌های مجاور یکسان در آرایه نمونه‌های گفتار
$PEAS_{Si}$	درصد نمونه‌های مجاور یکسان در آرایه نمونه‌های سکوت
τ	حد آستانه برای جداسازی نواحی گفتار و سکوت در دامنه توان نمونه
P	آرایه توان نمونه‌ها
P_i	توان آمین نمونه در آرایه توان
P_{max}	بالاترین توان در آرایه توان
P_{min}	کمترین توان در آرایه توان
\hat{P}	نرمال شده آرایه توان P
\hat{P}_i	آمین نمونه در آرایه توان نرمال شده
$G(x)$	تابع عضویت گاوسی
G_{avg}	میزان متوسط عضویت گاوسی در مقادیر PEAS گفتار و سکوت
m	مقدار میانگین
m_{Sp}	مقدار میانگین باند گفتار داده‌های آموزشی
m_{Si}	مقدار میانگین باند سکوت داده‌های آموزشی
σ	انحراف از معیار
σ_{Sp}	انحراف از معیار باند گفتار
σ_{Si}	انحراف از معیار باند سکوت
TP	مثبت صحیح
FN	منفی کاذب

۳-۱- مؤلفه‌های PEAS

روش نهان‌کاوی PEAS شامل سه مؤلفه است: جداسازی نمونه‌های گفتار و سکوت، استخراج ویژگی و دسته‌بندی نمونه (فایل) داده شده. در مؤلفه اول، جداسازی نمونه‌ها، دسته‌بندی نمونه‌ها که خود شامل چهار بخش "محاسبه توان نمونه‌ها"، "نرمال‌سازی توان"، "اعمال حد آستانه" و "دسته‌بندی نمونه‌ها" است، نمونه‌های داده‌شده در دسته‌های گفتار و یا سکوت جداسازی می‌شوند. دومین مؤلفه، "استخراج ویژگی"، مقادیر ویژگی PEAS را از نمونه‌های گفتار و سکوت استخراج می‌کند. نهایتاً، در سومین مؤلفه، با نام دسته‌بندی نمونه‌ها، فایل گفتاری

می‌دهند. نهان‌کاو ارائه‌شده توسط جایاسری و امریدا [۳۶] نه تنها بخش داده‌ای بسته‌های VoIP را مورد تحلیل قرار می‌دهد، بلکه بخش مرتبط با ساختار پروتکل را نیز برای کشف داده‌های نهان شده تحلیل می‌نماید. برای این منظور، از مجموعه‌ای از ویژگی‌های آماری معمول، مانند کوواریانس و میان‌ه استفاده می‌شود.

یک روش مؤثر و پیشرفته‌تر نهان‌کاوی صوتی، برای کشف نمونه‌های نهان‌نگاری شده VoIP، مبتنی بر ترکیب نهان‌نگاری و تحلیل گفتار است که بر اساس ویژگی‌های جایگاه انتشار پالس، توسط تیان و همکاران طراحی شده است [۳۷]. محققین برای دسته‌بندی گفتارهای VoIP در نرخ بیتی پایین، از ماشین بردار پشتیبان استفاده کردند. به‌عنوان عامل جداسازی، از احتمالات گذر مارکوف جایگاه پالس، ویژگی‌های نامتغیر زمان کوتاه سیگنال‌های گفتار، احتمالات انتشار جایگاه پالس در ویژگی‌های انتشار طولانی مدت و ماتریس‌های احتمال مشترک، برای شناسایی وابستگی‌های بین پالس‌ها استفاده شده است. ارزیابی نهان‌کاو ارائه‌شده در برابر تعداد زیادی از نمونه‌های گفتاری کدگذاری شده با قالب G.729a صورت گرفته و نتایج حاکی از دقت ۹۲٪ بر روی نمونه‌هایی است که با ظرفیت کامل نهان‌نگاری شده‌اند.

چهارلنگ و همکاران [۳۸] یک روش نهان‌کاوی بر مبنای محاسبات کوانتومی را ارائه کرده‌اند. این الگوریتم قادر به کشف نهان‌نگاری فایل‌های صوتی در شبکه‌های ارتباطی کوانتومی است. نهان‌کاو شامل یک بلوک استخراج ویژگی از قاب‌های سیگنال‌های صوتی و همچنین مدارهای کوانتومی، جهت پیاده‌سازی با الگوریتم K نزدیک‌ترین همسایه (KNN) و شرط فاصله همینگ است. مدارات کوانتومی با استفاده از شبیه‌سازی، در برابر فایل‌های صوتی متفاوتی مورد آزمون قرار گرفتند. کسب دقت بیشتر از ۸۰٪ در طی ارزیابی، حاکی از کاربردپذیری این الگوریتم در شبکه‌های کوانتومی می‌باشد.

۳- ساختار نهان‌کاو پیشنهادی

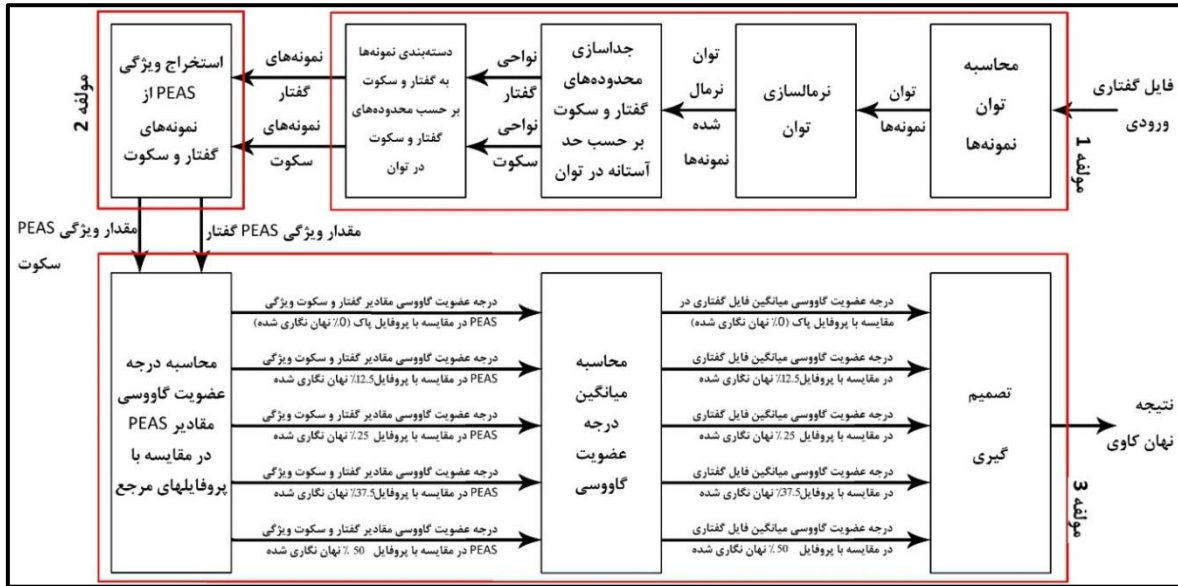
نهان‌کاو معرفی‌شده در این تحقیق، مبتنی بر ویژگی جدید ارائه‌شده با عنوان "درصد نمونه‌های مجاور یکسان"^۲ است که در ادامه، به اختصار، با عنوان PEAS مورد اشاره قرار خواهد گرفت.

به‌منظور ساده‌سازی درک علائم و اختصارات استفاده شده و همچنین حفظ یکپارچگی مقاله، علائم و اختصارات مورد توجه در ادامه مقاله، به‌همراه شرح مختصر آنها، در جدول (۱) ارائه‌شده‌اند.

^۱ K-nearest neighbor

^۲ Percent of Equal Adjacent Samples

داده‌شده به‌عنوان پاک یا حامل دسته‌بندی می‌شود. نمودار کلی روند کار PEAS در شکل (۲) نمایش داده شده است.



شکل (۲): ساختار نهان‌کاوی PEAS

توان نمونه‌های ورودی با استفاده از معادله (۳) محاسبه می‌گردد:

$$P_i = |D_i|^2 \quad \text{for } i=1,2,\dots, \text{length}(P) \quad (3)$$

توان نمونه‌ها همیشه تحت تأثیر کیفیت ضبط آن می‌باشد؛ تا جایی که دو ضبط مختلف از یک صدای یکسان ممکن است اساساً با یکدیگر متفاوت باشند. نرمال‌سازی توان فرآیندی است که با استفاده از آن، می‌توان بر این مشکل غلبه کرد و نمونه‌ها را به یک مقیاس یکسان تغییر داد. از آنجا که توان نمونه‌ها در قالب آرایه ذخیره می‌شود، برای نرمال‌سازی این آرایه از روش مقیاس‌پذیری ویژگی حداقلی-حداکثری (Min-Max)، بر مبنای معادله (۴) استفاده می‌گردد.

$$\hat{P}_i = \frac{P_i - P_{min}}{P_{max} - P_{min}} \quad (4)$$

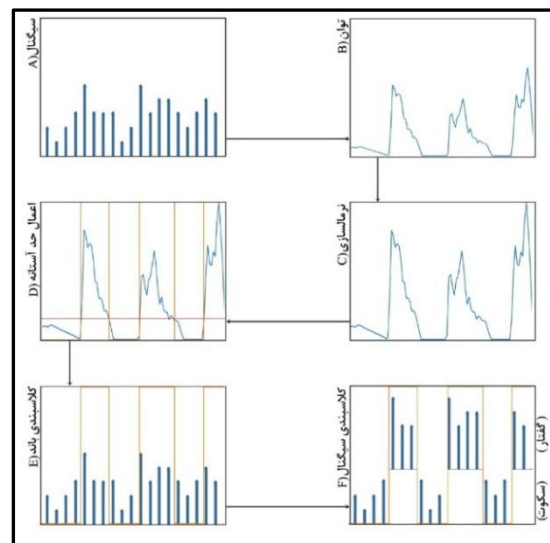
پس از نرمال‌سازی، نمونه‌های ورودی، بر اساس حد آستانه تعریف‌شده در باندهای گفتار و یا سکوت، تفکیک می‌شوند. این فرآیند، در شکل (۳)، با برجسب‌های D، E و F مشخص شده است. حد آستانه مؤثر برای این تفکیک در بخش ۴-۲ تعیین گردیده است.

آرایه‌های گفتار و سکوت، بر اساس یک نگاشت بین نمونه‌های ورودی و باندهای دسته‌بندی‌شده، تشکیل می‌گردند. معادله (۵) و شبه‌کد ارائه‌شده در شکل (۴) شرایط و گام‌های لازم برای این تفکیک را مشخص می‌کنند.

$$\begin{cases} \text{append}(\hat{P}_i, Sp) & \hat{P}_i \geq \tau \\ \text{append}(\hat{P}_i, Si) & \hat{P}_i < \tau \end{cases} \quad (5)$$

۳-۲- جداسازی نمونه‌های گفتار و سکوت

هدف اولین مؤلفه، جداسازی نمونه‌های ورودی در دسته‌های گفتار و سکوت است. به این منظور، در زیرمؤلفه اول، توان نمونه‌ها محاسبه شده و سپس در زیرمؤلفه دوم، نرمال‌سازی می‌شود. در سومین زیرمؤلفه، نواحی سکوت و گفتار، پس از اعمال حد آستانه بر روی توان‌های نرمال‌سازی شده، جداسازی می‌شوند. پس از آن، نمونه‌های ورودی به چهارمین زیرمؤلفه، بر اساس موقعیت دسته‌بندی‌شده آنها در آرایه توان نرمال شده، به‌عنوان گفتار و یا سکوت دسته‌بندی می‌شوند. بنابراین، خروجی‌های اولین مؤلفه، دو آرایه هستند که به ترتیب حاوی نمونه‌های گفتار و سکوت می‌باشند. گام‌های ذکرشده در شکل (۳) با برجسب‌های A تا E نام‌گذاری شده‌اند.



شکل (۳): فرآیند دسته‌بندی نمونه‌های ورودی به‌عنوان گفتار و سکوت

همچنین، شبه‌کد ارائه‌شده در شکل (۶) مراحل محاسبه PEAS برای نمونه‌های گفتار و سکوت را به تصویر کشیده است.

$$PEAS_{Sp} = \frac{\sum_{i=0}^{len(Sp)-2} [(Sp_i = Sp_{i+1}) \& (Sp_i \neq -1)] * 100}{Cnt(Sp)} \quad (۶)$$

$$PEAS_{Si} = \frac{\sum_{i=0}^{len(Si)-2} [(Si_i = Si_{i+1}) \& (Si_i \neq -1)] * 100}{Cnt(Si)} \quad (۷)$$

۳-۳-۲- کاهش وابستگی سیگنال

نهان‌سازی داده باعث افت هارمونی و هماهنگی بین نمونه‌های یک فایل صوتی گفتاری می‌شود. در نتیجه، مقادیر ویژگی PEAS برای نمونه‌های حامل، متفاوت از نمونه‌های پاک خواهد بود. به‌عبارت‌دیگر، هرچه حجم داده‌های نهان شده بیشتر باشد، اعوجاجات و ناهماهنگی‌ها در بین نمونه‌های مجاور افزایش می‌یابد. در نرخ‌های نهان‌نگاری پایین، تغییرات در نمونه‌ها به‌سختی قابل درک هستند؛ در حالی که در نرخ‌های بالا، فاصله بین نمونه‌های پاک و حامل افزایش یافته و با دقت بیشتری قابل جداسازی می‌باشند. در نهان‌کاوی گفتاری PEAS، میزان اعوجاجات ایجادشده بین نمونه‌های مجاور، به‌عنوان عامل جداسازی با استفاده از یادگیری ماشین، مورد استفاده قرار گرفته است. شکل (۷) چگونگی تأثیر نهان‌نگاری بر روی ویژگی PEAS را در نرخ‌های نهان‌نگاری مختلف نمایش می‌دهد. این شکل، بر اساس داده‌های به‌دست‌آمده از مقادیر PEAS، از باندهای گفتار و سکوت پنجاه نمونه فایل صوتی گفتاری، در پنج نرخ نهان‌نگاری ۰٪ (پاک)، ۱۲٫۵٪، ۲۵٪، ۳۷٫۵٪ و ۵۰٪ ترسیم شده است.

For $i=0$ to $len(Sp)-2$ do

If $(Sp_i = Sp_{i+1})$ and $(Sp_i \neq -1)$ then

SpCount += 1

PEAS-SP = $(SpCount * 100) / Cnt(Sp)$

A: محاسبه ویژگی PEAS در باند گفتار

For $i=0$ to $len(Si)-2$ do

If $(Si_i = Si_{i+1})$ and $(Si_i \neq -1)$ then

SiCount += 1

PEAS-SI = $(SiCount * 100) / Cnt(Si)$

B: محاسبه ویژگی PEAS در باند سکوت

شکل (۶): شبه‌کد محاسبه ویژگی PEAS

For $i=0$ to $len(\hat{P})-1$ do

If $\hat{P}_i \geq \tau$ then

append (\hat{P}, Sp)

append $(-1, Si)$

else

append (\hat{P}, Si)

append $(-1, Sp)$

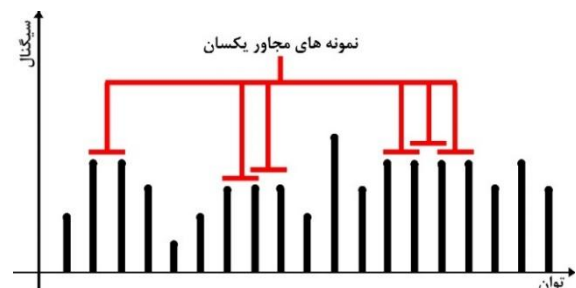
شکل (۴): شبه‌کد جداسازی نمونه‌های ورودی در آرایه‌های گفتار و سکوت، بر اساس اعمال حد آستانه بهینه بر روی آرایه توان نرمال‌سازی شده

۳-۳-۳- استخراج ویژگی

دومین مؤلفه، استخراج ویژگی، ویژگی PEAS را در آرایه‌های گفتار و سکوت محاسبه می‌کند. خروجی این مرحله دو مقدار برای ویژگی PEAS است که بر اساس آرایه‌ها محاسبه شده است.

۳-۳-۱- ویژگی PEAS

در این تحقیق، ویژگی "درصد نمونه‌های مجاور یکسان" (PEAS) به‌عنوان عامل تمایز نمونه‌های پاک و حامل طراحی شده است. برای محاسبه این ویژگی، نمونه‌های مجاور در فایل گفتاری داده‌شده، از اولین تا آخرین نمونه، با یکدیگر مقایسه می‌شوند. در صورتی که مقدار نمونه‌های مجاور یکسان باشد، شمارنده این ویژگی یک واحد افزایش می‌یابد. مفهوم ویژگی نمونه‌های مجاور یکسان، در شکل (۵) به تصویر کشیده شده است. از آنجا که روش نهان‌کاوی پیشنهادی نمونه‌ها را در دو دسته گفتار و سکوت بررسی می‌کند، این ویژگی برای هر کدام از دسته‌ها، باید به‌صورت جداگانه محاسبه شود.



شکل (۵): مفهوم ویژگی نمونه‌های مجاور یکسان

درصد نمونه‌های مجاور یکسان برای باندهای گفتار و سکوت، به‌ترتیب، با استفاده از معادلات (۶) و (۷) محاسبه می‌شود.

که نتیجه نهان‌کاوی را بر اساس نرخ نهان‌نگاری پروفایل تولیدکننده بالاترین میزان میانگین، تعیین می‌کند.

توابع گاوسی نوعی از توابع آماری هستند که در ریاضیات، پردازش سیگنال و تصویر استفاده می‌شوند و با شرح انتشار طبیعی داده‌ها در حل مسائل مرتبط با انتشار و دسته‌بندی داده‌ها، به‌صورت گسترده‌ای مورد استفاده قرار می‌گیرند. تابع عضویت گاوسی نوعی از این توابع است که در مسائل دسته‌بندی، با دو پارامتر مشخص می‌شود: میانگین (m) و انحراف از معیار (σ). در این تابع، برای هر عنصر، یک نگاهت به محدوده ۰ تا ۱ ایجاد می‌شود. این مقدار که درجه عضویت یا مقدار عضویت نامیده می‌شود، درجه عضویت هر عنصر را در قالب یک مجموعه فازی (۰ تا ۱) تعیین می‌کند.

برای محاسبه درجه عضویت گاوسی یک مقدار داده شده، به دو پارامتر m و σ نیاز است که در این الگوریتم، از داده‌های آموزشی استخراج می‌شوند. از آنجا که در این روش، نمونه‌های صوتی گفتاری داده‌شده در دو باند گفتار و سکوت و در پنج نرخ نهان‌نگاری مختلف مورد بررسی قرار می‌گیرند، ده مجموعه از مقادیر $\{m, \sigma\}$ مورد نیاز است تا پروفایل نهان‌کاوی، برای نمونه‌های صوتی داده‌شده در تمامی نرخ‌های نهان‌نگاری پشتیبانی شده ایجاد گردد. ساختار پایه پروفایل مرجع PEAS در جدول (۲) نمایش داده شده است. برای تشخیص پاک یا حامل بودن، مقادیر PEAS محاسبه‌شده برای نمونه فایل داده شده، با پروفایل مقایسه می‌شوند.

جدول (۲): ساختار پروفایل مرجع نهان‌کاوی PEAS

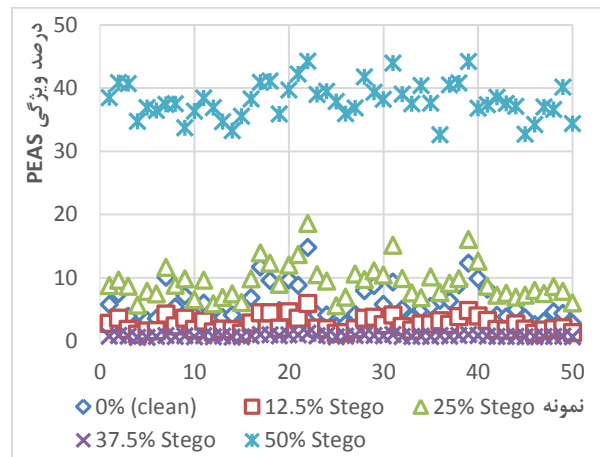
نرخ نهان‌نگاری										
حامل ۵۰٪		حامل ۳۷/۵٪		حامل ۲۵٪		حامل ۱۲/۵٪		پاک (۰٪)		گفتار
m	σ	m	σ	m	σ	m	σ	m	σ	
m	σ	m	σ	m	σ	m	σ	m	σ	سکوت

$$G(x) = \exp\left[\frac{-(x-m)^2}{2\sigma^2}\right] \quad (۸)$$

$$G(PEAS_{Sp}) = \exp\left[\frac{-(PEAS_{Sp} - m_{Sp})^2}{2\sigma_{Sp}^2}\right] \quad (۹)$$

$$G(PEAS_{Si}) = \exp\left[\frac{-(PEAS_{Si} - m_{Si})^2}{2\sigma_{Si}^2}\right] \quad (۱۰)$$

با استفاده از دو معادله (۹) و (۱۰)، دو مقدار درجه عضویت گاوسی به‌دست می‌آید. به‌منظور دست‌یابی به یک مقدار نهایی



شکل (۷): مقادیر ویژگی PEAS در باند سکوت پنجاه نمونه صوتی گفتاری در پنج نرخ نهان‌نگاری

۴-۳- دسته‌بندی نمونه‌ها

سومین مؤلفه، خود شامل سه بخش است که وظیفه دسته‌بندی نمونه‌های داده‌شده را به پاک یا حامل بر عهده دارند. این مؤلفه، ابتدا مقادیر PEAS را از مؤلفه دوم دریافت کرده و سپس درجه عضویت گاوسی آنها را در مقایسه با محدوده‌های استخراج شده پاک و حامل، در پروفایل‌های مرجع، محاسبه می‌نماید. خروجی اولین زیرمؤلفه شامل پنج آرایه متشکل از درجه‌های عضویت گاوسی است که بر اساس مقایسه مقادیر PEAS با پنج پروفایل مرجع ایجاد شده‌اند. دومین زیرمؤلفه میزان میانگین درجه عضویت گاوسی برای هر یک از آرایه‌های دریافتی را محاسبه کرده و آنها را به بخش تصمیم‌گیری ارسال می‌کند؛ یعنی بخشی

دسته‌بند انتخابی در نهان‌کاوی PEAS، تابع عضویت گاوسی است. درجه عضویت با استفاده از معادله (۱۰) محاسبه می‌شود که در آن، x متغیری است که مقدار ویژگی را در خود نگه می‌دارد و $G(x)$ درجه عضویت آن است. از آنجا که ویژگی انتخابی، مقادیر PEAS در باندهای گفتار و سکوت هستند، با جایگزینی x در معادله (۶) با مقادیر PEAS در باند گفتار و سکوت (به ترتیب، $PEAS_{Sp}$ و $PEAS_{Si}$)، به ترتیب معادلات (۹) و (۱۰) حاصل می‌شوند.

است که از ترکیب میانگین تابع عضویت گاوسی و فرمول‌های $PEAS_{Sp}$ و $PEAS_{Si}$ به دست آمده است. برای تولید نتیجه نهایی نهان‌کاوی، معادله (۱۲) باید پنج بار با استفاده از m و σ مرتبط با $PEAS_{Sp}$ و $PEAS_{Si}$ ، در پنج نرخ نهان‌نگاری پوشش داده شده، محاسبه گردد. نرخ نهان‌نگاری پروفایلی که بالاترین میزان درجه عضویت گاوسی را تولید کرده باشد، پاک و یا حامل بودن نمونه فایل داده شده را مشخص می‌کند. شبه‌کد تعیین نتیجه نهایی نهان‌کاوی، با استفاده از معادله (۱۲)، در شکل (۸) ارائه شده است.

$$G_{avg}(PEAS_{Sp}, PEAS_{Si})$$

$$= \frac{\exp\left[\frac{\sum_{i=0}^{len(Sp)-2} [(Sp_i = Sp_{i+1}) \& (Sp_i <> -1)] * 100 - m_{Sp}^2}{Cnt(Sp)} - 2\sigma_{Sp}^2\right] + \exp\left[\frac{\sum_{i=0}^{len(Si)-2} [(Si_i = Si_{i+1}) \& (Si_i <> -1)] * 100 - m_{Si}^2}{Cnt(Si)} - 2\sigma_{Si}^2\right]}{2} \quad (12)$$

۵۰٪ برای ارزیابی تقسیم گردید.

۳-۶- آموزش

همان‌گونه که پیش از این ذکر شد، نیمی از نمونه‌های موجود در پایگاه داده، برای آموزش دسته‌بند مورد استفاده قرار گرفته‌اند. تابع عضویت گاوسی، برای دسته‌بندی نمونه‌های جدید، نیازمند پارامترهای m و σ است. برای این منظور، این مقادیر از باندهای گفتار و سکوت نمونه‌های پاک و حامل، به ترتیب با نرخ‌های نهان‌نگاری ۰٪، ۱۲٪، ۲۵٪، ۳۷٪ و ۵۰٪، از پایگاه داده استخراج شدند.

پارامتر مهم دیگری که در این روش وجود دارد، یک حد آستانه برای جداسازی نمونه‌های گفتار و سکوت است. در آغاز، حد آستانه بهینه ناشناخته است و باید بر اساس مجموعه‌ای از آزمون‌ها بر روی آستانه‌های مختلف تعیین گردد. به همین منظور، پنج آزمایش با حدود آستانه ۰/۱، ۰/۱۵، ۰/۲، ۰/۲۵ و ۰/۳ صورت گرفتند. جدول (۳) زوج‌های m و σ به دست آمده در طی آزمون‌های صورت گرفته در فرآیند آموزش را نشان می‌دهد. بر اساس مقادیر حساسیت^۲ و ویژگی^۳ ارائه شده در جدول (۴)، بهترین نتایج زمانی به دست آمده است که حد آستانه بر روی عدد ۰/۱۵ تنظیم گردیده است. به همین دلیل، زوج m و σ و مقادیر به دست آمده این حد آستانه، در جدول (۳) به صورت ضخیم^۴ نشان داده شده است.

برای تصمیم‌گیری، میانگین مقادیر PEAS به دست آمده، توسط معادله (۱۱) محاسبه شده و در فرآیند تصمیم‌گیری مورد استفاده قرار می‌گیرد.

$$G_{avg}(PEAS_{Sp}, PEAS_{Si}) = \frac{G(PEAS_{Sp}) + G(PEAS_{Si})}{2} \quad (11)$$

به منظور دستیابی به یک معادله یکتا که قادر به محاسبه میزان نهایی درجه عضویت گاوسی باشد، معادله (۱۲) با توسعه معادله (۱۱) و با جایگزینی مقدار $G(PEAS_{Sp})$ و $G(PEAS_{Si})$ با معادل آن در معادله (۸) تولید می‌شود. معادله (۱۲) فرمول نهایی

Ratio=(0%, 12.5%, 25%, 37.5%, 50%)

for i=1 to 5 do

md[i]= G_{avg}(PEAS_{Sp}[ratio[i]], PEAS_{Si}[ratio[i]])

max_md=max(md[1]..md[5])

If max_md=md[1] then print('the given object is clean')

else print('the given object is stego')

شکل (۸): شبه‌کد فرآیند تصمیم‌گیری

۳-۵- پایگاه داده

برای آموزش و ارزیابی نهان‌کاوی PEAS، یک پایگاه داده متشکل از ۱۱۰۰ نمونه صوتی گفتاری تولید شد. به این منظور، ده کتاب صوتی به زبان فارسی، از قالب MP3 به قالب ۱۶ Wav بیتی تک باند^۱، با فرکانس نمونه‌برداری ۴۴۱۰۰ هرتز تبدیل و در قطعات ۱۰ ثانیه‌ای برش خوردند. در فرآیند تشکیل پایگاه داده، قطعاتی که دارای موسیقی زمینه بودند از پایگاه داده حذف شدند. پایگاه داده تولیدی، ترکیبی از صدای زنان و مردان و به ترتیب، با تعداد ۷۷۳ و ۳۲۷ است.

پس از تولید پایگاه داده، فایل‌های آن با استفاده از روش جایگزینی کم‌ارزش‌ترین بیت در نرخ‌های ۱۲٪، ۲۵٪، ۳۷٪ و ۵۰٪، با استفاده از ابزار Wavsteg در Python نهان‌نگاری شدند. بنابراین نسخه نهایی پایگاه داده متشکل از ۱۱۰۰ نمونه در پنج نسخه پاک (۰٪)، ۱۲٪، ۲۵٪، ۳۷٪ و ۵۰٪ (مجموعاً ۵۵۰۰ فایل) است. پایگاه داده در نهایت به دو بخش ۵۰٪ برای آموزش و

² Sensitivity

³ Specificity

⁴ Bold

¹ Mono

جدول (۳): مقادیر استخراج شده m و σ از باندهای گفتار و سکوت نمونه‌های گفتاری در پایگاه داده آموزشی در پنج نرخ نهان‌نگاری و پنج حد آستانه

نرخ نهان‌نگاری										آستانه	کانال
۰٪		۱۲٪		۲۵٪		۳۷٪		۵۰٪			
m	σ	m	σ	m	σ	m	σ	m	σ		
۰/۰۰۷	۰/۰۰۵۹	۰/۰۰۴۳	۰/۰۰۳۵	۰/۰۰۲۶۳	۰/۰۰۲۱۴	۰/۰۰۰۸۶	۰/۰۰۰۶۶	۰/۰۰۶۹۰۸	۰/۰۰۵۱۸۳	گفتار	۰/۱۰
۵/۶۱۵۷	۴/۰۹۶۳	۲/۱۹۳۶	۱/۱۵۵۵	۷/۲۶۷۸	۲/۷۶۳۲	۰/۶۶۴۵	۰/۱۶۴۲	۳۵/۷۱۵۶	۵/۰۱۴۷	سکوت	
۰/۰۰۳۳	۰/۰۰۳۱	۰/۰۰۲۱	۰/۰۰۱۹	۰/۰۰۱۲۹	۰/۰۰۱۱۱	۰/۰۰۰۴۵	۰/۰۰۰۳۷	۰/۰۰۳۵۲۱	۰/۰۰۲۹۵۵	گفتار	۰/۱۵
۵/۶۱۸۸	۴/۰۹۳	۲/۱۹۵۱	۱/۱۵۴۶	۷/۲۷۸۲	۲/۷۵۹۷	۰/۶۶۴۰	۰/۱۶۳۸	۳۵/۹۹۳۷	۴/۹۴۵۹	سکوت	
۰/۰۰۱۸	۰/۰۰۱۸	۰/۰۰۱۱	۰/۰۰۱۲	۰/۰۰۰۷۱	۰/۰۰۰۶۶	۰/۰۰۰۲۷	۰/۰۰۰۲۴	۰/۰۰۲۰۱	۰/۰۰۱۸۴۱	گفتار	۰/۲۰
۵/۶۲۰۱	۴/۰۹۲۷	۲/۱۹۵۷	۱/۱۵۴۴	۷/۲۸۲۵	۲/۷۵۸۳	۰/۶۶۳۶	۰/۱۶۳۶	۳۶/۱۱۳۲	۴/۹۲۰۱	سکوت	
۰/۰۰۱۱	۰/۰۰۱۲	۰/۰۰۰۷	۰/۰۰۰۸	۰/۰۰۰۴۴	۰/۰۰۰۴۲	۰/۰۰۰۱۷	۰/۰۰۰۱۵	۰/۰۰۱۲۷۴	۰/۰۰۱۲۳۱	گفتار	۰/۲۵
۵/۶۲۰۷	۴/۰۹۲۶	۲/۱۹۵۹	۱/۱۵۴۳	۷/۲۸۴۵	۲/۷۵۷۶	۰/۶۶۳۳	۰/۱۶۳۴	۳۶/۱۶۸۳	۴/۹۰۳۲	سکوت	
۰/۰۰۰۷	۰/۰۰۰۸	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۲۸	۰/۰۰۰۲۸	۰/۰۰۰۱۱	۰/۰۰۰۱۱	۰/۰۰۸۰۹	۰/۰۰۸۵۸	گفتار	۰/۳۰
۵/۶۲۱۰	۴/۰۹۲۴	۲/۱۹۶۱	۱/۱۵۴۲	۷/۲۸۵۵	۲/۷۵۷۲	۰/۶۶۳۱	۰/۱۶۳۴	۳۶/۲۰۲۸	۴/۸۹۳۳	سکوت	

گفتاری و موسیقی کلاسیک آغاز شده و با آزمایش‌هایی جهت تخمین طول پیام نهان‌نگاری شده ادامه می‌یابد.

۴- ارزیابی

در این بخش، ضمن معرفی معیارهای ارزیابی، به بررسی شرایط لازم برای ارزیابی نهان‌نگار و پیشنهادی و معرفی روند این ارزیابی، پرداخته خواهد شد.

۴-۱- معیارهای ارزیابی

مهم‌ترین معیار ارزیابی نهان‌نگار PEAS، توانایی آن برای کشف نمونه‌های حامل است و حساسیت، معیاری است که این توانایی را می‌سنجد. به عبارت ساده، حساسیت بیانگر درصد نمونه‌های حامل تشخیص داده شده نسبت به کل تعداد نمونه‌های حامل است. حساسیت را می‌توان با استفاده از معادله (۱۳) محاسبه کرد. از سوی دیگر، کشف نمونه‌های حامل تنها عامل مهم در ارزیابی یک روش نهان‌نگار نیست و درصد صحیح تشخیص نمونه‌های پاک نیز مهم است. این معیار ویژگی نامیده می‌شود و توسط معادله (۱۴) محاسبه می‌گردد.

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} * 100 \quad (13)$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} * 100 \quad (14)$$

۴-۲- روند ارزیابی

روال ارزیابی روش پیشنهادی با آماده‌سازی دو پایگاه داده صوتی

۴-۲-۱- پایگاه داده صوتی گفتاری

الگوریتم نهان‌نگار کوی ارائه‌شده، با استفاده از چهار هسته ماشین بردار پشتیبان و تابع عضویت گاوسی پیاده‌سازی شده است. چهار پیاده‌سازی ماشین بردار پشتیبان عبارت‌اند از هسته خطی، چند جمله‌ای حالت مقیاس، گاوسی حالت خودکار و گاوسی حالت مقیاس. جدول (۴) حاوی نتایج به‌دست‌آمده در پیاده‌سازی‌ها می‌باشد. با توجه به نتایج و حساسیت‌های ارائه‌شده، ماشین بردار پشتیبان نتایج بسیار مطلوبی را با توان تشخیصی بسیار بالا، حتی در نرخ‌های نهان‌نگاری بسیار پایین به‌دست آورده است. با این وجود، زمانی که معیار ویژگی نیز مورد توجه قرار می‌گیرد، تمام پیاده‌سازی‌های ماشین بردار پشتیبان، عملکردهای بسیار ضعیفی را نشان می‌دهند. آنچه که در عمل اتفاق می‌افتد این است که درصد بسیار بالایی از نمونه‌هایی که در شبکه‌ها در حال تبادل هستند، پاک می‌باشند و تولید تعداد زیاد هشدارهای اشتباه، الگوریتم ابداع‌شده را به یک الگوریتم غیرکاربردی بدل می‌نماید.

در فرآیند نهان‌نگار، مهم‌ترین هدف، کشف نمونه‌های حامل است. بنابراین، حساسیت اهمیتی بالاتر از ویژگی دارد. به‌منظور انتخاب دسته‌بند و حد آستانه بهینه، به ازای هر یک از حدود آستانه، حساسیت هر یک از دسته‌بندها سنجیده شد. برای هر

یک از دسته‌بندها، سطری که بالاترین میزان حساسیت را داشته است به‌صورت ضخیم نشان داده شده است.

همان‌طور که در جدول (۴) نشان داده شده است، میزان ویژگی در پیاده‌سازی‌های ماشین بردار پشتیبان که به‌صورت ضخیم مشخص شده‌اند، بین ۸/۴٪ تا ۳۵/۵۹٪ نوسان دارد. با در نظر گرفتن این محدوده، حتی بالاترین میزان ویژگی به‌دست آمده با ماشین بردار پشتیبان، قادر به ارائه یک نهان‌کاو کاربردی و قابل اعتماد نیست. زمانی که تابع عضویت گاوسی به‌عنوان

دسته‌بند استفاده می‌شود، اگرچه از نظر معیار حساسیت، عملکرد بهتری نسبت به پیاده‌سازی‌های مبتنی بر ماشین بردار پشتیبان ندارد، اما از نظر معیار ویژگی، با فاصله زیادی بهتر از آن عمل می‌کند. در حد آستانه ۰/۱۵، تابع عضویت گاوسی حساسیت بسیار مطلوبی را ارائه می‌کند، درحالی‌که از نظر ویژگی نیز میزان آن ۷۰/۲۵٪ بوده و عملکرد خوبی دارد. بنابراین، این دسته‌بند به‌عنوان موتور تشخیص نهان‌کاو PEAS انتخاب شده است و حد آستانه بهینه نیز ۰/۱۵ تعیین گردیده است.

جدول (۴): میزان حساسیت و ویژگی الگوریتم نهان‌کاو گفتار، PEAS، با استفاده از دسته‌بندهای ماشین بردار پشتیبان و تابع عضویت گاوسی، در پنج حد آستانه

دسته‌بند	حد آستانه	ویژگی				حساسیت			
		نرخ نهان‌نگاری				نرخ نهان‌نگاری			
		۰٪	۱۲/۵٪	۲۵٪	۳۷/۵٪	۵۰٪			
ماشین بردار پشتیبان - خطی	۰/۱۰	۳۵/۵۹٪	۸۶/۱۳٪	۸۱/۱۳٪	۱۰۰٪	۱۰۰٪			
	۰/۱۵	۴۰٪	۸۰/۸۷٪	۷۸/۱٪	۱۰۰٪	۱۰۰٪			
	۰/۲۰	۳۵/۷۱٪	۷۷/۴۷٪	۷۳/۳۶٪	۱۰۰٪	۱۰۰٪			
	۰/۲۵	۲۸/۰۷٪	۸۱/۶۴٪	۸۰/۷۸٪	۱۰۰٪	۹۹/۵۴۵٪			
	۰/۳۰	۳۲/۶۹٪	۸۱/۲۷٪	۷۳/۸۷٪	۱۰۰٪	۹۹/۵۴۷٪			
ماشین بردار پشتیبان - چندجمله‌ای (حالت Scale)	۰/۱۰	۵۳/۶۲٪	۹۳/۱۸٪	۱۶/۴۵٪	۱۰۰٪	۱۰۰٪			
	۰/۱۵	۶۱/۲۱٪	۹۰/۴۱٪	۱۱/۷۳٪	۱۰۰٪	۱۰۰٪			
	۰/۲۰	۸/۴٪	۹۵/۲۸٪	۹۵/۵۳٪	۱۰۰٪	۱۰۰٪			
	۰/۲۵	۵۷/۵۲٪	۹۰/۲۵٪	۱۳/۸٪	۱۰۰٪	۱۰۰٪			
	۰/۳۰	۱۱/۰۱٪	۹۳/۸٪	۹۶/۳۱٪	۱۰۰٪	۱۰۰٪			
ماشین بردار پشتیبان - گاوسی (حالت Auto)	۰/۱۰	۳۰/۶۱٪	۸۹/۹۶٪	۸۴/۳۳٪	۱۰۰٪	۱۰۰٪			
	۰/۱۵	۲۶/۷۹٪	۸۹/۴٪	۸۴/۰۳٪	۱۰۰٪	۱۰۰٪			
	۰/۲۰	۳۸٪	۸۴/۸۲٪	۸۲/۸۵٪	۱۰۰٪	۱۰۰٪			
	۰/۲۵	۲۹/۷٪	۹۰/۷۸٪	۸۵/۳۱٪	۱۰۰٪	۱۰۰٪			
	۰/۳۰	۲۹/۹۵٪	۸۶/۹۱٪	۸۴/۸۷٪	۱۰۰٪	۱۰۰٪			
ماشین بردار پشتیبان - گاوسی (حالت Scale)	۰/۱۰	۳۰/۴۸٪	۸۵/۴۸٪	۸۴/۷۲٪	۱۰۰٪	۱۰۰٪			
	۰/۱۵	۲۷/۸۳٪	۸۶/۸٪	۸۵/۳۲٪	۱۰۰٪	۱۰۰٪			
	۰/۲۰	۲۶/۳۸٪	۸۸/۴۹٪	۸۵/۳۴٪	۱۰۰٪	۹۹/۵۴٪			
	۰/۲۵	۲۹/۲٪	۹۸/۲۹٪	۸۷/۰۹٪	۱۰۰٪	۱۰۰٪			
	۰/۳۰	۲۶/۴٪	۹۰/۲۳٪	۸۵/۹۱٪	۱۰۰٪	۱۰۰٪			
تابع عضویت گاوسی	۰/۱۰	۷۱/۳۲٪	۶۲/۳۶٪	۶۷/۹۲٪	۹۱/۰۳٪	۹۹/۴۶٪			
	۰/۱۵	۷۰/۲۵٪	۶۴/۱۵٪	۶۷/۲٪	۸۹/۷۸٪	۹۹/۸۲٪			
	۰/۲۰	۶۶/۴۸٪	۶۵/۰۵٪	۶۵/۲۳٪	۸۹/۹۶٪	۹۹/۸۲٪			
	۰/۲۵	۶۶/۳٪	۶۵/۹۴٪	۶۱/۱۱٪	۸۹/۹۶٪	۱۰۰٪			
	۰/۳۰	۶۲٪	۶۷/۱۲٪	۶۱/۴۶٪	۹۰/۸۶٪	۱۰۰٪			

نیز، در حدود آستانه مختلف، در شکل (۹) به تصویر کشیده شده است. همان‌گونه که در شکل (۹) مشخص است، دسته‌بندی

علاوه بر میزان حساسیت‌ها و ویژگی‌های اعلام‌شده در جدول (۴)، منحنی عملیاتی دریافت‌کننده تابع عضویت گاوسی

جدول (۵) نشان داده شده، اگرچه دقت عمل الگوریتم در دسته‌بندی نمونه‌های موسیقی کلاسیک، کمتر از نمونه‌های گفتاری می‌باشد، اما کماکان، علی‌رغم تک‌بعدی بودن تحلیل‌ها، با ارائه حساسیت‌های ۳۱٪/۲، ۸۶٪/۳، ۵۱٪/۶ و ۸۱٪/۲ در نرخ‌های ۱۲٪/۵ تا ۵۰٪، عملکرد نسبتاً مطلوبی را ارائه می‌دهد. لازم به ذکر است که دسته‌بند استفاده شده، بر روی نمونه‌های گفتاری آموزش داده شده و سپس بر روی نمونه‌های موسیقی آزموده شده است. آموزش و آزمون بر روی پایگاه داده موسیقی می‌تواند منجر به بهبود حساسیت گردد.

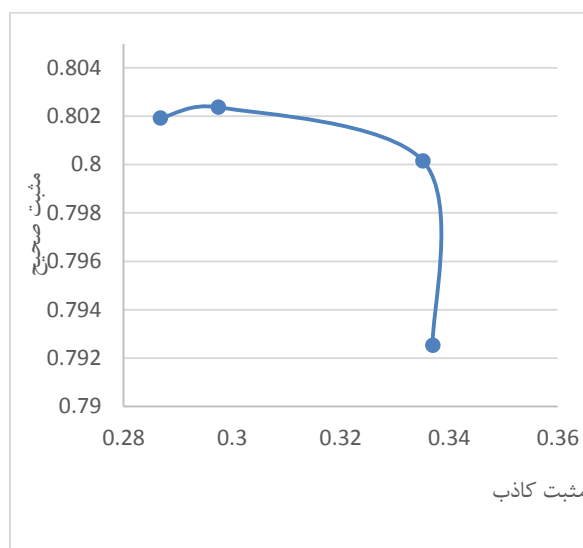
جدول (۵): ارزیابی عملکرد الگوریتم نهان‌کاوی گفتاری PEAS بر روی پایگاه داده موسیقی کلاسیک

نهان‌نگاری	نهان‌نگاری	نهان‌نگاری	نهان‌نگاری	
۵۰٪	۳۷٪/۵	۲۵٪	۱۲٪/۵	
۸۱٪/۲	۵۱٪/۶	۸۶٪/۳	۳۱٪/۲	حساسیت

۴-۲-۳- تخمین طول پیام

هدف اصلی الگوریتم‌های نهان‌کاوی، کشف وجود و یا عدم وجود پیام در نمونه‌های داده شده است. در صورتی که جمع‌بندی نهایی الگوریتم، حاکی از وجود داده‌های مخفی در فایل حامل باشد، برای ممانعت از تحویل پیام در مقصد، می‌توان با اعمال تغییرات در بدنه فایل حامل، پیام را تخریب نمود. به‌منظور بهینه‌سازی این فرآیند، نیاز است که الگوریتم نهان‌کاوی مورد استفاده قدرت تخمین حجم پیام پنهان شده را نیز دارا باشد. بر اساس تخمین الگوریتم، می‌توان تعداد بهینه بیت‌های کم ارزش را برای تغییر انتخاب نمود. در این زمینه، دقت الگوریتم نهان‌کاوی PEAS، با استفاده از پیاده‌سازی‌های مختلف ماشین بردار پشتیبان و همچنین تابع عضویت گاوسی، در حدود آستانه ۰/۱ تا ۰/۳، با گام‌های ۰/۰۵، مورد ارزیابی قرار گرفته و نتایج آن در جدول (۶) ارائه گردیده است. البته با توجه به اینکه در فرآیند نهان‌کاوی، دسته‌بند مبتنی بر تابع عضویت گاوسی در حد آستانه ۰/۱۵ انتخاب شده است، در ساختار نهایی الگوریتم، برای تخمین حجم پیام نیز از همان دسته‌بند و حد آستانه استفاده شده است. نتایج ارزیابی حاکی از دقت ۵۶٪/۲۷ و ۹۹٪/۲۸ در تخمین حجم پیام مخفی شده در نرخ‌های مختلف نهان‌نگاری می‌باشد.

به‌صورت نامتعادل بوده و حد آستانه ۰/۱۵، بالاترین میزان مثبت صحیح را در برابر میزان نسبتاً حداقلی مثبت کاذب ارائه می‌کند.



شکل (۹): نمودار منحنی عملیاتی دریافت‌کننده در تابع عضویت گاوسی

۴-۲-۴- پایگاه داده موسیقی کلاسیک

علاوه بر ارزیابی نهان‌کاوی گفتاری PEAS بر روی یک پایگاه داده صوتی گفتاری، این الگوریتم بر روی یک پایگاه داده صوتی، متشکل از ۱۰۰۰ قطعه ۱۰ ثانیه‌ای موسیقی سمفونی کلاسیک بتهون نیز ارزیابی گردیده است تا عملکرد آن بر روی نمونه‌های موسیقی نیز سنجیده شود. برای تولید این پایگاه داده، همانند پایگاه داده گفتاری تولیدشده، موسیقی‌های سمفونی کلاسیک بتهون، از قالب MP3 به قالب Wav با تنظیمات (۱۶بیتی، تک باند، ۴۴۱۰۰Hz) تبدیل شده و ۱۰۰۰ نمونه صوتی موسیقی را تشکیل دادند. سپس این قطعات، با استفاده از ابزار Wavsteg در زبان Python، در نرخ‌های ۱۲٪/۵، ۲۵٪، ۳۷٪/۵ و ۵۰٪، نهان‌نگاری شدند تا مجموعاً ۵۰۰۰ قطعه موسیقی برای ارزیابی عملکرد الگوریتم تولید شود. پایگاه داده نهایتاً به دو بخش مساوی برای آموزش و ارزیابی تقسیم گردید.

با توجه به اینکه بهترین عملکرد PEAS در حد آستانه ۰/۱۵ و با استفاده از تابع عضویت گاوسی به‌دست‌آمده است، بنابراین، از این حد آستانه و دسته‌بند، برای دسته‌بندی نمونه‌ها در پایگاه داده موسیقی کلاسیک نیز استفاده شده است. همان‌گونه که در

جدول (۶): دقت الگوریتم نهان‌کاوی گفتار PEAS در تخمین طول پیام با استفاده از دسته‌بندهای ماشین بردار پشتیبان و تابع عضویت گاوسی در پنج حد آستانه

نرخ نهان‌نگاری				حد آستانه	دسته‌بند
%۵۰	%۳۷/۵	%۲۵	%۱۲/۵		
%۱۰۰	%۹۸/۶۱	%۷۱/۶۹	%۷۱/۲۸	۰/۱۰	ماشین بردار پشتیبان - خطی
%۱۰۰	%۹۹/۵۵	%۶۷/۶۶	%۶۹/۸۶	۰/۱۵	
%۱۰۰	%۹۸/۷۵	%۶۳/۰۸	%۶۴/۸۶	۰/۲۰	
%۹۹/۵۴	%۹۳/۳۰	%۶۹/۹۵	%۶۷/۶۳	۰/۲۵	
%۹۹/۵۴	%۹۶/۹۵	%۶۴/۸۶	%۷۰/۷۷	۰/۳۰	
%۱۰۰	%۱۰۰	%۲/۹۵	%۴۶/۸۱	۰/۱۰	ماشین بردار پشتیبان - چندجمله‌ای (حالت Scale)
%۱۰۰	%۱۰۰	%۱۱/۳	%۴۶/۱۱	۰/۱۵	
%۱۰۰	%۱۰۰	%۸۷/۵	%۵۰	۰/۲۰	
%۱۰۰	%۱۰۰	%۱۲/۱۳	%۴۳/۶۴	۰/۲۵	
%۹۹/۵۴	%۱۰۰	%۸۷/۰۹	%۴۷/۱۴	۰/۳۰	
%۱۰۰	%۹۹/۴۹	%۷۷/۴۱	%۷۷/۲۲	۰/۱۰	ماشین بردار پشتیبان - گاوسی (حالت Auto)
%۱۰۰	%۹۹/۵۳	%۷۳/۲۳	%۷۱/۴۲	۰/۱۵	
%۱۰۰	%۹۹/۵۴	%۷۳/۸	%۶۷/۴۱	۰/۲۰	
%۱۰۰	%۹۸/۱۶	%۷۵/۳۹	%۷۹/۲۶	۰/۲۵	
%۱۰۰	%۹۸/۶۴	%۶۵/۳۶	%۷۰/۴۶	۰/۳۰	
%۱۰۰	%۱۰۰	%۷۷/۳۳	%۶۸/۵۴	۰/۱۰	ماشین بردار پشتیبان - گاوسی (حالت Scale)
%۱۰۰	%۹۹/۵۳	%۷۶/۱۴	%۷۰/۶۳	۰/۱۵	
%۹۹/۵۴	%۹۹/۰۳	%۷۴/۱۳	%۷۴/۷۷	۰/۲۰	
%۱۰۰	%۹۹/۱۸	%۷۸/۳۴	%۶۶/۶۶	۰/۲۵	
%۱۰۰	%۱۰۰	%۷۸/۴	%۶۶/۵۱	۰/۳۰	
%۹۹/۲۸	%۶۸/۹۹	%۶۴/۶۹	%۵۴/۴۸	۰/۱۰	تابع عضویت گاوسی
%۹۹/۲۸	%۷۰/۱۶	%۶۳/۰۸	%۵۶/۲۷	۰/۱۵	
%۹۸/۷۴	%۷۱/۸۶	%۶۰/۷۵	%۵۵/۹۱	۰/۲۰	
%۹۸/۹۲	%۷۲/۵۸	%۵۵/۹۱	%۵۷/۱۶	۰/۲۵	
%۹۸/۹۲	%۷۲/۷۵	%۵۶/۰۹	%۵۷/۸۸	۰/۳۰	

برخی نرخ‌ها، میزان حساسیت ارائه نشده است. در مورد برخی دیگر نیز، نرخ نهان‌نگاری واقعی ممکن است تفاوت بسیار اندکی داشته باشد.

PEAS از دید بار پردازشی، در مقایسه با تحقیقات مشابه، دارای کمترین ابعاد است. برخی از الگوریتم‌ها مانند [۳۴ و ۶] ابعاد پردازشی بالایی دارند و در روش‌هایی مانند [۳۵، ۳۶ و ۳۷]

۴-۳- بحث

به‌منظور ارزیابی و درک بهتر توانایی‌های PEAS، این الگوریتم در جدول (۷)، از نظر ابعاد پردازشی و بر اساس معیارهای حساسیت و ویژگی، با الگوریتم‌های مشابه مقایسه شده است. این نکته قابل ذکر است که برخی از کارهای مقایسه شده، دقیقاً در همان نرخ‌های نهان‌نگاری PEAS، نهان‌نگاری نشده‌اند. بنابراین، در

نیز، تعداد ابعاد حداقل دو برابر PEAS است. نکته مهم دیگر این است که هیچ کدام از تحقیقات مشابه، میزان دقت خود را در کشف فایل‌های پاک اعلام نکرده‌اند. با صرف نظر از این شرط و استفاده از ماشین بردار پشتیبان، با هسته چندجمله‌ای، حساسیت میانگین PEAS به $97/702\%$ و حساسیت آن برای پایین‌ترین نرخ نهان‌نگاری ($12/5$) به $95/28\%$ افزایش می‌یابد. البته در ازای این حساسیت بالا، میزان ویژگی به $8/4\%$ کاهش می‌یابد که البته به دلیل تعداد بسیار بالای هشدارهای اشتباه، این پیاده‌سازی قابل پذیرش نمی‌باشد.

بر اساس نتایج حساسیت گزارش شده در جدول (۶)، عملکرد نهان‌کاوی PEAS، بیش از 9% بهتر از [۳۶] در پایین‌ترین نرخ نهان‌نگاری ($12/5$) است. هیچ‌یک از تحقیقات مشابه میزان

نیز، تعداد ابعاد حداقل دو برابر PEAS است. نکته مهم دیگر این است که هیچ کدام از تحقیقات مشابه، میزان دقت خود را در کشف فایل‌های پاک اعلام نکرده‌اند. با صرف نظر از این شرط و استفاده از ماشین بردار پشتیبان، با هسته چندجمله‌ای، حساسیت میانگین PEAS به $97/702\%$ و حساسیت آن برای پایین‌ترین نرخ نهان‌نگاری ($12/5$) به $95/28\%$ افزایش می‌یابد. البته در ازای این حساسیت بالا، میزان ویژگی به $8/4\%$ کاهش می‌یابد که البته به دلیل تعداد بسیار بالای هشدارهای اشتباه، این پیاده‌سازی قابل پذیرش نمی‌باشد.

بر اساس نتایج حساسیت گزارش شده در جدول (۶)، عملکرد نهان‌کاوی PEAS، بیش از 9% بهتر از [۳۶] در پایین‌ترین نرخ نهان‌نگاری ($12/5$) است. هیچ‌یک از تحقیقات مشابه میزان

جدول (۷): جدول مقایسه عملکرد

نام	نوع	دسته‌بند	تعداد ابعاد تحلیلی	نرخ تشخیص				
				حساسیت (%)				
				نرخ نهان‌نگاری (%)				
ویژگی (%)	۱۲,۵	۲۵	۳۷,۵	۵۰	۱۰۰			
Kocal et al. [۳۴]	گفتار	SVM	۲۲	-	-	-	۸۴	۹۱
Miao et al. [۳۵]	گفتار	SVM	۲	-	۸۵	-	۹۵	-
PEAS steganalysis	گفتار	GMF	۱	۷۰,۲۵	۶۴,۱۵	۶۷,۲	۸۹,۷۸	۹۹,۸۲
Paulin et al. [۶]	صوتی	DBN	۲۵ ضریب + ۲۵۶ گاوسی	-	-	-	-	۸۲,۲۵
Jayasree & [۳۶] Amritha	VoIP	SVM	۲	-	۵۵	۷۳	-	-
Tian et al. [۳۷]	VoIP	SVM	۳	-	-	-	-	۹۲

توجه بسیاری از مهاجمان و قانون‌شکنان را به خود جلب نموده است. از سوی دیگر، تنوع پایین الگوریتم‌های نهان‌کاوی گفتار و مهم‌تر از آن، دستیابی به حساسیت پایین، با وجود استفاده از تعداد ابعاد پردازشی بالا، استفاده این الگوریتم‌ها را به‌ویژه در کاربردهای پردازش زمان واقعی، غیرعملی کرده است. در این مقاله، یک روش نوین برای کشف پیام‌های مخفی در نمونه‌های صوتی گفتاری ارائه شده است. این روش از ویژگی "درصد نمونه‌های مجاور یکسان" برای جداسازی نمونه‌های گفتاری پاک از حامل استفاده می‌کند. گام‌های نهان‌کاوی گفتاری PEAS

۵- نتیجه‌گیری

هدف نهان‌نگاری، مخفی نگه‌داشتن تبادل پیام‌های محرمانه و در نقطه مقابل آن، فلسفه نهان‌کاوی، کشف این تبادل‌های مخفیانه است. در حالی که تمامی رسانه‌های دیجیتال، پتانسیل استفاده در مخفی‌سازی اطلاعات را دارند، تصویر، ویدئو و صدا، ظرفیت و ویژگی‌های روشی مطلوب‌تری را برای این منظور ارائه می‌دهند. به‌عنوان یک زیردسته از نهان‌کاوی صوتی، نهان‌کاوی گفتار به دلیل استفاده گسترده در اینترنت و کاربردهایی مانند VoIP،

- [7] M. Shahrezaei, and N. Razmjoei, "A New Method for Image Steganography Using Discrete Wavelet Transforms," *Journal of electronical and cyber defence*, vol. 7, pp. 83-92, 2019. (In Persian)
- [8] W. Bender, D. Gruhl, N. Morimoto, and A Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, pp. 313-336, 1996.
- [9] W. Zeng, H. Ai, and R. Hu, "A novel steganalysis algorithm of phase coding in audio signal," presented at the Sixth International Conference on Advanced Language Processing and Web Information Technology, 2007.
- [10] R. M. Nugraha, "Implementation of direct sequence spread spectrum steganography on audio data," *International Conference on Electrical Engineering and Informatics*, 2011.
- [11] Y. C. Qi, L. Ye, and C. Liu, "Wavelet domain audio steganalysis for multiplicative embedding model," *International Conference on Wavelet Analysis and Pattern Recognition*, 2009.
- [12] I. Avciabas, "Audio steganalysis with content-independent distortion measures," *IEEE Signal Process. Lett.*, vol. 13, pp. 92-95, 2006.
- [13] M. K. Johnson, S. Lyu, and H. Farid, "Steganalysis of recorded speech," *Security, Steganography, and Watermarking of Multimedia Contents VII*, vol. 5681, pp. 664-672, 2005.
- [14] H. Ghasemzadeh and M. K. Arjmandi, "Universal audio steganalysis based on calibration and reversed frequency resolution of human auditory system," *IET Signal Process.*, vol. 11, pp. 916-922, 2017.
- [15] X. M. Ru, H. J. Zhang, and X. Huang, "Steganalysis of audio: Attacking the steghide," *International Conference on Machine Learning and Cybernetics*, 2005.
- [16] J. W. Fu, Y. C. Qi, and J. S. Yuan, "Wavelet domain audio steganalysis based on statistical moments and PCA," *International Conference on Wavelet Analysis and Pattern Recognition*, 2007.
- [17] O. H. Kocal, E. Yuruklu, and E. Dilaveroglu, "Speech steganalysis based on the delay vector variance method," *Turk. J. Elec. Eng. & Comp. Sci.*, vol. 24, pp. 4129-4141, 2016.
- [18] Q. Liu, A. H. Sung, and M. Qiao, "Temporal derivative-based spectrum and mel-cepstrum audio steganalysis," *IEEE Trans. Inf. Forensics Secur.*, vol. 4, pp. 359-368, 2009.
- [19] C. Kraetzer and J. Dittmann, "Mel-cepstrum based steganalysis for VoIP steganography," *Security, steganography, and watermarking of multimedia contents IV*, vol. 6505, 2007.
- [20] H. Ghasemzadeh, M. K. Arjmandi, "Reversed-Mel cepstrum based audio steganalysis", 4th International Conference on Computer and Knowledge Engineering, 2014.
- [21] C. Han, R. Xue, R. Zhang, and X. Wang, "A new audio steganalysis method based on linear prediction," *Multimed. Tools. Appl.*, vol. 77, pp. 15431-15455, 2018.
- [22] F. Djebbar, and B. Ayad, "Audio steganalysis based on lossless data-compression techniques," *International Conference on Information and Communications Security*, 2012.
- [23] Y. Tint, and K. T. Mya, "Audio steganalysis using features extraction and classification," *International Journal of Research and Reviews in Computer Science*, vol. 3, 2012.
- [24] H. Ghasemzadeh, M. T. Khass, and M. K. Arjmandi, "Audio steganalysis based on reversed psychoacoustic model of human hearing," *Digit. Signal Process.*, vol. 51, pp. 133-141, 2016.
- [25] U. Yavanoglu, B. Ozcakmak, and O. Milletsever, "A new intelligent steganalysis method for waveform audio files,"

عبارت‌اند از: محاسبه توان نرمال شده نمونه‌ها، جداسازی نواحی گفتار و سکوت، محاسبه ویژگی PEAS در نواحی گفتار و سکوت و نهایتاً دسته‌بندی نمونه فایل گفتاری به پاک یا حامل. روش ارائه‌شده، با استفاده از پنج دسته‌بند پیاده‌سازی شده که در بین آنها، تابع عضویت گاوسی بهترین نتایج را تولید کرده است. ارزیابی الگوریتم پیشنهادی، بر حسب دو معیار حساسیت و ویژگی، با تحلیل نمونه‌های پاک و حامل صورت گرفته است. بر اساس نتایج حاصل، قدرت ویژگی معادل $70/25\%$ و میزان حساسیت، به ترتیب برای نرخ‌های نهان‌نگاری $12/5\%$ ، 25% ، $37/15\%$ و 50% معادل $64/15\%$ ، $67/2\%$ ، $89/78\%$ و $99/82\%$ می‌باشند. هیچ‌یک از تحقیقات مشابه، اطلاعاتی در زمینه عملکرد خود در کشف نمونه‌های پاک ارائه نکرده‌اند. از این رو، عملکرد PEAS با آنها قابل مقایسه نیست. در زمینه حساسیت، PEAS عملکرد بهتری از تمامی الگوریتم‌های مشابه در نرخ‌های $12/5\%$ و 50% دارد. این نکته نیز قابل ذکر است که PEAS تنها بر مبنای یک بعد تحلیلی عمل می‌کند، درحالی‌که تعداد ابعاد پردازشی در الگوریتم‌های مشابه، بین ۲ تا ۲۸۱ متغیر است. علاوه بر کشف نمونه‌های حامل، PEAS قادر است در نرخ‌های نهان‌نگاری مختلف، با دقتی بین $56/27\%$ تا $99/28\%$ ، طول پیام پنهان شده را تخمین بزند. عملکرد این الگوریتم، علاوه بر نمونه‌های گفتاری، بر روی نمونه‌های یک پایگاه داده، متشکل از قطعات موسیقی کلاسیک نیز سنجیده شده که نتایج، حاکی از حساسیت‌های $31/2\%$ ، $3/86\%$ ، $51/6\%$ و $81/2\%$ در نرخ‌های $12/5\%$ تا 50% است.

۶- مراجع

- [1] H. Ghasemzadeh and M. H. Kayvanrad, "Comprehensive review of audio steganalysis methods," *IET Signal Process.*, vol. 12, pp. 673-687, 2018.
- [2] S. S. Chaeikar, M. Zamani, A. B. A. Manaf, and A. M. Zeki, "PSW statistical LSB image steganalysis," *Multimed. Tools Appl.*, vol. 77, pp. 805-835, 2018.
- [3] S. S. Chaeikar and A. Ahmadi, "Ensemble SW image steganalysis: A low dimension method for LSB detection," *Signal Process. Image Commun.*, vol. 70, pp. 233-245, 2019.
- [4] Y. Khorasani, J. Mazloum, and M. Shayesteh Far, "Increasing Insertion Capacity and Resistance in Image Steganography Based on Mapping and Bits Reduction," *Journal of electronical and cyber defence*, vol. 7, pp. 33-40, 2019. (In Persian)
- [5] S. S. Chaeikar, A. A. Manaf, A. Alarood, and M. Zamani, "PFW: Polygonal Fuzzy Weighted—An SVM Kernel for the Classification of Overlapping Data Groups," *Electronics*, vol. 9, p. 615, 2020.
- [6] M. Zamani, A. A. Manaf, R. Ahmad, F. Jaryani, H. Taherdoost, S. S. Chaeikar, and H. R. Zeidanloo, "A novel approach for genetic audio watermarking," *Journal of Information Assurance and Security*, vol. 5, pp. 102-111, 2010.

- [33] E. Yuruklu, O. H. Kocal, and E. Dilaveroglu, "A new approach for speech audio steganalysis using delay vector variance method," *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, vol. 19, pp. 27-36, 2014.
- [34] O. H. Koçal, E. Yuruklu, and I. & Avcibas, "Chaotic-type features for speech steganalysis," *IEEE Trans. Inf. Forensics Secur.*, vol. 3, pp. 651-661, 2008.
- [35] H. Miao, L. Huang, Y. Shen, X. Lu, and Z. Chen, "Steganalysis of compressed speech based on Markov and entropy," *International Workshop on Digital Watermarking*, 2013.
- [36] N. Jayasree and P. P. Amritha, "A Model for the Effective Steganalysis of VoIP," *Artificial Intelligence and Evolutionary Algorithms in Engineering Systems*, pp. 379-387, New Delhi, 2015.
- [37] H. Tian, Y. Wu, C. C. Chang, Y. Huang, J. Liu, T. Wang, and Y. Cai, "Steganalysis of analysis by synthesis speech exploiting pulse position distribution characteristics," *Security and communication networks*, vol. 9, pp. 2934-2944, 2016.
- [38] J. Chaharlang, M. Mosleh, and S. Rasouli-Heikalabad, "A novel quantum steganography-Steganalysis system for audio signals," *Multimed. Tools Appl.*, vol. 79, pp. 17551-17577, Feb. 2020.
- 11th International Conference on Machine Learning and Applications, 2012.
- [26] X. Yu, R. Wang, D. Yan, and J. Zhu, "MP3 audio steganalysis using calibrated side information feature," *J. Comput. Inf. Syst.*, vol. 8, pp. 4241-4248, 2012.
- [27] D. Yan, R. Wang, X. Yu, and J. Zhu, "Steganalysis for MP3Stego using differential statistics of quantization step," *Digit. Signal Process.*, vol. 23, pp. 1181-1185, 2013.
- [28] Y. Ren, T. Cai, M. Tang, and L. Wang, "AMR steganalysis based on the probability of same pulse position," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, pp. 1801-1811, 2015.
- [29] H. Miao, L. Huang, Z. Chen, W. Yang, and A. Al-Hawbani, "A new scheme for covert communication via 3G encoded speech," *Comput. Electr. Eng.*, vol. 38, pp. 1490-1501, 2012.
- [30] B. Geiser and P. Vary, "High rate data hiding in ACELP speech codecs," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.
- [31] Y. Ren, Q. Xiong, and L. Wang, "Steganalysis of AAC using calibrated Markov model of adjacent codebook," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2016.
- [32] H. Ozer, I. Avcibas, B. Sankur, and N. D. Memon, "Steganalysis of audio based on audio quality metrics," *Security and Watermarking of Multimedia Contents V*, vol. 5020, pp. 55-66, 2003.

Speech Steganalysis of Least Significant Bits Based on the Percentage of Equal Adjacent Samples

S. Yazdanpanah, M. Kheyrandish*, M. Mosleh

*Department of Computer Engineering, Dezful Branch, Islamic Azad University, Dezful, Iran

(Received: 06/04/2020, Accepted: 26/10/2020)

ABSTRACT

The popularity of audio formats usually attracts the attention of intruders and criminals to use this medium as a cover for establishing their secret communications. The extensive use of this formats, along with various modern techniques, designed for audio steganography, can cause the cyber spaces to be insecure environments. In order to deal with threats, some audio steganalysis techniques have been presented that statistically analyze various audio formats, such as music, MP3, and VoIP, efficiently. Among the presented approaches, combining the techniques of signal processing and machine learning has made possible the creation of steganalyzers that are highly accurate. However, since the statistical properties of audio files differ from purely speech ones, the current steganalysis methods cannot detect speech stego files, accurately. Another issue is the large number of analysis dimensions which increase the implementation cost, significantly. As response to these issues, this paper proposes the percentage of equal adjacent samples (PEAS) feature, as a one-dimensional feature for speech steganalysis. Using a classifier, based on the Gaussian membership function, on stego instances with 50% embedding ratio, the evaluation results for the designed steganalyzer, show a sensitivity of 99.82%. Additionally, it can efficiently estimate the length of a hidden message with the desirable accuracy. Also, the PEAS steganalysis was evaluated on a database, containing classic music instances, and the results show an 81.2% efficient performance.

Keywords: Speech Steganalysis, Audio Steganalysis, Digital Signal Processing, LSB, Steganography

* Corresponding Author Email: kheyrandish@iaud.ac.ir