نهان‌کاوی فایل‌های فشردهٔ صوتی با استفاده از یادگیری ماشین

محسن سلیمانی[1*] مهدی چهل امیرانی[2]، سیدجهانشاه کبودیان[3]

[1*] دانشجوی دکترای، دانشکدهٔ مهندسی برق و کامپیوتر، دانشگاه ارومیه، ارومیه

[2] استاد دانشکدهٔ مهندسی برق و کامپیوتر، دانشگاه ارومیه، ارومیه

[3] استادیار دانشکدهٔ مهندسی برق و کامپیوتر، دانشگاه رازی، کرمانشاه

چکیده

علم پنهان‌سازیِ پیام حاوی اطلاعات در یک رسانهٔ حامل را نهان‌نگاری و تلاش برای تشخیص وجود یا نبود پیام نهان‌شده در شیء پوششی را تحلیل نهان‌نگاری یا نهان‌کاوی می‌نامند. فرمت فشرده‌سازی MP3 در میان داده‌های صوتی به‌عنوان میزبانی مناسب و فراگیر برای نهان‌نگاری اطلاعات مورداستفاده قرار گرفته و شیوه‌های نهان‌نگاری مختلفی برای این منظور طراحی شده‌اند؛ در این پژوهش، هدف ارائهٔ الگوریتمی برای نهان‌کاوی به‌طور خاص برای فایل‌های فشردهٔ صوتی با قالب MP3 است که با نرم‌افزار MP3stego نهان‌نگاری شده‌اند. برای تهیهٔ دادگان نهان‌نگاری از فایل‌های متنی با متون تصادفی استفاده شده‌است. ابتدا با استفاده از اطلاعات جانبیِ مستخرج از فایل‌های MP3، ویژگی‌های لازم استخراج‌شده و دادگان صوتی که شامل دو دسته فایل‌های نهان‌نگاری‌شده و فایل‌های نهان‌نگاری‌نشده است، به دو بخش دادگان آموزش و دادگان آزمون تقسیم شده و در ادامه با استفاده از روش‌های یادگیری ماشین (ماشین بردار پشتیبان)، سامانه تشخیص فایل‌های آلوده و فایل‌های تمیز طراحی شده و درنهایت کارایی سامانه با استفاده از دادگان آزمون اندازه‌گیری می‌شود. در این مقاله، یک ویژگی جدید به‌نام قله‌داربودن طیف (SPK) از اطلاعات جانبی فایل MP3 استخراج می‌شود. سامانک پیشنهادی با استفاده از دادگان جداگانهٔ آزمون که شامل فایل‌های تمیز و فایل‌های نهان‌نگاری‌شده با ظرفیت‌های نهان‌نگاری متنوع است، آزمایش شده و با دقت ۱۰۰٪ و بدون خطا فایل‌های تمیز و آلوده را از هم متمایز می‌کند. نتایج حاصل حاکی از شناسایی دقیق موارد نهان‌نگاری‌شده درعین کاهش پیچیدگی محاسباتی و افزایش سرعت این نوع نهان‌کاوی نسبت به روش‌های ابداعی گذشته است.

واژگان کلیدی: فایل فشردهٔ صوتی، نهان‌کاوی فایل‌های صوتی، نهان‌نگاری در فایل‌های صوتی، MP3،MP3Stego

Steganalysis of Compressed Audio Files Based on Machine Learning

Mohsen Soleimani [1*], Mehdi Chehel Amirani [2], Seyed Jahanshah Kabudian [3]

PhD Student, Faculty of Electrical & Computer Engineering, Urmia University, Urmia, Iran[1]
Professor, Faculty of Electrical & Computer Engineering, Urmia University, Urmia, Iran[2]
Assistant Professor, Faculty of Electrical & Computer Engineering, Razi University, Kermanshah, Iran[3]

Abstract

The science of hiding a message containing information in a carrier medium is called steganography, and the attempt to detect the presence or absence of a hidden message in a cover medium is called steganalysis. The MP3 compression format has been used among audio data as a suitable and comprehensive host for information encryption, and various encryption methods have been designed for this purpose. In this research, the aim is to present an algorithm for audio ateganalysis, specifically for compressed audio files in MP3 format, in which some data has been embedded using MP3stego software. To prepare encrypted data, text files with random texts have been used. First, by using the side information extracted from MP3 files, the necessary features are extracted and the audio data, which includes two categories of stego files and clean files, is divided into two parts:

* Corresponding author

*نویسندهٔ عهده‌دار مکاتبات

• تاریخ ارسال مقاله: ۱۴۰۰/۷/۱۸ • تاریخ پذیرش: ۱۴۰۲/۱۲/۶ • تاریخ انتشار: ۱۴۰۳/۸/۸ • نوع مطالعه: پژوهشی

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.55]

training data and test data. And then, using machine learning techniques (support vector machine), the detection system of infected files and clean files is designed, and finally, the efficiency of the system is measured using the test data. In this paper, a new feature called spectral peakiness (SPK) is extracted from the side information of MP3 file. The proposed system was tested using separate test data, which includes clean files and stego files with various encryption capacities, and it distinguished clean and stego files with 100% accuracy and without error. The results indicate the perfect classification of stego and clean files while reducing the computational complexity and increasing the speed of steganalysis compared to other methods.

Instead of using the audio signal information stored in the MP3 file, the proposed method uses the side information of the MP3 file, which is less dependent on the audio content of the file. In this method, the MDB side information in the compressed audio file is assumed as a sequence, and then, using a feature extraction method, a new feature in the frequency domain called spectral peakiness is calculated. This simple yet powerful feature is combined with features such as temporal average and spectral average of the MDB sequence and forms a low-dimensional (three-dimensional) feature vector. This feature vector will then be classified by a support vector machine (SVM) classifier as a suspicious file or a normal file. The feature extraction method, while being simple and having very few calculations, has 100% accuracy (recognition without any error) for MP3 files, even when the amount of the hidden information in the audio file is very low.

Keywords: **Compressed Audio File, Audio Steganography, Audio Steganalysis, MP3, MP3stego.**

۱-مقدمه

امروزه اینترنت یکی از مهم‌ترین راه‌های ارتباطی بین افراد و سازمان‌هاست. گسترش اینترنت و وجود برخی افراد و نفوذگران که سعی در دست‌یابی به اطلاعات خصوصی افراد و سازمان‌ها را دارند و همچنین تلاش برخی از عناصر بیگانه در برقراری ارتباط با عوامل خود باعث ایجاد برخی روش‌های نهان‌نگاری شده‌است. نهان‌نگاری (استگانوگرافی[1]) هنر و علم پنهان‌سازی پیام یا اطلاعات طبقه‌بندی‌شده در داخل یک‌سری اطلاعات بدون طبقه‌بندی است[۱]؛ از طرفی مراقبت و کنترل بسترها و شبکه‌های داخلی و شبکه‌های مرتبط با اینترنت یکی از وظایف نهادها و سازمان‌های امنیتی، نظامی و انتظامی است. سهل‌انگاری و غفلت از این مهم، بستری کم‌هزینه، مطمئن، پرظرفیت و پنهان را برای سوءاستفاده‌کنندگان از این دانش و فن فراهم می‌آورد. نهان‌کاوی[2]، کشف و مقابله با عملیات نهان‌نگاری است و تحقیق‌وتوسعهٔ روش‌های آن می‌تواند به بهبود و ارتقای امنیت روش‌های موجود نهان‌نگاری نیز کمک کند[۲].

نهان‌نگاری و نهان‌کاوی از جهت بستر مورداستفاده، فن و الگوریتم به‌کاررفته و ویژگی‌های خاص دیگر تقسیم‌بندی و دسته‌بندی‌های متنوعی شده‌اند. نهان‌نگاری صوت، تصویر، متن، فیلم و شبکه‌ای، نهان‌نگاری LSB[3] (بدون توجه به بستر مورداستفاده)، نهان‌نگاری در حوزهٔ زمان و فرکانس، نهان‌نگاری در حوزهٔ فشرده و غیرفشرده از اقسام این دسته‌بندی‌ها در موضوع نهان‌نگاری است؛ همچنین نهان‌کاوی خاص‌وعام، نهان‌کاوی هدفمند و غیرهدفمند از اقسام دیگر تقسیم‌بندی در موضوع نهان‌کاوی است.

همان‌طور که اشاره شد یکی از انواع دسته‌بندی‌های نهان‌کاوی، خاص‌وعام‌بودن هدف نهان‌کاوی است. در نهان‌کاوی عام، هدف شناسایی و تحلیل هر نوع نهان‌نگاری و در نهان‌کاوی خاص، هدف شناسایی و تحلیل یک‌نوع خاص از نهان‌نگاری است.

روشن است که درصورت شناسایی و تحلیل نهان‌نگاری‌های متنوع به‌صورت خاص می‌توان یک نهان‌کاو عام بهینه تولید کرد؛ درصورتی‌که الگوریتم نهان‌نگاری را بدانیم، استفاده از نهان‌کاوی خاص به‌سرعت و دقت عملیات کمک می‌کند.

در بررسی‌های مختلف ثابت شده که در نهان‌نگاری فایل‌های فشردهٔ صوتی، بسته به‌نوع و روش نهان‌نگاری ممکن است، ردپای نهان‌نگاری، هم در مقایسه با سیگنال صوتی اولیه و هم در مشخصات آماری پارامترهای تولیدشده در جریان فشرده‌سازی از جمله اطلاعات جانبی[4] پدیدار شود. بسته به نوع روش‌های نهان‌نگاری، تغییراتی در اطلاعات جانبی، داده‌های اصلی و محتوای پوشانهٔ[5] سیگنال رخ می‌دهد و به‌طورمعمول ردپایی از نهان‌نگاری برجای می‌ماند و ازاین‌جهت است که می‌تواند موردشناسایی قرارگیرد.

روش‌های نهان‌کاوی فایل‌های صوتی را نیز می‌توان در دو دستهٔ روش‌های نهان‌کاوی عام و روش‌های نهان‌کاوی خاص قرارداد؛ در دستهٔ نخست، فرض براین است که هر

[1] Steganography
[2] steganalysis
[3] Least Significant Bit Matching
[4] Side Information
[5] Covertext

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.55]

روش نهان‌نگاری به‌طورتقریبی ویژگی‌های آماری سیگنال حامل را تاحدی خراب می‌کند. در این روش‌ها اختلاف بین سیگنال نهان‌نگاری‌شده و سیگنال حامل اندازه گرفته می‌شود؛ مانند نهان‌کاوی براساس معیارهای کیفی صوت، نهان‌کاوی براساس ویژگی‌های آشوب‌گونه، نهان‌کاوی مبتنی بر ویژگی‌های مستقل از محتوا و نظیر آن. در دستهٔ دوم، الگوریتم پیشنهادی تنها برای یک روش نهان‌نگاری خاص طراحی شده‌است؛ مثل نهان‌کاوی در فاز، نهان‌کاوی در پژواک و نهان‌کاوی MP3stego و غیره[۳].

در فرایند نهان‌کاوی ابتدا با بررسی خصوصیات فایل حامل، در حالت تمیز و نهان‌نگاری‌شده، باید به‌دنبال ویژگی‌هایی بود که از لحاظ آماری بهترین تمایز را بین دودسته سیگنال پوشانه و سیگنال نهانه[1] برقرار کنند؛ سپس با ویژگی‌های استخراج‌شده، بهترین طبقه‌بندی ممکن انجام می‌شود. تاکنون روش‌های مختلفی برای نهان‌کاوی فایل‌های فشردهٔ صوتی MP3 ارائه شد است[۳]. یکی از مهم‌ترین روش‌ها، نهان‌کاوی با استفاده از اطلاعات جانبی است. اطلاعات جانبی، بخشی از فایل‌های فشردهٔ صوتی MP3 است که در جریان فشرده‌سازی تولید می‌شود و شامل بخش‌های گوناگونی است. اصول تعبیه‌سازی در MP3Stego استفاده از توازن متغیرِ طول بلوک برای تعبیه است. فایل‌های کدگذاری و کدگشایی برای تعبیهٔ اطلاعات و یا استخراج پیام تعبیه‌شده در دسترس است و یکی از بخش‌هایی که درحین فرایند نهان‌نگاری دچار تغییر می‌شود، متغیر اشاره‌گر main_data_begin (mdb) است.

در این پژوهش، هدف ارائهٔ الگوریتمی برای نهان‌کاوی خاص فایل‌های فشردهٔ صوتی MP3 است که با نرم‌افزار MP3stego نهان‌نگاری شده‌اند. انتخاب نوع و تعداد ویژگی‌ها و تعداد نمونه‌های آزمایشی و همچنین الگوریتم استفاده‌شده برای یادگیری ماشین در بهبود نتیجه‌گیری در سامانه ارائه‌شده بسیار مؤثر است. در این‌گونه پژوهش‌ها هدف بهینه‌سازی تشخیص و تفکیک نمونه‌های آزمودنی با کمترین هزینه و زمان است؛ لذا درصورتی‌که بتوانیم ضمن افزایش دقت آزمایش، زمان نتیجه‌گیری را کاهش دهیم، موفق‌تر عمل کرده‌ایم.

برای انجام پژوهش از نرم‌افزار MP3Stego برای نهان‌نگاری، از بانک‌دادهٔ دورتموند برای تأمین فایل‌های استاندارد MP3، از یک برنامهٔ تولیدکنندهٔ فایل‌های تصادفی متنی برای تعبیه در فایل‌های فشردهٔ صوتی و از یک برنامهٔ آنالیز فایل‌های MP3 در محیط نرم‌افزار متلب استفاده شده‌است. همچنین مراحل آموزش و آزمون سامانه با استفاده از ماشین بردار پشتیبان SVM[2] انجام شده‌است.

باتوجه‌به قدمت نرم‌افزار MP3Stego آزمایش‌هایی درگذشته برای نهان‌کاوی خاص این نرم‌افزار صورت‌گرفته‌است. این نرم‌افزار هم روی اطلاعات جانبی و هم روی ضرایب MDCT[3] تأثیر می‌گذارد و از این تغییرات برای کشف نهان‌نگاری استفاده شده‌است. در آزمایش‌های انجام‌شده در مرجع [۴] مقدار میانگین دقت آزمون در حالت استفاده از متغیر اشاره‌گر main_data_begin عدد ۹۲٪/۷ گزارش شده‌است.

در پژوهش دیگر انجام‌شده در مرجع [۵] ادعا شده‌است که با تحلیل آمارگان متغیر طول part2_3_length می‌توان از وجود و یا نبود دادهٔ مخفی آگاه شد؛ لیکن دقت این کار مطلوب نیست.

در گزارش ارائه‌شدهٔ دیگری در مرجع [۶] از تفاضل واریانس متغیر big_value فایل نهانه و فایل پوشانه برای تمایز فایل‌های تمیز و آلوده استفاده شده‌است.

در پژوهش انجام‌شده در مرجع [۷] ادعا شده که با استفاده از ضرایب MDCT نهان‌کاوی درحالت نرخ پایین تعبیه با MP3Stego انجام شده‌است.

همچنین در مرجع [۱۵] از ویژگی‌های کپستروم استخراج‌شده از بانک فیلتر با مقیاس فرکانسی معکوس مل برای نهان‌کاوی فایل‌های صوتی استفاده شده‌است.

۲-مفاهیم و مبانی نظری

۲-۱-تعریف مفاهیم و اصطلاحات

استگانوگرافی: استگانوگرافی در اصل کلمه‌ای با ریشهٔ یونانی است که از دوکلمهٔ Steganos به‌معنی پوشیده و graphy به‌معنی نوشتن تشکیل‌شده‌است که معادل فارسی آن نهان‌نگاری است[۱].

نهانه: به پیغامی که از درج داده‌های پیام محرمانه در داخل داده‌های پوشانه ایجاد می‌شود، «نهانه» گویند[۱].

پوشانه: پیغامی که داده‌های پیام محرمانه در داخل داده‌های آن درج می‌شوند، «پوشانه» نامیده می‌شود[۱].

نهان‌نگاری: نهان‌نگاری پنهان‌کردن پیام از دید دشمن است و یا دقیق‌تر هنر و علم پنهان‌سازی پیام یا

[1] Stegotext

[2] Support Vector Machine

[3] Modified Discrete Cosine Transform

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.55]

اطلاعات طبقه‌بندی‌شده در داخل یک‌سری اطلاعات بدون طبقه‌بندی را نهان‌نگاری گویند[۱].

نهان‌کاوی: تلاش برای تشخیص وجود یا نبود پیام نهان‌شده در شئ پوششی را تحلیل نهان‌نگاری یا نهان‌کاوی گویند[۱].

نهان‌کاوی خاص‌وعام: در نهان‌کاوی عام، هدف شناسایی و تحلیل هر نوع نهان‌نگاری و در نهان‌کاوی خاص، هدف شناسایی و تحلیل یک نوع خاص از نهان‌نگاری است[۲].

آب‌نشان‌گذاری[1]: آب‌نشان‌گذاری یا نقش‌برآب‌زنی روشی برای حفظ حق مالکیت است که توسط آن نشانه‌ای از صاحب اثر به‌گونه‌ای در اثر تعبیه می‌شود که بدون ایجاد لطمه به کیفیت اثر به‌راحتی از آن قابل‌جداکردن نباشد و مشخص‌کنندهٔ منشأ و یا صاحب اثر باشد[۹].

۲-۲- چارچوب‌های نظری

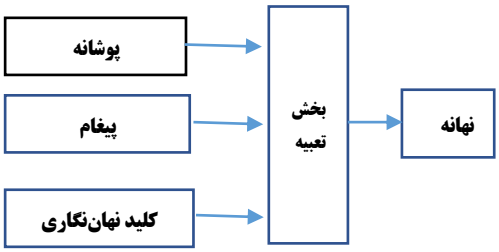
یک سامانه نهان‌نگاری دارای سه ویژگی اساسی شفافیت[2]، ظرفیت[3] و مقاومت[4] است. منظور از شفافیت غیرقابل‌تشخیص‌بودن نهانه از پوشانه از دید ناظر انسانی است؛ البته شفافیت شاخص‌های آماری نیز به‌عنوان یکی دیگر از ویژگی‌ها و ملزومات نهان‌نگاری مطرح شده‌است. ظرفیت، مقدار داده‌ای است که می‌تواند با استفاده از یک روش پنهان‌سازی اطلاعات تعبیه شود و مقاومت به این معنی است که نهانه درمقابل تهدیدات عمدی (نظیر فیلترکردن، فشرده‌سازی و ازاین‌قبیل) و تهدیدات غیرعمدی (نظیر اعوجاج کانال، تداخل و ازاین‌قبیل) مقاوم باشد. البته نهان‌نگاری ملزومات اساسی دیگری نیز از قبیل امنیت و برگشت‌پذیری و پیچیدگی عملیات دارد[۱]. این ویژگی‌ها در تقابل با یکدیگرند و در شرایط عادی باید به تمام این ویژگی‌ها توجه کرد؛ اما در بعضی کاربردها ناچاریم به‌بعضی از ویژگی‌ها اهمیت بیشتری دهیم. نهان‌نگاری و پنهان‌سازی اطلاعات کاربردهای متنوع و زیادی دارد و می‌تواند در پوشانه‌های مختلفی از قبیل پوشانه‌های نوشتاری، صوتی، تصویری و غیره انجام پذیرد.

۲-۲-۱- نهان‌نگاری صوتی

به‌دلیل محدودیت‌های سامانه شنوایی انسان، صوت نیز می‌تواند بستر مناسبی برای نهان‌نگاری باشد. سامانه شنوایی انسان می‌تواند محدودهٔ فرکانسی خاصی را درک کند؛ همچنین نسبت به پارازیت‌هایی که اضافه می‌شود بسیار حساس است. هرگونه مزاحمتی در یک فایل صوتی قابل‌درک است حتی اگر خیلی ناچیز باشد. به‌هرحال با اینکه سامانه شنوایی انسان حساسیت زیادی دارد، اما دربرابر بعضی تغییرات حساسیت خود را از دست می‌دهد. مثلاً صداهای بلند صداهای آرام را درخود جای می‌دهند[۱۰].

برای استفاده از پوشانهٔ صوتی برای مخفی‌کردن اطلاعات دو نکته را باید درنظرداشت: نخست باید از ضعف سامانه شنوایی انسان استفاده‌کرد و دوم توجه خاصی به‌حساسیت فوق‌العادهٔ آن داشت[۱۱].

نهان‌نگاری صوتی در دو حوزهٔ فشرده و غیرفشرده ممکن است؛ در حوزهٔ غیرفشرده به‌دلیل حجیم‌بودن فایل می‌توان بدون اینکه سیگنال صوت دارای اعوجاج قابل تشخیصی باشد، اطلاعات را نهان‌نگاری کرد؛ البته به‌همین‌میزان تشخیص آن نیز کار آسان‌تری است؛ زیرا تغییرات عمدی در سیگنال بدون اینکه اثر انگشت آماری قابل ردیابی درون سیگنال بگذارد تقریباً محال است، اما در حوزهٔ فشرده، به‌دلیل وجود ظرفیت‌های متنوع، کار نهان‌نگاری و بالطبع نهان‌کاوی پیچیده‌تر است.



(شکل-۱): شمای کلی نهان‌نگاری صوتی
(Figure-1): Overall Schematic for Audio Steganography

۲-۲-۲- فایل فشردهٔ صوتی MP3

الگوریتم‌های کدگذاری صوت یا فشرده‌سازی صوت برای به‌دست‌آوردن یک نمایش دیجیتال و فشرده از سیگنال‌های صوتی باکیفیت بالا استفاده می‌شوند تا درنهایت بتوانیم انتقال یا ذخیره‌سازی صوت را با کارایی بالاتری انجام دهیم. هدف اصلی در کدگذاری صوت، نمایش سیگنال با کمینه تعداد ممکن بیت است، درعین‌این‌که بتوان بازتولید سیگنال را با وضوح بالایی انجام داد؛ یعنی صوت خروجی را طوری تولید کنیم که حتی با گوش حساس نیز قابل‌تفکیک از سیگنال اولیهٔ ورودی نباشد. فرمت MPEG-2 و یا لایهٔ سوم MPEG-1 که در بین کاربران با نام MP3 شناخته می‌شود، مشهورترین فرمت انتشار صوت بر روی اینترنت است که دارای نرخ فشرده‌سازی بالا و کیفیتی مطلوب است.

[1] Watermarking
[2] Transparency
[3] Capacity
[4] Robustness

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.55]

به‌خاطر ویژگی‌های این نوع فایل‌ها می‌توان آن‌ها را یک حامل مناسب برای کاربردهای نهان‌نگاری درنظرگرفت. مهم‌ترین این ویژگی‌ها عبارت‌اند از محبوبیت و رواج بالای این فرمت، تأمین ظرفیت مناسب برای نهان‌نگاری و توجه به این مطلب که فشرده‌سازی MP3 همراه با اتلاف و افت است و این افت باعث می‌شود که سیگنال تولیدی از لحاظ آماری ویژگی‌های مشخصی نداشته باشد (البته درصورت دسترسی‌نداشتن به فایل حامل اصلی) و می‌تواند به‌عنوان یک حامل با امنیت بالا به‌کار رود[۸].

۲-۲-۳- آناتومی فایل فشردهٔ صوتی MP3

تمام فایل‌های MP3 به قطعه‌های کوچکی تقسیم می‌شوند که فریم نام دارند. هر فریم تعداد ۱۱۵۲ نمونهٔ صوتی را در خود ذخیره می‌کند و ۲۶ میلی‌ثانیه به‌طول می‌انجامد؛ علاوه‌برآن، یک فریم به دو گرانول تقسیم می‌شود که هر گرانول شامل ۵۷۶ نمونه است.

هر فریم شامل پنج بخش سرایند[۱]، CRC[۲]، اطلاعات جانبی[۳]، دادهٔ اصلی[۴] و دادهٔ کمکی[۵] است[۱۱].

Ancillary Data	Main Data	Side Information	CRC	Header

(شکل-۲): چیدمان یک فریم از فایل MP3 [۱۲]

(Figure-2): Information structure for one frame of MP3 file

سرایند ۳۲ بیت دارد و شامل یک کلمهٔ هم‌زمان‌سازی به‌همراه توصیفی از فریم است. اگر بیت حفاظت در سرایند به یک تنظیم شود بخش CRC وجود دارد و امکان بررسی داده‌های با حساسیت بالا را برای خطاهای انتقال فراهم می‌کند. بخش اطلاعات جانبی فریم شامل اطلاعات موردنیاز برای کدگشایی دادهٔ اصلی است. بخش‌های متفاوت اطلاعات جانبی در شکل (۳) نشان داده شده‌اند.

Main_data_bigin	Private_bits	scfsi	Side_info_gr.0	Side_info_gr.1

(شکل-۳): ساختار اطلاعات جانبی [۱۲]

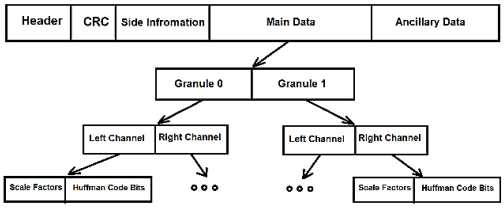
(Figure-3): Structure of Side Information

بخش دادهٔ اصلی شامل ضرایب مقیاس، بیت‌های کدشدهٔ هافمن و بخش دادهٔ کمکی اختیاری بوده و برای ثابت‌ماندن نرخ بیت فایل MP3 استفاده می‌شود.

۲-۲-۴-نهان‌نگاری در فایل‌های MP3

شیوه‌های نهان‌نگاری در فایل‌های MP3 به دو گروه از راه‌کارها تقسیم می‌شوند: نخست راهکارهایی که اطلاعات را در حوزهٔ فشرده و در رشته‌بیت MP3 پنهان می‌کنند، دوم راهکارهایی که اطلاعات را در طی مراحل فشرده‌سازی یک فایل به‌صورت MP3 در آن پنهان می‌کنند. گفتنی است دستهٔ دیگری از روش‌ها نیز وجود دارند که اطلاعات را در فایل صوتی فشرده‌نشده پنهان می‌کنند، اما باتوجه‌به اینکه مقاوم‌سازی دربرابر فشرده‌سازی نیازمند محاسبات و تدابیر پیچیده‌ای است، از این روش نهان‌نگاری به‌ندرت استفاده می‌شود[۱۱].

یکی از روش‌های معروف و در دسترس نهان‌نگاری فایل‌های MP3، روش MP3Stego است این نرم‌افزار، نهان‌نگاری را درحین فشرده‌سازی انجام می‌دهد[۳].



(شکل-۴): سازماندهی ضرایب مقیاس در گرانول‌ها و کانال‌ها [۱۲]

(Figure-4): Structure of scale factors in granules and channels

۲-۲-۵- نرم‌افزار MP3Stego

همان‌طور که اشاره شد، این نرم‌افزار پیغام محرمانه را هم‌زمان با روند فشرده‌سازی پنهان می‌کند. داده‌ها ابتدا رمزگذاری و فشرده‌سازی می‌شوند و سپس در رشته‌بیت MP3 پنهان می‌شوند. روند پنهان‌سازی اطلاعات در دل فرایند کدگذاری MP3، به‌نام حلقهٔ داخلی اتفاق می‌افتد. حلقهٔ داخلی داده‌های ورودی را چندی‌سازی کرده و گام چندی‌سازی را یک واحد افزایش می‌دهد، تا اینکه داده‌های چندی‌سازی‌شده با ارقام دودویی کدگذاری شوند. همچنین یک حلقهٔ بیرونی وجود دارد که نوفه چندی‌سازی را بررسی می‌کند تا از مقدار آستانهٔ تعیین‌شده توسط بخش روان‌شنیداری تجاوز نکند. روش رمزگذاری فشرده شامل دوحلقهٔ تودرتو است؛ به‌این‌ترتیب که حلقهٔ درونی مقدار دقیق داده‌ها را تعیین کرده و کوانتایزر مناسبی را که با بودجهٔ موجود در دسترس است، تعیین می‌کند و از سوی دیگر، حلقهٔ بیرونی، اعوجاج[۶] فرایند رمزگذاری را کنترل می‌کند. نهان‌نگار MP3Stego وضعیت نهایی حلقهٔ درونی را تغییر می‌دهد و پیام را در طول فرایند فشرده‌سازی پنهان می‌کند. همان‌طورکه اشاره شد نهان‌نگاری با این نرم‌افزار، هم بر روی اطلاعات جانبی و هم بر ضرایب تبدیل کسینوسی گسستهٔ اصلاح‌شده (MDCT[7]) تأثیر می‌گذارد.

[1] Header
[2] Cyclic Redundancy Check(CRC)
[3] Side Information
[4] Main Data
[5] Ancillary Data

[6] Distortion
[7] Modified Discrete Cosine Transform

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.55]

۲-۲-۶-نهان‌کاوی فایل‌های صوتی MP3 مبتنی بر متغیرهای اطلاعات جانبی

نرم‌افزار Mp3stego تعبیه‌سازی را با تغییر طول بلوک گرانول انتخاب‌شده انجام می‌دهد. این فرایند تعبیه‌سازی روی متغیرهای اطلاعات جانبی مانند متغیر main data begin، گام چندی‌سازی، طول بلوک، متغیر big_value و میانگین مخزن تأثیر می‌گذارد[۸]. در هر حلقهٔ تکراری که شرط تطابق توازن طول بلوک و بیت مخفی برقرار نباشد، گام چندی‌سازی یک‌واحد افزایش می‌یابد؛ افزایش گام چندی‌سازی باعث می‌شود ضرایب MDCT با گام درشت‌تر کوانتیزه شوند؛ بنابراین مقدار آن‌ها کوچک‌تر می‌شود و می‌توانند با تعداد بیت کمتری کد شوند. از طرفی سازوکار مخزن بیت در MP3 امکان ذخیره و قرض را برای فریم‌ها فراهم کرده‌است. در حالت نهان‌نگاری چون تعداد بیت موردنیاز برای کدکردن ضرایب تبدیل کسینوسی گسستهٔ اصلاح‌شده کاهش می‌یابد؛ بنابراین فضای کمتری از فریم صرف کدکردن دادهٔ اصلی می‌شود و مابقی بیت‌های دردسترسِ فریم جاری ذخیره می‌شوند. دادهٔ اصلی فریم بعدی بلافاصله بعد از دادهٔ اصلی فریم قبل‌تر شروع می‌شود[۳]؛ بنابراین متغیر اشاره‌گر (mdb) main_data_begin تغییر می‌کند.

متغیر main_data_begin که یک اشاره‌گر منفی است و به ابتدای دادهٔ اصلی فریم جاری اشاره می‌کند، در حالت نهانه درمقایسه با حالت پوشانه، دارای مقدار بزرگ‌تری است.

۳-روش‌شناسی

در این تحقیق، در ابتدا روش‌های فشرده‌سازی صوت و فشرده‌سازی MP3 مطالعه و بررسی شده و در ادامه معروف‌ترین روش نهان‌نگاری صوتی در دسترس Mp3stego موردمطالعه و ارزیابی قرارگرفت. سیگنال صوت بیش‌تر ماهیت دوره‌ای دارد و می‌توان گفت هرگونه دست‌کاری و تغییر غیرمجاز می‌تواند ردپاهای قابل ردگیری روی سیگنال بگذارد[۳]. نهان‌نگار MP3stego نیز در دوبخش اطلاعات جانبی سیگنال و اطلاعات اصلی تأثیر می‌گذارد که در این پژوهش نهان‌کاوی با استفاده از آنالیز بخش اطلاعات جانبی صورت‌گرفته‌است.

در ابتدا با بررسی و تحلیل قسمت‌های مختلف اطلاعات جانبی، ویژگی مناسب برای نهان‌کاوی انتخاب شده‌است و با استفاده از دانش یادگیری ماشین به آموزش سامانه تصمیم‌گیر پرداخته و در پایان به ارزیابی سامانه پیشنهادی پرداخته‌ایم.

۳-۱-استخراج ویژگی

هدف از این پژوهش، ارتقای دقت و افزایش سرعت شناسایی فایل‌های فشردهٔ صوتیMP3 است که با نرم‌افزار MP3Stego نهان‌نگاری شده‌اند. ازآنجاکه این نهان‌نگار تعبیهٔ پیغام را درحین کوانتزاسیون انجام می‌دهد، بخش‌هایی از اطلاعات جانبی دچار تغییرات می‌شوند و می‌توان متناسب با روش نهان‌نگاری از هریک از این متغیرها برای استخراج ویژگی و آشکارکردن نهان‌نگاری استفاده کرد. در این پژوهش پس از بررسی‌ها و آزمایش‌های مختلف بر روی متغیرهای اطلاعات جانبی، متغیر اشاره‌گر main_data_begin (mdb) برای این کار مناسب تشخیص داده شد. بدین صورت که دنبالهٔ mdb به ازای قاب‌های متوالی فایل صوتی به‌صورت یک سیگنال زمانی به‌نام $x_{mdb}[m]$ فرض می‌شود و یک بردار ویژگی سه‌بُعدی FeatVec از آن استخراج می‌شود (m شمارهٔ قاب است). عنصر نخست از این بردار ویژگی، یعنی m_{mdb} میانگین متغیر mdb در حوزهٔ زمان است. عنصر دوم از این بردار ویژگی، یعنی μ_X میانگین اندازهٔ طیف فوریهٔ متغیر mdb در حوزهٔ فرکانس است. عنصر سوم از این بردار ویژگی، یعنی SPK_{mdb} یک ویژگی ارائه‌شده و جدید در حوزهٔ طیف است و میزان قله‌داربودن طیف یا Spectral Peakiness نامیده می‌شود؛ درنهایت، بردار سه‌بعدی FeatVec استخراج‌شده از روی دنبالهٔ mdb فایل صوتی mp3 به‌یک ماشین بردار پشتیبان (مرزی) داده می‌شود و فایل صوتی به دو کلاس تمیز (Plain) و آلوده (Stego) دسته‌بندی می‌شود. روابط دقیق استخراج بردار ویژگی از روی متغیر اطلاعات جانبیِ mdb به‌صورت زیر است:

$$x_{\mathrm{mdb}}[m] = \text{Main Data Begin of } m\text{-th frame}$$

$$\mu_{\mathrm{mdb}} = \frac{1}{M}\sum_{m=0}^{M-1} x_{\mathrm{mdb}}[m] \qquad (۱)$$

$$x'_{\mathrm{mdb}}[m] = x_{\mathrm{mdb}}[m] - \mu_{\mathrm{mdb}} \qquad (۲)$$

$$X_{\mathrm{mdb}}[k] = \left|\sum_{m=0}^{M-1} x'_{\mathrm{mdb}}[m] e^{-j\frac{2\pi km}{M}}\right| \qquad (۳)$$

$$\mu_X = \frac{2}{M}\sum_{k=1}^{M/2} X_{\mathrm{mdb}}[k] \qquad (۴)$$

$$X'_{\mathrm{mdb}}[k] = X_{\mathrm{mdb}}[k] - \mu_X \qquad (۵)$$

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.55]

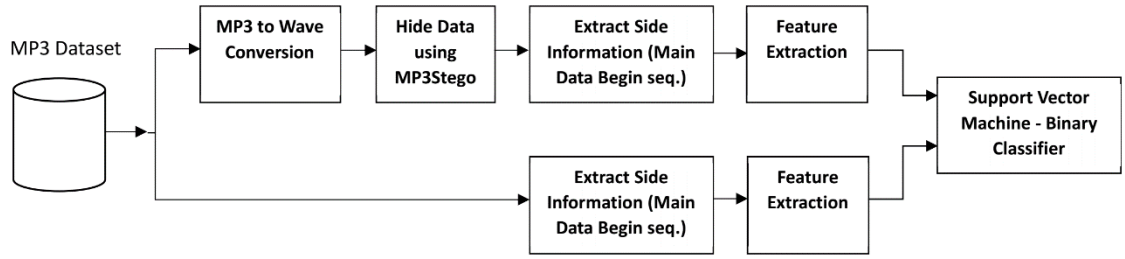
$$\mathrm{SPK}_{\mathrm{mdb}} = \frac{\sqrt{\frac{2}{M}\sum_{m=1}^{M/2} X'_{\mathrm{mdb}}[k]^{\,2}}}{\frac{2}{M}\sum_{m=1}^{M/2} \left|X'_{\mathrm{mdb}}[k]\right|} \qquad (۶)$$

$$\mathrm{FeatVec} = \left[\mu_{\mathrm{mdb}}\ \mu_X\ \mathrm{SPK}_{\mathrm{mdb}}\right]^{\mathrm{T}} \qquad (۷)$$

۳-۲-آموزش سامانه

ازآنجاکه هدف اصلی این پژوهش شناسایی فایل‌های نهان‌نگاری‌شده با استفاده از یادگیری ماشین است تهیهٔ بانک‌داده برای آموزش و آزمون ضروری است. به‌این‌منظور در ابتدا کلیهٔ فایل‌های انتخاب‌شدهٔ MP به فرمت wav تبدیل‌شده و در ادامه فایل‌های wav توسط نرم‌افزار MP3Stego با درصدهای متفاوتی از ظرفیت، نهان‌نگاری شده و پس از استخراج پارامترهای اطلاعات جانبی از هر دودسته فایل‌های MP3 آلوده و تمیز برای دو شمای کلی فاز آموزش سامانه طراحی‌شده در شکل (۵) نشان‌داده شده‌است.

(شکل-۵): بلوک دیاگرام مرحلهٔ آموزش سامانهٔ پیشنهادی
(Figure 5): Block diagram of the training stage for the proposed system



۳-۳- یادگیری سامانه

برای طبقه‌بندی داده‌ها از ماشین بردار پشتیبان (SVM) که یک الگوریتم با نظارت در یادگیری ماشین است، استفاده شده‌است. ماشین بردار پشتیبان از روشی که به آن روش مبتنی بر هسته (کرنل[1]) گفته می‌شود، استفاده می‌کند. در این روش در واقع توابعی وجود دارند که فضای ورودی بُعد پایین را دریافت کرده و آن را با تبدیلات غیرخطی به فضای بُعد بالاتر تبدیل می‌کنند. این تبدیل، یک مسئلهٔ غیرقابل جداسازی به‌صورت خطی در فضای بُعد پایین را به‌یک مسئله قابل جداسازی به‌صورت خطی در فضای بُعد بالا تبدیل می‌کند.

۳-۴- آزمون سامانه

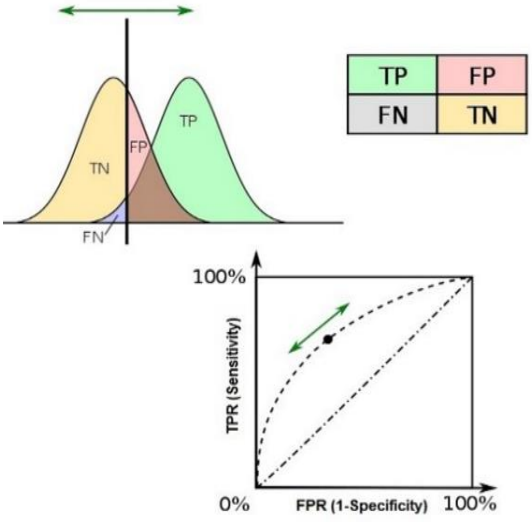
در مرحلهٔ آزمون یک فایل صوتی MP3 به سامانه آموزش‌دیده وارد می‌شود و از سامانه می‌خواهد که باتوجه‌به ویژگی‌های به‌دست‌آمده تشخیص دهد که این فایل، یک فایل تمیز است یا یک فایل نهان‌نگاری شده‌است. فایل‌های برچسب‌دار جدید به سامانه آموزش‌دیده داده شده و نتایج خروجی سامانه با برچسب مرجع فایل ورودی مقایسه می‌شود.

برای تبیین ارزیابی میزان موفقیت سامانه آموزش‌دیده از منحنی ROC و معیار AUC استفاده شده‌است.

[1] kernel

۳-۴-۱-منحنی راک[2]

منحنی راک که از آن به‌عنوان «منحنی مشخصهٔ عملکرد سامانه» نام برده می‌شود، یک نمودار گرافیکی است که برای تشخیص و ارزیابی کارایی یک سامانه طبقه‌بند دودویی استفاده می‌شود.



(شکل-۶): نمایش منحنی راک و حالات مختلف خروجی سامانه
(Figure-6): The ROC curve and types of system errors

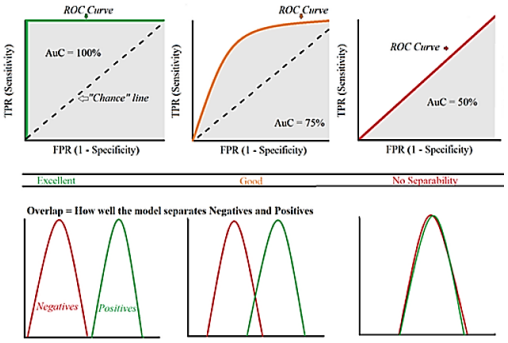
منحنی راک که به‌صورت شکل زیر است، برای بیان نرخ پذیرش صحیح (TPR) در برابر نرخ پذیرش غلط (FPR) به‌کار می‌رود. محور عمودی این منحنی حساسیت و محور افقی آن همان (ویژگی - ۱) است.

[2] Receiver Operating Characteristics

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.55]

۳-۴-۲- سطح زیر منحنی راک[1] (AUC)

مفهومی دیگر به‌نام سطح زیر منحنی راک که به‌طورمعمول به‌اختصار AUC نامیده می‌شود، می‌تواند رفتار سامانه را ارزیابی کند. شکل (۷) ارتباط بین سطح زیر منحنی راک و میزان درستی نتایج سامانه را در حالات متفاوت نشان می‌دهد.



(شکل-۷): نمایش ارتباط قدرت تشخیص سامانه و منحنی راک

(Figure-7): The relation between ROC curve and recognition accuracy

برای تفسیر مقادیر AUC می‌توان از قوانین کلی مطابق با جدول زیر استفاده کرد:

(جدول-۱): دسته‌بندی قدرت تشخیص سامانه

(Table-1): Recognition ability of the system according to AUC value

مقدار عددی AUC در منحنی ROC	قدرت تشخیص آزمون (درستی نتایج آزمون)
۰/۹ تا ۱	بسیار عالی
۰/۸ تا ۰/۹	عالی
۰/۷ تا ۰/۸	قابل‌قبول
۰/۵	بدون قدرت تشخیص

۴-آزمایش‌های انجام‌شده

در ادامه به‌تشریح مراحل اجرای روش پیشنهادی می‌پردازیم:

۴-۱-بانک داده‌های فرمت MP3

برای تأمین داده‌های آموزشی و آزمایشی از بانک دادهٔ دورتموند[2] [۱۳] استفاده شده‌است. این بانک شامل تعداد زیادی فایل‌های استاندارد با فرمت MP3 است که در ژانرهای مختلف دسته‌بندی شده‌اند. تمام فایل‌های صوتی استفاده‌شده در مجموعه‌دادگان، فایل‌های استریو دوثانیه‌ای با نرخ فشرده‌سازی 128kbps و نرخ نمونه‌برداری 44100Hz هستند. از میان داده‌های این بانک تعداد دویست‌نمونه از سبک‌های مختلف موسیقی انتخاب شد.

از این دویست نمونهٔ انتخاب‌شده صد نمونه برای آموزش و صد نمونه جهت‌بخش آزمون مشخص‌شده و از هر صد نمونه پنجاه نمونه توسط نهان‌نگار MP3Stego با ظرفیت‌های متنوع ۱٪ و ۵۰٪ و ۱۰۰٪ (بیشینهٔ ظرفیت پنهان‌سازی نرم‌افزار MP3Stego) نهان‌نگاری شده و برای مراحل آموزش و آزمایش استفاده شد؛ به‌منظور تأمین فایل‌های متنی مناسب جهت نهان‌نگاری از یک نرم‌افزار تولیدکنندهٔ متون تصادفی با ظرفیت‌های موردنیاز یک‌درصد، پنجاه‌درصد و صددرصد استفاده شده‌است.

۴-۲- تبدیل فرمت فایل‌های MP3 به Wav

ازآنجاکه نهان‌نگاری با نرم‌افزار MP3Stego درحین فشرده‌سازی انجام می‌شود ابتدا فایل‌های MP3 استخراج‌شده از بانک داده توسط نرم‌افزار مبدل [14] OGG به فرمت wav تبدیل می‌شوند؛ این فرایند در محیط ویندوز صورت می‌پذیرد.

۴-۳- تولید فایل نهان‌نگاری‌شده با فایل‌های متنی

مجدداً فایل‌های غیرفشرده wav توسط نرم‌افزار MP3Stego نهان‌نگاری می شوند، MP3Stego در خط فرمان ویندوز و جداگانه برای هر فایل انجام می‌شود؛ به‌منظور اطمینان از انجام صحیح نهان‌نگاری، فایل‌های نهانهٔ کدگشایی‌شده و سپس فایل‌های مشکل‌دار کنار گذاشته شده‌اند.

۴-۴- استخراج اطلاعات جانبی

مقادیر متغیر اشاره‌گر main_data_begin (mdb) برای فایل MP3 تمیز (پوشانه) و برای فایل‌های آلوده (نهانه) توسط برنامهٔ تهیه‌شده در محیط متلب[3] استخراج و ذخیره می‌شود؛ در واقع برای هر نمونه فایل MP3 پوشانهٔ فایل متناظر نهان‌نگاری شده (با ظرفیت‌های مختلف) ایجاد می‌شود.

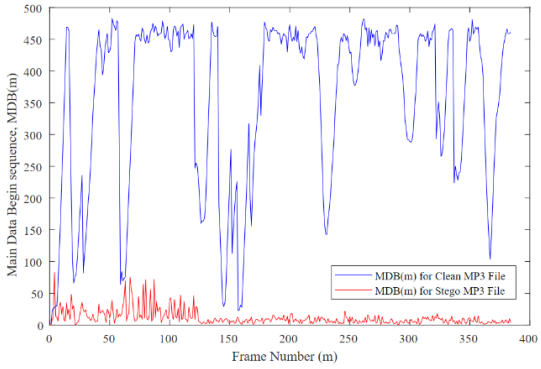
در شکل (۸) مقادیر متغیر mdb برای یک فایل صوتی پوشانه و یک فایل صوتی نهانه نشان‌داده شده‌است.

همان‌طورکه مشاهده می‌شود مقدار متوسط متغیر mdb برای فایل نهانه کاهش پیدا کرده‌است.

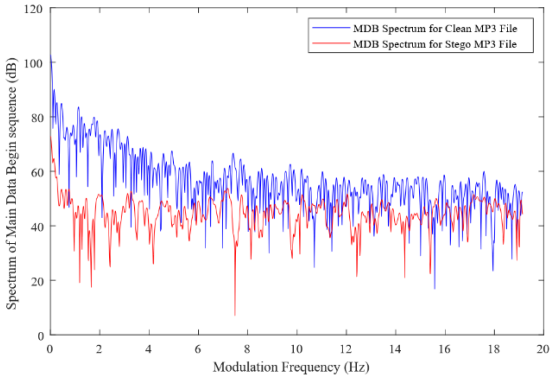
در این مرحله، تعداد صدنمونه از داده‌ها برای مرحلهٔ آموزش و تعداد صدنمونهٔ دیگر برای مرحلهٔ آزمون سامانه جداسازی می‌شود.

[1] Area Under the ROC Curve
[2] Dortmund
[3] MATLAB

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]
[DOI: 10.61186/jsdp.21.2.55]



(شکل- ۸-الف)



(شکل-۸-ب)

(شکل-۸): مقایسهٔ تأثیر نهان‌نگاری در شکل متغیر mdb : الف- مقایسهٔ دنبالهٔ mdb در حوزهٔ زمان برای فایل تمیز (آبی رنگ) و فایل آلوده (قرمز رنگ). ب- مقایسهٔ لگاریتم طیف فوریه دنبالهٔ mdb در حوزهٔ فرکانس برای فایل تمیز (آبی رنگ) و فایل آلوده (قرمز رنگ)

(Figure-8): Effects of steganography on the shape of MDB side information – A: MDB in time domain, B: Fourier spectrum of MDB in frequency domain

۴-۵- آموزش سامانه با تکنیک‌های یادگیری ماشین

برای دسته‌بندی فایل‌های صوتی به دو دسته نهان‌نگاری‌شده و نهان‌نگاری‌نشده، از یک دسته‌بند دودویی استفاده شده‌است. دسته‌بند موردنظر ماشین بردار پشتیبان (SVM) است.

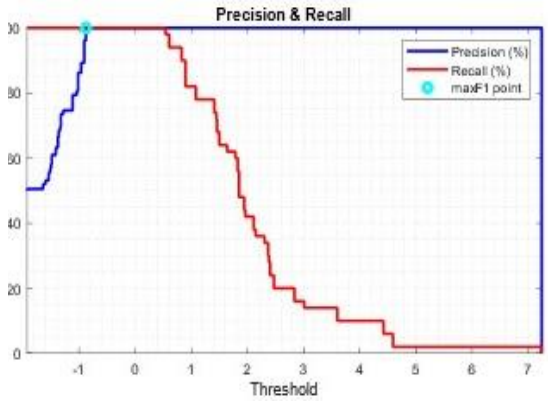
۴-۶-ارزیابی آزمون سامانه

برای آزمون سامانه طراحی‌شده فایل‌های جداشدهٔ آزمون به‌ترتیب به سامانه ارائه و صحت تشخیص برای صدنمونه برای هر ظرفیت نهان‌نگاری کنترل می‌شود. در جدول (۲) نتایج مرحلهٔ آزمون سامانه برای سه ظرفیت ۱٪، ۵۰٪ و ۱۰۰٪ ارائه شده‌است. ازآنجاکه نتایج حاصله برای هر سه‌دسته ظرفیت مشابه است، نتایج یک‌جا ارائه شده‌است.

در آزمایش‌های انجام‌شده در مرجع [۴] مقدار میانگین دقت تست برای متغیر main_data_begin عدد ۹۷.۶٪ گزارش شده‌است. در مرجع دیگر [۸]، حداقل خطای EER ۱.۱۱٪ ، حداکثر مقدار AUC برابر با ۹۹.۹۲٪ و حداکثر مقدار F-measure برابر با ۹۸.۹۹٪ بوده‌است؛ درواقع در الگوریتم پیاده‌شده تأثیر نهان‌نگاری بر روی متغیر mdb زیاد بوده و باعث ایجاد تمایز و دسته‌بندی مناسب بین سیگنال‌های آلوده و تمیز می‌شود.

باتوجه به‌حجم بسیار بالای استفاده از فایل‌های فشردهٔ صوتی با فرمت MP3 و دردسترس‌بودن نهان‌نگار MP3stego، استفاده از ویژگی mdb در فرایند نهان‌کاوی می‌تواند ابزار خوبی برای شناسایی فایل‌های نهان‌نگاری شده با این روش باشد.

جهت بررسی و ارزیابی بیشتر روش می‌توان آزمایش‌هایی را با نمونه‌های طولانی‌تر از فایل MP3 (بیشتر از ده ثانیه) انجام داد.



(شکل-۹): نمایی از منحنی دقت (Precision) در کنار منحنی یادآوری (Recall) برای سامانه پیشنهادی در حالت ظرفیت پنهان‌نگاری ۱٪ (سخت‌ترین حالت). همان‌طورکه دیده می‌شود، سامانهٔ پیشنهادی توانسته‌است به‌مقدار کارایی F1=100% دست یابد (دقت ۱۰۰٪).

(Figure-9): Precision versus Recall curves for the proposed system at 1% embedding capacity.

همان‌طور که در شکل (۹) منحنی راک مربوط به سامانه پیشنهادی نشان‌داده‌شده و باتوجه به‌جدول (۲) قدرت تشخیص سامانه پیشنهادی عالی ارزیابی می‌شود.

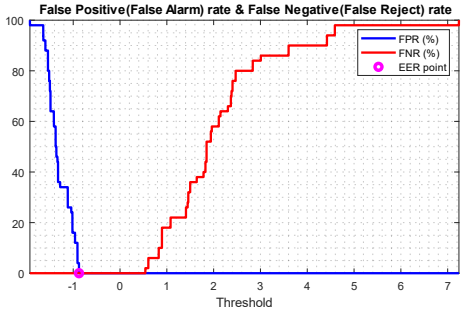
(جدول-۲): نتیجهٔ آزمایش‌ها برای ظرفیت‌های متفاوت نهان‌نگاری

(Table-2): Experimental results for different embedding rates (capacities)

ظرفیت نهان‌نگاری (٪)	AUC	F-measure	EER
Mp3stego ٪۱۰۰ - ٪۵۰ - ٪۱	۱	٪۱۰۰	٪۰

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.55]



(شکل-۱۰): نمایی از منحنی نرخ خطای پذیرش اشتباه (FPR) در کنار منحنی نرخ خطای رد اشتباه (FNR) برای سامانه پیشنهادی در حالت ظرفیت پنهان‌نگاری ۱٪ (سخت‌ترین حالت). همان‌طور که دیده می‌شود، سامانهٔ پیشنهادی توانسته‌است به‌مقدار خطای EER=0% دست یابد (بدون خطا).

(Figure-10): FPR versus FNR curves for the proposed system at 1% embedding capacity.

۵- نتیجه‌گیری

در سال‌های اخیر نهان‌نگاری اطلاعات در فایل‌های چندرسانه‌ای توجه زیادی را در میان پژوهشگران به‌خود جلب کرده‌است. از دید امنیت فضای تولید و تبادل اطلاعات، بالتبع انجام نهان‌کاوی و آشکارسازی و کشف محتویات چندرسانه‌ای که اطلاعاتی در آن پنهان شده‌است، اهمیت بسیار فراوانی دارد؛ دراین‌مقاله یک روش نهان‌کاوی جدید برای فایل‌های صوتی با قالب MP3 که در آن با استفاده از الگوریتم مشهور MP3Stego اطلاعاتی پنهان شده، ارائه شده‌است. روش پیشنهادی به‌جای اینکه از اطلاعات سیگنال صوت ذخیره‌شده در فایل MP3 استفاده کند، از اطلاعات جانبی فایل بهره می‌گیرد که وابستگی کمتری به محتویات صوتی فایل دارد؛ در این روش، اطلاعات جانبی MDB در فایل صوتی فشرده به‌عنوان یک دنباله فرض می‌شود و سپس با استفاده از یک روش استخراج ویژگی، یک ویژگی جدید در حوزهٔ فرکانس به نام قله‌داربودن طیف محاسبه می‌شود. این ویژگی ساده و درعین‌حال قوی با ویژگی‌هایی همچون میانگین زمانی و میانگین طیف دامنهٔ دنباله MDB تلفیق شده و یک بردار ویژگی با ابعاد کم (سه‌بُعدی) را تشکیل می‌دهد. این بردار ویژگی سپس به‌وسیلهٔ یک دسته‌بند از نوع ماشین بردار پشتیبان (SVM) به‌عنوان فایل نهان‌نگاری‌شده (مشکوک) یا فایل عادی (فاقد اطلاعات) شناخته خواهد شد. روش استخراج ویژگی درعین‌ساده‌بودن و داشتن محاسبات بسیار کم، برای نهان‌کاوی فایل‌های MP3 دقت ۱۰۰٪ دارد (بازشناسی بدون خطا)، حتی زمانی‌که اطلاعات پنهان‌شده در فایل صوتی بسیار کم و تُنُک باشد.

۶- مراجع 6-Refrences

[۱] شیخ‌زادگان، جواد، آشنایی با نهان‌نگاری و پنهان‌سازی اطلاعات، تهران، انتشارات پژوهشکدهٔ پردازش هوشمند علائم، ۱۳۹۲.

[2] H. Ghasemzadeh, and M. H. Kayvanrad, “Comprehensive review of audio steganalysis methods”, IET Signal Processing, vol. 12, pp. 673-687, 2018.

[۳] دالوند، الهام، «نهان کاوی فایل‌های صوتی MP3 با روش‌های مبتنی بر اطلاعات جانبی»، پایان‌نامهٔ کارشناسی ارشد، دانشگاه رازی، کرمانشاه، ۱۳۹۴.

[4] R. Kuriakose, and P. Premalatha, "A novel Method for MP3 Steganalysis", Proceedings of Intelligent Computing, Communication and Devices, ICCD, New Delhi, India, 2015, pp. 605-611.

[5] H. Song, T. Hu, Y. Huang, and M. Guo, "Detecting MP3Stego and Estimating the Hidden Size", in WSEAS International Conference on Computers, Athens, Greece, 2003, pp. 2881-2884.

[6] X. Yu, R. Wang, D. Yan, and J. Zhu, "MP3 Audio Steganalysis Using Calibrated Side Information Feature", Journal of Computational Information Systems, vol. 8, pp. 4241-4248, 2012.

[7] C. Jin, R. Wang, D. Yan Ma, P. and K. Yang, "A Novel Detection Scheme for MP3stego with Low Payload", in IEEE China Summit & International Conference on Signal and Information Processing, ChinaSIP, Xi'an, China, 2014, pp. 602-606.

[۸] دارابی، علیرضا، «نهان یابی فایل‌های صوتی MP3 با تاکید بر روش‌های پردازش سیگنال»، پایان نامهٔ کارشناسی ارشد، دانشگاه رازی، کرمانشاه، ۱۳۹۵.

[۹] گروه واژه‌گزینی انجمن رمز ایران، واژه‌نامه و فرهنگ امنیت فضای تولید و تبادل اطلاعات، ویرایش دوم، تهران، مؤسسه انتشارات علمی دانشگاه صنعتی شریف، ۱۳۹۴.

[۱۰] مهدوی جعفری، سمیه، "نهان‌نگاری اصوات دیجیتال"، پایان نامهٔ کارشناسی ارشد، دانشگاه شهید باهنر، کرمان، ۱۳۸۸.

[۱۱] سلیمیان ریزی، عاطفه، "روش‌های مخفی‌سازی اطلاعات در فایل‌های صوتی MP3". پایان نامهٔ کارشناسی ارشد، دانشگاه صنعتی اصفهان، اصفهان، ۱۳۹۳.

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.55]

[12] R. Raissi, "The Theory behind MP3", Personal Report, Dec. 2002.
[13] Music Audio Benchmark Data Set, Dortmund University, [Online], Available: https://www-ai.cs.tu-dortmund.de/audio.html [Accessed: Sept. 2, 2024]
[14] Ogg MP3 converter, [Online], Available: http://www.ogg-mp3.net [Accessed: Sept. 2, 2024]

[۱۵] اشعری، فاطمه، ریاحی، نوشین، «نهان‌کاوی صوت مبتنی بر همبستگی بین فریم و کاهش بازگشتی ویژگی»، پردازش علائم و داده‌ها، دورهٔ ۱۵، شمارهٔ ۳، صفحات ۱۱۳-۱۲۲، ۱۳۹۷.



محسن سلیمانی مدرک دکترای خود را در سال ۱۴۰۰ در رشتهٔ مهندسی برق-مخابرات از دانشگاه ارومیه دریافت کرد. زمینه‌های پژوهشی موردعلاقهٔ ایشان عبارت‌اند از: امنیت اطلاعات، نهان‌نگاری، نهان‌کاوی.
نشانی رایانامهٔ ایشان عبارت است از:
mohsen1soleimani@gmail.com



مهدی چهل امیرانی مدرک دکتری خود را در سال ۱۳۸۸ در رشتهٔ مهندسی برق-مخابرات از دانشگاه علم‌وصنعت ایران دریافت کرد. ایشان هم‌اکنون استاد دانشکدهٔ مهندسی برق و کامپیوتر در دانشگاه ارومیه هستند. زمینه‌های پژوهشی موردعلاقهٔ ایشان عبارت‌اند از: پردازش سیگنال، پردازش تصویر و پردازش سیگنال‌های زیست‌پزشکی.
نشانی رایانامهٔ ایشان عبارت است از:
m.amirani@urmia.ac.ir



سیدجهانشاه کبودیان مدرک دکترای خود را در سال ۱۳۸۹ در رشتهٔ مهندسی کامپیوتر-هوش مصنوعی و رباتیک از دانشگاه صنعتی امیرکبیر دریافت کرد. ایشان هم‌اکنون استادیار دانشکدهٔ مهندسی برق و کامپیوتر در دانشگاه رازی کرمانشاه هستند. زمینه‌های پژوهشی موردعلاقهٔ ایشان عبارت‌اند از: پردازش سیگنال دیجیتال، پردازش سیگنال‌های آکوستیکی، صوت، موسیقی و گفتار، داده‌کاوی، یادگیری ماشین، یادگیری عمیق و الگوریتم‌های فراابتکاری.
نشانی رایانامهٔ ایشان عبارت است از:
kabudian@razi.ac.ir

[Downloaded from jsdp.rcisp.ac.ir on 2026-11-04]

[DOI: 10.61186/jsdp.21.2.55]