

## دسته‌بندی کدگذارهای صوت در بستر شبکه‌های مخابراتی

فرزانه جعفری<sup>۱</sup>، کارشناسی ارشد؛ مهدی تیموری<sup>۲</sup>، استادیار

۱- دانشکده علوم و فنون نوین- دانشگاه تهران- تهران- ایران- farzaneh.jafari@ut.ac.ir

۲- دانشکده علوم و فنون نوین- دانشگاه تهران- تهران- ایران- mehditeimouri@ut.ac.ir

**چکیده:** داده‌های صوتی به‌طور گسترده‌ای در شبکه‌های مخابراتی در حال تبادل هستند. به‌دلیل محدودیت منابع شبکه، این داده‌ها معمولاً به‌صورت فشرده‌شده ارسال می‌شوند. روش‌های مختلفی برای فشرده‌سازی داده‌های صوت وجود دارد. لذا برای دسترسی غیرمجاز به اطلاعات صوتی، ابتدا باید نوع کدگذار به‌کاررفته در فشرده‌سازی صوت را شناسایی نمود. یکی از بهترین روش‌ها برای شناسایی نوع کدگذار داده‌های صوت، شناسایی بر اساس محتوای بسته‌های کدشده دریافتی است. در چنین روش‌هایی، نوع کدگذار به‌کاررفته برای تولید بسته‌های صوتی بر اساس ویژگی‌های آماری شناسایی می‌شود. در این مقاله، روشی برای دسته‌بندی کدگذارهای صوتی بر اساس دو ویژگی طولانی‌ترین زیررشته مشترک و طولانی‌ترین زیرتوالی مشترک پیشنهاد شده است. نتایج شبیه‌سازی نشان می‌دهد که عملکرد روش پیشنهادی (با صحت حدود ۹۷٪ برای بسته‌های ۸ کیلوبایتی) به‌شکل فراوانی بهتر از روش‌های موجود است.

**واژه‌های کلیدی:** کدگذارهای صوت، دسته‌بندی، طولانی‌ترین زیررشته مشترک، طولانی‌ترین زیرتوالی مشترک.

## Classification of Audio Codecs in Telecommunication Networks

Farzaneh Jafari<sup>1</sup>, M.Sc. Graduated Student; Mehdi Teimouri<sup>2</sup>, Assistant Professor

1- Faculty of New Science and Technology, University of Tehran, Tehran, Iran, Email: farzaneh.jafari@ut.ac.ir

2- Faculty of New Science and Technology, University of Tehran, Tehran, Iran, Email: mehditeimouri@ut.ac.ir

**Abstract:** Audio data are widely exchanged in telecommunications networks. Due to the limitation of network resources, these data are usually compressed before transmission. Various methods exist for compressing audio data. Hence, in order to unlawfully access these audio information, one needs to first identify the codec which is used for audio compression. One of the best approaches for audio codec identification is identification based on the contents of received packets. In these methods, statistical features of received packets are used for identification of employed codec. In this paper, a method of audio codec classification is proposed based on longest common substring and longest common subsequence features. Simulation results (with accuracy of 97% for 8 Kbytes packets) demonstrate the superiority of the proposed method compared to conventional methods.

**Keywords:** Audio codecs, classification, longest common substring, longest common subsequence.

تاریخ ارسال مقاله: ۱۳۹۶/۰۲/۲۰

تاریخ اصلاح مقاله: ۱۳۹۶/۰۷/۲۷

تاریخ پذیرش مقاله: ۱۳۹۶/۰۹/۱۴

نام نویسنده مسئول: مهدی تیموری

نشانی نویسنده مسئول: ایران - تهران - خیابان کارگر شمالی - دانشکده علوم و فنون نوین دانشگاه تهران.

## ۱- مقدمه

شناسایی کور شبکه‌های مخابراتی یکی از حوزه‌هایی است که اخیراً مورد توجه زیادی قرار گرفته است [۱، ۲]. شناسایی نوع داده‌های چندرسانه‌ای<sup>۱</sup> ارسالی در شبکه‌های مخابراتی یک زیرموضوع از علوم جرم‌شناسی است که عمدتاً با عنوان شناسایی قطعات فایل<sup>۲</sup> شناخته می‌شود. شناسایی نوع داده‌ها بیشترین اهمیت را در عملکرد مناسب سیستم‌های تشخیص نفوذ، سیستم‌های ضد ویروس و نرم‌افزارهای فیلتر شبکه دارد. از کاربردهای مهم شناسایی داده‌ها، می‌توان به نظارت شبکه، کشف جرم‌های کامپیوتری و دزدی دیجیتال اشاره کرد. زمانی که می‌خواهیم شبکه‌ای را مدیریت کرده و اطلاعات ردوبدل شده در آن را دیده‌بانی نماییم، هنگامی که می‌خواهیم اطلاعاتی را از رسانه‌های خراب، تکه‌تکه شده یا سانحه دیده استخراج کنیم و یا زمانی که می‌خواهیم به صورت غیرمجاز به یک شبکه دسترسی داشته باشیم و اطلاعات کاربران را شنود کنیم، باید ابتدا نوع داده را شناسایی نماییم [۳].

یکی از انواع مهم داده‌های ارسالی در شبکه‌های مخابراتی، داده‌های صوتی است. معمولاً از کدگذارهای صوت برای فشرده‌سازی داده‌های صوتی و ارسال آن‌ها در شبکه‌ها استفاده می‌شود. لذا برای دسترسی به محتوای اطلاعات صوتی، ابتدا لازم است که نوع کدگذار مورد استفاده برای فشرده‌سازی صوت، تعیین شود. روش‌های سنتی و عمدتاً ناکارآمد شناسایی نوع داده، شامل روش شناسایی بر اساس پسوند فایل<sup>۳</sup>، روش شناسایی بر اساس بایت جادویی<sup>۴</sup> و روش هیستوگرام<sup>۵</sup> هستند. روش‌های مدرن تشخیص نوع داده، روش‌هایی هستند که بر اساس محتوای داده عمل می‌کنند. در این روش‌ها، ما با استفاده از محتوای داده و بدون در نظر گرفتن سرآیند<sup>۶</sup> و دنباله<sup>۷</sup> داده‌ها، قادر به شناسایی داده‌های جعلی هستیم و می‌توانیم به ماهیت واقعی داده‌ها پی ببریم [۴]. این روش‌ها عمدتاً بر پایه استخراج ویژگی‌های مختلف از محتوای داده‌ها عمل می‌کنند.

شناسایی، تفکیک و دسته‌بندی داده‌های چندرسانه‌ای را می‌توان در سه سطح مختلف در نظر گرفت. در سطح اول شناسایی، دسته‌بندی انواع داده‌های چندرسانه‌ای در کلاس‌های کلی مانند گفتار، متن، تصویر و ... انجام می‌پذیرد. در این سطح از شناسایی، هدف اصلی تحلیل ترافیک کلی شبکه است. در یکی از کارهای انجام‌شده در این حوزه، دسته‌بندی داده‌های چندرسانه‌ای گفتار، متن و دورنگار مدنظر قرار گرفته است [۵]. در این تحقیق، برای استخراج ویژگی‌ها از N-گرم‌ها<sup>۸</sup> و برای انتخاب ویژگی از معیار تحلیل مؤلفه‌های اصلی (PCA<sup>۹</sup>) استفاده شده است. در نهایت، از شبکه عصبی مصنوعی (ANN<sup>۱۰</sup>) برای یادگیری و دسته‌بندی داده‌ها، استفاده شده است.

در سطح دوم شناسایی که موضوع اصلی مقاله پیش رو است، سامانه تحلیل گر انتظار دریافت کلاسی خاص از داده‌های چندرسانه‌ای را دارد که

در این مقاله به‌طور خاص داده صوتی است. این داده‌ها می‌توانند با استانداردهای مختلف، کدگذاری شده باشند. در چنین شرایطی، هدف سامانه شناسایی نوع کدگذار به‌کاررفته در هر بسته داده است. دسته‌بندی در این سطح، به معنی تعیین روش کدگذاری به‌کاررفته (مانند PCM<sup>۱۱</sup> و LPC<sup>۱۲</sup>) در هر بسته است. در یکی از تحقیقات انجام گرفته شده در این زمینه، دسته‌بندی بین کدگذارهای گفتار u-law، a-law و ADPCM<sup>۱۳</sup> های ۱۶، ۲۴، ۳۲ و ۶۴ کیلوبیت بر ثانیه مدنظر قرار گرفته است [۶]. الگوریتم دسته‌بندی مورد استفاده در این تحقیق قادر به دسته‌بندی بین این شش کلاس است. در این الگوریتم از سه ویژگی خودهمبستگی<sup>۱۴</sup>، ممان مرکزی<sup>۱۵</sup> مرتبه دوم و نسبت دودویی<sup>۱۶</sup> استفاده شده است. یکی دیگر از تحقیقات مرتبط، تفکیک بین دو فرمت MP3 و FLAC بر اساس کلمات هماهنگ‌ساز<sup>۱۷</sup> است [۷]. کلمه هماهنگ‌ساز یک ویژگی برای شناسایی داده‌های چندرسانه‌ای بوده و مبنای استفاده از آن‌ها ترکیبی از روش‌های ساختاری<sup>۱۸</sup> و آماری<sup>۱۹</sup> است. در این روش، در ابتدای هر قاب<sup>۲۰</sup> یک کلمه هماهنگ‌ساز قرار می‌گیرد. این کلمه با توجه به کدگذار تولیدکننده هر بسته متفاوت بوده و نشان‌دهنده کلاس مربوط به هر بسته است؛ اما در این روش مشکلاتی نیز وجود دارد و این امکان وجود دارد که این کلمات در داده‌هایی غیر از داده‌های مورد بررسی نیز وجود داشته باشند. برای حل این مشکل از روش‌هایی از قبیل اندازه‌گیری طول بین دو کلمه هماهنگ‌ساز استفاده شده است [۷]. یکی از مهم‌ترین تحقیقات این حوزه، توسط هیکسانمز<sup>۲۱</sup> و همکاران [۸] انجام شده است که در آن شناسایی انواع زیادی از کدگذارهای صوت بر اساس ویژگی‌های تصادفی<sup>۲۲</sup> و بی‌نظمی<sup>۲۳</sup> مدنظر قرار گرفته است. ویژگی‌های تصادفی در حوزه زمان، شامل دو دسته ویژگی‌های آماری ساده مانند میانگین<sup>۲۴</sup>، واریانس<sup>۲۵</sup>، خودهمبستگی، آنتروپی<sup>۲۶</sup> و دسته ویژگی‌های آماری مرتبه بالا مانند وابستگی دوگانه<sup>۲۷</sup>، چولگی<sup>۲۸</sup> و کشیدگی<sup>۲۹</sup> هستند. در حوزه فرکانس، طیف فرکانس به چهار زیر باند تقسیم شده و ویژگی‌های میانگین، واریانس و چولگی از هر زیر باند استخراج می‌شود. ویژگی‌های بی‌نظمی، شامل ویژگی‌های نزدیک‌ترین کاذب<sup>۳۰</sup> و نمای لپانوف<sup>۳۱</sup> است. پس از استخراج ویژگی‌ها، از الگوریتم جستجوی شناور ترتیبی رو به جلو (SFFS<sup>۳۲</sup>) برای انتخاب ویژگی و از ماشین بردار پشتیبانی (SVM<sup>۳۳</sup>) با هسته رادیال<sup>۳۴</sup> برای انجام عمل دسته‌بندی استفاده شده است [۸].

در سطح سوم شناسایی، سامانه تحلیل گر انتظار دریافت یک کلاس خاص از داده چندرسانه‌ای را دارد که توسط یک استاندارد کدگذاری مشخص کد شده باشد. در چنین حالتی، وظیفه سامانه تحلیل گر یافتن پارامترهای مجهول کدگذار است. در یکی از تحقیقات منتشرشده در این زمینه، دسته‌بندی کدگذار گفتار CVSD<sup>۳۵</sup> با نرخ‌های بیت متفاوت ۸، ۱۰، ۱۰، ۱۶، ۲۴ و ... کیلوبیت بر ثانیه مدنظر قرار گرفته است [۹]. در این

می‌شوند. علاوه بر این، کدگذارها از نظر اهداف طراحی نیز با یکدیگر متفاوت هستند. این اهداف شامل توانایی در برقراری تعادل بین نرخ فشرده‌سازی و کیفیت، اصلاح خطاهای شبکه، تطابق با پهنای باندهای مختلف و غیره است.

با وجود تنوع بسیار بالای کدگذارهای صوت و اهمیت این حوزه، تمامی روش‌های ذکر شده در [۵، ۶، ۷، ۹، ۱۱، ۱۲ و ۱۳] تعداد محدودی از کدگذارهای صوت را مورد بررسی قرار داده‌اند. در این بین، تاکنون تنها هیکسانمز و همکاران [۸] دسته‌بندی را در میان طیف گسترده‌تری از کدگذارهای صوت انجام داده‌اند. به عبارت دیگر، می‌توان روش هیکسانمز و همکاران [۸] را به‌عنوان مرجع اصلی جهت مقایسه در نظر گرفت؛ زیرا در [۸]، علاوه بر به‌کار بردن دامنه وسیعی از کدگذارهای صوت با تکنولوژی‌های مختلف، دقت دسته‌بندی کدگذارها دارای نتایج قابل‌توجه‌تری نسبت به کارهای انجام شده در این زمینه است. اکثر روش‌های استخراج ویژگی به کار برده شده در تحقیقات پیشین توسط هیکسانمز و همکاران [۸] استفاده شده است و روش استخراج ویژگی بی‌نظمی به‌عنوان روش پیشنهادی آن‌ها برای بهبود شناسایی کدگذارهای صوت معرفی شده است.

در این مقاله، روشی جدید برای استخراج ویژگی پیشنهاد می‌کنیم که نسبت به روش هیکسانمز و همکاران [۸] بهبود قابل‌توجهی در دسته‌بندی کدگذارهای صوت ایجاد می‌کند.

### ۳- مدل سیستم

برای درک بهتر مدل سیستم، یک شبکه ارسال صوتی را به‌صورت آنچه در شکل ۱ نمایش داده شده است در نظر بگیرید. در این شبکه، چندین کاربر وجود دارند و هر کاربر از کدگذار صوت مخصوص به خود استفاده می‌کند. هر کاربر ورودی خود را با استفاده از کدگذار مربوط به خود کد کرده و در قالب بسته‌هایی با طول مشخص در شبکه ارسال می‌کنند. جهت استفاده عادلانه از منابع شبکه، یک تسهیم‌گر<sup>۱</sup> ایفای نقش می‌کند. این بسته‌ها پس از دریافت (احتمالاً غیرمجاز) در اختیار بلوک دسته‌بند قرار داده می‌شوند. بلوک دسته‌بند صوت، ویژگی‌هایی را از هر بسته استخراج کرده و بر اساس این ویژگی‌ها نوع کدگذار مرتبط با آن بسته را مشخص می‌نماید.

این مدل، یک مدل ساده‌شده از سیستم واقعی شنود است که در آن ابتدا شکل موج فرکانس رادیویی توسط گیرنده دریافت، تحلیل و دمدوله می‌شود. همچنین اضافاتی مانند سرآیند بسته‌ها از داده جدا می‌شوند و در انتها داده‌های خام به دسته‌بند تحویل داده می‌شوند. به عبارت دیگر، فرض می‌کنیم، سربارهای لایه‌های قبل از روی داده‌ها برداشته شده و با داده‌های چندرسانه‌ای خالص روبرو هستیم. علاوه بر این، فرض می‌کنیم،

تحقیق از ممان‌های مرکزی مرتبه اول تا پنجم، خودهمبستگی و نسبت دودویی به‌عنوان ویژگی استفاده شده است. یکی دیگر از تحقیقات انجام‌شده در این حوزه، تخمین عرض تصویر کدشده توسط استاندارد JPEG<sup>۲</sup> است [۱۰].

در این مقاله، ما بر روی شناسایی کدگذارهای صوتی در سطح دوم شناسایی تمرکز می‌کنیم. در این راستا، روشی برای دسته‌بندی کدگذارهای صوتی بر اساس دو ویژگی طولانی‌ترین زیررشته مشترک و طولانی‌ترین زیرتوالی مشترک پیشنهاد شده است. نتایج شبیه‌سازی نشان می‌دهد که عملکرد روش پیشنهادی (با صحت حدود ۹۷٪ برای بسته‌های ۸ کیلوبایتی) به‌شکل قابل‌ملاحظه‌ای بهتر از روش‌های موجود است. ساختار این مقاله در ادامه به این شرح است. در بخش دوم، تحقیقات پیشین ارائه شده است. مدل سیستم و تعریف دقیق مسئله در بخش سوم ارائه می‌شود. در بخش چهارم، روش پیشنهادی برای دسته‌بندی کدگذارهای صوت ارائه می‌شود. در ادامه و در بخش پنجم، نتایج به‌دست‌آمده از شبیه‌سازی‌های مختلف مورد مقایسه و تحلیل قرار می‌گیرد. در انتها و در بخش ششم، نتیجه‌گیری ارائه می‌گردد.

### ۲- تحقیقات پیشین

در حوزه شناسایی کدگذارهای صوت تحقیقات نسبتاً اندکی انجام گرفته است. یکی از تحقیقات انجام شده در این حوزه، تلاش برای کشف و شناسایی کدگذارهای صوت در کانال‌های تلفنی است. در این کار، میانگین دقت دسته‌بندی برای کدگذارهای LPC، a-law و ADPCM حدود ۹۰٪ است [۱۱]. در یکی دیگر از تحقیقات انجام‌شده در این زمینه [۱۲]، از تجزیه طیف نوپز هارمونیک<sup>۳</sup> برای استخراج ویژگی استفاده شده است. نتایج حاصل از شناسایی شش کدگذار صوت ADPCM، AMR، GSM 6.60، GSM 6.20، GSM 7.23.1 و G.711 در این مقاله، نشان می‌دهد که اگرچه برخی از کدگذارها با خطای بسیار کمی شناسایی می‌شوند، اما برخی دیگر نیز با خطایی حدود ۲۰٪ شناسایی می‌شوند. در تحقیق مشابه دیگری [۱۳]، یک نمایه چندبعدی<sup>۴</sup> برای هر کدگذار ایجاد شده است که شامل ویژگی‌های به‌دست‌آمده از طیف نوپز<sup>۵</sup> و هیستوگرام دامنه زمان<sup>۶</sup> از سیگنال‌های گفتار کدشده است. نتایج حاصل از شناسایی هفت کدگذار صوت G.711، G.726، G.728، G.729، iLBC، AMR و Silk نشان می‌دهد که به‌غیر از کدگذارهای iLBC و Silk که با خطای کمی شناسایی می‌شوند، باقی کدگذارها با نرخ خطایی بین ۱۵٪ تا ۲۱٪ شناسایی می‌شوند.

تحقیقات انجام شده در این زمینه، هر کدام سهم مهمی در پیشبرد حوزه شناسایی کور داشته‌اند؛ اما باید به این نکته توجه کرد که بیش از یک هزار کدگذار صوت مختلف برای فشرده‌سازی داده‌های صوت وجود دارد. این کدگذارها برای انواع مختلفی از کاربرد مانند ارسال صوت بر روی شبکه‌های اینترنتی و یا انتقال صوت در شبکه‌های تلفن استفاده

پارامترها و ساخت مدل اولیه استفاده نشده و از آن صرفاً برای آزمایش دسته‌بند آموزش دیده استفاده می‌شود.

برای آموزش و آزمون (آزمایش) دسته‌بند باید ویژگی‌هایی از هر بسته استخراج شود. در یکی از کامل‌ترین تحقیقات این حوزه، هیکسانمز و همکاران [۸] ۶۵ ویژگی مختلف را مورد استفاده قرار داده‌اند. ویژگی‌های مورد استفاده در دو دسته ویژگی‌های تصادفی و ویژگی‌های بی‌نظمی دسته‌بندی می‌شوند. ویژگی‌های تصادفی به دو دسته، ویژگی‌های تصادفی در حوزه زمان و ویژگی‌های تصادفی در حوزه فرکانس تقسیم می‌شوند. ویژگی‌های تصادفی در حوزه زمان، شامل دو دسته ویژگی‌های آماری ساده مانند میانگین، واریانس، خودهمبستگی [۱۵]، آنتروپی [۱۵] و دسته ویژگی‌های آماری مرتبه بالا مانند وابستگی دوگانه [۱۶]، چولگی [۱۷] و کشیدگی [۱۷] هستند. وابستگی دوگانه با استفاده از جعبه‌ابزار HOSA<sup>۵۰</sup> محاسبه می‌شود [۱۸]. در ویژگی‌های تصادفی، ۲۱ ضریب اول از خودهمبستگی به‌عنوان ۲۱ ویژگی، در نظر گرفته می‌شوند. در کل، ویژگی‌های تصادفی در حوزه زمان شامل ۲۷ ویژگی هستند. در حوزه فرکانس، هر طیف فرکانس به چهار زیر باند تقسیم شده و ویژگی‌های میانگین، واریانس و چولگی از هر زیر باند استخراج می‌گردد که در کل برابر با ۱۲ ویژگی هستند. ویژگی‌های بی‌نظمی، شامل ویژگی‌های نزدیک‌ترین کاذب و نمای لیپانوف است [۸]. در نزدیک‌ترین کاذب سه ویژگی نسبت نزدیک‌ترین همسایگی کاذب یا FNF<sup>۵۱</sup>، میانگین از نزدیک‌ترین کاذب و اندازه ریشه میانگین مربعات از همسایگی، در بعدهای ۳ تا ۷ محاسبه شده که باعث تولید ۱۵ ویژگی می‌گردد. در نمای لیپانوف نیز، ۱۱ ویژگی در بعدهای ۱ تا ۱۱ استخراج می‌گردد. در مجموع تعداد ویژگی‌های بی‌نظمی برابر با ۲۶ است. برای به‌دست آوردن ویژگی‌های بی‌نظمی از نرم‌افزار TISEAN استفاده شده است [۱۹].

جدول ۱: کدگذارهای صوت مورد استفاده

تکنولوژی کدگذارها	کدگذارها	نرخ بیت پیش‌فرض (Kbps)	
با اتلاف	a-law [20]	۶۴	
	u-law [20]	۶۴	
	ADPCM [20]	۳۲	
	PCM [20]	G726 [21]	۳۲
		CVSD [22]	۱
		GSM 6.10 [20]	۱۳
		GSM (native)	۱۸
LPC [20]	AMR [23]	۱۲,۲	
	AWB [24]	۱۲,۶۵	
	G729 [25]	۸	

خطاهای احتمالی در داده‌های دریافتی توسط کد کانال تصحیح شده است.

کدگذارهای صوت مورد استفاده نیز از بین ۱۷ کدگذار نشان داده‌شده در جدول ۱ انتخاب می‌شوند. این کدگذارها از کدگذارهای محبوب و پر استفاده در ارتباطات مخابراتی و کامپیوتری هستند که به دو دسته با اتلاف و بدون اتلاف طبقه‌بندی می‌شوند. گروه کدگذارهای با اتلاف نیز به سه دسته PCM، LPC و MDCT قابل افراز هستند. مدولاسیون پالس کد یا PCM، فرآیندی برای پله‌ای کردن نمونه‌های سیگنال آنالوگ در واحد زمان است. روش کدگذار پیش‌بینی‌کننده خطی یا LPC، یک مدل بسیار ساده برای تولید گفتار است و سعی بر تقلید مکانیزم گفتار تولیدشده، توسط انسان را دارد. روش MDCT بهبود یافته روش DCT است. علاوه بر ۱۷ کدگذار جدول ۱، فرض می‌کنیم کاربران ممکن است داده‌های خام با فرمت wav نیز برای ارسال خود انتخاب کنند و به عبارتی از کدگذار صوت استفاده نمایند؛ بنابراین، در مجموع دسته‌بند باید دسته‌بندی را در یک فضای مسئله ۱۸ کلاسی انجام دهد.

مجموعه داده<sup>۴۲</sup> استفاده شده در این مقاله، شامل ۲۰۰۰ نمونه گفتار و ۵۰۰ نمونه موزیک<sup>۴۳</sup> با ژانرهای<sup>۴۴</sup> مختلف است. نمونه‌های گفتار از سایت VoxForge گرفته شده‌اند [۱۴]. این نمونه‌ها دارای طول‌های متنوع ۱ تا ۱۳ ثانیه و نرخ بیت ۲۵۶ کیلوبیت بر ثانیه هستند. نمونه‌های موزیک نیز دارای طول ۵ ثانیه و نرخ بیت ۱۴۱۱ کیلوبیت بر ثانیه (CD quality) هستند. کل این ۲۵۰۰ نمونه به ۱۸ زیرمجموعه با زیرمجموعه با اندازه تقریباً یکسان افراز شده است (اندازه هر زیرمجموعه ۱۳۸ و یا ۱۳۹ است). نمونه‌های هر زیرمجموعه توسط یکی از ۱۸ کدگذار ممکن کد می‌شوند به طوری که کدگذار زیرمجموعه‌های مختلف متمایز باشند (البته یکی از کلاس‌ها به صورت خام wav درمی‌آید). در انتها از هر داده کدشده یک بخش تصادفی به طول مشخص انتخاب می‌شود. در این مقاله چهار سناریوی مختلف برای طول بسته‌ها در نظر می‌گیریم: ۱، ۲، ۴ و ۸ کیلوبایتی.

به ازای هر طول بسته مشخص، یک مجموعه داده شامل ۲۵۰۰ بسته خواهیم داشت. نیمی از مجموعه داده به‌عنوان مجموعه آموزش<sup>۴۵</sup> و نیمی دیگر به‌عنوان مجموعه آزمون<sup>۴۶</sup> در نظر گرفته می‌شوند (این انتخاب نیز به صورت تصادفی انجام شده است). از مجموعه آموزش برای به‌دست آوردن بهترین پارامترهای دسته‌بندی با اعتبارسنجی متقابل<sup>۴۷</sup> ۵-فولد<sup>۴۸</sup> استفاده شده است. دسته‌بند مورد استفاده ماشین بردار پشتیبانی (SVM) با هسته رادیال است (جهت مقایسه، از درخت تصمیم<sup>۴۹</sup> نیز به‌عنوان یک گزینه مناسب برای دسته‌بند استفاده خواهیم نمود). پس از استخراج پارامترهای مناسب دسته‌بند، با استفاده از کل داده آموزش دسته‌بند را آموزش می‌دهیم. مجموعه آزمون نیز به‌هیچ‌عنوان برای به‌دست آوردن

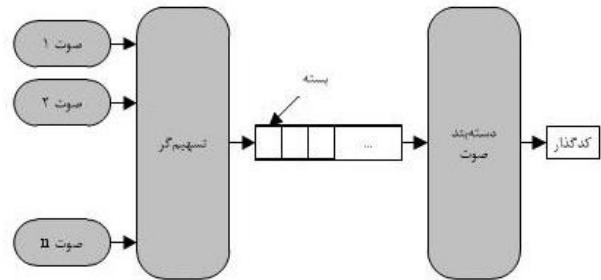
طولانی‌ترین زیرتوالی مشترک برای دسته‌بندی استفاده گردد. روش کار به این صورت است که از بسته‌های خروجی هر کدگذار،  $N$  بسته را به‌عنوان نماینده و به‌صورت تصادفی انتخاب نماییم. این نمونه‌های نماینده نباید در آموزش و آزمون دسته‌بند مورد استفاده قرار گیرند. نمونه‌های نماینده در حقیقت برای استخراج ویژگی از بسته‌های آموزش و آزمون به‌کار می‌روند. برای هر بسته آموزش و یا آزمون، طولانی‌ترین زیررشته مشترک و طولانی‌ترین زیرتوالی مشترک با هر یک از نماینده‌ها محاسبه می‌گردد. سپس میانگین نرمال شده مقدار طولانی‌ترین زیررشته مشترک بسته مورد نظر با تمام  $N$  نماینده هر کدگذار را به‌عنوان یک ویژگی در نظر می‌گیریم. همین کار را برای طولانی‌ترین زیرتوالی مشترک انجام می‌دهیم. با این کار، برای هر بسته ۳۶ ویژگی پیشنهادی به‌دست می‌آید. برای استفاده از ویژگی‌های پیشنهاد شده، دو راه وجود دارد. راه اول این است که آموزش و آزمون را صرفاً با استفاده از ۳۶ ویژگی پیشنهادی انجام دهیم. روش دومی که به ذهن متبادر می‌گردد این است که این ویژگی‌ها را در کنار ۶۵ ویژگی معرفی‌شده توسط هیکناسمز و همکاران [۸] مورد استفاده قرار دهیم (که جمعاً برابر با ۱۰۱ ویژگی می‌شود).

قبل از پرداختن به نتایج شبیه‌سازی، بهتر است مفاهیم طولانی‌ترین زیررشته مشترک و طولانی‌ترین زیرتوالی مشترک را به‌طور دقیق معرفی نماییم. برای یافتن طولانی‌ترین زیرتوالی مشترک دو رشته بایت، زیر توالی‌های مشترک بین این دو رشته بایتی با استفاده از حذف قسمت‌های غیرمشترک دو رشته به یکدیگر متصل می‌شوند. به‌طور مثال، اگر مشخص باشند، بزرگ‌ترین زیرتوالی مشترک این دو رشته  $\langle A, B, C, B, D, A, B \rangle$  و  $\langle B, C, D, B \rangle$  می‌باشد (هر یک از کاراکترهای زیرتوالی مشترک با خطی در زیرش مشخص شده است) [۳۴]. برای محاسبه طولانی‌ترین زیرتوالی مشترک بین دو رشته بایت به طول  $m$  و  $n$  از روش برنامه‌ریزی پویا<sup>۲۵</sup> با زمان اجرای  $O(mn)$  استفاده می‌کنیم [۳۵]. در شکل ۲ مثالی از نحوه به‌دست آوردن طولانی‌ترین زیرتوالی مشترک با استفاده از برنامه‌ریزی پویا نشان داده شده است.

		A	B	C	D	A
	0	0	0	0	0	0
A	0	1	1	1	1	1
C	0	1	1	2	2	2
B	0	1	2	2	2	2
D	0	1	2	2	3	3
E	0	1	2	2	3	3
A	0	1	2	2	3	4

شکل ۲: نحوه به‌دست آوردن طولانی‌ترین زیرتوالی مشترک با استفاده از برنامه‌ریزی پویا.

	Speex [26]	۲۲
	iLBC [27]	۱۳،۳۳
MDCT [29]	AAC [28]	۱۲۸
	MP3 [28]	۱۲۸
	OGG [30]	۱۲۸
	WMA	۱۲۸
بدون اتلاف	Flac [31]	-



شکل ۱: مدل کلی شناسایی و دسته‌بندی کدگذارهای صوت در شبکه‌های مخابراتی.

#### ۴ روش پیشنهادی

در تحقیقات پیشین، ویژگی‌های زیادی برای دسته‌بندی کدگذارهای صوت پیشنهاد شده است. استفاده از این ویژگی‌ها در کنار یکدیگر، به نتایج دسته‌بندی نسبتاً خوبی منجر شده است [۸]؛ اما یکی از حقایقی که از چشم این تحقیقات دور مانده است، اضافه‌سازی ذاتی موجود در خروجی کدگذارهای صوت مختلف است. در حقیقت، هرچند هدف همه کدگذارهای صوت کاهش اضافه‌سازی اطلاعاتی صوت ورودی است، اما در نهایت اضافه‌سازی بسیار کم در خروجی هر کدگذار وجود خواهد داشت. بیشتر این اضافه‌سازی جزئی به‌دلیل ساختار مشخص هر یک از این کدگذارها ایجاد می‌شود. طبق مشاهدات ما، برای بسته‌های مختلف تولیدشده توسط یک کدگذار خاص، الگوی اضافه‌سازی موجود تا حدی مشابه می‌باشد. در حقیقت، تکنولوژی کدگذارهای صوت به سمت تقویت فشرده‌سازی داده‌ها در حرکت است؛ بنابراین، ویژگی‌های استخراج‌شده در تحقیقات پیشین، به‌سختی قادر خواهند بود تا کدگذارهای صوت را با دقت بالایی از یکدیگر تفکیک کنند.

موضوع مهمی که هیچ‌یک از تحقیقات پیشین مربوط به دسته‌بندی کدگذارهای صوت آن‌ها را در نظر نگرفته‌اند، سکوت‌های موجود در صدا و الگوهای خاص متناظر با آن‌ها در خروجی کدگذار است. در حقیقت، هر کدگذار با توجه به ساختار مشخص خود می‌تواند الگوهای منحصر به فرد در رشته بیت خروجی تولید نماید که این الگوها متناظر با زمان‌های سکوت در صوت ورودی است. با توجه به توضیحات فوق، پیشنهاد می‌کنیم که از دو نوع ویژگی طولانی‌ترین زیررشته مشترک و

بیان کرده است و توانسته با دقت ۹۲/۹۱٪ سه نوع از فایل‌های مختلف از جمله PDF، RTF و DOC را از یکدیگر تفکیک کند [۳۳].

## ۵ نتایج شبیه‌سازی

برای پیاده‌سازی روش پیشنهادی، از هر کدگذار به صورت تصادفی ۱۰ بسته (یا همان نمونه) نماینده انتخاب می‌شود. ۵ نمونه از این ۱۰ نمونه را از مجموعه آموزش و ۵ نمونه دیگر را از مجموعه آزمون انتخاب می‌کنیم. این ۱۰ نمونه را از مجموعه آموزش و آزمون خارج می‌کنیم و تنها برای استخراج ۳۶ ویژگی پیشنهادی از آن‌ها استفاده می‌کنیم. مدل دسته‌بند را نیز یک مرتبه SVM و مرتبه دیگر درخت تصمیم در نظر می‌گیریم. برای هر مدل دسته‌بند، ابتدا با استفاده از داده آموزش مدل را آموزش داده و سپس با استفاده از داده آزمون صحت<sup>۴</sup> عملکرد مدل یا درصد درستی کلاس‌بندی را می‌سنجیم.

همان‌طور که قبلاً اشاره شد، در این مقاله چهار سناریوی مختلف برای طول بسته‌ها در نظر می‌گیریم: ۱، ۲، ۴ و ۸ کیلوبایتی. در جدول ۲ صحت عملکرد برای دسته‌بند SVM در سناریوهای مختلف نمایش داده شده است. سطر اول جدول ۲ نشان‌دهنده نتایج استفاده از ۶۵ ویژگی هیکسانمز و همکاران [۸] است؛ سطر دوم نشان‌دهنده نتایج استفاده از ۳۶ ویژگی پیشنهادی است؛ سطر سوم نیز نتیجه ترکیب ۶۵ ویژگی هیکسانمز و همکاران با ۳۶ ویژگی پیشنهادی ما در این مقاله (در مجموع ۱۰۱ ویژگی) است. همان‌طور که ملاحظه می‌شود، افزایش طول بسته‌ها صحت دسته‌بندی را افزایش می‌دهد. در بسته‌های با طول کوچک ۱ کیلوبایت استفاده از ترکیب ویژگی‌ها منجر به صحت بالاتری می‌شود. با این حال، با افزایش طول بسته‌ها می‌توان دید که استفاده صرف از ۳۶ ویژگی پیشنهادی دارای صحت بالاتری نسبت به روش ترکیبی است (صحت حدود ۹۷٪ برای بسته‌های ۸ کیلوبایتی). یک دلیل مشخص برای این رفتار این است که در بسته‌های با طول بزرگ، بخت یافتن زیررشته و زیرتوالی مشترک میان دو بسته مختلف مربوط به یک کدگذار خاص افزایش می‌یابد. در جدول ۳ نتیجه استفاده مجزا از ۱۸ ویژگی طولانی‌ترین زیررشته مشترک و همچنین ۱۸ ویژگی طولانی‌ترین زیرتوالی مشترک برای بسته‌های با طول ۸ کیلوبایت نشان داده شده است. همان‌طور که ملاحظه می‌شود، تأثیر طولانی‌ترین زیرتوالی مشترک برای دسته‌بندی داده‌ها بیشتر است و این ویژگی به‌تنهایی قادر به تفکیک داده‌ها با صحت ۹۰٪ است.

در جدول ۴ ماتریس درهم‌ریختگی<sup>۵۵</sup> برای دسته‌بندی بسته‌های ۸ کیلوبایتی با استفاده از ۳۶ ویژگی پیشنهادی نشان داده شده است. همان‌طور که مشاهده می‌شود، با استفاده از ویژگی‌های معرفی شده

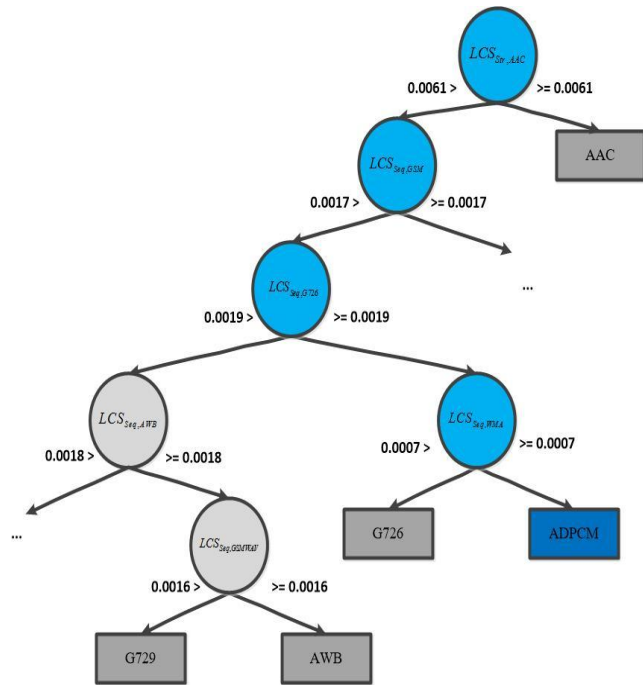
برای محاسبه طولانی‌ترین زیررشته مشترک بین دو رشته بایت، باید طولانی‌ترین زیررشته مشترک و پیوسته بین این دو رشته بیابیم. به‌طور مثال، اگر  $X = \langle A, B, C, B, D, C, B \rangle$  و  $Y = \langle B, D, C, A, B \rangle$ ، آنگاه  $\langle B, D, C \rangle$  بزرگ‌ترین زیررشته مشترک می‌باشد (هر یک از کاراکترهای زیررشته مشترک با خطی در زیرش مشخص شده است). برای محاسبه طولانی‌ترین زیررشته مشترک بین دو رشته بایت نیز از روش برنامه‌ریزی پویا استفاده می‌کنیم. در شکل ۳ مثالی از نحوه به‌دست آوردن طولانی‌ترین زیررشته مشترک با استفاده از برنامه‌ریزی پویا نشان داده شده است.

	A	B	C	X	Y	Z	A	Y
	0	0	0	0	0	0	0	0
X	0	0	0	1	0	0	0	0
Y	0	0	0	0	2	0	0	1
Z	0	0	0	0	0	3	0	0
A	0	1	0	0	0	0	4	0
B	0	0	2	0	0	0	0	0
C	0	0	0	3	0	0	0	0

شکل ۳: نحوه به‌دست آوردن طولانی‌ترین زیررشته مشترک با استفاده از برنامه‌ریزی پویا

لازم به‌ذکر است که روش پیشنهادی تنها در این مقاله استفاده نشده است و در حوزه‌های دیگر نیز مورد توجه واقع شده است. به‌عنوان مثال، در سال ۲۰۰۸، کلهون و کولز [۳۲] دو روش تفکیک‌کننده خطی<sup>۵۴</sup> فیشر و طولانی‌ترین زیررشته و زیرتوالی مشترک را با یکدیگر مقایسه کرده‌اند. ایده کلهون و کولز [۳۲] بر این اساس است که به احتمال زیاد دو فایل هم نوع (خروجی از یک کدگذار) دارای زیررشته‌های مشترک طولانی‌تری نسبت به دو فایل متفاوت (خروجی از دو کدگذار مختلف) می‌باشند. نتایج نشان می‌دهند که اگرچه ویژگی‌های طولانی‌ترین زیررشته و زیرتوالی مشترک نسبت به روش تفکیک‌کننده خطی، از لحاظ پردازش داده‌ها سریع‌تر نیستند اما داده‌ها را با دقت دسته‌بندی بالاتری از یکدیگر تفکیک می‌کنند [۳۲]. یکی دیگر از کارهای انجام شده در این زمینه شناسایی بسته‌های فایل با استفاده از طولانی‌ترین زیرتوالی مشترک است [۳۳]. هدف از این کار شناسایی فایل‌هایی است که به‌عنوان شواهد دیجیتال به مراجع قانونی ارائه می‌شوند. در این نوع از فایل‌ها معمولاً قسمت سرآیند فایل از دست‌رفته است و تنها راه ممکن برای بازیابی آن‌ها استفاده از محتوای فایل‌ها است؛ بنابراین نویسنده مقاله بهترین روش برای بازیابی این نوع فایل‌ها را استفاده از طولانی‌ترین زیرتوالی مشترک

ویژگی طولانی‌ترین زیررشته و زیرتوالی مشترک مربوط به ADPCM استفاده نشده است. با نگاهی دقیق‌تر به شکل ۴، می‌توان گفت که در مسیر منتهی به این کدگذار در درخت تصمیم، علاوه بر ویژگی‌های طولانی‌ترین زیرتوالی مشترک مربوط به کدگذارهای GSM، G726 و WMA، از ویژگی طولانی‌ترین زیررشته مشترک مربوط به کدگذار AAC استفاده شده است. توجه به این نکته ضروری است که دو کدگذار G726 و ADPCM از خانواده کدگذارهای PCM هستند و لذا ویژگی‌های طولانی‌ترین زیرتوالی و زیررشته مشترک هر یک می‌تواند به دسته‌بندی دیگری کمک نماید.



شکل ۴: بخشی از درخت تصمیم برای دسته‌بندی کدگذارها با استفاده از ۳۶ ویژگی پیشنهادی و بسته‌های ۱ کیلوبایتی.

### ۶ نتیجه‌گیری

تاکنون کارهای زیادی در زمینه دسته‌بندی داده‌های چندرسانه‌ای انجام گرفته شده است. تمامی کارهای انجام گرفته شده در این زمینه، سعی در استخراج ویژگی‌هایی از داده‌های کدگذاری شده دارند تا بتوانند با استفاده از ویژگی‌های مناسب، داده‌های چندرسانه‌ای را با صحت بالاتری از یکدیگر تفکیک نمایند. در این مقاله استفاده از ویژگی‌های طولانی‌ترین زیررشته و زیرتوالی مشترک برای تفکیک کدگذارهای صوتی پیشنهاد

بسته‌های مربوط به شش کدگذار Speex، G726، AMR، WAV، OGG و CVSD با صحت ۱۰۰٪ و بسته‌های مربوط به سایر کدگذارها با صحت بالای ۹۳٪ شناسایی شده‌اند.

جدول ۲: نتایج حاصل از دسته‌بندی کدگذارها با استفاده از الگوریتم SVM

SVM				ویژگی‌ها
۸ کیلوبایت	۴ کیلوبایت	۲ کیلوبایت	۱ کیلوبایت	
۸۹/۸۳٪	۸۹/۷۸٪	۸۱/۱۷٪	۷۴/۰۶٪	۶۵ ویژگی هیکسانمز و همکاران [۸]
۹۷/۹۶٪	۹۵/۰۵٪	۹۳/۷۵٪	۸۸/۰۴٪	۳۶ ویژگی پیشنهادی
۹۴/۳۸٪	۹۳/۵۹٪	۹۲/۹۴٪	۹۰/۵۷٪	ترکیب ویژگی‌های پیشنهادی و ویژگی‌های هیکسانمز

جدول ۳: مقایسه قدرت دسته‌بندی ویژگی‌های طولانی‌ترین زیررشته در مقایسه با ویژگی‌های طولانی‌ترین زیرتوالی مشترک برای بسته‌های با طول ۸ کیلوبایت و الگوریتم SVM

۸ کیلوبایت	ویژگی‌ها
۶۴/۱۷٪	طولانی‌ترین زیررشته مشترک
۹۰/۰۷٪	طولانی‌ترین زیرتوالی مشترک

استفاده از درخت تصمیم برای دسته‌بندی دارای یک مزیت بسیار مهم است. با آموزش یک درخت تصمیم می‌توان میزان اهمیت و اولویت ویژگی‌های مختلف در دسته‌بندی را شناسایی نمود. در جدول ۵ نتایج استفاده از دسته‌بند درخت تصمیم نشان داده شده است.

همان‌طور که ملاحظه می‌شود، در حل مسئله مدنظر ما، دسته‌بند درخت تصمیم عملکرد ضعیف‌تری نسبت به دسته‌بند SVM دارد. با این حال می‌توان از درخت تصمیم حاصل دیدگاه بهتری نسبت به ویژگی‌های مختلف به‌دست آورد. برای این منظور، باید نگاهی دقیق‌تر به درخت تصمیم حاصل از ۳۶ ویژگی پیشنهادی داشته باشیم تا به نحوه کارکرد این ویژگی‌ها پی ببریم.

در شکل ۴، بخشی از درخت تصمیم حاصل برای بسته‌های ۱ کیلوبایتی نمایش داده شده است. همان‌طور که در این شکل ملاحظه می‌شود، دو ویژگی طولانی‌ترین زیررشته و زیرتوالی مشترک هر کدگذار مشخص لزوماً تنها در تصمیم‌گیری در مورد آن کدگذار خاص ایفای نقش نمی‌کنند. به‌طور مثال، همان‌طور که در شکل ۴ دیده می‌شود، برای تفکیک کدگذار ADPCM از سایر کدگذارها، در حقیقت از هیچ‌یک از دو

شده و مورد ارزیابی قرار گرفته است. این ویژگی‌ها در حقیقت نشان‌دهنده جریان‌های بایستی مشخص در خروجی کدگذارهای مختلف هستند. چنین جریان‌های مشخصی وابسته به ساختار خاص هر کدگذار می‌باشند. نتایج شبیه‌سازی نشان می‌دهد که ترکیب ویژگی‌های پیشنهادی با ویژگی‌های هیگسامز و همکاران باعث افزایش صحت دسته‌بندی تا حدود ۹۰٪ در بسته‌های یک کیلوبایتی می‌شود. با این حال، با افزایش طول بسته‌ها، استفاده از ویژگی‌های پیشنهادی باعث حصول نتایج بهتر نسبت به روش ترکیبی می‌گردد.



جدول ۴: ماتریس درهم‌ریختگی دسته‌بندی با استفاده از ۳۶ ویژگی پیشنهادی برای بسته‌های با طول ۸ کیلوبایت و دسته‌بند SVM

Codecs	aac	mp3	ogg	wma	wav	a-law	u-law	adpcm	amr	awb	gsm	gsmwav	g729	g726	ilbc	speex	flac	cvsd
aac	۷/۱ ۹۵	.	.	.	.	.	.	.	.	۱/۴۳	.	.	.	.	۲/۸۶	.	.	.
mp3	.	۹۸/۶۱	.	۱/۳۹	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ogg	.	.	۱۰۰	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
wma	.	.	.	۹۷/۱	.	.	.	.	.	.	.	.	.	.	.	.	.	.
wav	.	.	.	.	۱۰۰	.	.	.	.	.	.	.	.	.	.	.	.	.
a-law	.	.	.	.	.	۹۸/۴۶	.	.	.	.	.	.	.	.	.	.	۱/۵۴	.
u-law	.	.	.	.	.	۹۸/۴۸	۱/۳۵	.	.	.	.	.	.	.	.	.	.	.
adpcm	.	.	.	.	.	.	.	۹۶/۶۷	.	.	۳/۳۳	.	.	.	.	.	.	.
amr	.	.	.	.	.	.	.	.	۱۰۰	.	.	.	.	.	.	.	.	.
awb	.	.	.	.	.	.	.	.	۱/۳۵	۹۸/۶۵	.	.	.	.	.	.	.	.
gsm	.	.	.	.	.	.	.	.	.	.	۹۲/۶۶	۲/۶۷	۲/۶۷	.	.	.	.	.
gsm wav	.	.	۱/۳۳	.	.	.	.	.	.	.	.	۹۸/۶۷	.	.	.	.	.	.
g729	.	.	.	.	.	.	.	.	.	۲/۱۳	.	.	۹۳/۷۵	.	۱/۵۶	.	.	.
g726	.	.	.	.	.	.	.	.	.	.	.	.	.	۱۰۰	.	.	.	.
ilbc	.	.	.	.	.	.	.	.	.	۱/۳	.	.	۲/۹	.	۹۴/۸	.	.	.
speex	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	۱۰۰	.	.
flac	.	.	.	.	.	.	.	.	.	.	.	.	۱/۴۹	.	.	.	۷۸	.
cvsd	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	۱۰۰

- [13] F. Jenner, A. Kwasinski, "Highly accurate non-intrusive speech forensics for codec identifications from observed decoded signals", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1737-1740, 2012.
- [14] *voxforge*. Available: <http://www.voxforge.org/>
- [15] S. Kant, V. Sharma, N. Verma, and B. Dass, "Identification scheme for romanized Indian languages from their plain and ciphered bit stream," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 13, pp. 329-345, 2010.
- [16] S. Elgar and R. Guza, "Statistics of bicoherence," *IEEE Transactions On Acoustics Speech And Signal Processing*, vol. 36, pp. 1667-1668, 1988.
- [17] L. Benetazzo, M. Bertocco, P. Paglierani, and E. Rizzi, "Speech/voice-band data classification for data traffic measurements in telephone-type systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, pp. 413-417, 2000.
- [18] A. Swami, J. M. Mendel, and C. L. M. Nikias, "Higher-order spectral analysis toolbox," *Technical support and product enhancement suggestions*, 1984.
- [19] R. Hegger, H. Kantz, and T. Schreiber, "Practical implementation of nonlinear time series methods: The TISEAN package," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 9, pp. 413-435, 1999.
- [20] W. C. Chu, *Speech coding algorithms: foundation and evolution of standardized coders*: John Wiley & Sons, 2004.
- [21] T. Equipments, "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," *ITU-T Recommendation, G*, vol. 7, p. 59, 1990.
- [22] M. K. Ganapathiraju, N. Balakrishnan, and R. Reddy, "Improving Recognition Accuracy on CVSD Speech under Mismatched Conditions," *School of Computer Science, Carnegie Mellon University (Pittsburgh, PA)(undated)*.
- [23] E. ETSI, "301 704 V7. 2.1 (2000-04)," *Digital cellular telecommunications system (Phase 2+)*, 2000.
- [24] G. Recommendation, "722.2: "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)", ed: ITU-T, 2003.
- [25] L. H. Swee, "Implementation of G. 729 on the TMS320C54x," *Texas Instrument Application Report*, 2000.
- [26] J.-M. Valin, "The Speex Codec Manual Version 1.2 Beta 3," *Xiph.org Foundation*, 2007.
- [27] S. Andersen, A. Duric, H. Astrom, R. Hagen, W. Kleijn, and J. Linden, "Internet low bit rate codec (iLBC)," 2070-1721, 2004.
- [28] K. Brandenburg, "MP3 and AAC explained," in *Audio Engineering Society Conference: 17th International Conference: High-Quality Audio Coding*, 1999.
- [29] V. Britanak and K. R. Rao, "A new fast algorithm for the unified forward and inverse MDCT/MDST computation," *Signal Processing*, vol. 82, pp. 433-459, 2002.
- [30] S. Pfeiffer, "The ogg encapsulation format version 0," 2003.
- [31] *FLAC: free lossless audio codec*. Available: <https://xiph.org/flac/format.html>
- [32] Calhoun, William C., and Drue Coles. "Predicting the types of file fragments." *Digital Investigation* 5,S14-S20, 2008.

## جدول ۵: نتایج حاصل از دسته‌بندی کدگذارها با استفاده از درخت

تصمیم				
۸	۴	۲	۱	ویژگی‌ها
کیلوبایت	کیلوبایت	کیلوبایت	کیلوبایت	
۷۹/۵۰٪	۷۵/۰۲٪	۷۱/۶۷٪	۶۹/۱۳٪	۶۵ ویژگی هیکسانمز و همکاران [۸]
۹۰/۴۸٪	۹۰/۳۵٪	۸۷/۵۰٪	۸۳/۰۲٪	۳۶ ویژگی پیشنهادی
۸۴/۵۴٪	۸۵/۳۳٪	۸۸/۶۴٪	۸۵/۱۳٪	ترکیب ویژگی‌های پیشنهادی و ویژگی‌های هیکسانمز

## مراجع

- [۱] مهدی تیموری، «آشکارسازی سیگنال لینک ۱۶»، *مجله مهندسی برق دانشگاه تبریز*، جلد ۴۶، شماره ۴، صفحات ۸۴-۸۷، ۱۳۹۵.
- [۲] مهدی تیموری، حمید رضا کاکایی مطلق و مرتضی حدادی، «شناسایی کور کدهای ضربی BCH»، *مجله مهندسی برق دانشگاه تبریز*، جلد ۴۷، شماره ۱، صفحات ۴۹-۵۴، ۱۳۹۶.
- [3] M. Karresand and N. Shahmehri, "File type identification of data fragments by their binary structure," in *2006 IEEE Information Assurance Workshop*, pp. 140-147, 2006.
- [4] M. C. Amirani, M. Toorani, and A. Beheshti, "A new approach to content-based file type detection," in *Computers and Communications, 2008. ISCC 2008. IEEE Symposium on*, pp. 1103-1108, 2008.
- [5] M. Din, R. Ratan, A. K. Bhateja, and A. Bhateja, "Multimedia Classification Using ANN Approach," in *Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012)*, pp. 905-910, December, 2012.
- [6] P. Tripathi, K. P. Raju, and N. S. Kumar, "A Novel Methodology for Identification of Unclassified Digital Voice," *International Journal of Application or Innovation in Engineering & Management*, vol. 2, pp. 297-302, 2013.
- [7] X. Jin and J. Kim, "Audio Fragment Identification System," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 9, pp. 307-320, 2014.
- [8] S. Hicsonmez, H. T. Sencar, and I. Avcibas, "Audio codec identification from coded and transcoded audios," *Digital Signal Processing*, vol. 23, pp. 1720-1730, 2013.
- [9] P. Tripathi, K. P. Raju, and V. R. Chandra, "A Novel Technique for Detection of CVSD Encoded Bit Stream," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, pp. 6035-6040, 2014.
- [10] W. Xianyan, H. Qi, L. Dan, and N. Xiamu, "A New Method for Estimating the Widths of JPEG Images," *arXiv preprint arXiv:1410.2100*, 2014.
- [11] D. Alley, "Automatic identification of voice band telephony coding schemes using neural networks", *Electron. Lett.* Vol. 29, No. 13, pp.1156-1157, 1993.
- [12] K. Scholz, L. Leutelt, U. Heute, "Speech-codec detection by spectral harmonic-plus-noise decomposition", *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 2295-2299, 2004.

- [35] A. Chaudhuri, "A Dynamic Algorithm for the Longest Common Subsequence Problem using Ant Colony Optimization Technique," *arXiv preprint arXiv:1307.1905*, 2013.
- [36] L. Allison and T. I. Dix, "A bit-string longest-common-subsequence algorithm," *Information Processing Letters*, vol. 23, pp. 305-310, 1986.
- [33] Rahmat, R. F., et al. "File Type Identification of File Fragments using Longest Common Subsequence (LCS)." *Journal of Physics: Conference Series*. Vol. 801. No. 1. IOP Publishing, 2017.
- [34] W .C. Calhoun and D. Coles, "Predicting the types of file fragments," *Digital Investigation*, vol. 5, pp. S14-S20, 2008.

## زیرنویس‌ها

- 
- <sup>1</sup> Multimedia data
  - <sup>2</sup> File fragment identification
  - <sup>3</sup> Identification based on extension
  - <sup>4</sup> Identification based on magic byte
  - <sup>5</sup> Histogram
  - <sup>6</sup> Header
  - <sup>7</sup> Footer
  - <sup>8</sup> N-gram
  - <sup>9</sup> Principal component analysis
  - <sup>10</sup> Artificial neural network
  - <sup>11</sup> Pulse code modulation
  - <sup>12</sup> Linear prediction coding
  - <sup>13</sup> Adaptive differential pulse code modulation
  - <sup>14</sup> Auto-correlation
  - <sup>15</sup> Central moment
  - <sup>16</sup> Binary ratio
  - <sup>17</sup> Sync word
  - <sup>18</sup> Structural
  - <sup>19</sup> Statistical
  - <sup>20</sup> Frame
  - <sup>21</sup> Hicsonmez
  - <sup>22</sup> Randomness
  - <sup>23</sup> Chaotic
  - <sup>24</sup> Mean
  - <sup>25</sup> Variance
  - <sup>26</sup> Entropy
  - <sup>27</sup> Bicoherence
  - <sup>28</sup> Skewness
  - <sup>29</sup> Kortusis
  - <sup>30</sup> False nearest
  - <sup>31</sup> Lyapunov exponent
  - <sup>32</sup> Sequential floating forward search
  - <sup>33</sup> Support vector machine
  - <sup>34</sup> Radial kernel
  - <sup>35</sup> Continuously variable slope delta modulation
  - <sup>36</sup> Joint photographic experts group
  - <sup>37</sup> Spectral harmonic-plus-noise decomposition
  - <sup>38</sup> Multi-dimensional profile
  - <sup>39</sup> Noise spectra
  - <sup>40</sup> Time-domain amplitude histogram
  - <sup>41</sup> Multiplexer
  - <sup>42</sup> Data-set
  - <sup>43</sup> Music
  - <sup>44</sup> Genre
  - <sup>45</sup> Train
  - <sup>46</sup> Test
  - <sup>47</sup> Cross validation
  - <sup>48</sup> Fold
  - <sup>49</sup> Decision tree
  - <sup>50</sup> Higher order spectral analysis
  - <sup>51</sup> False neighbor fraction
  - <sup>52</sup> Dynamic programming
  - <sup>53</sup> Linear discriminant
  - <sup>54</sup> Accuracy
  - <sup>55</sup> Confusion Matrix