

علوم و مهندسی کامپیوتر

نشریه علمی پژوهشی انجمن کامپیوتر ایران

صاحب امتیاز: انجمن کامپیوتر ایران
مدیر مسئول: دکتر جعفر حبیبی
سر دبیر: دکتر حمید سربازی آزاد

شورای علمی

- گ. آقا، استاد دانشگاه ایلینویز، امریکا
ف. ارباب، استاد سی.دبلیو.آی و دانشگاه لایدن، هلند
ن. باقرزاده، استاد دانشگاه کالیفرنیا- پروین، امریکا
ک. بدیع، دانشیار مرکز تحقیقات مخابرات، ایران
ب. بوزه، استاد دانشگاه ایالتی اورگان، امریکا
ب. پرهامی، استاد دانشگاه کالیفرنیا- سنتا باربارا، امریکا
ف. جهانیان، استاد دانشگاه میشیگان، امریکا
آ. زومایا، استاد دانشگاه سیدنی، استرالیا
ح. سربازی آزاد، دانشیار دانشگاه صنعتی شریف، ایران
ب. شیرازی، استاد دانشگاه ایالتی واشنگتن، امریکا
ا. کبیر، دانشیار دانشگاه تربیت مدرس، ایران
ر. صفابخش، استاد دانشگاه صنعتی امیرکبیر، ایران
ح. ر. عربنیا، استاد دانشگاه جورجیا، امریکا
ع. عدالت، استاد کالج سلطنتی لندن، انگلستان
م. فتحی، دانشیار دانشگاه علم و صنعت، ایران
م. ح. قاسمیان، استاد دانشگاه تربیت مدرس، ایران
م. قدسی، استاد دانشگاه صنعتی شریف، ایران
ع. هورسان، استاد دانشگاه ایالتی پنسیلوانیا، امریکا
ف. لو، استاد دانشگاه هنگ کنگ، چین
ع. موقر، استاد دانشگاه صنعتی شریف، ایران
ن. مهدوی امیری، استاد دانشگاه صنعتی شریف، ایران
م. ر. میبیدی، استاد دانشگاه صنعتی امیرکبیر، ایران
ک. ناکانو، استاد دانشگاه هیروشیما، ژاپن
م. ولدخوا، استاد دانشگاه گلاسگو، انگلستان

مشاوران

- ج. آرمسترانگ، دانشگاه نورث ایسترن، امریکا
ا. المقارم، دانشگاه پردو، امریکا
م. انواری، مرکز فناوری اطلاعات و ارتباطات پیشرفته شریف، ایران
ح. برنجی، ناسا، امریکا
د. بیورنر، دانشگاه دلفت، هلند
م. پدram، دانشگاه کالیفرنیا جنوبی، امریکا
ب. ذوقی، دانشگاه ای.اند.ام تگزاس، امریکا
پ. روگای، دانشگاه کالیفرنیا- دیویس، امریکا
م. قنبری، دانشگاه اسکس، انگلستان
م. صرافزاده، دانشگاه کالیفرنیا- لس آنجلس، امریکا

همکاران دفتر مجله: فاطمه قاسمی، شهیندخت شهنازی، لیلا نورانی، محمدرضا حسینی فرح آبادی

نشانی: تهران، خیابان آزادی، ضلع غربی دانشگاه صنعتی شریف، کوچه شهید ولی... صادقی، پلاک ۲۶، طبقه ۴، واحد ۱۶، دفتر انجمن کامپیوتر ایران،
نشریه علوم و مهندسی کامپیوتر. تلفن ۰۳۳۲-۲۲۲۸۰، دورنگار ۲۲۸۲۸۶۸۷
پست الکترونیکی: csi-jcse@ipm.ir، تور جهان گستر: <http://cs.ipm.ac.ir/jcse>

مقالات درج شده در این نشریه صرفاً بیانگر نظرات مؤلفین آنها است و مسئولیت صحت و سقم داده‌ها و نتایج بر عهده آنها است.

فهرست مقالات

- طرح دو فیلتر جدید برای بهبود کیفیت گفتار مبتنی بر توزیع احتمال پسین برای ضرایب موجک ۱
بهزاد زمانی دهکردی، احمد اکبری، بابک ناصرشریف
- چراغ راهنمای هوشمند ۱۶
امین‌اله مه‌آبادی
- استفاده از فناوری پردازش گفتار در ارزیابی خودکار روانی گفتار در سیستم‌های آموزش زبان خارجی ۳۰
بهزاد بختیاری، محمد مهدی همایون‌پور
- اجتماع نمونه‌دهی هستان‌شناسی و حاشیه‌نویسی معنایی متون فارسی در سیستم POPTA ۴۰
مهرنوش شمس‌فرد، بهاره صراف‌زاده
- یک الگوریتم جدید بر مبنای اتوماتاهای یادگیر برای حل مسئله درخت اشتاینر ۵۳
سمیرا نوفرستی، محمدرضا میبیدی
- استخراج ویژگی‌ها با استفاده از اطلاعات متقابل جهت طبقه‌بندی سیگنال‌های مغزی در سیستم‌های ارتباطی مغز با کامپیوتر ۶۰
فرید اویسی ارنگه، عباس عرفانیان امیدوار
- ارزیابی کمی سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد در نرم‌افزارهای حساس به ایمنی ۶۸
غلامرضا شاه محمدی، سعید جلیلی مهران

طرح دو فیلتر جدید برای بهبود کیفیت گفتار مبتنی بر توزیع احتمال پسین برای ضرایب موجک

بهزاد زمانی دهکردی^۱ احمد اکبری^۱ بابک ناصرشریف^{۱و۲}

^۱ دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران
^۲ دانشکده فنی و مهندسی، دانشگاه گیلان، گیلان، ایران

چکیده

روش‌های حذف نویز مبتنی بر موجک عموماً بر اساس آستانه‌گذاری استوارند که این امر نیز ریشه در فرض تمرکز انرژی سیگنال در تعداد کمی از ضرایب تبدیل موجک سیگنال دارد. عیب عمده روش‌های آستانه‌گذاری در سیگنال گفتار، تخریب گفتار در فرکانس‌های بالا و عملکرد ضعیف آنها در حضور نویزهای رنگی است. از این‌رو دو روش فیلتر کردن ضرایب تبدیل موجک به عنوان جایگزینی برای روش‌های آستانه‌گذاری مطرح می‌گردد. در مقاله حاضر ابتدا با انگیزه نشان دادن برتری روش‌های فیلتر کردن بر آستانه‌گذاری، دو روش حذف نویز آستانه‌گذاری نرم با حد آستانه SURE و فیلتر وینر با هم مقایسه شده‌اند. علاوه بر این دو فیلتر جدید بر مبنای توزیع احتمال پسین ضرایب تبدیل موجک پیشنهاد شده است. نتایج کار حاضر نشان می‌دهند که فیلتر وینر نسبت به روش آستانه‌گذاری، دارای عملکرد بهتر و کیفیت شنیداری بالاتری می‌باشد. به علاوه، مقایسه فیلترهای پیشنهادی با فیلتر وینر در حوزه موجک نشان می‌دهد که این فیلترها دارای کارایی بالاتری نسبت به فیلتر وینر هستند، همچنین فیلترهای پیشنهادی در حوزه موجک نسبت به همتایشان در حوزه فرکانس، نسبت سیگنال به نویز را بیشتر افزایش داده و اعوجاج کمتری بر روی سیگنال تمیز ایجاد می‌نمایند.

کلمات کلیدی: تبدیل موجک، فیلتر وینر، درخت موجک بسته‌ای، آستانه‌گذاری، MAP.

۱- مقدمه

محیط کار یا نویز مربوط به وزش باد در محیط آزاد نیز از جمله عوامل افت کیفیت صوت در سیستم‌های هدست می‌باشد. یکی از کاربردهای دیگر بهبود گفتار سیستم‌های بازشناسی گفتار می‌باشند. معمولاً سیستم‌های بازشناسی گفتار با سیگنال‌های صوتی ضبط شده در محیط آزمایشگاهی آموزش داده می‌شوند، در حالیکه در عمل باید در محیط‌های واقعی که در آن نویز وجود دارد مورد استفاده قرار گیرند. در این حالت برای جلوگیری از پایین آمدن دقت بازشناسی، نیاز به حذف نویز از گفتار احساس می‌شود. نمونه دیگری از کاربرد حذف نویز کدکننده‌هایی هستند که برای کد کردن گفتار بکار می‌روند. این کدکننده‌ها معمولاً برای کار با سیگنال تمیز طراحی می‌گردند حال آنکه در عمل برای شرایط واقعی مورد استفاده قرار می‌گیرند که سیگنال گفتار آغشته به نویز محیط می‌باشد.

استفاده از بهبود گفتار، در بسیاری از کاربردهایی که صدا در یک محیط نویزی دریافت می‌گردد، ضروری است. از جمله این موارد به دریافت صدا از طریق میکروفن تلفن، بدون استفاده از گوشی^۱ و نیز میکروفن‌های رومیزی همه‌جهته در یک کنفرانس صوتی می‌توان اشاره کرد. هدست‌هایی که روی گوش قرار می‌گیرند و با تکنولوژی بلوتوث^۲ با تلفن همراه ارتباط برقرار می‌کنند نیز برای حفظ کیفیت صوت به بهبود صوت نیاز دارند. در این دستگاه‌ها به علت فاصله چند سانتیمتری میکروفن با دهان گوینده صدای مربوط به نویز محیط اطراف نیز دریافت می‌گردد و موجب افت کیفیت صوت می‌شود. سر و صدای اضافی مانند همهمه افراد در

می‌شود و حد آستانه در باندهای بالایی برای فریم واکنش افزایش و برای فریم بی‌واک کاهش می‌یابد [۹]. در کل روش‌های آستانه در حذف نویزهای سفید و ایستا خوب عمل می‌کنند. در حالیکه عملکرد خوبی در نویزهای رنگی ندارند. برای حل این مشکل در سال ۲۰۰۳، Wang و Lu روش آستانه‌گذاری در باندهای بحرانی درخت تجزیه موجک را ارائه دادند. در این تکنیک ابتدا آستانه ماسک نویز، برای هر فریم تخمین زده شده و سپس این آستانه برای تنظیم آستانه نهایی بکار گرفته می‌شود. اندازه آستانه بطور اتوماتیک در باندهای بالا در طول یک سگمنت واکنش و در باندهای پائین در طول یک سگمنت بی‌واک افزایش می‌یابد [۱۰].

روش‌های آستانه‌گذاری از نظر نتایج شنوائی [۱۱، ۱۲] دارای عملکرد ضعیفی هستند [۱۳، ۱۴، ۳۴]. برای حل مشکل روش‌های حد آستانه، استفاده از فیلتر مطرح شده است که در آن پارامترهای فیلتر وابسته به شرایط آماری ضرایب تبدیل موجک تنظیم می‌گردد. فیلتر وینر و فیلتر تفاضل طیف و فیلتر کالمن نمونه‌هایی از فیلترهای بکار رفته در حوزه تبدیل موجک هستند [۱۵، ۱۶، ۱۷]. بطور عمده روش‌های طراحی فیلتر بهبود گفتار در حوزه موجک با الهام از اصول روش‌های طراحی فیلتر بهبود گفتار در حوزه فرکانس انجام می‌شوند. در مقاله حاضر، طراحی فیلتر در حوزه موجک و بر اساس بیشینه کردن احتمال پسین برای ضرایب موجک سیگنال تمیز با الهام از مرجع [۱۸، ۱۹] در حوزه تبدیل فوری مورد توجه قرار گرفته است. از آنجا که روش کمینه کردن خطای میانگین مربعات (MMSE) و میزان نویز غیر رنگی باقیمانده در مرجع [۱۸]، نیازمند محاسبات نمائی و توابع بسل است، از همین رو استفاده از فیلترهایی که با محاسبات کمتر عملکردی نزدیک به روش کمینه کردن خطای میانگین مربعات داشته باشند، در مرجع [۱۹] مد نظر قرار گرفته است. در مقاله حاضر بر همین روال عمل شده است. بر این اساس دو فیلتر در حوزه موجک در این مقاله پیشنهاد شده‌اند که در مقایسه با فیلتر وینر نتایج بهتری داشته‌اند. همچنین در مقایسه با فیلترهای مشابهشان در حوزه تبدیل فوری، دارای نتایج بهتری بوده‌اند.

ساختار ادامه مقاله به این صورت است که بخش دوم مقاله، به مروری کلی بر تبدیل موجک و معرفی انواع درخت تجزیه موجک اختصاص دارد. بخش سوم به بررسی ویژگیهای موجک و دو گرایش کلی حذف نویز در حوزه موجک از قبیل فیلتر کردن وینر و آستانه‌گذاری نرم پرداخته می‌شود. در ادامه در بخش چهارم با استفاده از دو معیار MAP-G و MAP-SG دو فیلتر در حوزه تبدیل موجک پیشنهاد شده است. بخش پنجم نیز به مقایسه دو روش فیلتر کردن وینر و روش آستانه‌گذاری نرم اختصاص دارد. سپس در بخش ششم نتایج بدست آمده از پیاده‌سازی دو فیلتر پیشنهادی در برابر فیلتر وینر در دو حوزه موجک و فرکانس آورده شده است. در آخر نیز به جمع‌بندی کلی اختصاص دارد.

۲- تبدیل موجک

تبدیل موجک پیوسته (که انتگرال موجک نیز نامیده می‌شود) توسط Grossman و Morlet پایه گذاری شده است [۲۰]. تئوری موجک بر مبنای یک مجموعه از فیلترهایی است که توسط انبساط و انقباض موجک مولد ایجاد می‌شوند. همه موجک‌ها از مقیاس موجک مادر بدست می‌آیند. بنابراین فقط کافی است که یک فیلتر طراحی شده و بقیه فیلترها با تغییر مقیاس در زمان و فرکانس روی این فیلتر بدست می‌آیند.

شکل ۱ ساختار ساده تبدیل موجک را نشان می‌دهد. در این تبدیل سیگنال ورودی به دو زیرباند بالا و پائین تفکیک می‌شود. در ادامه برای بازسازی از دو فیلتر دیگر استفاده می‌شود. آنچه در شکل دیده می‌شود درخت موجکی با یک سطح می‌باشد. $h_a(k)$ فیلتر پائین‌گذر و $h_d(k)$ فیلتر بالاگذر می‌باشند. پاسخ فرکانسی آنها وابسته به نوع تابع پایه موجک می‌باشد و برای تجزیه سیگنال به کار

در این حالت نیز بهبود گفتار و حذف نویز می‌تواند صحت عملکرد کدکننده را افزایش دهد.

برای حذف نویز از سیگنال گفتار، می‌توان طیف فرکانسی سیگنال نویزی را به گونه ای اصلاح نمود که پس از بازیابی سیگنال در حوزه زمان، صداهای پس زمینه گفتار حذف گردد و یا کاسته شود. به این منظور ابتدا باید سیگنال گفتار با تبدیلی نظیر تبدیل فوری به حوزه فرکانس انتقال یابد و سپس اصلاحات لازم بر روی طیف فرکانسی به شیوه‌ای صورت گیرد تا نویز کاهش یابد. از جمله‌ی شناخته شده‌ترین این روشها، روش تفاضل طیف توان و نیز فیلتر وینر^۲ اعمال شده بر طیف می‌باشد [۱]. گرچه این روش‌ها در عمل موفق بوده‌اند و پیاده سازی آنها نیز بسیار ساده می‌باشد. با این حال مشکلات و کاستی‌هایی نیز دارند. یکی از معایب این روشها اعوجاجی است که روی سیگنال مفید ایجاد می‌نمایند. مشکل دیگر ایجاد یک نوع نویز پس زمینه مصنوعی در این روشهاست که به نویز موسیقی مشهور شده است. علاوه بر این بنا به ماهیت تبدیل فوری و نایستانت بودن سیگنال گفتار، عملیات تفاضل طیف و یا فیلتر کردن باید بر روی طیف یک پنجره محدود از سیگنال صورت گیرد که کوچک بودن پنجره عموماً موجب محدودیت در رزولوشن فرکانسی طیف سیگنال صوتی می‌گردد و کارائی اصلاح طیف و یا فیلتر کردن پائین می‌آید.

از دیگر روش‌های بهبود گفتار استفاده از تبدیل موجک است. ویژگی اصلی تبدیل موجک بهره جستن از پنجره‌های زمانی با طول متفاوت برای باندهای فرکانسی مختلف است. به این ترتیب با استفاده از تبدیل موجک می‌توان به دقت فرکانسی بالا در باندهای فرکانسی پایین در عین حفظ دقت در رزولوشن زمانی دست یافت. به این ترتیب تبدیل موجک محدودیت‌های ناشی از کار با پنجره محدود در تبدیل فوری را ندارد. مهمترین روش برای بهبود گفتار با استفاده از تبدیل موجک، روش آستانه‌گذاری ضرایب تبدیل موجک است. این روش بر این اساس عمل می‌کند که سیگنال گفتار همانند بسیاری از سیگنالهای دیگر، انرژی سیگنال غالباً در تعداد کمی از ضرایب تبدیل موجک متمرکز است. این ضرایب نسبت به ضرایب دیگر تبدیل موجک خود سیگنال، یا ضرایب تبدیل موجک هر سیگنال دیگری (بویژه نویز) که انرژی آن بر روی تعداد زیادی از ضرایب گسترش می‌یابد، بزرگتر هستند. بنابراین می‌توان با صفر کردن ضرایب کوچکتر، علاوه بر حفظ اطلاعات مهم سیگنال اصلی، نویز را نیز محدود کرد. مطابق با این ویژگی، ضرایب تبدیل موجک سیگنال با یک حد آستانه^۳ مقایسه می‌شوند و مقدار ضرایب کوچکتر از این حد آستانه برابر صفر قرار داده می‌شوند. این عمل را می‌توان به نوعی، فیلتر کردن ضرایب تبدیل موجک دانست. روش‌های حد آستانه علاوه بر حذف نویز مستقیم از ضرایب موجک، برای تخمین طیف فرکانسی سیگنال گفتار نیز بکار گرفته شده‌اند [۲، ۳، ۴].

در آستانه‌گذاری ضرایب موجک دو مسئله مطرح می‌شود: شیوه تعیین حد آستانه و نحوه اعمال حد آستانه بر ضرایب. مسئله دوم با تعیین یک تابع موسوم به تابع آستانه‌گذاری حل می‌شود. توابع آستانه‌گذاری متفاوتی همچون توابع آستانه‌گذاری نرم، سخت، نیمه‌نرم و گاروت و گاروت پله‌ای^۵ در کارهای تحقیقاتی متفاوت تعریف شده‌اند [۵، ۳۴]. برای تعیین حد آستانه روش‌های متفاوتی وجود دارد که حد آستانه را مستقل یا وابسته به تابع آستانه‌گذاری تعیین می‌نمایند. روش‌هایی چون کمینه کردن نرخ یابش نادرست^۶ (MFDR) [۵] و روش تاپ^۷ [۶] حد آستانه را مستقل از نوع تابع اعمال حد آستانه محاسبه می‌نمایند. روش کمینه کردن تخمین مخاطره نا ارب استین^۸ (SURE) نیز روش دیگری برای تعیین حد آستانه است که به نوع تابع اعمال حد آستانه وابسته است.

رایج‌ترین روش تعیین حد آستانه، روشی است که Johnston و Donoho برای حذف نویز سفید گوسی پیشنهاد داده‌اند. این روش برای نویزهای رنگی به خوبی عمل نمی‌کند [۷، ۸]. شیخ‌زاده و ابوطالبی در سال ۲۰۰۱ روشی بهبود یافته پیشنهاد داده‌اند که در آن سیگنال به فریم‌های واکنش و بی‌واک تقسیم‌بندی

شکل ۴ یک درخت موجک بسته ای که دارای ۴ زیرباند یکنواخت می باشد را نشان می دهد. در درخت موجک بسته ای همه زیرباندها و نودهای داخلی درخت به یک میزان رشد می کنند، در نتیجه تحلیل فرکانسی در تمام محدوده فرکانسی یکسان خواهد بود.

۳- معرفی دو گرایش مختلف برای حذف نویز از ضرایب موجک

روش های حذف نویز از سیگنال گفتار در حوزه فرکانس، از این واقعیت استفاده می نمایند که طیف نویز جمع شونده ایستادن، در حوزه فرکانس قابل تخمین بوده و تکنیک پذیرای نویز و سیگنال گفتار در این حوزه بهتر انجام می گیرد. در این روش ها فیلتر حذف نویز با استفاده از تخمین طیف نویز محاسبه می گردد و بر سیگنال نویزی اعمال می شود. برای تخمین پارامترهای فیلتر بایستی سیگنال فریم بندی شود. از طرفی در تبدیل فوریه نسبت تعداد نمونه ها در حوزه زمان با تعداد زیرباندها برابر می باشد. بنابراین برای داشتن جزئیات فرکانسی بیشتر باید پنجره آنالیزی بزرگتری در نظر بگیریم. انتخاب پنجره آنالیزی بزرگ از یک طرف موجب می گردد جزئیات زمانی سیگنال در نظر گرفته نشود و قسمتهای گذرای سیگنال در طی پردازش فرکانسی آسیب ببینند و از طرف دیگر به علت غیر ایستادن بودن سیگنال در پنجره های بزرگ دقت آنالیز زمان فرکانس کم شود. به علاوه با میانگین گیری از نمونه های دور گذشته در نتایج شنوائی اثر آکو نیز ظاهر می گردد. کوچک کردن پنجره آنالیز نیز موجب کم شدن جزئیات فرکانسی سیگنال شده و موجب اعوجاج در سیگنال بهبود یافته می گردد. تبدیل موجک با فراهم کردن مقیاس زمان - فرکانس دلخواه این عیب را برطرف نموده است.

با استفاده از تبدیل موجک می توان زیرباندها را بطور دلخواه روی بازه فرکانس توزیع کرد و توزیع زیرباندها را به سیستم شنوائی انسان نزدیک نمود [۲۱، ۲۲]. بنابراین حوزه موجک برای پردازش گفتار مناسب می باشد [۲۳]. در ضمن می دانیم که انواع نویزها توزیع مختلفی روی زیرباندهای فرکانسی دارند، از همین رو می توان توزیع زیرباندها را نیز بر اساس نوع نویز تعیین کرد. تبدیل موجک می تواند از توابع پایه مختلف بهره بگیرد و تابعی را که دارای عملکرد بهتری است، بکار برد.

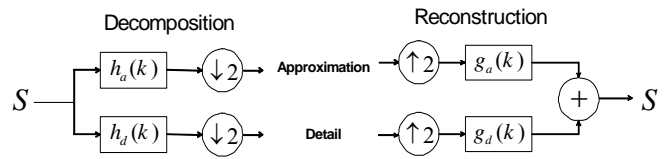
از مزایای دیگر موجک نسبت به فوریه بار محاسباتی کمتر آن می باشد [۷، ۲۴]. می دانیم که در سیگنال گفتار فاز از اهمیت خاصی برخوردار است، در تبدیل موجک فاز سیگنال جدا نمی گردد، بنابراین فاز نیز در پروسه بهبود سیگنال گفتار بکار برده می شود. در صورتی که در اغلب روش های حذف نویز در حوزه فرکانس فرض می کنند فاز سیگنال بهبود یافته با فاز سیگنال نویزی برابر است. با توجه به این ویژگی ها در این مقاله به روش های حذف نویز مبتنی بر موجک پرداخته شده است. بطور کلی می توان گفت، حذف نویز از ضرایب موجک به دو روش آستانه گذاری یا اعمال فیلتر نظیر فیلتر وینر در زیرباندها انجام می گیرد که هر یک از دو روش در ادامه شرح داده شده است.

۳-۱- گرایش فیلتر کردن ضرایب موجک - فیلتر وینر

برای محاسبه فیلتر وینر در حوزه موجک به صورت زیر عمل می شود. فرض کنید سیگنال گفتار تمیز S با نویز جمع شونده n آغشته شده باشد. در این صورت سیگنال نویزی x در حوزه زمان از رابطه (۱) تبعیت می کند.

$$x(i, m) = s(i, m) + n(i, m), \quad i = 0, 1, \dots, L-1 \quad (1)$$

می روند. خروجی فیلتر پائین گذر ضرایب تقریب و خروجی فیلتر بالا گذر ضرایب تفصیل را تشکیل می دهند.

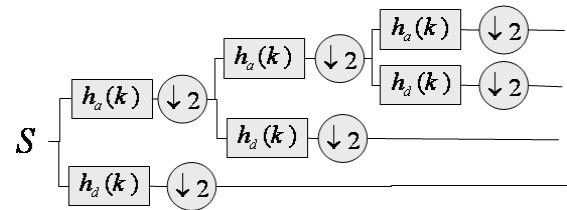


شکل ۱- ساختار تبدیل و عکس تبدیل موجک

مجموعه موجکها به عنوان یک بانک فیلتر برای آنالیز سیگنال می تواند در نظر گرفته شود. در این راستا، بانک فیلتر موجک را به صورت درخت نیز نمایش می دهند. می توان تبدیل موجک را بفرم درخت موجک گسسته دوتائی و درخت موجک بسته ای^۱ در نظر گرفت.

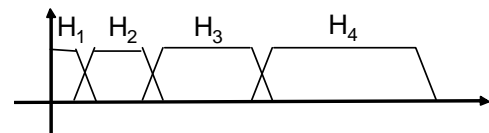
۲-۱- درخت موجک گسسته دوتائی و موجک بسته ای

شکل ۲ ساختار یک درخت بانک فیلتر را نشان می دهد که هم ارز با تبدیل موجک گسسته دوتائی روی یک دنباله می باشد. در این نوع درخت گسترش درخت از طرف زیرباندهای پائین یعنی باندهای فرکانس پائین می باشد. بنابراین تحلیل فرکانسی دقیق تری در این زیرباندها داریم.

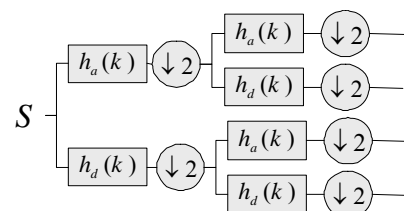


شکل ۲- درخت موجک گسسته دوتائی با ۴ زیرباند نامساوی

در مرحله اول طیف سیگنال ورودی به دو زیرباند بالائی و زیرباند پائینی تفکیک شده، سپس زیرباند پائینی دوباره به دو زیرباند پائینی و بالائی تفکیک می شود و این عمل برای زیرباند پائینی حاصل تکرار شده تا در مجموع ۴ زیرباند نهائی ایجاد شوند. شکل ۳ مشخصه طیفی فیلتربانک حاصل را نشان می دهد.



شکل ۳- مشخصه طیفی فیلتربانک موجک نشان داده شده در شکل ۲



شکل ۴- درخت موجک بسته ای با ۴ زیرباند یکنواخت

دیدگاه دیگر برای بدست آوردن فیلتر این است که برای هر نمونه از یک زیرباند از یک فریم، یک فیلتر محاسبه شود. که آنرا می‌توان در رابطه (۸) نشان داد:

$$\hat{S}_j(i, m) = H_j(i, m) \cdot X_j(i, m) \quad (8)$$

که با توجه به این امر فیلتر وینر را مجدداً بفرم رابطه (۹) تعریف می‌کنیم:

$$H_j(i, m) = \frac{R_{pr,j}(i, m)}{R_{pr,j}(i, m) + 1} \quad (9)$$

$R_{pr,j}(i, m)$ را بفرم رابطه (۱۰) که یک رابطه بازگشتی می‌باشد، تعریف می‌کنیم. $R_{pr,j}(i, m)$ در واقع نسبت واریانس سیگنال تمیز به واریانس سیگنال نویز در طی فریم‌های متوالی می‌باشد.

$$R_{pr,j}(i, m) = \alpha R_{pr,j}(i, m-1) + (1-\alpha) \cdot \max\left\{\frac{\hat{S}_j^2(i, m)}{\hat{N}_j^2(i, m)}, 0\right\} \quad (10)$$

برای تخمین $S_j^2(i, m)$ از تفاضل سیگنال نویزی و سیگنال نویز تخمین زده شده استفاده می‌شود. در این رابطه α فاکتور فراموشی می‌باشد که بر اساس روش واریانس بالانس شده مطرح شده در [۲۵] بدست می‌آید.

۳-۲- گرایش آستانه‌گذاری روی ضرایب موجک -

آستانه‌گذاری نرم

ایده اصلی آستانه‌گذاری روی ضرایب تبدیل موجک بر اساس این واقعیت است که در سیگنال گفتار، انرژی سیگنال غالباً در تعداد کمی از ضرایب تبدیل موجک متمرکز است. این ضرایب نسبت به ضرایب دیگر تبدیل موجک خود سیگنال، یا ضرایب تبدیل موجک هر سیگنال دیگری از جمله نویز که انرژی آن بر روی تعداد زیادی از ضرایب گسترش می‌یابد، بزرگتر هستند. بنابراین می‌توان با صفر کردن ضرایب کوچکتر، علاوه بر حفظ اطلاعات مهم سیگنال اصلی، نویز را نیز محدود کرد. مطابق با این ویژگی، ضرایب تبدیل موجک سیگنال با یک حد آستانه مقایسه می‌شوند و مقدار ضرایب کوچکتر از این حد آستانه برابر صفر قرار داده می‌شوند. از روش‌های آستانه‌گذاری می‌توان آستانه‌گذاری نرم [۸، ۲۶]، آستانه‌گذاری سخت [۸]، آستانه‌گذاری Quantile، آستانه‌گذاری نیمه نرم [۲۷]، آستانه‌گذاری با روش وفقی BayesShrink [۲۸، ۲۹] و آستانه‌گذاری سخت اصلاح شده [۲۹] را نام برد. در زیر به بررسی روش آستانه‌گذاری نرم می‌پردازیم که با رابطه (۱۱) نشان داده شده است.

$$\hat{S}_j(i, m) = \begin{cases} 0 & \text{if } |X_j(i, m)| < T_j(m) \\ \text{sgn}(X_j(i, m))(|X_j(i, m) - T_j(m)|) & \text{otherwise} \end{cases} \quad (11)$$

که در آن $X_j(i, m)$ نمونه \hat{m} از تبدیل موجک سیگنال گفتار نویزی در زیرباند \hat{m} از فریم \hat{m} و $\hat{S}_j(i, m)$ نمونه \hat{m} از تبدیل موجک سیگنال بهبودیافته در زیرباند \hat{m} از فریم \hat{m} است. همچنین $T_j(m)$ حد آستانه زیرباند \hat{m} از فریم \hat{m} می‌باشد، در ادامه جزئیات نحوه محاسبه انواع آستانه آمده است.

که در آن i اندیس زمان، L طول فریم و m اندیس فریم می‌باشد. در این صورت اگر بر رابطه (۱) تبدیل موجک را اعمال نمائیم، خواهیم داشت:

$$X_j(i, m) = S_j(i, m) + N_j(i, m), \quad i = 0, 1, \dots, L_j - 1 \quad (2)$$

که در آن $X_j(i, m)$ نمونه \hat{m} از تبدیل موجک سیگنال گفتار نویزی در زیرباند \hat{m} از فریم \hat{m} ، و $S_j(i, m)$ و $N_j(i, m)$ نمونه متناسب برای سیگنال تمیز و سیگنال نویز می‌باشند. و L_j طول زیرباند \hat{m} می‌باشد. می‌توان مطابق رابطه (۳) برای بدست آوردن سیگنال بهبود یافته سیگنال نویزی را فیلتر نمود، در این رابطه برای هر زیرباند از یک فریم یک فیلتر بدست می‌آید:

$$\hat{S}_j(i, m) = H_j(m) \cdot X_j(i, m) \quad (3)$$

در صورتی که از فیلتر وینر برای پیاده‌سازی رابطه بالا استفاده کنیم تابع تبدیل آن بصورت رابطه (۴) تعریف می‌شود:

$$H_j(m) = \frac{E\{S_j^2(i, m)\}}{E\{S_j^2(i, m)\} + \text{est}\{N_j^2(m)\}} \quad (4)$$

که در آن $H_j(m)$ اندازه فیلتر برای زیرباند \hat{m} فریم \hat{m} می‌باشد که در رابطه (۳) بکار می‌رود. $E\{S_j^2(i, m)\}$ انرژی زیرباند \hat{m} فریم \hat{m} سیگنال تمیز و $\text{est}\{N_j^2(m)\}$ تخمین انرژی نویز در زیرباند \hat{m} فریم \hat{m} می‌باشد که از رابطه (۵) بدست می‌آید.

$$\text{est}\{N_j^2(m)\} = \alpha \cdot \text{est}\{N_j^2(m-1)\} + (1-\alpha) \cdot E\{\hat{N}_j^2(i, m)\} \quad (5)$$

که $E\{\hat{N}_j^2(i, m)\}$ انرژی زیرباند \hat{m} فریم \hat{m} نویز تخمین زده شده، $\text{est}\{N_j^2(m-1)\}$ تخمین انرژی نویز در زیرباند \hat{m} فریم $(m-1)$ و α فاکتور فراموشی می‌باشد که بر اساس روش واریانس بالانس شده مطرح شده در [۲۵] بدست می‌آید. در این روش اندازه ضریب فراموشی به طول فریم بستگی دارد به این صورت که α برای فریم‌های بزرگ به یک نزدیک شده و برای فریم‌های کوچک به صفر نزدیک می‌شود. رابطه محاسبه α به در رابطه (۶) آمده است:

$$\alpha = \frac{T_{sm} \times f_s - L/2}{T_{sm} \times f_s + L/2} \quad (6)$$

که L طول فریم، f_s فرکانس نمونه‌برداری سیگنال ورودی و T_{sm} زمان تثبیت شدن تخمین توان سیگنال می‌باشد که معمولاً برابر 0.2 ثانیه در نظر گرفته می‌شود [۲۵]. $E\{S_j^2(i, m)\}$ انرژی سیگنال تمیز در زیرباند \hat{m} فریم \hat{m} از رابطه (۷) بدست می‌آید.

در حالت ایده‌آل برای تخمین‌ها از سیگنال‌های نویز و تمیزی که در دسترس هستند استفاده می‌شود. ولی در شرایط واقعی که فقط سیگنال نویزی در اختیار است، از مناطقی از سیگنال که فقط نویز وجود دارد برای به روز کردن تخمین انرژی نویز استفاده می‌شود و تا یافتن بخش دیگری از سیگنال که فقط نویز در آن حاضر باشد، از تخمین قبلی نویز استفاده می‌شود.

$$E\{S_j^2(i, m)\} = E\{X_j^2(i, m)\} - \text{est}\{N_j^2(m)\} \quad (7)$$

۳-۲-۱- انتخاب آستانه

روش‌های مختلفی را می‌توان برای انتخاب حد آستانه در نظر گرفت، اما در اغلب کاربردها، روش پیشنهاد شده توسط Donoho [۸] مورد استفاده قرار می‌گیرد که به فرم رابطه (۱۲) محاسبه می‌شود.

$$T = \hat{\sigma}_n \sqrt{2 \log(N)} \quad (12)$$

که N ، طول فریم جاری و $\hat{\sigma}_n$ تخمینی از انحراف معیار کلی نویز است. در برخی از منابع این روش تعیین حد آستانه، روش VisuShrink نامیده می‌شود. از آنجا که مقدار واقعی واریانس نویز σ_n^2 در دسترس نیست. در مورد نویز سفید گوسی می‌توان انحراف معیار $\hat{\sigma}_n$ را با توجه به رابطه (۱۳) از روی ضرایب موجک تخمین زد.

$$\hat{\sigma}_n = \frac{MAD_1}{0.6745} = \frac{\text{median}(|W_1^D| - \text{median}(W_1^D))}{0.6745} \quad (13)$$

در این رابطه MAD بیانگر Median Absolute Difference ضرایب موجک تفصیل مربوط به سطح یک است. اگر از تبدیل بسته موجک برای بدست آوردن ضرایب موجک استفاده کرده باشیم، مقدار حد آستانه از رابطه (۱۴) محاسبه می‌شود.

$$T = \hat{\sigma}_n \sqrt{2 \log(N \log_2 N)} \quad (14)$$

روش آستانه‌بندی پیشنهاد شده توسط Donoho را می‌توان توسعه داد و حد آستانه را با توجه به رابطه (۱۵) به طور مجزا برای هر سطح در تبدیل بسته موجک محاسبه کرد [۳۰].

$$T_j = \hat{\sigma}_{n_j} \sqrt{2 \log(N)} \quad (15)$$

انحراف معیار نویز برای هر سطح نیز از رابطه (۱۶) تخمین زده می‌شود.

$$\hat{\sigma}_{n_j} = \frac{MAD_j}{0.6745} \quad (16)$$

که در آن مقدار MAD برای هر سطح از روی ضرایب موجک تفصیل همان سطح محاسبه می‌شود.

$$MAD_j = \text{median}(|W_j^D| - \text{median}(W_j^D)) \quad (17)$$

این روش آستانه‌بندی در منابع مختلف با عنوان روش SureShrink شناخته می‌شود.

در این بخش دو دیدگاه کلی برای حذف نویز در حوزه موجک بررسی شدند. برای هر دیدگاه روش‌های متعددی در مقالات ارائه شده است که در اینجا دو روش مرسوم تر در حوزه موجک مطرح شدند، که در بخش پنجم مقاله به مقایسه این دو روش می‌پردازیم.

۴- معرفی فیلترهای مبتنی بر توزیع احتمال پسین ضرائب موجک

روش حذف نویز ارائه شده توسط Ephraim در حوزه فرکانس [۱۸]، علیرغم بهینه بودن از نقطه نظر کمینه کردن خطای میانگین مربعات (MMSE) روی دامنه سیگنال در حوزه فرکانس، نیازمند محاسبات نمایی و توابع بسل می‌باشد. از همین رو در سال‌های اخیر فیلترهای متعددی که دارای عملکرد مشابه و محاسبات کمتر بودند ارائه شده‌اند از جمله می‌توان به فیلترهای Joint-MAP، MAP-SAE⁺ و MMSE اشاره کرد که همگی در حوزه فرکانس می‌باشند [۱۹]. با توجه به ویژگی‌های مطرح شده برای تبدیل موجک در بخش سوم مقاله که نشان‌دهنده کارآمد بودن موجک برای پردازش گفتار می‌باشد، بر آن شدیم که از این تبدیل بجای تبدیل فوری استفاده کنیم. در اینجا با الهام از مرجع [۱۸] و نیز دو معیار تخمین MAP-G و MAP-SG دو فیلتر حذف نویز در حوزه موجک پیشنهاد می‌نمائیم. طریقه محاسبه فیلترهای پیشنهادی در ادامه توضیح داده می‌شوند. به علاوه در بخش‌های بعدی این فیلترها پیاده‌سازی و ارزیابی خواهند شد.

۴-۱- پیشنهاد معیار MAP-G برای محاسبه فیلتر

معیار MAP-G توزیع احتمال پسین سیگنال تمیز را بشرط داشتن سیگنال نویزی ماکزیمم می‌کند. اگر سیگنال گفتار و نویز دارای توزیع تصادفی گوسی با میانگین صفر و واریانس σ_s^2 و σ_n^2 باشند، می‌توان احتمال سیگنال تمیز را بشرط داشتن سیگنال نویزی بفرم رابطه (۱۸) نوشت. در [۳۱] با استفاده از معیار Kurtosis نشان داده شده است که ضرائب موجک در زیرباند‌های درخت تجزیه موجک بسته‌ای دارای توزیع گوسی با واریانس کم می‌باشند.

$$\begin{aligned} p(S_j(i, m) | X_j(i, m)) & \quad (18) \\ & \propto p(X_j(i, m) | S_j(i, m)) p(S_j(i, m)) \\ & \propto \frac{S_j(i, m)}{\pi^2 \sigma_{S_j(i)}^2 \sigma_{N_j(i)}^2} \exp\left(-\frac{|X_j(i, m) - S_j(i, m)|^2}{\sigma_{N_j(i)}^2} - \frac{S_j^2(i, m)}{\sigma_{S_j(i)}^2}\right) \end{aligned}$$

که $\sigma_{S_j(i)}^2 = E[S_j^2(i)]$ و $\sigma_{N_j(i)}^2 = E[\hat{N}_j^2(i)]$ به ترتیب واریانس سیگنال نویز و واریانس سیگنال تمیز برای نمونه i ام از زیرباند j ام در طول فریم‌های متوالی می‌باشند. از رابطه فوق لگاریتم طبیعی می‌گیریم، و آنرا J می‌نامیم، بنابراین:

$$J = -\frac{|X_j(i, m) - S_j(i, m)|^2}{\sigma_{N_j(i)}^2} - \frac{S_j^2(i, m)}{\sigma_{S_j(i)}^2} + \ln(S_j(i, m)) + cte \quad (19)$$

که در آن cte یک مقدار ثابت می‌باشد. اگر از رابطه فوق نسبت به $S_j(i, m)$ مشتق می‌گیریم، خواهیم داشت:

$$\frac{\partial}{\partial S_j(i, m)} J = -\frac{2(S_j(i, m) - X_j(i, m))}{\sigma_{N_j(i)}^2} - \frac{2S_j(i, m)}{\sigma_{S_j(i)}^2} + \frac{1}{S_j(i, m)} \quad (20)$$

با مساوی صفر قرار دادن رابطه (۲۰) خواهیم داشت:

که σ_S^2 و σ_N^2 به ترتیب واریانس سیگنال نویز و واریانس سیگنال تمیز برای نمونه نام از زیرباند نام در طول فریم‌های متوالی می‌باشند. ضرایب موجک سیگنال تمیز دارای توزیع super Gaussian بوده که $p(S_j(i, m))$ بفرم زیر خواهد بود:

$$p(S_j(i, m)) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{S_j^\nu(i, m)}{\sigma_{S_j(i)}^{\nu+1}} \exp\left(-\mu \frac{S_j(i, m)}{\sigma_{S_j(i)}}\right) \quad (27)$$

که Γ تابع گاما و نیز μ و ν پارامترهای توزیع هستند. با جایگذاری روابط (۲۶) و (۲۷) در رابطه (۲۵) داریم:

$$p(S_j(i, m)|X_j(i, m)) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{S_j^\nu(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2 \sigma_{S_j(i)}^{\nu+1}} \times \exp\left(-\frac{X_j^2(i, m) + S_j^2(i, m)}{\sigma_{N_j(i)}^2} - \mu \frac{S_j(i, m)}{\sigma_{S_j(i)}}\right) \times I_0\left(\frac{2S_j(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2}\right) \quad (28)$$

از سوئی رابطه بسل را می‌توان بفرم رابطه (۲۹) نیز حساب کرد:

$$I_0(|x|) \approx \frac{1}{\sqrt{2\pi}|x|} \exp(|x|) \quad (29)$$

اگر رابطه (۲۹) را در رابطه (۲۸) جایگزین کنیم، داریم:

$$p(S_j(i, m)|X_j(i, m)) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{S_j^{\nu-\frac{1}{2}}(i, m) X_j^{\frac{1}{2}}(i, m)}{2\sqrt{\pi} \sigma_{N_j(i)} \sigma_{S_j(i)}^{\nu+1}} \times \exp\left(-\frac{X_j^2(i, m) + S_j^2(i, m)}{\sigma_{N_j(i)}^2} - \mu \frac{S_j(i, m)}{\sigma_{S_j(i)}} + \frac{2S_j(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2}\right) = \frac{\mu^{\nu+1}}{\Gamma(\nu+1)} \frac{X_j^{\frac{1}{2}}(i, m)}{2\sqrt{\pi} \sigma_{N_j(i)} \sigma_{S_j(i)}^{\nu+1}} \exp\left(-\frac{X_j^2(i, m)}{\sigma_{N_j(i)}^2}\right) S_j^{\nu-\frac{1}{2}}(i, m) \times \exp\left(-\frac{S_j^2(i, m)}{\sigma_{N_j(i)}^2} - S_j(i, m) \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2}\right)\right) \approx S_j^{\nu-\frac{1}{2}}(i, m) \exp\left(-\frac{S_j^2(i, m)}{\sigma_{N_j(i)}^2} - S_j(i, m) \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2}\right)\right) \quad (30)$$

از رابطه فوق لگاریتم طبیعی می‌گیریم، و آنرا J می‌نامیم.

$$J = -\frac{S_j^2(i, m)}{\sigma_{N_j(i)}^2} \quad (31)$$

$$S_j(i, m) \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2}\right) + \left(\nu - \frac{1}{2}\right) \ln(S_j(i, m))$$

از رابطه بالا نسبت به $S_j(i, m)$ مشتق می‌گیریم، خواهیم داشت:

$$-2\left(\sigma_{S_j(i)}^2 + \sigma_{N_j(i)}^2\right) S_j^2(i, m) + 2\sigma_{S_j(i)}^2 X_j(i, m) S_j(i, m) + \sigma_{S_j(i)}^2 \cdot \sigma_{N_j(i)}^2 = 0 \quad (21)$$

معادله فوق را بر اساس متغیر $S_j(i, m)$ حل می‌کنیم، برای سادگی روابط بدست آمده، دو مقدار $R_{pr,j}(i, m)$ تعریف شده در رابطه (۱۰) و $R_{po,j}(i, m)$ بفرم رابطه (۲۲) را تعریف می‌کنیم. $R_{po,j}(i, m)$ در واقع نسبت سیگنال به نویز در فریم جاری می‌باشد.

$$R_{po,j}(i, m) = \frac{\hat{S}_j^2(i, m)}{\hat{N}_j^2(i, m)} \quad (22)$$

با این فرض و با حل معادله درجه دوم (۲۱) خواهیم داشت:

$$S_j(i, m) = \left\{ \frac{R_{pr,j}(i, m)}{2(1 + R_{pr,j}(i, m))} + \frac{\sqrt{R_{pr,j}^2(i, m) + 2(1 + R_{pr,j}(i, m)) \frac{R_{pr,j}(i, m)}{(R_{po,j}(i, m) + 1)}}}{2(1 + R_{pr,j}(i, m))} \right\} X_j(i, m) \quad (23)$$

بنابراین فیلتر بفرم رابطه (۲۴) در خواهد آمد:

$$H_j(i, m) = \frac{R_{pr,j}(i, m)}{2(1 + R_{pr,j}(i, m))} + \frac{\sqrt{R_{pr,j}^2(i, m) + 2(1 + R_{pr,j}(i, m)) \frac{R_{pr,j}(i, m)}{(R_{po,j}(i, m) + 1)}}}{2(1 + R_{pr,j}(i, m))} \quad (24)$$

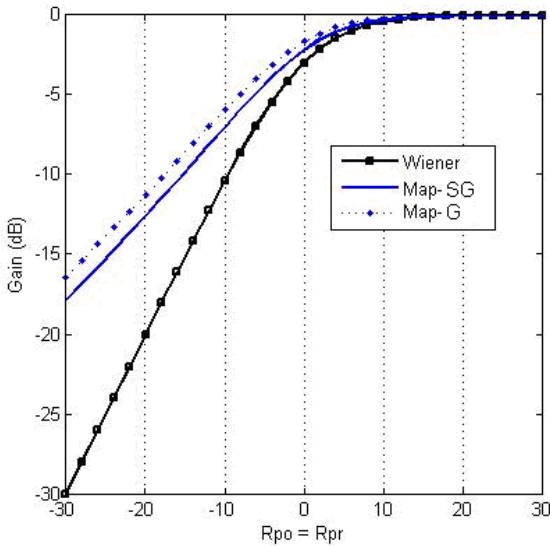
۴-۲- پیشنهاد معیار MAP-SG برای محاسبه فیلتر

در این قسمت نیز با معیار MAP-SG سعی می‌شود تا احتمال تخمین دامنه سیگنال بهبودیافته را با توجه به داشتن دامنه سیگنال نویزی ماکزیمم کنیم که مشابه با معیار MAP-G عمل می‌شود. در این حالت فرض می‌کنیم که ضرایب موجک سیگنال تمیز دارای توزیع super Gaussian بوده و احتمال ضرایب موجک سیگنال نویزی به شرط ضرایب موجک سیگنال تمیز دارای توزیع Rice باشد. با این فرض احتمال $p(S_j(i, m)|X_j(i, m))$ بفرم زیر تعریف می‌شود:

$$p(S_j(i, m)|X_j(i, m)) \propto p(X_j(i, m)|S_j(i, m)) p(S_j(i, m)) \quad (25)$$

ضرایب موجک سیگنال نویزی به شرط ضرایب موجک سیگنال تمیز دارای توزیع Rice می‌باشد که می‌توان در حوزه موجک $p(X_j(i, m)|S_j(i, m))$ بفرم رابطه (۲۶) نوشت: [۳۵]

$$p(X_j(i, m)|S_j(i, m)) = \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2} \times \exp\left(-\frac{X_j^2(i, m) + S_j^2(i, m)}{\sigma_{N_j(i)}^2}\right) I_0\left(\frac{2S_j(i, m) X_j(i, m)}{\sigma_{N_j(i)}^2}\right) \quad (26)$$

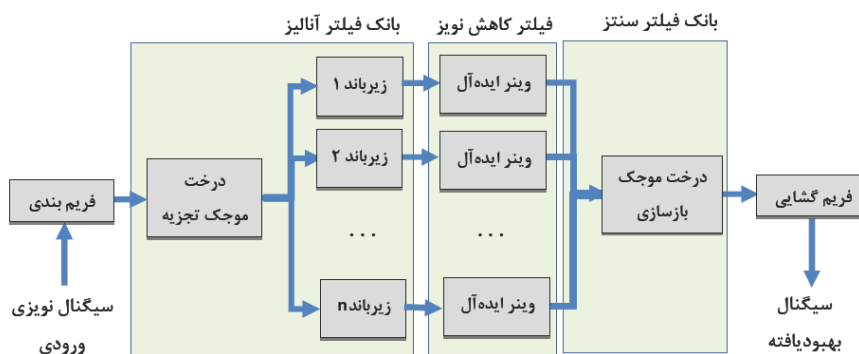


شکل ۵- اندازه سه فیلتر وینر، MAP-G و MAP-SG در حالت بهینه برای مقادیر مختلف R_{pr}

شکل ۵ اندازه سه فیلتر وینر، MAP-G و MAP-SG در حالت بهینه برای مقادیر مختلف R_{pr} را نشان می‌دهد. در این نمودار محور افقی اندازه R_{pr} را بر حسب دسی بل نشان می‌دهد که از -30dB تا +30dB تغییر می‌کند. محور عمودی نیز اندازه فیلتر را بر حسب دسی بل نشان می‌دهد. چنانچه در شکل ۵ دیده می‌شود، در سیگنال به نویزهای پائین، فیلتر وینر نویز بیشتری را نسبت به دو فیلتر دیگر حذف می‌کند. بنابراین انتظار می‌رود که تخریب بیشتری نیز صورت بگیرد. ولی برای نرخ سیگنال به نویز بیشتر از صفر اندازه این سه فیلتر بهم نزدیک می‌شود که انتظار عملکرد یکسانی از سه فیلتر می‌رود. در قسمت ششم مقاله به بررسی این سه فیلتر در دو حوزه فرکانس و موجک پرداخته می‌شود.

۵- مقایسه گرایش‌های حذف نویز از ضرائب موجک معرفی شده در بخش ۳

بمنظور ارزیابی و مقایسه دو روش آستانه‌گذاری و فیلترینگ، روش‌های آستانه‌گذاری نرم و فیلتر وینر با یکدیگر مقایسه شده‌اند. برای فیلتر وینر رابطه (۴) و برای روش آستانه‌گذاری نرم رابطه (۱۲) و تعیین حد آستانه SureShrink یعنی رابطه (۱۶) در شرایط ایده‌آل بکار رفته‌اند، به این صورت که برای بدست آوردن انرژی و انحراف معیار نویز یک زیرباند از خود سیگنال نویز استفاده شده است.



شکل ۶- بلوک دیاگرام سیستم کاهش نویز

$$\frac{\partial}{\partial S_j(i, m)} J = -\frac{2S_j(i, m)}{\sigma_{N_j(i)}^2} - \left(\frac{\mu}{\sigma_{S_j(i)}} - \frac{2X_j(i, m)}{\sigma_{N_j(i)}^2} \right) + \left(\nu - \frac{1}{2} \right) \frac{1}{S_j(i, m)} \quad (32)$$

با مساوی صفر قرار دادن معادله بالا و حل این معادله درجه دوم بر حسب متغیر $S_j(i, m)$ رابطه (۳۳) حاصل می‌شود.

$$S_j(i, m) = \frac{\mu\sigma_{N_j(i)}^2 - X_j(i, m)}{4\sigma_{S_j(i)}} + \sqrt{\left(-\frac{\mu\sigma_{N_j(i)}^2 - X_j(i, m)}{4\sigma_{S_j(i)}} + \frac{X_j(i, m)}{2} \right)^2 - \frac{\sigma_{N_j(i)}^2}{2} \left(\nu - \frac{1}{2} \right)} \quad (33)$$

با ساده‌سازی به رابطه زیر می‌رسیم:

$$S_j(i, m) = \left\{ \frac{\mu\sigma_{N_j(i)}^2}{4\sigma_{S_j(i)}X_j(i, m)} - \frac{1}{2} + \sqrt{\left(-\frac{\mu\sigma_{N_j(i)}^2}{4\sigma_{S_j(i)}X_j(i, m)} + \frac{1}{2} \right)^2 - \frac{\sigma_{N_j(i)}^2}{2X_j^2(i, m)} \left(\nu - \frac{1}{2} \right)} \right\} X_j(i, m) \quad (34)$$

برای سادگی روابط بدست آمده، دو مقدار $R_{pr,j}(i, m)$ و $R_{po,j}(i, m)$ تعریف شده در روابط (۱۰) و (۲۲) بکار می‌بریم. بنابراین فیلتر بفرم زیر در خواهد آمد:

$$H_j(i, m) = \frac{\mu}{4R_{pr,j}(i, m)R_{po,j}(i, m)} - \frac{1}{2} + \sqrt{\left(-\frac{\mu}{4R_{pr,j}(i, m)R_{po,j}(i, m)} + \frac{1}{2} \right)^2 - \frac{1}{2R_{po,j}^2(i, m)} \left(\nu - \frac{1}{2} \right)} \quad (35)$$

در این مقاله مقادیر دو پارامتر μ و ν ، با توجه به [36] بترتیب ۱.۵ و ۰.۰۱ در نظر گرفته شده‌اند. دو فیلتر پیشنهادی با معیارهای MAP-G و MAP-SG به میزان تخمین‌های R_{pr} و R_{po} وابسته می‌باشند. در حالت بهینه دو تخمین R_{pr} و R_{po} برابرند. شکل ۵ اندازه سه فیلتر وینر با رابطه (۹)، MAP-G با رابطه (۲۴) و MAP-SG با رابطه (۳۵) را در حالت بهینه برای مقادیر مختلف R_{pr} نشان می‌دهد.

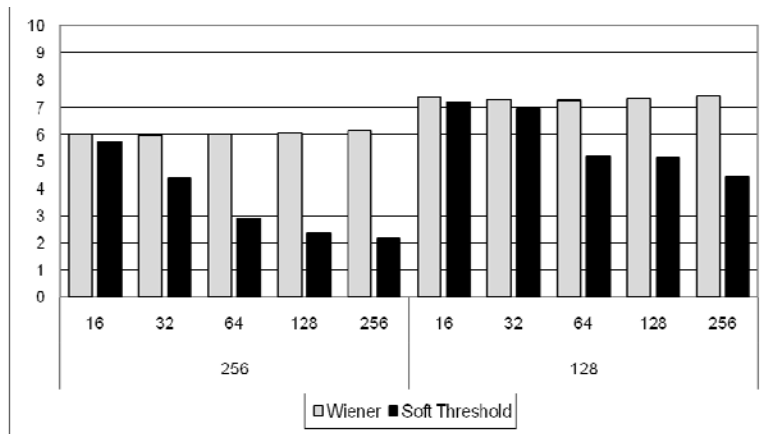
۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 0dB و 20dB نشان می‌دهد. نتایج نظری نشان می‌دهد که مقدار نسبت سیگنال به نویز بهبودیافته برای روش فیلتر کردن وینر بیشتر از روش آستانه‌گذاری نرم می‌باشد. تفاوت سطح نسبت سیگنال به نویز بهبودیافته دو روش برای نرخ سیگنال به نویز ورودی 0dB بیشتر از تفاوت سطح نسبت سیگنال به نویز بهبودیافته برای نرخ سیگنال به نویز ورودی 10dB می‌باشد، چراکه در نویز بیشتر عملکرد فیلتر وینر بهتر بوده و همه نویز را حذف می‌کند ولی در روش آستانه‌گذاری نویز زیادی باقی مانده و از طرف دیگر اعوجاج بیشتری در سیگنال مفید بوجود آمده است.

چنانچه در شکل ۷ دیده می‌شود، با افزایش تعداد زیرباند مقدار نسبت سیگنال به نویز برای روش آستانه‌گذاری کاهش می‌یابد، در واقع آستانه‌گذاری در تعداد زیرباندهای کمتر بهتر عمل می‌کند. ولی با افزایش تعداد زیرباند مقدار نسبت سیگنال به نویز برای روش وینر افزایش می‌یابد، زیرا جزئیات فرکانسی افزایش می‌یابد، بنابراین تخمین دقیقتری از فیلتر داشته و انتظار عملکرد بهتری برای وینر در زیرباندهای بیشتر داریم. از سویی با افزایش طول فریم از ۱۲۸ نمونه به ۲۵۶ نمونه مقدار نسبت سیگنال به نویز بهبودیافته برای روش وینر اندکی کاهش می‌یابد، چرا که با افزایش طول فریم رزلوشن زمانی سیگنال کاهش یافته است در حالیکه تعداد زیرباندها تغییری نکرده است و لذا دقت تخمین پارامترهای فیلتر وینر کم می‌شود.

بلوک دیاگرام سیستم کاهش نویز فیلتر کردن وینر مورد استفاده در این ارزیابی در شکل ۶ نشان داده شده است. همانطور که در شکل مشاهده می‌کنید، ابتدا سیگنال ورودی به فریم‌هایی با طول مشخص تقسیم می‌شود. در ادامه فریم توسط فیلتربانک موجک به N زیرباند تقسیم می‌گردد و روی هر زیرباند یک فیلتر وینر ایده آل اعمال می‌شود. پس از آن با استفاده از عکس تبدیل موجک فریم بهبود یافته بدست می‌آید و در نهایت فریم‌ها در کنار هم قرار می‌گیرند تا سیگنال بهبود یافته نهایی بدست آید.

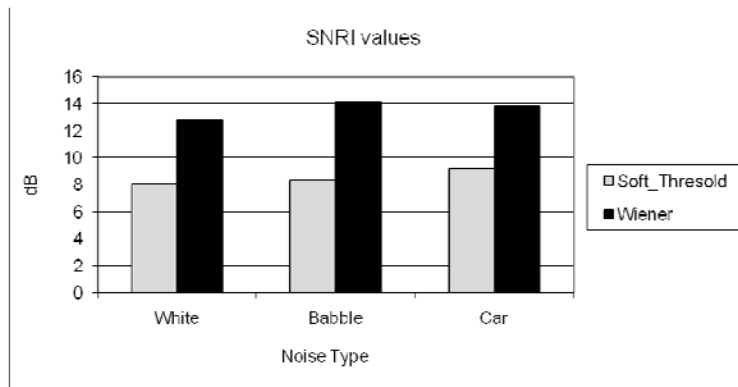
تخمین مؤلفه‌های فیلتر وینر ایده‌آل از روی دو سیگنال نویز و سیگنال گفتار انجام می‌گیرد. مقادیر برای پارامترهای سیستم با توجه به نتایج گزارش شده در مقاله [۳۷] انتخاب شده‌اند. که با توجه به مرجع [۳۸، ۳۷] از موجک بسته‌ای برای تجزیه سیگنال به زیرباندها استفاده شد، که در آن زیرباندها یکنواخت می‌باشند. همچنین طول فریم از ۱۲۸ تا ۱۰۵۴ و تعداد زیرباندها از ۱۶ تا ۲۵۶ در نظر گرفته شده است. تابع پایه موجک بکار رفته نیز $daub10^{11}$ می‌باشد. در پیاده‌سازی آستانه‌گذاری نرم برای آنکه تخمین بهتری از آستانه داشته باشیم، تخمین آستانه را روی فریم بزرگتری انجام می‌دهیم. این عمل به این خاطر انجام می‌گیرد که در واقع با فرض اولیه آستانه‌گذاری یعنی بینهایت بودن فریم تطابق بیشتری داشته باشد.

شکل ۷ مقادیر نسبت سیگنال به نویز بهبودیافته را برای دو روش آستانه‌گذاری نرم و وینر با طول فریم ۱۲۸ و ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴،



نرخ سیگنال به نویز ورودی 20dB

شکل ۷- مقادیر نسبت سیگنال به نویز بهبودیافته دو روش آستانه‌گذاری نرم و فیلتر وینر با طول فریم ۱۲۸ و ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴، ۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 0dB و 20dB



شکل ۸- مقادیر نسبت سیگنال به نویز بهبودیافته برای دو روش آستانه‌گذاری نرم و فیلتر وینر با طول فریم ۲۵۶ و نیز تعداد زیرباند ۶۴ در حضور نویزهای گوسی سفید، نویز همهمه و نویز ماشین با نرخ سیگنال به نویز ورودی 10dB

شکل ۱۰ نتایج تست شنوائی را برای دو روش آستانه‌گذاری نرم و فیلتر وینر زمانیکه طول فریم ۲۵۶ و تعداد زیرباند ۱۶ و ۲۵۶ نشان می‌دهد. نتایج شنوائی نشان می‌دهد که میزان تخریب گفتار و نیز میزان نویز باقیمانده در روش آستانه‌گذاری بسیار بیشتر از روش فیلتر وینر می‌باشد از طرفی روش فیلتر کردن دارای مقدار کمی نویز موزیکال می‌باشد ولی در مجموع تست شنوائی شنوندگان سیگنال‌های بهبودیافته بروش فیلتر کردن را به سیگنال‌های بهبودیافته بروش آستانه‌گذاری ترجیح داده‌اند. برای طول فریم‌های دیگر و تعداد زیرباندهای دیگر نیز نتایج مشابهی بدست آمده است که برای جلوگیری از طولانی شدن مقاله به ذکر همین نتایج بسنده می‌کنیم.

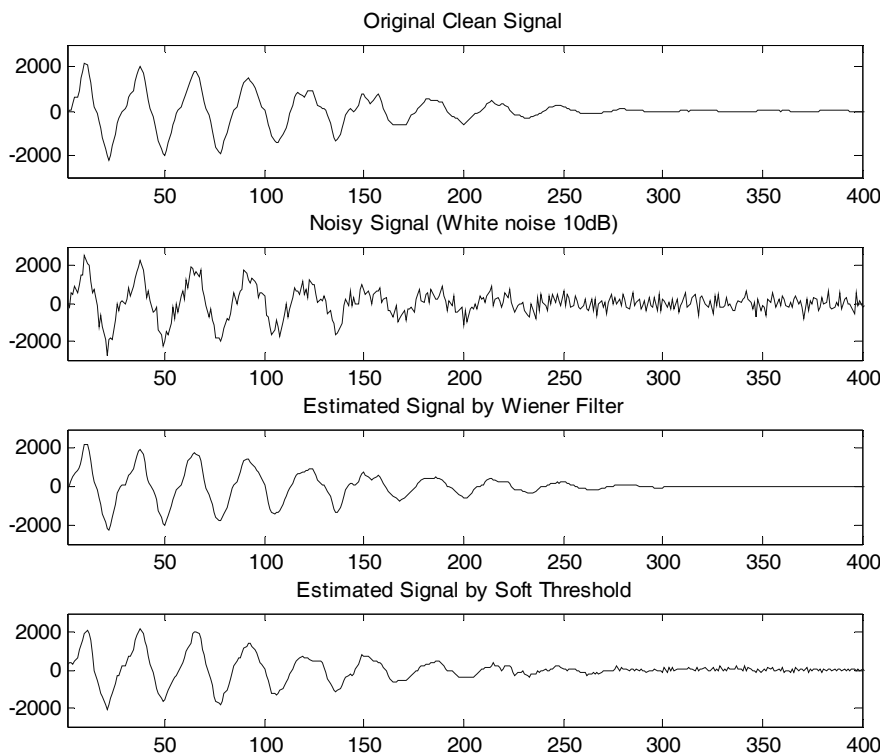
۶- مقایسه و ارزیابی فیلترهای پیشنهادی در بخش ۴

در این بخش به بررسی فیلترهای ارائه شده می‌پردازیم، سیستم حذف نویز مطابق با بلوک دیاگرام شکل ۶ در عمل پیاده‌سازی شده است. در این پیاده‌سازی بجای فیلتر وینر ایده‌آل سه فیلتر وینر غیر ایده‌آل، MAP-G و MAP-SG که روابط آنها بترتیب در (۱۱)، (۲۶) و (۳۷) آمده است، استفاده شده است. در واقع تخمین نویز از روی قسمتهایی از سیگنال نویزی که تنها شامل نویز می‌باشد صورت می‌گیرد. در این ارزیابی از موزیک بسته‌ای برای تجزیه سیگنال به زیرباندها استفاده شد. همچنین طول فریم ۲۵۶ نمونه و تعداد زیرباندها از ۱۶ تا ۲۵۶، تابع پایه موزیک نیز db10 در نظر گرفته شده است. جدول ۱ مقادیر نسبت سیگنال به نویز بهبودیافته را برای سه روش فیلتر کردن وینر، MAP-G و MAP-SG با طول فریم ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴، ۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB نشان می‌دهد.

شکل ۸ مقادیر نسبت سیگنال به نویز بهبودیافته برای دو روش آستانه‌گذاری نرم و فیلتر وینر با طول فریم ۲۵۶ و نیز تعداد زیرباند ۶۴ در حضور نویزهای گوسی سفید، نویز همپه و نویز ماشین با نرخ سیگنال به نویز ورودی 10dB نشان می‌دهد. چنانکه در شکل هم دیده می‌شود مقدار نسبت سیگنال به نویز بهبودیافته در حضور نویزهای مختلف برای روش وینر بیشتر از روش آستانه‌گذاری نرم می‌باشد. و روش فیلتر کردن نتایج بهتری نسبت به روش آستانه‌گذاری در حضور نویزهای مختلف دارد. این تفاوت برای نویز همپه که تا حد زیادی نالیستان است بیشتر مشخص می‌باشد. این امر کاملاً قابل پیش بینی بوده چون در شرایطی مانند همپه که نویز زمینه مشخصاتی شبیه به سیگنال اصلی دارد، توزیع سیگنال و نویز روی زیرباندهای مختلف شبیه به همدیگر خواهد بود و آستانه‌گذاری کارایی خود را تا حدی از دست می‌دهد، حال آنکه فیلتر بهینه با تخمین دقیق تر از نویز، کارا تر است.

در شکل ۹ قسمتی از سیگنال نویزی، سیگنال بهبود یافته بروش فیلتر کردن وینر و سیگنال بهبود یافته بروش آستانه‌گذاری نرم را نشان می‌دهد. سیگنال نویزی شامل نویز و سیگنال گفتار می‌باشد. چنانچه در شکل هم دیده می‌شود، نویز بیشتری در روش فیلتر کردن وینر قسمت اعظم نویز حذف می‌شود ولی در روش آستانه‌گذاری بدلیل غیر بهینه بودن روش آستانه‌گذاری مقداری از نویز باقی می‌ماند. همچنین روش فیلتر کردن در حذف نویز بهتر عمل کرده و سیگنال بهبودیافته به سیگنال تمیز شباهت بیشتری دارد.

تست شنوائی به این صورت انجام شده است که شنوندگان به سیگنال بهبود یافته گوش داده و به ۲ جنبه وجود نویز زمینه^{۱۲} و تخریب صدا^{۱۳} امتیازی بین ۱ تا ۳ داده‌اند. در این تست به سیگنال بهبود یافته ای که کیفیت خوبی دارند امتیاز یک، به سیگنالهای بهبودیافته ای که کیفیت پایین تری دارند ولی از نظر شنونده قابل قبول هستند، امتیاز دو و به سیگنالهایی که کیفیت نازلی دارند و از دید شنونده غیرقابل قبول هستند، امتیاز سه داده می‌شود.



شکل ۹- شکل موج ۴۰۰ نمونه از سیگنال نویزی، سیگنال بهبود یافته بروش فیلتر کردن وینر و سیگنال بهبود یافته بروش آستانه‌گذاری نرم با طول فریم ۲۵۶ و نیز تعداد زیرباند ۶۴ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

شنوایی نشان داد که کیفیت سیگنال بهبود یافته در تعداد زیرباندهای بیشتر، بهتر از کیفیت سیگنال بهبود یافته در تعداد زیرباندهای کمتر می‌باشد و میزان تخریب گفتار در آنها کمتر می‌باشد.

شکل ۱۱ اندازه سه فیلتر وینر، MAP-G و MAP-SG را برای یک فریم با ۶۴ زیرباند نشان می‌دهد. در این نمودار محور افقی فرکانس و محور عمودی اندازه فیلتر در مقیاس دسی بل را نشان می‌دهد. چنانچه که در نمودار دیده می‌شود اندازه فیلتر MAP-G بالاتر از دو فیلتر دیگر و اندازه فیلتر وینر پائینتر از دو فیلتر دیگر می‌باشد. از همینرو در جدول (۱) برای تعداد زیرباند ۶۴، مقدار نسبت سیگنال به نویز بهبودیافته برای روش فیلتر وینر بیشتر و برای فیلتر MAP-G کمتر می‌باشد.

در این مقاله علاوه بر تست شنوایی از مقدار MOS تخمینی با ابزار PESQ^۴ برای ارزیابی نتایج نیز استفاده شده است. قابل ذکر است این روش از طرف ITU به عنوان استاندارد ارزیابی کیفیت گفتار پیشنهاد شده است. این روش برای استفاده بر روی یک پهنه زیاد از شرایط شبکه شامل اتصالات آنالوگ، کدکننده‌ها، بسته‌های گم شده و پارامتر تأخیر تهیه شد. نویز زمینه و نویز پردازش نیز می‌تواند توسط معیار PESQ با استفاده از سیگنال تمیز بدست آید. روش PESQ به فایل بهبود یافته نمره‌ای بین صفر و ۴.۵ می‌دهد. که نمره صفر به سیگنال نویزی و تخریب شده و نمره ۴.۵ به سیگنال تمیز داده می‌شود. در این روش هر چه عدد بزرگتر باشد به معنی کیفیت بالاتر می‌باشد. [۳۲، ۳۳]

جدول ۱- مقادیر نسبت سیگنال به نویز بهبودیافته را برای سه روش فیلتر کردن وینر، MAP-G و MAP-SG با طول فریم ۲۵۶ و نیز تعداد زیرباند ۱۶، ۳۲، ۶۴، ۱۲۸ و ۲۵۶ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

Number of subbands	Methods		
	MAP-SG	MAP-G	Wiener
16	10.05	9.9	9.77
32	10.28	10.12	10.27
64	10.24	10.09	10.44
128	10.05	9.88	10.46
256	9.58	9.43	10.25

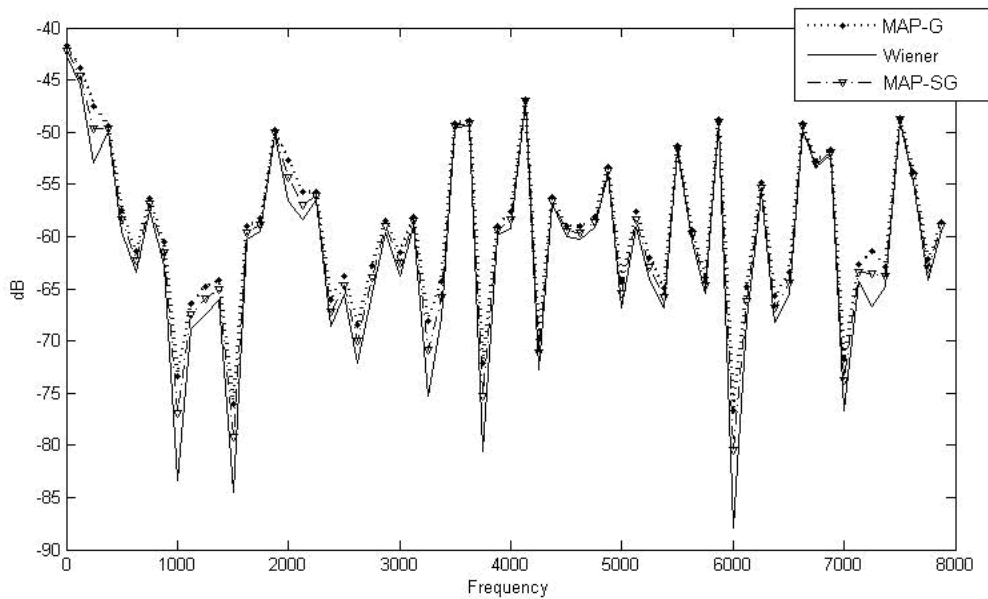
چنانچه در جدول دیده می‌شود مقادیر نسبت سیگنال به نویز بهبودیافته برای سه روش فوق برای یک زیرباند خاص دارای مقادیر نزدیک بهم می‌باشد و این اختلاف در حدود 1dB می‌باشد که می‌توان از آن صرفه نظر کرد. از سویی نیز مقدار نسبت سیگنال به نویز برای زیرباندهای مختلف نیز تقریباً یکسان می‌باشد. در واقع از لحاظ تست نظری تفاوت چندانی بین تعداد زیرباندها وجود ندارد، ولی تست



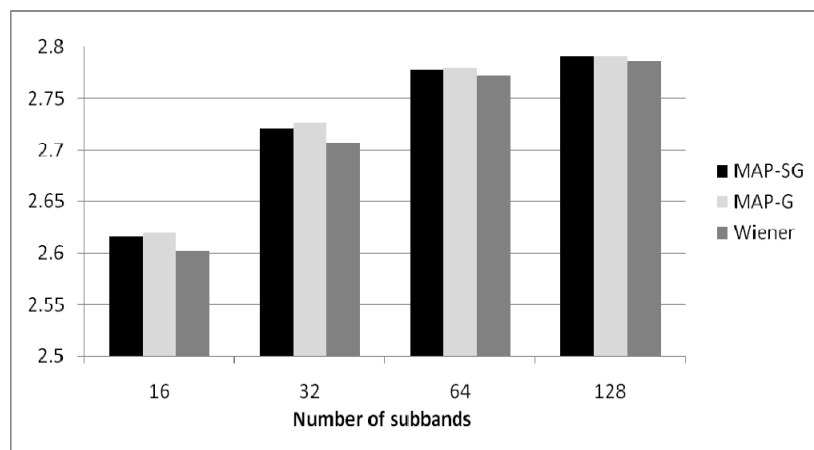
تعداد زیرباند = ۱۶



تعداد زیرباند = ۲۵۶



شکل ۱۱- اندازه سه فیلتر وینر، MAP-G و MAP-SG را برای یک فریم با طول ۲۵۶ نمونه و تعداد زیرباند ۶۴ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB



شکل ۱۲- مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای فریم ۲۵۶ نمونه‌ای و تعداد زیرباند ۱۶، ۳۲، ۶۴ و ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

به نویز در حوزه فرکانس می‌باشد، که این امر به دلیل خواص موجک می‌باشد. به این معنی که می‌توان در حوزه موجک مقیاس زمان- فرکانس مختلف داشته باشیم که این خاصیت موجک را ابزاری قدرتمند برای تحلیل سیگنال گفتار می‌کند.

جدول ۲- مقادیر نسبت سیگنال به نویز بهبودیافته را برای سه روش فیلتر کردن وینر، MAP-G و MAP-SG در دو حوزه موجک و فوریه با طول فریم ۲۵۶ و تعداد زیرباند ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

	Wavelet	FFT
MAP-SG	10.05	4.35
MAP-G	9.88	3.4
Wiener	10.46	6.44

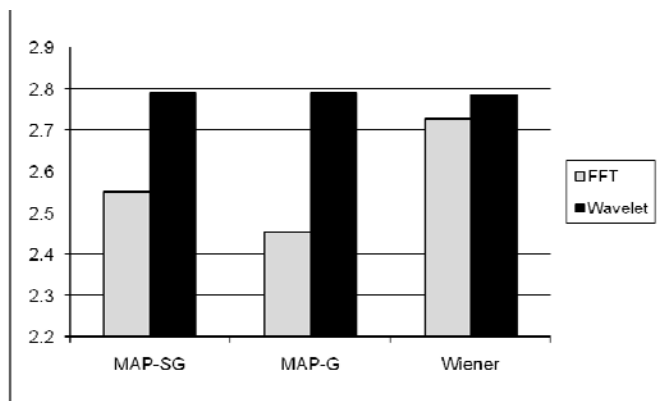
شکل ۱۲ مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای فریم ۲۵۶ نمونه‌ای و تعداد زیرباند ۱۶، ۳۲، ۶۴ و ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB نشان می‌دهد. چنانچه دیده می‌شود روش MAP-G دارای مقدار بیشتری نسبت به دو روش دیگر در طول فریم‌های مختلف دارد. همچنین نتایج بهتری برای تعداد زیرباند بیشتر نسبت به تعداد زیرباند کمتر بدست آمد که این امر بدلیل تحلیل فرکانسی دقیقتر در زیرباندهای بیشتر می‌باشد.

نتایج بدست آمده از آزمایشات با نتایج تئوری که در شکل ۵ نشان داده شده بود مطابقت دارد و دیده می‌شود که فیلتر وینر دارای کمی تخریب گفتار بیشتر نسبت به دو فیلتر پیشنهادی است. در ادامه سه فیلتر فوق را در حوزه فرکانس پیاده سازی شده و نتایج آنها با معادله‌شان در حوزه موجک مقایسه گردید. جدول ۲ نتایج تست نظری یعنی مقدار نسبت سیگنال به نویز بهبودیافته را برای این سه فیلتر در دو حوزه نشان می‌دهد. همانطور که در جدول دیده می‌شود مقدار نسبت سیگنال به نویز برای هر سه فیلتر در حوزه موجک بیشتر از مقدار نسبت سیگنال

شکل ۱۳ طیف نگار قسمتی از سیگنال بهبودیافته را که توسط روش‌های مختلف حذف نویز در دو حوزه موجک و فرکانس بدست آمده است را نشان می‌دهد. مطابق با شکل در تمامی طیفها فرکانس‌های اصلی سیگنال حفظ شده‌اند ولی در روش‌های حذف نویز در حوزه فرکانس، نویز زمینه بیشتری باقی مانده است و همین امر باعث کاهش کیفیت شنیداری آنها می‌شود. چنانچه در طیف سیگنالهای بهبودیافته‌ی بدست آمده از فیلترهای پیشنهادی در حوزه موجک نیز دیده می‌شود نویز بیشتری حذف شده است بدون اینکه مؤلفه‌های اصلی فرکانس مخدوش شوند. این مطلب برای هر سه فیلتر وینر، MAP-G و MAP-SG صادق می‌باشد.

شکل ۱۴ مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای دو حوزه فرکانس و موجک با طول فریم ۲۵۶ و تعداد زیرباند ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB را نشان می‌دهد. چنانچه دیده می‌شود مقادیر MOS تخمینی با PESQ برای هر سه فیلتر در حوزه موجک بیشتر از سه فیلتر فوق در حوزه فرکانس می‌باشد و بهبود حاصل شده توسط روش موجک در مقایسه با FFT می‌تواند به تجزیه مناسب تر سیگنال توسط موجک نسبت داده شود.

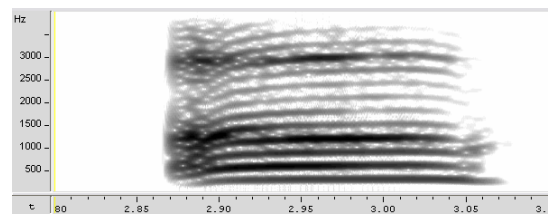
در مجموع می‌توان با استفاده از نتایج ذکر شده ادعا نمود دو فیلتر پیشنهادی و فیلتر وینر دارای کیفیت بهتری هم از نظر تست نظری و هم از نظر تست عملی و شنوائی نسبت به معادلشان در حوزه فرکانس می‌باشند.



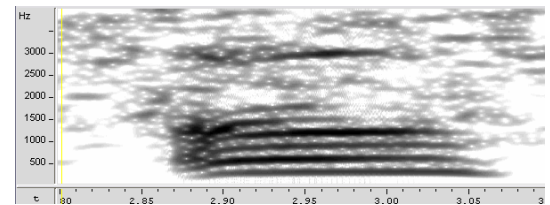
شکل ۱۴- مقادیر MOS تخمینی با PESQ برای سه فیلتر وینر، MAP-G و MAP-SG برای دو حوزه فرکانس و موجک با طول فریم ۲۵۶ و تعداد زیرباند ۱۲۸ در حضور نویز گوسی سفید با نرخ سیگنال به نویز ورودی 10dB

۷- نتیجه گیری

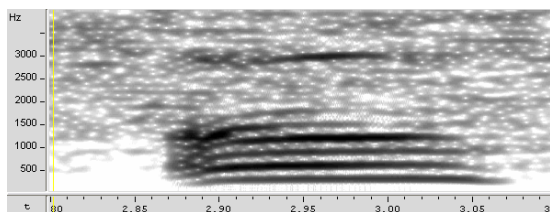
در کار حاضر، ابتدا نشان داده شد که روش فیلتر کردن ضرایب تبدیل موجک بر آستانه‌گذاری این ضرایب، چه از نظر عملی و چه از دیدگاه نظری برتری دارد. دلیل این امر می‌تواند در تخمین هوشمندانه نویز در روش‌های فیلترینگ باشد. در ادامه دو فیلتر جدید در حوزه موجک با الهام از روش‌های بهینه فیلتر کردن در حوزه فرکانس، پیشنهاد شده‌اند. این فیلترها، احتمال پسین ضرایب تبدیل موجک سیگنال بهبود یافته را بیشینه می‌کنند. فیلترهای پیشنهادی در حوزه موجک، از لحاظ ارزیابی نظری عملکردی همپا، و از لحاظ ارزیابی با معیار PESQ عملکردی بهتر نسبت به فیلتر وینر در حوزه موجک دارند. به عبارت دیگر اختلاف مقدار نسبت سیگنال به نویز بهبودیافته برای هر سه نوع فیلتر در حد 1dB می‌باشد که قابل صرف نظر کردن است. از طرفی نیز در ارزیابی PESQ فیلتر MAP-G عملکرد بهتری نسبت به دو فیلتر دیگر دارد.



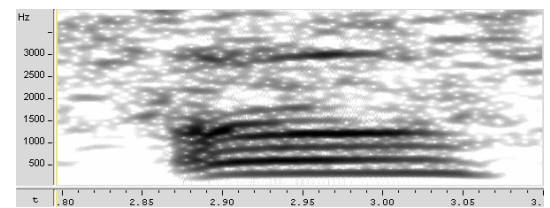
(الف) سیگنال تمیز



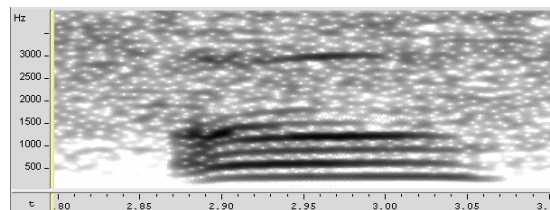
(ب) حذف نویز با فیلتر وینر در حوزه موجک



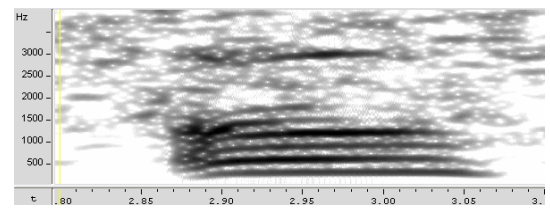
(ج) حذف نویز با فیلتر وینر در حوزه فرکانس



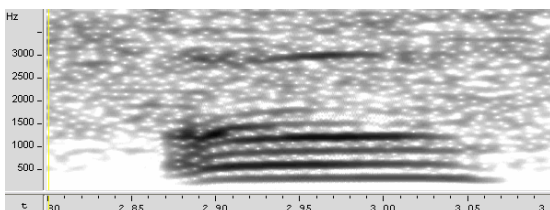
(د) حذف نویز با فیلتر MAP-G در حوزه موجک



(ه) حذف نویز با فیلتر MAP-G در حوزه فرکانس



(و) حذف نویز با فیلتر MAP-SG در حوزه موجک



(ز) حذف نویز با فیلتر MAP-SG در حوزه فرکانس

شکل ۱۳- طیف نگار قسمتی از سیگنال بهبودیافته بدست آمده از روش‌های مختلف حذف نویز در دو حوزه موجک و فرکانس

transform," *Speech Communication*, Vol. 41, pp. 409-427, 2003.

[11] E. Zwicker, and H. Fastl, *Psychoacoustics: Facts and Models*, New York, Springer-Verlag, 1990.

[12] N. Virag, "Single channel speech enhancement based on masking properties of the human auditory system", *IEEE Trans. Speech and Audio Processing*, Vol. 7, No. 2, pp. 126-137, 1999.

[13] Y. Shao, and C. H. Chang, "A versatile speech enhancement system based on perceptual wavelet denoising," *Proc, IEEE Int'l Symp. on Circuits and Systems*, pp. 864-867, 2005.

[14] S. H. Chen, and J. F. Wang, "Speech Enhancement Using Perceptual Wavelet Packet Decomposition and Teager Energy Operator", *Journal of VLSI Signal Processing*, Vol. 36, No. 2-3, pp. 125 - 139, 2004.

[15] G. Ju, and L. Lee, "Speech Enhancement and Improved Recognition Accuracy by Integrating Wavelet Transform and Spectral Subtraction Algorithm," *Proc, European Conf. on Speech Communication and Technology*, pp. 1370-1380, 2003.

[16] N. Ma, M. Bouchard, and R. A. Goubran, "A Wavelet Kalman Filter With Perceptual Masking For Speech Enhancement In Colored Noise," *Proc, IEEE Int'l Conf. on Acoustics, Speech, and Signal processing*, Vol. 1, pp. 149-152, 2005.

[17] I. Y. Soon, S. N. Kah, and C. K. Yeo, "Wavelet for speech denoising," *Proc, IEEE TENCON*, Vol. 2, pp. 479-482, 1997.

[18] Y. Ephraim, and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. 32, No. 6, pp. 1109-1121, 1984.

[19] P. J. Wolfe, and S. J. Godsill, "Simple alternatives to the Ephraim and Malah suppression rule for speech enhancement," *Proc, 11th IEEE Workshop Statistical Signal Processing*, pp. 496-499, 2001.

[20] A. Grossman, R. Martinet, and J. Morlet, "Reading and understanding continuous wavelet transform," *Proc, Int'l Conf. Wavelets*, pp. 2-20, 1989.

[21] Y. H. Huang, and T. D. Chiueh, "A new audio coding scheme using a forward masking model and perceptually weighted vector quantization," *IEEE Trans. Speech and Audio Processing*, Vol. 10, No. 5, pp. 325-335, 2002.

[22] Y. Shao, and C. H. Chang, "A Generalized Perceptual Time-Frequency Subtraction Method for Speech Enhancement," *Proc, IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2537-2540, 2006.

[23] C. Tantibundhit, J. R. Boston, C. C. Li, J. D. Durrant, S. Shaiman, K. Kovacyk, and A. El-Jaroudi, "Speech

در انتها نیز نتایج بیانگر آنند که فیلترهای پیشنهادی در حوزه موجک، از دیدگاه ارزیابی نظری دارای عملکرد برجسته‌ای نسبت به معادلشان در حوزه فرکانس هستند، این اختلاف در مقدار نسبت سیگنال به نویز بهبودیافته برای دو روش پیشنهادی در حوزه موجک در قیاس با مقدار نسبت سیگنال به نویز بهبودیافته برای فیلترهای معادلشان در حوزه فوریه حدود 6dB می‌باشد و این اختلاف برای فیلتر وینر حدوداً 4dB است. علاوه بر این معیار PESQ نیز مؤید عملکرد مناسب این فیلترها در حوزه موجک است. از نظر پیچیدگی محاسباتی، محاسبه فیلتر در روش MAP-G حدود ۵ برابر و در روش MAP-SG حدود ۷ برابر محاسبات لازم برای فیلتر وینر در هر زیربانند می‌باشد. پیچیدگی محاسباتی اعمال فیلترها یکسان می‌باشد، چراکه طول فیلترها یکسان است. قابل ذکر است در روشهای مورد استفاده محاسبات ناشی از تبدیل موجک بیشتر از محاسبات مربوط به فیلترهاست از اینرو افزایش محاسبات فوق‌الذکر در روشهای پیشنهادی اثر قابل توجهی در زمان اجرای کل فرایند حذف نویز ندارند.

مراجع

[1] A. Huang, A. Acero, and H. Hon, *Spoken Language processing*, Prentice Hall, 2001.

[2] Y. Hu, and P. C. Loizou, "Incorporating a psychoacoustical model in frequency domain speech enhancement," *Proc, IEEE Signal Processing Letters*, Vol. 11, pp. 270-273, 2004.

[3] A. C. Cristan, and A. T. Walden, "Multitaper power spectrum estimation and thresholding: Wavelet packets versus wavelets," *IEEE Trans. Signal Processing*, Vol. 50, pp. 2976 - 2986, 2002.

[4] Y. Hu, and P. C. Loizou, "Speech enhancement based on wavelet thresholding the multitaper spectrum," *IEEE Trans. Speech and Audio Processing*, Vol. 12, No. 1, pp. 59-67, 2004.

[5] K. P. Soman, and K. I. Ramachandran, *Insight into wavelets: From Theory to Practice*, India, Prentice-Hall, 2005.

[6] S. Jung, Y. Kwon, and S. Yang, "Speech Enhancement By Wavelet Packet Transform With Best Fitting Regression Line In Various Noise Environments," *IEEE Int'l Conf. on Acoustics, Speech, and Signal processing*, Vol. 1, pp. 469-472, 2006.

[7] D. L. Donoho, and I. M. Johnston, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* 81, pp. 425-455, 1994.

[8] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Information Theory*, pp. 613-627, 1995.

[9] H. Sheikzadeh, and H. R. Abutalebi, "An improved wavelet based speech enhancement system," *Proc, European Conf. on Speech Communication and Technology*, pp. 1855-1858, 2001.

[10] C. T. Lu, and H. C. Wang, "Enhancement of single channel speech based on masking property and wavelet

Speech Model," *EURASIP Journal on Applied Signal Processing*, No. 7, pp. 1110-1126, 2005.

[۳۷] ب. زمانی، ب. ناصرشریف، و ا. اکبری، "بررسی اثر باندهای بحرانی و پارامترهای بکارگیری تبدیل موجک در بهبود گفتار"، *مجموعه مقالات یازدهمین کنفرانس بین‌المللی کامپیوتر ایران*، ۱۳۸۴.

[۳۸] ب. زمانی، س. رسول‌زاده، ب. ناصرشریف، و ا. اکبری، "بکارگیری بانک‌فیلتر WOLA در بهبود گفتار بر مبنای تبدیل موجک به منظور کاهش پیچیدگی محاسباتی"، *مجموعه مقالات چهاردهمین کنفرانس بین‌المللی برق ایران*، ۱۳۸۵.



بهزاد زمانی دهکردی دانشجوی دوره دکترای هوش مصنوعی در دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت می‌باشد. او همچنین درجه کارشناسی‌ارشد هوش مصنوعی را از دانشگاه علم و صنعت در سال ۱۳۸۴ و کارشناسی سخت‌افزار را از دانشگاه صنعتی اصفهان در سال ۱۳۸۱ اخذ نموده است. در حال حاضر وی در آزمایشگاه پردازش صوت و گفتار در زمینه تشخیص گفتار و روش‌های متمایزسازی بر مبنای توابع هسته مشغول به تحقیق است. زمینه تحقیق وی شامل بهبود گفتار، بازشناسی گفتار، تبدیل موجک، تبدیلات متمایزساز ویژگی و توابع هسته می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

bzamani@iust.ac.ir



احمد اکبری درجه کارشناسی را در مهندسی الکترونیک و درجه کارشناسی‌ارشد را در مهندسی مخابرات از دانشگاه صنعتی اصفهان در سالهای ۱۳۶۶ و ۱۳۶۹ دریافت کرد. وی درجه دکترا را در پردازش سیگنال و مخابرات راه دور از دانشگاه رن فرانسه در سال ۱۳۷۴ اخذ نمود و در همان سال به عنوان عضو هیات علمی به دانشگاه علم و صنعت پیوست و در حال حاضر با عنوان دانشیار دانشکده مهندسی کامپیوتر مشغول به فعالیت است. وی سرپرست مرکز تحقیقات فناوری اطلاعات و آزمایشگاه پردازش صوت و گفتار در این دانشگاه است. در ضمن ایشان سمت سردبیری مجله علمی پژوهشی موسسه پردازش هوشمند علائم را بر عهده دارند و از سال ۱۹۹۹ عضو انجمن کامپیوتر ایران می‌باشند. زمینه تحقیق وی شامل بهبود گفتار، تشخیص گفتار، مقاوم‌سازی بازشناسی گفتار، انتقال صوت از طریق شبکه و پروتکل‌های مربوطه می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

akbari@iust.ac.ir



بابک ناصرشریف متولد ۱۳۵۳ در تهران است. وی درجه کارشناسی را در رشته مهندسی کامپیوتر گرایش سخت‌افزار از دانشگاه صنعتی امیرکبیر در سال ۱۳۷۶ دریافت نمود. وی موفق به اخذ درجه کارشناسی‌ارشد و دکترا در رشته مهندسی کامپیوتر گرایش هوش مصنوعی از دانشگاه علم و صنعت ایران در سال‌های ۱۳۷۹ و ۱۳۸۶ گردید. وی از سال ۱۳۸۶ تا کنون عضو هیئت علمی گروه مهندسی کامپیوتر در دانشکده فنی دانشگاه گیلان است. زمینه تحقیق وی شامل بهبود گفتار، بازشناسی گفتار و مقاوم‌سازی آن، تبدیل موجک، مدل‌سازی و بازشناسی الگو و واژه‌یابی گفتاری می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

nasser_s@iust.ac.ir

enhancement using transient speech components," *Proc, IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing*, pp. 833-836, 2006.

[24] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelet and Wavelet Transforms*, Prentice Hall Inc., 1998.

[25] R. Martin, "Noise Power Spectral Density Estimation Based on Optimal smoothing and Minimum Statistic," *IEEE Trans. on Speech and Audio Processing*, Vol. 9, No. 5, 2001.

[26] J. Weaver, D. M. Healy, and L. D. Cromwell, "Filtering noise from images with wavelet transforms," *Magnetic Resonance in Medicine*, Vol. 21, No. 2, pp. 288-95, 1991.

[27] J. W. Seok, and K. S. Bae, "Speech enhancement with reduction of noise components in the wavelet domain," *Proc, IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing*, pp. 1323-1326, 1997.

[28] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive Wavelet Thresholding for Image Denoising and Compression," *IEEE Trans. Image Processing*, Vol. 9, No. 9, pp. 1532-1546, 2000.

[29] I. J. Kim, S. I. Yang, and Y. Kwon, "Speech enhancement using adaptive wavelet shrinkage," *Proc, IEEE Int'l Symp. on Industrial Electronics*, Vol. 1, pp. 501-504, 2001.

[30] I. M. Johnston, and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," *Journal of the Royal Statistical Society, Series B*, Vol. 59, pp. 319-351, 1997.

[31] R. Moussaoui, J. Rouat, and R. Lefebvre, "Wavelet based independent component analysis for multi-channel source separation," *Proc, IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing*, pp. 645-648, 2006.

[32] "Perceptual Evaluation of Speech Quality (PESQ), an objective method for end-to-end speech quality assessment of narrow band telephone network and speech codec," ITU-T Recommend, pp. 862, Feb 2001.

[33] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra, "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs," *Proc, IEEE ICASSP*, Vol. 2, pp. 749-752, 2001.

[34] S. Ayat, M. Manzuri, and R. Dianat, "An Improved Wavelet-Based Speech Enhancement by Using Speech Signal features," *International Journal of Computers & Electrical Engineering*, Vol. 32, No. 6, pp. 411-425, 2006.

[35] R. McAulay, and M. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoustics Speech Signal Processing*, vol. 28, no. 2, pp. 137-145, 1980.

[36] T. Lotter, and P. Vary, "Speech Enhancement by MAP Spectral Amplitude Estimation Using a Super-Gaussian

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۶/۰۲/۲۳

تاریخ اصلاح: ۸۹/۰۱/۲۴

تاریخ قبول شدن: ۸۹/۰۳/۲۵

نویسنده مرتبط: بهزاد زمانی دهکردی، دانشکده مهندسی کامپیوتر، دانشگاه

علم و صنعت ایران، تهران، ایران.

-
- ¹ Hands Free
 - ² Bluetooth
 - ³ Wiener Filter
 - ⁴ Threshold
 - ⁵ Step Garrote
 - ⁶ Minimizing the False Discovery Rate
 - ⁷ Top
 - ⁸ Minimizing Stein's Unbiased Risk Estimate
 - ⁹ Wavelet Packet
 - ¹⁰ Maximum A Posteriori Spectral Amplitude Estimator
 - ¹¹ Daubechy Order 10
 - ¹² Background Artifact
 - ¹³ Speech Artifact
 - ¹⁴ Perceptual Evaluation of Speech Quality

چراغ راهنمای هوشمند

امین‌اله مه‌آبادی

دانشکده فنی و مهندسی، دانشگاه شاهد، تهران، ایران

چکیده

در این مقاله مدلی نو برای یک سیستم ویدیویی خودکار محاسبه‌گر حجم تردد خودروهای عبوری^۱ جهت زمان‌بندی و کنترل هوشمند چراغ راهنمایی^۲ ارائه شده است. ضمن محاسبه حجم ترافیک براساس خودروها و به تفکیک لاین‌ها، تردد چهارراه را زمان‌بندی می‌کند. به پارامتر انتظار اجتماعی راننده به عنوان زمان پیش‌فرض توجه دارد. در مقابل ازدحام، سایه‌ها، تغییرات محیطی و تغییرات آرام نور محیط مقاوم و تاحدودی تغییرات ناگهانی را تحمل کرده و به شعاع پوشش دوربین وابستگی ندارد. سایه و خطاهای موجود را جهت افزایش دقت، از تصاویر حذف می‌کند و نیازمند به تشخیص و تفکیک خودرو نیست. دارای دقت بالای تشخیص حجم خودروها، حذف سایه خودرو و تشخیص خودرو بصورت مجزا با روش جدیدی است. سیستم ۱۰ فریم در ۱ ثانیه برای هر دوربین را پردازش می‌کند و با افزایش فریم‌ها نتایج تشخیص بهبود می‌یابد. به دلیل تناسب سرعت خودروها با سرعت حرکت در چهارراه، سیستم بلادرنگ^۳ عمل می‌کند.

کلمات کلیدی: مدل چهارراه، چراغ راهنمایی هوشمند، حمل و نقل هوشمند^۴، بخش‌بندی تصویر^۵، فرکانس دایره‌ای^۶، شبکه پتری.

۱- مقدمه

این روش زیاد قابل اطمینان نیست و در کف جاده نصب می‌شود. تغییر لاین خودرو را نمی‌فهمد. در عبور دو خودرو در یک تقاطع به اشتباه یکی را از قلم می‌اندازد، مشکل توان مصرفی و شعاع کم پوشش دارد. امکان ایجاد چهارراه مجازی و قدرت تحلیل را ندارد و بیشتر برای جریان ترافیک بکار می‌رود. محدودیت زمانی ثبت در ثانیه را دارد [۶].

استفاده از حلقه‌های مغناطیسی، سنسورهای فشار، امواج ماوراء صوت و دوربین‌های ویدیویی در بسیاری از کشورها مرسوم هستند. در این میان بررسی تصاویر ویدیویی از خطا می‌کاهد و نتایج حاصل را دقیق‌تر می‌کند. از تصاویر می‌توان بعنوان اسناد مهم برای استنتاج‌های دیگر استفاده کرد که طبیعتاً ترکیب سیستم‌های مختلف دقت در محاسبه پارامترها و نتایج را افزایش می‌دهد [۷، ۱۰].

سیستم خودکار کنترل ترافیک ویدیویی نسبت به سیستم‌های دیگر امتیازات بسیاری دارد. اولاً دوربین با شعاع دید وسیعتر پارامترهای بیشتری (مانند طبقه بندی نوع خودرو، تعداد خودروها، حجم ترافیک و غیره) را بررسی و ارائه می‌دهد. ثانیاً، نصب دوربین‌های ویدیویی^۷ بسیار آسانتر و کم هزینه‌تر از حلقه‌های مغناطیسی^۸ و مانند آن است [۱۱]. ثالثاً، دیدگسترده دوربین می‌تواند جایگزین

سیستم‌های حمل و نقل هوشمند در دنیای آینده مسیر اصلی توسعه مدیریت ترافیک هستند. این سیستم‌ها به زیرساخت هوشمند و خودروهای هوشمند تقسیم می‌گردند. در این ساختار، سیستم‌های هوشمند کنترل ترافیک از بخش‌های کلیدی آن هستند. سیستم‌هایی که علاوه بر فعالیت اصلی خود بتوانند پارامترهای مربوط به حرکت خودروها را نیز ارائه دهند [۱]. این پارامترها می‌توانند تشخیص حجم ترافیک، نوع خودروها، تعداد خودروها، سرعت خودروها، تخلفات رانندگی، حوادث رانندگی و مانند آن باشند [۲]. به علت وجود تفاوت در پارامترهای مورد نظر، روشهای متفاوت مورد استفاده قرار می‌گیرند [۳، ۴]. یک روش، استفاده از نیروی انسانی مشاهده‌گر جهت جمع‌آوری اطلاعات و استنتاج پارامترهای مورد نظر اطلاعات بدست آمده است که در آن ابزارهای فیزیکی نیز می‌توانند مورد استفاده قرار گیرند [۵].

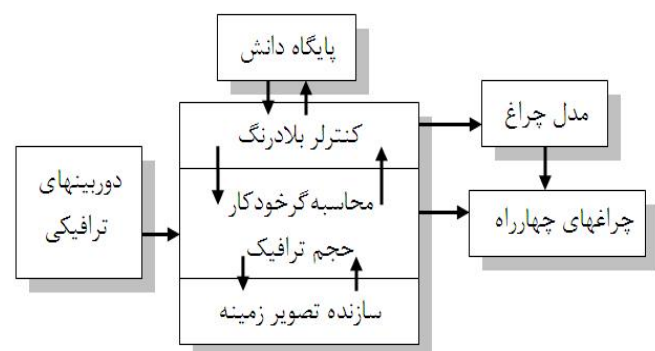
روشی که فعلاً بیشترین استفاده را در شمارش تعداد اتومبیل‌ها و تخمین سرعت آن‌ها دارد، استفاده از حلقه‌های مغناطیسی نصب شده کف خیابان است.

فاز، توالی فازها و پشتیبانی از وقفه‌های ترافیکی تولیدی کنترلر بلادرنگ (مانند بروز حادثه) براساس مدل جدید چند لایه پتری^{۱۴} است. بخش دوم، وظیفه استخراج تصویر زمینه یا مرجع^{۱۵} از فریم‌های ورودی را برعهده دارد. بخش سوم، برای محاسبه و نمایش حجم ترافیک با تصاویر ورودی جهت تولید پارامترهای زمانبندی چراغ‌های هرفاز و توالی فازها، تصویر زمینه را اصلاح و نهایتاً تصویر بخش‌بندی شده ورودی را بعنوان تصویر نهایی ارائه می‌دهد تا محاسبات ترافیکی تولید پارامترهای زمانبندی انجام شود.

بخش چهارم، کنترلر بلادرنگ در بردارنده پایگاه دانش که براساس نتایج استخراجی بخش سوم به بررسی وضعیت مداوم چهارراه، استخراج استثنایا و تولید وقفه‌های مربوطه برای بخش اول و دوم می‌پردازد تا بتواند هوشمندی خود را در نقش عملیات بلادرنگ یک چراغ در یک ساختار غیر مستقیم ایفاء کند. نتایج (مطابق روال شکل ۲) از تصویر بخش‌بندی شده بدست می‌آیند تا نهایتاً پارامترهای زمانبندی چراغ تولید شود. در اینجا کنترلر بلادرنگ با پایگاه دانش خود، وظیفه کنترل تطبیق حتی‌الامکان زمانبندی با ساختار بلادرنگ را بر عهده دارد و در صورت لزوم سیگنال‌های وقفه لازم آن را تولید می‌کند. نوآوری ما در این سیستم عبارتست از:

- ۱- ارائه یک مدل جدید تحلیل ترافیک چهارراه براساس تصاویر ویدئویی.
- ۲- مدل‌سازی کنترلر بلادرنگ یک چراغ راهنمایی هوشمند، ارائه مدل قابل تغییر با فازهای مختلف چراغ براساس متاپتری نت (بخش ۲ و ۳) و مدل‌سازی یک چهارراه چهار فاز عالی با جهت‌های مستقیم و مورب (بخش ۴).
- ۳- زمانبندی ترافیک چهارراه براساس حجم ترافیک در تمامی تقاطع‌ها (بخش ۵)، عدم وابستگی به شکل خودرو و مقاوم در برابر ترافیک سنگین فشرده (بخش ۶) و عدم استفاده از روش‌های ردیابی و مقاوم در برابر توقف طولانی مدت (بخش ۷).
- ۴- ارائه تکنیک جدید تحمل تغییرات آرام محیط و پویای نور (بخش ۸)، تکنیک جدید و دقیق حذف سایه از تصاویر ترافیکی خودرو و عدم نیاز به تطابق کامل شعاع پوشش دوربین‌ها یا تنظیم دقیق آنها، و انجام محاسبات ترافیک به تفکیک لاین‌ها و امکان توسعه آتی آنها (بخش ۹) و
- ۵- طراحی مناسب جهت تبدیل سیستم به یک سنسور پردازشگر هوشمند.

در ادامه ارائه مقاله، ابتدا در بخش ۳ اجزای سیستم تعریف شده و مدل جدید چهارراه ارائه می‌گردد؛ سپس در بخش ۴ روابط محاسباتی زمانبندی و تحلیل ترافیک مورد بحث قرار می‌گیرد؛ در بخش ۶ شیوه تخمین حرکت و تولید فرامین مورد نیاز ارائه می‌شود. سپس در بخش ۷ الگوریتم ساختن تصاویر زمینه معرفی می‌گردد. در بخش ۸ روش تشخیص حجم ترافیک، حذف سایه بیان می‌گردد. در بخش ۹ نتایج حاصل بیان و در بخش ۱۰ نتیجه‌گیری نهایی ارائه می‌شود.



شکل ۱- سیستم چراغ راهنمای هوشمند

چندین سنسور^۹ یا حلقه مغناطیسی گردد. رابعا، برخلاف امواج ماوراء صوت^{۱۱} یا حلقه‌های مغناطیسی استفاده از دوربین برای انسان کاملاً بی‌خطر است. خامسا تعویض لاین خودروها را می‌تواند درک کند درحالی‌که حلقه‌ها اشتباه دارند و مناسب خودروهای ساکن نیستند [۷، ۸، ۱۱، ۱۳].

سیستم‌های خودکار کنترل ترافیک ویدئویی می‌توانند در دو حالت بلادرنگ^{۱۱} و غیرمستقیم^{۱۲} مورد استفاده قرار گیرند [۱۴]. مثلاً برای تعیین زمان تعویض چراغ‌های راهنمایی یا تشخیص و ثبت تخلفات، سیستم در حالت غیرمستقیم استفاده می‌شود درحالی‌که کنترل ترافیک در محل‌های پرتراکم و همراه با تغییرات شبانه روز، بصورت بلادرنگ صورت می‌گیرد. تاکنون تشخیص خودکار ترافیک با استفاده از روش‌های ویدئویی، بدلیلی بطور گسترده مورد استفاده قرار نگرفته است.

اولا این سیستم‌ها دارای محدودیت‌کارایی هستند و مثلاً با ازدحام زیاد، سایه‌های بلند و تغییرات نور شب و روز مشکل دارند [۷]. ثانياً هزینه تولید آنها بسیار بالا است. ثالثاً وجود خطوط ارتباط سریع انتقال تصاویر دوربین‌ها به پردازشگر لازم است. رابعا برای عملیات بلادرنگ به سخت‌افزار پیشرفته نیاز دارند. اکنون توجه به این سیستم بسیار زیاد شده است. به دلیل آنکه می‌توان از قابلیت‌های دیگر آن مانند شمارش خودروها و تعیین کلاس آنها بهره‌برد، در حال گسترش و در آینده جایگزین روش‌های سنتی می‌گردد.

۲- چراغ هوشمند

چراغ هوشمند راهنمایی را چراغی می‌نامیم که با دریافت تصاویر دوربین‌ها، امر کنترل هوشمند ترافیک را برعهده دارد. ضمناً می‌تواند وقوع حادثه را به نوعی حدث بزند و دوباره برنامه‌ریزی شود. زمانبندی چهارراه را متناسب با ترافیک خیابان‌های ورودی و خروجی، تنظیم کند و بتواند به پارامتر رضایت اجتماعی رانندگان توجه کند. سیستم پیشنهادی با دارا بودن این ویژگی‌ها، باوجود داشتن ماهیت غیرمستقیم، با بهره‌بردن از کنترلر بلادرنگ تحلیل‌گر دانش خود، نقش یک سیستم بلادرنگ را بازی می‌کند تا بتواند با رفع موانع تغییرات محیطی، کارایی خود را افزایش دهد. درمقابل خرابی‌هایی مانند قطع ارتباط تصاویر دوربین مقاوم است و با اتصال مجدد، بلافاصله اصلاح لازم را صورت می‌دهد.

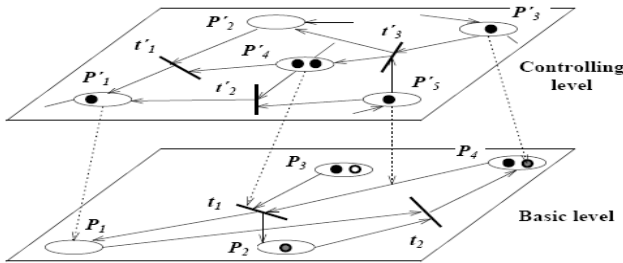
یک مدل نوین چندلایه کنترلر چراغ برای آن طراحی و بکارگرفته شده، و قابلیت انعطاف تغییر فاز چراغ از چهار فاز به سه فاز و غیره را دارد. با بهره‌از تکنیک‌های جدید و افزایش کارایی الگوریتم‌ها در مقابل ازدحام، سایه‌ها، و تغییرات محیطی مشکلی ندارد. نیازی به وجود فاصله بین خودروها جهت تشخیص آنها نیست و اصولاً وابسته به شکل خودرو نمی‌باشد. در صورت سکون و عدم حرکت خودروها، به راحتی حجم آنها را محاسبه می‌کند و وابسته به ردیابی خودروها نیست. روش‌هایی که تنها از تشخیص خودرو و روش شمارش خودروها استفاده می‌کنند نمی‌توانند حالت ترافیک سنگین با سکون طولانی مدت و فشرده‌گی خودروها را تشخیص دهند. برای ردیابی بکار می‌روند و مناسب مقایسه با محاسبات چراغ‌راهنمایی چهارراه نیستند [۱۵].

روش پیشنهادی ما ضمن حل مشکلات روش‌های قبلی در بخش حذف سایه، قادر است در صورت نیاز از قدرت تشخیص خودرو مشابه روش‌های ردیابی استفاده کند. سیستمی با ساختار مشابه جهت محاسبات تصاویر دریافتی و قابل کاربرد برای زمانبندی چهارراه مشاهده نشد و بیشتر تکنیک‌های موجود پردازش تصویر جهت ردیابی خودروها و سیستم‌های در حال حرکت بکار می‌روند.

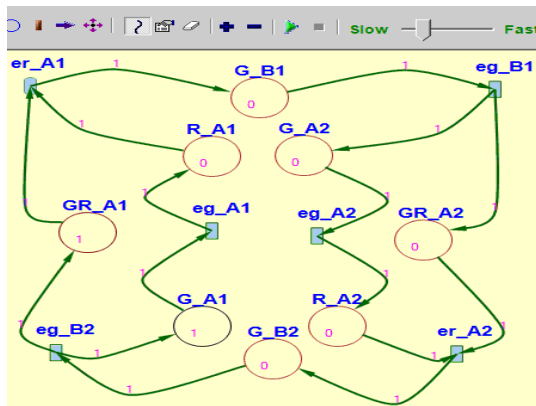
این سیستم (مطابق شکل ۱) از چهار بخش اصلی ۱- مدل کنترلر چراغ، ۲- سازنده تصویر زمینه، ۳- محاسبه‌گر حجم ترافیک و ۴- کنترلر بلادرنگ سیستم^{۱۳} تشکیل می‌گردد. وظیفه بخش اول، کنترل عملیات چراغ‌های مختلف هر

۳- مدل چهارراه

به مدل پتری نت زمانی، لایه‌هایی اضافه شده که با تولید سیگنال‌های پایان زمان چراغ، آن‌را به مدل اعمال و امکان اتصال گذار^{۲۳} به گذار را (که در مدل پتری نبود)، فراهم می‌آورد. مدل چراغ ارائه شده دارای یک ساختار دوسطحی (پایه و کنترل) با چندلایه بنام متا^{۲۴}، پایه^{۲۵}، مدل^{۲۶} و متامدل^{۲۷} همراه یک مولد سیگنال وقفه^{۲۸} است. در لایه متا می‌توان مکان‌های فازهای مورد نظر را روشن یا خاموش^{۲۹} نمود که با اینکار آن مکان و اتصالاتش (یعنی یک فاز) از شبکه پویای مدل^{۳۰} حذف می‌شود.



شکل ۵- سطوح پایه و کنترل مدل متاپتری نت [۱۷]



شکل ۶- بخشی از مدل پتری کنترلر چهارراه ۴ فاز و ابزار مدل‌سازی آن

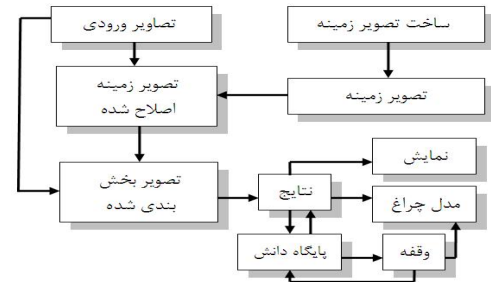
لایه مدل پایه، حاوی مدل پایه یا چهارفاز با ساختار پتری نت زمانی است. لایه مدل، شامل مدل نهائی آماده کار (مثلا سه فاز) و زیرمجموعه‌ای از لایه مدل پایه است. مولد سیگنال وظیفه تولید سیگنال‌های خاص وقفه (زمان چراغ) را دارد. با ایجاد لایه‌های جدید و دانش مدل‌های متانتورک‌های منطقی، مدل ترکیبی ارائه شده تا ضمن پشتیبانی از قطع عملیات جاری بتواند تغییرات بلادرنگ محیط را به سیستم منتقل کند [۱۸].

فرامین مولد سیگنال وقفه به لایه مدل اعمال می‌شود. در این مدل می‌توان فازهای مختلف را فعال یا غیرفعال^{۳۱} کرد. مدل چراغ ارائه شده یک مدل منحصر بفرد است که ضمن امکان کنترل، اجازه تغییر و انعطاف مدل کامل پایه به مدل‌های دیگر را فراهم می‌کند. مثلاً "در لایه متا با خاموش کردن B1، میتوان این فاز را از مدل حذف کرد و به یک مدل سه فاز تبدیل نمود.

در این سیستم با قطع ارتباط دوربین‌ها، چراغ با زمان تقریبی تعیین شده به عملیات خود ادامه می‌دهد و در مقابل خرابی از این نوع مقاوم است. لذا شیوه ارتباط راه‌های چهارراه، کنترل و تصمیم‌گیری چراغ‌ها با متاپتری نت مدل می‌شود و روتین‌های تصمیم‌گیری در گذارهای شبکه می‌آیند که با دریافت ژتون‌های^{۳۲} پایان، عملیات زمانی جاری را خاتمه داده و از نو زمان بندی می‌شوند.

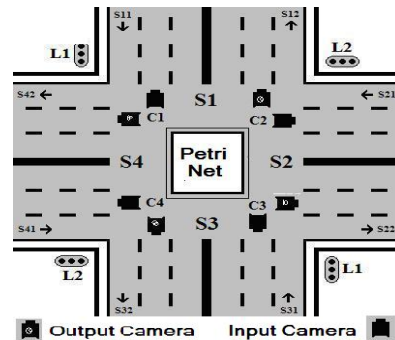
در حین عملیات تشخیص حجم خودروها یا بروز سانحه و یا خروجی اشغال شده، می‌توان سیگنال‌های پایان زمان هر فاز (چراغ‌ها) را به سیستم اعمال کرد.

چهارراه ایزوله (چهار فاز)^{۱۶} و محل دوربین‌های ورودی و خروجی سیستم (مطابق شکل ۳) مدل شده و به شرح شکل ۴ فازهای آن ارائه گردیده است. ساختار کنترل و ارتباط خطوط ورودی و خروجی به چهارراه و مسیر حرکت خودروها توسط یک مدل جدید چندلایه، آمیخته مدل چند لایه پتری نت زمانی^{۱۷} و متاپتری نت^{۱۸} (براساس متانتورک‌های منطقی^{۱۹}) به نام پتری نت ترکیبی^{۲۰} مدل شده است (شکل ۵). مدل اصلی چراغ چهارفاز، براساس چند لایه در شکل ۶ آمده است [۱۶، ۱۸].

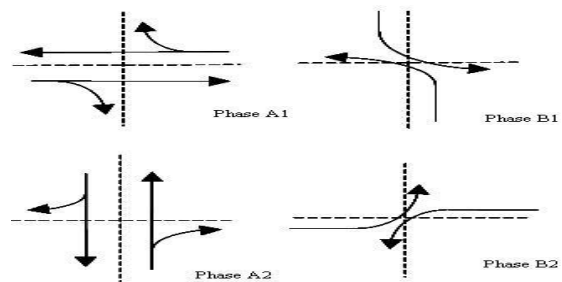


شکل ۲- فرآیند عملیات سیستم

سطح متای کنترلی، یک پتری سنتی و لایه پایه آن می‌تواند سنتی، رنگ‌دار، زمانی و غیره باشد. سیستم، سیگنال‌های سبز^{۳۱} و قرمز^{۳۲}، مدت زمان تقریبی (چراغ‌های) هر فاز و چشم‌ک‌زن بودن با رنگ زرد و قرمز را تعیین می‌کند. ضمناً بنابر دلایلی (مانند سانحه و خروجی پر) می‌تواند سیگنال پایان این زمان‌ها را زودتر از موعد مقرر به مدل ارسال کند تا برای ایفاء نقش بلادرنگ بودن، اصلاح زمان بندی داشته باشد.

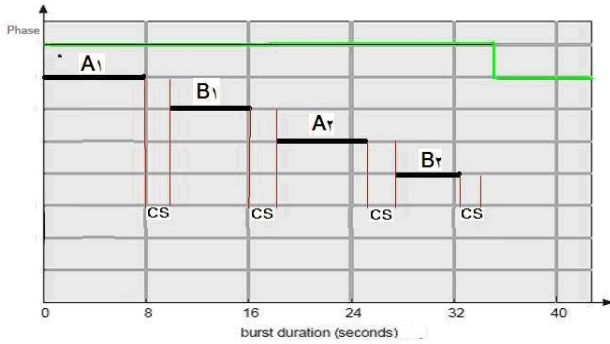


شکل ۳- چهارراه ایزوله و موقعیت دوربین‌ها



شکل ۴- مسیرهای حرکت چهارراه نوع عالی چهارفاز

در ترافیک چهارراه دو زمان، یعنی زمان هر چراغ و زمان تعویض آن چراغ محدود کننده هم دیگر هستند. اگر سرباری زمان تعویض زیاد باشد ممکن است ترافیک چهارراه با بن بست روبرو شود. زمان حداقل چراغ یا بین چراغ‌ها، برای غلبه بر بن بست ارائه شده است. در این حالت زمان چراغ باید از زمان تعویض چراغ (که برای حرکت از صفر مناسب است و محاسبه شده) بیشتر یا مساوی باشد. در صورت کاهش این زمان در فازهای A_i و B_i در دراز مدت مشکل داریم که با سیگنال‌های پایان، آنرا تقلیل می‌دهیم.



شکل ۸- نمونه دوره عملیاتی، زمان فازها و تعویض چراغ‌ها

$$T_{TurnAroundTime_i} \cong \sum_{i=1}^2 (T_{A_i} + T_{B_i}) + T_{ContextSwitching_i} \quad (1)$$

$$0 \leq T_{TurnAroundTime_i} \leq T_{Social} \quad (2)$$

$$4 \times T_{ContextSwitching_{Phase_i}} \leq T_{TurnAroundTime} \leq T_{Social} \quad (3)$$

$$T_{ContextSwitching_{Phase_i}} \geq ((T_{Processing} + T_{Command})/4) \quad (4)$$

$$T_{TurnAroundTime} \geq \left(\sum_{i=1}^2 (T'_{A_i} + T'_{B_i}) \right)$$

$$T'_{A_i} = T_{A_i} + T_{ContextSwitching_{A_i}}$$

$$T'_{B_i} = T_{B_i} + T_{ContextSwitching_{B_i}}$$

روابط زمان فازهای مختلف در قبل آمده‌اند. حداقل زمان تعویض $T_{MinContextSwitching}$ وابسته به کمترین زمان لازم برای شروع حرکت خودروها از حالت سکون، و حداقل تعداد خودروهای عبوری از چهارراه $T_{StartTime}$ (بر حسب طول مجموع خودروها بر متر)، با سرعت مورد انتظار $N_{NoOfVehicles}$ (متر بر ثانیه) است. به تجربه، با فرض طول متوسط خودرو ۷.۵ متر (با احتساب فواصل جلو و عقب) و حداقل عبور ۶ ردیف خودرو با سرعت 14.4 km/h (یا 4 m/s)، حداقل زمان آن ۱۱.۲۵ ثانیه بدست آمده است و نباید از ۱۸٪ حجم اسمی آن فاز کمتر باشد (رابطه ۵). در حالت چراغ چشمک‌زن مساله فرق می‌کند.

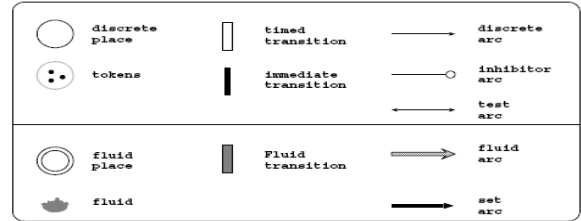
$$T_{MinContextSwitching} = T_{StartTime} + N_{NoOfVehicles} / W_{Speed}$$

$$T_{ContextSwitching} = (V_{vehicleLeng_{thAvg}} \times 6) / S_{PeedAvg} \quad (5)$$

$$= 7.5 \text{ m} \times 6 / 4 \text{ m/s} = 11.25 \text{ s}$$

$$T_{capacity_{Phase_i}} \times 0.18 \leq T_{ContextSwitching_{Phase_i}}$$

در این شبکه با مقاردهی اولیه مدل، فازهای A_1, B_1, A_2, B_2 مطابق مدل و به ترتیب اجراء می‌شوند (شکل ۴). در شکل ۶ مکان‌های G_{A_i} موید سیگنال سبز فاز A_i, R_{A_i} سیگنال قرمز فاز A_i و گذارهای eg_{A_i} پایان سیگنال سبز فاز A_i, er_{A_i} پایان سیگنال قرمز فاز A_i هستند و برای فاز B_i نیز چنین است. عناصر مدل در شکل ۷ آمده است [۲۰، ۲۴].



شکل ۷- عناصر مدل پترنت ترکیبی [۱۹]

از آنجا که تعیین رنگ چراغ تنها برای خودروهای وارد شونده به چهارراه اهمیت دارد، در هر خیابان منتهی به چهارراه، بخش ورودی به آن چهارراه در نظر گرفته و با نصب یک دوربین بر روی هر ورودی، ترافیک خودروهای در حال حرکت به سمت دوربین یا دبی ورودی محاسبه می‌شود (شکل ۳). به دلیل نیاز به برآورد مناسب باید برای کلیه خروجی‌ها نیز دوربین فرض کرد تا از سرریزی^{۳۳} در خروجی پر و ایجاد بن بست^{۳۴} جلوگیری گردد (البته می‌توان از همان دوربین‌های ورودی بهره برد).

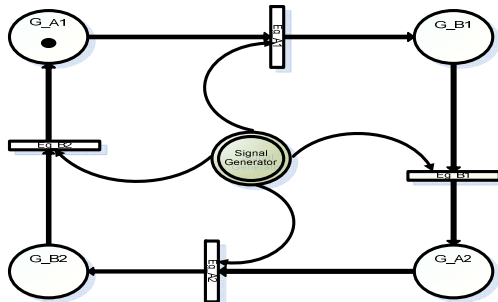
ترافیک ورودی، به تعداد لاین‌های^{۳۵} تشکیل دهنده خیابان، تقسیم‌بندی شده و حجم ترافیک در هر لاین بصورت جداگانه محاسبه و ارائه می‌شود. سپس با توجه به دوربین‌های خروجی، محاسبات تخلیه یا دبی خروجی را نیز خواهد داشت تا با توجه به آن از بروز حوادث احتمالی جلوگیری گردد و از جریان خروجی مناسب بهره برد.

۴- زمان بندی ترافیک

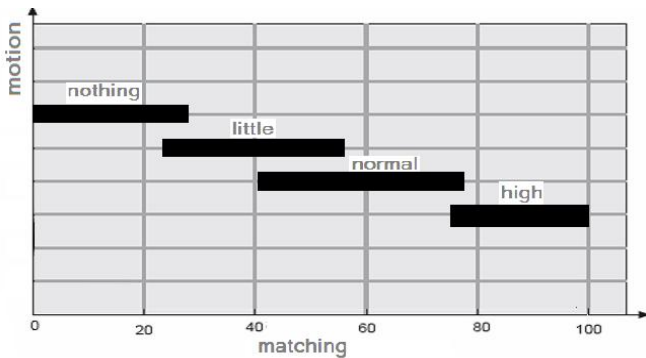
در دنیای واقعی بدلیل نیاز به تغییر زمان بندی یا ساختار چراغ‌ها، هیچگاه نمی‌توان یک چراغ با زمان بندی ثابت داشت. برای مدل سازی این سیستم ابتدا یک دوره عملیاتی^{۳۶} بنام $T_{TurnAroundTime_i}$ تعریف می‌گردد که با ایجاد حوادثی^{۳۷} مانند تصادف، زودتر از اتمام دوره زمانی مقرر، تغییر می‌کند. این زمان همیشه ثابت نیست و با عوامل اصلی تغییر ترافیک مانند آهنگ حرکت، محاسبه مجدد آن لازم است. این زمان عبارت از مجموع زمان‌های مربوط به هر کدام از فازها و زمان لازم بین فازها برای حرکت^{۳۸} و عبور خودروها از آن فاز بنام تعویض فازها یا سویچ بین آنها ($T_{ContextSwitching_i}$) می‌باشد (رابطه ۱).

این دوره زمانی، با زمان پاسخ سیستم برای رانندگان متناسب است و نمی‌تواند از حوصله آنها یعنی T_{Social} خارج باشد. پس این متغیر، زمانی از یک پارامتر اجتماعی^{۳۹} متناسب با میزان آرامش و انتظار (جامعه استفاده کننده آن) است لذا در بدترین حالت مقدار غیر صفر دارد (رابطه ۲). انتظار می‌رود بسمت صفر میل کند یعنی چراغ همیشه سبز، این امر ممکن نیست زیرا هیچ عبور و مروری از چهارراه (یا تقاطع) انجام نمی‌شود. حتی در چهارراه خالی نیز زمان تعویض چراغ‌ها وجود دارد. مینیمم میزان این پارامتر وابسته به تکنولوژی چراغ است و مطابق شکل ۸، چهار واحد از این زمان تعویض فازها را در مدل داریم (مطابق شکل ۸ و رابطه ۳). لذا این زمان نمی‌تواند از زمان لازم برای محاسبه ترافیک موجود، پردازش تصاویر، صدور فرامین و ارسال آن به چراغ کمتر باشد (رابطه ۴).

سیستم را کنترل کند. این وقفه‌ها عبارتند از: ۱- **حادثه**: رخداد اتفاقی مانند تصادف که باید زمان‌بندی تغییر کند. ۲- **تغییر آهنگ حرکت**: افزایش یا کاهش ناگهانی سرعت خودروها که باعث تغییرات زمانی لازم می‌شود و اجتناب‌ناپذیر است. ۳- **روز بدون ترافیک**: حجم ترافیک قابل توجه نیست و سیستم دامنه زمانی حداقل را اعمال می‌کند و در بعضی مواقع حالت چشم‌گزن (زرد و قرمز) دارد.



شکل ۹- مولد سیگنال وقفه برای لایه مدل چراغ



شکل ۱۰- تصمیم‌گیری فازی یافتن آهنگ حرکت ترافیک

سیستم دانش حرکت‌های سریع، متعادل، کند و توقف کامل ترافیک را براساس تغییرات در فریم‌های متوالی به صورت فازی کسب می‌کند. سیستم با تفاضل فریم‌های متوالی، در چهار توالی، به‌آهنگ حرکت ترافیک هرلین یا هر فاز پی می‌برد. حالت‌ها عبارت از توقف کامل **nothing**، حرکت‌کند **little**، حرکت متعادل **normal** و حرکت سریع **high** مطابق شکل ۱۰ است. در رابطه ۹ بعضی قوانین صدور فرمان یا ژتون پایان سیگنال‌های تنظیم زمانی مربوط به مدل آمده است.

۷- تصویر زمینه

در سیستم پیشنهادی، از ترکیب روش‌های پردازش تصویر و تحلیل تصاویر متوالی ترافیکی استفاده شده و اصلاحاتی در تکنیک‌های موجود و استفاده از الگوریتم‌های جدید نیز صورت گرفته است. استفاده از تصویر زمینه در سیستم‌هایی که از روش زمینه برای تشخیص و جداسازی حرکت استفاده می‌کنند، کاملاً ضروری است [۲۵].

$$R_{A1} = R_{11} + R_{31} \quad (۸)$$

$$R_{A2} = R_{21} + R_{41}$$

$$R_{B1} = R_{11B1} + R_{31B1}$$

$$R_{B2} = R_{21B2} + R_{41B2}$$

زمان انتظار راننده نمی‌تواند از زمان کارسیستم کمتر باشد. عواملی مانند زمان اخذ تصویر، پردازش، تشخیص و صدور فرمان به چراغ در این امر مؤثرند و از بعضی آنها بدلیل کم بودن می‌توان صرف‌نظر کرد. زمان انتظار راننده تابعی از تکنولوژی مورد استفاده در سخت‌افزار، نرم‌افزار، دوربین، و کیفیت تصاویر است (رابطه ۶).

$$T_{Social} = T_{Acquisition} + T_{Processing} + T_{Command} \quad (۶)$$

$$T_{Social} \subseteq T_{Technology}$$

نهایتاً زمان هر فاز چراغ T_{Phase_i} ، به میزان β_i درصد، از زمان $T_{TurnAroundTime}$ است. β_i همان نسبت حجم خودرو R_{Phase_i} به حجم اسمی $R_{Phase_iCapacity}$ آن فاز (رابطه ۷) است.

$$T_{Phase_i} \approx \beta_i \times T_{TurnAroundTime} \quad (۷)$$

$$\beta_i = \left(R_{Phase_i} / R_{Phase_iCapacity} \right)$$

۵- حجم ترافیک چهارراه

در مدل ارائه شده خیابان‌های مرتبط، مانند یک خیابان واحد در یک چهارراه تلقی شده و دارای چراغ راهنمایی مشترک هستند. حجم ورودی‌ها محاسبه و با مقادیر متناظر خیابان متقاطع مقایسه می‌گردد. با توجه به شکل ۳، تصاویر هر یک از خیابان‌های ورودی S_{41A1} ، S_{31A2} ، S_{21A1} ، S_{11A2} ، S_{41A1} ، S_{31A2} ، S_{21A1} ، S_{11A2} و جداگانه توسط دوربین‌های C_4 ، C_3 ، C_1 ، C_4 ، C_3 ، C_2 ، C_1 و C_2 ارسال و پس از محاسبه حجم ترافیک، مجموع حجم‌های S_{31A2} و S_{11A2} فاز $A2$ ، مجموع حجم‌های S_{41A1} و S_{21A1} فاز $A1$ ، مجموع لاین‌های $S_{11Lane-m}$ و $S_{31Lane-m}$ فاز $B1$ و مجموع لاین‌های $S_{21Lane-m}$ و $S_{41Lane-m}$ فاز $B2$ بررسی می‌گردد تا نهایتاً وضعیت چراغ‌های L_1 مربوط به خیابان S_1S_3 ، L_2 مربوط به خیابان S_2S_4 ، چراغ $L_3 \equiv L_1$ مربوط به جهت‌های S_1S_2 و S_3S_4 فاز $B1$ و چراغ $L_4 \equiv L_2$ مربوط به جهت‌های S_4S_1 و S_2S_3 فاز $B2$ مشخص گردد.

با فرض آنکه حجم خودرو مربوط به ورودی‌های S_{41} ، S_{31} ، S_{21} ، S_{11} ، $S_{41Lane-m}$ ، $S_{31Lane-m}$ ، $S_{21Lane-m}$ ، $S_{11Lane-m}$ به ترتیب R_{41} ، R_{31} ، R_{21} ، R_{11} ، R_{41B2} ، R_{31B2} ، R_{21B2} و R_{11B2} باشد، حجم ترافیک هر فاز $(Phase_A \vee Phase_B)$ مطابق رابطه ۸ است.

در رابطه ۸، R_{A1} حجم ترافیک دو ورودی روبروی هم S_{31} و S_{11} (یعنی خیابان L_1 چهارراه)، R_{A2} حجم ترافیک دو ورودی روبروی هم S_{41} و S_{21} (یعنی خیابان L_2 چهارراه)، R_{B1} حجم ترافیک دو لاین ورودی فاز $B1$ و R_{B2} حجم ترافیک دو لاین ورودی فاز $B2$ می‌باشند. رنگ چراغ‌های L_1 و L_2 برای فازهای مختلف از مدل تعیین شده و با توجه به حجم ترافیک و طبق مدل شکل ۶ ادامه می‌یابد.

۶- زمان بندی بلادرنگ

در اینجا، لزوم تغییر در زمان محاسبات، به‌عنوان سیگنال‌های وقفه تلقی می‌شود. کنترلر بلادرنگ با داشتن پایگاه دانش سیستم، توسط مولد سیگنال (شکل ۹) به تولید سیگنال وقفه مربوطه می‌پردازد تا در صورت بروز آنها، با تولید ژتون پایانی،

این حالت روشنایی اشیاء و سطوح بطور همزمان تغییر می‌کند. ۳- سایه‌ها^{۴۴} توسط اشیاء زمینه که بدون حرکتند و یا اشیاء متحرک که متعلق به زمینه نیستند، ایجاد می‌شوند. این تغییرات قسمتی از محیط را تحت تأثیر قرار می‌دهند. همه سطوح واقع در سایه، کاهش روشنایی داشته و با توجه به جهت و فاصله نسبت به منبع نور تغییرات آرام روشنایی خواهند داشت.

ب- تغییرات پویا: این نوع تغییرات در اثر عواملی غیر از روشنایی و نور محیط ایجاد شده و عبارتند از: ۱- حرکت در بخشی از زمینه که ممکن است در اثر حرکت برگ درختان به علت وزش باد یا حرکت امواج آب و یا برف و باران ایجاد شود یعنی روشنایی بخش‌هایی از تصویر بصورت تصادفی تغییر کند. ۲- حرکت سراسری که بر اثر تغییر کوچکی در موقعیت دوربین بر اثر وزش باد یا حرکت وسایل نقلیه سنگین یا باران که در محیط‌های آزاد اتفاق می‌افتد. ۳- تغییرات فیزیکی شامل هر تغییری در ظاهر محیط و زمینه بر اثر تغییر موقعیت اشیائی که قبلاً متعلق به زمینه بوده‌اند است [۱۴، ۲۶، ۲۸].

۷-۲- تصاویر زمینه

روش پیشنهادی ساخت و اصلاح تصویر زمینه، مدتی نوین، ترکیبی از روش تفاضل پویا^{۴۳} و بکارگیری الگوریتم‌های جدید (مطابق شکل ۱۱) است [۲۹، ۳۰]. پس از ساخت تصویر زمینه، این تصویر به عنوان مرجع، ذخیره و سپس محاسبه‌گر خودکار حجم تردد، از این تصویر استفاده و عمل اصلاح تصویر زمینه را (که بیشتر برای هماهنگی با تغییرات روشنایی و حرکات جزئی زمینه در طول مدت استفاده از سیستم است) صورت می‌دهد. ورودی آن، دنباله تصاویر متوالی از یک خیابان منتهی به چهارراه و خروجی آن نیز تصویر زمینه مربوط به آن ورودی است.

از تفاضل تصویر فریم جاری و زمینه جاری و آستانه‌گیری از نتیجه تفاضل، یک مقدار باینری بعنوان نقاب شیء^{۴۴} (مطابق رابطه ۱۰) بدست می‌آید. از تفاضل فریم جاری و فریم قبلی و آستانه‌گیری^{۴۵} از نتیجه تفاضل، نقاب حرکت^{۴۶} (مطابق رابطه ۱۱) بدست می‌آید. در روابط ۱۰ و ۱۱، i و j مختصات پیکسل، CF تصویر فریم جاری، CB تصویر زمینه جاری، PF تصویر فریم قبلی و T مقدار آستانه می‌باشد. با AND کردن نقاب‌های شیء و حرکت، نقاب نهایی بدست می‌آید (رابطه ۱۲) که این نقاب می‌تواند (مطابق رابطه ۱۳) تصویر زمینه موقت را بدست آورد. IB تصویر زمینه موقت است و تصویر زمینه نهایی مطابق رابطه ۱۴ بدست می‌آید.

در رابطه ۱۴، IB تصویر زمینه موقتی، CB در سمت راست رابطه، تصویر زمینه جاری مرحله قبل و CB در سمت چپ رابطه، تصویر جدید زمینه جاری می‌باشد. متغیر α یک پارامتر یادگیری^{۴۷} است که تأثیر زیادی در نتیجه نهایی دارد. α می‌تواند به‌عنوان یک مقدار ثابت و بدون تغییر مورد استفاده قرار گیرد اما ثابت بودن آن مشکلاتی ایجاد می‌کند. مثلاً مقادیر بزرگتر α ، تغییرات را به‌سرعت منعکس می‌کند، بنابراین ممکن است پس از حصول یک تصویر زمینه خوب، با عبور یک شیء یا خودروی دارای سرعت کم، تأثیر آن در زمینه برای مدتی به‌نمایش درآید. از طرفی مقدار کم α ، تغییرات سریع مانند تغییر ناگهانی نور محیط، مدت زمان و تعداد فریم‌های بیشتری را برای تطابق زمینه طلب می‌کند ولی در تصویر نهایی زمینه، کیفیت بهتری ارائه می‌دهد [۳۱]. در تعیین مقدار α ، باید تناسبی بین کیفیت بهتر و سرعت بیشتر برقرار گردد. با استفاده از تصاویر ورودی در شرایط مختلف به رابطه‌ای برای α دست یافته‌ایم (رابطه ۱۷). وقتی شرایط نور محیط به سرعت تغییر می‌کند، مقدار α بزرگتر شده و در مواقع فقدان تغییر سریع نور، مقدار α کوچکتر می‌شود. با استفاده از تغییرات نور، یک مقدار قابل تطبیق برای α تعیین کرده‌ایم. میانگینی از فریم جاری و فریم قبلی برای محاسبه α بکار می‌رود.

یک راه‌حل ساده، استفاده از تصویر اولیه ذخیره شده با زمینه ثابت و از قبل تهیه شده (در محیط خیابان بدون وجود اشیاء متحرک مانند خودرو) است. گرچه این روش در محیط‌های بسته و بدون تغییرات نور کاملاً جواب می‌دهد، اما در محیط‌های باز و در معرض نور خورشید با مشکل مواجه است و تصویر زمینه ثابت (بدون تغییر و اصلاح) کاربرد ندارد. مشکل دیگر گرفتن تصویر بدون حضور اشیاء متحرک در محیط‌های شلوغ خصوصاً خیابان می‌باشد که بعضاً غیرممکن است.

RL1 :

$$\text{if } (R_{Ai} < R_{Bi}) \& (R_{Ai} \rightarrow 0 \vee R_{Bi} \rightarrow 1) \& (\text{State} = G_Ai) \\ \text{:: EndGreenAi} \approx \text{eg_Ai}$$

RL2 :

$$\text{if } (R_{Ai} > R_{Bi}) \& (R_{Ai} \rightarrow 1 \vee R_{Bi} \rightarrow 0) \& (\text{State} = G_Bi) \\ \text{:: EndGreenBi} \approx \text{eg_Bi}$$

RL3 :

$$\text{if } (R_{A1} < R_{A2}) \& (R_{A1} \rightarrow 0 \vee R_{A2} \rightarrow 1) \& (\text{State} = G_A1) \\ \text{:: EndGreenA1} \approx \text{eg_A1}$$

RL4 :

$$\text{if } (R_{A1} > R_{A2}) \& (R_{A2} \rightarrow 1 \vee R_{A1} \rightarrow 0) \& (\text{State} = G_A2) \\ \text{:: EndGreenA2} \approx \text{eg_A2}$$

(۹)

لذا چاره‌ای جز ساختن بلادرنگ تصویر زمینه از تصاویر ورودی اشیاء و اجزاء متحرک و بروز کردن مداوم آن نیست. البته پیاده‌سازی این روش با مشکلاتی همراه است ولی بدلیل مذکور، تنها راه حل موجود می‌باشد که سیستم پیشنهادی آن استفاده کرده است. حصول تصویر کیفی زمینه دارای محدودیت‌هایی است و باید:

- ۱- تعداد فریم‌های مورد استفاده در ساخت تصویر مرجع زیاد باشد.
- ۲- خودروها با فاصله از هم حرکت کنند و مدت زمان زیادی توقف نداشته باشند.
- ۳- خودروها و اشیاء متحرک با رنگهای متفاوت و گوناگون در قسمت‌های مختلف تصاویر ظاهر شوند.
- ۴- بعلاوه در محیط‌های بسته و باز همواره تغییراتی در محیط ایجاد می‌شود که می‌توان از آنها به عنوان تغییرات زمینه نام برد، این تغییرات می‌توانند محلی باشند و بخشی از محیط را شامل شوند، یا سراسری بوده و کل تصویر محیط را تحت تأثیر قرار دهند. مدل پیشنهادی ارائه شده می‌تواند این تغییرات را تحمل کند.

۷-۱- تغییرات در زمینه

عمدتاً دو نوع تغییر در محیط یا زمینه ممکن است رخ دهد: الف- تغییرات نور^{۴۰} یا روشنایی و ب- تغییرات پویا^{۴۱} [۲۵]. در محیط‌های عملی،

الف- تغییرات نور: یا روشنایی همواره به سه صورت در محیط اتفاق می‌افتد که عبارتند از: ۱- تغییرات تدریجی نور (در محیط‌های آزاد یا بسته زیر سقف) به علت تغییر موقعیت خورشید ایجاد و به عنوان تغییر کلی در نظر گرفته می‌شوند. نتیجه تغییر نور، تغییر سطح روشنایی در اشیاء و سطوح در معرض روشنایی است و بصورت آرام و تدریجی خواهد بود. این تغییر نور از نوع سراسری است. ۲- تغییرات ناگهانی نور می‌تواند در محیط‌های بسته مثلاً با روشن یا خاموش کردن لامپ و در محیط‌های باز با عبور یک ابرتیره از محیط اتفاق بیفتد. در

$$\text{Area_Unmoved} = \sum_{i,j} (1 - \text{Motion_Mask}(i, j)) \quad (15)$$

$$\text{Sum_Unmoved} = \sum_{i,j} |\text{CF}(i, j) - \text{PF}(i, j)| / 256 \quad (16)$$

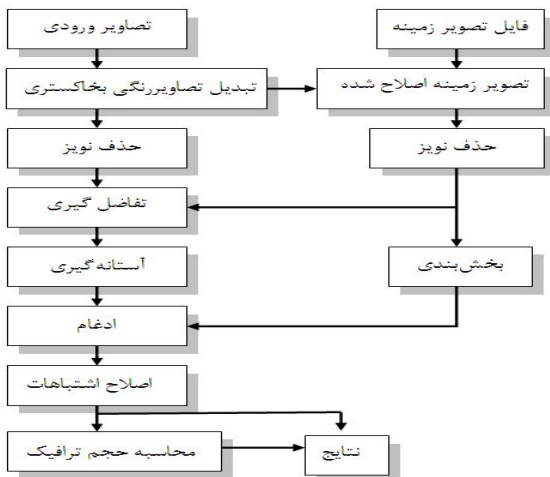
$$\text{Motion_Mask}(i, j) = 0$$

$$\beta = \frac{\text{Sum_Unmoved}}{\text{Area_Unmoved}} \quad (17)$$

$$\alpha = 0.9\alpha + 0.1\beta$$

۸- محاسبه حجم ترافیک

الگوریتم محاسبه حجم ترافیک مطابق شکل ۱۳ است. ورودی‌های سیستم چهار رشته تصویر متوالی جداگانه و چهار تصویر زمینه متناظر با رشته‌های تصاویر ورودی می‌باشد. خروجی آن نیز نمایش حجم ترافیک بدست آمده هر یک از چهار ورودی^{۴۹} و چهار خروجی^{۵۰} بطور جداگانه (به تفکیک لاین‌ها) و خیابان‌های ایجاد کننده چهارراه است.



شکل ۱۳- الگوریتم محاسبه حجم ترافیک

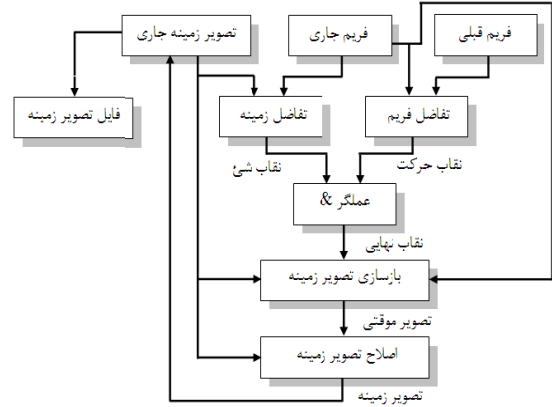


شکل ۱۴- چهار ورودی به چهارراه به شعاع دید متفاوت

در اولین مرحله، چهار سری تصویر از چهار ورودی به چهارراه (مانند شکل ۱۴ با میادین دید ۳۰، ۴۵، ۶۰ و ۱۰۰ متر)، به سیستم وارد می‌شوند. سپس چهار تصویر زمینه مربوط به ورودی‌ها (بدون وجود خودروها و اشیاء متحرک) و قبلاً ساخته

$$\text{Object_Mask}(i, j) = \begin{cases} 1 & \text{if } |\text{CF}(i, j) - \text{CB}(i, j)| > T \\ 0 & \text{else} \end{cases} \quad (10)$$

$$\text{Motion_Mask}(i, j) = \begin{cases} 1 & \text{if } |\text{CF}(i, j) - \text{PF}(i, j)| > T \\ 0 & \text{else} \end{cases} \quad (11)$$



شکل ۱۱- اصلاح مستقیم تصویر زمینه

$$\text{Mask}(i, j) = \text{Motion_Mask}(i, j) \& \text{Object_Mask}(i, j) \quad (12)$$

$$\text{IB}(i, j) = \begin{cases} \text{CF}(i, j) & \text{if } (\text{Mask}(i, j) = 0) \\ \text{CB}(i, j) & \text{if } (\text{Mask}(i, j) = 1) \end{cases} \quad (13)$$

$$\text{CB}(i, j) = \alpha \cdot \text{IB}(i, j) + (1 - \alpha) \cdot \text{CB}(i, j) \quad (14)$$

خودروهای در حال حرکت نباید به‌عنوان تغییرات نور در نظر گرفته شوند لذا میانگین‌گیری در قسمت‌های غیرمتحرک انجام می‌گیرد (روابط ۱۵ تا ۱۷). در روابط ۱۵ تا ۱۷، Area_Unmoved مشخص‌کننده بخش‌های بدون حرکت بین فریم جاری و فریم قبلی، Sum_Unmoved مجموع نرمال شده قدرمطلق اختلاف سطح روشنایی فریم جاری و فریم قبلی در نقاط بدون حرکت تصویر، و β از تقسیم این دو مقدار بدست می‌آید. بنابر این تغییرات سریع در روشنایی محیط بین دو فریم متوالی با توجه به رابطه ۹، مقادیر بزرگتری از β را تولید می‌نماید که نتیجه آن α بزرگتر است. ثابتهای ۰/۱ و ۰/۹ حساسیت مدل زمینه^{۴۸} به تغییرات نور را کنترل می‌کنند و به تجربه بدست آمده‌اند [۳۲، ۳۳]. شکل ۱۲ نمایش نتیجه روش و تصاویر حاصل پس از به‌ترتیب ۵، ۱۰ و ۱۵ فریم ورودی است که افزایش تعداد فریم‌ها موجب کیفیت بهتر تصویر زمینه است.



شکل ۱۲- تصویر زمینه حاصل (الف) بعد ۱ فریم، (ب) بعد ۵ فریم، (پ) بعد ۱۰ فریم، (ت) بعد ۱۵ فریم

(مانند شدت نور و انباشتگی آن) به مرور زمان، نتایج نادرستی را تولید نماید. کاهش اثر نویز با اعمال نقاب 3×3 رابطه ۲۱ است (شکل ۱۶) [۳۴، ۳۵].

$$R = 1/9 \sum_{i=1}^9 Z_i \quad (21)$$

۸-۲- بخش‌بندی تصویر

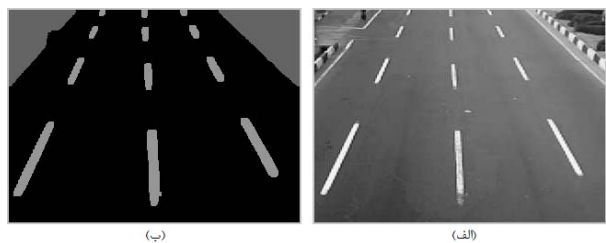
از این پس، عملیات بصورت موازی در دو بخش اصلاح زمینه و محاسبه ترافیک پیش می‌رود. یک بخش، بخش‌بندی تصویر است که روی تصویر زمینه اعمال می‌شود (و هدف اینکار جداکردن محدوده سمت چپ خیابان از پیاده‌رو و قسمت‌های غیرمرتبط مانند جداول و سمت راست خیابان است) [۳۶، ۴۰]. تنها خودروهایی که به چهارراه وارد می‌شوند و جهت حرکت آنها بطرف دوربین است، برای محاسبه حجم ترافیک در نظر گرفته می‌شوند. لذا محدوده سمت چپ خیابان (یعنی مسیر عبور خودروها به سمت چهارراه) از بقیه قسمت‌ها جدا می‌گردد تا حجم پردازش کاهش یابد و اشیاء متحرک غیر از خودروهای مورد نظر (مانند عابرین پیاده‌رو یا خودروهای عبوری خلاف جهت چهارراه)، از محاسبات حذف شوند [۲۸، ۳۰، ۴۰].

برای بخش‌بندی تصویر از روش ناحیه‌گرایی استفاده شده است. البته یک روش کلی که برای همه انواع تصاویر درست عمل کند وجود ندارد و دیگر اینکه هیچ روشی بطور دقیق و عالی عمل نمی‌کند [۲۸، ۴۲]. سه ناحیه جداگانه در تصویر بدست می‌آید. ناحیه اول، محدوده خیابان ورودی به چهارراه است. ناحیه دوم، خطوط جداکننده لاین‌ها در وسط خیابان ورودی است. ناحیه سوم نیز بخش‌های غیرقابل پردازش در تصویر هستند که از محاسبات و اندازه‌گیری حذف می‌شوند. نتیجه بخش‌بندی تصویر زمینه در شکل ۱۷ قابل مشاهده است.



شکل ۱۶- (الف) تصویر زمینه قبل میانگین، (ب) بعد میانگین

از وظایف سیستم، تقسیم‌بندی خیابان ورودی به لاین‌های جداکننده و اندازه‌گیری حجم بطور جداگانه در هر لاین و نمایش آن است. حداکثر تعداد هفت لاین برای هر خیابان در نظر گرفته شده که در عمل نیز بیشتر از این تعداد مشاهده نمی‌شود.



شکل ۱۷- (الف) تصویر زمینه، (ب) تصویر بخش‌بندی شده

شده مانند شکل ۱۵ را بار می‌کند و محاسبه ترافیک را انجام می‌دهد (در ادامه نحوه پردازش تصاویر و محاسبه حجم ترافیک بیان و با سایر روش‌ها مقایسه می‌شود).



شکل ۱۵- تصاویر زمینه چهار ورودی سیستم

۸-۱- اصلاح تصویر زمینه

سیستم قادر است با دوربین‌های رنگی و تکرنگ کار کند. روش تبدیل تصاویر رنگی با فرمت RGB به سطح خاکستری، مدل YIQ می‌باشد. مؤلفه Y مدل، اطلاعات روشنایی و مؤلفه‌های I و Q، اطلاعات رنگ را شامل می‌شود (ماتریس رابطه ۱۸) و مقدار سطح هر پیکسل در مؤلفه Y مطابق رابطه ۱۹ محاسبه می‌گردد [۲۸].

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.29 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (18)$$

$$\text{Gray}(i, j) = 0.299 \times R(i, j) + 0.587 \times G(i, j) + 0.114 \times B(i, j) \quad (19)$$

برای افزایش کارایی، سرعت در پردازش و اجتناب از مشکلات قبلی، از روش آستانه‌گیری مطابق رابطه ۲۰ برای بازسازی و اصلاح تصویر زمینه استفاده می‌شود [۱۴].

$$\text{CB}(i, j) = \begin{cases} \text{CF}(i, j) & \text{if } |\text{CF}(i, j) - \text{CB}(i, j)| < T \\ \text{CB}(i, j) & \text{else} \end{cases} \quad (20)$$

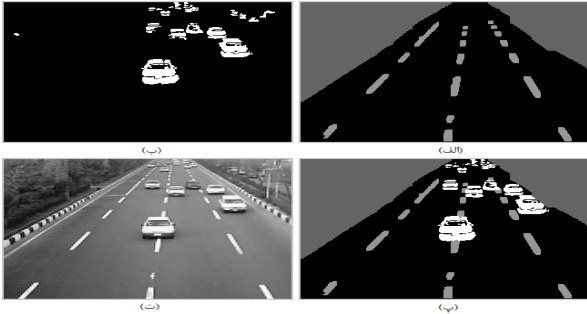
در رابطه ۲۰، CB تصویر زمینه، CF تصویر ورودی و T مقدار آستانه است. تغییرات کوچک مانند تغییرات آرام نور و سایه‌های اجسام ثابت در اثر تغییر جهت تابش خورشید، سرعت در تصویر زمینه منعکس شده و تصویر زمینه را به مرور زمان با این تغییرات هماهنگ و اصلاح می‌نماید. در این روش میزان تطابق با تغییرات ناگهانی، به مقدار T بستگی دارد. مقادیر بزرگ آستانه موجب می‌شود که تغییر مکان اجسام متحرک و خودروها در زمینه منعکس گردند و مورد نظر سیستم نیست.

همچنین مقادیر کوچکتر T، احتمال عدم تطابق با تغییرات ناگهانی نور را بیشتر می‌کند. این تغییرات در استفاده از دوربین‌های ارزان قیمت در طول مدت تصویربرداری بسیار اتفاق می‌افتد. با مقادیر کوچک آستانه، ممکن است تغییرات در تصویر زمینه ثبت نشده و پس از مدت طولانی و افزایش حجم تغییرات محیطی

۸-۳- تفاضل گیری

$$W = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad Z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 127 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

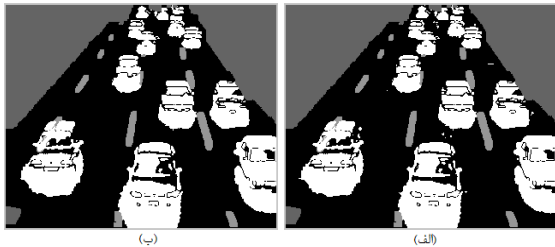
$$R = \sum_{i=1}^9 W_i Z_i \quad \text{if } (R = 0) \quad Z_5 = 0 \quad (26)$$



شکل ۱۹- (الف) تصویر بخش بندی زمينه، (ب) تصویر باينري ورودی، (پ) ادغام تصاویر (الف) و (ب)، (ت) تصویر اصلی

۹- اصلاح اشتباهات

بخشی از اشتباهات بر اثر ایجاد خطای سایه در تصویر خودرو و بخش دیگر خطای بخش بندی یا تفاضل است. این مرحله حاوی اصلاح اشتباهات و نقاطی است که به اشتباه در قسمت بخش بندی یا تفاضل در یکدسته قرار گرفته اند (رابطه ۲۶). برای حذف پیکسل های اشتباه از نقاب هایی با ابعاد ۳، ۴، ۵، ۶، ۷ استفاده شده تا بیشترین میزان اصلاح در خطاهای مراحل بخش بندی و تفاضل گیری صورت گیرد. در شکل ۲۰ نتایج بهتری حاصل شده ولی هنوز مشکل خطای حاصل از سایه حل نشده است.



شکل ۲۰- پیکسل های اشتباه (الف) اصلاح نشده، (ب) اصلاح شده

۹-۱- حذف سایه

سایه در تعریف متداول بعنوان تفاوت وضوح ثابت بین تصویر زمينه و تصویر موجود فعلی است. تعریف سایه در سیستم ارائه شده چراغ عبارت است از ۱- تعریف طبیعی سایه در روز که زمينه را (باضرب ω_d) تاریک تر می کند، و ۲- میزان تابش نور چراغ خودروها در شب که زمينه را (باضرب ω_n) روشن تر می کند. برای حذف سایه در این سیستم از روش های متداول استفاده نشده زیرا بیشتر آن روش ها با شرایط کنترل شده و فضای بسته کاربرد دارند و برای فضای باز دارای مشکل است [۴۶، ۴۵]. بعضی از آن روش ها براساس کانتور، رنگ و یا یافتن نقاط با ویژگی خاص است [۴۷، ۴۹].

در سیستم پیشنهادی از این ایده استفاده شده که "سایه اشیاء زمينه را از بین نمی برد و فقط رنگ آنها را تیره تر می سازد؛ وسط دو چراغ^{۵۵}، محور تقارن

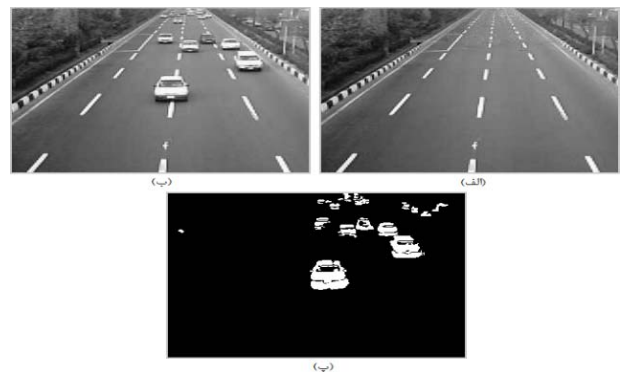
خروجی این بخش، تصویر حاصل از تفاضل^{۵۳} دو تصویر ورودی و زمينه مطابق رابطه ۲۲ است و در آن DF تصویر خروجی تفریق شده، CF تصویر ورودی و CB تصویر زمينه هستند [۲۸]. متد استخراج اشیاء از تصویر زمينه، آستانه گیری سراسری^{۵۴} است [۱۴]. مقدار آستانه، تجربی بدست آمده و نتایج خوبی را تولید می کند (رابطه ۲۳) [۴۳]. در رابطه ۲۳، BF تصویر باينري حاصل از آستانه گیری و DF تصویر خروجی از بخش تفاضل تصاویر ورودی و زمينه می باشد.

$$DF(i, j) = CF(i, j) - CB(i, j) \quad (22)$$

$$BF(i, j) = \begin{cases} 1 & \text{if } |DF(i, j)| > T \\ 0 & \text{if } |DF(i, j)| \leq T \end{cases} \quad (23)$$

$$BF(i, j) = \begin{cases} 1 & \text{if } |CF(i, j) - CB(i, j)| > T \\ 0 & \text{if } |CF(i, j) - CB(i, j)| \leq T \end{cases} \quad (24)$$

به این ترتیب اشیاء متحرک و خودروها، از زمينه و قسمت های بدون حرکت تصویر جدا می شود. عملاً برای کاهش حجم پردازش ها عمل تفاضل گیری و آستانه گیری با یکدیگر بصورت رابطه ۲۴ ادغام شده اند (نتیجه در شکل ۱۸). بدلیل وجود سایه در تصویر باينري خودرو در این مرحله، محاسبات نهایی حجم ترافیک دارای خطا خواهد بود و باید این خطا اصلاح شود [۴۴].



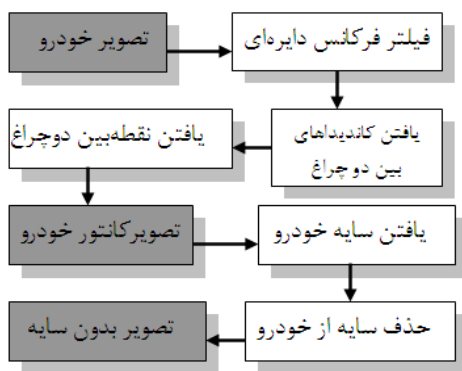
شکل ۱۸- (الف) تصویر زمينه، (ب) تصویر ورودی، (پ) تصویر باينري

۸-۴- ادغام

مرحله بعد ادغام دو تصویر حاصل از بخش تفاضل گیری شامل زمينه و تصویر زمينه بخش بندی شده برای بخش های غیرقابل پردازش، خطوط منقطع جداکننده لاین ها، و لاین ها می باشد. تصویر بدست آمده در قسمت تفاضل تصویر و آستانه گیری شامل همه اشیاء متحرک بود. در این مرحله خودروهایی که در خیابان ورودی به چهارراه و در جهت دوربين در سمت چپ در حرکتند از سایر اشیاء مانند عابرین پیاده در پیاده رو یا خودروهای سمت راست (طبق رابطه ۲۵) جدا می گردند (شکل ۱۹).

$$SF(i, j) = \begin{cases} BF(i, j) & \text{if } (BF(i, j) = 255 \ \& \ SB(i, j) \neq 100) \\ SB(i, j) & \text{else} \end{cases} \quad (25)$$

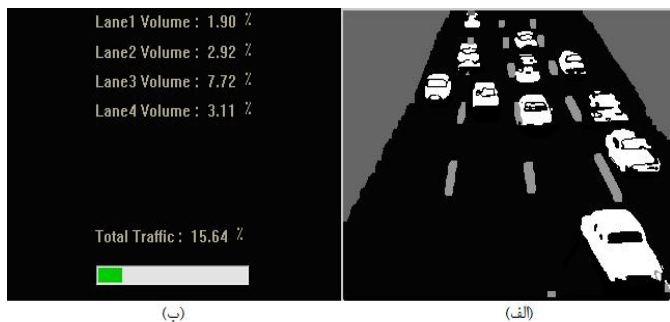
این تفاضل کمتر از حدی باشد حذف می‌شود. میزان سایه با طول و عرض خودرو نسبت دارد و در روش پیشنهادی این امر مهم نیست. تقریباً تمامی پیکسل‌های سایه حذف می‌گردد. برای حذف سایه نوع دوم یا نور چراغ، مشکل وجود دارد که این سایه به حجم تصویر خودرو افزوده می‌شود. فقط در فضای روبروی دوربین وجود دارد. آن را نمی‌توان به راحتی حذف کرد و خطای زیادی به محاسبات وارد نمی‌سازد.



شکل ۲۴- الگوریتم حذف سایه طرفین خودرو

۹-۲- حجم ترافیک

با داشتن تصویر بخش‌بندی شده ورودی و تصویر زمینه بخش‌بندی شده متناظر، که هر لاین در آن با یک سطح خاکستری علامت‌گذاری شده، می‌توان حجم کل ترافیک خودروهای ورودی به یک چهارراه را به تفکیک هر لاین محاسبه کرد [۴۲، ۱۴]. حجم ترافیک، نسبت حجم خودروها به سطح خیابان است. حجم ترافیک در هر لاین با شمارش تعداد پیکسل‌های آن که سطح روشنایی‌اش در تصویر بخش‌بندی شده زمینه مشخص است، به عنوان C_i و تعداد پیکسل‌های سطح خیابان به عنوان B_i و تقسیم C_i ها بر B_i ها محاسبه می‌شود (رابطه ۲۷). شکل ۲۵ نتایج یک ورودی شامل تصویر نهایی حاصل از ادغام، درصد حجم ترافیک هر لاین، درصد مجموع حجم لاین‌ها (یعنی حجم ترافیک در یک ورودی) همراه نوار گرافیکی پیشرفت را نشان می‌دهد.

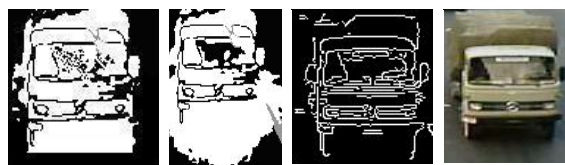


شکل ۲۵- (الف) تصویر خروجی، (ب) میزان ترافیک

$$\text{TrafficCondition} = \sum_{j=1}^3 \sum_{i=1}^{\text{Lanes}} \frac{C_i}{B_i} \quad (27)$$

خروجی سیستم هشت بخش دارد که نمایش هفت قسمت آن نتیجه محاسبات است و آخرین خروجی به عنوان سیگنال کنترل چراغ بکار می‌رود که عبارتند از: ۱- درصد حجم ترافیک هر ورودی و هر لاین بصورت جداگانه و به

خودرو است؛ تفاضل کمترین فاصله از طرفین این محور، مرز سایه را تعیین می‌کند و با بکارگیری آن سایه در طرفین حذف می‌شود؛ سایه جلوی خودرو با استفاده از تشخیص لبه جلوی خودرو و صرف‌نظر از سپرمجازی حذف می‌گردد. از این روش تاکنون استفاده نشده است. در زاویه دید دوربین، سایه خودروهای ورودی به کادر تصویر (ابتدای حوزه دید دوربین) از خودروهای خروجی از کادر (انتهای حوزه دید دوربین) سایه کمتری دارند. هرچه میدان دید دوربین وسیع‌تر باشد میزان سایه ایجاد می‌کند، و خطای تعیین محور تقارن بیشتر است. با اعمال فیلترکنی^۶ به تصویر خودرو، سایه به یک فضای غیر بسته (در خودروهای نزدیک دوربین) و فضای بسته جلوی ماشین (در خودروهای دور از دوربین) تبدیل شده که دارای تاریک‌ترین رنگ است (شکل ۲۱). رسم یک محور در راستای محور تقارن، از ابتداء تا انتهای تقاطع این خط با خودرو، طول خودرو را می‌دهد که: الف- کلیه اضافات طرفین این محور به عنوان سایه طرفین تلقی شده و حذف می‌گردد، ب- محل برخورد ابتدایی‌ترین نقطه خودرو با این محور برای حذف سایه جلوی خودرو بکار می‌رود. اگر خودرو نزدیک دوربین باشد بطور خودکار سایه از خودرو حذف شده است. سایه در تصاویر خودروهای دورتر باعث ایجاد سپر مجازی است (شکل ۲۱).



الف- خودرو ب- فیلترکنی ج- سایه د- حذف سایه

شکل ۲۱- حذف سایه جلوی خودرو

سپر مجازی، سایه جلوی خودرو تلقی و حذف می‌گردد. البته به دلیل آنکه در صورت فقدان سایه، سپر واقعی در تصویر سه بعدی، عمودی است لذا حذف آن خطایی به سیستم وارد نمی‌کند. بقای این بخش برای محاسبات، دارای خطا بوده و حذف آن صحیح‌تر است (شکل ۲۲). برای یافتن محور تقارن خودرو از فیلتر فرکانس دایره‌ای استفاده شده که پس از یافتن کاندیداها و بررسی آنها، مرکز دوچراغ بدست آمده است (شکل ۲۳) [۵۰، ۵۱].



شکل ۲۲- سایه جلو خودرو



الف- نقاط کاندیدا ب- محور بین چراغ‌ها ج- خودرو

شکل ۲۳- نقاط کاندیدا در فرکانس دایره‌ای و محور بین چراغ‌ها

اکنون (مطابق شکل ۲۴) عمود بر این محور و روی تصویر کانتور و در راستای خطوط لاین خیابان، کوتاه‌ترین فاصله از طرفین (چپ و راست و فقط در یک طرف) مرز تصویر خودرو تلقی می‌گردد و تفاضل آن با طرف مقابل مرز سایه است. اگر

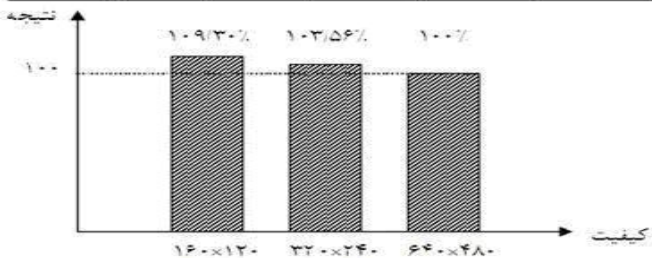
صحيح و با خطا در جدول ۵ آمده است. در جداول، ستون اول بیانگر درجه کیفیت تصویر است. ستون دوم، تعداد پیکسل‌های سطح خیابان ورودی به تقاطع می‌باشد. ستون سوم، تعداد پیکسل‌های سفید نشان‌دهنده خودروهای در حال حرکت به سمت چهارراه است. ستون چهارم، پیکسل‌های غیر قابل پردازش تصویر است که از نتایج نهایی حذف شده‌اند. ستون پنجم، میزان اشغال خیابان یا نسبت تعداد پیکسل‌های سفید به سطح خیابان است. در ستون آخر، نتایج برحسب تصویر درجه بالا و به صورت درصد محاسبه شده است.

چون تصاویر با درجه متوسط، نتایج خوب و قابل قبولی در مقایسه با تصاویر با کیفیت بالا ارائه می‌کنند و با توجه به اینکه کیفیت این دسته تصاویر، به تصاویر دوربین‌های ترافیکی نزدیکتر است، تصاویر با درجه متوسط به عنوان تصاویر پیش فرض در سیستم در نظر گرفته شدند. همچنان که در جدول ۱ قابل مشاهده می‌شود نتایج از ۱۰٪ بیشتر شده است. این امر به دلیل ورود تشخیص خطا به محاسبات می‌باشد و نتیجه حذف نادرست سایه در ترافیک بسیار فشرده است که بروز می‌کند و در دوربین‌های کم کیفیت بارز است.

جدول ۵ نتایج حاصل از تکرار آزمایشات را نشان می‌دهد. در این جدول نتایج تشخیص صحیح، تشخیص خطا، درصد تشخیص صحیح شی خودرو بطور مجزا و در ترافیک کم (خلوت)، ترافیک متوسط (معمولی) و ترافیک زیاد (شلوغ) را نشان می‌دهد. کاهش درجه ترافیک باعث افزایش دقت تشخیص خودروها بصورت مجزا است ولی به جهت آنکه در روش پیشنهادی اثری ندارد، درجه تشخیص تغییرات چندانی نمی‌کند.

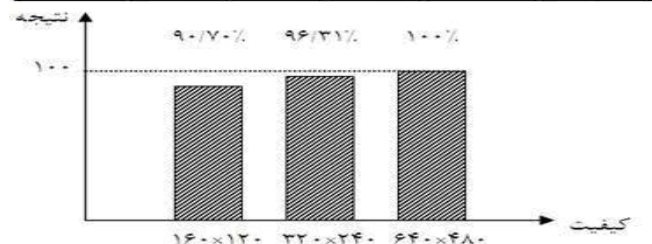
جدول ۱- ورودی ۱ و نتیجه آن

نتیجه	حجم	غیر قابل پردازش	پیکسل سفید	سطح خیابان	کیفیت
۱-۰٪	۲۱/۰۶٪	۳-۹۲۲	۹۶۵-	۴۵۸۳۱	۶۴۰×۴۸۰
۱-۳/۵۶٪	۲۱/۸۱٪	۳-۵۱۵	۱-۰۹۲	۴۶۲۶۵	۳۲۰×۲۴۰
۱-۹/۱۳٪	۲۲/۰۳٪	۲۷۶۸۱	۱۱۳-۷	۴۹۱۱۳	۱۶۰×۱۲۰



جدول ۲- ورودی ۲ و نتیجه آن

نتیجه	حجم	غیر قابل پردازش	پیکسل سفید	سطح خیابان	کیفیت
۱۰۰٪	۱۳/۰۲٪	۳۷۵۳۵	۵۱۱۵	۳۹۲۶۸	۶۴۰×۴۸۰
۹۶/۳۱٪	۱۲/۵۴٪	۳۴۶۷۱	۵۲۸۳	۴۲۱۲۵	۳۲۰×۲۴۰
۹۰/۷۰٪	۱۱/۸۱٪	۳۱۰۳۸	۵۴۰۵	۴۵۷۲۹	۱۶۰×۱۲۰



تفکیک تعداد لاین‌ها، ۲- درصد حجم کل ترافیک در هر چهار ورودی، ۳- میزان حجم کل ترافیک به صورت نوار پیشرفت برای هر چهار ورودی، ۴- درصد حجم ترافیک هر خیابان گذرنده از چهارراه برای خیابان‌های متقاطع، ۵- میزان حجم ترافیک در هر خیابان گذرنده از چهارراه به صورت نوار سه رنگ قرمز، سبز و زرد برای هر دو خیابان، ۶- تصویر ورودی بصورت بخش‌بندی شده و تمایز خودروها در محدوده ورودی به چهارراه برای هر چهار ورودی، ۷- سرعت متوسط خودروها با تعداد آنها و فاصله نسبی خودروها در هر لاین، و ۸- تولید فرامین کنترل مدل چراغ برای سوئیچ بین چراغ‌های چهارراه. نتایج چهار ورودی چهارراه و در واقع خروجی کامل و نهایی سیستم در شکل ۲۶ آمده است.



شکل ۲۶- حجم ترافیک نهایی و لاین‌های ورودی‌های چهارراه

۱۰- نتایج

این سیستم با زبان برنامه‌نویسی ++C پیاده‌سازی شده و با چهار رشته تصویر از چهار خیابان منتهی به چهارراه، هر یک شامل ۱۰۰ فریم با سه درجه تفکیک مختلف روی یک کامپیوتر Pentium IV-1.2GHZ با حافظه RAM 128MB و چهار ورودی همزمان تست گردیده است. سیستم ۱۰ فریم بر ۳ ثانیه برای هر دوربین را پردازش می‌کند و مدل پتری کنترلر چراغ تحت ابزار خود و در غالب یک کد اجرایی اجرا می‌شود.

نمونه: نتایج از نظر کیفیت تصاویر و تشخیص در ترافیک متفاوت با استفاده از تصاویر با کیفیت‌های پیکسل مختلف ۶۴۰×۴۸۰، بالا، ۳۲۰×۲۴۰ متوسط و ۱۶۰×۱۲۰ ضعیف با میدان دیدهای ۳۰، ۴۵، ۶۰ و ۱۰۰ متر بدست آمده که نتایج تصاویر با درجه بالا و متوسط با میدان دید بیشتر نزدیک بهم بوده و با تصاویر ضعیف و میدان دید کم اختلاف بیشتر دارند.

در آزمون‌های ابتدایی، از نظر کیفیت دوربین قابل استفاده به دلیل آنکه سیستم در تصمیم‌گیری به صورت فازی عمل می‌گردد و نمی‌توانستیم نتایج بسیار دقیق و قابل مقایسه رقمی از آن بدست آوریم لذا با توجه به نتایج خوب حاصل از تصاویر کیفی درجه بالا، از آنها بعنوان تصاویر مرجع استفاده شد و نتایج آن بعنوان معیار انتخاب و سایر تصاویر با این نتایج سنجیده می‌شد. از نظر آزمون‌های نهایی تشخیص با ارزیابی معیارها اولیه، سیستم در ترافیک‌های مختلف بررسی و معیارهای دیگری برای نتایج ارائه شد که از آن جمله می‌توان به تشخیص‌های صحیح، خطا، میزان اثر سایه و تشخیص خودرو به عنوان میزان اثر الگوریتم‌ها اشاره کرد.

آزمون: سیستم با چهار رشته تصویر از چهار خیابان منتهی به چهارراه، هر یک شامل ۱۰۰ فریم و سه درجه تفکیک مختلف با ۷۰ آزمون متوالی تست گردیده و نتایج حاصل شده‌اند. این نتایج به همراه نمودارهای مربوطه در رابطه با کیفیت مختلف تصاویر دوربین‌ها در جداول ۱ تا ۴ و نتایج حاصل از تشخیص

دقت تشخیص بالای خودروها و حذف سایه برخوردار است. در حال بهبود پایگاه دانش و افزودن بخش‌های دیگر مانند شمارش خودرو به سیستم و افزایش سرعت پردازش بخش‌های مختلف آن بر روی یک معماری موازی هستیم تا بتواند به عنوان یک سیستم بلادرنگ، با تغییر مکان دوربین و در چارچوب‌های مختلف عمل کند.

مراجع

[1] B. M. C. Basnayake, Y. K. Rajapakse, and P. P. M. Dabare, "Coordination of Closely located Traffic Signals," Final Year Project Report, Department of Civil Engineering, University of Moratuwa, Srilanka, 2000.

[2] F. Dellart, D. Pomerlain, and C. Thorpe, "Model-based Car Tracking with a Road Follower," *Proc. Int'l Conf. on Robotics and Automation*, 1998.

[3] R. Y. Tsai, and T. S. Huang, "Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. ASSP-29, No. 6, pp. 1147-1152, 1981.

[4] D. Koller, K. Danilidis, and H. H. Nagel, "Model-based Object Tracking in Monocular Image Sequences of Road Traffic Scenes," *International Journal of Computer Vision*, Vol. 10, No. 3, pp.257-281, 1993.

[5] D. Beymer, "Traffic Surveillance Home Page," <http://www.cs.berkeley.edu/~beymer/traffic-surveill.html>.

[6] B. Coifman, S. Dhoorjaty, and Z. H. Lee, "Estimating median velocity instead of mean velocity at single loop detectors," *Transportation Research*, Vol. 11C, No. 3-4, pp. 211-222, 2003.

[7] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A Real-time Computer Vision System for Measuring Traffic Parameters," *Proc. Intelligent Transport Systems*, 2003.

[8] J. A. K. S. Jayasinghe, and G. G. D. nishantha, "A CCTV Image Grabber for Vehicular Parameter Detection, Symposium on Research for Industry," Engineering Research Unit, University of Moratuwa, SriLanka, pp. 51-59, 1999.

[9] A. Rosenfeld, "Image Analysis: Problems, Progress and Prospects," *Pattern Recognition*, Vol. 17, pp. 3-12, 1984.

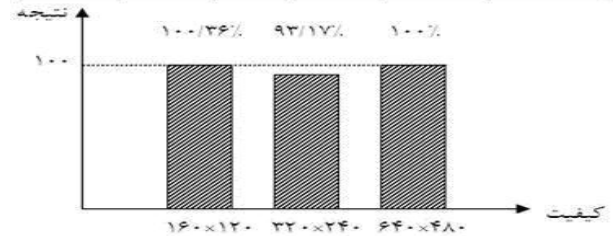
[10] T. O. Binford, "Survey of Model-Based Image Analysis Systems," *The International Journal of Robotics Research*, Vol. 1.1, No. 1, 1982.

[11] B. Abdulbahai, and S. M. Tabib, "Spatio temporal inductive pattern recognition for vehicle re-identification," *Transportation Research*, Vol. 11C, No. 3-4, pp. 223-240, 2003.

[12] D. R. Magee, "Tracking multiple vehicles using foreground, background and motion models, Image and Vision Computing," *Image and Vision Computing*, Vol. 22, pp. 143-155, 2004.

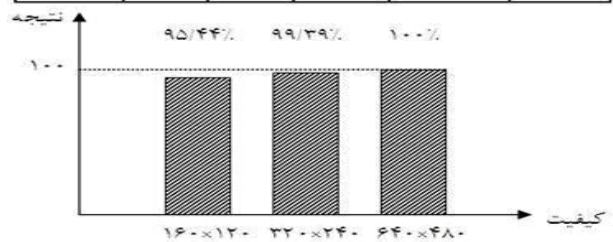
جدول ۳- ورودی ۳ و نتیجه آن

نتیجه	حجم	غیر قابل پردازش	پیکسل سفید	سطح خیابان	کیفیت
۱۰۰٪	۱۶/۵۵٪	۳۵۵۸۳	۶۸۲۱	۴۱۲۰۸	۶۴۰×۴۸۰
۹۳/۱۷٪	۱۵/۴۲٪	۳۱۹۴۶	۶۹۱۸	۴۴۸۵۱	۳۲۰×۲۴۰
۱۰۰/۳۶٪	۱۶/۶۱٪	۳۴۷۶۹	۶۹۸۰	۴۲۰۱۵	۱۶۰×۱۲۰



جدول ۴- ورودی ۴ و نتیجه آن

نتیجه	حجم	غیر قابل پردازش	پیکسل سفید	سطح خیابان	کیفیت
۱۰۰٪	۲۶/۳۲٪	۲۴۶۸۱	۱۳۷۲۱	۵۲۱۱۸	۶۴۰×۴۸۰
۹۹/۳۹٪	۲۶/۱۶٪	۲۳۸۴۱	۱۳۸۵۴	۵۲۹۵۰	۳۲۰×۲۴۰
۹۵/۴۴٪	۲۵/۱۲٪	۲۰۷۱۵	۱۴۰۸۵	۵۶۰۶۸	۱۶۰×۱۲۰



جدول ۵- کارایی الگوریتم در ترافیک متفاوت

درصد	ترافیک کم	ترافیک متوسط	ترافیک سنگین
تشخیص صحیح	٪۹۹	٪۹۶.۵	٪۹۳.۲
تشخیص خطا	٪۱	٪۳.۵	٪۶.۸
حذف سایه	٪۹۸	٪۹۵	٪۹۰
تشخیص خودرو	٪۹۷	٪۹۵	٪۹۰

۱۱- نتیجه‌گیری

در این پژوهش، هدف ارائه مدل جدیدی برای چراغ راهنمای هوشمند براساس تصاویر دریافتی دوربین‌ها ارائه بوده است که بتواند نقش یک سیستم بلادرنگ را برای یک چهارراه چهارفاز بازی کند. با بهره‌گیری از تکنیک‌های جدید و افزایش کارایی الگوریتم‌ها در مقابل ازدحام، سایه‌ها، و تغییرات محیطی نسبتاً مقاوم است. تغییرات آرام نور محیط^{۵۷} و تا حدودی تغییرات ناگهانی نور را تحمل کرده و نسبت به شعاع پوشش دوربین نیز وابستگی ندارد. با بهره‌گیری از پایگاه دانش، و حدس حادثه، مجدداً برنامه‌ریزی می‌شود. زمان‌بندی را براساس کل ترافیک چهارراه تنظیم می‌کند و به پارامتر انتظار اجتماعی رانندگان به عنوان میانگین انتظار توجه دارد. سایه و خطاهای موجود را برای افزایش دقت محاسبه ترافیک، به تفکیک لاین‌ها از تصاویر حذف می‌کند. با افزایش کیفیت تصاویر، دقت نتایج بهتر می‌شود. اشیاء متحرک در پیاده‌رو و سمت دیگر خیابان اثری بر نتایج ندارند. سیستم از

- Computing: Techniques and Applications*, pp. 469-478, 2003.
- [26] D. Gutchessy, M. Trajković, E. Cohen-Solalz, D. Lyonsz, and A. K. Jainy, "A Background Model Initialization Algorithm for Video Surveillance," *Proc. 8th IEEE Int'l Conf. on Computer Vision*, Vol. 1, pp. 744-740, 2001.
- [27] W. Zhang, X. Z. Fang, and W. Y. Lin, "Moving vehicles segmentation based on Gaussian motion model," *Proc. Int'l Society for Optical Engineering*, Vol. 5960, pp. 141-148, 2005.
- [28] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, New Jersey, Prentice-Hall, 2001.
- [29] R. Jain, "Extraction of Motion Information from Peripheral Processes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 5, pp. 489-503, 1981.
- [30] E. Y. Kim, S. H. Pank, K. Jung, and H. J. Kim, "Genetic Algorithm-based Segmentation of Video Sequences," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-7, No. 4, pp. 367-373, 1985.
- [31] R. L. Kashyap, and R. Chellappa, "Stochastic Models for Closed Boundary Analysis," *IEEE Trans. Information Theory*, pp. 627-637, 1981.
- [32] J. K. Aggrawal, and R. O. Duda, "Computer Analysis of Moving Polygonal Images," *IEEE Trans. Computers*, Vol. C-24, No. 10, pp. 966-976, 1975.
- [33] Artificial Intelligence Laboratory at Kyungpook National University, Vehicle Detection and Tracking in a Traffic Scene, <http://ailab.knu.ac.kr/TrafficSurveillanceSystem/tracking.htm>.
- [34] T. J. Broida, and R. Chellappa, "Estimation of Object Motion Parameters from Noisy Images," *Proc. IEEE Conf. Computer Vision Pattern Recognition*, pp. 82-88, 1985.
- [35] C. Mukesh. R. Gadiya, and C. Motwani, "Survey of Image Denoising Techniques," *Proc. GSPx*, pp. 27-30, 2004.
- [36] R. Chellappa, and A. A. Sawchuck, "Digital Image Processing and Analysis," *Digital Image Processing*, Vol. 2, pp. 220-225, 1985.
- [37] P. C. Chen, and T. Pavlidis, "Image Segmentation as an Estimation Problem," *Comput. Graph. Image Processing*, Vol. 12, pp. 153-172, 1980.
- [38] P. C. Chen, and T. Pavlidis, "segmentation by Texture Using Correlation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, pp. 64-69, 1983.
- [39] R. Ohlander, K. Price, and D. R. Reddy, "Picture Segmentation Using a Recursive Region Splitting Method," *Computer Graphics and Image Processing*, Vol. 8, pp. 313-333, 1978.
- [13] L. Lin, X. B. Han, R. Ding, G. Li, S. C. Y Lu, and Q. Hong, "A New Rechargeable Intelligent Vehicle Detection Sensor," *Journal of Physics, Conference International Symposium on Measurement Technology and Intelligent Instruments*, Vol. 1, pp. 102-106, 2005.
- [14] A. S. K. Chakravarthy, V. D. S. Phaneendhar, and S. Sadhukhan, "A Report on Traffic Vision (Traffic Surveillance Using Computer Vision)," International Institute of Information Technology, 2002.
- [15] J. Nascimento, and J. Marques, "Performance evaluation of object detection algorithms for video surveillance," *IEEE Trans. Multimedia*, Vol. 8, No. 9, pp. 761-774, 2006.
- [16] C. Ramchandani, "Analysis of Asynchronous Concurrent Systems by Time Petri Nets," Project MAC-TR 120, MIT, 1974.
- [17] V. Savolainen, and V. Terziyan, "Metapetrinets for Controlling Complex and Dynamic Processes," *International Journal of Information and Management Sciences*, Vol. 10, No.1, pp.13-32, 1999.
- [18] V. Terziyan, and S. Puuronen, "Reasoning with Multilevel Contexts in Semantic Metanetworks," *Proc. Formal Aspects in Context*, pp. 107-126, 2000.
- [19] A. Bobbio, M. Gribaudo, and A. Horváth, "Modeling a Car Safety Controller in Road Tunnels using Hybrid Petri Nets," *Proc. the Intelligent Transportation Systems Conf.*, pp.1436-1441, 2006.
- [20] M. Gribaudo, M. Sereno, A. Horváth, and A. Bobbio, "Fluid stochastic Petri nets augmented with flush-out arcs," *Proc. Modelling and analysis Discrete Event Dynamic Systems*, Vol. (1/2), No. 11, pp. 97-117, 2001.
- [21] G. Horton, V. Kulkarni, D. Nicol, and K. Trivedi, "Fluid stochastic Petri nets, Theory, application and solution techniques," *European Journal of Operational Research*, Vol. 1, No. 10, pp. 184-201, 1998.
- [22] Y. Papadopoulos, J. McDermid, A. Mavrides, C. Scheidler, and M. Maruhn, "Model-based semiautomatic safety analysis of programmable systems in automotive applications," *Proc. IEEE Int'l Conf. Advanced Driver Assistance Systems*, pp. 53-57, 2001.
- [23] G. Sala, R. Brignolo, E. Carrubba, U. Jallasse, and D. Shinar, "Improvement of accident prevention in road tunnels through intelligent infrastructures and intelligent vehicles operation," *Proc. Intelligent Transport Systems*, 2003.
- [24] K. Trivedi, and V. Kulkarni, "FSPNs: Fluid Stochastic Petri nets," *Proc. Int'l Conf on Application and Theory of Petri Nets*, Vol. 691 of LNCS, pp. 24-31, 1993.
- [25] Y. Ming, J. Jiang, and J. Ming, "Background Modeling and Subtraction Using a Local-Linear-Dependence-Based Cauchy Statistical Model," *Proc. of Digital Image*

آدرس پست الکترونیکی نامبرده عبارتست از:

mahabadi@shahed.ac.ir

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۶/۱۲/۰۵

تاریخ اصلاح: ۸۹/۰۳/۰۹

تاریخ قبول شدن: ۸۹/۰۳/۲۵

نویسنده مرتبط: امین‌الله مه‌آبادی، دانشکده فنی و مهندسی، دانشگاه شاهد، تهران، ایران.

- 1 Traffic Measurement
- 2 Intelligent Traffic Signal
- 3 Real Time
- 4 Intelligent Transportation Systems
- 5 Segmentation
- 6 Circle Frequency
- 7 Video Camera
- 8 Inductive Loop
- 9 Sensors
- 10 Ultrasound
- 11 Real Time
- 12 Offline
- 13 Real Time Controller
- 14 Petri Net
- 15 Reference Image
- 16 High Type - 4 Phase Intersection
- 17 Time Petri Net
- 18 Meta Petri Net
- 19 Semantic Meta network
- 20 Hybrid Petri Net
- 21 Green Signal
- 22 Red Signal
- 23 Transition
- 24 Meta Level
- 25 Basic Model Level
- 26 Model Level
- 27 Meta Model Level
- 28 Interrupt Signal Generator
- 29 Place off-on
- 30 Dynamic Network
- 31 Active/Deactive
- 32 Token
- 33 Overflow
- 34 Deadlock
- 35 Lane
- 36 Turn Around Time
- 37 Events
- 38 Context Switching Time
- 39 Social Behavior
- 40 Variations of Lighting
- 41 Dynamic Varieties
- 42 Shadow
- 43 Dynamic Deference Method
- 44 Object Mask
- 45 Thresholding
- 46 Movement Mask
- 47 Learning Parameter
- 48 Background Model Sensitivity
- 49 Input Section
- 50 Output Section
- 51 Median
- 52 Regional
- 53 Differential
- 54 Global Thresholding
- 55 Between Lights
- 56 Canny Filter
- 57 Light

[40] A. M. Nazif, and M. D. Levine, "Low Level image Segmentation: An Expert System," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI- 6, No. 5, pp. 555-577, 1984.

[41] S. M. Smith, "ASSET-2: Real-Time Motion Segmentation and Shape Tracking," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 17, No. 8, pp. 814-82, 1995.

[42] G. C. de Silva, "Automation of Traffic Flow Measurement Using Video Images," Department of Computer Science and Engineering, University of Moratuwa, SriLanka, 2001.

[43] I. E. Abdou, and W. K. Pratt, "Quantitative Design and Evaluation of Enhancement/ Thresholding Edge Detectors," *Proceedings of the IEEE*, Vol. 67, No. 5, pp. 753-763, 1979.

[44] P. L. Rosin, and T. Ellis, "Image difference threshold strategies and shadow detection," *Information Science and Engineering Journal*, Vol. 8, pp. 345-354, 2005.

[45] H. Y. Cheng, Q. Z. Wu, K. C. Fan , and B. S. Jeng, "Binarization Method Based on Pixel-level Dynamic Thresholds for Change Detection in Image Sequences," *Information Science and Engineering Journal*, Vol. 22, pp. 545-557, 2006.

[46] C. Jiang, and M. O. Ward, "Shadow identification," *Proc, IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 606-612, 1992.

[47] X. Jay, Z. Wei, and F. Yewei, "Effective vehicle detection technique for traffic surveillance systems," *Journal of Visual Communication & Image Representation*, Vol. 17, pp. 647-658, 2006.

[48] X. Li, Z. Liu, and K. Leung, "Detection of vehicle from traffic scenes using fuzzy integrals," *The Journal of the pattern recognition society*, Vol. 35, pp. 967-980, 2004.

[49] J. W. Roach, and J. K. Aggrawal, "Determining the Movement of Objects from a Sequence of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 6, pp. 554-562, 1980.

[50] S. Kawato, and J. Ohya, "Real-time detection of nodding and head-shaking by directly detection and tracking "between-eyes", " *Proc, IEEE 4th Int'l Conf. on Automatic Face and Gesture Recognition*, pp. 40-45, 2000.



امین‌الله مه‌آبادی مدرک کارشناسی سخت‌افزار کامپیوتر خود را از دانشکده مهندسی برق دانشگاه علم و صنعت ایران در سال ۱۳۶۸ اخذ نمود همچنین مدرک کارشناسی ارشد معماری کامپیوتر خود را از دانشکده مهندسی کامپیوتر دانشگاه صنعتی امیرکبیر در سال ۱۳۷۵ دریافت کرد. وی از سال ۱۳۷۶ عضو هیات علمی گروه سخت‌افزار دانشکده فنی و مهندسی دانشگاه شاهد است. تاکنون پنج کتاب در زمینه علوم کامپیوتر ترجمه و تالیف نموده و مقالاتی نیز در مجلات علمی پژوهشی منتشر کرده است. زمینه‌های پژوهشی مورد علاقه نامبرده پردازش موازی، پردازش تصویر و شبیه‌سازی است.

استفاده از فناوری پردازش گفتار در ارزیابی خودکار روانی گفتار در سیستم‌های آموزش زبان خارجی

محمد مهدی همایون پور

بهزاد بختیاری

دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران

چکیده

در این مقاله روشی جدید برای ارزیابی خودکار روانی گفتار ارائه خواهد شد. از کاربردهای ارزیابی خودکار روانی گفتار می‌توان به آموزش زبان خارجی به کمک کامپیوتر^۱ اشاره نمود. بدین صورت که گفتار زبان آموز ارزیابی شده و سطح روان و سلیس بودن گفتار وی مشخص می‌گردد. با اختیار داشتن ارزیابی کننده خودکار روانی گفتار، زبان آموز می‌تواند از سطح روان بودن گفتار خود آگاه شده و با تکرار و تمرین سعی در بهتر نمودن سطح روانی گفتار خود نماید. برای این منظور معیارهای کمی مختلفی ارائه شده است که می‌توان به کمک آنها روانی گفتار را پیش‌بینی نمود. در اینجا با بررسی مناسب بودن این معیارها، روشی برای ارزیابی خودکار روانی گفتار ارائه می‌شود. برای بررسی صحت ارزیابی خودکار ارائه شده دادگانی از ترک زبانان تهیه شده و آزمایشات بر روی آن انجام گرفت. همبستگی بین نمرات داوران انسانی و ارزیابی کننده خودکار در نهایت نزدیک به ۰/۹۵ رسید.

کلمات کلیدی: آموزش زبان خارجی به کمک کامپیوتر، ارزیابی خودکار روانی گفتار، شناسایی خودکار گفتار.

۱- مقدمه

هزینه‌های گزافی دارند، دسترسی به سیستم‌های خودکار آموزش زبان سهل الحصول و کم هزینه می‌باشد.

عموماً آن مختصه‌های آوایی که پس از تقسیم یک زنجیره به کوچکترین واحد های تشکیل دهنده آن بدست می‌آیند را واحدهای زنجیری^۲ می‌گویند. واکه‌ها و همخوانها واحدهای زنجیری گفتار را تشکیل می‌دهند. آن دسته از مشخصات آوایی که در هیچ یک از واحدهای زنجیری نمی‌گنجند (زیرا این مشخصات اولاً همزمان با بیش از یک واحد زنجیری تولید می‌شوند، ثانیاً جایگاه آنها در زنجیره گفتار متغیر است) را واحدهای زیر زنجیری یا مشخصات نوایی^۳ گفتار^۴ گویند. اگر بخواهیم در یک تقسیم‌بندی کلی از مشخصه‌های نوایی^۴ گفتار نام ببریم، باید از تکیه^۵، زیر و بمی^۶، سلیس و روانی^۷ گفتار، درنگ و وزن طبیعی^۸ گفتار یا ریتم^۹ نام ببریم. در اینجا از بین مشخصه‌های نوایی به ارزیابی سلیس و روان بودن گفتار به کمک کامپیوتر خواهیم پرداخت. این مشخصه در ارتباطات شفاهی از اهمیت خاصی برخوردار است. سطح مربوط به آموزش این مشخصه در سطوح آموزش

امروزه با توجه به اهمیت بالای آموزش زبان، از کاربردهای مهم و مفید فناوری شناسایی خودکار گفتار می‌توان به استفاده از این فناوری در آموزش زبان خارجی اشاره نمود. از این فناوری می‌توان در دو زمینه آموزش تلفظ و آموزش نوای گفتار استفاده نمود. در این مقاله آموزش روانی گفتار مورد بررسی قرار خواهد گرفت. بطور کلی آموزش خودکار زبان دارای مزیت‌هایی نسبت به روشهای آموزش سنتی می‌باشد. به عنوان مثال در کلاس‌های سنتی، معلم دقت و انرژی خود را صرف همه زبان آموزان می‌کند و بر روی یک زبان آموز خاص تمرکز ندارد. در حالیکه در آموزش خودکار زبان، سیستم به سادگی بر روی یک زبان آموز تمرکز می‌کند. از مشکلات دیگری که در کلاسهای سنتی آموزش زبان وجود دارد، خجالت کشیدن زبان آموز در سؤال کردن‌های متوالی و یا کم حوصلگی و خستگی معلم می‌باشد. همچنین از آنجائیکه کلاسهای سنتی همیشه در دسترس نیستند و

در این مقاله نیز هدف از بررسی روانی گفتار در آموزش زبان، تأکید بر روی جنبه‌های زمانی روانی گفتار خواهد بود. در جهت تعریف جنبه‌های زمانی روانی گفتار اغلب فرض می‌کنیم که هدف آموزش زبان است که شامل تولید "گفتار گویندگان بومی با توجه به پارامترهای زمانی که حدالمقدور شامل وقفه‌های سکوت^{۱۴}، درنگ^{۱۵}، وقفه‌های پر شده^{۱۶}، شروع دوباره و اصلاح مجدد، ... نباشد [۵]. قابل ذکر است که مطالعات کمی صورت گرفته در مورد وقفه‌های واجی برای گویندگان بومی نتایج هموار و پیوسته‌ای را همیشه به دنبال ندارد. در واقع این مطلب نشان می‌دهد که درنگ‌های واجی برای تمایز بین گفتار بومی و غیربومی نمی‌تواند کافی باشد. آقای مول^{۱۷} برای حل این مشکل پیشنهاد کرد، توزیع تعداد وقفه‌ها در نظر گرفته شود [۱۳]. مطالعاتی که بر روی مقایسه گفتار دو زبان با گویندگان یکسان صورت گرفت نشان می‌دهد که می‌توان تفاوتی بین دو گفتار قائل شد [۱۲، ۱۳].

از ابزارهای ارائه شده برای این منظور می‌توان به کارهای ارائه شده در [۱، ۱۴] اشاره کرد. در [۱۴]، روانی در گفتار بصورت بیان همه کلمات در محدوده کشش^{۱۸} از پیش تعریف شده‌ای که از میانگین گفتار بومی بدست می‌آید، تعریف می‌شود. در [۱، ۹] ویژگی مختلف شامل کشش بخش‌ها با یکدیگر ترکیب می‌شوند تا معیاری را برای روانی گفتار تشکیل دهند. در این کار نشان داده شده که بین ارزیابی خودکار روانی گفتار و ارزیابی داوران انسانی در روانی گفتار همبستگی^{۱۹} وجود داشته و در نتیجه می‌توان به ارزیابی مبتنی بر کامپیوتر اعتماد نمود. در هر دو کار ذکر شده معیارهایی برای سلیس و روان بودن گفتار معرفی می‌شود ولی روشی برای ارزیابی خودکار میزان روان بودن گفتار زبان آموزان ارائه نمی‌شود.

همچنین در دانشگاه کالیفرنیا جنوبی تحقیقاتی در زمینه ارزیابی خودکار روانی گفتار صورت گرفته که به کمک آن کودکان مکزیکی به آموختن زبان انگلیسی پرداختند [۱۵]. در [۱۶] اشاره شده است که برای ارزیابی خودکار روانی گفتار دو دسته الگوریتم وجود دارد. در دسته اول هدف پیدا کردن مکان و زمان دقیق وقفه گفتار زبان آموز می‌باشد. از جمله این کارها می‌توان به [۱۷، ۱۸] اشاره نمود. در این کارها سیستم مبتنی بر شناسایی کننده‌های خودکار می‌باشد که فرض بر این است که نگارش آوایی گفتار نیز در دسترس می‌باشد. اما نوع دوم سعی می‌کند به میزان روانی گفتار زبان آموز امتیازی تخصیص دهند که از این جمله می‌توان به [۱۹، ۲۰] اشاره نمود. در این مقاله سعی گردید که هر دو هدف محقق شود به نحوی که کار صورت گرفته علاوه بر اینکه مکان‌های وقفه گفتار زبان آموز مشخص می‌شود در کل نمره‌ای نیز به میزان روانی گفتار وی داده می‌شود.

نمای کلی روش ارزیابی خودکار روانی گفتار را در شکل ۱ مشاهده می‌کنید. در ادامه، در بخش ۲ به دادگان غیربومی مورد استفاده خواهیم پرداخت. در بخش ۳، معیارهای کمی روانی گفتار را تعریف خواهیم نمود و به بررسی رابطه آنها با ارزیابی‌های انسانی خواهیم پرداخت. در بخش ۴، روشی برای ارزیابی خودکار روانی گفتار به کمک معیارهای کمی تعریف شده خواهیم پرداخت. بخش ۵ نیز به خلاصه و نتیجه‌گیری می‌پردازد.

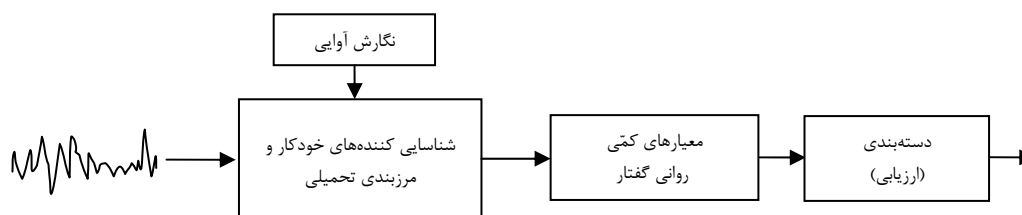
زبان، به زبان آموزانی با سطح بالای متوسط پیشنهاد می‌شود. بدیهی است که زبان آموزان مبتدی حتی در قسمت بیان یک کلمه و تلفظ صحیح آن نیز دچار اشکال می‌گردند و در نتیجه تمرین برای روانی گفتار یک جمله کاری دشوار برای زبان آموزان مبتدی خواهد بود. اما زبان آموزانی که سطح مبتدی را پشت سر گذاشته و به سطح متوسط می‌رسند می‌توانند با تمرین بیان جمله‌های مختلف در روان صحبت کردن مهارت لازم را کسب کنند.

به طور معمول اساتید و محققان، "روانی گفتار" را بصورت کارایی زبان بومی و غیربومی توصیف می‌کنند [۱]. در [۲، ۳، ۴، ۵، ۶] مروری بر روی واژه روانی گفتار و جنبه‌های مختلف آن شده است. در اینجا نشان خواهیم داد که بین "تعریف روانی گفتار با توجه به کارایی زبان بومی" و "روانی گفتار در محتوای آموزش زبان خارجی" تفاوت‌هایی وجود دارد. در حالت دوم روانی بصورت معیار مهمی از کارایی افراد غیربومی محسوب می‌شود که می‌توان آن را مورد ارزیابی و داوری قرار داد [۸]. این مطلب از این واقعیت ناشی می‌شود که روانی گفتار اغلب در ارزیابی‌ها و تست‌ها گنجانده می‌شود.

اگر کارایی گویندگان بومی در نظر گرفته شود، می‌توان از آن برای تعیین مشخصات کارایی گویندگان استفاده نمود. ولی در حقیقت نمی‌تواند نشان دهنده یک معیار ارزیابی باشد. همچنین از طرف دیگر واژه غیر روان بودن معمولاً به اتصال قسمتی از گفتار ناخواسته مانند، وقفه‌ها به گفتار اطلاق می‌شود. در نتیجه غیر روان بودن بصورت تعداد تکرارهای غیر عادی زیاد وقفه‌ها در جریان گفتار شناخته می‌شود [۹].

آقای فیلمور^۹ در [۳] چهار توانایی مختلفی را برای واژه روانی ذکر نموده است:

- ۱- توانایی برای صحبت کردن در طولی از زمان با تعداد وقفه‌های کم
 - ۲- توانایی صحبت کردن بطوری که بین جملات بیان شده ارتباط منطقی برقرار باشد
 - ۳- توانایی بکار گیری چیزهایی مناسب در صحبت کردن در یک زمینه وسیع از نظر محتوایی
 - ۴- توانایی در خلق و تصوّر چیزهایی در کاربردهای زبانی
- اما در مقوله آموزش زبان خارجی تعاریف دیگری نیز وجود دارد. از جمله، در آموزش زبان ارتباطی^{۱۰}، تأکید بر روی روانی نقطه مقابل دقت بشمار می‌آید. بر اساس تعریف آقای پرمفیت^{۱۱} روانی، تأثیر مضاعف عملیات سیستم آموزش زبان است که موجب اکتساب توسط زبان آموزان می‌گردد [۴]. با توجه به این تعریف، کارایی مشابه بودن به گویندگان بومی^{۱۲} هدفی نیست که بایستی به آن دست یابیم. همچنین کارایی شباهت به بومی بودن را می‌توان بصورت هدف نهایی تفسیر روانی در نظر گرفت. این تعریف می‌تواند در کاربردهای روزانه زبان توسعه پیدا کرده و نشان دهنده مهارت و کارایی زبان گردد [۵، ۷]. بالاخره، در حالت‌های خاص و محدود شده‌ای، واژه روانی برای مهارت در گفتار شفاهی و به طور خاص تر در جنبه‌های زمانی^{۱۳} بکار می‌رود [۵، ۶، ۸، ۱۰، ۱۱، ۱۲].



شکل ۱- دیگرام مراحل ارزیابی خودکار روانی گفتار

۲- دادگان غیربومی

۲-۲- ارزیابی روانی دادگان غیربومی توسط داوران انسانی و سازگاری آنها در ارزیابی

۲-۱- تهیه و طراحی دادگان غیربومی

بعد از تهیه دادگان مذکور بایستی جملات بیان شده توسط داوران انسانی مورد ارزیابی قرار گیرند تا نتایج ارزیابی خودکار و ارزیابی انسانی با هم مقایسه گردند. داوران مورد استفاده ۹ تن از دانشجویان کارشناسی ارشد و دکتری کامپیوتر بوده اند که زمینه فعالیت آنها پردازش گفتار می‌باشد. این دانشجویان آشنایی کامل به آواشناسی و مفاهیم پردازش گفتار دارند. برای داوران دادگان غیربومی نرم افزاری تهیه شد که فایل‌های گفتار را به ترتیب برای داور پخش می‌نمود. داوران به صورت انفرادی در محیط آرام توسط گوشی به گفتار گویندگان غیربومی گوش داده و سطح روانی گفتار مربوطه را به کمک نرم افزار ذخیره می‌کنند. این ارزیابی‌ها برای هر گوینده غیربومی، در دامنه ۱ تا ۵ صورت می‌گرفت که ۱ مربوط به روانترین و ۵ مربوط به بدترین گفتار از لحاظ سلیسی و روانی گفتار می‌باشد. مفهوم ۵ کلاس مذکور و امتیازات مربوطه در جدول ۲ آمده است.

از آنجائیکه با مقایسه نتایج ارزیابی داوران انسانی و نتایج سیستم ارزیابی خودکار، کارایی و صحت سیستم ارزیابی خودکار محاسبه خواهد شد و با توجه به اینکه داور امری شخصی می‌باشد، بایستی ابتدا نحوه ارزیابی داوران انسانی بررسی شود. بدیهی است در حالتیکه نتایج یک داور نسبت به دیگر داوران به دلایلی مانند، عدم دقت کافی، سخت‌گیری، در نظر گرفتن جنبه‌های دیگری از روانی گفتار و غیره فاصله بسیار زیادی داشته باشد، بایستی از آن داور در ارزیابی‌ها استفاده نشود. در این قسمت نتایج ارزیابی داوران در ارزیابی یک مجموعه گفتار غیربومی مشترک مورد بررسی قرار می‌گیرد. برای اینکه سازگاری داوران در ارزیابی روانی گفتار مشخص گردد، امتیازات داده شده توسط هر داور نسبت به امتیازات دیگر داوران مورد بررسی قرار می‌گیرد.

جدول ۳ همبستگی بین یک داور و دیگر داوران را نشان می‌دهد. برای هر داور همبستگی آن با تک تک داوران محاسبه شده و در نهایت میانگین همبستگی‌ها در جدول منظور شده است. با دقت در این جدول مشاهده می‌شود که ارزیابی روانی گفتار توسط داوران انسانی دارای همبستگی بالایی می‌باشد که نشان دهنده سازگار بودن داوران در این ارزیابی می‌باشد. تمام مقدار همبستگی‌ها بالای ۰/۹۵ می‌باشند. این همبستگی بالا شاید به این دلیل باشد که اولاً مشخصه روانی گفتار، مشخصه پیچیده‌ای نیست و ارزیابی آن ساده و بدون از مشکلات خاص می‌باشد. ثانیاً می‌توان نتیجه گرفت داوران تقریباً به یک اندازه سخت‌گیری نموده‌اند و همچنین جنبه‌های یکسانی از روانی گفتار مدنظر همه داوران بوده است.

به منظور بررسی این مطلب که امتیازات مربوط به ارزیابی داوران در گفتار بومی و غیربومی متفاوت بوده و علاوه بر آن برای بررسی تفاوت روانی گویندگان بومی و گویندگان غیربومی، میانگین امتیازات دادگان غیربومی مذکور با میانگین یک مجموعه گفتار بومی مورد مقایسه قرار می‌گیرد.

جدول ۲- سطوح مختلف روانی گفتار و امتیازات مربوطه

امتیاز	سطح روانی گفتار
۱	بسیار خوب
۲	خوب
۳	متوسط
۴	بد
۵	بسیار بد

در این مقاله چون هدف آموزش زبان مقصد به زبان آموزشی است که زبان مادری‌شان با زبان مقصد متفاوت است، بایستی دادگان غیربومی از گویندگان زبان مبدأ تهیه شود تا بتوان کارایی الگوریتم‌های ارائه شده را بررسی نمود. برای این منظور زبان مبدأ، زبان ترکی و زبان مقصد، زبان فارسی انتخاب شد. دلیل انتخاب زبان فارسی این بوده است که الگوریتم‌های ارائه شده در آموزش زبان به کمک کامپیوتر در آموزش زبان فارسی تا کنون کمتر مورد بررسی قرار گرفتند و همچنین زبان ترکی به دلیل دسترس بودن ترک زبانان در کشور انتخاب شد.

با توجه به اینکه در این مقاله هدف آموزش زبان به کمک کامپیوتر در زمینه روانی گفتار می‌باشد، بایستی جملاتی انتخاب شود تا ویژگی‌های خاصی ارضاء گردد. از جمله ویژگی‌هایی که این جملات بایستی آن را دارا باشند می‌توان به پوشش دادن تمام واج‌های موجود در زبان فارسی اشاره نمود. همچنین مناسب بودن طول جملات نیز از ویژگی‌های مطلوب محسوب می‌شود. چرا که این جملات قرار است توسط افراد غیربومی بیان شود و بعنوان مثال، بلند بودن جمله ممکن است زبان آموز را در بیان آن دچار مشکل سازد. در اینصورت آموزش بدرستی انجام نمی‌گیرد و تعامل درستی بین سیستم و زبان آموز انجام نمی‌شود. علاوه بر این بهتر است جملات پیچیده، حاوی کلمات مهجور، کلمات ادبی سنگین و مشابه اینها نباشد تا زبان آموز به سادگی جملات را بیان نموده و از مشکلات سلیسی و روانی خود آگاه گردد. برای این منظور ۴۰ جمله از دادگان فارس دات ۱۰۰ انتخاب شدند. این دادگان شامل ۶۰۰ جمله می‌باشد که اکثر جملات به گونه‌ای طراحی شده‌اند که تمام واج‌های زبان فارسی را بپوشانند. لازم به ذکر است از این دادگان برای آموزش مدل‌ها نیز استفاده شده است که گفتار استفاده شده از این دادگان جملاتی به غیر از این ۴۰ جمله می‌باشد. برای تهیه دادگان مناسب سعی شد از ترک زبانان افرادی با تحصیلات، جنسیت و سنین مختلف انتخاب شود. در کل گفتار ۱۱ گوینده ترک زبان ضبط شد که جدول ۱ جنسیت، سن، میزان تحصیلات و تعداد جمله بیان شده را نمایش می‌دهد.

جدول ۱- اطلاعات مربوط به دادگان گویندگان غیربومی

ردیف	شناسه گوینده	جنس	حدود سنی	میزان تحصیلات	تعداد جمله بیان شده
۱	۳	مرد	۲۵	کارشناسی	۱۰
۲	۴	زن	۵۵	زیر دیپلم	۶
۳	۵	مرد	۶۰	دیپلم	۱۱
۴	۹	زن	۴۰	دیپلم	۱۲
۵	۱۲	مرد	۲۵	کارشناسی	۱۸
۶	۱۷	زن	۳۵	دیپلم	۶
۷	۲	مرد	۲۵	کارشناسی ارشد	۱۰
۸	۶	مرد	۳۰	کارشناسی ارشد	۱۰
۹	۷	مرد	۳۰	کارشناسی	۱۰
۱۰	۱۳	مرد	۳۰	کارشناسی	۱۰
۱۱	۱۵	مرد	۲۵	کارشناسی	۱۰

جدول ۳- همبستگی داوران انسانی در ارزیابی روانی گفتار دادگان غیربومی

داور	۱	۲	۳	۴	۵	۶	۷	۸	۹
همبستگی	۰/۹۶	۰/۹۵	۰/۹۵	۰/۹۶	۰/۹۵	۰/۹۷	۰/۹۶	۰/۹۵	۰/۹۶

$$\#P = \text{تعداد وقفه‌های درون جمله‌ای بیشتر از } ۰/۲ \text{ ثانیه} \quad (۴)$$

کل زمان وقفه‌ها (TDP)^{۲۴}:

$$TDP = \text{کل زمان وقفه‌های درون جمله‌ای بیشتر از } ۰/۲ \text{ ثانیه} \quad (۵)$$

میانگین طول وقفه‌ها (MLP)^{۲۵}:

$$MLP = \text{میانگین کل زمان وقفه‌های درون جمله‌ای بیشتر از} \quad (۶)$$

۰/۲ ثانیه

میانگین طول تداوم (MLR)^{۲۶}:

$$MLR = \text{میانگین تعداد واج‌هایی که بین دو وقفه بیشتر از } ۰/۲ \quad (۷)$$

ثانیه رخ می‌دهد

هفت معیار مذکور، از معیارهای معروفی هستند که محققان زیادی بر روی آنها مطالعاتی انجام داده اند. در کارهای ابتدایی، این معیارها بصورت دستی توسط یک فرد خبره از گفتار استخراج می‌شد. اما در [۱] این معیارها بصورت خودکار توسط شناسایی کننده‌های خودکار از گفتار گویندگان استخراج می‌شود. در اینجا نیز با توجه به اینکه هدف ارزیابی خودکار روانی به کمک کامپیوتر می‌باشد از شناسایی کننده‌های خودکار برای این منظور استفاده خواهد شد. در [۱] این معیارها برای هلندی زبانان مورد بررسی قرار گرفت که در واقع گویندگان بومی هلندی زبان و گویندگان غیربومی دارای زبان‌های مختلفی بودند که هلندی صحبت می‌کنند. در اینجا معیارهای مذکور با دادگان غیربومی (آذری زبانان) و بومی (فارسی زبانان) مورد بررسی قرار می‌گیرد.

۳-۱- شناسایی کننده‌های خودکار و استخراج معیارهای کمی روانی گفتار

از آنجائیکه روش‌های ارزیابی خودکار ارائه شده مبتنی بر مدل مخفی مارکوف می‌باشند از شناسایی کننده‌های مبتنی بر مدل مخفی مارکوف استفاده خواهیم نمود. مدل مخفی مارکوف از ابزارهای قوی آماری برای شناسایی گفتار بشمار می‌رود که در مسائل شناسایی گفتار به کثرت مورد استفاده قرار می‌گیرد. برای شناسایی در سطح واج، هر واج با یک مدل مخفی مارکوف سه حالتی چپ به راست مدل می‌شود. از آنجائیکه بخش آزمایشات در این مقاله مربوط به آموزش زبان فارسی به ترک زبانان می‌باشد، بایستی شناسایی کننده‌هایی برای واج‌های زبان فارسی طراحی شود. برای این منظور ۲۹ واج زبان فارسی و یک برچسب برای سکوت در نظر گرفته شد.

مجموعه گفتار بومی از دادگان فارس دات انتخاب شد. برای این منظور از ۲۰ گوینده با لهجه تهرانی استفاده شد. جدول ۴ میانگین امتیازات دادگان غیربومی و امتیازات مجموعه گفتار بومی را نمایش می‌دهد. همانطور که در این جدول مشاهده می‌شود گویندگان بومی امتیازات بهتری از گویندگان غیربومی کسب نموده اند. پس می‌توان نتیجه گرفت که اولاً تفاوتی بین روانی گفتار گویندگان بومی و غیربومی وجود دارد. ثانیاً داوران نیز در ارزیابی روانی گویندگان بومی و غیربومی تفاوت قائل می‌شوند.

جدول ۴- میانگین امتیازات داوران انسانی در روانی گفتار دادگان بومی و دادگان غیربومی

دادگان بومی	۲/۵
دادگان غیربومی	۴/۱

۳-۲- معیارهای کمی روانی گفتار

در این قسمت معیارهایی روانی گفتار معرفی خواهند شد. این معیارها بایستی به نحوی باشند که بتوان آنها را بصورت خودکار استخراج نمود. هدف از این معیارها اینست که بتوان به کمک آنها روانی گفتار را پیش‌بینی (ارزیابی) نمود. محققان زیادی بر روی معیارهای مختلف کار نموده اند، از جمله می‌توان به [۵، ۸، ۱۰، ۱۱، ۱۲] اشاره نمود. معیارهای معرفی شده در این کارها، جنبه زمانی روانی گفتار را در نظر گرفته اند و معیارهایی کمی^{۲۰} می‌باشند. از بین معیارهای ارائه شده در این کارها، هفت معیار پایه و مهم زیر انتخاب شدند:

سرعت گفتار (ROS)^{۲۱}:

$$ROS = \frac{\text{تعداد واج‌های بیان شده}}{\text{کل زمان گفتار با در نظر گرفتن وقفه‌های درون جمله}} \quad (۱)$$

نسبت آوا به زمان (PTR)^{۲۲}:

$$PTR = \frac{\text{کل زمان گفتار بدون وقفه‌ها}}{\text{کل زمان گفتار با در نظر گرفتن وقفه‌های درون جمله}} \times ۱۰۰\% \quad (۲)$$

سرعت تعامل (ART)^{۲۳}:

$$ART = \frac{\text{تعداد واج‌های بیان شده}}{\text{کل زمان گفتار بدون وقفه‌ها}} \quad (۳)$$

تعداد وقفه‌های سکوت (#P):

با توجه به اینکه ضرایب کپستراتال مبتنی بر معیار مل در محیط‌های عملیاتی و نویری از سایر ضرایب بهتر عمل می‌کنند، این ضرایب به همراه لگاریتم انرژی و مشتقات اول و دوم آن جهت بردارهای ویژگی انتخاب شده‌اند. بُعد بردارهای فوق ۳۹ می‌باشد. جهت استخراج ویژگی از فریم‌هایی با طول ۳۰ میلی‌ثانیه و همپوشانی ۲۰ میلی‌ثانیه استفاده شده است که بر روی آنها پنجره در طرفین کاهنده همینگ اعمال شده است تا اثر ناپیوستگی در ابتدا و انتهای فریم کمتر شود. پس از فریم‌بندی، طیف سیگنال گفتار به کمک یک فیلتر پیش‌تأکید که دارای ضریب ۰/۹۷۵ می‌باشد، هموار شده است. تعداد فیلترهای مورد استفاده در بانک فیلتر مل برابر ۱۸ بوده است.

برای آموزش مدل‌ها از دادگان فارس دات نسخه ۱۰۰ استفاده شده است. این دادگان شامل ۳۰۰ گوینده و ۶۰۰ جمله می‌باشد که هر گوینده ۱۰ جمله در محیطی کاملاً آرام و بدون نویز بیان نموده است. گویندگان دارای سن، جنس و تحصیلات مختلفی هستند. لهجه‌های مختلف زبان فارسی در این دادگان عبارتند از: تهرانی، اصفهانی، یزدی، شمالی، جنوبی، لوری، خراسانی، بلوچی، کردی، ترکی. در اینجا ۴۵ گوینده با لهجه تهرانی از این دادگان برای آموزش انتخاب شد. این گویندگان کاملاً از لحاظ جنسی و سنی مختلف انتخاب شدند. به کمک این شناسایی‌کننده‌ها و نگارش آوایی گفتار بیان شده، مرزبندی تحمیلی به کمک الگوریتم ویتربی انجام می‌گیرد و مرزهای واج‌ها و سکوت تعیین می‌شود. با داشتن مرزهای مربوط به واج‌ها و سکوت‌ها معیارهای کمی مذکور محاسبه می‌شوند. از آنجائیکه محل وقفه‌ها در گفتار زبان آموز مشخص نیست، در نگارش آوایی بعد از هر کلمه یک برچسب سکوت گذاشته می‌شود. بدیهی است در صورتیکه وقفه‌ای وجود نداشته باشد، طول تعیین شده توسط مرزبندی تحمیلی برای سکوت در نظر گرفته شده کوچک خواهد بود و در صورتیکه وقفه‌ای رخ دهد طول دقیق آن مشخص می‌گردد.

۲-۳- معیارهای کمی روانی گفتار و پیش‌بینی روانی گفتار

در این قسمت به بررسی ارتباط بین معیارهای کمی تعریف شده و روانی گفتار زبان آموزان خواهیم پرداخت. برای این منظور دادگان غیربومی که در بخش ۲ توضیح داده شد مورد استفاده قرار گرفته و معیارهای مذکور از این دادگان استخراج می‌شود، سپس مقادیر این معیارها با نمراتی که ۹ داور انسانی به روانی گفتار زبان آموزان داده‌اند مقایسه می‌شوند. جدول ۵ همبستگی بین نمرات داوران انسانی و معیارهای کمی روانی گفتار را نشان می‌دهد.

شکل ۲ نیز میانگین قدر مطلق همبستگی هر معیار را با ۹ داور انسانی نمایش می‌دهد. این شکل در واقع قدر مطلق ستون میانگین در جدول ۵ را نمایش می‌دهد. مقادیر منفی در این جدول نشان دهنده وجود ارتباط معکوس معیارها با روانی گفتار (نمرات داوران انسانی) می‌باشد. بعنوان مثال تعداد وقفه‌ها (#P) هر چقدر بیشتر باشد، گفتار بیان شده از روانی و سلیسی کمتری برخوردار است. با دقت در این جدول مشاهده می‌شود که میانگین همبستگی هر معیار با نمرات داوران، قابل قبول بوده و بین معیارها و ارزیابی‌های داوران انسانی همبستگی بالایی وجود دارد. بیشترین همبستگی مربوط به معیار سرعت گفتار (ROS) می‌باشد. شاید این معیار با ذهنیت ما در مورد روان صحبت کردن نیز مطابقت داشته باشد. چرا که معمولاً اگر فردی روان و سلیس صحبت کند، سرعت گفتار وی بایستی مناسب باشد و در صورتیکه بیان فرد روان نباشد گفتار او شامل مکث‌ها، وقفه‌ها، مین کردن‌ها و چیزهای مشابه اینها خواهد بود و در نتیجه سرعت گفتار وی کاهش پیدا خواهد کرد.

در نتیجه در این مثال معیار میانگین طول وقفه‌ها نمی‌تواند معیار مناسبی باشد. بطور کلی اگر چه معیار سرعت گفتار (ROS) معیار مناسب‌تری می‌باشد، اما نمی‌تواند به تنهایی پوشش دهنده تمام خصوصیات روانی گفتار باشد. شاید با در نظر گرفتن تعداد وقفه‌ها در کنار سرعت گفتار اطلاعات بیشتری بدست آید. سطر آخر جدول ۵ میانگین قدر مطلق همبستگی بین هر داور و معیارها را نشان می‌دهد. با دقت در این سطر مشاهده می‌شود که میانگین قدر مطلق همبستگی بین معیارهای کمی تعریف شده با هر داور مقدار قابل قبولی می‌باشد. بیشترین همبستگی مربوط به داوران ۲، ۵ و ۸ می‌باشد. میانگین سطر آخر برابر ۰/۹۰ می‌باشد این مقدار که در واقع میانگین میانگین‌های قدر مطلق همبستگی بین معیارها و هر داور می‌باشد، نشان‌دهنده این مطلب است که میانگین همبستگی که بین کل معیارها و هر داور انسانی وجود دارد برابر ۰/۹۰ می‌باشد که مقدار قابل قبولی می‌باشد. نتیجه مشابهی با میانگین‌گیری از ستون آخر جدول ۵ بدست می‌آید. میانگین این ستون نشان دهنده این واقعیت است که میانگین همبستگی بین هر معیار با داور انسانی برابر ۰/۹۰ می‌باشد.

میانگین قدر مطلق همبستگی هر معیار با امتیازات داوران در شکل ۲ نمایش داده شده است که بطور کلی می‌توان گفت تقریباً تمام همبستگی‌ها قابل قبول هستند و این معیارها با روانی گفتار به نوعی در ارتباط می‌باشند و در نتیجه می‌توان از این معیارها برای پیش‌بینی و ارزیابی خودکار روانی گفتار استفاده نمود.

در جدول ۶ همبستگی بین هر دو معیار آمده است. با توجه به جدول، مشاهده می‌شود که بین همه معیارها نسبت به هم تقریباً همبستگی بالایی وجود دارد. در این جدول، بالاترین همبستگی‌ها مربوط به معیار سرعت گفتار (ROS) و کمترین همبستگی‌ها مربوط به معیار کل زمان وقفه‌ها (TDP) می‌باشد. قابل ذکر است علی‌رغم وجود همبستگی‌های بالا بین معیارها، این معیارها متفاوت می‌باشند.

به منظور بررسی بیشتر، میانگین و انحراف معیار معیارهای کمی روانی گفتار و امتیازات داوران انسانی مربوط به دادگان غیربومی و مجموعه گفتار بومی بخش ۲-۱ محاسبه شد و نتایج آن در جدول ۷ آمده است. در این جدول NS مربوط به گویندگان بومی و NNS مربوط به گویندگان غیربومی می‌باشد. با دقت در این جدول مشاهده می‌شود که معیارهای کمی روانی گفتار برای گویندگان بومی و غیربومی مقادیر متفاوتی دارند. به عبارت دیگر، گویندگان بومی گفتاری سریع‌تری داشته و در گفتار خود وقفه‌های کمتری تولید می‌کنند.

۴- ارزیابی خودکار روانی گفتار

ستون آخر (میانگین) در جدول ۹ و شکل ۳ همبستگی کلی ارزیابی انسانی و ارزیابی خودکار را نمایش می‌دهد. با توجه به نتایج بدست آمده مشاهده می‌شود که معیار سرعت گفتار (ROS) بهترین ویژگی برای پیش‌بینی روانی گفتار می‌باشد. این مطلب با نتایج مقایسه همبستگی معیارهای کمتی روانی گفتار و نمرات داوران انسانی که در جدول ۵ آمده است، مطابقت دارد. کمترین همبستگی نیز مربوط به معیار، میانگین طول تداوم (MLR) می‌باشد. با توجه به جدول ۳، مشاهده می‌شود که ۹ داور انسانی در ارزیابی یک دادگان غیربومی همبستگی‌هایی نزدیک ۰/۹۵ داشتند. در واقع اگر یک ارزیابی کننده دیگر با داوران مذکور در ارزیابی دادگان مشابه نزدیک به این همبستگی را داشته باشد، می‌توان نتیجه گرفت که ارزیابی کننده مذکور مانند ارزیابی کننده‌های انسانی عمل نموده است. در جدول ۸ برای معیار سرعت گفتار (ROS) همبستگی ۰/۹۳ بدست آمد که نزدیک به ارزیابی انسانی می‌باشد.

می‌توان برای ارزیابی خودکار روانی گفتار از ترکیب معیارهای مختلف کمتی تعریف شده نیز استفاده نمود. در واقع در این حالت برای تخمین روانی گفتار از ویژگی‌های بیشتری استفاده می‌شود و بعد ورودی افزایش می‌یابد. برای دسته‌بندی، میانگین معیارهای استفاده شده بصورت برداری در نظر گرفته می‌شود و بردار ناشناخته به دسته‌ای تعلق می‌گیرد که فاصله کمتری با آن دسته دارد. بنابراین خواهیم داشت:

$$\min_i \sum_{j=1}^n |V_j - M_{i,j}| \quad 1 \leq i \leq 5 \quad (9)$$

که n تعداد معیارهای استفاده شده و $M_{i,j}$ میانگین معیار j ام مربوط به دسته i ام می‌باشد (جدول ۸). بردار V نیز مربوط به ویژگی‌های استخراج شده از گفتار می‌باشد. در اینصورت بردار V به دسته i ام تعلق خواهد داشت. در جدول ۱۰ ترکیب دوتایی معیارها و نتایج آنها آمده است. در این جدول معیارهای هر ردیف با معیاری از هر ستون ترکیب شده و نتیجه همبستگی ارزیابی انسانی و ارزیابی خودکار به کمک دو معیار مربوطه آمده است. در اینجا نیز ابتدا همبستگی‌های بین ارزیابی‌های خودکار و تک تک داوران انسانی محاسبه شده و سپس میانگین آنها برای هر مورد منظور می‌شود.

در بخش قبل معیارهای کمتی تعریف شده در [۱] مورد بررسی قرار گرفت و مناسب بودن آنها برای تخمین و ارزیابی روانی گفتار برای زبان فارسی مشاهده شد. در این بخش سعی خواهد شد با استفاده از معیارهای مذکور به ارزیابی روانی گفتار گویندگان غیربومی بپردازیم. با توجه به دیاگرام انجام مراحل ارزیابی خودکار روانی گفتار در شکل ۱، بعد از استخراج معیارهای کمتی روانی گفتار، توسط دسته‌بندی کننده ارزیابی روانی خودکار انجام می‌گیرد. خروجی ارزیابی خودکار روانی مشابه ارزیابی‌های داوران انسانی دارای ۵ کلاس خواهد بود که در جدول ۲ آمده است. برای تعیین سطوح آستانه و دسته‌بندی سطح روانی گفتار با توجه به معیارهای کمتی می‌توان از میانگین هر دسته مربوط به معیار کمتی استخراج شده استفاده نمود. بدین صورت که دادگانی که توسط داوران انسانی مورد ارزیابی قرار گرفته است، در نظر گرفته می‌شود و سپس میانگین هر دسته محاسبه می‌شود. در نتیجه برای هر دسته از سطح روانی گفتار مقدار میانگین معیار مربوط به آن دسته در اختیار خواهد بود. حال برای دسته‌بندی مقدار یک معیار به سطح روانی مربوطه کفایت فاصله مقدار معیار مذکور با تک تک میانگین هر دسته محاسبه شود، در نهایت به دسته‌ای تعلق خواهد گرفت که به آن نزدیک‌تر باشد. بنابراین خواهیم داشت:

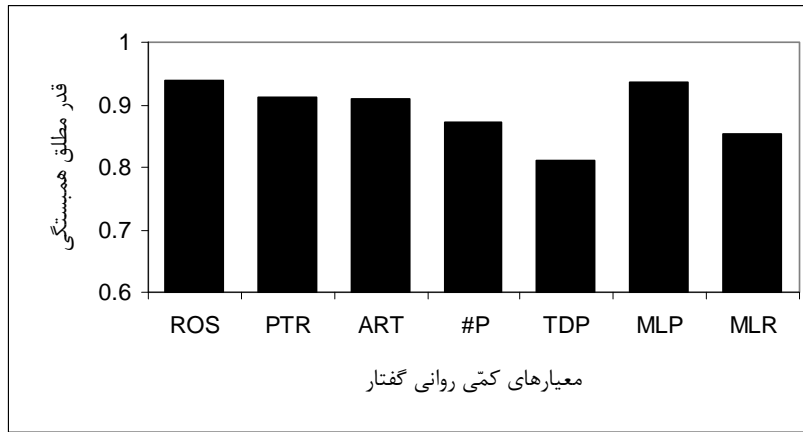
$$\min_i |V_x - M_{x,i}| \quad 1 \leq i \leq 5 \quad (8)$$

که V_x و $M_{x,i}$ به ترتیب برابر میانگین معیار x مربوط به دسته i ام و مقدار معیار x که از گفتار زبان آموز استخراج شده است، می‌باشد. بعنوان مثال، $\min_i |V_{ROS} - M_{ROS,i}|$ دسته‌بندی را برای معیار سرعت گفتار (ROS) انجام می‌دهد. مقدار i نتیجه دسته‌بندی را مشخص می‌نماید. جدول ۸ میانگین دسته‌های مختلف روانی گفتار را برای هر معیار نشان می‌دهد.

همانطور مشاهده می‌شود در اینجا برای تخمین و ارزیابی روانی گفتار فقط از یک معیار استفاده شده است. برای بدست آوردن مقادیر میانگین مربوط به هر معیار تقریباً نیمی از دادگان غیربومی استفاده شد. جدول ۹ نتیجه دسته‌بندی نیمی دیگر از دادگان و همبستگی آن را با امتیازات داوران انسانی نمایش می‌دهد.

جدول ۵- همبستگی بین نمرات داوران انسانی و معیارهای کمتی روانی گفتار

	داور ۱	داور ۲	داور ۳	داور ۴	داور ۵	داور ۶	داور ۷	داور ۸	داور ۹	میانگین
ROS	۰/۸۹۷۳	۰/۹۹۰۵	۰/۹۳۵۹	۰/۸۶۴۰	۰/۹۹۰۵	۰/۹۷۵۷	۰/۹۱۱۸	۰/۹۹۰۵	۰/۸۹۷۳	۰/۹۳۹۳
PTR	۰/۹۷۵۲	۰/۹۴۱۰	۰/۸۷۷۴	۰/۸۵۲۲	۰/۹۴۱۰	۰/۸۷۳۹	۰/۸۴۰۵	۰/۹۴۱۰	۰/۹۷۵۲	۰/۹۱۳۰
ART	۰/۸۰۰۵	۰/۹۶۷۵	۰/۹۱۵۲	۰/۸۵۴۸	۰/۹۶۷۵	۰/۹۷۹۴	۰/۹۲۴۹	۰/۹۶۷۵	۰/۸۰۰۵	۰/۹۰۸۶
#P	-۰/۹۳۱۹	-۰/۸۹۹۷	-۰/۸۳۹۱	-۰/۸۱۱۳	-۰/۸۹۹۷	-۰/۸۳۷۲	-۰/۸۰۲۳	-۰/۸۹۹۷	-۰/۹۳۱۹	-۰/۸۷۲۵
TDP	-۰/۸۵۵۸	-۰/۸۲۴۳	-۰/۷۴۸۶	-۰/۸۲۸۵	-۰/۸۲۴۳	-۰/۷۵۰۳	-۰/۷۸۸۱	-۰/۸۲۴۳	-۰/۸۵۵۸	-۰/۸۱۱۱
MLP	-۰/۹۷۴۰	-۰/۹۶۱۶	-۰/۸۸۴۵	-۰/۹۰۸۶	-۰/۹۶۱۶	-۰/۸۹۸۸	-۰/۸۹۵۰	-۰/۹۶۱۶	-۰/۹۷۴۰	-۰/۹۳۵۵
MLR	۰/۸۰۶۵	۰/۹۱۴۰	۰/۸۶۲۰	۰/۶۸۹۷	۰/۹۱۴۰	۰/۹۵۶۳	۰/۸۰۸۱	۰/۹۱۴۰	۰/۸۰۶۵	۰/۸۵۲۳
میانگین قدر مطلق	۰/۸۹۱۶	۰/۹۲۸۴	۰/۸۶۶۲	۰/۸۲۹۹	۰/۹۲۸۴	۰/۸۹۶۰	۰/۸۵۳۰	۰/۹۲۸۴	۰/۸۹۰۰	



شکل ۲- میانگین قدر مطلق همبستگی هر معیار با ۹ داور انسانی

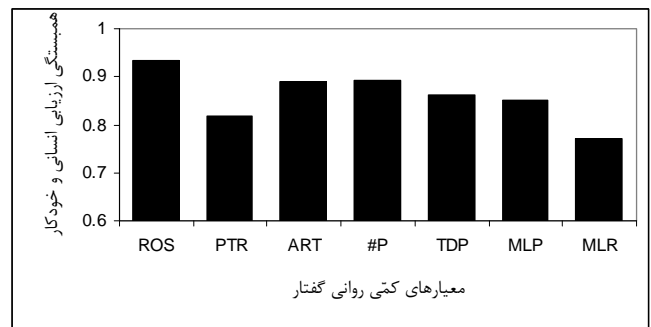
جدول ۶- همبستگی بین معیارهای مختلف روانی گفتار

MLR	MLP	TDP	#P	ART	PTR	ROS	
۰/۹۵۳۶	-۰/۹۳۹۴	-۰/۸۲۲۱	-۰/۹۰۸۴	۰/۹۸۱۸	۰/۹۳۶۳	۱	ROS
۰/۸۲۷۳	-۰/۹۷۸۳	-۰/۹۳۷۰	-۰/۹۸۵۲	۰/۸۵۸۶	۱		PTR
۰/۹۵۴۱	-۰/۸۷۳۸	-۰/۷۴۷۱	-۰/۸۳۱۶	۱			ART
-۰/۷۹۸۴	۰/۹۳۲۵	۰/۹۷۰۰	۱				#P
-۰/۶۵۸۸	۰/۸۸۲۳	۱					TDP
-۰/۸۲۱۴	۱						MLP
۱							MLR

برای ترکیب سه معیار مشخص شده است. نتایج بدست آمده نشان می‌دهد که افزودن معیار کل زمان وقفه‌ها (TDP) به دو معیار اخیر فقط اندکی کارایی را افزایش می‌دهد. افزودن معیارهای دیگر نیز یا موجب کاهش کارایی می‌گردند و یا کارایی را تغییر نمی‌دهند.

جدول ۷- میانگین و انحراف معیار امتیازات داوران انسانی و معیارهای کمی روانی گفتار در دادگان غیربومی و مجموعه گفتار بومی

sd NNS	\bar{x} NSS	sd NS	\bar{x} NS	
۰/۹۸	۴/۱	۰/۳۵	۲/۵	داوران انسانی
۳/۸۷	۸/۱۷	۲/۲۳	۱۲/۳۳	ROS
۱۷/۸۵	۷۴/۱۲	۸/۷۷	۸۵/۴۵	PTR
۲/۸۲	۱۰/۶۱	۱/۱۹	۱۲/۰۱	ART
۱/۳۲	۱/۴۵	۰/۲۲	۰/۷۸	#P
۱/۸۷	۱/۳۹	۰/۱۲	۰/۲۱	TDP
۰/۵۵	۰/۵۸	۰/۲۵	۰/۳۲	MLP
۱۵/۰۳	۲۰/۵۶	۹/۹۶	۳۴/۲۶	MLR



شکل ۳- همبستگی ارزیابی انسانی و ارزیابی خودکار توسط معیارهای مختلف روانی گفتار

در جدول ۱۰ مشاهده می‌شود که ترکیب دو معیار سرعت گفتار (ROS) و سرعت تعامل (ART) بیشترین کارایی را نتیجه می‌دهد. این مقدار (۰/۹۴۱۳) از همبستگی‌های بدست آمده در ارزیابی خودکار با استفاده از یک معیار که در جدول ۹ آمده است، بیشتر می‌باشد.

از آنجائیکه در ترکیب دو تایی معیارها، بیشترین کارایی برای دو معیار سرعت گفتار (ROS) و سرعت تعامل (ART) بدست آمد، حال با افزودن معیارهای دیگر به این دو معیار نتایج را بررسی می‌کنیم. نتیجه بدست آمده از این ترکیب در جدول ۱۱ آمده است. در این جدول معیارهای دیگر به این دو اضافه شده و نتایج

نتایج حاصل شده از ارزیابی خودکار روانی نشان می‌دهد که ارزیابی کننده خودکار نیز می‌تواند مشابه یک داور انسانی عمل نموده و روانی گفتار را با دقتی مشابه ارزیابی نماید. چرا که همبستگی‌های بدست آمده برای ارزیابی خودکار نزدیک به همبستگی‌های بین داوران انسانی می‌باشد.

جدول ۱۰- همبستگی بین ارزیابی انسانی و ارزیابی خودکار روانی گفتار با استفاده از ترکیب دو به دو معیارهای کمی روانی گفتار

MLR	MLP	TDP	#P	ART	PTR	+
۰/۸۸۰۵	۰/۹۳۷۳	۰/۹۳۷۳	۰/۹۳۷۲	۰/۹۴۱۳	۰/۷۲۴۶	ROS
۰/۸۵۹۷	۰/۸۱۸۶	۰/۸۱۸۶	۰/۸۱۸۶	۰/۸۶۲۳		PTR
۰/۸۶۹۸	۰/۸۷۵۶	۰/۸۹۷۵	۰/۸۹۷۵			ART
۰/۸۳۹۷	۰/۹۰۷۷	۰/۹۰۳۵				#P
۰/۸۳۹۷	۰/۸۹۷۳					TDP
۰/۸۳۹۵						MLP

جدول ۱۱- همبستگی بین ارزیابی انسانی و ارزیابی خودکار روانی گفتار با استفاده از ترکیب سه تایی معیارهای کمی روانی گفتار

MLR	MLP	TDP	#P	PTR	+
۰/۸۳۲۷	۰/۹۴۱۳	۰/۹۴۳۵	۰/۹۳۸۸	۰/۷۴۵۴	ART.ROS

مراجع

[1] C. Cucchiari, H. Strik, and L. Boves, "Quantitative Assessment of Second Language Learners' Fluency by Means of Automatic Speech Recognition Technology," *Acoustical Society of America*, Vol. 107, pp. 989-999, 2000.

[2] R. Leeson, *Fluency and Language Teaching*, London, Longman, 1975.

جدول ۸- میانگین دسته‌های مختلف روانی گفتار برای هر معیار روانی گفتار

دسته ۱	دسته ۲	دسته ۳	دسته ۴	دسته ۵	
۱۴/۸۶	۱۱/۹۴	۹/۰۲	۷/۴۰	۴/۱۴	ROS
۹۲/۸۳	۹۰/۲۸	۷۷/۹۰	۷۸/۶۰	۵۳/۸۳	PTR
۱۶/۰۱	۱۳/۲۰	۱۱/۵۹	۹/۴۴	۷/۸۶	ART
۰	۰/۳۳	۱/۱۴	۱/۲۵	۳	#P
۰	۰/۱۰	۰/۵۱	۱/۵۱	۳/۵۸	TDP
۰	۰/۰۸	۰/۴۲	۰/۴۱	۱/۲۲	MLP
۴۱/۴۱	۳۹/۲۹	۲۰/۸۵	۱۱/۳۷	۸/۶۲	MLR

۵- نتیجه گیری

در این مقاله به ارزیابی روانی خودکار گفتار با استفاده از معیارهای زمانی روانی گفتار پرداختیم. همانطور که در بخش ۱ توضیح داده شد، تعاریف مختلفی برای روانی گفتار وجود دارد. در اینجا جنبه‌های زمانی روانی گفتار در آموزش زبان مورد بررسی قرار گرفت. معیارهای کمی زمانی متنوعی برای روانی گفتار توسط بسیاری از محققین تعریف شده است که هفت مورد از معروفترین آنها برای زبان فارسی مورد بررسی قرار گرفت. نتایج ارزیابی داوران انسانی در مورد روانی گفتار دادگان بومی و غیربومی نشان داد که داوران در ارزیابی این دادگان با هم سازگاری مناسبی داشتند و گویندگان بومی نمرات بهتری از گویندگان غیربومی اخذ نمودند. این ارزیابی در ۵ سطح صورت گرفت. مناسب بودن معیارهای روانی تعریف شده برای زبان فارسی، با مقایسه امتیازات داوران انسانی مشاهده شد. به کمک این معیارها توانستیم روانی گفتار را پیش‌بینی نموده و ارزیابی روانی گفتار را بصورت خودکار انجام دهیم.

معیار سرعت گفتار (ROS) کارایی بالاتری را از دیگر معیارها از خود نشان داد. مشاهده شد که افزودن یک یا دو معیار دیگر باعث افزایش کارایی این معیار شد تا حدی که به نتایج همبستگی بین داوران انسانی بسیار نزدیک‌تر شد. همچنین مشاهده شد که افزودن معیارهای بیشتر موجب افزایش کارایی ارزیابی خودکار نگردد. که می‌توان نتیجه گرفت در بررسی جنبه‌های زمانی روانی گفتار، معیار سرعت گفتار ابعاد بسیاری از این مسأله را تحت پوشش خود قرار می‌دهد و معیار مناسبی می‌باشد.

جدول ۹- همبستگی بین نمرات داوران انسانی و ارزیابی خودکار روانی گفتار با استفاده از معیارهای مختلف کمی روانی گفتار

میانگین	داور ۱	داور ۲	داور ۳	داور ۴	داور ۵	داور ۶	داور ۷	داور ۸	داور ۹
۰/۹۳۲۱	۰/۹۲۱۲	۰/۹۴۲۳	۰/۹۳۳۵	۰/۹۳۶۷	۰/۹۳۸۸	۰/۹۱۴۴	۰/۹۲۹۸	۰/۹۳۱۱	۰/۹۳۴۵
۰/۸۵۳۴	۰/۸۲۸۵	۰/۸۳۲۱	۰/۷۸۳۵	۰/۷۷۳۷	۰/۸۴۲۳	۰/۷۹۵۶	۰/۸۰۳۳	۰/۸۱۳۲	۰/۸۱۸۶
۰/۸۹۲۳	۰/۸۸۰۱	۰/۸۸۵۵	۰/۸۹۲۳	۰/۹۰۶۵	۰/۹۲۶۴	۰/۹۱۴۵	۰/۸۷۴۵	۰/۹۰۷۶	۰/۸۹۰۲
۰/۸۸۲۳	۰/۸۷۸۹	۰/۸۹۳۴	۰/۹۰۳۴	۰/۸۸۴۳	۰/۹۲۸۷	۰/۹۱۵۶	۰/۸۹۳۳	۰/۹۰۱۲	۰/۸۹۲۱
۰/۸۵۲۳	۰/۸۶۹۹	۰/۸۴۷۳	۰/۸۵۱۲	۰/۸۷۲۳	۰/۸۸۱۲	۰/۹۰۴۳	۰/۸۴۲۳	۰/۸۸۳۴	۰/۸۶۳۱
۰/۸۶۴۵	۰/۸۶۲۳	۰/۸۸۳۴	۰/۸۲۱۱	۰/۸۵۶۶	۰/۸۵۳۴	۰/۸۴۸۷	۰/۸۴۵۵	۰/۸۵۴۴	۰/۸۵۰۶
۰/۷۸۳۳	۰/۷۶۹۹	۰/۷۹۵۵	۰/۸۰۲۲	۰/۷۷۷۱	۰/۷۶۲۲	۰/۷۵۲۲	۰/۷۶۳۳	۰/۷۶۲۲	۰/۷۷۲۰

[18] K. Audhkhasi, K. Kandhway, O. D. Deshmukh, and A. Verma, "Formant-based technique for automatic filled-pause detection in spontaneous spoken English," *Proc, IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, pp. 4857-4860, 2009.

[19] O. D. Deshmukh, K. Kandhway, A. Verma, and K. Audhkhasi, "Automatic evaluation of spoken English fluency," *Proc, IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, pp. 4829-4832, 2009.

[20] K. Zechner, and I. Bejar, "Towards automatic scoring of non-native spontaneous speech," *Proc, Human Language Technologies Conf.*, pp. 216-223, 2006.



بهزاد بختیاری در سال‌های ۱۳۷۸ تا ۱۳۸۲ به تحصیل در مقطع کارشناسی مهندسی کامپیوتر گرایش نرم‌افزار مشغول و در سال‌های ۱۳۸۲ تا ۱۳۸۴ در دانشگاه صنعتی امیرکبیر در مقطع کارشناسی‌ارشد مهندسی کامپیوتر گرایش هوش مصنوعی به تحصیل پرداخته است. نامبرده هم‌اکنون عضو هیات علمی گروه کامپیوتر موسسه آموزش عالی غیردولتی - غیرانتفاعی سجاد است. زمینه‌های مورد علاقه ایشان پردازش سیگنال گفتار، شبکه‌های عصبی و پردازش تکاملی می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

bakhtiari@sadjad.ac.ir



محمد مهدی همایون‌پور در سال ۱۳۳۹ در شهر شیراز متولد شد. تحصیلات تا مقطع دیپلم را در شهر شیراز سپری و دیپلم متوسطه خود را در سال ۱۳۵۸ دریافت کرد. وی تحصیلات خود در مقطع کارشناسی را در رشته مهندسی برق (الکترونیک) در دانشگاه صنعتی امیرکبیر (سال ۱۳۶۶)، کارشناسی‌ارشد را در رشته برق (مخابرات)، از دانشگاه خواجه‌نصیرالدین طوسی (سال ۱۳۶۹)، کارشناسی‌ارشد دوم خود را در زمینه فونیتیک (۱۳۷۴) در دانشگاه سوربون جدید در فرانسه و همزمان دوره دکترا خود را در دانشگاه پاریس ۱۱ در زمینه مهندسی برق (۱۳۷۴) بیابان رسانید. نامبرده از سال ۱۳۷۴ در سمت عضو هیات علمی دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه صنعتی امیرکبیر به تدریس و تحقیق مشغول می‌باشد. زمینه‌های تخصصی مورد علاقه ایشان شامل پردازش سیگنال‌های دیجیتال، بازشناسی گفتار و گوینده، تبدیل متن به گفتار، کدینگ گفتار، پردازش زبان طبیعی، تشخیص نفوذ در سیستم‌ها و شبکه‌های کامپیوتری، اتوماسیون صنعتی، چند رسانه‌ای و طراحی سخت‌افزار می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

homayoun@ce.aut.ac.ir

[3] C. J. Fillmore, "On Fluency in Individual Differences in Language Ability and Language Behavior," *Academic: New York*, pp. 85-101, 1979.

[4] C. Brumfit, *Communicative Methodology in Language Teaching: The Roles of Fluency and Accuracy*, Cambridge, Cambridge University Press, 1984.

[5] P. Lennon, "Investigating Fluency In EFL: A Quantitative Approach," *Language Learning* 3, pp. 387-417, 1990.

[6] R. Schmidt, "Psychological Mechanisms Underlying Second Language Fluency," *Stud. Second Language Acquisition*, Vol. 14, pp. 357-385, 1992.

[7] F. Chambers, "What do we mean by fluency," *System*, Vol. 25, No. 4, pp. 535-544, 1997.

[8] H. Riggenbach, "Toward an Understanding Of Fluency: A Microanalysis Of Non-Native Speaker Conversations," *Discourse Process*, Vol. 14, No. 4, pp. 423-441, 1991.

[9] T. J. Peters, and B. Guitar, *Stuttering. An Integrated Approach to its Nature and Treatment William and Wilkins*, Baltimore, 1991.

[10] P. Nation, "Improving Speaking Fluency," *System*, Vol. 17, No. 3, pp. 377-384, 1989.

[11] B. F. Freed, "What Makes Us Think That Students Who Study Abroad Become Fluent?," *Proc, Second Language Acquisition in a Study-Abroad Context*, pp. 123-148, 1995.

[12] R. Towell, R. Hawkins, and N. Bazergui, "The Development of Fluency in Advanced Learners of French," *Appl. Linguistics*, Vol. 17, No. 1, pp. 84-119, 1996.

[13] D. Mo'hle, "A Comparison Of The Second Language Speech Production Of Different Native Speakers," *Proc, Second Language Productions*, pp. 26-49, 1984.

[14] M. Eskenazi, "Using Automatic Speech Processing for Foreign Language Pronunciation Tutoring: Some and Prototype," *Language Learning and Technology*, Vol. 2, No. 2, pp. 62-76, 1999.

[15] M. Black, J. Tepperman, S. Lee, P. Price, and S. Narayanan, "Automatic detection and classification of disfluent reading miscues in young children's speech for the purpose of assessment," *Proc, Interspeech ICSLP*, pp. 206-209, 2007.

[16] K. Audhkhasi, "Automatic Evaluation of Fluency in Spoken Language," *IETE Technical Review*, Vol. 26, pp. 108-114, 2009.

[17] Y. Liu, E. Shriberg, A. Stolcke, D. Hillard, M. Ostendorf, and M. Harper, "Enriching speech recognition with automatic detection of sentence boundaries and disfluencies," *IEEE Trans. Speech and Audio Processing*, Vol. 14, No. 5, pp. 1526-40, September 2006.

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۶/۰۶/۰۴

تاریخ اصلاح: ۸۹/۰۱/۲۴

تاریخ قبول شدن: ۸۹/۰۳/۲۳

نویسنده مرتبط: دکتر محمد مهدی همایون‌پور، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران.

-
- ¹ CALL (Computer-Assisted Language Learning)
 - ² Segmental
 - ³ Supra Segmental
 - ⁴ Prosody
 - ⁵ Stress
 - ⁶ Intonation
 - ⁷ Fluency
 - ⁸ Rhythm
 - ⁹ Fillmore
 - ¹⁰ Communicative Language Teaching
 - ¹¹ Burmfit
 - ¹² Native-Speaker-Like Performance
 - ¹³ Temporal
 - ¹⁴ Silent Pauses
 - ¹⁵ Hesitations
 - ¹⁶ Filled Pauses
 - ¹⁷ Mohle
 - ¹⁸ Duration
 - ¹⁹ Correlation
 - ²⁰ Quantitative Measures
 - ²¹ Rate of Speech
 - ²² Phonation Time Ratio
 - ²³ Articulation Rate
 - ²⁴ Total Duration of Pauses
 - ²⁵ Mean Length of Pauses
 - ²⁶ Mean Length of Runs

اجتماع نمونه‌دهی هستان‌شناسی و حاشیه‌نویسی معنایی متون فارسی در سیستم POPTA

مهرنوش شمس‌فرد بهاره صراف‌زاده

دانشکده مهندسی برق و کامپیوتر، دانشگاه شهید بهشتی، تهران، ایران

چکیده

نمونه‌دهی خودکار هستان‌شناسی به استخراج نمونه‌های کلاس‌های یک هستان‌شناسی از متن و افزودن آنها به هستان‌شناسی می‌پردازد. از سوی دیگر وب معنایی با افزودن حاشیه‌نویسی به اسناد وب کنونی محتوای ساخت‌یافته‌ای را فراهم می‌نماید. حاشیه‌نویسی‌های معنایی داده‌های نمونه کلاس‌های هستان‌شناسی را برچسب‌گذاری نموده و آنها را بر کلاس‌های هستان‌شناسی منطبق می‌نمایند. بنابراین فعالیت نمونه‌دهی به هستان‌شناسی می‌تواند با ایجاد حاشیه‌نویسی‌های معنایی همراه باشد. در این مقاله یک سیستم نمونه‌دهی و حاشیه‌نویسی میان‌زبانی^۱ به نام POPTA^۲ معرفی می‌شود که متون فارسی را با توجه به یک هستان‌شناسی (با مداخل لغوی انگلیسی) حاشیه‌نویسی کرده و این هستان‌شناسی را نیز با توجه به این متون نمونه‌دهی می‌نماید. این سیستم از ترکیبی از روش‌های آماری و مبتنی بر الگو به همراه متدهای مبتنی بر وب، موتورهای جستجو و دایرالمعارف آنلاین ویکی‌پدیا به منظور نمونه‌دهی به هستان‌شناسی و حاشیه‌نویسی متون بهره می‌گیرد.

کلمات کلیدی: نمونه‌دهی هستان‌شناسی، حاشیه‌نویسی متون، پردازش زبان فارسی، یادگیری هستان‌شناسی، ویکی‌پدیا، وب معنایی.

۱- مقدمه

بنابراین فعالیت نمونه‌دهی به هستان‌شناسی با ایجاد حاشیه‌نویسی‌های معنایی همراه است. ساخت کاملاً خودکار حاشیه‌نویسی‌های معنایی مسئله‌ای حل نشده است. در عوض سیستم‌های امروزی بر ساخت نیمه‌خودکار حاشیه‌نویسی‌ها متمرکز شده‌اند [۱].

به عبارت دیگر با داشتن یک هستان‌شناسی دامنه (مفاهیم و روابط) و استخراج نمونه‌های این مفاهیم یا کلاس‌ها از یک متن، می‌توان این نمونه‌ها را در متن برچسب‌گذاری معنایی کرد (حاشیه‌نویسی) و یا این نمونه‌ها را تحت کلاس پدر در هستان‌شناسی متصل نمود (نمونه‌دهی).

تا کنون رویکردهایی برای حاشیه‌نویسی و یا نمونه‌دهی ارائه شده است. بطور کلی حاشیه‌نویسی معنایی، به یک هستان‌شناسی دارای دامنه و زبان یکسان با متن نیاز دارد. با وجود این که هستان‌شناسی‌ها پایگاه‌های دانش مفهومی هستند و بطور نظری باید مستقل از زبان باشند، هنگامی که از نمادها برای بازنمایی مفاهیم استفاده می‌کنیم، معمولاً از واژگان یک زبان برای نشان دادن مفاهیم بهره می‌بریم و بنابراین ویژگی استقلال از زبان را از دست می‌دهیم. به بیان دیگر عناصر

بعدی وب - که وب معنایی نامیده می‌شود - برای بهبود معنای داده‌ها از هستان‌شناسی‌ها بهره می‌گیرد. ساخت دستی این هستان‌شناسی‌ها بسیار زمان‌بر و پرهزینه است. به علاوه هستان‌شناسی‌های ساخته شده توسط انسان‌ها به شیوه تفکر و سلیق فردی آنها وابسته است. از این رو به دنبال بهره‌گیری از روش‌های خودکار یادگیری هستان‌شناسی (OL) و استخراج دانش می‌باشیم. یکی از جنبه‌های ساخت خودکار هستان‌شناسی‌ها تکمیل هستان‌شناسی با نمونه‌های مفاهیم موجود در آن است. این فعالیت نمونه‌دهی هستان‌شناسی^۳ (OP) نامیده می‌شود. از سوی دیگر تحقق یافتن وب معنایی به در دسترس بودن همه جانبه حاشیه‌نویسی معنایی برای اسناد موجود و جدید بر روی وب نیازمند است. حاشیه‌نویسی‌های معنایی داده‌های نمونه کلاس‌های هستان‌شناسی را برچسب‌گذاری نموده و آنها را بر کلاس‌های هستان‌شناسی منطبق می‌نمایند.

می‌شوند، حامل یک رابطه معنایی خاص هستند. این سیستم از الگوهای هیرست^۹ [۷] استفاده می‌کند سیستم معرفی شده در [۸] با استفاده از الگوهای رابطه خاص قلمرو دست‌ساز - که تعمیم یافته الگوهای هیرست هستند - به نمونه‌دهی یک هستان‌شناسی می‌پردازد. این الگوریتم از نمونه‌های یک کلاس که توسط گوگل برگردانده شده‌اند، برای یافتن نمونه‌های کلاس‌های دیگر بهره می‌گیرد.

بانسکو و پاسکا [۹] از ویکی‌پدیا به منظور ایجاد یک روش رفع ابهام موجودیت‌های نامدار استفاده کرده و با استفاده از یک دیکشنری، اسامی خاص را به موجودیت‌های نامدار مربوط به آنها متصل می‌نمایند. سیستم آنها برای استخراج موجودیت‌های نامدار از صفحه‌های تغییر مسیر و رفع ابهام و نیز سیستم رده‌ها و پیوندهای ویکی‌پدیا بهره می‌گیرد. چرنو و همکارانش [۱۰] از این ایده استفاده می‌کنند که طبقات یا رده‌هایی که ارتباطات و پیوندهای زیادی با هم دارند، دارای روابط معنایی با هم هستند.

معیار دوم: نمونه‌های استخراج شده آنچه تحت عنوان نمونه از منبع ورودی استخراج و به مفاهیم متناظر در هستان‌شناسی متصل می‌شود، ممکن است در سیستم‌های مختلف متفاوت باشد. دو گونه متداول از این نمونه‌ها صفحات وب و موجودیت‌های نامدار متون زبان طبیعی هستند. در گونه اول در یک قلمرو مشخص، صفحات وب نمونه‌هایی از هستان‌شناسی آن دامنه هستند. از این رو ما نیازمند ابزارهایی نیمه‌خودکار برای یافتن این نمونه‌ها و روابط میان آنها و مقداردهی به ویژگی‌های آنها می‌باشیم. برای مثال سیستم OILSW [۱۱] از صفحات وب به عنوان نمونه استفاده می‌نماید و هر صفحه را به مفهوم معادلش در یک هستان‌شناسی پیش فرض متصل می‌نماید.

در گونه دوم نمونه‌ها، عناصر متنی (مانند کلمات و عبارات) هستند که متناظر با نمونه‌های جهان خارجند و باید به مفهوم معادل خود در هستان‌شناسی مربوطه متصل شوند. برای مثال در سیستم KMi رویدادها به عنوان کلاس‌های هستان‌شناسی تعریف شده‌اند. در این سیستم ابتدا توپولوژی رویداد مستقیماً در هستان‌شناسی تعریف می‌شود. سپس برای هر رویداد، اسلات‌هایی تعریف می‌گردد که ممکن است با یک مولفه استخراج اطلاعات نمونه‌دهی شوند. هدف پرکردن خودکار بیشترین تعداد ممکن از حفره‌هاست.

معیار سوم: رویکرد نمونه‌دهی - رویکردهای نمونه‌دهی به دسته‌های اصلی مبتنی بر الگو، آماری و مبتنی بر مکاشفه تقسیم می‌شوند. در رویکردهای مبتنی بر الگو، نمونه‌دهی هستان‌شناسی با استفاده از الگوها و یا برپایه ساختار ترم‌ها انجام می‌شود. این رویکردها بدنبال عباراتی هستند که بوضوح وجود یک رابطه is-a را بین دو کلمه نشان می‌دهند، مانند "The ant is an insect" یا "ants and other insects". الگوهای هیرست [۷] نمونه‌های معروفی برای استخراج روابط is-a از متن هستند که در بسیاری سیستم‌ها [۶، ۱۲] مورد استفاده قرار گرفته‌اند.

در رویکردهای مبتنی بر مکاشفه که در بسیاری موارد به صورت ترکیبی با رویکردهای دیگر مورد استفاده قرار می‌گیرند، قوانین مکاشفه‌ای برای استخراج روابط شمول بکار می‌روند. برای مثال [۱۰] از یک مکاشفه انطباق هسته^{۱۰} استفاده می‌کند که بر اساس آن میان هسته یک عبارت و کل عبارت می‌توان یک رابطه is-a برقرار نمود؛ مثلاً "Christmas tree" یک نوع از "tree" است.

در روش‌های آماری از ویژگی‌های وابسته به متن استفاده می‌شود [۱۴]. این رویکردها از یک پیکره برای استخراج ویژگی‌هایی از متن که یک کلاس معنایی را آشکار می‌سازد، استفاده می‌کنند. این ویژگی‌ها می‌توانند ظاهری یا ساختاری باشند. ارزیابی انجام‌شده برای مقایسه نشان داده است که ویژگی‌های ساختاری به کارایی بهتری منجر می‌شوند. همچنین در بسیاری موارد سیستم‌های نمونه‌دهی از ترکیبی از رویکردهای فوق بهره می‌گیرند. سیستم معرفی شده در [۱۲] از الگوهای هیرست، سلسله مراتب وردنت^{۱۱}، مکاشفه‌های انطباق هسته، استنتاج

هستان‌شناسی با یک زبان خاص (معمولاً انگلیسی) واژه‌سازی^۵ می‌شوند و نمی‌توانند مستقیماً برای پردازش متون فارسی بکار روند.

بنابراین از سویی برای حاشیه‌نویسی متون فارسی نیاز به هستان‌شناسی‌های واژه‌سازی شده با واژه‌های فارسی (یا هستان‌شناسی‌های فارسی) داریم که در حال حاضر موجود نیستند و لازم است بطور دستی یا خودکار تولید شوند. از سوی دیگر یکی از راه‌های یادگیری و نمونه‌دهی هستان‌شناسی‌های فارسی استفاده از پیکره‌های برچسب خورده معنایی است که متأسفانه برای زبان فارسی تهیه نشده‌اند. این مسئله منجر به ایجاد حلقه‌ای شده است که برای رفع آن، در این مقاله به ارائه یک رویکرد موازی میان زبانی^۶ می‌پردازیم.

در این رویکرد از وب نیز به عنوان بزرگترین و گسترده‌ترین پیکره موجود، از گوگل^۷ به عنوان قدرتمندترین موتور جستجوی کنونی و از ویکی‌پدیا^۸ به عنوان یک پایگاه اطلاعاتی جامع که به تازگی صفحات فارسی آن نیز در حال گسترش است به منظور افزایش کارایی و دقت سیستم و نیز غلبه بر گلوگاه اکتساب دانش بهره گرفتیم.

POPTA نه تنها نخستین سیستم حاشیه‌نویسی متون فارسی و نمونه‌دهی هستان‌شناسی از روی متون فارسی است، بلکه جزء معدود سیستم‌هایی در جهان است که دو فرآیند نمونه‌دهی و حاشیه‌نویسی را به صورت میان زبانی و بطور توأم انجام می‌دهد.

علاوه بر ویژگی‌های گفته شده، سیستم POPTA خود را بر همه فرمت‌های کاراکترهای موجود در کلمات فارسی منطبق نموده و برای کار با هر منبع، فرمت کلمات ورودی را با فرمت کلمات موجود در آن منبع تطبیق می‌دهد.

در ادامه این مقاله در بخش ۲ مروری بر کارهای صورت گرفته برای نمونه‌دهی و حاشیه‌نویسی هستان‌شناسی‌ها خواهیم داشت، در بخش ۳ سیستم POPTA را به عنوان یک سیستم نمونه‌دهی و حاشیه‌نویسی فارسی معرفی نموده و در بخش ۴ به ارزیابی آن می‌پردازیم. در آخر نیز نگاهی به رویکردهای آینده خواهیم داشت.

۲- کارهای مرتبط

در این بخش به معرفی فعالیت‌های انجام شده در حوزه نمونه‌دهی و حاشیه‌نویسی معنایی خواهیم پرداخت. با مطالعه فعالیت‌های انجام شده در حوزه نمونه‌دهی هستان‌شناسی‌ها چهار معیار نوع منبع ورودی، نمونه‌های استخراج شده، رویکرد نمونه‌دهی و محصول نهایی را جهت دسته بندی روش‌های مختلف استخراج نمونه‌ایم [۲]. در ادامه این بخش ابتدا به دسته بندی روشهای موجود با توجه به این معیارها پرداخته و در هر دسته مثال هایی از سیستم های نمونه‌دهی ارائه خواهیم نمود.

معیار اول: منبع ورودی - منظور از منبع ورودی منبعی است که نمونه‌ها از آن استخراج می‌شوند. این منبع می‌تواند یک پیکره متنی (مانند منبع ورودی در سیستم KMi [۳]) و یا صفحات موجود در وب (مانند منبع ورودی در [۴، ۵]) باشد. در میان سیستم‌هایی که عمل نمونه‌دهی و حاشیه‌نویسی را به کمک وب انجام می‌دهند، گروهی از سیستم‌ها از سرویس‌های موجود در سطح وب - مانند موتور جستجوی گوگل - بهره گرفته و گروهی دیگر از منابع اطلاعاتی موجود در سطح وب - مانند دایره‌المعارف آنلاین ویکی‌پدیا - برای استخراج دانش مورد نیاز خود استفاده کرده‌اند. OntoGenie [۴] یک ابزار نیمه اتوماتیک است که هستان‌شناسی‌های خاص قلمرو و داده‌های غیر ساختاریافته در وب را به عنوان ورودی دریافت کرده و نمونه‌های هستان‌شناسی را تولید می‌نماید. این ابزار از هستان‌شناسی زبانی وردنت به عنوان پلی بین هستان‌شناسی‌های خاص قلمرو و داده‌های وب بهره می‌گیرد. PANKOW [۶] سیستم حاشیه‌نویسی و نمونه‌دهی دیگری است که معتقد است الگوهای ساختارهای مشخص که در متن پیدا

مولفه استخراج اطلاعات حاشیه‌نویسی معنایی با استفاده از مولفه‌های ابزار GATE اعمال می‌شود.

سیستم دیگر، کار ارائه شده توسط کاسادو و همکارانش [۲۲، ۲۳] است. در این کار روشی برای حاشیه‌نویسی خودکار روابط معنایی مختلف در ویکی‌پدیا ارائه شده است. این فرآیند مبتنی بر کشف و تعمیم خودکار الگوهای واژگانی بوده و امکان تشخیص روابط موجود میان مفاهیم را می‌دهد. بدین منظور از یک رویکرد مبتنی بر یادگیری خودکار الگوهای واژی- نحوی^{۱۴} بهره گرفته شده. این فرآیند با فهرست اولیه‌ای که شامل زوج‌هایی از آیتم‌های مرتبط است، آغاز می‌گردد. سپس جملات بسیاری که دربرگیرنده این زوج‌ها هستند بطور خودکار از وب جمع‌آوری شده و توسط ابزارهای NLP مانند قطعه‌بند^{۱۵}، برچسب‌ن مقله نحوی، ریشه‌یاب و تشخیص‌دهنده موجودیت‌های نامدار پردازش می‌شوند. اطلاعاتی که از این پردازش بدست می‌آیند، می‌توانند برای تحقیق در مورد لغات، ساختارها و موجودیت‌هایی که عموماً هنگامی که رابطه‌ای بین دو مفهوم بیان می‌شود در زبان طبیعی استفاده می‌شود، بکار روند. این سیستم بر روی ۸ رابطه سال تولد شخص، سال مرگ شخص، مکان تولد شخص، بازیگر - فیلم، نویسنده - کتاب، بازیکن فوتبال - تیم، کشور - رئیس دولت، کشور - پایتخت تست گردیده است.

همچنین در سال‌های اخیر سیستم‌های مختلفی برای حاشیه‌نویسی متون خاص قلمرو جهت گسترش وب معنایی بوجود آمده‌اند که از جمله آنها می‌توان به [۲۴] برای حاشیه‌نویسی متون حوزه زیست پزشکی اشاره نمود.

۳- معرفی سیستم POPTA

POPTA یک سیستم موازی نمونه‌دهی و حاشیه‌نویسی است که بر روی منابع زبانی چندگانه کار می‌کند. ورودی‌های این سیستم یک هستان‌شناسی موجود با واژه سازی انگلیسی (مانند SUMO یا وردنت)، یک پیکره متنی از متون زبان فارسی که برچسب مقله نحوی دارند و نیز صفحات وب به زبان فارسی خواهد بود. هدف این سیستم یافتن نمونه‌های مناسب از پیکره و نیز از روی وب برای نمونه‌دهی به هستان‌شناسی ورودی و در نهایت حاشیه‌نویسی معنایی متون فارسی می‌باشد.

الگوریتم کلی کار به این ترتیب است: ابتدا کلیه اسامی خاص موجود در پیکره متنی فارسی به عنوان نمونه استخراج شده و با توجه به روش‌هایی که در بخش‌های بعدی به آنها خواهیم پرداخت، مجموعه‌ای از برچسب‌های اولیه به آنها تخصیص می‌یابد. سپس این نمونه‌ها به همراه مجموعه برچسب‌های کاندیدای آنها وارد مولفه کار با وب (Googling) شده و با استفاده از موتور جستجوی گوگل (در صورت امکان) مناسبترین برچسب ممکن به هر یک از این نمونه‌ها تخصیص می‌یابد. در صورت موفقیت‌آمیز بودن عملیات انتخاب مناسب‌ترین برچسب، این برچسب‌ها توسط یک واسط میان‌زبانی - که از دو دیکشنری دو زبانه فارسی به انگلیسی و انگلیسی به فارسی استفاده می‌کند - به زبان انگلیسی ترجمه شده و وارد فاز انتخاب مفهوم متناظر در هستان‌شناسی مقصد (وردنت) می‌شوند. پس از یافتن مفهوم متناظر با هر برچسب، از این مفاهیم برای حاشیه‌نویسی متن ورودی استفاده می‌شود. در صورت عدم موفقیت در انتخاب مناسب‌ترین برچسب به کمک گوگل، نمونه ورودی به عنوان یک پرس و جو به ویکی‌پدیا داده می‌شود تا نوع آن تعیین گردد و ادامه کار از مولفه یافتن مفهوم متناظر در وردنت و مانند بخش قبل دنبال می‌گردد.

لازم به ذکر است که از آنجایی که کاراکترهای فارسی دارای کدگذاری‌های مختلفی می‌باشند و در برخی نوشته‌ها از کدهای فارسی و در برخی از کدهای عربی استفاده می‌شود، در متون فارسی شاهد برخی حروف مشابه با کدهای متفاوت هستیم. این مشکل بیشتر برای حروف 'ی' و 'ک' مشهود است. در برخی

مبتنی بر پیکره و دیگر منابع موجود برای یادگیری روابط طبقه‌ای استفاده می‌نمایند.

همچنین رویکردهای اخیر را می‌توان با توجه به استفاده متفاوت از داده‌های آموزشی، به دو گروه تقسیم نمود: رویکردهای بدون نظارت و رویکردهای با نظارت که از داده‌های آموزشی برچسب‌گذاری شده دستی استفاده می‌کنند. در حالی که روش‌های بدون نظارت کارایی پایینی دارند، رویکردهای با نظارت به دقت بالاتری دست یافته‌اند، ولی به ساخت دستی یک مجموعه آموزشی که در عین حال آنها را از کاربردهای در مقیاس بزرگ باز می‌دارد، نیاز دارند [۱۴].

معیار چهارم: محصول نهایی - سیستم‌های موجود را بسته به نوع خروجی‌ای که تولید می‌نمایند، می‌توان به دو گروه سیستم‌های نمونه‌دهی و سیستم‌های پشتیبان تقسیم نمود. گروه اول سیستم‌هایی هستند که برای نمونه‌دهی به هستان‌شناسی‌ها کاربرد دارند و خروجی آنها نمونه‌های تولید شده برای مفاهیم هستان‌شناسی ورودی است (مانند [۴، ۸، ۱۱، ۱۴، ۱۵]). گروه دوم نیز ملزومات کار گروه اول را فراهم می‌نمایند. در واقع این سیستم‌ها به نمونه‌دهی هستان‌شناسی‌ها نمی‌پردازند ولی پیش‌نیازهای لازم برای فعالیت‌های سیستم‌های نمونه‌دهی را فراهم می‌نمایند. این سیستم‌ها عموماً به کار استخراج موجودیت‌های نام‌دار و یا زیرمجموعه‌ای خاص از نمونه‌ها می‌پردازند (مانند [۱۶]).

برخی از سیستم‌های نمونه‌دهی [۶] در واقع سیستم‌های حاشیه‌نویسی هستند که عمل نمونه‌دهی را نیز انجام می‌دهند. در مقابل سکوها‌های حاشیه‌نویسی معنایی (SAP^{۱۲}) وظیفه اصلی یافتن موجودیت‌های نامدار متن و الصاق برچسب معنایی مناسب (با توجه به یک هستان‌شناسی موجود) به آنها را دارا می‌باشند. این سکوها در معماری، ابزارها و روش‌های استخراج اطلاعات، هستان‌شناسی اولیه، میزان کار دستی لازم برای انجام حاشیه‌نویسی، کارایی و دیگر ویژگی‌ها مانند مدیریت حافظه با هم متفاوت هستند. این سکوها می‌توانند بر مبنای نوع روش حاشیه‌نویسی مورد استفاده به دو گروه اصلی مبتنی بر الگو و مبتنی بر یادگیری ماشینی و یا ترکیبی از هر دو دسته‌بندی شوند [۱]. SAP‌های مبتنی بر الگو ممکن است قادر به کشف الگوهای جدید و یا استفاده از الگوهای از پیش تعریف شده باشند. بیشتر روش‌های کشف الگو یک مجموعه اولیه از موجودیت‌ها را تعریف و پیکره را به منظور یافتن الگوهایی که این موجودیت‌ها در آنها وجود دارند، پیمایش می‌کنند. موجودیت‌های جدید همراه الگوهای جدید کشف می‌شوند. این فرآیند بصورت بازگشتی تا جایی ادامه می‌یابد که دیگر هیچ موجودیتی کشف نشود یا این‌که کاربر فرآیند را متوقف نماید. حاشیه‌نویسی‌ها می‌توانند توسط قوانینی دستی برای یافتن موجودیت‌ها در متن نیز تولید شوند.

SAP‌های مبتنی بر یادگیری ماشینی از دو رویکرد بهره می‌گیرند: احتمال و استنتاج. SAP‌های احتمالاتی از مدل‌های آماری برای پیش‌بینی مکان موجودیت‌ها در داخل متن استفاده می‌کنند. برای مثال در الگوریتم DATAMOLD [۱۷] برای یافتن نمونه‌های داده‌ای در داخل صفحات HTML، از مدل مخفی مارکوف استفاده می‌شود. همچنین الگوریتم LP^۲ [۱۸] هسته الگوریتم استخراج اطلاعات (IE) در ابزار Amilcare [۱۹] است که برای اعمال استنتاج پوششی^{۱۳} توسط SAP‌های Armadillo [۲۰] و Ont-O-Mat [۱۹] استفاده می‌شود [۱].

سکوی KIM [۲۱] نمونه‌ای مطرح از سکوها‌های حاشیه‌نویسی است که شامل یک هستان‌شناسی (KIMO)، یک پایگاه دانش، یک حاشیه‌نویس معنایی، یک سرور شاخص‌گذاری و بازیابی به همراه پایانه‌هایی برای واسطه‌گری با سرور می‌باشد. KIMO یک مجموعه پایه از کلاس‌های موجودیت، روابط و محدودیت ویژگیها را تعریف می‌کند. پایگاه دانش به کمک ۸۰۰۰۰ موجودیت، شامل موقعیت‌ها و سازمان‌ها که از یک بدنه خبری عمومی جمع‌آوری شده‌اند، گسترش می‌یابد. موجودیت‌های دارای نامی که در طی روند حاشیه‌نویسی پیدا شده‌اند، با نوعشان در آنتولوژی و همچنین با یک مرجع در پایگاه دانش تطبیق داده می‌شوند.

وجود ندارد، لازم است وظایف چنین سیستمی در داخل سیستم نمونه‌دهی به انجام رسد. بطور کلی برای یافتن کاندیداهای مناسب برای نمونه‌ها و یا مفاهیم جدید - که باید حاشیه‌نویسی شوند - در یک صفحه وب، ابتدا لازم است محتوای متنی صفحات وب را استخراج کرده (برچسب‌های HTML را حذف کنیم) و یک برچسب زن مقوله نحوی بر روی آن اعمال کنیم تا به هر قطعه^{۱۷} رده دستوری متناظر با آن، تخصیص یابد. این برچسب زن می‌تواند اسامی خاص را تعیین نماید. در متون لاتین معمولاً می‌توان نمونه‌ها را از طریق کلماتی که با حروف بزرگ نوشته شده‌اند و برچسب اسم خاص^{۱۸} دارند، جستجو کرد. برای یافتن مفاهیم نیز می‌توان از کلمات با حروف کوچک و با برچسب اسم عام^{۱۹} بهره برد. در زبان‌هایی مانند فارسی، عربی و ... که فاقد حروف بزرگ هستند باید اطلاعات ساخت - واژی^{۲۰} و بویژه اطلاعات مقوله نحوی در مورد هر کلمه فراهم باشد. در POPTA برای استخراج نمونه‌ها، از متون یک پیکره فارسی حاوی برچسب‌های مقوله نحوی به عنوان ورودی استفاده شده است. این سیستم کلمات مفرد دارای برچسب اسم خاص (PR) را به عنوان نمونه استخراج می‌نماید. انواع دیگری از اسامی خاص وجود دارند که در این پیکره با برچسب اسم خاص مشخص نشده‌اند. از جمله این اسامی می‌توان به نام کتاب‌ها، مجلات، سریال‌ها، فیلم‌ها و سایر محصولات فرهنگی - هنری اشاره نمود. این سیستم برای بازیابی این اسامی از برخی علائم نگارشی خاص مانند قرار گرفتن نام در "" و یا عباراتی مانند "کلمه ۱: کلمه ۲" که می‌تواند نشان‌دهنده رابطه "کلمه ۲ نوعی (is-a) کلمه ۱ است" باشد، استفاده می‌نماید. همچنین وجود برخی کلمات کلیدی در همسایگی این اسامی می‌تواند در انتخاب آنها به عنوان اسامی خاص مؤثر باشند. در گام‌های بعدی می‌توان محدوده نمونه‌ها را گسترش داد و انواع اسامی از جمله اسامی عام و نیز برخی از افعال را نیز به عنوان نمونه در نظر گرفت. بدیهی است با توسعه سیستم‌های استخراج اطلاعات برای متون فارسی می‌توان از این سیستم‌ها برای استخراج خودکار نمونه‌ها از سایر متون فارسی و نیز متون موجود در سطح وب - که فاقد برچسب‌های مقوله نحوی هستند - بهره برد.

۳-۳- انتخاب برچسب‌های معنایی^{۲۱}

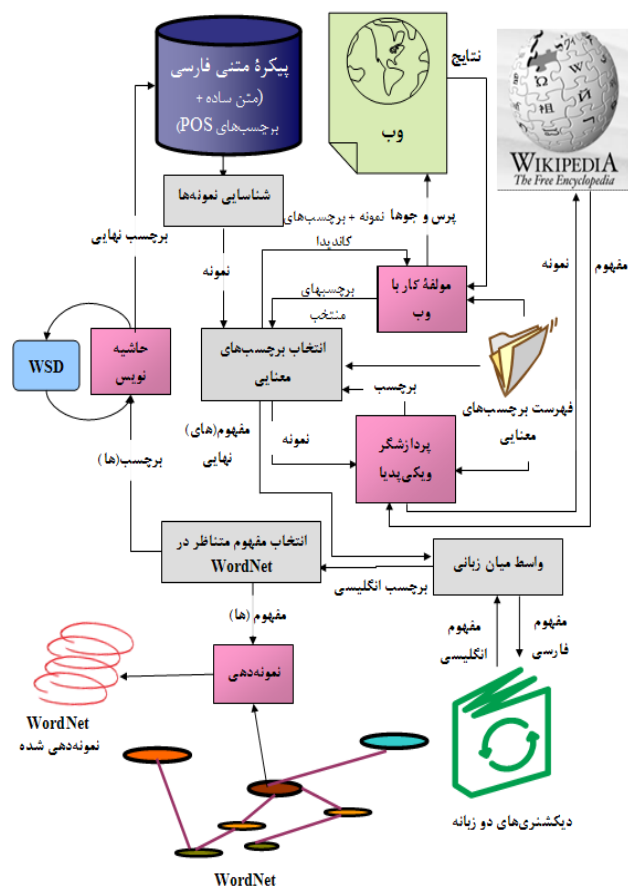
در سیستم POPTA، به منظور برچسب‌گذاری معنایی متون، مجموعه‌ای از برچسب‌های معنایی تهیه شده است. از جمله این برچسب‌ها می‌توان به انواع مکان‌های جغرافیایی، انسان‌ها، انواع محصولات فرهنگی - هنری، دوره‌های تاریخی (سلسله‌ها، دودمان‌ها، خاندان‌ها)، فرقه‌ها، اقوام، مذاهب و ادیان، زبان‌ها، ماه‌های سال، سیارات و ... اشاره نمود.

همانطور که مشاهده می‌شود تعداد و تنوع این برچسب‌های معنایی بسیار زیاد است (در نسخه فعلی سیستم ۵۴ برچسب از پیش تعریف شده دارد) و برای تعیین برچسب مناسب برای هر نمونه باید به نحوی تعداد برچسب‌های کاندیدا برای آن نمونه را بوسیله یک الگوریتم برچسب‌زنی کاهش داد تا به حد ایده‌آل یک برچسب برای هر نمونه نزدیک شویم.

این سیستم از مجموعه‌ای از قوانین مکاشفه‌ای (مانند قانون انطباق هسته^{۲۲} در گروه‌های اسمی) به همراه ترکیبی از روش‌های آماری از جمله روش‌های N-gram و روش‌های مبتنی بر الگو (براساس الگوهای ارائه شده در [۲۵]) برای محدود کردن لیست مفاهیم مرتبط با هر نمونه بهره می‌گیرد.

در بخشی از این رویکرد ترکیبی، سیستم، هم وقوعی کلمات کلیدی نشان‌دهنده یکی از کلاس‌های هستان‌شناسی با نمونه کاندیدای موردنظر را بررسی می‌کند. برای یافتن این کلمات کلیدی می‌توان به سه طریق عمل کرد: (۱) ساخت خودکار فهرستی از ترم‌هایی که در وب هم‌وقوعی بالایی با این نمونه‌های کاندیدا دارند، (۲) انتخاب مفاهیم موجود در سطوح پایین‌تر مفاهیم عمومی که در

از کلمات این حروف به شکل نرمال 'ی' و 'ک' و در برخی دیگر بصورت 'ی' و 'ک' نوشته شده‌اند. بنابراین در هر کجای سیستم که به نوعی با متون فارسی موجود سر و کار داریم (مانند متون پیکره، اطلاعات فارسی موجود در وب و یا دیکشنری دوزبانه) لازم است با انجام عملیات نرمال‌سازی فرمت‌های غیرهمسان را همسان نماییم و یا کلیه فرمت‌های ممکن را جهت انطباق بیازماییم. این کار سرعت پردازش را پایین می‌آورد ولی در مقابل درصد خطا را نیز کاهش می‌دهد. معماری این سیستم در شکل ۱ نمایش داده شده است. در ادامه به شرح عملکرد هر یک از این مولفه‌ها در سیستم POPTA می‌پردازیم.



شکل ۱- معماری سیستم POPTA

۳-۱- پیکره

این سیستم از یک پیکره متنی فارسی برچسب خورده (پیکره دکتر بیجن‌خان) استفاده می‌کند. در این پیکره برای هر کلمه اطلاعاتی از قبیل مقوله نحوی، مفرد یا جمع بودن، کسره اضافه، ضمائر متصل، اسامی خاص یا عام و ... فراهم است. POPTA از برچسب‌های مقوله نحوی^{۱۶} مفرد یا جمع بودن و خاص یا عام بودن اسامی استفاده می‌نماید.

۳-۲- شناسایی نمونه‌ها

بسیاری از سیستم‌های نمونه‌دهی به هستان‌شناسی‌ها در گام اول از یک سیستم IE برای استخراج موجودیت‌های نام‌دار موجود در متن بهره می‌گیرند. از آنجایی که در حال حاضر سیستم استخراج موجودیت‌های نام‌دار برای متون زبان فارسی

نمونه دیگر قرار گرفتن چند اسم خاص در کنار یکدیگر است که می‌تواند در محدوده مورد بررسی نشاندهنده یکی از سه وضعیت زیر باشد:

- همه اسمی انسان هستند؛ مانند: محمد حسین شهریار
- همه اسمی مکان هستند؛ مانند: ملبورن استرالیا
- اسم اول اثری متعلق به اسم دوم است؛ مانند: شاهنامه فردوسی

در چنین مواردی وجود کلمات کلیدی در همسایگی این عبارات می‌تواند به تعیین قطعی و یا بالا بردن احتمال وجود یک برچسب خاص کمک نماید. برای مثال در موارد فوق (وجود اسمی خاص متوالی) در صورت وجود القابی مانند آقا، خانم، امام، دکتر، مهندس، خاله، سید، استاد، جناب، سرکار و ... می‌توان فهمید که این اسمی خاص همگی انسان هستند و بنابراین برچسب Human به آنها تخصیص می‌یابد. در غیر اینصورت اگر کلماتی کلیدی مانند کتاب، نشریه، سریال، مجله، نمایش، تئاتر، فیلم و ... در همسایگی و قبل از اسم خاص اول یافت شود احتمال وضعیت سوم (نام اثر + نام صاحب اثر) بالا رفته و برچسب HPPL تخصیص می‌یابد که در آن H برای Human، P برای Property (جهت تاکید P دوبار ظاهر شده و این یعنی احتمال برچسب Property بیشتر از بقیه است) و L برای Location بکار رفته است.

در حالت مشابه در صورت وجود کلماتی کلیدی مانند قاره، شهر، کشور، استان، ایالت، روستا، رودخانه، دریا، دریاچه، اقیانوس، فلات، جلگه، کویر، دشت، جزیره، سد و ... احتمال وضعیت دوم (مکان‌های جغرافیایی) بالا می‌رود و برچسب HPLL تخصیص می‌یابد. در صورتی که هیچ کلمه کلیدی در همسایگی این اسمی خاص وجود نداشته باشد که به ما در تعیین نوع این اسمی کمک نماید، برچسب کلی HPL تخصیص می‌یابد که نشان می‌دهد احتمال رخداد هر یک از سه وضعیت فوق یکسان است.

این سیستم همچنین برای استخراج نمونه‌های موجود در متن و برچسب‌دهی به آنها کلمات هم‌پایه را نیز در نظر می‌گیرد. هنگامی که چند اسم خاص با حروف ربط (و، یا، و، ...) از هم جدا می‌شوند می‌توان آن‌ها را نمونه‌هایی هم‌پایه دانست. در این صورت با برچسب‌دهی به یکی از این اسمی خاص می‌توان برچسب سایر اسمی هم‌پایه را نیز مشخص نمود. در واقع می‌توان گفت این کلمات همگی نمونه‌های یک مفهوم هستند. کلیه این موارد در متن با برچسب‌های مجزا مشخص می‌شوند.

با انتخاب این برچسب‌های اولیه برای همه نمونه‌ها فاز برون خط پایان می‌یابد و نمونه‌های برچسب‌دهی شده وارد فاز برخط می‌شوند. در این فاز نیازمند آن هستیم تا برچسب‌های اولیه را به برچسب‌های نهایی که تنها نشانگر یک مفهوم واحد باشند، تبدیل نماییم. در فاز برخط، POPTA از وب و موتور جستجوی گوگل برای رفع ابهام در دسته‌بندی و انتخاب مناسب‌ترین مفهوم از میان مفاهیم کاندیدا و نیز استفاده از نمونه‌های حاشیه‌نویسی شده برای یافتن نمونه‌های دیگر هم‌پایه استفاده می‌نماید. این فرآیند در بخش ۴-۳ تشریح شده است. همچنین در صورت مطلوب نبودن نتایج حاصل از بکارگیری گوگل، POPTA با بهره‌گیری از دایره المعارف ویکی‌پدیا به جستجوی مفهوم متناظر با نمونه مورد بحث می‌پردازد. شرح این روش در بخش ۵-۳ آمده است.

۳-۴- مولفه کار با وب ۲۴

همانطور که در بخش قبل گفته شد، برای برخی از نمونه‌ها در مرحله اول لیستی از مفاهیم مرتبط ایجاد می‌شود. برای انتخاب مناسب‌ترین مفهوم از میان این مفاهیم کاندیدا، مجموعه‌ای از پرس و جوها به زبان طبیعی با استفاده از یک سری الگوهای معرفی شده تولید می‌شوند. این الگوها برای زبان فارسی تولید و یا تطبیق یافته‌اند.

هستان‌شناسی‌های خاص آن حوزه و یا در هستان‌شناسی وردنت موجودند و (۳) استفاده از فهرست‌های موجود در ویکی‌پدیا.

راه‌حل اول که در واقع یک روش آماری است، نیازمند حذف کلمات زائد^{۲۳} از این فهرست و اولویت دادن به کلماتی می‌باشد که در مفاهیم یک هستان‌شناسی وجود دارند و نشان‌دهنده کلمات عمومی‌تر هستند. راه دوم با مفاهیم موجود در ساختار هستان‌شناسی‌ها کار می‌کند و نیازمند پیمایش گراف هستان‌شناسی و یافتن فرزندان مناسب‌تر به عنوان نماینده‌هایی برای کلاس‌های در بر گیرنده نمونه‌های کاندیدا است.

راه حل سوم راه حل بهتری به نظر می‌رسد و پیچیدگی‌های راه‌حل‌های قبلی را ندارد. تنها مشکل این راه این است که صفحات فارسی ویکی‌پدیا هنوز به گستردگی صفحات دیگر زبان‌ها و بویژه صفحات انگلیسی نیستند و این امر باعث می‌شود که فهرست کاملی از تمامی کلمات کلیدی مورد نیاز برای ایجاد برچسب‌های معنایی را بدست ندهد.

از آنجایی که مجموعه کلمات کلیدی مورد نیاز این سیستم برای تشخیص انواع اسمی خاص محدود و از پیش تعریف شده می‌باشد، در حال حاضر سیستم POPTA از فهرست کلمات کلیدی - که بصورت دستی ایجاد شده است - استفاده می‌کند. در روش پیشنهادی سیستم در میان کلماتی که در پنجره‌ای به طول ۴ و قبل از نمونه کاندیدای موردنظر رخ داده‌اند به جستجوی کلمات کلیدی فوق می‌پردازد و در صورت یافتن، آنها را به عنوان مفاهیم مرتبط کاندیدا انتخاب می‌نماید.

بطور کلی کارایی روش‌های N-gram بستگی زیادی به سائز پنجره و ویژگی‌های زبان طبیعی مورد استفاده دارند، بنابراین برای برخی نمونه‌ها (مانند اسمی کشورها، شهرها و سایر مکان‌های جغرافیایی) سائز پنجره را به اندازه طول یک گروه اسمی در نظر گرفته و از مکاشفه‌های انطباق هسته برای کار با هسته گروه‌های اسمی بهره گرفته‌ایم.

روش‌های مبتنی بر الگو از کارایی بیشتری نسبت به روش‌های آماری برخوردارند و در صورت تطابق بخشی از متن با الگوهای مشخص می‌توان روابط معنایی مفیدی را استخراج نمود. از این رو برای بهره‌مندی از این روش‌ها الگوهای مختلفی که نشان‌دهنده روابط is-a و یا instance-of می‌باشند را فرموله کردیم.

به عبارت دیگر این سیستم از سهولت و عمومیت روش‌های آماری همراه با دقت بالا و کارایی بیشتر روش‌های مبتنی بر الگو بهره گرفته است و بر خطاهای ناشی از روش‌های آماری و رخداد پایین الگوهای معنایی در متون پیکره و نیز در سطح وب غلبه می‌نماید.

در مرحله تعیین برچسب‌های معنایی برای هر نمونه اغلب با بیش از یک برچسب کاندیدا برای نمونه‌ها روبرو هستیم. در واقع روش‌های آماری و الگویی به همراه مکاشفه‌های بکار رفته به تنهایی نمی‌توانند مناسب‌ترین برچسب را برای یک نمونه تعیین نمایند. از این رو در این مرحله یک سری برچسب اولیه به نمونه‌ها تخصیص داده می‌شود که نشان‌دهنده تمامی برچسب‌های کاندیدای مجاز برای آن نمونه می‌باشند. در این مجموعه می‌توان احتمال رخداد برچسب‌ها را براساس شواهد و کلمات کلیدی که در همسایگی نمونه رخ داده اند تقویت یا تضعیف نمود.

برای مثال دو جمله "امروز، قدرت حکومت در دست سلطان زنگبار است." و "با فتوحات سلطان محمود غزنوی، این منطقه به دست غزنویان افتاد." را در نظر بگیرید. در اولی اسم خاص "زنگبار" که در همسایگی کلمه کلیدی سلطان آمده است، نام یک مکان جغرافیایی است، حال آن‌که در دومی اسم خاص "محمود غزنوی" که باز هم در همسایگی کلمه کلیدی سلطان قرار گرفته است به نام یک انسان اشاره دارد. بنابراین به این دسته از کلمات برچسب اولیه HLOC تخصیص می‌یابد که نشان‌دهنده یکی از دو نوع انسان (H) و یا مکان (LOC) می‌باشد.

می‌کند. برای مثال اگر در مرحله اول "ایران" به عنوان نمونه‌ای از مفهوم "کشور" انتخاب شود، در مرحله دوم با استفاده از الگوهایی مانند "ایران"، * و سایر کشورها، کشورهای دیگری (که جایگزین رشته * می‌شوند) نیز از وب استخراج می‌گردند و یا در سطح پیکره نمونه‌های هم‌پایه دیگر را که با حروف ربط (از جمله 'و'، 'یا'، 'و' ...) به نمونه شناخته شده (در اینجا ایران) مرتبط شده‌اند را شناسایی می‌نماید.

با اتمام این بخش برای بسیاری از نمونه‌های برچسب خورده از مرحله قبل مناسب‌ترین برچسب معنایی نهایی انتخاب شده و هر نمونه با یک برچسب معنایی به زبان فارسی که اشاره به مفهومی خاص دارد، برچسب‌گذاری می‌گردد. البته هنوز مواردی باقی می‌مانند که در این مرحله نیز برچسب مناسبی برایشان انتخاب نمی‌شود. در این موارد برای برچسب‌گذاری از دایره‌المعارف ویکی‌پدیا کمک می‌گیریم که شرح فرآیند آن در بخش بعد به تفصیل آمده است.

۳-۵- پردازشگر ویکی‌پدیا

همانطور که گفته شد، پس از اتمام کار فازهای قبل تعدادی از اسامی خاص باقی می‌مانند که با استفاده از روش‌های مطرح شده قادر به تعیین برچسبی برای آنها نبوده‌ایم.

برای برچسب‌دهی به این اسامی باقیمانده یک راه این است که تمامی برچسب‌های معنایی موجود را به عنوان برچسب‌های کاندیدا در نظر گرفته و سپس با موتور جستجوی گوگل مناسب‌ترین برچسب را که دارای بالاترین تعداد hit می‌باشد برای آن اسم انتخاب نماییم.

از آنجایی که تعداد این برچسب‌های کاندیدا بسیار زیاد است برای انتخاب هر برچسب بطور متوسط ۲۰۰ پرس و جو تولید می‌شود (برای هر نمونه لازم است حدود ۵۰ برچسب کاندیدا و برای هر یک به طور متوسط ۴ پرس و جو آزموده شود تا از میان آنها برچسب مناسب انتخاب گردد). لذا تعیین مناسب‌ترین برچسب برای هر اسم خاص با این روش عملاً غیرممکن است. چرا که این کار کارایی سیستم را بسیار پایین می‌آورد و استفاده از چنین سیستمی را در عمل ناممکن می‌سازد.

یک راه‌حل مناسب برای این مشکل استفاده از ویکی‌پدیا برای تعیین برچسب مناسب می‌باشد. از آنجایی که ویکی‌پدیا یک دایره‌المعارف جامع و فراگیر است که موضوعات متنوع زیادی را پوشش می‌دهد و مطالب آن به بیش از ۱۹۰ زبان مختلف دنیا از جمله زبان فارسی موجود است، می‌تواند منبع مهمی برای تامین اطلاعات مورد نیاز برای بسیاری از کلمات و بویژه اسامی خاص باشد و فعالیت استخراج دانش را برای سیستم‌های نمونه‌دهی و حاشیه‌نویسی تسهیل نماید و بر کارایی این سیستم‌ها بیفزاید.

در این سیستم در دو مرحله به پردازش اطلاعات بازیابی شده از ویکی‌پدیا می‌پردازیم. در مرحله اول از ویکی‌پدیا تنها به منظور برچسب‌دهی نمونه‌های موجود در متن استفاده می‌کنیم. در مرحله دوم که از خروجی‌های مرحله اول بهره می‌گیرد به استخراج نمونه‌های دیگری که در متن صفحات بازیابی شده وجود دارند، می‌پردازیم.

الف) تعیین مفهوم متناظر با نمونه مستخرج از متن پیکره

در این مرحله کلیه اسامی خاص موجود در متن که بدون برچسب باقی مانده‌اند به عنوان پرس و جو به ویکی‌پدیا داده می‌شوند و با پردازش جملات اولیه صفحات بازیابی شده برچسب مناسب برای آنها تعیین می‌گردد.

با بررسی مقالات موجود در این دانشنامه مشخص می‌گردد که در بیشتر صفحات ویکی‌پدیا که عنوان آن یک اسم خاص است در چند جمله اول متن نام مفهوم متناظر با آن اسم خاص به صراحت ذکر شده است. از آنجایی که در این

نمونه‌ای از پرس و جوهای مناسب برای رفع ابهام برچسب LOC در جدول ۱ آورده شده است که در آن <طبقه> و یا <Category> به یکی از گونه‌های مکان‌های جغرافیایی مانند شهر، استان، کشور، جلگه و ... اشاره دارد.

جدول ۱- چند پرس و جو برای رفع ابهام گونه location برای نمونه X

فارسی	انگلیسی
X و سایر <طبقه>ها	X and other <Category>s
<طبقه>هایی مانند X	<Category>s such as X
...	...
<طبقه> X	The <Category> of X

در این قسمت باید توجه داشت برخی الگوها مانند الگوهای هیرست، دقت^{۲۵} بالایی دارند ولی معمولاً فراخوان^{۲۶} آنها بویژه در یک پیکره متنی پایین است. مثلاً الگوهایی مانند "استان‌هایی از جمله ایلام" و یا "جلگه‌هایی مانند آراگون" رخداد بسیار کمی در پیکره‌ها و حتی در سطح وب دارند، از این رو این الگوها برای تعیین نوع نمونه‌هایی که رخداد پایینی در وب دارند، مناسب نیستند. از طرف دیگر الگوهای دیگری (در زبان فارسی) وجود دارند مانند "نام مفهوم + نام نمونه" (مانند "کشور ایران" و یا "استان ایلام") که فراخوان بسیار بالایی دارند و برای نمونه‌هایی که اسامی آنها کمتر در متون متداول فارسی به چشم می‌خورد (مانند "جلگه آراگون") مناسب هستند ولی در مقابل این الگوها دقت پایینی دارند و عدم توجه به این مطلب می‌تواند درصد خطا را بالا ببرد. مثلاً یک پرس و جو مانند "تهران بزرگترین شهر ایران است" می‌تواند منجر به در نظر گرفتن "ایران" به عنوان نمونه‌ای از مفهوم "شهر" شود. (این مسئله به دلیل مورد خاص ساختار کسره اضافه در زبان فارسی رخ می‌دهد و معادلی در زبان انگلیسی ندارد. در فارسی "city of Iran" دقیقاً مانند "Iran city" نوشته می‌شود). بنابراین همواره نیازمند یک مصالحه بین دو معیار دقت و فراخوان خواهیم بود. در واقع در هنگام استفاده از الگوهایی با دقت پایین باید از ویژگی‌های دیگری نیز برای اطمینان از صحت برچسب معنایی انتخابی بهره برد.

بطور کلی این سیستم ۳ گروه از پرس و جوها را تولید می‌نماید:

- ۱) پرس و جوهای که نشان‌دهنده یکی از مکان‌های جغرافیایی می‌باشد؛ مانند [نام نمونه] و سایر [نام مفهوم]‌ها
 - ۲) پرس و جوهای که نشان‌دهنده نام یکی از انسان‌ها می‌باشد؛ مانند [لقب] + [نام نمونه]
 - ۳) پرس و جوهای که نشان‌دهنده نام یکی از محصولات فرهنگی- هنری می‌باشد؛ مانند [نام مفهوم] + -- + [نام نمونه]
- در اینجا نیاز به یک مرحله نرمال‌سازی پرس و جوها به منظور انطباق با اطلاعات فارسی موجود در سطح وب داریم. در این مرحله پرس و جوهای مورد نظر باید به ازای همه جایگشت‌های حروف چندگانه (۲ حالت برای ۲ حرف ی و ک) تولید شوند. این امر می‌تواند حجم پرس و جوها را تا ۴ برابر بالا ببرد و تا حدی از کارایی سیستم می‌کاهد ولی تا هنگامی که کاراکترهای فارسی دارای کدهای استاندارد یکتا نشوند، این سیستم باید بر هرگونه فرمت ورودی منطبق گردد.

پس از تولید پرس و جوهای مناسب، هر یک از طریق Google API در وب جستجو می‌شوند و نتایج بازیابی شده حاصل، توسط برنامه پردازش می‌شود و مفهومی که بالاترین تعداد hit را دارد به عنوان مناسب‌ترین برچسب معنایی انتخاب می‌گردد. این سیستم همچنین هر بار در یک گردش از حلقه، از نمونه‌هایی که نوع آنها تعیین شده برای یافتن نمونه‌های دیگر هم‌پایه در سطح وب و پیکره استفاده

در چنین مواردی به عنوان راهکار مکمل در POPTA ۴ کلمه کلیدی نزدیکتر به نمونه را به عنوان مفاهیم کاندیدا در نظر گرفته و سپس با موتور جستجوی گوگل میان این مفاهیم به رفع ابهام پرداخته و مفهومی با بالاترین تعداد hit را به عنوان مفهوم مناسب برای نمونه مورد نظر انتخاب می‌نماییم. عدد ۴ به صورت تجربی و پس از آزمون های متعدد بدست آمده است.

ب) یافتن نمونه‌های جدید از متون ویکی‌پدیا

در گام دوم فراتر از نمونه‌دهی هستان‌شناسی مقصد با استفاده از نمونه‌های موجود در متن رفته و به منظور یافتن سایر نمونه‌های موجود در متن و تعیین مفهوم مورد نظر آنها به پردازش کلی صفحات بازبایی شده و نیز ساختارهای موجود در این صفحات می‌پردازیم.

اسامی خاص موجود در صفحات ویکی‌پدیا که عناوین مقالات دیگری در این دایره‌المعارف هستند، همگی به صفحات مربوط به خود پیوند یافته‌اند. بنابراین می‌توان بسیاری از نمونه‌های موجود در متن و حتی برخی از مفاهیم عمومی‌تر مانند شهر، کشور، زبان، صنعت، انقلاب و ... را استخراج نمود. بیشتر این کلمات پیوند یافته را اسامی خاص تشکیل می‌دهند و مفاهیم عمومی‌تر به ندرت به چشم می‌خورند.

کلمات مهم متن مانند عنوان مقاله (که در متن نیز تکرار می‌شود) و یا عناوین سرفصل‌های یک مقاله نیز می‌توانند در استخراج بخش‌های مهم متن بسیار کمک کننده باشند.

دلیل ساختار غیر استاندارد ویکی‌پدیا پردازش کلیه متون صفحه و استخراج اطلاعات مورد نیاز بسیار دشوار است. تاکنون تلاش‌هایی برای توسعه یک واسط برنامه کاربردی (API) برای ویکی‌پدیا صورت گرفته است [۲۶، ۲۷]. بیشتر این APIها برای زبان‌های انگلیسی، ژاپنی و چینی توسعه یافته است. متأسفانه در حال حاضر واسطی برای دسترسی به صفحات فارسی ویکی‌پدیا وجود ندارد.

به دلیل مشکلات موجود برای پردازش و تجزیه صفحات فارسی ویکی‌پدیا توجه خود را به ساختارهای موجود در این صفحات معطوف کردیم. سیستم POPTA پنج ساختار پیوندها، تغییر مسیر، رده‌ها، ابهام زدایی و جعبه‌های اطلاعاتی را مورد بررسی قرار داده است [۲].

• صفحات تغییر مسیر بیشتر برای تهیه فهرستی از لغات معادل (مترادف، صورت‌های جمع یا مفرد، روش‌های نگارشی مختلف، مخفف‌ها و نیز اشتباهات املائی مصطلح) کاربرد دارند و می‌توان در پایان فاز نمونه‌دهی به هستان‌شناسی برای هر نمونه فهرستی از لغات معادل آن را نیز ذخیره نمود.

• پیوندها در ویکی‌پدیا دو نوع هستند: داخلی و خارجی.

(۱) پیوندهای داخلی که به موجودیتی در داخل ویکی‌پدیا - که عموماً عنوان مقاله‌ای در خود این دایره‌المعارف است - اشاره دارند. همانطور که ذکر شد، سیستم POPTA از این پیوندها برای استخراج نمونه‌های موجود در صفحات بازبایی شده ویکی‌پدیا بهره می‌گیرد. استفاده دیگر این پیوندها ایجاد مجموعه لغات هم‌وقوع با عنوان صفحه و نیز یافتن موجودیت‌های مرتبط با نمونه مورد نظر است؛ ساختار مشابه دیگر "جستارهای وابسته" است که به مفاهیم موجودیت‌های مرتبط با نمونه جستجو شده اشاره دارد.

(۲) پیوندهای خارجی (یا "پیوند به بیرون") که به صفحاتی در خارج از سایت ویکی‌پدیا اشاره دارند. این سیستم در حال حاضر به این دسته از پیوندها نپرداخته است.

• رده‌های موجود در ویکی‌پدیا ساختارهای بسیار مفیدی برای تعیین حوزه دربرگیرنده مقاله بازبایی شده است و هنگامی که این مقاله یک اسم خاص را توصیف می‌کند، این رده‌ها می‌توانند تعیین کننده مفهوم متناظر با آن نمونه باشند. خوشبختانه این رده‌ها و ساختار پیوندهای آنها بصورت جدول‌های SQL موجودند. از این رو می‌توان بدون تجزیه و پردازش مقالات موجود در ویکی‌پدیا به

سیستم برای تعیین نوع اسامی خاص از ویکی‌پدیا استفاده می‌کنیم، می‌توانیم با یافتن الگوهایی که بیانگر رابطه یک نمونه با مفهوم متناظر آن است، به استخراج مفاهیم دربرگیرنده نمونه‌های ورودی بپردازیم. مثلاً جملاتی مانند "ایران کشوری در جنوب غربی آسیا در منطقه‌ای مشهور به خاورمیانه است." و یا "بینواییان نام رمان معروفی نوشته ویکتور هوگو نویسنده سرشناس فرانسوی است." به روشنی بیانگر مفهوم مرتبط با نمونه ورودی است.

اسامی انسان‌ها معمولاً با کلماتی مانند انواع مشاغل (نویسنده، جامعه‌شناس، کارگردان، بازیگر، فیلمنامه‌نویس و ...)، انواع مسئولیت‌های دولتی - حکومتی (رهبر، رئیس جمهور، نخست وزیر، پیشوا، صدراعظم و ...)، کلماتی که نشانه‌دهنده ارتباط دو یا چند انسان با هم هستند (یاران، صحابه، پدر، مادر، فرزند، همفکری کردن و ...) و سایر کلمات کلیدی که عموماً برای یک انسان بکار می‌روند (شخصیت، رئیس، نوشته، اثر، متولد و ...) هم‌وقوعی دارند. مثلاً:

"هیتلر دارای مقاماتی همچون صدر اعظم آلمان، رئیس دولت، و ریاست ایالات، یک دیکتاتور مطلق و سخنرانی با استعداد می‌باشد"، "تیکو کریمی متولد ۱۹ آبان ۱۳۵۰ در تهران بازیگر، کارگردان و یک مترجم ایرانی است."، "وئسان ویلم ون گوگ نقاشی نامدار زاده هلند بود."

با پردازش و تجزیه صفحات فارسی ویکی‌پدیا و با توجه به راهنماهایی که در کد این صفحات قرار داده شده‌اند (مانند تعیین اولین رخداد عنوان مقاله در متن - که همان نمونه مورد نظر است -، اولین رخداد هر کلمه‌ای از متن که عنوان صفحه دیگری در ویکی‌پدیا است، کلماتی که هنوز مقاله متناظر با آنها در ویکی‌پدیا ایجاد نشده است و ...) و با پی‌گیری کلمات کلیدی موجود در همسایگی نمونه مورد جستجو (که در مثال‌ها بصورت زیرخطدار مشخص شده‌اند) می‌توان به نوع برچسب معنایی متناظر با مفهوم دربرگیرنده آن نمونه پی برد (برای اطلاعات بیشتر در مورد ساختار ویکی‌پدیا و نحوه تجزیه صفحات آن رجوع کنید به [۲]).

در این بخش نیز همانند بخش واسط زبانی نیازمند انجام نرمال‌سازی جهت انطباق برچسب‌های فارسی ورودی و کلمات فارسی موجود در صفحات ویکی‌پدیا می‌باشیم.

اگرچه در بسیاری موارد بررسی کلمات کلیدی موجود در همسایگی نمونه مورد نظر مارا به مفهوم متناظر آن می‌رساند، ولی همیشه نمی‌توان نزدیکترین کلمه کلیدی همسایه را به عنوان برچسب معنایی متناظر انتخاب نمود. چرا که (۱) ممکن است بجای کلمه کلیدی نیازمند بررسی عبارات کلیدی باشیم مثلاً اگر فرض کنیم کشور، شهر، استان و ایالت کلمات کلیدی هستند و بدنبال مفهوم متناظر با نمونه "شیراز" یا "تهران" می‌گردیم در جملات "شیراز مرکز استان فارس است" و یا "تهران پایتخت کشور ایران است." وجود کلمه کلیدی استان در همسایگی نمونه "شیراز" و کلمه کلیدی کشور در همسایگی نمونه "تهران" می‌تواند منجر به خطا شود در چنین شرایطی معرفی عبارات کلیدی "پایتخت کشور" و "مرکز استان" که همیشه معادل کلمه کلیدی "شهر" هستند مشکل را حل می‌کند. در POPTA پایگاهی از عبارات کلیدی و کلمات کلیدی معادل آنها معرفی شده که در کاهش خطا بسیار موثر است. واضح است که از آنجا که قادر به تعیین همه عبارات کلیدی موجود در زبان فارسی نیستیم و همچنین به دلیل آن که افزایش تعداد کلمات کلیدی موجود میزان خطا در تخصیص نمونه به نزدیکترین کلمه کلیدی را افزایش خواهد داد، لازم است راهکار مکملی برای این مشکل پیشنهاد نماییم.

(۲) ممکن است نزدیکترین کلمه کلیدی همسایه مورد نظر به نمونه دیگری اشاره کند. مثلاً در جمله "تهران در کشور ایران واقع شده است" اساساً کلمه کلیدی برای نمونه تهران وجود ندارد و یا در جمله "تهران در کشور ایران بزرگترین شهر است" کلمه کلیدی مناسب برای تهران کلمه شهر است که گرچه در جمله وجود دارد ولی نزدیکترین کلمه کلیدی به نمونه نیست.

در این رویکرد به منظور ترجمه برچسب‌های فارسی به زبان انگلیسی در مرحله اول از یک دیکشنری فارسی به انگلیسی که بصورت یک پایگاه داده دو زبانه موجود است، بهره بردیم. پس از انجام نرمال سازی برچسب‌ها جهت انطباق با فرمت دیکشنری، با دادن هر کلمه فارسی به عنوان ورودی به این دیکشنری مجموعه‌ای از لغات انگلیسی معادل بازیابی می‌شوند. این مجموعه از لغات بصورت الفبایی مرتب هستند؛ بنابراین ممکن است برخی از این لغات ترجمه‌های مصطلح واژه فارسی ورودی نبوده و وجود آنها در لیست لغات انگلیسی می‌تواند تعداد مفاهیم معادل انتخاب شده در هستان‌شناسی مقصد شده را افزایش دهد و این امر منجر به افزایش ابهام در نمونه‌دهی خواهد شد.

برای حل این مشکل این بار در جهت عکس عمل کرده و کلیه لغات انگلیسی بازیابی شده از دیکشنری اول را به یک دیکشنری انگلیسی به فارسی دیگر می‌دهیم تا برای هر کلمه انگلیسی لغات معادل آن به زبان فارسی بدست آیند. برای این کار نمی‌توان از نسخه انگلیسی به فارسی همان دیکشنری اول استفاده نمود؛ چرا که همان مجموعه لغات فارسی اولیه بدست می‌آیند و این امر کمکی به ما نمی‌کند. بنابراین در این بخش از دیکشنری انگلیسی به فارسی آنلاین آریان‌پور استفاده کردیم (که در آن معانی به ترتیب مصطلح بودن مرتب شده‌اند) و برای هر لغت انگلیسی بدست آمده از دیکشنری قبلی مجموعه واژه‌های معادل آن را در زبان فارسی بازیابی نمودیم.

نکته قابل توجه این است که همانطور که گفته شد این بار این مجموعه لغات بازیابی شده به ترتیب مصطلح بودن و ارتباط بیشتر با مفهوم مورد نظر مرتب هستند و این امر به ما در انتخاب بهترین ترجمه‌های معادل از میان این مجموعه لغات بازیابی شده کمک می‌نماید.

در این مرحله برای هر لغت انگلیسی بدست آمده از ترجمه اول مجموعه‌ای از لغات فارسی معادل داریم که از ترجمه دوم بدست آمده و به ترتیب مربوط بودن و کاربرد مرتب هستند. حال کفایت از بین این کلمات انگلیسی اولین کلمه‌ای را انتخاب نماییم که کلمه فارسی اولیه (برچسب معنایی ورودی) در میان کلمات فارسی معادل آن کلمه انگلیسی زودتر رخ داده باشد (در اولویت بالاتری قرار گرفته باشد). در واقع معمولاً کلمه‌ای مناسب است که کلمه فارسی اولیه، اولین معادل آن باشد و در صورتی که این کلمه فارسی اولین معادل هیچ یک از کلمات انگلیسی بدست آمده از ترجمه اول نباشد به سراغ معادل‌های بعدی می‌رویم و در صورتی که این کلمه برای بیش از یک کلمه انگلیسی در اولین مکان واقع شده باشد، همه آن لغات انگلیسی انتخاب خواهند شد.

برای روشن تر شدن موضوع برچسب کشور را در نظر بگیرید. در مرحله اول این کلمه به دیکشنری فارسی به انگلیسی ذکر شده داده می‌شود. کلمات معادل انگلیسی زیر به ترتیب الفبایی بدست می‌آیند:

Commonwealth, Country, Kingdom, Nation, Soil, State, Territory
 حال هریک از این کلمات را به دیکشنری دوم می‌دهیم، نتایج حاصل در جدول ۲ نمایش داده شده‌اند.

جدول ۲- نتایج حاصل از ترجمه لغت "کشور" با استفاده از دو دیکشنری

فارسی	انگلیسی
-	Commonwealth
کشور، دیار، بیرون شهر، دهات، بیلاق	Country
پادشاهی، کشور، قلمرو پادشاهی	Kingdom
ملت، قوم، امت، خانواده، طایفه، کشور	Nation
خاک، کثیف کردن، ...، زمین، کشور، سرزمین، ...	Soil
حالت، کشور، ایالت	State
سرزمین، خاک، خطه، زمین، ملک، کشور، قلمرو	Territory

آنها دسترسی داشت. پیاده‌سازی این بخش نیز در فهرست کارهای آینده این سیستم قرار می‌گیرد.

• ساختارهای ابهام‌زدایی نیز ساختارهای بسیار مفیدی برای رفع ابهام معنایی لغات و یافتن معانی مختلف هر کلمه^{۲۸} (WSD) است. این ویژگی ویکی‌پدیا برای بخش WSD حاشیه‌نویسی می‌تواند بکار رود. در فرایند WSD یا رفع ابهام سعی می‌شود تشخیص داده شود از میان معانی متعدد یک کلمه کدام معنی در رخدادهای حاضر آن مدنظر بوده است. یافتن مجموعه لغات هم‌وقوع با هر یک از معانی نمونه مورد نظر (با استفاده از پیوندهای موجود در متن مقاله مربوط به هر یک از این نمونه‌ها) و محاسبه میزان اشتراک و هم‌پوشانی این لغات با لغات موجود در همسایگی نمونه استخراج شده از متن می‌تواند به تعیین معنی نمونه استخراج شده کمک نماید. در مرحله بعد می‌توان مفهوم متناظر با هر معنی نمونه جستجو شده را تعیین نمود.

• جعبه‌های اطلاعاتی یا Infoboxها از پرکاربردترین ساختارهای موجود در ویکی‌پدیا می‌باشند. Infoboxها کلیه اطلاعات مهم موجود در متن را بصورت یک جدول خلاصه می‌سازند. در میان Infoboxهای موجود برای انواع مقالات در ویکی‌پدیا، اطلاعات موجود در جداول صفحات مربوط به کشورها، شهرها و انسان‌ها اهمیت بیشتری دارند. سایر موجودیت‌های موجود در ویکی‌پدیا معمولاً فاقد جدول می‌باشند.

در حال حاضر سیستم POPTA اطلاعات موجود در Infoboxهای مربوط به کشورها و شهرها را استخراج می‌نماید. پیاده‌سازی استخراج اطلاعات جداول موجود در صفحات مربوط به شخصیت‌های انسانی در فهرست کارهای آینده این سیستم قرار می‌گیرد.

۳-۶- واسط میان زبانی

همانطور که پیش از این گفته شد، POPTA یک سیستم میان زبانی است. دلیل این نامگذاری آن است که این سیستم متون فارسی را با استفاده از یک هستان‌شناسی انگلیسی حاشیه‌نویسی کرده و هستان‌شناسی انگلیسی را نیز با استفاده از متون فارسی نمونه‌دهی می‌نماید. بنابراین نیازمند واسطی هستیم تا بتواند برچسب‌های انتخاب شده را که به زبان فارسی هستند، به مفهوم معادل آنها در هستان‌شناسی ورودی که به زبان انگلیسی است، متصل نماید.

می‌دانیم که هر کلمه فارسی می‌تواند معادل چندین کلمه انگلیسی باشد و برعکس. در نتیجه بسیار محتمل است که هیچ تناظر یک به یکی بین مفاهیم انگلیسی موجود در هستان‌شناسی و مفاهیم فارسی مرتبط با نمونه‌های انتخاب شده وجود نداشته باشد.

برای یافتن مناسب‌ترین تناظر، نیازمند اعمال روش‌های رفع ابهام معنایی کلمات (WSD) هستیم. به این منظور دو راه حل وجود دارد: (۱) ترجمه دستی بخشی از مجموعه‌های مترادف^{۲۹} وردنت و اتصال خودکار مفاهیم انتخاب شده فارسی به این مجموعه‌های مترادف ترجمه شده با استفاده از ترکیبی از معیارهای شباهت زبانی (مانند معیار edit-distance) و ساختاری، (۲) ترجمه خودکار مفاهیم انتخاب شده توسط یک دیکشنری دو زبانه و انتخاب مجموعه‌های مترادف معادل وردنت.

مشکل ابهامات معنایی به انتخاب معنی^{۳۰} مناسب کلمه (یا عبارت) نمونه و نیز انتخاب ترجمه مناسب آن (مفهوم متناظر) در زبان مقصد (انگلیسی) توجه دارد و در بخش حاشیه‌نویسی جدی‌تر از بخش نمونه‌دهی است. برای رفع ابهام در ترجمه و انتخاب معنی، یک رویکرد جدید مکاشفه‌ای مبتنی بر دیکشنری معرفی نموده‌ایم.

در اینجا باید توجه داشت برچسب‌هایی که در حوزه کار این سیستم قرار دارند عموماً دارای معانی متفاوتی نیستند و می‌توان گفت مصطلح‌ترین معنی هر یک از این برچسب‌ها همان معنی مورد نظر است.

از آنجایی که وردنت برای هر کلمه عددی نگه می‌دارد که نشان‌دهنده فرکانس ارجاع آن کلمه به یک مجموعه مترادف خاص است [۲۸]، در واقع می‌توان مجموعه مترادفی که بالاترین فرکانس ارجاع را دارد مصطلح‌ترین و پرکاربردترین مجموعه مترادف برای کلمه ورودی دانست و مجموعه مترادف با بالاترین فرکانس ارجاع را به عنوان مجموعه مترادف مناسب برای برچسب معنایی مورد نظر انتخاب نمود. در صورت وجود بیش از یک مجموعه مترادف با ماکسیمم فرکانس ارجاع همه این مجموعه‌های مترادف به عنوان مفهوم معادل نمونه مورد نظر انتخاب می‌گردند. برای مثال برای برچسب سد معادل انگلیسی dam انتخاب می‌گردد که در وردنت دارای سه مجموعه مترادف متناظر زیر می‌باشد:

- a barrier constructed to contain the flow of water or to keep out the sea
 - a metric unit of length equal to ten meters
 - female parent of an animal especially domestic livestock
- از آنجایی که فرکانس ارجاع به مجموعه مترادف اول بیشتر از دو مجموعه دیگر است، این مجموعه به عنوان مفهوم متناظر در وردنت انتخاب شده و نمونه‌های استخراج شده به آن متصل می‌گردند.
- نتایج ارزیابی این روش انتخاب مجموعه مترادف در بخش ۵ آورده شده است. به این ترتیب در پایان مرحله نمونه‌دهی، هر نمونه به یک یا چند مفهوم معادل که می‌توانند معادل یک یا چند مجموعه مترادف از وردنت باشند، متصل می‌گردد.

۳-۸ - حاشیه‌نویسی معنایی

همانطور که مطرح شد، در این سیستم حاشیه‌نویسی معنایی همزمان با نمونه‌دهی به هستان‌شناسی انجام می‌گیرد. به عبارت دیگر هنگامی که مفهوم معادل نمونه استخراج شده از متن مشخص گردید و محل آن در هستان‌شناسی ورودی تعیین شد، بخشی از متن که دربرگیرنده آن نمونه است با برچسبی همنام با شماره مفهوم معادل در هستان‌شناسی حاشیه‌نویسی می‌شود.

در این مرحله برای هر نمونه موجود در متن، برچسب معنایی تعیین شده، شماره مجموعه مترادف (یا مجموعه‌های مترادف) معادل در وردنت و تعریف زبان طبیعی این مجموعه‌های مترادف ذخیره می‌شود.

بنابراین تا این جا از یک پیکره فارسی دارای برچسب‌های مقوله نحوی برای نمونه‌دهی به هستان‌شناسی و نیز حاشیه‌نویسی معنایی همان پیکره بهره بردیم. در مرحله بعد می‌توانیم از همین پیکره که در حال حاضر برچسب‌های معنایی را نیز علاوه بر برچسب‌های مقوله نحوی داراست، برای حاشیه‌نویسی معنایی اسناد موجود در وب استفاده نماییم.

یک راه انجام این کار بکارگیری عملیات یادگیری ماشینی است تا با ترکیب روش‌های آماری و مبتنی بر الگو، قوانین موجود در این پیکره را بیاموزد و با کمک این قوانین جدید و اضافه نمودن آنها به مجموعه قوانین مکاشفه‌ای قبلی به حاشیه‌نویسی معنایی اسناد موجود بر روی وب بپردازد.

۴ - آزمون و ارزیابی

این سیستم برای ارزیابی نتایج حاصل از فاز استخراج برچسب‌های معنایی و فاز نمونه‌دهی به هستان‌شناسی، بر روی ۱۵ متن از متون پیکره فارسی موجود - شامل ۸۰۸۷۹ کلمه - آزمایش گردید و دو معیار دقت و فراخوان برای دو

همانطور که دیده می‌شود، کلمه کشور به عنوان اولین ترجمه کلمه Country، دومین ترجمه کلمات Kingdom و State، ششمین ترجمه کلمات Nation و Territory و ... رخ داده است. بنابراین کلمه "Country" به عنوان مناسب‌ترین ترجمه برای برچسب فارسی کشور انتخاب می‌گردد. نتایج ارزیابی این روش در بخش ۵ آورده شده است.

۳-۷ - نمونه‌دهی به هستان‌شناسی

پس از ترجمه کلیه برچسب‌های معنایی بدست آمده از مرحله برچسب‌زنی معنایی نوبت به یافتن مفهوم معادل هر برچسب (ترجمه شده) در هستان‌شناسی مقصد است.

در حال حاضر هستان‌شناسی وردنت به عنوان هستان‌شناسی مقصد سیستم POPTA انتخاب شده است و مفاهیم آن با استفاده از نمونه‌های موجود در متن نمونه‌دهی می‌شوند. با توجه به وجود نگاشت‌هایی میان وردنت و هستان‌شناسی‌های دیگر (مثلاً SUMO) امکان برقراری اتصال میان متون فارسی و سایر هستان‌شناسی‌ها نیز وجود خواهد داشت.

از آنجایی که هر مفهوم در هستان‌شناسی وردنت از مجموعه‌ای از لغات هم‌معنی تشکیل شده است، عملیات یافتن مفهوم معادل برچسب مورد نظر تسهیل می‌گردد و نیازی به در نظر گرفتن معادل‌های این برچسب برای تطابق با نام یکی از مفاهیم هستان‌شناسی وجود ندارد. از سوی دیگر یک کلمه بسته به معانی مختلفی که می‌تواند داشته باشد ممکن است در بیش از یک مجموعه مترادف ظاهر شود. این امر موجب می‌شود برای هر نمونه بیش از یک مفهوم متناظر در وردنت وجود داشته باشد و لذا مجدداً نیاز به روش‌های WSD خواهیم داشت. برای مثال با استفاده از واسط میان زبانی، معادل انگلیسی issue برای مفهوم "نشریه" انتخاب می‌گردد. با توجه به معانی مختلف این کلمه دو مجموعه مترادف متناظر زیر در وردنت انتخاب می‌شوند:

- an important question that is in dispute and must be settled;
- one of a series published periodically;

حال آنکه تنها مجموعه مترادف دوم مورد نظر ماست.

در حالت عمومی می‌توان برای هر یک از برچسب‌های معنایی موجود در سیستم یک تعریف زبان طبیعی در نظر گرفت. برای یافتن این تعاریف زبان طبیعی می‌توان از دایره‌المعارف آزاد ویکی‌پدیا بهره برد. اغلب هنگامی که یک مفهوم عمومی در ویکی‌پدیا جستجو می‌شود، صفحه بازبایی شده حاوی یک تعریف زبان طبیعی کوتاه و کامل برای آن مفهوم خواهد بود. به علاوه می‌توان لیست کلمات هم‌وقوع با این مفهوم را نیز با تجزیه صفحه بازبایی شده و پردازش کلماتی که به صفحات دیگر این دایره‌المعارف پیوند یافته‌اند، استخراج نمود.

با داشتن یک تعریف زبان طبیعی برای هر برچسب و یا کلمات هم‌وقوع با آن برچسب و ترجمه آنها به زبان انگلیسی می‌توان با توجه به شباهت معنایی میان تعریف هر برچسب با تعریف بازبایی شده از وردنت برای هر مجموعه مترادف و میزان هم‌پوشانی این دو تعریف، مناسب‌ترین مجموعه مترادف را از میان مجموعه‌های مترادف کاندیدا انتخاب نمود. مثلاً برای مفهوم کتاب تعاریف زیر ارائه شده:

- ویکی‌پدیا:

"کتاب مجموعه‌ای از صفحه‌های کاغذی است که متنی روی آن‌ها نوشته شده‌است و در یکی از طرف‌ها به هم بسته شده‌اند و مجلد است."

- وردنت:

"a written work or composition that has been published"
"physical objects consisting of a number of pages bound together"

از سوی دیگر از آنجایی که سبزه متن‌های مختلف مورد استفاده در آزمون‌ها و در نتیجه میزان رخداد کلمات کلیدی و نسبت آن‌ها به کل متن متغیر بوده است، بر آن شدیم تا نتایج ارزیابی را یکبار به صورت میانگین عادی و یک بار به صورت میانگین وزندار محاسبه کنیم. در میانگین‌گیری وزندار به هر سند وزنی متناسب با حجم آن تخصیص داده‌ایم. در ادامه در ارائه نتایج ارزیابی‌ها میانگین‌های وزندار داخل پراتر و در مقابل میانگین‌های عادی نوشته شده‌اند. قابل مشاهده است که میانگین‌گیری وزندار تغییر عمده‌ای در میزان دقت و فراخوان ایجاد نمی‌کند

جدول ۳- نتایج ارزیابی تولید برچسب‌های معنایی در سیستم POPTA

(۱) فاز اول - بدون استفاده از ویکی‌پدیا

ورودی	نمونه‌دهی به هستان‌شناسی		حاشیه‌نویسی معنایی	
	دقت	فراخوان	دقت	فراخوان
۱	٪۹۵	٪۷۶	٪۹۹	٪۹۲
۲	٪۹۷	٪۶۳	٪۹۹	٪۹۱
۳	٪۹۱	٪۵۱	٪۸۹	٪۳۶
۴	٪۱۰۰	٪۶۰	٪۱۰۰	٪۵۲
۵	٪۸۲	٪۵۶	٪۸۰	٪۷۷
۶	٪۱۰۰	٪۶۴	٪۱۰۰	٪۸۷
۷	٪۱۰۰	٪۶۹	٪۱۰۰	٪۶۹
۸	٪۸۹	٪۸۶	٪۹۲	٪۹۳
۹	٪۹۴	٪۵۲	٪۹۴	٪۹۰
۱۰	٪۹۵	٪۴۳	٪۹۶	٪۳۱
۱۱	٪۱۰۰	٪۵۲	٪۱۰۰	٪۸۰
۱۲	٪۸۱	٪۶۳	٪۵۶	٪۳۰
۱۳	٪۹۲	٪۳۸	٪۹۳	٪۳۴
۱۴	٪۸۷	٪۴۶	٪۹۰	٪۱۶
۱۵	٪۹۳	٪۵۵	٪۹۷	٪۷۴

(۲) فاز دوم - با استفاده از ویکی‌پدیا

ورودی	نمونه‌دهی به هستان‌شناسی		حاشیه‌نویسی معنایی	
	دقت	فراخوان	دقت	فراخوان
۱	٪۹۶	٪۹۷	٪۹۵	٪۹۷
۲	٪۹۳	٪۹۲	٪۹۴	٪۹۲
۳	٪۸۸	٪۹۲	٪۸۵	٪۵۰
۴	٪۹۷	٪۹۴	٪۹۸	٪۸۷
۵	٪۸۶	٪۷۵	٪۸۵	٪۹۶
۶	٪۹۵	٪۹۳	٪۱۰۰	٪۸۷
۷	٪۹۷	٪۹۱	٪۹۵	٪۹۵
۸	٪۹۰	٪۹۳	٪۹۴	٪۹۶
۹	٪۹۷	٪۸۷	٪۹۶	٪۹۷
۱۰	٪۹۳	٪۸۳	٪۹۳	٪۸۲
۱۱	٪۱۰۰	٪۹۴	٪۹۸	٪۸۷
۱۲	٪۸۷	٪۹۴	٪۸۹	٪۹۲
۱۳	٪۹۴	٪۸۱	٪۹۷	٪۹۵
۱۴	٪۹۲	٪۶۳	٪۹۶	٪۵۸
۱۵	٪۹۳	٪۷۰	٪۹۷	٪۸۸

پیمانه نمونه‌دهی به هستان‌شناسی و حاشیه‌نویسی معنایی در هر فاز بطور جداگانه محاسبه گردید. در این بخش نتایج ارزیابی را برای سه بخش (الف) انتخاب برچسب معنایی یا مفهوم متناظر با نمونه‌ها، (ب) ترجمه صحیح مفهوم موردنظر به انگلیسی و (ج) یافتن مفهوم (مجموعه مترادف) متناظر با مفهوم یافت شده در وردنت ارائه نموده‌ایم.

الف) انتخاب برچسب معنایی یا مفهوم متناظر با نمونه‌ها

جدول ۳ نتایج حاصل از ارزیابی سیستم (۱) تولید برچسب‌های معنایی بدون استفاده از ویکی‌پدیا و (۲) تولید برچسب‌های معنایی - پس از استفاده از ویکی‌پدیا را نشان می‌دهد.

همانطور که در جدول ۳ فاز اول مشاهده می‌شود این سیستم از دقت بالایی برخوردار است. بیشتر خطاهای رخ داده را می‌توان به ۲ دسته کلی تقسیم‌بندی نمود:

- خطاهای ناشی از استفاده از الگوهای به شکل "نام مفهوم + اضافه + نام نمونه" که همانطور که قبلاً گفته شد فراخوان بالا ولی دقت پایین دارند. مثلاً در عبارتی مانند "یزد، شهر ادیبان و عارفان" با استفاده از این الگو کلمه "ادیبان" را به عنوان نمونه‌ای از مفهوم "شهر" استخراج می‌کند.
- این دسته از خطاها را می‌توان با استفاده از ویژگی‌های دیگری که برای هر مفهوم قابل استخراج است، کاهش داد. این ویژگی‌ها می‌توانند از توصیف‌های زبان طبیعی موجود برای هر مفهوم در هستان‌شناسی ورودی و یا در سطح وب و نیز از منابع اطلاعاتی دیگر مانند انواع دایره‌المعارف‌های موجود در سطح وب (از جمله ویکی‌پدیا) استخراج شوند. راه‌حل دیگر آن است که از دایره‌المعارف ویکی‌پدیا علاوه بر برچسب‌دهی نمونه‌های باقیمانده برای بررسی میزان صحت نتایج حاصل از جستجوی گوگل بهره ببریم. بدین منظور می‌توان نتایجی که از جستجوی پرس و جوهای با الگوی ذکر شده بدست آمده را بار دیگر در ویکی‌پدیا جستجو نماییم تا درستی یا نادرستی آن مشخص گردد.

- خطاهای ناشی از WSD که تنها بر بخش حاشیه‌نویسی تاثیر می‌گذارد. این خطاها به دلیل عملیات رفع ابهام برخی کلمات توسط موتور جستجوی گوگل رخ می‌دهند. از آنجایی که گاه این جستجو مستقل از متن ورودی و در سطح وب صورت می‌گیرد و نمونه موجود در متن به مفهومی با بیشترین تعداد رخداد در سطح وب تخصیص می‌یابد، علیرغم این‌که مفهوم انتخاب شده مفهومی مناسب برای نمونه استخراج شده است و خطایی در بخش نمونه‌دهی به هستان‌شناسی رخ نداده است، ممکن است اتصال این مفهوم به آن بخش از متن صحیح نباشد؛ چرا که نمونه موجود در متن می‌تواند معنی دیگری از مفهوم یافت شده در وب باشد. در این بخش نیز می‌توان از دایره‌المعارف ویکی‌پدیا جهت رفع ابهام معنایی این نمونه‌ها بهره برد. با توجه به ویژگی‌ها و امکانات منحصر بفرد این دانشنامه و بهره‌برداری سیستم‌های مختلف از آن به‌منظور استخراج دانش و معنا می‌توان استفاده از این دایره‌المعارف را جهت ارتقاء بخش حاشیه‌نویسی در فهرست کارهای آینده این سیستم قرار داد.

همانطور که در جدول ۳ فاز دوم مشاهده می‌شود، استفاده از ویکی‌پدیا به میزان قابل توجهی فراخوان سیستم را افزایش داده و بر کارایی سیستم افزوده است. میزان کاهش دقت سیستم در این فاز در مقابل این افزایش قابل چشم‌پوشی است.

تعداد خطاهای این فاز کمتر از فاز قبل است. خطاهای موجود در این فاز عمدتاً به ناکامل بودن مجموعه کلمات کلیدی و مکاشفه‌های بکار رفته مربوط می‌شود. با گسترش مجموعه لغات هم‌وقوع با هر یک از برچسب‌ها و افزایش مکاشفه‌های مناسب برای یافتن مرتبط‌ترین کلمه کلیدی همسایه می‌تواند میزان خطاهای این بخش را کاهش دهد.

خروجی‌های POPTA یک پیکره فارسی حاشیه‌نویسی شده برای انجام حاشیه‌نویسی‌های معنایی بیشتر مبتنی بر پیکره برای زبان فارسی خواهد بود. POPTA در حال حاضر برای نمونه‌دهی هستان‌شناسی وردنت با در نظر گرفتن اسامی خاص به عنوان نمونه‌هایی از مفاهیم این هستان‌شناسی پیاده‌سازی شده است. گام بعدی برای توسعه این سیستم در نظر گرفتن اسامی عام و افعال به عنوان نمونه‌ها و نیز اتصال آنها به هستان‌شناسی‌های دیگر از جمله SUMO می‌باشد.

ارتقاء مؤلفه WSD نیز یکی دیگر از بخش‌هایی است که در لیست فعالیت‌های آینده قرار دارد.

برای پیشرفت هرچه بیشتر سیستم‌های استخراج‌گر دانش و حرکت به سوی داشتن یک وب معنایی برای زبان فارسی، در مرحله اول نیازمند توسعه مؤلفه‌های آماده پردازش زبان فارسی و نیز مؤلفه‌های استخراج اطلاعات (IE) می‌باشیم. این امر نمونه‌دهی به هستان‌شناسی‌ها و نیز حاشیه‌نویسی معنایی متون فارسی را تسهیل می‌نماید.

ایجاد استاندارد یکتا برای کدگذاری‌های مربوط به کاراکترهای زبان فارسی نیز می‌تواند به اعتبار اطلاعات استخراج شده از سطح وب بیفزاید.

استفاده از وب و Google API کمک شایانی به عملیات رفع ابهام در دسته‌بندی و غلبه بر گلوگاه استخراج دانش کرده است ولی مشکلاتی نیز در این میان وجود دارد. از قبیل اطلاعات نادرست موجود در سطح وب که می‌تواند منجر به استخراج دانش نادرست از متون گردد و نیز غیر قابل اعتماد بودن گوگل که گاهی باعث می‌گردد نتایج حاصل از اجرای چند باره سیستم با ورودی‌های یکسان، متفاوت باشد. برای رفع این مشکل نیازمند اجرای چند باره سیستم برای هر متن ورودی می‌باشیم.

همانطور که در بخش ارزیابی دیدیم، استفاده از ویکی‌پدیا به میزان قابل توجهی فراخوان سیستم را افزایش داده و بر کارایی سیستم افزوده است. به علاوه خطاهای این فاز نیز از فاز قبل بسیار کمتر و سرعت آن نیز بالاتر است. بنابراین مهمترین کاری که در فهرست کارهای آینده این سیستم قرار می‌گیرد افزایش نقش ویکی‌پدیا در برچسب‌دهی به نمونه‌های استخراج شده از متن و همچنین بررسی میزان درستی برچسب‌های تخصیص داده شده در فاز اول می‌باشد.

برای گسترش بخش ویکی‌پدیا در این پروژه نیازمند بهبود بخشی به روش‌های استخراج اطلاعات از صفحات این دایرالمعارف و نیز آشنایی بیشتر با ساختارهای موجود در آن هستیم. بنابراین در اولین گام به توسعه یک تجزیه‌گر صفحات ویکی‌پدیا پرداخته و در گام دوم از ساختارهای موجود در این دایرالمعارف برای استخراج اطلاعات معنایی در حوزه‌ای گسترده‌تر بهره می‌گیریم.

مراجع

[1] L. Reeve, and H. Han, "Survey of Semantic Annotation Platforms," *Proc. ACM Symp. on Applied Computing*, PP. 1634-1638, 2005.

[۲] ب. صرافزاده، نمونه‌دهی هستان‌شناسی و حاشیه‌نویسی خودکار متون فارسی، پایان‌نامه کارشناسی، دانشکده مهندسی برق و کامپیوتر، دانشگاه شهید بهشتی، ۱۳۸۷.

[3] M. Vargas-Vera, and D. Celjuskja, "Event Recognition on News Stories and Semi-Automatic Population of an Ontology," *Proc. IEEE/WIC/ACM Int'l Conf. on Web Intelligence*, pp. 615-618, 2004.

بطور میانگین این سیستم قبل از استفاده از ویکی‌پدیا به دقت ۹۳٪ (۹۳/۴) و فراخوان ۵۸۸/۲٪ (۵۶/۱) در بخش نمونه‌دهی به هستان‌شناسی و دقت ۹۲۳/۳٪ (۹۲/۴) و فراخوان ۶۳۳/۴٪ (۵۹) در بخش حاشیه‌نویسی معنایی دست یافته است. خروجی نهایی سیستم که پس از استفاده از ویکی‌پدیا بدست آمده است دقت ۹۳/۲٪ (۹۳/۲) و فراخوان ۸۶/۶٪ (۸۷) در بخش نمونه‌دهی به هستان‌شناسی و دقت ۹۴٪ (۹۴) و فراخوان ۸۶/۶٪ (۸۶/۲) در بخش حاشیه‌نویسی معنایی را نشان می‌دهد.

مطابق نتایج فوق استفاده از ویکی‌پدیا منجر به افزایش کارایی سیستم با حذف موارد نامعلوم در فازهای قبلی می‌شود. برای مثال اسامی خاصی که هیچ کلمه کلیدی مناسبی در همسایگی آنها قرار ندارند و مکاشفه‌ای نیز برای تشخیص نوع آنها در دست نداریم در فاز اول بدون برچسب باقی می‌مانند. مثلاً در جمله "در آن هنگام بوشهر مرکز عبور بازرگانی بود که از هند و هندوچین و چین به سوی عمان و آفریقا و اروپا کالا حمل می‌کردند"، کلمات "بوشهر"، "هند"، "هندوچین"، "چین"، "عمان"، "آفریقا" و "اروپا" که همگی اسامی خاص هستند در فاز اول بدون برچسب باقی می‌مانند؛ حال آن‌که با دادن هر یک از این کلمات به عنوان پرس و جو به ویکی‌پدیا به ترتیب برچسب‌های "شهر"، "کشور"، "منطقه"، "کشور"، "کشور"، "قاره" و "قاره" به این هفت نمونه تخصیص می‌یابد.

البته در این مرحله نیز به علت وجود کلماتی با معانی گوناگون، ممکن است در بخش حاشیه‌نویسی دچار مشکل شویم، در حالی که عملیات نمونه‌دهی به خوبی صورت می‌گیرد. مثلاً در جمله "داستان‌های کهن ایران را از پهلوی به فارسی دری ترجمه کرده‌اند." کلمات "ایران"، "پهلوی" و "فارسی" در فاز اول بدون برچسب باقی می‌مانند. پس از جستجو در ویکی‌پدیا به ترتیب برچسب‌های "کشور"، "دودمان" و "زبان" به این سه نمونه تخصیص می‌یابد. حال آنکه مفهوم متناظر با نمونه پهلوی، در این جمله، "زبان" است و در این مرحله با خطای بخش حاشیه‌نویسی روبرو هستیم (گرچه در بخش نمونه‌دهی خطایی نداریم). کلماتی که عنوان هیچ مقاله‌ای در ویکی‌پدیا نیستند (مانند "بهارستان") و یا در صفحه مربوط به آنها کلمه کلیدی مناسبی به کار نرفته است (مانند "خورشید")، در انتهای فاز دوم نیز بدون برچسب باقی می‌مانند.

ب) ترجمه صحیح مفهوم موردنظر به انگلیسی

با ارزیابی روش مطرح شده در بخش ۶-۳ مشاهده گردید که این روش برای ۹۶.۲٪ برچسب‌های فارسی معادل انگلیسی مناسبی ارائه داده است.

ج) یافتن مفهوم (مجموعه مترادف) متناظر با مفهوم یافت شده در وردنت

پس از انتخاب ترجمه مناسب برای هر برچسب به مرحله یافتن مفهوم متناظر با هر برچسب در هستان‌شناسی وردنت می‌رسیم. با ارزیابی نتایج حاصل از روش مطرح شده در بخش ۷-۳ مشاهده گردید برای ۸۱٪ برچسب‌ها، مناسب‌ترین مفهوم همان پرارجاع‌ترین مفهوم می‌باشد؛ برای ۹/۲٪ از برچسب‌ها پرارجاع‌ترین مفهوم گرچه مناسب است ولی مناسب‌ترین مفهوم نیست و لذا خطایی نیز رخ نداده است و برای ۹/۸٪ از برچسب‌ها انتخاب پرارجاع‌ترین مفهوم ایجاد خطا می‌نماید.

۵- نتیجه‌گیری

در این مقاله سیستم POPTA را به عنوان نخستین سیستم نمونه‌دهی و حاشیه‌نویسی موازی برای متون فارسی معرفی نمودیم. POPTA با هدف ایجاد و توسعه سیستم‌های مورد نیاز با استفاده از منابع اولیه محدود ایجاد شده است. این سیستم هستان‌شناسی‌های انگلیسی را با استفاده از متون فارسی نمونه‌دهی نموده و متون فارسی را نیز با توجه به این هستان‌شناسی حاشیه‌نویسی می‌نماید. یکی از

- Learning And Population: Bridging The Gap Between Text And Knowledge*, pp. 26-32, 2006.
- [17] V. Borkar, K. Deshmukhy, and S. Sarawagiz, "Automatic Segmentation of Text into Structured Records," *Proc, ACM SIGMOD Int'l Conf. on Management of Data*, pp. 175-186, 2001.
- [18] M. Vargas-Vera, E. Motta, J. Domingue, M. Lanzoni, A. Stutt, and F. Ciravegna, "MnM: Ontology Driven Semi-Automatic and Automatic Support for Semantic Markup," *Proc, 13th Int'l Conf. on Knowledge Engineering and Management*, pp. 379-391, 2002.
- [19] S. Handschuh, S. Staab, and F. Ciravogna, "S-CREAM- - Semi-automatic CREATION of Metadata in SAAKM 2002," *proc, Semantic Authoring, Annotation & Knowledge Markup Workshop*, pp. 27-33, 2002.
- [20] A. Dingli, F. Ciravegna, and Y. Wilks, "Automatic Semantic Annotation using Unsupervised Information Extraction and Integration," *Proc, K-CAP 2003 Workshop on Knowledge Markup and Semantic Annotation*, pp. 2-9, 2003.
- [21] B. Popov, A. Kiryakov, A. Kirilov, D. Manov, D. Ognyanoff, and M. Goranov, "KIM – Semantic Annotation Platform," *Proc, 2nd Int'l Semantic Web Conf.*, pp. 834-849, 2003.
- [22] M. Ruiz-Casado, E. Alfonseca, and P. Castells, "From Wikipedia to Semantic Relationships: a Semi-automated Annotation Approach," *Proc, 1st Workshop on Semantic Wikis: From Wiki to Semantics, at the 3rd European Semantic Web Conf.*, Vol. 206 of Workshop on Semantic Wikis, 2006.
- [23] M. Ruiz-Casado, E. Alfonseca, M. Okumura, and P. Castells, "Information Extraction and semantic annotation of Wikipedia," *Ontology Learning and Population: bridging the gap between text and knowledge*, pp. 145- 169, 2008.
- [24] C. Jonquet, N. H. Shah, and M. A. Musen, "The Open Biomedical Annotator," *Proc, AMIA Summit on Translational Bioinformatics*, pp. 56-60, 2009.
- [25] م. شمس‌فرد، طراحی مدل یادگیر هستان‌شناسی: نمونه‌سازی در یک سیستم درک متن فارسی، رساله دکتری، دانشگاه صنعتی امیرکبیر، ۱۳۸۱.
- [26] T. Riddle, Parse::mediawikidump, URL: <http://search.cpan.org/~triddle/Parse-MediaWikiDump-0.40/>, 2006.
- [27] T. Zesch, I. Gurevych, and M. Mühlhäuser, "Analyzing and Accessing Wikipedia as a Lexical Semantic Resource," *Proc, Biannual Conf. of the Society for Computational Linguistics and Language Technology*, pp. 213-221, 2007.
- [28] F. Suchanek, G. Kasneci, and G. Weikum, "YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia," *Proc, 16th Int'l Conf. on World Wide Web*, pp. 697-706, 2007.
- [4] Ch. Patel, K. Supekar, and Y. Lee, "OntoGenie: Extracting Ontology Instances from WWW," *Human Language Technology for the Semantic Web and Web Services*, pp. 123-126, 2003.
- [5] V. De Boer, M. Van Someren, and B. J. Wielinga, "Relation instantiation for ontology population using the web," *Lecture Notes in Artificial intelligence (LNAI)*, Vol. 4314, pp. 202-213, 2007.
- [6] P. Cimiano, and S. Staab, *Learning by Googling*, ACM press, 2004.
- [7] A. Hearst, "Automatic Acquisition of Hyponyms from Large Text Corpora," *Proc, 14th Int'l Conf. on Computational Linguistics*, pp. 539-545, 1992.
- [8] G. Geleijnse, and J. Korst, "Automatic Ontology Population By Googling," *Proc, 7th Belgium-Netherlands Conf. on Artificial Intelligence*, pp. 120-126, 2005.
- [9] R. Bunescu, and M. Pasca. "Using encyclopedic knowledge for named entity disambiguation," *Proc, the European Conf. of the Association for Computational Linguistics*, pp. 9-16, 2006.
- [10] S. Chernov, T. Iofciu, W. Nejdl, and X. Zhou, "Extracting Semantic Relationships between Wikipedia Categories," *Proc, 1st Int'l Workshop: SemWiki2006 — From Wiki to Semantics*, Vol. 206 of Workshop on Semantic Wikis, 2006.
- [11] S. Soltani, and A. Abdollahzadeh Barforoush, "OILSW: A New System for Ontology Instance Learning in Semantic Web," *Proc, Int'l Conf. on Semantic Web and Digital Libraries*, pp. 54-63, 2007.
- [12] P. Cimiano, A. Pivk, L.S. Thieme, and S. Staab, "Learning Taxonomic Relations from Heterogeneous Sources of Evidence," *Workshops on Ontology Learning from Text: Methods, Evaluation and Applications*, Vol. 123, pp. 59-73, 2004.
- [13] P. Velardi, R. Navigli, A. Cuchiarelli, and F. Neri. "Evaluation of Ontolearn, a Methodology for Automatic Population of Domain Ontologies," *Workshops on Ontology Learning from Text: Methods, Evaluation and Applications*, pp. 92-106, 2005.
- [14] H. Tanev, and B. Magnini, "Weakly Supervised Approaches for Ontology Population," *11th Conf. the European Chapter of the Association for Computational Linguistics*, pp. 17-24, 2006.
- [15] D. Celjuska, and M. Vargas-Vera, "Ontosophie: A Semi-Automatic System for Ontology Population from Text," *Proc, Int'l Conf. on Natural Language Processing*, 2004.
- [16] B. Magnini, E. Pianta, O. Popescu, and M. Speranza, "Ontology Population from Textual Mentions: Task Definition and Benchmark," *Proc, Workshop On Ontology*

²⁹ Synsets
³⁰ Sense



مهرنوش شمس فرد دانش آموخته‌ی کارشناسی و کارشناسی ارشد رشته‌ی مهندسی کامپیوتر گرایش نرم‌افزار از دانشگاه صنعتی شریف و دکتری در رشته‌ی مهندسی کامپیوتر گرایش هوش مصنوعی از دانشگاه صنعتی امیرکبیر است. وی از سال ۱۳۸۲ با عنوان استادیار و سرپرست آزمایشگاه پردازش زبان طبیعی در دانشگاه شهید بهشتی مشغول به فعالیت‌های آموزشی - پژوهشی در زمینه‌های پردازش زبان طبیعی، مهندسی هسته‌شناسی، کاوش متن و وب معنایی است. آدرس پست‌الکترونیکی ایشان عبارت است از:

m-shams@sbu.ac.ir



بهاره صرافزاده مدرک کارشناسی خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار از دانشگاه شهید بهشتی تهران در سال ۸۷ دریافت نمود. او در حال حاضر در حال گذراندن دوره کارشناسی ارشد در دانشگاه یورک (تورنتو، انتاریو) می‌باشد و بر روی ترجمه ماشینی برای زبان فارسی - انگلیسی کار می‌کند. وی عضو آزمایشگاه پردازش زبان طبیعی دانشگاه شهید بهشتی است و علاقه اصلی او رشته‌ی هوش مصنوعی با تمرکز بر وب معنایی و درک متن می‌باشد. آدرس پست‌الکترونیکی ایشان عبارت است از:

bahareh.sarrafzadeh@gmail.com

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۷/۰۶/۲۹

تاریخ اصلاح: ۸۹/۰۳/۰۹

تاریخ قبول شدن: ۸۹/۰۳/۲۶

نویسنده مرتبط: دکتر مهرنوش شمس فرد، دانشکده مهندسی برق و کامپیوتر، دانشگاه شهید بهشتی، تهران، ایران.

- ¹ Cross-Lingual
- ² Parallel Ontology Population and Text Annotation
- ³ Ontology Population (or Instantiation)
- ⁴ Lexemes
- ⁵ lexicalize
- ⁶ Cross-Language
- ⁷ Google
- ⁸ Wikipedia
- ⁹ Hearst
- ¹⁰ Head-Matching
- ¹¹ WordNet
- ¹² Semantic Annotation Platform
- ¹³ Wrapper Induction
- ¹⁴ Lexico-Syntactic
- ¹⁵ Segmenter
- ¹⁶ POS tags
- ¹⁷ Token
- ¹⁸ Proper Noun
- ¹⁹ Common Noun
- ²⁰ Morphological
- ²¹ Semantic Tags
- ²² Head-Matching
- ²³ Stop Word
- ²⁴ Googling
- ²⁵ Precision
- ²⁶ Recall
- ²⁷ Application Program Interface
- ²⁸ Word Sense Disambiguation

یک الگوریتم جدید بر مبنای اتوماتاهای یادگیر برای حل مسئله درخت اشتاینر

سمیرا نوفرستی^۱ محمدرضا میبیدی^۲

^۱ دانشکده مهندسی برق و کامپیوتر، دانشگاه سیستان و بلوچستان، زاهدان، ایران
^۲ دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران

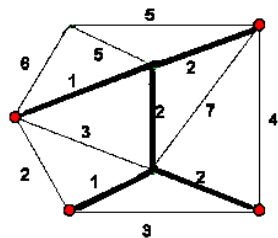
چکیده

مسئله درخت اشتاینر در گراف عبارت است از پیدا کردن کم هزینه‌ترین درختی که شامل تعدادی گره خاص به نام ترمینال باشد. این مسئله از جمله مسائل NP-hard است و به همین دلیل الگوریتم‌های تقریبی متعددی برای حل آن پیشنهاد شده است. اغلب این الگوریتم‌ها درخت‌های اشتاینر نزدیک به بهینه تولید می‌کنند اما از سرعت همگرایی مناسبی برخوردار نیستند. در این مقاله یک الگوریتم تکرار شونده مبتنی بر اتوماتاهای یادگیر برای حل مسئله اشتاینر ایستا و پویا پیشنهاد می‌شود. نتایج شبیه‌سازی‌های انجام گرفته کارایی الگوریتم پیشنهادی را هم از لحاظ کیفیت جواب‌های تولید شده و هم از لحاظ سرعت همگرایی به جواب در مقایسه با الگوریتم‌های گزارش شده نشان می‌دهد.

کلمات کلیدی: درخت اشتاینر ایستا، درخت اشتاینر پویا، اتوماتاهای یادگیر، مسایل مشکل.

۱- مقدمه

غیرترمینال نیز استفاده می‌شود. این گره‌ها نقاط اشتاینر نامیده می‌شوند. مثالی از درخت اشتاینر در یک گراف نمونه در شکل ۱ نشان داده شده است. در این شکل گره‌های ترمینال و یال‌های درخت اشتاینر به صورت پررنگ رسم شده‌اند.



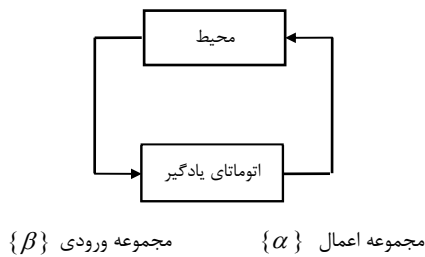
شکل ۱- درخت اشتاینر در گراف

برای پیاده‌سازی انتقال چندپخشی در شبکه‌های کامپیوتری ایستا و شبکه‌های کامپیوتری سیار نیاز به ساخت درخت‌های چندپخشی می‌باشد. یکی از متداولترین راه‌های ساخت درخت‌های چندپخشی استفاده از الگوریتم‌های ساخت درخت اشتاینر^۱ در گراف است.

مسئله درخت اشتاینر در گراف به صورت زیر تعریف می‌شود: گراف $G(V, E)$ را در نظر بگیرید که V مجموعه گره‌ها و $E \subseteq V * V$ مجموعه یال‌های گراف را نشان می‌دهد. به هر یال (i, j) در گراف یک هزینه $c(i, j)$ نسبت داده می‌شود. یک زیرمجموعه از گره‌ها به نام T نیز به عنوان مجموعه ترمینال‌ها تعریف می‌شود. هدف مسئله اشتاینر یافتن درختی در گراف است که مجموعه ترمینال‌ها را در برگیرد و هزینه آن حداقل باشد. این درخت، درخت اشتاینر کمینه نامیده می‌شود. اگر درخت حاصل فقط شامل ترمینال‌ها باشد، درخت پوشای کمینه^۲ نامیده می‌شود و الگوریتم‌های چندجمله‌ای متعددی مانند الگوریتم پریم^۳ برای حل آن وجود دارد. با این وجود در حالت کلی برای کاهش هزینه درخت از گره‌های

در شبکه‌های سیار، توپولوژی شبکه در طی زمان به دلیل اضافه و یا حذف شدن گره‌ها تغییر می‌کند که در این شرایط مسئله درخت اشتاینر پویا مطرح می‌شود. در این موارد هدف یافتن یک درخت اشتاینر بعد از تغییرات انجام گرفته

این مقادیر در طی زمان تغییر می‌کنند. شکل ۲ ارتباط بین اتوماتای یادگیر و محیط را نشان می‌دهد.



شکل ۲- ارتباط بین اتوماتای یادگیر و محیط

اتوماتاهای یادگیر به دو گروه اتوماتای یادگیر با ساختار ثابت و اتوماتای یادگیر با ساختار متغیر تقسیم می‌شوند. در ادامه به شرح مختصری درباره اتوماتاهای یادگیر با ساختار متغیر که در این مقاله استفاده شده است می‌پردازیم.

۲-۱- اتوماتای یادگیر با ساختار متغیر

اتوماتای یادگیر با ساختار متغیر توسط δ تایی $LA \equiv \{\alpha, \beta, p, T, c\}$ نشان داده می‌شود که $\alpha \equiv \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ مجموعه عمل‌ها، $\beta \equiv \{\beta_1, \beta_2, \dots, \beta_m\}$ مجموعه ورودی‌های اتوماتای یادگیر، $p \equiv \{p_1, p_2, \dots, p_r\}$ بردار احتمال انتخاب عمل‌ها، $T \equiv p(n+1) = T[\alpha(n), \beta(n), p(n)]$ الگوریتم یادگیری و $c \equiv \{c_1, c_2, \dots, c_r\}$ احتمال جریمه شدن هر عمل می‌باشد. اگر در اتوماتای یادگیر عمل α_i در مرحله n ام انتخاب شود و پاسخ مطلوب از محیط دریافت نماید، احتمال $p_i(n)$ افزایش یافته و سایر احتمال‌ها کاهش می‌یابند و برای پاسخ نامطلوب احتمال $p_i(n)$ کاهش یافته و سایر احتمال‌ها افزایش می‌یابند. در هر حال، تغییرات به گونه‌ای صورت می‌گیرد تا حاصل جمع $p_i(n)$ همواره ثابت و مساوی یک باقی بماند.

رابطه (۱) نحوه بروزرسانی بردار احتمالات یک اتوماتای یادگیر هنگامی که این اتوماتا پاسخ مطلوب از محیط دریافت می‌کند را نشان می‌دهد.

$$\begin{aligned} p_i(n+1) &= p_i(n) + a[1 - p_i(n)] \quad \forall j \quad j \neq i \\ p_j(n+1) &= (1 - a)p_j(n) \end{aligned} \quad (1)$$

در رابطه (۲) بروزرسانی بردار احتمالات اتوماتای یادگیر هنگامی که پاسخ نامطلوب از محیط دریافت می‌کند، نشان داده شده است.

$$\begin{aligned} p_i(n+1) &= (1 - b)p_i(n) \\ p_j(n+1) &= \frac{b}{r-1} + (1-b)p_j(n) \quad \forall j \quad j \neq i \end{aligned} \quad (2)$$

در روابط فوق، a پارامتر پاداش و b پارامتر جریمه می‌باشد. با توجه به مقادیر a و b سه حالت مختلف را می‌توان در نظر گرفت. زمانیکه a و b با هم برابر باشند، الگوریتم را L_{RP} می‌نامند. زمانیکه b از a خیلی کوچکتر باشد، الگوریتم را L_{RE} و زمانیکه b مساوی صفر باشد، الگوریتم را L_{RI} می‌نامند. برای مطالعه بیشتر در رابطه با اتوماتاهای یادگیر با ساختار ثابت و متغیر می‌توان به [۱۲، ۱۳] مراجعه نمود.

در گراف می‌باشد. ساخت درخت‌های چندپختی در شبکه‌های سیار به مراتب دشوارتر از ساخت این درخت‌ها در شبکه‌های ایستا می‌باشد. یک الگوریتم کارا برای ساخت درخت‌های چندپختی در شبکه‌های سیار بایستی بتواند درخت‌های مناسب را سریع تولید کند حتی اگر چند تغییر همزمان در گروه چندپختی رخ دهد.

به دلیل اهمیت مسئله درخت اشتاینر تاکنون الگوریتم‌های متعددی برای حل آن گزارش شده است. الگوریتم‌های دقیق که معمولاً از روشهای برنامه‌نویسی پویا و برش‌وانشعب استفاده می‌کنند [۱، ۲، ۳] در بدترین حالت دارای پیچیدگی نمایی هستند و برای استفاده در کاربردهای عملی مناسب نمی‌باشند. به همین دلیل الگوریتم‌های تقریبی متعددی برای حل مسئله اشتاینر در گراف گزارش شده است [۴، ۵، ۶].

تلاشهای زیادی نیز در جهت حل تقریبی مسئله اشتاینر پویا انجام گرفته است [۷، ۸]. در سال‌های اخیر نیز الگوریتم‌های تکرار شونده مانند الگوریتم ژنتیکی [۹] و الگوریتم کلونی مورچه‌ها [۱۰] برای حل مسئله درخت اشتاینر ایستا و پویا پیشنهاد شده است. در این روش‌ها رسیدن به یک پاسخ بهینه تضمین نمی‌شود اما در اغلب موارد جواب‌های تقریبی قابل قبولی تولید می‌کنند.

در این مقاله ابتدا یک الگوریتم تکرار شونده مبتنی بر اتوماتاهای یادگیر برای حل مسئله اشتاینر ایستا پیشنهاد می‌شود و سپس با استفاده از این الگوریتم، الگوریتم دیگری برای حل مسئله اشتاینر پویا ارائه می‌گردد. در الگوریتم حل مسئله اشتاینر ایستا، هر گره غیرترمینال از گراف به یک اتوماتای یادگیر مجهز است. هر اتوماتای یادگیر دارای دو عمل می‌باشد که حضور یا عدم حضور گره متناظر در درخت اشتاینر را مشخص می‌کند. در هر تکرار از الگوریتم اتوماتاهای یادگیر به صورت همزمان فعال شده و یک عمل را انتخاب می‌کنند. سپس بر روی گره‌هایی که اتوماتاهای یادگیر آنها عمل حضور در درخت اشتاینر را انتخاب کرده‌اند یک درخت پوشای کمینه ایجاد می‌شود. با توجه به هزینه درخت بدست آمده از هر تکرار، بردار احتمالات اتوماتاهای یادگیر بروز می‌شود. این روند به دفعات تکرار می‌گردد و در پایان بهترین درخت بدست آمده به عنوان جواب نهایی انتخاب می‌شود. نتایج شبیه‌سازی‌های انجام گرفته کارایی الگوریتم پیشنهادی را هم از لحاظ کیفیت جوابهای تولید شده و هم از لحاظ سرعت همگرایی به جواب در مقایسه با تعدادی از الگوریتم‌های گزارش شده نشان می‌دهد.

ادامه مقاله بدین صورت سازماندهی شده است. ابتدا در بخش ۲ اتوماتاهای یادگیر به صورت اجمالی معرفی می‌گردد. در بخش ۳ الگوریتم پیشنهادی برای حل مسئله درخت اشتاینر ایستا و نتایج شبیه‌سازی‌های انجام گرفته ارائه می‌شود. در بخش ۴ الگوریتم پیشنهادی برای حل به مسئله درخت اشتاینر پویا و نتایج شبیه‌سازی‌ها ارائه می‌شود. بخش پایانی مقاله نتیجه‌گیری می‌باشد.

۲- اتوماتاهای یادگیر

اتوماتای یادگیر یک مدل انتزاعی است که می‌تواند تعداد محدودی عمل را انجام دهد. هر عمل انتخاب شده توسط محیطی تصادفی ارزیابی می‌گردد و پاسخی به اتوماتای یادگیر داده می‌شود. اتوماتای یادگیر از این پاسخ استفاده نموده و عمل خود را برای مرحله بعد انتخاب می‌کند. محیط تصادفی را می‌توان با سه‌تایی $E \equiv \{\alpha, \beta, c\}$ تعریف نمود که $\alpha \equiv \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ مجموعه ورودی‌ها، $\beta \equiv \{\beta_1, \beta_2, \dots, \beta_m\}$ مجموعه خروجی‌ها و $c \equiv \{c_1, c_2, \dots, c_r\}$ مجموعه احتمال‌های جریمه شدن می‌باشد [۱۱].

هرگاه β_i دو مقداری باشد $\beta_i = 1$ به عنوان جریمه و $\beta_i = 0$ به عنوان پاداش در نظر گرفته می‌شود. c_i احتمال اینکه عمل α_i نتیجه نامطلوب داشته باشد است. در محیط پایدار مقادیر c_i بدون تغییر باقی می‌مانند، حال آنکه در محیط ناپایدار

از آنجا که عمل کاهش وقت گیر است و منجر به افزایش زمان اجرای الگوریتم می‌شود، این عمل می‌تواند بر روی تعداد کمی از درخت‌های بدست آمده انجام شود.

۳-۱- نتایج شبیه‌سازی

الگوریتم پیشنهادی بر روی گرافهای مجموعه B از OR-Library آزمایش شده است. این مجموعه شامل ۱۸ گراف است که برای هر گراف تعداد گره ها، تعداد یالها و هزینه درخت اشتاینر بهینه در گراف مشخص است [۱۴].

در آزمایش اول برای پیدا کردن کاراترین الگوریتم یادگیری، الگوریتم پیشنهادی با الگوریتم‌های یادگیری L_{RP} و L_{REP} پیاده‌سازی شده است. در L_{RP} مقدار پارامترهای a و b برابر ۰/۳ و در L_{REP} مقدار پارامتر a برابر ۰/۳ و مقدار پارامتر b برابر ۰/۱ در نظر گرفته شده است. نتایج میانگین ۲۰ بار اجرای این الگوریتم‌ها با ۵۰۰ تکرار در هر اجرا در جدول ۱ ارائه شده است. ستون اول جدول شماره گراف در مجموعه B مسائل بیسلی، ستون دوم هزینه درخت اشتاینر بهینه برای هر گراف که در [۱۴] گزارش شده است و ستونهای بعدی درصد خطای نسبی حاصل از اجرای الگوریتم پیشنهادی با الگوریتم‌های یادگیر مختلف را نشان می‌دهد. برای محاسبه خطای نسبی از رابطه زیر استفاده می‌شود:

$$relative - error = \frac{C - C_{opt}}{C_{opt}} \quad (5)$$

که C هزینه درخت اشتاینر حاصل از الگوریتم و C_{opt} هزینه درخت اشتاینر بهینه در گراف است. عمل کاهش تنها در ۳۰ تکرار انتهایی انجام گرفته است. همان‌طور که از جدول ۱ مشخص است استفاده از اتوماتای یادگیر L_{REP} منجر به نتایج بهتری نسبت به L_{RP} می‌شود. الگوریتم یادگیری L_{REP} برای ۱۰ گراف اول مجموعه B بیسلی به جواب بهینه دست می‌یابد. به همین دلیل در آزمایشات بعدی از این روش یادگیری استفاده شده است.

جدول ۱- نتایج الگوریتم NLA با الگوریتم‌های یادگیری L_{REP} و L_{RP}

شماره گراف	هزینه بهینه	L_{RP}	L_{REP}
۱	۸۲	۰	۰
۲	۸۳	۰	۰
۳	۱۳۸	۰	۰
۴	۵۹	۲/۵۴	۰
۵	۶۱	۱/۶۴	۰
۶	۱۲۲	۰	۰
۷	۱۱۱	۵/۴۹	۰
۸	۱۰۴	۲/۸۸	۰
۹	۲۲۰	۰	۰
۱۰	۸۶	۲/۰۹	۰

در جدول ۲ کیفیت جواب‌های بدست آمده در روش پیشنهادی با تعدادی از الگوریتم‌های حل مسئله اشتاینر مقایسه شده است. ستون اول جدول شماره گراف در مجموعه B بیسلی، ستون دوم هزینه درخت اشتاینر بهینه و سایر ستون‌ها درصد خطای نسبی تعدادی از الگوریتم‌های گزارش شده برای حل مسئله درخت اشتاینر شامل SDG، SPH، ADH و ACS را نشان می‌دهد. الگوریتم SPH برای ساخت درخت اشتاینر یک گره ترمینال را به تصادف انتخاب می‌کند. سپس تا قرار

۳- حل مسئله اشتاینر ایستا با استفاده از اتوماتاهای یادگیر

در الگوریتم پیشنهادی که آنرا NLA^y می‌نامیم، گراف مسئله با یک شبکه از اتوماتاهای یادگیر مدل می‌شود. هر گره غیر ترمینال از گراف به یک اتوماتای یادگیر با دو عمل حضور گره در درخت اشتاینر و عدم حضور گره در درخت اشتاینر مجهز است. در ابتدای الگوریتم، احتمال انتخاب این اعمال مساوی و برابر ۰/۵ می‌باشد.

در هر تکرار از الگوریتم همه اتوماتاهای یادگیر به صورت همزمان فعال شده و هر یک از آنها طبق بردار احتمالات انتخاب اعمال خود یکی از دو عمل خود را انتخاب می‌کند. سپس با استفاده از الگوریتم پریم یک درخت پوشای کمینه بر روی گره‌هایی که انتخاب شده‌اند ایجاد می‌شود. در پایان درخت حاصل هرس می‌شود به این صورت که گره‌های غیر ترمینالی که در برگ‌های درخت واقع شده‌اند، حذف می‌گردند. در این روش با احتمال $1-q$ هر اتوماتای یادگیر گره z را با احتمال $p(j)$ انتخاب و با احتمال q گره z را به صورت حریصانه انتخاب می‌کند. برای تعیین مقدار پارامتر q از یک اتوماتای یادگیر دیگر استفاده می‌شود [۱۵].

در هر تکرار از هزینه درخت بدست آمده برای ارزیابی عمل انتخابی اتوماتاها استفاده می‌شود. اگر هزینه درخت فعلی از هزینه بهترین درخت بدست آمده تا این مرحله کمتر باشد، اعمال انتخاب شده توسط اتوماتاهای یادگیر از طریق افزایش احتمال انتخاب آنها طبق رابطه زیر پاداش داده می‌شود. فرمول یادگیری برای پاداش عمل انتخابی بصورت زیر می‌باشد:

$$p(t+1) = p(t) + \theta.a.(1 - p(t)) \quad (3)$$

$$\theta = \left| \frac{f(c) - c_{min}}{f(c)} \right|$$

در رابطه فوق $f(c)$ هزینه درختی است که در تکرار t تولید شده و c_{min} هزینه بهترین درخت بدست آمده تا این مرحله است. a پارامتر پاداش نامیده می‌شود.

اگر هزینه درخت فعلی از هزینه بهترین درخت اشتاینر بدست آمده تاکنون بیشتر باشد اعمال انتخاب شده توسط اتوماتاهای یادگیر گره‌های شرکت کننده در بهترین درخت حاصل شده تا این مرحله پاداش و اعمال انتخاب شده توسط اتوماتاهای یادگیر دیگر گره‌ها جریمه می‌شوند. برای جریمه عمل انتخابی از فرمول زیر استفاده می‌شود.

$$p(t+1) = p(t) - \theta.b.p_i(n) \quad (4)$$

که b پارامتر جریمه نامیده می‌شود. برای بهبود نتایج حاصل شده، در بعضی از تکرارها بر روی درخت‌های بدست آمده عمل کاهش انجام می‌گیرد. در این عمل برای هر یال از درخت به صورت زیر عمل می‌شود:

۱. با حذف یال از درخت، دو زیردرخت مجزا حاصل می‌شود.
۲. با اجرای الگوریتم جستجوی عمق اول نقاط اشتاینری که در برگ‌های دو زیردرخت واقع شده‌اند، حذف می‌شوند.
۳. با استفاده از الگوریتم دایکسترا کوتاهترین مسیر بین گره‌های دو زیردرخت مشخص می‌شود.

۴. اگر مجموع هزینه‌های دو زیردرخت و هزینه مسیر یافت شده کمتر از هزینه درخت قبلی باشد دو زیردرخت از طریق کوتاهترین مسیر متصل شده و درخت حاصل به عنوان درخت فعلی در نظر گرفته می‌شود.

در آزمایش پایانی برای بررسی کارایی الگوریتمهای پیشنهادی در شبکه‌های واقعی از شبیه‌ساز شبکه ns2^۱ استفاده شده است. با استفاده از اصلاح شده روش واگسمن [۱۵] تعدادی توپولوژی شبکه تولید شده است. توپولوژی شبکه با توزیع تصادفی n گره در ناحیه مستطیل شکل با طول و عرض ۳۰۰ متر ایجاد شده است. برای تولید این گرافها مقدار پارامتر α در روش واگسمن برابر ۰/۵ و مقدار پارامتر β برابر ۰/۱ در نظر گرفته شده است. مشخصات این گرافها در جدول ۳ نشان داده شده است.

جدول ۳- مشخصات شبکه‌های ایجاد شده

شماره گراف	تعداد گره‌ها	تعداد یال‌ها	تعداد ترمینال‌ها
۱	۱۰	۴۳	۳
۲	۲۰	۵۶	۶
۳	۳۰	۱۰۷	۸
۴	۴۰	۱۹۳	۱۱
۵	۵۰	۲۶۷	۱۳
۶	۶۰	۴۶۴	۱۶
۷	۷۰	۵۸۹	۱۸
۸	۸۰	۱۶۸	۲۵
۹	۹۰	۲۰۵	۲۵

در شکل ۴ نتیجه اجرای الگوریتم پیشنهادی بر روی شبکه‌های جدول ۳ با الگوریتمهای تکرار شونده کلونی مورچه‌ها [۱۰] و الگوریتم ژنتیکی [۹] مقایسه شده است. این الگوریتمها بر روی یک کامپیوتر محلی با مشخصات اینتل پنتیوم ۴، ۳.۰۶ گیگاهرتز، ۴۸۰ مگابایت حافظه و ویندوز XP سرویس پک ۲ اجرا شده‌اند. این نتایج حاصل میانگین‌گیری از ۲۰ بار اجرای الگوریتمهاست. تعداد تکرار الگوریتم NLA و ACS در هر اجرا برابر ۵۰۰ در نظر گرفته شده است. در شکل ۴- الف درصد خطای نسبی الگوریتمهای NLA و ACS با یکدیگر مقایسه شده است. از آنجا که هزینه درخت اشتاینر بهینه در گرافهای تولید شده در دسترس نیست برای مقایسه الگوریتمها با یکدیگر، در معیار خطای نسبی بجای هزینه بهینه از هزینه حاصل از اجرای الگوریتم ژنتیک با تعداد تکرارهای زیاد استفاده شده است.

در شکل ۴- ب زمان اجرای الگوریتمهای NLA، GA و ACS نشان شده است. همان‌طور که در شکل ۴ مشاهده می‌شود، الگوریتم GA در تعداد تکرارهای زیاد جوابهای بهتری نسبت به دیگر الگوریتمها تولید می‌کند با این وجود زمان اجرای این الگوریتم در مقایسه با دیگر الگوریتمها به مراتب بیشتر است. الگوریتم پیشنهادی در مقایسه با الگوریتم کلونی مورچه‌ها به نتایج بهتری رسیده است. به علاوه این الگوریتم زمان محاسباتی کمتری نیاز دارد.

۴- حل مسئله درخت اشتاینر پویا

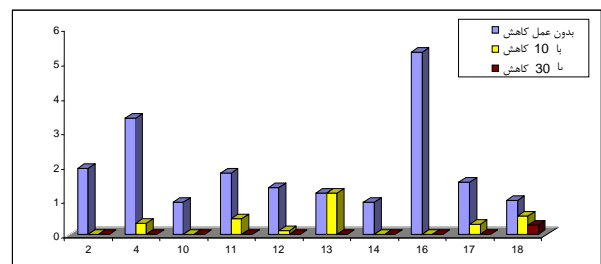
مسئله درخت اشتاینر پویا به صورت زیر تعریف می‌شود: گراف همبند $G(V,E)$ که V مجموعه متناهی از گره‌های گراف و E مجموعه یالها (اتصالات بین گره‌ها) است و یک دنباله از درخواستهای تغییر در گراف را در نظر بگیرید. هدف از مسئله اشتاینر پویا یافتن درختی با حداقل هزینه است که مجموعه ترمینالها را در برگیرد و به هر درخواست رسیده نیز بدون اطلاع از درخواستهای بعدی پاسخ دهد. درخواستهای در نظر گرفته شده عبارتند از افزودن یا حذف یک گره ترمینال، افزودن یا حذف یک گره غیرترمینال، افزودن یا حذف یک یال و تغییر وزن یک یال.

گرفتن کلیه گره‌های ترمینال در درخت، در هر مرحله یال با کمترین هزینه را به درخت اضافه می‌کند [۴]. الگوریتم ADH از گره‌های ترمینال شروع می‌کند. هر گره ترمینال نزدیکترین گره به خود را انتخاب می‌کند. با این کار مجموعه‌ای از جنگلها ایجاد می‌گردد که در پایان الگوریتم از ادغام آنها یک درخت ایجاد می‌شود [۵]. الگوریتم ACS ابتدا بر روی هر گره ترمینال یک مورچه قرار می‌دهد. در هر تکرار از الگوریتم کلیه مورچه‌ها با توجه به غلظت فرمون یال‌های مجاور یک یال را انتخاب و به سمت آن حرکت می‌کنند. در صورت برخورد یک مورچه به مسیر مورچه دیگر درختهای آنها ادغام می‌شود. این مراحل آنقدر تکرار می‌شود تا نهایتاً یک درخت باقی بماند [۱۰]. همان‌طور که در جدول ۲ نشان داده شده است الگوریتم پیشنهادی برای ۱۴ گراف از ۱۸ گراف مجموعه B بیسلی به جواب بهینه دست یافته است. در مورد ۴ گراف دیگر نیز در صد خطای نسبی الگوریتم NLA کمتر از درصد خطای نسبی الگوریتمهای SDG، SPH و ADH است.

جدول ۲- مقایسه درصد خطای نسبی الگوریتمهای حل مسئله اشتاینر

شماره گراف	هزینه بهینه	SDG %	SPH %	ADH %	ACS %	NLA %
۱	۸۲	۰	۰	۰	۰	۰
۲	۸۳	۸۴/۳	۰	۰	۰	۰
۳	۱۳۸	۱/۴۵	۰	۰	۰	۰
۴	۵۹	۸/۴۳	۵/۰۸	۵/۰۸	۲/۷	۰
۵	۶۱	۴/۹۲	۰	۰	۰	۰
۶	۱۲۲	۴/۹۲	۳/۲۸	۱/۶۴	۱/۸	۰
۷	۱۱۱	۰	۰	۰	۰	۰
۸	۱۰۴	۰	۰	۰	۰	۰
۹	۲۲۰	۲۷/۲	۰	۰	۰/۰۵	۰
۱۰	۸۶	۱۳/۹۵	۴/۶۵	۴/۶۵	۳/۸۴	۰
۱۱	۸۸	۲/۲۷	۲/۲۷	۲/۲۷	۱/۵۹	۱/۸
۱۲	۱۷۴	۰	۰	۰	۰/۱۱	۰
۱۳	۱۶۵	۶/۰۶	۷/۱۸	۲/۲۴	۰	۴/۲
۱۴	۲۳۵	۱/۲۸	۲/۵۵	۰/۴۳	۰/۰۹	۰
۱۵	۳۱۸	۲/۲۰	۰	۰	۰	۰
۱۶	۱۲۷	۷/۸۷	۳/۱۵	۰	۰/۱۶	۱/۲
۱۷	۱۳۱	۳۴/۵	۳/۸۲	۳/۰۵	۱/۲۲	۱/۳
۱۸	۲۱۸	۴/۵۹	۱/۸۳	۰	۱/۴۲	۰

برای بررسی تأثیر عمل کاهش، الگوریتم پیشنهادی یکبار بدون عمل کاهش و بار دیگر با اجرای عمل کاهش بر روی تکرارهای نهایی الگوریتم (به تعداد ۱۰ و ۳۰) اجرا شده است. شکل ۳ درصد خطای نسبی الگوریتم NLA با تعداد عمل‌های کاهش متفاوت را نشان می‌دهد. همان‌طور که مشاهده می‌شود عمل کاهش باعث بهبود نتایج حاصل از الگوریتم می‌شود.



شکل ۳- درصد خطای نسبی الگوریتم NLA با عمل‌های کاهش متفاوت

اقدامها قبل از تغییرات در نظر گرفته شود، با تعداد تکرارهای بسیار کمتر از تعداد تکرارهای حالتی که مقادیر بردار احتمالات اقدامها به طور تصادفی انتخاب شده باشند به جواب مطلوب و نزدیک به بهینه رسید. ماهیت تصادفی الگوریتم باعث می شود نسبت به نوع تغییر حساسیت زیادی نداشته باشد.

جدول ۴- درصد خطای نسبی وزن درخت اشتاینر پویا با تعداد تکرارهای مختلف

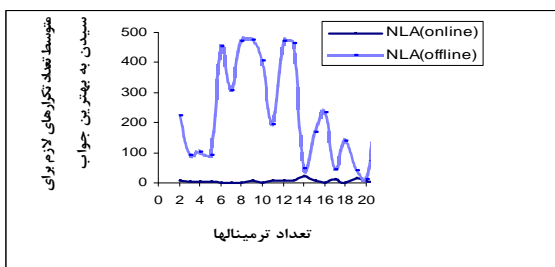
الف) درصد خطای نسبی وزن درخت اشتاینر پویا با تعداد تکرار ۲۰

شماره گراف	تعداد گره	تغییر ۲٪	تغییر ۵٪	تغییر ۱۰٪
۱	۵۰	۰	۰	۰
۳	۵۰	۰	۰	۰
۹	۷۵	۰	۰	۰
۱۵	۱۰۰	۰	۰	۰
۱۸	۱۰۰	۰/۴۴	۰	۰/۸۸

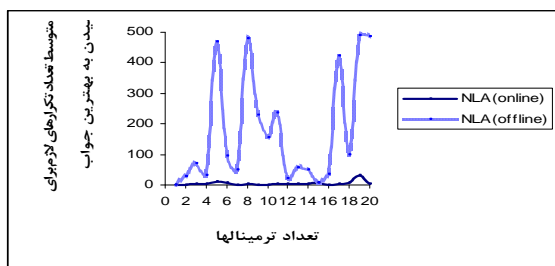
ب) درصد خطای نسبی وزن درخت اشتاینر پویا با تعداد تکرار ۱۰

شماره گراف	تعداد گره	تغییر ۲٪	تغییر ۵٪	تغییر ۱۰٪
۱	۵۰	۰	۰	۰
۳	۵۰	۰	۰	۰
۹	۷۵	۰	۰	۰
۱۵	۱۰۰	۰	۰	۰/۰۶
۱۸	۱۰۰	۰/۸۸	۰	۰/۹۷

برای مقایسه زمان اجرای الگوریتم ایستا و پویا نمودار متوسط تعداد تکرارهای لازم برای رسیدن به بهترین جواب بر حسب تعداد ترمینالها، برای گرافهای شماره ۲ و ۱۱ از مجموعه B مسائل بیسلی در شکل ۵ نشان داده شده است. منظور از بهترین جواب، کمترین هزینه بدست آمده از اجرای الگوریتم ایستا در تکرارهای بالاست. همان طور که در شکل دیده می شود در اکثر موارد الگوریتم NLA پویا نیاز به تعداد تکرارهای بسیار کمتری دارد.

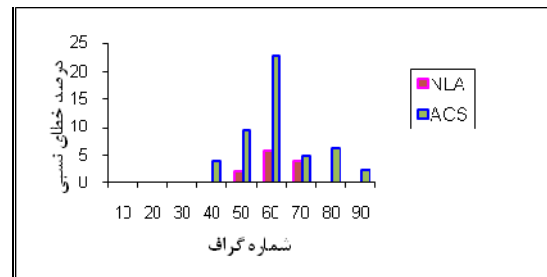


الف- گراف شماره ۲

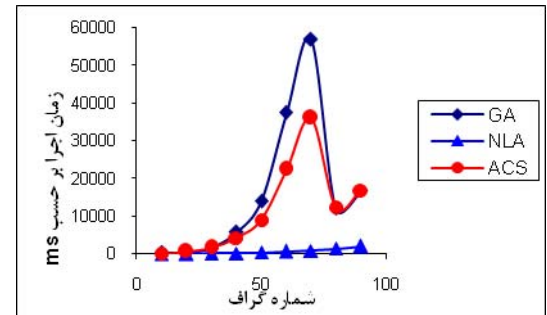


ب- گراف شماره ۱۱

شکل ۵- متوسط تعداد تکرارهای لازم در الگوریتم NLA ایستا و پویا برای رسیدن به جواب بر حسب تعداد ترمینالها



الف) مقایسه درصد خطای نسبی الگوریتمهای ACS و NLA



ب) مقایسه زمان اجرای الگوریتمهای ACS، NLA و GA

شکل ۴- مقایسه الگوریتمهای ACS، NLA و GA

در بخش قبل با استفاده از شبکه‌های یادگیر همکار روشی برای حل مسئله اشتاینر ایستا پیشنهاد شد. در الگوریتم حل مسئله اشتاینر پویا در ابتدا الگوریتم ایستا بر روی گراف قبل از تغییرات اجرا و بردارهای احتمالهای اقدامهای اتوماتاهای یادگیر محاسبه می شود. بعد از تغییرات (حذف و یا اضافه کردن گره به گراف)، الگوریتم درخت اشتاینر ایستا با در نظر گرفتن آخرین مقادیر بردارهای احتمالهای اقدامهای اتوماتاهای یادگیر اجرا می شود. از این طریق با تعداد تکرارهای کمتری درخت اشتاینر مورد نظر تعیین می شود.

۴-۱- نتایج شبیه سازی

جدول ۴ نتیجه الگوریتم NLA را بر روی گرافهای ۱، ۳، ۹، ۱۵ و ۱۸ از مجموعه B از دسته مسائل بیسلی با تعداد تکرارهای ۱۰ و ۲۰ برای هر درخواست را نشان می دهد. در این آزمایشها، ابتدا الگوریتم ایستا بر روی گرافهای اصلی با تعداد تکرار ۵۰ اجرا شده است. سپس هر یک از این گرافها به میزان ۲٪، ۵٪ و ۱۰٪ تعداد گرههای گراف تغییر داده شده و برای هر درخواست الگوریتم حل مسئله اشتاینر ایستا اجرا شده است. برای محاسبه درصد خطای نسبی، الگوریتم ایستا با تعداد تکرارهای زیاد بر روی گراف تغییر یافته اعمال و کم هزینه ترین درخت بدست آمده به عنوان جواب در نظر گرفته شده است.

جدول ۴- الف نتیجه اجرای الگوریتم NLA با تعداد تکرار ۲۰ برای هر درخواست را نشان می دهد. در این جدول متوسط درصد افت کیفیت (بدتر شدن هزینه کل درخت اشتاینر) ۰/۰۹ درصد ثبت شده است. جدول ۴- ب نتیجه اجرای الگوریتم را با تعداد تکرار ۱۰ برای هر درخواست نشان می دهد. در این مرحله کیفیت درخت اشتاینر به میزان کمی بدتر شده است. متوسط خطای نسبی بدست آمده ۰/۱۳ درصد است. در این مرحله نیز متوسط خطا بیشتر شده است ولی نتایج همچنان قابل مقایسه می باشد و اختلاف قابل توجهی دیده نمی شود. بنابراین جدول ۴ نشان می دهد که پس از هر تغییر یا تغییرات، می توان با استفاده از الگوریتم حل مسئله درخت اشتاینر ایستا به شرطی که مقادیر بردار احتمالات

۵- نتیجه گیری

در این مقاله ابتدا یک الگوریتم مبتنی بر اتوماتاهای یادگیر برای حل مسئله اشتاینر ایستا پیشنهاد شد و سپس این الگوریتم برای حل مسئله اشتاینر پویا استفاده گردید. نتایج بدست آمده از آزمایش‌ها نشان داد که الگوریتم پیشنهاد شده در این مقاله برای حل هر دو مسئله درخت اشتاینر ایستا و درخت اشتاینر پویا از کارایی بالایی برخوردار است. نتایج شبیه‌سازی‌های انجام گرفته کارایی الگوریتم پیشنهادی را هم از لحاظ کیفیت جواب‌های تولید شده و هم از لحاظ سرعت همگرایی به جواب در مقایسه با تعدادی از الگوریتم‌های گزارش شده نشان می‌دهد. به علاوه از آنجا که زمان اجرای این روش متناسب با تعداد گره‌های گراف است برای گراف‌های شلوغ (گراف با تعداد یال‌های زیاد) سرعت همگرایی بسیار مناسبی دارد.

مراجع

- [1] S. E. Dreyfuss, and R. A. Wagner, "The Steiner Problem in Graphs," *Networks*, Vol. 1, pp. 195-207, 1971.
- [2] S. Chopra, E. R. Gorres, and M. R. Rao, "Solving the Steiner tree problem on a graph using branch and cut," *ORSA Journal on Computing*, Vol. 4, pp. 320-335, 1992.
- [3] S. Milan, and V. Mirko, "An Exact Algorithm for Steiner Tree Problem on Graphs," *International Journal of Computers, Communications & Control*, Vol. I, No. 1, pp. 41-46, 2006.
- [4] H. Takahashi, and A. Matsuyama, "An Approximate Solution for the Steiner Problem in Graphs," *Mathematica Japonica*, Vol. 24, No. 6, pp. 573-577, 1980.
- [5] V. J. Rayward-Smith, and A. Clare, "On Finding Steiner Vertices," *Networks*, Vol. 16, pp. 283-294, 1986.
- [6] L. Kou, G. Markowsky, and L. Berman, "A Fast Algorithm for Steiner Trees," *Acta Informatica*, Vol. 15, pp. 141-145, 1981.
- [7] Y. T. Tsai, CH. Tang, and Y. Y. Chen, "An Average Case Analysis of a Greedy Algorithm for the On-Line Steiner Tree Problem," *Computre Math. Applic.*, Vol. 31, No. 11, pp. 121-131, 1996.
- [8] B. M. Waxman, "Routing of multipoint connections," *IEEE Journal on Selected Areas in Communications*, Vol. 6, No. 9, pp. 1611-1622, 1998.
- [9] S. Ding, and N. Ishii, "An Online Genetic Algorithm for Dynamic Steiner Tree Problem," *Proc, Symposium on Computational Geometry*, pp. 337-343, 1995.
- [10] م. تشکری هاشمی، پ. ادیبی، ع. جهانیان، و ع. نوراله، "حل مسئله درخت اشتاینر پویا به کمک سیستم کولونی مورچه‌ها،" *مجموعه مقالات نهمین کنفرانس سالانه انجمن کامپیوتر ایران*، ۱۳۸۲.
- [11] R. S. Sutton, and A.G. Barto, *Reinforcement Learning: An introduction*, Cambridge, MA: MIT Press, 1998.

[12] M. A. L. Thathachar, and P. S. Sastry, "Varieties of Learning Automata: An Overview," *IEEE Trans. Systems, Man and Cybernetics – Part B: Cybernetics*, Vol. 32, No. 6, 2002.

[۱۳] م. ر. شیرازی، و م. ر. میبیدی، "به کارگیری اتوماتای یادگیر در سیستم‌های جندعامله همکار،" *مجموعه مقالات اولین کنفرانس بین المللی فناوری اطلاعات و دانش*، ص. ۳۳۸-۳۴۹، تهران، ۱۳۸۲.

[14] J. E. Beasley, "OR-Library: Distributing Test Problems by Electronic Mail," *Operational Research. SOC.*, Vol. 41, No. 11, pp. 1096-1072, 1990.

[15] B. M. Waxman, "Routing of multipoint connections," *IEEE J. Select. Areas Commun.* Vol. 6, No. 9, pp. 1617-1622, 1988.



سمیرا نوفرستی در حال حاضر عضو هیات علمی گروه فناوری اطلاعات، دانشکده مهندسی برق و کامپیوتر دانشگاه سیستان و بلوچستان است. وی مدرک کارشناسی خود را در رشته کامپیوتر گرایش نرم‌افزار از دانشگاه صنعتی شریف در سال ۱۳۸۲ و مدرک کارشناسی‌ارشد را در همین رشته از دانشگاه صنعتی امیرکبیر در سال ۱۳۸۴ اخذ کرده است. پس از فارغ‌التحصیلی به مدت یک سال عضو هیات علمی دانشگاه آزاد اسلامی واحد بیرجند بوده و از آذر ماه ۱۳۸۵ به عنوان هیات علمی دانشگاه سیستان و بلوچستان مشغول به کار شده است. هم‌اکنون نیز سمت معاونت کتابخانه مرکزی و مرکز اسناد دانشگاه را دارا است. زمینه‌های تحقیقاتی وی عبارتند از سیستم‌های یادگیر، مهندسی نرم‌افزار و هوش مصنوعی.

آدرس پست‌الکترونیکی ایشان عبارت است از:

snoferesti@ece.usb.ac.ir



محمدرضا میبیدی در سال ۱۳۳۱ در اراک بدنیا آمد. وی تحصیلات دانشگاهی خود را در دانشگاه شهید بهشتی در رشته اقتصاد آغاز و تا سطح کارشناسی‌ارشد در آنجا به تحصیل پرداخت و در سال ۱۳۵۶ موفق به اخذ مدرک کارشناسی‌ارشد شد. وی در سال ۱۳۶۲ موفق به اخذ دکتری رشته علوم کامپیوتر از دانشگاه اوکلاهما ی آمریکا شد و هم‌اکنون عضو هیات علمی و استاد دانشگاه صنعتی امیرکبیر می‌باشد. مشاغل و سمت‌های اداری و مدیریتی دکتر محمدرضا میبیدی به ترتیب زیر است: رئیس و عضو کمیته مهندسی کامپیوتر در وزارت علوم تحقیقات و فناوری از سال ۱۳۷۵ تا کنون، عضو کمیته داور برای چهاردهمین جشنواره خوارزمی در وزارت علوم تحقیقات و فناوری از سال ۱۳۷۵ تا کنون، عضو کمیته علمی در مرکز تحقیقات ریاضیات و فیزیک نظری وزارت علوم تحقیقات و فناوری از سال ۱۳۷۷ تا سال ۱۳۷۹، عضو کمیته چند رشته‌ای وزارت علوم تحقیقات و فناوری از سال ۱۳۷۵ تا کنون، عضو کمیته پژوهشی وزارت علوم تحقیقات و فناوری از سال ۱۳۸۲ تا سال ۱۳۸۳، عضو کمیته علمی در مرکز تحقیقات علوم پایه وزارت علوم تحقیقات و فناوری از سال ۱۳۸۱ تا کنون، عضو کمیته انتخاب بهترین کتاب ایرانی وزارت فرهنگ و ارشاد اسلامی از سال ۱۳۷۵ تا کنون.

فعالیت‌های آموزشی و عنوان دروس ارائه شده توسط محمدرضا میبیدی به قرار زیر است: نظریه زبان‌ها و ماشین‌ها، تحلیل و طراحی الگوریتم‌های موازی، ساختمان داده، تحلیل و طراحی الگوریتم، محاسبات موازی. آدرس پست‌الکترونیکی ایشان عبارت است از:

mmeybodi@aut.ac.ir

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۶/۰۷/۲۴

تاریخ اصلاح: ۸۹/۰۳/۰۹

تاریخ قبول شدن: ۸۹/۰۳/۲۳

نویسنده مرتبط: سمیرا نوفرستی، دانشکده مهندسی برق و کامپیوتر، دانشگاه سیستان و بلوچستان، زاهدان، ایران.

¹ Steiner Tree

² Minimum Spanning Tree

³ Prim

⁴ Linear Reward Penalty

⁵ Linear Reward Epsilon Penalty

⁶ Linear Reward Inaction

⁷ Network of Learning Automata

⁸ Network Simulator

استخراج ویژگی‌ها با استفاده از اطلاعات متقابل جهت طبقه‌بندی سیگنال‌های مغزی در سیستم‌های ارتباطی مغز با کامپیوتر

فرید اویسی ارنگه عباس عرفانیان امیدوار

دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران

چکیده

از عوامل مهم در دقت و کارایی طبقه‌بندی کننده‌ها، کاهش ابعاد فضای ویژگی است. کاهش هزینه محاسباتی و دقت طبقه‌بندی، دو دلیل عمده کاهش بعد فضای ویژگی است. به طور کلی دو روش برای کاهش ابعاد فضای ویژگی وجود دارد: انتخاب ویژگی و استخراج ویژگی. در روش انتخاب ویژگی‌ها، ویژگی‌هایی که در طبقه‌بندی مفید هستند، از یک مجموعه کامل ویژگی انتخاب می‌شوند. در روش استخراج ویژگی، ویژگی‌های جدیدی با استفاده از ویژگی‌های اولیه تولید می‌شود. در این تحقیق از یک الگوریتم استخراج ویژگی با استفاده از اطلاعات متقابل برای طبقه‌بندی سیگنال‌های مغزی در سیستم‌های ارتباطی مغز با کامپیوتر مورد استفاده قرار گرفته است. هدف در این سیستم، تشخیص تصور بستن پنجه دست و حالت استراحت ذهنی با استفاده از سیگنال‌های مغزی است. برای این منظور، آزمایشات مختلفی بر روی چهار فرد سالم در روزهای مختلف انجام شده است. نتایج نشان می‌دهد که عملکرد طبقه‌بندی کننده با استفاده از ویژگی‌های استخراج شده به وسیله اطلاعات متقابل نسبت به عملکرد طبقه‌بندی کننده با مجموعه کامل ویژگی بهبود قابل توجهی داشته است. علاوه بر این، عملکرد استخراج ویژگی مبتنی بر اطلاعات متقابل با روش استخراج ویژگی آنالیز مولفه‌های اصلی در طبقه‌بندی سیگنال‌های مغزی مقایسه شده، نتایج، برتری کامل استخراج ویژگی مبتنی بر اطلاعات متقابل را نسبت به روش آنالیز مولفه‌های اصلی نشان می‌دهد.

کلمات کلیدی: استخراج ویژگی، اطلاعات متقابل، طبقه‌بندی، ارتباط مغز با کامپیوتر، سیگنال‌های مغزی.

۱- مقدمه

پتانسیل‌های آهسته مغزی [۴]، مولفه‌های نوسانی EEG [۵]، [۶]، [۷]، و انواع مختلف پتانسیل‌های برانگیخته شامل پتانسیل آمادگی [۸]، پتانسیل برانگیخته بینایی [۹]، و P300 [۱۰] در سیستم‌های مختلف BCI بکار گرفته شده است. استخراج ویژگی برای طبقه‌بندی سیگنال‌های مغزی ناشی از فعالیت‌های ذهنی، نقش مهمی را در کارایی سیستم‌های BCI ایفا می‌کند. ویژگی‌های بی‌ربط یا اضافی، باعث پیچیدگی فرآیند طبقه‌بندی و وجود خطا در فرآیند پیش‌بینی می‌شود. حتی زمانی که ویژگی‌های ارائه شده دارای اطلاعات کافی درباره کلاس خروجی باشند، ممکن است به علت بزرگی ابعاد فضای ویژگی، خروجی به درستی پیش‌بینی نشود. هم چنین بعضی از آزمایشات نشان داده است که با اضافه کردن یک ویژگی بی‌ربط عملکرد سیستم‌های طبقه‌بندی کننده بدتر می‌شود [۱۱].

در سال‌های اخیر کوشش‌های زیادی جهت استفاده از سیگنال‌های مغزی به عنوان یک کانال جدید ارتباطی بین انسان و کامپیوتر تحت عنوان ارتباط مغز با کامپیوتر^۱ انجام گرفته و مقالات فراوانی در این رابطه انتشار یافته است [۱]، [۳]. بسیاری از این کوشش‌ها در جهت افزایش کارایی سیستم‌های BCI انجام گرفته است. کارایی چنین سیستم‌هایی به چند عامل مهم بستگی دارد. این عوامل شامل نوع سیگنال مغزی است که به عنوان ورودی به سیستم BCI استفاده می‌شود، روش‌های پردازش سیگنال جهت استخراج و یا انتخاب ویژگی، روش طبقه‌بندی ویژگی‌ها، فعالیت ذهنی، و آموزش شخص است. انواع مختلف سیگنال‌های مغزی جهت تشخیص فعالیت ذهنی بکار گرفته شده است. تاکنون از

بهینه تمیز دهنده است. این روش، بیشتر از $N_c - 1$ قادر به استخراج ویژگی نیست، در حالیکه N_c تعداد کلاس‌ها است.

اخیراً روش تحلیل مولفه‌های مستقل^۸ [۱۹]، که برای جداسازی کور منابع اندیشه شده است، به علت کاربرد گسترده آن، بسیار مورد توجه قرار گرفته است. ICA یک روش آماری است که داده‌های تصادفی مشاهده شده را به مولفه‌هایی تبدیل می‌کند که تا حد امکان از هم مستقل باشند. بل و سچونسکی [۱۹] یک الگوریتم یادگیری بدون معلم برای تحلیل مولفه‌های مستقل ارائه کرده است. این الگوریتم براساس بیشینه کردن آنترپپی در یک شبکه عصبی پیشرو تک لایه است. خروجی‌های ICA یک مجموعه از بردارهای مستقل است که ترکیب خطی از داده‌های مشاهده شده می‌باشد. این روش در جداسازی کور منابع و حذف انواع آرتیفک بطور گسترده استفاده شده است. اما ICA معمول برانزده استخراج ویژگی برای مسائل طبقه‌بندی نیست. زیرا که آن یک روش یادگیری بدون معلم است و اطلاعات کلاس‌ها را در تخمین ماتریس ترکیب بکار نمی‌برد.

محققین سعی کرده‌اند با استفاده از تحلیل مولفه‌های مستقل روش‌هایی برای استخراج ویژگی ارائه دهند که ویژگی‌ها علاوه بر اینکه از یکدیگر مستقل باشند، دارای حداکثر اطلاعات درباره کلاس باشند. در این راستا، در مرجع [۱۱] روشی برای استخراج ویژگی بر مبنای تحلیل مولفه‌های مستقل ارائه شده است. در این روش، متغیرهای ورودی با یک تبدیل خطی به دو دسته ویژگی تبدیل می‌شوند. یک دسته دارای حداکثر اطلاعات در مورد کلاس خروجی و دسته دیگر تا حد امکان نسبت به کلاس مستقل هستند.

این روش محدود به دو کلاس است. از این روش برای استخراج ویژگی از سیگنال‌های مغزی در حین تصور حرکت دست در سیستم‌های ارتباطی مغز با کامپیوتر استفاده شده است [۲۰]. هدف اصلی در استخراج ویژگی، تولید ویژگی‌هایی است که حداقل اطلاعات متقابل را بین هم داشته باشند و در عین حال حداکثر اطلاعات متقابل با کلاس خروجی. به این منظور در مرجع [۲۱]، ابتدا با استفاده از روش ICA ویژگی‌های مستقل ایجاد شده، سپس ویژگی‌های با حداکثر اطلاعات متقابل با کلاس خروجی انتخاب شده است. اما باید توجه داشت که کمینه کردن اطلاعات مشترک بین ویژگی‌ها با استفاده از ICA بطور کامل انجام نمی‌شود [۲۲].

روش‌هایی که از آمارگان درجه دوم استفاده می‌کنند، رابطه خطی بین متغیرهای تصادفی را مقایسه می‌کنند، در حالیکه روش‌های مبتنی بر نظریه اطلاعات رابطه غیرخطی بین متغیرهای تصادفی را در نظر می‌گیرند، به عنوان مثال بین یک بردار ویژگی و کلاس. یک شاخص برای روش‌های مبتنی بر نظریه اطلاعات، استفاده از اطلاعات متقابل است. اطلاعات متقابل یک روش غیرپارامتریک برای اندازه‌گیری میزان وابستگی بین دو متغیر تصادفی است. چنانچه دو متغیر تصادفی مستقل باشند، مقدار اطلاعات متقابل صفر خواهد بود. هدف از استخراج ویژگی با استفاده از اطلاعات متقابل، تولید ویژگی‌هایی است که بطور مشترک دارای حداکثر وابستگی را به کلاس خروجی داشته باشند. علی‌الرغم مزیت‌های اطلاعات متقابل، تخمین اطلاعات متقابل با بعد بالا بطور دقیق غیر ممکن است.

تخمین اطلاعات متقابل بین دو متغیر از طریق هیستوگرام قابل انجام است. اما به علت پراکندگی توزیع داده‌ها در فضای با بعد بالا، تخمین زنده‌های مبتنی بر هیستوگرام قابل اعتماد نیستند [۲۳]. به منظور رفع این مشکل، از تخمین آنترپپی معرفی شده توسط رینی همراه با تخمین پنجره پارزن برای تخمین اطلاعات متقابل استفاده شده است [۱۲]، [۲۴].

ما در این مقاله از این روش تخمین اطلاعات متقابل به منظور استخراج ویژگی از سیگنال‌های مغزی برای سیستم‌های ارتباطی مغز با کامپیوتر استفاده کرده‌ایم. اطلاعات متقابل بین کلاس‌های خروجی و ویژگی‌های استخراج شده به

به طور کلی دو روش برای کاهش بعد فضای ویژگی وجود دارد. روش اول که انتخاب ویژگی^۲ نامیده می‌شود، از میان بردار کامل ویژگی‌ها، تعدادی که در طبقه‌بندی کلاس‌ها مفید هستند انتخاب می‌شوند. در روش دوم که استخراج ویژگی^۳ نامیده می‌شود ویژگی‌های جدیدی با استفاده از مجموعه ویژگی‌های اصلی ساخته می‌شوند. روش‌های کاهش بعد را می‌توان از جنبه‌ای دیگر به دو دسته فیلتر و پوشه‌ای^۴ تقسیم کرد. در روش‌های فیلتر، انتخاب یا استخراج ویژگی با ارزیابی یک شاخص انجام می‌شود [۱۲]. اما در روش‌های پوشه‌ای [۱۳]، کاهش بعد ویژگی با طبقه‌بندی کننده ترکیب می‌شود. در این روش ویژگی‌هایی انتخاب یا استخراج می‌شوند که خطای طبقه‌بندی کننده را حداقل سازند. روش‌های پوشه‌ای دو مشکل عمده دارند. اول اینکه این روش‌ها دارای بار محاسباتی زیادی هستند. دوم اینکه ویژگی‌های بدست آمده تنها برای طبقه‌بندی کننده خاص استفاده شده دارای حداقل خطا است [۱۴].

از روش‌های انتخاب و استخراج ویژگی در سیستم‌های ارتباطی مغز با کامپیوتر استفاده است. در مرجع [۱۵]، مولفه‌های فرکانسی به عنوان بردار ویژگی برای طبقه‌بندی سیگنال‌های مغزی در هنگام تصور حرکت دست راست و چپ استفاده شده است. در این کار از الگوریتم DSLVQ^۵ برای تعیین میزان وابستگی هر مولفه فرکانسی نسبت به کلاس استفاده شده است، سپس مولفه‌ای که بیشترین وابستگی را به کلاس داشته باشد، بطور دستی انتخاب شده است. نشان داده شده است که به طور کلی دقت طبقه‌بندی با ویژگی‌های انتخاب شده، بهتر از دقت طبقه‌بندی با تمام ویژگی‌ها است.

در مرجع [۸]، با استفاده از نظریه اطلاعات، جهت انتخاب ویژگی از یک بانک ویژگی شامل ۱۰۰۰ ویژگی استفاده شده است. اما این الگوریتم نتیجه قابل قبولی نداده است. از روش‌های پوشه‌ای نیز برای انتخاب ویژگی در سیستم‌های ارتباطی مغز با کامپیوتر استفاده شده است [۱۶]، [۸].

روش پوشه‌ای استفاده شده در این مقالات شامل SVM به عنوان طبقه‌بندی کننده و الگوریتم ژنتیک جهت جستجوی زیرمجموعه ویژگی، است. در مرجع [۱۶]، از متوسط توان در باندهای فرکانسی دلتا (۴-۲)، تتا (۸-۴)، آلفا در فرکانس‌های پایین (۱۲-۱۰)، آلفا در فرکانس‌های بالا (۱۲-۱۰)، بتا (۳۵-۱۲)، گاما (۵۰-۳۵) هرگز مربوط به ۳۰ کانال مغزی به عنوان مجموعه کامل ویژگی استفاده شده است. سپس با استفاده از الگوریتم ژنتیک تعدادی ویژگی انتخاب شده است.

هدف ما در این مقاله، استخراج ویژگی‌های مناسب از سیگنال‌های الکتروانسفالوگرام با استفاده از اطلاعات متقابل برای سیستم‌های ارتباطی مغز با کامپیوتر است. فعالیت ذهنی در نظر گرفته شامل حالت تصور بستن دست و حالت استراحت است. هدف، تشخیص این دو حالت ذهنی با استفاده از سیگنال‌های مغزی است.

۲- استخراج ویژگی

کاهش بعد فضای متغیرهای ورودی، یک مرحله مهم در فرآیند شناسایی و طبقه‌بندی الگوها است. یکی از روش‌های معمول کاهش بعد فضای داده‌ها و استخراج ویژگی، روش تحلیل مولفه‌های اصلی^۶ [۱۷] است. این روش به دنبال یک بازنمایی از داده‌ها بر حسب کوچکترین متوسط مربعات خطا بین داده‌های بازنمایی شده و اصلی است. لذا، PCA کاری به استخراج ویژگی‌های بهینه به منظور طبقه‌بندی ندارد. به عبارتی، این روش به منظور استخراج ویژگی و کاهش بعد، تنها به داده‌های ورودی توجه می‌کند و کلاس داده‌ها را در نظر نمی‌گیرد.

از روش‌های معمول و رایج استخراج ویژگی، تحلیل خطی تفکیک کننده^۷ [۱۸] است. این روش یک تبدیل خطی است که برای حالت‌های خاصی به طور

۴- استخراج ویژگی بر مبنای اطلاعات متقابل

فرض کنید مجموعه ای از داده‌های آموزشی $\{x_i, c_i\}$ به عنوان نمونه‌هایی از یک متغیر تصادفی پیوسته $X(x_i \in R^D)$ ، و برچسب کلاس‌ها به عنوان نمونه‌های یک متغیر تصادفی گسسته $C(c_i \in \{1, 2, \dots, N_c\}, i \in \{1, 2, \dots, N\})$ در اختیار باشد. هدف از استخراج ویژگی، پیدا کردن یک تبدیل g ، به طوری که $f_i = g(w, x_i)$ ، $f_i \in R^d$ ، $d < D$ ، تبدیل شده F و برچسب کلاس‌ها C ، $I(F, C)$ بیشینه شود. به منظور تحقق این، لازم است که $I(F, C)$ به صورت تابعی از مجموعه داده‌ها و شکل قابل مشتق پذیر تخمین زده شود. سپس از این می‌توان با مشتق گرفتن از $I(F, C)$ نسبت به پارامترهای تبدیل و استفاده از الگوریتم کمترین شیب^{۱۱}، قاعده یادگیری استخراج ویژگی را بدست آورد:

$$w_{t+1} = w_t + \eta \frac{\partial I}{\partial w} = w_t + \eta \sum_{i=1}^N \frac{\partial I}{\partial f_i} \frac{\partial f_i}{\partial w} \quad (5)$$

این روال استخراج ویژگی در شکل ۱ نشان داده شده است.

۴-۱- تخمین اطلاعات متقابل

برای محاسبه اطلاعات متقابل احتیاج به تخمین تابع چگالی احتمال^{۱۱} وجود دارد. در بسیاری از مقالات از روش هیستوگرام برای تخمین تابع چگالی احتمال استفاده می‌شود که در آن عمل انتگرال برای سادگی به عمل جمع تبدیل شده است، اما دقت اکثر تخمین زنده‌های هیستوگرام به علت کمبود داده و پراکندگی توزیع داده در بعد بالا، ضعیف است. بنابراین تخمین زنده‌های هیستوگرام تنها در داده‌های با بعد پایین استفاده می‌شوند [۲۶]. در این مقاله برای محاسبه اطلاعات متقابل از تخمین انتروپی معرفی شده توسط رینی استفاده شده است [۱۲]. با جایگزینی انتروپی رینی در رابطه (۳)، تخمین اطلاعات متقابل به صورت ذیل حاصل می‌شود:

$$I(F; C) \cong H_2(C) - H_2(C|F) \quad (6)$$

در این رابطه، $H_2(X)$ انتروپی Quadratic رینی است:

$$H_2(X) = -\log \frac{1}{N} \sum_{n=1}^N G(x(n) - x(n-1), 2\sigma^2) \quad (7)$$

همچنین می‌توان با استفاده از رابطه (۱۲) رابطه (۱۵) را بیان کرد.

$$H(X|C) = -\sum_{j=1}^{N_c} \left(\frac{N_j}{N_T} \log \frac{1}{N_j} \sum_{n=1}^{N_j} G(x_j(n) - x_j(n-1), 2\sigma^2) \right) \quad (8)$$

در این روابط، N_j تعداد الگوهای آموزشی مربوط به کلاس j ، N_T تعداد الگوهای آموزشی و N_c تعداد کلاس‌ها است. با قرار دادن رابطه (۶) و (۷) در رابطه (۵) برای $I(X, C)$ رابطه زیر برحسب متغیر X حاصل می‌شود.

عنوان یک شاخص جهت گسترش قانون یادگیری برای استخراج کننده ویژگی بکار گرفته شده است. استخراج کننده ویژگی می‌تواند یک تبدیل خطی باشد که خروجی آن ویژگی‌های جدید باشند. قبل از ارائه روش استخراج ویژگی، در بخش بعد، مفاهیم پایه نظریه اطلاعات بطور مختصر بیان می‌شود.

۳- مفاهیم پایه در نظریه اطلاعات

در فرآیند طبقه‌بندی، ویژگی‌هایی که اطلاعاتی در مورد کلاس داشته باشند می‌توانند به تفکیک کلاس‌ها کمک کنند، اما ویژگی‌هایی که اطلاعاتی در مورد کلاس ندارند، تنها سبب افزایش بعد فضای ویژگی شده و باعث کاهش کارایی طبقه‌بندی کننده می‌شود. برای اندازه‌گیری اطلاعات موجود در یک ویژگی مربوط به یک کلاس نیاز به ابزار جدیدی است. نظریه اطلاعات شانون^{۱۲}، یک روش برای اندازه‌گیری اطلاعات متغیرهای تصادفی با استفاده از انتروپی و اطلاعات متقابل ارائه می‌کند [۲۵]. روش‌هایی که از آمارگان درجه دوم استفاده می‌کنند، رابطه خطی بین متغیرهای تصادفی را مقایسه می‌کنند، در حالیکه روش‌های مبتنی بر نظریه اطلاعات رابطه غیرخطی بین متغیرهای تصادفی را در نظر می‌گیرند، به عنوان مثال بین یک بردار ویژگی و کلاس.

چنانچه احتمال کلاس‌های مختلف $p(c); c = 1, \dots, N_c$ باشد، میزان عدم قطعیت کلاس خروجی با انتروپی اندازه‌گیری می‌شود:

$$H(C) = -\sum_{c=1}^{N_c} p(c) \log p(c) \quad (1)$$

عدم قطعیت C با معلوم بودن بردار ویژگی f با استفاده از انتروپی شرطی به صورت ذیل اندازه‌گیری می‌شود.

$$H(C|F) = -\int p(f) \left(\sum_{c \in C} p(c|f) \log(p(c|f)) \right) df \quad (2)$$

در حالیکه $p(c|f)$ احتمال شرطی متغیر C با فرض معلوم بودن f است. بطور کلی انتروپی شرطی کمتر و یا مساوی انتروپی اولیه است، هنگامی برابر است که دو متغیر C و f مستقل باشد. میزانی که عدم قطعیت کلاس کاهش می‌یابد، اطلاعات متقابل نامیده می‌شود و برابر است با

$$I(F; C) = H(C) - H(C|F) \quad (3)$$

اطلاعات متقابل را می‌توان بعد از اعمال مقادیر $p(c, f) = p(c|f)p(f)$ و $p(c) = \int p(c, f) df$ به صورت ذیل بیان کرد:

$$I(F; C) = \sum_{c \in C} \int p(c, f) \log \frac{p(c, f)}{p(c)p(f)} df \quad (4)$$

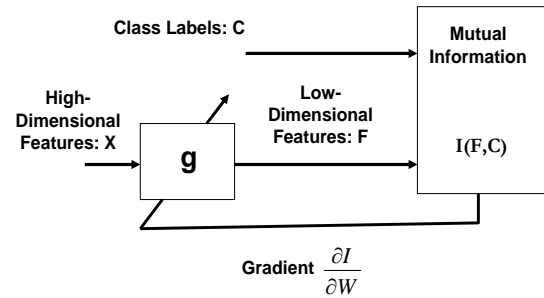
اگر اطلاعات متقابل بین دو متغیر بزرگ باشد، نشان دهنده وابسته بودن آن دو متغیر به یکدیگر خواهد بود. اطلاعات متقابل فقط در صورت مستقل بودن دو متغیر تصادفی صفر خواهد بود.

در این تحقیق، قدر مطلق دامنه (MAV)، واریانس، توان در هر یک از فرکانس‌های ۹-۲۸ Hz، ضرایب مدل AR از مرتبه ۵، تعداد عبور از صفر و تعداد اکسترمم‌های محلی به عنوان ویژگی برای طبقه‌بندی سیگنال‌های مغزی در حین تصور حرکت در نظر گرفته شده است. بعد از محاسبه ویژگی‌ها، ویژگی‌های مناسب با استفاده از روش ارائه شده استخراج شده و سپس از آنها برای آموزش و آزمون طبقه‌بندی کننده استفاده شده است. در این مقاله از طبقه بندی کننده معروف LDA قطری نرم افزار MATLAB [۲۹] برای طبقه‌بندی استفاده شده است. در هر روز آزمایشی، برای هر یک از دو حالت آرامش و تصور حرکت ۱۰۰ الگو و در مجموع ۲۰۰ الگو وجود دارد. هر ویژگی برای ۲۰۰ آزمون نرمالیزه شده است. الگوهای مربوط به حالت باز کردن و بستن پنجه دست در یک کلاس و الگوهای مربوط به حالت آرامش در کلاس دیگر قرار داده شده است. هر بار ۵۰ الگو بطور تصادفی از بین الگوهای آرامش و ۵۰ الگو بطور تصادفی از میان الگوهای اتفاق انتخاب می‌شود.

این الگوها برای آموزش به طبقه‌بندی کننده داده می‌شود و از سایر الگوها برای ارزیابی کارایی طبقه‌بندی کننده استفاده می‌شود. این عمل ۱۰ بار انجام می‌شود و نتایج آن میانگین‌گیری می‌شود. این کار بر روی ۷ کانال به صورتی انجام می‌شود که ترتیب الگوها برای آموزش و آزمون طبقه‌بندی کننده‌های ۷ کانال در هر بار یکسان باشد. سپس بر روی نتایج کانال‌ها رای‌گیری صورت گرفته است. رای‌گیری به این صورت انجام می‌شود که خروجی طبقه‌بندی کننده‌ها با هم جمع می‌شود، اگر جواب نهایی مثبت باشد به معنای تصور حرکت و اگر منفی باشد به معنای وقوع آرامش است. در آموزش طبقه‌بندی کننده مقدار +۱ برای حالت بستن و یا باز کردن دست و -۱ برای حالت آرامش در نظر گرفته شده است (شکل ۲).

$$I(X, C) = -\log \frac{1}{N_T} \sum_{n=1}^{N_T} G(x(n) - x(n-1), 2\sigma^2) + \sum_{j=1}^{N_C} \left(\frac{N_j}{N_T} \log \frac{1}{N_j} \sum_{n=1}^{N_j} G(x_j(n) - x_j(n-1), 2\sigma^2) \right) \quad (9)$$

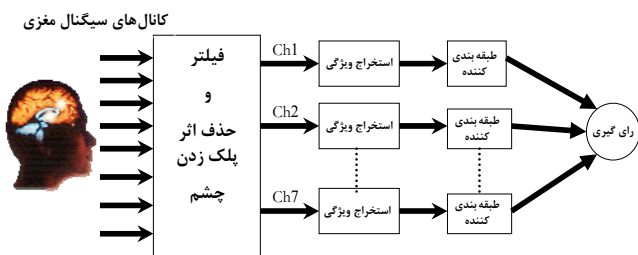
با استفاده از قانون یادگیری تصادفی کمترین شیب (۵) و اطلاعات متقابل بر حسب آنتروپی رینی، می‌توان تابع تبدیل بین متغیرهای ورودی و ویژگی‌های جدید را تخمین زد. این الگوریتم بطور اختصار MRMI-SIG¹⁷ [۲۴] نامیده شده است.



شکل ۱- استخراج ویژگی با استفاده از اطلاعات متقابل

۵- روال آزمایش

برای ارزیابی روش، از داده‌های اخذ شده در آزمایشگاه پردازش سیگنال‌های بیولوژیک گروه مهندسی پزشکی دانشگاه علم و صنعت ایران استفاده شده است [۲۷]. برای اخذ سیگنال‌های مغزی آزمایشات مختلفی بر روی انسان انجام شده است. شخص بر روی صندلی در مقابل صفحه نمایش کامپیوتر می‌نشیند و با ظاهر شدن یک نشانه بر روی صفحه نمایش کامپیوتر، شخص تصور حرکت را انجام می‌دهد. با توجه به نشانه، شخص تصور حرکت بستن و یا باز کردن پنجه دست را انجام می‌شود. چنانچه نشانه‌ای بر روی صفحه کامپیوتر ظاهر نشود، شخص عمل خاصی را انجام نمی‌دهد. سیگنال‌های EEG با نصب الکترودهای از جنس Ag/AgCl در موقعیت‌های Fp1, T5, Cz, Pz, Fz, F4, F3, F4, F3 مطابق استاندارد ۱۰-۲۰ از سر افراد سالم ثبت شده است. همه کانال‌ها نسبت به الکترود مرجعی در گوش راست با فرکانس نمونه‌برداری ۲۵۶ هرتز ثبت شده‌اند. مدت زمان هر آزمون برابر با پنج ثانیه بوده و برای هر حالت ۵۰ آزمون انجام شده است. لازم به ذکر است که آزمایشات در حالت چشم باز انجام شده است. سیگنال‌های اخذ شده از یک فیلتر پایین گذر ۴۵ هرتز عبور کرده‌اند و متوسط داده‌های هر آزمون از تمام داده‌های آن کم شده است. از یک الکترود که بر روی پیشانی سمت چپ قرار داده شده، برای اخذ سیگنال‌های چشمی استفاده شده است. یکی از مشکلات مهم تحلیل سیگنال‌های EEG، وجود آرتیفکت پلک زدن است. در این تحقیق از روش تحلیل مولفه‌های مستقل و با استفاده از الگوریتم Infomax آرتیفکت پلک زدن حذف شده است [۲۸]. لازم به ذکر است که از هر فرد در ۴ روز متفاوت آزمایش به عمل آمده است.



شکل ۲- مراحل طبقه‌بندی تصور حرکت پنجه دست

در این تحقیق، همچنین برای استخراج ویژگی از روش PCA استفاده شده است و کارائی آن با روش پیشنهادی استخراج ویژگی و استفاده از مجموعه کامل ویژگی مقایسه شده است. در شکل ۳ متوسط طبقه‌بندی تصور حرکت پنجه دست ناشی از ۴ روز آزمایش برای ۴ فرد (AA, ME, MA, AE) نشان داده شده است. در این شکل‌ها، درصد طبقه‌بندی به ازاء انتخاب ۳، ۵، ۱۰ و ۲۰ ویژگی با استفاده از الگوریتم MRMI-SIG و PCA آورده شده است.

درصد طبقه‌بندی به ازاء ۲۹ ویژگی (مجموعه کامل ویژگی‌ها)، بدون استفاده از الگوریتم استخراج ویژگی نوشته شده است. همانطور که از شکل ۳ مشخص است، برای فرد AA با استفاده از الگوریتم MRMI-SIG به ازاء ۳ ویژگی بهترین دقت طبقه‌بندی، ۷۱/۲۵٪ بدست آمده است که حدود ۱۱٪ نسبت به نتیجه طبقه‌بندی با استفاده از مجموعه تمام ویژگی‌ها بهبود یافته است. اما این بهبودی به ازاء ۱۰ و ۲۰ ویژگی استخراج شده با استفاده از الگوریتم MRMI-SIG حدود ۳/۵ درصد است. این در حالی است که نتایج الگوریتم PCA همواره کمتر از نتایج طبقه‌بندی با استفاده از مجموعه کامل ویژگی‌ها است. در مورد فرد AE نیز متوسط نتیجه طبقه‌بندی با استفاده از ۳ ویژگی استخراج شده توسط الگوریتم

۶- طبقه‌بندی سیگنال‌های مغزی در هنگام تصور

برای طبقه‌بندی سیگنال‌های مغزی در سه حالت آرامش، باز و بستن پنجه دست، ویژگی‌های مربوط به داده‌های هر آزمون و برای هر کانال مغزی محاسبه شده است.

MRMI-SIG حدود ۵ درصد نسبت به نتیجه طبقه‌بندی با استفاده از تمام ویژگی‌ها بهتر است.

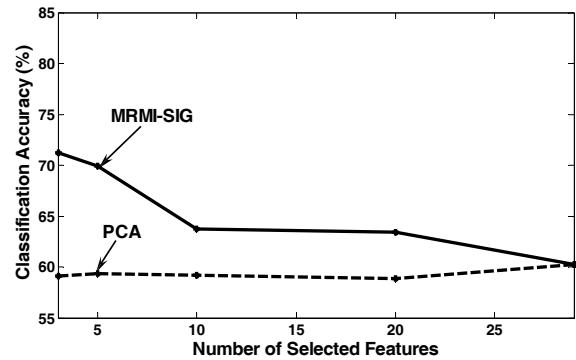
این در حالی است که دقت طبقه‌بندی با استفاده از همین الگوریتم به ازاء ۱۰ و ۲۰ ویژگی حدود ۲ درصد نسبت به استفاده از تمام ویژگی‌ها کمتر شده است. در مورد این فرد بهترین نتیجه حاصل از PCA ۳ درصد کمتر از دقت طبقه‌بندی با استفاده از مجموعه کامل ویژگی‌ها بوده است. در فرد ME نیز دقت طبقه‌بندی حدود ۳ درصد با استفاده از ۳ ویژگی استخراج شده توسط الگوریتم MRMI-SIG نسبت به استفاده از تمام ویژگی‌ها افزایش و حدود ۱ درصد به ازاء ۱۰ و ۲۰ ویژگی کاهش یافته است. اما توسط الگوریتم PCA، بهترین نتیجه با ۲۰ ویژگی بدست آمده که ۲ درصد از نتیجه بدست آمده با مجموعه کامل ویژگی‌ها کمتر بوده است. در مورد فرد MA دقت طبقه‌بندی با استفاده از ۵ ویژگی حدود ۳/۵ افزایش یافته است.

در شکل ۴ متوسط نتایج طبقه‌بندی برای ۴ نفر با استفاده از ۳، ۵، ۱۰ و ۲۰ ویژگی استخراج شده با استفاده از الگوریتم MRMI-SIG و PCA و تمام ویژگی‌ها (۲۹ ویژگی) آورده شده است. همانطور که ملاحظه می‌شود، بهترین نتیجه با استفاده از ۳ ویژگی با استفاده از الگوریتم MRMI-SIG حاصل شده است. در این حالت دقت طبقه‌بندی حدود ۵/۵ درصد نسبت به حالت مجموعه کامل ویژگی‌ها افزایش نشان می‌دهد. در مورد الگوریتم PCA حداکثر دقت طبقه‌بندی با استفاده از ۲۰ ویژگی رخ داده است. دقت در این حالت حدود ۱ درصد نسبت به حالت استفاده از تمام ویژگی‌ها کمتر است.

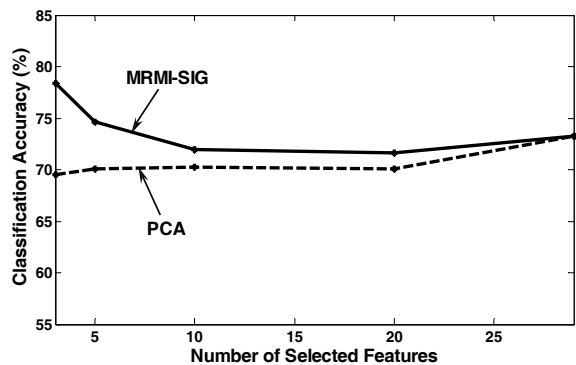
در شکل ۵، دقت طبقه‌بندی به ازاء ۳، ۵، ۱۰ و ۲۰ ویژگی استخراج شده با استفاده از الگوریتم MRMI-SIG و PCA و همچنین دقت طبقه‌بندی با مجموعه کامل ویژگی‌ها برای روزهای مختلف فرد AA نشان داده شده است. ملاحظه می‌شود در روز اول این فرد، حدود ۹ درصد و در روز دوم حدود ۱۴ درصد، دقت طبقه‌بندی با استفاده از الگوریتم MRMI-SIG و انتخاب سه ویژگی نسبت به تمام ویژگی‌ها بیشتر است. این افزایش برای روز سوم حدود ۱۵ درصد و برای روز چهارم حدود ۱۱ درصد است. اما با الگوریتم PCA برای تمام روزها بجز روز سوم، دقت طبقه‌بندی کمتر از دقت طبقه‌بندی با استفاده از تمام ویژگی‌ها بوده است.

در روز سوم با استفاده از ۵ ویژگی استخراج شده توسط PCA حدود ۱ درصد نسبت به حالت تمام ویژگی‌ها بهبود مشاهده می‌شود. میانگین دقت طبقه‌بندی برای افراد مختلف و بازای تعداد ویژگی‌های مختلف با استفاده از الگوریتم PCA و تمام ویژگی‌ها (۲۹ ویژگی) در جدول ۱ آورده شده است.

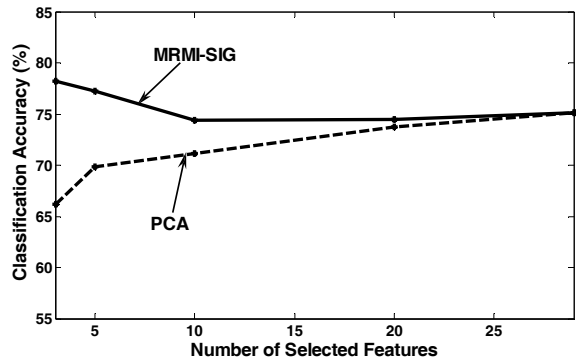
نتایج نشان می‌دهد که دقت طبقه‌بندی با استفاده از الگوریتم MRMI-SIG، بیشتر از دقت طبقه‌بندی با استفاده از الگوریتم PCA و استفاده از مجموعه تمام ویژگی‌ها بوده است. به طور متوسط، بهترین نتیجه طبقه‌بندی با استفاده از الگوریتم MRMI-SIG با سه ویژگی و حدود ۷۴٪ حاصل شده، در حالیکه با استفاده از الگوریتم PCA، بهترین نتیجه با ۲۰ ویژگی و حدود ۶۷٪ بدست آمده است. میانگین دقت طبقه‌بندی با تمام ویژگی‌ها حدود ۶۸٪ است.



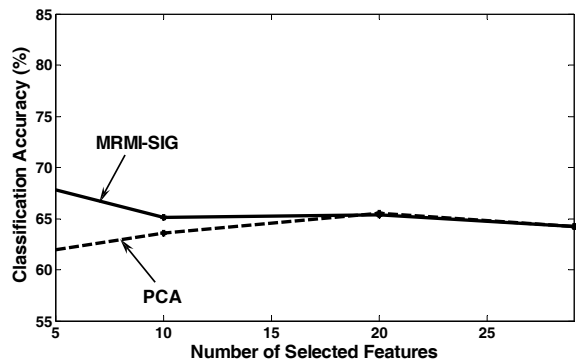
(الف)



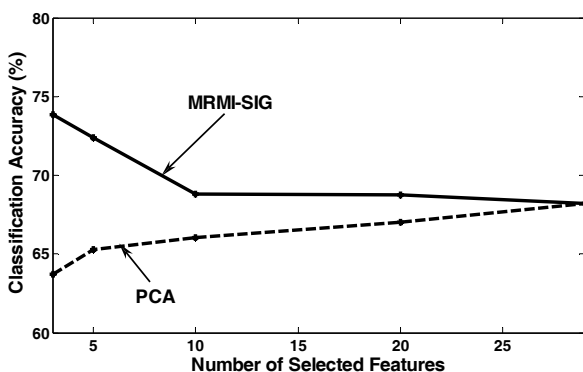
(ب)



(ج)



(د)



شکل ۴- متوسط طبقه‌بندی تصور حرکت پنجه دست برای تمام افراد

شکل ۳- متوسط طبقه‌بندی تصور حرکت پنجه دست برای فرد الف) AA (ب) AE (ج) ME (د) MA

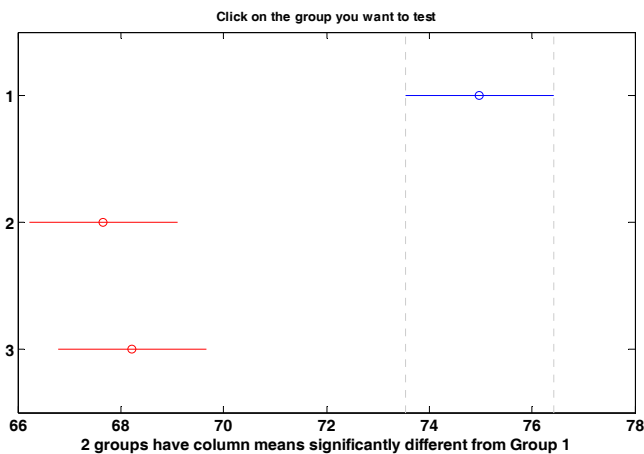
۷- نتیجه‌گیری

یکی از مهمترین پیش پردازش‌ها به منظور بهبود عملکرد سیستم‌های طبقه‌بندی، کاهش بعد فضای ویژگی است. کاهش بعد فضای ویژگی باعث کاهش پیچیدگی فرآیند طبقه‌بندی و در نتیجه کاهش وقوع خطا می‌شود. یکی از روش‌هایی که برای کاهش بعد فضای ویژگی معرفی شده است، استخراج ویژگی نامیده می‌شود. در این روش‌ها ویژگی‌های جدیدی با استفاده از مجموعه ویژگی‌های اصلی ساخته می‌شود. تاکنون روش‌های زیادی برای استخراج ویژگی‌های مناسب معرفی شده است. یکی از روش‌های استخراج ویژگی استفاده از تحلیل مولفه‌های اصلی است.

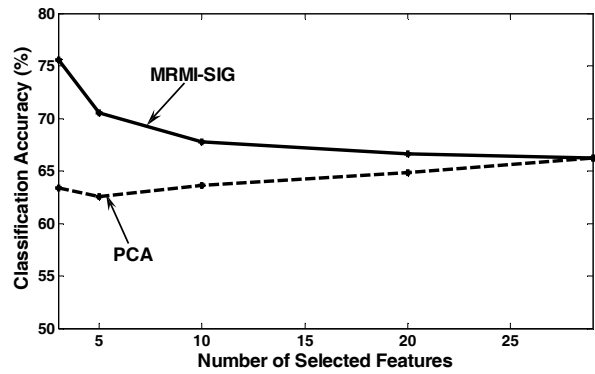
اشکال اصلی این روش این است که برای استخراج ویژگی، تنها به داده‌های ورودی توجه می‌کند و کلاس خروجی سیستم طبقه‌بندی کننده را مورد ملاحظه قرار نمی‌دهد. یکی دیگر از روش‌هایی که به طور گسترده به منظور استخراج ویژگی به کار رفته، روش تحلیل مولفه‌های مستقل است. در این روش سعی می‌شود، ویژگی‌هایی استخراج شود که تا حد امکان از هم مستقل باشند. اما در این روش نیز مشکل عدم توجه به کلاس خروجی وجود دارد. اخیراً استخراج ویژگی با استفاده از نظریه اطلاعات مورد توجه زیاد قرار گرفته است. در این روش‌ها سعی می‌شود ویژگی‌هایی استخراج شوند که دارای حداکثر اطلاعات را در مورد کلاس خروجی داشته باشند.

در این مقاله، از یک روش استخراج ویژگی مبتنی بر اطلاعات متقابل برای طبقه‌بندی سیگنال‌های مغزی در حین تصور حرکت برای سیستم‌های ارتباطی مغز با کامپیوتر مورد استفاده قرار گرفته شد و کارایی آن با روش استخراج ویژگی PCA و همچنین با مجموعه کامل ویژگی‌ها مقایسه گردید. نتایج نشان می‌دهد که استخراج ویژگی‌های با استفاده از اطلاعات متقابل باعث افزایش دقت طبقه‌بندی شده است.

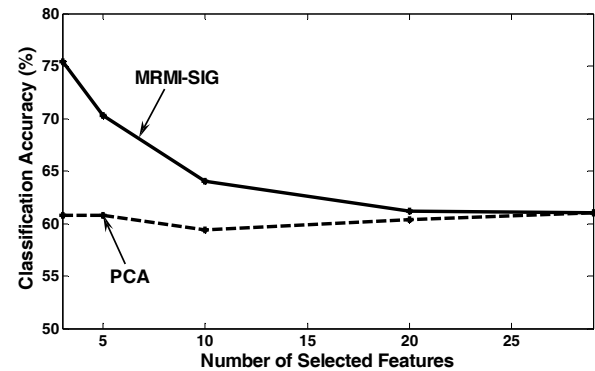
در شکل ۶ نتیجه مقایسه دقت طبقه‌بندی سیگنال‌های مغزی با استفاده از دو روش استخراج ویژگی، MRMI-SIG، PCA و با استفاده از مجموعه کامل ویژگی با استفاده از آنالیز پراش^{۱۳} نشان داده شده است. نتایج نشان می‌دهد که روش استخراج ویژگی MRMI-SIG بطور قابل ملاحظه‌ای باعث بهبود طبقه‌بندی کننده شده است ($p=5 \times 10^{-4}$).



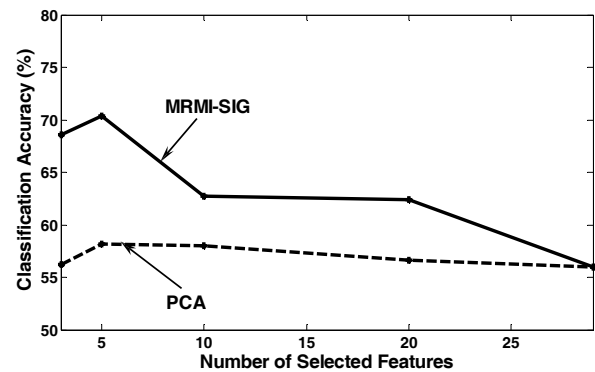
شکل ۶- نتیجه آزمون آماری بین نتایج طبقه‌بندی با استفاده از تمام ویژگی‌ها و ویژگی‌های استخراج شده با استفاده از الگوریتم MRMI-SIG و PCA



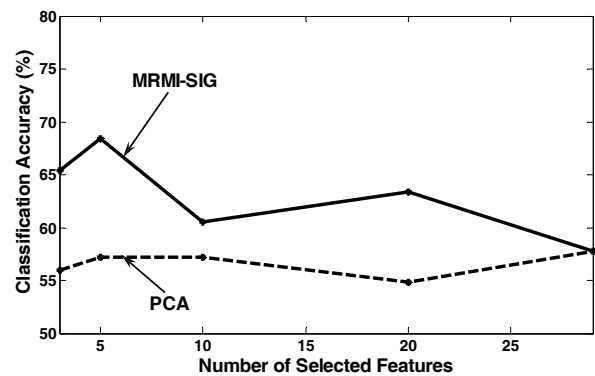
(الف)



(ب)



(ج)



(د)

شکل ۵- طبقه‌بندی تصور حرکت پنجه دست برای فرد AA (الف) روز اول (ب) روز دوم (ج) روز سوم (د) روز چهارم

جدول ۱- دقت طبقه‌بندی با تعداد ویژگی استخراج شده مختلف با استفاده از الگوریتم MRMI-SIG و PCA برای افراد مختلف

تعداد ویژگی	۳ ویژگی		۵ ویژگی		۱۰ ویژگی		۲۰ ویژگی		۲۹ ویژگی
	MRMI-SIG	PCA	MRMI-SIG	PCA	MRMI-SIG	PCA	MRMI-SIG	PCA	Full Set
فرد AA	۷۱/۲۵	۵۹/۱	۶۹/۹	۵۹/۳۵	۶۳/۷۵	۵۹/۱۵	۶۳/۴	۵۸/۹	۶۰/۲۵
فرد AE	۷۸/۳۵	۶۹/۵۵	۷۴/۶	۷۰/۰۵	۷۱/۹۵	۷۰/۲۵	۷۱/۶۵	۷۰/۰۵	۷۳/۲۵
فرد ME	۷۸/۲	۶۶/۲	۷۷/۲	۶۹/۸	۷۴/۴	۷۱/۱	۷۴/۵	۷۳/۷	۷۵/۱۵
فرد MA	۶۷/۷	۶۰/۰۵	۶۷/۸	۶۱/۹۵	۶۵/۱۵	۶۳/۶	۶۵/۴	۶۵/۵	۶۴/۲۵
میانگین	۷۳/۸۷	۶۳/۷۲	۷۲/۳۷	۶۵/۲۸	۶۸/۸۱	۶۶/۰۲	۶۸/۷۳	۶۷/۰۳	۶۸/۲۲

مراجع

- [11] N. Kwak, and C.-H. Choi, "Feature extraction based on ICA for binary classification problems," *IEEE Trans. Knowledge and Data Eng.*, Vol. 15, No. 6, pp. 1374,1388, 2003.
- [12] K. Torkkola, "Feature Extraction by non-parametric Mutual Information Maximization," *Journal of Machine Learning Research*, Vol. 3, pp. 1415–1438, 2003.
- [13] R. Kohavi, and G. John, "Wrapper for Feature Subset Selection," *Artificial Intelligence*, Vol. 97, Nos. 1-2, pp. 273-324, 1997.
- [14] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Analysis*, Vol. 27, No. 8, pp. 1226-1238, 2005.
- [15] M. Pregenzer, and G. Pfurtscheller, "Frequency Component Selection for an EEG-Based Brain to Computer Interface," *IEEE Trans. Rehab. Eng.*, Vol. 7, No. 4, pp. 413-419, 1999.
- [16] D. Peterson, J. Knight, M. Kirby, C. Anderson, and M. Thaut, "Feature Selection and Blind Source Separation in an EEG-Based Brain-Computer Interface," *EURASIP Journal on Applied Signal Processing*, pp. 3128-3140, 2005.
- [17] H. Li, T. Jiang, and K. Zhang, "Efficient and Robust Feature Extraction by Maximum Margin Criterion," *IEEE Trans. Neural Networks*, Vol. 17, No. 1, pp. 157–1165, 2006.
- [18] R. O. Duda, P. E. Hart, and D. Stork, *Pattern Classification*, Wiley, 2000.
- [19] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
- [20] A. Erfanian, and A. Erfani, "EEG-based brain-computer interface for hand grasp control: feature extraction by using ICA," *Proc. 9th Annual Conf. Int'l. Functional Electrical Stimulation Society*, 2004.
- [21] T. Trappenberg, J. Ouyang, and A. Back "Input Variable Selection: Mutual Information and Linear Mixing Measures," *IEEE Trans. Knowledge and Data Eng.*, Vol. 15, No. 1, pp. 37-46, 2006.
- [1] E. Haselsteiner, and G. Pfurtscheller, "Using time-dependent neural networks for EEG classification," *IEEE Trans. Rehab. Eng.*, Vol. 8, No. 4, pp. 457-463, 2000.
- [2] B. Mahmoudi, and A. Erfanian, "Single-channel EEG-based prosthetic hand grasp control for amputee subjects," *Proc. 24th Annual Int'l Conf. IEEE Engineering in Medicine and Biology Society*, pp. 2406 – 2407, 2002.
- [3] C. Guger, H. Ramoser, and G. Pfurtscheller, "Realtime EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI)," *IEEE Trans. Biomed. Eng.* Vol. 8, No. 4, pp. 447-456, 2000.
- [4] T. Hinterberger, S. Schmidt, N. Neumann, J. Mellinger, B. Blankertz, G. Curio, and N. Birbaumer, "Brain – computer communication and slow cortical potentials," *IEEE Trans. Biomed Eng.*, Vol. 51, pp. 1011-1018, 2004.
- [5] S. G. Mason, and G. E. Birch, "A brain-controlled switch for asynchronous control applications," *IEEE Trans. Biomed Eng.*, Vol. 47, pp. 1297-1307, 2000.
- [6] G. Pfurtscheller, C. Neuper, A. Schlogl, and K. Lugger, "Motor imagery and direct brain-computer communication," *Proc. the IEEE*, Vol. 89, pp. 1123-1134, 2001.
- [7] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain-computer interface research at the Wadsworth Center," *IEEE Trans Rehab Eng.*, Vol. 8, pp. 222–226, 2000.
- [8] E. Yom-Tov, and G. F. Inbar, "Feature selection for the classification of movements from single movement-related potentials," *IEEE Trans. Rehab. Eng.*, Vol. 10, No. 3, pp. 170-177, 2002.
- [9] G. Pfurtscheller, C. Neuper, A. Schlogl, and K. Lugger, "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters," *IEEE Trans. Rehab. Eng.*, Vol. 6, No. 3, pp. 316-325, 1998.
- [10] H. Serby H, E. Yom-Tov, and G. F. Inbar, "An improved P300-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, Vol. 13, pp. 89–98, 2005.

دانشگاه کیس وسترن آمریکا و مرکز پزشکی VA در شهر کلیولند ایالت اوهایو آمریکا گذرانده است. وی از سال ۱۳۷۴ به بعد، عضو هیأت علمی گروه مهندسی پزشکی دانشکده برق دانشگاه علم و صنعت ایران بوده و از سال ۱۳۷۹ تا ۱۳۸۷ مدیریت گروه مهندسی پزشکی دانشکده برق دانشگاه علم و صنعت ایران را به عهده داشته است. در حال حاضر دانشیار مهندسی پزشکی این دانشگاه است.

زمینه‌های تحقیقاتی وی شامل شبکه‌های عصبی، پردازش سیگنال‌های بیولوژیک، مدل‌سازی سیستم‌های بیولوژیک، مهندسی عصبی، سیستم‌های کنترلی عصبی-عضلانی، نظریه آشوب و فرکتال و کاربرد آن در حل مسائل پزشکی می‌باشد. یکی از مهمترین زمینه‌های تحقیقاتی وی کنترل سیستم‌های عصبی-عضلانی است. از سال ۱۳۶۹ تا کنون، وی مسؤل چندین پروژه تحقیقاتی در زمینه کنترل سیستم‌های عصبی-عضلانی بوده است. از نتایج این پروژه ها، گسترش پروتزهای عصبی حرکتی قابل حمل، برای اولین بار در ایران بوده است. با استفاده از این سیستم‌های الکترونیکی بیماران ضایعه نخاعی قادر به ایستادن و گام برداشتن می‌باشند.

عرفانیان، تا کنون دارای بیش از ۱۲۰ مقاله چاپ شده در داخل و خارج از کشور است و در سال ۱۳۷۷ و ۱۳۸۵ به عنوان پژوهشگر نمونه دانشگاه انتخاب شده است. در حال حاضر، وی دروس کارشناسی ارشد و دکترا شامل شبکه‌های عصبی، پردازش سیگنال‌های بیولوژیک، مدل‌سازی سیستم‌های بیولوژیک، سیستم‌های کنترل عصبی-عضلانی، تکنولوژی توانبخشی و پردازش سیگنال‌های دیجیتال پیشرفته را برای دانشجویان فوق لیسانس و دکترا ارائه می‌کند. عرفانیان عضو جامعه بین‌المللی (International Functional Electrical Stimulation Society) FES و IEEE است.

آدرس پست الکترونیکی ایشان عبارت است از:

erfanian@iust.ac.ir

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۶/۰۵/۲۲

تاریخ اصلاح: ۸۹/۰۱/۲۲

تاریخ قبول شدن: ۸۹/۰۴/۰۱

نویسنده مرتبط: دکتر عباس عرفانیان امیدوار، دانشکده مهندسی برق، دانشگاه علم و صنعت ایران، تهران، ایران.

[22] A. Hyvarinen, P. O. Hoyer, and M. Inki, "Topographic Independent Component Analysis," *Neural computation*, pp. 1527-1558, 2001.

[23] D. Huang, and T. Chow, "Effective feature selection scheme using mutual information," *Neurocomputing*, pp. 325-343, 2005.

[24] K. Hild, D. Erdogmus, K. Torkkola, and J. Pirinice, "Feature Extraction Using Information-Theoretic Learning," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 28, No. 9, pp. 1385-1392, 2006.

[25] N. Kwak, and C.-H. Choi, "Input feature selection for classification problems," *IEEE Trans. Neural Netw.*, Vol. 13, No. 1, pp. 143-159, 2002.

[26] D. Huang, and T. Chow, "Effective feature selection scheme using mutual information," *Neurocomputing*, pp. 325-343, 2005.

[27] B. Mahmoudi, and A. Erfanian, "Electro-encephalogram based brain-computer interface: improved performance by mental practice and concentration skills," *Med. & Biol. Eng & Compu*, Vol. 44, pp. 959-969, 2006.

[28] H. Ghandeharion, and A. Erfanian, "A fully Automatic method for ocular artifact suppression from EEG Data using wavelet transform and independent component analysis," *Proc, 28th Annual Int'l. Conf. IEEE Engineering in Medicine and Biology Society*, pp. 5265-5268, 2006.

[29] W. J. Krzanowski, *Principles of multivariate analysis: a user's perspective*, Oxford University Press, Oxford, 2000.



فرید اویسی ارنکه در سال ۱۳۶۱ در شهر تهران متولد گردید. ایشان مدارک کارشناسی و کارشناسی ارشد خود را در سال‌های ۱۳۸۳ و ۱۳۸۶ به ترتیب از دانشگاه آزاد اسلامی و دانشگاه علم و صنعت ایران دریافت کرد. زمینه کارهای تحقیقاتی او تاکنون پردازش سیگنال‌های مغزی، پردازش سیگنال‌های صوتی و روش‌های بازشناخت الگو بوده است. آدرس پست الکترونیکی ایشان عبارت است از:

faridoveisi@gmail.com



عباس عرفانیان امیدوار مدرک لیسانس خود را در سال ۱۳۶۴ در رشته مهندسی کامپیوتر از دانشگاه شیراز، مدرک فوق لیسانس را در سال ۱۳۶۸ در مهندسی برق و الکترونیک از دانشگاه صنعتی شریف تهران و مدرک دکترای تخصصی

خود را در مهندسی پزشکی در سال ۱۳۷۴ از دانشگاه تربیت مدرس تهران، ایران اخذ کرد. وی در دوره دکترای خود، به عنوان دانشجوی نمونه دانشگاه تربیت مدرس برگزیده شد. از سال ۱۳۶۵ الی ۱۳۶۸ به عنوان مهندس الکترونیک در مرکز تحقیقات مخابرات ایران مشغول بکار بوده و در زمینه طراحی و ساخت آزمایشگاه‌های میانمداری فعالیت داشته است. از سال ۱۳۶۸ الی ۱۳۷۰ به عنوان مربی آموزشی در دانشکده کامپیوتر دانشگاه علم و صنعت ایران مشغول بکار بوده است. او در سال ۱۳۷۲، یک دوره تحقیقاتی در زمینه کنترل سیستم‌های عصبی-عضلانی و رفع مشکلات حرکتی در افراد دچار ضایعه نخاعی با استفاده از فناوری عصبی در

¹ Brain-Computer Interface (BCI)

² Feature Selection

³ Feature Extraction

⁴ Wrapper

⁵ Distinctive Sensitive Learning Vector Quantization

⁶ Principle Component Analysis

⁷ Linear Discriminant Analysis

⁸ Independent Component Analysis

⁹ Shannon

¹⁰ Gradient ascent

¹¹ Probability Density Function

¹² Minimum Renyi's Mutual Information-Stochastic Information Gradient

¹³ ANOVA

ارزیابی کمی سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد در نرم‌افزارهای حساس به ایمنی

غلامرضا شاه محمدی سعید جلیلی

دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران

چکیده

در فرایند طراحی معماری تصمیمات متعددی اتخاذ می‌شود که تاثیر این تصمیمات در سطح سیستم نرم‌افزاری می‌باشد. یکی از این تصمیم‌های مهم، انتخاب سبک معماری مناسب است. با توجه به نقش حیاتی قابلیت اعتماد در سیستم‌های نرم‌افزاری حساس به ایمنی، یکی از معیارهای اصلی در انتخاب سبک معماری نرم‌افزار قابلیت اعتماد بالا می‌باشد. در فرایند طراحی از سبک‌های معماری نرم‌افزار، که راه‌حل‌های آزموده شده‌ای می‌باشند، زیاد استفاده می‌شود. اما با توجه به اینکه در هر سیستمی ترکیب خاصی از صفات کیفی مورد توجه است، عدم تحلیل کمی تاثیر سبک‌های معماری نرم‌افزار بر روی صفات کیفی مانع استفاده موثر از سبک‌های معماری می‌گردد. این تحقیق گامی به سوی کمی سازی تاثیر سبک‌های معماری نرم‌افزار بر روی صفت کیفی قابلیت اعتماد می‌باشد. به این ترتیب با استفاده از نتایج این تحقیق می‌توان تعیین کرد که در سیستمی که قابلیت اعتماد از اهمیت بالایی برخوردار است، بکارگیری چه سبک‌هایی قابلیت اعتماد بالاتری را بدنبال خواهد داشت.

در این مقاله با استفاده از دو رویکرد دیاگرام بلوکی قابلیت اعتماد و رویکرد مدل مبتنی بر حالت مارکوف، قابلیت اعتماد سبک‌های معماری نرم‌افزار محاسبه می‌گردد. آنگاه نتایج حاصل از دو رویکرد روی سبک‌های معماری نرم‌افزار، مقایسه می‌گردد. دستاورد اصلی این مقاله کمی‌سازی و رتبه‌بندی ارزش سبک‌های معماری نرم‌افزار از دید صفت کیفی قابلیت اعتماد در مرحله انتخاب سبک معماری نرم‌افزار می‌باشد.

کلمات کلیدی: معماری نرم‌افزار، ارزیابی قابلیت اعتماد سبک‌های معماری نرم‌افزار، دیاگرام بلوکی قابلیت اعتماد، مدل مارکوف.

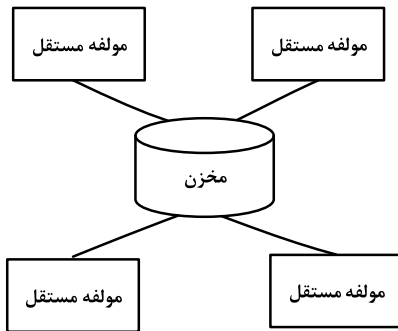
۱- مقدمه

سیستم می‌باشد [۲]. یکی از تصمیم‌های مهم مرحله طراحی نرم‌افزار، انتخاب سبک معماری مناسب است. از آنجا که وظیفه‌مندی‌های نرم‌افزار با استفاده از ساختار(های) متعددی قابل تامین است [۳]، بنابراین عامل تعیین کننده در انتخاب سبک معماری نرم‌افزار، میزان حمایت از صفات کیفی مورد نظر می‌باشد. به همین دلیل تحقیقات متعددی در این زمینه انجام شده است [۴، ۵، ۶، ۷، ۸]. با توجه به نقش اساسی قابلیت اعتماد در سیستم‌های نرم‌افزاری، خصوصاً سیستم‌های نرم‌افزاری حساس به ایمنی، یکی از ملاک‌های اصلی در انتخاب سبک معماری نرم‌افزار، میزان حمایت سبک معماری از قابلیت اعتماد به نرم‌افزار می‌باشد. از مشکلات عمده در استفاده از سبک‌های معماری عدم بررسی تاثیر کمی آنها روی صفات کیفی می‌باشد [۹]. این تحقیق گامی به سوی کمی سازی

نقش نرم‌افزارهای حساس به ایمنی در زندگی بشر آنقدر مهم است، که شکست آنها باعث بروز خسارت‌های مالی و جانی جبران ناپذیری می‌گردد. در نتیجه توجه به صفات کیفی نرم‌افزار در مراحل اولیه چرخه حیات نرم‌افزار اهمیت فراوانی دارد. قابلیت اعتماد یکی از صفات کیفی حیاتی نرم‌افزار تلقی می‌شود و در نرم‌افزارهای حساس به ایمنی تامین آن یک ضرورت اجتناب ناپذیر می‌باشد. طراحی معماری نرم‌افزار اولین مرحله در توسعه نرم‌افزار است که ارزیابی چگونگی تامین نیازهای نرم‌افزار، خصوصاً نیازهای غیروظیفه‌ای، در آن انجام می‌شود [۱]. در فرایند طراحی معماری تصمیمات متعددی اتخاذ می‌شود که تاثیر این تصمیمات در سطح

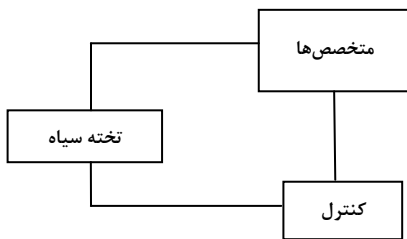
می‌دهد [۱۶].

سبک مخزن (RPS): سبک مخزن (شکل ۱) شامل یک انباره مرکزی به نام مخزن (C_{TPS}) و مجموعه‌ای از مولفه‌ها (C_i) است که اعمال ذخیره، بازیابی و بروز رسانی اطلاعات را بر روی انباره مرکزی انجام می‌دهند. پایگاه داده، نمونه‌ای از این سبک می‌باشد [۱۵]. در سبک مخزن تعامل هر مولفه با مخزن یک تراکنش محسوب می‌شود.



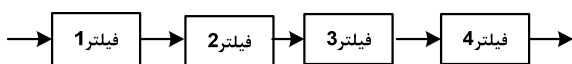
شکل ۱- سبک مخزن

سبک تخته سیاه^۸ (BKB): مولفه‌های این سبک (شکل ۲)، تخته‌سیاه (C_{bkb})، منابع‌دانش^۹ (متخصص‌ها) (C_{kr}) و مولفه کنترل (C_c) است. مولفه کنترل، یک حلقه را اجراء و ضمن نظارت بر تغییرات تخته‌سیاه و ارزیابی متخصص‌ها، یکی از آنها را برای اجراء فعال می‌کند [۱۴]. بنابراین انجام یک تراکنش در این سبک مستلزم اجراء برنامه مولفه کنترل و بررسی وضعیت مولفه تخته‌سیاه توسط مولفه کنترل و انتخاب متخصص مناسب و تعامل متخصص با تخته‌سیاه می‌باشد.



شکل ۲- سبک تخته سیاه

سبک لوله و فیلتر (P/F): این سبک (شکل ۳) از تعدادی مولفه (C_i) تشکیل شده است. هر مولفه در نقش یک فیلتر است و تعدادی ورودی و تعدادی خروجی دارد. خروجی هر مولفه، ورودی مولفه بعدی است [۱۴]. در انجام یک تراکنش کلیه فیلترها موثر می‌باشند.



شکل ۳- سبک لوله و فیلتر

سبک لایه‌ای (LYD): در سبک لایه‌ای (شکل ۴)، تمرکز روی سطوح مختلف انتزاع در نرم‌افزار است. سبک لایه‌ای بصورت سلسله مراتبی سازماندهی می‌شود. هر لایه (C_i) خدمتی را برای لایه بالاتر فراهم و از خدمت لایه پایین‌تر استفاده می‌کند [۱۴]. در انجام یک تراکنش کلیه لایه‌ها موثر می‌باشند.

تاثیر سبک‌های معماری نرم‌افزار بر صفت کیفی قابلیت اعتماد می‌باشد.

روش‌های متعددی برای ارزیابی قابلیت اعتماد نرم‌افزار ارائه شده است [۱۰، ۱۱، ۱۲]. یادآوری می‌گردد در مرحله انتخاب سبک معماری نرم‌افزار در فرایند ساخت نرم‌افزار، وظیفه‌مندی‌های مولفه‌های سبک‌های معماری نرم‌افزار مشخص نیست. در نتیجه تنها برخی از این روش‌ها مانند دی‌گرام بلوکی قابلیت اعتماد^۱ (RBD) [۱۲] و روش مبتنی بر حالت مارکوف [۱۰] برای ارزیابی قابلیت اعتماد سبک‌های معماری نرم‌افزار، قابل استفاده می‌باشند. در [۱۳]، سبک‌های معماری نرم‌افزار را از دید قابلیت اعتماد براساس روش RBD ارزیابی کردیم و بازای اندازه‌های مختلف نرم‌افزار (تعداد مختلف مولفه‌ها)، قابلیت اعتماد سبک‌های معماری نرم‌افزار را محاسبه کردیم. در این تحقیق: (۱) به جای اندازه نرم‌افزار، با توجه به مفهوم تراکنش و تعداد مولفه‌های موثر در انجام تراکنش در هر یک از سبک‌های معماری نرم‌افزار، قابلیت اعتماد سبک‌های معماری نرم‌افزار ارزیابی شده است. (۲) مولفه سرویس‌گر در سبک‌های متقاضی/سرویس‌گر و کارگزار به جای مولفه واحد پیچیده، مرکب از m مولفه در نظر گرفته شده‌اند. (۳) ارزیابی سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد با رویکردهای RBD و مبتنی بر حالت مارکوف انجام و ضمن مقایسه نتایج دو رویکرد، سبک‌های معماری نرم‌افزار رتبه‌بندی شده‌اند.

روش RBD، یک نمایش بلوکی از مولفه‌های نرم‌افزار و وضعیت قابلیت اعتماد آنها ارائه می‌دهد. با توجه به اینکه سبک‌های معماری نرم‌افزار، مدل‌های مختلفی برای حل مسایل معماری ارائه می‌دهند، RBD روش مناسبی برای تعیین قابلیت اعتماد سبک‌های معماری و تخمین قابلیت اعتماد آنها و در نتیجه رتبه‌بندی سبک‌های معماری از دید قابلیت اعتماد در مرحله انتخاب سبک معماری برای نرم‌افزار می‌باشد. در رویکرد مبتنی بر حالت، معماری نرم‌افزار به صورت زنجیر مارکوف زمان گسسته مدل شده و قابلیت اعتماد نرم‌افزار به صورت تحلیلی، تخمین زده می‌شود. بنابراین با تبدیل هر سبک معماری به ماشین حالت، امکان ارزیابی قابلیت اعتماد آن به صورت تحلیلی وجود دارد.

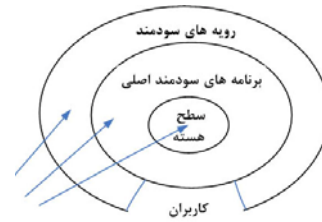
ما در این مقاله با استفاده از دو رویکرد RBD و روش مبتنی بر حالت مارکوف سبک‌های معماری نرم‌افزار را از دید قابلیت اعتماد بطور کمی ارزیابی می‌کنیم. این مقاله در بخش‌های زیر تنظیم شده است. در بخش دوم، سبک‌های معماری نرم‌افزار، در بخش سوم مدل‌های کیفی نرم‌افزار و در بخش چهارم قابلیت اعتماد نرم‌افزار و روشهای ارزیابی آن مطرح می‌گردد. در بخش پنجم ارزیابی کمی و در بخش ششم رتبه‌بندی سبک‌های معماری نرم‌افزار و در بخش هفتم بررسی موردی انجام شده، ارائه می‌گردد. در بخش هشتم نتیجه‌گیری می‌گردد.

۲- سبک‌های معماری نرم‌افزار

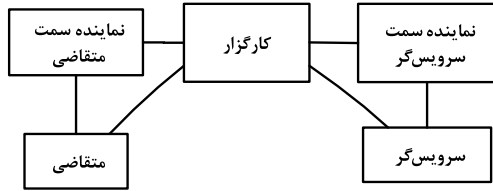
سبک‌های معماری نرم‌افزار، مدل‌هایی برای حل مسایل طراحی نرم‌افزار ارائه می‌دهند. هر مدل، مولفه‌های تشکیل‌دهنده، مسئولیت مولفه‌ها، روابط و نحوه همکاری آنها را توصیف می‌کند [۱۴]. استفاده از الگوها در بسیاری از اصول مهندسی رایج است. الگوها براساس مقیاس به دو دسته سبک‌های معماری^۲ و الگوهای طراحی^۳ تقسیم می‌شوند. Shaw و همکارش [۱۵] هفت سبک معماری رایج شامل لوله و فیلتر^۴، شی‌گرا^۵ یا نوع داده انتزاعی، فراخوانی ضمنی^۶، لایه‌ای^۷، مخزن^۸، مفسر و کنترل فرایند را معرفی کرده‌اند. Buschmann و همکارانش [۱۴] نیز در کتاب خود، الگوها را در سطوح مختلف بررسی کرده‌اند.

در این بخش با توجه به اینکه قابلیت اعتماد سبک‌های معماری نرم‌افزار را با توجه به مولفه‌های موثر در انجام تراکنش بررسی می‌کنیم، ضمن اشاره به مفهوم تراکنش، هشت سبک معماری بطور اجمالی معرفی می‌شود. یادآوری می‌گردد، تراکنش، مجموعه‌ای از اعمال است که یک واحد منطقی از کار را تشکیل

معرفی و سرویس‌های خود را از طریق واسط متد، در دسترس متقاضیان قرار می‌دهند. متقاضیان از خدمات سرویس‌گرها از طریق کارگزار استفاده می‌کنند. کارگزار، محل سرویس‌گر مناسب را تعیین، درخواست را به او ارسال، و نتیجه را به متقاضی برمی‌گرداند [۱۴]. تراکنش در این سبک مشابه سبک متقاضی/سرویس‌گر است، اما مولفه‌های بیشتری در آن موثر می‌باشند.

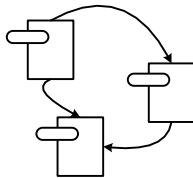


شکل ۴- سبک لایه‌ای



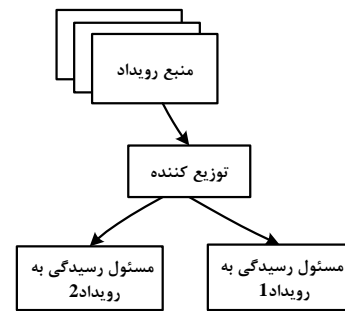
شکل ۷- سبک کارگزار

سبک شی‌گرا^{۱۴} (OO): در سبک شی‌گرا (شکل ۸)، نمایش داده و عملیات مرتبط با آن در یک شی محصور می‌شود. مولفه‌های این سبک، اشیاء و تعامل اشیاء از طریق فراخوانی تابع می‌باشد [۱۵]. تراکنش در سبک شی‌گرا به سناریو معروف است. هر سناریو زنجیری از اجرای متدی از تعدادی شی را شامل می‌شود.



شکل ۸- سبک شی‌گرا

سبک فراخوانی ضمنی (I/I): این سبک (شکل ۵) شامل تعدادی مولفه (C_i) (که مولد رویداد، علاقه‌مند به رویداد و یا هم مولد و هم علاقه‌مند به رویداد می‌باشد) است و یک مولفه در نقش هماهنگ‌کننده یا توزیع‌کننده (C_d)، با وقوع رویداد، از طریق جدول نگاشت رویداد به مولفه علاقه‌مند مراجعه، ضمن تعیین مولفه علاقه‌مند، پیامی به او ارسال یا رویه‌ای از آن مولفه را فراخوانی می‌کند [۱۵]. تراکنش در این سبک با وقوع رویداد شروع شده و مولفه توزیع‌کننده، مولفه علاقه‌مند را فعال می‌کند. با اتمام فعالیت مولفه علاقه‌مند تراکنش خاتمه می‌یابد.



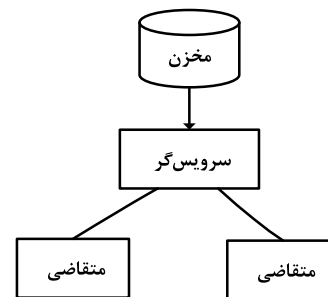
شکل ۵- سبک فراخوانی ضمنی

۳- مدل‌های کیفی نرم‌افزار

هدف اصلی از ساخت نرم‌افزار، پیاده‌سازی نیازهای وظیفه‌ای آن است. بایستی توجه داشت که در فرایند ساخت نرم‌افزار، کیفیت نرم‌افزار براساس معماری آن پایه‌گذاری می‌شود. یادآوری می‌گردد که طراحی مجدد نرم‌افزارها اغلب به سبب نقص در انجام وظایف محوله به آنها نیست، بلکه به دلیل عدم برخورداری از کیفیت لازم می‌باشد [۱۷]. بنابراین نیرو محرکه‌ی پشتیبان صفات کیفی مهم نرم‌افزار، تقسیم وظایف میان ساختارهای (مولفه‌ها) معماری است.

معماری نرم‌افزار [۱۸] باید دو گروه از نیازهای کیفی را تامین کند. نیازهای کیفی توسعه و نیازهای کیفی عملیاتی. نیازهای کیفی توسعه، مانند قابلیت نگهداری و قابلیت استفاده مجدد، مورد توجه تولیدکننده نرم‌افزار می‌باشد. نیازهای کیفی عملیاتی، مانند کارایی و قابلیت اعتماد مورد توجه کاربران سیستم می‌باشند. مدل‌های کیفی مختلفی برای نرم‌افزار ارائه شده که مهمترین آنها مدل MacCall [۱۹]، مدل Boehm [۲۰] و مدل ISO9126 [۲۱] می‌باشد. یک مدل کیفی سلسله مراتبی سه سطحی از صفات کیفی نرم‌افزار ارائه می‌دهد که قابلیت اعتماد در اولین سطح آن قرار دارد. مدل کیفی Boehm نیز کیفیت نرم‌افزار را مشابه مدل MacCall به صورت سلسله مراتبی از مشخصه‌ها^{۱۵}، در سه سطح تعریف می‌کند و قابلیت اعتماد در سطح دوم این مدل قرار دارد. مدل ISO 9126 که توسط سازمان بین‌المللی استاندارد، برای سنجش کیفیت نرم‌افزار ارائه شده است نیز سلسله مراتبی از مشخصه‌ها در دو سطح است و قابلیت اعتماد در سطح اول آن قرار دارد. با توجه به اینکه مدل کیفی ISO 9126 براساس مدل های MacCall و Boehm تعریف شده است و بسیاری از مشخصه‌های آنها را دارد، بنابراین در این تحقیق صفت کیفی قابلیت اعتماد نرم‌افزار بر مبنای تعریف این مدل مورد سنجش

سبک متقاضی/سرویس‌گر (C/S): مولفه‌های این سبک (شکل ۶) متقاضیان (C_c)، سرویس‌گر (ها) (C_s) و مخزن (C_{TPS}) می‌باشند. متقاضیان برای استفاده از خدمات سرویس‌گرها، باید از نام و خدمات ارائه شده توسط آنها اطلاع داشته باشند. تراکنش در این سبک با ارسال درخواست متقاضی به سرویس‌گر آغاز می‌شود و به طور معمول نیاز به مراجعه سرویس‌گر به مخزن دارد و در نهایت با ارسال نتیجه توسط سرویس‌گر به متقاضی به پایان می‌رسد.



شکل ۶- سبک متقاضی/سرویس‌گر

سبک کارگزار^{۱۱} (BRK): مولفه‌های این سبک (شکل ۷) متقاضیان (C_c)، سرویس‌گر (ها) (C_s)، کارگزار (C_{brk})، نماینده سمت متقاضی^{۱۲} (C_{esp})، نماینده سمت سرویس‌گر^{۱۳} (C_{ssp}) و مخزن (C_{TPS}) است. کارگزار مسئول هماهنگی ارتباط و مجزا کردن متقاضیان از سرویس‌گرها می‌باشد. سرویس‌گرها خود را به کارگزار

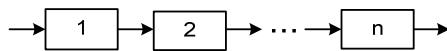
قرار می‌گیرد.

نرم‌افزار می‌توان استفاده کرد. روش دیاگرام بلوکی قابلیت اعتماد (RBD) به عنوان یک روش مبتنی بر ساختار نرم‌افزار، برای ارزیابی قابلیت اعتماد سیستم‌های نرم‌افزاری ارائه شده است. Tripathi و همکارش [۲۳] از RBD برای تخمین قابلیت اعتماد سیستم‌های نرم‌افزاری مرکب از سلسله مراتبی از پیمان‌ها استفاده کرده‌اند. Leblanc و همکارش [۲۴] براساس RBD مدلی برای نمایش مسایل دنیای واقعی و الگوریتمی برای تحلیل این مدل در مرحله اولیه توسعه نرم‌افزار ارائه کرده‌اند.

در ادامه با استفاده از روش RBD و روش مدل مبتنی بر حالت مارکوف، سبک‌های معماری نرم‌افزار مورد ارزیابی قرار می‌گیرند.

۴-۱- مدل سازی و ارزیابی قابلیت اعتماد سبک‌های معماری نرم‌افزار با روش دیاگرام بلوکی قابلیت اعتماد (RBD)

با توجه به اینکه هر سبک معماری، مولفه‌های تشکیل دهنده معماری و نحوه تعامل آنها با یکدیگر را تعیین می‌کند، روش RBD برای ارزیابی سبک‌های معماری نرم‌افزار مناسب می‌باشد. روش دیاگرام بلوکی قابلیت اعتماد، قابلیت اعتماد سبک‌های معماری نرم‌افزار را براساس ساختار آنها تعیین می‌کند. پیش‌بینی قابلیت اعتماد سبک معماری نرم‌افزار از طریق بررسی قابلیت اعتماد مولفه‌های تشکیل دهنده آن سبک معماری، قابل انجام است. برای ساخت یک RBD باید پیکربندی مولفه‌های سبک معماری نرم‌افزار را تعیین کرد (شکل ۹).



شکل ۹- دیاگرام بلوکی پیکربندی ترتیبی

یک پیکربندی ترتیبی از مولفه‌ها را نشان می‌دهد. در پیکربندی ترتیبی، شکست هر مولفه منجر به شکست نرم‌افزاری می‌گردد. قابلیت اعتماد نرم‌افزار با پیکربندی ترتیبی برابر احتمال این است که همه مولفه‌های سبک معماری آن نرم‌افزار درست عمل کنند. در این صورت، قابلیت اعتماد نرم‌افزار از رابطه (۱) محاسبه می‌گردد.

$$R_s = P(X_1 \cap X_2 \cap \dots \cap X_n) = P(X_1) P(X_2|X_1) P(X_3|X_1 X_2) \dots P(X_n|X_1 X_2 \dots X_{n-1}) \quad (1)$$

در این رابطه R_s قابلیت اعتماد نرم‌افزار، X_i رخداد فعال بودن مولفه i -ام و $P(X_i)$ احتمال فعال بودن مولفه i -ام می‌باشد. اگر مولفه‌ها مستقل از یکدیگر در نظر گرفته شوند رابطه (۱) به رابطه (۲) تبدیل می‌شود. در این رابطه R_i قابلیت اعتماد مولفه i -ام می‌باشد.

$$R_s = \prod_{i=1}^n P(X_i) = \prod_{i=1}^n R_i \quad (2)$$

در نرم‌افزارهای با پیکربندی موازی (شکل ۱۰)، برای موفقیت نرم‌افزار حداقل باید یکی از مولفه‌ها درست عمل کند. احتمال شکست نرم‌افزاری با n مولفه موازی برابر احتمال این است که همه مولفه‌های نرم‌افزار دچار شکست شوند. به عبارت دیگر، اگر حداقل یک مولفه درست عمل کند، آنگاه نرم‌افزار درست عمل خواهد کرد. قابلیت اعتماد نرم‌افزار از رابطه (۳) محاسبه می‌گردد.

برطبق مدل ISO 9126 قابلیت اعتماد، مجموعه‌ای از زیر مشخصه‌ها است که قابلیت نرم‌افزار را برای حفظ کارایی نرم‌افزار تحت شرایط بیان شده در دوره زمانی معین، مشخص می‌کند [۲۱]. به عبارت دیگر قابلیت اعتماد، مدت زمانی را نشان می‌دهد که نرم‌افزار برای استفاده در دسترس است. زیرمشخصه‌های قابلیت اعتماد در مدل ISO 9126 به شرح زیر است.

- بلوغ^{۱۶}: فرکانس شکست نرم‌افزار بر اثر عیوب در نرم‌افزار را مشخص می‌کند. هر قدر فرکانس شکست کمتر باشد، میزان بلوغ نرم‌افزار بیشتر است.
 - تحمل خطا: قابلیت حفظ سطح کارایی مشخص شده در صورت بروز عیب در نرم‌افزار را نشان می‌دهد.
 - قابلیت ترمیم: برقراری مجدد سطح کارایی و ترمیم داده‌ها در موقع بروز شکست، و زمان و تلاش مورد نیاز برای آن را نشان می‌دهد.
 - بررسی زیرمشخصه‌های قابلیت اعتماد، موارد زیر را نشان می‌دهد.
 - فرکانس شکست نرم‌افزار با تعداد مولفه‌های حیاتی سبک معماری، ارتباط مستقیم دارد. زیرا هر قدر تعداد مولفه‌های حیاتی سبک معماری بیشتر باشد، امکان بالقوه شکست نرم‌افزار بیشتر می‌گردد.
 - تحمل خطای نرم‌افزار نیز با تعداد مولفه‌های حیاتی سبک معماری، ارتباط معکوس دارد، زیرا هر قدر تعداد مولفه‌های حیاتی سبک معماری بیشتر باشد، تحمل خطای نرم‌افزار کمتر می‌باشد.
 - قابلیت ترمیم با تعداد مولفه‌های حیاتی سبک معماری، ارتباط معکوس دارد. زیرا اصولاً قابلیت ترمیم برای مولفه‌هایی مطرح است که در صورت شکست آنها، کارایی نرم‌افزار دچار تنزل جدی می‌شود.
- با توجه به مطالب فوق‌الذکر، رویکرد RBD و رویکرد مدل مبتنی بر حالت مارکوف که ارزیابی قابلیت اعتماد را براساس مولفه‌های نرم‌افزار و نحوه تعامل آنها انجام می‌دهند، روش‌های مناسبی برای ارزیابی قابلیت اعتماد و زیر مشخصه‌های آن در سطح سبک‌های معماری نرم‌افزار می‌باشند.

۴- قابلیت اعتماد نرم‌افزار و روش‌های ارزیابی آن

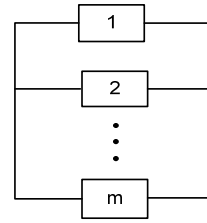
در [۱]، روش‌های ارزیابی قابلیت اعتماد در یک انتزاع سطح بالا، به روش‌های کمی و کیفی تقسیم شده‌اند. گونه‌های مختلفی از روش‌های کمی وجود دارد که برخی قبل و برخی بعد از پیاده‌سازی سیستم نرم‌افزاری قابل استفاده می‌باشند. روش‌های مبتنی بر سنجش که روی شکست و از کار افتادگی سیستم متمرکز می‌باشند، قبل و بعد از پیاده‌سازی سیستم نرم‌افزاری قابل استفاده می‌باشند.

روش‌های مدل‌سازی و ارزیابی قابلیت اعتماد معماری نرم‌افزار [۲۲] به سه رویکرد مبتنی بر حالت، مبتنی بر مسیر و افزایشی دسته بندی می‌شود. در رویکرد مبتنی بر حالت، معماری نرم‌افزار به صورت زنجیر مارکوف زمان گسسته یا پیوسته، یا فرایند نیمه مارکوف مدل شده و قابلیت اعتماد نرم‌افزار به صورت تحلیلی روی معماری با رفتار شکست، تخمین زده می‌شود. در رویکرد مبتنی بر مسیر، برای محاسبه قابلیت اعتماد نرم‌افزار، مسیرهای ممکن اجراء، به صورت تجربی، توسط آزمون یا به صورت الگوریتمی، در نظر گرفته می‌شود. در رویکرد افزایشی، معماری نرم‌افزار بطور صریح در نظر گرفته نمی‌شود، بلکه تمرکز این روش بر تخمین چگالی شکست وابسته به زمان سیستم کاربردی، با استفاده از داده‌های شکست مولفه، می‌باشد. با توجه به این که سبک‌های معماری نرم‌افزار انتزاعی هستند و در نتیجه فاقد مسیرهای ممکن اجراء و تخمین چگالی شکست وابسته به زمان هستند، روش‌های مبتنی بر رویکرد مسیر و رویکرد افزایشی برای ارزیابی قابلیت اعتماد سبک‌های معماری نرم‌افزار، قابل استفاده نمی‌باشند. اما از مدل مبتنی بر حالت مارکوف برای ارزیابی قابلیت اعتماد سبک‌های معماری

۲. یک ماتریس تصادفی $T=(P_{ij})$ ، با ابعاد $n \times n$ ، که $\sum_{j=1}^n P_{ij} = 1$ و $P_{ij} \geq 0$ ، با فرض اینکه سیستم در حالت s_i باشد، احتمال گذار به حالت s_j است. P_{ij} با فرض اینکه سیستم در حالت s_i باشد، احتمال گذار به حالت s_j است.

۳. بردار $\pi^0 = (\pi_1^0, \dots, \pi_n^0)$ که نشانگر احتمال آن است که سیستم ابتدا در حالت s_i باشد.

تعریف ۱- حالت s_i گذرا است اگر گذار $s_i \rightarrow s_j$ موجود باشد اما گذار $s_j \rightarrow s_i$ موجود نباشد، یعنی s_i به بعضی از s_j ها دسترسی دارد، اما s_j ها به s_i دسترسی ندارند. در غیر اینصورت حالت s_i ارگودیک است.



شکل ۱۰- دیاگرام بلوکی پیکربندی موازی

$$R_s = 1 - P(X_1 \cap X_2 \cap \dots \cap X_n) = 1 - [P(X_1)P(X_2|X_1)P(X_3|X_1X_2) \dots P(X_n|X_1(X_2 \dots X_{n-1}))] \quad (3)$$

تعریف ۲- کلاس یک زنجیر مارکوف، کلاس هم ارزی استنتاج شده^{۱۷} توسط رابطه ارتباط روی مجموعه‌ای از حالت‌ها است. یک کلاس α به کلاس β دسترسی دارد اگر برای بعضی از $s_i \in \alpha$ و $s_j \in \beta$ ، $s_i \rightarrow s_j$ باشد. کلاسی که به کلاس دیگری دسترسی نداشته باشد، کلاس نهایی نامیده می‌شود. یک کلاس هم ارزی ارگودیک، یک کلاس نهایی است که همه حالت‌های آن ارگودیک باشد.

قضیه ۱- اگر $a_{ij}^{(q)}$ عنصر (i, j) از A^q را مشخص کند، ماتریس A کاهش‌ناپذیر است اگر و فقط اگر برای هر (i, j) ، یک عدد طبیعی q موجود باشد، بطوریکه $a_{ij}^{(q)} > 0$ باشد.

در این رابطه R_s قابلیت اعتماد نرم‌افزار، X_i پیشامد شکست مولفه i -ام و $P(X_i)$ احتمال شکست مولفه i -ام می‌باشد. اگر مولفه‌ها از یکدیگر مستقل باشند، رابطه (۳) به رابطه (۴) تبدیل می‌شود. در این رابطه R_i قابلیت اعتماد مولفه i -ام می‌شود.

مدل قابلیت اعتماد نرم‌افزار Cheung [۱۰] که به مدل کاربر-گرا معروف است، و روی مدل مارکوف ارائه شده است، در ادامه معرفی می‌گردد.

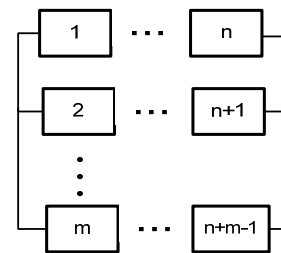
$$R_s = 1 - P(X_1)P(X_2) \dots P(X_n) = 1 - \prod_{i=1}^n P(X_i) = 1 - \prod_{i=1}^n Q_i = 1 - \prod_{i=1}^n (1 - R_i) \quad (4)$$

۴-۲-۱- مدل Cheung

در این رابطه Q_i عدم اطمینان مولفه i -ام می‌باشد. در پیکربندی موازی، مولفه با بیشترین قابلیت اعتماد، بیشترین اثر را روی قابلیت اعتماد نرم‌افزار دارد، زیرا قابل اعتمادترین مولفه، آخرین مولفه‌ای است که احتمال شکست دارد.

این مدل از یک ماتریس تصادفی استاندارد $n \times n$ بنام T تشکیل شده است (رابطه ۵). $n = k + 2$ به صورت رابطه (۶) است. در ماتریس T ، یک ماتریس داخلی بنام M قرار دارد. این ماتریس، یک ماتریس گذار حالت $k \times k$ است که s_1 حالت ورود و s_k حالت خروج می‌باشد. $M(i, j)$ عنصر i -امین ردیف و j -امین ستون، R_i قابلیت اعتماد مولفه نرم‌افزاری c_i و $P_{i,j}$ احتمال گذار از مولفه c_i به c_j می‌باشد که $1 \leq i, j \leq k$ در حالت ارگودیک S و F بترتیب حالت موفق و حالت شکست نرم‌افزار را نشان می‌دهند. بنابراین، اجرای موفق نرم‌افزار در نهایت به حالت ارگودیک S و در غیر این صورت به حالت F ختم می‌شود.

شکل ۱۱، پیکربندی ترتیبی- موازی را نشان می‌دهد. برای محاسبه قابلیت اعتماد نرم‌افزار با پیکربندی ترتیبی- موازی، ابتدا قابلیت اعتماد بخش‌های ترتیبی را محاسبه کرده و در رابطه (۴)، یعنی رابطه قابلیت اعتماد نرم‌افزار با پیکربندی موازی، درج می‌کنیم. به این ترتیب قابلیت اعتماد نرم‌افزار با پیکربندی ترتیبی- موازی بدست می‌آید.



شکل ۱۱- دیاگرام بلوکی پیکربندی ترتیبی- موازی

از آنجا که حالات S و F ارگودیک می‌باشند، احتمال گذار از حالت S به S و احتمال گذار از حالت F به F هر یک معادل ۱ می‌باشند. دو بردار داخلی B_1 و B_2 نیز وجود دارد که B_1 فقط یک عنصر مخالف ۰ دارد که احتمال رسیدن به حالت S از حالت خروج s_k است. B_2 احتمال‌های شکست از حالت‌های s_1 تا s_k به حالت F را ذخیره می‌کند.

بر اساس معادله‌های (۵) و (۶)، این ماشین حالت محدود یک مدل مارکوف زمان گسسته است که مجموع هر ردیف ماتریس استاندارد T برابر ۱ می‌باشد.

$$T = \begin{matrix} & S & F & s_1 \dots s_k \\ \begin{matrix} S \\ F \\ s_1 \dots s_k \end{matrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ B_1 & B_2 & M \end{bmatrix} & & \end{matrix}, B_1 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ R_k \end{bmatrix}, B_2 = \begin{bmatrix} 1 - R_1 \\ \vdots \\ 1 - R_{k-1} \\ 1 - R_k \end{bmatrix} \quad (5)$$

۴-۲-۲- مدل سازی و ارزیابی قابلیت اعتماد سبک‌های معماری نرم‌افزار با روش مبتنی بر حالت مارکوف

$M(i, j) = \begin{cases} R_i P_{ij} & \text{state } s_i \text{ reaches } s_j \text{ and } i \neq k \\ 0, & \text{otherwise,} \end{cases}$ (6)
for $1 \leq i, j \leq k$

مدل مارکوف، یک ماشین حالت محدود است که احتمال گذار به حالت بعدی فقط به حالت جاری بستگی دارد. در مدل مارکوف زمان گسسته، گذارها در فواصل زمانی گسسته رخ می‌دهند و احتمال گذار، از یک توزیع گسسته تبعیت می‌کند [۲۵]. مدل مارکوف زمان گسسته شامل:

۱. مجموعه محدودی از n حالت $S = \{S_1, \dots, S_n\}$

سبک‌های معماری نرم‌افزار محاسبه و در پایان، نتایج حاصل از دو روش با یکدیگر مقایسه می‌شود.

۵-۱- سنجش سبک‌های معماری نرم‌افزار با روش RBD

در این بخش براساس روش RBD، سبک‌های معماری نرم‌افزار مورد سنجش قرار می‌گیرد و نتایج ارزیابی در جدول ۱ ارائه می‌شود. یادآوری می‌گردد که قابلیت اعتماد مولفه‌ها مستقل از یکدیگر در نظر گرفته می‌شود.

سبک مخزن (RPS): در سبک مخزن انجام تراکنش مستلزم تعامل مولفه مستقلی مانند C_i با مخزن است. در نتیجه قابلیت اعتماد سبک مخزن از رابطه (۸) محاسبه می‌گردد. در این رابطه R_{RPS} ، قابلیت اعتماد مولفه مخزن و R_i قابلیت اعتماد هر یک از مولفه‌های مستقل می‌باشد.

$$R = R_i \cdot R_{RPS} \quad (۸)$$

سبک تخته سیاه (BKB): در سبک تخته‌سیاه، انجام تراکنش مستلزم اجرای برنامه مولفه کنترل (C_c) و بررسی وضعیت مولفه تخته‌سیاه (C_{BKB}) توسط مولفه کنترل و انتخاب متخصص مناسب (C_{kr}) و تعامل متخصص با تخته‌سیاه می‌باشد. در نتیجه قابلیت اعتماد سبک تخته‌سیاه از رابطه (۹) محاسبه می‌گردد. در این رابطه R_c قابلیت اعتماد مولفه کنترل، R_{BKB} قابلیت اعتماد مولفه تخته‌سیاه و R_{kr} قابلیت اعتماد متخصص می‌باشد.

$$R = R_c \cdot R_{BKB} \cdot R_{kr} \quad (۹)$$

سبک لوله و فیلتر (P/F): در سبک لوله و فیلتر (شکل ۳) برای انجام یک تراکنش، همه مولفه‌ها باید فعال باشند. با توجه به اینکه در هر تراکنش m مولفه موثر می‌باشد، قابلیت اعتماد این سبک از رابطه (۱۰) محاسبه می‌شود.

$$R = \prod_{i=1}^m R_i \quad (۱۰)$$

سبک لایه‌ای (LYD): در سبک لایه‌ای (شکل ۴) برای تامین یک سرویس مشخص، همه مولفه‌ها باید فعال باشند. بنابراین قابلیت اعتماد این سبک نیز از رابطه (۱۰) محاسبه می‌شود. در این رابطه R_i قابلیت اعتماد هر لایه می‌باشد.

سبک فراخوانی ضمنی (I/I): در این سبک تراکنش با وقوع رویداد شروع شده و منجر به فعال‌سازی مولفه علاقه‌مند (C_i) توسط مولفه توزیع کننده رویداد (C_d) شده و با اتمام فعالیت مولفه علاقه‌مند خاتمه می‌پذیرد. بنابراین قابلیت اعتماد این سبک از رابطه (۱۱) محاسبه می‌شود. در این رابطه، R_d قابلیت اعتماد مولفه توزیع کننده رویداد و R_i قابلیت اعتماد مولفه علاقه‌مند به رویداد می‌باشد.

$$R = R_d \cdot R_i \quad (۱۱)$$

سبک متقاضی / سرویس‌گر (C/S): در این سبک انجام تراکنش با ارسال درخواست مولفه متقاضی به سرویس‌گر آغاز شده و مولفه‌های سرویس‌گر و مخزن در آن موثر می‌باشند. پس از انجام فرایند مورد نظر توسط سرویس‌گر که معمولاً به تعامل با مولفه مخزن نیاز دارد، نتیجه به مولفه متقاضی ارسال می‌شود. نکته قابل

در مدل Cheung [۱۰]، قابلیت اعتماد نرم‌افزار (R) از رابطه (۷) محاسبه می‌شود.

$$R = (-1)^{k+1} \cdot R_k \frac{|E|}{|I-M|} \quad (۷)$$

$|I-M|$ دترمینان ماتریس (I-M) و $|E|$ دترمینان ماتریس حاصل از حذف آخرین ردیف و اولین ستون ماتریس (I-M) می‌باشد. از آنجا که T استاندارد است، $|I-M| \neq 0$ است. بنابراین قابلیت اعتماد نرم‌افزار دارای جواب می‌باشد. R_k ، قابلیت اعتماد آخرین مولفه در اجراء می‌باشد.

برای ساخت مدل مارکوف برای ارزیابی قابلیت اعتماد، باید سبک معماری نرم‌افزار به یک ماشین حالت تبدیل شود. از قابلیت اعتماد مولفه و احتمال گذار میان مولفه‌ها برای محاسبه احتمال گذار میان حالت‌ها استفاده می‌شود. در موقع نگاشت سبک معماری نرم‌افزار به ماشین حالت، یک حالت، مجموعه غیر تهی از مولفه‌های نرم‌افزاری در حال اجراء در مدت زمان معینی از شروع تا انتهای اجرای مولفه را نشان می‌دهد. گذار حالت زمانی رخ می‌دهد که مدت زمان اجراء مولفه پایان یابد.

۴-۲-۲- محدودیت مدل سازی قابلیت اعتماد مارکوف

فرض مدل مارکوف این است که انتقال کنترل میان مولفه‌های نرم‌افزاری یک فرایند مارکوف است. به عبارت دیگر این مدل برای مدل‌سازی نرم‌افزار با رفتار احتمالی مستقل از تاریخچه قابل استفاده می‌باشد. رفتار قطعی نرم‌افزار به معنی زنجیر اجراء با احتمال گذار ۱ بین هر دو مولفه متوالی می‌باشد. زنجیر اجراء قطعی یکی از خصوصیات زیر را دارد.

- استقلال از تاریخچه اجراء. در این مورد هر مولفه در زنجیر دقیقاً یکبار ظاهر می‌شود.
- وابستگی به تاریخچه اجراء. در این مورد حداقل یک مولفه بیشتر از یکبار اجراء می‌شود.

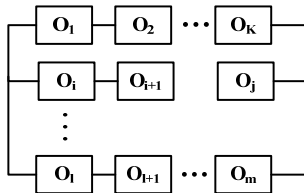
مورد اول، یک فرایند مارکوف است که احتمال همه گذارهایش ۱ است. این مورد با روش‌های معمول قابل مدل‌سازی است. در مورد دوم یک مولفه بیشتر از یک جانشین یا مقدم بلافاصل دارد، و گذار از این مولفه به مولفه بعدی به مولفه قبلی، یعنی تاریخچه اجراء آن بستگی دارد. ویژگی ماتریس تصادفی مارکوف این است که مجموع احتمال گذار از هر مولفه به همه مولفه‌های بعدی بلافاصل آن برابر ۱ می‌باشد. در مورد دوم مولفه‌هایی وجود دارند که این ویژگی را نقض می‌کنند، و زنجیر اجراء، یک فرایند مارکوف نیست. در صورتیکه زنجیر، مستقل از تاریخچه اجراء باشد، برای اجرای یک مولفه همیشه از حالت واحدی استفاده می‌شود. بنابراین تعداد حالت‌ها برای فرایند مارکوف، بدون توجه به تعداد دفعات اجرای هر مولفه، برابر تعداد مولفه‌های اجراء شده می‌باشد. اما در صورتی که زنجیر به تاریخچه اجراء وابسته باشد، برای هر حضور مولفه در زنجیر، به یک شاخص مجزا برای تفکیک نتیجه اثر تاریخچه نیاز است. بنابراین برای نمایش اجراء‌های مختلف مولفه باید از حالت‌های مجزا استفاده کرد [۲۶].

۵- ارزیابی کمی قابلیت اعتماد سبک‌های معماری

در بخش ۴-۱، روش دیگرام بلوکی قابلیت اعتماد (RBD) و در بخش ۴-۲، مدل مبتنی بر حالت مارکوف برای محاسبه قابلیت اعتماد سبک‌های معماری نرم‌افزار معرفی شد. در این بخش به تفکیک این دو روش، قابلیت اعتماد هر یک از

شی‌گزار شامل موردهای کاربری متعددی است و هر مورد کاربری مسیر اجراء خاص خود را دارد. با شکست یک مورد کاربری، سیستم به کار خود با توان عملیاتی کمتر ادامه می‌دهد. بنابراین پیکربندی سبک شی‌گزار مطابق شکل ۱۳ می باشد که هر مسیر، اجراء یک مورد کاربری را نشان می دهد. بنابراین قابلیت اعتماد سبک شی‌گزار از رابطه (۱۵) محاسبه می شود.

$$R_u = \prod_{i=1}^k R_i \quad (14)$$



شکل ۱۳- دیاگرام بلوکی اجراء تراکنش‌های مختلف

$$R = 1 - \prod_{u=1}^p (1 - R_u) \quad (15)$$

در این رابطه، p تعداد موردهای کاربری و R_u قابلیت اعتماد هر مورد کاربری می‌باشد. با جایگذاری R_u از رابطه (۱۴)، رابطه (۱۵) به رابطه (۱۶) تبدیل می‌شود.

$$R = 1 - \prod_{u=1}^p (1 - \prod_{j=1}^k R_j) \quad (16)$$

از آنجا که تعدادی از کلاس‌ها در اغلب موردهای کاربری مشترک می باشند، قابلیت اعتماد سبک شی‌گزار مطابق رابطه (۱۷) می باشد که c تعداد کلاس‌های مشترک در موردهای کاربری می‌باشد.

$$R = \prod_{l=1}^c R_l \cdot (1 - \prod_{u=1}^p (1 - \prod_{j=1}^k R_j)) \quad (17)$$

بررسی قابلیت اعتماد بخش موازی $(1 - \prod_{j=1}^k R_j)$ ، با تعداد مختلف کلاس‌ها نشان داد که قابلیت اعتماد این بخش نزدیک به ۱ است. در نتیجه قابلیت اعتماد سبک شی‌گزار معادل قابلیت اعتماد بخش ترتیبی $(\prod_{l=1}^c R_l)$ می‌باشد. بنابراین قابلیت اعتماد این سبک از رابطه (۱۸) محاسبه می‌شود.

$$R = \prod_{l=1}^c R_l \quad (18)$$

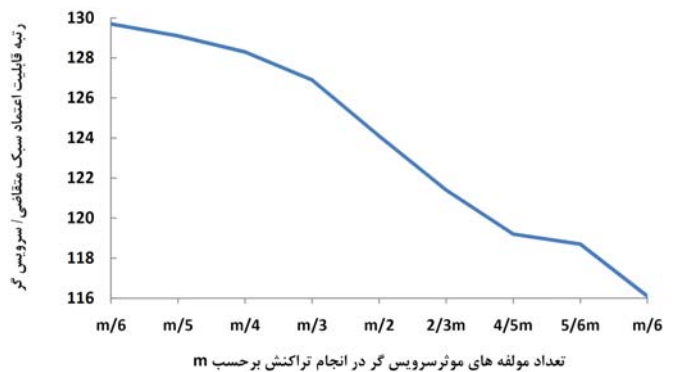
شکل ۱۴، نمودار تغییر رتبه قابلیت اعتماد سبک شی‌گزار برحسب درصد کلاس‌های مشترک در موردهای کاربری برای $n_0=15$ است. مطابق نمودار با افزایش این درصد، رتبه قابلیت اعتماد سبک کاهش می‌یابد. مولفین براساس تجارب خود در تولید سیستم‌های نرم‌افزاری، تعداد کلاس‌های مشترک در موردهای کاربری را ۲۰ درصد تعداد کلاس‌های سیستم $(0.2n_0)$ در نظر گرفتند.

توجه این است که سرویس‌گر، یک مولفه مرکب است و خود شامل چند مولفه است. به عبارت دیگر اجراء سرویس‌گر به منزله اجراء m مولفه می‌باشد. در نتیجه در انجام تراکنش، $m+1$ مولفه، یعنی یکبار مخزن و m مولفه در سرویس‌گر، هر یک در تراکنش موثر می‌باشند. بنابراین قابلیت اعتماد این سبک از رابطه (۱۲) محاسبه می‌شود. در این رابطه، R_{tps} قابلیت اعتماد مولفه مخزن و $R_1 \dots R_m$ قابلیت اعتماد مولفه‌های تشکیل دهنده سرویس‌گر می‌باشد.

$$R = R_m \dots R_2 R_1 R_{tps} \quad (12)$$

در مولفه سرویس‌گر سبک‌های متقاضی/سرویس‌گر و کارگزار، در انجام تراکنش بخشی از m مولفه موثر می‌باشند. شکل ۱۲، نمودار تغییر رتبه قابلیت اعتماد سبک متقاضی/سرویس‌گر را برحسب تعداد مولفه‌هایی از سرویس‌گر که در انجام تراکنش موثر می باشند را برحسب m برای $m=5$ نشان می‌دهد. مطابق نمودار با افزایش تعداد این مولفه‌ها، رتبه قابلیت اعتماد سبک کاهش می‌یابد. این مقدار نیز براساس تجارب مولفین در تولید سیستم‌های نرم‌افزاری، $m/2$ در نظر گرفته شد.

تولیدکنندگان نرم‌افزار براساس تجارب گذشته خود در تولید نرم‌افزار، پارامترهای درصد کلاس‌های مشترک در موردهای کاربری و تعداد مولفه‌های موثر سرویس‌گر در انجام تراکنش را تعیین می‌کنند.



شکل ۱۲- نمودار تغییر رتبه قابلیت اعتماد سبک متقاضی/سرویس‌گر برحسب تعداد مولفه‌های موثر سرویس‌گر در انجام تراکنش برای $m=5$

سبک کارگزار (BRK): در سبک کارگزار انجام تراکنش با ارسال درخواست مولفه متقاضی، آغاز شده و مولفه‌های کارگزار (C_{brk}) ، نماینده سمت متقاضی (C_{esp}) و نماینده سمت سرویس‌گر (C_{ssp}) ، سرویس‌گر (C_s) و مخزن (C_{tps}) در انجام تراکنش موثر می‌باشند.

مشابه سبک متقاضی/سرویس‌گر اجراء سرویس‌گر به منزله اجراء $m/2$ مولفه است. در نتیجه در انجام تراکنش، $m/2+7$ مولفه، موثر می‌باشند. بنابراین قابلیت اعتماد این سبک از رابطه (۱۳) محاسبه می‌شود. در این رابطه R_{esp} ، قابلیت اعتماد مولفه نماینده سمت متقاضی، R_{brk} قابلیت اعتماد مولفه کارگزار، R_{ssp} قابلیت اعتماد مولفه نماینده سمت سرویس‌گر و R_{tps} قابلیت اعتماد مولفه مخزن می‌باشد.

$$R = R_{esp} R_{brk} R_{ssp} R_{m/2} \dots R_2 R_1 R_{tps} \quad (13)$$

سبک شی‌گزار (OO): در سبک شی‌گزار (شکل ۸)، به ازای k کلاس موثر در مورد کاربری، قابلیت اعتماد مورد کاربری از رابطه (۱۴) محاسبه می‌شود. سبک

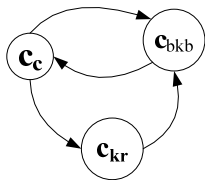
احتمال گذار از هر حالت به حالت بعدی (P_{ij}) قطعی و برابر ۱ است. ماتریس گذار M و $|I - M|$ به صورت زیر می‌باشند.

$$M = \begin{bmatrix} 0 & R_i \\ 0 & 0 \end{bmatrix}, \quad |I - M| = \begin{vmatrix} 1 & -R_i \\ 0 & 1 \end{vmatrix} = 1$$

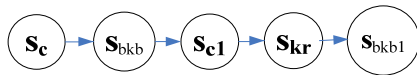
رابطه (۱۹) محاسبه می‌شود. براساس رابطه (۷)، قابلیت اعتماد سبک مخزن از رابطه (۱۹) محاسبه می‌شود.

$$R = R_i \cdot R_{tps} \quad (19)$$

سبک **تخته سیاه (BKB)**: مطابق شکل ۱۶، مولفه کنترل براساس وضعیت تخته سیاه، متخصص مناسب را فراخوانی می‌کند. متخصص فراخوانی شده با تخته سیاه تعامل می‌کند. احتمال از مولفه کنترل به مولفه تخته سیاه و متخصص ها هر یک برابر ۱ می‌باشد. مولفه بعدی مولفه کنترل از نظر اجراء، در موقع بررسی وضعیت مولفه تخته سیاه، مولفه تخته سیاه و بعد از بررسی وضعیت تخته سیاه، یک متخصص است. بنابراین ماشین حالت معادل این سبک، شرط مدل مارکوف را ندارد، و رفتار وابسته به تاریخچه دارد.



شکل ۱۶- تعامل مولفه‌ها در سبک تخته سیاه



شکل ۱۷- ماشین حالت معادل سبک تخته سیاه

در نتیجه باید برای هر احتمال گذار از هر حالت به حالت بعدی (P_{ij}) قطعی و برابر ۱ می‌باشد. ماتریس گذار M و $|I - M|$ به صورت زیر می‌باشند.

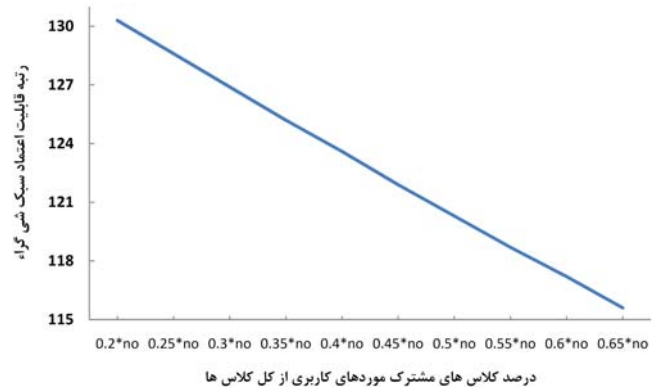
$$M = \begin{bmatrix} 0 & R_c & 0 & 0 & 0 \\ 0 & 0 & R_{bkb} & 0 & 0 \\ 0 & 0 & 0 & R_{c1} & 0 \\ 0 & 0 & 0 & 0 & R_{kr} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad |I - M| = \begin{vmatrix} 1 & -R_c & 0 & 0 & 0 \\ 0 & 1 & -R_{bkb} & 0 & 0 \\ 0 & 0 & 1 & -R_{c1} & 0 \\ 0 & 0 & 0 & 1 & -R_{kr} \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} = 1$$

سبک از رابطه زیر محاسبه می‌شود. $R_K = R_{bkb1}$ و $|E| = R_c R_{bkb} R_{c1} R_{kr}$ براساس رابطه (۷)، قابلیت اعتماد این سبک از رابطه زیر محاسبه می‌شود.

$$R = (-1)^{5+1} \cdot R_{bkb1} \frac{R_c R_{bkb} R_{c1} R_{kr}}{1} = R_c R_{bkb} R_{c1} R_{kr} R_{bkb1}$$

با توجه به اینکه مقادیر R_{bkb} و R_{bkb1} ، R_c و R_{c1} یکسان فرض می‌شود، قابلیت اعتماد این سبک از رابطه (۲۰) محاسبه می‌شود.

$$R = R_c^2 R_{bkb}^2 R_{kr} \quad (20)$$



شکل ۱۴- نمودار تغییر رتبه قابلیت اعتماد سبک شی گراء بر حسب درصد کلاس های مشترک مورد های کاربری برای no=15

جدول ۱، قابلیت اعتماد هر یک از سبک‌های معماری نرم‌افزار را براساس روش RBD نشان می‌دهد.

جدول ۱- قابلیت اعتماد سبک‌های معماری نرم‌افزار براساس روش RBD

نام سبک	نماد	قابلیت اعتماد
مخزن	RPS	$R_i \cdot R_{tps}$
تخته سیاه	BKB	$R_c R_{bkb} R_{kr}$
لوله و فیلتر	P/F	$\prod_{i=1}^m R_i$
لایه‌ای	LYD	$\prod_{i=1}^m R_i$
فراخوانی ضمنی	I/I	$R_d R_i$
متقاضی / سرویس گر	C/S	$R_{m/2} \dots R_2 R_1 R_{tps}$
کارگزار	BRK	$R_{csp} R_{brk} R_{ssp} R_{m/2} \dots R_2 R_1 R_{tps}$
شی گراء	OO	$\prod_{l=1}^c R_l$

۲-۵- سنجش سبک‌های معماری نرم‌افزار براساس مدل مارکوف

در این بخش با استفاده از روش Cheung [۱۰]، که مبتنی بر مدل مارکوف است، قابلیت اعتماد سبک‌های معماری نرم‌افزار را مورد سنجش قرار می‌دهیم.

سبک مخزن (RPS): در سبک مخزن، مولفه‌های مستقل با مخزن از طریق خواندن/نوشتن از/در مخزن، تعامل دارند. درخواست مولفه c_i به مولفه مخزن ارسال و مولفه مخزن درخواست را انجام می‌دهد. ماشین حالت معادل این سبک با توجه به تعامل یک مولفه در انجام یک تراکنش به شکل ۱۵ است. s_i حالت مولفه c_i در موقع ارسال درخواست به مولفه مخزن و s_{tps} حالت معادل مولفه مخزن می‌باشد.



شکل ۱۵- ماشین حالت معادل سبک مخزن

سبک لوله و فیلتر (P/F): شکل ۱۸، ماشین حالت معادل این سبک را با ۳ مولفه نشان می‌دهد. هر حالت متناظر با اجرای یک مولفه (فیلتر) است. یک گذار حالت زمانی رخ می‌دهد که اجرای یک فیلتر کامل شده و کنترل به فیلتر بعدی



شکل ۱۸- ماشین حالت معادل سبک لوله و فیلتر

محول می‌گردد. گذار از هر حالت به حالت بعدی (P_{ij}) قطعی و برابر ۱ می‌باشد، در نتیجه ماتریس گذار M و |I - M| به صورت زیر می‌باشد.

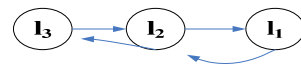
$$M = \begin{bmatrix} 0 & R_1 & 0 \\ 0 & 0 & R_2 \\ 0 & 0 & 0 \end{bmatrix}, \quad |I - M| = \begin{vmatrix} 1 & -R_1 & 0 \\ 0 & 1 & -R_2 \\ 0 & 0 & 1 \end{vmatrix} = 1$$

رابطه (۲۱) محاسبه می‌شود. |E|=R₁R₂ و R_K=R₃ براساس رابطه (۷)، قابلیت اعتماد این سبک از رابطه

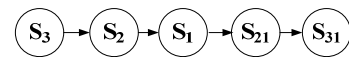
$$R = (-1)^{3+1} \cdot R_3 \frac{R_1 R_2}{1} = R_1 R_2 R_3 \quad (21)$$

در حالت کلی و به ازاء m مولفه موثر در تراکنش، قابلیت اعتماد این سبک از رابطه $\prod_{i=1}^m R_i$ محاسبه می‌شود.

سبک لایه‌ای (LYD): مطابق شکل ۱۹ که نحوه تعامل ۳ مولفه در این سبک را نشان می‌دهد، احتمال گذار از مولفه I₂ به مولفه‌های I₁ و I₃ هر یک برابر ۱ می‌باشد. با توجه به اینکه مولفه بعدی مولفه I₂ از نظر اجراء، یکبار مولفه I₁ و یکبار مولفه I₃ است. بنابراین ماشین حالت معادل این سبک، شرط مدل مارکوف را نداشته و رفتار وابسته به تاریخچه دارد. در نتیجه برای هر اجرای مولفه‌ها، باید یک حالت مجزا در نظر گرفت. شکل ۲۰، ماشین حالت معادل شکل ۱۹ است. حالت S₂₁ بعد از برگشت کنترل از مولفه I₁ به مولفه I₂ و حالت S₃₁ بعد از برگشت کنترل از مولفه I₂ به مولفه I₃ می‌باشد.



شکل ۱۹- سبک لایه‌ای با سه لایه



شکل ۲۰- ماشین حالت معادل سبک لایه‌ای

گذار از هر حالت به حالت بعدی (P_{ij}) قطعی و برابر ۱ می‌باشد، در نتیجه ماتریس گذار M و |I - M| به صورت زیر می‌باشد.

$$M = \begin{bmatrix} 0 & R_3 & 0 & 0 & 0 \\ 0 & 0 & R_2 & 0 & 0 \\ 0 & 0 & 0 & R_1 & 0 \\ 0 & 0 & 0 & 0 & R_{21} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad |I - M| = \begin{vmatrix} 1 & -R_3 & 0 & 0 & 0 \\ 0 & 1 & -R_2 & 0 & 0 \\ 0 & 0 & 1 & -R_1 & 0 \\ 0 & 0 & 0 & 1 & -R_{21} \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} = 1$$

$$R = (-1)^{5+1} \cdot R_{31} \frac{R_3 R_2 R_1 R_{21}}{1} = R_3 R_2 R_1 R_{21} R_{31}$$

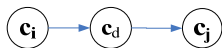
با توجه به اینکه قابلیت اعتماد مولفه‌ها در حالت‌های مختلف یکسان در نظر گرفته می‌شود، بنابراین R₂₁ با R₂ و R₃₁ با R₃ یکسان می‌باشد و قابلیت اعتماد این سبک از رابطه (۲۲) محاسبه می‌شود.

$$R = R_3^2 R_2^2 R_1 \quad (22)$$

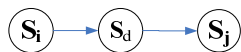
در حالت کلی و به ازاء m مولفه موثر در تراکنش، قابلیت اعتماد این سبک از رابطه (۲۳) محاسبه می‌شود.

$$R = R_m^2 \dots R_2^2 R_1 \quad (23)$$

سبک فراخوانی ضمنی (I/I): با توجه به مفهوم تراکنش در این سبک، وقوع رویداد توسط مولفه C_i، منجر به فعال‌سازی مولفه علاقه‌مند (C_j) توسط مولفه توزیع کننده رویداد (C_d) شده و با اتمام فعالیت مولفه علاقه‌مند تراکنش خاتمه می‌یابد. بنابراین تعامل مولفه‌ها به شکل ۲۱ و ماشین حالت معادل آن به شکل ۲۲ می‌باشد.



شکل ۲۱- تعامل مولفه‌ها در سبک فراخوانی ضمنی



شکل ۲۲- ماشین حالت معادل سبک فراخوانی ضمنی

احتمال گذار از هر حالت به حالت بعدی (P_{ij}) قطعی و برابر ۱ می‌باشد. ماتریس گذار M و |I - M| به صورت زیر می‌باشد.

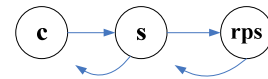
$$M = \begin{bmatrix} 0 & R_i & 0 \\ 0 & 0 & R_d \\ 0 & 0 & 0 \end{bmatrix}, \quad |I - M| = \begin{vmatrix} 1 & -R_i & 0 \\ 0 & 1 & -R_d \\ 0 & 0 & 1 \end{vmatrix} = 1$$

محاسبه می‌شود. |E|=R_iR_d و R_K=R_j براساس رابطه (۷)، قابلیت اعتماد از رابطه (۲۴)

$$R = R_i R_d R_j \quad (24)$$

سبک متقاضی/سرویس‌گر (C/S): تعامل مولفه‌ها در این سبک به شکل ۲۳ می‌باشد. احتمال گذار از مولفه سرویس‌گر (S) به مولفه متقاضی (C) و مخزن (RPS) هر یک برابر ۱ می‌باشد. مطابق شکل، مولفه بعدی سرویس‌گر از نظر اجراء یکبار مولفه مخزن و یکبار مولفه متقاضی است. بنابراین این سبک، شرط مدل مارکوف را نداشته و رفتار وابسته به تاریخچه دارد. در نتیجه برای هر اجرای مولفه، باید یک حالت مجزا در نظر گرفت. بنابراین ماشین حالت معادل آن به شکل ۲۴ می‌باشد.

مطابق نتایج سنجش این سبک به روش RBD، $m/2$ از مولفه‌های سرویس‌گر در تراکنش موثر می‌باشند، قابلیت اعتماد این سبک از رابطه (۲۸) محاسبه می‌شود.



شکل ۲۳- تعامل مولفه‌ها در سبک متقاضی/ سرویس‌گر

$$R = R_c^2 R_{csp}^2 R_{brk}^2 R_{ssp}^2 R_{m/2}^2 \dots R_2^2 R_1^2 R_{rps} \quad (28)$$

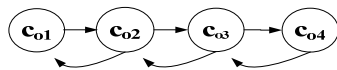
سبک شی گراء (OO): تعامل ۴ کلاس موثر در انجام مورد کاربری به شکل ۲۷ است. مطابق شکل، کلاس بعدی کلاس C_{03} از نظر اجراء یکبار کلاس C_{04} و یکبار کلاس C_{02} است. بنابراین این سبک نیز شرط مدل مارکوف را نداشته و رفتار وابسته به تاریخچه دارد. با توجه به اینمورد و تحلیل قابلیت اعتماد این سبک به روش RBD، کلاس مشترک در انجام مورد کاربری موثر می‌باشند و براساس رابطه (۲۹)، قابلیت اعتماد آن از رابطه (۲۹) محاسبه می‌شود.



شکل ۲۴- ماشین حالت معادل سبک متقاضی/ سرویس‌گر

احتمال گذار از هر حالت به حالت بعدی (P_{ij}) قطعی و برابر ۱ است. بنابراین ماتریس گذار M و $|I-M|$ به صورت زیر می‌باشد.

$$M = \begin{bmatrix} 0 & R_c & 0 & 0 & 0 \\ 0 & 0 & R_s & 0 & 0 \\ 0 & 0 & 0 & R_{rps} & 0 \\ 0 & 0 & 0 & 0 & R_s \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad |I-M| = \begin{bmatrix} 1 & -R_c & 0 & 0 & 0 \\ 0 & 1 & -R_s & 0 & 0 \\ 0 & 0 & 1 & -R_{rps} & 0 \\ 0 & 0 & 0 & 1 & -R_s \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = 1$$



شکل ۲۷- تعامل کلاس‌ها در انجام تراکنش

$$R = R_1^2 R_2^2 \dots R_{c-1}^2 R_c \quad (29)$$

از رابطه (۲۵) محاسبه می‌شود. $|E| = R_c R_s R_{rps} R_s$ و $R_K = R_c$ در نتیجه قابلیت اعتماد این سبک براساس رابطه (۲۵)، از رابطه (۲۵) محاسبه می‌شود.

جدول ۲، رابطه قابلیت اعتماد سبک‌های معماری براساس مدل مارکوف را نشان می‌دهد.

$$R = (-1)^{5+1} \cdot R_c \frac{R_c R_s R_{rps} R_s}{1} = R_c^2 R_s^2 R_{rps} \quad (25)$$

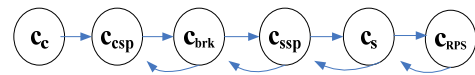
جدول ۲- قابلیت اعتماد سبک‌های معماری نرم‌افزار براساس مدل مارکوف

نام سبک	نماد	قابلیت اعتماد
مخزن	RPS	$R_i R_{rps}$
تخته سیاه	BKB	$R_c^2 R_{bkb}^2 R_{kr}$
لوله و فیلتر	P/F	$\prod_{i=1}^m R_i$
لایه‌ای	LYD	$R_{m-1}^2 \dots R_2^2 R_1$
فراخوانی ضمنی	I/I	$R_i R_d R_j$
متقاضی/ سرویس‌گر	C/S	$R_c^2 R_{m/2}^2 \dots R_2^2 R_1^2 R_{rps}$
کارگزار	BRK	$R_c^2 R_{csp}^2 R_{brk}^2 R_{ssp}^2 R_{m/2}^2 \dots R_2^2 R_1^2 R_{rps}$
شی گراء	OO	$R_1^2 R_2^2 \dots R_{c-1}^2 R_c$

مطابق نتایج سنجش این سبک به روش RBD، $m/2$ از مولفه‌های سرویس‌گر در تراکنش موثر می‌باشند. بنابراین قابلیت اعتماد این سبک از رابطه (۲۶) محاسبه می‌شود.

$$R = R_c^2 R_{m/2}^2 \dots R_2^2 R_1^2 R_{rps} \quad (26)$$

سبک کارگزار (BRK): در سبک کارگزار، تعامل مولفه‌ها به شکل ۲۵ می‌باشد. احتمال گذار از مولفه کارگزار به سرویس‌گر و متقاضی، هریک برابر ۱ می‌باشد. مطابق این شکل و مشابه سبک متقاضی/سرویس‌گر، این سبک، شرط مدل مارکوف را نداشته و رفتار وابسته به تاریخچه دارد. در نتیجه برای هر اجرای مولفه، باید یک حالت مجزا در نظر گرفت. بنابراین ماشین حالت این سبک به شکل ۲۶ می‌باشد.



شکل ۲۵- تعامل مولفه‌ها در سبک کارگزار

۵-۳- مقایسه نتایج ارزیابی دو رویکرد

در بخش ۵-۱ و ۵-۲، سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد براساس روش دیاگرام بلوکی قابلیت اعتماد و روش مبتنی بر حالت مارکوف ارزیابی شدند و قابلیت اعتماد آنها محاسبه شد. بررسی ارزش قابلیت اعتماد در دو روش نشان می‌دهد که:



شکل ۲۶- ماشین حالت معادل سبک کارگزار

۱- در رابطه قابلیت اعتماد سبک‌های متقاضی/سرویس‌گر، کارگزار و فراخوانی ضمنی، در روش مبتنی بر حالت مارکوف، قابلیت اعتماد مولفه متقاضی نیز موثر است. یادآوری می‌گردد، منشاء مدل مبتنی بر حالت مارکوف، روش Cheung است، که به مدل کاربرگرا معروف است، یعنی دریافت نتیجه برای کاربر اهمیت دارد. با این دیدگاه، مولفه متقاضی نیز در قابلیت اعتماد اثر دارد.

براساس رابطه (۲۷)، قابلیت اعتماد این سبک از رابطه (۲۷) محاسبه می‌شود.

$$R = R_c^2 R_{csp}^2 R_{brk}^2 R_{ssp}^2 R_s^2 R_{rps} \quad (27)$$

جدول ۴- مقادیر قابلیت اعتماد سبک‌های معماری نرم‌افزار براساس مدل مارکوف

نماد سبک	m=7	m=6	m=5	m=4	m=3	m=2
	n ₀ =21	n ₀ =18	n ₀ =15	n ₀ =12	n ₀ =9	n ₀ =6
RPS	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶
BKB	۰.۹۲	۰.۹۲	۰.۹۲	۰.۹۲	۰.۹۲	۰.۹۲
P/F	۰.۸۷	۰.۸۹	۰.۹	۰.۹۲	۰.۹۴	۰.۹۶
LYD	۰.۸۲	۰.۸۴	۰.۸۷	۰.۸۹	۰.۹۲	۰.۹۵
I/I	۰.۹۴	۰.۹۴	۰.۹۴	۰.۹۴	۰.۹۴	۰.۹۴
C/S	۰.۸۶	۰.۸۷	۰.۸۸	۰.۸۹	۰.۹۱	۰.۹۲
BRK	۰.۷۸	۰.۷۹	۰.۸۱	۰.۸۲	۰.۸۳	۰.۸۴
OO	۰.۹۱	۰.۹۲	۰.۹۳	۰.۹۴	۰.۹۵	۰.۹۷

برای رتبه‌بندی گزینه‌ها از دید یک معیار در روش تصمیم‌گیری چند معیاره فرآیند تحلیل سلسله‌مراتبی^{۱۸}، از نرمال سازی مقادیر معیار استفاده می‌شود. در اینجا نیز برای محاسبه رتبه سبک‌های معماری از دید قابلیت اعتماد از این شیوه استفاده می‌کنیم. بنابراین رتبه قابلیت اعتماد سبک‌های معماری، در دو رویکرد، از رابطه (۳۰) محاسبه و در جدول‌های ۵ و ۶ ارائه شده است.

در این رابطه، x_{ij} ارزش قابلیت اعتماد سبک در ردیف I-ام و ستون j-ام جدول (۴)۳، و مخرج رابطه مجموع ارزش قابلیت اعتماد ستون j-ام است. از ضریب ۱۰۰۰ در رابطه (۳۰) برای وضوح رتبه سبک‌ها استفاده شده است.

$$R_{ij} = \frac{x_{ij}}{\sum_{i=1}^s x_{ij}} * 1000 \quad (30)$$

ضمن اینکه رتبه قابلیت اعتماد سبک‌های معماری در دو رویکرد اندکی متفاوت است، و این اختلاف ناشی از تفاوت رابطه قابلیت اعتماد سبک‌های معماری در دو رویکرد است، با اینحال روند تغییر رتبه سبک‌های معماری با توجه به تغییر تعداد مولفه‌های موثر در انجام تراکنش در دو رویکرد یکسان است. روش مبتنی بر حالت مارکوف تمایز بیشتری را میان رتبه قابلیت اعتماد سبک‌های معماری نشان می‌دهد.

جدول ۵- رتبه سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد براساس روش RBD

نماد سبک	m=7	m=6	m=5	m=4	m=3	m=2
	n ₀ =21	n ₀ =18	n ₀ =15	n ₀ =12	n ₀ =9	n ₀ =6
RPS	۱۳۱.۷	۱۳۰.۴	۱۳۱.۷	۱۲۸.۳	۱۲۷.۳	۱۲۶
BKB	۱۲۸.۹	۱۲۷.۷	۱۲۸.۹	۱۲۵.۷	۱۲۴.۷	۱۲۳.۴
P/F	۱۱۹.۳	۱۲۰.۹	۱۲۳.۵	۱۲۳	۱۲۴.۷	۱۲۶
LYD	۱۱۹.۳	۱۲۰.۹	۱۲۳.۵	۱۲۳	۱۲۴.۷	۱۲۶
I/I	۱۳۱.۷	۱۳۰.۴	۱۳۱.۷	۱۲۸.۳	۱۲۷.۳	۱۲۶
C/S	۱۲۴.۸	۱۲۵	۱۲۷.۶	۱۲۵.۷	۱۲۶	۱۲۶
BRK	۱۱۸	۱۱۸.۲	۱۲۰.۷	۱۱۹	۱۱۸	۱۱۸.۱
OO	۱۲۶.۲	۱۲۶.۴	۱۱۲.۵	۱۲۷	۱۲۷.۳	۱۲۸.۶

شکل ۲۸، نمودار تغییر رتبه قابلیت اعتماد سبک‌های معماری را برحسب تعداد مولفه‌های موثر در انجام تراکنش در رویکرد مبتنی بر حالت مارکوف نشان می‌دهد. مطابق این نمودار، با افزایش تعداد مولفه‌های موثر در انجام تراکنش رتبه

۲- روش RBD، قابلیت اعتماد سبک‌های معماری را براساس مولفه‌های موثر در انجام تراکنش محاسبه می‌کند، اما به تعداد دفعات حضور این مولفه‌ها در انجام تراکنش توجه ندارد. روش مبتنی بر حالت مارکوف، ماشین حالت معادل هر سبک معماری را در نظر گرفته و براساس آن قابلیت اعتماد هر سبک معماری محاسبه می‌شود.

در نتیجه به تعداد دفعات حضور هر مولفه در انجام تراکنش توجه دارد. در اغلب سبک‌های معماری مانند سبک لایه‌ای، حضور مجدد مولفه در انجام تراکنش اغلب به سبب برگشت از فراخوانی است و در عمل مولفه کار خاصی انجام نمی‌دهد و زمان اندکی صرف برگشت از فراخوانی می‌شود. در نتیجه قابلیت اعتماد مولفه‌ها در این حالت ۰.۹۹ در نظر می‌گیریم.

۴-۵- ارزیابی سیستم‌های بزرگ

در دسته‌ای از سیستم‌های نرم‌افزاری نه تنها از یک سبک معماری به عنوان مبنای اصلی ساختاردهی به سیستم استفاده می‌شود، بلکه هر یک از اجزاء آنها نیز از سبک معماری خاصی استفاده می‌کنند. با توجه به تعیین قابلیت اعتماد سبک‌های معماری در این تحقیق، امکان ارزیابی و تعیین قابلیت اعتماد چنین سیستم‌هایی نیز وجود دارد.

۶- رتبه‌بندی سبک‌های معماری نرم‌افزار

در بخش پنجم، قابلیت اعتماد سبک‌های معماری نرم‌افزار محاسبه گردید. در رابطه قابلیت اعتماد اغلب سبک‌های معماری نرم‌افزار، پارامتر m که نشانگر تعداد مولفه‌های موثر در انجام تراکنش است، وجود دارد. بنابراین به منظور بررسی تاثیر تعداد مولفه‌های موثر در انجام تراکنش در قابلیت اعتماد سبک‌های معماری نرم‌افزار، مقدار m، ۲، ۳، ۴، ۵، ۶ و ۷ در نظر گرفته می‌شود. مولفه‌های سبک شی‌گرا، یعنی کلاس‌ها نسبت به مولفه‌های سایر سبک‌های معماری ریزدانه‌تر می‌باشند، با این تقریب که به طور متوسط در هر مولفه ۳ کلاس در انجام تراکنش موثر باشند، متناظر با تعداد مولفه‌های موثر در انجام تراکنش، در سایر سبک‌های معماری، تعداد کلاس‌های موثر در انجام تراکنش ۶، ۹، ۱۲، ۱۵، ۱۸ و ۲۱ در نظر گرفته می‌شود. قابلیت اعتماد هر مولفه/کلاس را ۰.۹۸ در نظر می‌گیریم.

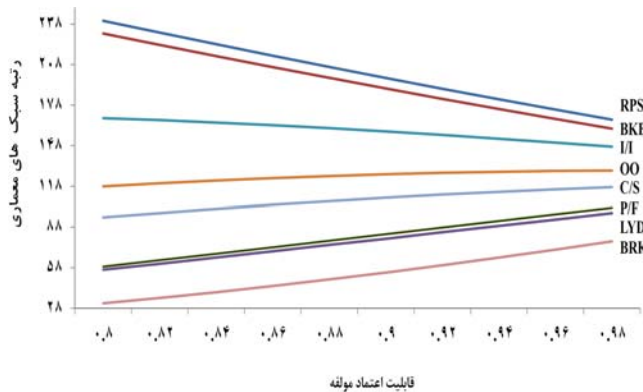
به ازاء تعداد مختلف مولفه‌های موثر در انجام تراکنش و براساس رابطه قابلیت اعتماد سبک‌های معماری در جدول‌های ۱ و ۲، و با توجه به مقدار قابلیت اعتماد هر مولفه/کلاس، مقادیر قابلیت اعتماد هر یک از سبک‌های معماری نرم‌افزار در دو رویکرد محاسبه و در جدول‌های ۳ و ۴ درج شده است.

جدول ۳- مقادیر قابلیت اعتماد سبک‌های معماری نرم‌افزار براساس روش RBD

نماد سبک	m=7	m=6	m=5	m=4	m=3	m=2
	n ₀ =21	n ₀ =18	n ₀ =15	n ₀ =12	n ₀ =9	n ₀ =6
RPS	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶
BKB	۰.۹۴	۰.۹۴	۰.۹۴	۰.۹۴	۰.۹۴	۰.۹۴
P/F	۰.۸۷	۰.۸۹	۰.۹	۰.۹۲	۰.۹۴	۰.۹۶
LYD	۰.۸۷	۰.۸۹	۰.۹	۰.۹۲	۰.۹۴	۰.۹۶
I/I	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶	۰.۹۶
C/S	۰.۹۱	۰.۹۲	۰.۹۳	۰.۹۴	۰.۹۵	۰.۹۶
BRK	۰.۸۶	۰.۸۷	۰.۸۸	۰.۸۹	۰.۸۹	۰.۹
OO	۰.۹۲	۰.۹۳	۰.۸۲	۰.۹۵	۰.۹۶	۰.۹۸

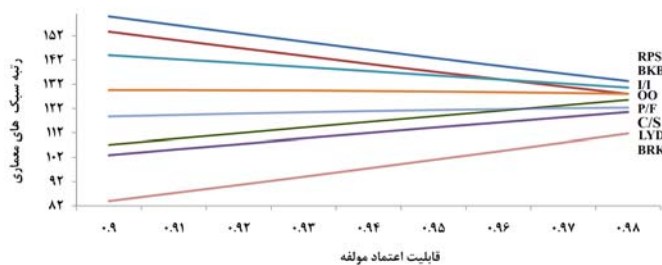
در نظر می‌گیریم.

در این بخش با تغییر قابلیت اعتماد مولفه‌های طراحی شده از ۰.۸ تا ۰.۹۸، تاثیر این تغییر را بر قابلیت اعتماد سبک‌های معماری به ازاء $m=5$ بررسی می‌کنیم. مطابق نمودار شکل ۲۹، با افزایش قابلیت اعتماد مولفه‌های طراحی شده، تفاوت رتبه قابلیت اعتماد سبک‌های معماری کمتر می‌شود. بجز $R=0.8$ ، سبک کارگزار کمترین و سبک مخزن بالاترین رتبه را از نظر قابلیت اعتماد دارد. با افزایش قابلیت اعتماد مولفه‌های طراحی شده، روند تغییر رتبه قابلیت اعتماد سبک‌های مخزن، تخته سیاه و فراخوانی ضمنی کاهشی و روند تغییر رتبه قابلیت اعتماد سایر سبک‌های معماری افزایشی است.



شکل ۲۹- نمودار تغییر رتبه سبک‌های معماری برحسب تغییر قابلیت اعتماد مولفه‌های طراحی شده

به منظور نمایش واضح تر تغییر رتبه سبک‌های معماری، در نمودار شکل ۳۰، روند تغییر قابلیت اعتماد مولفه‌های طراحی شده در محدوده ۰.۹ تا ۰.۹۸ نشان داده شده است.



شکل ۳۰- نمودار تغییر رتبه سبک‌های معماری برحسب تغییر قابلیت اعتماد مولفه‌های طراحی شده

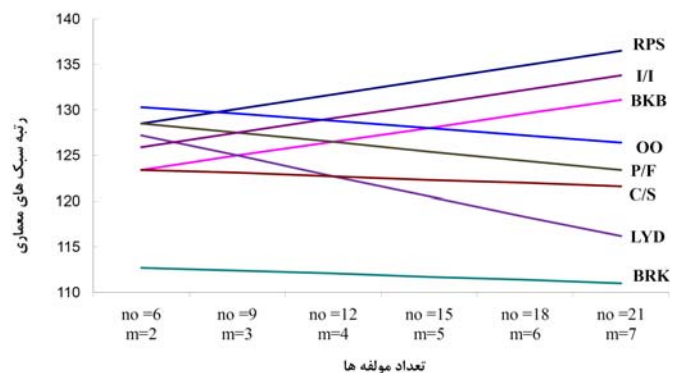
۷- بررسی موردی

سیستم نرم‌افزاری مورد بررسی، سیستم کنترل ترافیک هوایی^{۱۹} (ATC) [۳] می‌باشد. سیستم کنترل ترافیک هوایی تضمین حرکت ایمن هواپیما در خطوط هوایی را بر عهده دارد. هواپیماها در هر پرواز با تعدادی موجودیت ATC سرو کار دارند که اطلاعات خود را به مرکز کنترل ترافیک هوایی ارسال می‌کنند. این سیستم باید در فاصله زمانی تعیین شده واکنش نشان دهد. بنابراین از نوع بیدرنگ سخت^{۲۰} است و حساس به ایمنی تلقی می‌شود، زیرا زندگی انسان‌ها به آن بستگی دارد. بنابراین چنین سیستمی باید از قابلیت اعتماد بالایی برخوردار باشد. با فرض آنکه تعداد مولفه‌های موثر در انجام هر تراکنش را ۴ مولفه در نظر بگیریم، و با توجه به اینکه قابلیت اعتماد مولفه‌ها در چنین سیستمی باید بالا باشد، مطابق

سبک‌های مخزن (RPS)، فراخوانی ضمنی (I/I)، و تخته سیاه (BKB) افزایش و رتبه سایر سبک‌های معماری کاهش می‌یابد. از نظر شهودی این نتایج درست است. در سبک‌های مخزن (RPS)، فراخوانی ضمنی (I/I)، و تخته سیاه (BKB) تعداد مولفه‌های موثر در انجام تراکنش، ثابت است. در سایر سبک‌های معماری مانند سبک‌های لایه‌ای (LYD)، لوله و فیلتر (P/F)، با افزایش تعداد مولفه‌های موثر در انجام تراکنش، رتبه قابلیت اعتماد سبک‌ها کاهش می‌یابد. در سبک کارگزار (BRK) و سبک متقاضی/ سرویس‌گر (C/S) علاوه بر m مولفه، مولفه‌های ثابت دیگری در تراکنش موثر می‌باشند که باعث کاهش رتبه قابلیت اعتماد آنها می‌گردد.

جدول ۶- رتبه سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد براساس مدل مارکوف

نماد سبک	m=7 n ₀ =21	m=6 n ₀ =18	m=5 n ₀ =15	m=4 n ₀ =12	m=3 n ₀ =9	m=2 n ₀ =6
RPS	۱۳۶.۵	۱۳۴.۹	۱۳۳.۳	۱۳۱.۷	۱۳۰.۱	۱۲۸.۵
BKB	۱۳۱.۱	۱۲۹.۶	۱۲۸	۱۲۶.۵	۱۲۵	۱۲۳.۴
P/F	۱۲۳.۴	۱۲۴.۴	۱۲۵.۴	۱۲۶.۵	۱۲۷.۵	۱۲۸.۵
LYD	۱۱۶.۲	۱۱۸.۳	۱۲۰.۵	۱۲۲.۷	۱۲۵	۱۲۷.۲
I/I	۱۳۳.۸	۱۳۲.۲	۱۳۰.۶	۱۲۹.۱	۱۲۷.۵	۱۲۵.۹
C/S	۱۲۱.۶	۱۲۲	۱۲۲.۳	۱۲۲.۷	۱۲۳.۱	۱۲۳.۴
BRK	۱۱۱	۱۱۱.۴	۱۱۱.۷	۱۱۲.۱	۱۱۲.۴	۱۱۲.۷
OO	۱۲۶.۴	۱۲۷.۲	۱۲۸	۱۲۸.۸	۱۲۹.۶	۱۳۰.۳



شکل ۲۸- نمودار قابلیت اعتماد سبک‌های معماری نرم‌افزار برحسب تعداد مولفه‌های موثر در تراکنش در رویکرد مبتنی بر حالت مارکوف

۶-۱- تحلیل رتبه قابلیت اعتماد سبک‌ها با توجه به قابلیت اعتماد مولفه‌های طراحی شده

در اغلب سبک‌های معماری، وظایف برخی از مولفه‌ها بدون توجه به نرم‌افزار خاص مشخص است. سایر مولفه‌ها با توجه به وظایف نرم‌افزار خاص طراحی می‌شوند. به این ترتیب مولفه‌های هر سبک معماری به دو نوع با وظایف مشخص و مولفه‌های طراحی شده تقسیم می‌شوند. بطور مثال در سبک کارگزار، مولفه‌های کارگزار، نماینده سمت متقاضی و نماینده سمت سرویس‌گر و مخزن از نوع با وظایف مشخص و سایر مولفه‌ها از نوع طراحی شده می‌باشند. قابلیت اعتماد مولفه‌های با وظایف مشخص را ۰.۹۸ در نظر می‌گیریم. همانگونه که در بخش ۵-۳ اشاره گردید، قابلیت اعتماد هر مولفه در زمان برگشت از فراخوانی مولفه دیگر، را ۰.۹۹

[2] A. G. Jansen, and J. Bosch, "Software Architecture as a set of Architectural Design Decisions," *Proc. 5th IEEE/IFIP Working Conf. on Software Architecture*, pp. 109–119, 2005.

[3] L. Bass, P. Clements, and R. Kazman, *Software Architecture in Practice*, 2nd Edition, Addison-Wesley, 2003.

[4] L. Chung, B. A. Nixon, E. Yu, and J. Mylopoulos, *Non-Functional Requirements in Software Engineering*, Kluwer Academic Publishing, 2000.

[5] F. Losavio, L. Chirinos, A. Matteo, N. Lévy, and A. Ramdane-Cherif, "ISO Quality standards for measuring architectures," *The Journal of System and Software*, Vol. 72, No. 2, pp. 209-223, 2004.

[6] J. Shao, and BH. Far, "Development of an Intelligent System for Architecture Design and Analysis," *Proc. Of the Canadian Electrical and Computer Engineering Conference*, PP.539-542, 2004.

[7] H. Reza, and E. Grant, "Quality-Oriented Software Architecture," *Proc. of the IEEE Conf on Information Technology: Coding and Computing*, pp. 140- 145, 2005.

[8] M. AlSharif, W. P. Bond, and T. Al-Otaiby, "Assessing the complexity of software architecture," *Proc. of the 42nd annual ACM Southeast regional conference*, pp. 98–103, 2004.

[9] N. B. Harrison, and P. Avgeriou, "Leveraging Architecture Patterns to Satisfy Quality Attributes," *First European Conf. on Software Architecture*, pp. 263-270, 2007.

[10] R. C. Cheung, "A user-oriented software reliability model," *IEEE Trans. on Software Engineering*, Vol.6, No. 2, pp. 118–125, 1980.

[11] Y. Meng-Lai, C. L. Hyde, and L. E. James, "A Petri-net approach for early-stage system-level software reliability estimation," *Proc. of the IEEE Annual Reliability and Maintainability Symposium*, pp. 100-105, 2000.

[12] W. Farr, *Software Reliability Modeling Survey*, Chapter 3, *Handbook of Software Reliability Engineering*, McGraw-Hill, New York, 1996.

[۱۳] غ. ر. شاه محمدی، س. جلیلی، "ارزیابی کمی سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد با رویکرد مدل‌سازی دیاگرام بلوکی،" *مجموعه مقالات سیزدهمین کنفرانس بین‌المللی انجمن کامپیوتر ایران*، ۱۳۸۶.

[14] F. Buschmann, R. Meunier, H. Rohnert, P. Sommerlad, and M. Stal., *Pattern-Oriented Software Architecture-A system of Patterns*, John Wiley & Sons, 1996.

[15] M. Shaw, and D. Garlan, *software architecture: Perspectives on an Emerging Discipline*, Prentice Hall, 1996.

[16] A. Silberschatz, H. F. Korth, and S. Sudarshan, *Database System Concepts*, McGraw-Hill Series in

جدول ۶ سبک‌های شی‌گرا (OO)، مخزن (RPS)، و فراخوانی ضمنی (I/I) برای این سیستم توصیه می‌شود.

۸- نتیجه‌گیری

در این مقاله ابتدا مروری اجمالی بر سبک‌های معماری نرم‌افزار، دسته‌بندی‌ها و مدل‌های کیفی نرم‌افزار، داشتیم. آنگاه صفت کیفی قابلیت اعتماد در مدل ISO 9126 تبیین گردید. با توجه به فقدان بررسی کمی تاثیر سبک‌های معماری روی صفات کیفی خصوصاً قابلیت اعتماد و نظر به اهمیت سنجش قابلیت اعتماد سبک‌های معماری نرم‌افزار، به منظور انتخاب سبک معماری مناسب، روش‌های ارزیابی قابلیت اعتماد نرم‌افزار بررسی و مشخص گردید که در سطح سبک‌های معماری نرم‌افزار، دو رویکرد RBD و رویکرد مبتنی برحالت مارکوف برای ارزیابی سبک‌های معماری نرم‌افزار از دید قابلیت اعتماد قابل استفاده می‌باشند. آنگاه قابلیت اعتماد سبک‌های معماری با استفاده از این دو رویکرد محاسبه و نتایج آنها با هم مقایسه گردید.

نتایج رتبه‌بندی سبک‌های معماری در دو رویکرد نشان داد که ضمن اینکه مقادیر قابلیت اعتماد سبک‌های معماری در دو رویکرد اندکی متفاوت است، و این اختلاف ناشی از تفاوت رابطه قابلیت اعتماد سبک‌های معماری در دو رویکرد است، با اینحال روند تغییر رتبه سبک‌های معماری با تغییر تعداد مولفه‌های موثر در انجام تراکنش در دو رویکرد یکسان است. ضمن اینکه تمایز بیشتری بین رتبه سبک‌های معماری در رویکرد مبتنی بر مدل مارکوف مشاهده شد. به منظور بررسی تاثیر تعداد مختلف مولفه‌های موثر در انجام تراکنش روی رتبه صفت کیفی قابلیت اعتماد سبک‌های معماری، با تغییر تعداد مولفه‌ها از ۲ تا ۷، ارزش قابلیت اعتماد سبک‌های معماری محاسبه گردید.

مهمترین دستاورد این مقاله، کمی کردن صفت کیفی قابلیت اعتماد در سطح سبک‌های معماری نرم‌افزار و در مرحله انتخاب سبک معماری نرم‌افزار است. کار مشابهی که به ارزیابی، رتبه‌بندی و مقایسه قابلیت اعتماد سبک‌های معماری نرم‌افزار پرداخته باشد، مشاهده نگردیده است. Wang و همکاران [۲۷]، با اصلاح مدل قابلیت اعتماد مبتنی بر مارکوف پیشنهادی Cheung [۱۰]، ساخت یک ماشین حالت را که قادر به بررسی مدل‌های همگن و ناهمگن باشد، ارائه داده‌اند. روش آنها برای ارزیابی سبک‌های معماری ناهمگن قابل استفاده می‌باشد.

Shao و همکارش [۶] یک سیستم هوشمند برای انتخاب سبک معماری نرم‌افزار ارائه داده‌اند، اما چگونگی ارزیابی و رتبه‌بندی سبک‌های معماری نرم‌افزار از دید صفات کیفی را مشخص نکرده‌اند.

سپاسگزاری

این تحقیق با حمایت مالی مرکز تحقیقات مخابرات ایران طرح شماره ۸۵-۱۲-۸۵ TMU انجام شده است.

مراجع

[1] A. Immonen, and E. Niemelä, "Survey of reliability and availability prediction methods from the viewpoint of software architecture," *Journal of Software and Systems Modeling*, Vol. 7, No.1, pp. 49-65, 2008.



سعید جلیلی کارشناسی‌ارشد خود را در سال ۱۳۶۴ از دانشگاه صنعتی شریف و درجه دکترا را در سال ۱۳۷۰ از دانشگاه برادفورد انگلستان اخذ نموده است. زمین‌های تحقیقاتی مورد علاقه ایشان، آزمون نرم‌افزار، واریسی حین اجرای نرم‌افزار، ارزیابی کمی معماری نرم‌افزار و یادگیری ماشین می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

sjalili@modares.ac.ir

اطلاعات بررسی مقاله:

تاریخ ارسال: ۸۷/۰۶/۲۶

تاریخ اصلاح: ۸۹/۰۴/۲۹

تاریخ قبول شدن: ۸۹/۰۵/۱۱

نویسنده مرتبط: دکتر سعید جلیلی، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران.

- ¹ Reliability Block Diagrams
- ² Architecture Styles
- ³ Design Patterns
- ⁴ Pipe and Filter
- ⁵ Implicit Invocation
- ⁶ Layered
- ⁷ Repository
- ⁸ Blackboard
- ⁹ Knowledge Resources
- ¹⁰ Dispatcher
- ¹¹ Broker
- ¹² Client-Side Proxy
- ¹³ Server-Side Proxy
- ¹⁴ Object Oriented
- ¹⁵ Characteristics
- ¹⁶ Maturity
- ¹⁷ Induced
- ¹⁸ Analytic Hierarchy Process
- ¹⁹ Air Traffic Control
- ²⁰ Hard Real-Time

Computer Science, 1997.

[17] L. Bass, P. Clements, and R. Kazman, *Software Architecture in Practice*, Addison-Wesley, 1998.

[18] IEEE std 610.12-1990 (n.d.), IEEE Standard Glossary of Software Engineering Terminology, 1990.

[19] J. A. McCall, P. K. Richards, and G. F. Walters, "Factors in Software Quality," *Nat'l Tech. Information Service*, Vol. 1, 2 and 3, 1977.

[20] B. W. Boehm, J. R. Brown, H. Kaspar, M. Lipow, G. McLeod, and M. Merritt, *Characteristics of Software Quality*, North Holland, 1978.

[21] ISO, "ISO 9126-1:2001, Software engineering – Product quality, Part 1: Quality model," 2001.

[22] K. Goseva-Popstojanova, and K. S. Trivedi, "Architecture-based approach to reliability assessment of software systems," *Performance Evaluation*, Vol 45, pp. 179–204, 2001.

[23] R. Tripathi, and R. Mall, "Early Stage Software Reliability and Design Assessment," *Proc. of the 12th Asia-Pacific Software Engineering Conference*, pp. 619-628, 2005.

[24] S. P. Leblanc, and P. A. Roman, "Reliability Estimation of Hierarchical Software Systems," *Proc. of Annual Reliability and Maintainability Symposium*, pp. 249-253, 2002.

[25] A. Berman, and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press Inc., New York, 1979.

[26] W. Wang, and M. Chen, "Heterogenous Software Reliability Modeling," *Proc. of the 13th International IEEE Symposium on Software Reliability Engineering*, pp. 41-52, 2002.

[27] W. Wang, D. Pan, and M. Chen, "Architecture-based software reliability modeling," *Journal of Systems and Software*, Vol. 79, No. 1, PP. 132-146, 2006.



غلامرضا شاه محمدی دانشجوی دوره دکترا مهندسی کامپیوتر گرایش نرم‌افزار دانشکده مهندسی برق و کامپیوتر دانشگاه تربیت مدرس می‌باشد. نامبرده در طراحی و ارزیابی معماری نرم‌افزار مبتنی بر سبک‌های معماری پژوهش می‌کند. وی کارشناسی‌ارشد و کارشناسی خود را در رشته مهندسی کامپیوتر گرایش نرم‌افزار به ترتیب از دانشگاه تربیت مدرس و دانشگاه فردوسی مشهد اخذ نموده است. زمینه‌های تحقیقاتی مورد علاقه ایشان ارزیابی کمی معماری نرم‌افزار، طراحی مبتنی بر سبک معماری نرم‌افزار، روش‌های مبتنی بر سنجش برای ارزیابی سیستم، کاربرد روش‌های رسمی در مهندسی نرم‌افزار و تخمین هزینه سیستم می‌باشد.

آدرس پست‌الکترونیکی ایشان عبارت است از:

shahmohamadi@modares.ac.ir

راهنمای تهیه مقالات



نسخه نهائی مقالات ارسالی برای چاپ در نشریه علوم و مهندسی کامپیوتر باید بر طبق اصول مطرح شده در این راهنما تهیه شده باشد. رعایت این اصول در نسخه اولیه نیز قویاً توصیه می‌شود. مقالات می‌تواند به یکی از دو زبان فارسی یا انگلیسی ارسال گردد.

۱- ساختار مقاله

- عنوان: کوتاه و معرف محتوای مقاله باشد و از ۱۵ کلمه تجاوز نکند.
- نام نویسندگان و مؤسسه محل اشتغال آنان: از ذکر عناوین خودداری شود.
- چکیده فارسی: حاوی تعریف مسأله، روش حل، و نتایج مهم باشد و از ۱۵۰ کلمه تجاوز نکند.
- واژه‌های کلیدی: حداکثر ۱۰ کلمه
- بدنه اصلی مقاله: بدنه اصلی با «مقدمه» شروع و با «نتیجه‌گیری» خاتمه می‌یابد. بخش‌ها و زیربخش‌های بدنه اصلی باید شماره‌گذاری شوند. شماره «مقدمه» یک خواهد بود.
- شکر و قدردانی (در صورت نیاز).
- مراجع: مراجع به ترتیبی که در متن به آنها رجوع می‌شود آورده شوند. نام مؤلفان مراجع در صورت لزوم در متن بصورت فارسی آورده شود. رجوع به مراجع با ذکر شماره آنها در داخل کروشه ([]) انجام پذیرد.
- پیوست‌ها (در صورت نیاز)
- واژه نامه (در صورت نیاز)
- برای مقالات فارسی، عنوان مقاله، نام نویسندگان، مؤسسه محل اشتغال، چکیده، و واژه‌های کلیدی به زبان انگلیسی نیز در صفحه‌ای جداگانه داده شود. برای مقالات انگلیسی این اطلاعات به زبان فارسی نیز در صفحه‌ای جداگانه داده شود.

۲- معادله‌ها، شکل‌ها، جدول‌ها، و عکس‌ها

- معادله‌ها باید با فاصله کافی از بالا و پائین تایپ و به صورت متوالی شماره‌گذاری شوند. شماره معادله در پیرامون در انتهای سمت راست سطر حاوی معادله قرار داده شود. معادلات دستنویس به هیچ شکل قابل قبول نیستند.
- شکل‌ها و جدول‌ها باید دارای شماره و عنوان باشند. در شکل‌ها شماره و عنوان در زیر شکل و در جدول‌ها در بالای شکل قرار می‌گیرد. اعداد و متون روی شکل‌ها و جدول‌ها باید دارای اندازه مناسب و کاملاً خوانا باشند.
- اعداد و کلمات روی شکل‌ها و جدول‌ها در مقالات فارسی به زبان فارسی باشند.
- عکس‌ها سیاه و سفید، برقی، و با کیفیت عالی باشند.

۳- نحوه نگارش مراجع

در لیست مراجع انواع مختلف مرجع‌ها به شکل زیر نوشته شوند:

- [۱] ب. مقدم، ا. تقوی، و ن. طاهری، **آشنائی با شبکه‌های کامپیوتری**، چاپ دوم، انتشارات نصر، تهران، ۱۳۷۵.
- [۲] ی. براون، **مقدمه‌ای بر شبکه‌های عصبی**، ترجمه م. ع. آرام، انتشارات فجر، مشهد، ۱۳۷۰.
- [۳] **راهنمای کاربران حسابرس**، شرکت پردازش رایانه‌ای ایران، تهران، ۱۳۶۵.
- [۴] ج. عارف، **استنتاج فازی بوسیله شبکه‌های عصبی**، پایان‌نامه کارشناسی ارشد هوش مصنوعی، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شهر، ۱۳۷۴.
- [۵] ج. حسینی، و ح. ربانی، «تشخیص چهره انسان در تصویر»، **نشریه امیرکبیر**، سال هشتم، شماره ۴۲، ص ۱۲۵-۱۴۷، ۱۳۷۷.

[۶] ج. حسینی، و.ح. ربانی، «تشخیص چهره انسان در تصویر»، در مجموعه مقالات هفتمین کنفرانس سالانه انجمن کامپیوتر ایران، ص ۲۲۴-۲۳۲، ۱۳۸۰.

[7] M. A. Ahmadi, and M. H. Rahimi, *Fuzzy Set Theory*, New Jersey: Prentice-Hall, 1995.

[8] M. A. Ahmadi, M. H. Rahimi, and A. Fatemi, "Evidence-Based Recognition of 3D Objects," *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 12, no. 10, pp. 811-835, 1994.

[9] A. Taheri, "On-Line Fingerprint Verification," *Proc. IEEE Intl Conf. Pattern Recognition*, pp. 752-758, 1992.

[10] M. A. Ahmadi, *On-line Fingerprint Verification*, Ph. D. Dissertation, MIT, Cambridge, MA, 1982.

[11] A. J. Washington, "The Fingerprint of MalcomX," <http://www.dermatoglyphics.com>, June 2003.

[12] International Biometrics Group, <http://www.biometricgroup.com>, May 2003.

۴- واژه‌نامه

هر واژه خارجی در واژه‌نامه در انتهای مقاله با شماره‌ای مشخص شود و شماره در معادل فارسی آن واژه در متن، بصورت بالانویس آورده شود.

۵- آماده‌سازی مقاله

- مقاله را با نرم افزار WORD XP تحت ویندوز XP تایپ و با چاپگر لیزری روی کاغذ A4 چاپ نمائید.
- متن چکیده به صورت تک ستونی با طول سطر ۱۸ سانتیمتر و متن مقاله به صورت دو ستونی با طول هر ستون ۸۷ میلیمتر و فاصله دو ستون ۶ میلیمتر تایپ شود. حاشیه‌ها از بالا و پائین برابر ۲۰ میلیمتر و از طرفین برابر ۱۵ میلیمتر اختیار شود.
- فاصله عنوان مقاله در صفحه اول از بالای صفحه برابر ۸۵ میلیمتر باشد و عنوان وسط چین شود.
- کلیه عناوین بصورت پررنگ با قلم «**Nazanin** یا **B Nazanin**» تایپ شوند، اندازه قلم عنوان مقاله ۱۸، عناوین سطح اول ۱۴، و عناوین سطح دوم و سوم ۱۲ انتخاب شوند.
- متن چکیده ها با قلم «**Nazanin** یا **B Nazanin**» اندازه ۹، متن مقاله با قلم «**Nazanin** یا **B Nazanin**» اندازه ۱۰، و کلمات و متن انگلیسی با قلم **Times New Roman** اندازه ۹ تایپ شوند.
- تمام متن بصورت تک فاصله تایپ شود. اسامی نویسندگان از عنوان مقاله و اسامی نویسندگان از عناوین محل اشتغال نویسندگان دو خط فاصله داشته باشد. بالای هر عنوان یک سطر فاصله قرار داده شود.
- سعی شود تعداد صفحات مقاله از ۳۰ صفحه بیشتر نباشد.

۶- نحوه ارسال مقاله

- شش نسخه از مقاله چاپ شده با چاپگر لیزری و دارای کیفیت عالی به دفتر نشریه ارسال گردد. ارسال اولیه مقاله می‌تواند از طریق پست الکترونیکی نیز به آدرس سردبیر انجام شود، اما از طریق دورنگار قابل قبول نیست.
- مقاله ارسالی برای نشریه علوم و مهندسی کامپیوتر نباید در جای دیگری به چاپ رسیده باشد و یا در زمان بررسی توسط نشریه برای چاپ به نشریه دیگری ارسال گردد. نویسندگان باید به‌مراه مقاله نامه‌ای دال بر این مطلب نیز ارسال دارند.
- پس از قبول مقاله، نسخه نهائی تصحیح شده مقاله باید هم بصورت چاپ شده و هم در فایل کامپیوتری در قالب‌های DOC و PDF روی یک دیسکت به نشریه ارسال گردد. همچنین فرم «**اختصاص حق انتشار**» باید توسط نویسندگان تکمیل، امضاء و ارسال شود.
- در نسخه نهائی باید بیوگرافی و عکس کلیه مؤلفین نیز در انتهای مقاله قرار داده شود.

- [2] M. A. Ahadi, M. H. Rahimi, and A. Fatemi, "Evidence-Based Recognition of 3D Objects," *IEEE Trans. Pattern Analysis and Mach. Intell.*, Vol. 12, No.10, pp. 18-25, 1994.
- [3] A. Taheri, "On-line Fingerprint verification," *Proc, IEEE Int'l Conf. Pattern Recognition*, pp. 752-759, 1992.
- [4] M. A. Ahadi, *On-line Fingerprint verification*, Ph.D. Dissertation, University of Tehran, Tehran, Iran, 1994.
- [5] M. A. Washington, "The Fingerprint of Malcom X," <http://www.dermatologyphicisc.com>, June 2004.
- [6] International Biometrics Group, <http://www.Biometrics.com>, May 2003.

4. Manuscript preparation

Typing and Printout: The manuscript should be prepared using Microsoft Word and should be printed on A4 size paper using a laser printer.

Fonts: Use Times New Roman font type. The font size must be 9 for the Abstract and 10 for the main text. All the titles should be made bold. The paper title must have a font size of 18, first-level sub-titles a size of 14, second-and third-level sub-titles a size of 12.

Layout: The paper should be typed in 2-column single space format. The top margin for the first page should be equal to 85mm, each column length should be 87mm, and the spacing between columns 6mm. Only the abstract is to be typed in a single column format. Two space line must be allowed between the paper title and authors' names, and one space line should be allowed above each section title or sub-title.

5. Paper Submission

Exclusive Submission: Submission of a paper to JCSE implies that it has not been published previously, that it is not under consideration for publication elsewhere, and that if accepted, it will not be published elsewhere in any language. Explicit announcement of this matter must be made in a letter to the editor.

Paper Length: The paper is expected to be no longer than 30 pages.

Submission Process: Authors are requested to submit their papers electronically in PDF format to the editor-in-chief at csi-jcse@ipm.ir with message subject of "Paper Submission". All other relevant correspondence should also be addressed to the this email.

Review Process: Each manuscript will be reviewed by at least 3 reviewers and their constructive criticism will be forwarded to the authors.

Final Manuscript: After the acceptance of a paper, the authors should provide the Journal with two copies of the final manuscript and its Word computer file on a diskette. In addition, the authors should sign and submit the form for the transfer of copyright to the Computer Society of Iran.

INFORMATION FOR AUTHORS

The final manuscript of the English papers accepted for publication in the CSI JCSE should be prepared in accordance with this guide. It is, however, strongly recommended that preparation of the initial manuscript also follows the instructions given here. Manuscript may be in English or Farsi.

1. Paper Organization

- Title: The title should be short (at most 15 words) and indicative of the paper contents.
- Authors: the authors' names (initial and last name) and their affiliations should appear next.
- Abstract: the abstract should include the problem explanation, methods used for solution, and the significant results; and should not be longer than 150 words.
- Keywords: The keywords should be relatively independent and together optimally characterize the paper. Include 5 to 10 keywords.
- Text: the main body of the paper should start with Introduction and end with concluding remarks. All sections and sub-sections should be numbered. The number for introduction is 1.
- Acknowledgement: If required, Acknowledgements appear after the concluding remarks.
- References: All publications cited in the text should be presented in the order they are cited in a numbered list of References following Acknowledgements. In the text refer to references by the reference number in square brackets on the line.
- Appendices: If appendices are necessary, they are placed after the list of References.
- Farsi Section: the title, authors' names and affiliations, abstract, and keywords should also be given in Farsi a separate sheets (for papers written in English). For the non-Farsi speaking authors, the journal will supply this section.

2. Figures, Tables, Photographs and Equations

- Figures and Tables: Each figure or table must have a number and a caption. In figures, the number and the caption appear under the figure while in tables, they appear over it. The size of text and numbers in tables and figures must be suitable to allow high legibility. Do not use any type of shading in computer generated illustrations.
- Photographs: High quality glossy black and white photographs must be supplied as they are to be reproduced.
- Equations: Equations are to be numbered consecutively. The number of each equation should appear in parentheses in the right-most end of the equation line. Sufficient space should be allowed above and below each equation.

3. References

The format for various types of references should be as follows:

- [1] M. A. Ahadi, and M. H. Rahimi, *Fuzzy Set Theory*, New Jersey, Prentice-Hall, 1995.

Quantitative Evaluation of Architectural Styles from the Reliability Viewpoint in Safety Critical Software

Gholamreza Shahmohammadi

Saeed Jalili

Department of Computer Engineering, Tarbiat Modares University, Tehran, Iran

ABSTRACT

In the process of software architecture design, different decisions are made that have system-wide impact. One important decision is the selection of a suitable architecture style. The main problem in using architecture styles is the lack of investigation of their quantitative effect on software quality attributes. Concerning the vital role of reliability in safety critical software, one of the main criteria in choosing architecture style of these systems is high reliability. The purpose of this study is to quantify the impact of architecture styles on software reliability that is a desired quality of safety critical software. We evaluate styles through two approaches of Reliability Block Diagram (RBD) and Markov State-based Model, and quantify their impact on software reliability. Then the results of two approaches will be compared. The main contribution of this paper is quantification and ranking of software architecture styles from the perspective of reliability quality attribute at stage of architectural style selection.

Keywords: Software Architecture, Evaluation of Software Architecture Styles, RBD, Markov State-Based Model.

Feature Extraction Using Mutual Information for Classification of Electroencephalogram in Brain Computer Interface

Farid Oveisi

Abbas Erfanian Omidvar

Department of Biomedical Engineering, Iran University of Science and Technology, Tehran, Iran

ABSTRACT

Reducing the number of features is essential to improve the accuracy, efficiency and scalability of a classification process. There are two main reasons to keep the dimensionality of the input features: computational cost and classification accuracy. Reducing the number of input features can be done by selecting relevant features (i.e., feature selection) or extracting new features containing maximal information about the class label from the original ones (i.e., feature extraction). In this work, we use a mutual information based feature extraction (MIFX) algorithm for classification of electroencephalogram (EEG) in brain-computer interface (BCI). The tasks to be discriminated are the imaginative hand movement and the resting state. The experiments were conducted on four healthy subjects on different days. The results show that the classification accuracy obtained by MIFX is higher than that achieved by full feature set. Moreover, the results indicate that the performance obtained using MIFX is higher than that obtained using principle component analysis.

Keywords: Feature Extraction, Mutual Information, Classification, Brain-Computer Interface, Brain Signal.

A New Approach Based on Learning Automata for Solving Steiner Tree Problem

Samira Noferesti¹

Mohammad Reza Meybodi²

¹Department of Electrical and Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran

²Department of Computer Engineering and Information Technology, Amirkabir University of Technology, Tehran, Iran

ABSTRACT

The Steiner tree problem in a weighted graph is to find a least cost sub tree in the graph such that it contains special vertices called terminal points. The Steiner tree problem on graphs is one of the best-known NP-hard problems. For this reason, many approximation algorithms for solving steiner tree problem have been proposed. The traditional algorithms achieve near optimal results, but cost more computational time. In this paper we propose an iterative algorithm based on learning automata for solving dynamic Steiner tree problem. Simulation results show that our proposed algorithm generates Steiner trees of low cost in reasonable running time.

Keywords: Wavelet Static Steiner Tree Problem, Dynamic Steiner Tree Problem, Learning Automata, Hard Problems.

Ontology Population and Semantic Annotation of Persian Texts in POPTA

Mehrnoush Shamsfard

Bahareh Sarrafzadeh

Department of Electrical and Computer Engineering, Shahid Beheshti University, Tehran, Iran

ABSTRACT

Ontology population refers to extracting the instances of classes (concepts) of an ontology from text. On the other hand Semantic Web provides structured content by adding annotations to documents. Semantic annotation tags class instances and map them into classes in an ontology. Thus annotation and population are two sides of a coin. This paper introduces an inter-language population and annotation system called POPTA. It annotates Persian texts according to an English ontology and populates the English ontology by instances extracted from Persian texts. POPTA exploits a combination of statistical and pattern based approaches besides web based methods, search engines, online encyclopedia and Wikipedia to perform the annotation and population tasks.

Keywords: Ontology Population, Semantic Annotation, Persian Text Processing, Semantic Web, Wikipedia.

Automatic Fluency Evaluation Using Speech Processing Applied in CALL

Behzad Bakhtiyari

Mohammad Mehdi Homayounpour

Department of Computer Engineering and Information Technology, Amirkabir University of Technology, Tehran, Iran

ABSTRACT

This paper presents a new method for evaluation of speech fluency. Speech fluency evaluation is needed in computer-assisted foreign language learning, where the fluency of speech uttered by a new language learner should be evaluated. Fluency evaluation helps the language learner to know his/her degree of speech fluency, so he/she can do more exercise and try to ameliorate his pronunciation fluency. To do this some quantitative measures are presented and a method has been proposed to automatically evaluate speech fluency using the best measures. In order to study the precision of the evaluation of the proposed method, a database of speech data from Turkish speakers was recorded and used as the benchmark in our studies. The correlation between the scores presented by human judges and the proposed automatic technique for measuring the fluency of uttered speech is about 95% which shows that the proposed technique performs as similar as the human judges. The proposed technique can be applied to other languages.

Keywords: Computer-Assisted Language Learning, CALL, Fluency Evaluation, Speech Processing.

Intelligent Traffic Light

Aminollah Mahabadi

Faculty of Engineering, Shahed University, Tehran, Iran

ABSTRACT

This paper presents a new model for an automatic intelligent traffic light system to measure traffic intersection from video images and controls the light timing program with calculating the occupation of inputs and outputs roads as a reliable sensor. It is based on background image differential and updating that with three stages as creating real time background images, occupation calculation with eliminating shadows and traffic lights phases timings with high accuracy by a new method. System plays real time behaviors with accepting knowledge from the traffic knowledge based section and movement recalculating with commands generating. Control model of lights and phases is based on metapeternet with timing depend on total capacity for all inputs and outputs camera's images of the four phase intersection. Automatically moving objects in another road sides against the movement direction of each road isn't effected on the results. The system processes 10 frames in 1 second for each camera. It is fault tolerant in congestion and shadow situations in smooth or suddenly light environments changes.

Keywords: Intelligent Traffic Light, Intelligent Transportation Systems (ITS), Petri Net, Image Processing, Traffic Measurement, Intersection Modeling, Shadow.

Speech Enhancement Using Two New Filters Based on a Posteriori Probability Distribution of Wavelet Coefficients

Behzad Zamani Dehkordi¹ Ahmad Akbari¹ Babak Nasesharif^{1,2}

¹Department of Computer Engineering, Iran University of Science and Technology, Tehran, Iran

²Faculty of Engineering, University of Guilan, Guilan, Iran

ABSTRACT

Wavelet-based denoising methods can be divided into thresholding and filtering based methods. The thresholding methods, which are based on the concentration of signal energy in a few number of wavelet coefficients, usually remove weak speech components. This tends to poor speech quality after removing colored noises due to speech distortion. Filtering based methods apply a filter to wavelet coefficients in order to denoise them and produce less speech distortion. In this paper, we first show the superiority of filtering methods in comparison to thresholding methods. For this purpose, we compare SURE soft thresholding method with Wiener filter. Then, we propose two new filters based on a posteriori distribution of wavelet coefficients. These new filters are compared to Wiener filter in the wavelet domain. Results show better performance of proposed filters. We also compare our filters with a maximum a posteriori filter in the frequency domain. Experimental results show that proposed filters improve signal to noise ratio and give a less distorted enhanced speech signal.

Keywords: Wavelet Transform, Wiener Filter, Packet Wavelet Tree, Threshold, MAP.

The CSI Journal on Computer Science and Engineering

Vol. 6

No. 3 (a)

2008

ABSTRACTS

- **Speech Enhancement Using Two New Filters Based on a Posteriori Probability Distribution of Wavelet Coefficients** 1
Behzad Zamani Dehkordi, Ahmad Akbari and Babak Nasesharif

- **Intelligent Traffic Light** 2
Aminollah Mahabadi

- **Automatic Fluency Evaluation Using Speech Processing Applied in CALL** 3
Behzad Bakhtiyari and Mohammad Mehdi Homayounpour

- **Ontology Population and Semantic Annotation of Persian Texts in POPTA** 4
Mehrnoush Shamsfard and Bahareh Sarrafzadeh

- **A New Approach Based on Learning Automata for Solving Steiner Tree Problem** 5
Samira Noferesti and Mohammad Reza Meybodi

- **Feature Extraction Using Mutual Information for Classification of Electroencephalogram in Brain Computer Interface** 6
Farid Oveisi and Abbas Erfanian Omidvar

- **Quantitative Evaluation of Architectural Styles from the Reliability Viewpoint in Safety Critical Software** 7
Gholamreza Shahmohammadi and Saeed Jalili

The CSI Journal on Computer Science and Engineering

A quarterly publication of Computer Society of Iran (CSI)

Editor-in-Chief

H. Sarbazi-Azad, Associate Professor, Sharif University of Technology, and IPM, Tehran, Iran.

Editorial Board

G. Agha, Professor, University of Illinois at Urbana-Champaign, USA.
H. Arabnia, Professor, University of Georgia, USA.
F. Arbab, Professor, CWI and Leiden University, The Netherlands.
K. Badie, Associate Professor, Iran Telecommunication Research Center, Iran.
N. Bagherzadeh, Professor, University of California at Irvine, USA.
B. Bose, Professor, Oregon State University, USA.
A. Edalat, Professor, Imperial College, UK.
M. Fathi, Associate Professor, Iran University of Science and Tech., Iran.
M. H. Ghassemian, Professor, Tarbiat Modarres University, Iran.
M. Ghodsi, Professor, Sharif University of Technology, Iran.
A. R. Hurson, Professor, Pennsylvania State University, USA.
F. Jahanian, Professor, University of Michigan, USA.
E. Kabir, Associate Professor, Tarbiat Modarres University, Iran.
F. C. M. Lau, Professor, University of Hong Kong, Hong Kong.
A. Movaghar, Professor, Sharif University of Technology, Iran.
N. Mahdavi-Amiri, Professor, Sharif University of Technology, Iran.
R. Meybodi, Professor, Amirkabir University of Technology, Iran.
K. Nakano, Professor, Hiroshima University, Japan.
M. Ould-Khaoua, Professor, University of Glasgow, UK.
B. Parhami, Professor, University of California at Santa Barbara, USA.
R. Safabakhsh, Professor, Amirkabir University of Technology, Iran.
H. Sarbazi-Azad, Associate Professor, Sharif University of Technology, and IPM, Iran.
B. Shirazi, Professor, Washington State University, USA.
A. Zomaya, Professor, The University of Sydney, Australia

International Advisory Board

A. Almagarmid, Purdue University, USA.
M. Anvari, University of California at Berkeley, USA.
J. Armstrong, Northeastern University, USA.
H. Berenji, NASA, USA.
D. Bjorner, Delft University, The Netherlands.
M. Ghanbari, University of Essex, UK.
M. Pedram, University of Southern California, USA.
P. Rogaway, University of California at Davis, USA.
M. Sarrafadeh, Northwestern University, USA.
B. Zoghi, Texas A&M University, USA.

Assistants

F. Ghassemi (Editorial Assistant)
S. Shahnazi (Publication Assistant)
L. Nourani (Publication Assistant)
M.R. Hoseiny-Farahabadi (Webmaster)

Disclaimer: Publication of papers in CSI-JCSE does not imply that the editorial board, reviewers, or CSI-JCSE accept, approve or endorse the data and conclusions of authors.