



دانشگاه صنعتی شاهرود



قطب علمی
سامانه های پردازش دیجیتال

سومین کنفرانس پردازش سیگنال و

سیستم های هوشمند ایران

۲۹ و ۳۰ آذر ماه ۱۳۹۶ دانشگاه صنعتی شاهرود، شاهرود

تنظیم و صفحه آرایی: حامد غضنفری

دبیرخانه کنفرانس

شاهرود، دانشگاه صنعتی شاهرود، دانشکده مهندسی کامپیوتر

کد پستی: ۳۶۱۹۹۵۱۶۱ تلفن: ۰۲۳۳۲۳۰۰۳۳۸

تارنما: www.icspis.shahroodut.ac.ir و www.icspis.ir

پست الکترونیک: icspis@shahroodut.ac.ir



مجموعه مقالات فارسی

سومین کنفرانس

پردازش سیگنال و سیستم‌های هوشمند ایران (ICSPIS)

دانشگاه صنعتی شاهرود، شاهرود

۲۹ و ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود



قطب علمی
سامانه‌های پردازش دیجیتال



ردیف	عنوان	نویسندگان	صفحه
۱	پیاده‌سازی و شبیه‌سازی استراتژی تجمع ربات‌ها با استفاده از تئوری کنترل سوپروایزری در رباتیک جمعی	فائزه میرزائی، علی‌اکبر پویان، سعیده فردوسی	۷
۲	ارائه روش هوشمند برای طراحی سیستم‌های ANFIS	امیر سلطانی محبوب، سید حمید ظهیری	۱۵
۳	بخش‌بندی خودکار تصاویر سی‌تی اسکن ریه به کمک یادگیری تقویتی	پرینا قیصری، منصور فاتح، محسن رضوانی	۲۲
۴	نشان کاوی کور بردار حرکت ویدئوی فشرده‌سازی شده H.264	الهه سادات سادات، کریم فائز	۲۸
۵	به کارگیری الگوریتم آداپوست و درخت تصمیم در تشخیص بیماری قلبی عروق کرونر	زینب حسنی، نجمه صمدیانی	۳۴
۶	بهینه‌سازی استقرار در شبکه‌های حسگر بیسیم جهت دستیابی به اتصال و پوشش کامل	امید رضا معروضی، مریم رشمه کریم	۳۹
۷	بهبود الگوریتم HAU-Miner (الگوریتم کشف مجموعه اقلام با متوسط / ارزش بالا)، با هدف کاهش حافظه مصرفی	آزاده سلطانی، محمود سلطانی	۴۵
۸	ارائه یک سیستم تشخیص نفوذ سبک برای تشخیص حملات انکار سرویس در شبکه‌های حسگر بیسیم	مهدی صادقی زاده، امیدرضا معروضی، علی‌اکبر پویان	۵۰
۹	شناسایی ترافیک شبکه در مسیر یاب‌ها با حفظ حریم خصوصی کاربران	محمد رضا گندمی، حمید حسن پور	۵۶
۱۰	حذف نویز جهتی با استفاده از آرایه‌های میکروفونی تفاضلی	حمید ولی‌نژاد طرقله، حمیدرضا ابوطالبی	۶۳
۱۱	فشرده‌سازی تصاویر رنگی مبتنی بر شبکه عصبی عمیق	علی کارگران، هادی گرایلو	۶۹
۱۲	ارائه مجموعه‌ای از ویژگی‌های آماری جهت تعیین انسجام متن در سیستم‌های پرسش و پاسخ تعاملی	محمد مهدی حسینی، مرتضی زاهدی	۷۵
۱۳	یادگیری متریک هیستوگرامی l_2 با تنظیم‌کننده dropout	حمید صادقی، ابوالقاسم-اسدالله راعی	۸۱
۱۴	بررسی تأثیر کاهش ویژگی بر افزایش نرخ دقت تشخیص صفحات وب هرز	علی سلیمانی ایوری، فائزه اصدقی	۸۷
۱۵	شناسایی ندول‌های ریوی در تصاویر CT اسکن قفسه سینه با استفاده از خوشه‌بندی سه‌بعدی و الگوریتم RUSBoost	مهدیه سادات حسینی ورکی، کریم فائز	۹۳
۱۶	ارائه مدل ترکیبی متشکل از الگوریتم خفاش و اتوماتای یادگیر برای بهینه‌سازی در محیط‌های ایستا	آرش ملکی نودهی، حمیدرضا کرماجانی، مانی زارعی	۱۰۰
۱۷	ثبت خودکار تصاویر رادار روزنه مصنوعی بر روی تصاویر مرئی	سوسن غلامی، علیرضا احمدی فرد	۱۰۷
۱۸	ارائه یک الگوریتم جدید جهت حذف نویز تصویر براساس یادگیری واژه‌نامه ناهمدوس	سمیرا مودتی	۱۱۴
۱۹	بهبود امنیت لایه فیزیکی با استفاده از تخصیص توان تطبیقی	مهدی محمدی، محمد ترابی	۱۲۰
۲۰	تخصیص عادلانه منابع در سیستم‌های چندبخشی مبتنی بر MIMO-OFDM با استفاده از الگوریتم ژنتیک	علیرضا غلامرضائی، حمید فرخی	۱۲۶
۲۱	شناسایی نواحی برجسته تصویر با تلفیق ویژگی‌های رنگ، بافت و موقعیت مکانی پیکسل‌ها	سکینه اسدی امیری	۱۳۳
۲۲	ردیابی بلادرنگ هدف با استفاده از پردازش چند هسته‌ای و دو ویژگی مبتنی بر رنگ و مبتنی بر بافت	مهدی مقیمی، حسین خسروی	۱۳۹
۲۳	بهبود سامانه ردیابی تصویری هدف مبتنی بر ویژگی‌های استخراج شده از شبکه عصبی عمیق	احسان عبداللهی، محمد تقی صادقی	۱۴۴
۲۴	روشی سریع و کارآمد برای حذف سایه خودروهای متحرک بمنظور تخمین محدوده دقیق خودرو در تصاویر دریافتی از دوربین جاده‌ای	رسول عسگریان دهکردی، حسین خسروی	۱۵۰
۲۵	تعیین میزان بهینه تبادل برق در نیروگاه‌های تلمبه ذخیره‌ای شبکه الکتریکی هوشمند با استفاده از الگوریتم کرم شب تاب	محسن خنجری، محمد حسینی ابرده	۱۵۸
۲۶	دسته‌بندی قطعی پاسخ‌ها در یک سیستم پرسش و پاسخ تعاملی	مولود آیت، محمد مهدی حسینی، مرتضی زاهدی	۱۶۴
۲۷	ارزیابی امنیت لایه فیزیکی در شبکه‌های مخابرات بیسیم با استفاده از معیارهای ظرفیت کانال امن و احتمال قطع امن	پریناز کاوندی، محمد ترابی	۱۷۰

۱۷۶	خشایار نوریان، حسین خسروی	تعیین هویت بلادرننگ چهره با استفاده از ویژگی‌های مکان فرکانس چهره در مجموعه‌های بزرگ	۲۸
۱۸۳	جواد برسلانی، علی دستفان، جواد قالیبافان	طراحی و بکارگیری فیلتر وینر در جداسازی سیگنال‌های EMI ناشی از مبدل‌های DC/DC افزاینده در حالت موازی	۲۹
۱۸۸	محمود معلم، علی‌اکبر پویان	تشخیص افعال روزمره با استفاده از تلفن همراه و یادگیری عمیق	۳۰
۱۹۵	سید مجتبی حسینی راد	خوشه‌بندی هوشمند شبکه‌های حسگر بی‌سیم با رویکرد الگوریتم فراابتکاری	۳۱



پیاده‌سازی و شبیه‌سازی استراتژی تجمع ربات‌ها با استفاده از تئوری کنترل سوپروایزری در رباتیک جمعی

فائزه میرزائی^۱، علی اکبر پویان^۲، سعیده فردوسی^۳

^۱ آزمایشگاه هوش مصنوعی و سیستم‌های توزیعی، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود، fmirzaei@shahroodut.ac.ir

^۲ آزمایشگاه هوش مصنوعی و سیستم‌های توزیعی، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود، apouyan@shahroodut.ac.ir

^۳ دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، شاهرود، sferdowski@shahroodut.ac.ir

چکیده

امروزه نرم‌افزارهای کنترلی سیستم‌های مبتنی بر رباتیک جمعی با تکیه بر دست‌یابی به اهداف مشخص طراحی می‌شوند. آزمون، آنالیز و تصدیق وجود ویژگی‌های موردنیاز در این سیستم‌ها به دشواری صورت می‌گیرد. از طرفی، به دلیل تولید دستی کد، نمی‌توان تضمین کرد که پیاده‌سازی انجام شده منطبق با طراحی اولیه باشد. برای مقابله با این چالش‌ها، استفاده از تئوری کنترل سوپروایزری پیشنهاد می‌شود. در این مقاله، تئوری کنترل سوپروایزری بر روی سکوی نوظهور آرگوس پیاده‌سازی شده است. پس از محاسبه سوپروایزر، کد کنترلی موردنیاز به صورت خودکار تولید شده و در نهایت استراتژی تجمع ربات‌ها با محدودیت‌های بیشتر نسبت به کارهای پیشین، طراحی و توسعه داده شده است. محدودیت‌هایی که در این مقاله پیشنهاد شده‌اند شامل استفاده از سنسورهای ضعیف‌تر و با برد کمتر و جمعیت بسیار بیشتر ربات‌ها می‌باشند. نتایج آزمایش‌های صورت گرفته و مقایسه آن با روش‌های پیشین، نشان از عملکرد مناسب استراتژی پیشنهادی دارد.

کلمات کلیدی

رباتیک جمعی، تئوری کنترل سوپروایزری، استراتژی تجمع ربات‌ها، سکوی آرگوس

الهام می‌گیرند [۱]. رباتیک جمعی نیز در تلاش است تا طرز فکر و رفتار حیوانات در موقعیت‌های مختلف را تقلید نماید.

رباتیک جمعی دارای ویژگی‌های خاصی است که آن را از سیستم‌های چندرباته^۲ متمایز می‌کند. ویژگی‌هایی همچون استقلال^۳، تعداد زیاد^۴، توانایی محدود، مقیاس‌پذیری^۵، مقاومت^۶، انعطاف، هماهنگی توزیع‌شده^۷ و همگن بودن از مهم‌ترین موارد هستند [۲]. در

۱- مقدمه

رباتیک جمعی^۱ (SR) یکی از مورد علاقه‌ترین موضوعات تحقیق در حوزه رباتیک است که حوزه‌های هوش مصنوعی، تئوری کنترل، رباتیک، مهندسی سیستم و بیولوژی را در کنار یکدیگر قرار می‌دهد. رباتیک جمعی به مطالعه جمع بزرگی از ربات‌ها که برای انجام یک هدف مشترک باهم همکاری می‌کنند، می‌پردازد. برخی از محققان علم رباتیک، از هوش انسان و برخی دیگر از رفتار حیوانات در طبیعت

² Multi-robot Systems

³ Autonomy

⁴ Large Number

⁵ Scalability

⁶ Robustness

⁷ Distributed Coordination

¹ Swarm Robotic

طراحی ماشین حالت متناهی توسط روش‌های بهینه‌سازی تعیین می‌شوند.

شبکه پتری^{۱۴} روشی دیگر برای طراحی کنترل‌کننده در رباتیک جمعی است [۷]. کنترل‌کننده‌های مبتنی بر پتری نت توسط کامپیوتری مرکزی با ربات‌ها در ارتباط هستند؛ در صورتی که در رباتیک جمعی نباید کنترل‌کننده‌ی مرکزی وجود داشته باشد، زیرا این امر موجب افزایش هزینه و پیچیدگی طراحی می‌شود. منطق زمانی^{۱۵} نیز برای مدل کردن و آنالیز سیستم‌های رباتیک جمعی استفاده شده است [۸]. این روش می‌تواند با روش بررسی مدل ترکیب شده، ارزیابی فرمال سیستم را انجام دهد [۸].

استفاده از تئوری کنترل سوپروایزری^{۱۶} (SCT) در رباتیک جمعی اولین بار در سال ۲۰۱۶ مطرح شد [۳]. چند مورد از مزایای SCT در مقابل استفاده از روش‌های فرمال^{۱۷} موجود عبارتند از: کاهش راه‌حل‌های تک‌منظوره^{۱۸}، تولید خودکار کدهای کنترلی به منظور مدل‌سازی نیازهای^{۱۹} سیستم، شکستن طراحی یک سیستم پیچیده به بخش‌های کوچک و قابل فهم، اثبات وجود رفتارها در کد کنترلی، قابلیت استفاده‌ی مجدد^{۲۰} کنترل‌کننده‌های طراحی شده در سکوی رباتیک مختلف. در بخش بعد به این موارد بیشتر پرداخته خواهد شد.

در این مقاله، چند نوآوری صورت گرفته است. در گام اول، نرم‌افزاری برای طراحی رفتارها و مشخصه‌های کنترلی سیستم به روش‌های فرمال پیاده‌سازی شده است. این نرم‌افزار به صورت خودکار و با استفاده از رفتارها و مشخصه‌های کنترلی سیستم، کنترل‌کننده را بدست می‌آورد؛ آنگاه وجود ویژگی‌های مانند کنترل‌پذیری و عدم وجود بن‌بست را به روش‌های فرمال بررسی کرده و تضمین می‌کند. در گام دوم، کد کنترلی معادل با کنترل‌کننده‌ی بدست آمده در مرحله قبل به صورت خودکار تولید می‌گردد. در گام سوم، تئوری کنترل سوپروایزری در سکوی نرم‌افزاری آرگوس^{۲۱} [۹] پیاده‌سازی می‌شود. به این وسیله، کد بدست آمده در مرحله قبل به صورت مستقیم در سکوی آرگوس قابل اجرا خواهد بود. در پایان، یک روش جدید برای تجمع ربات‌ها پیشنهاد شده و با استفاده از گام‌های قبل در محیط آرگوس شبیه‌سازی گشته است.

در ادامه مقاله و در بخش دوم، پیاده‌سازی تئوری کنترل سوپروایزری در رباتیک جمعی تشریح گشته است. در بخش سوم به جزئیات پیاده‌سازی تجمع ربات‌ها با استفاده از SCT پرداخته شده است. بخش چهارم شامل نتایج شبیه‌سازی‌های صورت گرفته می‌باشد. نتیجه‌گیری و کارهای آینده نیز در بخش پایانی ارائه گشته است.

رباتیک جمعی، ربات‌ها باید بتوانند به صورت فیزیکی با محیط تعامل داشته و بر آن اثر بگذارند. در این سیستم‌ها مشابه اجتماعات حیوانی مانند مورچه و زنبور عسل، شاهد تجمع تعداد زیادی ربات هستیم؛ علاوه‌براین، رهبر کنترل‌کننده‌ای نیز مشاهده نمی‌کنیم. توانایی‌های هر یک از ربات‌ها باید کم باشد؛ به گونه‌ای که توانایی انجام وظایف بزرگ را به تنهایی نداشته باشند. این مسئله به کاهش هزینه‌ی ربات نیز کمک می‌کند که از قیود اصلی رباتیک جمعی است. سیستم باید با افزایش تعداد ربات‌ها به درستی عمل کند و حتی در شرایط مناسب، کارایی کل مجموعه ارتقا یابد. ربات‌ها توانایی تولید راه‌حل‌های ماژوله برای وظایف مختلف را دارند. به این معنی که برای انجام عملیات مختلف، نیاز به تنظیمات سخت‌افزاری زیاد یا تغییر سکو-های نرم‌افزاری نباشد. مجموعه ربات‌ها در شرایطی که تعدادی از ربات‌ها دچار خسارت شوند یا محیط آشوبی شود، باید به کار خود ادامه دهد؛ حتی اگر لازم باشد کارایی تا حدی کاهش یابد. در واقع با وجود فراوانی^۹ در این سیستم‌ها، خسارت یک ربات، توسط ربات‌های دیگر جبران می‌شود. در رباتیک جمعی، هماهنگی بین ربات‌ها توزیع شده است و هر ربات حسگرها و ارتباطات محدود و محلی دارد. همچنین، تعداد محدودی گروه وجود دارد که ربات‌های هر گروه از نظر فیزیکی و نرم‌افزاری مشابه‌اند و نقش یکسان در عملیات دارند.

طراحی منطق کنترل‌کننده ربات‌ها یک مسئله چالش‌برانگیز است؛ زیرا منطق در سطح ربات‌ها و به صورت منفرد طراحی می‌شود ولی نتیجه آن به شکل جمعی مورد بررسی قرار می‌گیرد. از طرفی، همه ربات‌ها یک کد کنترلی یکسان را اجرا می‌کنند. به همین دلیل، نرم‌افزار کنترل‌کننده ربات‌ها معمولاً به روش‌های خاص منظوره و اغلب با استفاده از روش‌های مبتنی بر مهندسی نرم‌افزار حاصل می‌شود. سیستم‌های طراحی شده با این روش‌ها به سادگی قابل انتقال به کاربردهای دنیای واقعی نیستند. علاوه‌براین، تجزیه و تحلیل وجود ویژگی‌های موردنیاز کنترل‌کننده در کدهای بدست آمده از این روش‌ها دشوار بوده و نگهداری از این کدها نیز به سادگی صورت نمی‌پذیرد [۳]. حتی در صورت استفاده از روش‌های فرمال برای طراحی منطق کنترل‌کننده، هیچ تضمینی برای انطباق کدهای نوشته شده با طراحی صورت گرفته وجود ندارد [۴].

مرجع [۵] روشی با عنوان طراحی ویژگی محور^۱ ارائه کرده است. این روش از چهار مرحله تشکیل می‌شود. ابتدا نیازها به صورت فرمال تعیین می‌گردند. در بخش دوم مدل ماکروسکوپی^{۱۱} با استفاده از زنجیره‌ی مارکو^{۱۲} طراحی و توسط روش بررسی مدل^{۱۳} ارزیابی می‌شود. سپس از مدل برای پیاده‌سازی و شبیه‌سازی استفاده می‌گردد. مرجع [۶] روشی برای طراحی خودکار ماشین حالت متناهی احتمالاتی ارائه داده است. در این روش، از ماژول‌های از پیش تعریف شده‌ای برای طراحی استفاده می‌گردد بطوریکه حالت‌های مناسب

¹⁴ Petri Net

¹⁵ Temporal Logic

¹⁶ Supervisory Control Theory

¹⁷ Formal

¹⁸ Ad Hoc

¹⁹ Specification

²⁰ Reusability

²¹ ARGOS

⁸ Platform

⁹ Redundancy

¹⁰ Property-driven Design

¹¹ Macroscopic

¹² Markov Chain

¹³ Model Checking

۲- پیاده‌سازی تئوری کنترل سوپروایزری در رباتیک جمعی

روش پیشنهادی، تئوری کنترل سوپروایزری را در رباتیک جمعی به کار گرفته و در این مسیر چندین نوآوری نیز ارائه کرده است. بنابراین در ابتدا به تشریح SCT پرداخته، سپس مراحل موردنیاز برای محاسبه خودکار سوپروایزر و کد کنترلی توصیف شده است؛ در پایان این بخش، به لایه‌های نرم‌افزاری موردنیاز برای پیاده‌سازی در سکوی آرگوس اشاره شده است.

۲-۱- تئوری کنترل سوپروایزری

تئوری کنترل سوپروایزری یک فریم‌ورک تئوری برای سنتز کنترل‌کننده‌ها است که به آن‌ها سوپروایزر می‌گویند. فرض می‌شود که سیستم مورد بررسی را می‌توان به صورت یک سیستم رخداد گسسته^{۲۲} (DES) بیان کرد. یک سیستم رخداد گسسته از تعدادی حالت گسسته تشکیل شده و جابجایی بین حالات توسط رخدادها صورت می‌گیرد که گذار^{۲۳} نام دارد. SCT بین رخدادهای قابل کنترل و رخدادهای غیرقابل کنترل تمایز قائل می‌شود. رخدادهای غیرقابل کنترل، سیگنال‌های واکنشی همچون سیگنال ارسال شده توسط حسگرها را نشان می‌دهند. رخدادهای قابل کنترل، سیگنال‌های فرمان هستند که توسط کنترل‌کننده ایجاد می‌شوند.

در SCT، طراح دو مسئله را مدل می‌کند: (۱) رفتارها: سیستم چه کاری می‌تواند انجام دهد و (۲) نیازها: سیستم چه کاری را باید انجام دهد. در بخش اول توانایی‌های سیستم توسط تعدادی مدل رفتار آزاد^{۲۴} توصیف می‌شوند و در بخش دوم مشخصه‌های کنترلی^{۲۵} تعیین می‌گردند؛ و هر دو بخش مذکور توسط یک زبان فرمال تعریف می‌شوند. زبان مجموعه‌ای از کلمات است و کلمات نتیجه‌ی الحاق الفبای زبان می‌باشند. هر حرف الفبا به یک رویداد در DES نسبت داده می‌شود؛ بنابراین دنباله‌ی مطلوب رویداد، کلمات زبان را تشکیل می‌دهند. SCT همه‌ی مدل‌های رفتار آزاد و مشخصات کنترلی را در یک زبان ترکیب می‌نماید. ناظر SCT با محدود کردن مجموعه‌ای از رخدادهای قابل کنترل که سیستم انتخاب می‌نماید؛ تضمین می‌کند که در هر زمان تنها کلمات معتبر یا پیش‌وند کلمات معتبر اتفاق بیفتد.

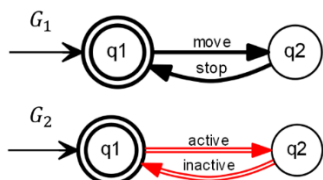
فرض کنید رباتی می‌خواهد یک شیشه شیر را از یخچال بردارد. ربات در ابتدا رویداد قابل کنترل "باز کردن در یخچال" را انتخاب می‌کند. حال اگر رویداد غیرقابل کنترل "شیر در یخچال وجود دارد" اتفاق بیفتد؛ SCT مجموعه رویدادها را به "شیر را از یخچال بردار" و "درب یخچال را ببند" محدود می‌کند. در صورتی که رویداد غیرقابل کنترل "شیر در یخچال وجود ندارد" اتفاق بیفتد؛ SCT مجموعه رویدادها را به "درب یخچال را ببند" محدود می‌کند. در هر دو سناریو، ربات تا بسته شدن درب یخچال کار دیگری نمی‌تواند انجام دهد.

۲-۱-۱- مولدها

زبان‌های منظم^{۲۶} رایج‌ترین زبان‌های استفاده شده در SCT هستند که تحت عنوان زبان‌های نوع سوم نیز شناخته می‌شوند. کلمات زبان منظم را می‌توان توسط مولد^{۲۷} ایجاد نمود. مولد ساختاری شبیه به آتاماتای متناهی داشته و ماشین حالت متناهی^{۲۸} خوانده می‌شود. با این حال آتاماتا کلمات یک زبان را با «پذیرش» یا «رد» آن‌ها تشخیص می‌دهد ولی مولد کلمات یک زبان را تولید می‌نماید. یک مولد G به صورت $(Q, \Sigma, \delta, q_0, Q_m)$ تعریف می‌شود که در آن Q مجموعه‌ی حالات، Σ الفبای زبان و δ تابع انتقال است که به صورت $\delta: Q \times \Sigma \rightarrow Q$ تعریف می‌شود. $q_0 \in Q$ حالت اولیه و $Q_m \subseteq Q$ مجموعه حالات علامت‌دار هستند. حالات علامت‌دار به حالت‌های امن سیستم گفته می‌شوند. به عنوان مثال پایان یک وظیفه را می‌توان با حالت علامت‌دار نشان داد ولی این حالت‌ها لزوماً به معنی پایان عملیات نبوده و ممکن است سیستم پس از این حالت‌ها به کار خود ادامه دهد. زبان تولید شده توسط مولد G را به صورت $L(G)$ نمایش می‌دهند. رویدادهای سیستم که معادل الفبای زبان هستند به دو نوع قابل کنترل (Σ_c) و غیرقابل کنترل (Σ_{nc}) تقسیم می‌شوند، به صورتی که $\Sigma = \Sigma_c \cup \Sigma_{nc}$ و $\Sigma_c \cap \Sigma_{nc} = \emptyset$. رویداد قابل کنترل $e_c \in \Sigma_c$ در حالت $q \in Q$ تنها در صورتی فعال می‌شود که تابع $\delta(q, e_c)$ تعریف شده باشد.

۲-۱-۲- مدل رفتار آزاد

در تئوری کنترل سوپروایزری، هر یک از توانایی‌های فیزیکی سیستم (رفتارها) با استفاده از یک مدل رفتار آزاد بیان می‌شود. در نتیجه برای یک سیستم m مولد G_i ($i \in \{1, 2, \dots, m\}$) تعریف شده که هر مولد یک مدل رفتار آزاد را توصیف می‌کند. به صورت پیش فرض مدل‌های رفتار آزاد را مستقل از یکدیگر در نظر می‌گیرند. نوار نقاله‌ای را در یک کارخانه‌ی تولیدی در نظر بگیرید که محصولات را جابجا می‌کند. در کنار نوار نقاله، حسگری وجود دارد که وجود و عدم وجود محصول روی نوار نقاله را تشخیص می‌دهد. شکل ۱ مدل رفتار آزاد برای نوار نقاله و حسگر را نشان می‌دهد. رخدادهای $stop$ و $move$ قابل کنترل و رخدادهای $active$ و $inactive$ غیرقابل کنترل هستند. معمولاً تنها حالت اولیه‌ی مدل‌های رفتار آزاد را علامت‌دار قرار می‌دهند. این کار به معنی برگشت سوپروایزر به شرایط اولیه است. کمان‌های یک‌خطی و دوخطی به ترتیب به معنی رویدادهای قابل کنترل و غیرقابل کنترل هستند.



شکل ۱: مثالی از مدل رفتار آزاد نوار نقاله (G1) و حسگر (G2)

²⁶ Regular Languages

²⁷ Generator

²⁸ Finite State Machine

²² Discrete Event System

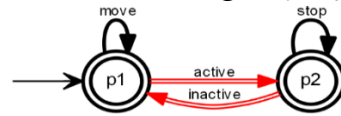
²³ Transition

²⁴ Free Behavior Model

²⁵ Control Specification

۲-۱-۳- مشخصه‌های کنترلی

رفتار مطلوب سیستم با n مشخصه کنترلی نمایش داده می‌شود. هر مشخصه کنترلی E_j ($j \in \{1, 2, \dots, n\}$) ارتباط دو یا چند مدل رفتار آزاد را نشان می‌دهد. شکل ۲ مشخصه‌ی تعیین‌کننده‌ی ارتباط بین مدل‌های نوار نقاله و حسگر را نمایش می‌دهد. قانون موجود در این مشخصه به صورت "هنگامی که یک محصول در مقابل حسگر قرار گیرد، نوار نقاله متوقف شده و در غیر این صورت باید حرکت نماید" است. SCT از وقوع برخی رخدادهای قابل کنترل در بعضی از حالت‌ها جلوگیری می‌کند. در مثال نوار نقاله و حسگر، در حالت q_1 رخداد $stop$ غیرفعال و رخداد $move$ فعال است. بنابراین هنگامی که حسگر غیرفعال است، نوار حرکت می‌کند.



شکل ۲: مشخصه کنترلی که موجب می‌شود نوار نقاله تنها در صورتی حرکت کند که محصولی در مقابل حسگر نباشد

۲-۲- محاسبه خودکار سوپروایزر

برای بدست آوردن سوپروایزر، مولدهای مربوط به مدل رفتار آزاد و مشخصه‌های کنترلی با هم ترکیب می‌شوند. در فرآیند ترکیب، مولد مربوط به مدل رفتار آزاد با توجه به مشخصه‌های کنترلی محدود می‌گردد. در واقع ربات تنها قادر به اجرای اعمالی است که توسط مشخصه‌های کنترلی مجاز اعلام گشته است. در این مقاله، نرم‌افزار طراحی شده که همه مراحل موردنیاز برای بدست آوردن سوپروایزر را به صورت خودکار انجام می‌دهد.

برای ترکیب دو مولد G_a و G_b با الفبای $\{a, b\}$ ، $i \in \sum_i$ از عملگر ترکیب موازی^{۲۹} به صورت نشان داده شده در فرمول ۱ استفاده می‌شود. این عملگر با نماد \parallel نمایش داده شده است.

$$G_a \parallel G_b = (Q_a \times Q_b, \sum_a \cup \sum_b, \delta_{a \parallel b}, (q_{0_a}, q_{0_b}), Q_{m_a} \times Q_{m_b}) \quad (1)$$

تابع انتقال $\delta_{a \parallel b}$ به صورت زیر محاسبه می‌گردد.

$$\delta_{a \parallel b}((q_a, q_b), e) = \begin{cases} (\delta_a(q_a, e), \delta_b(q_b, e)) & \text{if } \delta_a(q_a, e) \text{ and } \delta_b(q_b, e) \text{ defined} \\ (\delta_a(q_a, e), q_b) & \text{if } \delta_a(q_a, e) \text{ defined and } e \notin \sum_b \\ (q_a, \delta_b(q_b, e)) & \text{if } \delta_b(q_b, e) \text{ defined and } e \notin \sum_a \\ \text{undefined} & \text{otherwise} \end{cases} \quad (2)$$

فرمول (۲) اجازه می‌دهد رویدادهایی که بین دو مولد مشترک نیستند، به صورت غیرهمزمان عمل کنند. با ترکیب موازی مدل رفتار آزاد G و مشخصه کنترلی E ، مولد زبان هدف (K) یا سوپروایزر حاصل می‌شود:

$$K = G \parallel E \quad (3)$$

در این مرحله زبان $L(K)$ لزوماً کنترل‌شونده نیست. زبان K با الفبای \sum و مجموعه رویدادهای غیرقابل کنترل $\sum_u \subseteq \sum$ در صورتی کنترل‌شونده است که [۱۰]:

$$\forall s \in \overline{L(K)}, \forall e_u \in \sum_u, se_u \in L(G) \rightarrow se_u \in \overline{L(K)} \quad (4)$$

نماد \bar{L} بیانگر زبان پیشوند بسته شده‌ی زبان L است و به صورت

زیر تعریف می‌شود:

$$\bar{L} = \{s \in \sum^* : \exists t \in \sum^* \wedge st \in L\} \quad (5)$$

به بیانی دیگر، اگر s پیشوند یکی از رشته‌های تولید شده توسط مولد K باشد و e_u یک رویداد غیرقابل کنترل که از نظر فیزیکی امکان رخ دادن آن وجود دارد ($se_u \in L(G)$)، آنگاه زبان $L(K)$ در صورتی کنترل‌شونده است که رشته se_u نیز پیشوند یکی از رشته‌های تولید شده توسط مدل K باشد. در صورتی که اینچنین نباشد، زبان $L(K)$ غیرقابل کنترل و حالتی که در آن رویداد غیرقابل کنترل e_u رخ داده است، حالت بد نامیده می‌شود. برای استخراج زیرمجموعه‌ای از زبان $L(K)$ که کنترل‌شونده است، باید همه حالت‌های بد و سایر حالت‌هایی که توسط رویدادهای غیرقابل کنترل به حالت‌های بد می‌رسند، حذف گردند. به زبان حاصل، بزرگترین زیرمجموعه کنترل‌شونده زبان $L(K)$ گفته می‌شود [۱۱].

پس از بدست آوردن زبان کنترل‌شونده، نوبت به حذف بن‌بست‌ها می‌رسد. در صورتی که زبان دارای یکی از دو نوع بن‌بست مرده^{۳۰} و زنده^{۳۱} باشد، بلاک‌شونده است. برای حذف بن‌بست‌ها، حالت‌های غیرقابل دسترس از حالت ابتدایی باید حذف شوند. علاوه‌براین، حالت‌هایی که مسیری بین آن‌ها و حداقل یکی از حالت‌های علائتمدار وجود ندارد نیز باید حذف شوند. حذف این دو مجموعه از حالت‌ها، بلاک‌شونده نبودن زبان $L(K)$ را تضمین می‌کند [۱۱]. بنابراین، سوپروایزر نهایی به صورت نشان داده شده در فرمول ۶ بدست می‌آید. عملگر $SupC$ ، با توجه به مدل رفتار آزاد G ، ابتدا حالت‌های بد و سپس حالت‌های غیرقابل دسترس را از مولد K حذف می‌کند. مولد نهایی حاصل S کنترل‌شونده و بلاک‌نشونده است.

$$S = SupC(K, G) \quad (6)$$

تا اینجا، نحوه محاسبه سوپروایزر تشریح شد. در صورتی که تعداد مولدهای مدل رفتار آزاد و یا مشخصه‌های کنترلی بیش از یکی باشد، سه رویکرد برای محاسبه سوپروایزر مطرح می‌شود [۱۲]. در ادامه سه روش مربوطه تشریح داده شده است. بر اساس نتایج و مقایسه‌های ارائه شده در مرجع [۳]، روش سوم از جهت اندازه ماشین حالت سوپروایزر، بهتر از دو روش دیگر عمل می‌کند. به همین دلیل، در این مقاله نیز از این روش برای طراحی سوپروایزر استفاده شده است. پس از پایان طراحی مدل‌های رفتار آزاد و مشخصه‌های کنترلی موردنیاز، نرم‌افزار پیشنهادی به صورت خودکار، سوپروایزرهای ماژولار محلی را محاسبه کرده و کد موردنیاز برای اجرای آن‌ها در سکوی آرگوس را تولید می‌کند.

۱) سوپروایزر یکپارچه^{۳۲}

اگر همه مولدهای مربوط به مدل‌های رفتار آزاد و مشخصه‌های کنترلی باهم ترکیب شده و نتیجه یک سوپروایزر واحد گردد؛ سوپروایزر حاصل، یکپارچه نامیده می‌شود. این روش در چهار مرحله صورت می‌گیرد. ابتدا همه m مدل رفتار آزاد با هم ترکیب شده و یک مولد مدل رفتار آزاد واحد G^{mon} را تولید می‌کنند (فرمول ۷). سپس، n

³⁰ Deadlock

³¹ Livelock

³² Monolithic

²⁹ Parallel Composition

$$S_i^{loc} = SupC(K_i^{loc}, G_i^{loc}) \quad \forall i \in \{1, \dots, n\} \quad (16)$$

۲-۳- پیاده‌سازی سازی تئوری کنترل سوپروایزری بر روی سکوی آرگوس

برای فراهم آوردن امکان اجرای سوپروایزر در سکوی آرگوس، نیاز به پیاده‌سازی سه لایه نرم‌افزاری بر روی ربات است (شکل ۳). منطق لایه‌ها مشابه روند طی شده در [۳] می‌باشد. در پایین‌ترین لایه، رویه‌های عملیاتی^{۳۳} قرار می‌گیرد. وظیفه این لایه تبدیل سیگنال‌ها به رویدادها و برعکس است. در واقع این لایه، لایه سخت‌افزاری را از سایر لایه‌های نرم‌افزاری پنهان می‌کند. لایه دوم، اجراکننده ماشین حالت^{۳۴} است. این لایه یک ماشین حالت را دریافت کرده و در هر گام ربات، یک مرحله از ماشین حالت را اجرا می‌نماید. در بالاترین لایه، سوپروایزر قرار دارد که از یک یا چند ماشین حالت و تعدادی تکه‌کد تشکیل شده است. هر تکه‌کد در زمان رخداد یک رویداد قابل کنترل، اجرا می‌گردد.



شکل ۳: لایه‌های پیاده‌سازی شده بر روی لایه سخت‌افزار توسط روش پیشنهادی در سکوی آرگوس.

۳- پیاده‌سازی تجمع ربات‌ها با استفاده از تئوری کنترل سوپروایزری

استراتژی تجمع به گروهی از ربات‌ها اجازه می‌دهد تا در یک محیط همگن، گرد هم آیند [۱۳]. این استراتژی معمولاً در وظیفه‌های بزرگ‌تر و پیچیده‌تر، مورد استفاده قرار می‌گیرد. مرجع [۱۳] یک کنترل‌کننده واکنشی برای تجمع ربات‌ها پیشنهاد کرده است که نیاز به محاسبات بسیار کمی دارد و در نتیجه برای رباتیک جمعی بسیار مناسب است. برای پیاده‌سازی این روش، هر ربات باید به یک سنسور دودویی I مجهز باشد که در صورت مشاهده سایر ربات‌ها در مسیر دید ربات، مقدار یک و در غیراینصورت مقدار صفر را برمی‌گرداند. در صورتی که مقدار سنسور I برابر با صفر باشد، ربات در یک مسیر دایره‌ای شکل، با سرعت $(-0.7, -1) = (v_{l0}, v_{r0})$ و در جهت ساعتگرد رو به عقب حرکت می‌کند؛ و در صورتی که مقدار سنسور برابر با یک باشد، ربات با سرعت $(1, -1) = (v_{l1}, v_{r1})$ در جهت ساعتگرد به دور خود می‌چرخد. v_{r0} و v_{l0} به ترتیب، سرعت‌های موتور مربوط به چرخ‌های چپ و راست ربات هستند. مقادیر گزارش شده برای سرعت به بازه $[0, 1]$ نگاشت شده‌اند. این مقادیر در مرحله پیاده‌سازی به بازه واقعی سرعتی که ربات پشتیبانی می‌کند، برگردانده

مشخصه کنترلی با هم ترکیب شده و یک مولد مشخصه کنترلی واحد E^{mon} را تولید می‌کنند (فرمول ۸). آنگاه مولد K^{mon} محاسبه شده (فرمول ۹) که پس از اعمال عملگر $SupC$ منجر به تولید سوپروایزر نهایی S^{mon} خواهد شد (فرمول ۱۰).

$$G^{mon} = G_1 \parallel \dots \parallel G_m \quad (7)$$

$$E^{mon} = E_1 \parallel \dots \parallel E_n \quad (8)$$

$$K^{mon} = G^{mon} \parallel E^{mon} \quad (9)$$

$$S^{mon} = SupC(K^{mon}, G^{mon}) \quad (10)$$

۲ سوپروایزرهای ماژولار

به دلیل استفاده از عملگر ترکیب موازی، تعداد حالت‌های مولد حاصل از ترکیب، ممکن است بسیار بزرگ باشد. در نتیجه، انتقال چنین ماشین حالتی به ربات‌ها، نیاز به مقدار زیادی حافظه دارد که در بسیاری از موارد، امکان‌پذیر نیست. برای رفع این مشکل، ایده سوپروایزرهای ماژولار مطرح شد [۱۱]. این روش به ازای هر مشخصه کنترلی، یک سوپروایزر بدست می‌آورد که به صورت موازی اجرا می‌شوند. برای اعمال این روش، مانند روش قبل، ابتدا همه m مدل رفتار آزاد با هم ترکیب شده و یک مولد مدل رفتار آزاد واحد G^{mod} را تولید می‌کنند (فرمول ۱۱). سپس مولدهای K_i^{mod} به ترتیب برای تک‌تک مشخصه‌های کنترلی E_i محاسبه می‌شود (فرمول ۱۲). در نهایت، پس از اعمال عملگر $SupC$ به هر مولد K_i^{mod} سوپروایزر نهایی S_i^{mod} حاصل خواهد شد. (فرمول ۱۳).

$$G^{mod} = G^{mon} = G_1 \parallel \dots \parallel G_m \quad (11)$$

$$K_i^{mod} = G^{mod} \parallel E_i \quad \forall i \in \{1, \dots, n\} \quad (12)$$

$$S_i^{mod} = SupC(K_i^{mod}, G^{mod}) \quad \forall i \in \{1, \dots, n\} \quad (13)$$

برای استفاده از این روش، مشخصه‌های کنترلی نباید با یکدیگر تداخل داشته باشند. دو مولد S_1 و S_2 را متداخل می‌نامیم اگر مولد حاصل از ترکیب موازی آن‌ها دارای یکی از انواع حالت‌های غیرقابل دسترس باشد. برای رفع تداخل، باید دو مولد مربوطه را حذف کرده و مولد $S_1 \parallel S_2$ را جایگزین آن دو کرد [۱۲].

۳ سوپروایزرهای ماژولار محلی

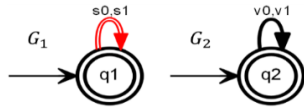
این روش هم در سطح مدل‌های رفتار آزاد و هم در سطح مشخصه‌های کنترلی به صورت ماژولار عمل می‌کنند. در روش قبل، هر مشخصه کنترلی با کل مدل‌های رفتار آزاد ترکیب می‌شد. در این روش، هر مشخصه کنترلی تنها با مدل‌های رفتار آزاد مرتبط با خودش، ترکیب می‌شود (اشتراک مجموعه رویدادهای شان تهی نباشد). استفاده از این روش نیز به عدم تداخل مشخصه‌های کنترلی وابسته است. برای محاسبه سوپروایزرهای ماژولار محلی، ابتدا مدل‌های رفتار آزاد مرتبط با هر مشخصه کنترلی E_i با هم ترکیب شده و مدل رفتار آزاد واحد G_i^{loc} را بدست می‌دهند (فرمول ۱۴). سپس مولدهای K_i^{loc} به ترتیب برای تک‌تک مشخصه‌های کنترلی E_i محاسبه می‌شود (فرمول ۱۵). در نهایت، پس از اعمال عملگر $SupC$ به هر مولد K_i^{loc} سوپروایزر نهایی S_i^{loc} بدست خواهد آمد. (فرمول ۱۶).

$$G_i^{loc} = G_{k_1} \parallel \dots \parallel G_{k_r} \quad \forall i \in \{1, \dots, n\}, \forall k_i \in \{1, \dots, m\} \wedge (\sum G_{k_i} \cap \sum E_i) \neq \emptyset \quad (14)$$

$$K_i^{loc} = G_i^{loc} \parallel E_i \quad \forall i \in \{1, \dots, n\} \quad (15)$$

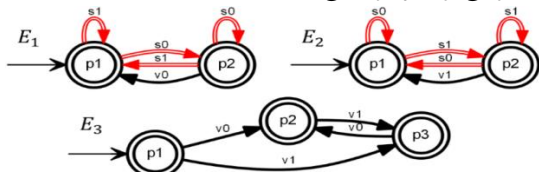
³³ Operational Procedures

³⁴ Automata Player



شکل ۵: مدل‌های رفتار آزاد مربوط به استراتژی تجمع ربات‌ها

شکل ۶، مشخصه‌های کنترلی مورد نیاز برای پیاده‌سازی استراتژی تجمع ربات‌ها را نشان داده است. مشخصه کنترلی E_1 ، رویداد v_0 را تنها در صورتی فعال^{۳۵} می‌کند که هیچ شیء‌ای در معرض دید ربات وجود نداشته باشد. مشخصه کنترلی E_2 ، رویداد v_1 را تنها در صورتی فعال^{۳۶} می‌کند که شیء‌ای در معرض دید ربات وجود داشته باشد. وقتی سرعت موتور ربات تنظیم می‌شود، ربات تا بی‌نهایت بر اساس آن سرعت حرکت می‌کند. بنابراین نیازی به اتفاق افتادن مداوم رویدادهای v_0 یا v_1 وجود ندارد. هدف از مشخصه کنترلی E_3 همین امر است؛ یعنی اجرای نوبتی رویدادهای v_0 و v_1 . عدم وجود مشخصه کنترلی E_3 موجب عملکرد اشتباه ربات‌ها نخواهد شد و تنها از یک امر بی‌اثر جلوگیری می‌کند.



شکل ۶: مشخصه‌های کنترلی مربوط به استراتژی تجمع ربات‌ها

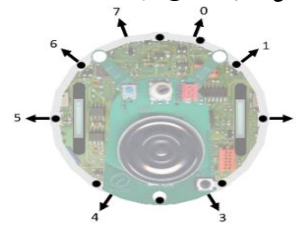
شکل ۷، سوپروایزر یکپارچه و سوپروایزرهای ماژولار محلی بدست آمده از ترکیب شکل ۵ و شکل ۶ را به تصویر کشیده است. این سوپروایزرها و تصاویر نمایش داده شده توسط نرم‌افزار پیشنهادی محاسبه شده‌اند. برای اجرای استراتژی تجمع ربات‌ها در محیط شبیه‌سازی و یا در دنیای واقعی، کافی است که یکی از این دو سوپروایزر به عنوان نرم‌افزار کنترلی ربات‌ها مورد استفاده قرار گیرند. نتیجه استفاده از هر یک از دو نوع سوپروایزر برابر است، هرچند مجموع تعداد حالت‌ها و رویدادهای سوپروایزرهای ماژولار محلی بسیار کمتر است. این تفاوت در مولدهای بزرگتر، بارزتر خواهد شد.

۴- نتایج آزمایش‌ها

تمامی آزمایش‌های انجام شده در این مقاله بر روی سکوی آرگوس صورت گرفته است. سکوی آرگوس یک شبیه‌ساز برای سیستم‌های چندرباته است که تاکید ویژه‌ای بر روی مقیاس جمعیت ربات‌ها دارد. این سکو می‌تواند جمعیت‌های چند هزار رباته را به شکلی کاملاً بهینه و با سرعت مناسب شبیه‌سازی کند. ویژگی‌هایی از قبیل پشتیبانی از ربات‌های متنوع، ماژولار بودن، شبیه‌سازی چندمنجه^{۳۷} و امکان سفارشی‌سازی بالا، موجب انتخاب آن برای انجام شبیه‌سازی‌های این مقاله شده است. یکی از قابلیت‌های منحصر بفرد آرگوس، امکان کامپایل کد کنترلی برای ربات‌های واقعی است؛ بنابراین نیازی به پیاده‌سازی دو برنامه متفاوت، یکی به هدف شبیه‌سازی و یکی به هدف ربات واقعی وجود ندارد.

خواهند شد. مرجع [۱۳] این منطق ساده را برای تجمع ۴۰ ربات به کار گرفته و نتیجه مناسبی بدست آورده است. این مرجع از ربات E-Puck [۱۴] برای انجام آزمایش بهره برده است. برای پیاده‌سازی سنسور I ، از دوربین VGA در این ربات استفاده شده است. این دوربین تصاویر با وضوح $640 * 480$ برمی‌گرداند؛ حافظه E-Puck به قدری کوچک است که امکان ذخیره‌سازی یک تصویر با این ابعاد را ندارد. به همین دلیل، مرجع [۱۳] پس از کاهش ابعاد تصویر، تنها یک ستون به ارتفاع ۱۵ پیکسل و عرض ۱ پیکسل را استخراج کرده و با رای‌گیری از رنگ پیکسل‌های موجود در این ستون، در مورد وجود یا عدم وجود شیء در مقابل ربات تصمیم می‌گیرد.

در این مقاله، روش توضیح داده شده در بالا با مقداری تغییر و با استفاده از تئوری کنترل سوپروایزر پیاده‌سازی شده است. روش پیاده‌سازی شده در این مقاله از سنسور مجاورت سنج به جای دوربین بهره برده است. محدودیت استفاده از سنسور بسیار بیشتر از دوربین است. زیرا بازه دید دوربین E-Puck در حدود ۱۴۰ سانتی‌متر است، در صورتی که سنسور مجاورت سنج دارای برد ۸ سانتی‌متری است. علاوه بر این، با استفاده از دوربین می‌توان وجود اشیاء رنگی و یا خاکستری را تشخیص داد. در صورتی که سنسور مجاورت سنج تنها وجود اشیاء در مجاورت ربات را گزارش می‌کند و هیچ اطلاعاتی از نوع شیء مربوطه در اختیار قرار نمی‌دهد. بنابراین پیاده‌سازی تجمع ربات‌ها تنها با استفاده از سنسور مجاورت سنج کار بسیار دشواری است که در این مقاله مورد بررسی قرار گرفته است. E-Puck دارای هشت سنسور مجاورت است که در شکل ۴ نمایش داده شده‌اند. روش پیشنهادی از ترکیب سنسورهای شماره صفر و هفت برای تشخیص سایر ربات‌ها بهره می‌برد. خروجی این سنسورها مقداری در بازه [۰-۱] است؛ صفر به معنی عدم وجود شیء و یک به معنی وجود شیء در نزدیک‌ترین فاصله ممکن به سنسور. برای ترکیب دو سنسور شماره صفر و هفت، بیشینه مقدار خروجی آن‌ها مورد استفاده قرار گرفته است. در ضمن، طراحی سوپروایزر کاملاً بر اساس مراحل توضیح داده شده در بخش‌های قبل صورت می‌گیرد.



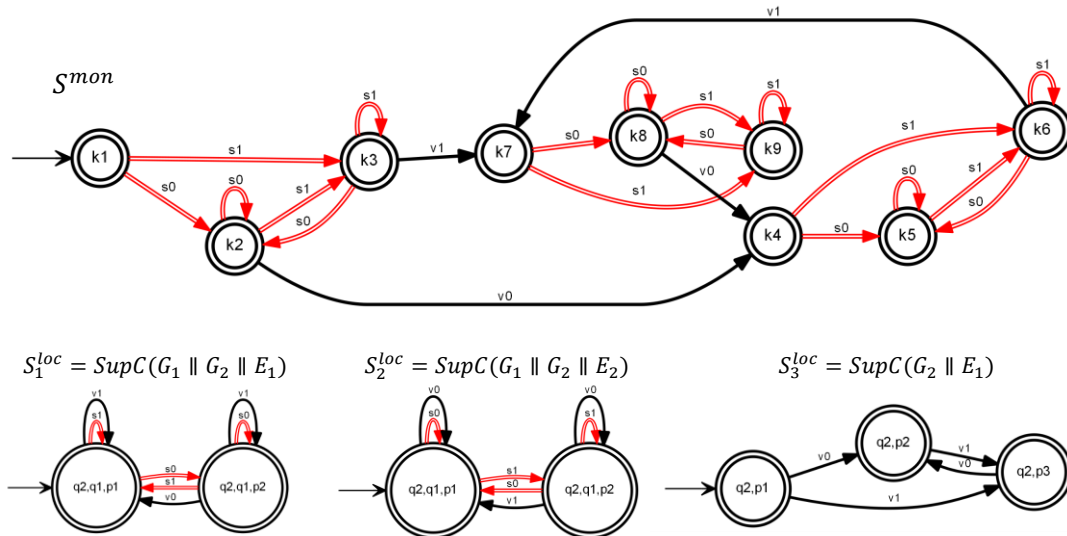
شکل ۴: مکان سنسورهای مجاورت سنج ربات E-Puck که از نمای بالا نمایش داده شده‌اند

در شکل ۵، مدل‌های رفتار آزاد مربوط به تجمع ربات‌ها را نمایش داده است. مدل رفتار آزاد G_1 بیانگر سنسور دودویی مورد بحث است. رویدادهای غیرقابل کنترل s_0 و s_1 به ترتیب معادل دو مقدار خروجی $I = 1$ و $I = 0$ می‌باشند. مدل رفتار آزاد G_2 حرکت‌های ممکن ربات را توصیف می‌کند. رویدادهای قابل کنترل v_0 و v_1 به ترتیب معادل دو سرعت موتورهای ربات (v_{r0}, v_{r1}) و (v_{l0}, v_{l1}) می‌باشند. هر حالت در این دو مولد، هم حالت ابتدایی و هم علامت‌دار است.

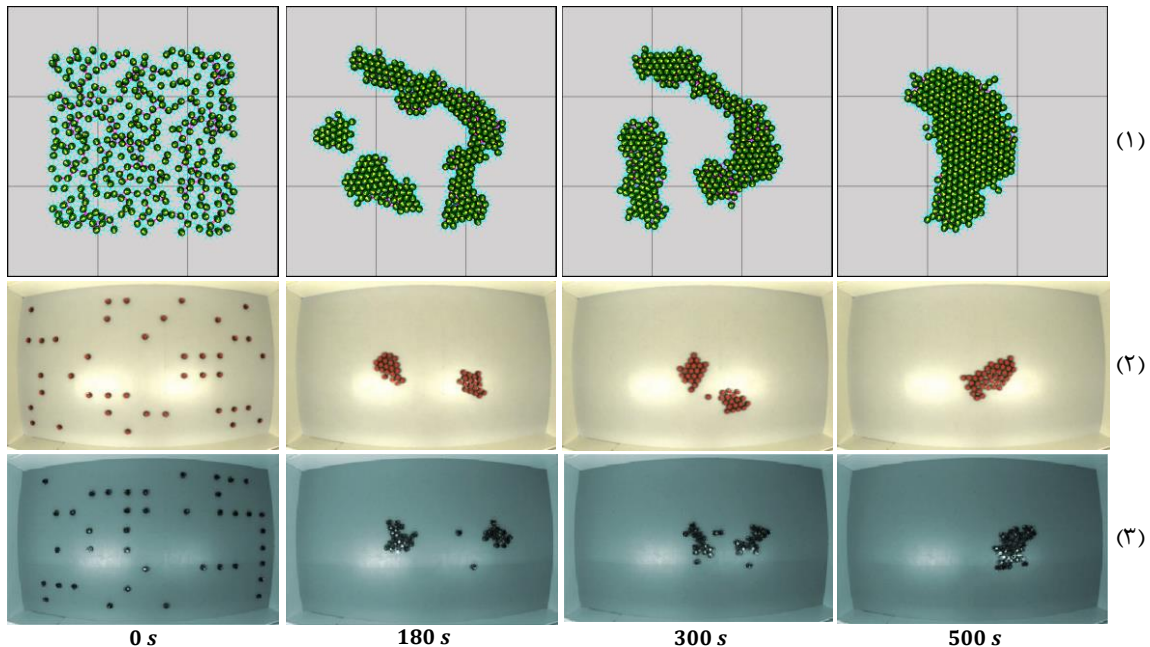
³⁵ Enabled

³⁶ Enabled

³⁷ Multi-thread



شکل ۷: سوپروایزر یکپارچه (S^{mon}) و سوپروایزرهای ماژولار محلی مربوط به استراتژی تجمع ربات‌ها



شکل ۸: مقایسه دنباله تصاویر گرفته شده در زمان‌های صفر، ۱۸۰، ۳۰۰ و ۵۰۰ ثانیه از آزمایش تجمع ربات‌ها. (۱) روش پیشنهادی این مقاله با ۳۰۰ ربات، (۲) روش پیشنهادی در [۱۳] با ۴۰ ربات و (۳) روش پیشنهادی در [۳] با ۴۰ ربات.

۳۰۰ سانتی‌متر و ارتفاع نیم متر. برای افزایش سرعت شبیه‌سازی، در هر ثانیه، ۱۰ دور از الگوریتم اجرا می‌شود. دوربین در ارتفاع دو و نیم متری از محیط شبیه‌سازی قرار داده شده و در زمان‌های مشخص، تصاویری تهیه می‌کند.

نتیجه آزمایش انجام شده در شکل ۸ نمایش داده شده است. نتیجه بدست آمده با دو روش از جدیدترین روش‌های موجود مورد مقایسه قرار گرفته است [۳، ۱۳]. هر دو روش مورد مقایسه از دوربین VGA ربات استفاده کرده‌اند که برد چند ده برابری نسبت به سنسور مجاورت‌سنج ربات دارد. با این حال، مشاهده می‌شود که نتیجه بدست آمده توسط روش پیشنهادی کاملاً مناسب است. علاوه بر این، تعداد ربات‌های مورد استفاده در آزمایش انجام شده توسط این مقاله تقریباً هفت برابر بیشتر از سایر روش‌های مقایسه شده، می‌باشد.

برای انتقال سوپروایزرها به حافظه داخلی ربات‌ها از روش پیشنهاد شده توسط [۱۵] استفاده شده است. این روش کمترین مقدار از حافظه ربات را اشغال کرده و قابل پیاده‌سازی بر روی انواع ربات‌های موجود است. مقدار حافظه اشغالی برای یک مولد با s حالت و t گذار برابر است با $s + 3t$ بایت. بنابراین برای نگهداری سوپروایزرهای یکپارچه و ماژولار محلی محاسبه شده در بخش قبل به ترتیب به ۷۵ بایت و ۶۱ بایت حافظه نیاز است.

برای انجام شبیه‌سازی‌ها از ۳۰۰ ربات E-Puck بهره برده شده است. به ورودی سنسور مجاورت‌سنج، نویز گوسی با انحراف معیار ۰٫۳ اعمال شده است تا به رفتار دنیای واقعی نزدیک‌تر شود. نویز مشابهی به عملکرد موتورهای ربات نیز اعمال گشته است. ابعاد محیط مورد استفاده برای انجام آزمایش‌ها برابر است با ۳۰۰ سانتی‌متر در

- [4] Knight, J.C., DeJong, C.L., Gibble, M.S., Nakano, L.G. "Why Are Formal Methods Not Used More Widely?", Fourth NASA formal methods workshop, pp. 1-12, NASA, 1997.
- [5] Brambilla, M., Dorigo, M., Birattari, M. "Property-Driven Design for Robot Swarms: A Design Method Based on Prescriptive Modeling and Model Checking.", *ACM Transactions on Autonomous and Adaptive Systems.*, vol. 9, no. 4, pp. 17, 2015.
- [6] Francesca, G., Brambilla, M., Brutschy, A., Trianni, V., Birattari, M. "AutoMoDe: A Novel Approach to the Automatic Design of Control Software for Robot Swarms.", *Swarm Intelligence.*, vol. 8, no. 2, pp. 89-112, 2014.
- [7] King, J., Pretty, R.K., Gosine, R.G. "Coordinated Execution of Tasks in a Multiagent Environment.", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans.*, vol. 33, no. 5, pp. 615-619, 2003.
- [8] Winfield, A.F., Sa, J., Fernández-Gago, M.-C., Dixon, C., Fisher, M. "On Formal Specification of Emergent Behaviours in Swarm Robotic Systems.", *International journal of advanced robotic systems.*, vol. 2, no. 4, pp. 39, 2005.
- [9] Pinciroli, C., Trianni, V., O'Grady, R., Pini, G., Brutschy, A., Brambilla, M., Mathews, N., Ferrante, E., Di Caro, G., Ducatelle, F., Birattari, M., Gambardella, L.M., Dorigo, M. "ARGoS: A Modular, Parallel, Multi-Engine Simulator for Multi-Robot Systems.", *Swarm Intelligence.*, vol. 6, no. 4, pp. 271-295, 2012.
- [10] Cassandras, C.G., Lafortune, S. *Introduction to discrete event systems.*, Springer Science & Business Media, 2009.
- [11] Ramadge, P.J., Wonham, W.M. "Supervisory Control of a Class of Discrete Event Processes.", *SIAM journal on control and optimization.*, vol. 25, no. 1, pp. 206-230, 1987.
- [12] Lopes, Y.K., Leal, A.B., Rosso, R.S.U., Harbs, E. "Local Modular Supervisory Implementation in Microcontroller.", *Proceedings of the 9th International Conference of Modeling, Optimization and Simulation*, 2012.
- [13] Gauci, M., Chen, J., Li, W., Dodd, T.J., Groß, R. "Self-Organized Aggregation without Computation.", *The International Journal of Robotics Research.*, vol. 33, no. 8, pp. 1145-1161, 2014.
- [14] Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klaptocz, A., Zufferey, J., Floreano, D., Martinoli, A. "The E-Puck, a Robot Designed for Education in Engineering.", *Proceedings of the 9th conference on autonomous robot systems and competitions*, pp. 59-65, 2009.
- [15] Lopes, Y.K., Trenkwalder, S.M., Leal, A.B., Dodd, T.J., Groß, R. "Probabilistic Supervisory Control Theory (pSCT) Applied to Swarm Robotics.", *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pp. 1395-1403, 2016.

همان‌طور که در شکل ۸ مشاهده می‌شود، ربات‌ها ابتدا دسته‌های مجتمع کوچک تشکیل داده و سپس این دسته‌ها در هم ادغام شده و دسته‌های بزرگتر را ساخته‌اند. این رفتار در [۱۳] نیز گزارش شده است. این آزمایش ۲۰ مرتبه تکرار شده است تا یک میانگین از رفتار ربات‌ها بدست آید. در ۱۷ مرتبه از ۲۰ آزمایش صورت گرفته، ربات‌ها بعد از ۵۰۰ ثانیه در یک دسته واحد تجمع کردند؛ در سه مرتبه باقیمانده، دو دسته از ربات‌ها تشکیل شد. به دلیل برد بسیار کم سنسور مجاورت‌سنج، در صورتی که فاصله ربات‌ها از هم زیاد باشد، تجمع آن‌ها نیاز به زمان بسیار بیشتری نسبت به روش‌های ارائه شده در [۳، ۱۳] دارد.

۵- نتیجه‌گیری

در این مقاله، یکی از کم‌هزینه‌ترین استراتژی‌های تجمع ربات‌ها با استفاده از تئوری کنترل سوپروایزری پیاده‌سازی و شبیه‌سازی شد. رباتیک جمعی دارای محدودیت‌های زیادی است که حل مسائل به کمک آن را دشوار می‌سازد. یکی از این موارد، قدرت محاسباتی و ارتباطی پایین ربات‌ها است. مورد دوم، جمعیت بالای ربات‌ها است. هر دوی این موارد در روش پیشنهادی کاملاً رعایت شده‌اند. استفاده از سنسور مجاورت‌سنج به جای دوربین VGA، بار محاسباتی را بیش از پیش کاهش داده است. استفاده از ۳۰۰ ربات برای انجام آزمایش‌ها نیز دشواری مسئله را افزایش داد؛ این تعداد تقریباً هفت برابر تعداد ربات‌هایی بود که در کارهای پیشین مورد آزمایش قرار گرفت. برای انجام شبیه‌سازی‌ها از سکوی قدرتمند آرگوس بهره گرفته شد. بنابراین نیاز به پیاده‌سازی SCT بر روی این سکو وجود داشت که به درستی انجام شد. آنگاه نرم‌افزاری برای سنتز سوپروایزری پیاده‌سازی شد که تمامی مراحل سنتز را به صورت خودکار انجام می‌داد و در نهایت کد کنترلی موردنیاز برای اجرای شبیه‌سازی در محیط آرگوس را تولید می‌کرد. نتایج آزمایش‌ها روشنگر عملکرد مناسب روش پیشنهادی است که با وجود محدودیت‌های بیشتر نسبت به روش‌های پیشین، کارایی برابری نسبت به آن‌ها دارد. برد کمتر سنسور مورد استفاده برای تشخیص وجود اشیاء در روبروی ربات، دارای یک ضعف نسبت به روش‌های پیشین است. در مواردی که فاصله ربات‌ها از هم زیاد باشد، نیاز به زمان بیشتری برای تجمع کامل ربات‌ها وجود دارد. برای پوشش این ضعف، می‌توان از دوربین همه‌جانبه ربات E-Puck بهره برد که دارای برد چند برابری نسبت به سنسور مجاورت است. استفاده از این ویژگی می‌تواند در کارهای آینده مورد بررسی قرار گیرد. انجام این آزمایش بر روی ربات‌های واقعی نیز در اولویت‌های کارهای آینده قرار دارد.

مراجع

- [1] Prasad, S., Rawool, S. "Swarm Robotics: Nature Inspired Systems.", *International Journal of Engineering Research and General Science.*, vol. 4, no. 5, pp. 168-174, 2016.
- [2] Dorigo, M., Roosevelt, A.F.. "Swarm Robotics.", *Autonomous Robots.*, no. Special Issue, 2014.
- [3] Lopes, Y.K., Trenkwalder, S.M., Leal, A.B., Dodd, T.J., Groß, R. "Supervisory Control Theory Applied to Swarm Robotics.", *Swarm Intelligence.*, vol. 10, no. 1, pp. 65-97, 2016.



ارائه روش هوشمند برای طراحی سیستم‌های ANFIS

امیر سلطانی محبوب^۱، سید حمید ظهیری^۲

^۱ دانشجوی کارشناسی ارشد، دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند

amir.s.m@birjand.ac.ir

^۲ استاد، دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند

hzahiri@birjand.ac.ir

چکیده

مهم‌ترین موضوع در طراحی سیستم‌های ANFIS، تعیین متغیرهای بهینه‌ی فازی اعم از نوع و مکان توابع عضویت و روش آموزش آن است. در واقع این‌ها پارامترهای ساختاری یک سیستم ANFIS هستند که طراح سعی می‌کند با یافتن مقادیر بهینه آن‌ها و روش آموزش مناسب، به بهترین عملکرد سیستم ANFIS دست یابد. در این مقاله، یک روش جدید مبتنی بر الگوریتم بهینه‌سازی صفحات شیب‌دار (IPO) برای اولین بار به جهت طراحی بهینه و هوشمند یک طبقه بند ANFIS پیشنهاد و پیاده‌سازی شده است. روش پیشنهادی برای طراحی طبقه بند ANFIS در مواجهه با مجموعه داده‌هایی با تعداد کلاس‌های مرجع متفاوت، طول بردارهای ویژگی متفاوت که از پیچیدگی قابل قبولی برخوردارند، به محک آزمایش سپرده شده است. نتایج مقایسه‌ای به دست آمده حاصل از این پژوهش نشان می‌دهد که روش ارائه شده برای طراحی بهینه و هوشمند یک طبقه بند ANFIS به وسیله‌ی الگوریتم بهینه‌سازی صفحات شیب‌دار (IPO)، نسبت به یک طبقه بند ساده ANFIS از دقت بالاتر و عملکرد بهتری برخوردار می‌باشد.

کلمات کلیدی

بازشناسی الگو، طبقه‌بندی، سیستم‌های ANFIS، الگوریتم بهینه‌ساز صفحات شیب‌دار

تغییر در عملکرد سیستم‌های ANFIS خواهد شد. از این رو، شیوه‌های متفاوتی برای یافتن مقادیر بهینه پارامترهای مذکور و روش آموزش آن به منظور رسیدن به بهترین عملکرد سیستم‌های ANFIS پیشنهاد شده است [4-5].

اخیرا در طراحی سیستم‌های ANFIS برای یافتن این پارامترهای ساختاری از روش‌های تکاملی استفاده می‌شود [6-7] با توجه به موارد ذکر شده در این مقاله، از الگوریتم ابتکاری و هوشمند بهینه‌سازی صفحات شیب‌دار (IPO) [8] که بر پایه دینامیک حرکت توپ‌ها بروی صفحات شیب‌دار بدون اصطکاک است برای طراحی بهینه و هوشمند (آموزش، انتخاب هوشمند و بهینه توابع عضویت) طبقه بند داده ANFIS به جهت بالا بردن دقت و عملکرد بهتر نسبت به روش طراحی کلاسیک در سیستم طبقه بند ANFIS استفاده شده و سپس مقایسه‌ای میان عملکرد روش پیشنهادی (برای طراحی هوشمند و بهینه طبقه بند ANFIS توسط روش IPO) با طبقه بند ساده ANFIS که با روش‌های هوشمند (IPO+PSO [9] و GA [10-11]) ترکیب شده توسط اعمال بر چهار مجموعه داده‌های محک گل زنیق، مجموعه داده‌های سرطان

(۱) مقدمه

با توجه به اینکه سیستم‌های ANFIS از قواعد شفاهی برای توصیف مسئله استفاده می‌کنند، قابلیت استفاده در کاربردهای گوناگون را دارند. یکی از مهم‌ترین اهداف در طراحی سیستم‌های ANFIS روش آموزش و همچنین توابع عضویت کارآمد است. این کار در کاربردهای معمولی به صورت دستی و به روش سعی و خطا انجام می‌پذیرد، اما با افزایش ابعاد فضای پاسخ مسئله دستیابی به بهترین پارامترهای ساختاری برای طراحی سیستم‌های ANFIS به صورت دستی و با استفاده از روش‌های سعی و خطا معمولاً کاری طاقت‌فرسا و گاه غیرممکن است. به‌ویژه که با افزایش فضای پاسخ مسئله، تعداد بهترین نوع تابع عضویت و محل استقرار آن‌ها به صورت چشم‌گیری افزایش پیدا می‌کند [1-2].

ساختار سیستم‌های ANFIS [3] شامل قواعد فازی (اگر - آنگاه) و توابع عضویت می‌باشد و پارامترهای ساختاری آن تاثیر چشمگیری بر عملکرد آن دارد، در واقع، تغییر هر یک از پارامترهای ساختاری منجر به

به‌طوری‌که، x_i^d موقعیت توپ i ام در بعد d ام در فضای n بعدی است. هدف پیدا کردن موقعیت $f(x_i^d)$ مینیمم تابع هدف می‌باشد که بروی فضای جستجو تعریف شده است. در یک زمان مشخص مانند t ، زاویه بین توپ i ام و توپ j ام در بعد d یعنی θ_{ij}^d به صورت زیر حساب می‌گردد:

$$\theta_{ij}^d = \tan^{-1} \left(\frac{f_i(t) - f_j(t)}{x_i^d - x_j^d} \right), \quad d = 1, 2, \dots, n \quad (2)$$

$i, j = 1, 2, \dots, N \quad i \neq j$

به‌گونه‌ای که $f_i(t)$ مقدار تابع هدف (ارتفاع) برای توپ i ام در زمان t است. بنا بر تجربه و آزمایش و برای کاهش حجم محاسباتی، جهت رسیدن به جواب‌های بهتر، شتاب هر توپ در هر بعد به صورت جداگانه از دیگر ابعاد محاسبه می‌شود. برای اینکه توپ‌ها به سمت پایین سرازیر شوند، در محاسبه شتاب آن‌ها، توپ‌هایی که در ارتفاع پایین‌تر قرار دارند در نظر گرفته می‌شوند. طبق قانون دوم نیوتن داریم:

$$\sum \vec{f} = m \cdot \vec{a} \quad (3)$$

برای شتاب هم داریم:

$$a = g \cdot \sin(\theta) \quad (4)$$

که در آن g ثابت گرانش زمین است. در الگوریتم IPO با اعمال یکسری تغییرات، شتاب توپ i ام در زمان t ام در بعد d ام به صورت زیر محاسبه می‌شود:

$$a_i^d(t) = \sum_{j=1}^N U(f_j(t) - f_i(t)) \cdot \sin(\theta_{ij}^d(t)) \quad (5)$$

بطوریکه U تابع پله واحد است. در هر تکرار الگوریتم برای بروز شدن و تعیین موقعیت جدید هر توپ پس از محاسبه شتاب آن از فرمول (۵) و این نکته که حرکت از نوع حرکت با شتاب ثابت است، می‌توان از فرمول موقعیت به صورت زیر استفاده کرد که در آن ضرایب k_1 ، k_2 ضرایب متغیر بازمان برای کنترل بهتر عملکرد IPO است.

$$x_i^d(t+1) = k_1 \cdot \text{rand}1 \cdot a_i^d(t) \cdot \Delta t^2 + \quad (6)$$

$$k_2 \cdot \text{rand}2 \cdot v_i^d(t) \cdot \Delta t + x_i^d(t)$$

$$k_1(t) = \frac{c_1}{1 + \exp((t - \text{shift}_1) \times \text{scale}_1)} \quad (7)$$

$$k_2(t) = \frac{c_2}{1 + \exp((t - \text{shift}_2) \times \text{scale}_2)} \quad (8)$$

در فرمول شماره (۶) $\text{rand}1$ و $\text{rand}2$ ضرایب وزنی به صورت تصادفی در بازه $[0,1]$ هستند تا به الگوریتم خاصیت تصادفی داده شود و در فرمول‌های (۷) و (۸) مقادیر c_1 ، c_2 ، shift_1 ، shift_2 ، scale_1 و scale_2 ثوابتی هستند که برای هر تابع به صورت تجربی به دست می‌آیند.

مقدار $v_i^d(t)$ سرعت توپ i ام در بعد d و در زمان t است و به صورت زیر تعریف می‌شود:

$$v_i^d(t) = \frac{x_{\text{best}}^d(t) - x_i^d(t)}{\Delta t} \quad (9)$$

۳) سیستم‌های ANFIS، معرفی توابع عضویت، طراحی بهینه و هوشمند طبقه بند ANFIS توسط روش IPO

در این بخش در ابتدا به صورت خلاصه ساختار سیستم استنتاج عصبی-فازی (ANFIS) مرور خواهد شد. سپس در بخش بعد، به

سینه، داده‌های Pima و داده‌های Bupa Liver از مخزن پایگاه داده (UCI) انجام گرفته است. سازمان‌دهی کلی مقاله به شکل زیر است: در بخش ۲ به معرفی الگوریتم بهینه‌سازی صفحات شیب‌دار پرداخته می‌شود. در بخش ۳ سیستم‌های ANFIS، معرفی توابع عضویت، روش طراحی بهینه و هوشمند (انتخاب بهینه نوع و مکان توابع عضویت و همچنین روش آموزش) طبقه بند ANFIS توسط الگوریتم IPO توصیف می‌شود. در بخش ۴ عملکرد طبقه بند ANFIS ارزیابی و نهایتاً در بخش ۵ نتیجه‌گیری بیان می‌شود.

۲) الگوریتم بهینه‌سازی صفحات شیب‌دار

روش الگوریتم بهینه‌سازی سیستم صفحات شیب‌دار که به اختصار IPO می‌نامیم، الهام گرفته از نحوه حرکت دینامیکی اجسام کروی روی سطح شیب‌دار بدون اصطکاک است که همگی تمایل دارند تا به پایین‌ترین نقطه سطح برسند. در IPO همان‌طور که در شکل (۱) نمونه‌ای از فضای جستجو به همراه سه توپ نشان داده شده است، تعدادی توپ کوچک فضای جستجو را برای یافتن نقطه بهینه (در اینجا مینیمم) جستجو می‌کنند. این توپ‌های کوچک در IPO عوامل جستجو هستند مانند ذرات در الگوریتم بهینه‌سازی اجتماع ذرات [9] (PSO). ایده اصلی الگوریتم مزبور نسبت دادن ارتفاع نسبت به یک نقطه مرجع به هر توپ است که میزان ارتفاع بر اساس تابع برازندگی به دست می‌آید. این مقادیر ارتفاع در واقع تخمینی برای انرژی پتانسیل توپ‌ها که در ارتفاعات مختلف قرار دارند و با پایین آمدن آن‌ها، این انرژی به انرژی جنبشی تبدیل شده و باعث شتاب گرفتن توپ‌ها می‌شود. در واقع، توپ‌ها تمایل دارند تا انرژی پتانسیل خود را برای رسیدن به نقطه مینیمم از دست دهند. لذا توپ‌ها به صورت مکرر بروی فضای جستجو برای یافتن پاسخ بهتر حرکت کرده و به تناسب شتابی را به خود اختصاص می‌دهند.

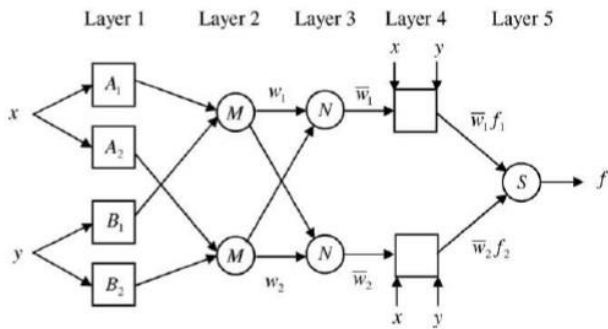


شکل ۱: نمونه‌ای از فضای جستجو به همراه سه توپ

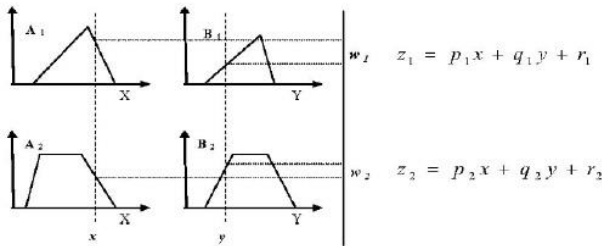
هر توپ در فضای جستجو دارای سه مختصات می‌باشد: موقعیت، ارتفاع و زاویه‌ای که با دیگر توپ‌ها می‌سازد. موقعیت هر توپ یک جواب در فضای جستجو و ارتفاع آن به وسیله تابع برازندگی به دست می‌آید. برای داشتن تخمینی از صفحه شیب‌دار که توپ‌ها بروی آن قرار دارند از خطوطی مستقیم استفاده می‌شوند که مراکز توپ‌ها را به هم وصل می‌کنند. این خطوط با خط راست افقی عبوری از مرکز توپ، زاویه‌ای را تشکیل می‌دهند که به وسیله آن و میزان شتاب اختصاص یافته به هر توپ، تعیین شود. سیستمی با N توپ را در نظر بگیرید. موقعیت i ام به وسیله فرمول زیر تعریف می‌گردد:

$$\vec{x}_i = (x_i^1, x_i^2, \dots, x_i^n) \quad i = 1, 2, \dots, N \quad (1)$$

$$x_j^{\min} < x_j < x_j^{\max}, \quad 1 < j < n$$



شکل ۲: ساختار ANFIS [3]



شکل ۳: منطق فازی ANFIS [3]

همان‌طور که مشاهده شد در ساختار ANFIS دولایه تطبیقی (لایه اول ولایه چهارم) وجود دارد که در لایه اول پارامترهای قابل اصلاح (c_i, b_i, a_i) مرتبط با ورودی‌های تابع عضویت وجود دارند و در لایه چهارم پارامترهای قابل اصلاح (q_i, r_i, p_i) مرتبط با چندجمله‌ای مرتبه اول مدل Sugeno می‌باشد.

۲-۳- معرفی توابع عضویت

در این بخش قبل از اینکه وارد مرحله طراحی بهینه و هوشمند طبقه بند ANFIS شویم نوع توابع عضویت که طبق جدول شماره (۱) شامل توابع عضویت: مثلثی (triangle)، ذوزنقه‌ای (trapezoid)، گوسی (gaussian)، ناقوس شکل (bell) و حلقوی از چپ یا راست (sig) و تابع عضویت (Pi) شکل است توصیف می‌شوند.

جدول ۱: توابع عضویت

TriM	TrapM	GbellM	Gauss	PiM	DsigM	PsigM
F	F	F	MF	F	F	F

۳-۳- به‌کارگیری روش IPO برای طراحی بهینه و هوشمند طبقه بند ANFIS

در استفاده از روش‌های بهینه‌سازی، برای طراحی بهینه و هوشمند طبقه بند ANFIS موضوع قابل توجه تعریف تابع ارزیابی است که روش بهینه سازی مورد استفاده، مقدار بهینه آن (ماکزیمم یا مینیمم) را می‌یابد. درواقع، اهمیت تابع ارزیابی از این جهت است که، مسیر جستجو را هدایت می‌کند.

۱-۳-۳- تعریف تابع ارزیابی

در طبقه بند Anfis مورد نظر هدف رسیدن به کمترین خطای طبقه‌بندی و بیشترین قابلیت تفسیرپذیری می‌باشد. این هدف در این مقاله، با

معرفی توابع عضویت و به‌کارگیری روش IPO برای طراحی بهینه و هوشمند طبقه بند ANFIS توضیح داده می‌شود.

۳-۱- ساختار سیستم استنتاج عصبی-فازی (ANFIS) تطبیقی

از ترکیب موفقیت‌آمیز شبکه عصبی، منطق فازی و محاسبات تکاملی مدل جدیدی بنام ANFIS به وجود آمده است. درواقع این مدل یک سیستم ترکیبی است که در آن هم از توانایی تصمیم‌گیری منطق فازی و هم قابلیت یادگیری شبکه عصبی استفاده شده است. به همین دلیل ANFIS در زمینه آموزش، ایجاد و طبقه‌بندی دارای توانایی مناسبی است. این مدل شامل قسمت‌هایی از سیستم فازی است که محاسبات در هر گام به وسیله نرون‌های لایه‌ی پنهان و قدرت یادگیری شبکه عصبی تولید می‌شود تا باعث افزایش دانش سیستم گردد [12].

برای ارائه ساختار شبکه ANFIS، دو قاعده فازی بر مبنای مدل مرتبه اول (sugeno) (اگر-آنگاه) به صورت زیر در نظر گرفته شده است:

$$\text{قاعده ۱: اگر } (A_1 \text{ در } x) \text{ و } (B_1 \text{ در } y) \text{ آنگاه } (f_1) \quad (f_1 = p_1x + q_1y + r_1)$$

$$\text{قاعده ۲: اگر } (A_2 \text{ در } x) \text{ و } (B_2 \text{ در } y) \text{ آنگاه } (f_2) \quad (f_2 = p_2x + q_2y + r_2)$$

x و y ورودی‌ها، (A_i) و (B_i) مجموعه فازی، (f_i) خروجی‌ها با ناحیه فازی مشخص شده با قواعد فازی، (p_i) ، (q_i) و (r_i) پارامترهای طراحی شبکه که در طول پروسه آموزش تخمین زده می‌شوند. ساختار کلی این دو قاعده در شکل‌های (۲) و (۳) نشان داده شده است.

در لایه اول، همه نودها، نودهای تطبیقی است. خروجی لایه ۱، درجه عضویت ورودی‌ها می‌باشد که به صورت زیر بیان شده است:

$$O_i^1 = m_{A_i}(x) \quad i = 1, 2 \quad (10)$$

$$O_i^1 = m_{B_{i-2}}(y) \quad i = 3, 4 \quad (11)$$

$m_{A_i}(x)$ و $m_{B_{i-2}}(y)$ می‌توانند با هر نوع تابع عضویت فازی در نظر گرفته شوند. اگر تابع عضویت زنگی شکل را در نظر بگیریم، می‌توان $m_{A_i}(x)$ را به صورت زیر در نظر گرفت:

$$m_{A_i}(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2bi}} \quad (12)$$

a_i ، b_i و c_i پارامترهای توابع عضویت می‌باشند.

در لایه دوم، نودها ثابت می‌باشند. عملکرد این نودها به صورت عملگر ضرب ساده می‌باشد. خروجی این لایه‌ها به صورت زیر به دست می‌آید:

$$O_i^2 = w_i = m_{A_i}(x) m_{B_i}(y) \quad (13)$$

در لایه سوم، نودها باز هم ثابت می‌باشند و برای نرمالیزه کردن خروجی لایه‌های قبلی استفاده می‌شوند:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \quad (14)$$

در لایه چهارم، نودها به صورت انطباقی است. خروجی آن ضرب ساده وزن‌های نرمالیزه شده در چندجمله‌ای مرتبه اول مدل Sugeno می‌باشد:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad i = 1, 2 \quad (15)$$

• خروجی نتایج به دست آمده

گام چهارم: ارزیابی دوباره مقدار تابع هدف که اگر، به معیار شرط توقف یعنی رسیدن به طراحی در طبقه بند ANFIS (تعیین نوع توابع عضویت و آموزش آن) درازای کمترین مقدار MSE رسیده باشد شرایط خاتمه لحاظ می‌شود.

در طراحی بهینه و هوشمند این طبقه بند ANFIS مقادیر، کران بالا و پایین متغیر X (موقعیت هر توپ) به ترتیب 10 و -10 است و در جدول شماره (۲)، مقادیر پارامترهای انتخابی برای الگوریتم بهینه‌سازی IPO که برای محاسبه مقادیر k1 و k2 بکار می‌رود بیان شده است.

جدول ۲: پارامترهای به کار گرفته شده

C1=1	Shift1=100	Scale1=0.01	N Pop=20
C2=2	Shift1=1000	Scale2=2	Max It=50

۴) نتایج تجربی

برای ارزیابی و مقایسه عملکرد طراحی بهینه و هوشمند در طبقه بند ANFIS توسط روش IPO، چهار پایگاه داده در جدول شماره (۳) با ابعاد ویژگی متفاوت بکار گرفته شده است و نتایج به دست آمده به صورت مقایسه‌ای با روش‌های دیگر هوشمند PSO، GA، گزارش شده است.

جدول ۳: توصیف مجموعه داده‌های مورد آزمایش [13]

تعداد کلاس‌ها	تعداد ویژگی	تعداد داده	مجموعه داده
2	9	699	Breast Cancer
3	4	150	Iris
2	6	345	Bupa Liver
2	8	768	Pima indians

۴-۱- ارزیابی نتایج

جهت به دست آوردن دقت روش پیشنهادی در طبقه‌بندی داده‌های گل زنبق به ۳ کلاس (زنبق خاردار، زنبق رنگارنگ و زنبق ورجینیکا)، داده‌های سرطان سینه به ۲ کلاس (خوش خیم و بد خیم)، داده‌های شراب در ۳ کلاس، داده‌های Bupa liver در ۲ کلاس و داده‌های Pima در ۲ کلاس، قسمتی از داده‌ها را طبق جدول شماره (۴) به عنوان داده‌های آموزشی جهت آموزش طبقه بند مبتنی بر سیستم استنتاج عصبی - فازی تطبیقی اعمال شده و مابقی به عنوان داده آزمایش بکار گرفته شده است.

جدول ۴: تعداد داده‌های آموزش و آزمایش

تعداد داده‌های آموزش	تعداد داده‌های آزمایش	داده
135	15	Iris
629	70	Breast Cancer
310	35	Bupa Liver
692	76	Pima indians

تنظیم پارامترهای سیستم فازی پایه با توجه به تابع خطای مدل‌سازی، توسط الگوریتم‌های بهینه‌سازی صورت می‌گیرد که به صورت ریاضی مدل می‌شود:

$$\min_{\theta} \text{Error} = \frac{1}{N} \sum_{i=1}^n e_i^2 \quad (16)$$

$$e_i = t_i - f(x_i|\theta) \quad (17)$$

N: تعداد نمونه‌های ورودی سیستم و جهت نرمالیزه کردن

e_i : مقدار خطاهای سیستم

x_i : مقدار داده‌های ورودی

θ : پارامترهای قابل تنظیم سیستم استنتاج عصبی - فازی تطبیقی به ازای ورودی‌های مختلف x_i

و به طور کلی تابع ارزیابی به صورت معادله زیر بیان می‌شود:

$$\text{Fitness} = \min \text{Error}(\theta) \quad (18)$$

$$\theta = (\theta_1, \theta_2, \dots, \theta_n) \quad (19)$$

از آنجایی که، الگوریتم‌های IPO، GA و PSO توانایی یافتن مینیمم تابع هدف را دارند تابع ارزیابی رابطه (۱۸) خواهد بود.

۳-۴- طراحی بهینه و هوشمند طبقه بند ANFIS با استفاده از روش IPO

شکل (۴) فلوچارت کلی، طراحی بهینه و هوشمند (انتخاب هوشمند و بهینه‌ی توابع عضویت و آموزش) طبقه بند ANFIS، از جدول پروب توابع عضویت، با استفاده از روش IPO را نشان می‌دهد.

گام اول: شروع، با توجه به نوع پایگاه داده انتخابی، ابعاد ویژگی آن پایگاه داده در نظر گرفته می‌شود. به طور مثال، برای داده‌های iris که دارای ۴ ویژگی از x_1, x_2, \dots, x_4 است، ابعاد ویژگی برابر با ۴ در نظر گرفته می‌شود.

گام دوم: شکل‌گیری پروب توابع عضویت بر اساس تعداد ویژگی پایگاه داده است. در شکل (۵) شکل کلی پروب توابع عضویت یا ساختار یک توپ در روش IPO برای پایگاه داده با n ویژگی را نشان می‌دهد و تعداد تکرار و $scale1, scale2, shift1, shift2$ مقداردهی می‌شوند.

گام سوم: ارزیابی اولیه تابع هدف که برابر است با:

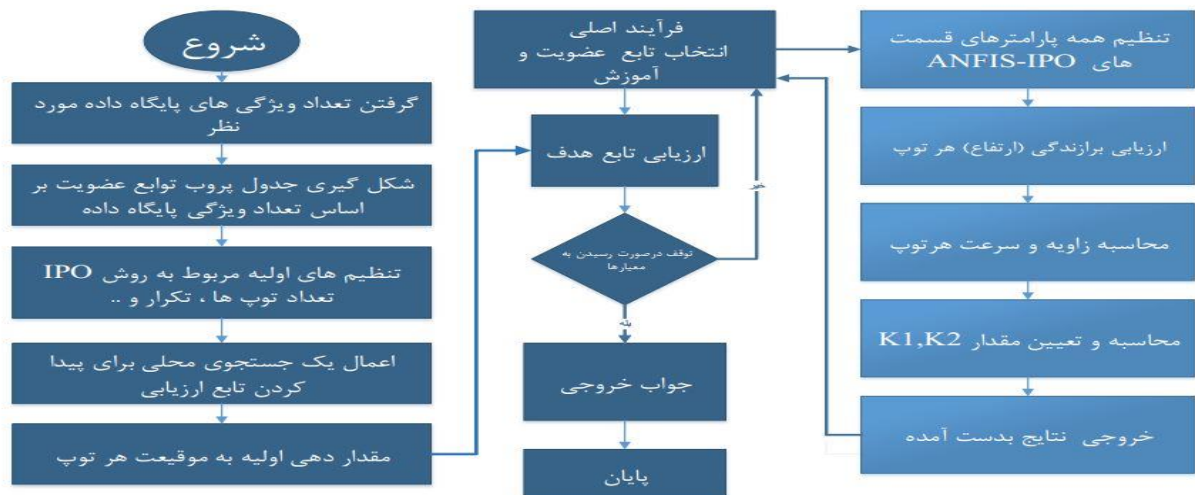
$$\min_{\theta} \text{Error} = \frac{1}{N} \sum_{i=1}^n e_i^2 \quad (20)$$

$$\text{Fitness} = \min \text{Error}(\theta) \quad (21)$$

که در آن θ ، همان مقادیر پروب توابع عضویت است.

حال اگر به معیار شرط توقف یعنی، رسیدن به بهترین طراحی طبقه بند ANFIS (تعیین نوع توابع عضویت و آموزش آن)، درازای کمترین مقدار MSE رسیده باشد شرایط خاتمه لحاظ می‌شود. در غیر این صورت وارد مرحله زیر می‌شود:

- تنظیم همه پارامترهای قسمت ANFIS-IPO
- ارزیابی برازندگی (ارتفاع) هر توپ است
- بر اساس روابط ذکر شده در قسمت توضیحات روش IPO محاسبه زاویه و سرعت هر توپ
- محاسبه و تعیین مقادیر K1, K2



شکل ۴. فلوچارت کلی طراحی بهینه و هوشمند طبقه بند ANFIS توسط روش IPO

$$Ball = \{(a_1, b_1), (c_1, \sigma_1), \dots, (a_1, b_1, c_1, d_1), (a_2, b_2), (c_2, \sigma_2), \dots, (a_2, b_2, c_2, d_2), \dots, (a_n, b_n), (c_n, \sigma_n), \dots, (a_n, b_n, c_n, d_n)\}$$

شکل ۵. یک پروب توابع عضویت، که از یک پایگاه داده انتخابی با n ویژگی تشکیل شده

جدول ۷: مقایسه دقت طبقه‌بندی داده‌های آزمایش (طبقه بند ساده

(ANFIS) برای مجموعه داده Iris

روش ترکیب طبقه بند ANFIS	میزان دقت داده آزمایش داده Iris %
PSO	95.97
GA	95.33
IPO	96.67

جدول ۸: مقایسه دقت طبقه‌بندی داده‌های آزمایش (برای طراحی

بهینه و هوشمند طبقه بند ANFIS) برای مجموعه داده Iris

بهترین نوع تابع عضویت یافت شده توسط روش IPO برای داده Iris	روش ترکیب طبقه بند ANFIS	میزان دقت داده آزمایش %
Gbell MF	PSO	97.33
	GA	98
	IPO	98

جدول ۹: مقایسه دقت طبقه‌بندی داده‌های آزمایش (طبقه بند ساده

(ANFIS) برای مجموعه داده Cancer

روش ترکیب طبقه بند ANFIS	میزان دقت داده آزمایش داده Cancer %
PSO	96.14
GA	96.14
IPO	96.28

جهت محاسبه دقت طبقه بند ANFIS، 10 مرتبه اجرا شده و میانگین جواب‌های به دست آمده به عنوان دقت طبقه بند لحاظ شده و روش پیشنهادی در نرم‌افزار (Matlab R2015b) و سیستمی با مشخصات Cori5 4210U و حافظه رم پیاده‌سازی شده است.

۴-۲- مقایسه با روش‌های موجود

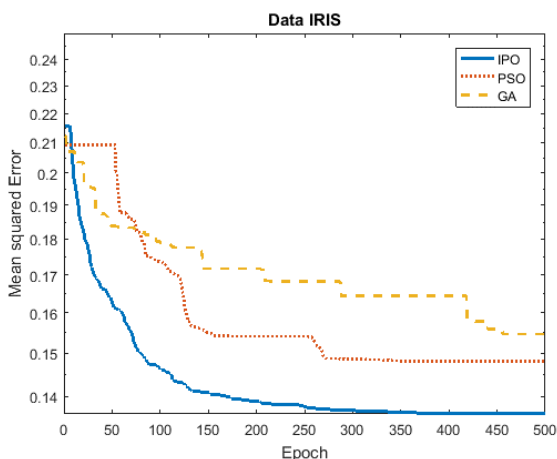
کارایی روش طراحی بهینه و هوشمند طبقه بند ANFIS در مقایسه با سایر روش‌های هوشمند در طبقه بند مبتنی بر سیستم استنتاج عصبی-فازی تطبیقی نشان از بهبود دقت طبقه‌بندی درست در مقایسه با طبقه‌بندی مذکور می‌باشد. در جدول‌های شماره (۵) و (۶) مقادیر انتخاب شده برای الگوریتم‌های مذکور آمده است، که در الگوریتم PSO پارامترهای C1 و C2 ضرایب یادگیری جمعی و شخصی است و w، وزن اینرسی برای به تعادل آوردن جستجوی سراسری و محلی است. در الگوریتم GA، پارامترهای مقدار احتمال جهش (Mutation Probability) و احتمال بازترکیب (Crossover Probability) است

جدول ۵: مقادیر پارامترهای انتخاب شده برای الگوریتم‌ها

PSO	GA
C1=2	Mutation=0.1
C2=2	CrossOver= 0.8
W=0.9	-

جدول ۶: مقادیر پارامترهای انتخاب شده برای الگوریتم‌ها

Var min	-10
Var max	10
NPop	20
MatIt	50



شکل ۶. نتایج همگرایی MSE برای مجموعه داده Iris

همان‌طور که از جدول (۷)، مقایسه دقت طبقه‌بندی داده‌های آزمایشی، برای مجموعه داده Iris مشخص است، در روش ترکیب طبقه بند ANFIS (نوع توابع عضویت به صورت پیش فرض در طبقه بند ساده ANFIS، گوسی می‌باشد) با روش‌های هوشمند کم‌ترین میزان دقت برابر با 95.33 درصد برای روش GA و بیشترین دقت 96.67 درصد برای روش IPO برای 10 بار تکرار گزارش شده است. این نتایج بیانگر این مسئله است که، انتخاب بهترین نوع توابع عضویت برای پایگاه داده مورد نظر در بالا بردن کارایی، دقت و کم کردن خطای طبقه بند ANFIS تاثیر چشمگیری دارد. لذا با توجه به نتایج و مطالب اشاره شده با استفاده از روش پیشنهادی (طراحی بهینه و هوشمند طبقه بند ANFIS با استفاده از روش IPO) در ابتدا بهترین نوع تابع عضویت برای پایگاه داده مورد نظر انتخاب می‌شود و بعد از طریق روش‌های ترکیب طبقه بند ANFIS آموزش داده می‌شود و همان‌طور که از جدول (۸) مشخص است توسط روش IPO بهترین نوع تابع عضویت برای مجموعه داده IRIS، تابع Gbell MF یافت شده است که کمترین دقت برابر با 97.33 و بیشترین دقت 98 برای روش‌های ترکیب طبقه بند ANFIS برای ۱۰ بار تکرار گزارش شده است. در شکل (۶) نتایج همگرایی مقدار MSE برای مجموعه داده IRIS که از طریق ترکیب طبقه بند ANFIS با روش‌های هوشمند است نشان داده شده است و برای مقایسه بهتر هر کدام از این روش‌های هوشمند با یکدیگر مقدار تکرار برابر 500 در نظر گرفته شده است.

مقایسه انجام شده، برای چهار مجموعه داده انتخاب شده، بیانگر عملکرد بهتر و دقت بالاتر، روش IPO برای طراحی بهینه و هوشمند طبقه بند ANFIS، در این مقاله، در هر چهار مجموعه داده نسبت به طبقه بند ساده (غیر هوشمند) ANFIS می‌باشد و این مؤید این مطلب است که روش IPO برای طراحی بهینه و هوشمند طبقه بند ANFIS، البته در این مسئله خاص، نسبت به روش‌های دیگر بهتر عمل می‌کند. با توجه به نتایج به دست آمده می‌توان گفت که روش پیشنهادی در طبقه بند مورد نظر، کارایی قابل قبولی در مسئله بازشناسی الگو دارد.

جدول ۱۰: مقایسه دقت طبقه‌بندی داده‌های آزمایشی (برای طراحی بهینه و هوشمند طبقه بند ANFIS) برای مجموعه داده Cancer

میزان دقت داده آزمایشی %	روش ترکیب طبقه بند ANFIS	بهترین نوع تابع عضویت یافت شده توسط روش IPO برای داده Cancer
97	PSO	Psig MF
97.14	GA	
97.14	IPO	

جدول ۱۰: مقایسه دقت طبقه‌بندی داده‌های آزمایشی (طبقه بند ساده ANFIS) برای مجموعه داده Pima

میزان دقت داده آزمایشی % داده Pima	روش ترکیب طبقه بند ANFIS
76.56	PSO
77.05	GA
77.68	IPO

جدول ۱۱: مقایسه دقت طبقه‌بندی داده‌های آزمایشی (برای طراحی بهینه و هوشمند طبقه بند ANFIS) برای مجموعه داده Pima

میزان دقت داده آزمایشی %	روش ترکیب طبقه بند ANFIS	بهترین نوع تابع عضویت یافت شده توسط روش IPO برای داده Pima
78.64	PSO	Dsig MF
78.35	GA	
78.32	IPO	

جدول ۱۲: مقایسه دقت طبقه‌بندی داده‌های آزمایشی (طبقه بند ساده ANFIS) برای مجموعه داده Bupa Liver

میزان دقت داده آزمایشی % داده Bupa Liver	روش ترکیب طبقه بند ANFIS
71.35	PSO
68.24	GA
71.43	IPO

جدول ۱۳: مقایسه دقت طبقه‌بندی داده‌های آزمایشی (برای طراحی بهینه و هوشمند طبقه بند ANFIS) برای مجموعه داده Bupa Liver

میزان دقت داده آزمایشی %	روش ترکیب طبقه بند ANFIS	بهترین نوع تابع عضویت یافت شده توسط روش IPO برای داده Bupa
73.63	PSO	Dsig MF
73.55	GA	
73.84	IPO	

(۵) نتیجه‌گیری

در این پژوهش یک روش هوشمند برای طراحی بهینه‌ی یک طبقه بند ANFIS به وسیله الگوریتم بهینه‌سازی صفحات شیب‌دار (IPO) پیشنهاد شده است. نتایج تجربی نشان می‌دهد، که طراحی بهینه و هوشمند طبقه بند ANFIS توسط روش IPO بر چهار مجموعه داده مشهور نتایج خوب و رضایت بخشی را حاصل کرد و در هر چهار مسئله جواب‌های بهتری نسبت به روش‌های طراحی دیگر در طبقه بند مبتنی بر سیستم استنتاج عصبی-فازی ارائه کرده است. یک محدودیت مهم که ممکن است استفاده از روش طراحی بهینه و هوشمند طبقه بند ANFIS توسط الگوریتم بهینه‌سازی صفحات شیب‌دار (IPO) را محدود کند، وابستگی به پارامترهای اولیه الگوریتم جهت رسیدن به همگرایی سریع الگوریتم است. استفاده از IPO با کنترل‌کننده‌های کمکی به منظور کنترل پارامترهای آن می‌تواند سبب همگرایی سریع‌تر طبقه بند گردد. این توانایی در تحقیقات آینده می‌تواند مورد بررسی قرار گیرد.

مراجع

- [1] C. C. Lee, "Fuzzy logic in control systems: fuzzy logic controller. II," *IEEE Trans. Syst. Man. Cybern.*, vol. 20, no. 2, pp. 404-418, 1990.
- [2] T. J. Ross, *Fuzzy Logic with Engineering Applications: Third Edition*, vol. 222. New Delhi: Tata McGraw-Hill Publishing Company limited, 2010.
- [3] J. S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," *IEEE Trans. Syst. Man Cybern.*, vol. 23, no. 3, pp. 665-685, 1993.
- [4] S.-H. Zahiri, "Swarm Intelligence and Fuzzy Systems (Computer Science, Technology and Applications): Seyed-Hamid Zahiri: March 1, 2011," 2010
- [5] "Identification using ANFIS with intelligent hybrid stable learning algorithm approaches and stability analysis of training methods," *Appl. Soft Comput.*, vol. 9, no. 2, pp. 833-850, Mar. 2009.
- [6] D. Karaboga and E. Kaya, "An adaptive and hybrid artificial bee colony algorithm (aABC) for ANFIS training," *Appl. Soft Comput.*, vol. 49, pp. 423-436, 2016.
- [7] K. Thangavel and A. Kaja Mohideen, "Mammogram Classification Using ANFIS with Ant Colony Optimization Based Learning," Springer Singapore, 2016, pp. 141-152.
- [8] M. H. Mozaffari, H. Abdy, and S. H. Zahiri, "IPO: An inclined planes system optimization algorithm," *Comput. Informatics*, vol. 35, no. 1, pp. 222-240, 2016.
- [9] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Neural Networks, 1995. Proceedings., IEEE Int. Conf.*, vol. 4, pp. 1942-1948 vol.4, 1995.
- [10] T. Bäck, *Evolutionary algorithms in theory and practice : evolution strategies, evolutionary programming, genetic algorithms*. Oxford University Press, 1996.
- [11] R. Chelouah and P. Siarry, "A Continuous Genetic Algorithm Designed for the Global Optimization of Multimodal Functions," *J. Heuristics*, vol. 6, no. 2, pp. 191-213, 2000.
- [12] K. sugun. S.Eswari, P.N.Raghunath, "Ductility performance of HyFRC," *Am. J. Appl. Sci.*, vol. 5, no. 9, pp. 1257-1262, 2008.
- [13] K. Bache and M. Lichman, "UCI Machine Learning Repository," *University of California Irvine School of Information*, vol. 2008, no. 14/8. 2013.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

بخش بندی خودکار تصاویر سی تی اسکن ریه به کمک یادگیری تقویتی

پرنیا قیصری^۱، منصور فاتح^۲، محسن رضوانی^۳

^۱ دانشجوی کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود،

P.gheysari@shahroodut.ac.ir

^۲ استادیار دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود،

mansoor_fateh@shahroodut.ac.ir

^۳ استادیار دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود، شاهرود،

mrezvani@shahroodut.ac.ir

چکیده

یکی از مراحل اصلی در تحلیل خودکار تصاویر پزشکی، بخش بندی تصاویر است. در صورت بخش بندی مناسب این تصاویر، امکان تحلیل دقیق تر و در نتیجه تشخیص صحیح بیماری میسر می شود. آمار نشان می دهد که ۱۳٪ تلفات انسانی ناشی از سرطان در سرتاسر جهان، به دلیل بیماری های ریوی هستند. بدیهی است که با تشخیص به موقع این بیماری ها می توان گام مناسبی در کمک به بیماران انجام داد. در این مقاله با استفاده از روش یادگیری Q-learning، نواحی مختلف تصاویر سی تی ریه تعیین و تمام نواحی تصویر توسط مرزهای اصلی تعیین شده است. در نهایت رنگ هر ناحیه مشخص و تصویر بخش بندی شده به عنوان خروجی نشان داده شده است. در این تحقیق، آزمایش هایی جهت بررسی نتایج روش پیشنهادی بر روی داده های واقعی انجام شده است. نتایج این آزمایش ها نشان دهنده دقت کیفی بالاتر روش پیشنهادی در مقایسه با دیگر روش های بخش بندی است.

کلمات کلیدی

بخش بندی، یادگیری تقویتی، ریه، Q-learning

بیماری در مسائل پزشکی را تحت تاثیر قرار می دهد [5]. بهره گیری از تکنیک های بخش بندی تصویر در کنار تجربه و دانش انسان می تواند تاثیر شایسته ای در شناسایی بیماری ها داشته باشد [5]. بیماری های ریوی از سخت ترین و خطرناک ترین بیماری ها هستند. ۱۳٪ از مرگ و میر افراد سرطانی، مختص بیماری های ریوی است. از این رو، استفاده از روش های تشخیص در مراحل اولیه پیشرفت آن، می تواند در روند درمان بیمار بسیار موثر و حائز اهمیت باشد [6,7].

در تشخیص بیماری های ریه، گرفتن تصویر ریه بسیار ضروری است. از تکنیک های تصویربرداری مدرن مانند تصویربرداری رزونانس مغناطیسی (NMR)، سونوگرافی، اسکن CT، X-Ray، MRI، SPECT و غیره برای ثبت این تصاویر استفاده می شود. تصویربرداری CT به دلیل ثبت تصاویر از مقاطع مختلف ریه با عمق های دلخواه، از قدرتمندترین

۱- مقدمه

بخش بندی تصویر به فرآیند تفکیک اجزای تصویر اطلاق می شود. به طور دقیق تر، بخش بندی تصویر فرآیند نسبت دادن یک برچسب به هر یک از پیکسل های تصویر است [1]. هدف اصلی بخش بندی، تقسیم تصویر به اجزاء اصلی و استخراج قسمت های مورد علاقه است. این فرآیند، تفکیک و جداسازی تصویر به نواحی مختلف، بر اساس مجموعه ای از ویژگی های خاص است. از اساسی ترین ویژگی های بخش بندی تصویر لبه، بافت و رنگ تصویر هستند [2,3]. بخش بندی مناسب، تاثیر قابل ملاحظه ای بر استخراج صحیح ویژگی ها و طبقه بندی صحیح دادگان دارد [3,4]. به طور مثال، بخش بندی مناسب تصویر، تشخیص ردیابی خودکار در کاربردهای نظامی و تشخیص

می‌شود. وزن یال بین دو گره بر اساس یک معیار شباهت همچون رنگ یا بافت، بین دو پیکسل متناظر با این دو گره در تصویر تعریف می‌شود. عملیات بخش‌بندی توسط هرس کردن یال‌های وزن‌دار و افراز گراف اولیه به N زیرگراف جدا از هم انجام می‌شود [12].

در بخش‌بندی تصویر می‌توان از یادگیری تقویتی جهت ترکیب روش‌های آستانه‌گذاری مختلف برای بخش‌بندی خودکار و یافتن آستانه بهینه استفاده کرد. عامل از آستانه‌های مختلف جهت یافتن بهترین آستانه استفاده می‌کند. مقادیر آستانه‌های حاصل از روش‌های مختلف، به عامل تحویل داده می‌شوند. سپس آستانه‌ها توسط عامل تقویتی وزن-دهی می‌شوند. در ابتدا وزن‌ها می‌توانند به صورت دلخواه یا تصادفی انتخاب شوند. البته مجموع وزن‌ها باید یک شود. سپس این وزن‌ها توسط عامل یادگیری تقویتی تغییر کرده تا به مقادیر بهینه همگرا شوند [2].

به طور کلی، یادگیری تقویتی در بخش‌بندی تصاویر پزشکی [13,14]، شناسایی و دنبال کردن شی در تصاویر متوالی [15,16]، برطرف کردن نویز تصویر [17]، شناسایی اشیاء و اعمال تغییرات مورد نظر در آن‌ها [18] به کارگرفته می‌شود.

در ادامه به تحقیقات انجام شده جهت تشخیص بیماری‌های ریوی به کمک روش‌های هوش مصنوعی و بخش‌بندی تصویر، اشاره می‌شود. همانطور که قبلاً مطرح شد، تصاویر CT اطلاعات مناسبی از وضعیت بیماری‌های ریوی را در اختیار پزشک قرار می‌دهند. ثبت این تصاویر با اجرای مداوم چرخه تنفسی همراه است که باعث ایجاد نویز در تصاویر در حال ثبت می‌شود. بنابراین بهینه‌سازی و نویزدایی از این تصاویر، اولین مرحله قابل اجرا است. در مرجع [5] با ترکیب دو رویکرد تطبیق تصویر و بخش‌بندی تصویر، یک فرمول جدید ارائه شده است. نتایج این مقاله در مقایسه با روش‌های انطباق تصویر و روش‌های بخش‌بندی، بهتر شده است. در این روش، در گام اول عمل تطبیق تصویر به منظور استخراج نتایج قابل توجه از نمایش بهتر تصاویر ثبت شده، انجام می‌شود. در گام بعدی از الگوریتم‌های بخش‌بندی برای تقسیم‌بندی قسمت‌های مختلف استفاده می‌شود.

در مرجع [6] از الگوریتم آب‌پخشان^۶ و آستانه‌گذاری جهت بخش‌بندی و شناسایی سلول‌های سرطانی، استفاده شده است. بخش‌بندی و شناسایی سلول سرطانی با استفاده از ترکیب این دو روش، سرعت تجزیه و تحلیل تصویر را بالا برده است. پس از آستانه‌گذاری تصویر، الگوریتم آب‌پخشان با سرعت و دقت بیشتری تومور سرطانی را از پس‌زمینه شناسایی کرده است.

در مرجع [19] از روش انتقال میانگین^۸ به همراه یک روش ماشین‌بردار پشتیبان (SVM)، جهت بخش‌بندی تصاویر در زمینه ژئومورفولوژی استفاده شده است. نتایج این مقاله نشان می‌دهد که ترکیب انتقال میانگین و طبقه‌بندی مبتنی بر SVM، یک چارچوب موثر برای طبقه‌بندی یک شکل خاص را ارائه می‌دهد.

۳- یادگیری تقویتی

یادگیری تقویتی، یادگیری قوی و بدون ناظر بر مبنای تعامل با محیط است. این یادگیری با استفاده از سیگنال تقویتی، قادر به آموزش عامل^۹

ابزار برای ارزیابی بیماری‌های ریوی است [6-8]. از CT برای ارزیابی میزان پیشرفت بیماری‌های ریوی مانند اندازه‌گیری تغییرات حجم تومور در طی زمان، تشخیص توابع اولیه ریوی مانند ارزیابی تپه‌ریه (چرخه تنفسی) و به طور کلی نمایش دقیق و سه بعدی از قسمت‌های خاص، بافت‌های متفاوت و آناتومی ریه استفاده می‌کنند [5,7,8]. امروزه با توجه به حجم زیاد داده‌های فراهم شده توسط دستگاه‌های تصویر برداری CT و تحلیل دستی این تصاویر، تشخیص بیماری تبدیل به یک کار خسته‌کننده و چالش برانگیز شده است [5]. به همین دلیل استفاده از سیستم‌های تشخیص خودکار کامپیوتری، برای تشخیص بیماری‌هایی مانند سرطان ریه، انسداد ریه، بیماری‌های بافتی و غیره همواره مورد توجه بوده است [6].

یکی از مراحل مهم و ابتدایی در اکثر این سیستم‌ها، بخش‌بندی تصاویر ریه است [6,9]. این مرحله یک پیش نیاز برای انجام عملیات تحلیل محتوای تصویر و تشخیص الگوهای مورد جستجو در تصاویر است [9]. هدف از بخش‌بندی تصویر، شناسایی خودکار مناطق مورد نظر است [5,6]. روش‌های مختلف بخش‌بندی شامل آستانه‌گذاری^۱، خوشه‌بندی^۲، بخش‌بندی مبتنی بر ناحیه^۳ و بخش‌بندی مبتنی بر لبه^۴ هستند. در این تحقیق از تصاویر CT ریه برای اعمال بخش‌بندی خودکار استفاده شده است.

در ادامه، در بخش دوم تحقیقات مرتبط با بخش‌بندی ارائه شده است. در بخش سوم مقاله، مبانی یادگیری تقویتی^۵ شرح داده شده است. در بخش چهارم روش پیشنهادی توصیف شده است. در بخش پنجم آزمایش‌ها، معرفی پایگاه داده، نتیجه و ارزیابی آورده شده است. در بخش نهایی، نتیجه‌گیری مباحث و پیشنهادات شرح داده شده است.

۲- کارهای مرتبط

همانطور که ذکر شد، بهترین راه استخراج اطلاعات از تصویر، بخش‌بندی آن است. در این راستا، الگوریتم‌های بسیاری پیشنهاد شده که با روش‌ها و اهداف متفاوت، بخش‌بندی را پیاده‌سازی کرده‌اند. در ادامه برخی از این الگوریتم‌ها معرفی می‌شوند.

بخش‌بندی تصاویر اسکن شده‌ی نقشه‌های فرش، از جمله الگوریتم‌های کاربردی بخش‌بندی تصاویر است. این روش، به منظور کاهش رنگ نقشه‌های فرش به صورت خودکار ارائه شده است [10].

در مرجع [11] بخش‌بندی تصویر مبتنی بر لبه‌یابی انجام شده است. با پیدا کردن خطوط بسته به وسیله لبه‌ها می‌توان قطعه‌ای از تصویر را شناسایی کرد. پیش از پیدا کردن لبه اشیاء و ادغام آن‌ها با هم برای یافتن یک شی کامل، باید تصویر رفع نویز شود. امکان از بین رفتن مرزهای ضعیف بعد از نویز زدایی بیشتر خواهد بود. به همین دلیل برای نویز زدایی باید روش مناسبی در نظر گرفته شود تا در نتیجه‌ی آن لبه‌های کامل‌تری را بتوان استخراج کرد.

از جمله روش‌های قوی بخش‌بندی، روش‌های مبتنی بر گراف^۶ هستند. روش‌های مبتنی بر گراف بر روی تصاویر با نواحی همگن نتایج مناسبی دارند. علاوه بر آن، اطلاعات مکانی و فضای ویژگی پیکسل‌ها به طور همزمان برای قطعه‌بندی تصویر با یکدیگر ترکیب می‌شوند. در این روش‌ها، هر پیکسل تصویر به عنوان یک گره از گراف در نظر گرفته

مقادیر جدول گسترش پیدا کرده و در نهایت به مقادیر بهینه همگرا خواهند شد [28].

الگوریتم Q-learning را در ساده‌ترین شکل به صورت زیر می‌توان بیان کرد [13,17,25,30]:

۱. مشخص کردن متغیر γ و ماتریس پاداش r .
۲. مقداردهی اولیه ماتریس Q با صفر.
۳. برای هر تکرار داریم:
 - a) حالت اولیه به صورت تصادفی انتخاب می‌شود.
 - b) تا زمانی که به حالت هدف نرسیده‌ایم:
 - I. از میان تمام عملیات ممکن برای حالت فعلی یکی انتخاب می‌شود.
 - II. با یک عمل به حالت بعدی می‌رود.
 - III. بیشترین مقدار Q برای حالت بعدی بر اساس تمام اعمال بدست می‌آید (محاسبه رابطه ۱).
 - IV. حالت بعدی به عنوان حالت جاری در نظر گرفته می‌شود.

۴- روش پیشنهادی

برای بخش‌بندی تصویر از روش‌های مختلف یادگیری می‌توان استفاده کرد. در این مقاله تصاویر ریه به عنوان ورودی به الگوریتم Q-learning داده می‌شود. در این الگوریتم، از سیاست $\epsilon - greedy$ استفاده می‌شود. همچنین مقادیر γ و α در رابطه (۱)، 0.8 انتخاب شده و عمل بخش‌بندی، روی تصاویر CT انجام می‌شود. در ادامه تمام نواحی تصویر توسط مرزهای اصلی تعیین می‌شود. سپس رنگ هر ناحیه با توجه به نوع بافت (برای نمونه ماهیچه، تومور، رگ و لخته خون) مشخص می‌گردد.

۴-۱- Q-learning

در روش Q-learning هر تکرار معادل یک دوره آموزش است. در هر دوره آموزش، عامل محیط را کاوش می‌کند. دریافت پاداش و ضرر تا رسیدن به حالت هدف ادامه دارد. هدف از آموزش، ساخت مغز عامل برای بخش‌بندی تصویر است [2,8,5].

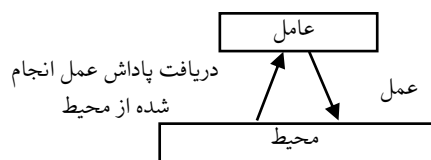
حال با توجه به فرمول Q-learning و مقادیر آن، بخش‌بندی تصاویر CT، با یادگیری تقویتی مدل می‌شود.

حالت: در این مقاله هر پیکسل از تصویر، به عنوان حالت در نظر گرفته می‌شود. در این صورت به تعداد حاصل ضرب تمام پیکسل‌های افقی و عمودی، حالت وجود دارد.

عمل: ۹ عمل برای هر حالت وجود دارد. این حالت‌ها به صورت زیر تعریف می‌شوند.

- ۱) عدم انجام عمل (N) (۳ جنوب (S) (۴ شرق (E) (۵ غرب (W) (۶ شمال غرب (NW) (۷ شمال شرق (NE) (۸ جنوب غرب (SW) (۹ جنوب شرق (SE).
- همچنین احتمال رفتن به هر خانه متناسب با عمل انجام شده ۱۰۰٪ در نظر گرفته می‌شود. حالت‌های بعدی با توجه به این احتمال،

در محیط‌های پیچیده، غیرقطعی و تصادفی است [20,21]. یادگیری تقویتی به مسئله‌هایی با یک عامل مستقل می‌پردازد. عامل حالت‌هایی را درک کرده و مطابق با آن ادراک، اعمال بهینه‌ای را برای رسیدن به اهدافش انجام می‌دهد [10-22]. در یادگیری، عامل عملی را در محیط انجام می‌دهد و متناسب با حالت و عمل^{۱۱} انجام شده به آن پاداش^{۱۱} یا تنبیه^{۱۲} تعلق می‌گیرد. یادگیری از طریق این پاداش‌ها انجام می‌شود. هدف اصلی یادگیری تقویتی حداکثر کردن پاداش‌های دریافتی از محیط در طول زمان است [1,2]. مقادیر پاداش می‌توانند کمی یا کیفی باشند. عامل با عمل در محیط تاثیر می‌گذارد و تنها با تعامل با محیط از حالت‌ها آگاهی می‌یابد. حالات در واقع پارامترهای شرح‌دهنده محیط هستند [8-25]. به شیوه رفتار عامل در طول زمان سیاست^{۱۳} گویند [17,23]. سیاست بهترین عمل در هر حالت را مشخص می‌کند [14,17]. در شکل تعامل بین عامل و محیط در یادگیری تقویتی نشان داده شده است [17].



شکل ۱: تعامل بین عامل و محیط

در این مقاله، بر روی یکی از الگوریتم‌های یادگیری تقویتی، به نام Q-learning تمرکز شده است. این الگوریتم حتی در حالت عدم اطلاع عامل از محیط، درست کار می‌کند. در این روش، عامل، رسیدن به هدف و انتخاب بهترین عمل در هر حالت را می‌آموزد [18]. در این الگوریتم، یادگیری از طریق تجربیات یا آموزش انجام می‌شود. هر تکرار معادل با یک دوره آموزش است [23,24]. یک اپیزود (یک مرحله از یادگیری تقویتی با رسیدن به هدف) الگوریتم با رسیدن عامل به وضعیت نهایی پایان می‌یابد. هدف از آموزش، ساخت مغز عامل است و با ماتریس Q نمایش داده می‌شود [24,25]. آموزش بیشتر، منجر به ماتریس Q بهتر می‌شود [26,22]. با داشتن ماتریس Q ، عامل در عوض کاوش و جستجوی متعدد، با رجوع به ماتریس حالت‌ها و انتخاب گزینه ماکزیمم، بهترین حالت را انتخاب می‌نماید [25,22,27].

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (1)$$

رابطه ۱، معادله Q-learning را در یادگیری تقویتی نشان می‌دهد [25,22,27]. تمام مقادیر ماتریس Q ابتدا صفر هستند. عامل بطور متناوب وضعیت فعلی S را تشخیص داده و عمل a را انجام می‌دهد. سپس پاداش حاصل و حالت جدید ناشی از انجام عمل را مشاهده می‌کند [28]. $Q(s, a)$ ، در واقع بهترین عمل در موقعیت جدید به هر حالت و عمل نسبت داده می‌شود [22,27,29]. این مقدار، مجموع پاداش‌های دریافت شده با شروع از حالت S و انجام عمل a توسط عامل است [29]. سیستم در مرحله یادگیری، جدول Q ها را پر می‌کند. در مراحل بعدی از این جدول استفاده خواهد شد. یعنی در گذر از هر حالت به حالت دیگر، جریانی با بیشترین مقدار Q انتخاب می‌شود [27,29]. با رسیدن عامل به هدف، اپیزود تمام می‌شود و یک اپیزود جدید شروع می‌گردد. با افزایش یادگیری، مقادیر غیر صفر به سایر

۴-۲- تعیین نواحی تصویر

در تصاویر اسکن شده از ریه، هر ناحیه دارای بافت جدا و مشخص است که باعث تفاوت رنگی در هر ناحیه شده است. در این مرحله تصویر را از فضای رنگی به فضای خاکستری برده و تمام مرزها به کمک فیلتر Canny مشخص می‌شود. به علت گسسته بودن مرزهای تصویر، تفکیک نواحی با مشکل مواجه خواهد بود. برای کاهش گسستگی مرزها، از تار کردن مرزها استفاده می‌شود. در این بخش شدت روشنایی هر پیکسل با میانگین شدت روشنایی آن و هشت عدد از پیکسل‌های اطرافش جایگزین شود. در ادامه، شدت روشنایی پیکسل‌های بزرگتر از صفر با ۲۵۵ جایگزین خواهد شد. برای جداسازی مرزها از بخش‌های دیگر تصویر، شدت روشنایی نقاط مرزی تار شده ۲۵۵ لحاظ می‌شود. حال که مرز هر بخش مشخص شده، به پیکسل‌های هر ناحیه برچسب یکسان اعمال می‌شود.

۴-۳- تعیین رنگ هر ناحیه

پس از مشخص شدن مرز هر بخش تصویر و برچسب زدن بر پیکسل‌ها، باید رنگ تمام پیکسل‌هایی یک ناحیه تعیین شود. ابتدا میانگین رنگ هر ناحیه، توسط رابطه (۵) محاسبه و به تمام پیکسل‌های آن نسبت داده می‌شود. اگر تفاوت میانگین محاسبه شده با رنگ هر پیکسل بیش از ۱۰ باشد، آن پیکسل کنار گذاشته خواهد شد و رنگ ناحیه با پیکسل‌های باقی مانده محاسبه می‌شود.

$$dist = \sqrt{\frac{(R(x,y) - Rmean)^2 + \dots + (G(x,y) - Gmean)^2 + \dots + (B(x,y) - Bmean)^2}{3}} \quad (5)$$

۵- آزمایش‌ها

در این بخش ابتدا پایگاه داده معرفی شده و سپس نتایج بخش بندی روش پیشنهادی، با دو روش کاربردی بخش بندی تصاویر، مقایسه و ارزیابی خواهد شد.

۵-۱- پایگاه داده

پایگاه داده تهیه شده برای این مقاله، تصاویر CT از ریه ۱۰ هزار انسان است. برخی از این نمونه‌ها از بیماران و برخی دیگر از افراد سالم تهیه شده است. این تصاویر با صرف زمان زیاد در بیمارستان امام حسین شاهرود تهیه شده است. این تصاویر با دادن تعهد جهت فاش نشدن اطلاعات بیماران، در اختیار نویسندگان مقاله قرار گرفته است.

۵-۲- ارزیابی

در این بخش روش پیشنهادی با دو روش آب‌پخشان و انتقال میانگین مقایسه شده است. جهت مقایسه‌ی نتایج بخش بندی، چندین تصویر از بخش‌های مختلف ریه افراد متفاوت، استفاده شده است. نتایج نهایی چند تصویر به عنوان نمونه در شکل ۱، شکل ۲، شکل ۳ و شکل ۴ آورده شده است. این تصاویر به دو متخصص^{۱۴} رادیولوژی و سونوگرافی نشان داده شده است. هر دوی این متخصصان تصویر حاصل از الگوریتم پیشنهادی را مناسب‌تر از دو روش دیگر دانسته‌اند. هر دو متخصص

در رابطه‌ی (۲) آورده شده است. حالت‌های بعدی دیگر نیز به همین صورت قابل تعریف هستند.

$$\begin{aligned} P_r(S_{t+1} = S_{x+1,y} | S_t = S_{x,y}, E) &= 1 \quad \forall S_t \in S - S_E \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, E) &= 1 \quad \forall S_t \in S_E \\ P_r(S_{t+1} = S_{x-1,y} | S_t = S_{x,y}, W) &= 1 \quad \forall S_t \in S - S_W \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, W) &= 1 \quad \forall S_t \in S_W \\ P_r(S_{t+1} = S_{x,y+1} | S_t = S_{x,y}, N) &= 1 \quad \forall S_t \in S - S_N \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, N) &= 1 \quad \forall S_t \in S_N \\ P_r(S_{t+1} = S_{x,y-1} | S_t = S_{x,y}, S) &= 1 \quad \forall S_t \in S - S_S \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, S) &= 1 \quad \forall S_t \in S_S \\ P_r(S_{t+1} = S_{x+1,y+1} | S_t = S_{x,y}, NE) &= 1 \quad \forall S_t \in S - S_E - S_N \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, NE) &= 1 \quad \forall S_t \in S_E + S_N \\ P_r(S_{t+1} = S_{x-1,y-1} | S_t = S_{x,y}, SE) &= 1 \quad \forall S_t \in S - S_E - S_S \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, SE) &= 1 \quad \forall S_t \in S_E + S_S \\ P_r(S_{t+1} = S_{x-1,y+1} | S_t = S_{x,y}, NW) &= 1 \quad \forall S_t \in S - S_W - S_N \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, NW) &= 1 \quad \forall S_t \in S_W + S_N \\ P_r(S_{t+1} = S_{x-1,y-1} | S_t = S_{x,y}, SW) &= 1 \quad \forall S_t \in S - S_W - S_S \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, SW) &= 1 \quad \forall S_t \in S_W + S_S \\ P_r(S_{t+1} = S_{x,y} | S_t = S_{x,y}, NA) &= 1 \quad \forall S_t \in S \end{aligned} \quad (2)$$

اگر عمل انتخابی باعث ایجاد حالتی خارج از فضای حالت شود، رابطه‌ی (۳) لحاظ می‌شود.

$$S_{t+1} = S_t, R_{SS'}^a = -3 \quad (3)$$

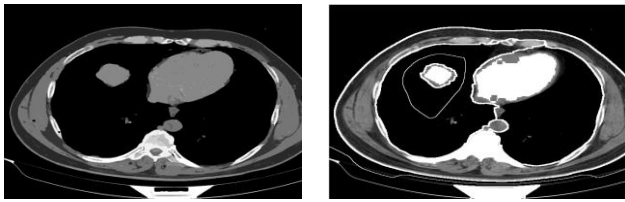
پاداش و ضرر: با توجه به رنگی بودن تصاویر اسکن شده از بیماران، پاداش و ضرر بر اساس فاصله اقلیدسی محاسبه خواهد شد. در غیر این صورت، پاداش یا ضرر بر مبنای معیار زیر تعریف می‌شود. این معیار بر اساس مرجع [10] تدوین شده است و متناسب با کاربرد بخش بندی در تصاویر سی تی ریه بهینه شده است. در این معیار به ازای هر عمل در هر مؤلفه رنگ، اختلاف میزان رنگ بین حالت فعلی و حالت بعدی مقایسه می‌شود و بر مبنای آن میزان ضرر یا پاداش معین می‌شود.

رنگ حالت فعلی - رنگ حالت بعدی = Dif.color

در رابطه بالا Dif.color اختلاف میزان رنگ بین حالت فعلی و حالت بعدی در هر مؤلفه رنگ به طور مجزا به ازای هر عمل است.

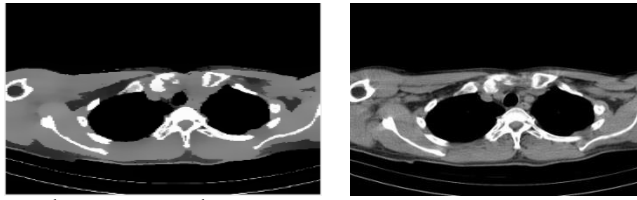
هر پیکسل در اکثر تصاویر سی تی رنگی معادل با یکی از پیکسل‌های همسایه خود دارد. از این رو، معیار پاداش و ضرر در راستای هم‌رنگ شدن هر پیکسل با یکی از پیکسل‌های اطراف خود تعریف می‌شود. در این راستا، به ازای عملی با تفاوت رنگی بیش از آستانه بین حالت فعلی و حالت بعدی، ضرر زیادی به عمل تعلق می‌گیرد. به ازای تفاوت رنگی کمتر از آستانه، پاداش به عمل تعلق می‌گیرد. ۲۰، مقدار آستانه اولیه تفاوت برای هر مؤلفه رنگ است. لازم به ذکر است که آستانه مشخص شده برای تصاویر سی تی طبق آزمایش‌های متعدد محاسبه شده است. این معیار به صورت رابطه (۴) تعریف می‌شود.

$$\begin{aligned} \text{if } Dif.color > threshold \text{ or } Dif.color < -threshold \rightarrow \\ R_{SS}^a &= E(r_{t+1} | S_t = S, S_{t+1} = S', a_t = a) = -500 \quad \forall S_t \in S \\ \text{if } -threshold \leq Dif.color \leq threshold \rightarrow \\ R_{SS}^a &= E(r_{t+1} | S_t = S, S_{t+1} = S', a_t = a) \\ &= ||Dif.color| - threshold| - threshold \quad \forall S_t \in S \end{aligned} \quad (4)$$

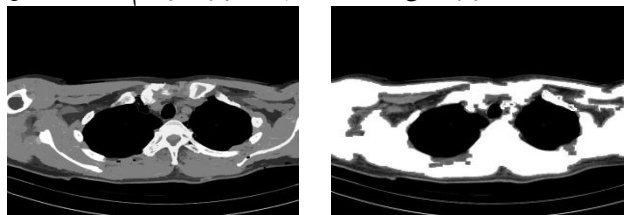


(د) تصویر روش پیشنهادی (ج) تصویر الگوریتم آب‌پخشان

شکل ۴: نتیجه پیاده‌سازی ۳ روش روی تصویر ریه فرد سوم



(ب) تصویر الگوریتم انتقال میانگین (الف) تصویر اصلی



(د) تصویر روش پیشنهادی (ج) تصویر الگوریتم آب‌پخشان

شکل ۵: نتیجه پیاده‌سازی ۳ روش روی تصویر ریه فرد چهارم

در این راستا الگوریتم‌های مختلف بخش‌بندی معرفی شده و روش پیشنهادی با دو روش مهم بخش‌بندی مورد مقایسه قرار گرفته است. نتایج بخش‌بندی به دو متخصص رادیولوژی و سونوگرافی نشان داده شد. طبق نظر هر دو متخصص روش پیشنهادی مقاله از دقت، وضوح و تفکیک بیشتر اجزا، برخوردار است. در این مقاله تحقیقات مناسبی در زمینه بخش‌بندی تصاویر پزشکی ریه و نیاز پزشکان در تشخیص بهتر صورت گرفت. این تحقیقات نشان داد که نیاز بخش‌بندی با قابلیت تفکیک عضلات با دقت بالا، وجود دارد. لازم به ذکر است که روش پیشنهادی بخوبی بافت عضلانی و استخوانی را از چربی، تومور و اجزای دیگر قابل تفکیک می‌کند.

سپاسگزاری

نویسندگان این مقاله از مسئولین محترم بیمارستان امام حسین شاهرود جهت در اختیار گذاشتن عکس‌های بیماران برای انجام این تحقیق تشکر و قدردانی می‌کنند. همچنین از دکتر حبیب ذاکریان و دکتر لاله ابراهیم‌پور بخاطر در اختیار گذاشتن وقت و دادن نظرات سازنده و مفیدشان قدردانی ویژه می‌شود.

مراجع

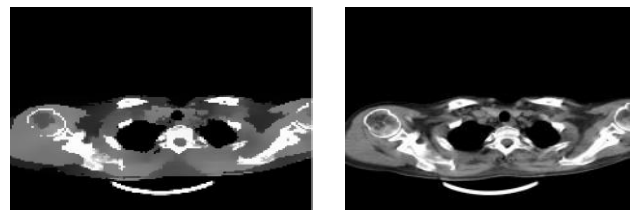
- [1] Liu, Ji, Matthew R. Whiteway, Daniel A. Butts, and Patrick O. Kanold. "Spatial organization of the mouse auditory cortex to sound dynamics revealed using automated image segmentation." *bioRxiv* (2017): 139659.

[۲] سید ابراهیمی، سیدمحمد، پورقاسم، حسین، کشاورز، احمد، "بخش‌بندی ملانوما و دیگر عارضه‌های رنگی پوست در تصاویر درموسکپی با استفاده از

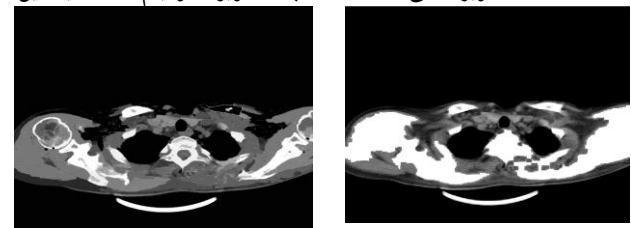
دلیل بهتر بودن این الگوریتم را تفکیک بافتی بیشتر، مرزبندی بهتر اجزا و در نتیجه وضوح بیشتر تصویر دانستند. از دید تخصصی آن‌ها تراکم بافتی دو عضله مختلف با دانسیته‌های^{۱۰} متفاوت در روش پیشنهادی از شفافیت بیشتری برخوردار است. برطرف شدن تاری و نویز از تصویر CT، کمک شایانی به شناسایی اجزا و تفکیک عضلات از بافت چربی و استخوان کرده است.

۶- نتیجه‌گیری و پیشنهاد برای کارهای آتی

در این مقاله ابتدا مرور مناسبی بر روش‌های مختلف بخش‌بندی انجام شد. در ادامه به معرفی روش‌های یادگیری تقویتی در این حوزه پرداخته شد. سپس عمل بخش‌بندی با استفاده از یادگیری تقویتی، روی تصاویر CT گرفته شده از ۱۰ هزار نفر، اعمال شده است.

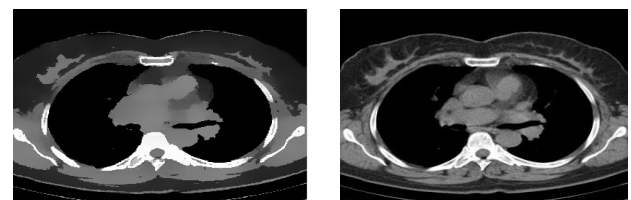


(ب) تصویر الگوریتم انتقال میانگین (الف) تصویر اصلی

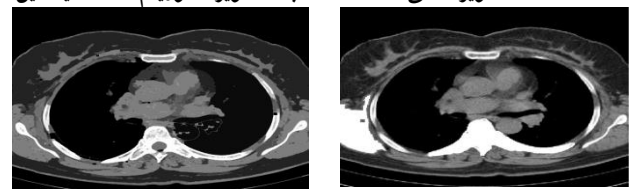


(د) تصویر روش پیشنهادی (ج) تصویر الگوریتم آب‌پخشان

شکل ۲: نتیجه پیاده‌سازی ۳ روش روی تصویر ریه فرد اول

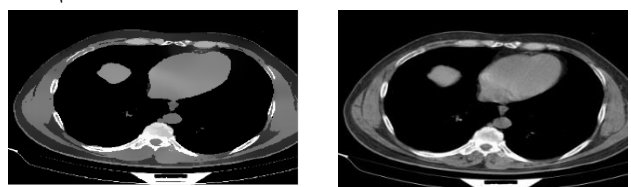


(ب) تصویر الگوریتم انتقال میانگین (الف) تصویر اصلی



(د) تصویر روش پیشنهادی (ج) تصویر الگوریتم آب‌پخشان

شکل ۳: نتیجه پیاده‌سازی ۳ روش روی تصویر ریه فرد دوم



(ب) تصویر الگوریتم انتقال میانگین (الف) تصویر اصلی

- specific image segmentation using reinforcement learning." *IEEE transactions on medical imaging* 32, no. 5 (2013): 943-956.
- [21] Aljabar, Paul, Robin Wolz, Latha Srinivasan, Serena J. Counsell, Mary A. Rutherford, A. David Edwards, Joseph V. Hajnal, and Daniel Rueckert. "A combined manifold learning analysis of shape and appearance to characterize neonatal brain development." *IEEE transactions on medical imaging* 30, no. 12 (2011): 2072-2086.
- [22] Yan, Jun, Haibo He, Xiangnan Zhong, and Yufei Tang. "Q-learning-based vulnerability analysis of smart grid against sequential topology attacks." *IEEE Transactions on Information Forensics and Security* 12, no. 1 (2017): 200-210.
- [23] Wei, QingLai, Ruizhuo Song, Benkai Li, and Xiaofeng Lin. "A Novel Policy Iteration-Based Deterministic Q-Learning for Discrete-Time Nonlinear Systems." In *Self-Learning Optimal Control of Nonlinear Systems*, pp. 85-109. Springer Singapore, 2018.
- [24] Wei, Qinglai, Frank L. Lewis, Qiuye Sun, Pengfei Yan, and Ruizhuo Song. "Discrete-time deterministic \$ Q \$-learning: A novel convergence analysis." *IEEE transactions on cybernetics* 47, no. 5 (2017): 1224-1237.
- [25] Morozs, Nils, Tim Clarke, and David Grace. "Cognitive spectrum management in dynamic cellular environments: A case-based Q-learning approach." *Engineering Applications of Artificial Intelligence* 55 (2016): 239-249.
- [26] Shi, Guang, Derong Liu, and Qinglai Wei. "Echo state network-based Q-learning method for optimal battery control of offices combined with renewable energy." *IET Control Theory & Applications* 11, no. 7 (2017): 915-922.
- [27] Duan, Jiajun, Hao Xu, and Wenxin Liu. "Q-learning based Damping Control of Wide-area Power Systems under Cyber Uncertainties." *IEEE Transactions on Smart Grid* (2017).
- [28] Luo, Biao, Derong Liu, Tingwen Huang, and Ding Wang. "Model-free optimal tracking control via critic-only Q-learning." *IEEE transactions on neural networks and learning systems* 27, no. 10 (2016): 2134-2144.
- [29] Díaz, Henry, Leopoldo Armesto, and Antonio Sala. "Improvement of LMI controllers of Takagi-Sugeno models via Q-learning." *IFAC-PapersOnLine* 49, no. 5 (2016): 67-72.
- [30] de Freitas, Nando. *Learning to Communicate to Solve Riddles with Deep Distributed Recurrent Q- Networks*. No. arXiv: 1602.02672. 2016.
- ترکیب روش‌های آستانه گذاری مبتنی بر الگوریتم یادگیری تقویتی" مجله مهندسی برق، دوره ۴، شماره ۱۶، ۳۷-۴۸، زمستان ۹۲.
- [3] Pont-Tuset, Jordi, et al. "Multiscale combinatorial grouping for image segmentation and object proposal generation." *IEEE transactions on pattern analysis and machine intelligence* 39.1 (2017): 128-140.
- [4] Liang-Chieh, Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan Yuille. "Semantic image segmentation with deep convolutional nets and fully connected crfs." In *International Conference on Learning Representations*. 2015.
- [5] Swierczynski, Piotr, Bartłomiej W. Papież, Julia A. Schnabel, and Colin Macdonald. "A Level-set Approach to Joint Image Segmentation and Registration with Application to CT Lung Imaging." *Computerized Medical Imaging and Graphics* (2017).
- [6] Kumar, S. Logesh, M. Swathy, S. Sathish, J. Sivaraman, and M. Rajasekar. "Identification of lung cancer cell using watershed segmentation on CT images." *Indian Journal of Science and Technology* 9, no. 1 (2016).
- [7] Chen, Chii-Jen. "Image Segmentation for Lung Lesions Using Ant Colony Optimization Classifier in Chest CT." In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 283-289. Springer, Cham, 2017.
- [8] Alilou, Mehdi, Niha Beig, Mahdi Orooji, Prabhakar Rajiah, Vamsidhar Velcheti, Sagar Rakshit, Niyoti Reddy et al. "An integrated segmentation and shape-based classification scheme for distinguishing adenocarcinomas from granulomas on lung CT." *Medical Physics* (2017).
- [9] Jithy, P. K., and Philomina Simon. "Review on Automatic Segmentation Techniques in Medical Images." (2017).
- [۱۰] فاتح، منصور، کبیر، احسان اله، نیلی احمدآبادی، مجید، "کاهش رنگ در نقشه چاپی فرش به کمک یادگیری تقویت شده"، مجله مهندسی برق و مهندسی کامپیوتر، دوره ۹، شماره ۳، ۱۳۳-۱۴۲، پاییز ۱۳۹۰.
- [11] Arbelaez, Pablo, Michael Maire, Charless Fowlkes, and Jitendra Malik. "Contour detection and hierarchical image segmentation." *IEEE transactions on pattern analysis and machine intelligence* 33, no. 5 (2011): 898-916.
- [12] Bampis, Christos G., Petros Maragos, and Alan C. Bovik. "Graph-Driven Diffusion and Random Walk Schemes for Image Segmentation." *IEEE Transactions on Image Processing* 26, no. 1 (2017).
- [13] Fakhry, Ahmed, Tao Zeng, and Shuiwang Ji. "Residual deconvolutional networks for brain electron microscopy image segmentation." *IEEE transactions on medical imaging* 36, no. 2 (2017): 447-456.
- [14] Wang, Lichao, Robert Merrifield, and Guang-Zhong Yang. "Reinforcement learning for context aware segmentation." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 627-634. Springer, Berlin, Heidelberg, 2011.
- [15] Parikh, Mehul C., and Kishor G. Maradia. "Comparative analysis of motion base image segmentation using machine learning techniques." *Indian Journal of Science and Technology* 9, no. 9 (2016).
- [16] Sahba, Farhang. "Object Segmentation in Image Sequences Using Reinforcement Learning." In *Computational Science and Computational Intelligence (CSCI), 2016 International Conference on*, pp. 1416-1417. IEEE, 2016.
- [17] Dumay, Adrie, Mariëlle NAJ Claessens, Cees Roos, Jan J. Gerbrands, and Johan HC Reiber. "Object delineation in noisy images by a modified policy-iteration method." *IEEE transactions on pattern analysis and machine intelligence* 14, no. 9 (1992): 952-958.
- [18] Branson, Steve, Oscar Beijbom, and Serge Belongie. "Efficient large-scale structured learning." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1806-1813. 2013.
- [19] Pipaud, Isabel, and Frank Lehmkühl. "Object-based delineation and classification of alluvial fans by application of mean-shift segmentation and support vector machines." *Geomorphology* (2017).
- [20] Wang, Lichao, Karim Lekadir, Su-Lin Lee, Robert Merrifield, and Guang-Zhong Yang. "A general framework for context-

زیر نویس‌ها

- 1 Thresholding
- 2 Clustering
- 3 Region based segmentation
- 4 Edge based segmentation
- 5 Reinforcement learning
- 6 Graph based segmentation
- 7 Watershed
- 8 Mean shift
- 9 Agent
- 10 Action
- 11 Reward
- 12 Punishment
- 13 Policy
- 15 Density

۱۴ دکتر حبیب ذاکریان و دکتر لاله ابراهیم‌پور



نهان کاوی کور بردار حرکت ویدئوی فشرده‌سازی شده H.264

الهه سادات سادات^۱، کریم فائز^۲

^۱ دانشجوی دکتری، دانشگاه صنعتی امیرکبیر، e.sadat@aut.ac.ir

^۲ استاد دانشکده برق، دانشگاه صنعتی امیرکبیر، kfaez@aut.ac.ir

چکیده

در این مقاله، روشی برای نهان کاوی کور بردار حرکت ویدئوی فشرده‌سازی شده H.264 ارائه می‌شود. نهان کاو ارائه شده از دو دسته ویژگی استفاده می‌کند. دسته اول ویژگی‌ها از میزان جابه‌جایی بردار حرکت از مقدار بهینه‌ی آن و همچنین تفاضل تابع هزینه بردار حرکت انتخاب شده و مقدار کمینه‌ی محلی آن استخراج شده است و دسته دوم ویژگی‌ها از هیستوگرام تفاضل تابع هزینه بردار حرکت انتخاب شده و مقدار کمینه‌ی محلی آن (در هر مکان از همسایگی بردار حرکت انتخاب شده) به دست آمده است همچنین به منظور آشکار سازی درج پیام از تبدیل فوریه ویژگی‌های دسته اول نیز استفاده شده است. برای اصلاح و بهبود نتایج نهان کاوی، با توجه به احتمال آلوده بودن و یا پاک بودن هر بردار حرکت ضریبی برای هر بردار ویژگی در نظر گرفته می‌شود. نتایج نشان می‌دهد نهان کاو ارائه شده توانسته است دقت در نهان کاوی را نسبت به روش‌های قبل بهبود بخشد و با وجود افزایش بعد ویژگی با تغییر دسته بند، پیچیدگی زمانی ناشی از افزایش بعد را کاهش دهد.

کلمات کلیدی

نهان کاوی، بردار حرکت، ویدئو H.264، پنهان نگاری، تخمین حرکت

به رسانه‌ی مناسب درج پیام به منظور ارتباطات پنهان تبدیل شده است. رمزگذاری^۱ ویدئو امکان درج پیام را بیشتر از تصویر ایجاد می‌کند. پیام را می‌توان در پیکسل‌ها و یا ضرایب DCT ویدئو مانند تصویر درج شود و یا در پارامترهایی مانند بردار حرکت، نوع ماکرو بلاک، اندازه ماکرو بلاک و ... درج نمود. در بین پارامترهای فوق، درج پیام در بردار حرکت متداول‌تر است و بیشتر مورد توجه قرار گرفته است.

۱-۱ - روش‌های نهان نگاری در بردار حرکت

ویدئو

روش‌های نهان نگاری در بردار حرکت ویدئو را می‌توان به سه نسل تقسیم کرد [2]. در روش‌های نسل اول نهان نگاری، بردار حرکت مناسب برای درج پیام بر اساس اندازه بردار حرکت انتخاب می‌شود و جایگذاری پیام در کم ارزش‌ترین بیت بردار حرکت انجام می‌شود [3]. این روش‌ها ویژگی‌های آماری بردار حرکت را حفظ نمی‌کنند و با در نهان کاو با استخراج ویژگی‌های آماری قابل آشکار سازی هستند.

۱ - مقدمه

نهان نگاری اساس ارتباطات پنهان است. هدف نهان نگاری، درج پیام در رسانه و پنهان کردن وجود پیام است به صورتی که به لحاظ ادراکی و آماری وجود پیام قابل تشخیص نباشد. نهان کاوی به معنای تشخیص وجود پیام در رسانه است. به رسانه‌ی بدون پیام رسانه پاک و به رسانه‌ای که در آن پیام درج شده است رسانه‌ی آلوده گفته می‌شود. به تشخیص بین رسانه‌ی پاک و رسانه‌ی آلوده نهان کاوی گفته می‌شود. رسانه می‌تواند صوت، تصویر و یا ویدئو باشد. اگر طول پیام زیاد باشد ویدئو مناسب‌ترین رسانه برای درج پیام است. با این وجود تاکنون روش‌های نهان نگاری در تصویر به دلیل عمومیت آن، بیشتر مورد توجه پژوهشگران قرار گرفته است. تحقیقات نشان می‌دهد در ۶ سال اخیر اشتراک گذاری ویدئو ۸ برابر شده است و در سال ۲۰۱۶، ۷۳ درصد از ترافیک اینترنت برای تبادل ویدئو بوده است و پیش بینی می‌شود تا سال ۲۰۲۱ به ۸۶ درصد برسد [1]. با رشد اشتراک گذاری این رسانه، ویدئو

این دسته هستند که با بازسازی روابط تعیین حرکت و استخراج ویژگی از بردارهای حرکت بازسازی شده و SAD مربوط به آن و مقادیر رمز شده‌ی متناظرشان، سعی در تشخیص درج پیام دارند [14,15,16].

در این مقاله با ارائه‌ی ویژگی‌های حساس به درج پیام، دقت آشکارسازی افزایش داده شده است. برای غلبه بر پیچیدگی زمانی در دسته بند SVM که با افزایش تعداد داده‌های آموزشی و افزایش بعد ویژگی‌ها ایجاد می‌شود از دسته بند فیشر^۲ استفاده شد. نتایج نشان می‌دهد دسته بند فیشر در مقایسه با دسته بند SVM با کرنل گوسی برای داده‌های نهان کاوی مناسب‌تر است زیرا که ویژگی‌ها مؤثر را استخراج می‌کند و با حفظ دقت، سرعت بیشتر و پیچیدگی محاسبات کمتری دارد. در بخش دوم به تعریف پیش زمینه‌های مورد نیاز جبران سازی حرکت در رمزنگاری و روابط آن در رمزنگاری ویدئو می‌پردازیم، در بخش سوم به تعریف پارامترهای تخمین کامل حرکت و بازسازی روابط جبران سازی حرکت در کدگشا پرداخته می‌شود، در بخش چهارم روش پیشنهادی، ویژگی‌های استخراج شده جهت نهان کاوی، نوع وزن دهی آن‌ها و دسته بند انتخاب شده، بیان می‌شود. در بخش پنجم نتایج پیاده سازی و مقایسه با روش‌های قبل آورده شده است و در بخش پایانی نتیجه گیری بیان شده است.

۲- جبران سازی حرکت

استانداردهایی که بعد از H.261 ارائه شده است به منظور کاهش داده‌های تکراری در زمان، دارای تخمین برون قابی هستند. برای کنترل میزان اغتشاش و نرخ بیت به طور همزمان در کدگذاری غالباً از روش بهینه سازی لاگرانژ استفاده می‌شود. دلیل استفاده از معادله‌ی لاگرانژ سادگی و کارآمدی آن است.

هدف از تخمین حرکت، یافتن بهترین تطبیق بلوک جاری در قاب مرجع است و با استفاده از مقایسه‌ی بین بلوک جاری و بلوک‌هایی که در ناحیه جستجوی قاب مرجع هستند انجام می‌شود. منظور از قاب مرجع، آخرین قاب مرجعی است که رمز شده است. بهترین تطبیق با کمینه کردن معادله لاگرانژ نرخ بیت-اغتشاش به دست می‌آید. تابع هزینه لاگرانژ، به منظور یافتن بهترین تطبیق برای بلوک S به صورت رابطه‌ی (۱) است.

$$j = D + \lambda R \quad (1)$$

D میزان اغتشاش و R نرخ بیت را بیان می‌کنند. پارامتر λ به منظور کنترل و مصالحه بین نرخ بیت و اغتشاش تعریف شده است. پس بردار حرکتی که تابع هزینه لاگرانژ را کمینه می‌کند به صورت رابطه (۲) تعریف می‌شود.

$$mv^* = \arg \min \{D(S, S_{mv}^{ref}) + \lambda R_{motion}(mv, ref)\} \quad (2)$$

S_{mv}^{ref} ، بلوک تخمین زده شده در قاب مرجع است که جابه‌جایی مکانی آن نسبت به بلوک S به اندازه بردار حرکت mv است. تابع محاسبه کننده‌ی اغتشاش $D(S, S_{mv}^{ref})$ می‌تواند SAD یا SATD باشد. در رمزگذاری و در بخش تخمین حرکت با دقت کمتر از یک پیکسل، از SATD به جای SAD استفاده می‌شود و از رابطه (۳) و (۴) قابل محاسبه است.

روش‌های نسل دوم با تعریف کردن تابع اغتشاش مناسب، سعی در کمینه کردن این تابع با درج پیام دارند. در این روش‌ها، هدف درج تعداد بیت بالاتری از پیام به ازای تغییر یک بیت از بردار حرکت است [4]. این روش‌ها شامل کد ترلیس [5] و کد کاغذ خیس [6] و... می‌باشد. از معایب روش‌های نسل دوم، عدم حفظ بهینه‌ی محلی بودن بردار حرکت است و با دسته ویژگی‌های Aoso [7] قابل تشخیص هستند.

روش‌های نسل سوم شامل روش‌هایی است که بردار حرکت را غالباً به نواحی که در آن SAD^۲ به صورت محلی بهینه باشد، تغییر می‌دهند. از این دسته روش‌ها می‌توان MVMPLO را نام برد [8]. همچنین روش‌هایی که زاویه بردار حرکت را تغییر می‌دهند از این نسل هستند. این روش‌ها در ناحیه متناسب با پیام، بردار حرکت بهینه‌ی محلی انتخاب می‌کنند و در نتیجه بهینه بودن بردار حرکت را تا حدی حفظ می‌کنند [9,10,11]. نهان کاوی این روش‌ها در زمانی که نرخ بیت پایین باشد و یا طول پیام کم باشد، حتی با دسته ویژگی‌های Aoso نیز دشوار است. البته این روش‌ها تضمین نمی‌کنند که بردار حرکت را به بهینه‌ی محلی دیگری تغییر دهند پس اگر تخمین بردار محلی بهینه با دقت بالاتری انجام شود این روش‌ها نیز قابل آشکارسازی هستند، به این دلیل روش تخمین نزدیک به کامل بهینه‌ی محلی^۳ [۲] توانسته است با دقت مناسبی، این نسل از نهان نگاری‌ها را آشکار کند. در این روش از هر دو پارامتر نرخ بیت و اغتشاش به طور همزمان، برای بازسازی معادله لاگرانژ مربوط به جبران سازی حرکت استفاده شده است.

۱-۲- روش‌های نهان کاوی بردار حرکت ویدئو

در مقابل روش‌های نهان نگاری، روش‌های نهان کاوی با هدف تشخیص وجود پیام ارائه شده است. روش‌های نهان کاوی در بردار حرکت ویدئو را می‌توان به دو دسته کلی تقسیم نمود، دسته اول شامل روش‌هایی است که با استخراج ویژگی‌های آماری بردار حرکت سعی در تشخیص وجود پیام دارند و فرآیند درج پیام را مانند سیگنال نویز جمع شونده که به مولفه‌های عمودی و افقی بردار حرکت اضافه می‌شود مدل میکنند [12]. همچنین این روش‌ها با بررسی پیوستگی مقادیر بردار حرکت به صورت افقی، عمودی و زمانی و استخراج ویژگی‌ها بر همین اساس سعی در تشخیص وجود پیام دارند. در [13] مدل غنی شامل ۴۴۸۷۵ ویژگی ارائه شده است که این ویژگی‌ها با اعمال فیلترهای مختلف بر صفحات مکانی، افقی-زمانی و عمودی-زمانی هر دو مولفه‌ی بردار حرکت استخراج شده است. این روش به دلیل استفاده از تعداد بالای ویژگی‌ها، پیچیدگی محاسباتی بالایی دارد. این دسته روش‌ها در زمانی که ظرفیت درج پایین باشد و یا مقدار تغییر بردار حرکت کم باشد دقت پایینی دارند و دلیل آن عدم استفاده از پارامتر SAD در نهان کاوی است.

دسته دوم نهان کاوها شامل روش‌هایی است که از اطلاعات بردار حرکت، SAD و روابط تعیین بردار حرکت استفاده می‌کنند. این روش‌ها نسبت به روش‌های دسته اول دقت بالاتری دارند. دلیل دقت بالاتر این روش‌ها استخراج ویژگی‌های موثرتر و حساس تر نسبت به درج پیام است. دسته ویژگی Aoso [7] با ۱۸ ویژگی و NPE [2] با ۳۶ ویژگی از

$$\Omega(mv_x, mv_y) = \{(mv_x + i, mv_y + j) | i, j = 0, \pm 1, \dots, \pm n\} \quad (9)$$

این همسایگی و تابع هزینه به ازای هر مکان از همسایگی در شکل ۱ نمایش داده شده است.

$J_{\text{Motion}}^{\text{Motion}}(m_1^1)$	$J_{\text{Motion}}^{\text{Motion}}(m_1^2)$	$J_{\text{Motion}}^{\text{Motion}}(m_1^3)$
m_1^1 ($x_i - 1, y_i - 1$)	m_1^2 ($x_i, y_i - 1$)	m_1^3 ($x_i + 1, y_i - 1$)
$J_{\text{Motion}}^{\text{Motion}}(m_1^4)$	$J_{\text{Motion}}^{\text{Motion}}(m_1^5)$	$J_{\text{Motion}}^{\text{Motion}}(m_1^6)$
m_1^4 ($x_i - 1, y_i$)	m_1^5 (V_i) (x_i, y_i)	m_1^6 ($x_i + 1, y_i$)
$J_{\text{Motion}}^{\text{Motion}}(m_1^7)$	$J_{\text{Motion}}^{\text{Motion}}(m_1^8)$	$J_{\text{Motion}}^{\text{Motion}}(m_1^9)$
m_1^7 ($x_i - 1, y_i + 1$)	m_1^8 ($x_i, y_i + 1$)	m_1^9 ($x_i + 1, y_i + 1$)

(ب)

(الف)

شکل ۱: همسایگی ۳×۳ تعریف شده برای بردار حرکت V_i (الف) مکان متناسب با هر بردار حرکت (ب) تابع هزینه‌ی بردارهای حرکت نمایش داده در الف [2]

۴- روش پیشنهادی

در روش پیشنهادی از دو دسته ویژگی مورد استفاده قرار گرفته است. دسته اول ویژگی‌ها، اصلاح شده‌ی ویژگی‌های روش‌های ارائه شده در قبل است. دسته دوم، ویژگی‌های جدید و حساس به تغییر بردار حرکت و درج پیام هستند. در دو دسته ویژگی با استفاده از بازسازی معادله لاگرانژ در یک همسایگی $n \times n$ ($\Omega(mv_i)$) از بردار حرکت انتخاب شده، محاسبه می‌شوند.

۴-۱- ویژگی‌های دسته‌ی اول

بخش اول این ویژگی‌ها، بر مبنای احتمال قرار گرفتن بردار حرکت در هر مکان از همسایگی $\Omega(mv_i)$ با مقدار تابع هزینه‌ی کمینه ($J_{\text{min}}^{\text{SAD}}(\Omega(mv_i))$) تعریف شده است. بخش دوم بر مبنای مقدار تفاضل بین مقدار تابع هزینه‌ی کمینه و تابع هزینه مربوط به بردار حرکت انتخاب شده ($J^{\text{SAD}}(mv_i)$) تعریف شده است که مشابه ویژگی‌های ارائه شده در [2,7] است و با وزن دهی مناسب اصلاح شده‌اند و از رابطه‌ی (۱۰) و (۱۲) به دست می‌آیند.

$$f^1(k) = \frac{1}{N} \sum_{i=1}^N W(mv_i) \delta(J^{\text{SAD}}(m_i^k), J_{\text{min}}^{\text{SAD}}(\Omega(mv_i))) \quad (10)$$

($k = 1, 2, \dots, 9$)

که در آن i شماره‌ی بردار حرکت، k مکان در همسایگی Ω از بردار حرکت mv_i ، W وزن مربوط به هر بردار حرکت را نشان می‌دهد و از رابطه‌ی (۱۱) قابل محاسبه است.

$$W(mv) = \begin{cases} 0 & |mv_x| + |mv_y| = 0 \\ 1 & |mv_x| + |mv_y| > 0 \end{cases} \quad (11)$$

وزن دهی هر بردار ویژگی بر اساس احتمال درج پیام در بردار حرکت است. از آنجایی که اغلب روش‌های نهان نگاری یک مرحله انتخاب بردار حرکت دارند و در این مرحله بردار حرکت با مقدار صفر با احتمال کمتری مورد تغییر قرار می‌گیرد، همچنین در روش‌هایی که از زاویه برای درج پیام حرکت استفاده می‌کنند نیز بردار حرکت با مقدار صفر دارای پیام نیست پس وزن‌ها به صورت رابطه‌ی (۱۱) تعریف شده‌اند. بخش دوم این ویژگی‌ها به صورت رابطه‌ی (۱۲) تعریف شده‌اند.

$$SAD(S, S_{mv}^{\text{ref}}) = \sum_i \sum_j |S(i, j) - S_{mv}^{\text{ref}}(i, j)| \quad (3)$$

$$SATD(S, S_{mv}^{\text{ref}}) = \sum_i \sum_j HT(|S(i, j) - S_{mv}^{\text{ref}}(i, j)|) \quad (4)$$

R ، مجموع تعداد بیت‌های لازم برای کد کردن mv است. در صورتی که از روش قاب مرجع چندگانه استفاده کرده باشیم تعداد بیت‌های لازم برای کد کردن شاخص قاب مرجع نیز به آن اضافه می‌شود. رمزگذار H.261، اختلاف بردار حرکت جاری و بردار حرکت مرجع را به صورت EXP_golomb و یا CABAC^۱ انجام می‌دهد و توسط پرچم آن نوع رمزگذاری مشخص می‌شود. CABAC نسبت به رمزنگاری با طول متغیر بهتر می‌تواند تعداد بیت اختصاص داده شده را کاهش دهد و به لحاظ محاسباتی پیچیده‌تر است و زمان محاسبات بالاتری دارد. به همین دلیل در نرم‌افزارهای X264 و JM برای سادگی از EXP_golomb برای تخمین بیت استفاده شده است. در EXP_golomb همه مقادیر به صورت معادله (۵) کد می‌شوند. سپس هر کد به صورت باینری و با طول بیان شده در رابطه‌ی (۶) ذخیره می‌شود. بردار حرکت دارای دو جهت افقی و عمودی است و هر کدام به روش بیان شده کد می‌شوند و در نهایت مجموع طول بیت در هر دو جهت به صورت رابطه‌ی (۷) محاسبه می‌شود.

$$\text{map}(x) = \begin{cases} 2|x| & x \leq 0 \\ 2|x| - 1 & x > 0 \end{cases} \quad (5)$$

$$R_x = 2 \lceil \log_2(\text{map}(mv_x - mv_x^{\text{ref}})) \rceil + 1 \quad (6)$$

$$R = R_x + R_y \quad (7)$$

بردار حرکت به دست آمده توسط معادله لاگرانژ (۱) در نرخ بیت و میزان اغتشاش بهینه‌ی محلی است.

۳- بازسازی معادله لاگرانژ

داده‌ها در کدگذاری به صورت با اتلاف فشرده می‌شوند و در نتیجه بازسازی دقیق آن‌ها در کدگشا ممکن نیست. داده‌های بازسازی شده با داده‌های اصلی (قبل از کدگذاری)، یکسان نیستند و اطلاعات اغتشاش در معادله لاگرانژ به صورت کامل قابل بازسازی نیست. با این حال در [7] اثبات شده است که داده‌های بازسازی شده تخمین قابل اعتمادی از داده‌های اصلی هستند پس معادله لاگرانژ پس از کدگشایی به صورت نزدیک به کامل بازسازی می‌شود.

ضریب λ در معادله لاگرانژ کنترل بین نرخ بیت و اغتشاش را بر مبنای ضریب چندی سازی^۷ انجام می‌دهد. هرچه مقدار λ بیشتر باشد، معادله لاگرانژ نرخ بیت را بیشتر بهینه می‌کند و در نتیجه کیفیت ویدئو کمتر می‌شود. به منظور تخمین ضریب ثابت (λ) معادله لاگرانژ، در فشرده‌سازی H.264 می‌توان از معادله‌ی (۸) استفاده نمود [17].

$$\lambda = \sqrt{0.85 \times 2^{\left(\frac{Q-12}{3}\right)}} \quad (8)$$

که در آن Q ضریب چندی سازی است. با استفاده از روابطی که در این بخش و بخش دوم آورده شد، می‌توان معادله لاگرانژ را به صورت تقریبی بازسازی کرد. پس از بازسازی برای هر بردار حرکت معادله لاگرانژ، باید به صورت محلی در یک همسایگی Ω که به صورت رابطه‌ی (۹) تعریف شده است، کمینه باشد.

به واسطه درج پیام ایجاد شده‌اند را آشکار کند، ابعاد این دسته ویژگی‌ها ۱۸ است. در مجموع از ۷۲ ویژگی در روش پیشنهادی استفاده شده است.

۴-۳- دسته بند

یکی از بهترین دسته بندها برای داده‌های دو کلاسی، دسته بنده SVM است و در اغلب نمان کاوهای پایه مورد استفاده قرار می‌گیرد. پیچیدگی زمانی دسته بند SVM با کرنل گوسی، با افزایش تعداد داده‌های آموزشی از مرتبه‌ی ۳ افزایش می‌یابد و برای داده‌های آموزشی بزرگ مناسب نیست. معمولاً در نمان کاوها با بعد ویژگی بالا، به دلیل پیچیدگی محاسباتی مورد استفاده قرار نمی‌گیرد. دلیل پیچیدگی محاسباتی بالا علاوه بر بعد ویژگی، انتخاب و تنظیم پارامترهای تابع کرنل است. دسته بند فیشر پیچیدگی محاسباتی کمتری نسبت به SVM با کرنل دارد و برای داده‌های نمان کاو با ویژگی‌های تعریف شده (که دارای ماتریس کوواریانس تقریباً یکسانی در همه‌ی ابعاد است) مناسب‌تر است. همچنین دسته بند فیشر با نگاشت داده‌ها به فضایی که ماتریس کوواریانس برون کلاسی بیشینه و درون کلاسی کمینه شود، ویژگی‌های مؤثر را استخراج می‌کند. به عبارت دیگر این دسته بند مانند یک مرحله استخراج ویژگی عمل می‌کند و علاوه بر کاهش زمان محاسبات، به دلیل تناسب با داده‌های نمان کاو دقت را نیز حفظ می‌کند.

۵- نتایج پیاده سازی

برای پیاده سازی از نرم افزار JM19.0 استفاده شد. از پایگاه داده با ۲۸۴ ویدئو با طول تقریبی ۴۰ قاب و قالب CIF (مجموعاً ۱۱۳۶۰ قاب) استفاده شده است که همه ۳۰ قاب در ثانیه نمونه برداری شده‌اند. نرخ بیت ۳ (mb/s) قرار داده شده است. ۶۰ درصد از داده‌ها برای آموزش و ۴۰ درصد برای آزمایش مورد استفاده قرار گرفته است. هر آزمایش ۱۰ بار تکرار شده است و میانگین نتایج در این بخش ارائه شده است. برای هر ویدئو ویژگی‌های بیان شده در بخش ۴ به دست آمده است. برای مقایسه با روش‌های قبلی و ظرفیت درج یکسان، از معیار نسبت تعداد بردار حرکت‌هایی که تغییر داده شده به کل بردارهای حرکت ویدئو (CMVR^۱) استفاده شده است. معیار دقت در نمان کاوها، نسبت تشخیص ویدئوهای آلوده و پاک به کل ویدئوها است. دقت آشکارسازی از رابطه‌ی (۱۸) قابل محاسبه است.

$$P_A = 1 - (P_F + P_M) \quad (18)$$

که در آن P_F احتمال تشخیص ویدئوی پاک به عنوان ویدئوی آلوده است و P_M احتمال تشخیص ویدئوی آلوده به عنوان ویدئوی پاک است. نمان کاو ارائه شده در مقابل روش‌های مختلفی از نمان نگاری آزمایش شده است. در آزمایش ۱ داده‌های آموزشی به صورت جفت ویدئوی پاک و آلوده به دسته بند آموزش داده می‌شوند. در قسمت دوم داده‌های آموزش و آزمایش به صورت غیر جفت آموزش داده می‌شوند.

۵-۱- آزمایش اول

در این آزمایش، داده‌های آموزشی به صورت جفت آلوده و پاک به نمان کاو آموزش داده شده است و با چند روش نمان نگاری نتایج مورد

$$f^2(k) = \frac{1}{Z} \sum_{i=1}^N W(mv_i) \cdot \exp\left\{\frac{J^{SAD}(mv_i) - J_{\min}^{SAD}(\Omega(mv_i))}{J^{SAD}(mv_i)}\right\} \cdot \delta(J^{SAD}(m_i^k), J_{\min}^{SAD}(\Omega(mv_i))) \quad (12)$$

$$(k=1,2,\dots,9)$$

که در آن Z پارامتری برای نرمال کردن دسته ویژگی است و به صورت رابطه‌ی (۱۳) تعریف می‌شود.

$$Z = \sum_{k=1}^9 \sum_{i=1}^N W(mv_i) \cdot \exp\left\{\frac{J^{SAD}(mv_i) - J_{\min}^{SAD}(\Omega(mv_i))}{J^{SAD}(mv_i)}\right\} \cdot \delta(J^{SAD}(m_i^k), J_{\min}^{SAD}(\Omega(mv_i))) \quad (13)$$

تابع اغتشاش در تابع هزینه می‌تواند SAD یا SATD باشد. تعداد ویژگی‌های دسته اول ۱۸ است.

۴-۲- ویژگی‌های دسته‌ی دوم

بردار ویژگی جدید طراحی شده، بر اساس هیستوگرام تفاضل تابع هزینه بردار حرکت انتخاب شده و مقدار کمینه‌ی محلی برای هر مکان از همسایگی آن $(\Omega(mv_i))$ تعریف شده است. ویژگی‌های دسته‌ی اول به دلیل استفاده مستقیم از مقدار $J^{SAD}(mv_i) - J_{\min}^{SAD}(\Omega(mv_i))$ و در نظر نگرفتن عدم امکان بازسازی دقیق تابع هزینه، احتمال خطا دارد و ممکن است اختلاف در محاسبه‌ی تابع اغتشاش را به جای درج پیام آشکار کند. ویژگی‌های دسته دوم به دلیل در نظر گرفتن بازه‌هایی برای مقدار $J^{SAD}(mv_i) - J_{\min}^{SAD}(\Omega(mv_i))$ ، می‌تواند خطای ناشی از عدم بازسازی کامل تابع هزینه کمتر کند. ویژگی‌های (f^3) به صورت هیستوگرام نرمال شده‌ی $J^{SAD}(mv_i) - J_{\min}^{SAD}(\Omega(mv_i))$ در هر مکان از همسایگی $(\Omega(mv_i))$ به صورت رابطه‌ی (۱۴) تعریف شده است و در نتیجه برای هر مکان از همسایگی، یک بردار هیستوگرام به عنوان ویژگی استخراج می‌شود. پس برای یک همسایگی 3×3 و هیستوگرام با ۴ بازه^۸، ۳۶ ویژگی استخراج می‌شود. مجموعاً ویژگی‌های دسته اول و دوم برای دسته بندی مورد استفاده قرار می‌گیرد.

$$f^3(k) = \frac{1}{N} \cdot \text{histogram}\{W(mv_i) \cdot \exp\left(\frac{J^{SAD}(mv_i) - J_{\min}^{SAD}(\Omega(mv_i))}{J^{SAD}(mv_i)}\right) \cdot \delta(J^{SAD}(m_i^k), J_{\min}^{SAD}(\Omega(mv_i))) \mid i=1,2,\dots,N\} \quad (14)$$

$$(k=1,2,\dots,9)$$

ویژگی دیگری که استفاده شده است با تبدیل فوریه‌ی یک بعدی از بردار ویژگی‌های $f^2(k)$ به صورت رابطه‌ی (۱۵) و (۱۶) به دست می‌آید.

$$f^4 = \left| \text{FFT}(\{f^1(1), f^1(2), f^1(3), f^1(6), f^1(9), f^1(8), f^1(7)\}) \right| \quad (15)$$

$$f^5 = \left| \text{FFT}(\{f^2(1), f^2(2), f^2(3), f^2(6), f^2(9), f^2(8), f^2(7)\}) \right| \quad (16)$$

$$\text{FFT}(x) = \sum_{n=0}^{N-1} x_n \exp\left(\frac{2\pi i}{N} kn\right) \quad k=1,\dots,N-1 \quad (17)$$

این دسته ویژگی با استخراج ویژگی‌های فرکانسی تغییرات $J^{SAD}(mv_i) - J_{\min}^{SAD}(\Omega(mv_i))$ حول بردار حرکت mv_i می‌تواند تغییراتی که

۷۴,۳۸	Aoso[7]
۸۶,۵۲	NPE [2]
۸۷,۰۱۷	روش پیشنهادی

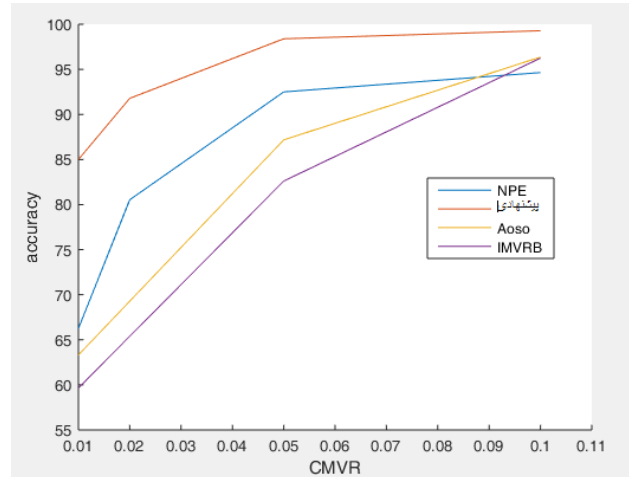
۶- نتیجه گیری

با افزایش اشتراک گذاری ویدئو، مناسب‌ترین رسانه برای درج پیام با طول زیاد است و نهان‌کاری ویدئو بیشتر مورد توجه قرار گرفته است. در این مقاله، ویژگی‌های جدید و حساس به درج پیام ارائه شد. نتایج نشان می‌دهد نهان‌کاری ارائه شده توانسته است علاوه بر بهبود دقت آشکارسازی، دقت آشکارسازی نهان‌کاری کور (با ویدئوهای آلوده توسط روش‌های متفاوت نهان‌کاری و نرخ درج‌های متفاوت به صورت غیر جفت) را نیز حفظ کند. در این مقاله ویدئو با استاندارد H.264 را مورد توجه قرار گرفت، در پژوهش‌های آینده می‌توان نهان‌کاری بردار حرکت در ویدئو با استاندارد HEVC را نیز مورد بررسی قرار داد. همچنین می‌توان با افزایش تعداد ویژگی‌ها و انتخاب دسته بند مناسب برای داده‌ها با ابعاد ویژگی بالا (مانند دسته بند گروهی) در پژوهش‌های آینده نتایج نهان‌کاری کور را بهبود بخشید.

مراجع

- [1] <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>
- [2] H. Zhang, Y. Cao and X. Zhao, "A Steganalytic Approach to Detect Motion Vector Modification Using Near-Perfect Estimation for Local Optimality," in *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 2, pp. 465-478, Feb. 2017
- [3] H. A. Aly, "Data hiding in motion vectors of compressed video based on their associated prediction error," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 14-18, Mar. 2011.
- [4] Y. Yao, W. Zhang, N. Yu, and X. Zhao, "Defining embedding distortion for motion vector-based video steganography," *Multimedia Tools Appl.*, vol. 74, no. 24, pp. 11163-11186, Dec. 2015.
- [5] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 920-935, Sep. 2011.
- [6] J. Fridrich, M. Goljan, P. Lisonek, and D. Soukal, "Writing on wet paper," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3923-3935, Oct. 2005.
- [7] K. Wang, H. Zhao, and H. Wang, "Video steganalysis against motion vector-based steganography by adding or subtracting one motion vector value," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 5, pp. 741-751, May 2014.
- [8] H. Zhang, Y. Cao, and X. Zhao, "Motion vector-based video steganography with preserved local optimality," *Multimedia Tools Appl.*, vol. 75, no. 21, pp. 13503-13519, 2016.
- [9] F. Pan, L. Xiang, X.-Y. Yang, and Y. Guo, "Video steganography using motion vector and linear block codes," in *Proc. IEEE Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, pp. 592-595, Jul. 2010.
- [10] X. He and Z. Luo, "A novel steganographic algorithm based on the motion vector phase," in *Proc. Int. Conf. Comput. Sci. Softw. Eng.*, vol. 3, pp. 822-825, Dec. 2008.
- [11] D.-Y. Fang and L.-W. Chang, "Data hiding for digital video with phase of motion vector," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, p. 4, May 2006.

ارزیابی قرار گرفته است. شکل ۲ شکل، نمودار دقت آشکارسازی نهان‌کاری ارائه شده نسبت به ظرفیت‌های درج پیام متفاوت (CMVR) را نشان می‌دهد، همچنین دقت آشکارسازی نهان‌کاری Aoso [7]، NPE [2] و IMVRB[14] در مقابل روش نهان‌کاری Yao [12] نشان داده شده است.



شکل ۲: دقت آشکارسازی روش پیشنهادی و روش‌های Aoso[7]، IMVRB[14] و NPE [2]

در جدول 1، مقایسه زمان اجرای روش پیشنهادی با دسته بند فیشر و روش SVM با کرنل گوسی که غالباً برای نهان‌کاری کور مورد استفاده قرار می‌گیرد آورده شده است.

جدول 1

روش	زمان (ثانیه)
SVM با کرنل گوسی	۲,۷۶
فیشر	۰,۰۹۲

۵-۲- آزمایش دوم

این آزمایش برای آزمودن نهان‌کاری ارائه شده نسبت به روش‌های نهان‌کاری ای است که آموزش ندیده است و اصطلاحاً آزمودن نهان‌کاری کور است. آموزش داده‌ها به صورت همزمان جهت آموزش استفاده نمی‌شود. در این آزمایش ویدئوها با ظرفیت درج‌های مختلف به صورت در هم شده آموزش داده می‌شود. در مرحله آزمایش نیز به همین صورت عمل می‌شود و نهان‌کاری نسبت به یک ظرفیت درج ثابت آموزش داده نمی‌شود. در مجموع از ۲۲۷۲ ویدئوی پاک و آلوده (تعداد متوسط ۴۰ قاب برای هر ویدئو) برای آموزش و آزمایش استفاده شده است. نتایج آزمایش ۲ برای روش پیشنهادی و چند روش نهان‌کاری دیگر در جدول 2 آورده شده است.

جدول 2: دقت آشکار سازی در آموزش غیر جفت

روش	دقت آشکارسازی
IMVRB[14]	۷۴,۹۰

- [12] Y. Su, C. Zhang, and C. Zhang, "A video steganalytic algorithm against motion-vector-based steganography," *Signal Process.*, vol. 91, no. 8, pp. 1901–1909, Aug. 2011.
- [13] K. Tasdemir, F. Kurugollu and S. Sezer, "Spatio-Temporal Rich Model-Based Video Steganalysis on Cross Sections of Motion Vector Planes," in *IEEE Transactions on Image Processing*, vol. 25, no. 7, pp. 3316–3328, July 2016.
- [14] P. Wang, Y. Cao, X. Zhao, and B. Wu, "Motion vector reversion-based steganalysis revisited," in *Proc. IEEE China Summit Int. Signal Inf. Process. (ChinaSIP)*, pp. 463–467, Jul. 2015.
- [15] Y. Cao, X. Zhao, and D. Feng, "Video steganalysis exploiting motion vector reversion-based features," *IEEE Signal Process. Lett.*, vol. 19, no. 1, pp. 35–38, Jan. 2012.
- [16] Y. Deng, Y. Wu, and L. Zhou, "Digital video steganalysis using motion vector recovery-based features," *Appl. Opt.*, vol. 51, no. 20, pp. 4667–4677, Jul. 2012.
- [17] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 688–703, Jul. 2003.

زیرنویس‌ها

- 1 encoding
- 2 Sum of absolute differences
- 3 Near-Perfect Estimation for Local Optimality
- 4 fisher
- 5 index
- 6 Contextbased Adaptive Binary Arithmetic Coding
- 7 Quantization parameter
- 8 bin
- 9 corrupted MV ratio



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه تبریز

به کارگیری الگوریتم آدابوست و درخت تصمیم در تشخیص بیماری قلبی عروق کرونر

زینب حسنی^۱، نجمه صمدیانی^۲

^۱ مربی، گروه کامپیوتر، دانشگاه کوثر بجنورد، بجنورد، hassani@kub.ac.ir

^۲ مربی، گروه کامپیوتر، دانشگاه کوثر بجنورد، بجنورد، nsamadiani@kub.ac.ir

چکیده

امروزه، ابزارهای کامپیوتری نقش موثری را در حوزه‌های مختلف علم پزشکی مانند تشخیص و درمان بیماری‌ها ایفا می‌کنند. این مقاله، با استفاده از دو الگوریتم یادگیری ماشین و داده‌کاوی، به تشخیص بیماری قلبی عروق کرونر می‌پردازد. سیستم طراحی شده، دو تکنیک درخت تصمیم و الگوریتم آدابوست را برای شناسایی افراد بیمار قلبی و سالم مربوط به پایگاه داده‌ی کلیولند به کار می‌گیرد. از آنجایی که تعداد مشخصه‌های پایگاه داده زیاد است، درخت تصمیم مشارکت هر مشخصه را برای تشخیص بهتر مشخص می‌کند و یک ابزار مفید برای این نوع مسائل به حساب می‌آید. در این سیستم، توانستیم بعد از کاهش مشخصه‌ها از ۱۳ به ۸ مورد با استفاده از درخت تصمیم، با دقت ۹۳,۰۷٪ بیماری عروق کرونر را تشخیص دهیم. هم‌چنین، الگوریتم آدابوست یکی از الگوریتم‌های بوستینگ نیز توانست افراد مبتلا به این بیماری را با دقت ۹۰,۳۷ درصد مشخص کند. نتایج حاصل، برتری سیستم پیشنهادی را نسبت به سایر پژوهش‌های موجود در این حوزه نشان می‌دهد.

کلمات کلیدی

درخت تصمیم، الگوریتم‌های گروهی، آدابوست، بیماری عروق کرونر، بیماری قلبی.

کرونری نمی‌توانند بر اساس نیاز اکسیژن این عضلات، به آنها خون و اکسیژن برسانند و در نتیجه شخص دچار درد سینه و آنژین قلبی می‌گردد. در حالت پیشرفته، اگر به علت مسدود شدن سرخرگ تاجی توسط یک لخته خون، به طور کامل جلوی خونرسانی آن گرفته شود، قسمتی از عضله قلب که دیگر خون به آن نمی‌رسد، خواهد مرد و این به سکتة قلبی می‌انجامد [۳].

چندین روش تشخیصی برای بیماری عروق کرونر مانند تست ورزش، الکتروکاردیوگرافی (ECG)، آنژیوگرافی یا کاتتریزاسیون قلبی وجود دارند. اما تمام این روش‌ها محدودیت‌هایی مانند درد بیمار و عدم دقت کافی در تشخیص را دارند که پزشکان را به استفاده از روش‌های کامپیوتری ترغیب می‌کنند [۴]. روش‌های مبتنی بر کامپیوتر

۱ - مقدمه

بیماری عروق کرونر نه تنها عمومی‌ترین نوع بیماری قلبی است بلکه عامل رو به رشد مرگ در میان زنان و مردان است [۱]. این بیماری زندگی ۱۷ میلیون از افراد سراسر جهان را می‌گیرد و بیشترین علت مرگ نسبت به سایر بیماری‌های قلبی عروقی است [۲]. در بیماری عروق کرونر، سرخرگ‌های تاجی تنگ و باریک می‌شوند (استنوسیس) و عضلات قلب از رسیدن خون و اکسیژن کافی محروم می‌گردند. در این صورت ممکن است در حالت استراحت اشکالی برای فرد ایجاد نشود، اما وقتی که قلب مجبور باشد کار بیشتری انجام دهد، سرخرگ‌های

۲- پایگاه داده

در این مقاله، از داده‌های قلبی پایگاه داده‌ی دانشگاه کالیفرنیا استفاده می‌کنیم. این داده‌ها توسط مرکز بیماری قلبی مجارستان و داده‌های کلینیک کلیولند جمع‌آوری شده است [۱۹]. این داده‌ها، شامل ۷۶ مشخصه هستند که در تمام آزمایشات منتشر شده تنها از ۱۳ مورد آنها به عنوان مشخصه‌های ورودی و از ۱ مورد به عنوان خروجی استفاده شده است. مشخصه‌های ورودی عبارتند از سن، فشارخون در حال استراحت، کلسترول، حداکثر ضربان قلب، جنسیت، نوع درد قفسه سینه، میزان قند خون ناشتا، نتایج نوار قلب در حال استراحت، درد ناشی از ورزش، پایین‌افتادگی فاصله ST در نوار قلب ناشی از تمرین نسبت به مرحله استراحت، انحراف در شیب ST در زمان ورزش، تعداد عروق اصلی درگیر در فلوروسکوپی و اسکن تالیوم. خروجی نیز وضعیت آنژیوگرافی شخص را نشان می‌دهد. از آنجایی که با یک مساله باینری روبرو هستیم، خروجی دو مقدار "صفر" برای افراد سالم و "یک" برای شخص بیمار قلبی را دارد.

۳- درخت تصمیم

الگوریتم درخت تصمیم یکی از پرکاربردترین الگوریتم‌های داده‌کاوی است. درخت تصمیم، یک الگوریتم یادگیری بر اساس داده‌های آموزشی است که از مزایای آن می‌توان به سادگی، شفافیت و توانایی استخراج قوانین اشاره کرد. یک درخت تصمیم، یک طبقه‌بند است که به عنوان یک بخش بازگشتی از فضای نمونه تشریح می‌شود [۲۰]. سیستم یادگیری یک درخت تصمیم معمول از یک استراتژی بالا به پایین پیروی می‌کند که یک درخت ساده اما نه الزاماً ساده‌ترین درخت ساخته خواهد شد. هر گره داخلی یا غیر برگ، با یک مشخصه مشخص می‌شود. این مشخصه سوالی را در رابطه با مثال ورودی مطرح می‌کند. در هر گره داخلی به تعداد جواب‌های ممکن با این سوال شاخه وجود دارد که هر یک با مقدار آن جواب مشخص می‌شوند. هم‌چنین، برگ‌های این درخت یک کلاس و یا یک دسته از جواب‌ها را نشان می‌دهند. هر مسیر به گرهی برگ در درخت تصمیم، یک قانون کلاس‌بندی را نشان می‌دهد. کلید ساختن یک درخت تصمیم موثر انتخاب معیار جداسازی خوب است. در این مقاله، شاخص گوناگونی جینی^۱ به عنوان معیار جداسازی انتخاب شده است که مقدار ناخالصی جینی $d(t)$ در گره t به صورت زیر محاسبه می‌شود:

$$i(t) = 1 - S \quad (1)$$

که $S = \sum p^2(j|t)$, $for = 0, 1, 2, \dots, k$ (معیار ناخالصی) است. K تعداد کلاس‌های موجود در آن گره است و $p(j|t)$ به تکرار کلاس j در گره t اشاره می‌کند. شاخص گوناگونی جینی یک گره زمانی به مقدار ماکسیمم خود می‌رسد که همه‌ی کلاس‌ها در گره با احتمال مساوی رخ می‌دهد و زمانی حداقل است که گره شامل تنها یک کلاس باشد [۲۱].

۴- الگوریتم آدابوست

الگوریتم‌های گروهی یا به عبارت دیگر "فرا الگوریتم‌ها"، چندین الگوریتم یادگیری را با هم ترکیب می‌کنند تا مدل قوی‌تری بسازند. در

که مشخصه‌های موثر را استخراج می‌کنند و از آنها در کلاس‌بندها برای تشخیص زودهنگام بیماری‌ها استفاده می‌کنند بر این مشکلات غالبند. به طور کلی، تعدادی از ریسک فاکتورها مسبب بیماری عروق کرونر هستند. بررسی منابع مختلف نشان می‌دهد که ریسک فاکتورهای موثر بیماری عروق کرونر شامل: مصرف سیگار، فشارخون بالا، اختلالات چربی خون (کلسترول تام بالا، تری‌گلیسیرید بالا، LDL بالا و HDL پایین)، دیابت، عدم فعالیت فیزیکی، چاقی، سن، جنسیت و سابقه خانوادگی می‌باشد [۵]. بر اساس این ریسک فاکتورها، الگوریتم‌های مختلفی مانند درخت تصمیم [۶-۸]، بردار پشتیبان خطی [۴-۹] و انواع شبکه‌های عصبی برای تشخیص و پیشگیری بیماری عروق کرونر ارائه و توسعه داده شده‌اند.

در [۱۰]، با استفاده از سیستم بازشناسی هوش مصنوعی و نزدیکترین همسایگی (KNN) سیستمی برای تشخیص بیماری عروق کرونر با صحت ۸۷ درصد پیشنهاد شده است. نویسندگان کار [۱۱]، داده‌های تست ورزش و ماشین بردار پشتیبان را به کار گرفتند و به صحت ۸۱٫۴۶ درصد در شناسایی بیماری عروق کرونر دست یافتند. با میان‌گیری از عملکرد چندین شبکه عصبی مختلف، یک سیستم تشخیص بیماری عروق کرونر با صحت ۸۹٫۰۱ درصد پیشنهاد شده است [۱۲]. یک روش ترکیبی تشخیص بیماری عروق کرونر در [۱۳] پیشنهاد شده است. این روش قادر است قدرت یادگیری و عملکرد شبکه عصبی را با بهینه‌سازی وزن‌های آن توسط الگوریتم ژنتیک ۱۰ درصد افزایش دهد. شبکه پیشنهادی صحت ۸۸٫۲۵٪ را در تشخیص بیماری عروق کرونر به دست آورده است. ابراهیمی و همکاران، شبکه عصبی پرسپترون چند لایه با مجموعه داده‌های ماشین بردار پشتیبان و الگوریتم بیز هسته برای مجموعه داده‌های بیماران عروق کرونر را مطالعه نموده‌اند. بهترین عملکرد برای مدل شبکه عصبی پرسپترون چند لایه برابر با ۸۰ درصد و بهترین عملکرد در الگوریتم بیز هسته با مقدار ۸۵ درصد را بدست آورده‌اند [۱۴]. خسروانیا و همکار، اطلاعات ۱۵۲ نفر را با استفاده از یک شبکه عصبی احتمالی، به گروه‌های مبتلایان به بیماری عروق کرونر و سالم با دقت ۹۰ درصد دسته‌بندی کردند [۱۵]. هم‌چنین، روش‌های مختلف انتخاب ویژگی مانند روش فیلترینگ [۱۶]، الگوریتم ژنتیک [۱۷] و سایر روش‌های انتخاب ویژگی [۱۸] برای پیش بینی بیماری عروق کرونر استفاده شده‌اند.

در این مقاله، روشی برای تشخیص بیماری عروق کرونر پیشنهاد می‌دهیم. در این روش، ابتدا با استفاده از درخت تصمیم و حذف مشخصه‌های کم اهمیت، بیماری عروق کرونر را پیش‌بینی می‌کنیم. سپس پس از حذف داده‌های پرت، در الگوریتم آدابوست داده‌های بیماران را طبقه بندی کرده و با دقت بالا افراد بیمار و سالم را تشخیص می‌دهیم.

در ادامه، در بخش دوم پایگاه داده‌ی مورد استفاده معرفی می‌شود. در بخش سوم، چهارم و پنجم به توضیح درخت تصمیم، الگوریتم آدابوست و kmeans می‌پردازیم. جزئیات روش کار در بخش ششم شرح داده می‌شود. در بخش هفتم، نتایج حاصل از اعمال روش پیشنهادی بر پایگاه داده گزارش می‌شود. نتیجه‌گیری پایان بخش مقاله خواهد بود.

مشابه نیستند. الگوریتم K-means با هدف حداقل کردن کمترین مربعات خطا، توسط مک کوئین ارائه شد. ابتدا k عضو (که k تعداد خوشه‌ها است) به صورت تصادفی از میان n عضو به عنوان مراکز خوشه‌ها انتخاب می‌شود. سپس n-k عضو باقیمانده به نزدیک‌ترین خوشه (بر اساس معیار فاصله) تخصیص می‌یابند. بعد از تخصیص همه اعضا مراکز خوشه مجدداً محاسبه می‌شوند و با توجه به مراکز جدید به خوشه‌ها تخصیص می‌یابند و این کار تا زمانی که مراکز خوشه-ها ثابت بماند ادامه می‌یابد [۲۴].

۶- روش کار

در این بخش مراحل انجام کار توضیح داده می‌شود. ابتدا برای کنترل دامنه، مقادیر کلیه ویژگی‌ها در بازه [0, 1] نرمال شده است. فرمول نرمال‌سازی به صورت زیر است:

$$X = (x - x_{MIN}) / (x_{MAX} - x_{MIN}) \quad (5)$$

که در آن x_{MAX} و x_{MIN} به ترتیب بیشترین مقدار و کمترین مقدار مربوط به هر مشخصه است.

از آنجایی که می‌خواهیم افراد را به دو دسته‌ی افراد سالم و بیمار تقسیم کنیم، تعداد کلاس‌های هدف ما دو کلاس خواهد بود. داده‌های نرمال شده را برای کلاس‌بندی به درخت تصمیم داده و با استفاده از معیار جداساز جینی، مشخصه‌های پراهمیت و کم‌اهمیت را مشخص می‌کنیم. پس از حذف مشخصه‌های کم‌اهمیت که تاثیر مثبتی در طبقه‌بندی ندارند، علاوه بر بهبود دقت طبقه‌بندی، به دلیل کم شدن بُعد داده‌ها، سرعت طبقه‌بند افزایش و از پیچیدگی سیستم کاسته می‌شود. سپس برای حذف داده‌های پرت و نویزی خوشه‌بند k-means فعالیت خود را آغاز می‌کند. بدین ترتیب با کاهش تعداد نمونه‌ها مجموعه داده جدید به عنوان ورودی به الگوریتم آدابوست داده می‌شود. یادگیرنده‌های پایه‌ی الگوریتم آدابوست، ۳۰۰ درخت هستند که با نرخ یادگیری ۰,۰۱ آموزش خود را شروع می‌کنند. تعداد داده‌های تست و آموزش در این الگوریتم، با پیاده‌سازی اعتبارسنجی ضربدری k-fold در هر مرحله مشخص می‌شود.

۷- نتایج

ابتدا داده‌های قلبی موجود را با ۱۳ مشخصه به درخت تصمیم می‌دهیم. دقت دسته‌بندی در این حالت برابر ۸۳,۸۷ درصد به دست می‌آید. با نمایش مشخصه‌های پراهمیت، اقدام به حذف مشخصه‌های کم‌اهمیت می‌کنیم. مطابق شکل ۱، تنها هشت مشخصه هستند که در دسته‌بندی، تاثیر مثبت دارند. بنابراین در آزمایش دوم، تعداد ۵ نمونه از مشخصه‌ها را حذف کرده و داده‌ها را با ۸ مشخصه‌ی پراهمیت به درخت تصمیم می‌دهیم. در نهایت، دقت طبقه‌بندی برابر ۹۳,۰۷ درصد به دست آمد. از مجموع ۳۰۳ داده‌ی موجود، تعداد ۲۸۲ نمونه به صورت صحیح در کلاس درست قرار گرفتند و ۲۱ نمونه به کلاس نادرست تعلق یافتند.

این الگوریتم‌ها، دقت مدل ترکیبی از هر کدام از مدل‌های اولیه بیشتر است. در این مقاله، برای تشخیص افراد سالم و بیمار از الگوریتم آدابوست، یکی از محبوب‌ترین الگوریتم‌های بوستینگ، استفاده می‌کنیم. بوستینگ یک فرا الگوریتم ترکیبی در حوزه یادگیری ماشینی است که برای کاهش عدم توازن و همچنین واریانس به کار می‌رود. این الگوریتم یادگیرنده‌های ضعیف را به صورت تکرار شونده آموزش داده و به مجموعه قبلی اضافه می‌نماید تا در نهایت به یک طبقه‌بند قوی دست پیدا کند. یادگیرنده‌های ضعیف در حین اضافه شدن به مجموعه، وزن دهی می‌شوند که این وزن دهی معمولاً بر اساس میزان دقت در طبقه‌بندی نمونه‌هاست. الگوریتم آدابوست، کوتاه شده‌ی عبارت "بوستینگ سازگار"، الگوریتمی برای ساختن طبقه‌بندهای قوی با ترکیب خطی از یادگیرنده‌های ضعیف است که یادگیرنده‌ها را به طور پیوسته آموزش می‌دهد. برای هر یادگیرنده با اندیس t، آدابوست خطای کلاس‌بندی وزن‌دار را محاسبه می‌کند:

$$\epsilon_t = \sum_{n=1}^N d_n^{(t)} I(y_n \neq h_t(x_n)) \quad (2)$$

که در آن، x_n یک بردار از مقادیر پیش‌بینی‌کننده‌ها با n نمونه است. Y_n برچسب کلاس درست و h_t مقدار پیش‌بینی یادگیرنده با اندیس t است. I تابع پیش‌بینی‌کننده و $d_n^{(t)}$ وزن نمونه n در مرحله t است.

سپس آدابوست، وزن‌ها را برای نمونه‌هایی که توسط یادگیرنده t درست کلاس‌بندی نشده‌اند، افزایش و وزن نمونه‌هایی که صحیح دسته‌بندی شده‌اند را کاهش می‌دهد تا در مراحل بعدی (توسط یادگیرنده‌های جدید) بیشتر مورد توجه بوده و با دقت بیشتری طبقه‌بندی گردند. یادگیرنده‌ی بعدی، $t+1$ ، داده‌ها را با وزن‌های به روزرسانی شده $d_n^{(t+1)}$ آموزش می‌دهد. زمانی که آموزش به اتمام می‌رسد، آدابوست پیش‌بینی برای داده‌های جدید را با استفاده از رابطه (۲) محاسبه می‌کند:

$$f(x) = \sum_{t=1}^T \alpha_t h_t(x) \quad (3)$$

که $\alpha_t = \frac{1}{2} \log \frac{1-\epsilon_t}{\epsilon_t}$ وزن یادگیرنده‌ی ضعیف در الگوریتم گروهی است.

آموزش آدابوست به عنوان یک حداقل‌سازی exponential loss به حساب می‌آید:

$$\sum_{n=1}^N w_n \exp(-y_n f(x_n)) \quad (4)$$

که $y_n \in \{-1, +1\}$ برچسب کلاس درست است. W_n وزن‌های نرمال شده‌ی نمونه‌ها به علاوه یک و $f(x_n) \in (-\infty, +\infty)$ امتیاز کلاس‌بندی پیش‌بینی شده است [۲۲-۲۳].

۵- الگوریتم خوشه‌بند Kmeans

خوشه‌بندی یکی از روش‌های داده‌کاوی است که بر اساس یک معیار تشابه نمونه‌های مختلف را در خوشه‌ها قرار می‌دهد به طوری که نمونه‌های یک گروه مشابهند و نمونه‌های متعلق به گروه‌های مختلف

می‌شود دقت دسته‌بندی توسط کلاسه‌بندهای پیشنهادی بیشتر از سایر روش‌های موجود است. هم‌چنین، سادگی و سریع بودن روش پیشنهادی، برتری آن را به سایر روش‌ها نشان می‌دهد.

جدول ۱- مقایسه روش پیشنهادی با سایر پژوهش‌های موجود

مرجع	روش	دقت %
[9]	ترکیب الگوریتم‌های هوش مصنوعی و KNN	۸۷
[11]	میانگین‌گیری چندین شبکه عصبی	۸۹,۰۱
[12]	شبکه عصبی + الگوریتم ژنتیک	۸۸,۲۵
[13]	الگوریتم بیز هسته	۸۵
روش پیشنهادی	درخت تصمیم با کاهش ویژگی	۹۳,۰۷
روش پیشنهادی	الگوریتم آداپوست	۹۰,۳۷

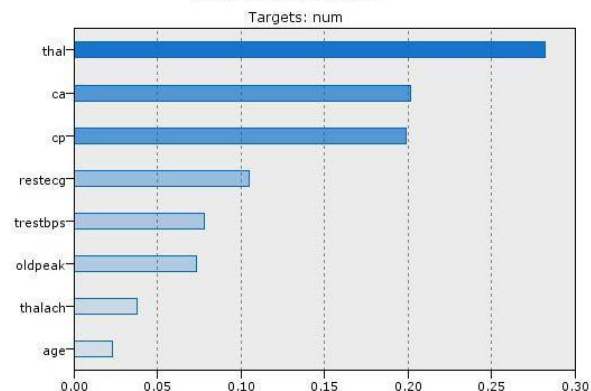
۸- نتیجه‌گیری

بیماری عروق کرونر یکی از عمومی‌ترین بیماری‌های قلبی است که سالیانه تعداد زیادی از افراد جهان به دلیل ابتلا به آن جان خود را از دست می‌دهند. به دلیل دشواری روش‌های واقعی مانند آنژیوگرافی و عدم دقت کافی در نتایج حاصل، استفاده از سیستم‌های کامپیوتری برای تشخیص این بیماری مرسوم شده است. در این مقاله، با استفاده از درخت تصمیم و الگوریتم آداپوست توانستیم با دقت ۹۳,۰۷ و ۹۰,۳۷ درصد این بیماری را تشخیص دهیم. مقایسه نتایج حاصل از آزمایش با سایر پژوهش‌های موجود، برتری روش پیشنهادی را نشان می‌دهد.

مراجع

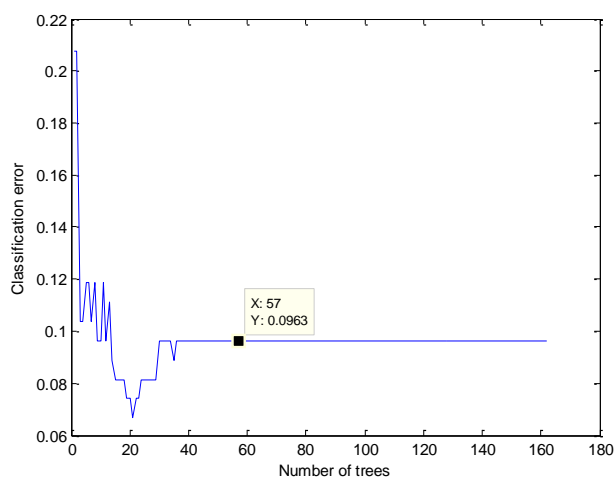
1. Coronary Artery Disease: MedlinePlus. <<https://www.nlm.nih.gov/medlineplus/coronaryarterydisease.html>> (accessed: 2015-11-18).
2. Wong, N.D., "Epidemiological studies of CHD and the evolution of preventive cardiology," Nat. Rev. Cardiol, Vol.11, pp. 276-289, 2014.
3. Buchan, K., Filannino, M., Uzuner, O., "Automatic prediction of coronary artery disease from clinical narratives," Journal of biomedical informatics, Vol.72, pp.23-32, 2017.
4. Davari Dolatabadi, A., Esmael Zadeh, S., Mohammadzadeh, B., "Automated diagnosis of coronary artery disease (CAD) patients using optimised SVM," Vol.138, pp. 117-126, 2017.
5. KarimiZarchi A., Naghie M.R., "Prevalence of Risk Factors of Coronary Heart Disease and Effect of Life-Style Modification Guides," Kowsar Medical Journal., Vol.14, pp. 157-162, 2009.
6. Shouman M, Turner T, Stocker R., "Using Decision Tree for Diagnosing Heart Disease Patients," AusDM '11 Proceedings of the Ninth Australasian Data Mining Conference, Vol.121, pp. 23-30, 2011.
7. Kochurani O.G, Aji S, Kaimal M.R., "A Neuro Fuzzy Decision Tree Model for Predicting the Risk in Coronary Artery Disease," IEEE 22nd International Symposium on Intelligent Control (ISIC), 2007.

Variable Importance



شکل ۱- اهمیت مشخصه‌های موجود در داده‌ها که اسکن تالیوم پراهمیت‌ترین مشخصه و سن کم‌اهمیت‌ترین مشخصه است.

همانطور که گفته شد داده‌ها به روش اعتبارسنجی ضربدری به الگوریتم آداپوست اعمال می‌شوند. این روش مشخص می‌کند نتایج بر روی مجموعه داده تا چه اندازه قابل تعمیم و مستقل از داده‌های آموزشی است. در این نوع اعتبارسنجی داده‌ها به K زیرمجموعه افزای می‌شوند. از این K زیرمجموعه، هر بار یکی برای اعتبارسنجی و 1-K تالی دیگر برای آموزش به کار می‌روند. این روال K بار تکرار می‌شود و همه داده‌ها دقیقاً یکبار برای آموزش و یکبار برای اعتبارسنجی استفاده می‌شوند. در نهایت، میانگین نتیجه این K بار اعتبارسنجی به عنوان یک تخمین نهایی برگزیده می‌شود. به طور معمول از 10-fold استفاده می‌شود. در این مقاله هم از 10-fold بهره بردیم. شکل ۲، کمترین خطای طبقه‌بندی به دست آمده را به ازای یادگیرنده‌های پایه نشان می‌دهد. بنابراین، دقت طبقه‌بندی برابر با ۹۰,۳۷ درصد است.



شکل ۲- کمترین خطای دسته‌بندی به دست آمده به ازای یادگیرنده‌های پایه

برای نشان دادن عملکرد بهتر روش پیشنهادی، آن را با سایر پژوهش‌های موجود در تشخیص بیماری عروق کرونر مقایسه کرده‌ایم. جدول ۱، نتایج این مقایسه را نشان می‌دهد. همانطور که مشاهده

زیرنویس‌ها

- ¹ Gini Diversity Index
² K-fold cross validation

8. El-Bialy R, Salamay M, Karam O, Essam M., "Feature Analysis of Coronary Artery Heart Disease Data sets," *Procedia Computer Science*, Vol. 65, pp. 459-468, 2015.
9. Wang J, He Q, Yao K, Rong W, Xing Y, Yue Z, "Support Vector machine (SVM) and Traditional Chinese Medicine: Syndrome Factors Based an SVM from Coronary Heart Disease Treated by Prominent Traditional Chinese Medicine Doctors," *Fifth International Conference on Natural Computation (ICNC)*, 2009.
10. Polat, K., Sahan, S. and Gunes, S., "Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn (nearest neighbour) based weighting preprocessing," *Computers in Biology and Medicines*, Vol. 32, pp. 625-631, 2007.
11. Babaoglu, I., Findik, O. and Bayrak, M., "Effects of principle component analysis on assessment of coronary artery diseases using support vector machine," *Pattern Recognition*, Vol. 37, pp. 2182-2185, 2010.
12. Das, R., Turkoglu, I. and Sengur, A., "Effective diagnosis of heart disease through neural networks ensembles," *Expert Systems with Applications*, Vol. 36, pp. 7675-7680., 2009.
13. Arabasadi, Z., Alizadehsani, R., Roshanzamir, M., Moosaei, H., Yarifard, A., "Computer aided decision making for heart disease detection using hybrid neural network- genetic algorithm," *Ccomputer methods and programs in biomedicine*, Vol. 141, pp. 19-26, 2017.
14. Shafiee H, Ebrahimi M., "Accurate Prediction of Coronary Artery Disease Using Bioinformatics Algorithm," *Qom Univ Med Sci.*, Vol.10, pp. 22-35, 2016.
15. Khosravianian A, Ayat S., "Presenting an Intelligent System for Diagnosis of Coronary Heart Disease by Using Probabilistic Neural Network," *Health Inf Manage.*, Vol.12, No.1, pp. 13-23, 2015.
16. Lee, H.G., Noh, K.Y. and Ryu, K.H., "A Data Mining Approach for Coronary Heart Disease Prediction using HRV Features and Carotid Arterial Wall Thickness," in *Proceedings of International Conference on Biomedical Engineering and Informatics*, Sanya, USA, pp. 200 - 206, 2008.
17. Anbarasi, M., Anupriya, E. and Iyengar, N., "Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm," *International Journal of Engineering Science and Technology*, Vol. 2, pp. 5370-5376, 2010.
18. Niklas, N., Halling, A., Freitag, M. and Odeberg, J., "Classifying the Severity of an Acute Coronary Syndrome by Mining Patient Data," *Computer methods and programs in biomedicine*, Vol. 8, pp. 14-24, 2009.
19. Newman, D. J., Hettich, S., Blake, C. L. S., & Merz, C. J., *UCI repository of machine learning database*. Irvine, CA: University of California, 1998.
20. Rokach, L., & Maimon, O., "Data mining with decision trees theory and applications," Singapore: World Scientific Publishing, 2008.
21. Weihong, W. et al., "A preliminary study on constructing decision tree with gene expression programming," In *Proc. 1st international conference on innovative computing, information and control*, pp. 222-225, 2006.
22. Freund, Y., Schapire, R.E., "A desicion-theoretic generalization of on-line learning and an application to boosting", In *Springer European conference on computational learning theory*, pp. 23-37, 1995.
23. Schapire, R.E., "The boosting approach to machine learning: An overview", In *Springer Nonlinear estimation and classification*, pp. 149-171, 2003.
24. Sun, Y. C., & Clark, O. G., "Implementing an intuitive reasoner with a large weather database," *International Journal of Information and Mathematical Sciences*, Vol. 5, pp. 85-102, 2009.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

بهینه‌سازی استقرار در شبکه‌های حسگر بیسیم جهت دستیابی به اتصال و پوشش کامل

امید رضا معروضی^۱، مریم رشمه کریم^۲

^۱ استادیار، گروه مهندسی برق، دانشگاه صنعتی شاهرود، شاهرود،
ormzi@yahoo.com

^۲ دانشجوی کارشناسی ارشد، گروه مهندسی برق، دانشگاه صنعتی شاهرود، شاهرود،
maryam137200@gmail.com

چکیده

شبکه‌های حسگر بیسیم مجموعه‌ای از حسگرهای متصل است که در آن هر گره حسگر قادر به جمع‌آوری برخی از اطلاعات، پردازش، ذخیره‌سازی و در صورت لزوم ارتباط برقرار کردن آن از طریق حسگرهای دیگر است تا زمانی که داده‌های جمع‌آوری شده به یک ایستگاه پایه (Sink) برسد که در آن تصمیمات لازم در نظر گرفته می‌شود. یکی از جنبه‌های طراحی بسیار حیاتی در شبکه‌های حسگر بیسیم، مسئله استقرار حسگرها می‌باشد. استقرار به دو قسمت استقرار تصادفی و استقرار قطعی (معین) تقسیم می‌شود. استقرار مناسب باید هر دو پوشش و اتصال را در نظر بگیرد. در این پژوهش به مقایسه الگوهای استقرار منظم شامل شش ضلعی، مربع، متوازی الاضلاع و مثلث متساوی الاضلاع جهت دستیابی به پوشش و اتصال پرداخته شده و در نهایت به مقایسه نتایج حاصل با نتایج بدست آمده از استقرار شبکه حسگر بیسیم توسط الگوریتم بهینه‌سازی فراابتکاری ژنتیک پرداخته و نتایج حاکی از آن است که الگوریتم فراابتکاری ژنتیک به تعداد حسگرهای کمتری برای استقرار با حفظ پوشش کامل و اتصال نیاز دارد.

کلمات کلیدی

شبکه حسگر بیسیم، استقرار، الگوریتم ژنتیک، پوشش، اتصال.

استقرار حسگر از آن جایی که هزینه و توانایی تشخیص یک شبکه حسگر بیسیم را بازتاب می‌دهد، یک مسئله حیاتی است. همچنین استقرار حسگرها در حوزه هدف برای ایجاد شبکه بیسیم، یک معیار مهم در بسیاری از برنامه‌های کاربردی است. استقرار یک شبکه حسگر بیسیم، تقریباً تمام عملکرد آن، مانند اتصال بین حسگرها، پوشش موثر شبکه و طول عمر شبکه را تحت تأثیر قرار می‌دهد. در حال حاضر در نظر گرفتن پوشش به تنهایی در هنگام مستقر شدن شبکه‌های حسگر بیسیم کافی نیست، اتصال نیز باید در نظر گرفته شود. به دلیل عوامل فراتر از کنترل انسانی، حسگرها پس از استقرار ممکن است دچار خرابی‌های غیرمنتظره شوند، بنابراین متصل بودن شبکه حسگر بیسیم به تنهایی کافی نیست، بلکه باید k -انصال باشد. برای انتقال موفقیت آمیز

۱- مقدمه

در طول دهه گذشته، شبکه‌های حسگر بیسیم به یک حوزه پژوهشی غنی، با معرفی طیف گسترده‌ای از برنامه‌های کاربردی جدید همچنان انگیز و به تازگی به یک بخش جدایی ناپذیر در مفهوم اینترنت اشیاء تبدیل شده است. گره‌های حسگر به شدت از لحاظ ذخیره‌سازی، منابع، قابلیت‌های محاسباتی، پهنای باند ارتباطی و منبع تغذیه محدودیت دارند. شبکه حسگر بیسیم با توجه به محدودیت‌های بسیاری مانند طول عمر شبکه، کارایی و عملکرد شبکه با چالش‌های بسیاری مواجه است. این چالش‌ها می‌تواند یک روش استقرار کارآمد را مدیریت کند.

بررسی استقرار تصادفی در شبکه حسگر بیسیم پرداختند. آنها همچنین عملکرد ترکیبی از مدل انتشار ساده که تعداد زیادی از گره‌ها را در اطراف سینک قرار داده و انتشار پیوسته که میزان پوشش و نرخ اتصال بالایی را فراهم می‌کند، بررسی کردند [10]. لیو در سال ۲۰۱۵ انواع روش‌های استقرار در شبکه حسگر بیسیم و محدودیت‌های آن را ارائه نمود [11]. ریبا و همکاران در سال ۲۰۱۵ به بهینه‌سازی استقرار برای دستیابی به پوشش و اتصال در شبکه حسگر بیسیم پرداختند. آنها در این پژوهش از مدل خطی عدد صحیح و همچنین الگوریتم‌های فراابتکاری جستجوی محلی و ژنتیک برای بهینه‌سازی استفاده نمودند و در نهایت به مقایسه نتایج بدست آمده از روش‌های مختلف پرداختند [12]. گوپتا و همکاران در سال ۲۰۱۶ به طبقه بندی مسائل پوشش در شبکه حسگر بیسیم و روش‌های حل و طراحی آن پرداختند [13]. در این پژوهش به بررسی تعداد حسگرهای مورد نیاز جهت استقرار، برای دستیابی به پوشش و اتصال کامل با استفاده از الگوریتم فراابتکاری ژنتیک و الگوهای منظم پرداخته شده است.

۲- پارامترهای موثر بر طراحی شبکه حسگر

بیسیم

پارامترهای مختلفی بر طراحی شبکه حسگر بیسیم و تعداد گره‌های مورد نیاز جهت استقرار وجود دارند.

۱-۲- یکنواختی

حسگرهای نزدیک به سینک معمولاً در طول انتقال داده بیش از دیگران شرکت می‌کنند. بنابراین، برای حفظ تقریباً یک سطح انرژی در کل شبکه، تعدادی گره را می‌توان در نزدیکی سینک قرار داد. تراکم گره یکنواخت، احتمال تجمع و سوراخ پوشش را کاهش می‌دهد. یکنواختی را می‌توان به عنوان میانگین انحراف استاندارد محلی از فاصله بین گره تعریف کرد:

$$U = \frac{1}{N} \sum_{i=1}^N U_i \quad (1)$$

$$U_i = \left(\frac{1}{K_i} \sum_{j=1}^{K_i} (d_{ij} - M_i) \right)^{1/2} \quad (2)$$

N تعداد کل گره‌ها است، K_i تعداد همسایگی گره i ام است. d_{ij} فاصله بین گره‌های i و j و M_i میانگین فاصله‌های بین مودال، بین گره i و همسایگان آن است [14].

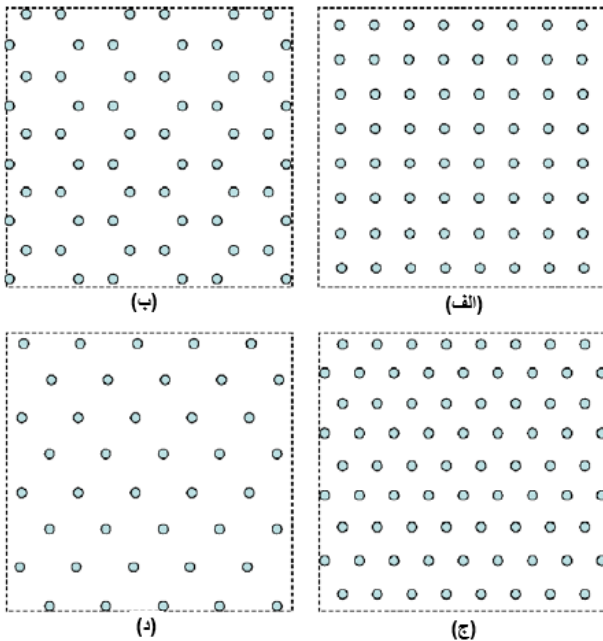
۲-۲- چگالی گره

تراکم پذیری به دلیل تاثیر بر پوشش، هزینه و عملکرد شبکه یکی از پارامترهای مهم طراحی در شبکه‌های حسگر بیسیم است. مناطق با چگالی بالا باعث افزایش هزینه شبکه و هزینه بالای محاسبات می‌شوند، درحالی‌که مناطق با تراکم کم می‌توانند مسئله سوراخ‌های پوشش یا تفکیک کردن شبکه را به وجود آورند. توزیع گره یکسان یا غیر یکنواخت است. حسگر نزدیک به سینک در انتقال داده استفاده می‌شود، بنابراین تراکم گره در نزدیکی سینک باید بیشتر باشد. تراکم گره یکنواخت شانس خوسه بندی و سوراخ پوشش را کاهش می‌دهد. تعداد

داده‌های حس شده، هر گره از شبکه باید با هر مسیر ممکن از یک یا چند هاپ به سینک متصل شود. اتصال برای حفظ توپولوژی شبکه مهم است. توپولوژی بر روی بسیاری از ویژگی‌های شبکه مانند تأخیر، توان و ظرفیت تاثیر می‌گذارد. پیچیدگی مسیریابی و پردازش داده نیز به توپولوژی بستگی دارد. اصطلاح K -اتصال ($K \geq 1$) بدان معنی است که حداقل k مسیر مجزا بین هر جفت از حسگر در یک شبکه گیرنده بیسیم وجود دارد. اگر همواره اتصال شبکه (احتمالاً بیش از چند بار متوقف می‌شود) بین هر دو گره وجود داشته باشد، شبکه گفته شده است که متصل است. اتصال یک مسئله باز است که در هنگام استقرار در نظر گرفته می‌شود. از دست دادن پوشش در شبکه ممکن است توسط برنامه‌های کاربردی قابل تحمل باشد، اما از دست رفتن اتصال می‌تواند غیر قابل جبران باشد زیرا می‌تواند تمام یا بخش‌هایی از شبکه را بی‌فایده کند و داده‌های منتجش نمی‌تواند به ایستگاه پایانی برسند. بنابراین وجود درجه بالایی از اتصال در شبکه‌های حسگر بیسیم مطلوب است. هر حسگر دارای یک شعاع ارتباطی (r_c) و یک شعاع حسی (r_s) است. تعیین الگوی استقرار بهینه برای هر دو پوشش و k -اتصال برای مقادیر کلی k و r_c/r_s یک مسئله بسیار پیچیده است [1]. الگوهای منظم در شبکه‌های حسگر بیسیم با توجه به دو دلیل آسان بودن استقرار و دستیابی به درجه بالاتری از اتصال حائز اهمیت است. چهار الگو شش ضلعی، مربع، متوازی الاضلاع و مثلث متساوی الاضلاع برای استقرار به طور منظم وجود دارد. در سال‌های اخیر پژوهش‌های بسیاری در زمینه استقرار شبکه‌های حسگر بیسیم انجام شده است، در زیر به برخی از آنها اشاره می‌کنیم.

در سال ۲۰۰۳ وانگ و همکاران به بررسی پوشش جامع و پیکربندی اتصال در شبکه‌های حسگر بیسیم پرداختند [2]. در سال ۲۰۰۳ ژو و چاکرابرتی در طی پژوهشی به مسئله استقرار حسگر و محلی‌سازی هدف با استفاده از روش نیروی مجازی پرداختند [3]. وانگ و همکاران در سال ۲۰۰۵ طی پژوهشی به مطالعه و بررسی مسئله بهینه‌سازی استقرار در شبکه حسگر بیسیم با در نظر گرفتن محدودیت طول عمر و حداقل هزینه پرداختند [4]. در سال ۲۰۰۶ بای و همکاران در پژوهشی دیگر به بررسی استقرار حسگرها به منظور دستیابی به پوشش و اتصال پرداختند. آنها در این پژوهش یک الگوی استقرار بهینه برای دستیابی به پوشش و ۲-اتصال ارائه کرده و ثابت کردند که این الگو برای تمام مقادیر r_c/r_s بهینه است [5]. در سال ۲۰۰۷ عزیز و همکاران به بهینه‌سازی مسئله پوشش با استفاده از الگوریتم فراابتکاری ازدحام ذرات و دیاکرام ورونی پرداختند [6]. ژو و همکاران در سال ۲۰۱۲ مسئله پوشش و اتصال در شبکه‌های حسگر بیسیم را مورد بررسی قرار دادند. آنها در این پژوهش به اثبات رابطه‌ای بین پوشش و اتصال و همچنین به مقایسه ابزارهای شبیه‌سازی پرداختند [7]. در سال ۲۰۱۲ اکور و تاکور در مطالعه‌ای دیگر مسئله استقرار شبکه‌های حسگر بیسیم و پارامترهای موثر بر طراحی شبکه حسگر را مورد بررسی قرار دادند [8]. در پژوهشی دیگر رانی سوبانا شلا و دواراجان در سال ۲۰۱۲، به بهینه‌سازی استقرار گره‌های حسگر با استفاده از الگوریتم بهینه‌سازی ازدحام ذرات چند هدفه و شبکه فازی عصبی پرداختند [9]. دف و یاسر در سال ۲۰۱۴ در مقاله‌ای به طبقه بندی مسائل استقرار در شبکه حسگر بیسیم و روش‌های حل آن پرداختند [1]. سنوچی و همکاران در سال ۲۰۱۴ به

کاربرد دارند. الگوی شبکه مثلثی حداقل ۶ اتصال، شبکه مربعی حداقل ۴ اتصال، متوازی الاضلاع بسته به شکل آن حداقل ۴ یا ۶ اتصال و شش ضلعی حداقل ۳ اتصال را فراهم می‌کند. این مقادیر از اتصال برای فواصل میانگرمه مناسب در چهار الگو به طوری است که هر حسگر می‌تواند به طور مستقیم با نزدیکترین همسایه خود ارتباط برقرار کند. هنگامی که $r_c \geq \sqrt{3} r_s$ ، الگوی مبتنی بر مثلث برای رسیدن به پوشش و اتصال بهینه است. در واقع، الگوی شبکه مثلثی ۶-اتصال را برای محدوده r_c/r_s فراهم می‌کند.



شکل (۱): چهار الگوی منظم رایج استقرار-الف) مربع، ب) شش ضلعی، ج) مثلث، د) متوازی الاضلاع

۳-۱- کارایی برخی از الگوهای استقرار منظم

برای محاسبه تعداد حسگرهای مورد نیاز جهت استقرار شبکه حسگر بیسیم از APN (Area Per Node) استفاده می‌شود. همانطور که می‌دانید از APN برای نشان دادن اندازه منطقه یک چند ضلعی ورونی استفاده می‌شود.

اگر $\{a_1, a_2, \dots, a_p\}$ مجموعه ای از نقاط p در یک سطح اقلیدسی s باشد. چند ضلعی ورونی $V(a_i)$ ، مجموعه ای از تمام نقاط در s است که نسبت به هر نقطه دیگری به a_i (از لحاظ فاصله اقلیدسی) نزدیکتر است، به عبارتی دیگر:

$$V(a_i) := \{x \in s : \forall j \in [1, p], d(x, a_i) \leq d(x, a_j)\} \quad (6)$$

در یک توپولوژی منظم که از الگوهای همگن تشکیل شده است، مساحت APN که توسط γ مشخص می‌شود، را می‌توان به صورت زیر محاسبه کرد:

$$\gamma = \frac{A_p}{N_p} \cdot N_n \quad (7)$$

در رابطه اخیر A_p مساحت الگو N_p تعداد گره‌هایی که الگو را ایجاد می‌کنند، می‌باشد و همچنین N_n نشان دهنده تعداد بلوک‌های الگو است که یک گره را به اشتراک می‌گذارند. به عنوان مثال در یک الگوی شش

گره‌ها باید حداقل باشد و همچنین بتواند حداکثر پوشش را بدست آورد.

۲-۳- محدوده سنجش و ارتباطات

هر شبکه حسگر دارای یک شعاع ارتباطی (r_c) و یک شعاع سنجش (r_s) است. نسبت r_c/r_s باید برای استقرار در نظر گرفته شود. اگر $r_c > r_s$ باشد، مسئله پوشش و اگر $r_c < r_s$ مسئله اتصال وجود دارد، بنابراین هر دو پوشش و اتصال باید تضمین شود. نسبت r_c/r_s در بیشتر روش‌های مبتنی بر الگو در نظر گرفته شده است.

۲-۴- مدل سنجش گره

به طور کلی مدل سنجش حسگرها، یک مدل ریاضی است که احتمال تشخیص هدف را با فرض اینکه هدف یا رویداد در یک نقطه P_j در منطقه مورد نظر رخ می‌دهد، توصیف می‌کند. احتمال تشخیص هدف یا رویداد توسط حسگر S_i با P_{ij} تعریف شده که تابعی از پارامترهای مختلف است. فاصله اقلیدسی بین آنها d_{ij} ، جهت گیری حسگر، پارامترهای مختلف محیطی، پارامترهای سخت افزار حسگر و غیره از جمله پارامترهای موثر در مدل سنجش گره می‌باشند.

۲-۴-۱- مدل سنجش باینری

در مدل باینری فرض می‌شود که یک حسگر دارای یک محدوده سنجش ثابت r_s است. اگر یک رویداد در نقطه P_j ، در فاصله کمتر یا مساوی r_s از محل حسگر S_i رخ دهد، این رویداد توسط S_i شناسایی شده است.

$$P_{ij} = \begin{cases} 1 & d_{ij} \leq r_s \\ 0 & \text{در غیر این صورت} \end{cases} \quad (3)$$

اگر چه این مدل با توجه به سادگی آن به طور گسترده ای در مقالات استفاده شده، اما غیر واقعی است.

۲-۴-۲- مدل سنجش احتمالاتی

با الهام از مدل‌های حسگر احتمالاتی مطرح شده در زمینه ناوبری ربات، مدل حسگر احتمالاتی به صورت زیر است:

$$P_{ij} = \begin{cases} 1 & r_s - r_e \geq d_{ij} \\ e^{-\alpha d_{ij}^\beta} & r_s + r_e \geq d_{ij} \geq r_s - r_e \\ 0 & r_s + r_e \leq d_{ij} \end{cases} \quad (4)$$

$$a = d_{ij} - (r_s - r_e) \quad (5)$$

در رابطه (۴) α و β پارامترهای حسگر با مقادیر متغیر بین ۰ و ۱، با توجه به ویژگی‌های فیزیکی حسگر است. در رابطه (۵) r_e مقیاس عدم قطعیت در تشخیص حسگر است [1].

۳- الگوهای منظم

در این بخش، چهار الگو از الگوهای منظم استقرار شش ضلعی، مربعی، متوازی الاضلاع و مثلث متساوی الاضلاع (که در شکل (۱) نشان داده شده است) را در نظر می‌گیریم [5]. این الگوها اغلب برای آسان بودن استقرار یا برای دستیابی به درجه بالایی از قابلیت اتصال در عمل

که حداکثر مقادیر APN برای چهار الگوی منظم، y_{max}^R ، y_{max}^S ، y_{max}^H و y_{max}^T در روابط (۸)، (۹)، (۱۰) و (۱۱) بیان شده است [5].

۴- الگوریتم بهینه سازی ژنتیک

الگوریتم ژنتیک روشی برای بهینه سازی مسائل محدود و نامحدود می باشد و بر اساس فلسفه انتخاب اصلح در طبیعت بنا شده است. الگوریتم ژنتیک به صورت تکرار، جمعیتی از راه حل‌ها را اصلاح می کند. در هر مرحله، الگوریتم ژنتیک تعدادی از افراد را به صورت تصادفی از جمعیت کنونی انتخاب می کند تا والدینی برای فرزندان مرحله بعد باشند. در نسل‌های متوالی، جمعیت به سوی راه حل بهینه پیش می رود. از الگوریتم ژنتیک برای حل بسیاری از مسائل بهینه سازی که با الگوریتم‌های استاندارد هم خوانی ندارد، استفاده می شود. الگوریتم ژنتیک برای تولید نسل بعدی از روی نسل کنونی از سه قانون اصلی تبعیت می کند:

- قوانین انتخاب: افرادی را که در جمعیت نسل بعدی شرکت میکنند، انتخاب می کند.
- قوانین تقاطع: برای تشکیل فرزندان نسل بعد، دو والد را با هم ترکیب می کنند.
- قوانین جهش: برای تشکیل فرزندان، تغییرات تصادفی در والدین به وجود می آورند.

جعبه ابزار الگوریتم ژنتیک کمینه تابع برازندگی را محاسبه می کند. الگوریتم با تولید یک جمعیت اولیه تصادفی شروع می شود. سپس الگوریتم، رشته ای از جمعیت‌های جدید را به وجود می آورد. در هر مرحله الگوریتم از افراد نسل کنونی استفاده می کند تا جمعیت بعدی را بسازد. برای تولید جمعیت جدید، الگوریتم مراحل زیر را انجام می دهد.

- با محاسبه مقدار برازندگی، به هر عضو جمعیت کنونی نمره می دهد.
- نمرات خام تابع را تغییر مقیاس می دهد تا در محدوده ای قابل استفاده تر قرار بگیرد.
- والدین را بر اساس مقدار برازندگی آنها انتخاب می کند.
- اعدادی از افراد جمعیت کنونی که برازندگی کمتری دارند، به عنوان نخبه انتخاب و مستقیماً به نسل بعد منتقل می شوند.
- از روی والدین فرزندان را تولید می کند. فرزندان یا به صورت جهش یا به صورت تقاطع تولید می شوند.
- جمعیت کنونی را با فرزندان جایگزین می کند تا نسل بعد را تولید کند.

و هنگامی که یکی از معیارهای توقف به وقوع بپیوندد الگوریتم متوقف می شود. فلوجارت الگوریتم ژنتیک در شکل (۲) نشان داده شده است. یکی از مهمترین مزایای الگوریتم ژنتیک به عنوان یک ابزار بهینه سازی، توانایی آن برای مقابله با مسائل بهینه سازی چند هدفه و ترکیبی است. این ویژگی الگوریتم ژنتیک، استفاده از آنها در تدوین الگوریتم‌های استقرار برای شبکه‌های حسگر بیسیم چند هدفه را برانگیخت [15].

ضلعی، $N_n = 3$ و $N_p = 6$ ، در الگو شبکه مثلثی، $N_n = 6$ و $N_p = 3$ و در یک الگوی مربعی، $N_n = 4$ و $N_p = 4$ است.

از آنجا که APN میانگین سهم هر گره برای پوشش در یک توپولوژی مبتنی بر الگوی منظم را نشان می دهد، از آن برای تعیین کارایی یک الگو استفاده می شود. برای هر الگو، هدف حداکثر کردن APN است به طوری که هر نقطه در منطقه استقرار با حفظ اتصال، پوشش داده شود. بر اساس منابع موجود، حداکثر APN برای توپولوژی منظم مبتنی بر شش ضلعی، مربعی و شبکه مثلثی، به واسطه y_{max}^T ، y_{max}^S ، y_{max}^H به ترتیب توسط روابط زیر بیان می شود:

$$y_{max}^H = \frac{3}{4} \sqrt{3} (\min\{r_c, r_s\})^2 \quad (۸)$$

$$y_{max}^S = 2 \left(\min\left\{\frac{\sqrt{2}}{2} r_c, r_s\right\} \right)^2 \quad (۹)$$

$$y_{max}^T = \frac{3}{2} \sqrt{3} \left(\min\left\{\frac{\sqrt{3}}{3} r_c, r_s\right\} \right)^2 \quad (۱۰)$$

می دانیم که متوازی الاضلاع از دو مثلث تشکیل شده است. همچنین زمانی که زاویه تند متوازی الاضلاع $\frac{\pi}{2}$ باشد، یک مربع بدست می آید. بنابراین توپولوژی شبکه مربع حالت خاصی از توپولوژی شبکه متوازی الاضلاع است. به طور مشابه زمانی که زاویه تند متوازی الاضلاع برابر $\frac{\pi}{3}$ باشد دو مثلث متساوی الاضلاع بدست می آید. در این مورد توپولوژی مبتنی بر متوازی الاضلاع تبدیل به توپولوژی شبکه مثلثی می شود. بنابراین متوازی الاضلاع را می توان به دو مثلث با زاویه حاده تقسیم کرد و برای حداکثر کردن مساحت متوازی الاضلاع، ماکزیمم کردن مساحت هر یک از مثلث‌ها کافی است.

در توپولوژی مبتنی بر متوازی الاضلاع با زاویه تند θ هر دو پوشش و اتصال را با شعاع ارتباطی r_c و شعاع حسی r_s فراهم می کند، تعریف شده است.

$$\gamma^R(r_c, r_s, \theta) = \alpha^2(r_c, r_s, \theta) \sin\theta \quad (۱۱)$$

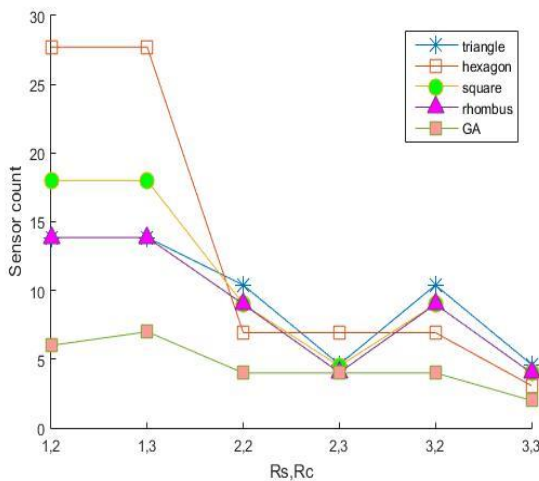
که:

$$\alpha(r_c, r_s, \theta) = \min\left\{r_c, 2r_s \frac{\cos\theta}{2}\right\} \quad (۱۲)$$

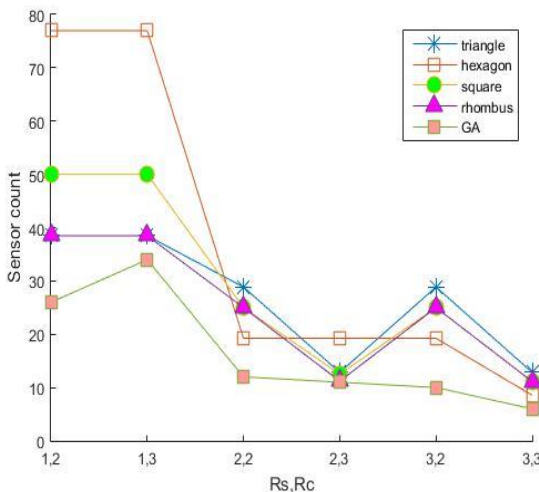
حداکثر مقدار γ_{max}^R ، $\gamma^R(r_c, r_s, \theta)$ در $\theta = \frac{\pi}{2}$ زمانی که $\frac{r_c}{r_s} \leq \sqrt{2}$ و در $\theta = \frac{\pi}{3}$ زمانی که $\sqrt{2} \leq \frac{r_c}{r_s} \leq \sqrt{3}$ و در $\theta = \frac{\pi}{3}$ زمانی که $\frac{r_c}{r_s} \geq \sqrt{3}$ رخ می دهد.

با توجه به مراجع مشاهده شده است که در شبکه ای از حسگرهای همگن با شعاع حسی r_s و شعاع ارتباطی r_c مستقر شده بر روی ناحیه مربع شکل که برای دستیابی به پوشش کامل و اتصال مورد نیاز است، حداکثر APN برای الگوی استقرار منظم شش ضلعی، مربعی، متوازی الاضلاع و مثلثی توسط رابطه زیر بدست می آید:

$$\gamma_{max} = \begin{cases} y_{max}^H & 0 \leq \frac{r_c}{r_s} \leq \frac{1}{2} 3^{3/4} \\ y_{max}^S & \frac{1}{2} 3^{3/4} \leq \frac{r_c}{r_s} \leq \sqrt{2} \\ y_{max}^R & \sqrt{2} \leq \frac{r_c}{r_s} \leq \sqrt{3} \\ y_{max}^T & \sqrt{3} \leq \frac{r_c}{r_s} \end{cases} \quad (۱۳)$$



شکل (۳): تعداد حسگرهای مورد نیاز در شبکه حسگر مربعی $6 * 6$

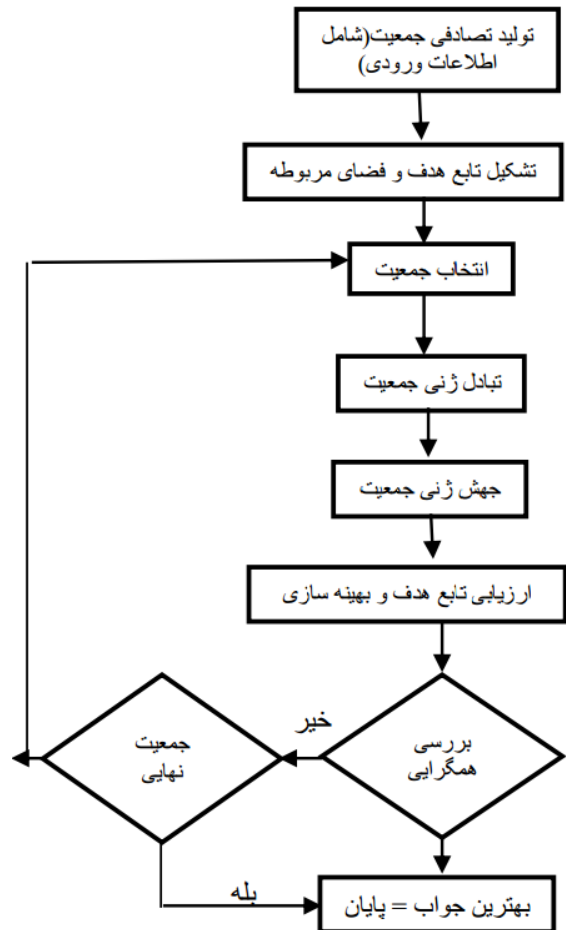


شکل (۴): تعداد حسگرهای مورد نیاز در شبکه حسگر مربعی $10 * 10$

۶- نتیجه گیری

پوشش در شبکه‌های حسگر بیسیم بستگی به جنبه‌هایی مانند هندسه شبکه، مدل سنجش حسگر و بطور قابل توجهی استقرار دارد. استقرار مناسب گره می‌تواند منجر به کاهش پیچیدگی مشکلات، کاهش مصرف انرژی، کاهش هزینه و موجب گسترش طول عمر شبکه حسگر بیسیم شود. در شبکه‌های حسگر بیسیم در نظر گرفتن هر دو پوشش و اتصال الزامی است. متصل نبودن شبکه باعث می‌شود تا اطلاعات به طور صحیح به ایستگاه پایه نرسد و موجب افت کارایی در شبکه می‌شود. تعیین الگوی استقرار بهینه برای هر دو پوشش و k -اتصال برای مقادیر کلی r_c/r_s و یک مسئله بسیار پیچیده است. یکی از روش‌های رایج در استقرار شبکه‌های حسگر بیسیم، استفاده از الگوهای منظم می‌باشد. در این مقاله استقرار شبکه حسگر بیسیم از دو روش الگوهای منظم و الگوریتم فراابتکاری ژنتیک مقایسه شده است. در این پژوهش ابتدا

از آنجایی که این الگوریتم تنها می‌تواند در مد متمرکز مستقر شود، نیاز به یک گره سینک یا ایستگاه پایه قدرتمند دارند. تابع هدف الگوریتم ژنتیک ابتدا پوشش را در شبکه حسگر بیسیم مربعی بررسی می‌کند به طوری که هر نقطه در شبکه توسط حداقل یک حسگر پوشش داده شود. پس از اینکه پوشش کامل در شبکه برقرار بود، تابع هدف اتصال را در شبکه بررسی می‌کند. مختصات $(2,2)$ به عنوان مکان گره سینک فرض شده است. برای بررسی اتصال در شبکه حسگر بیسیم مربعی، هر نقطه در شبکه باید به نقطه سینک متصل باشد.



شکل (۲): فلوچارت الگوریتم ژنتیک

۵- شبیه سازی

در این پژوهش تعداد حسگرهای مورد نیاز در یک شبکه حسگر بیسیم برای استقرار با حفظ اتصال و پوشش کامل، توسط الگوریتم فراابتکاری ژنتیک و الگوهای استقرار منظم، شبیه سازی شده است. شبکه حسگر بیسیم به عنوان یک شبکه مربعی با اندازه $M * N$ تعریف شده است. شکل (۳)، تعداد حسگرهای مورد نیاز در یک شبکه مربعی $6 * 6$ با $1 \leq r_s \leq 3$ و $2 \leq r_c \leq 3$ را نشان می‌دهد. همچنین شکل (۴)، تعداد حسگرهای مورد نیاز در یک شبکه مربعی $10 * 10$ با همان رنج حسی و ارتباطی را نشان می‌دهد.

همواره تعداد گره‌های مورد نیاز توسط الگوریتم فراابتکاری ژنتیک کمتر مساوی تعداد گره‌های مورد نیاز توسط الگوهای منظم برای دستیابی به پوشش و اتصال، است.

- [7] Akewar, M. C., & Thakur, N. V. (2012). A study of wireless mobile sensor network deployment. *International Journal of Computer and Wireless Communication*, 2(4).
- [8] Rani, K. S. S., & Devarajan, N. (2012). Optimization model for sensor node deployment. *European Journal of Scientific Research*, 70(4), 491-498.
- [9] Senouci, M. R., Mellouk, A., & Aissani, A. (2014). Random deployment of wireless sensor networks: a survey and approach. *International Journal of Ad Hoc and Ubiquitous Computing*, 15(1-3), 133-146
- [10] Liu, X. (2015). A deployment strategy for multiple types of requirements in wireless sensor networks. *IEEE transactions on cybernetics*, 45(10), 2364-2376.
- [11] Rebai, M., Snoussi, H., Hnaïen, F., & Khoukhi, L. (2015). Sensor deployment optimization methods to achieve both coverage and connectivity in wireless sensor networks. *Computers & Operations Research*, 59, 11-21.
- [12] Gupta, N., Kumar, N., & Jain, S. (2016, October). Coverage problem in wireless sensor networks: A survey. In *Signal Processing, Communication, Power and Embedded System (SCOPEs)*, 2016 International Conference on (pp. 1742-1749). IEEE.
- [13] Sharma, K. P., & Sharma, T. P. A Comparative Study of Deployment Strategies in Wireless Sensor Networks.
- [14] Yoon, Y., & Kim, Y. H. (2013). An efficient genetic algorithm for maximum coverage deployment in wireless sensor networks. *IEEE Transactions on Cybernetics*, 43(5), 1473-1483.
- [15] Sharma, K. P., & Sharma, T. P. A Comparative Study of Deployment Strategies in Wireless Sensor Networks.
- [16] Yoon, Y., & Kim, Y. H. (2013). An efficient genetic algorithm for maximum coverage deployment in wireless sensor networks. *IEEE Transactions on Cybernetics*, 43(5), 1473-1483.

پوشش در شبکه حسگر بیسیم بررسی می‌شود و اگر نقاط مورد نظر در شبکه پوشش داده شده بودند، اتصال به نقطه ایستگاه پایه بررسی و تعداد گره‌های مورد نیاز برای استقرار شبکه مشخص می‌شود. تعداد گره‌های مورد نیاز جهت استقرار در دو شبکه حسگر مربعی $6 * 6$ و $10 * 10$ و مقادیر یکسان r_s و r_c شبیه سازی شده است. برای شبکه‌هایی با اندازه بزرگتر، زمان اجرای برنامه طولانی می‌شود. مشاهده شده است که تعداد گره‌های مورد نیاز توسط الگوریتم فراابتکاری ژنتیک کمتر مساوی تعداد گره‌های مورد نیاز توسط الگوهای منظم برای دستیابی به پوشش و اتصال، است. به عبارتی دیگر الگوریتم فراابتکاری ژنتیک کارایی بهتری نسبت به الگوهای منظم دارد. همچنین هیچ الگوی منظمی برای تمام مقادیر r_s/r_c بهینه نیست. برای فواصل $0.5 \leq r_s/r_c < 1$ الگوی شبکه متوازی الاضلاع و مثلثی به تعداد گره کمتری نسبت به الگوهای دیگر جهت دستیابی به پوشش و اتصال، نیاز دارد. در $r_s/r_c = 1$ ، به عبارتی در مختصات ایستگاه پایه (سینک)، الگوی شش ضلعی بهینه است. برای $\frac{2}{3} \leq r_s/r_c < \frac{3}{2}$ ، الگوی مبتنی بر متوازی الاضلاع نسبت به سایر الگوهای منظم بهینه است.

مراجع

- [1] Deif, D. S., & Gadallah, Y. (2014). Classification of wireless sensor networks deployment techniques. *IEEE Communications Surveys & Tutorials*, 16(2), 834-855.
- [2] Wang, X., Xing, G., Zhang, Y., Lu, C., Pless, R., & Gill, C. (2003, November). Integrated coverage and connectivity configuration in wireless sensor networks. In *Proceedings of the 1st international conference on Embedded networked sensor systems* (pp. 28-39). ACM.
- [3] Zou, Y., & Chakrabarty, K. (2003, March). Sensor deployment and target localization based on virtual forces. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications*. IEEE Societies (Vol. 2, pp. 1293-1303). IEEE.
- [4] Xu, K., Wang, Q., Hassanein, H., & Takahara, G. (2005, August). Optimal wireless sensor networks (WSNs) deployment: minimum cost with lifetime constraint. In *Wireless and Mobile Computing, Networking and Communications*, 2005. (WiMob'2005), IEEE International Conference on (Vol. 3, pp. 454-461). IEEE.
- [5] Bai, X., Kumar, S., Xuan, D., Yun, Z., & Lai, T. H. (2006, May). Deploying wireless sensors to achieve both ptimization. In *Intelligent and Advanced Systems*, 2007. ICIAS 2007. International Conference on (pp. 961-965).
- [6] Zhu, C., Zheng, C., Shu, L., & Han, G. (2012). A survey on coverage and connectivity issues in wireless sensor networks. *Journal of Network and Computer Applications*, 35(2), 619-632.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی تبریز

بهبود الگوریتم HAU-Miner (الگوریتم کشف مجموعه اقلام با متوسط ارزش بالا)، با هدف کاهش حافظه مصرفی

آزاده سلطانی^۱، محمود سلطانی^۲

۱ - استادیار، گروه مهندسی کامپیوتر، دانشگاه بجنورد

a.soltani@ub.ac.ir

۲ - مربی، گروه مهندسی کامپیوتر، دانشگاه مهندسی فناوری‌های نوین قوچان

m.soltani@qiet.ac.ir

چکیده

کشف مجموعه اقلام با ارزش به علت وجود کاربردهای فراوان امروزه مورد توجه بسیار زیادی قرار گرفته است. در الگوریتم‌های ارائه شده برای این مسئله، ارزش سود یک مجموعه قلم داده، مجموع ارزش اقلام موجود در آن مجموعه در نظر گرفته می‌شود. این تعریف سبب می‌شود، مجموعه‌هایی با تعداد اقلام زیاد نیز به عنوان مجموعه‌های با ارزش معرفی شوند که در دنیای واقعی ارزش چندانی ندارند. برای کاهش اثر طول مجموعه، الگوریتم‌های کشف مجموعه اقلام با متوسط ارزش بالا معرفی شده‌اند که ارزش نهایی هر مجموعه را بر طول آن تقسیم می‌کنند. الگوریتم HAU-Miner از جمله این الگوریتم‌هاست که ابتدا پایگاه داده را در ساختاری در حافظه اصلی ذخیره می‌کند و سپس با الگوریتمی بازگشتی، مجموعه‌هایی با متوسط ارزش بالا را می‌یابد. از آنجاییکه این روش وابسته به ذخیره سازی کل داده‌ها در حافظه اصلی می‌باشد، بهینه بودن این ساختار از اهمیت بالایی برخوردار است. در این تحقیق تلاش شده است، با حذف بخش‌های غیر ضروری در این ساختار و همچنین تغییراتی در الگوریتم، حافظه مصرفی آن کاهش داده شود. آزمایش‌های انجام شده بر روی مجموعه داده‌های واقعی نشان می‌دهد تغییرات اعمال شده توانسته است میزان حافظه مصرفی را کاهش دهد.

کلمات کلیدی

کشف اقلام با ارزش، متوسط ارزش مجموعه قلم داده، الگوریتم HAU-Miner، ماکزیمم ارزش تراکنش.

۱ - مقدمه

معرفی شده است [8]-[4]. در این مسئله اقلام می‌توانند بیش از یک بار در تراکنش ظاهر شوند و همچنین ارزش و سود متفاوتی داشته باشند. بنابراین، دو مفهوم سود درونی و سود بیرونی تعریف می‌شود که به ترتیب معادل سود واقعی آن قلم و تعداد تکرار آن در تراکنش هستند. هدف، یافتن مجموعه اقلامی است که سود بیشتری در کل مجموعه داده برای ما ایجاد می‌کنند. سود یک مجموعه قلم داده با جمع کردن حاصلضرب سودهای درونی و بیرونی تک تک اقلام موجود در آن به دست می‌آید. همین تعریف ساده، نقطه ضعف این روش‌ها محسوب می‌شود؛ زیرا هر چه مجموعه، طولانی‌تر باشد ارزش آن بیشتر می‌گردد. این نکته سبب می‌شود، این روش‌ها، مجموعه‌های طولانی را به عنوان مجموعه‌های با ارزش بالا تولید

کشف مجموعه اقلام پرتکرار (FIM^۱) از جمله وظایف مهم داده-کاوی است که اولین بار در سال ۱۹۹۴ توسط آگراوال ارائه شد [۱]. هدف این مسئله یافتن اقلامی^۲ است که مکرراً، به طور همزمان در تراکنش‌ها ظاهر شده باشند [۲]. [۳]. محدودیت اصلی FIM این است که فرض می‌کند هر قلم تنها یکبار در هر تراکنش ظاهر می‌شود و همچنین تمام اقلام دارای ارزش یکسانی هستند. در نتیجه، این الگوریتم‌ها، امکان یافتن مجموعه اقلام مفید در بعضی پایگاه داده‌ها (مثل سبدهای خرید فروشگاه که سود اقلام متفاوت است) را ندارند. به منظور مقابله با این مشکل، مسئله کشف مجموعه اقلام با ارزش

داده باشد. یک تراکنش دلخواه را با $T_q = \{i_1, i_2, \dots, i_k\}$ نشان می‌دهیم ($T_q \in D$). هر تراکنش دارای یک شناسه‌ی یکتا (q) می‌باشد که TID نامیده می‌شود. همچنین مجموعه‌ای از k قلم را، یک مجموعه قلم k تایی می‌نامیم ($X = \{i_1, i_2, \dots, i_k\}$). اگر مجموعه قلم X زیرمجموعه‌ای از یک تراکنش دلخواه باشد ($X \subseteq T_q$) آنگاه می‌گوییم T_q شامل X است و یا X در T_q ظاهر شده است. برای هر قلم i_j ، دو مقدار ارزش بیرونی $pr(i_j)$ (مثال: سود واقعی یک قلم در فروشگاه) و ارزش درونی $q(i_j, T_q)$ (مثال: تعداد قلم خریداری شده در سبد خرید) تعریف شده است. تعریف‌هایی که در زیر آورده شده است در ادامه‌ی مقاله مورد نیاز است.

تعریف ۱: متوسط ارزش قلم i_j

متوسط ارزش قلم i_j در تراکنش T_q به صورت $au(i_j, T_q)$ نمایش داده می‌شود و برابر است با حاصلضرب ارزش بیرونی و ارزش درونی این قلم داده در تراکنش T_q که با فرمول ۱ محاسبه می‌شود.

$$au(i_j, T_q) = q(i_j, T_q) \times pr(i_j) \quad (1)$$

تعریف ۲: متوسط ارزش مجموعه قلم k تایی در تراکنش T_q

متوسط ارزش مجموعه قلم k تایی X در تراکنش T_q ، با $au(X, T_q)$ نمایش داده می‌شود و به کمک فرمول ۲ محاسبه می‌گردد.

$$au(X, T_q) = \frac{\sum_{i_j \in X \wedge X \subseteq T_q} q(i_j, T_q) \times pr(i_j)}{k} \quad (2)$$

در این فرمول، k تعداد اقلام موجود در X است (به عبارت دیگر $|X| = k$).

تعریف ۳: متوسط ارزش مجموعه قلم k تایی در پایگاه داده D

متوسط ارزش مجموعه قلم k تایی X در پایگاه داده D ، برابر است با مجموع متوسط ارزش X در تمام تراکنش‌های موجود در D که به صورت $au(X)$ نشان داده می‌شود و به کمک فرمول ۳ محاسبه می‌گردد.

$$au(X) = \sum_{X \subseteq T_q \wedge T_q \in D} au(X, T_q) \quad (3)$$

تعریف ۴: کشف اقلام با متوسط ارزش بالا

در مسئله کشف اقلام با متوسط ارزش بالا، به دنبال یافتن تمام مجموعه اقلامی ($HAUIs$) هستیم که متوسط ارزش آنها در کل پایگاه داده D ، از مینیمم از پیش تعریف شده ($minUtil$) کمتر نباشد.

$$HAUIs = \{X \mid au(X) \geq minUtil\} \quad (4)$$

۲-۲ - الگوریتم $HAUI-Miner$

الگوریتم $HAUI-Miner$ از یک ساختار به نام ساختار لیست متوسط ارزش^۲ برای نگهداری پایگاه داده استفاده می‌کند. همچنین برای کاهش فضای جستجو از ماکزیمم ارزش هر تراکنش استفاده می‌کند. این ویژگی (ماکزیمم ارزش) دارای خاصیت بستار رو به پایین است. در مسئله کشف مجموعه اقلام پرتکرار، مقدار پشتیبان^۴ مجموعه اقلام دارای خاصیت بستار رو به پایین بود و از همین خاصیت برای کاهش فضای جستجو استفاده می‌شد. به عبارت دیگر اگر یک مجموعه قلم، دارای تعداد تکرار کمتر از مینیمم باشد تمام مجموعه‌هایی که از روی آن ساخته می‌شود نیز حتما دارای تعداد تکرار کمتر از مینیمم خواهند بود و می‌توان از گسترش دادن آن مجموعه قلم در پروسه کشف صرف

کنند که معمولا در دنیای واقعی ارزش چندانی ندارند. برای پوشش این نقطه ضعف روش‌های مختلفی ارائه شده است که اکثر آنها با تقسیم کردن ارزش مجموعه قلم بر طول آن، سعی می‌کنند تاثیر مستقیم تعداد اقلام یک مجموعه، در سود نهایی آن را از بین ببرند [9]-[14]. یک راه بدیهی این است که مشابه قبل مجموعه‌های با ارزش تولید شوند و سپس آنهایی که ارزش متوسط بالایی دارند به عنوان خروجی جدا شوند. این روش بدیهی، زمان‌بر است؛ این در حالیست که در بسیاری از مواقع معیار متوسط ارزش می‌تواند فضای جستجو را کاهش داده و زمان الگوریتم را کمتر نماید. بنابراین در روش‌های ارائه شده همواره هدف این بوده است که متوسط ارزش جایگزین مجموع ارزش شود و در پروسه‌ی کشف الگوهای با ارزش نیز، به کاهش فضای جستجو کمک کند. الگوریتم‌های اولیه‌ی ارائه شده در این زمینه بر اساس *apriori* کار میکردند از جمله [9]TPAU. این الگوریتم‌ها نیاز به چندین بار اسکن پایگاه داده دارند و همچنین تعداد زیادی از مجموعه‌های کاندید را تولید می‌کنند که در گذر بعدی نیاز به شمارش خواهند داشت. به منظور رفع این نقطه ضعف‌ها، روش مبتنی بر درخت ارائه شده است [۱۳] که پایگاه داده را در قالب درختی ذخیره کرده و سپس با پیمایش این درخت به طور بازگشتی و با اعمال تکنیک‌های خاص، برای هرس کردن فضای جستجو، مجموعه‌های با ارزش متوسط را می‌یابد. این شیوه نیز نیازمند تولید درخت‌های شرطی فراوانی است که زمان بر است [۱۴]. دسته دیگری از الگوریتم‌ها سعی کرده اند مجموعه داده را در لیست‌هایی ذخیره کرده و سپس با پردازش این لیست‌ها مجموعه‌های با ارزش بالا را بیابند [۱۰]، [۱۲]. نکته قابل توجه در این دسته روش‌ها، حافظه مصرفی است؛ زیرا قرار است که مجموعه داده را در حافظه‌ی اصلی ذخیره کنند. بنابراین باید تا جای ممکن این فضا را کاهش داد.

روش $HAUI-Miner$ یکی از جدیدترین روش‌هایی است که برای کشف اقلام با متوسط ارزش بالا ارائه شده است که مبتنی بر لیست است. هدف از این تحقیق بهبود این الگوریتم از نظر حافظه مصرفی می‌باشد. در روش بهبود یافته، سعی شده است، هم در ساختار لیست-ها تغییراتی اعمال شود و هم با اصلاحاتی، تعداد مراحل کاهش داده شود.

ساختار ادامه مقاله به صورت زیر است: در بخش دوم به بیان دقیق مسئله و معرفی الگوریتم $HAUI-Miner$ پرداخته می‌شود. فصل سوم، روش پیشنهادی را معرفی کرده و آزمایش‌های انجام شده در بخش ۴ تحلیل می‌گردد و در نهایت فصل ۵ شامل نتیجه گیری است.

۲- بیان مسئله و معرفی ادبیات لازم

در این بخش ابتدا به بیان دقیق مسئله می‌پردازیم سپس روش $HAUI-Miner$ را به طور کامل توضیح خواهیم داد.

۲-۱- بیان دقیق مسئله

فرض کنید $D = \{T_1, T_2, \dots, T_n\}$ مجموعه‌ی تراکنش‌ها و $I = \{i_1, i_2, \dots, i_m\}$ مجموعه‌ی کل اقلام مجزای موجود، در مجموعه

برای ساخت لیست‌های AU مربوط به مجموعه اقلام k تایی ($k \geq 2$)، نیاز به اسکن پایگاه داده نیست و می‌توان این لیست‌ها را به کمک اشتراک گرفتن از روی لیست‌های کوچکتر ساخت. برای این منظور کافی است بین tid های دو مجموعه اشتراک گرفته شود و به ازای هر tid جدید یک المان ایجاد کرد. سپس ارزش و tmu را برای این المان مقداردهی نمود. اگر $k = 2$ ، یعنی دو قلم داده قرار است در کنار هم مجموعه‌ی جدیدی را ایجاد کنند و پیشوند مشترکی ندارند. بنابراین ارزش مجموعه‌ی جدید برابر خواهد بود با، مجموع ارزش این دو قلم در دو ساختار قبلی و از آنجاییکه tmu برای هر دو، یک مقدار است (هر دو مربوط به یک تراکنش هستند) با یکی از tmu ها مقدار می‌گیرد. در صورتیکه اگر $k > 2$ باشد دو مجموعه دارای پیشوند مشترکی هستند که اگر مثل قبل ارزش‌های دو لیست را با هم جمع کنیم ارزش پیشوند دو بار محاسبه می‌شود. بنابراین، همه مراحل مانند قبل است با این تفاوت که ارزش پیشوند مشترک را از حاصل جمع ارزش‌ها کم می‌کنیم.

الگوریتم HAU-Miner، یک الگوریتم بازگشتی است که با لیست‌های اولیه که بر اساس پایگاه داده تصویر شده ساخته شده‌اند شروع می‌کند و در هر مرحله با ترکیب لیست‌ها، لیست‌های جدیدی ساخته و در نتیجه می‌تواند تمام مجموعه اقلام با ارزش را تولید کند. در این مراحل برای کاهش فضای جستجو از tmu کم می‌گیرد. به عبارت دیگر اگر مجموع tmu ها در یک لیست کمتر از حداقل ارزش باشد آنگاه تمام گسترش‌هایی که از روی آن مجموعه ساخته خواهند شد نیز کم ارزش خواهند بود. بنابراین می‌توان از گسترش آن مجموعه صرف نظر کرد و در نتیجه فضای جستجو کاهش یافته و زمان کمتری صرف خواهد شد.

۳- روش پیشنهادی

الگوریتم HAU-Miner پایگاه داده را در قالب ساختاری به نام لیست‌های AU ذخیره می‌کند و سپس با اعمال یک پروسه‌ی بازگشتی و به کمک ترکیب لیست‌ها مجموعه اقلام بزرگتر را تولید کرده و ارزش آنها را تنها به کمک همین لیست‌ها به دست می‌آورد. حجم این ساختار دارای اهمیت بالایی خواهد بود؛ زیرا قرار است در حافظه اصلی جای گیرد و در نتیجه محدود کننده حجم پایگاه داده‌ای می‌شود که قرار است مورد پردازش قرار گیرد. در این تحقیق به دنبال کاهش حافظه‌ی مصرفی در الگوریتم HAU-Miner هستیم. برای این منظور دو راهکار ارائه نموده‌ایم. اول، بهبود ساختار لیست‌ها و دوم، کاهش مراحل بازگشت که منجر به نگهداری ساختارهای کمتر در حافظه خواهد شد.

همانطور که در بخش ۲-۲ گفته شد در ساختار AU مربوط به مجموعه قلم X ، لیستی از المان‌ها ذخیره می‌شود که در آن، به ازای هر تراکنشی که حاوی X است، tid آن تراکنش، ارزش X در تراکنش و tmu آن تراکنش ذخیره می‌شود. این در حالیکه پس از مشخص شدن tmu برای تراکنش‌های یک پایگاه داده تصویر شده، مقدار آنها تغییری نخواهد کرد. لیست AU برای مجموعه‌های با طول یک با همان tmu ها ساخته می‌شود و در تمام مراحل کشف، از همان tmu

نظر کرد. در مسئله کشف مجموعه اقلام با متوسط ارزش بالا، این خاصیت دیگر صدق نمی‌کند. اگر مجموعه قلمی دارای متوسط ارزش کمتر از مینیمم باشد ممکن است توسعه‌ای از آن، دارای متوسط ارزش بالاتر از مینیمم شود. به عبارت دیگر قلم داده‌ای با ارزش بالا به آن اضافه شود که ارزش کم اقلام قبلی را جبران کند. بنابراین دیگر نمی‌توان با این شیوه فضای جستجو را کاهش داد. در این دسته از الگوریتم‌ها، از ویژگی جدیدی به نام ماکزیمم ارزش تراکنش^۵ استفاده شده است. این مقدار، حداکثر مقدار ارزش در یک تراکنش می‌باشد. بدیهی است که مقدار متوسط ارزش هر زیر مجموعه‌ای از تراکنش، کمتر یا مساوی مقدار ماکزیمم ارزش تراکنش است. این ویژگی دارای خاصیت بستانار رو به پایین نیز هست؛ بنابراین می‌توان از آن برای هرس کردن فضای جستجو استفاده نمود.

تعریف ۵: ماکزیمم ارزش تراکنش

ماکزیمم ارزش تراکنش T_q ، با $tmu(T_q)$ نشان داده می‌شود و برابر است با بزرگترین ارزش اقلام در تراکنش T_q . این مقدار به کمک فرمول ۵ به دست می‌آید.

$$tmu(T_q) = \max(\{u(i_j) | i_j \in T_q\}) \quad (5)$$

تعریف ۶: حد بالای متوسط ارزش مجموعه قلم X

حد بالای متوسط ارزش مجموعه قلم X با $AUUB(X)$ نشان داده می‌شود. و برابر است با مجموع ماکزیمم ارزش تراکنش‌هایی که X در آنها ظاهر شده است و با فرمول ۶ محاسبه می‌شود.

$$AUUB(X) = \sum_{X \in T_q \wedge T_q \in D} tmu(T_q) \quad (6)$$

الگوریتم HAU-Miner در اولین اسکن پایگاه داده ماکزیمم ارزش هر تراکنش را پیدا می‌کند؛ سپس حد بالای متوسط ارزش را برای هر یک از اقلام محاسبه کرده و تمام اقلامی که دارای $AUUB$ کمتر از مینیمم ارزش باشند را از پایگاه داده D حذف می‌کند. سپس، تمام تراکنش‌ها بر اساس ترتیب $AUUB$ اقلام مرتب می‌شوند. پایگاه داده‌ی جدید، پایگاه داده‌ی اصلاح شده نامیده می‌شود و با D' نشان داده می‌شود. در مرحله بعد پایگاه داده D' به ازای هر قلم تصویر می‌شود و پایگاه داده تصویر شده به صورت یک لیست ذخیره می‌گردد.

در این مرحله، $AUUB$ اقلام در پایگاه داده‌های تصویر شده نیز مجدداً محاسبه شده و اقلامی که $AUUB$ آنها کمتر از مینیمم باشد از پایگاه داده‌های تصویر شده، حذف می‌گردند. سپس لیست‌های متوسط ارزش (لیست-AU)، از روی پایگاه داده‌های تصویر شده‌ای که دو بار اصلاح شده‌اند، ساخته می‌شوند. به ازای هر قلم داده/مجموعه X در یک پایگاه داده تصویر شده، ساختاری ساخته می‌شود که شامل X به همراه لیستی از المان‌هاست. به ازای هر تراکنش T_q که شامل X باشد یک المان متناظر وجود دارد. هر المان شامل سه فیلد است:

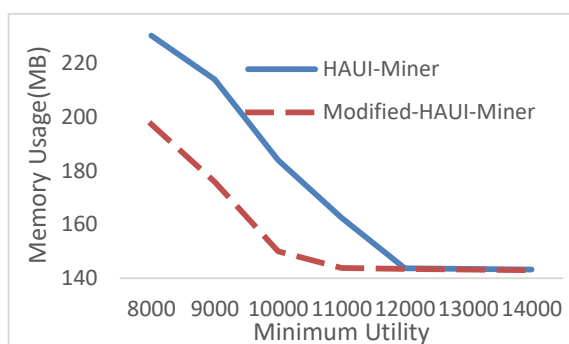
- tid : مشخصه تراکنش T_q است.
- iu : ارزش واقعی X را در تراکنش T_q نشان می‌دهد.
- tmu : حداکثر ارزش تراکنش T_q را نشان می‌دهد.

اگر مجموع $U(X)$ (حاصلضرب ارزش درونی در ارزش بیرونی) تمام المان‌های یک لیست، کمتر از حداقل نباشد آنگاه نشان می‌دهد که آن مجموعه قلم با ارزش است و به عنوان یک مجموعه قلم با ارزش مستقیماً به خروجی می‌رود.

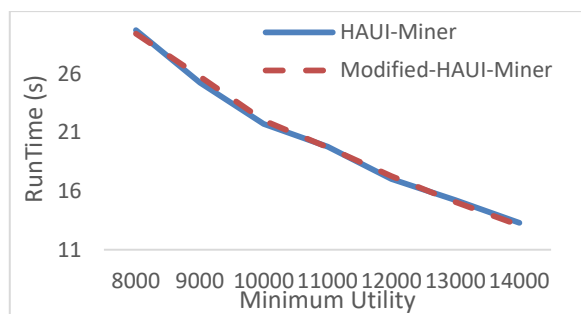
واقعی retail و kosarak انجام گرفته است. این دو مجموعه داده برای مسئله کشف اقلام باارزش آماده شده اند. به عبارت دیگر ارزش‌های درونی و بیرونی به طور تصادفی برای اقلام در نظر گرفته شده است. این مجموعه‌های آماده از سایت www.philippe-fourmier.com و viger.com/spmf بارگذاری گردیده است.

۴-۱- آزمایش اول

در این آزمایش دو الگوریتم HAU1-Miner و Modified-HAU1-Miner را بر روی مجموعه داده‌ی retail اعمال کرده و نتایج را بر اساس دو معیار زمان اجرا و حافظه مصرفی مقایسه کرده‌ایم. این مجموعه داده دارای ۸۸۱۶۲ تراکنش است. حداقل ارزش (Minimum Utility) را بین ۹۰۰۰ تا ۱۴۰۰۰ تغییر داده‌ایم. شکل-های ۱ و ۲ نتایج این آزمایش را نشان می‌دهند.



شکل ۱- میزان حافظه مصرفی اجرای دو الگوریتم HAU1-Miner و Modified-HAU1-Miner بر روی مجموعه داده retail



شکل ۲- زمان اجرای دو الگوریتم HAU1-Miner و Modified-HAU1-Miner بر روی مجموعه داده retail

با افزایش حد آستانه‌ی حداقل ارزش (Minimum Utility)، تعداد مجموعه‌های باارزش، کمتر شده و بازگشت‌ها زودتر متوقف می‌شوند؛ بنابراین انتظار داریم که زمان اجرا و حافظه مصرفی نیز کاهش یابد. شکل‌های ۱ و ۲ نیز این کاهش را نشان می‌دهند.

همانطور که در شکل ۱ مشهود است، الگوریتم بهبود یافته توانسته است میزان حافظه مصرفی را کاهش دهد این درحالیست که زمان اجرا با توجه به شکل ۲ تغییری نکرده است.

ها استفاده می‌شود. بنابراین برای کاهش حافظه مصرفی، در این تحقیق پیشنهاد می‌شود که tmu ها را به صورت تکراری در لیست‌ها ذخیره نکنیم و به ازای هر پایگاه داده تصویر شده، لیست tmu ها را فقط یک بار و جدا از ساختار لیست AU ذخیره کنیم. به عبارت دیگر میتوان آرایه‌ای به طول تعداد تراکنش‌های موجود در پایگاه داده‌ی تصویر شده ایجاد کرد، که در خانه متناظر با هر تراکنش، ماکزیمم ارزش آن تراکنش ذخیره شود. به این ترتیب هم در حافظه مصرفی کاهش داده‌ایم و هم دسترسی به tmu ها با سرعت مناسبی انجام خواهد شد. در نتیجه المان‌های ذخیره شده در لیست‌ها، فقط حاوی tid و ارزش مجموعه در تراکنش خواهند بود و tmu از آنها حذف می‌شود. از آنجاییکه تراکنش‌ها در این لیست‌ها تکرار می‌شوند، حذف tmu از آنها و ذخیره سازی آن در آرایه‌ای جدا، می‌تواند به کاهش حافظه‌ی مصرفی کمک شایانی کند؛ خصوصاً زمانی که تعداد بازگشت‌ها زیاد می‌شود.

علاوه بر تغییرات فوق که باعث می‌شود لیست‌ها، حافظه‌ی کمتری مصرف کنند؛ می‌توان با کاهش مراحل بازگشت نیز به کمتر کردن فضای مورد نیاز کمک کرد. از آنجاییکه الگوریتم HAU1-Miner یک الگوریتم بازگشتی با پیمایش عمقی است و در فراخوانی‌های مجدد، لیست‌های قبلی را لازم دارد؛ در هر مرحله، لیست‌ها باید در حافظه‌ی استک ذخیره شوند. بنابراین هر چه تعداد مراحل کمتر باشد نیاز به فضای کمتری است. در این تحقیق سعی شده است که از این منظر نیز حافظه مصرفی بهبود داده شود. همانطور که قبلاً گفته شد، در الگوریتم HAU1-Miner، ابتدا بر اساس هر تصویر از پایگاه داده، برای تمام اقلام لیست AU ساخته می‌شود و سپس با یک پروسه‌ی بازگشتی ابتدا مجموعه‌های دوتایی کاندید از روی این لیست‌ها ساخته می‌شود و سپس بر اساس هر کدام از این مجموعه‌های کاندید، مجموعه‌هایی با طول بزرگتر از دو ساخته می‌شوند. این جداسازی به این علت است که نحوه محاسبه ارزش، برای مجموعه‌های جدید دو عضوی با مجموعه‌هایی با بیش از دو عضو متفاوت است.

همانطور که مشخص است، تنها قلم داده‌ای که لیست AU آن واقعی است، تنها همان قلم داده‌ی z_i ای است که پایگاه داده بر اساس آن تصویر شده است. و برای بقیه‌ی قلم‌ها، لیست AU را در صورت وجود قلم z_i تولید کرده‌ایم. بنابراین می‌توان این مرحله را تغییر داد. به این صورت که فقط برای قلم داده‌ی z_i لیست AU را بدست آورد و برای سایر اقلام z_k لیست AU مربوط به مجموعه دوتایی $\{z_i, z_k\}$ را تولید نمود. با این کار می‌توان گفت، بدون اینکه به الگوریتم خدشه‌ای وارد شود، مرحله اول حذف می‌شود و در نتیجه دیگر نیاز به حافظه‌ای که قبلاً برای نگهداری لیست‌های AU مربوط به مجموعه‌های یک عضوی استفاده می‌شد نخواهیم داشت.

۴- ارزیابی روش پیشنهادی

در این بخش، روش پیشنهادی از نظر زمان و حافظه‌ی مصرفی با الگوریتم HAU1-Miner مقایسه می‌گردد. کد الگوریتم HAU1-Miner از پکیج [15] SPMF گرفته شده است. در این پکیج بیش از ۱۳۰ الگوریتم در زمینه‌ی کشف الگوهای پرتکرار و قوانین انجمنی به زبان جاوا پیاده‌سازی شده است. آزمایش‌ها بر روی دو مجموعه داده‌ی

استفاده شده است اول کوچک کردن لیست‌های ارزش و دوم کاهش تعداد مراحل بازگشت در الگوریتم. آزمایش‌های انجام گرفته شده به خوبی نمایانگر کاهش میزان حافظه مصرفی در الگوریتم بهبود یافته است.

مراجع

- [1] R. Agrawal and R. Srikant, *Fast Algorithms for Mining Association Rules in Large Databases*, San Francisco, CA, USA, 1994.
- [2] J. P. 46] J. Han and Y. Yin, *Mining frequent patterns without candidate generation*, vol. 29. 2000.
- [3] M. J. Zaki and K. Gouda, "Fast vertical mining using diffsets," in *Proc. 9th ACM SIGKDD*, 2003, pp. 326–335.
- [4] Y. Liu, W. Liao, and A. N. Choudhary, "A Two-Phase Algorithm for Fast Discovery of High Utility Itemsets," in *PAKDD*, 2005, vol. 3518, pp. 689–695.
- [5] H.-F. Li, H.-Y. Huang, Y.-C. Chen, Y.-J. Liu, and S.-Y. Lee, "Fast and memory efficient mining of high utility itemsets in data streams," in *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, 2008, pp. 881–886.
- [6] V. S. Tseng, B.-E. Shie, C.-W. Wu, and S. Y. Philip, "Efficient algorithms for mining high utility itemsets from transactional databases," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 8, pp. 1772–1786, 2013.
- [7] W. Song, Y. Liu, and J. Li, "BAHUI: fast and memory efficient mining of high utility itemsets based on bitmap," *Int. J. Data Warehous. Min. IJDWM*, vol. 10, no. 1, pp. 1–15, 2014.
- [8] G.-C. Lan, T.-P. Hong, and V. S. Tseng, "An efficient projection-based indexing approach for mining high utility itemsets," *Knowl. Inf. Syst.*, vol. 38, no. 1, pp. 85–107, 2014.
- [9] T.-P. Hong, C.-H. Lee, and S.-L. Wang, "Mining high average-utility itemsets," in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, 2009, pp. 2526–2530.
- [10] G.-C. Lan, T.-P. Hong, V. S. Tseng, and others, "A projection-based approach for discovering high average-utility itemsets," *J. Inf. Sci. Eng.*, vol. 28, no. 1, pp. 193–209, 2012.
- [11] G.-C. Lan, T.-P. Hong, and V. S. Tseng, "Efficiently mining high average-utility itemsets with an improved upper-bound strategy," *Int. J. Inf. Technol. Decis. Mak.*, vol. 11, no. 5, pp. 1009–1030, 2012.
- [12] T. Lu, B. Vo, H. T. Nguyen, and T.-P. Hong, "A new method for mining high average utility itemsets," in *IFIP International Conference on Computer Information Systems and Industrial Management*, 2014, pp. 33–42.
- [13] C.-W. Lin, T.-P. Hong, and W.-H. Lu, "Efficiently Mining High Average Utility Itemsets with a Tree Structure," in *Intelligent Information and Database Systems*, 2010, pp. 131–139.
- [14] J. C.-W. Lin, T. Li, P. Fournier-Viger, T.-P. Hong, J. Zhan, and M. Voznak, "An efficient algorithm to mine high average-utility itemsets," *Adv. Eng. Inform.*, vol. 30, no. 2, pp. 233–243, 2016.
- [15] P. Fournier-Viger, A. Gomariz, T. Gueniche, A. Soltani, C.-W. Wu, and V. S. Tseng, "SPMF: a java open-source pattern mining library," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 3389–3393, 2014.

زیرنویس‌ها

¹ Frequent Itemset Mining

² Items

³ Average-Utility-List Structure

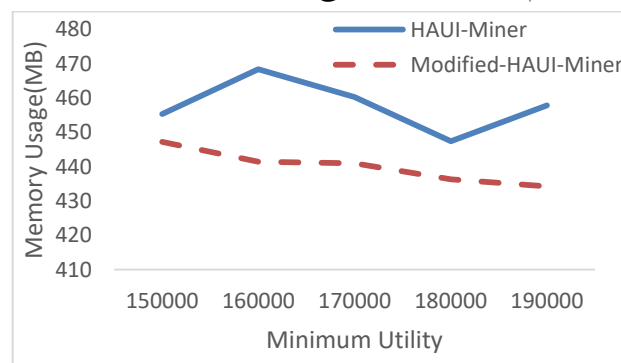
⁴ Support

⁵ Transaction-Maximum Utility

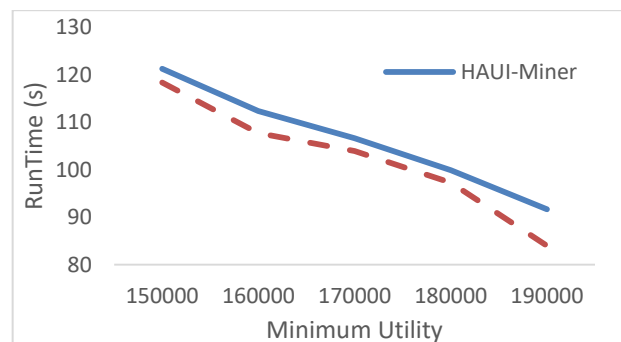
⁶ Average- Utility Upper Bound (AUUB)

۲-۴- آزمایش دوم

در این آزمایش دو الگوریتم HAU-Miner و Modified-HAU-Miner را بر روی مجموعه داده‌ی Kosarak اعمال کرده و نتایج را بر اساس دو معیار زمان اجرا و حافظه مصرفی مقایسه کرده ایم. مجموعه داده‌ی Kosarak دارای حدوداً ۹۹۰۰۰۰ تراکنش است. حداقل ارزش (Minimum Utility) را بین ۱۵۰۰۰۰ تا ۱۹۰۰۰۰ تغییر داده ایم. شکل‌های ۳ و ۴ نتایج این آزمایش را نشان می‌دهند.



شکل ۳- میزان حافظه مصرفی اجرای دو الگوریتم HAU-Miner و Modified-HAU-Miner بر روی مجموعه داده Kosarak



شکل ۴- زمان اجرای دو الگوریتم HAU-Miner و Modified-HAU-Miner بر روی مجموعه داده Kosarak

در این آزمایش نیز شکل ۳ نشان دهنده‌ی کاهش حافظه مصرفی در الگوریتم پیشنهادی است و با توجه به شکل ۴، در این آزمایش زمان نیز در روش پیشنهادی کاهش یافته است. تعداد بالاتر تراکنش‌ها در این آزمایش تفاوت‌ها را بیشتر نشان داده است.

۵- نتیجه

کشف اقلام با متوسط ارزش بالا، یک وظیفه پرکاربرد در داده کاوی است که الگوریتم‌های متنوعی برای آن ارائه شده است. HAU-Miner از کاراترین الگوریتم‌ها برای حل این مسئله است. این روش نیازمند به ذخیره‌ی کل مجموعه داده در قالب لیست‌های ارزش در حافظه اصلی است. بنابراین حافظه مصرفی، تاثیر مستقیم بر اندازه‌ی مجموعه داده‌هایی دارد که الگوریتم قابلیت پردازش آنها را خواهد داشت. لذا، در این تحقیق سعی کرده‌ایم، که الگوریتم HAU-Miner را از نظر حافظه مصرفی بهبود دهیم. برای این منظور از دو راهکار



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

ارائه یک سیستم تشخیص نفوذ سبک برای تشخیص حملات انکار سرویس در شبکه‌های حسگر بی سیم

مهدی صادقی زاده^۱، امیدرضا معروضی^۲، علی اکبر پویان^۳

^۱ دانشجوی دکتری، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود

m.sadeghizadeh@shahroodut.ac.ir

^۲ استادیار، گروه مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، شاهرود

marouzi@shahroodut.ac.ir

^۳ استادیار، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود

apouyan@shahroodut.ac.ir

چکیده

شبکه‌های حسگر بی سیم یکی از فناوری‌های کاربردی و جذاب است که در سالهای اخیر بسیار مورد توجه محققان قرار گرفته است. این شبکه‌ها به دلیل ویژگی‌هایی همچون سهولت استفاده و هزینه پایین راه‌اندازی و بکارگیری، در زمینه‌های متنوعی بکار گرفته شده‌اند. اما این شبکه‌ها دارای محدودیت‌هایی در محاسبات، حافظه و انرژی هستند. با توجه به بحرانی بودن اغلب کاربردهای این شبکه‌ها، امنیت به عنوان یکی از پارامترهای اساسی کیفیت سرویس در آنها مطرح می‌باشد. ما در این مقاله قصد داریم تا با شبیه سازی و تحلیل رفتار حمله انکار سرویس که یکی از حملات رایج در این شبکه‌ها است، تاثیرات نامطلوب آن را بر روی کارایی و عملکرد شبکه بررسی نماییم. هدف از این بررسی ارائه یک سیستم تشخیص نفوذ سبک بر روی شبکه‌های حسگر بی سیم در مقابله با حمله انکار سرویس با توجه به خصوصیات استخراج شده می باشد. نتایج حاصله حاکی از آن هستند که با بکارگیری این سیستم تشخیص نفوذ، بهبود زیادی در امنیت شبکه در برابر حمله انکار سرویس ایجاد شده است.

کلمات کلیدی

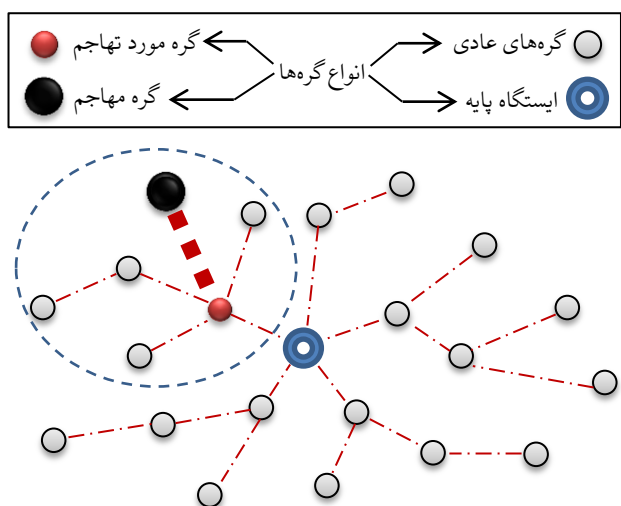
شبکه‌های حسگر بی سیم، حمله انکار سرویس، سیستم‌های تشخیص نفوذ

و متنوعی از کاربردها مانند کنترل و نظارت نظامی تا کنترل آتش سوزی جنگلها و نظارت بر ایمنی سازه‌ها و ساختمانها، بسیار مطلوب و مقرون به صرفه می‌باشند. با این وجود گره‌های موجود در این شبکه‌ها به دلیل توان پردازشی پایین، حافظه و انرژی محدودشان دارای محدودیتهای منابع شدیدی هستند. [۱]

با توجه به این که این شبکه‌ها معمولاً در مکانهای دور و فاقد حفاظت و یا اغلب در جاهایی که شرایط عملیاتی نامطلوب و یا حتی خصمانه دارند به کار گرفته می‌شوند، برای تهاجم و حملات امنیتی بسیار مستعد می‌باشند [۲]، که این امر با توجه به منابع محدود آنها

۱- مقدمه

پیشرفت‌های اخیر در زمینه الکترونیک و مخابرات بی سیم این امکان را برای ما فراهم کرده است که بتوانیم حسگرهایی را با توان مصرفی پایین، اندازه کوچک، قیمت مناسب و کاربری‌های گوناگون طراحی و ایجاد نماییم. شبکه‌های حسگر بی سیم که مرهون این پیشرفت‌ها می‌باشند در واقع از صدها و یا هزاران دستگاه کوچک تشکیل شده‌اند که دارای قابلیت‌های حسگری، پردازشی و ارتباطی برای نظارت بر محیط‌های واقعی هستند. این شبکه‌ها برای ایفای نقش در طیف وسیع



شکل ۱: حمله انکار سرویس

همانطور که در شکل ۱ مشاهده می‌گردد، گره مهاجم با ارسال سیلی از پیامها به سمت گره هدف، باعث می‌شود گره مورد تهاجم امکان سرویس دهی به گره‌های دیگر شبکه را نداشته باشد و در نتیجه کار کل خوشه مشخص شده به طور کامل مختل می‌گردد. گاهی اوقات این حمله با تهاجم به چندین گره مهم در شبکه حتی می‌تواند کار کل شبکه را نیز مختل نماید.

۲-۲- سیستم‌های تشخیص نفوذ

در یک شبکه یا یک سیستم، هر نوعی از فعالیت‌های غیر مجاز و نامطلوب، نفوذ نامیده می‌شوند. یک سیستم تشخیص نفوذ یک مجموعه از ابزار، متدها و منابع برای کمک به شناسایی، ارزیابی و گزارش نفوذها می‌باشد. تشخیص نفوذ یک واحد حفاظتی منفرد و جداگانه نیست بلکه معمولاً بخشی از یک سیستم حفاظت کلی‌تر است که در کنار یک سیستم یا دستگاه نصب می‌گردد. نفوذ به صورت: "هر مجموعه از فعالیت‌ها که تلاش می‌کند تا اصلت، محرمانگی و یا موجودیت یک منبع را به خطر بیاندازد" تعریف می‌شود [۹] و تکنیک‌های پیشگیری از نفوذ (مانند رمزنگاری، تأیید هویت، کنترل دسترسی، مسیریابی امن و غیره) به عنوان اولین خط تدافعی در برابر نفوذها ارائه می‌شوند.

با این وجود، باید توجه داشت که در هیچ سیستم امنیتی، نمی‌توان به طور کامل از نفوذها پیشگیری نمود. نفوذ و تسخیر یک گره منجر به افشای اطلاعات محرمانه مانند کلیدهای امنیتی برای نفوذگرها می‌شود. این امر منجر به شکست مکانیزم امنیتی پیشگیرانه می‌گردد. بنابراین سیستم‌های تشخیص نفوذ به جهت شناسایی و آشکار سازی نفوذها طراحی شده‌اند، قبل از این که آنها بتوانند منابع سیستم و اطلاعات امنیتی را فاش نمایند. در هر طرح امنیتی، برخی یا همه اطلاعات زیر توسط سیستم‌های تشخیص نفوذ برای حمایت و پشتیبانی سیستم‌های دیگر تامین می‌گردد:

- شناسایی مهاجم (نفوذگر)
- تعیین محدوده مهاجم (برای نمونه یک گره تک و یا یک ناحیه)

باعث کاهش شدید عملکرد و کارایی آنها می‌گردد. بنابراین تامین امنیت در شبکه‌های حسگر بی‌سیم در برابر مهاجمان و حملات مختلف به یک موضوع مهم مبدل شده است، به ویژه اگر این شبکه‌ها در فرآیندهای بحرانی نیز دخیل باشند. شبکه‌های حسگر بی‌سیم امن در کاربردهای نظامی (تاکتیکی) دارای اهمیت بحرانی و حساسی هستند، به گونه‌ای که یک شکاف امنیتی در شبکه می‌تواند باعث تحریک و تضعیف نیروهای خودی در میدان جنگ گردد. [۳]

یکی از حملاتی که در شبکه‌های حسگر بی‌سیم رایج بوده و کارایی شبکه را به شدت کاهش می‌دهد حمله انکار سرویس می‌باشد. ما در این مقاله قصد داریم با شبیه سازی حمله انکار سرویس، نحوه عملکرد و کارایی شبکه را در مواجهه با آن ارزیابی نماییم. این ارزیابی و بررسی تاثیرات شبکه حسگر بی‌سیم از حمله انکار سرویس می‌تواند ما را در طراحی و ایجاد یک مکانیزم امنیتی مناسب به جهت افزایش کارایی شبکه کمک نماید. در حقیقت هدف نهایی ما از این شبیه‌سازی، بررسی و تحلیل رفتار، طراحی یک سیستم تشخیص نفوذ سبک در شبکه‌های حسگر بی‌سیم می‌باشد به گونه‌ای که عملکرد شبکه را در برابر حمله انکار سرویس به خوبی حفظ نماید.

ادامه این مقاله به این صورت سازماندهی شده است: در بخش ۲، ابتدا حمله انکار سرویس را معرفی نموده و سپس سیستم‌های تشخیص نفوذ را تشریح می‌نماییم. در انتها نیز، مهمترین کارهای انجام شده در زمینه تشخیص نفوذ حمله انکار سرویس را به همراه مزایا و معایب آنها بررسی می‌کنیم. در بخش ۳ نیز سیستم تشخیص نفوذ پیشنهادی خود را ارائه می‌نماییم. در بخش ۴ به شبیه سازی و ارائه نتایج خواهیم پرداخت که در ابتدا به شبیه سازی حمله انکار سرویس پرداخته و عملکرد شبکه مفروض را در معرض تهاجم آن ارزیابی می‌نماییم. سپس به شبیه سازی و ارائه نتایج حاصل از سیستم تشخیص نفوذ پیشنهادی می‌پردازیم. در انتهای مقاله نیز در بخش ۵، نتیجه گیری کلی را ارائه خواهیم نمود.

۲- کارهای مرتبط

در این بخش ابتدا حمله انکار سرویس را در شبکه‌های حسگر بی‌سیم معرفی می‌کنیم و سپس سیستم‌های تشخیص نفوذ به همراه انواع آنها مورد تشریح قرار می‌گیرند. در انتها نیز مهمترین سیستم‌های تشخیص نفوذ موجود را در تشخیص حمله انکار سرویس مورد بررسی و ارزیابی قرار خواهیم داد.

۲-۱- معرفی حمله انکار سرویس

حمله انکار سرویس به عنوان یکی از رایج ترین حملات در شبکه‌های حسگر بی‌سیم در مراجع زیادی [۴-۸] مورد توجه محققان بوده است. نحوه عملکرد آن به این صورت است که گره مهاجم که معمولاً گره‌ای با قدرت و انرژی بالا است، با سرعت خیلی زیاد بسته‌های داده را در شبکه به گره‌های دیگر ارسال می‌نماید. این کار موجب می‌گردد که گره‌های مورد تهاجم، به دلیل حجم بالای پیام‌های رسیده از گره مهاجم، درگیر این امر شده و عملاً از سرویس دهی به بقیه گره‌های شبکه و انجام وظایف دیگر خود بازمانند.

دهی معمول خود را از دست بدهند. بنابراین در شکل ۲ با بررسی فاصله زمانی بین بسته‌های دریافتی (IRP) می‌توان این مهاجم را شناسایی نمود. علاوه بر این، در اغلب موارد مهاجم با توان بالایی ارسال بسته‌ها را انجام می‌دهد که از روی قدرت سیگنال دریافتی (RSP) نیز می‌توان آن را شناسایی کرد.

```
If (IRP < ThresholdIRP || RSP > ThresholdRSP)
{
  Create (alert);
  Send (alert, node-ID, malicious-ID);
  //send to cluster-head
}
```

شکل ۲: شبه کد تشخیص حمله انکار سرویس

به جهت بررسی فاصله زمانی بین بسته‌های دریافتی می‌توان از کد AWK شکل ۳ بر روی فایل trace خروجی از شبیه‌سازی انجام شده در NS2 استفاده کرد.

```
for ( node-id in Node ){
  for ( packet in Packets_received [node-id] ){
    src = packet.source;
    time = packet.time;

    prev [src] = current [src];
    current [src] = time;
    interval = current [src] - prev [src];

    if ( interval < threshold ) {
      Create (alert); // src is malicious node
      Send (alert, node-id, src);
      //send alert to cluster-head
    }
  }
}
```

شکل ۳: شبه کد AWK برای بررسی فاصله بین پیامها

هر وقت پیامی مبنی بر هشدار وجود مهاجم از گره‌های دیگر به سرخوشه ارسال می‌گردد، سرخوشه با بروز رسانی وضعیت هشدار صادر شده و مقایسه آن با حد آستانه تصمیم‌گیری نهایی را انجام می‌دهد. در صورت تجاوز هشدارها از حد آستانه، گره مفروض به عنوان مهاجم شناسایی شده و آن را در لیست مهاجمان قرار می‌دهد و از طریق ارسال پیام به سایر گره‌های موجود در خوشه، لیست آنها را نیز بروز رسانی می‌نماید. این امر در شکل ۴ ارائه شده است.

```
Receive (alert);
If ( Looking (alert, intrusion alert) )
{
  Attacker_Count [Node-ID] ++;
  If (Attacker_Count [Node-ID] > ThresholdAlarm)
  {
    Insert(Blacklist, Node-ID);
    Propagate (Blacklist);
  }
}
```

شکل ۴: شبه کد عملیات سرخوشه

- تعیین زمان نفوذ (مثلا تاریخ رخداد آن)
- نحوه فعالیت مهاجم (این که نفوذگر فعال یا غیر فعال است)
- تعیین نوع نفوذ (حملاتی مانند کرم چاله، سیاه چاله و ...)
- تعیین لایه‌ای که نفوذ در آن واقع شده است (برای مثال لایه شبکه)

۲-۳- سیستم‌های تشخیص نفوذ موجود

به جهت پیشگیری از نفوذ و حملات، پروتکل‌های مسیریابی^۱ مختلفی مانند [۱۰] SPINS، [۷] SAR، [۱۱] UDSR و [۱۲] S-LEACH برای شبکه‌های حسگر بی‌سیم ارائه شده است که همه آنها محدودیتهایی را به جهت افزایش امنیت، به شبکه اعمال می‌کنند. بنابراین با توجه به احتمال بسیار پایین انجام حمله، در شرایطی که حمله‌ای وجود نداشته باشد مصرف انرژی شبکه را افزایش خواهند داد.

در مرجع [۱۳] روشی به نام MOM برای تشخیص حملات انکار سرویس ارائه شده است که در آن حملاتی را که درون خوشه اتفاق می‌افتد شناسایی می‌نماید. روال کار آن به این صورت است که دو لیست از پیام‌های عادی و پیام‌های نامعتبر ایجاد می‌کند و برای هر پیام جدید ورودی، ابتدا آن را با لیست پیام‌های نامعتبر مقایسه کرده و عدم اعتبار آن را تشخیص می‌دهد. همچنین در صورت معتبر بودن پیام آن را با لیست پیام‌های عادی نیز مقایسه می‌کند تا تکراری بودن آن را تشخیص دهد. دومشکل اساسی این روش یکی نیازمندی به حافظه بالا به جهت نگهداری لیست پیام‌های عادی و نامعتبر و دیگری عدم تشخیص پیام‌های نامعتبر جدید و غیر تکراری می‌باشد.

در مقاله [۱۴] مدلی مبتنی بر شناسایی الگو برای حملات فرسودگی منابع ارائه شده که می‌تواند برای حملات انکار سرویس نیز بکارگرفته شود. در این مقاله از یک سری گره‌های خاص به جهت نظارت بر ترافیک شبکه و تشخیص ترافیک عادی از ترافیک غیر عادی بر اساس مدل شناسایی الگو ارائه شده، استفاده شده است.

در مرجع [۱۵] یک سیستم تشخیص نفوذ مبتنی بر خوشه بندی برای حمله انکار سرویس ارائه شده است که در آن یک مجموعه گره ویژه به نام گره‌های محافظ^۲، عملیات بررسی و تحلیل ترافیک شبکه و کشف و گزارش حملات انکار سرویس را به گره‌های سرخوشه برعهده دارند. مشکل اساسی این روش عدم توجه به پارامتر انرژی و مصرف آن است که در شبکه‌های حسگر دارای اهمیت بالایی است. در این مقاله گره‌های محافظ به صورت ثابت در کل زمان کاری شبکه در نظر گرفته شده‌اند و تنها به منظور انجام فرایند تشخیص نفوذ استفاده می‌شوند.

در مرجع [۱۶] به جهت اصلاح روش فوق گره‌های محافظ به صورت دوره‌ای و بر اساس انرژی باقیمانده انتخاب می‌شوند و علاوه بر عملیات تحلیل ترافیک و تشخیص نفوذ، اعمال معمول گره‌های شبکه حسگر را نیز انجام می‌دهند تا در نهایت یک راه حل پویا برای افزایش طول عمر شبکه ارائه شود. البته خود الگوریتم انتخاب و تعویض گره‌های محافظ دارای سربار انرژی و محاسبات خواهد بود.

۳- سیستم تشخیص نفوذ پیشنهادی

در این حمله مهاجم با توجه به سرعت بالا در ارسال بسته‌ها به سایر گره‌ها قصد دارد تا بار کاری آنها را به حدی برساند که امکان سرویس

۴- شبیه‌سازی‌ها و ارائه نتایج

در این بخش ابتدا ما به شبیه‌سازی حمله انکار سرویس می‌پردازیم که در بخش ۲ معرفی شد و در ادامه نیز به شبیه‌سازی سیستم تشخیص نفوذ پیشنهادی خود و ارائه نتایج بدست آمده از آن و مقایسه با کارهای دیگران خواهیم پرداخت.

۴-۱- شبیه‌سازی حمله انکار سرویس

ما به جهت شبیه‌سازی‌های خود از شبیه‌ساز NS2 استفاده کرده‌ایم که به عنوان یکی از معتبرترین شبیه‌سازهای شبکه می‌باشد. شبیه‌ساز NS2 به طور ساده یک ابزار شبیه‌سازی مبتنی بر رخداد است که برای مطالعه و بررسی ماهیت پویای شبکه‌های ارتباطی مطرح شده است.

در این شبیه‌سازی، پارامترهای شبکه پایه خود را با توجه به ماهیت شبکه‌های حسگر بی‌سیم و نیازمندیهای موجود و بررسی کاربردهای معمول در این شبکه‌ها تعیین کردیم. در این سناریو ما شبکه‌ای متشکل از ۷۰ گره در ۴ خوشه با محیطی به وسعت ۱۰۰ * ۱۰۰ مترمربع، ارائه کردیم که با ترافیک CBR و اندازه بسته‌های ۷۰ بایت در مدت زمان ۱۰۰ ثانیه تغذیه می‌گردد. در جدول ۱ لیست کامل پارامترها همراه با مقادیر مربوطه ارائه شده است.

جدول ۱: پارامترهای شبیه‌سازی شبکه حسگر بی‌سیم

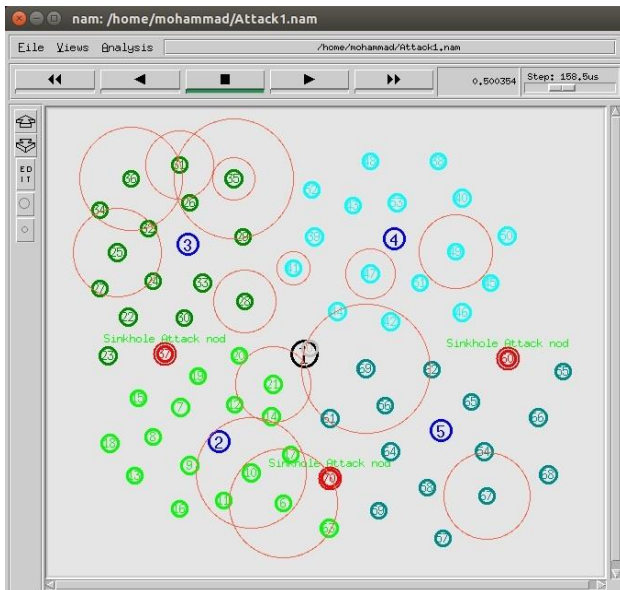
No	Parameters	Values
1	Number of nodes	70
2	Size of network	100 * 100 m ²
3	Routing protocol	AODV
4	MAC protocol	802.11
5	Type of traffic	CBR
6	Packet size	70 byte
7	Number of Cluster	4
8	Interface Queue Type	Drop Tail
9	Queue Length	50
10	Simulation Time	100 sec
11	Type of nodes	Mica2
12	Sensing Power	0.015 w
13	Processing Power	0.024 w
14	Sleep Power	0.0001 w
15	RX Power	0.024 w
16	TX Power	0.036 w
17	Initial Energy of nodes	1 j

با توجه به این که حمله انکار سرویس با ارسال سریع پیامهای داده به یک یا چند گره مشخص در آنها ایجاد ازدحام کرده و موجب اختلال در سرویس دهی آنها در شبکه می‌گردد، بنابراین ما به جهت شبیه‌سازی آن در NS2، گره‌ای با انرژی بیشتر و محدوده ارسال بالاتر ایجاد کردیم و نرخ ارسال ترافیک آن را چندین برابر گره‌های عادی شبکه قرار دادیم. در جدول ۲ پارامترهای مربوط به شبیه‌سازی حمله انکار سرویس ارائه شده‌اند.

جدول ۲: پارامترهای شبیه‌سازی حمله انکار سرویس

No	Parameters	Values
1	Number of attacker	1 / 2 / 3
2	Antenna range	100 m
3	Initial Energy of nodes	10 j
4	transfer rate of packets	Between 0.01 to 0.1 sec
5	Attacker location	Random / manual

شکل ۵ که از نرم افزار NAM گرفته شده است، مربوط به شبیه‌سازی شبکه حسگر با پارامترهای جدول ۱ در حضور گره‌های مهاجم با پارامترهای جدول ۲ می‌باشد که در آن گره‌های آبی پر رنگ نشان‌دهنده سرخوشه‌ها، گره مشکی رنگ معرف ایستگاه پایه و گره‌های قرمز رنگ نشان دهنده مهاجمین هستند. همچنین به جهت نشان دادن گره‌های معمولی مربوط به هر خوشه از رنگ یکسان استفاده شده است. ما در شبیه‌سازی‌ها برای توزیع گره‌های شبکه حسگر در محیط، هر دو روش توزیع تصادفی و توزیع دستی را ارائه و ارزیابی کردیم که شکل ۵ نمونه‌ای از توزیع دستی است که به نسبت توزیع تصادفی دارای چگالی توزیع یکنواخت‌تری در محیط می‌باشد.



شکل ۵: نمای گرافیکی شبکه حسگر بی‌سیم با NAM.

نتایج شبیه‌سازی حملات در بخش بعد همراه با نتایج حاصل از بکارگیری سیستم تشخیص نفوذ پیشنهادی بر روی شبکه حسگر بی‌سیم ارائه شده‌اند.

۴-۲- شبیه‌سازی سیستم تشخیص نفوذ پیشنهادی

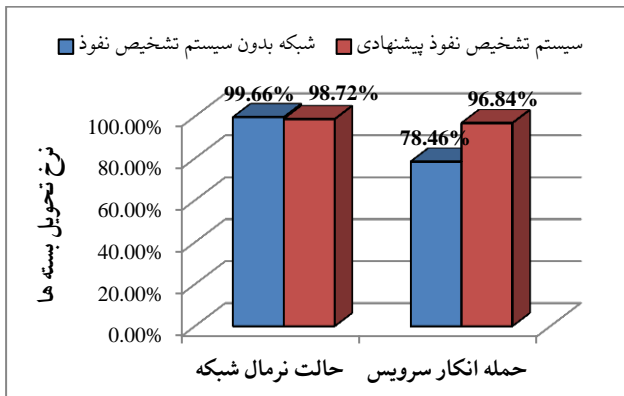
به جهت ارزیابی کارایی سیستم تشخیص نفوذ پیشنهادی، معیارهای زیر در نظر گرفته شده‌اند:

- **نرخ تشخیص:** نرخ تشخیص یا دقت تشخیص درصد حملات تشخیص داده شده را نسبت به کل حملات مشخص می‌نماید.

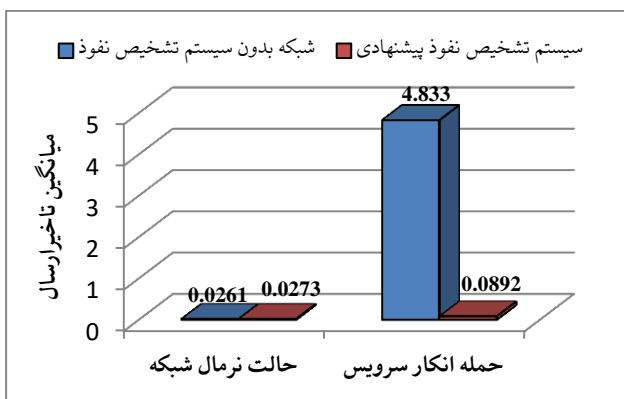
$$\text{Detection rate} = \frac{\text{No. of Detected Attacks}}{\text{No. of Attacks}} * 100\% \quad (1)$$

- **نرخ تشخیص نادرست:** این معیار میزان تشخیص نادرست را در تشخیص حملات نشان می‌دهد. به عبارت دیگر این معیار مشخص می‌کند که چه درصدی از حملات تشخیص داده شده حمله نبوده‌اند و سیستم تشخیص نفوذ اشتباهات آنها را به عنوان یک حمله تشخیص داده است.

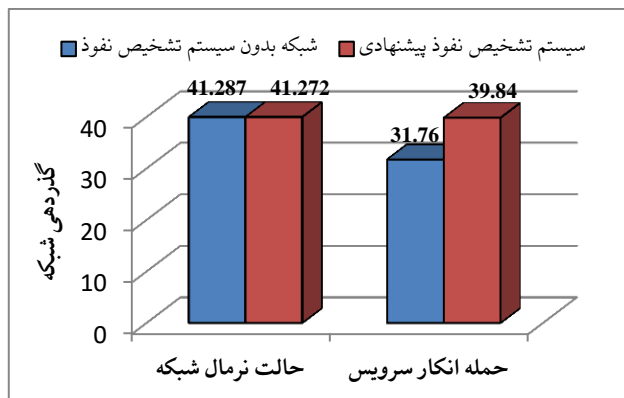
$$\text{False positive rate} = \frac{\text{No. of misdetected Attacks}}{\text{No. of Normal connections}} * 100\% \quad (2)$$



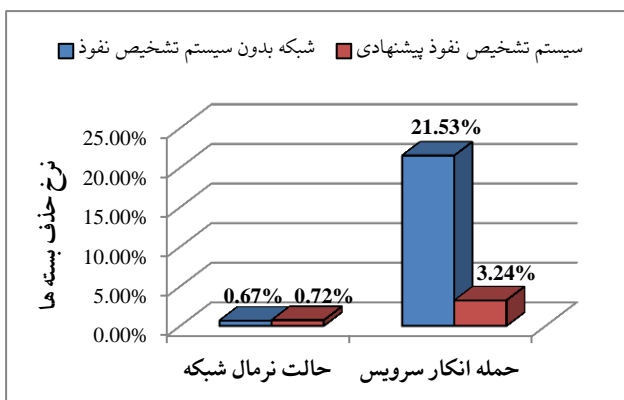
شکل ۷: نمودار نرخ تحویل بسته‌ها در شبکه



شکل ۸: نمودار میانگین تاخیر انتها به انتها در ارسال بسته‌ها



شکل ۹: نمودار میزان گزردهی شبکه



شکل ۱۰: نمودار نرخ حذف بسته‌ها در شبکه

- میانگین مصرف انرژی: این معیار میانگین انرژی مصرف شده در گره‌های شبکه را توسط سیستم تشخیص نفوذ پیشنهادی نشان می‌دهد.

$$\frac{\sum_{i=1}^{No. of nodes} Initial Energy_i - Residual Energy_i}{No. of nodes} \quad (3)$$

- تاخیر انتها به انتها در ارسال: این معیار زمان صرف شده برای ارسال یک بسته در طول شبکه از مبدا به مقصد است.
- میزان گزردهی شبکه: این معیار میزان داده‌های دریافت شده در کل شبکه را در واحد زمان بیان می‌نماید و توسط فرمول زیر محاسبه می‌گردد.

$$\sum_{f=1}^{Max Flow} \frac{No. of received packets * Packet size * 8}{flow time} \quad (4)$$

- نرخ تحویل بسته‌ها: این معیار میزان داده‌های دریافت شده را به نسبت داده‌های ارسال شده در کل شبکه مشخص می‌نماید و از تقسیم گزردهی شبکه بر نرخ ارسال ترافیک حاصل می‌شود.

$$Packet Delivery Ratio = \frac{No. of received packets}{No. of sent packets} * 100 \quad (5)$$

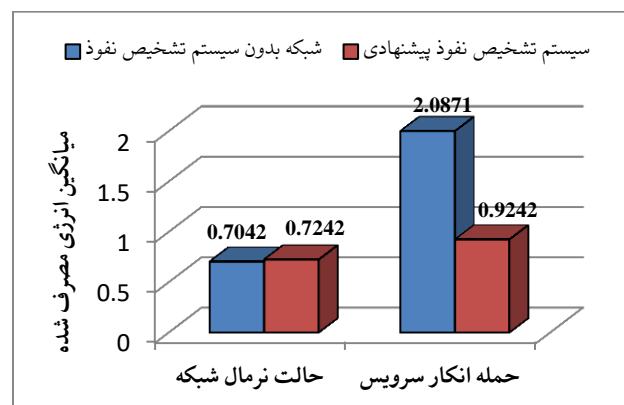
- نرخ حذف بسته‌ها: این معیار درصد بسته‌های حذف شده را نسبت به کل بسته‌های ارسال شده تعیین می‌نماید.

$$Packet Loss Ratio = \frac{Packet Loss}{No. of sent packets} * 100 \quad (6)$$

در ادامه نتایج حاصل از شبیه سازی ها در جدول ۳ و نمودارهای شکل ۳ تا ۱۰ ارائه شده‌اند.

جدول ۳: نتایج و مقایسه تشخیص نفوذ پیشنهادی با کارهای پیشین

سیستم تشخیص نفوذ	نرخ تشخیص	نرخ تشخیص خطا
سیستم تشخیص پیشنهادی	۸۹%	۰/۴%
سیستم مرجع [۴]	۸۶/۴%	---
سیستم مرجع [۱۴]	۸۴%	۰/۳۲%
سیستم مرجع [۱۶]	۸۲%	---
سیستم مرجع [۸]	۶۵%	۱۵%



شکل ۶: نمودار میانگین مصرف انرژی در شبکه

۵- نتیجه

- in International Conference on Communications and Mobile Computing, CMC'09, pp. 507-511, 2009.
- [13] Y. ZHANG, X. LI, and Y. LIU, "The detection and defence of DoS attack for wireless sensor network", The Journal of China Universities of Posts and Telecommunications, vol. 19, pp. 52-56, 2012.
- [14] Z. Baig, "Pattern recognition for detecting distributed node exhaustion attacks in wireless sensor networks", Computer Communications, vol. 34, pp. 468-484, 2011.
- [15] G.H. Lai and C.M. Chen, "Detecting Denial of Service Attacks in Sensor Networks", Journal of Computers, Vol.18, No.4, January 2008.
- [16] M. Guechhari, L. Mokdad, and S. Tan, "Dynamic solution for detecting denial of service attacks in wireless sensor networks", in IEEE International Conference on Communications (ICC), pp. 173-177, 2012.

زیرنویس‌ها

- 1 Routing Protocols
- 2 Pattern Recognition
- 3 Guarding Nodes
- 4 Throughput

ما در این مقاله ابتدا حمله انکار سرویس را بر روی شبکه‌های حسگر بی‌سیم معرفی کردیم. در ادامه انواع سیستم‌های تشخیص نفوذ موجود را مورد ارزیابی قرار دادیم. در بخش بعدی یک سیستم تشخیص نفوذ سبک برای تشخیص حمله انکار سرویس با توجه به خصوصیات مربوط به آن ارائه نمودیم که این خصوصیات را نیز بر اساس تحلیل رفتار حمله انکار سرویس در بخش ۴-۱ استخراج نمودیم. در انتها نیز نتایج حاصل از شبیه‌سازی حمله انکار سرویس و سیستم تشخیص نفوذ پیشنهادی روی آن را ارائه کردیم. با توجه به نتایج ارائه شده در نمودارهای شکل ۶ تا ۱۰، مشاهده می‌گردد که حمله انکار سرویس کارایی شبکه را به شدت تنزل می‌دهد و با بکارگیری سیستم تشخیص نفوذ پیشنهادی به خوبی می‌توان کارایی شبکه را در حد مطلوب نگهداری کرد. همانطور که در جدول ۳ نیز مشاهده می‌شود، سیستم پیشنهادی ما در مقایسه با کارهای موجود دارای نرخ تشخیص بالاتر همراه با نرخ خطای ناچیز می‌باشد، همچنین با توجه به انرژی مصرفی پایین آن روشی نسبتاً کارا و سبک وزن برای تشخیص حملات انکار سرویس می‌باشد.

مراجع

- [1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks", IEEE Communications magazine, vol. 40, no. 8, pp. 102-114, 2002.
- [2] J. Sen, "A Survey on Wireless Sensor Network Security", International Journal of Communication Networks and Information Security. (IJCNIS), vol. 1, no. 2, pp. 55-78, 2009.
- [3] Y. Zhou, Y. Fang, and Y. Zhang, "Securing wireless sensor networks: a survey", IEEE Communications Surveys and Tutorials, vol. 10, no. 3, pp. 6-28, 2008.
- [4] S. Shin, T. Kwon, G.Y. Jo, Y. Park, H. Rhy, "An experimental study of hierarchical intrusion detection for wireless industrial sensor networks", IEEE Transactions on Industrial Informatics, vol. 6, no 4, pp. 744-757, 2010.
- [5] A. Nadeem, M. Howarth, "Adaptive intrusion detection & prevention of denial of service attacks in MANETs", In Proceedings of the 2009 international conference on wireless communications and mobile computing: Connecting the world wirelessly, pp. 926-930. ACM, 2009.
- [6] S. Ghildiyal, A. K. Mishra, A. Gupta, N. Garg, "Analysis Of Denial Of Service (Dos) Attacks In Wireless Sensor Networks", International Journal of Research in Engineering and Technology (IJRET), vol. 3, June 2014.
- [7] A. Agah, K. Basu, and S. K. Das, "Enforcing security for prevention of DoS attack in wireless sensor networks using economical modeling", IEEE International Conference on Mobile Adhoc and Sensor Systems, pp. 1-8, 2005.
- [8] A. Agah and S.K. Das, "Preventing DoS attacks in wireless sensor networks: A repeated game theory approach", International Journal of Network Security, vol. 5, no. 2, 145-153, 2007.
- [9] Y. Zhang, W. Lee, and Y.A. Huang, "Intrusion detection techniques for mobile wireless networks", Journal of Wireless Networks, vol. 9, no. 5, pp. 545-556, 2003.
- [10] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "SPINS: security protocols for sensor networks", Wireless Networks, vol. 8, no. 5, pp. 521-534, 2002.
- [11] A. Agah, K. Basu, S.K. Das, "Preventing DoS attack in sensor networks: A game theoretic approach", Proceedings of IEEE International Conference on Communications, Seoul, South Korea, 2005.
- [12] M. Mohi, A. Movaghar, and P. M. Zadeh, "A Bayesian game approach for preventing DoS attacks in wireless sensor networks",



شناسایی ترافیک شبکه در مسیریاب‌ها با حفظ حریم خصوصی کاربران

محمد رضا گندمی^۱، حمید حسن پور^۲

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود - ایران
ga.mohamadreza@gmail.com

^۲ دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات - دانشگاه صنعتی شاهرود - شاهرود - ایران
h.hassanpour@shahroodut.ac.ir

چکیده

به منظور بالا بردن کیفیت سرویس دهی و شناسایی ناهنجاری‌ها در شبکه، شناسایی و طبقه‌بندی سریع و با دقت ترافیک شبکه نیاز مبرم مدیران شبکه می‌باشد. تاکنون روش‌های متعددی جهت شناسایی ترافیک شبکه ارائه شده است. از جمله این روش‌ها می‌توان به شناسایی مبتنی بر شماره درگاه، شناسایی از طریق بررسی محتوی، و شناسایی برپایه ویژگی‌های آماری اشاره نمود. استفاده از شماره درگاه‌های پویا و عدم امکان دسترسی به محتوی هنگام استفاده از پروتکل‌های رمزنگاری و همچنین نیاز به پردازش بسیار جهت بازگشایی محتوی، موجب عدم کارآمدی روش‌های شناسایی برپایه شماره درگاه و بررسی محتوی شده است. در روش‌های شناسایی ترافیک مبتنی بر ویژگی‌های آماری، عملیات شناسایی ترافیک شبکه بدون نیاز به بررسی محتوی بسته و با هدف حفظ حریم خصوصی ارسال‌کننده بسته‌ها انجام می‌شود. این مقاله ویژگی‌های آماری تاثیرگذار در شناسایی ترافیک شبکه و رفتار آنها را مورد تحلیل و بررسی قرار می‌دهد. برای ارزیابی ویژگی‌های ارائه شده، داده‌های دو پایگاه داده UNIBS و پایگاه داده جمع‌آوری شده بر روی مسیریاب، مورد استفاده و تحلیل قرار گرفته است. جهت ارزیابی نتایج، میزان صحت تشخیص درست برنامه کاربردی توسط الگوریتم‌های Decision Table, Bagging و Random Tree مورد بررسی قرار گرفته است. نتایج ارزیابی نشان می‌دهد که استفاده از ویژگی‌های ارائه شده میزان دقت شناسایی برنامه کاربردی را به بیش از ۹۷٪ افزایش می‌دهد.

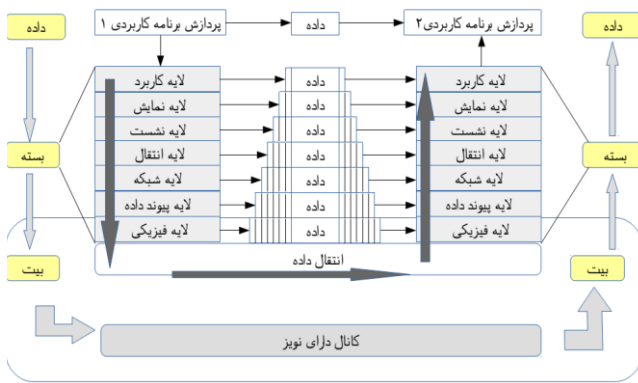
کلمات کلیدی

ترافیک شبکه، شناسایی ترافیک شبکه، یادگیری ماشین، برنامه‌های کاربردی

سوی دیگر، سرویس دهندگان اینترنت را بر آن می‌دارد تا در پی یافتن راهکارهایی جهت بهبود کیفیت بهره‌برداران از منابع شبکه باشند. یکی از این راهکارها، تخصیص پهنای باند مشخص به هر یک از برنامه‌ها و سرویس‌های موجود در شبکه و اولویت‌دهی به آن‌ها در استفاده از ظرفیت خطوط ارتباطی و منابع پردازشی تجهیزات شبکه است. در حقیقت جهت تحقق مدیریت شبکه و کنترل امنیت، مدیران شبکه بایستی هویت و وضعیت برنامه‌های کاربردی را توسط طبقه‌بندی ترافیک معلوم نمایند. همچنین با توجه به افزایش بدافزارها و تلاش آن‌ها برای پنهان‌سازی ترافیک خود به منظور گریز از سیستم‌های تشخیص نفوذ و دور زدن دیوارهای آتش، دسته‌بندی ترافیک به عنوان یک گام اولیه در تشخیص نفوذ و تأمین امنیت شبکه در مقابل

۱- مقدمه

امروزه، با افزایش و رشد برنامه‌های کاربردی، حجم ترافیک ارسالی بر روی بستر اینترنت و شبکه‌های مرتبط با آن روز به روز در حال افزایش می‌باشد. در برخورد با این حجم انبوه از ترافیک شبکه، مدیریت ترافیک شبکه اهمیت بسیار بالایی دارد. گام نخست در نیل به این هدف، شناسایی و طبقه‌بندی ترافیک شبکه است. طبقه‌بندی ترافیک، تکنیک ابتدایی می‌باشد که توسط ارائه دهندگان سرویس‌های اینترنتی، جهت مدیریت منابع شبکه و ضمانت امنیت شبکه مورد استفاده قرار می‌گیرد. از طرفی دیگر، تقاضای رو به رشد برای استفاده از پهنای باند بیشتر از یک سو و محدودیت ظرفیت فیزیکی خطوط ارتباطی شبکه از



شکل ۱: ساختار بسته و ارسال آن در مدل هفت لایه شبکه

۲- شناسایی ترافیک شبکه

۲-۱- تبادل داده در شبکه

در هنگام ارسال داده‌ها، هر لایه پیام‌های لایه بالاتر را دریافت می‌کند، و پس از اجرای برخی پردازش‌ها بر روی آن، مقداری داده به عنوان سرآیند به آن اضافه کرده و آن را به سوی لایه پایین‌تر رهسپار می‌کند. در نهایت در لایه فیزیکی پیام نهایی به صورت امواجی به سوی گیرنده رهسپار می‌گردد. شکل (۱) ساختار تشکیل بسته و نحوه ارسال آن در مدل هفت لایه شبکه‌های رایانه‌ای را نشان می‌دهد. در حالت کلی داده‌های در اختیار تحلیل‌گر ترافیک نظیر برنامه کاربردی X ، به دو گروه ویژگی‌های مرتبط با سرآیند (محتوی بسته در لایه‌های مختلف) و داده‌های مرتبط با ویژگی‌های آماری (زمان تولید بسته، اندازه بسته) تقسیم می‌شوند.

۲-۱-۱- ویژگی‌های مرتبط با محتوی بسته

ویژگی‌های مرتبط با محتوی بسته، می‌تواند شامل محتویات بسته (در صورت عدم استفاده از پروتکل رمزنگاری)، اندازه هر قسمت از بسته در لایه‌های مختلف و ویژگی‌های آماری هر بخش از بسته باشد. در این دسته از داده‌ها برخی مرتبط با برنامه کاربردی مورد نظر است، و برخی هیچ ارتباطی با برنامه کاربردی ندارد. به عنوان مثال، شناسه IP فرستنده جزء داده‌هایی است که هیچ ارتباطی با برنامه کاربردی اجرا شده در سمت فرستنده ندارد. البته از داده‌های حاصل از بازگشایی لایه‌های مختلف یک بسته، تنها بخشی از آن‌ها مفید بوده و به نوعی به برنامه کاربردی تولید کننده آن وابسته می‌باشد.

۲-۱-۲- ویژگی‌های مرتبط با جریان داده

در نظر بگیرید که برنامه کاربردی X بسته λ م خود را در لحظه t_i از لایه کاربرد، رهسپار لایه انتقال می‌کند. این بسته بعد از گذراندن u_i واحد زمانی، توسط تحلیل‌گر ترافیک در زمان d_i دریافت می‌شود و $(d_i = u_i + t_i)$ در حالت کلی یک متغیر تصادفی است که تحت تاثیر عوامل بسیاری نظیر تاخیرهای ناشی از انتظار در صف مسیریاب‌ها و یا کارت‌های شبکه، مدت زمان پردازش بسته در گره‌های شبکه قرار می‌گیرد. زمان مابین تولید دو بسته متوالی T_i به

تهدیدات سایبری دارای اهمیت است. براین اساس، می‌توان طبقه‌بندی ترافیک را مقدماتی لازم برای بسیاری از وظایف امنیتی، مدیریتی و کنترلی در شبکه دانست.

همزمان با پیشرفت و توسعه شبکه‌های رایانه‌ای به خصوص اینترنت، عواملی همچون استفاده از سازوکارهایی مانند پروتکل TLS^1 ، به منظور تامین امنیت لایه کاربرد، استفاده از شماره درگاه‌های پویا، بالا رفتن ترافیک شبکه ناشی از برنامه‌های کاربردی P2P و بالا رفتن نرخ تبادل اطلاعات در شبکه موجب چالش برانگیزتر شدن مبحث طبقه‌بندی ترافیک شده است. تاکنون روش‌های بسیاری برای شناسایی ترافیک شبکه ارائه شده است [7-2]. از جمله این روش‌ها می‌توان به روش‌های شناسایی از طریق شماره درگاه [8] و همچنین شناسایی بر مبنای محتوی بسته [9] اشاره نمود. روش کار در روش‌های شناسایی از طریق شماره درگاه، به این صورت است که تحلیل‌گر با بررسی سرآیند لایه انتقال بسته‌ها، برنامه کاربردی را با توجه به شماره درگاه تخصیص یافته شناسایی می‌نماید. در روش‌های مبتنی بر بررسی محتوی تحلیل‌گر با بازگشایی قسمت داده مربوط به بسته‌ها، با توجه به امضای مخصوص آن برنامه کاربردی، آن را شناسایی می‌کند. این روش‌ها به دلیل پویایی شماره‌های درگاه و همچنین استفاده از پروتکل‌های رمزنگاری دقت خود را از دست داده و به درستی عمل نمی‌کنند. از اینرو در سال‌های اخیر استفاده از روش‌های شناسایی بر پایه استفاده از ویژگی‌های آماری بسته‌ها، به دلیل عدم نیاز به بررسی محتوی بسته که موجب حریم خصوصی کاربران و عدم نشر اطلاعات انتقال یافته می‌شود و همچنین دقت و سرعت بالای آنها، مورد توجه بسیاری از محققین قرار گرفته است.

در این مقاله ضمن مرور بر روش‌های موجود جهت شناسایی ترافیک برنامه‌های کاربردی، به بررسی روش‌های شناسایی بر پایه ویژگی‌های آماری و رفتار این ویژگی‌ها پرداخته می‌شود. سپس ویژگی‌های آماری دو پایگاه داده UNIBS و پایگاه داده جمع آوری شده استخراج شده و توسط الگوریتم‌های Decision Table, Bagging و RandomTree مورد تحلیل قرار می‌گیرد.

در ادامه مقاله، در بخش دوم مروری بر ساختار تبادل داده در شبکه انجام می‌شود. در بخش سوم مدل سامانه شناسایی ترافیک و همچنین فرضیات لازم جهت انجام عملیات شناسایی و در بخش چهارم مروری بر روش‌های شناسایی ترافیک صورت گرفته است. در بخش پنجم، روش‌های شناسایی بر پایه ویژگی‌های آماری مطرح شده و ویژگی‌های موثر مورد تحلیل قرار گرفته است. در بخش ششم، پایگاه داده UNIBS و پایگاه داده جمع‌آوری شده مطرح شده و با استفاده از الگوریتم‌های یادگیری ماشین به تحلیل نتایج پرداخته شده است و در بخش هفتم، به جمع بندی مسائل پرداخته می‌شود.

¹ Transport Layer Security

۲-۲- مروری بر روش‌های موجود

روش‌های موجود برای شناسایی ترافیک شبکه را می‌توان در چهار دسته کلی تقسیم‌بندی نمود [12].

طبقه‌بندی بر مبنای پروتکل و شماره درگاه: این روش طبقه‌بندی از این ایده منتج می‌شود که برنامه‌های کاربردی مختلف از پروتکل‌ها و شماره درگاه‌های متفاوتی استفاده می‌کنند. با توجه به اینکه امروزه استفاده از شماره درگاه‌های پویا رواج بسیار یافته است، این روش طبقه‌بندی چندان مفید فایده نخواهد بود [8].

طبقه‌بندی مبتنی بر بررسی محتوی: در این شیوه محتوای بسته‌های شبکه مورد بازرسی قرار می‌گیرد. این روش طبقه‌بندی، به نوبه خود به دو زیردسته DPI^5 و SPI^6 طبقه‌بندی می‌شود. در DPI ، سعی می‌شود که محتوی با یک امضای ثابت مقایسه گردد. طبقه‌بندی بر مبنای پروتکل و شماره درگاه زیرمجموعه DPI محسوب می‌گردد. DPI به صورت گسترده در نرم‌افزارها و دیوار آتش مورد استفاده قرار می‌گیرد [9]. پیچیدگی بالا و عدم دستیابی به محتوای رمزشده برخی لایه‌ها، از بزرگترین اشکالات این روش است. در SPI ، ویژگی‌های آماری محتوی در نظر گرفته می‌شود. به عنوان مثال، نویسندگان مقاله در [13] یک روش شناسایی ترافیک بر مبنای خواص آماری تعدادی از بیت‌های اول بسته UDP ارایه داده‌اند.

روش‌های مبتنی بر رفتار: طبقه‌بندی روش‌های مبتنی بر رفتار بسیار گسترده‌تر از دیگر روش‌ها است. در این شیوه، سعی می‌شود که در یک بازه زمانی بزرگ تمامی ترافیک رسیده از کاربر مورد بررسی قرار گیرد و الگوی کلی ترافیک او استخراج شود. برخی از کارهای تحقیقاتی در این نوع طبقه‌بندی در مراجع [14, 15] قابل مشاهده می‌باشند.

طبقه‌بندی مبتنی بر ویژگی‌های آماری: در این شیوه به ویژگی‌های آماری زمان و اندازه بسته‌های دریافتی، پرداخته می‌شود. مهم‌ترین گام در این قسمت، استخراج تمایزهای آماری بین برنامه‌های کاربردی مختلف است. مقالات بسیاری نیز در این حوزه مطرح شده است [4, 16, 17].

لازم به ذکر است که وجه تمایز اصلی بین طبقه‌بندی آماری و SPI این است که در طبقه‌بندی آماری، ویژگی‌های آماری زمان تولید بسته‌ها و طول آن‌ها، بدون در نظر گرفتن محتوا، مورد بررسی قرار می‌گیرد؛ بدین‌سان این روش‌ها از لحاظ محاسباتی بسیار سبک هستند، چراکه دیگر نیازی به بازگشایی محتوی بسته ندارد. در این نوع از تحلیل، علم تئوری صف نقش بسیار برجسته‌ای دارد، چرا که بسیاری از اتفاقاتی که در مسیر ارسال یک بسته تا رسیدن آن به مقصد رخ می‌دهد را می‌توان با یک صف مدل نمود. برخی از قوانین حاکم بر صف در این مورد می‌تواند مورد استفاده قرار گیرد، به عنوان مثال قانون پایستگی جریان، که بر طبق آن در یک سامانه کلی نرخ ورود بسته‌ها به صف برابر با نرخ خروج از صف است [18].

صورت رابطه (۱) تعریف می‌شود و زمان مابین دریافت دو بسته متوالی توسط تحلیل‌گر ترافیک، با پارامتر D_i نشان داده می‌شود، و به صورت رابطه (۲) تعریف می‌شود.

$$T_i = t_i - t_{i-1} \quad (1)$$

$$D_i = d_i - d_{i-1} = T_i + u_i - u_{i-1} \quad (2)$$

فرض کنید، $f(T_{j1}, T_{j2}, \dots, T_{jn})$ به عنوان تابع چگالی توام مرتبه n بین متغیر تصادفی $T_{j1}, T_{j2}, \dots, T_{jn}$ در نظر گرفته شود. لازم به ذکر است که اگر متغیر تصادفی T_i از یک توزیع iid^2 پیروی کند، با دانستن تابع توزیع مرتبه اول که با $f_T(t)$ نشان داده می‌شود، می‌توان تابع توزیع تمامی مراتب را بدست آورد [2]. ذکر این نکته ضروری است که اگر تابع توزیع زمان مابین تولید بسته‌ها در هر مرتبه‌ای موجود باشد، می‌توان تمامی اطلاعات آماری مربوط به برنامه کاربردی X را استخراج نمود. به ندرت دو برنامه کاربردی را می‌توان یافت که تمامی اطلاعات آماریشان مانند هم باشد. لذا به نظر می‌رسد که اطلاعات آماری، یک امضای منحصر به فرد هر برنامه کاربردی خواهد بود. اما متأسفانه این کار به این سادگی نخواهد بود. به طور کلی با در نظر گرفتن رابطه (۲) و داشتن $f(T_{j1}, T_{j2}, \dots, T_{jn})$ نمی‌توان توزیع $f(D_{j1}, D_{j2}, \dots, D_{jn})$ را بدست آورد. تحلیل‌گر نیاز به دانستن تابع توزیع u_i در هر مرتبه‌ای و به ازای هر i دارد. از سوی دیگر بدست آوردن تابع توزیع u_i کار بسیار دشواری است. اگر T_i دارای توزیع iid نباشد، استخراج تمامی توابع توزیع در تمامی مراتب (تا بی‌نهایت مرتبه) به دلیل محدودیت پردازشی، امری غیر ممکن است. حتی اگر T_i دارای توزیع iid باشد، لزوماً D_i دارای توزیع iid نخواهد بود [10, 11]. در شبکه اینترنت برای اطلاع از اینکه بسته‌ها با موفقیت به مقصد رسیده‌اند یا خیر، از سازوکار ARQ^3 استفاده می‌شود.

ARQ موجب می‌شود که زمان ارسال بسته‌ها از سمت فرستنده، تابعی از زمان دریافت ACK^4 بسته‌های قبلی باشد. بدین‌سان توزیع زمان ورود بسته‌ها نه تنها به نوع کاربرد، بلکه به تمامی ساختار شبکه‌ای موجود بین فرستنده و گیرنده و میزان ترافیک شبکه وابسته خواهد بود. ذکر این نکته حائز اهمیت می‌باشد که تحلیل‌گر ترافیک می‌تواند به شرط عدم رمزنگاری، بسته‌ها را تا لایه کاربرد بازگشایی کند و از محتوای آن آگاه شود. وظیفه اصلی مسیریاب تعیین مسیر بسته‌ها با استفاده از جدول ارسال می‌باشد. در شبکه نیز تعیین مسیر توسط آدرس IP صورت می‌گیرد. در نتیجه مسیر یاب‌های شبکه، کار خود را در لایه شبکه انجام داده و بسته‌ها تا لایه شبکه بازگشایی می‌شوند.

این بدان معناست که، بررسی محتویات بسته تا لایه شبکه سربرار چندانی نخواهد داشت. اما اگر تحلیل‌گر ترافیک، لایه‌های بالاتر را نیز بازگشایی کند، بایستی جهت بازگشایی هر لایه و خواندن سرآیند و محتوی بسته‌ها، متحمل سربرار محاسباتی و زمانی زیادی شود.

² Independent Identically Distributed

³ Automatic Repeat Request

⁴ Acknowledgement

⁵ Deep Packet Inspection

⁶ Stochastic Packet Inspection

۳- مدل سیستم و فرضیات

در نظر بگیرید که سه کاربر به نام‌های A, B, C به شبکه اینترنت متصل هستند (شکل (۲)) و اطلاعات آن‌ها از طریق مسیریاب‌های مختلف به مقصد برسد. تحلیلگر ترافیک در مسیر اتصال دو مسیریاب واقع شده و شروع به ضبط داده‌های این اتصال می‌کند. اگر تحلیلگر بخواهد شناسایی ترافیک را به صورت کامل انجام دهد، پیش از هرکار ترافیک کاربران مختلف باید از یکدیگر تمییز داده شوند. در نگاه نخست استفاده از آدرس IP به عنوان معیاری برای تمایز کاربران مفید به نظر می‌آید. لیکن این موضوع با وجود زیرشبکه‌های NAT^۷ که موجب پنهان شدن آدرس IP کاربران خود می‌گردد، دیگر به سادگی مقدر نیست. در مرحله بعد و پس از تفکیک ترافیک یک کاربر از دیگر کاربران، باید بین ترافیک برنامه‌های کاربردی مختلف کاربر تمایز و جداسازی انجام شود. در انتها اگر هدف فیلترینگ (تصفیه سازی) است، باید ترافیکی که با ترافیک هدف تطابق دارد شناسایی و برچسب زده شود، اگر هدف این باشد که تک‌تک جریان‌های شناسایی شوند باید نگاهی مناسب بین ترافیک‌ها و برنامه‌های کاربردی نسبت داده شود. بایستی دقت داشت که اطلاعات محدود به محتوای بسته‌ها نبوده و هر نوع اطلاعاتی که به نوعی با برنامه کاربردی در حال اجرا مرتبط است، می‌تواند برای تحلیلگر شبکه مفید باشد. به عنوان نمونه می‌توان به نوع پروتکل برنامه کاربردی اشاره نمود.

۴- شناسایی بر پایه ویژگی‌های آماری

در بسیاری از مراجع ترافیک برنامه‌های کاربردی در حال اجرا بر روی شبکه به دو دسته ترافیک برنامه‌های نظیر به نظیر (P2P) و ترافیک برنامه‌های وب (Client-Server) تقسیم می‌شوند. در روش شناسایی بر پایه ویژگی‌های آماری، محتوی بسته‌ها (لایه انتقال به بالا) در نظر گرفته نمی‌شوند، لذا محتوی تولید شده توسط کاربر کاملاً مخفی مانده و حریم خصوصی آنها کاملاً حفظ می‌گردد. در نتیجه رفتار برنامه‌های کاربردی با توجه به خصوصیات لایه شبکه بررسی می‌شوند. در جدول (۱) خصوصیات رفتاری ترافیک برنامه‌های کاربردی P2P و Client-Server با یکدیگر مقایسه شده‌اند.

جدول (۱) خصوصیات رفتاری برنامه‌های کاربردی P2P و

Client-Server		خصوصیات
ترافیک (Client-Server)	ترافیک (P2P)	
کم ۰,۷-۰,۸	زیاد <0.5	میزان بایت ارسالی پارامتر خاصیت خودشبهی توزیع داده‌ها
Heavy Tail, On/Off <100 ms	Long Tail >100 ms	مدت زمان جریانها نرخ ارسال داده bit/sec
زیاد >0.3 ms	کم <0.3 ms	زمان بین ارسال و دریافت بسته‌ها
< 2.5 * 10 ⁴ byte	> 2.5 * 10 ⁴ byte	اندازه پنجره ارسال
کم	زیاد	آنتروپی

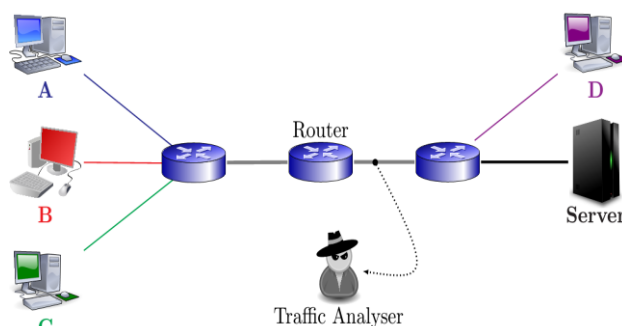
با توجه به ویژگی‌های رفتاری برنامه‌های کاربردی و پس از بررسی‌های به عمل آمده، چهار ویژگی آماری ترتیب اندازه پنجره ارسال، مدت زمان بین ارسال بسته‌ها، مدت زمان جریان‌ها و تعداد بسته‌ها در یک جریان در لایه شبکه استخراج گردید. شکل (۳) نمودار توزیع احتمالاتی این ویژگی‌ها برای شش برنامه کاربردی Skype, Telegram, Idm, Firefox, Chrome و IE در مجموعه داده جمع آوری شده که در بخش ۱-۵ معرفی شده است را نشان می‌دهد. همانطور که مشاهده می‌شود رفتار برنامه‌های کاربردی P2P و Client-Server در این ویژگی‌ها کاملاً متفاوت می‌باشد. به طور مثال در قسمت توزیع اندازه پنجره ارسال، همانطور که مشاهده می‌شود برنامه‌های کاربردی Skype, Telegram و IDM دارای رفتاری مشابه در این ویژگی می‌باشند ولی در خصوص برنامه‌های کاربردی Chrome, Firefox و IE تابع توزیعی شبیه به هم ولی کاملاً متفاوت با برنامه‌های کاربردی P2P می‌باشند. این تغییر رفتار را می‌توان در تغییر پارامتر تابع نمایی دانست.

۵- نتایج ارزیابی

در بخش قبل چهار ویژگی آماری بر اساس رفتار برنامه‌های کاربردی استخراج گردید. در این بخش در ابتدا دو پایگاه داده Unibs و پایگاه جمع آوری شده بر روی مسیر یاب بررسی شده و با استفاده از الگوریتم‌های یادگیری ماشین عملکرد این ویژگی‌ها در شناسایی شش برنامه کاربردی بررسی شده است.

۵-۱- پایگاه داده UNIBS

این داده‌ها از مسیریاب درون دانشگاه برسیکا در سال ۲۰۰۹ جمع‌آوری شده است. Ethernet دانشگاه شامل چندین 1000Base-TX (که یکی از رایج‌ترین و سریعترین اشکال اتترنت فعلی است)



شکل ۲: اجرای برنامه‌های کاربردی مختلف بر روی شبکه

⁷ Network Address Translation

داده‌ها از لینک 100Mb/s در مسیریاب به کمک نرم‌افزار Tcpcdump جمع‌آوری شده است. از آنجایی که دسترسی کامل به روتر وجود داشته است، ۲۵۰ بایت اول هر فریم ذخیره شده است. برچسب زنی بر داده‌ها توسط الگوریتم ground truth و با بررسی محتوای بسته‌ها انجام شده است. به همین جهت برچسب جریان‌های مربوط به برنامه‌های کاربردی با اطمینان بالایی تهیه شده است. جدول (۲) مشخصات و تعداد نمونه‌های زیرمجموعه‌ای حاوی شش برنامه کاربردی از این پایگاه داده را نشان داده است [19].

جدول (۲). پایگاه داده UNIBS

تعداد نمونه	برنامه کاربردی
۳۵۱۶۴۱	Firefox
۶۵۶۶	Thunderbird
۱۸۱۵۵	Skype
۳۴۸۲۶۵	Safari
۱۲۴۴۲۴	Bitorrent
۱۰۰۷۷۹	Mail

۵-۲- پایگاه داده جمع‌آوری شده

این داده‌ها از مسیریاب یک شبکه محلی در سال ۲۰۱۷ جمع‌آوری شده است. ترافیک آخرین نسخه شش نرم‌افزار کاربردی در سال ۲۰۱۷ با استفاده از نرم‌افزار Tcpcdump و از مسیریاب با لینک 1gb/s ضبط شده است. نحوه ضبط ترافیک شبکه بدین صورت بوده است که ۵۰ نود بر روی شبکه محلی قرار داشته و ترافیک این نودها از یک مسیریاب عبور کرده است. در نودهای خاص برنامه‌های کاربردی خاصی اجرا شده و نحوه برچسب زنی بر داده‌ها بدین صورت انجام شده است. جدول (۳) مشخصات و تعداد نمونه‌های ضبط شده را نشان می‌دهد. ما با استفاده از لاگ نرم افزار Tcpcdump نشست‌های ترافیک شبکه را مورد بررسی قرار داده‌ایم و نتیجه آن بدین صورت می‌باشد که تنها دو درصد از بسته‌ها از دست رفته‌اند.

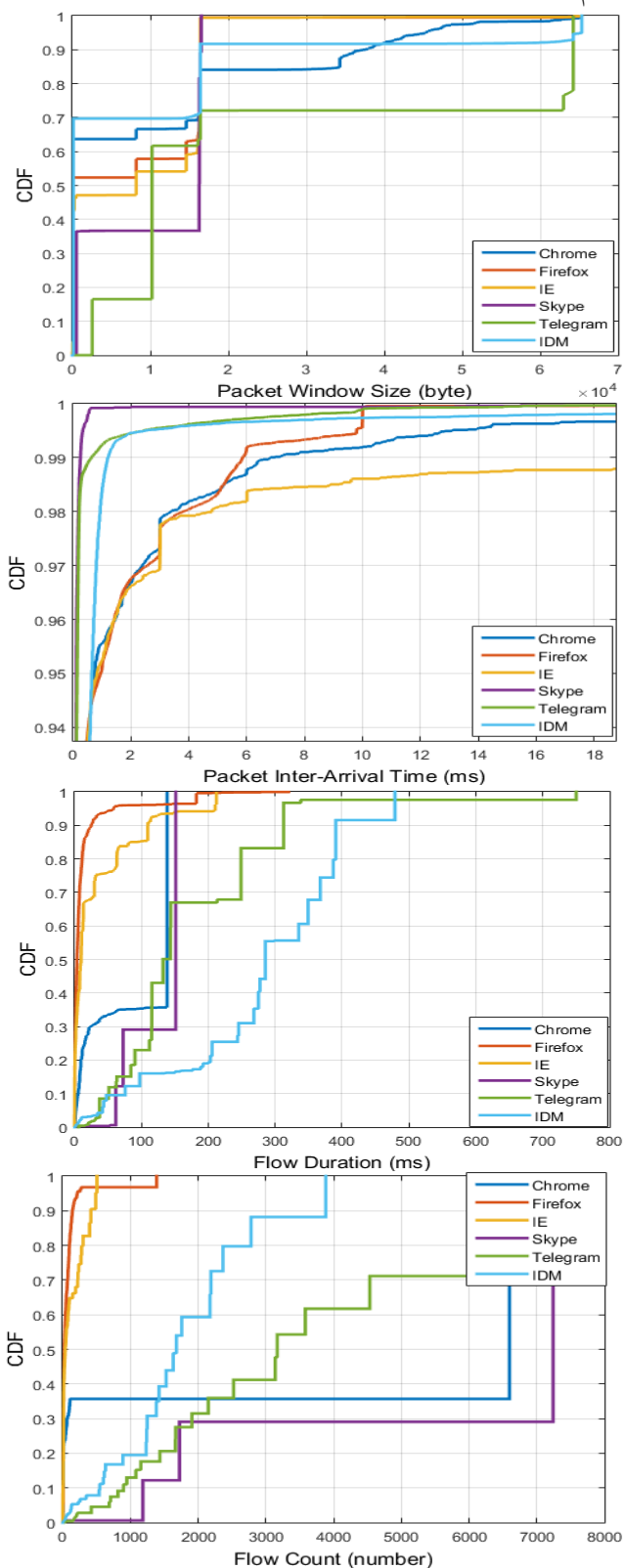
جدول (۳). پایگاه داده جمع‌آوری شده

تعداد نمونه	برنامه کاربردی
۲۱۸۹۷۰	Firefox
۲۱۷۴۰۵	Skype
۲۱۶۸۲۷	Telegram
۲۱۷۱۳۵	IDM
۲۱۲۴۲۴	Chrome
۲۱۱۴۲۴	IE

۵-۳- نتایج شناسایی

جهت شناسایی برنامه‌های کاربردی، با استفاده از ویژگی‌های استخراج شده در بخش قبل، عملیات شناسایی برنامه‌های کاربردی توسط سه الگوریتم Decision Table، Bagging و RandTree انجام شد. شکل (۴) و (۵) دقت شناسایی برنامه‌های کاربردی با استفاده از اجرای سه الگوریتم یادگیری ماشین بر روی دو پایگاه داده را نشان می‌دهد.

تشکیل شده که مسیریابی توسط پردازشگرهای دو هسته‌ای لینوکس انجام شده است.

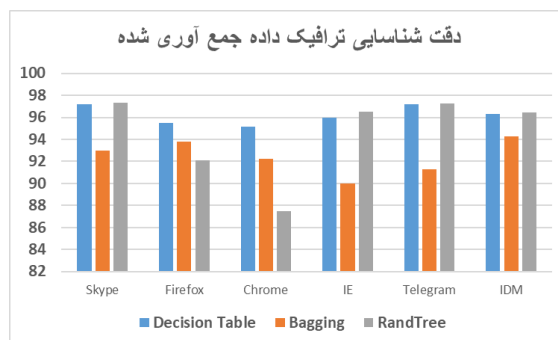


شکل ۳: توزیع احتمالاتی (CDF) ویژگی‌های استخراج شده برای شش برنامه کاربردی، Skype, Telegram, Idm, Firefox, Chrome و IE

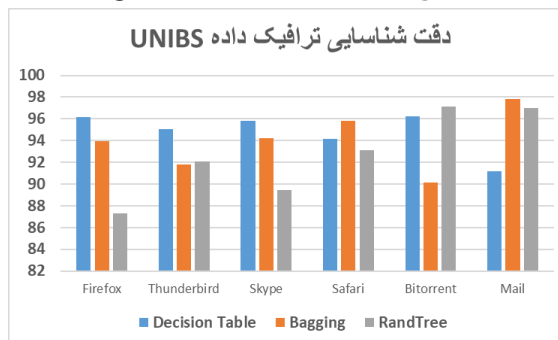
بالاتری نسبت به روش‌های دیگر را نشان می‌دهد. دلیل استفاده از دو پایگاه داده که در سال‌های متفاوت جمع‌آوری شده و دارای ترافیک برنامه‌های کاربردی با دو نسخه قدیمی و جدید هستند، این امر می‌باشد که ویژگی‌های استخراج شده از رفتار برنامه‌های کاربردی، قابلیت تمایزدهنگی خود را در نسخه‌های متفاوت برنامه‌های کاربردی حفظ کرده و همچنان توسط این ویژگی‌ها می‌توان ترافیک برنامه‌های کاربردی را در داده‌های مختلف با دقت بالایی شناسایی نمود.

۶- جمع‌بندی

در این مقاله مروری بر مفاهیم شناسایی ترافیک شبکه و همچنین انواع روش‌های شناسایی ترافیک برنامه‌های کاربردی انجام گرفت. در این میان روش‌های شناسایی بر پایه ترکیب ویژگی‌های آماری و الگوریتم‌های یادگیری ماشین، به دلیل عدم بررسی محتوی بسته و حفظ حریم خصوصی و همچنین دقت و سرعت بالای شناسایی برنامه‌های کاربردی، مورد توجه بسیاری از محققین قرار گرفته است. از اینرو در این مقاله رفتار شش برنامه کاربردی شامل Skype, Firefox, Chrome, IE, Telegram و IDM مورد بررسی قرار گرفت و چهار ویژگی با قابلیت تمایزدهنگی بالا استخراج گردید. جهت ارزیابی ویژگی‌ها داده‌های دو پایگاه داده UNIBS و پایگاه داده جمع‌آوری شده مورد بررسی قرار گرفت. در این مقاله از سه الگوریتم یادگیری ماشین Decision Table, Bagging و RandomTree جهت تحلیل نتایج استفاده شده است. نتایج حاصل از شناسایی برنامه‌های کاربردی با استفاده از ویژگی‌های استخراج شده نشان می‌دهد که ویژگی‌های استخراج شده با توجه به رفتار برنامه‌های کاربردی دارای تمایزدهنگی بالایی بوده و دقت شناسایی کننده را به مقدار قابل قبولی افزایش می‌دهد.



شکل (۴). نتایج بدست آمده برای پایگاه داده جمع‌آوری شده



شکل (۵). نتایج بدست آمده برای پایگاه داده UNIBS

تعداد نمونه‌ها برای برنامه‌های کاربردی متفاوت یکسان نیست، به همین دلیل هر کدام از برنامه‌های کاربردی به صورت مجزا مورد بررسی قرار گرفته است.

با توجه به نتایج بدست آمده جدول (۴) به مقایسه نتایج روش پیشنهادی با کارهای مهم انجام شده در این زمینه می‌پردازد. همانطور که مشاهده می‌شود، ویژگی‌های استخراج شده از رفتار برنامه‌های کاربردی برای داده‌های دو پایگاه داده از نسخه‌های مختلف برنامه‌های کاربردی، موجب دقت شناسایی ۹۷,۵ درصد شده و این نتایج دقت

جدول (۴). مقایسه نتایج روش پیشنهادی و کارهای انجام شده

مقاله مرجع	روش شناسایی استفاده شده	ویژگی‌های مورد استفاده	برنامه‌های کاربردی	دقت شناسایی	پایگاه داده
[13]	Baysian Techniques	مدت زمان جریان، شماره درگاه، زمان مابین ارسال بسته‌ها، اندازه بسته، آنتروپی پهنای باند	P2P, Bulck, Mail, Services	۹۶,۲۹%	Proprietary Hand Classified Traces
[16]	Naive Bayes and Pearson's Chi-Squate test	تبدیل فوری مدت زمان مابین ارسال بسته‌ها اندازه بسته در لایه انتقال، میانگین مدت زمان مابین ارسال بسته‌ها	Skype traffic	97%	Two self collected datasets
[20]	C4.5 Decision tree	تعداد بسته‌ها، میزان بایت (ارسال و دریافت)، اندازه بسته (کمترین، بیشترین، میانگین، انحراف معیار)، زمان ارسال بسته‌ها (کمترین، بیشترین، میانگین، انحراف معیار)	torrent, dns, http, smtp, ssh, ssl	90%	UTM campus network
[15]	Feed-forward Neural Network	Fast Correlation-Based Filter Algorithm	http, p2p, imap, pop3,smtp, mysql	95%	Hand Classified Traces
روش پیشنهادی	Decision Table, Bagging, RandomTree	اندازه پنجره ارسال، مدت زمان بین بسته‌ها، زمان جریان، تعداد بسته در جریان	Skype, Firefox, Chrome, IE, Telegram, IDM	۹۷,۵%	پایگاه داده جمع آوری شده، UNIBS

مراجع

- [1] J. Zhang, Y. Xiang, Y. Wang, W. Zhou, Y. Xiang, and Y. Guan, "Network Traffic Classification Using Correlation Information," *IEEE Transactions on Parallel and Distributed Systems*, vol.24, pp.104–117, Jan. 2015.
- [2] P. Foremski, "On different ways to classify Internet traffic: a short review of selected publications," *Theoretical and Applied Informatics*, vol.25, no.2, pp.119–136, 2015.
- [3] Y. Wang, Y. Xiang, and J. Zhang, "Network traffic clustering using Random Forest proximities," in *IEEE International Conference on Communications (ICC)*, no. 60970146, pp.2058–2062, June 2014.
- [4] D. Adami, C. Callegari, S. Giordano, M. Pagano, and T. Pepe, "Skype-Hunter: A real-time system for the detection and classification of Skype traffic," *International Journal of Communication Systems*, vol.25, pp.386–403, Mar. 2012.
- [5] A. Finamore, M. Mellia, M. Meo, and D. Rossi, "KISS: Stochastic Packet Inspection Classifier for UDP Traffic," *IEEE/ACM Transactions on Networking*, vol.18, pp.1505–1515, Oct. 2010.
- [6] D. Bonfiglio, M. Mellia, M. Meo, D. Rossi, and P. Tofanelli, "Revealing skype traffic: when randomness plays with you," in *ACM SIGCOMM Computer Communication Review*, vol.37, pp.37–48, ACM, 2007.
- [7] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," in *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, (Amsterdam, The Netherlands, The Netherlands), pp.3–24, IOS Press, 2007.
- [8] T. Abuhmed, A. Mohaisen, and D. Nyang, "A Survey on Deep Packet Inspection for Intrusion Detection Systems," *Magazine of Korea Telecommunication Society*, vol.24, no.11, pp.25–36, 2008.
- [9] D. Daley, "Revisiting queueing output processes: a point process viewpoint," *Queueing Systems*, vol.68, no.3–4, pp.395–405, 2011.
- [10] D. J. Daley and D. N. Shanbhag, "Independent InterDeparture Times in M/G/1/N Queues," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol.37, no.2, pp.259–263, 1975.
- [11] S. Valenti, D. Rossi, A. Dainotti, A. Pescapè, A. Finamore, M. Mellia, and A. Pescapè, "Reviewing Traffic Classification," in *Data Traffic Monitoring and Analysis SE - 6* (E. Biersack, C. Callegari, and M. Matijasevic, eds.), vol.7754 of *Lecture Notes in Computer Science*, pp.123–147, Springer Berlin Heidelberg, 2013.
- [12] A. Finamore, M. Mellia, M. Meo, and D. Rossi, "KISS: Stochastic packet inspection Classifier for UDP traffic," *IEEE/ACM Transactions on Networking*, vol.18, no.5, pp.1505–1515, 2010.
- [13] T. Karagiannis, A. Broido, M. Faloutsos, and K. Claffy, "Transport layer identification of P2P traffic," in *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement - IMC '04*, (New York, New York, USA), p.121, ACM Press, 2004.
- [14] T. Karagiannis, K. Papagiannaki, N. Taft, and M. Faloutsos, "Profiling the End Host," in *Proceedings of the 8th International Conference on Passive and Active Network Measurement, PAM'07*, (Berlin, Heidelberg), pp.186–196, Springer-Verlag, 2007.
- [15] V. Carela-Español, P. Barlet-Ros, M. Solé-Simó, A. Dainotti, W. de Donato, and A. Pescapè, "K-dimensional Trees for Continuous Traffic Classification," in *Proceedings of the Second International Conference on Traffic Monitoring and Analysis, TMA'10*, (Berlin, Heidelberg), pp.141–154, Springer-Verlag, 2010.
- [16] N. Williams, S. Zander, and G. Armitage, "A preliminary performance comparison of five machine learning algorithms for practical ip traffic flow classification," *ACM SIGCOMM Computer Communication Review*, vol.36, no.5, pp.5–16, 2006.
- [17] G. Kesidis, "An Introduction to Communication Network Analysis," June 2007.
- [18] The UNIBS anonymized Internet traces. <http://www.ing.unibs.it/ntw/tools/traces/> (Accessed on 30th Sep. 2017)
- [19] J. Zhang, C. Chen, Y. Xiang, W. Zhou, and Y. Xiang, "Internet traffic classification by aggregating correlated naive bayes predictions," *Information Forensics and Security, IEEE Transactions on*, vol.8, no.1, pp.5–15, 2013.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه یزد

حذف نویز جهتی با استفاده از آرایه‌های میکروفونی تفاضلی

حمید ولی نژاد طریقه^۱، حمیدرضا ابوطالبی^۲

^۱ دانشجوی کارشناسی ارشد، دانشکده مهندسی برق، دانشگاه یزد، یزد
hamid.valinejad@yahoo.com

^۲ استاد، دانشکده مهندسی برق، دانشگاه یزد، یزد
habutalebi@yazd.ac.ir

چکیده

سیگنال گفتار یکی از منابع صوتی مهم و حاوی اطلاعات به شمار می‌رود که توسط عوامل محیطی فراوانی تخریب می‌شود. در این راستا آرایه‌های میکروفونی به منظور پردازش سیگنال‌های صوتی در محیط‌های نویزی مورد توجه بسیار قرار گرفته‌اند. در این میان، آرایه‌های میکروفونی تفاضلی به دلیل ویژگی‌های جذابی که دارند در این مقاله بررسی می‌شوند. هدف این مقاله حذف نویز جهتی با استفاده از آرایه‌های میکروفونی تفاضلی است. بدین منظور با ترکیب آرایه تفاضلی و روش‌های مبتنی بر فیلتر وینر ساختاری برای حذف نویز جهتی از سیگنال گفتار نویزی ارائه می‌شود. در این ساختار از سه حسگر (یک حسگر همه‌جهته و یک آرایه تفاضلی مرتبه اول) استفاده شده است. سیگنال نویزی توسط حسگر همه‌جهته دریافت شده و آرایه تفاضلی وظیفه دریافت سیگنال نویز خالص را بر عهده دارد. نتایج بدست آمده از شبیه‌سازی‌ها نشان می‌دهد که روش پیشنهادی عملکرد قابل ملاحظه‌ای در راستای حذف نویز جهتی نسبت به روش پایه داشته است.

کلمات کلیدی

بهسازی گفتار، آرایه‌های میکروفونی تفاضلی، تفریق طیفی، فیلتر وینر، نویز سفید گوسی

[1]. یکی از انواع تقسیم‌بندی‌های آرایه‌های میکروفونی بر اساس نحوه پاسخ آرایه در یک میدان صوتی است که بر پایه آن، آرایه‌های میکروفونی به دو گروه: (۱) آرایه‌های میکروفونی جمعی^۱ و (۲) آرایه‌های میکروفونی تفاضلی (DMA) تقسیم‌بندی می‌شوند [2,3].

آرایه‌های جمعی آرایه‌هایی هستند که با سنکرون‌سازی و اعمال کردن یک سری قواعد، به سیگنال بهسازی شده دست پیدا می‌کنند. در چنین آرایه‌هایی فشار صوتی در هر حسگر به یک سیگنال الکتریکی تبدیل می‌شود. با پردازش بر روی خروجی این حسگرها، سیگنال مطلوب بیرون کشیده شده و نویز و تداخل تضعیف می‌گردد [۲]. در حالت کلی، آرایه‌های جمعی ابعاد بزرگی دارند و اگرچه ممکن است در فرکانس‌های بالا به خوبی عمل کنند، اما در فرکانس‌های پایین چندان مناسب نیستند [4].

DMAها مرکب از تعدادی حسگر فشار صوتی بوده و به خاطر داشتن شکل الگوی تشعشی ثابت (نسبت به فرکانس) مورد توجه واقع شده‌اند. یکی از فرض‌های اساسی در ساخت DMAها این است که حسگرها باید آن‌قدر به هم نزدیک باشند که مشتق فشار آکوستیکی - که به وسیله تفاضل‌های محدود بین خروجی حسگرها محاسبه می‌گردد- به طور صحیح تقریب زده شود. این یکی از دلایل اصلی کوچک بودن

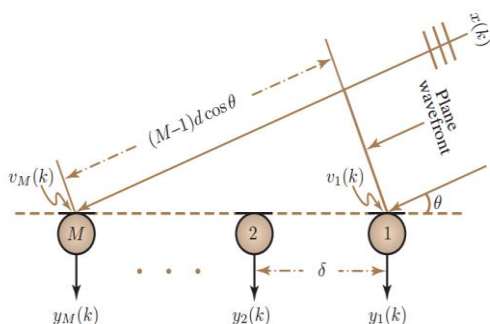
۱- مقدمه

در محیط‌های واقعی وجود نویز و سایر عوامل مخرب محیطی، اجتناب‌ناپذیر بوده و موجب کاهش کیفیت و قابلیت فهم سیگنال گفتار می‌شود [1]. در سیستم‌های قدیمی، سیگنال‌های صوتی نوعاً با یک حسگر ضبط شده و پردازش‌های لازم بر روی خروجی همان تک‌حسگر صورت می‌پذیرفت؛ اما این روش از محدودیت‌هایی نظیر از دست رفتن حالت طبیعی سیگنال صوتی، انعطاف ناپذیری سیستم در پردازش سیگنال ضبط شده و ... رنج می‌برد. برای مقابله با این معایب، مفهوم آرایه‌های میکروفونی مطرح شد. آرایه‌های میکروفونی توانایی بالقوه‌ای در حل مسایل مهمی مانند: کاهش نویز، حذف پژواک، تفکیک منابع و مکان‌یابی منابع، دارند. با این وجود طراحی و پیاده‌سازی آرایه‌های میکروفونی و الگوریتم‌های شکل‌دهنده پرتو آنها برای پردازش سیگنال‌های پهن‌بند مثل گفتار کار آسانی نیست [1,2].

آرایه میکروفونی شامل تعدادی حسگر صوتی است که به گونه‌ای خاص کنار هم قرار گرفته و با توجه به تنوع موجود در سیگنال‌های دریافتی حسگرها، امکان انجام پردازش‌های مختلف را فراهم می‌آورد

۲- مدل سیگنال

یک منبع سیگنال با امواج تخت که در یک محیط آکوستیکی بدون پژواک و با فرض پردازشی میدان دور با سرعت $c = 340\text{m/s}$ منتشر می‌شود را در نظر می‌گیریم. این امواج مطابق شکل ۱ به یک آرایه یکنواخت با M میکروفون با فاصله یکسان δ برخورد می‌کنند [1].



شکل ۱: آرایه خطی با فاصله یکنواخت بین میکروفون‌ها و منبع سیگنال در فاصله دور [1]

اگر از تضعیف انتشار صرف نظر کنیم سیگنال دریافتی در زمان‌های گسسته k و در میکروفون m ام به صورت زیر است:

$$\begin{aligned} y_m(k) &= x_m(k) + v_m(k) \\ &= x(k - t - \tau_m) + v_m(k) \end{aligned} \quad (1)$$

که t زمان انتشار از منبع $x(k)$ تا میکروفون اول (نقطه مرجع) است، τ_m تأخیر نسبی بین میکروفون اول و m ام، $x_m(k)$ و $v_m(k)$ به ترتیب سیگنال مطلوب و نویز مشاهده شده در میکروفون m ام است. فرض می‌شود که نویز و سیگنال مطلوب ناهمبسته هستند. τ_m در رابطه مذکور را به صورت زیر می‌توان بیان کرد:

$$\tau_m = \frac{(m-1)\delta \cos\theta}{c}, \quad m = 1, 2, \dots, M \quad (2)$$

که δ فاصله بین دو میکروفون همسایه و c سرعت صوت در هوا و θ هم زاویه برخورد سیگنال است. برای سادگی روابط را در حوزه فرکانس بازنویسی می‌کنیم [1]:

$$\begin{aligned} Y_m(\omega) &= X_m(\omega) + V_m(\omega) \\ &= x(\omega) e^{-j\omega(t+\tau_m)} + v_m(\omega) \end{aligned} \quad (3)$$

که $\omega = 2\pi f$ فرکانس زاویه‌ای و $j = \sqrt{-1}$ است. به منظور استفاده از فرض میدان دور بردار هدایت با طول M به صورت زیر تعریف می‌شود:

$$\mathbf{d}(\omega, \cos\theta) = [1 e^{-j\omega\delta \cos\theta/c} \dots e^{-j(M-1)\omega\delta \cos\theta/c}]^T \quad (4)$$

که منظور از بالانویس T عملگر ترانپوز است و $\tau_0 = \delta/c$ تأخیر بین دو حسگر مجاور در زاویه $\theta = 0^\circ$ است. در DMAها همیشه فرض $\delta \ll \lambda$ را داریم؛ در نتیجه [7,8]:

$$\frac{\omega\delta}{c} = \omega\tau_0 \ll 2\pi \quad (5)$$

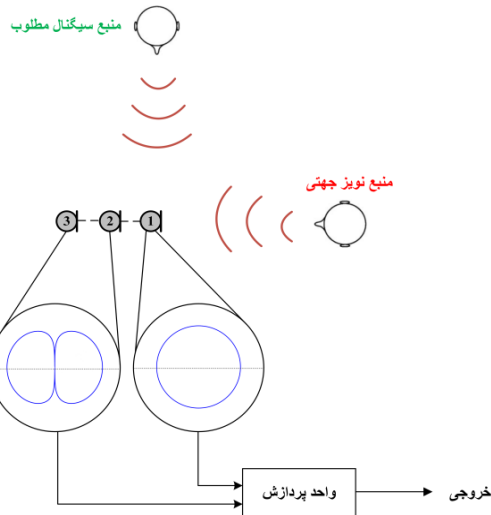
این شرط فقط برای مقادیر کوچک δ و فرکانس‌های پایین برقرار می‌شود. با این شرط همپوشانی فضایی 2 - که باعث تولید لوب‌هایی شبیه لوب اصلی در راستای سیگنال مطلوب می‌شود- خنثی می‌شود.

اندازه DMA است. یک DMA مرتبه N ام پاسخی متناسب با ترکیب خطی سیگنال‌های بدست‌آمده از مشتقات فضایی مرتبه صفر تا N میدان صوتی دارد. DMA مرتبه N از $N+1$ حسگر تشکیل شده و خروجی آن از ترکیب تفریق‌های خروجی دو DMA مرتبه $N-1$ بدست می‌آید [4,5]. آرایه‌های میکروفونی تفاضلی مزایا و معایبی نسبت به آرایه‌های جمعی دارند که در ادامه به اجمال به بعضی از آنها اشاره می‌کنیم [5]: مزایا: الف) DMAها نسبت به سایر آرایه‌ها اندازه کوچک‌تری دارند و این باعث می‌شود که بتوان از آنها در ابزارهای مختلف مخابراتی استفاده کرد. ب) DMAها دارای شکل‌الگوهای تغییرناپذیر با فرکانس بوده و به همین دلیل برای پردازش باند وسیع سیگنال‌های گفتار مناسب هستند. ج) DMAها نه فقط در فرکانس‌های بالا بلکه در فرکانس‌های پایین هم مناسب هستند. د) برای هر تعداد میکروفون، DMAها توانایی دارند به حداکثر بهره جهتی دسترسی پیدا کنند. معایب: الف) در حضور نویز سفید پاسخ فرکانسی DMA مرتبه N ام رفتاری شبیه یک فیلتر بالاگذر با شیب 6Ndb/oct دارد، بنابراین باید پاسخ فرکانسی آرایه برای پردازش باند وسیع سیگنال گفتار به طور صحیح جبران شود. ب) پاسخ فرکانسی و مرتبه DMA به مکان و جهت آرایه نسبت به منبع صدا، حساس است؛ از این رو باید پاسخ فرکانسی و مرتبه را مطابق با محدوده (مکان) منبع صوت و زاویه برخورد آن با آرایه متناسب کنیم. این فرآیند جبران‌سازی اگرچه ممکن است از لحاظ تئوری کار ساده‌ای باشد، اما عملاً برای DMAهای با مرتبه بالاتر از ۲ بسیار مشکل است. ج) DMAها در فرکانس‌های پایین از زیاد بودن نویز سفید رنج می‌برند؛ هرچند تحقیقات زیادی روی این مسأله انجام شده اما هنوز به موفقیت کامل نرسیده است.

در مقاله [6] یک روش جدید تفکیک گفتار در تلفن همراه، در محیط‌هایی که دو یا چند گوینده به صورت همزمان در حال صحبت کردن هستند ارائه شده است. روش پیشنهادی از دو لایه پردازشی تشکیل شده است. لایه اول شامل شکل‌دهنده‌های پرتو متفاوت و انتخاب کانال با توان حداقل بوده و در لایه دوم از روش تفریق طیفی، استفاده می‌شود. در ادامه دو شکل‌الگوی تشعشعی متفاوت برای آرایه در نظر گرفته شده و سیگنال‌های مورد نظر توسط آرایه در این دو حالت دریافت می‌شوند. با داشتن این دو سیگنال در لایه دوم و استفاده از روش تفریق طیفی، تفکیک منابع صورت می‌پذیرد.

در این مقاله، با توجه به کاربرد روزافزون آرایه‌های میکروفونی تفاضلی، به بهسازی گفتار در این آرایه‌ها با رویکرد حذف نویز جهتی با استفاده از روش‌های مبتنی بر فیلتر وینر می‌پردازیم.

این مقاله شامل بخش‌های زیر است. در بخش دوم مدل سیگنال مورد استفاده توضیح داده شده است. در بخش سوم ساختار کلی ترکیب آرایه‌های میکروفونی تفاضلی و فیلتر وینر ارائه می‌شود. در بخش چهارم عملکرد واحد پردازش مورد بررسی قرار می‌گیرد. در بخش پنجم ضمن بررسی کارایی روش پیشنهادی به مقایسه آن با روش پایه خواهیم پرداخت. بخش ششم نیز خلاصه و نتیجه‌گیری خواهد بود.



شکل ۳: ساختار استفاده شده برای حذف نویز جهتی به منظور بهسازی گفتار با استفاده از آرایه میکروفونی تفاضلی

۴- نحوه عملکرد واحد پردازش در ساختار پیشنهادی

در ساختار پیشنهادی، پس از دریافت سیگنال نویزی و اتمام پردازش‌های لازم توسط آرایه‌های تفاضلی بر روی این سیگنال، یکی از دو روش تفریق طیفی و فیلتر وینر تک کاناله به عنوان هسته اصلی واحد پردازش، روی سیگنال خروجی آرایه تفاضلی اعمال می‌شود. در ادامه به معرفی اجمالی این دو روش می‌پردازیم.

۴-۱- تفریق طیفی

تفریق طیفی یکی از اولین روش‌های ارائه شده برای کاهش نویز است. توصیف‌های متنوع و متفاوتی از این روش وجود دارد که اساس همه آنها یک قاعده ساده است. با فرض نویز جمع می‌توان تخمینی از طیف سیگنال تمیز را با تفریق طیف نویز تخمین زده شده از طیف سیگنال نویزی بدست آورد. طیف نویز در خلال زمان‌هایی که سیگنال تمیز حضور ندارد تخمین زده شده و بروز رسانی می‌شود. با فرض این که نویز ایستاد، یا تغییرات آن بسیار آهسته باشد و این که طیف نویز بین دوره‌های بروز رسانی تغییرات زیادی نداشته باشد. سیگنال گفتار بهسازی شده با محاسبه عکس تبدیل فوریه طیف سیگنال تخمین زده شده و با استفاده از فاز سیگنال نویزی بدست می‌آید. این روش بدلیل نیاز به تنها تبدیل فوریه گسسته و عکس آن بسیار ساده است. البته به منظور جلوگیری از اعوجاج گفتار فرآیند تفریق باید خیلی دقیق انجام شود. در صورتی که تفریق دقیق نباشد ممکن است بعضی از اطلاعات گفتار حذف و گفتار دچار اعوجاج شود و یا این که نویز در خروجی باقی بماند [9]. بدین منظور از تفریق طیفی تعمیم یافته استفاده می‌شود که رابطه آن بصورت زیر است [10]:

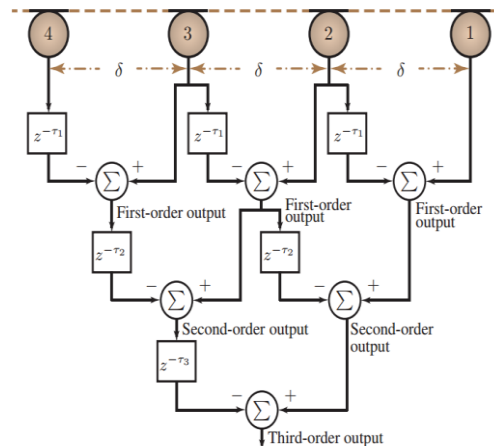
$$|\hat{X}(n, \omega)|^p = \max\{|Y(n, \omega)|^p - \beta E\{|V(\omega)|^p\}, \alpha E\{|V(\omega)|^p\}\} \quad (7)$$

که در این رابطه $\hat{X}(n, \omega)$ تخمین سیگنال تمیز، $Y(n, \omega)$ سیگنال نویزی و $V(\omega)$ سیگنال نویز است. p نمای قدرت، β عامل زیر تفریق

با توجه به شکل ۱ و با فرض میکروفون اول به عنوان مرجع و $\theta = 0^\circ$ (زاویه ورود سیگنال مطلوب به آرایه صفر باشد) می‌توان رابطه (۳) را به صورت زیر نوشت:

$$y(\omega) = [Y_1(\omega) Y_2(\omega) \dots Y_M(\omega)]^T \\ = \mathbf{d}^H(\omega, \cos 0^\circ) X(\omega) + \mathbf{v}_m(\omega) \quad (6)$$

شکل ۲ ساختار کلی DMA را نشان می‌دهد که در این مقاله از DMA مرتبه اول با دو میکروفون استفاده می‌شود.



شکل ۲: ساختار عمومی DMA مرتبه اول، دوم و سوم [1]

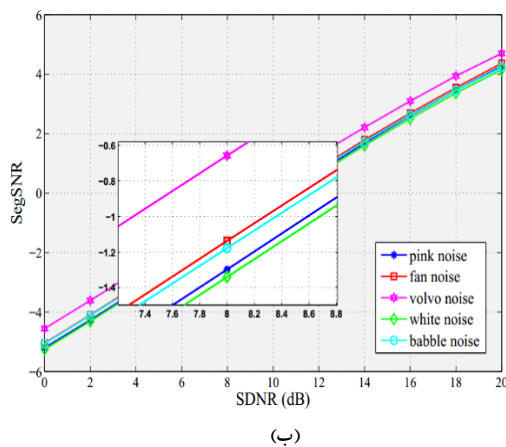
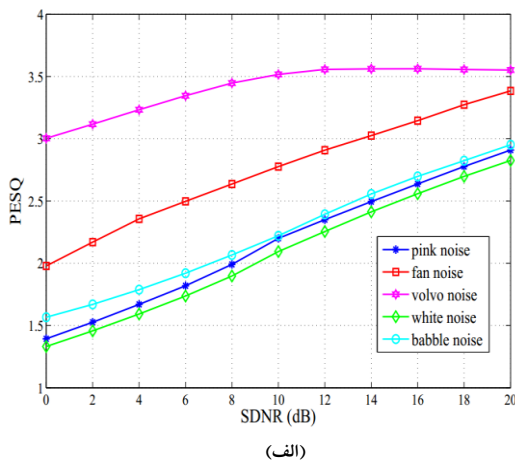
همان‌طور که در شکل ۲ مشخص است یک DMA مرتبه سه، از چهار میکروفون و از ترکیب دو DMA مرتبه دو تشکیل شده است.

۳- ترکیب آرایه‌های میکروفونی تفاضلی و فیلتر

وینر

در این بخش ساختار کلی پیشنهادی این مقاله برای بهسازی گفتار - که در آن از آرایه‌های تفاضلی استفاده شده - معرفی می‌شود. هدف از انتخاب این ساختار، سادگی، پیچیدگی کم و سهولت در پیاده‌سازی عملی بوده است. بدین منظور از ساختار سه میکروفونی پیشنهاد شده در مقاله [6] استفاده کردیم. شکل ۳ ساختار مذکور را نشان می‌دهد. این ساختار از یک تک میکروفون و یک آرایه تفاضلی دو میکروفونی که هر سه آنها در یک راستا قرار گرفته‌اند، تشکیل شده است. سیگنال مطلوب و نویز جهتی توسط تک حسگر همه‌جهته که شکل‌الگوی تشعشعی دایره‌ای شکل دارد دریافت می‌شوند. بدلیل همه‌جهته بودن تک حسگر، بهره سیگنال نویزی دریافتی توسط آن در همه راستاها - با نظر گرفتن فاصله منبع مطلوب یا نویز جهتی تا حسگر - یکسان است. DMA دو میکروفون نیز ترکیب سیگنال منبع مطلوب و سیگنال نویز جهتی را دریافت کرده، اما بدلیل دو قطبی بودن شکل‌الگوی تشعشعی آرایه تفاضلی، منبع سیگنال مطلوب که در زاویه 90° درجه نسبت به محور افقی قرار دارد حذف شده و خروجی آرایه تفاضلی سیگنال منبع نویز جهتی است که در راستای صفر درجه نسبت به محور آرایه قرار دارد. در نهایت خروجی تک حسگر و آرایه تفاضلی به واحد پردازش وارد می‌شوند. در واحد پردازش از روش‌های تفریق طیفی و فیلتر وینر استفاده می‌شود.

چون نویز ولوو، نویزی پایین‌گذر است و اکثر مؤلفه‌های طیفی آن در باند فرکانسی سیگنال‌گفتار حضور ندارند بنابراین سیگنال‌گفتار را کمتر تخریب می‌کند. بدلیل شباهت طیفی زیاد نویز سفید و صورتی تفاوت آن‌ها در تخریب سیگنال‌گفتار اندک است. همان‌طور که پیداست میزان حذف نویز جهتی سفید نسبت به بقیه نویزها در این روش کمتر بوده است. هرچند حذف نویز همهمه بدلیل طیف نسبتاً مشابه آن با گفتار کاری مشکل است، اما این روش در حذف آن موفق ظاهر شده است. روند افزایشی در نمودار SegSNR نیز قابل مشاهده است. دو نمودار نمایان‌گر قابل قبول بودن عملکرد این روش در حذف نویز جهتی می‌باشد.



شکل ۴: عملکرد روش تفریق طیفی (الف) PESQ (ب) SegSNR

جداول ۱ و ۲ مقادیر PESQ و SegSNR ورودی را به ازای نویزهای مختلف به نمایش گذاشته تا بتوان مقایسه بهتری با نمودارهای آورده شده انجام داد.

جدول ۳: مقادیر PESQ ورودی برای نویزهای مختلف

نویز	صورتی	فن	ولوو	سفید	همهمه
۰dB	۱,۳۶	۱,۹۳	۲,۹۷	۱,۳۰	۱,۵۵
۴dB	۱,۶۳	۲,۳۱	۳,۲۰	۱,۵۶	۱,۷۷
۸dB	۱,۹۴	۲,۶۰	۳,۴۲	۱,۸۵	۲,۰۴
۱۲dB	۲,۳۱	۲,۸۸	۳,۵۴	۲,۲۱	۲,۳۷
۱۶dB	۲,۶۰	۳,۱۱	۳,۵۶	۲,۵۲	۲,۶۸
۲۰dB	۲,۸۸	۳,۳۵	۳,۵۶	۲,۷۹	۲,۹۳

α ضریبی است برای تعیین میزان حداقل نویز که هیچ‌گاه نویز، منفی در نظر گرفته نشود. در سیستم پیشنهادی این تحقیق، طیف نویز مورد نیاز برای روش تفریق طیفی، مستقیماً از سیگنال نویز خالص که خروجی DMA است بدست می‌آید.

۴-۲- فیلتر وینر تک کاناله

فیلتر وینر یک تخمین‌زن خطی است که در پردازش گفتار به صورت‌های متفاوتی مورد استفاده قرار می‌گیرد. فیلتر وینر بهینه را می‌توان به صورت زیر بدست آورد [11]:

$$H_W(n, \omega) = \frac{E[|X(n, \omega)|^2]}{E[|Y(n, \omega)|^2]} = 1 - \frac{E[|V(n, \omega)|^2]}{E[|Y(n, \omega)|^2]} \quad (8)$$

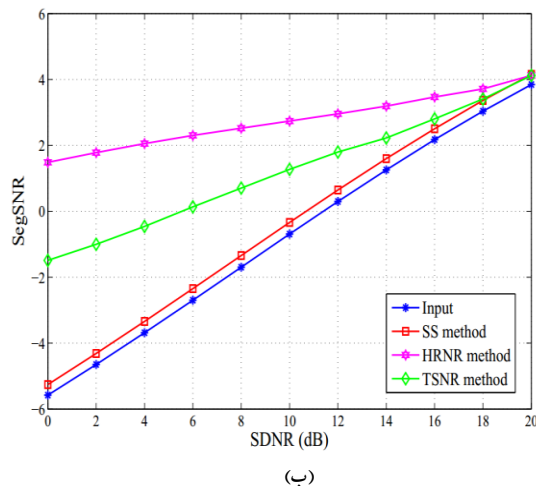
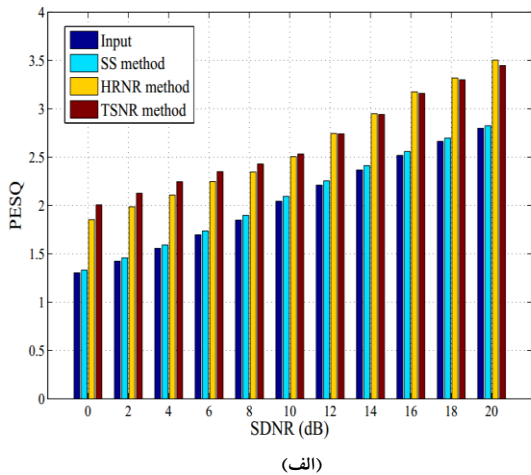
در تحقیق حاضر، از پیاده‌سازی فیلتر وینر در قالب روش‌های ترکیبی HRNR و TSNR بهره گرفته شده که توضیحات مربوط به این روش‌ها در مرجع [12] بیان گردیده است.

۵- ارزیابی عملکرد

در این‌جا محیط شبیه‌سازی یک اتاق با ابعاد $3 \times 5 \times 6$ متر (طول \times عرض \times ارتفاع) در نظر گرفته شده و آرایه در وسط اتاق قرار دارد. در شبیه‌سازی‌ها از سیگنال‌گفتار تمیز یک گوینده مرد به طول ۱۲ ثانیه به عنوان منبع سیگنال مطلوب استفاده شده است. سیگنال‌های دریافتی توسط میکروفون‌ها با فرکانس $16kHz$ نمونه‌برداری شده‌اند. پنجره استفاده شده، همینگ با طول $32ms$ ، و همپوشانی بین هر دو پنجره متوالی ۵۰ درصد در نظر گرفته شده است. در این آزمایش از فرض میدان دور برای منبع سیگنال مطلوب و منبع نویز جهتی استفاده می‌شود. در این فرض از بردار هدایت که نشان‌دهنده تأخیر سیگنال رسیده به هر میکروفون است استفاده می‌شود و اثر پژواک در این محیط نادیده گرفته می‌شود. فاصله بین حسگرها $21cm$ در نظر گرفته می‌شود (لازم به ذکر است با توجه به فرکانس نمونه‌برداری $16kHz$)، برای جلوگیری از ایجاد تأخیر کسری فاصله بین حسگرها در شبیه‌سازی، فاصله بین حسگرها $21.4375cm$ فرض شده است).

نویز جمع شونده با حسگرها شامل دو مؤلفه است: (۱) نویز جهتی که در راستای صفر درجه نسبت به محور آرایه قرار دارد و نسبت سیگنال به نویز آن بین $0dB$ و $20dB$ تغییر می‌کند. (۲) نویز خودی حسگر با نسبت سیگنال به نویز ثابت $35dB$.

در این آزمایش، هدف بررسی میزان حذف نویز جهتی در ساختار پیشنهادی است. به عبارتی در اینجا بررسی عملکرد این ساختار با استفاده از روش تفریق طیفی در حذف نویز جهتی مد نظر بوده است. شکل ۴ عملکرد ساختار سه‌میکروفونه را با استفاده از روش تفریق طیفی در حضور انواع نویزهای جهتی مختلف به نمایش می‌گذارد. شکل ۴ (الف) مقادیر PESQ بر حسب $SDNR^5$ و شکل ۴ (ب) مقادیر SegSNR را بر حسب $SDNR$ نشان می‌دهد. همان‌طور که مشخص است در $SDNR$ های بالا تفاوت مقادیر در نویز ولوو نسبت به سایر نویزها کمتر شده است. این مشاهده را می‌توان این‌گونه تحلیل نمود که



شکل ۵: مقایسه عملکرد روش‌های تفریق طیفی، HRNR و TSNR (الف) PESQ (ب) SegSNR

سپاسگزاری

نویسندگان مقاله بر خود لازم دانستند تا از زحمات و کمک‌های آقای سید مصطفی علویان شهری در این مقاله نهایت قدردانی به عمل آید.

مراجع

- [1] J. Benesty and J. Chen, *Study and Design of Differential Microphone Arrays*. Berlin, Germany: Springer-Verlag, 2012.
- [2] J. Benesty, J. Chen and I. Cohen, *Design of Circular Differential Microphone Arrays*. Berlin, Germany: Springer-Verlag, 2015.
- [3] J. Benesty, M. Souden and Y. Huang, "A Perspective on Differential Microphone Arrays in the Context of Noise Reduction," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 2, pp. 699–704, Feb. 2012.
- [4] L. Zhao, J. Benesty, and J. Chen, "Design of Robust Differential Microphone Arrays," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 10, pp. 1455-1466, Oct 2014.
- [5] J. Chen and J. Benesty, "A general approach to the design and implementation of linear differential microphone arrays," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific*, vol., no., pp.1-7, Oct. 29 2013-Nov. 2013.
- [6] S. Takada, S. Kanba, T. Ogawa, K. Akagiri, and T. Kobayashi, "Sound source separation using null-beamforming and spectral subtraction for mobile devices," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, pp.30–33, Oct 2007.

جدول ۲: مقادیر SegSNR ورودی برای نویزهای مختلف

نویز	صورتی	فن	ولوو	سفید	همهجه
dB۰	-۵,۴۷	-۵,۳۱	-۴,۷۴	-۵,۵۸	-۵,۲۵
dB۴	-۳,۵۸	-۳,۴۲	-۲,۸۵	-۳,۶۹	-۳,۳۶
dB۸	-۱,۶۰	-۱,۴۴	-۰,۸۸	-۱,۷۰	-۱,۳۹
dB۱۲	۰,۴۰	۰,۵۲	۱,۰۴	۰,۲۹	۰,۵۷
dB۱۶	۲,۲۷	۲,۳۵	۲,۸۰	۲,۱۷	۲,۳۹
dB۲۰	۳,۹۲	۳,۹۸	۴,۳۳	۳,۸۵	۳,۹۷

همان‌طور که بیان شد عملکرد روش تفریق طیفی در حذف نویز جهت سفید نسبت به سایر نویزها ضعیف‌تر بود. به همین دلیل، در آزمایش دوم، نویز سفید مبنای بررسی قرار گرفت. شکل ۵ مقادیر PESQ و SegSNR را برای مقایسه عملکرد ترکیب آرایه سه میکروفون و روش‌های HRNR، TSNR و تفریق طیفی نسبت به نویز جهت سفید نشان می‌دهد. همان‌طور که در شکل مشخص است عملکرد دو روش HRNR، TSNR بسیار بهتر از روش تفریق طیفی بوده و مقدار PESQ رشد قابل ملاحظه‌ای داشته است. روش TSNR در SDNRهای ۱۰ dB به پایین بهتر بوده و در SDNRهای ۱۰ dB به بالا روش HRNR بهتر عمل کرده است. شکل ۵ (ب) نشان می‌دهد که با افزایش SDNR مقدار SegSNR در همه روش‌ها تقریباً یک روند افزایشی ثابت را طی کرده و در انتها به دلیل کاهش شدید سطح نویز روش تفریق طیفی از روش‌های پیشنهادی به مقدار بسیار اندکی بهتر بوده است. شایان ذکر است که سیگنال نویز در روش فیلتر وینر نیز مانند روش تفریق طیفی از خروجی آرایه تفاضلی بدست می‌آید.

۶- نتیجه گیری

در این مقاله با استفاده از یک آرایه میکروفونی تفاضلی و تنظیم نول آن در راستای سیگنال مطلوب، منبع سیگنال مطلوب را حذف کرده و تخمینی از سیگنال نویز، بدست آوردیم. به موازات آن، مجموع سیگنال مطلوب و نویز توسط یک میکروفون دیگر دریافت می‌شود. از این دو سیگنال بدست آمده در الگوریتم‌های بهسازی گفتار استفاده می‌شود. ابتدا از تفریق طیفی به عنوان روش پایه در واحد پردازش استفاده کردیم. در بهبود این روش از الگوریتم‌های HRNR و TSNR در واحد پردازش استفاده شد. نتایج حاصله بیانگر بهبود قابل ملاحظه‌ای نسبت به روش پایه است به گونه‌ای که به طور میانگین در اکثر SNRها مقدار PESQ، ۰,۷ رشد داشته است.

- [7] G. W. Elko, J. E. West, and S. Thompson, "Differential and gradient microphone arrays," *J. Acoust. Soc. Am.*, vol.114, pp.2426–2426, Oct 2003.
- [8] G. W. Elko and J. Meyer, *Microphone Arrays*, chap. 50, Part I, pp.1021–1041. Eds. Berlin, Germany: Springer-Verlag, 2008.
- [9] P. C. Loizou. *Speech enhancement: theory and practice*. Boca Raton, FL: CRC Press, 2013.
- [10] S. V. Vaseghi, *Spectral Subtraction*, pp.242–260. 2nd ed. New York: Wiley, 2002.
- [11] J. Benesty, J. Chen, and E. A. Habets. *Speech enhancement in the STFT domain*. Berlin, Germany: Springer-Verlag, 2011, Springer Briefs in Electrical and Computer Engineering.
- [12] C. Plapous, C. Marro, and P. Scalart, "Improved signal-to-noise ratio estimation for speech enhancement," *IEEE Trans. Audio, Speech, and Lang. Process.*, vol.14, pp.2098–2108, Nov 2006.

زیر نویس‌ها

- ¹ Additive Microphone Array
- ² Spatial Aliasing
- ³ Under Subtraction Factor
- ⁴ Microphone Self Noise
- ⁵ Signal to Background Noise Ratio



فشرده سازی تصاویر رنگی مبتنی بر شبکه عصبی عمیق

علی کارگران^۱، هادی گرایلو^۲

^۱ دانشجوی کارشناسی ارشد دانشگاه صنعتی شاهرود
ali.kargaran@shahroodut.ac.ir

^۲ استادیار دانشکده مهندسی برق و رباتیک دانشگاه صنعتی شاهرود
grailu@shahroodut.ac.ir

چکیده

هر روز مقدار زیادی اطلاعات در سراسر جهان ذخیره، پردازش و انتقال داده می شود. شبکه های عصبی به عنوان یک راه حل برای انجام پردازش تصویر به سرعت در حال توسعه و گسترش می باشد. در این مقاله یک شبکه عصبی عمیق (DNN) و فیلتر حذف نویز برای فشرده سازی تصاویر رنگی ارائه شده است. ابتدا یک شبکه عصبی عمیق با الگوریتم فشرده سازی اتوانکودر برای فشرده سازی تصاویر رنگی در این مقاله ارائه شده است. همچنین نتایج تجربی نشان می دهد نه تنها شبکه عصبی عمیق کیفیت تصاویر بازسازی شده را بهبود بخشیده، بلکه محاسبات پیچیده و زمان آموزش کاهش یافته است. دوم از یک کانال حذف نویز برای از بین بردن نویزهای ناشی از فشرده سازی بکار گرفته شده است که موجب بهبود کیفیت تصویر بازسازی و همچنین پارامترهای ریاضی شده است. در روش فشرده سازی پیشنهادی در این مقاله، ابتدا تصویر ورودی توسط لایه های رمزگذار میانی فشرده و سپس با لایه های رمزگشای میانی تصویر بازسازی می شود. که نتایج بدست آمده نشان از برتری و کارآمدی الگوریتم فشرده سازی روش پیشنهادی دارد.

کلمات کلیدی

فشرده سازی، شبکه عصبی عمیق، اتوانکودر، بازسازی تصویر.

های عصبی چند لایه در مقایسه با شبکه های عصبی تک لایه عملکرد بهتری دارند [۱-۳].

شبکه های عصبی برای حل مسائل فشرده سازی تصویر برای چندین دهه مورد استفاده قرار گرفته اند. یک مرور خوب از فشرده سازی تصویر با استفاده از شبکه های عصبی در [۴] ارائه شده است. محققان در [۵و۶] شبکه عصبی با یک لایه مخفی را برای فشرده سازی تصویر پیشنهاد داده اند. در [۷] نویسندگان یک مقایسه از توابع انتقال شبکه عصبی عمیق برای انجام فشرده سازی تصویر معرفی نموده اند. با این حال، آن ها به این نتیجه رسیدند که راه طولانی ای وجود دارد تا فشرده سازی با شبکه عصبی بتواند به خوبی روش های فشرده سازی بدون شبکه عصبی همانند (DCT) و (SPIHT) شود. تبدیل کسینوسی گسسته (DCT) الگوریتم پردازش تصویر بیشتر اغلب برای فشرده سازی تصویر در سالهای گذشته مورد استفاده قرار می گرفت. در [۸و۹] الگوریتم های DCT و DWT برای فشرده سازی و بازسازی

۱- مقدمه

روش های فشرده سازی و انتقال داده ها به طور موثر و ایمن در حال تبدیل شدن به کاربردی قابل توجه و تجارتي سودمند می باشد. تکنیک های فشرده سازی تصویر امکان ایجاد یک مصالحه بین کیفیت خوب تصاویر دیجیتال و به حداقل رساندن پهنای باند و نیازهای فضای ذخیره سازی را فراهم می کند. یک شبکه عصبی مصنوعی (ANNs) یک سیستم پردازش اطلاعات است که می تواند در فشرده سازی تصویر و کنترل خطای کانال (حذف نویز) که الهام گرفته شده از سیستم عصبی بیولوژیکی می باشد استفاده کرد. اخیراً، محققان به نتایجی دست یافته اند که برخی موفقیت ها با آموزش شبکه های عصبی عمیق (DNNs)، که شبکه عصبی مصنوعی با تعداد لایه های مخفی بسیاری بین لایه های ورودی و خروجی قرار دارد. تحقیقات پیشین نشان می دهد شبکه

تصویر بدون اتلاف عملکرد مدل‌های یادگیری عمیق به دست آمده است [۱۳]. برای فشرده سازی تصاویر با اتلاف Toderici و همکاران [۱۴] یک شبکه عصبی برگشتی (RNN) برای فشرده کردن تصویر $۳۲*۳۲$ ارائه داده اند. آنها بعد یک مجموعه ای از روش های فشرده سازی با دقت کامل برای رمزگذاری و رمزگشایی پیشرفته تصاویر معرفی نمودند [۱۵].

این روش ها یاد می گیرند اعوجاج تکنیک های فشرده سازی برای نرخ فشرده سازی داده شده را به حداقل برسانند. Balle و همکاران [۱۶] از نرمالسازی تقسیم کننده کلی (GDN) و جایگزینی گرد کردن کوانتیزه با افزودن نویز یکنواخت برای آرامش مداوم اعوجاج و کاهش میزان آنتروپی استفاده نموده اند. الگوریتم های رمزنگاری مبتنی بر بی نظمی راه های کارآمد برای توسعه رمزنگاری تصویر مناسب را پیشنهاد کرده اند. این الگوریتم ها به شرایط اولیه شان حساس هستند، هر تغییر کوچک می تواند پاسخ های بسیار متفاوتی را که کارایی طرح های رمزنگاری را تضمین می کند ارائه کند. نگاشت لجیستیکی یکی از آنهاست، که در بسیاری از زمینه های پردازش اطلاعات اهمیت زیادی دارد. چنین زمینه هایی شامل موارد زیر می شوند: زیست شناسی جمعیت، شیمی، رمزنگاری، ارتباطات و محیط زیست. پایداری و شاخه نگاشت لجیستیکی مانند شبکه های عصبی کوهن گروسبرگ با تاخیر [۱۷] و Neimkar Sacker با تاخیر [۱۸] مورد مطالعه قرار گرفته اند.

تصاویر مورد بررسی قرار گرفته است. در ضمن روش فشرده سازی SPIHT برای تصاویر تولید شده مصنوعی توسعه نمی یابد. همچنین تکنیک های مختلفی برای حل مسئله کنترل خطای کانال (حذف نویز) با استفاده از شبکه های عصبی مصنوعی پیشنهاد شده است، که این تکنیک ها را می توان به دو گروه تقسیم نمود. در گروه اول شبکه عصبی، تکنیک اصلاح خطا توانایی خود تست را ندارند. در [۱۰] شبکه های عصبی پیشنهاد شده بدون توانایی خود تست تنها می تواند تعدادی محدود از خطا را اصلاح کند. در گروه دوم، شبکه عصبی با توانایی خود تست برای انجام تکنیک های اصلاح خطا است. با این حال در این مقاله برای افزایش ظرفیت تصحیح خطا و قابلیت ذخیره سازی شبکه از فیلتر حذف نویز هم استفاده شده است. هدف این مقاله ایجاد الگوریتم اتوانکودر برای فشرده سازی و بازسازی تصویر می باشد.

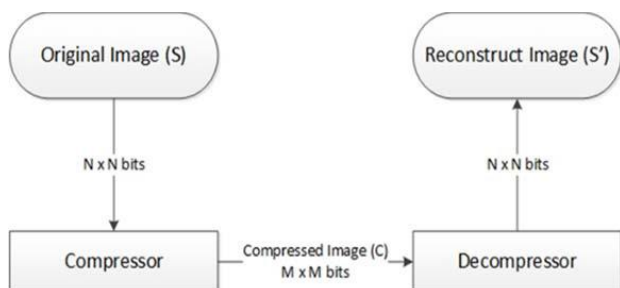
مدل های شبکه عمیق توسعه یافته مدل های شبکه عصبی برای یادگیری تبدیل غیرخطی روی داده ها هستند. این مدل ها در سال های اخیر مورد توجه فراوان قرار گرفته اند. در این مدل ها علاوه بر کدگذار در هر لایه یک کدگشا نیز وجود دارد. به عبارت دیگر تفاوت اساسی مدل های شبکه عمیق با شبکه عصبی این است که در هر لایه سعی می کنند قابلیت بازسازی داده را حفظ کنند.

ادامه مقاله بدین صورت است که در بخش ۲ فشرده سازی تصویر با شبکه عصبی عمیق ارائه شده است. بخش ۳ به تشریح جزئیات روش پیشنهادی پرداخته و بخش ۴ نتایج بررسی و ارزیابی روش پیشنهادی را ارائه خواهد نمود. و همچنین در بخش ۵ به نتیجه گیری پرداخته شده است.

۲- مروری بر روش های موجود

تصاویر دیجیتال بسیار متمرکز هستند و از این رو حجم زیادی از حافظه برای ذخیره سازی و زمان زیادی برای انتقال صرف می شود. با استفاده از تکنیک های فشرده سازی تصویر، ممکن است برخی از اطلاعات افزونگی موجود در تصاویر نیاز به فضای ذخیره سازی کمتر و زمان کمتری برای انتقال داشته باشد. شکل (۱) یک بلوک دیاگرام از فرآیند فشرده سازی تصویر با شبکه عصبی عمیق را نشان می دهد. سیستم فشرده سازی تصویر شامل یک ماژول فشرده سازی و یک ماژول بازسازی می باشد. S نشان دهنده تصویر اصلی، C نشان دهنده تصویر فشرده شده و S نشان دهنده تصویر بازسازی شده است.

برای استانداردهای تصویر موجود مانند JPEG و JPEG2000 کدک ها در عمل بطور جداگانه بهینه شده اند. در مرحله رمزگذاری آنها ابتدا یک تبدیل خطی را بر یک تصویر انجام می دهند. کوانتیزاسیون و رمزگذاری آنتروپی بدون اتلاف برای کم کردن نرخ فشرده سازی مورد استفاده قرار می گیرد [۱۱]. JPEG2000 از تجزیه موجک متعامد چندگانه برای تبدیل یک تصویر و رمزگذاری کدهای کوانتیزه شده با رمزگذاری بلوک جاسازی شده با قطع شدن بهینه استفاده می کند. در مرحله رمزگشایی، الگوریتم رمزگشایی تبدیل معکوس برای کاهش اعوجاج طراحی شده اند [۱۲]. اخیرا چندین مدل های فشرده سازی تصویر مبتنی بر یادگیری عمیق توسعه یافته اند. برای فشرده سازی



شکل ۱: بلوک دیاگرام فشرده سازی تصویر با شبکه عصبی

۲-۱- یادگیری عمیق در شبکه های عصبی

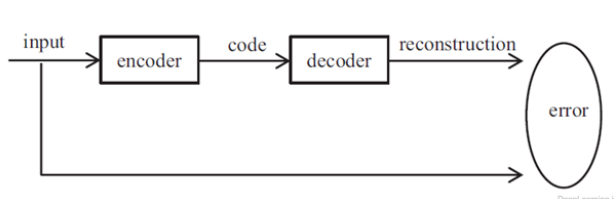
یادگیری عمیق یا به عبارت دیگر، یادگیری ساختار عمیق یا یادگیری سلسله مراتبی یک زیرشاخه از یادگیری ماشین است که اساس آن بر یادگیری نمایش دانش و ویژگی ها در لایه های مدل است. انگیزه ی اولیه این حوزه از طریقه بررسی ساختار عصبی در مغز انسان الهام گرفته شده است که در آن یاخته های عصبی با ارسال پیام به یکدیگر ادراک را امکان پذیر می کنند. بسته به فرض های متفاوت در مورد نحوه ی اتصال این یاخته های عصبی، مدل ها و ساختار های مختلفی در این حوزه پیشنهاد و بررسی شده اند، هرچند که این مدل ها به صورت طبیعی در مغز انسان وجود ندارد و مغز انسان پیچیدگی های بیشتری را دارا است. این مدل ها نظیر شبکه عصبی عمیق پیشرفت های خوبی را در حوزه های پردازش زبان های طبیعی، پردازش تصویر ایجاد کرده اند.

۳- روش پیشنهادی

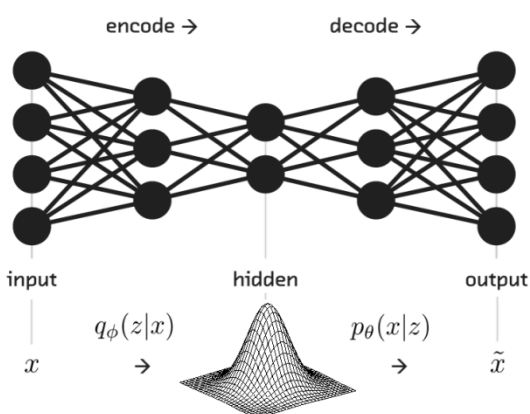
یکی از الگوریتم‌های کاربردی و جدید برای فشرده سازی تصویر اتوانکودر می باشد که در این مقاله پیشنهاد شده است. اتوانکودر نوعی خاصی از شبکه عصبی مصنوعی است که برای encode کردن بهینه یادگیری مورد استفاده قرار می گیرد [۱۹]. بجای آموزش شبکه و پیش بینی مقدار هدف y در ازای ورودی x ، یک اتوانکودر آموزش می بیند تا ورودی x خود را بازسازی کند. بنابراین بردارهای خروجی همان ابعاد بردار ورودی را خواهند داشت. فرآیند کلی یک اتوانکودر در شکل (۳) نشان داده شده است.

در حین این فرآیند اتوانکودر با کمینه سازی خطای نوسازی (reconstruction error) بهینه می شود. بطور کلی، یک لایه ی تنها، قادر به دریافت ویژگی های متمایز از داده خام نیست. محققان در حال حاضر از اتوانکودر عمیق استفاده میکنند که کد یاد گرفته شده از اتوانکودر قبلی را به اتوانکودر بعدی جهت به انجام رساندن کار خود ارسال میکنند. ساختار روش پیشنهادی مربوط به فشرده سازی و بازسازی تصویر در شکل (۴) نمایش داده شده است.

یک اتوانکودر اغلب با استفاده یکی از تعداد دیگری از پس انتشار آموزش می بیند. اگرچه اغلب عاملی موثر است اما مشکلات اساسی با استفاده از پس انتشار برای آموزش شبکه با لایه های میانی بسیاری وجود دارد. یکبار خطای پس انتشار در لایه های اول گرفته می شود که آنها مقدارشان کوچک و کاملاً بی اثر هستند.



شکل ۳: فرآیند کلی یک اتوانکودر



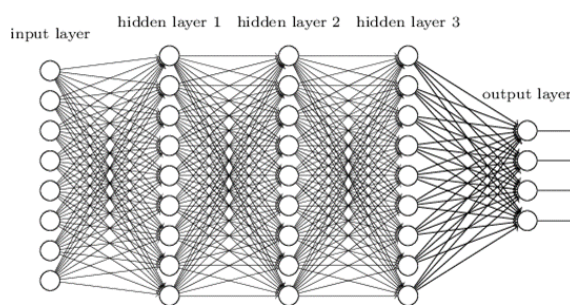
شکل ۴: ساختار روش پیشنهادی

هدف یک اتوانکودر، یادگیری بازنمایش فشرده شده ای برای یک مجموعه داده است. یعنی معمولاً اتوانکودر برای کاهش بعد، اندازه مورد استفاده قرار می گیرد. بر اساس تعریف [۲] اتوانکودر تابعی است که عمدتاً از دو بخش تشکیل می گردد: الف) انکدر: که یک تابع

مهندسی ویژگی (feature engineering) استخراج الگوهای مناسب از داده ها که به مدل‌های یادگیری ماشین کمک می نماید تا بهتر بتواند بین دسته ها یا کلاس های مختلف موجود تمایز قائل شود. مهندسی ویژگی مهمترین مهارتی هست که برای بدست آوردن نتایج خوب در کارهای مربوط به پیش بینی، فشرده سازی، تشخیص و... تو حوزه یادگیری ماشین نیاز است. الگوریتم های یادگیری ویژگی برای پیدا کردن و استخراج خودکار الگوهای مشترک استفاده می شوند تا از ویژگیهای استخراج شده در فرآیندهای دسته بندی و رگرسیون استفاده شوند. یادگیری ویژگی را می توان بعنوان همان مهندسی ویژگی ای دید که بصورت خودکار توسط یک الگوریتم انجام می شود.

در یادگیری ویژگی سلسله مراتبی ویژگی های غیرخطی چندین لایه را استخراج می نماییم و سپس آنها را به یک دسته بندی کننده می دهیم تا تمام ویژگی ها را ترکیب نموده و بتواند پیش بینی ای را انجام دهد. هرچقدر این سلسله مراتب لایه ها بیشتر (عمیقتر) باشد در نتیجه ویژگی های غیرخطی بیشتری هم بدست میاد برای همین هم هست که ما علاقه داریم از تعداد لایه های بیشتری در یادگیری عمیق استفاده کنیم، قسمت قبل از همین اصل نشأت گرفته و اشاره به سلسله مراتب عمیقی دارد که تو یادگیری ویژگیها به ما کمک می نماید. از طرف دیگر این ویژگی های پیچیده رو نمی توان بصورت مستقیم از تصویر ورودی بدست آورد. از لحاظ ریاضی بهترین ویژگی ای که از یک تصویر با یک لایه می توان بدست آورد فقط لبه ها (edge) و توده ها (blob) هست. دلیلش هم این می باشد که نهایت اطلاعاتی است که می توانیم از یک تبدیل غیرخطی از تصویر ورودی بدست آوریم. برای بدست آوردن و تولید ویژگی هایی که شامل اطلاعات بیشتری هستند نمی توان بصورت مستقیم روی تصویر ورودی کار کرد و لازمه برای این کار ویژگی های اولیه مثل لبه ها و توده ها را دوباره تبدیل کنیم تا ویژگی های پیچیده تری که شامل اطلاعات بیشتری برای تمایز بین کلاس ها مورد نیاز هست بدست آید.

همانطور که در شکل (۲) می بینید ساختار شبکه عصبی عمیق را به خوبی نشان می دهد، که نشان دهنده عمیق بودن شبکه می باشد.



شکل ۲: ساختار شبکه عصبی عمیق

هرچه تعداد لایه های میانی عمیق تر باشد، دقت و اطلاعاتی که در خروجی سیستم دریافت می شود بهتر و مناسب تر است. برای آموزش شبکه، تصاویر آموزشی ابتدا به یک ماتریس نرمالیزه شده پردازش شده و سپس به لایه ورودی شبکه عصبی عمیق اعمال می شود.

۴- نتایج شبیه سازی

گرچه روش پیشنهادی روی هر نوع تصاویری قابل اعمال است، اما جهت نمایش کارایی روش فشرده سازی روش پیشنهادی و نیز با توجه به تاثیر ماهیت تصویر روی این کارایی، تصاویر رنگی برای انجام این ارزیابی انتخاب شده اند. برای ارزیابی از سه نوع تصویر مختلف رنگی استفاده شده است. اندازه همه تصاویر تست شده برای ارزیابی ۵۱۲*۵۱۲ می باشد.

در جدول (۱) نتایج حاصل از فشرده سازی قبل از اعمال فیلتر حذف نویز نشان داده شده است.

جدول ۱: نتایج حاصل از الگوریتم قبل از اعمال فیلتر

عنوان تصویر	مقدار MSE	مقدار PSNR
lena	۹۵,۴۷۷۸	۱۷,۱۲۰۴
peppers	۴۲,۶۸۳۴	۱۴,۲۵۵۰
ball	۶۹,۵۷۱۳	۱۷,۷۱۹۸

در شکل (۵) به ترتیب تصویر اصلی و بازسازی شده را قبل اعمال فیلتر نشان می دهد.



شکل ۵ الف: تصویر اصلی peppers

استخراج ویژگی بوده و بردار ویژگی را از روی وردی ها محاسبه می نماید. بنابراین اگر بردار ویژگی را با h ، آنکدر را با f و مجموعه داده ها را x_i نمایش دهیم، رابطه (۱) برقرار خواهد بود:

$$h^{(t)} = f_{\theta}(x^{(t)}), \quad x^{(t)} = \{x^{(1)}, \dots, x^{(T)}\}$$

که h بردار ویژگی یا نمایش یا کد محاسبه شده از x می باشد.

ب) دیکدر: تابعی است که آن را با g نمایش می دهند و با استفاده از رابطه (۲) نگاشتی از فضای ویژگی به فضای ورودی انجام می دهد. رابطه (۲):

$$r = g_{\theta}(h)$$

در واقع از یک تابع احتمال مشخص، مدل های احتمالی تعریف شده و برای حداکثر نمودن (اغلب بصورت تقریبی) تشابه داده ها، آموزش داده می شوند. اتوانکودرها با استفاده از اصول آموزشی متفاوتی، آموزش داده می شوند. مجموعه پارامترهای آنکدر و دیکدر بطور مشابه برای بازسازی ورودی اصلی یلد گرفته می شوند. یعنی تلاش می شود خطای بازسازی $L(x, r)$ ، تا حد امکان حداقل شود. این کار با بازسازی r از روی x با پیش آموزش و اندازه گیری اختلاف میان x و r انجام می شود. به طور خلاصه آموزش اتوانکودر شامل پیدا نمودن بردار پارامتر تا برای حداقل نمودن خطای بازسازی می باشد:

رابطه (۳):

$$\mathcal{J}_{AE}(\theta) = \sum_t L(x^{(t)}, g_{\theta}(f_{\theta}(x^{(t)})))$$

به حداقل رساندن این مقدار نیز معمولا به روش گرادینان نزولی تصادفی مشابه روش آموزش پرسپترون چند لایه انجام می شود. عمومی سازی خوب در اینجا به معنی خطای بازسازی کوچک برای نمونه های تست است در حال که این خطا برای سایر پیکربندی های دیگر x حداکثر باشد. در [۲۰] ساختار جدیدی از اتوانکودر چند لایه مبتنی بر ساختار متقارن، نسبت به اتوانکودر مرسوم، برای کاهش بعد استفاده شده است. این ساختار جدید تعداد وزن های لازم برای تنظیم را کاهش داده و در نتیجه توانسته است هزینه محاسباتی را متعاقبا کاهش دهد. همچنین تعمیم کارایی را بهبود داده است.

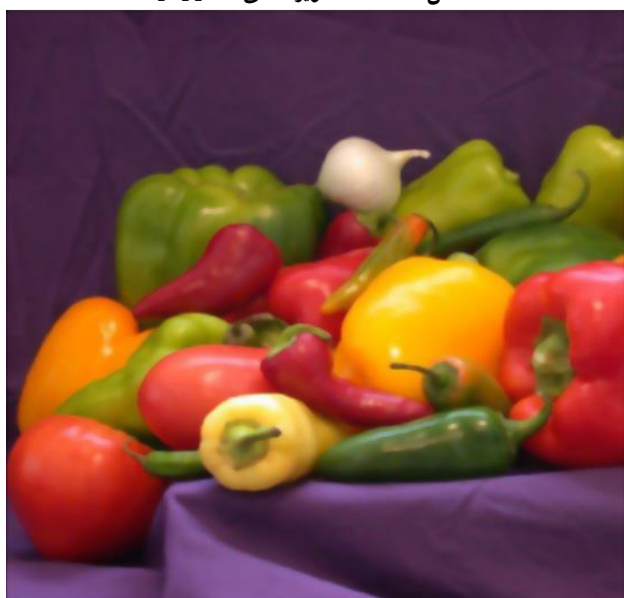
روشی که در این مقاله ارائه شده است بدین گونه است، ابتدا تصویر ورودی خوانده شده، سپس توسط لایه های میانی کدگذاری شده و تصویر فشرده می شود و سپس تصویر فشرده شده توسط لایه های میانی کدگشایی شده و تصویر اصلی بازسازی می شود. با توجه به نتایج بدست آمده، روش پیشنهادی روش موثر و کارآمدی می باشد. روشی پیشنهادی در این مقاله به دلیل ساختاری ساده ولی توانمند و کارا توانسته عملکرد خوبی را داشته باشد. همچنین در ادامه برای از بین بردن نویز از فیلتر حذف نویز برای فشرده سازی تصویر استفاده شده است، که با این تکنیک توانستیم پارامترهای فشرده سازی و بازسازی را بهبود بخشیم.



شکل ۶ الف: تصویر اصلی peppers



شکل ۵ ب: تصویر بازسازی شده peppers



شکل ۶ ب: تصویر بازسازی شده peppers

همانطور که در شکل (۵ ب) می‌بینید تصویر بازسازی شده دارای نویز می‌باشد که برای رفع این نقص از فیلتر حذف نویز استفاده شده است تا کارایی و همچنین کیفیت دیداری تصویر بازسازی شده بهتر شود. در جدول (۲) نتایج حاصل از الگوریتم فشرده سازی بعد از اعمال فیلتر حذف نویز نشان داده شده است.

جدول ۲: نتایج حاصل از الگوریتم بعد از اعمال فیلتر

عنوان تصویر	مقدار MSE	مقدار PSNR
lena	۵۰,۵۲۸۰	۱۹,۷۹۹۰
peppers	۲۶,۶۳۷۷	۱۶,۳۰۲۶
ball	۵۹,۱۵	۱۸,۴۲۴۳

همانطور که در جدول بالا مشاهده می‌شود با اضافه نمودن فیلتر حذف نویز پرامترهای مربوطه بهبود پیدا کرده است. همچنین کیفیت تصویر بازسازی شده از لحاظ دیداری مناسب تر شده است.

در شکل (۶) به ترتیب تصویر اصلی و بازسازی شده را بعد از اعمال فیلتر نشان داده شده است. یکی از نکات مثبت در روش پیشنهادی کاهش زمان آموزش می‌باشد، چون اغلب یکی از معایب در فشرده سازی تصویر با شبکه عصبی، زمان زیاد آموزش بوده است، که در الگوریتم روش پیشنهادی مطرح شده این موضوع برطرف شده است.

۵- نتیجه گیری

با توجه به میزان فشرده سازی در افزایش کیفیت تصویر، الگوریتم اتوانکودر برای فشرده سازی تصویر رنگی مبتنی بر شبکه عصبی عمیق در این مقاله ارائه شده است. نتایج بدست آمده نشان از عملکرد مناسب و کارآمد روش پیشنهادی دارد. یکی از نقاط ضعف روش پیشنهادی وجود نویز در تصویر بازسازی شده بود، که برای از بین بردن این ضعف از فیلتر حذف نویز استفاده شده است. که در جدول (۲) نتایج حاصل از الگوریتم بعد از اعمال فیلتر درستی این کار را نشان داده است. همچنین در روش پیشنهادی ارائه شده زمان آموزش شبکه نسبتاً کاهش یافته است که می‌تواند نکته قوتی در روش مطرح شده باشد. با توجه به گسترده و وسعت در شبکه های عصبی عمیق در آینده شاهد توسعه این حوزه در تمام حوزه ها و بخصوص در فشرده سازی تصویر و پیش بینی خواهیم بود.

مراجع

- [16]- J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-end optimized image compression," arXiv preprint arXiv:1611.01704, 2016.
- [17]- R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich "feature hierarchies for accurate object detection and semantic segmentation" In CVPR, pages 580–587, 2014.
- [18]- K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition" In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 770–778, 2016.
- [19]- Liou C-Y, Cheng W-C, Liou J-W, Liou D-R, "Autoencoder for words" Neurocomputing. 2014;139:84-96
- [20]- Alex K, Sutskever I, Geoffrey EH, "ImageNet Classification with Deep Convolutional Neural Networks". 2012:1097–105.
- [1]- Rafael C. Gonzalez, Richard E. Woods. "Digital Image Processing". Pearson Education. Press 2003.
- [2]- David MacKay. "Information Theory, Inference, and Learning Algorithm". Cambridge Univ. Press 2003.
- [3]- G. Hinton, L. Deng, et al., "Deep neural networks for acoustic modeling in speech recognition," IEEE Signal Processing Magazine, vol. 29(6), pp. 82-97, 2012.
- [4]- Vaddella, V., Rama, K.: 'Artificial neural networks for compression of digital images: a review', Int. J. Rev. Comput., 2010, 3, pp. 75–82.
- [5]- Dimililer, K.: 'Backpropagation neural network implementation for medical image compression', J. Appl. Math., 2013.
- [6]- Patel, B., Agrawal, S.: 'Image compression techniques using artificial neural network', Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET), 2013, 2, (10), pp. 2725–2729.
- [7]- F. Hussain and J. Jeong, "Exploiting deep neural networks for digital image compression," Web Applications and Networking (WSWAN), 2015 2nd World Symposium on, Sousse, 2015, pp. 1-6.
- [8]- Benchikh, Salam; Corinthios, Michael, "A hybrid image compression technique based on DWT and DCT transforms," in Advanced Infocom Technology 2011 (ICAIT 2011), International Conference on , vol., no., pp.1-8, 11-14 July 2011.
- [9]- Lakshmi, A.; Rakshit, S., "Gaussian Restoration pyramid: Application of image restoration to Laplacian pyramid compression," in Advance Computing Conference (IACC), 2010 IEEE 2nd International, vol., no., pp.66-71, 19-20 Feb. 2010.
- [10]- Minnix, J.I., "Fault tolerance of the backpropagation neural network trained on noisy inputs," in Neural Networks, 1992. IJCNN., International Joint Conference on , vol.1, no., pp.847-852 vol.1, 7-11 Jun 1992.
- [11]- G. K. Wallace, "The jpeg still picture compression standard" IEEE transactions on consumer electronics, 38(1):xviii–xxxiv, 1992.
- [12]- A. Skodras, C. Christopoulos, and T. Ebrahimi, "The jpeg 2000 still image compression standard" IEEE Signal processing magazine, 18(5):36–58, 2001.
- [13]- A. v. d. Oord, N. Kalchbrenner, and K. Kavukcuoglu, "Pixel recurrent neural networks" arXiv preprint arXiv:1601.06759, 2016.
- [14]- G. Toderici, S. M. O'Malley, S. J. Hwang, D. Vincent, D. Minnen, S. Baluja, M. Covell, and R. Sukthankar, "Variable rate image compression with recurrent neural networks" arXiv preprint arXiv:1511.06085, 2015.
- [15]- G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, and M. Covell, "Full resolution image compression with recurrent neural networks," arXiv preprint arXiv:1608.05148, 2016.



ارائه مجموعه‌ای از ویژگی‌های آماری جهت تعیین انسجام متن در سیستم‌های پرسش و پاسخ تعاملی

محمد مهدی حسینی^۱، مرتضی زاهدی^۲

^۱ دانشجوی دکترا، دانشکده کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود، hosseini_mm@shahroodut.ac.ir

^۲ استادیار، دانشکده کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود، zahedi@ganjineh.co.ir

چکیده

یکی از معیارهایی که در زمینه ارزیابی سیستم‌های پرسش و پاسخ تعاملی دارای اهمیت می‌باشد اندازه‌گیری اتوماتیک میزان انسجام متن خروجی سیستم از روی سوال‌ها و پاسخ‌های رد و بدل شده، بین کاربر و سیستم است. در این مقاله یک راه‌حل اتوماتیک برای تعیین میزان انسجام متن تولید شده، ارائه شده است. در روش پیشنهادی، مجموعه‌ای از ویژگی‌های آماری که بر اساس N -گرم‌های مختلف و بزرگترین رشته مشترک تعریف شده، پیشنهاد شده است که نتایج حاصل حاکی از دقت بالای ویژگی‌های پیشنهادی در تعیین انسجام متن می‌باشد. برای ایجاد پایگاه داده از مکالمات رد و بدل شده بین کاربران و سیستم‌ها، از چهار سیستم پرسش و پاسخ تعاملی موجود استفاده گردید. بر اساس خروجی سیستم‌ها، تعداد ۵۴۰ نمونه به عنوان داده مناسب در نظر گرفته شد تا مجموعه تست و آموزش بر اساس آن ایجاد گردید. سپس پیش پردازش بر روی مکالمات صورت پذیرفت و در مرحله بعد استخراج ویژگی صورت پذیرفت و بر اساس آن ماتریس ویژگی تشکیل گردید. در نهایت با استفاده از دسته‌بندی کننده‌های مختلف به طبقه‌بندی متن‌ها پرداخته شد که دسته‌بندی کننده SVM بر اساس معیار F با 0.81 درصد به عنوان بهترین دسته‌بند کننده انتخاب گردید.

کلمات کلیدی

ارزیابی، سیستم پرسش و پاسخ تعاملی، بردار ماشین پشتیبان، ویژگی آماری.

خاصی از اطلاعات سند به عنوان پاسخ بازگردانده می‌شود. بنابراین پاسخ ارائه شده ممکن است یک کلمه، یک جمله یا یک پاراگراف باشد. یک سیستم QA از سه بخش اصلی پردازش پرسش، تحلیل متن و پردازش پاسخ تشکیل می‌شود. فقدان تعامل دو طرفه بین سیستم و کاربر یکی از مهمترین مشکلات سیستم‌های QA محسوب می‌شود. با توجه به ساختار این سیستم‌ها، زمانیکه پرسش کاربر دارای ابهام بوده، یا اینکه پاسخ سیستم مطلوب کاربر نبوده یا کاربر نیازمند دریافت اطلاعات بیشتری باشد، راهکاری برای رفع این معضلات ارائه ننموده‌اند. که این مشکل با اضافه شدن سطح تعامل در سیستم‌های پرسش و پاسخ تعاملی^۲ رفع شده است. سیستم‌های موجود در زمینه IQA می‌توانند با توجه به شرایط و کاربردهایشان در سه گروه مختلف شامل مدیریت محدودیت، QA ارتقاء یافته و سوالات متوالی قرارگیرند. پر

۱- مقدمه

سیستم پرسش و پاسخ^۱ به عنوان سیستمی با پتانسیل بالا شناخته می‌شود که کاربران را قادر می‌سازد تا به منابع علمی (از طریق پرسش با استفاده از زبان طبیعی) دسترسی داشته باشند و یک پاسخ مرتبط، مناسب و مختصر را دریافت کنند. با این حال، همچنان مشکلات چالش برانگیز فراوانی جهت مرتفع نمودن در این سیستم‌ها موجود است. سیستم‌های QA شکل پیچیده‌تر سیستم‌های بازیابی اطلاعات هستند که در این سیستم‌ها به جای ارائه‌ی کل سند، تنها بخش‌های

۲- سابقه تحقیق

همانطور که قبلاً ذکر شد بیشتر روش‌های ارزیابی موجود برای سیستم‌های IQA مبتنی بر روش‌های موجود برای سیستم‌های QA و سیستم‌های دیالوگ می‌باشد بنابراین ابتدا مروری بر روی این روش‌ها صورت پذیرفته سپس در مورد روش‌های اندازه‌گیری انسجام در متن صحبت خواهد شد. یکی از روش‌های ارزیابی مورد استفاده در سیستم‌های QA استفاده از مجموعه‌ای از سوالات و پاسخ‌ها به نام «مجموعه استاندارد طلایی» است. در این روش با استفاده از میزان منطبق بودن سیستم با این مجموعه استاندارد طلایی توانایی یک سیستم مورد سنجش قرار می‌گیرد. البته این روش برای سوالات پیچیده و مبهم هنوز تقویت نشده است. در ارزیابی سیستم‌های QA با استفاده از کاربران واقعی تحقیقات قابل توجهی وجود دارد. بیشتر ارزیابی‌های صورت پذیرفته در این حوزه توسط TREC^۳ انجام شده که اکثر این ارزیابی‌ها به جای اینکه مبتنی بر سیستم باشد بر اساس نظرات کاربران صورت گرفته است. همچنین بیشتر کارهای صورت گرفته در حوزه ارزیابی در زمینه استخراج پاسخ و نحوه تعامل و استفاده از آن انجام شده است. بیشتر روش‌های پیاده‌سازی شده در زمینه ارزیابی سیستم‌های QA از معیارهایی همانند MRR^۴، C@1، CWS^۵، K1 و... استفاده نمودند که هر کدام از این روش‌ها خود دارای نقاط ضعف بوده و قابلیت تعمیم به همه سیستم‌های مختلف QA را نداشتند. به عنوان مثال معیار MRR زمانی بکار گرفته می‌شد که سیستم برای پاسخ به سوال مطرح شده، چندین جواب را ارائه نماید. اما در سیستم‌هایی که در مجموعه داده‌های خود تنها یک پاسخ برای هر سوال ارائه می‌نمودند از روش ارزیابی C@1 استفاده می‌شد. بنابراین این یکی از معضلات استفاده از این روش‌ها در سیستم‌های IQA بود و از طرفی این معیارها بیشتر در جهت انتخاب پاسخ بکار گرفته می‌شدند و توانایی سیستم را در این راستا مورد ارزیابی قرار می‌دادند. سان [۱]، روشی را برای ارزیابی سیستم‌های IQA معرفی نمود که X-EVAL نامیده می‌شد. این روش ارزیابی به میزان مشارکت عوامل موثر در یک سیستم IQA برای رسیدن یک کاربر به نتایج نهایی می‌پرداخت. مطالعه صورت گرفته توسط آن‌ها، شامل گزارش تجربی از سه سیستم QA تعاملی و یک سیستم پایه بود. آن‌ها تاکید داشتند که هدف از این گزارش ارزیابی سیستم نیست بلکه این کار را به منظور ارزیابی اثر X-EVAL در تشخیص پاسخ صحیح در میان این چهار سیستم بکار گرفته‌اند. کوارترونی و ماناندهار [۲] روشی را ارائه نمودند که شامل یک ارزیابی کیفی از سیستم‌های پرسش و پاسخ تعاملی بود. آن‌ها در روش خود تعدادی پرسش مطرح کردند و از کاربران خواستند با دادن امتیازی بین یک (حداقل امتیاز) تا پنج (حداکثر امتیاز) کیفیت تعامل را اندازه‌گیری نمایند. منصور و حسن‌پور [۳] برای یک سیستم QA از دانش موجود در سوالاتی که قبلاً در این سیستم بین کاربران و سیستم رد و بدل شده بود برای پاسخ‌دهی به سوالات استفاده نمودند. آن‌ها با ارائه یک الگوریتم، مجموعه‌ای از قطعات کارآمد را ایجاد کردند تا با استفاده مجدد از این قطعات به بهبود بازدهی پاسخ به سوالات بعدی برای سیستم فراهم نمایند. کلی [۴] به ارزیابی عملکرد چهار سیستم IQA با کاربر واقعی در مقاله خود پرداخته است. آن‌ها به دنبال شناسایی

واضح است که وجود یک سیستم ارزیابی استاندارد می‌تواند نقش بسیار مهمی در ارتقای این سیستم‌ها ایفا نماید. با این وجود تقریباً هیچ روش استاندارد در زمینه ارزیابی سیستم‌های IQA طراحی نشده است و بیشتر روش‌های ارزیابی موجود بر مبنای روش‌های استفاده شده در QA و سیستم‌های دیالوگ می‌باشد. دلیل اینکه یک سیستم IQA از دو موجودیت سیستم و کاربر تشکیل شده است و ممکن است یک کاربر تحت تاثیر عوامل مختلفی با سیستم کار نماید لذا کار ارزیابی بسیار سخت و پیچیده است و به ناچار عملکرد بخش‌های از سیستم‌ها توسط ناظرین انسانی ارزیابی می‌شود. در حال حاضر روش‌های استاندارد وجود دارند که می‌توانند اطلاعات مربوط به عملکرد سیستم از قبیل زمان، دقت و یا بازیابی را با استفاده از آن‌ها به دست آورد اما همچنان برای ارزیابی دقیق‌تر، نیاز به شناسایی سهم سیستم و کاربران در عملکرد مطلوب یک سیستم IQA می‌باشد. بنابراین اکثر ارزیابی‌های موجود، از ارزیابی انسانی بهره می‌گیرند که در نتیجه، عملکرد یک سیستم از کاربری به کاربر دیگر متفاوت خواهد بود. از طرفی در ارزیابی سیستم‌های IQA، علاوه بر ارزیابی کمی از ارزیابی کیفی نیز استفاده می‌شود که نیازمند مشارکت کاربران در فرآیند ارزیابی برای تعیین میزان موفقیت تعامل بین سیستم و کاربر می‌باشد. بر اساس مطالعات صورت گرفته در فرآیند ارزیابی انسانی، پارامترهای مختلفی جهت ارزیابی مورد توجه قرار می‌گیرد. بنابراین برای جایگزینی یک مدل به جای انسان، نیازمند اندازه‌گیری اتوماتیک این ویژگی‌ها خواهیم بود که این مسئله خود یکی از چالش‌های روبروی محققین این حوزه می‌باشد. یکی از ویژگی‌های که در ارزیابی سیستم‌های IQA مورد توجه قرار می‌گیرد میزان انسجام در متن سوالات و پاسخ‌های رد و بدل شده بین کاربر و سیستم می‌باشد. یک متن غیرمنسجم می‌تواند خروجی سیستم‌های تولید کننده متن مانند خلاصه‌سازی، امتیازدهی خودکار مقالات، متن تولید شده توسط یک سیستم پرسش و پاسخ و یا حتی متن تولید شده توسط یک فرد اما با دانش نگارشی پایین باشد. رویکردهای کلی تعیین و ارزیابی انسجام به سه دسته کلی تقسیم می‌شود. دسته اول رویکردهایی هستند که ساختار هدفمند نامیده می‌شوند. در این رویکرد بیشتر بدنه اصلی و هدف نهایی متن مورد توجه قرار دارد. دسته دوم از این رویکردها به بخش بندی سخن‌ها می‌پردازند و بیشتر روش‌های ارائه شده در این دسته به انسجام ساختاری درون جمله‌ای می‌پردازند و گاهی این روش‌ها را مدل‌های ساختار زبانی نیز می‌نامند. دسته سوم، ساختار تمرکزی نامیده می‌شوند که بیشتر روش‌های این دسته تمرکز موضوعی بر روی موضوع اصلی متن را ارزیابی می‌کنند. بنابراین در این مقاله مجموعه‌ای از ویژگی‌های جدید آماری ارائه گردیده است تا بتوانیم نقش موثری در اندازه‌گیری اتوماتیک انسجام متن خروجی تولید شده در سیستم‌های IQA، از سوالات رد و بدل شده بین کاربر و سیستم داشته باشیم. ساختار مقاله پیشنهادی بدین صورت است که در بخش دوم مروری بر تحقیقات صورت گرفته در زمینه ارزیابی سیستم‌های QA، IQA پرداخته شده است. در بخش سوم مجموعه ویژگی‌های پیشنهاد شده، سیستم پایه تولید شده در آزمایشگاه و مدل دسته‌بندی تشریح خواهد شد. در بخش چهارم نتایج بدست آمده آورده شده است و در بخش آخر به نتیجه‌گیری و پیشنهادات پرداخته شده است.

پیشنهادی بر روی معرفی ویژگی‌های جدید برای تعیین میزان انسجام می‌باشد. شکل ۱ فلوچارت روش پیشنهادی را نمایش می‌دهد.

۳-۱- پیش‌پردازش

مراحل صورت گرفته جهت پیش پردازش اطلاعات در چهار گام صورت پذیرفت که این مراحل عبارتند از:

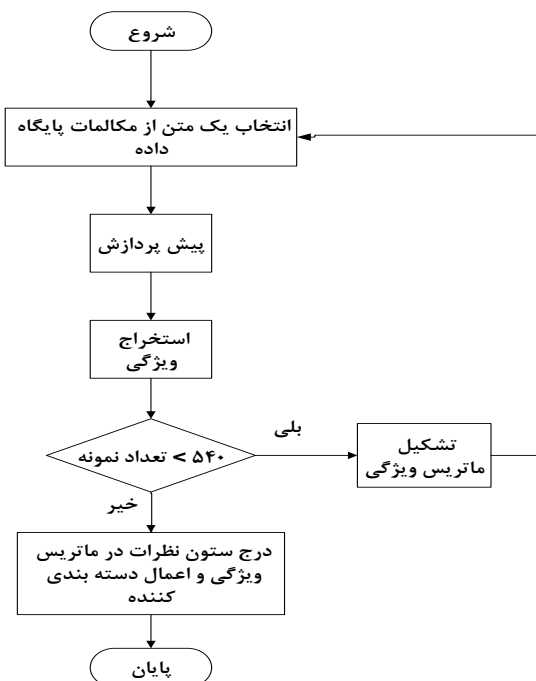
۱- مشخص کردن مرز جمله‌ها: برای مشخص نمودن مرزها از شمارش حروف و کلمه برای جداسازی استفاده کردیم.

۲- ریشه‌یابی: از روش مبتنی بر حذف پسوندها و پیشوندها برای تبدیل یک کلمه به شکل عمومی خودش استفاده گردید.

۳- حذف کلمات و واژه‌های غیر مهم: لیستی مشتمل بر ۲۰۰ کلمه بی تاثیر در متن آماده گردید و متناسب با آن‌ها از متن‌های ورودی حذف گردید.

۴- یکسان‌سازی متن‌ها: در متون انگلیسی تمامی کلماتی که با حروف بزرگ بودند به حروف کوچک تبدیل شدند و در متون فارسی یکسان‌سازی حروف صورت پذیرفت.

تمامی این کارها به صورت اتوماتیک صورت گرفت و جواب نهایی توسط ناظر انسانی کنترل شد.



شکل ۱: فلوچارت روش پیشنهادی

۳-۲- استخراج ویژگی

در این مرحله، تعدادی ویژگی بر اساس n -گرم‌ها و بزرگترین رشته مشترک^۶ و ویژگی‌های مرسوم دیگر در پردازش متن تعریف گردید که در ادامه به معرفی هر کدام از این ویژگی‌ها خواهیم پرداخت. اساس برخی از روابط استخراج شده برای ویژگی‌های تعریف شده در این مقاله، روابطی هستند که در زمینه ارزیابی اتوماتیک متن‌های خلاصه‌سازی

پتانسیل معیارهای ارزیابی برای سیستم‌های IQA، با استفاده از تجزیه و تحلیل نظرات ارزیابی ساخته شده توسط کاربران، برای چنین سیستم‌هایی بودند. آن‌ها در کار خود از داده‌هایی کیفی که از تحلیلگران اطلاعاتی در طول مصاحبه‌ها (که از یک کارگاه سه روزه ارزیابی فشرده از یک سیستم تعاملی پاسخ به پرسش کیفیت بالا گردآوری شده بود) جمع‌آوری نموده بودند بهره جستند. گینودو و استرودو [۵] روشی مبتنی بر ترکیب تئوری گراف و مدل مبتنی بر نهاد ارائه دادند. در روش پیشنهاد شده تعاملات بین جملات و عوامل انسجامی موجود در آن‌ها به یک گراف دو قسمتی مدل می‌شدند. بخش اول گراف دو قسمتی ایجاد شده، حاوی جملات و بخش دوم شامل عوامل انسجامی استخراج شده از جملات بود. هر نود جمله به یک نود عامل ارتباط برقرار می‌کرد اگر و فقط اگر در جمله، عامل مورد نظر وجود می‌داشت. گراف مورد نظر جهتدار بود و همیشه جهت آن از بخش جملات به بخش‌ها از عوامل انسجامی بود. وزن هر ارتباط با توجه به نقش گرامری عامل مورد نظر در جمله تعیین می‌شدند. شیواکومار و سومیا [۶] روشی مبتنی بر خوشه بندی ارائه نمودند. در روش پیشنهادی عقیده آن‌ها بر این بود که با انتخاب جملات مرتبط به هم و قرار دادن آن‌ها درون خوشه‌هایی جداگانه، بزرگترین خوشه به عنوان موضوع غالب متن تشخیص داده خواهد شد. لذا با قرار دادن جملات موجود در آن خوشه در متن خلاصه مرتبط‌ترین جملات انتخاب و منسجم‌ترین خلاصه تولید خواهد شد. بعدها از خوشه بندی در ترکیب با سایر الگوریتم‌ها برای ایجاد خلاصه‌های منسجم استفاده شد. یکی از مهمترین این ترکیب‌ها استفاده همزمان خوشه بندی و ماشین‌های بردار پشتیبان بود. پوجا و همکارانش از رویکردهای مبتنی بر زنجیره‌های لغوی در مقاله خود استفاده کردند. در این روش ایده بر این است که ابتدا کلمات مرتبط به هم استخراج و در یک زنجیره لغوی قرار گیرند سپس با استفاده از این زنجیره مرتبط‌ترین جملات انتخاب و در متن خلاصه قرار گیرند. در نتیجه خلاصه ایجاد شده شامل مرتبط ترین و منسجم‌ترین جملات خواهد بود. جانارا و همکارانش [۷] روشی برای تولید خلاصه‌هایی منسجم در خلاصه سازی چند سندی معرفی کردند. این روش موجب شد که انسجام خلاصه تولید شده در هر دو حوزه انتخاب جملات مبهم و مرتبط و همچنین چینش صحیح جملات به دنبال هم رعایت شود. یکی از بزرگترین چالش‌ها در خلاصه‌های چند سندی قرار دادن جملات انتخاب شده در مکانی صحیح در متن خروجی است. گراف G-Flow انسجام جملات انتخابی را با وزنی که به یالهای هر گراف داده است تخمین زده و در نهایت جملات با ارتباطی قوی‌تر را انتخاب و در متن خلاصه قرار می‌دهد. در روش پیشنهاد شده توسطی و ژانگ با بکارگیری مدلی شناختی جملات مهمتر متن تشخیص داده شده و خلاصه‌های منسجم‌تری تولید گردیده است. در این روش با استفاده از مجموعه متن‌های روایی و تئوری شناختی، ارتباطات بین جملات کشف و مهمترین آنان استخراج گردیده است.

۳-۳ روش پیشنهادی

در این مقاله مجموعه‌ای از ویژگی‌های جدید برای تعیین اتوماتیک انسجام متن خروجی سیستم‌های IQA پیشنهاد شده است. اساس روش

$$F_{LCS} = \frac{(1 + \beta^2)R_{LCS}P_{LCS}}{R_{LCS} + \beta^2P_{LCS}} \quad (5)$$

که در آن $\beta = \frac{P_{LCS}}{R_{LCS}}$ و M تعداد جفت سوال-پاسخ هر مکالمه می-باشد.

- ویژگی چهارم:

در این ویژگی، ما بزرگترین زیر رشته مشترک بین هر سوال و مجموعه جواب‌ها را یافته و درون یک مجموعه قرار داده، سپس در بین همه اعضای این مجموعه بزرگترین زیر رشته حاصل شده را انتخاب می-نماییم بر این اساس روابط به صورت زیر تعریف گردید:

$$R_{LCS} = \frac{1}{N} \times \sum_{i=1}^N \max \left(\sum_{j=1}^P \frac{LCS(Q_i, A_j)}{L_{Q_i}} \right) \quad (6)$$

$$P_{LCS} = \frac{1}{N} \times \sum_{i=1}^N \max \left(\sum_{j=1}^P \frac{LCS(Q_i, A_j)}{L_{A_j}} \right) \quad (7)$$

$$F_{LCS} = \frac{(1 + \beta^2)R_{LCS}P_{LCS}}{R_{LCS} + \beta^2P_{LCS}} \quad (8)$$

که در این رابطه مقدار $\beta = 1$ در نظر گرفته شد.

- ویژگی پنجم:

در این ویژگی پیشنهاد شده، n-گرم‌های مشترک بین مجموعه سوالات و پاسخ‌ها را بدست آورده و بعد از نرمال‌سازی به عنوان ارزش یک مکالمه گزارش می‌کنیم. در این ویژگی، با فرض اینکه ما دو مجموعه از سوالات و جواب‌ها داریم براساس روابط زیر میزان امتیاز هر Q_i با مجموعه جواب‌ها محاسبه و در نهایت توسط رابطه ۱۱ امتیاز مکالمه را محاسبه می‌نماییم. اینکار برای $n=3, 2$ که بر روی مجموعه داده ما دارای نتایج بهتری بود انجام گردید.

$$R_{skip_n} = \frac{1}{t} \times \frac{1}{k} \times \sum_{i=1}^t \sum_{j=1}^k \frac{skip_n(Q_i, A_j)}{C(m, n)} \quad (9)$$

$$P_{skip_n} = \frac{1}{t} \times \frac{1}{k} \times \sum_{i=1}^t \sum_{j=1}^k \frac{skip_n(Q_i, A_j)}{C(L, n)} \quad (10)$$

$$F_{skip_n} = \frac{1 + \beta^2 \times R_{skip_n} \times P_{skip_n}}{R_{skip_n} + \beta^2 \times P_{skip_n}} \quad (11)$$

که در آن t تعداد سوالات، k تعداد پاسخ‌ها، n اندازه n-گرم، m طول سوال Q_i ، L طول پاسخ A_j و $B=1$ در نظر گرفته شد.

- ویژگی ششم (فاصله همینگ):

یکی دیگر از ویژگی‌هایی که پیشنهاد گردید استفاده از فاصله همینگ برای اندازه‌گیری تشابه بین جملات یک مکالمه بود. هر جمله در مکالمه، از تعدادی کلمه تشکیل شده است با بدست آوردن میزان شباهت بین دو کلمه، میزان شباهت بین جملات را محاسبه می‌کنیم. برای محاسبه شباهت عدد حاصل را بر طول کلمات تقسیم و از عدد یک کسر می‌نماییم. این کار را برای کل کلمات جمله انجام داده و در نهایت بر اساس اعداد بدست آمده میزان شباهت بین دو جمله محاسبه می‌شود. روابط زیر معیار شباهت بین دو کلمه و یک جمله را نشان می-دهند.

شده [۸] بکار رفته‌اند اما متناسب با کار زمینه ارزیابی سیستم‌های پرسش و پاسخ تعاملی این روابط بازنویسی و به روزرسانی شده‌اند. - ویژگی اول:

N-گرم‌ها یکی از مشهورترین مدل‌های آماری زبان هستند. در این مدل‌ها ارتباطات زنجیره کلمات در نظر گرفته می‌شود به عبارت دیگر، مدل‌های n-گرم براساس هم‌پیوندی و کنار هم قرار گرفتن کاراکترهای لغات در پردازش متن عمل می‌نمایند. این ویژگی برای $n=1, 2, 3$ و برای مجموعه‌های پرسش و پاسخ (Q-A)، مجموعه سوال‌ها (Q-Q) و مجموعه جواب‌ها (A-A) طبق رابطه زیر محاسبه گردید.

$$Count_N = \sum_{S_i \in conv} \frac{\sum_{ngram \in S_i} Count_{match}(gram_n)}{\sum_{ngram \in S_i} Count(gram_n)} \quad (1)$$

که در این رابطه S_i i-امین جمله از هر مکالمه و n طول هر n-گرم می‌باشد.

- ویژگی دوم:

در یک مکالمه هر چه برای n‌های بزرگتر، تعداد n-گرم‌های مشترک بیشتر باشد امتیاز آن مکالمه بیشتر خواهد بود و احتمال پیوستگی متن مکالمه بیشتر خواهد شد [۹]. بر این اساس در این ویژگی پیشنهادی، هر کدام از n-گرم‌ها، بر اساس ارزش یک ضریب وزنی برای هر n-گرم به ارزش W_i با یکدیگر جمع می‌شوند تا مقدار این ویژگی بدست آید (رابطه ۲).

$$Count_Weight_N = \frac{1}{M} \times \sum_{i=1}^M \frac{\sum_{ngram \in S_i} W_k \times Count_{match}(gram_n)}{\sum_{ngram \in S_i} Count(gram_n)} \quad (2)$$

که در آن M تعداد عضوهای مجموعه و W_k ضریب تاثیر هر n-گرم و مقدار آن متناسب با عدد n می‌باشد. این ویژگی نیز برای $n=1, 2, 3$ محاسبه گردید.

- ویژگی سوم:

یکی دیگر از ویژگی‌هایی که در زمینه ارزیابی متون خلاصه‌سازی شده پیشنهاد گردید، استفاده از بزرگترین زیر رشته مشترک بود. از این ویژگی در ترجمه ماشینی و خلاصه‌سازی استفاده شده است که نتایج مناسبی را ارائه کرده است [۸]. بنابراین معادلات استفاده شده به صورت روابط زیر بروزرسانی گردید. در رابطه تعریف شده، برای یک مکالمه ابتدا یک جفت سوال-پاسخ را در نظر گرفته، سپس برای هر جفت بازایی و دقت را محاسبه و برای تمامی جفت سوال-پاسخ اینکار را انجام می‌دهیم در نهایت پاسخ بدست آمده را در رابطه معیار F قرار داده و امتیاز هر مکالمه را محاسبه می‌کنیم.

$$R_{LCS} = \frac{1}{M} \times \sum_{i=1}^M \frac{LCS(Q_i, A_i)}{L_{Q_i}} \quad (3)$$

$$P_{LCS} = \frac{1}{M} \times \sum_{i=1}^M \frac{LCS(Q_i, A_i)}{L_{A_i}} \quad (4)$$

این روش‌ها می‌توان به الگوریتم‌های نزدیکترین همسایگی، دسته‌بندی کننده بیز، ماشین بردار پشتیبان، آدابوست و درخت تصمیم اشاره نمود. جهت پیاده‌سازی این الگوریتم‌ها از نرم افزار وکا استفاده گردید. همچنین این نکته قابل ذکر است که مجموعه آموزش و مجموعه تست هر یک به ترتیب به میزان ۸۵٪ و ۱۵٪ از مجموعه داده به صورت تصادفی انتخاب شدند.

۴- نتایج آزمایشات

همانطور که قبلاً اشاره شد به جهت انتخاب بهترین دسته‌بندی کننده چندین دسته‌بندی کننده مورد استفاده قرار گرفت که نتایج حاصل از SVM بسیار مناسب‌تر نسبت به دیگر دسته‌بندی کننده‌ها بود نتایج حاصل از دسته‌بندی کننده SVM در جدول ۱ آورده شده است.

۴-۱- معیار ارزیابی

بنابراین برای ارزیابی نتایج حاصل از دسته‌بندی بدست آمده با داده‌های واقعی معیار F مورد استفاده قرار گرفت. رابطه مربوط به این معیار در زیر نشان داده شده است.

$$F - measure = \frac{(\beta^2 + 1) \times P \times R}{R + \beta^2 \times P} \quad (19)$$

که $P = \frac{TP}{TP + FP}$ ، $R = \frac{TP}{TP + FN}$ و $\beta = \frac{1-\alpha}{\alpha}$ می‌باشد

۴-۲- پایگاه داده

بدلیل نبود پایگاه داده استاندارد در زمینه ارزیابی سیستم‌های IQA، نیاز به تشکیل یک پایگاه داده مناسب از سوالات رد و بدل شده بین سیستم و کاربر با برچسب گذاری مناسب بود. بر این اساس، علاوه بر سیستم تعاملی پایه طراحی شده، سه سیستم دیگر تعاملی موجود در نظر گرفته شد. تعداد ۱۲۰ کاربر برای ۵ موضوع مختلف با سیستم کار نمودند و با توجه به موضوع مکالمات هر یک به صورت جداگانه ذخیره گردید. از این مجموعه ۶۰۰ تایی، ۵۴۰ نمونه توسط فرد خبره به عنوان نمونه مناسب‌تر انتخاب شد. جدول ۲، یک نمونه از خروجی متن ذخیره شده (بدون فرمت) از تعامل با سیستم پایه را نشان می‌دهد.

جدول ۱: نتایج حاصل از دسته‌بندی کننده‌ها

معیار F	دسته بندی کننده
۰/۷۶	درخت تصمیم
۰/۸۰	بیز
۰/۶۸	آدابوست
۰/۷۲	نزدیکترین همسایه
۰/۸۱	ماشین بردار پشتیبان

همچنین جهت ارزیابی مناسب‌تر یک پایگاه داده از مکالمات انگلیسی روزمره با موضوعات مختلف شکل گرفت. تعداد مکالمات در نظر گرفته ۱۰۰۰ مکالمه با ۵ موضوع مختلف بود که همگی مکالمات دارای انسجام متنی بودند که مکالمات واقعی به عنوان مکالمات مناسب

$$Similarity_{words} = 1 - \frac{Hamming_Distance(A,B)}{Max(|A|,|B|)} \quad (12)$$

$$Similarity_{sen} = \frac{1}{C(m,n)} \times \sum_{j=1}^{C(m,n)} Similarity_{words_j} \quad (13)$$

که در آن m طول سوال و n طول پاسخ‌ها می‌باشد.

- ویژگی هفتم (Tf-Idf):

پارامتر TF-idf یکی از ویژگی‌های آماری است که در سیستم‌های بازیابی اطلاعات، بر اساس آن می‌توان میزان شباهت میان کلمات منتخب با یک سند را محاسبه کرد. در روش پیشنهادی، برای این ویژگی میزان تکرار یک کلمه در مکالمه در مقابل تعداد تکرار آن در کل مکالمات محاسبه می‌شود. سپس مجموع این امتیازات برای کلمات یک مکالمه محاسبه، نرمالیزه شده و به عنوان امتیاز یک مکالمه گزارش می‌شود رابطه زیر نحوه محاسبه وزن هر کلمه حاصل را نشان می‌دهد.

$$W_{word(i)} = F_w \times \log \frac{N_{sentence}}{NC_{sentence}} \quad (14)$$

که در آن F_w تعداد تکرار هر کلمه در جملات یک مکالمه، $N_{sentence}$ تعداد جملات یک مکالمه و $NC_{sentence}$ تعداد جملاتی است که شامل کلمه iam می‌باشد.

- ویژگی هشتم:

استفاده از میزان شباهت با استفاده از روابط ۱۵ تا ۱۸ یکی دیگر از ویژگی‌های پیشنهادی است. در این ویژگی، دو مجموعه S_1 و S_2 تعریف شد که مجموعه S_1 ، مجموعه N-گرم‌های مربوط به سوال و مجموعه S_2 ، مجموعه N-گرم‌های مربوط به جواب می‌باشد. بنابراین مقدار شباهت بین این دو مجموعه توسط رابطه ۱۸ به صورت جداگانه برای هر جفت پرسش و پاسخ محاسبه شد و در نهایت میانگین امتیاز بدست آمده برای هر مکالمه محاسبه و به عنوان یک ویژگی در ماتریس ویژگی‌ها ذخیره گردید. مقدار این ویژگی برای n-گرم، ۱ تا ۳ محاسبه گردید.

$$Likeness(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|} \quad (15)$$

ویژگی ۱۴، ۱۵، ۱۶ همانند ویژگی سیزدهم می‌باشد اما معیار در نظر گرفته شده توسط روابط زیر تعریف گردید.

$$Cosine_dis = \frac{|S_1 \cap S_2|}{\sqrt{|S_1| \times |S_2|}} \quad (16)$$

$$Containment(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_2|} \quad (17)$$

$$Overlap(S_1, S_2) = \frac{|S_2|}{|S_1|} \quad (18)$$

۳-۳- دسته‌بندی متن خروجی سیستم IQA

برای دسته‌بندی ویژگی‌های استخراج شده از متون ذخیره شده در پایگاه داده از روش‌های موجود در زمینه داده کاوی استفاده گردید که از جمله

- [7] Christensen, Janara, Stephen Soderland Mausam, Stephen Soderland, and Oren Etzioni. "Towards Coherent Multi-Document Summarization." In HLT-NAACL, pp. 1163-1173. 2013.
- [8] L. C. Yew. "Rouge: A package for automatic evaluation of summaries." In Text summarization branches out: Proceedings of the ACL-04 workshop, vol. 8. 2004.
- [9] R. Barzilay, M. Lapata, "Modeling local coherence: An entity-based approach", Computational Linguistics, Vol 34, pp. 1-34, March 2008.

زیرنویس‌ها

- ¹ Question Answering System (QAs)
² Interactive Question Answering system (IQA)
³ Text Retrieval Evaluation Conference
⁴ Mean Reciprocal Rank
⁵ Confidence Weighted Score
⁶ Longest Common Substring(LCS)

در نظر گرفته شد. سپس به صورت تصادفی بعضی از خطوط مکالمات با یکدیگر جابجا گردید و به عنوان مکالمه نامناسب در پایگاه داده ذخیره گردید تا روابط حاصل نیز بر روی این پایگاه داده مورد آزمون قرار گیرد.

جدول ۲: نمونه تولید شده از یک مکالمه در سیستم IQA

سیستم	کاربر
اگر معدل دانشجوی کمتر از ۱۴ شود چه باید کرد؟	دانشجو مربوط به چه مقطعی است؟
دانشجو کارشناسی ارشد	منظور شما معدل کل هست؟
خیر، معدل یک ترم	در صورت کمتر شدن معدل از ۱۴ در هر نیمسال دانشجو مشروط تلقی می‌گردد.
دانشجوی کارشناسی چطور؟	در صورت کمتر شدن معدل از ۱۲ در هر نیمسال دانشجو مشروط تلقی می‌گردد.
چند ترم متوالی مشروط شدن دانشجو اخراج می‌شود؟	دانشجو کارشناسی یا ارشد؟
دانشجو کارشناسی	در صورتیکه دانشجو کارشناسی سه ترم متوالی یا چهار ترم غیر متوالی مشروط شود در آستانه اخراج قرار می‌گیرد.

۵- نتیجه‌گیری

در این مقاله مجموعه‌ای از ویژگی‌های آماری اتوماتیک برای تعیین انسجام بر اساس متن خروجی یک سیستم IQA ارائه شد که در هیچ یک از کارهای قبلی این حوزه مشاهده نگردید. سپس از روی متن‌های موجود در پایگاه داده ویژگی‌ها استخراج و براساس آن ماتریس ویژگی تشکیل گردید. برای دسته‌بندی نظرات از دسته‌بندی کننده SVM بهره گرفته شد. نتایج ارائه شده برای ارزیابی دسته‌بندی کننده بر اساس معیارهای ارزیابی F حاکی از موفقیت ویژگی‌های پیشنهادی در تعیین انسجام بود.

مراجع

- [1] S. Ying, P. B. Kantor and E. L. Morse. "Using cross-evaluation to evaluate interactive QA systems." Journal of the Association for Information Science and Technology 62, no. 9: 1653-1665, 2011.
- [2] Quarteroni, Silvia and S. Manandhar. "Designing an interactive open-domain question answering system." Natural Language Engineering 15, no. 1: 73-95, 2009.
- [3] M. Mansoori, and H. Hassanpour. "Boosting passage retrieval through reuse in question answering." International Journal of Engineering 25, no. 3:187-196, 2012.
- [4] Kelly, Diane, P. B. Kantor, E. L. Morse, J. Scholtz, and Y. Sun. "Questionnaires for eliciting evaluation data from users of interactive question answering systems." Natural Language Engineering 15, no. 1: 119-141, 2009.
- [5] C. Guinaudeau, M. Strube, "Graph-based Local Coherence Modeling", Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, pp. 93-103, 2013.
- [6] ShivaKumar, K. M., and R. Soumya. "Text Summarization using Clustering Technique and SVM Technique." International Journal of Applied Engineering Research 10, no. 12, PP.28873-28881, 2015.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی تبریز

یادگیری متریک هیستوگرامی \mathcal{H} با تنظیم‌کننده dropout

حمید صادقی^۱، ابوالقاسم-اسدالله راعی^۲

^۱ دانشکده مهندسی برق، دانشگاه صنعتی امیرکبیر، تهران
hamid.sadeghi@aut.ac.ir

^۲ دانشکده مهندسی برق، دانشگاه صنعتی امیرکبیر، تهران
raie@aut.ac.ir

چکیده

یادگیری متریک و ارائه یک معیار فاصله مناسب می‌تواند راه‌گشای بسیاری از مسائل بازشناسی الگو، بینایی ماشین و غیره باشد. در این مقاله، یک روش یادگیری متریک جدید برای استفاده در کاربردهای بینایی ماشین ارائه می‌شود. بسیاری از ویژگی‌هایی که در بینایی ماشین مورد استفاده قرار می‌گیرند، ماهیت هیستوگرامی دارند؛ حال آنکه روش‌های موجود یادگیری متریک اغلب بر پایه فاصله ماله‌لانوبیس توسعه داده شده‌اند که این معیار متناسب با فضای غیرمتعامد هیستوگرامی نیست. برای حل این نقطه ضعف اساسی، نوآوری این مقاله، طراحی یک روش یادگیری متریک جدید بر اساس معیار فاصله مربع کای (\mathcal{H}) است. فاصله \mathcal{H} یکی از معروف‌ترین و پرکاربردترین معیارهای مقایسه ویژگی‌های هیستوگرامی است. در روش پیشنهادی، ابتدا معیار \mathcal{H} به نحو مناسبی برای انجام یادگیری اصلاح شده است؛ سپس یک تابع هزینه محذب برای بهینه‌سازی این رابطه بر اساس دسته‌بندی‌کننده KNN تعریف شده است. از دیگر نوآوری‌های این مقاله، استفاده از یک روش dropout جدید در یادگیری متریک برای جلوگیری از بیش‌تعلیم بر روی مجموعه آموزشی و افزایش دقت بهینه‌سازی بوده است. نتایج آزمایش‌ها بر روی چهار پایگاه داده هیستوگرامی شناخته شده webcam، amazon، caltech و نشان می‌دهد که الگوریتم پیشنهادی در مقایسه با سایر روش‌های موجود دقت بسیار بالاتری در دسته‌بندی داده‌های هیستوگرامی دارد.

کلمات کلیدی

یادگیری متریک، فاصله \mathcal{H} ، فاصله هیستوگرامی، دسته‌بندی هیستوگرام، dropout.

مختلف در مسائلی همچون خوشه‌بندی و دسته‌بندی داده‌ها بوده است. در این مقاله یک روش یادگیری متریک جدید برای استفاده در داده‌های هیستوگرامی پیشنهاد شده است.

هدف اصلی یادگیری متریک، بهبود معیار مقایسه بین داده‌ها بر اساس یک مجموعه آموزشی است [2]. بسیاری از روش‌های یادگیری متریک، سعی در بهبود معیار فاصله ماله‌لانوبیس^۲ دارند. فاصله ماله‌لانوبیس بین دو بردار D بعدی \vec{x}_i و \vec{x}_j به صورت رابطه (۱) تعریف می‌شود.

$$D_A(\vec{x}_i, \vec{x}_j) = (\vec{x}_i - \vec{x}_j)^T A (\vec{x}_i - \vec{x}_j) \quad (1)$$

۱- مقدمه

یادگیری متریک و سنجش مناسب فاصله داده‌ها یکی از مسائل بنیادین در یادگیری ماشین است. به عنوان مثال، استفاده از یک معیار مقایسه مناسب، در مسائلی که در آنها تنها یک داده تعلیم برای هر دسته^۱ وجود دارد (مانند شناسایی چهره بر اساس یک تصویر) و نیز مسائلی که داده‌های تعلیم و تابع هدف بسیار پیچیده دارند، می‌تواند راه‌حلی بسیار مناسب باشد. در دهه گذشته تحقیقات بسیاری در زمینه یادگیری متریک انجام شده است [1,2]؛ این توجه محققین به دلیل کاربردهای

ایجاد می‌کند. برای بهینه‌سازی تابع هزینه پیشنهادی نیز از یک الگوریتم مبتنی بر کاهش گرادیان تصادفی^{۱۰} (SGD) استفاده شده است.

یکی از چالش‌های یادگیری متریک، تعداد بالای پارامترهای یادگیری است که به همراه تعداد کم داده آموزشی، باعث به وجود آمدن بیش‌تعلیم بر روی مجموعه آموزشی می‌شود. یکی از راه‌های پیشگیری از بیش‌تعلیم در یادگیری متریک، استفاده از تنظیم‌کننده است. نوآوری دیگر این مقاله استفاده از یک تکنیک dropout جدید به‌عنوان تنظیم‌کننده است. روش dropout در اصل برای مقابله با بیش‌تعلیم در آموزش شبکه‌های عصبی عمیق معرفی شده است [14] و در رگرسیون [15, 16] و برای نخستین بار در [17] برای مسأله یادگیری متریک نیز مورد استفاده قرار گرفته است. در این مقاله نیز یک رابطه dropout جدید برای یادگیری متریک فاصله^۲ پیشنهاد شده است و نشان داده می‌شود که این روش از نظر ریاضی، معادل با استفاده از یک تنظیم‌کننده نسبتاً پیچیده در حین یادگیری است.

ادامه مطالب موجود در این مقاله بدین شرح است: در بخش بعدی، روش پیشنهادی یادگیری متریک شرح داده می‌شود. در بخش سوم به شرح پایگاه‌های داده و آزمایش‌های انجام شده پرداخته می‌شود. نهایتاً بخش چهارم مربوط به جمع‌بندی این مقاله است.

۲- روش پیشنهادی

در این بخش، ابتدا روش یادگیری متریک LMNN به‌صورت مختصر معرفی می‌شود؛ در ادامه روش یادگیری متریک پیشنهادی براساس فاصله^۲ بیان می‌شود. سپس به شرح تکنیک dropout به‌عنوان تنظیم‌کننده در الگوریتم یادگیری پرداخته می‌شود.

۲-۱- یادگیری متریک LMNN

فرض کنید $X = \{(\vec{x}_i, y_i) \mid i = 1, 2, \dots, N\}$ مجموعه آموزشی از N داده باشد که در آن \vec{x}_i نشان‌دهنده داده i ام با برچسب y_i از فضای D بعدی $(\vec{x}_i \in \mathbb{R}^D)$ باشد؛ با فرض اینکه $A = L^T L$ باشد، می‌توان فاصله ماحالانویس بین دو بردار \vec{x}_i و \vec{x}_j (رابطه (۱)) را به‌صورت رابطه (۲) نوشت.

$$D_A(\vec{x}_i, \vec{x}_j) = (\vec{x}_i - \vec{x}_j)^T L^T L (\vec{x}_i - \vec{x}_j) = \left\| L(\vec{x}_i - \vec{x}_j) \right\|^2 \quad (2)$$

که در آن، L ماتریسی است که با یادگیری بر روی مجموعه آموزشی به‌دست می‌آید. روش LMNN به‌منظور یادگیری فاصله ماحالانویس در دسته‌بندی‌کننده KNN ارائه شده است [10]. در این روش، یادگیری L با دو هدف انجام می‌شود: (الف) K همسایه هر داده دارای برچسب مشابه با آن داده باشند و (ب) داده‌های نامشابه با یک حاشیه بزرگ قابل تفکیک باشند.

در LMNN، برای داده آموزشی \vec{x}_i ، تعداد K نزدیک‌ترین همسایه (براساس فاصله اقلیدسی) با برچسب مشابه، به‌عنوان داده‌های هدف^{۱۱} در نظر گرفته می‌شود [10]. در این مقاله، برای نمایش اینکه داده \vec{x}_j جزء داده‌های هدف \vec{x}_i است، از نشانه‌گذاری $i \sim j$ استفاده می‌شود. برای موفقیت‌آمیز بودن دسته‌بندی KNN، لازم است که در

که در آن، A یک ماتریس مربعی با ابعاد $D \times D$ است و با یادگیری بر روی مجموعه آموزشی به‌نحوی به‌دست می‌آید که قابلیت تفکیک‌پذیری دسته‌های مختلف را بهبود دهد.

اولین تلاش انجام شده در یادگیری فاصله ماحالانویس، بر اساس یک فرمول‌نویسی محدب و بدون تنظیم‌کننده انجام شده است [3]. پس از آن تلاش‌های دیگری تا به امروز در این زمینه انجام شده است [4-10]. در [4] روشی بر اساس رابطه $A = M^T W M$ ارائه شده است که در آن W ماتریس یادگیری به صورت قطری و ماتریس M ثابت و معلوم است. با این تعریف، ماتریس A مثبت نیمه معین^۳ (PSD) بوده و نیازی به اعمال قید PSD بودن ماتریس یادگیری نیست. نقطه ضعف این روش، تعیین دستی ماتریس M و نیز قابلیت تعمیم‌پذیری کمتر آن نسبت به یادگیری فاصله ماحالانویس متداول است؛ چراکه در این روش، ماتریس W قطری بوده و در فرآیند یادگیری تنها یک وزن‌دهی ویژگی‌ها آموزش داده می‌شود و ارتباط بین ویژگی‌ها در نظر گرفته نمی‌شود.

یکی دیگر از تحقیقات مهم در زمینه یادگیری فاصله ماحالانویس ITML^۴ [5] است؛ این روش برای نخستین بار تنظیم‌کننده LogDeg divergence (که در بسیاری از تحقیقات دیگر همچون [6] مورد استفاده قرار گرفته است) را معرفی نموده است. از محدودیت‌های این روش، مقداردهی دستی اولیه در ماتریس‌های یادگیری است که کیفیت یادگیری را تحت تأثیر قرار می‌دهد.

روش LMNN^۵، یکی دیگر از تحقیقات پر استناد در یادگیری متریک است [10]. این روش براساس دسته‌بندی‌کننده K -نزدیک‌ترین همسایگی^۶ (KNN) و فاصله ماحالانویس طراحی شده است. در سال‌های بعد، روش LMNN برای استفاده در داده‌های هیستوگرامی براساس فاصله مربع کای^۷ (χ^2) توسعه داده شد [11]. می‌توان گفت این روش، معروف‌ترین الگوریتم یادگیری متریک برای فاصله^۲ است؛ اما از نقاط ضعف آن می‌توان به عدم استفاده از تنظیم‌کننده^۸ در فرآیند آموزش و وقوع بیش‌تعلیم^۹ بر روی داده‌های آموزشی و همچنین محدب نبودن تابع هزینه تعریف شده اشاره کرد.

معیار فاصله ماحالانویس متناسب با فضای متعامد تعریف شده است؛ اما ویژگی‌هایی که در مسائل بینایی ماشین و پردازش تصویر به‌کار می‌روند، معمولاً داده‌های هیستوگرامی هستند. استفاده از معیار فاصله ماحالانویس در فضای غیرمتعامد هیستوگرامی کارایی چندانی نخواهد داشت. فاصله^۲ شناخته‌شده‌ترین معیار برای مقایسه داده‌های هیستوگرامی است و در کاربردهای مختلف همچون شناسایی چهره نتایج مناسبی داشته است [12, 13]؛ بنابراین، در این مقاله روش جدیدی برای یادگیری متریک براساس فاصله^۲ پیشنهاد می‌شود.

در روش پیشنهادی این مقاله، رابطه ریاضی فاصله^۲ به نحو مناسبی برای یادگیری متریک اصلاح می‌شود. رابطه پیشنهادی، همچون معیار ماحالانویس، دارای ماتریسی است که با یادگیری بر روی مجموعه آموزشی به‌دست می‌آید. در ادامه، یک تابع هزینه محدب برای یادگیری این ماتریس ارائه شده است. این تابع هزینه براساس بهبود دسته‌بندی‌کننده KNN طراحی شده است. محدب بودن تابع هزینه پیشنهادی، برتری قابل توجهی را نسبت به روش LMNN برای روش پیشنهادی

همان‌طور که ملاحظه می‌شود، روابط به‌دست آمده مشابه با روابط یادگیری متریک ماهالانوبیس هستند. به عبارت دیگر، یادگیری متریک بر اساس معیار فاصله χ^2 با یک تغییر متغیر به یک مسئله یادگیری متریک ماهالانوبیس تبدیل شده است. بر اساس تابع هزینه‌ای که برای LMNN در رابطه (۳) بیان شد، تابع هزینه پیشنهادی برای یادگیری فاصله χ^2 (رابطه (۸)) به صورت رابطه (۹) تعریف شده است.

$$\ell(L, X) = (1 - \mu) \sum_{i,j-i} \|\mathbf{L}(\bar{x}_{ioj} - \bar{x}_{joi})\|^2 + \mu \sum_{i,j-i; L: y_i \neq y_j} [1 + \|\mathbf{L}(\bar{x}_{ioj} - \bar{x}_{joi})\|^2 - \|\mathbf{L}(\bar{x}_{ioi} - \bar{x}_{ioi})\|^2]_+ \quad (9)$$

که در این رابطه، μ مانند رابطه (۳) بوده و سایر نمادها نیز پیش‌تر معرفی شده‌اند. برای کمینه‌سازی این تابع هزینه می‌توان از روش کاهش گرادین استفاده کرد؛ اما این تابع در عناصر ماتریس تبدیل خطی L محدب نیست [10]. بنابراین نتیجه نهایی وابستگی بالایی به مقادیر اولیه ماتریس L داشته و به کمینه‌های موضعی می‌رسد. برای حل این مشکل، رابطه (۹) می‌تواند به صورت یک مسئله semidefinite programming (SDP) [18] بازنویسی شود. در یک مسئله SDP، بهینه‌سازی با PSD بودن ماتریس انجام می‌شود. برای این کار، بر اساس تعریف $A = L^T L$ و رابطه (۷)، بهینه‌سازی به شکل رابطه (۱۰) و با قید PSD بودن ماتریس A انجام شده است. رابطه (۱۰) یک مسئله بهینه‌سازی محدب بوده و گرادین آن قابل محاسبه است.

$$\ell(A, X) = (1 - \mu) \sum_{i,j-i} \chi_A^2(\bar{x}_i, \bar{x}_j) + \mu \sum_{i,j-i; L: y_i \neq y_j} [1 + \chi_A^2(\bar{x}_i, \bar{x}_j) - \chi_A^2(\bar{x}_i, \bar{x}_i)]_+, A \succeq 0 \quad (10)$$

حال باید رابطه (۱۰) را به یک شکل استاندارد مسئله SDP نوشت. برای این کار متغیر slack (ξ_{ij}) برای \bar{x}_i و داده هدف $j \sim i$ و فریبکار \bar{x}_i آن تعریف شده است [10]. این متغیر نشان می‌دهد که داده فریبکار \bar{x}_i تا چه اندازه وارد محیط پیرامون \bar{x}_i با داده هدف $j \sim i$ و حاشیه امنیت آن شده است. با این تعریف، مسئله بهینه‌سازی برای یادگیری متریک χ^2 به صورت رابطه (۱۱) تعریف شده است.

$$\min (1 - \mu) \sum_{i,j-i} (\bar{x}_{ioj} - \bar{x}_{joi})^T A (\bar{x}_{ioj} - \bar{x}_{joi}) + \mu \sum_{i,j-i; L: y_i \neq y_j} \xi_{ij} \quad (11) \begin{cases} (\bar{x}_{ioi} - \bar{x}_{ioi})^T A (\bar{x}_{ioi} - \bar{x}_{ioi}) - (\bar{x}_{ioj} - \bar{x}_{joi})^T A (\bar{x}_{ioj} - \bar{x}_{joi}) \geq 1 - \xi_{ij} \\ \xi_{ij} \geq 0 \\ A \succeq 0 \end{cases}$$

۲-۳- بهینه‌سازی با تنظیم‌کننده dropout

در بخش پیش، یک مسئله بهینه‌سازی برای متریک χ^2 شرح داده شد. الگوریتم SGD، که به صورت کلی در الگوریتم ۱ نشان داده شده است، یکی از روش‌های پرکاربرد در حل مسائل بهینه‌سازی در یادگیری متریک است [1, 17, 19, 20]. در این بخش به معرفی روش بهینه‌سازی پیشنهادی براساس الگوریتم SGD و تنظیم‌کننده dropout پرداخته می‌شود.

روش dropout که در اصل برای تعلیم شبکه‌های عصبی عمیق معرفی شده است، برای یادگیری متریک نیز مورد استفاده قرار گرفته

بین K نزدیک‌ترین همسایه داده \bar{x}_j ، داده‌ای با برچسب نامشابه وجود نداشته باشد. بنابراین در روش LMNN، برای داده \bar{x}_i یک محیط در نظر گرفته می‌شود که در بردارنده داده‌های هدف به همراه یک حاشیه امنیت است. با این تعریف، داده‌های نامشابه‌ی که در این محیط قرار گیرند به عنوان داده‌های فریبکار^{۱۲} در نظر گرفته شده‌اند. در LMNN، باید تعداد داده‌های فریبکار پس از یادگیری، حداقل شود. تابع هزینه معرفی شده برای یادگیری LMNN به صورت رابطه (۳) است.

$$\ell(L, X) = (1 - \mu) \sum_{i,j-i} \|\mathbf{L}(\bar{x}_i - \bar{x}_j)\|^2 + \mu \sum_{i,j-i; L: y_i \neq y_j} [1 + \|\mathbf{L}(\bar{x}_i - \bar{x}_j)\|^2 - \|\mathbf{L}(\bar{x}_i - \bar{x}_i)\|^2]_+ \quad (3)$$

که در این رابطه، $[u]_+ = \max(u, 0)$ تابع هزینه لولایی^{۱۳} و μ ضریب ثابت تعیین‌کننده تأثیر عبارت اول و دوم است. عبارت اول این رابطه فواصل بزرگ بین \bar{x}_i و داده‌های هدف آن را جریمه می‌کند و عبارت دوم، داده‌های فریبکار را برای \bar{x}_i جریمه می‌کند. استفاده از تابع هزینه لولایی باعث می‌شود که داده‌های نامشابه‌ی که در فاصله‌ای دور از \bar{x}_i قرار دارند، در فرآیند آموزش بی‌تأثیر باشند.

۲-۲- یادگیری متریک پیشنهادی براساس فاصله χ^2

معیار فاصله χ^2 برای دو بردار هم‌سنگرامی \bar{x}_i و \bar{x}_j به صورت رابطه (۴) تعریف می‌شود.

$$\chi^2(\bar{x}_i, \bar{x}_j) = \sum_{b=1}^D \frac{(\bar{x}_i)_b - (\bar{x}_j)_b}{(\bar{x}_i)_b + (\bar{x}_j)_b} \quad (4)$$

که در آن، $(\bar{x}_i)_b$ نشان دهنده عنصر b ام از هم‌سنگرام \bar{x}_i است. رابطه (۴) می‌تواند به صورت ماتریسی به شکل رابطه (۵) بازنویسی شود.

$$\chi^2(\bar{x}_i, \bar{x}_j) = (\bar{x}_i - \bar{x}_j)^T (\text{diag}(\bar{x}_i + \bar{x}_j))^{-1} (\bar{x}_i - \bar{x}_j) = \left((\bar{x}_i - \bar{x}_j) \sqrt{(\text{diag}(\bar{x}_i + \bar{x}_j))^{-1}} \right)^T \left((\bar{x}_i - \bar{x}_j) \sqrt{(\text{diag}(\bar{x}_i + \bar{x}_j))^{-1}} \right) \quad (5)$$

که در آن، $\sqrt{\cdot}$ جذر عناصر ماتریس را محاسبه می‌کند. در این مقاله، رابطه (۶) برای یادگیری متریک در فاصله χ^2 پیشنهاد شده است.

$$\chi_A^2(\bar{x}_i, \bar{x}_j) = \left((\bar{x}_i - \bar{x}_j) \sqrt{(\text{diag}(\bar{x}_i + \bar{x}_j))^{-1}} \right)^T A \left((\bar{x}_i - \bar{x}_j) \sqrt{(\text{diag}(\bar{x}_i + \bar{x}_j))^{-1}} \right) \quad (6)$$

اگر بردارهای \bar{x}_{ioj} و \bar{x}_{joi} به ترتیب به صورت $\bar{x}_i \sqrt{(\text{diag}(\bar{x}_i + \bar{x}_j))^{-1}}$ و $\bar{x}_j \sqrt{(\text{diag}(\bar{x}_i + \bar{x}_j))^{-1}}$ تعریف شوند، می‌توان رابطه (۶) را به صورت رابطه (۷) ساده‌نویسی نمود.

$$\chi_A^2(\bar{x}_i, \bar{x}_j) = (\bar{x}_{ioj} - \bar{x}_{joi})^T A (\bar{x}_{ioj} - \bar{x}_{joi}) \quad (7)$$

با فرض $A = L^T L$ ، رابطه (۷) به شکل رابطه (۸) قابل بازنویسی است.

$$\chi_A^2(\bar{x}_i, \bar{x}_j) = \chi_L^2(\bar{x}_i, \bar{x}_j) = \|\mathbf{L}(\bar{x}_{ioj} - \bar{x}_{joi})\|^2 \quad (8)$$

یادگیری را به یک شکل جریمه می‌کنند. اما واضح است که عناصر قطر اصلی ماتریس یادگیری از اهمیت بالاتری نسبت به سایر عناصر برخوردار هستند؛ زیرا این عناصر بیان‌کننده اهمیت هر یک از ویژگی‌ها بوده و سایر عناصر بیان‌کننده تعامل بین ویژگی‌ها هستند. در ادامه، روشی پیشنهاد می‌شود که تنظیم‌کننده متفاوتی را برای عناصر قطر اصلی ماتریس یادگیری و سایر عناصر در نظر می‌گیرد.

ماتریس تصادفی B که عناصر قطر اصلی این ماتریس از یک توزیع یکنواخت در بازه $[0, 0.5]$ و عناصر غیر قطر اصلی از یک توزیع یکنواخت در بازه $[0, 1]$ هستند را در نظر بگیرید. تابع توزیع تجمعی (CDF) برای عناصر این ماتریس در بازه $[0, 0.5]$ به صورت رابطه (۱۸) است؛ این توابع توزیع تجمعی در شکل ۱ نشان داده شده‌اند.

$$F(\beta) = \begin{cases} 2\beta & \text{for } B_{ij} : i = j \\ \beta & \text{for } B_{ij} : i \neq j \end{cases} \quad (18)$$

احتمال dropout بر اساس ماتریس B به صورت رابطه (۱۹) تعریف می‌شود.

$$p(\delta_{ij} = 0) = p(B_{ij} \leq \beta) = \begin{cases} 2\beta & i = j \\ \beta & i \neq j \end{cases} \quad (19)$$

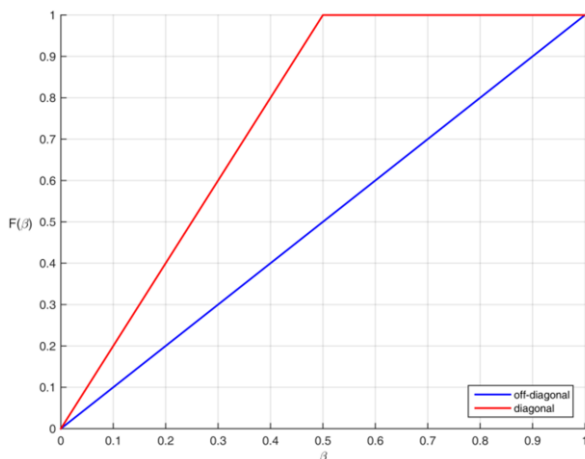
با توجه به رابطه (۱۹) و رابطه (۱۲) می‌توان نوشت:

$$E[\hat{A}_{t-1}] = \begin{cases} A_{i,j}^{t-1}(1-2\beta) & i = j \\ A_{i,j}^{t-1}(1-\beta) & i \neq j \end{cases} \quad (20)$$

استفاده از روش dropout معرفی شده در رابطه (۱۹)، معادل با حل بهینه‌سازی نشان داده شده در رابطه (۲۱) است.

$$\min_{\text{constraints}} \ell(A, X) + \frac{\beta}{2\eta} (\|A\|_F^2 + \sum_{i=1}^D A_{i,j}^2) \quad (21)$$

همان‌طور که در رابطه (۲۱) ملاحظه می‌شود، تنظیم‌کننده استفاده شده، معادل با نرم فروبنیوس ماتریس A به علاوه نرم ۱ قطر اصلی این ماتریس است. بنابراین، قطر اصلی ماتریس A ، یک‌بار در تنظیم‌کننده نرم فروبنیوس و بار دیگر در تنظیم‌کننده نرم ۱ شرکت داشته است، اما سایر عناصر فقط در تنظیم‌کننده نرم فروبنیوس حضور داشته‌اند. مسأله بهینه‌سازی رابطه (۲۱) به‌ظاهر پیچیده است، اما با روش dropout



شکل ۱: تابع توزیع تجمعی برای عناصر قطر اصلی و غیر قطر اصلی ماتریس تصادفی B

است [17]. این روش برای جلوگیری از وقوع بیش‌تعلیم الگوریتم بر روی داده‌های آموزشی انجام می‌شود. روش dropout برای یادگیری متریک می‌تواند قبل از مرحله ۵ در الگوریتم ۱ اعمال شود. برای این کار، طبق رابطه (۱۲)، عناصر ماتریس یادگیری در متغیرهای تصادفی برنولی ضرب می‌شوند.

$$\hat{A}_{ij} = \delta_{ij} A_{ij}, \quad i, j = 1, 2, \dots, d \quad (12)$$

که δ_{ij} یک متغیر تصادفی برنولی با شرط $\delta_{ij} = \delta_{ji}$ است. این شرط، خاصیت متقارن بودن ماتریس \hat{A}_{ij} را حفظ می‌کند. در این حالت، به‌روزرسانی ماتریس A در مرحله ۵ الگوریتم SGD به صورت رابطه (۱۳) تعریف می‌شود.

$$A_t = \hat{A}_{t-1} - \eta \nabla \ell \quad (13)$$

نرم فروبنیوس ۱۴ یکی از تنظیم‌کننده‌های متداول در یادگیری متریک است [1, 2, 17, 20] که برای جلوگیری از بیش‌تعلیم بر روی داده‌های آموزشی استفاده می‌شود. در ادامه نحوه پیاده‌سازی این تنظیم‌کننده با استفاده از تکنیک dropout بیان می‌شود.

شکل استاندارد یک مسأله یادگیری متریک با تنظیم‌کننده نرم فروبنیوس به صورت رابطه (۱۴) است.

$$\min_{\text{constraints}} \{ \ell_F(A, X) = \ell(A, X) + \frac{\beta}{2\eta} \|A\|_F^2 \} \quad (14)$$

که در آن، η گام تصحیح گرادیان و β ضریب تأثیر تنظیم‌کننده است. در حل مسأله بهینه‌سازی رابطه (۱۴) با استفاده از روش کاهش گرادیان، به‌روزرسانی A به صورت رابطه (۱۵) تعریف می‌شود.

$$A_t = A_{t-1} - \eta \nabla \ell_F \quad (15)$$

با توجه به $\nabla \|A\|_F^2 = 2A$ خواهیم داشت: $\nabla \ell_F = \frac{\beta}{\eta} A + \nabla \ell$ و می‌توان رابطه (۱۵) را به صورت رابطه (۱۶) بازنویسی کرد.

$$A_t = (1-\beta)A_{t-1} - \eta \nabla \ell \quad (16)$$

در روش dropout، اگر $p(\delta_{ij} = 0) = \beta$ ، خواهیم داشت:

$$E[\hat{A}_{t-1}] = (1-\beta)A_{t-1} \quad (17)$$

با توجه به رابطه (۱۷)، به‌روزرسانی A در الگوریتم SGD با روش dropout بیان شده در رابطه (۱۳) معادل استفاده از تنظیم‌کننده نرم فروبنیوس در بهینه‌سازی (رابطه (۱۶)) است. لذا تنظیم‌کننده نرم فروبنیوس با استفاده از dropout به‌سادگی پیاده‌سازی شده است. تنظیم‌کننده نرم فروبنیوس، به صورت ساده تمام عناصر ماتریس

الگوریتم ۱: یادگیری متریک با الگوریتم SGD

- ۱: ورودی‌ها: پایگاه داده آموزشی $X \in \mathbb{R}^{D \times N}$ و گام یادگیری η
- ۲: مقداردهی اولیه A_0 و T
- ۳: حلقه تکرار $t = 1$ تا T
- ۴: انتخاب نمونه تصادفی $\{\bar{x}_t, \bar{x}_t, \bar{x}_t\}_{j=1}^d$
- ۵: به‌روزرسانی $A_t = A_{t-1} - \eta \nabla \ell$
- ۶: $\prod_{psd} (A_t)$
- ۷: اتمام حلقه تکرار
- ۸: خروجی: \bar{A}

ارزیابی در سایر روش‌های موجود در جدول ۱ نیز توسط نویسندگان رعایت شده است و در نتیجه امکان انجام یک مقایسه عادلانه را فراهم می‌کند. تعداد همسایگی K در دسته‌بندی کننده KNN از الگوریتم پیشنهادی و همچنین در سایر روش‌ها برابر با ۳ قرار داده شده است. پیاده‌سازی QCS و QCN^{۱۰}، ITML^{۱۱}، LMNN^{۱۲}، GB-^{۱۸} LMNN منتشر شده و در دسترس است؛ همچنین نتایج این روش‌ها با معیار ارزیابی که ذکر شد، در [11] و [24] نیز گزارش شده است. در روش‌های LMNN- χ^2 و CDML، نتایج ارائه شده در جدول ۱ توسط خود نویسندگان گزارش شده است.

همان‌طور که در جدول ۱ ملاحظه می‌شود، خطای الگوریتم یادگیری متریک پیشنهادی بر روی پایگاه‌های داده dslr، webcam، amazon و caltech به ترتیب با خطای ۱۰/۸، ۵/۱، ۱۹/۸ و ۴۰/۶ درصد است که این مقادیر در مقایسه با سایر روش‌های موجود کمتر است. نتایج بدست آمده از آزمایش الگوریتم پیشنهادی نشان دهنده کارایی این الگوریتم در دسته‌بندی داده‌های هیستوگرامی است.

۴- جمع‌بندی

در این مقاله یک روش یادگیری متریک جدید برای دسته‌بندی داده‌های هیستوگرامی با استفاده از دسته‌بندی کننده KNN پیشنهاد شد. برخلاف اکثر روش‌های یادگیری متریک که بر مبنای معیار ماهالانوبیس (که برای فضای متعامد تعریف می‌شود) توسعه داده شده‌اند، در روش پیشنهادی، از فاصله χ^2 استفاده شده است. در نتیجه در ویژگی‌های هیستوگرامی دارای برتری قابل توجهی نسبت به روش‌های موجود است. در نوآوری دیگری، به منظور جلوگیری از بیش‌تعلیم در یادگیری، از یک تنظیم‌کننده dropout جدید استفاده شده است. استفاده از این روش، که معمولاً در یادگیری عمیق به کار می‌رود، باعث می‌شود که با یک پیاده‌سازی نسبتاً ساده بتوان قدرت تعمیم الگوریتم را افزایش داد. روش یادگیری متریک پیشنهادی داده‌های هیستوگرامی پایگاه‌های داده شناخته شده dslr، webcam، amazon و caltech را به ترتیب با خطای ۱۰/۸، ۵/۱، ۱۹/۸ و ۴۰/۶ درصد دسته‌بندی نموده است که بسیار پایین‌تر از دیگر روش‌های موجود بوده است.

مراجع

- [1] B. Kulis, "Metric learning: A survey," Foundations and Trends® in Machine Learning, vol. 5, no. 4, pp. 287-364, 2013.
- [2] A. Bellet, A. Habrard, M. Sebban, "A survey on metric learning for feature vectors and structured data," arXiv preprint arXiv:1306.6709, 2013.
- [3] E. P. Xing, M. I. Jordan, S. J. Russell, A. Y. Ng, "Distance metric learning with application to clustering with side-information," In Advances in neural information processing systems, pp. 521-528, 2003.
- [4] M. Schultz, T. Joachims, "Learning a distance metric from relative comparisons," In Advances in neural information processing systems, pp. 41-48, 2004.
- [5] J. V. Davis, B. Kulis, P. Jain, S. Sra, I. S. Dhillon, "Information theoretic metric learning," In Proceedings of the 24th International Conference on Machine Learning (ICML), pp. 209-216, 2007.
- [6] P. Jain, B. Kulis, I. S. Dhillon, K. Grauman, "Online Metric Learning and Fast Similarity Search," In Advances in Neural Information Processing Systems (NIPS), vol. 21, pp. 761-768, 2008.

معرفی شده در الگوریتم ۲، به راحتی قابل حل است.

نکته دیگری که در الگوریتم ۲ وجود دارد، تصویر کردن A روی مخروط PSD است. این کار که در الگوریتم ۱ در هر تکرار انجام می‌شود، هزینه محاسباتی زیادی دارد. نتایج آزمایش‌ها نشان می‌دهد که اگر بجای انجام این کار در هر تکرار، فقط ماتریس یادگیری نهایی روی PSD تصویر شود، علاوه بر کاهش هزینه محاسباتی، دقت آموزش نیز حفظ می‌شود [21]. بنابراین در الگوریتم ۲ از این روش استفاده شده است.

۳- پایگاه‌های داده و آزمایش‌ها

در این بخش به شرح پایگاه‌های داده و آزمایش‌های انجام شده بر روی آنها برای ارزیابی الگوریتم پیشنهادی پرداخته می‌شود.

۳-۱- پایگاه‌های داده

چهار پایگاه داده هیستوگرامی dslr، webcam، amazon و caltech که در [22] معرفی شده‌اند، در اصل برای بازشناسی اشیاء مورد استفاده قرار می‌گیرند و در منابع یادگیری متریک نیز پر استفاده بوده‌اند. این پایگاه‌های داده از ۱۰ شیء معمولی شامل کوله‌پشتی، دوچرخه، ماشین حساب، هدفون، صفحه‌کلید، رایانه لپ‌تاپ، نمایشگر، ماوس، لیوان و پروژکتور تشکیل شده است. داده‌های موجود در این پایگاه‌های داده، ویژگی‌های هیستوگرامی SURF به طول ۸۰۰ هستند. پایگاه‌های داده dslr، webcam، amazon و caltech به ترتیب دارای ۱۵۷، ۲۹۵، ۹۵۸ و ۱۱۲۳ نمونه هستند. از این چهار پایگاه داده برای ارزیابی الگوریتم یادگیری متریک پیشنهادی استفاده شده است.

۳-۲- آزمایش‌ها

در این بخش، الگوریتم پیشنهادی مطرح شده در الگوریتم ۲، بر روی پایگاه‌های داده هیستوگرامی شرح داده شده آزمایش شده است. درصد خطای دسته‌بندی الگوریتم پیشنهادی در مقایسه با معیارهای فاصله اقلیدسی و χ^2 ، معیار فواصل χ^2 درجه دوم (QCS و QCN) [23]، روش‌های ITML [5]، LMNN [10]، LMNN [11]، GB- χ^2 LMNN [11] و CDML [24] در جدول ۱ نشان داده شده است. در آزمایش‌های انجام شده در جدول ۱ از ارزیابی 5-fold cross validation با تقسیم‌بندی تصادفی استفاده شده است. این نحوه

الگوریتم ۲: یادگیری متریک با تنظیم‌کننده dropout

- ۱: ورودی‌ها: پایگاه داده آموزشی $X \in \mathbb{R}^{D \times N}$ و گام یادگیری η
- ۲: مقداردهی اولیه A_0 و T
- ۳: حلقه تکرار $t = 1$ تا T
- ۴: انتخاب نمونه تصادفی $\{\vec{x}_i, \vec{x}_j, \vec{x}_l\}_{j-1}$
- ۵: ایجاد ماتریس تصادفی B
- ۶: ایجاد پارامتر δ با استفاده از رابطه (۱۹)
- ۷: $\hat{A}_{ij}^{t-1} = \delta_{ij} A_{ij}^{t-1}$, $i, j = 1, 2, \dots, d$
- ۸: $A_t = \hat{A}_{t-1} - \eta \nabla \ell$
- ۹: اتمام حلقه تکرار
- ۱۰: خروجی: $\prod_{psd}(\bar{A})$

جدول ۱: مقایسه خطای دسته‌بندی روش‌های مختلف بر روی پایگاه‌های داده هیستوگرامی

dslr	webcam	amazon	caltech	روش
$60/6 \pm 3/1$	$43/8 \pm 1/7$	$33/7 \pm 0/7$	$53/8 \pm 1/3$	فاصله اقلیدسی
$22/2 \pm 1/8$	$13/0 \pm 1/2$	$34/3 \pm 1/0$	$58/8 \pm 1/1$	فاصله χ^2
$25/6 \pm 2/7$	$19/4 \pm 1/1$	$33/9 \pm 2/0$	$57/2 \pm 1/2$	[23] QCS
$27/8 \pm 4/1$	$17/5 \pm 2/1$	$34/5 \pm 1/5$	$56/1 \pm 1/2$	[23] QCN
$25/0 \pm 3/0$	$12/4 \pm 1/6$	$31/6 \pm 1/2$	$52/2 \pm 2/1$	[5] ITML
$28/9 \pm 1/6$	$15/8 \pm 3/0$	$31/8 \pm 1/4$	$50/9 \pm 1/4$	[10] LMNN
$20/6 \pm 1/1$	$8/3 \pm 0/9$	$23/7 \pm 0/8$	$46/5 \pm 1/1$	[11] χ^2 -LMNN
$22/9 \pm 2/7$	$12/4 \pm 0/9$	$29/6 \pm 1/7$	$49/8 \pm 1/0$	[11] GB-LMNN
$16/7 \pm 4/2$	$5/9 \pm 2/3$	$20/8 \pm 3/2$	$42/2 \pm 2/4$	[24] CDML
$10/8 \pm 2/5$	$5/1 \pm 1/2$	$19/8 \pm 1/2$	$40/6 \pm 4$	روش پیشنهادی

- [23] O. Pele, M. Werman, "The quadratic-chi histogram distance family," in Computer Vision—ECCV 2010, vol. 6312 of Lecture Notes in Computer Science, pp. 749–762, Springer, Berlin, Germany, 2010.
- [24] W. Yang, L. Xu, X. Chen, F. Zheng, Y. Liu, "Chi-squared distance metric learning for histogram data," Mathematical Problems in Engineering, vol. 2015, 2015.

زیر نویس‌ها

- 1 class
- 2 Mahalanobis
- 3 Positive Semi-Definite
- 4 Information-Theoretic Metric Learning
- 5 Large Margin Nearest Neighbors
- 6 K-Nearest Neighbors
- 7 Chi-square
- 8 regularizer
- 9 over-fitting
- 10 Stochastic Gradient Descent
- 11 targets neighbors
- 12 impostors
- 13 hinge loss
- 14 Frobenius norm
- 15 <http://www.ariel.ac.il/sites/ofirpele/QC/code/>
- 16 <http://www.cs.utexas.edu/users/pjain/itml/>
- 17 <http://www.cse.wustl.edu/~kilian/code/files/mLMMN2.4.zip>
- 18 <http://www.cs.cornell.edu/~kilian/code/code.html>

- [7] G. Kunapuli, J. Shavlik, "Mirror Descent for Metric Learning: A Unified Approach," In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Database (ECML/PKDD), pp. 859–874, 2012.
- [8] Q. Qian, R. Jin, J. Yi, L. Zhang, S. Zhu, "Efficient distance metric learning by adaptive sampling and mini-batch stochastic gradient descent (SGD)," Machine Learning, vol. 99, no. 3, pp. 353-372, 2015.
- [9] B. Nguyen, C. Morell, B. D. Baets, "Supervised distance metric learning through maximization of the Jeffrey divergence," Pattern Recognition, vol. 64, pp. 215-225, 2017.
- [10] K. Q. Weinberger, K. S. Lawrence, "Distance metric learning for large margin nearest neighbor classification," Journal of Machine Learning Research, vol. 10, pp. 207-244, 2009.
- [11] D. Kedem, S. Tyree, F. Sha, G. R. Lanckriet, K. Q. Weinberger, "Non-linear metric learning," In Advances in Neural Information Processing Systems, pp. 2573-2581, 2012.
- [12] T. Ahonen, A. Hadid, M. Pietikäinen, "Face recognition with local binary patterns," In Computer vision-eccv 2004, pp. 469-481, Springer Berlin Heidelberg, 2004.
- [13] G. Zhang, X. Huang, S. Z. Li, Y. Wang, X. Wu, "Boosting local binary pattern (LBP)-based face recognition," In Advances in biometric person authentication, pp. 179-186, Springer Berlin Heidelberg, 2004.
- [14] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," arXiv preprint arXiv:1207.0580 (2012).
- [15] L. v. d. Maaten, M. Chen, S. Tyree, K. Q. Weinberger, "Learning with marginalized corrupted features," In ICML, pp. 410-418, 2013.
- [16] S. Wager, S. Wang, and P. Liang, "Dropout training as adaptive regularization," In NIPS, pp. 351-359, 2013.
- [17] Q. Qian, J. Hu, R. Jin, J. Pei, S. Zhu, "Distance metric learning using dropout: a structured regularization approach," In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 323-332, 2014.
- [18] S. Boyd, L. Vandenberghe, "Convex Optimization," Cambridge University Press, 2004.
- [19] Q. Qian, R. Jin, J. Yi, L. Zhang, S. Zhu, "Efficient distance metric learning by adaptive sampling and mini-batch stochastic gradient descent (SGD)," Machine Learning, vol. 99, no. 3, pp. 353-372, 2015.
- [20] Q. Qian, R. Jin, S. Zhu, Y. Lin, "Fine-grained visual categorization via multi-stage metric learning," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3716-3724, 2015.
- [21] G. Chechik, V. Sharma, U. Shalit, S. Bengio, "Large scale online learning of image similarity through ranking," JMLR, vol. 11, pp. 1109-1135, 2010.
- [22] K. Saenko, B. Kulis, M. Fritz, T. Darrell, "Adapting visual category models to new domains," Computer Vision—ECCV 2010, pp. 213–226, 2010.



بررسی تأثیر کاهش ویژگی بر افزایش نرخ دقت تشخیص صفحات وب هرز

علی سلیمانی ایوری^۱، فائزه اصدقی^۲

^۱ دانشیار، دانشکده مهندسی برق، دانشگاه صنعتی شاهرود، شاهرود،
solimani_ali@shahroodut.ac.ir

^۲ دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود،
asdaghi@shahroodut.ac.ir

چکیده

وب اسپم روشی غیرقانونی و غیراخلاقی به منظور افزایش رتبه صفحات اینترنتی توسط فریب الگوریتم‌های موتورهای جستجو است. از آنجا که کیفیت نتایج برای موتورهای جستجو اهمیت بسیاری دارد، ویژگی‌های زیادی جهت تشخیص صفحات وب هرز پیشنهاد شده است. چالشی که تنوع ویژگی‌ها با آن روبرو است افزایش ابعاد ویژگی و در نتیجه کاهش نرخ تشخیص است. به همین دلیل استفاده از روش‌های کاهش ویژگی به عنوان یک پیش پردازش امری اجتناب‌ناپذیر به نظر می‌رسد. در این مقاله سعی داریم با بررسی ویژگی‌های مستخرج از صفحات وب و انتخاب زیرمجموعه‌ای مناسب از آنها، نرخ تشخیص صفحات وب هرز را افزایش دهیم. بدین منظور ابتدا با استفاده از ۱۱ روش جستجو و هشت معیار ارزیابی، ۲۶ زیرمجموعه مختلف از مجموع کلیه ویژگی‌ها انتخاب شد. سپس با استفاده از الگوریتم Naïve Bayes میزان کارایی و مؤثر بودن هر یک از این زیرمجموعه‌ها در تشخیص صفحات وب هرز با استفاده از معیاری به نام IBA اندازه‌گیری شد. نتایج این بررسی که با استفاده از ابزار weka و بر روی مجموعه داده معتبر WEBSpAM-UK2007 صورت گرفت نشان داد که از مجموع ۲۷۵ ویژگی موجود در این پایگاه داده، ۳۲ ویژگی که با روش جستجوی رتبه‌بندی و معیار χ^2 انتخاب شده‌اند باعث بهبود عملکرد طبقه‌بندی و افزایش آن از ۰/۳۳۷ به ۰/۳۶۹ شد.

کلمات کلیدی

وب اسپم، انتخاب ویژگی، روش‌های جستجو، ارزیابی ویژگی.

۱- مقدمه

جستجو است. وب اسپم به عنوان یک معضل رایج بین موتورهای جستجو شناخته می‌شود که از زمان پیدایش موتورهای جستجو همواره با آن درگیر بوده‌اند. هدف وب اسپم تغییر رتبه صفحات وب در نتایج جستجو است. به این صورت که رتبه صفحات را بیش از آنچه که باید نشان دهد تا ترجیحاً در میان ۱۰ نتیجه اول ظاهر شود. وب اسپم‌ها باعث کاهش کیفیت نتایج جستجو و در نتیجه اتلاف وقت کاربر می‌شوند. افزایش تعداد این صفحات، افزایش میزان منابع مصرفی توسط خزشگرها و اندیس‌سازها را در پی دارد که این امر موجب اتلاف منابع موتور جستجو و افزایش زمان پاسخگویی به کاربر می‌شود. به همین دلیل تحقیقات زیادی در جهت ارائه روشی مؤثر به منظور تشخیص صحیح این صفحات صورت گرفته است. از آنجا که یکی از عوامل مؤثر در افزایش نرخ تشخیص، انتخاب ویژگی‌های درست و کاهش ابعاد داده است، لذا در این مقاله قصد

امروزه با توجه به رشد اطلاعات در وب، موتورهای جستجو به عنوان درگاهی برای ورود به آن در نظر گرفته می‌شوند. تحقیقاتی که در مرجع [1] صورت گرفت نشان می‌دهد که تقریباً ۸۵٪ کاربران تنها از سه یا پنج پیوند اولیه صفحه نخست نتایج بازدید می‌کنند. در نتیجه حضور صفحه‌ای در نتایج بالای موتورهای جستجو به معنای بازدیدکننده بیشتر و نیز درآمد بیشتر است. به همین علت موتورهای جستجو برای آنکه بهترین نتایج ممکنه را در اختیار کاربران قرار دهند از الگوریتم‌های بسیار پیچیده‌ای استفاده می‌کنند.

در این میان وب اسپم روشی غیرقانونی و غیراخلاقی به منظور افزایش رتبه صفحات اینترنتی توسط فریب الگوریتم‌های موتورهای

است که اضافه یا حذف یک ویژگی باعث کاهش مقدار تابع ارزیابی شود. همچنین این الگوریتم می‌تواند بر اساس پیمایش فضا از یک سو به سوی دیگر و ذخیره ترتیب ویژگی‌های انتخاب شده، لیستی از ویژگی‌ها ایجاد کند که بر اساس میزان مفید بودنشان رتبه‌بندی شده‌اند.

۲-۱-۲ اول بهترین

روش جستجوی اول بهترین^۶ یک روش انتخاب حریصانه است که تلاش می‌کند کارایی زیرمجموعه فعلی را با افزودن یا حذف کلیه ویژگی‌های ممکن به مجموعه بهتری تبدیل کند. این الگوریتم می‌تواند با مجموعه‌ای خالی از ویژگی شروع به جستجو نماید و به جلو رود یا با مجموعه‌ای کامل از ویژگی‌های شروع کند و عقب‌گرد داشته باشد یا از هر نقطه‌ای شروع کند و در دو جهت جلو و عقب ادامه دهد. در صورتی که هیچ محدودیتی برای ادامه الگوریتم وجود نداشته باشد، این روش یک جستجوی جامع محسوب می‌شود [2].

۲-۱-۳ خطی رو به جلو

الگوریتم خطی رو به جلو^۷ نوعی روش اول بهترین است که ابتدا بر اساس یک ارزیاب، ویژگی‌ها را به طور انفرادی بررسی کرده و آن‌ها را رتبه‌بندی می‌کند. سپس K ویژگی اول را انتخاب کرده و بر روی آن‌ها عمل جستجو را انجام می‌دهد به این صورت که با یک مجموعه خالی شروع می‌کند، سپس در هر تکرار یک ویژگی را با استفاده از تابع ارزیابی به مجموعه جواب اضافه می‌کند. این کار را تا زمانی تکرار می‌شود که تعداد ویژگی لازم انتخاب شود [2].

۲-۱-۴ روبه‌جلو با تعیین اندازه زیرمجموعه

این روش توسعه یافته الگوریتم خطی رو به جلو است. در این نسخه از الگوریتم، در قسمت جستجو، یک اعتبارسنجی داخلی انجام می‌شود (نقطه شروع و تعداد دسته‌ها می‌تواند از قبل مشخص شود) تا اندازه زیرمجموعه بهینه مشخص شود (با استفاده از ارزیابی کننده‌ای که در ابتدا مشخص شده است). در نهایت یک بار دیگر الگوریتم خطی رو به جلو بر اساس اندازه زیرمجموعه بهینه بر روی تمام داده‌ها اعمال می‌شود [3].

۲-۱-۵ رتبه‌بندی

در این روش ارزش ویژگی‌ها به تنهایی و بدون در نظر گرفتن رابطه بین آن‌ها سنجیده می‌شود. خروجی این الگوریتم یک لیست رتبه‌بندی شده از ویژگی‌های به ترتیب اهمیت است. به همین علت این روش بیشتر در کنار توابع ارزیابی مانند ریلیف^۸، نرخ سود^۹ و انتروپی^{۱۰} استفاده می‌شود [2].

۲-۱-۶ جستجوی پراکنده

جستجوی پراکنده^{۱۱} یک الگوریتم تکاملی است و شامل مجموعه راه‌حلی‌هایی است که بر اساس مکانیزم ترکیبی بین آن‌ها تکامل یافته‌اند. این الگوریتم در بین زیرمجموعه‌های ویژگی یک جستجوی پراکنده

داریم تا انواع روش‌های انتخاب ویژگی را بررسی و از بین زیرمجموعه‌های پیشنهادی، موثرترین مجموعه را انتخاب کنیم. بدین منظور ابتدا با استفاده از هر یک از مجموعه ویژگی‌ها و طبقه‌بند بیز ساده صفحات وب را به دو دسته هرز و غیرهرز تقسیم می‌کنیم. سپس نتیجه به دست آمده را با معیار جدیدی که به همین منظور معرفی شده، ارزیابی می‌کنیم.

برای آشنایی بیشتر با موضوع در بخش بعد روش‌های کلاسیک انتخاب ویژگی شامل انواع روش‌های جستجو و توابع ارزیابی بررسی شده‌اند. در بخش سه مجموعه داده و معیارهای ارزیابی معرفی شده است. در نهایت نتایج در بخش چهارم و نتیجه‌گیری در بخش پنجم مطرح شده است.

۲- انتخاب ویژگی

مسئله انتخاب ویژگی تلاش می‌کند تا از میان 2^N زیرمجموعه ممکن از کل ویژگی‌ها، زیرمجموعه‌ای را انتخاب کند که بتواند مقدار یک تابع ارزیابی را بهینه کند. روش‌های مختلف انتخاب ویژگی شامل دو جزء اصلی الگوریتم جستجو^۱ و تابع ارزیابی کننده^۲ هستند. در بعضی از الگوریتم‌های جستجو تمام فضای ممکن و در برخی دیگر که می‌تواند مکاشفه‌ای یا جستجوی تصادفی باشد، در ازای از دست دادن مقداری از کارایی، فضای کوچکتری جستجو می‌شود. همچنین توابع ارزیابی کننده متعددی نیز وجود دارد. یک تابع ارزیابی، میزان خوب بودن زیرمجموعه تولید شده را محاسبه کرده و حاصل را با مقدار بهترین زیرمجموعه قبلی مقایسه می‌کند. اگر زیرمجموعه جدید، بهتر از زیرمجموعه‌های قدیمی باشد، زیرمجموعه جدید به عنوان زیرمجموعه بهینه، جایگزین قبلی می‌شود [2]. در این مقاله از ۱۱ روش جستجو و هشت تابع ارزیاب استفاده شده است که در ادامه به توضیح مختصر آن‌ها می‌پردازیم.

۲-۱- روش‌های جستجو

وظیفه الگوریتم جستجو که به آن تابع تولید کننده نیز گفته می‌شود، ایجاد زیرمجموعه‌های مختلف از کل ویژگی‌ها است. الگوریتم‌های جستجو یا بدون ویژگی و یا با مجموعه تمام ویژگی‌ها شروع به کار می‌کنند. در حالت اول ویژگی‌ها به ترتیب به مجموعه اضافه می‌شوند و زیرمجموعه‌های جدید را تولید می‌کنند. به این دسته روش‌ها، روش‌های پائین به بالا می‌گویند.

در حالت دوم از یک مجموعه شامل تمام ویژگی‌ها، شروع می‌کنیم و به مرور و در طی اجرای الگوریتم، ویژگی‌ها را حذف می‌کنیم تا به زیرمجموعه دلخواه برسیم. روش‌هایی که به این صورت عمل می‌کنند، روش‌های بالا به پائین نام دارند. در ادامه به شرح برخی از این الگوریتم‌ها می‌پردازیم.

۲-۱-۱ حریصانه

یکی از استراتژی‌های جستجو، روش حریصانه^۳ است که فضای ویژگی را به صورت جلو رفت^۴ یا عقب‌گرد^۵ جستجو می‌کند. این الگوریتم عمل جستجو را از یک نقطه دلخواه آغاز می‌کند و نقطه توقف آن زمانی

۲-۱-۱۱ الگوریتم زنبورعسل

این الگوریتم شبیه‌سازی رفتار جستجوی غذای گروه‌های زنبور عسل است. برای انتخاب ویژگی توسط این الگوریتم ابتدا زنبورها d گام به جلو می‌روند و در هر گام تصمیم می‌گیرند یک ویژگی را انتخاب کنند یا خیر. پس از اتمام d مسیر زنبورها به کندو بازگشته و نتایج خود را ارزیابی می‌کنند. در این مرحله زنبورها به دو دسته متعهد و غیرمتعهد تقسیم می‌شوند. زنبورهایی که برازندگی‌شان پس از ارزیابی از مقدار آستانه بیشتر بود به عنوان متعهد و بقیه غیرمتعهد هستند. این مفهوم به معنی وفاداری زنبور عسل به شاهد کشف شده است. زنبورهای غیرمتعهد می‌بایست بر اساس چرخ رولت به سمت راه‌حل زنبور متعهد حرکت کنند و پس از حرکت مسیر با d گام آزادند که مسیر جدیدی را انتخاب کنند [8].

۲-۲-۲ ارزیابی ویژگی

یک تابع ارزیابی، میزان خوب بودن یک زیرمجموعه تولید شده را بررسی کرده و یک مقدار به عنوان میزان خوب بودن زیرمجموعه مورد نظر بازمی‌گرداند. این مقدار با بهترین زیرمجموعه قبلی مقایسه می‌شود. اگر زیرمجموعه جدید، بهتر از زیرمجموعه‌های قبلی باشد، زیرمجموعه جدید به عنوان زیرمجموعه بهینه، جایگزین قبلی می‌شود. توابع ارزیابی که در این مقاله استفاده شده است در ادامه توضیح داده شده است.

۲-۲-۱ همبستگی

ضریب همبستگی^{۱۵} ابزاری آماری برای تعیین نوع و درجه رابطه یک متغیر کمی با متغیر کمی دیگر است. ضریب همبستگی، یکی از معیارهای استفاده شده در تعیین همبستگی دو متغیر است. ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می‌دهد. این ضریب بین ۱ تا -۱ است و در نبود وجود رابطه بین دو متغیر، برابر صفر است. همبستگی بین دو متغیر تصادفی X و Y به صورت زیر تعریف می‌شود که در آن E عملگر امید ریاضی، cov به معنای کوواریانس، corr نماد معمول برای همبستگی پیرسان و سیگما نماد انحراف معیار است [9].

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (1)$$

۲-۲-۲ مربع کای

در این روش ارزش هر ویژگی با استفاده از محاسبه مقدار آماری مربع کای^{۱۶} و با توجه به کلاس مد نظر ارزیابی می‌گردد. این مقدار با استفاده از آزمون مربع کای که جزء آزمون‌های غیر پارامتری است محاسبه می‌شود. برای این کار ابتدا مربع کای را برای هر کلاس و با استفاده از فرمول (۲) محاسبه می‌کنیم. در این فرمول O_i مقدار مشاهده ویژگی i و E_i مقدار مورد انتظار آن ویژگی است.

$$X^2 = \sum [(O_i - E_i)^2 / E_i] \quad (2)$$

انجام می‌دهد. شروع جستجو با تعداد زیادی زیرمجموعه‌های متنوع و حائز اهمیت است و الگوریتم زمانی پایان می‌یابد که نتیجه از مقدار آستانه مدنظر بیشتر شود یا امکان بهبود بیشتری نباشد [4].

۲-۱-۷ الگوریتم ژنتیک

در این روش یک دسته ویژگی از زیرمجموعه ویژگی‌های کاندید شده تولید می‌کنیم. در هر بار تکرار الگوریتم، با استفاده از عملگرهای جهش و بازترکیبی بر روی عناصر دسته قبلی، عناصر جدیدی تولید می‌کنیم. با استفاده از یک تابع ارزیابی، میزان شایستگی عناصر دسته فعلی را مشخص کرده و عناصر بهتر را به عنوان دسته نسل بعد انتخاب می‌کنیم. اگرچه یافتن بهترین جواب در این روش تضمین نمی‌شود؛ ولی همیشه یک جواب نسبتاً بهینه با توجه به مدت زمانی که به الگوریتم اجازه اجرا داده باشیم، پیدا می‌شود [2].

۲-۱-۸ الگوریتم توده ذرات

روش توده ذرات^{۱۲} از رفتار اجتماعی دسته‌های پرندگان الهام گرفته شده است. در این سیستم عامل‌ها به طور محلی با یکدیگر همکاری می‌نمایند و رفتار جمعیت باعث همگرایی در نقطه‌ای نزدیک به جواب بهینه سراسری می‌شود. برای نیل به این هدف لازم است تا در بعد اول همه اعضای گروه موقعیت خود را با تابعیت از بهترین فرد گروه تغییر دهند و در بعد دوم لازم است تا تک تک اعضا بهترین موقعیتی که تاکنون تجربه کرده‌اند را در حافظه نگهداری کرده، میل به سوی بهترین موقعیت خود در گذشته داشته باشند. سرعت و موقعیت ذرات تا زمان برآورده شدن شرط خاتمه (رسیدن به حد آستانه تعیین شده) با توجه به این دو مورد به طور مداوم به روز می‌شود [5].

۲-۱-۹ الگوریتم کلونی مورچه

برای پیاده‌سازی مسأله انتخاب ویژگی با الگوریتم کلونی مورچه^{۱۳} دو رشته صفر و یک به طول تعداد متغیرها در نظر می‌گیریم و الگوریتم را برای تشکیل رشته‌ای از صفر و یک‌ها اجرا می‌کنیم. مسیری که هر مورچه طی می‌کند با استفاده از توابع فرومون و ابتکاری مشخص می‌شود. در نهایت بهترین مسیری که توسط مورچه‌ها انتخاب شده است مشخص می‌کند که کدام ویژگی می‌بایست انتخاب شود [6].

۲-۱-۱۰ الگوریتم جستجوی خفاش

الگوریتم جستجوی خفاش^{۱۴} الهام گرفته از نحوه مسیریابی خفاش‌ها است. این موجودات قدرت بینایی ضعیفی داشته و بر اساس انعکاس صدای خود مسیر حرکت را تعیین می‌کنند. در این روش جستجوی فضای ویژگی بر اساس معیارهای سرعت و مکان صورت گرفته و از تنظیم فرکانس به منظور جهش به حالت بعد استفاده می‌شود. میزان این جهش با توجه به تغییرات بلندی صدا و پالس انتشار متفاوت است. این تغییرات قابلیت تمرکز خودکار را فراهم می‌کند به طوری که باعث استخراج راه‌حل‌های قوی‌تر می‌شود [7].

۲-۲-۶- طبقه‌بند OneR

این تابع ارزیابی ویژگی‌های را بر اساس نتیجه طبقه‌بند OneR ارزیابی می‌کند. این طبقه‌بند ویژگی‌ها را بر اساس نرخ خطا (در مجموعه آموزشی) رتبه‌بندی می‌کند. به این صورت که بر اساس هر یک از ویژگی‌ها یک درخت تصمیم سطح یک ایجاد می‌کند. سپس داده‌ها را بر این اساس طبقه‌بندی کرده و نرخ خطای به دست آمده را اندازه‌گیری می‌کند [10].

۲-۲-۷- Relief

یک تابع ارزیابی مبتنی بر فاصله است که بر اساس تولید مکاشفه‌ای عمل می‌کند. این روش از یک راه‌حل آماری برای انتخاب ویژگی استفاده می‌کند. همچنین یک روش مبتنی بر وزن است که از الگوریتم‌های مبتنی بر نمونه الهام گرفته است. روش کار به این صورت است که از میان مجموعه نمونه‌های آموزشی، یک زیرمجموعه نمونه انتخاب می‌کند. کاربر بایستی تعداد نمونه‌ها در این زیرمجموعه را مشخص کرده باشد و آن را به عنوان ورودی به الگوریتم ارائه دهد. الگوریتم به صورت تصادفی یک نمونه از این زیرمجموعه را انتخاب می‌کند. سپس برای هر یک از ویژگی‌های این نمونه، نزدیک‌ترین برخوردار^{۲۱} و نزدیکترین شکست^{۲۱} را بر اساس معیار اقلیدسی پیدا می‌کند. نزدیک‌ترین برخوردار، نمونه‌ای است که کمترین فاصله اقلیدسی را در میان سایر نمونه‌های هم‌کلاس با نمونه انتخاب شده دارد. نزدیک‌ترین شکست نیز نمونه‌ای با کمترین فاصله اقلیدسی در میان نمونه‌های غیر هم‌کلاس با نمونه انتخاب شده است [9].

۲-۲-۸- عدم قطعیت متقارن

در این روش همبستگی بین دو ویژگی X و Y بر مبنای تئوری اطلاعات اندازه‌گیری می‌شود. در واقع هر چه میزان عدم قطعیت متقارن^{۲۲} یک ویژگی بیشتر باشد، ارزش آن ویژگی بیشتر است. برای به دست آوردن این مقدار از فرمول (۶) استفاده می‌شود که در آن $IG(X)$ میزان بهره اطلاعاتی ویژگی X و $H(X)$ آنتروپی آن است [11].

$$SU(X, Y) = 2 \frac{IG(X|Y)}{H(X) + H(Y)} \quad (6)$$

۳- مجموعه داده و معیارهای ارزیابی

همان طور که گفته شده به منظور بررسی میزان اثربخشی روش‌های انتخاب ویژگی، مجموعه داده را با استفاده از ویژگی‌های پیشنهادی طبقه‌بندی کرده، میزان دقت به دست آمده را اندازه‌گیری می‌کنیم؛ اما پیش از آن در این بخش به معرفی مجموعه داده، الگوریتم طبقه‌بندی و معیارهای ارزیابی خواهیم پرداخت.

۳-۱- مجموعه داده

دادگان استفاده شده در این مقاله WEBSpam-UK2007 [12]، معتبرترین مجموعه داده در زمینه شناسایی صفحات وب اسپم است. این دادگان شامل مجموعه صفحاتی است که از نتایج یک خزشگر در دامنه uk به دست آمده و شامل ۱۰۵/۹ میلیون صفحه از ۱۱۴۵۲۹

سپس درجه آزادی دو متغیر را با استفاده از فرمول (۳) به دست می‌آوریم. در این فرمول ثابت‌های r و c به ترتیب تعداد دسته‌های موجود برای دو متغیر مورد آزمایش هستند.

$$P = (r - 1)(c - 1) \quad (3)$$

حال با توجه به جدول توزیع احتمال مربع کای، اگر مقدار P بیشتر از ۰,۰۵ است، فرضیه مورد قبول و دو متغیر در ارتباط هستند. در غیر این صورت متغیرها هیچ وابستگی به هم ندارند [9].

۲-۲-۳- پایداری

معیار پایداری^{۱۸} قابلیت پیشگویی مقدار یک متغیر به وسیله یک متغیر دیگر را اندازه‌گیری می‌کند. ضریب (Coefficient) یکی از معیارهای وابستگی کلاسیک است و می‌توانیم آن را برای یافتن همبستگی بین یک ویژگی و یک کلاس به کار ببریم. اگر همبستگی ویژگی X با کلاس C بیشتر از همبستگی ویژگی Y با کلاس C باشد، در این صورت ویژگی x برای قرار گرفتن در کلاس C بر y اولویت دارد. با یک تغییر کوچک، می‌توانیم وابستگی یک ویژگی با ویژگی‌های دیگر را اندازه‌گیری کنیم. این مقدار درجه افزونگی این ویژگی را نشان می‌دهد. همه توابع ارزیابی بر پایه معیار وابستگی را می‌توانیم بین دو دسته معیارهای مبتنی بر فاصله و اطلاعات تقسیم کنیم؛ اما به خاطر اینکه این روش‌ها با دیدی متفاوت از دید اصلی، به مسأله نگاه می‌کنند، این کار را انجام نمی‌دهیم [9].

۲-۲-۴- نرخ بهره

در این روش ارزش یک ویژگی با توجه به میزان نرخ بهره آن با توجه به کلاس‌ها به دست می‌آید. نرخ بهره در واقع نسبت میزان بهره اطلاعاتی ویژگی به اطلاعات انشعاب^{۱۸} آن ویژگی است. ارزش اطلاعاتی انشعاب، اطلاعات بالقوه‌ای که در صورت تقسیم مجموعه داده D به v بخش ایجاد می‌شود را نشان می‌دهد که v مقادیر دیده شده ویژگی A است. لذا اطلاعات انشعاب ویژگی A با توجه به مجموعه داده D از فرمول (۴) به دست می‌آید [9].

$$SplitInfo_A(D) = - \sum_{j=1}^v \frac{|D_j|}{|D|} \times \log_2 \left(\frac{|D_j|}{|D|} \right) \quad (4)$$

۲-۲-۵- بهره اطلاعاتی

در این روش ارزش یک ویژگی به وسیله بهره اطلاعاتی^{۱۹} آن به دست می‌آید. در بحث تئوری اطلاعات و یادگیری ماشین، بهره اطلاعاتی معادل واگرایی Kullback-Leibler است. فرض کنید که T یک مجموعه از داده‌های آزمایشی باشد که هر یک از این داده‌ها به شکل $(X, y) = (x_1, x_2, \dots, x_k, y)$ است. به نحوی که x_a عضو $vals(a)$ مقدار a امین ویژگی داده X است که به کلاس y تعلق دارد. در این صورت اگر $H()$ میزان آنتروپی باشد، بهره اطلاعاتی از فرمول (۵) محاسبه می‌شود [9].

$$IG(T, a) = H(T) - \sum_{v \in vals(a)} \frac{|\{X \in T | x_a = v\}|}{|T|} \cdot H(\{X \in T | x_a = v\}) \quad (5)$$

۴- پیاده‌سازی

ابتدا با استفاده از روش‌های جستجو و توابع ارزیابی اشاره شده در بخش سه، زیرمجموعه‌های مختلفی از ویژگی‌ها ایجاد شد. این کار با استفاده از نرم‌افزار weka [14] صورت گرفت. سپس به منظور تعیین بهترین روش انتخاب ویژگی، مجموعه داده بر اساس هر دسته از ویژگی‌های انتخاب شده و با استفاده از طبقه‌بند بیز ساده^{۲۰} دسته‌بندی شد. دلیل استفاده از طبقه‌بند بیز ساده این است که در [15] نویسندگان این مقاله انواع روش‌های طبقه‌بندی را بر روی این مجموعه داده بررسی کردند و مشخص شد این طبقه‌بند بهترین نرخ تشخیص را در بین دیگر الگوریتم‌های طبقه‌بندی دارد. نتایج این طبقه‌بندی در جدول ۱ نشان داده شده است. همان‌طور که مشاهده می‌شود استفاده از تابع ارزیابی مربع کای و روش جستجوی رتبه‌بندی، بهترین میزان IBA را به دست آورده است.

جدول ۱: نتایج حاصل از طبقه‌بندی مجموعه داده WEBSpam-UK2007 با ویژگی‌های انتخاب شده توسط روش‌های مختلف

Search Method	Attribute Evaluator	Selected Attributes	IBA
Forward Selection	Correlation	۴۴	۰/۳۳۷
	Consistency	۱۷	۰/۱۰۹
Ranker	Chi Square	۳۲	۰/۳۶۹
	InfoGain	۲۹	۰/۳۶۳
	Gain Ratio	۳۹	۰/۳۵
	SymUncertainty	۳۰	۰/۳۲۵
	ReliefF	۵۱	۰/۲۰۴
	OneR	۱۹	۰/۰۲۷
Best first	Correlation	۶۶	۰/۳۶۲
	Consistency	۸	۰/۲۴۵
LForward Selection	Correlation	۱۲	۰/۲۳۲
	Consistency	۲۱	۰/۱۸۳
Scatter Search	Correlation	۲۰	۰/۳
	Consistency	۲۸	۰/۱۹۹
Greedy Stepwise	Correlation	۶۶	۰/۳۶۲
	Consistency	۵	۰/۱۶۲
Rank Search	Correlation	۶۲	۰/۳۲۶
	Consistency	۱۳۱	۰/۲۵
Genetic Search	Correlation	۶۳	۰/۲۷۳
	Consistency	۱۵	۰/۲۰۸
Ant Search	Correlation	۲۴	۰/۱۶۶
	Consistency	۳۴	۰/۰۶۷
Bat Search	Correlation	۳۳	۰/۲۳۸
	Consistency	۲۴	۰/۳۰۳
Bee Search	Correlation	۳۳	۰/۳۰۸
	Consistency	۲۷	۰/۱۴۴

بررسی ویژگی‌های انتخاب شده نشان می‌دهد که از این ۳۲ ویژگی، هفت ویژگی پیوندمحور، هفت ویژگی پیوندمحور تغییر شکل یافته و ۱۸ ویژگی دیگر محتوامحور هستند که نشان از اهمیت بیشتر این دسته از ویژگی‌ها دارد.

میزبان^{۲۳} است. از این صفحات سه دسته ویژگی با عناوین ویژگی‌های مبتنی بر محتوا، ویژگی‌های مبتنی بر پیوند و ویژگی‌های مبتنی بر پیوند تبدیل یافته استخراج شده است.

- ویژگی‌های مبتنی بر محتوا

این دسته از ویژگی‌ها بر روی محتوای صفحات وب تمرکز دارند. ویژگی‌هایی از قبیل تعداد کلمات، میانگین طول کلمه، نرخ فشرده‌سازی و آنتروپی صفحه جزو این دسته از ویژگی‌ها محسوب می‌شوند. در این مجموعه داده ۹۶ ویژگی مبتنی بر محتوا استخراج شده است.

- ویژگی‌های مبتنی بر پیوند

این دسته از ویژگی‌های بر روی پیوندهای موجود در صفحات وب تمرکز دارند. از جمله این ویژگی‌های می‌توان تعداد پیوندهای خروجی صفحه اصلی، تعداد پیوندهای ورودی صفحه اصلی و نسبت تعداد پیوندهای خروجی داخلی به کل پیوندهای خروجی را نام برد. در مجموعه داده استفاده شده از هر میزبان ۴۱ ویژگی مبتنی بر پیوند استخراج شده است.

- ویژگی‌های مبتنی بر پیوند تبدیل یافته

این دسته از ویژگی‌های شامل تبدیلات عددی ساده و ترکیب ویژگی‌های مبتنی بر پیوند می‌شود. از جمله این ویژگی‌های می‌توان به لگاریتم پیوندهای ورودی به صفحه اصلی، لگاریتم پیوندهای خروجی از صفحه اصلی و نسبت بین رتبه صفحه و پیوندهای ورودی به صفحه اصلی اشاره کرد. در مجموعه داده استفاده شده از هر میزبان ۱۳۸ ویژگی مبتنی بر پیوند تبدیل یافته استخراج شده است.

۳-۲- معیارهای ارزیابی

از آنجا که مجموعه داده مورد استفاده توزیع نامتوازنی دارد (۹۴٪ غیرهرز و ۶٪ هرز) به منظور ارزیابی هر چه بهتر ویژگی‌های از معیار دیگری به نام IBA استفاده کردیم. این معیار تلاش دارد تا توازن بین میزان دقت کلی بایاس نشده و میزان غلبه کلاس‌ها بر یکدیگر ایجاد کند. بدین منظور از مفهومی به نام معیار غلبه استفاده می‌شود. این معیار به رابطه بین کلاس‌ها از لحاظ میزان غلبه اشاره دارد و عددی در بازه $[-1, +1]$ است.

$$Dom = TP_{rate} - TN_{rate} \quad (7)$$

شاخص دقت متوازن (IBA) با استفاده از مفهوم معیار نفوذ^{۲۴} معرفی شده و برای محیط‌های نامتوازن مفید است. فاکتور وزن نتایج را مدنظر قرار می‌دهد که دارای نرخ طبقه‌بندی نسبتاً بهتری بر روی کلاس اقلیت است. فرمول IBA به صورت زیر است:

$$IBA_{\alpha}(M) = (1 + \alpha \cdot Dom)M \quad (8)$$

(M) فاکتوری به منظور وزن‌دار کردن معیار کارایی

است که از فرمول زیر محاسبه می‌شود:

$$M = TP \times (1 - FP) \quad (9)$$

منظور از محاسبه این مقدار این است که بدانیم به ازای هر حدس درست در تشخیص صفحات هرز، چه میزان صفحه غیرهرز را به غلط هرز تشخیص داده‌ایم [13].

زیرنویس‌ها

۵- نتیجه

در این مقاله از ترکیب یازده روش جستجو و هشت تابع ارزیابی به منظور انتخاب ویژگی‌های مناسب از مجموعه داده WEBSpam-UK2007 استفاده شد. روش جستجوی رتبه‌بندی با استفاده از تابع ارزیابی مربع کای توانست علاوه بر کاهش قابل توجه ویژگی‌ها از ۲۷۵ به ۳۲ باعث افزایش میزان IBA دقت تشخیص از ۰/۳۳۷ به ۰/۳۶۹ شود. در کارهای بعدی بر آن خواهیم بود که با معرفی خصیصه‌های جدید در کنار این ویژگی‌ها، دقت تشخیص صفحات وب هرز را افزایش دهیم.

مراجع

- 1 Search Method
 - 2 Evaluator Function
 - 3 Greedy
 - 4 Feed Forward
 - 5 Backward
 - 6 Best First
 - 7 Linear Forward Selection
 - 8 Relief
 - 9 Gain Ratio
 - 10 Entropy
 - 11 Scatter
 - 12 Particle Swarm Intelligence (PSO)
 - 13 Ant Colony
 - 14 Bat Search
 - 15 Correlation
 - 16 Chi Squared
 - 17 Consistency
 - 18 Spilt Information
 - 19 Information Gain
 - 20 Nearest Hit
 - 21 Nearest Miss
 - 22 Symmetrical Uncertainty
 - 23 Host
 - 24 Dominance
 - 25 Naïve Bayes
- [1] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay, "Accurately interpreting click through data as implicit feedback," in *28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil, 2005.
 - [2] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 16-28, 2014.
 - [3] M. Gütlein, E. Frank, M. A. Hall and A. Karwath, "Large-scale attribute selection using wrappers," in *The IEEE Symposium on Computational Intelligence and Data Mining*, 2009.
 - [4] F. G. López, M. G. Torres, B. M. Batista, J. M. Pérez and M. Moreno-vega, "Solving Feature Subset Selection Problem by a Parallel Scatter Search," *European Journal of Operational Research*, vol. 169, no. 2, pp. 477-489, 2006.
 - [5] A. Moraglio, C. Di Chio and R. Poli, "Geometric Particle Swarm Optimisation," in *European Conference on Genetic Programming*, 2007.
 - [6] M. Deriche, "Feature Selection using Ant Colony Optimization," in *6th International Multi-Conference on Systems, Signals and Devices*, Djerba, Tunisia, 2009.
 - [7] A. S. S. Rani and R. R. Rajalaxmi, "Unsupervised feature selection using binary bat algorithm," in *2nd International Conference on Electronics and Communication Systems*, Coimbatore, India, 2015.
 - [8] M. Schiezarò and H. Pedrini, "Data feature selection based on Artificial Bee Colony algorithm," *Image and Video Processing*, no. 1, p. 47, 2013.
 - [9] L. C. Molina, L. Belanche and À. Nebot, "Feature selection algorithms: A survey and experimental evaluation," in *International Conference on Data Mining*, 2002.
 - [10] J. Novakovic, "The Impact of Feature Selection on the Accuracy of Bayes Classifier," in *18th Telecommunications forum*, 2010.
 - [11] G.-A. Madara, P. Inese and A. Ludmila, "The Impact of Feature Selection on the Information Held in Bioinformatics Data," *Information Technology and Management Science*, vol. 18, no. 1, pp. 115-121, 2015.
 - [12] C. Castillo, D. Donato, L. Becchetti, P. Boldi, S. Leonardi, M. Santini and S. Vigna, "A Reference collection for Web Spam," *ACM SIGIR Forum*, vol. 40, no. 2, pp. 11-14, 2006.
 - [13] V. García, R. A. Mollineda and J. S. Sánchez, "Index of Balanced Accuracy: A Performance Measure for Skewed Class Distributions," in *4th Iberian Conference on Pattern Recognition and Image Analysis*, Portugal, 2009.
 - [14] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. H. Witten, "The WEKA data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10-18, 2009.
 - [۱۵] سلیمانی. ع و اصدقی. ف، "بررسی تاثیر روش های طبقه‌بندی بر میزان تشخیص صفحات وب اسپم،" در سومین کنفرانس بین المللی پژوهش‌های کاربردی در علوم کامپیوتر و فناوری اطلاعات، تهران، ۱۳۹۴.



شناسایی ندول‌های ریوی در تصاویر CT اسکن قفسه‌سینه با استفاده از خوشه‌بندی سه‌بعدی و الگوریتم RUSBoost

مهدیه سادات حسینی ورکی^۱، کریم فائز^۲

^۱ دانشجوی کارشناسی ارشد، دانشگاه آزاد اسلامی، واحد قزوین، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، قزوین، ایران

Ir.ma.hosseini@gmail.com

^۲ استاد، دانشگاه صنعتی امیرکبیر، دانشکده مهندسی برق، تهران، ایران

kfaez@aut.ac.ir

چکیده

یکی از شایع‌ترین نشان‌های سرطان ریه ندول‌های ریوی می‌باشد که شناسایی آن در مراحل اولیه بسیار حیاتی است. هدف این مقاله پیاده‌سازی سیستمی جهت تشخیص ندول‌های ریوی در تصاویر CT اسکن قفسه‌سینه می‌باشد. بدین منظور در گام نخست پس از استانداردسازی تصاویر، حجم ریه از قفسه‌سینه تقطیع می‌گردد. سپس لازم است کلیه نواحی مشکوک به ندول شناسایی شوند. روش پیشنهادی برای شناسایی این نواحی دارای دو مرحله اصلی می‌باشد. در مرحله اول الگوریتم خوشه‌بندی K-Means با رویکردی متفاوت جهت شناسایی سه‌بعدی نواحی مشکوک به ندول بکار گرفته شده است و سپس در مرحله دوم تبدیل شاخص شکل به منظور بهبود فرآیند خوشه‌بندی استفاده شده است. در گام نهایی کاهش نواحی که به اشتباه به عنوان ندول استخراج شده‌اند، مدنظر می‌باشد. از آنجایی که تعداد نمونه‌های مثبت نادرست به مراتب بیشتر از تعداد نمونه‌های مثبت درست می‌باشد، با مشکل اساسی عدم تعادل مجموعه داده روبرو خواهیم بود. برای رویارویی با این مشکل پس از استخراج ویژگی‌ها از الگوریتم RUSBoost بهره برده‌ایم. روش پیشنهادی روی ۱۳۹ اسکن پایگاه داده LIDC-IDRI، تست گردیده است که نتیجه نهایی دارای حساسیت ۸۹٪ و میانگین تعداد مثبت نادرست ۴ به ازای هر اسکن می‌باشد. با توجه به تنوع اسکن‌ها در پایگاه داده، نتایج حاصل کارایی روش پیشنهادی را تأیید می‌نماید.

کلمات کلیدی

انتخاب ندول‌های کاندید، خوشه‌بندی سه‌بعدی، سیستم کمک تشخیص کامپیوتری (CAD)، عدم تعادل مجموعه داده، کاهش مثبت نادرست (FP).

۱- مقدمه

تکنولوژی توموگرافی رایانه‌ای^۱ (CT) به علت توانایی بالا در مشخص کردن ساختارهای روی هم افتاده ابزاری مناسب برای تشخیص ندول‌های ریوی می‌باشد. اما به دلیل حجم بالای تصاویر CT استفاده از آن به طور مستقیم امری خسته‌کننده است و نیاز به دقت و تجربه پزشک رادیولوژیست دارد. همچنین به دلیل اشکال متنوع ندول‌ها و اتصال یا عدم اتصال آن‌ها به ساختارهای اطراف، خطاهای عدم تشخیص و یا سوءتعبیر ندول در این روش تشخیص می‌تواند اتفاق بیافتد [2]. از طرفی بررسی‌های اخیر نشان می‌دهد، در فرآیند شناسایی ندول میان رادیولوژیست‌ها اختلاف نظر وجود دارد [3]. همین امر

سرطان ریه به عنوان یکی از رایج‌ترین سرطان‌ها همه ساله باعث مرگ بسیاری از افراد در سرتاسر جهان می‌شود که می‌توان با تشخیص در مراحل اولیه بسیاری از بیماران را از مرگ نجات داد. سرطان ریه رابطه‌ی مستقیمی با وجود یا عدم وجود ندول‌های ریوی دارد [1]. ندول ریوی یک کدورت کوچک و گرد یا بیضی‌شکل، با اندازه قطر کمتر از سه سانتی‌متر است که در ریه‌ها رشد می‌کند [1]. از منظر پزشکی، ندول یک دانسیته نسج نرم است که توسط هوای ریه احاطه شده است.

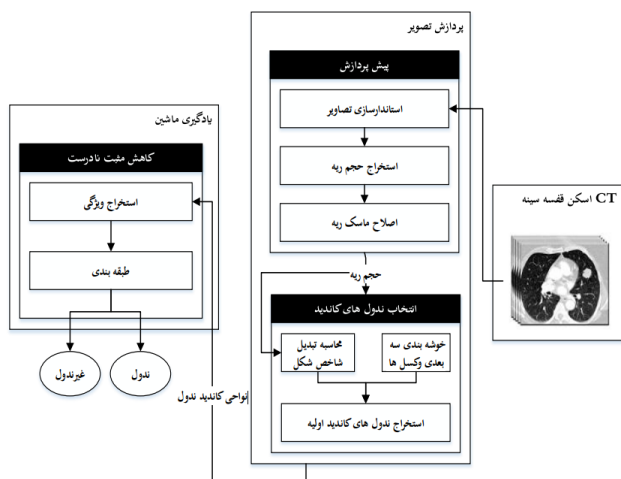
¹ Computed Tomography

۱-۲- نمای کلی

دیاگرام مربوط به سیستم CAD پیشنهادی در شکل ۱ قابل مشاهده است. با توجه به شکل، روند تشخیص در دو فاز پردازش تصویر و یادگیری ماشین، شامل سه گام اساسی به شرح زیر است:

۱. پیش‌پردازش: این مرحله شامل استانداردسازی تصاویر CT اسکن، جداسازی حجم ریه از پس‌زمینه و اصلاح ماسک ریه می‌باشد.
۲. انتخاب ندول‌های کاندید: در این مقاله، یک روش جدید جهت شناسایی ندول‌های کاندید پیشنهاد شده است.
۳. کاهش FP: در رویکرد پیشنهاد شده ابتدا تعدادی ویژگی از نواحی کاندید استخراج می‌گردد و در گام طبقه‌بندی برای رویارویی با مشکل عدم تعادل مجموعه داده از الگوریتم‌های مبتنی بر نمونه‌برداری و یادگیری جمعی استفاده خواهیم کرد.

ادامه مقاله به صورت زیر سازماندهی شده است. در بخش ۲ الگوریتم‌ها و روش‌های بکاررفته به تفصیل شرح داده می‌شود. در بخش ۳ نتایج آزمایشات، تحلیل‌ها و مقایسه با کار دیگران ارائه می‌شود. در بخش ۴ نیز به نتیجه‌گیری و پیشنهادات اختصاص یافته است.



شکل ۱ دیاگرام مربوط به سیستم CAD پیشنهادی برای شناسایی خودکار ندول‌های ریوی در تصاویر CT اسکن قفسه‌سینه.

۲- روش‌ها

۱-۲- پیش‌پردازش

برای رسیدن به نتایج بهتر و دقیق‌تر، ابتدا لازم است پیش‌پردازش‌هایی بر روی تصاویر اعمال شود. یکی از مشکلات تصاویر سوژه‌های مختلف، تفاوت در فاصله بین برش‌ها و در نتیجه تفاوت در رزولوشن در راستای Z می‌باشد [10]. همچنین رزولوشن تصاویر سوژه‌های مختلف در راستای X و Y نیز با هم متفاوت می‌باشد [10]، بنابراین نیاز است تصاویر از نظر سایز با هم یکسان شوند. بدین منظور توسط روش درون‌یابی [11]، سایز وکسل‌های تصاویر در همه نمونه‌ها یکسان و در هر سه راستا برابر ۱/۲۵ میلی‌متر می‌شود.

موجب شده است که در دو دهه اخیر سیستم‌های کمک تشخیص کامپیوتری^۲ (CAD) در جهت تشخیص ندول‌های ریوی توسعه یابند.

۱-۱- کارهای پیشین بر روی تشخیص ندول‌ها

تاکنون پژوهش‌های بسیاری در زمینه توسعه سیستم‌های CAD به منظور تشخیص ندول‌های ریوی توسط محققان علوم کامپیوتر انجام شده است. به طور کلی فرآیند تشخیص ندول در سیستم‌های CAD در دو فاز پردازش تصویر و یادگیری ماشین انجام می‌شود [4]. در فاز پردازش تصویر پس از جداسازی حجم ریه هدف اصلی استخراج نواحی مشکوک به ندول می‌باشد. یکی از مشکلات فرآیند انتخاب ندول‌های کاندید که در اکثر سیستم‌ها وجود دارد، تعداد بسیار زیاد نواحی کاندید که اکثر آن‌ها مثبت نادرست^۳ (FP) است، می‌باشد. به عنوان مثال در [3]، [5] و [6] به ترتیب تعداد ۲۶۹، ۱۵۲۶ و ۳۹۷ به ازای هر اسکن گزارش شده است. این نقص علاوه بر کاهش کارایی سیستم و افزایش زمان محاسبات بعدی، استفاده کلینیکی از راهکار ارائه شده را غیرممکن می‌کند.

در فاز یادگیری ماشین هدف اصلی شناسایی ندول‌ها و رفع FP ها از طریق استخراج ویژگی‌های ندول و طبقه‌بندی در دو گروه ندول و غیرندول می‌باشد. تاکنون کلاس‌های مختلفی جهت کاهش FP ها به کار گرفته شده است که می‌توان به مواردی همچون طبقه‌بندی‌کننده مبتنی بر قانون [5]، ماشین بردار پشتیبان [2]، و ... اشاره نمود. یک مسئله اساسی در این مرحله عدم تعادل مجموعه داده است. از آنجایی که تعداد نمونه‌ها در کلاس مثبت درست^۴ (TP) نسبت به کلاس FP بسیار کمتر است این مشکل رخ می‌دهد. این مسئله با وجود اهمیت بسیار در طبقه‌بندی کاندیدهای ندول کمتر مورد توجه قرار گرفته است و پژوهش‌های محدودی برای حل آن ارائه شده است [7]. به عنوان مثال پنگ کائو و همکاران [7] از ترکیب ماشین بردار پشتیبان حساس به هزینه و الگوریتم بهینه‌سازی ازدحام ذرات جهت حل این مشکل استفاده کرده‌اند. در [8] نیز روش یادگیری بوس‌تینگ تطبیقی^۵ (AdaBoost) با در نظر گرفتن وزن‌های مختلف برای طبقه‌بندی غلط دو کلاس موجود به کار گرفته شده است.

معمولاً گزارشات نهایی در مطالعات موجود براساس میزان حساسیت در تشخیص ندول و نرخ FP به ازای هر اسکن ارائه می‌شود. یکی از مشکلات عمده سیستم‌های موجود بالا بودن نرخ FP می‌باشد. میزان این معیار در پژوهش‌های مختلف نسبت به عملکرد رادیولوژیست‌ها بسیار بالا است. به عنوان مثال در [9] و [10] به ترتیب مقادیر ۲۵/۳ و ۱۰/۳ به ازای هر سوژه گزارش شده است. این در حالی است که نرخ FP در گزارشات رادیولوژیست‌ها حتی کمتر از ۱ به ازای هر سوژه می‌باشد. بنابراین افزایش حساسیت و کاهش نرخ FP یک مسئله مهم و حیاتی است که مطالعات بیشتر و عمیق‌تری را طلب می‌کند.

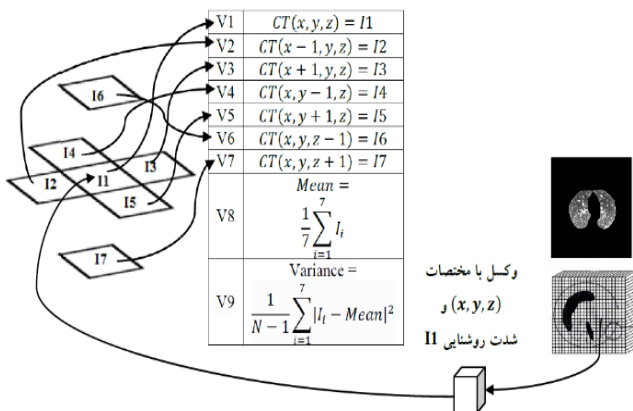
² Computer-Aided Detection

³ False Positive

⁴ True Positive

⁵ Adaptive Boosting

ویژگی مشکل از ساختار همسایگی مرتبه اول سه‌بعدی [5] و میانگین و انحراف معیار شدت روشنایی و کسل‌های موجود در این ساختار همسایگی می‌باشد. چگونگی تشکیل بردار ویژگی برای یک و کسل با موقعیت (x, y, z) و شدت روشنایی I_1 در شکل ۳ قابل مشاهده است. مسئله مهم دیگر استفاده از روش K-Means، انتخاب مناسب مقدار k می‌باشد. براساس بررسی‌های انجام شده در [5] در هر تصویر CT اسکن قفسه‌سینه تعداد چهار خوشه مختلف شامل پارانشیم ریه با فرکانس بالا، پارانشیم ریه با فرکانس پایین، رگ‌های خونی و ندول‌های کاندید وجود دارد. بنابراین در این پژوهش نیز مقدار k برابر با چهار در نظر گرفته شده است. همچنین با توجه به خوشه‌های موجود می‌توان به صراحت بیان کرد که میانگین شدت روشنایی برای خوشه ندول‌های کاندید از بقیه بیشتر است.



شکل ۳ چگونگی تشکیل بردار ویژگی برای یک و کسل با موقعیت (x, y, z) و شدت روشنایی I_1 به منظور خوشه‌بندی سه‌بعدی و کسل‌ها.

۲-۲-۲- محاسبه شاخص شکل

برای هر و کسل $p(x, y, z)$ شاخص شکل به صورت معادله (۱) تعریف می‌شود [10]:

$$SI(p) = 1/2 - 1/\pi \operatorname{atctan} \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \quad (1)$$

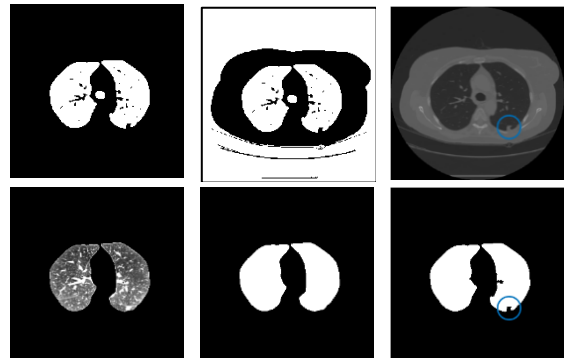
از طریق مقادیر ویژه ماتریس هسین قابل محاسبه است. جهت محاسبه ماتریس هسین نیز می‌توان از معادله (۲) استفاده کرد [10]:

$$H = \begin{bmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{bmatrix} \quad (2)$$

که F ماتریس سه‌بعدی شدت روشنایی هر و کسل و F_{ab} مشتقات جزئی مرتبه دوم در جهت‌های a و b می‌باشد. لازم به ذکر است به منظور کاهش تاثیرات ناشی از نویز و همچنین در نظر گرفتن اندازه‌های مختلف شی، قبل از محاسبه مشتقات مرتبه دوم، تصویر با هسته‌های گوسی فیلتر می‌گردد [14]. بزرگترین و کوچکترین مقدار ویژه در هر و کسل p به ترتیب جایگزین $k_1(p)$ و $k_2(p)$ در معادله (۱) می‌شود.

هر شکل مجزا به یک مقدار یکتا در مقادیر $SI(p)$ مرتبط می‌باشد. برای مثال شاخص شکل برای شکل‌های کروی ۱ و برای شکل-

گام بعدی در پیش‌پردازش تقطیع ریه‌ها از سایر ارگان‌ها و بافت‌ها برای حداکثر نمودن کارایی الگوریتم‌ها می‌باشد. ما در این مقاله از الگوریتم ارائه شده در [12] استفاده کردیم. این الگوریتم در چهار مرحله اصلی شامل آستانه‌گذاری ساده، برچسب‌گذاری سه‌بعدی اجزای متصل [11]، حذف حفره‌ها و اصلاح مرز ریه انجام می‌شود. مقدار آستانه براساس [13] $HU = 500$ در نظر گرفته شده است و اصلاح مرز براساس آنالیز کدهای زنجیره‌ای [12-14] انجام شده است. هدف از اصلاح مرز حفظ ندول‌های متصل به دیواره ریه می‌باشد. شکل ۲ مثالی از روند تقطیع ریه را نشان می‌دهد.



شکل ۲ روند تقطیع ریه براساس [12]. (۱) یک برش از تصویر CT اسکن در صفحه XY که محل ندول نیز مشخص شده است. (۲) نتیجه حاصل از آستانه‌گذاری ساده، (۳) استخراج ناحیه ریه، (۴) برکردن حفره‌ها، (۵) بازسازی مرز ریه (۶) استخراج ناحیه.



۲-۲- انتخاب ندول‌های کاندید

به طور کلی دو تکنیک اصلی جهت شناسایی نواحی مشکوک به ندول معرفی شده است: ۱. تکنیک‌های مبتنی بر شدت روشنایی مانند کوانتیزاسیون برداری [5] و آستانه‌گذاری چندگانه [12]، و ۲. تکنیک‌های مبتنی بر ویژگی‌های شکل مانند شاخص شکل [10-14]. در روش‌های مبتنی بر شدت روشنایی با وجود سادگی محاسبات تعداد FP‌های بسیار زیادی حاصل می‌شود از طرفی در روش‌های مبتنی بر شکل تعداد FP‌ها نسبتاً کم است اما تشخیص ندول‌های غیرکروی دشوار می‌باشد. بنابراین هدف ما در این مقاله ارائه راهکاری ترکیبی از شدت روشنایی و ویژگی شکل جهت بهبود فرآیند انتخاب ندول کاندید با هدف افزایش حساسیت و کاهش نرخ FP می‌باشد.

۲-۲-۱- خوشه‌بندی سه‌بعدی و کسل‌ها

در این مرحله الگوریتم خوشه‌بندی K-Means که یک روش ساده و سریع می‌باشد، به کار گرفته شده است. ایده اصلی استفاده از K-Means برای انتخاب ندول کاندید، تقسیم‌بندی و کسل‌ها براساس توزیع شدت روشنایی محلی به جای استفاده از شدت روشنایی یک و کسل واحد می‌باشد. بدین منظور در این مطالعه یک بردار ویژگی سه‌بعدی برای توصیف هر و کسل به کار گرفته شده است. این بردار

یکی از روش‌هایی که به منظور بهبود عملکرد طبقه‌بندی در مجموعه داده‌های متعادل یا نامتعادل به کار می‌رود، روش‌های یادگیری تجمیعی می‌باشد [15]. این روش‌ها یک مجموعه از دسته‌بندی‌های ضعیف (مانند درخت تصمیم) را گرفته و خروجی آن‌ها را با یکدیگر ترکیب می‌نماید تا دسته‌بندی نهایی را به گونه‌ای بسازد که کارایی آن از

جدول ۱ لیست ویژگی‌های استخراج‌شده از نواحی کاندید.

شماره	ویژگی	گروه
۳-۱	مساحت، اندازه قطر و نرخ مدور بودن	هندسی دوبعدی
۷-۴	حجم، میزان کروی بودن، کشیدگی هندسی و تراکم	هندسی سه‌بعدی
۸	میانگین شاخص شکل	هندسی سه‌بعدی
۱۴-۹	میانگین، انحراف معیار، حداکثر، حداقل، چولگی و کشیدگی	شدت روشنایی دوبعدی
۲۰-۱۵	میانگین، انحراف معیار، حداکثر، حداقل، چولگی و کشیدگی	شدت روشنایی سه‌بعدی
۲۵-۲۱	کنتراست، انرژی، آنتروپی، همگنی و همبستگی	بافت دوبعدی
۳۰-۲۶	کنتراست، انرژی، آنتروپی، همگنی و همبستگی	بافت سه‌بعدی
۳۲-۳۱	میانگین و انحراف معیار گرادیان	گرادیان دوبعدی
۳۴-۳۳	میانگین و انحراف معیار گرادیان	گرادیان سه‌بعدی

کارایی تک تک دسته‌بندی‌های استفاده شده در الگوریتم بیشتر باشد. روش Boosting یک ایده پرکاربرد برای ترکیب دسته‌بندی‌ها در الگوریتم‌های تجمیعی می‌باشد. ترکیب الگوریتم‌های مبتنی بر Boosting و نمونه‌برداری مانند SMOTEBoost [16] و RUSBoost [17] موجب بهبود عملکرد طبقه‌بندی در مجموعه داده‌های نامتعادل می‌گردد. بنابراین با توجه به نامتعادل بودن مجموعه ندول‌های کاندید و همچنین ضرورت استفاده از چندین دسته‌بند به جای یک دسته‌بند واحد در طبقه‌بندی ندول‌ها، در این مقاله از الگوریتم RUSBoost استفاده خواهیم کرد. این الگوریتم که توسط سایفرت و همکاران [17] در سال ۲۰۱۰ پیشنهاد شده است، ترکیب دو الگوریتم داده‌کاهی تصادفی (RUS) و AdaBoost [18] می‌باشد. این الگوریتم T دسته‌بند ضعیف را به صورت تکرارشونده آموزش می‌دهد و مدل نهایی $H(x)$ از طریق رأی‌گیری وزن‌دار T دسته‌بند ضعیف مطابق رابطه (۳) حاصل می‌شود:

$$H(x) = \arg \max_{y \in Y} \sum_{t=1}^T h_t(x, y) \log \frac{1}{\alpha_t} \quad (3)$$

که h_t فرضیه ضعیف و α_t پارامتر به‌روزرسانی وزن برای دسته‌بند t ام است که به صورت رابطه (۴) قابل محاسبه می‌باشد:

$$\alpha_t = \frac{\varepsilon_t}{1 - \varepsilon_t} \quad (4)$$

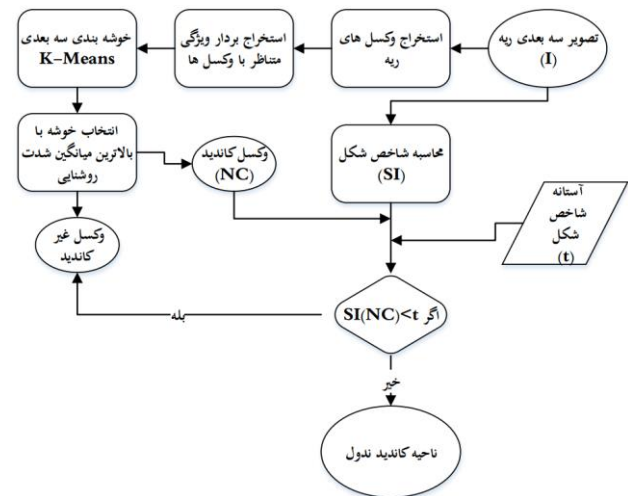
که ε_t خطای طبقه‌بندی وزن‌دار دسته‌بند t است و به اصطلاح pseudo-loss نامیده می‌شود [18].

تاکنون الگوریتم RUSBoost برای تشخیص تومورهای سینه [19] و کلیه [20] به کار گرفته شده است، اما استفاده از آن جهت تشخیص ندول‌های ریوی برای اولین بار می‌باشد.

های استوانه‌ای ۰/۷۵ است. بنابراین شاخص شکل می‌تواند بین شکل‌های مختلف تمایز قائل نماید که قابل استفاده بر روی تصاویر ریه نیز می‌باشد و مقدار شاخص شکل برای ندول‌ها نزدیک به ۱ و برای رگ‌ها در حوالی ۰/۷۵ است.

۲-۲-۳- الگوریتم پیشنهادی جهت شناسایی ندول‌های کاندید

اگر I یک تصویر سه‌بعدی ریه و t مقدار آستانه برای شاخص شکل باشد، مراحل اجرای الگوریتم پیشنهادی برای شناسایی ندول‌های کاندید در شکل ۴ قابل مشاهده می‌باشد. مقدار آستانه t برای شاخص شکل ۰/۹ در نظر گرفته شده است. لازم به ذکر است، در پایان الگوریتم از نواحی کاندید با تعداد وکسل کمتر از سه صرف نظر کرده‌ایم. مقدار ۰/۹ و سه به صورت تجربی تعیین گردیده است.



شکل ۴ الگوریتم پیشنهادی برای شناسایی ندول‌های کاندید با استفاده از خوشه‌بندی سه‌بعدی و تبدیل شاخص شکل.

۲-۲-۳- کاهش مثبت نادرست

۲-۳-۱- استخراج ویژگی

در این مقاله ویژگی‌های استخراج‌شده بر مبنای پژوهش‌های گذشته [2-5-14-10-12] شامل چهار گروه ویژگی‌های هندسی، شدت روشنایی، بافت و گرادیان به صورت دوبعدی و سه‌بعدی می‌باشد. در حالت دوبعدی بزرگترین برش ناحیه کاندید و در حالت سه‌بعدی حجم متعلق به ناحیه تقطیع شده بررسی گردیده است. در جدول ۱ لیست ویژگی‌های استخراج‌شده از هر ناحیه قابل مشاهده می‌باشد.

۲-۳-۲- طبقه‌بندی

در مرحله طبقه‌بندی ندول‌های کاندید دو مشکل عمده وجود دارد: ۱. به دلیل اختلاف بسیار زیاد در حجم، شکل و ظاهر ندول‌ها استفاده از یک طبقه‌بند واحد جهت توصیف و مدل‌سازی داده‌ها کاری دشوار محسوب می‌شود. ۲. دو کلاس موجود نامتوازن است و هزینه طبقه‌بندی نادرست برای این دو کلاس نابرابر است [7].

۳- نتایج آزمایشات و تحلیل‌ها

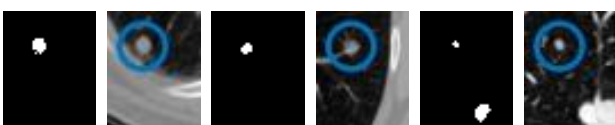
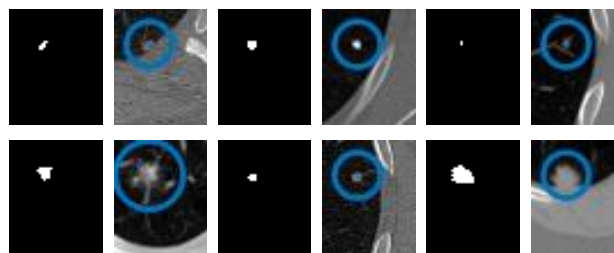
مجموعه آزمایشی ما در این مطالعه شامل ۱۳۹ تصویر CT اسکن قفسه‌سینه با فرمت دایکام (DCM) می‌باشد که از پایگاه داده LIDC-IDRI [21] گرفته شده است. در این مجموعه تعداد اسکن‌های متناظر با چهار دستگاه GE Medical Systems LightSpeed، Philips Brilliance، Siemens و Toshiba به ترتیب برابر با ۱۴، ۸۲، ۳۱ و ۱۲ می‌باشد. میانگین تعداد و ضخامت برش‌ها نیز به ترتیب برابر با ۱۸۱ و ۲/۱۶ میلی‌متر است. اندازه ماتریس تصویر برای هر برش ۵۱۲×۵۱۲ و اندازه پیکسل‌ها در محدوده ۰/۵۳ تا ۰/۹۴ میلی‌متر می‌باشد. محل ندول‌ها در هر اسکن توسط چهار رادیولوژیست مختلف گزارش شده است و مجموعه داده ما شامل تعداد ۲۱۹ ندول با اندازه قطر بین سه تا ۳۰ میلی‌متر می‌باشد که صحت وجود این ندول‌ها توسط هر چهار رادیولوژیست تایید گردیده است. در واقع از سطح توافق حداقل چهار رادیولوژیست استفاده شده است [22].

الگوریتم مطابق مطالب تشریح شده بر روی همه تصاویر مجموعه داده اجرا شده است. در مرحله پیش‌پردازش برای اصلاح مرز ریه در صورتی که فاصله جفت نقاط حاصل از آنالیز کدهای زنجیره‌ای [12] کمتر از ۱۹ میلی‌متر باشد، آن نقاط را به هم متصل می‌کنیم. مقدار ۱۹ میلی‌متر به صورت تجربی تعیین گردیده است. در مرحله محاسبه شاخص شکل هسته‌های گوسی با مقادیر انحراف معیار (Sigma) برابر با ۰/۷۵، ۱/۶، ۳/۴۸ و ۷/۵ میلی‌متر به کار برده شده است. محاسبه این مقادیر بر اساس [14] می‌باشد.

۳-۱- نتایج مراحل انتخاب ندول‌های کاندید

برای ارزیابی الگوریتم انتخاب ندول‌های کاندید از دو معیار درصد هم‌پوشانی و اصابت مرکز ثقل استفاده کرده‌ایم. هر ناحیه کاندید شناسایی شده در صورتی TP تلقی می‌شود که مرکز ثقل آن در ناحیه یکی از ندول‌های واقعی وجود داشته باشد، و یا میزان هم‌پوشانی آن با یکی از ندول‌های واقعی بیش از ۵۰٪ باشد. در صورتی که حداقل یکی از دو شرط بیان شده برقرار نباشد ناحیه کاندید به‌عنوان FP در نظر گرفته می‌شود.

به طور کلی از مجموع ۲۱۹ ندول موجود در پایگاه داده پنج ندول را در مرحله شناسایی ندول‌های کاندید از دست داده‌ایم و مقدار حساسیت ۹۸/۱۷٪ حاصل شده است. در شکل ۵ نمونه‌هایی از ندول‌های شناسایی شده و نتیجه حاصل از شناسایی نواحی کاندید توسط الگوریتم پیشنهادی قابل مشاهده می‌باشد.



شکل ۵ نمونه‌هایی از ندول‌های شناسایی شده و نتیجه حاصل از شناسایی نواحی کاندید توسط الگوریتم پیشنهادی.

از آنجا که روش پیشنهادی ما برای انتخاب ندول‌های کاندید براساس ویژگی شدت روشنایی و شاخص شکل می‌باشد، بهبود قابل ملاحظه‌ای در نرخ FP نسبت به کارهای گذشته حاصل شده است. در جدول ۲ مقایسه‌ی روش پیشنهادی و چند مورد از کارهای گذشته از نظر میانگین تعداد FP به ازای هر اسکن قابل مشاهده می‌باشد.

جدول ۲ مقایسه‌ی روش پیشنهادی جهت شناسایی ندول‌های کاندید و چند مورد از کارهای گذشته از نظر میانگین تعداد FP به ازای هر سوژه.

روش	براساس	FP به ازای هر اسکن
مارفی و همکاران [23]	ویژگی شکل	۸۸۱
نو و همکاران [6]	ویژگی شکل	۳۹۷
هان و همکاران [5]	ویژگی شدت روشنایی	۱۵۲۶
روش پیشنهادی	ویژگی شکل و شدت روشنایی	۹۶

۳-۲- نتایج مراحل کاهش مثبت نادرست

در فاز یادگیری ماشینی جهت آموزش و ارزیابی دسته‌بندی از روش اعتبارسنجی متقابل k (kF-CS) استفاده شده است. چهار معیار حساسیت، ویژه بودن، G-mean [7] و AUC [7] نیز جهت ارزیابی عملکرد طبقه‌بندی به کار گرفته شده است. در جدول ۳ نتایج حاصل از کاهش FP با استفاده از سه شیوه مقابله با دسته‌های نامتعادل [15] قابل مشاهده است. لازم به ذکر است، در همه الگوریتم‌ها از درخت تصمیم برای آموزش مدل استفاده شده است.

با توجه به جدول ۳، نتایج حاکی از آن است که استفاده از روش ترکیبی نمونه‌برداری و یادگیری جمعی بهتر از استفاده تنها از یکی از این دو متد می‌باشد. علیرغم عملکرد تقریباً یکسان RUSBoost و SMOTEBoost، زمان آموزش این دو تفاوت چشمگیری دارد. از آنجا که SMOTE یک تکنولوژی نمونه‌برداری تقویتی است پس با افزایش تعداد نمونه‌ها، SMOTEBoost از طریق مجموعه داده بزرگتری آموزش می‌بیند که موجب طولانی‌تر شدن زمان آموزش شده است. همچنین RUSBoost براساس نمونه‌برداری تصادفی است، پس نسبت به SMOTEBoost پیچیدگی محاسباتی کمتری دارد.

در جدول ۴ نتایج حاصل از الگوریتم پیشنهادی و چند نمونه از رویکردهای مطرح شده جهت کاهش مثبت نادرست بررسی شده است. نمودار آنالیز ROC روش‌های جدول نیز در شکل ۶ قابل مشاهده می‌باشد.

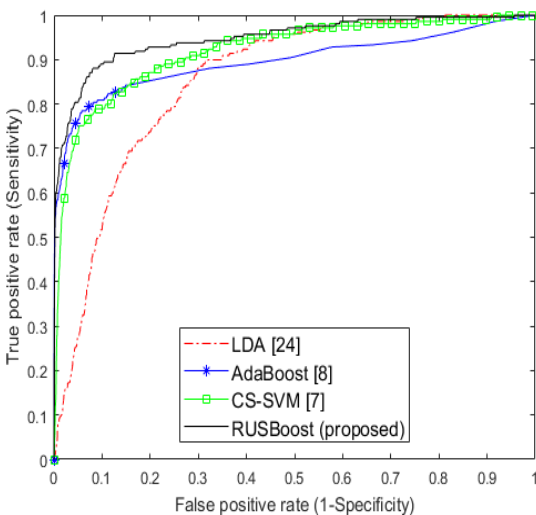
با توجه به جدول ۴ و شکل ۶ الگوریتم پیشنهادی نسبت به سایر روش‌های موجود عملکرد بهتری در کاهش مثبت نادرست ارائه نموده است. پس از روش پیشنهادی نتایج حاصل از CS-SVM [7] نسبت به بقیه بهتر است.

جدول ۳ نتایج حاصل از کاهش مثبت نادرست با استفاده از سه شیوه مقابله با دسته‌های نامتعادل [15].

گروه	الگوریتم	حساسیت	ویژه بودن	G-mean	AUC	زمان آموزش
نمونه‌برداری	RUS [15]	۷۶/۷۹	۹۰/۰۶	۸۳/۸۵	۸۱/۹۱	-
	SMOTE [16]	۵۸/۳۷	۵۱/۱۵	۵۴/۶۴	۵۳/۳۰	-
Boosting	AdaBoost [18]	۶۹/۸۵	۹۸/۹۲	۸۳/۱۳	۸۴/۳۹	-
نمونه‌برداری + Boosting	SMOTEBoost [16]	۸۰/۷۶	۹۵/۴۶	۸۰/۸۷	۸۴/۳۹	۵۱۰/۴
	RUSBoost (روش پیشنهادی)	۸۹/۰۷	۹۵/۴۴	۹۲/۱۱	۹۶/۱۵	۲۸/۶۱

روش‌های طبقه‌بندی دسته‌های نامتعادل در کاهش مثبت نادرست توانا تر است. نتایج بدست آمده با توجه به وجود ندول‌های مختلف و جمع‌آوری داده‌ها متناسب با چهار دستگاه تصویربرداری متفاوت، قابل قبول می‌باشد.

با توجه به عملکرد بسیار خوب الگوریتم RUSBoost در مقابله با دسته‌های نامتعادل، ارائه راهکاری جهت تنظیم اتوماتیک پارامترهای آن می‌تواند به منظور توسعه سیستم پیشنهادی انجام شود. همچنین یکی از نقص‌های سیستم‌های CAD از جمله سیستم پیشنهادی، بالا بودن نرخ FP نسبت به عملکرد رادیولوژیست‌ها می‌باشد. به منظور حل این مشکل استفاده از ترکیب چند طبقه‌بندی کننده تجمیعی در برنامه فعالیت‌های آتی این پروژه می‌باشد.



شکل ۶ نمودار آنالیز ROC جهت مقایسه الگوریتم پیشنهادی و چند نمونه از رویکردهای مطرح شده جهت کاهش مثبت نادرست.

جدول ۵ نتایج حاصل از شناسایی ندول در سیستم CAD پیشنهادی و چند مورد از کارهای گذشته که از پایگاه داده LIDC-IDRI [21] استفاده کرده‌اند.

سیستم CAD	تعداد سوژه	سطح توافق	حساسیت (درصد)	FP به ازای هر اسکن
مسی و همکاران [25]	۹۳	۱	۸۲/۶۶	۳
نمین و همکاران [10]	۶۳	-	۸۸	۱۰/۳
هان و همکاران [5]	۲۰۵	۱	۸۲/۷	۴
نو و همکاران [6]	۱۲۰	۴	۸۸/۶۵	۱۰۷
هوانگ و همکاران [26]	۹۹	۲	۹۰	۵
روش پیشنهادی	۱۳۹	۴	۸۹/۰۷	۴/۱۳

۳-۳- مقایسه نتایج

در جدول ۵ نتایج حاصل از شناسایی ندول در سیستم CAD پیشنهادی و چند مورد از کارهای گذشته که از پایگاه داده LIDC-IDRI [21] استفاده کرده‌اند، قابل مشاهده می‌باشد.

با توجه به جدول ۵، روش پیشنهادی نسبت به روش‌های [5-25] و [6] به ترتیب از لحاظ نرخ حساسیت و FP عملکرد بهتری ارائه نموده است. اگرچه میزان کارایی روش پیشنهادی و روش [10] تقریباً یکسان می‌باشد اما تعداد و تنوع نمونه‌ها در روش ما بیشتر از آن است. روش ارائه شده در [26] مبتنی بر یادگیری عمیق است که دقت آن با روش پیشنهادی تقریباً برابر می‌باشد ولی روش ما دو مزیت نسبت به آن دارد: ۱. تعداد سوژه‌های بیشتری نسبت به [26] مورد بررسی قرار گرفته است (۱۳۹ نسبت به ۹۹)، ۲. ندول‌های متصل به دیواره‌ی ریه در [26] در نظر گرفته نشده است ولی در روش پیشنهادی ما اینگونه نیست. به طور کلی با توجه به تنوع در مجموعه داده‌ها از نظر تعداد سوژه‌ها، سطح توافق، ضخامت برش، اندازه پیکسل‌ها، نوع و اندازه ندول‌ها و تفاوت در روش‌های ارزیابی و اعتبارسنجی، مقایسه بین سیستم‌های CAD بسیار دشوار است [25].

جدول ۴ مقایسه نتایج حاصل از الگوریتم پیشنهادی و چند نمونه از رویکردهای مطرح شده جهت کاهش مثبت نادرست.

روش‌ها	G-mean (%)	AUC (%)
LDA [24]	۷۶/۵۷	۸۵/۳۹
AdaBoost [8]	۸۳/۱۳	۸۴/۳۹
CS-SVM [7]	۸۶/۱۵	۹۰/۱۴
RUSBoost (روش پیشنهادی)	۹۲/۱۱	۹۶/۱۵

۴- نتیجه‌گیری و پیشنهادات

هدف از این مطالعه ارائه یک سیستم کمک تشخیص کامپیوتری برای شناسایی ندول‌های ریوی در تصاویر CT اسکن قفسه سینه می‌باشد. در رویکرد ارائه شده یک روش جدید براساس ویژگی شدت روشنایی و شاخص شکل جهت بهبود فرآیند انتخاب ندول‌های کاندید پیشنهاد گردید. روش پیشنهادی علاوه بر عدم پیچیدگی و افزایش حساسیت، موجب کاهش قابل توجه مثبت نادرست گردیده است. در فاز طبقه‌بندی نیز برای اولین بار الگوریتم RUSBoost به منظور مقابله با عدم تعادل مجموعه داده‌ی ندول‌ها به کار گرفته شده است. براساس معیارهای G-mean [7] و AUC [7]، این الگوریتم نسبت به دیگر

سیاسگزاری

Computer Science and Network (IJCSN), vol. 2, no. 1, 2013 May.

[16] N. Chawla, and et al., "Smoteboost: Improving prediction of the minority class in boosting," *7th European Conference on Principles and Practice of Knowledge Discovery in Databases*, Croatia, 2003 Sep, pp. 107-119.

[17] C. Seiffert, and et al., "RUSBoost: A Hybrid Approach to Alleviating Class Imbalance," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 40, no. 1, 2010 Jan, pp. 185-197.

[18] Y. Freund, and R. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, 1997 August, pp. 119-139.

[19] E. Kozegar, and et al., "Breast cancer detection in automated 3D breast ultrasound using iso-contours and cascaded RUSBoosts," *Ultrasonics*, vol. 79, 2017, pp. 68-80.

[20] A. Skalski, J. Jakubowski, and T. Drewniak, "Kidney tumor segmentation and detection on Computed Tomography data," *IEEE International Conference on Imaging Systems and Techniques (IST)*, Greece, 2016 oct, pp. 238-242.

[21] S. Armato, and et al., "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans," *international journal of medical physics research and practice*, vol. 38, no. 2, 2011 Feb, pp. 915-931.

[22] R. Ochs, and et al., "Forming a reference standard from LIDC data: Impact of reader agreement on reported CAD performance," *Medical Imaging 2007: Computer-Aided Diagnosis*, 2007 March.

[23] K. Murphy, and et al., "Automated detection of pulmonary nodules from low-dose computed tomography scans using a two-stage classification system based on local image features," *SPIE*, vol. 6514, 2007 March, pp. 1-12.

[24] T. Aggarwal and et al., "Feature Extraction and LDA based Classification of Lung Nodules in Chest CT scan Images," *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, india, 2015 Aug, pp. 1189-1193.

[25] T. Messay, R. Hardie, and S. Rogers, "A new computationally efficient CAD system for pulmonary nodule detection in CT imagery," *Medical Image Analysis*, vol. 14, no. 3, 2010 June, pp. 390-406.

[26] X. Huang, J. Shan, and V. Vaidya, "Lung nodule detection in CT using 3D convolutional neural networks," *IEEE 14th International Symposium on Biomedical Imaging (ISBI)*, Australia, 2017 April, pp. 379-383.

نویسندگان مقاله، مراتب قدردانی خود را از پایگاه داده LIDC-IDRI [21] برای در اختیار قرار دادن پایگاه داده استفاده شده در این مقاله به این گروه اعلام می‌دارند.

مراجع

[1] G. Kikano, A. Fabien, and R. Schilz, "Evaluation of the Solitary Pulmonary Nodule," *American Family Physician*, vol. 92, no. 12, 2015 December, pp. 1084-1091.

[2] A. Filho, and et al., "Automatic detection of solitary lung nodules using quality threshold clustering, genetic algorithm and diversity index," *Artificial Intelligence in Medicine*, vol. 60, no. 3, 2013 March, pp. 165-177.

[3] A. Setio, and et al., "Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, 2016 May, pp. 1160-1169.

[4] L. Bogoni, and et al., "Impact of a computer-aided detection (cad) system integrated into a picture archiving and communication system (pacs) on reader sensitivity and efficiency for the detection of lung nodules in thoracic ct exams," *J Digital Imaging*, vol. 25, no. 6, 2012 Dec, pp. 771-781.

[5] H. Han, and et al., "Fast and Adaptive Detection of Pulmonary Nodules in Thoracic CT Images Using a Hierarchical Vector Quantization Scheme," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, 2015 March, pp. 648-659.

[6] J. Novo, and et al., "3D lung nodule candidate detection in multiple scales," *IAPR International Conference on Machine Vision Applications (MVA)*, Tokyo, 2015 May, pp. 61-64.

[7] P. Cao, D. Zhao, and O. Zaiane, "Measure Oriented Cost-Sensitive SVM for 3D nodule detection," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Japan, 2013 July, pp. 3981-3984.

[8] M. Dolejsi, and et al., "Reducing false positive responses in lung nodule detector system by Asymmetric AdaBoost," *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2008 May, France, pp. 656-659.

[9] Y. Lee, and et al., "Automated detection of pulmonary nodules in helical ct images based on an improved template-matching technique," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, 2001 July, pp. 595-604.

[10] S. Namin, and et al., "Automated Detection and Classification of Pulmonary Nodules in 3D Thoracic CT Images," *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Turkey, 2010 Oct, pp. 3774-3779.

[11] RC. Gonzalez, and RE. Woods, *Digital Image Processing (3rd Edition)*, Prentice-Hall, Inc. Upper Saddle River, NJ, 2006.

[12] W. Choi, and T. Choi, "Genetic programming-based feature transform and classification for the automatic detection of pulmonary nodules on computed tomography images," *Information Sciences*, vol. 212, 2012 Dec, pp. 57-78.

[13] M. Brown, and et al., "Method for segmenting chest CT image data using an anatomical model: preliminary results," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, 1997 Dec, pp. 828-839.

[14] X. Ye, and et al., "Shape-based computer-aided detection of lung nodules in thoracic CT images," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 7, 2009 July, pp. 1810-1820.

[15] R. Longadge, S. Dongre, and L. Malik, "Class Imbalance Problem in Data Mining: Review," *International Journal of*



ارائه مدل ترکیبی متشکل از الگوریتم خفاش و اتوماتای یادگیر برای بهینه‌سازی در محیط‌های ایستا

آرش ملکی نودهی^۱، حمیدرضا کرماجانی^۲، مانی زارعی^۳

^۱ دانشجوی، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد صفادشت، تهران
maleki1368@hotmail.com

^۲ مربی، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد تویسرکان، همدان
hr.kermajani@gmail.com

^۳ استادیار، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد شهر قدس، تهران
m.zarei@qodsiau.ac.ir

چکیده

هرچه قدر پیچیدگی مسائل بهینه‌سازی دنیای واقعی افزایش می‌یابد، نیاز بشر به روش‌ها و تکنیک‌های بهینه‌سازی قوی‌تر به منظور حل این مسائل نیز بیشتر می‌شود. در این راستا علاوه بر ارائه روش‌های نوین بهینه‌سازی، بخش عمده‌ای از پژوهش‌ها بر بهبود روش‌های جاری تمرکز دارند. الگوریتم‌های تکاملی و هوش جمعی دسته‌ای از روش‌های بهینه‌سازی هستند که به دلیل مزایای چشم‌گیرشان نسبت به روش‌های کلاسیک، مورد توجه بسیاری از محققان و مراکز تحقیقاتی قرار گرفته‌اند. از آنجا که بهبود الگوریتم‌های ملهم از طبیعت با استفاده از تکنیک تنظیم پارامتر یک حوزه‌ی بسیار پر طرفدار در پژوهش‌های هوش مصنوعی می‌باشد، در این مقاله قصد داریم تا کارایی الگوریتم خفاش را با استفاده از این تکنیک بهبود دهیم. به این منظور از اتوماتای یادگیر بعنوان یک ابزار تصمیم‌سازی وقتی استفاده می‌کنیم تا با استفاده از یادگیری تقویتی رفتار الگوریتم خفاش را در محیط کنترل کنیم.

کلمات کلیدی

الگوریتم خفاش، تنظیم پارامتر، اتوماتا، بهینه‌سازی

با این وجود، از آنجا که پیچیدگی مسائل بهینه‌سازی دنیا بطور مداوم در حال افزایش است، بهبود روش‌های بهینه‌سازی نیز امری اجتناب ناپذیر است. در این راستا علاوه بر ارائه روش‌های نوین بهینه‌سازی، بخش عمده‌ای از پژوهش‌ها بر بهبود روش‌های جاری تمرکز دارند.

بطور کلی روش‌های ارائه شده برای بهبود الگوریتم‌های اتفافی مبتنی بر جمعیت را می‌توان به ۷ گروه مختلف تقسیم بندی کرد:

۱. تغییر الگوی مقداردهی اولیه افراد الگوریتم‌ها [1] [2]
۲. افزودن عملگرها و پارامترهای جدید به الگوریتم‌ها یا تغییر عملگرها و پارامترهای موجود [3] [4]
۳. بهبود الگوریتم‌ها با استفاده از تکنیک تنظیم پارامتر [5]

۱- مقدمه

بهینه‌سازی در بسیاری از شاخه‌های علوم و مهندسی کاربردهای گسترده‌ای دارد. از نمونه‌های کاربردی بهینه‌سازی در دنیای واقعی می‌توان به بهینه‌سازی توابع ریاضی، طراحی و بهینه‌سازی مدارات چاپی، طراحی مدل‌های محاسباتی مخازن نفتی، مهندسی کنترل اشاره کرد. طیف وسیعی از مسائل بهینه‌سازی با تکنیک‌های ریاضی دقیق قابل حل نیستند، از این رو روش‌های دیگری برای حل این دسته از مسائل لازم است. در این میان، پژوهش‌ها حاکی از آن هستند الگوریتم ملهم از طبیعت^۱ و فرآیندهای فیزیکی قادرند تا با موفقیت به بهینه‌سازی این دسته از مسائل بپردازند.

- همه خفاش‌ها از انعکاس صدا برای تشخیص فاصله استفاده می‌کنند و تفاوت بین مواد غذایی و موانع پیش روی خود را می‌دانند.
- پرواز خفاش‌ها به طور تصادفی با سرعت v_i در مکان x_i با فرکانس ثابت f_{min} و طول موج مختلف λ و بلندی صوت A به منظور شکار طعمه صورت می‌گیرد. همچنین آنها می‌توانند به طور خودکار امواج پخش شده و نرخ پالس‌های ارسالی خود را ($r \in [0,1]$) با توجه به نزدیکی شکارشان تنظیم کنند.
- با توجه به اینکه ممکن است بلندی صدا در بسیاری از جهات مختلف متفاوت باشد لذا فرض می‌کنیم که بلندی صدا از R_{max} (بیشترین مقدار) تا R_{min} (کمترین مقدار) متغیر می‌باشد.

طبق قوانین بیان شده، سرعت v_i^t و مکان x_i^t برای خفاش مجازی i ام در زمان t و همچنین فرکانس f_i با استفاده از روابط زیر محاسبه می‌شود:

$$f_i = f_{min} + (f_{max} - f_{min}) \times \beta \quad (1)$$

$$v_i^{t+1} = v_i^t + (x_i^t - x^*) \times f_i \quad (2)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (3)$$

β در رابطه (۱) یک بردار تصادفی با توزیع یکنواخت می‌باشد و x^* بهترین مکان فعلی است که در هر تکرار پس از مقایسه‌ی موقعیت خفاش‌ها مشخص می‌شود. معمولاً فرکانس f با $f_{min} = 0$ و $f_{max} = 100$ در نظر گرفته می‌شود و بصورت یکنواخت در بازه $[f_{min}, f_{max}]$ مشخص می‌شود. در هر تکرار، در بخش جستجوی محلی یکی از جواب‌ها از میان بهترین جواب‌ها انتخاب شده و موقعیت جدید هر خفاش به طور محلی با گام تصادفی به صورت زیر به روز می‌شود:

$$x_{new} = x_{old} + \epsilon A^{(t)} \quad (4)$$

که در آن $\epsilon \in [-1,1]$ یک عدد تصادفی، و $A^{(t)} = \langle A_i^t \rangle$ میانگین بلندی صدای خفاش‌ها در گام زمانی t می‌باشد. از دیدگاه پیاده‌سازی، بهتر است تا یک پارامتر کنترلی بمنظور کنترل اندازه‌ی قدم‌ها در نظر گرفته شود. به همین دلیل فرمول (۴) به شکل زیر ویرایش می‌شود:

$$x_{new} = x_{old} + \sigma \epsilon_t A^{(t)} \quad (5)$$

در فرمول (۵) ϵ_t یک عدد تصادفی با توزیع نرمال $N(0,1)$ ، و σ فاکتور مقیاس^۱ است. پارامتر σ در پیاده‌سازی یانگ برابر با ۰٫۰۱ در نظر گرفته شده است.

همچنین بلندی صدای A_i و نرخ پالس ارسالی r_i در هر تکرار به صورت زیر به روز می‌شود:

$$A_i^{t+1} = \alpha A_i^t, \quad r_i^{t+1} = r_i^0 [1 - \exp(-\gamma t)] \quad (6)$$

که در آن α و γ مقادیر ثابت هستند و برای هر $0 < \alpha < 1$ و $0 < \gamma$ وقتی $t \rightarrow \infty$ میل کند روابط زیر برقرار می‌باشد:

$$A_i^t \rightarrow \infty, \quad r_i^t \rightarrow r_i^0 \quad (7)$$

با توجه به مقدمه‌ای که بیان شد می‌توان الگوریتم خفاش را بصورت زیر بیان کرد:

۴. بهبود الگوریتم‌ها با استفاده از تکنیک چندجمعیتی [6]
 ۵. الگوریتم‌های ترکیبی [7]
 ۶. استفاده از مکانیزم‌های تبادل اطلاعات میان افراد الگوریتم [8]
 ۷. استفاده از ایده‌ی موازی‌سازی [9]
- خوانندگان محترم میتوانند برای اطلاعات بیشتر به مرجع [10] مراجعه کنند.

۱-۲- بهبود الگوریتم‌های تکاملی و هوش

جمعی با استفاده از تکنیک تنظیم پارامتر

یکی از روش‌های رایج بمنظور بهبود کارایی الگوریتم‌های تکاملی و هوش جمعی، برقراری تعادل میان قابلیت اکتشاف^۲ و استخراج^۳ این روش‌ها از طریق تنظیم پارامترهای کنترلی آنها می‌باشد. بطور کلی می‌توان روش‌های تنظیم پارامتر در الگوریتم‌های مختلف را به ۳ دسته تقسیم کرد که در ادامه هر یک از آنها بطور جداگانه بررسی می‌شوند.

۱-۲-۱- در نظر گرفتن مقادیر ثابت یا تصادفی برای پارامترها

ابتدایی ترین روش بمنظور یافتن مقدار (یا دامنه مقادیر) مناسب برای پارامترهای یک الگوریتم، استفاده از روش آزمون و خطا بر مبنای انجام آزمایشات مختلف است. این کلاس از روش‌ها شامل استراتژی‌هایی هستند که در آنها مقادیر پارامترها در حین فرآیند جستجو یا ثابت است و یا بصورت تصادفی از یک دامنه‌ی از پیش تعریف شده انتخاب می‌شود. گروهی از پژوهشگران با تکیه بر این روش تلاش کرده‌اند تا مقادیر مناسبی را برای پارامترهای مختلف الگوریتم‌های تکاملی و هوش جمعی مشخص کنند.

۱-۲-۲- استفاده از استراتژی پارامترهای وابسته به زمان

یکی دیگر از روش‌های معمول برای تنظیم پارامترهای الگوریتم‌های تکاملی و هوش جمعی، تغییر مقادیر پارامترها بر اساس تابعی وابسته به زمان یا تعداد تکرارهای الگوریتم می‌باشد.

۱-۲-۳- تنظیم پارامتر به روش تطبیقی

آخرین دسته از روش‌های تنظیم پارامتر استراتژی‌هایی هستند که وضعیت جستجو را تحت نظر می‌گیرند و مقادیر پارامترهای الگوریتم را بر اساس یک یا چند پارامتر بازخورد^۴ با وضعیت جدید تطبیق می‌دهند.

۲- الگوریتم خفاش

الگوریتم خفاش یکی از الگوریتم‌های فراابتکاری الهام گرفته شده از طبیعت است که در سال ۲۰۱۰ توسط یانگ معرفی گردید [11].

الگوریتم بهینه‌سازی خفاش، از شیوه ردیابی خفاش‌های کوچک با استفاده از انعکاس صدا الهام گرفته شده است. خفاش‌های کوچک می‌توانند در تاریکی مطلق با انتشار صوت و دریافت انعکاس آن طعمه‌های خود را مکان‌یابی کنند. بمنظور سهولت توسعه این الگوریتم یانگ سه فرض ایده‌آل زیر را در نظر گرفته است:

الگوریتم ۱. الگوریتم خفاش

۱. جمعیت اولیه خفاش‌ها را با x_i و v_i ($i = 1, 2, \dots, n$)، بصورت تصادفی در فضای جستجو مقداردهی کنید.
۲. برای تمام خفاش‌ها فرکانس f_i ، نرخ پالس r_i و بلندی صدای A_i مقداردهی کنید.
۳. t را برابر با ۱ قرار دهید.
۴. خفاش‌ها را مرتب کرده و پاسخ بهینه x^* را بیابید.
۵. موقعیت‌های جدید موقتی را با تنظیم فرکانس و به روز کردن سرعت همه خفاش‌ها ایجاد کنید (معادلات (۱) تا (۳)).
۶. اگر $r_i < rand$ ، آنگاه یک جواب در میان بهترین جواب‌ها انتخاب کرده و با استفاده از معادله (۵) یک جواب محلی در اطراف جواب انتخاب شده ایجاد کنید.
۷. یک پاسخ جدید با پرواز تصادفی تولید کنید.
۸. اگر $A_i < rand$ و $f(x_i) < f(x^*)$ ، آنگاه جواب‌های جدید را بپذیرید و طبق معادله (۶) مقادیر r_i و A_i را بروزرسانی کنید.
۹. اگر t به ماکزیمم مقدار خود رسید الگوریتم را متوقف کنید در غیر این صورت قرار دهید $t = t + 1$ و به ۴ بروید.

این الگوریتم تاکنون با موفقیت در کاربردهایی مانند بهینه‌سازی، دسته‌بندی، پردازش تصویر، گزینش ویژگی، زمان بندی، داده‌کاوی و بسیار کاربردهای علمی و مهندسی دیگر بکار برده شده است [12].

۳- اتوماتا یادگیر

مفهوم اتوماتای یادگیر نخستین بار توسط سیتلین^۸ مطرح شد. او یک اتوماتای قطعی را در تعامل با محیطی تصادفی بعنوان مدلی برای یادگیری معرفی کرد. نارندرا^۹ و ویس واناتان^{۱۰} هدف سیستم یادگیر را به صورت مسئله یافتن اقدام بهینه از میان مجموعه‌ای متناهی از اقدام‌های مجاز یک اتوماتای اتفاقی در نظر گرفته‌اند.

۳-۱- اتوماتای اتفاقی

هر اتوماتای اتفاقی بصورت پنج‌تایی $SLA = \{\alpha, \beta, F, G, \phi\}$ تعریف می‌شود که در آن $\alpha = \{\alpha_1, \dots, \alpha_r\}$ مجموعه اعمال اتوماتا، $\beta = \{\beta_1, \dots, \beta_r\}$ مجموعه ورودیهای اتوماتا، $F: \phi \times \beta \rightarrow \phi$ تابع تولید حالت جدید اتوماتا بر اساس حالت فعلی و ورودی اتوماتا، $G: \phi \rightarrow \alpha$ تابع خروجی اتوماتا و $\phi = \{\phi_1, \dots, \phi_n\}$ مجموعه حالت‌های داخلی اتوماتا می‌باشد. اگر نگاشته‌های F و G قطعی باشند، اتوماتای یک اتوماتای قطعی^{۱۱} نامیده می‌شود در غیر این صورت اتوماتای یک اتوماتای اتفاقی نامیده می‌شود.

در اتوماتای یادگیر با ساختار متغیر، تغییر بردار احتمال اعمال اتوماتا با استفاده از الگوریتم یادگیری انجام می‌شود. در اتوماتای یادگیر با ساختار متغیر، حالت داخلی اتوماتا ϕ توسط احتمالات اعمال اتوماتا نشان داده می‌شوند. وضعیت داخلی اتوماتا در گام زمانی n یعنی $\phi(n)$ با بردار احتمال اعمال اتوماتا $P(n)$ مشخص می‌شود:

بطوریکه شرط زیر باید برقرار باشد:

$$\sum_{i=1}^r p_i(n), \forall n, p_i(n) = \text{Probability}[\alpha(n) = \alpha_i] \quad (۸)$$

در ابتدا احتمال کلیه اعمال اتوماتا با هم برابر و مساوی با $\frac{1}{r}$ می‌باشد [14]. (که r تعداد اعمال اتوماتا است.)

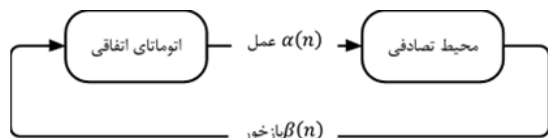
۳-۲- محیط

محیط را می‌توان با استفاده از یک سه تایی $E = \{\alpha, \beta, c\}$ نشان داد که در آن $\alpha = \{\alpha_1, \dots, \alpha_r\}$ مجموعه ورودی‌های محیط، $\beta = \{\beta_1, \dots, \beta_r\}$ مجموعه خروجی‌های محیط و $c = \{c_1, \dots, c_r\}$ مجموعه احتمالات جریمه متناظر می‌باشند [14].

ورودی محیط یکی از r عمل انتخاب شده توسط اتوماتا است. خروجی (بازخورد) محیط به هر عمل i اتوماتا توسط بازخورد β_i مشخص می‌شود. در چنین محیطی $\beta_i(n) = 1$ به عنوان پاسخ نامطلوب^{۱۲} یا شکست^{۱۳} و $\beta_i(n) = 0$ به عنوان پاسخ مطلوب^{۱۴} یا موفقیت در نظر گرفته می‌شوند. مجموعه c احتمالات جریمه پاسخ‌های محیطی را مشخص می‌کند که به صورت زیر تعریف می‌شود:

$$c_i = \text{Probability} [\beta(n) = 1 | \alpha(n) = \alpha_i] \quad i = \{1, \dots, r\}$$

در محیط‌های ایستا^{۱۵} مقادیر احتمالات جریمه (c_i ها) ثابت هستند. ارتباط اتوماتای تصادفی و محیط را بصورت شماتیک نشان می‌دهد.



شکل ۱ اتوماتای یادگیر تصادفی

۳-۳- الگوریتم‌های یادگیر

الگوریتم یادگیری T بصورت $p(n+1) = T[\alpha(n), \beta(n), p(n)]$ نشان داده می‌شود. ماهیت الگوریتم یادگیری تقویت بر اساس الگوریتم یادگیری T مشخص می‌شود، به این صورت که اگر T یک عملگر خطی باشد، الگوریتم یادگیری تقویتی، خطی نامیده می‌شود و در غیر این صورت الگوریتم یادگیری غیرخطی نامیده می‌شود.

اگر اتوماتای یادگیر در گام زمانی n م، یک عمل از میان مجموعه اعمال خود مانند α_i را انتخاب و به محیط اعمال کند و یک پاسخ مطلوب^{۱۶} $\beta_i(n) = 0$ از محیط دریافت نماید، $p_i(n)$ (احتمال عمل α_i) افزایش و احتمال سایر اعمال کاهش می‌یابد. در مقابل اگر پاسخ دریافتی از محیط نامطلوب باشد $\beta_i(n) = 1$ ، احتمال عمل α_i کاهش و احتمال سایر اعمال دیگر اتوماتا افزایش می‌یابد. تمامی این تغییرات به گونه ای صورت می‌گیرد که همواره حاصل جمع احتمالات تک تک اعمال برابر با یک باشد $\sum_{i=1}^r p_i(n) = 1$. احتمالات اعمال اتوماتا نیز به شکل زیر به روز می‌شوند:

[۱] پاسخ مطلوب از محیط

$$p_i(n+1) = p_i(n) + \sum_{j=1}^r f_j [p_j(n)] \quad (۹)$$

$$P(n) = \{p_1(n), p_2(n), \dots, p_r(n)\}$$

می‌کند. در ادامه‌ی این فصل دو الگوریتم پیشنهادی با جزئیات کامل شرح داده می‌شوند. در نهایت کارایی روش‌های پیشنهادی در مقایسه با الگوریتم پایه، با استفاده از توابع محک معروف در حوزه‌ی بهینه‌سازی مورد ارزیابی قرار خواهد گرفت.

۴-۱- الگوریتم خفاش مبتنی بر تنظیم پارامتر بصورت گروهی

در روش پیشنهادی از دو آتوماتای یادگیر به نام‌های LA_A و LA_r برای تنظیم پارامترهای A و r در الگوریتم خفاش استفاده شده است. هر دو اتوماتای فوق دارای ۳ عمل $\alpha_A = \alpha_r = \{increase, decrease, noChange\}$ هستند. این اعمال به ترتیب موجب افزایش، کاهش و یا عدم تغییر مقدار فعلی پارامترهای A و r می‌شوند. در روش پیشنهادی فوق در ابتدای فرآیند بهینه‌سازی مقادیر پارامترهای A و r با مقادیر اصلی آن‌ها یعنی $A=0.25$ و $r=0.5$ تنظیم می‌شوند. پس از آن در هر تکرار، ابتدا اتوماتون‌های متناظر یکی از اعمال خود را بصورت تصادفی و بر اساس توزیع یکنواخت انتخاب می‌کنند و الگوریتم خفاش بر اساس پارامترهای انتخاب شده فرآیند بهینه‌سازی را دنبال می‌کند. سپس، اگر مقدار بهترین خفاش جمعیت نسبت به تکرار قبل بهبود یافته باشد هریک از اتوماتون‌های LA_A و LA_r یک سیگنال تقویتی مطلوب را از محیط دریافت می‌کنند. رابطه زیر این امر را نشان می‌دهد:

$$\begin{aligned} \text{Reinforcement signal} & \quad (15) \\ = \begin{cases} 0 & \text{if } (f(\vec{X}_{best,G}) < f(\vec{X}_{best,G+1})) \\ 1 & \text{otherwise} \end{cases} \end{aligned}$$

در رابطه فوق $f(\vec{X}_{best,G})$ مقدار برآزش بهترین خفاش در نسل G و $f(\vec{X}_{best,G+1})$ مقدار برآزش بهترین خفاش در نسل $G+1$ می‌باشد. پس از آن اتوماتون‌ها بر اساس سیگنال تقویتی دریافتی بردار احتمالات اعمال خود را با استفاده از روابط زیر به‌هنگام می‌کند:

$$p_j(n+1) = \begin{cases} p_j(n) + a \times (1 - p_j(n)) & i = j \\ p_j(n) \times (1 - a) & \forall j \neq i \end{cases} \quad (16)$$

$$p_j(n+1) = \begin{cases} p_j(n) \times (1 - b) & i = j \\ \frac{b}{r-1} + (1 - b) \times p_j(n) & \forall j \neq i \end{cases} \quad (17)$$

فرآیند فوق تا رسیدن به شرط خاتمه ادامه خواهد یافت. شبه‌کد الگوریتم فوق در شکل (۲) نشان داده شده است.

Algorithm 1. Improved bat algorithm with shared parameters (IBWSP)

1. Initialize the algorithm and problem parameters: population size NP , fitness evaluations counter $fe = 0$, maximum fitness evaluations FE_{max} .
2. Initialize a set of LA_A with action set $a = \{Increase, Decrease, No_change\}$, action probability vector $p = \{(1/4), (1/4), (1/2)\}$, $\alpha = 0.1$, and $\beta = 0.01$.
3. Initialize a set of LA_r with action set $a = \{Increase, Decrease, No_change\}$, action probability vector $p = \{(1/4), (1/4), (1/2)\}$, $\alpha = 0.1$, and $\beta = 0.01$.
4. Generate an initial population of bats $P_0 = \{\vec{X}_{1,0}, \dots, \vec{X}_{NP,0}\}$ in the D -dimensional search space

$$p_j(n+1) = p_j(n) - f_j[p_j(n)] \quad \forall j, j \neq i \quad [2] \text{ پاسخ نامطلوب از محیط}$$

$$p_i(n+1) = p_i(n) - \sum_{j=1}^r g_j[p_j(n)] \quad (10)$$

$$p_j(n+1) = p_j(n) + g_j[p_j(n)] \quad \forall j, j \neq i$$

توابع f_j و g_j در روابط فوق دو تابع نامنفی هستند که به ترتیب توابع پاداش و جریمه نامیده می‌شوند. در یک الگوریتم یادگیری تقویتی خطی با تعداد اعمال r ، توابع f_j و g_j به صورت زیر تعریف می‌شوند:

$$f_j[p_j(n)] = ap_j(n) \quad (11)$$

$$0 < a < 1$$

$$g_j[p_j(n)] = \frac{b}{r-1} - bp_j(n) \quad (12)$$

$$0 \leq a < 1$$

در روابط فوق a پارامتر پاداش و b پارامتر جریمه می‌باشد. با استفاده از روابط فوق می‌توان شکل عمومی الگوریتم یادگیری بصورت زیر نوشت:

$$p_j(n+1) = \begin{cases} p_j(n) + a(1 - p_j(n)) & \text{if } i = j \\ (1 - a)p_j(n) & \text{if } i \neq j \end{cases} \quad (13)$$

[۳] پاسخ مطلوب از محیط

$$p_j(n+1) = \begin{cases} (1 - b)p_j(n) & \text{if } i = j \\ \frac{b}{r-1} + (1 - b)p_j(n) & \text{if } i \neq j \end{cases} \quad (14)$$

[۴] پاسخ نامطلوب از محیط

با توجه به مقادیر پاداش و جریمه a و b در روابط فوق، سه حالت مختلف را می‌توان در نظر گرفت:

- اگر مقادیر a و b برابر باشند، اتوماتای یادگیر L_{RP} نامیده می‌شود.
- زمانی که b مساوی با صفر باشد اتوماتای یادگیر L_{RI} نامیده می‌شود.
- اگر $a \ll b$ باشد، اتوماتای یادگیر L_{REP} نامیده می‌شود [14].

۴- الگوریتم پیشنهادی

در هر دو الگوریتم پیشنهادی روند جستجوی الگوریتم خفاش زیر نظر گرفته می‌شود و پارامترهای A و r که در بخش ۳ معرفی شدند بر اساس بازخوردی که از پیشرفت الگوریتم دریافت می‌شود با گذشت زمان تغییر می‌کنند. تفاوت دو الگوریتم پیشنهادی در نوع مکانیزم تنظیم پارامتر آن‌ها می‌باشد به‌گونه‌ای که در اولین الگوریتم مقدار A و r برای تمامی خفاش‌ها به صورت گروهی تنظیم می‌شود (به این معنی که تمام خفاش‌ها مقادیر یکسانی برای پارامترهای خود اختیار می‌کنند)، ولی در الگوریتم دوم هر خفاش مقدار مستقلی را برای پارامترهای خود انتخاب

(IBWDP)

1. Initialize the algorithm and problem parameters: population size NP , fitness evaluations counter $fe = 0$, maximum fitness evaluations FE_{max} .
2. Initialize the set of $LA_A = \{LA_A(1), \dots, LA_A(NP)\}$ with parameters: $a = \{Increase, Decrease, No_change\}$, action probability vector $p = \{(1/4), (1/4), (1/2)\}$, $\alpha = 0.1$, and $\beta = 0.01$.
3. Initialize the set of $LA_r = \{LA_r(1), \dots, LA_r(NP)\}$ with parameters: $a = \{Increase, Decrease, No_change\}$, action probability vector $p = \{(1/4), (1/4), (1/2)\}$, $\alpha = 0.1$, and $\beta = 0.01$.
4. Let $\alpha_j(Z) \mid j \in \{1, 2, 3\}, Z \in [1, \dots, NP]$ be the selected action of $LA_A(Z)$.
5. Let $\alpha_k(Z) \mid k \in \{1, 2, 3\}, Z \in [1, \dots, NP]$ be the selected action of $LA_r(Z)$.
6. Let A_i and r_i be the values for parameters A and r for i -th bat.
7. Generate an initial population of bats $P_0 = \{\vec{X}_{1,0}, \dots, \vec{X}_{NP,0}\}$ in the D -dimensional search space according to uniform random distribution.
8. Let f be the fitness value.
9. Evaluate the objective function values $f(\vec{X}_1), \dots, f(\vec{X}_{NP})$.
10. $fe := NP$
11. **while** ($fe < FE_{max}$)
12. **for** each bat $i \in [1, \dots, NP]$ in current population
13. Select and action for $LA_A(i)$ and $LA_r(i)$ from their action set based on their probability vector p , according to a random uniform distribution.
14. Update A_i and r_i based on selected action.
15. Evolve the i -th bats according to bat algorithm using the selected A_i and r_i .
16. Evaluate the objective function value $f(\vec{x}_i)$.
17. Calculate the reinforcement signal β according to Eq. (18).
18. Update the probability vector $LA_A(i)$ and $LA_r(i)$ using the generated β and Eq. (16) and Eq. (17).
19. **end-for**
20. $G := G + 1$
21. $fe := fe + NP$
22. **end-while**

شکل ۳ شبه کد الگوریتم IBWDP

۴-۳- آزمایشات

در این بخش کارایی استراتژی‌های پیشنهادی بر الگوریتم بهینه‌سازی خفاش در حل مسائل بهینه‌سازی مورد ارزیابی قرار خواهد گرفت. این مسائل شامل مجموعه مسائل شناخته شده و کاملی هستند که به‌طور گسترده در مطالعات حوزه بهینه‌سازی مورد استفاده قرار گرفته‌اند. برخی از این مسائل تک‌قله‌ای و برخی دیگر چندقله‌ای هستند. هدف از این بخش ارزیابی سرعت همگرایی و بررسی قابلیت الگوریتم پیشنهادی در بهینه‌سازی مسائلی با بهینه‌های محلی می‌باشد. علاوه بر این مسائل با ابعاد مختلف مورد آزمایش قرار گرفته‌اند تا تاثیر افزایش ابعاد بر کارایی روش‌های پیشنهادی نیز مورد بررسی قرار گیرد. در

- according to uniform random distribution.
5. Let f be the fitness value.
6. Evaluate the objective function values $f(\vec{X}_1), \dots, f(\vec{X}_{NP})$.
7. $fe := NP$
8. **while** ($fe < FE_{max}$)
9. Select an action for each LA_A and LA_r from their action set based on their probability vector p , according to a random uniform distribution.
10. Evolve the population of bats according to bat algorithm using the selected A and r .
11. Evaluate the objective function value f for all bats.
12. Calculate the reinforcement signal β according to Eq. (15).
13. Update the probability vector LA_A and LA_r using the generated β and Eq. (16) and Eq. (17).
14. $fe := fe + NP$;
15. **end-while**

شکل ۲ شبه‌کد الگوریتم IBWSP

۴-۲- الگوریتم خفاش مبتنی بر تنظیم پارامتر بصورت انفرادی

در این روش هر یک از خفاش‌ها از یک اتوماتای یادگیر مختص خود به منظور تنظیم پارامترهای A و r خودش استفاده می‌کند. این به آن معناست که تعداد اتوماتون‌های یادگیر در این روش دو برابر تعداد خفاش‌ها می‌باشد. در روش پیشنهادی در ابتدا برای هر خفاش i از جمعیت دو اتوماتای یادگیر به نام‌های $LA_A(i)$ و $LA_r(i)$ تعریف می‌شود که وظیفه آن‌ها تنظیم پارامترهای A و r برای خفاش مذکور می‌باشد. هر دو اتوماتای فوق دارای ۳ عمل $\alpha_A = \alpha_r = \{increase, decrease, noChange\}$ هستند که این اعمال به ترتیب موجب افزایش، کاهش و یا عدم تغییر مقدار فعلی پارامترهای A و r برای خفاش i می‌شوند. در آغاز فرآیند بهینه‌سازی مقادیر پارامترهای A و r برای هر خفاش با مقادیر اصلی آن‌ها یعنی $A=0.25$ و $r=0.5$ تنظیم می‌شوند. پس از آن در هر تکرار، ابتدا اتوماتون‌های متناظر با هر خفاش یکی از اعمال خود را بصورت تصادفی و بر اساس توزیع یکنواخت انتخاب می‌کنند و خفاش فوق بر اساس پارامترهای انتخاب شده فرآیند بهینه‌سازی را با استفاده از الگوریتم خفاش دنبال می‌کند. سپس، اگر مجموعه اعمال انتخابی $\langle r_i, A_i \rangle$ موجب بهبود مقدار شایستگی خفاش نام نسبت به تکرار قبل شده باشد هریک از اتوماتون‌های $LA_A(i)$ و $LA_r(i)$ یک سیگنال تقویتی مطلوب را از محیط دریافت می‌کنند. رابطه زیر این امر را نشان می‌دهد:

$$\text{Reinforcement signal} \quad (18)$$

$$= \begin{cases} 0 & \text{if } (f(\vec{X}_{i,G}) < f(\vec{X}_{i,G+1})) \\ 1 & \text{otherwise} \end{cases}$$

در رابطه فوق $f(\vec{X}_{i,G})$ مقدار برآزش خفاش نام در نسل G و $f(\vec{X}_{i,G+1})$ مقدار برآزش خفاش نام در نسل $G+1$ می‌باشد. پس از آن اتوماتون‌های خفاش نام بر اساس سیگنال تقویتی دریافتی بردار احتمالات اعمال خود را با استفاده از روابط (۱۶) و (۱۷) به‌هنگام می‌کنند. فرآیند فوق تا رسیدن به شرط خاتمه ادامه خواهد یافت. شبه‌کد الگوریتم فوق در شکل (۳) نشان داده شده است.

Algorithm 2. Improved bat algorithm with distinct parameters

۱. نتایج بدست آمده توسط الگوریتم IBWDP در ۴ تابع نسبت به الگوریتم پایه بهتر است و نتایج این الگوریتم در هیچ‌یک از توابع نسبت به الگوریتم پایه بدتر نشده است.
۲. روش پیشنهادی IBWSP نسبت به روش پایه در بهینه‌سازی ۴ تابع عملکرد بدتری از خود نشان می‌دهد.
۳. از میان الگوریتم‌های پیشنهادی و الگوریتم پایه الگوریتم IBWDP بهترین و الگوریتم IBWSP بدترین کارایی را در بهینه‌سازی مجموعه توابع محک دارا می‌باشند.
- علاوه بر مقادیر میانگین و انحراف از معیار، به منظور مطالعه بهتر الگوریتم‌های پیشنهادی بهترین و بدترین مقادیر یافت شده توسط آن‌ها را در اجزای مختلف در جدول (۱) گزارش شده است.
- همانطور که می‌توان در جدول (۱) مشاهده کرد، روش IBWDP در اغلب موارد قادر است به نتایج بهتری در بهینه‌سازی توابع محک دست یابد.

۵- نتیجه‌گیری و کارهای آتی

در این مقاله دو رویکرد مختلف به منظور تنظیم پارامترهای بلندی صدای A و نرخ پالس ارسالی r در الگوریتم خفاش بر اساس اتوماتای یادگیر معرفی شده است. در رویکرد نخست، از یک اتوماتای یادگیر به منظور تنظیم پارامترهای کل جمعیت استفاده شده است. در رویکرد دوم، هر خفاش از جمعیت پارامترهای خود را به صورت مستقل تنظیم می‌کند. نتایج آزمایشات انجام شده بر روی برخی از توابع ریاضی معروف نشان دهنده آن است که رویکرد دوم در اغلب موارد به نتایجی قابل قبول و مطلوب‌تر از الگوریتم پایه دست یافته است. به عنوان روال آتی تحقیقات می‌توان از دیگر مدل‌های یادگیری در اتوماتای یادگیر، استفاده از اعمال فازی در اتوماتای یادگیر، و کاربرد روش پیشنهادی در حل یک مسئله مهندسی اشاره کرد.

منابع

- [1] S. Rahnamayan, H. R. Tizhoosh, and M. M. Salama, "Opposition-based differential evolution," *IEEE Trans. Evol. Comput.*, vol. 12, no. 1, pp. 64-79, 2008.
- [2] M. Ali, M. Pant, and A. Abraham, "Unconventional initialization methods for differential evolution," *Appl. Math. Comput.*, vol. 219, no. 9, pp. 4474-4494, 2013.
- [3] R. Thomsen, "Multimodal optimization using crowding-based differential evolution," presented at the Evolutionary Computation, 2004. CEC2004. Congress on, 2004, vol. 2, pp. 1382-1389.
- [4] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," presented at the Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, 1998, pp. 69-73.
- [5] B. Panigrahi, V. R. Pandi, and S. Das, "Adaptive particle swarm optimization approach for static and dynamic economic load dispatch," *Energy Convers. Manag.*, vol. 49, no. 6, pp. 1407-1415, 2008.
- [6] S. Yang and C. Li, "A clustering particle swarm optimizer for locating and tracking multiple optima in dynamic environments," *IEEE Trans. Evol. Comput.*, vol. 14, no. 6, pp. 959-974, 2010.
- [7] J. Sun, Q. Zhang, and E. P. Tsang, "DE/EDA: A new evolutionary algorithm for global optimization," *Inf. Sci.*, vol. 169, no. 3, pp. 249-262, 2005.
- [8] F. Van Den Bergh, "An analysis of particle swarm optimizers," 2007.
- [9] X. Xu, Z. Ji, F. Yuan, and X. Liu, "A Novel Parallel Approach of Cuckoo Search using MapReduce," presented at the 2014 International Conference on Computer, Communications and Information Technology (CCIT 2014), 2014.

تمامی آزمایشات مقادیر پارامترهای الگوریتم خفاش پایه، تعداد خفاش‌ها برابر $Q_{min}=0, Q_{max}=2, A=0.25, r=0.5$ است. در هر دو الگوریتم پیشنهادی میزان پاداش $a=0.1$ و جریمه $b=0.01$ در نظر گرفته شده است. علاوه بر این، گام افزایش یا کاهش پارامترهای الگوریتم خفاش برابر با $0,01$ می‌باشد. به دلیل ماهیت تصادفی الگوریتم‌های مبتنی بر جمعیت، هر آزمایش ۳۰ بار تکرار و از آزمون Wilcoxon rank sum برای بررسی معنادار بودن تفاوت میانگین پاسخ‌های نهایی استفاده شده است. شرط خاتمه‌ی فرآیند بهینه‌سازی نیز رسیدن به $D \times 1000$ ارزیابی تابع هدف در نظر گرفته شده است که D تعداد ابعاد مسئله می‌باشد. نتایج آزمایشات در جدول (۱) گزارش شده است. لازم به ذکر است که هر سه روش بر روی یک کامپیوتر با پردازنده مرکزی Intel(R) Core(TM) i7-6500U @ 2.50 GHz با حافظه 8GB و بر روی سیستم عامل ویندوز ۱۰ اجرا شده‌اند. هر سه الگوریتم در نرم‌افزار Matlab R2016a کدنویسی و اجرا شده‌اند.

۴-۴- بهینه‌سازی توابع معروف

جدول (۱) نتایج عددی روش‌های پیشنهادی در مقایسه با الگوریتم خفاش پایه را نشان می‌دهد. مقادیر کارایی بهترین الگوریتم به صورت توپر نمایش داده شده‌اند. در مواردی که نتایج آزمون Wilcoxon rank sum اختلاف معناداری را میان دو یا چند الگوریتم نشان نمی‌دهد مقادیر کارایی تمام آن‌ها به صورت توپر نمایش داده شده است. جدول ۱ میانگین، انحراف از معیار استاندارد، بهترین مقادیر، بدترین مقادیر کسب شده توسط روش‌های پیشنهادی در مقایسه با الگوریتم پایه بر روی برخی از توابع بهینه‌سازی معروف.

Functions(Value)	Bat		
	Mean	Best	Worst
Sphere	1.03E-03 (1.17E-04)	7.80E-04	1.30E-03
Rastrigin	9.40E+01 (2.91E+01)	5.10E+01	1.75E+02
Rosenbrock	2.57E+01 (2.03E+00)	2.22E+01	2.95E+01
Ackley	1.62E+01 (1.19E+00)	1.27E+01	1.84E+01
Griewank	2.00E+01 (1.58E+01)	1.73E+00	6.54E+01
Hyper-Ellipsoid	3.04E-03 (5.61E-04)	1.69E-03	3.95E-03
Norwegian	1.99E-01 (9.25E-02)	9.57E-06	3.00E-01
Levy	1.76E+01 (1.02E+01)	5.17E+00	5.03E+01
ShiftedGriewank	4.96E+02 (1.04E+02)	2.73E+02	7.45E+02
IBWSP			
Functions(Value)	Mean	Best	Worst
Sphere	3.52E+02 (1.15E+03) ↓	7.92E-04	4.98E+03
Rastrigin	1.13E+02 (3.78E+01) ↓	5.19E+01	2.11E+02
Rosenbrock	2.61E+01 (2.14E+00) ~	2.22E+01	2.95E+01
Ackley	1.63E+01 (1.22E+00) ~	1.37E+01	1.82E+01
Griewank	4.84E+01 (4.70E+01) ↓	7.44E-03	1.43E+02
Hyper-Ellipsoid	1.11E+03 (2.91E+03) ↓	1.96E-03	1.35E+04
Norwegian	1.84E-01 (1.15E-01) ~	1.22E-05	3.28E-01
Levy	1.39E+01 (5.99E+00) ~	4.71E+00	2.49E+01
ShiftedGriewank	5.55E+02 (2.42E+02) ~	1.57E+01	1.10E+003
IBWDP			
Functions(Value)	Mean	Best	Worst
Sphere	1.05E-03 (1.21E-04) ~	7.45E-04	1.25E-03
Rastrigin	9.41E+01 (2.94E+01) ~	3.80E+01	1.51E+02
Rosenbrock	2.60E+01 (2.29E+00) ~	2.19E+01	2.95E+01
Ackley	1.62E+01 (1.06E+00) ~	1.38E+01	1.78E+01
Griewank	7.09E+00 (9.57E+00) ↑	5.19E-05	3.55E+01
Hyper-Ellipsoid	2.65E-03 (5.76E-04) ↑	1.70E-03	4.32E-03
Norwegian	2.94E-03 (9.64E-03) ↑	2.33E-07	4.74E-03
Levy	1.49E+01 (6.19E+00) ~	6.62E+00	3.25E+01
ShiftedGriewank	3.43E+02 (1.70E-02) ↑	5.71E-05	5.94E+02

با بررسی نتایج گزارش شده در جدول (۱) می‌توان به نتیجه‌گیری‌های زیر دست یافت:

[۱۱] آ. ملکی نودهی و پ. غلام نژاد، "طبقه بندی روش های بهبود الگوریتم های تکاملی و هوش جمعی"، دومین کنفرانس بین المللی و سومین همایش ملی کاربرد فناوری های نوین در علوم مهندسی، ۱۳۹۴.

- [11] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," *Nat. Inspired Coop. Strateg. Optim. NICSO 2010*, pp. 65–74, 2010.
- [12] X.-S. Yang and X. He, "Bat algorithm: literature review and applications," *Int. J. Bio-Inspired Comput.*, vol. 5, no. 3, pp. 141–149, 2013.
- [13] A. H. Gandomi, X.-S. Yang, A. H. Alavi, and S. Talatahari, "Bat algorithm for constrained optimization tasks," *Neural Comput. Appl.*, vol. 22, no. 6, pp. 1239–1255, 2013.
- [14] K. S. Narendra and M. A. Thathachar, *Learning automata: an introduction*. Courier Corporation, 2012.

زیرنویس ها

- ¹ Nature-inspired algorithms
- ² Exploration
- ³ Exploitation
- ⁴ Iteration number
- ⁵ Feedback
- ⁶ Scaling factor
- ⁷ Feature selection
- ⁸ Tsetlin
- ⁹ Narendra
- ¹⁰ Viswanatan
- ¹¹ Deterministic Automata
- ¹² Unfavorable
- ¹³ Failure
- ¹⁴ Favorable
- ¹⁵ Stationary
- ¹⁶ Favorable
- ¹⁷ Linear Reward Inaction
- ¹⁸ Linear Reward Epsilon Penalty



ثبت خودکار تصاویر رادار روزنه مصنوعی بر روی تصاویر مرئی

سوسن غلامی^۱، علیرضا احمدی فرد^۲

^۱ دانشکده برق و ریاضیات، دانشگاه صنعتی شاهرود، شاهرود
Sue.gholami@gmail.com

^۲ دانشکده برق و ریاضیات، دانشگاه صنعتی شاهرود، شاهرود
ahmadyfard@shahroodut.ac.ir

چکیده

ثبت خودکار تصاویر SAR به اپتیک یک مشکل چالش برانگیز است زیرا که ویژگی‌های هندسی و رادیمتری دو تصویر با جنس متفاوت، ناسازگار هستند. روش‌های مبتنی بر تطابق ویژگی‌های هندسی عملکرد خوبی ارائه می‌کنند، ولی توانایی آنها به استخراج ویژگی‌های تغییرناپذیر با تبدیل‌های هندسی و روش تطابق بکار گرفته شده بستگی دارد. در این مقاله، روشی جدید برای تطابق تصاویر SAR بر تصویر اپتیکی ارائه می‌شود این روش بر پایه استخراج خطوط تصویر به کمک تبدیل رادون می‌باشد. بر اساس خطوط استخراج شده سلول‌های ورونوی در هریک از دو تصویر تشکیل می‌گردد. برای تطبیق دو تصویر سلول‌های ورونوی متناظر جستجو می‌گردد برای این منظور از فاصله هاسدورف استفاده شده است. نتایج این تحقیق نشان می‌دهد که روش پیشنهادی عملیات تطابق دو تصویر را بهبود می‌بخشد. در مقایسه با الگوریتم‌های پیشین، روش ارائه شده نقاط دقیق و مقاوم تری را استخراج می‌کند این موجب دقت بالاتر در ثبت تصاویر می‌گردد.

کلمات کلیدی

ثبت خودکار تصاویر، تبدیل رادون، نقاط تلاقی، دیاگرام ورونوی، تطابق (جفت کردن).

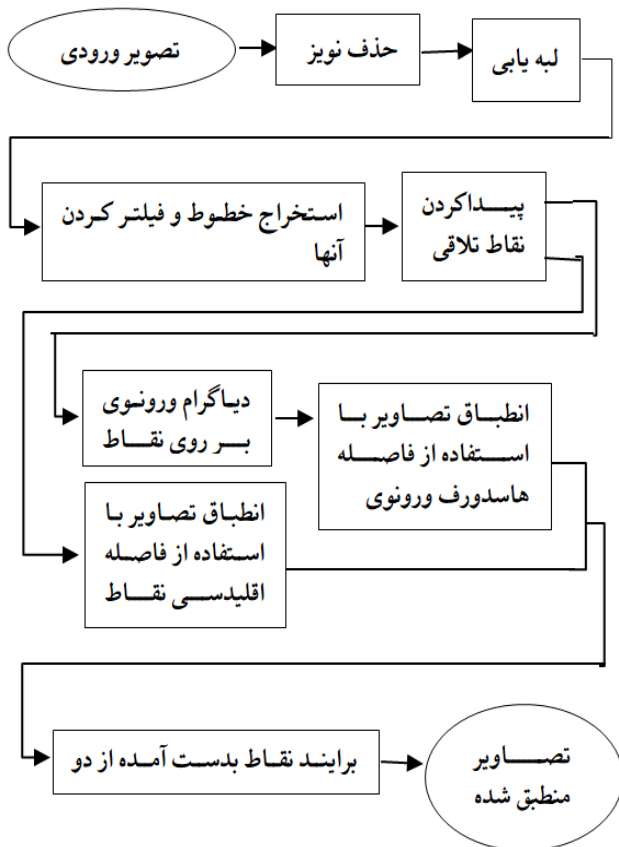
(معمولاً امواج طیف مرئی خورشید) که توسط زمین بازتاب می‌شود را دریافت می‌کنند.

برخلاف سامانه‌های تصویربرداری نوری غیرفعال که به شدت نور خورشید وابسته اند در تصویربرداری رادارهای SAR که از نوع فعال می‌باشد امواج توسط سنجنده (هواپیما) هنگام تصویربرداری به زمین گسیل می‌شود. تصاویر اپتیکی در نور مرئی اطلاعات مهمی را از سطح زمین ثبت می‌کند اما براحتی از جو زمین تأثیر می‌پذیرد و در طول روز امکان پذیر است. در صورتیکه امکان عکس برداری SAR تحت شرایط جوی سخت، روز یا شب وجود دارد اما نویز اسپکل که غیر اجتناب است وضوح این تصویربرداری را محدود می‌کند [2-4]. به منظور استفاده از اطلاعات همزمان تصاویر اپتیکی و SAR لازم است این تصاویر در یک صحنه با هم تطابق یابند. یکی از اهداف تطابق این

۱- مقدمه

ثبت تصاویر، پردازشی است که یک انطباق هندسی بین دو تصویر گرفته شده از دیدهای مختلف یا زمانهای مختلف و یا استفاده از سیگنالهای مختلف و قیدهای متفاوت فراهم می‌نماید [1]. در تصویربرداری توسط سنجنده‌های ماهواره ای و هوایی دو روش کلی وجود دارد: روش فعال و روش غیرفعال. در روش فعال، سنجنده امواج را به سطح زمین می‌تاباند سپس بازتاب آن را دریافت می‌کند این در حالی است که در روش غیرفعال سنجنده موجی ارسال نمی‌کند بلکه تنها امواج موجود در محیط

جهت تطابق تصویر SAR با تصویر اپتیکی ابتدا می‌بایست تصویر SAR حذف نویز می‌شود. الگوریتم مورد استفاده در این مقاله برای حذف نویز استفاده از فیلتر Lee است که به خوبی اثر این نویز را کم می‌نماید. در الگوریتم پیشنهادی پس از حذف نویز تصویر، تصویر لبه یابی شده و جهت استخراج خطوط آن را تبدیل رادون گرفته می‌شود. چون تعداد زیادی خط از صحنه استخراج می‌شود که بسیاری از آنها در یک راستا هستند از بین خطوط با فاصله نزدیک که زوایای کمی نسبت به هم دارند خطوط کوتاهتر حذف می‌شوند. پس از انتخاب خطوط مهم نقاط تلاقی خطوط بدست می‌آید. به هر نقطه تلاقی یک سلول ورونی نسبت داده می‌شود. سلول ورونی هر نقطه یک چند ضلعی پیرامون نقطه است که اضلاع آن عمود منصف خطوطی هستند که آن نقطه را به نقاط همجوار متصل می‌سازد. بعد از محاسبه ورونی نقاط تلاقی، فاصله هاسدورف چندضلعی‌ها سنجیده شده و از ماتریس بدست آمده مقادیر و بردارهای ویژه استخراج می‌شود. تمامی مراحل بالا را برای تصویر نوری نیز به کار می‌رود. پس از آن تطبیق نقاط و چند ضلعی‌های ورونی متناظر از دو تصویر انجام می‌شود.



شکل ۱ بلوک دیاگرام الگوریتم پیشنهادی

دو نوع تصویر این است که چنانچه امکان استفاده از سیستم موقعیت یاب (مثلاً GPS) وجود نداشته باشد با استفاده از تصویر اپتیکی موقعیت تصویر SAR را در نقشه‌های جغرافیایی تشخیص دهیم.

در حوزه ثبت تصاویر تحقیقات زیادی صورت گرفته که در ادامه بطور اجمالی به آن می‌پردازیم.

در سال ۲۰۱۰ آقایان سوری و رنارتز با استفاده از روش رایج و مقایسه شدت روشنایی توانستند تصاویر راداری ماهواره را بر هم منطبق نمایند [5].

در سال ۲۰۱۲ آقایان محمود حسن و همکاران با استفاده از روش آنتروپی باقی مانده متقاطع تجمعی (cross-cumulative residual entropy) تصاویری که از حسگرهای متفاوتی جمع آوری شده بود را بر هم منطبق نمودند [6].

در سال ۲۰۱۴ آقایان گونگ و همکاران با استراتژی تطبیق حدودی تا دقیق روشی ارائه نمودند که در آن ابتدا تطبیق حدودی دو تصویر صحنه انجام می‌شود و سپس تطبیق دقیق بین تصاویر انجام می‌پذیرد [7].

در سال ۲۰۱۰ آقایان ونگ و کلاؤسی بر روی تطبیق تصاویر گرفته شده از حسگرهای متفاوت با باندهای فرکانسی متفاوت کار نمودند. در این روش با استفاده از تبدیل موجک از تصاویر استخراج ویژگی نمودند. این الگوریتم بر روی تصاویر راداری، اپتیکی و لیداری تست شدند که نتایج بهتری نسبت به روش‌های مشابه را در بر داشت [8].

در سال ۱۹۹۲ آقایان شپیرو و بردی براساس نوع توصیف تصویر و ویژگی‌های نقاط، تصاویر را با یکدیگر منطبق کردند [9].

در سال ۲۰۱۵ نیز هایگنگ سو و همکارانش یک روش برای ثبت خودکار تصاویر SAR و اپتیک ارائه نمودند. این روش برپایه ۳ گام مهم بنا شده است. (۱) پروسه تکرار انتخاب خطوط و نقاط تلاقی آنها. (۲) چند مرحله‌ای بودن، از ثبت حدودی تصاویر به ثبت دقیق (۳) دیاگرام ورونی جهت بهبود انطباق تصاویر. که استفاده از سلول‌های ورونی از نوآوری‌های این تحقیق بشمار می‌رود [10].

در این مقاله سعی شده است که با الهام گرفتن از مقالات مرور شده روشی کارا برای ثبت خودکار تصاویر SAR و اپتیکی ارائه شود. در روش پیشنهادی برای استخراج خطوط از تبدیل رادون استفاده شده است. خطوط بدست آمده به کمک این تبدیل بسیار دقیق‌تر از تبدیل هاف که در سایر مقالات پیشنهاد شده است می‌باشند. همچنین در ادامه از تابعی استفاده شده است که تلفیقی از توابع ثبت می‌باشد که به مراتب نتایج دقیق و صحیح‌تری را عرضه خواهد کرد.

۲- تئوری

همانطور که در مقدمه آورده شد مساله پیش رو تطابق دو تصویر از جنس متفاوت است که باید به بررسی آنها پرداخته شود. تصاویر نوری قابلیت تفسیر خوبی دارند اما متاسفانه تحت تأثیر شرایط جوی هستند در حالیکه تصاویر SAR در طول شبانه روز قابل ثبت است اما وجود نویز اسپیکل که ذاتاً هنگام تصویر برداری SAR وجود دارد وضوح این تصاویر را کاهش می‌دهد.



شکل ۴ تصویر ورودی اپتیکی



شکل ۵ تصویر حذف نویز شده اپتیکی با فیلتر Lee

۲-۲- لبه یابی با GGS

لبه یابی یکی از مراحل اصلی پردازش تصاویر SAR می باشد. بیشتر آشکارسازهای مبتنی بر نسبت، از مستطیل دو پنجره استفاده می کنند و لبه های ضخیم را با آستانه گیری ساده ای استخراج می کنند. بسیاری از کاربردهای تصاویر SAR احتیاج به لبه های نازک با عرض یک پیکسل دارند.

ساختار پنجره در GGS دو پنجره تانسور از تابع گاوسی در طول جهت لبه و تابع گاما در جهت عمود به لبه است [12]. مقیاس تابع گاوسی برای تنظیم طول پنجره، و دو پارامتر از تابع گاما برای تنظیم عرض پنجره و فضای بین دو پنجره استفاده می شود. بانک فیلتر با استخراج تغییرات شدت اثر زیادی را در عملکرد لبه یابی می گذارد.

نتایج حاصل از ارزیابی این روش نشان می دهد که لبه یابی با استفاده از GGS دو پنجره نسبت به یکبار استفاده از مستطیل دو پنجره لبه های نادرست کمتری در اطراف لبه های صحیح ایجاد می نماید علاوه بر این، در روش GGS پیوستگی لبه های استخراج شده حفظ می شود [12].

در شکل ۶ و ۷ تصاویر لبه یابی شده SAR و اپتیک را به ترتیب با استفاده از الگوریتم GGS مشاهده می نمایید.

۱-۲- فیلتر حذف نویز Lee

پردازش تصویر یکی از علوم پرکاربرد مهندسی است که در سالهای اخیر پیشرفت قابل توجهی داشته است، یکی از مسائل مورد مطالعه در پردازش تصویر حذف اختلالات ناخواسته نظیر انواع نویز از تصویر است که اغلب به عنوان مرحله پیش پردازش شناخته شده است.

در اثر تداخل سازنده و مخرب امواج بازگشتی در سیستم تصویر برداری فراصوت و یا در تصاویر ماهواره ای، الگویی نقطه ای ضرب شونده ناخواسته که به نویز اسپکل مشهور است، تشکیل می شود.

فیلتر Lee [11] در مناطقی از تصویر که واریانس زیاد است (جزئیات و نزدیک لبه ها) موجب همواری تصویر نخواهد شد. بعکس اگر واریانس در یک منطقه از تصویر پایین باشد جهت حذف نویز تصویر ورودی هموار می شود. فیلتر Lee بر اساس مدل نویز اسپکل ضرب شونده طراحی شده است. این فیلتر از فرض توزیع گاوسی برای نویز در تصویر SAR استفاده می کند. ابتدا تصویر SAR را با یک مدل خطی تقریب زده و سپس با معیار حداقل میانگین مربعات خطا ارزش پیکسل های تصویر اصلاح می شود [11]. الگوریتم فیلتر عبارت است از:

$$\hat{x}_{i,j} = \bar{x}_{i,j} + k_{i,j} (z_{i,j} - \bar{u}_{i,j} \bar{x}_{i,j} - \bar{w}_{i,j})$$

که $u_{i,j}$ و $w_{i,j}$ نویز سفید مستقل هستند و $z_{i,j}$ تصویر نویزی تخریب شده است و

$$k_{i,j} = \frac{\bar{u}_{i,j} Q_{i,j}}{\bar{x}_{i,j}^2 \sigma_2^2 + \bar{u}_{i,j}^2 Q_{i,j} + \sigma_1^2}$$

که $Q_{i,j}$ واریانس پیشین $x_{i,j}$ است.

در شکل ۲ و ۳ تصویر اصلی و تصویر حذف نویز شده SAR و در شکل ۴ و ۵ تصویر اصلی و حذف نویز شده اپتیک را مشاهده می نمایید.



شکل ۲ تصویر ورودی SAR



شکل ۳ تصویر حذف نویز شده با فیلتر Lee

$$Rf(\alpha, s) = \int_{-\infty}^{\infty} f(x(z), y(z)) dz = \int_{-\infty}^{\infty} f((z \sin \alpha + s \cos \alpha), (-z \cos \alpha + s \sin \alpha)) dz \quad (۳)$$

به صورت کلی، در فضای اقلیدسی n بُعدی R^n ، تبدیل رادون یک تابع پیوسته کاملاً پشتیبانی شده f برابر است با تابعی بر روی فضای Σ_n کل ابرصفحه‌ها در R^n (Rf) این تبدیل برای $\xi \in \Sigma_n$ و وقتی که انتگرال با توجه به اندازه طبیعی ابرسطح، $d\sigma$ ، گرفته می‌شود (تعمیم عبارت $|dx|$ از مسئله دو بعدی) بصورت زیر تعریف می‌شود:

$$Rf(\xi) = \int_{\xi} f(x) d\sigma(x) \quad (۴)$$

توجه کنید که هر عنصر Σ_n به صورت مکان هندسی راه‌حلی برای یک معادله توصیف می‌شود:

$$x \cdot \alpha = s \quad (۵)$$

که $\alpha \in S^{n-1}$ یک بردار واحد است و $s \in R$ می‌باشد. بنابراین، تبدیل رادون n بُعدی می‌تواند به صورت تابعی بر روی $S^{n-1} * R$ نوشته شود:

$$Rf(\alpha, s) = \int_{x \cdot \alpha = s} f(x) d\sigma(x) \quad (۶)$$

در شکل ۸ و ۹ به ترتیب اعمال تبدیل رادون را بر تصاویر لبه یابی شده SAR و اپتیک مشاهده می‌شود.



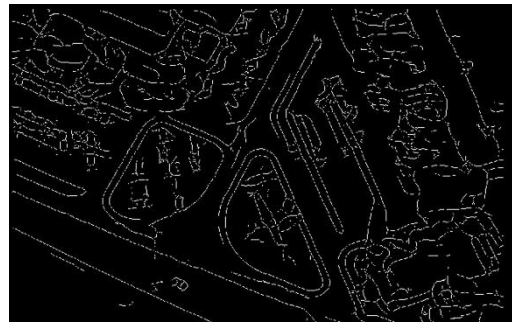
شکل ۸ اعمال تبدیل رادون بر تصویر لبه یابی شده SAR



شکل ۹ اعمال تبدیل رادون بر تصویر لبه یابی شده Optic

۲-۴- فیلتر کردن خطوط

با استخراج خطوط تصاویر توسط تبدیل رادون نوبت به پردازش بر روی خطوط می‌رسد.



شکل ۶ تصویر لبه یابی شده SAR با استفاده از GGS



شکل ۷ تصویر لبه یابی شده اپتیک با استفاده از GGS

۲-۳- تبدیل رادون

در ریاضیات، تبدیل رادون^۲ دو بُعدی، که نام خود را از ریاضی‌دان ژوهان رادون گرفته است، تبدیلی است که از محاسبه انتگرال سیگنال ورودی بر روی خطوط مستقیم در راستاهای مختلف بدست می‌آید. این تبدیل در سال ۱۹۱۷ توسط رادون معرفی شد [13]. وی فرمولی را نیز برای معکوس این تبدیل ارائه نمود.

به خروجی تبدیل رادون، اغلب سینوگرام^۳ گفته می‌شود زیرا تبدیل رادون یک تابع دلتای دیراک، به شکل توزیعی بر روی گراف یک موج سینوسی می‌باشد. در نتیجه، تبدیل رادون تعدادی از اشیاء، ظاهر گرافیکی‌ای همچون چندین موج سینوسی با فاز و دامنه مختلف دارد.

فرض کنید $f(x) = f(x, y)$ تابع پیوسته کاملاً پشتیبانی شده (compactly supported) بر روی R^2 باشد. تبدیل رادون، Rf ، تابع برابر خواهد بود با [13]:

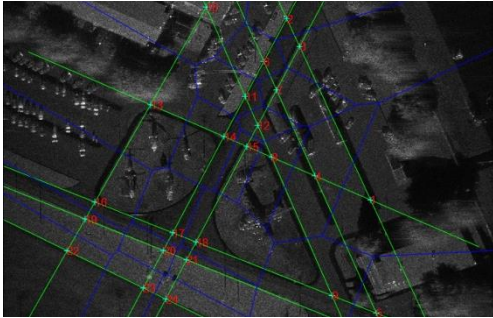
$$Rf(L) = \int_L f(x) |dx| \quad (۱)$$

پارامتریزه کردن هر خط مستقیم L با توجه به طول کمان z همواره می‌تواند به صورت زیر نوشته شود:

$$(x(z), y(z)) = ((z \sin \alpha + s \cos \alpha), (-z \cos \alpha + s \sin \alpha)) \quad (۲)$$

که s فاصله L از مبدا و α زاویه‌ای است که بردار نرمال L با محور x می‌سازد. مقادیر (α, s) می‌توانند به عنوان مختصات روی فضای تمام خطوط در R^2 در نظر گرفته شوند، و تبدیل رادون در این مختصات توسط معادله زیر بیان می‌شود [13]:

در شکل ۱۲ و ۱۳ و ررونوی نقاط تلاقی را به نمایش گذاشته است به گونه ای که خطوط و نقاط تلاقی آنها با خط سبز و ررونوی آن با آبی به تصویر کشیده شده است.

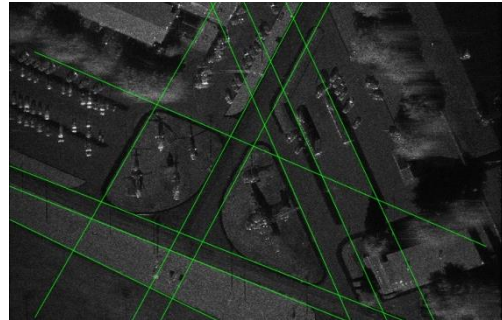


شکل ۱۲ و ررونوی نقاط تلاقی خطوط (آبی)، خطوط منتخب (سبز)



شکل ۱۳ و ررونوی نقاط تلاقی (آبی)، خطوط منتخب (سبز)

در این مرحله باید خطوط نزدیک به هم را فیلتر کرد. در شکل ۱۰ و ۱۱ خطوط باقی مانده پس از فیلتر کردن نشان داده شده است. منظور از خطوط نزدیک به هم، خطوطی است که ρ و θ آنها در دستگاه قطبی اختلاف اندکی با یکدیگر داشته باشند.



شکل ۱۰ خطوط منتخب نهایی در تصویر SAR



شکل ۱۱ خطوط منتخب نهایی در تصویر اپتیکی

۳- تطبیق تصاویر

برای تطابق دو تصویر به یکدیگر الگوریتم هایی پیشنهاد شده اند که در این مقاله سعی بر بهبود تطابق این روشها داریم. در گام اول از تطبیق تصاویر همانند روش بردارهای ویژه عمل کرده، فاصله اقلیدسی نقاط تلاقی را محاسبه نموده و پس از آن ماتریس مجاورت آن $[H_{ij}]$ بدست می آید. به گونه ای که:

$$(7) \quad r_{ij} = \|x_i - x_j\|^2$$

$$(8) \quad H_{ij} = e^{-r_{ij}/2\sigma^2}$$

x ویژگی های هر تصویر است که در مسئله تطابق تصاویر SAR و اپتیکی نقاطی از دو تصویر را در نظر می گیرد و r فاصله اقلیدسی بین دو نقطه از تصویر H همان ماتریس مجاورت است. سپس به محاسبه بردارهای ویژه پرداخته می شود و اگر هر سطر از بردارهای ویژه F نامیده شود با کم کردن اندازه سطرها از یکدیگر تابعی بدست خواهد آمد که بخشی از خروجی را تشکیل می دهد، اگر کل تابع خروجی بوسیله این الگوریتم محاسبه شود تعداد نقاط تطبیق داده شده درست خیلی کم خواهد بود [9].

تعداد جفت نقاط تطبیق یافته در تصویر (۱۴) ۹ جفت است که همه غلط می باشند. زیرا همانگونه که در خود الگوریتم گفته شده است، این الگوریتم برای تصاویری که تغییرات زیادی دارند (در اینجا هم مقیاس متفاوت است و هم چرخش وجود دارد) ناموفق

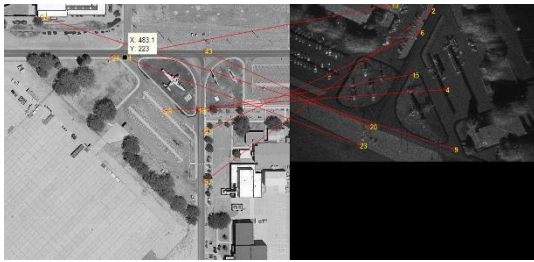
۲-۵- سلول های ورونوی

در علم ریاضیات دیاگرام ورونوی روشی برای تقسیم فضا ویژگی به تعدادی ناحیه می باشد. در این دیاگرام به هر مجموعه ای از نقاط (که دامنه ها، سایتها یا ژنراتورها نامیده می شوند) ناحیه ای اختصاص داده می شود. این نواحی سلول های ورونوی نامیده می شود. برای یک مجموعه از نقاط دیاگرام ورونوی سطح را به مناطقی تقسیم بندی می کند که برای هر نقطه از مجموعه نقاط یک منطقه تعریف می شود. به طوری که تمام نقاط این منطقه به نقطه تولید کننده آن منطقه نزدیکتر می باشد. که در اینجا این نقاط همان نقاط تلاقی خطوط منتخب هستند. از کاربردهای این دیاگرام در مثلث بندی دیلانی می باشد [14].

این دیاگرام به افتخار یوهان پتر گوستاف لوژون دیریکله به نام موزاییک کاری دیریکله، و بعد از گریگوری ورونوی به نام موزاییک کاری ورنوی یا تجزیه ورنوی نامیده شد [14].

در این الگوریتم هدف بدست آوردن افزای از فضای \mathbb{R}^d است، به قسمی که اگر نماینده ی هر ناحیه ی R_i نقطه ی $c_i \in S$ باشد، آنگاه برای تمام نقاط واقع در ناحیه R_i ، نقطه c_i تقریباً نزدیک ترین نقطه در شبکه ی P به آنها شود:

$$V_P^* = \left\{ x \in \mathbb{R}^d \mid x - c_i \leq (1 + \varepsilon) x - c_i, \forall c_j \in S - \{c_i\} \right\}$$



شکل ۱۶ نقاط جفت شده توسط الگوریتم پیشنهادی

یعنی در الگوریتم پیشنهادی تعداد نقاط جفت شده صحیح ۶۲ درصد و در دو الگوریتم بالا به ترتیب صفر و ۱۲ درصد بوده اند.

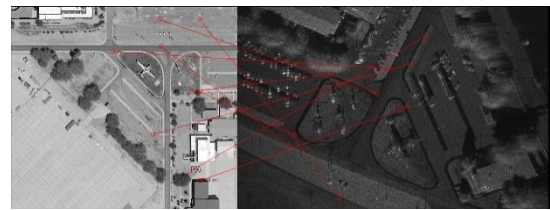
۴- نتیجه

در این مقاله سعی بر آن شد که نقاط تطابق مناسبی را از تصاویر SAR و اپتیکی استخراج شود. ابتدا تصاویر ورودی لبه یابی می شوند. برای استخراج خطوط از لبه های بدست آمده استفاده از تبدیل رادون پیشنهاد می گردد. در حالیکه در سایر روشها از تبدیل هاف استفاده شده است. نتایج حاصل از تبدیل رادون به مراتب بهتر از روش هاف می باشد. در ادامه از سلول های ورونوی مربوط به نقاط تلاقی خطوط بدست آمده استفاده نموده با بررسی تشابه چند ضلعی های حاصل از ورونوی تطابق دو تصویر انجام می شود. نتایج تجربی نشان می دهد تعداد نقاط صحیح جفت شده در روش پیشنهادی مناسب تر از سایر روش ها می باشد.

مراجع

- [1] Zitová, B. and Flusser, J., 2003. *Image registration: a survey*. *Image and Vision Computing*, 21(11), pp. 977-1000.
- [2] H. M. Chen, P. K. Varshney, and M. K. Arora, "Performance of mutual information similarity measure for registration of multitemporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2445-2454, Nov. 2003.
- [3] Y. F. Ban and A. Jacob, "Object-based fusion of multitemporal multiangle ENVISAT ASAR and HJ-1B multispectral data for urban land-cover mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 1998-2006, Apr. 2013.
- [4] D. Brunner, G. Lemoine, L. Bruzzone, and H. Greidanus, "Building height retrieval from VHR SAR imagery based on an iterative simulation and matching technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1487-1504, Mar. 2010.
- [5] S. Suri and P. Reinartz, "Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 939-949, Feb. 2010.
- [6] M. Hasan, M. R. Pickering, and X. P. Jia, "Robust automatic registration of multimodal satellite images using CCRE with partial volume interpolation," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 4050-4061, Oct. 2012.
- [7] M. G. Gong, S. M. Zhao, L. C. Jiao, D. Y. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4328-4338, Jul. 2014.
- [8] A. Wong and D. A. Clausi, "AISIR: Automated inter-sensor/inter-band satellite image registration using robust complex wavelet feature representations," *Pattern Recognit. Lett.*, vol. 31, no. 10, pp. 1160-1167, Jul. 2010.
- [9] Larry S Shapiro and J Michael Brady. "Feature-based

است [9]. بنابراین برای تصحیح آن باید الگوریتم دیگری اضافه شود تا خطای آن را پوشش بدهد.



شکل ۱۴ انطباق دو تصویر با استفاده از بردارهای ویژه

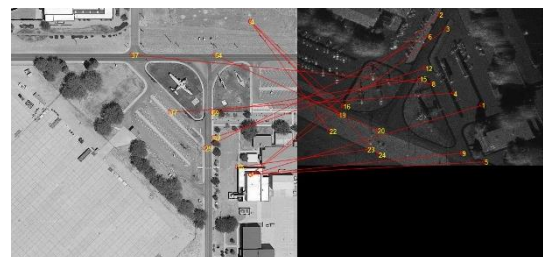
در گام دوم برای تطابق تصاویر به یکدیگر به سراغ ورونوی رفته و هر نقطه تلاقی یک چندضلعی بر آن محاط شده است که فاصله هاسدورف آنها از یکدیگر را سنجیده و هر کدام فاصله های کمتری داشتند یعنی انطباق بیشتری با یکدیگر دارند. فاصله هاسدورف دو مجموعه نقاط متناهی I (یک تصویر را ارائه می دهد) و M (یک مدل از چند هدف که می خواهد در تصویر جای گیری کند را ارائه می دهد) در صفحه به صورت معادله زیر تعریف می شود:

$$H(M, I) = \max(h(M, I), h(I, M))$$

که

$$h(M, I) = \max_{m \in M} \min_{i \in I} m - i$$

در این قسمت نیز اگر این الگوریتم به تنهایی پیاده شود نقاط جفت شده صحیح کمی می دهد. نقاط جفت شده در تصویر (۱۵)، ۱۶ جفت است که فقط ۲ جفت از آنها درست است و ۱۴ جفت نقطه غلط دارد.



شکل ۱۵ نقاط جفت شده با استفاده از فاصله هاسدورف

ورونوی ها

الگوریتم پیشنهاد شده در این مقاله عملکرد به مراتب بهتر از دو الگوریتم ذکر شده در بالا خواهد داشت.

در اینجا تلفیقی از گام اول و دوم انجام می پذیرد به گونه ای که مقدار خروجی را درگام اول نرمالیزه کرده و در گام دوم فاصله های هاسدورف بین ورونوی های دو تصویر سنجیده می شود. سپس چندضلعی های متناظر با توجه به فاصله آنها در فضای ویژگی برگزیده می شود، به گونه ای که چندضلعی هایی که شباهت بیشتری با یکدیگر دارند فاصله هاسدورف کمتری دارند، زیرا فاصله هاسدورف نسبت به دوران چندضلعی و تصویر حساس نیست. در آخر مجموع گام اول و دوم در نظر گرفته می شود.

در تصویر (۱۶) دیده می شود که از ۸ نقطه جفت شده ۵ نقطه صحیح و ۳ نقطه غلط وجود دارد.

- correspondence: an eigenvector approach*". Robotics Research Group, Department of Engineering Science, Oxford University. 7 October 1991. revisrtl paper received: 14 January 1992.
- [10] Haigang Sui, Chuan Xu, Junyi Liu, and Feng Hua. "Automatic Optical-to-SAR Image Registration by Iterative Line Extraction and Voronoi Integrated Spectral Point Matching". IEEE Transactions on Geoscience and Remote Sensing, Vol. 53, NO. 11, November 2015.
- [11] Jong-Sen Lee. "Digital Image Enhancement and Noise Filtering by Use of Local Statistics". IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. PAMI-2, NO. 2, MARCH 1980.
- [12] Peng-Lang Shui and Dong Cheng. "Edge Detector of SAR Images Using Gaussian-Gamma-Shaped Bi-Windows". IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 9, NO. 5, SEPTEMBER 2012.
- [13] Sigurdur Helgason, Radon Transform, Second Edition, Cambridge, 1999.
- [14] N. Ahuja, "Dot pattern processing using Voronoi neighborhoods," *IEEETrans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 3, pp. 336-343, 1982.

زیر نویس ها

- ¹ Gaussian-Gamma-Shaped
² Radon Transform
³ Sinogram
⁴ Voronoi



ارائه یک الگوریتم جدید جهت حذف نویز تصویر براساس یادگیری واژه‌نامه ناهمدوس

سمیرا مودتی^۱

^۱ استادیار دانشگاه مازندران، دانشکده فنی و مهندسی، دانشگاه مازندران، بابلسر، ایران
s.mavaddati@umz.ac.ir

چکیده

در این مقاله الگوریتم کاهش نویز سیگنال تصویر جدیدی براساس یادگیری واژه‌نامه فراکامل مبتنی بر حداقل همدوسی مورد بررسی قرار می‌گیرد. روال یادگیری واژه‌نامه ناهمدوس طی سه مرحله و به صورت بازنمایی تنک مبتنی بر حداکثر همدوسی داده و اتم، اصلاح اتم‌های همدوس در واژه‌نامه و ناهمدوسی اتم‌ها براساس تصحیح ماتریس گرام انجام می‌شود. در نظر گرفتن معیار همدوسی میان فریم داده و اتم به همراه حداکثرسازی ناهمدوسی میان اتم‌ها، موجب دستیابی به تصویر حذف نویز شده با دقت بیشتر می‌شود. نتایج شبیه‌سازی‌های انجام شده بر موفقیت روش پیشنهادی از نظر دقت نتایج بهسازی تصویر و دستیابی به همدوسی مطلوب در اتم‌های واژه‌نامه نسبت به الگوریتم پایه در این زمینه تاکید می‌کند.

کلمات کلیدی

حذف نویز تصویر، یادگیری واژه‌نامه، بازنمایی تنک، همدوسی.

۱- مقدمه

ویژگی‌های سیگنال تصویر انجام می‌شود. همچنین با توجه به ساختار تصویر یک معیار توقف برای کدگذاری تنک به منظور بدست آوردن تصویر با مشخصات بهتر، در نظر گرفته می‌شود. استفاده از آموزش واژه‌نامه در این روش موجب کیفیت بهتر تصویر حذف نویز شده و پیچیدگی کمتر محاسباتی می‌گردد. روش حذف نویز محلی-تطبیقی براساس خوشه‌بندی تصویر نویزی در نواحی دارای ساختار هندسی مشابه در [۹] پیشنهاد شده است. به منظور محاسبه مشخصه وزن محلی در روال خوشه‌بندی مطرح شده، از توابع مشتق و رگرسیون توابع کرنل استفاده می‌گردد. این وزن‌ها در انتقال قابل اطمینان داده بسیار مقاوم هستند. اساس این شیوه خوشه‌بندی، بر تکنیک آنالیز مولفه‌های اساسی^۲ استوار است [۹]. استفاده از تکنیک بازنمایی تنک و آموزش واژه‌نامه به منظور حذف نویز تصویر برای اولین بار در [۱۰] ارائه گردید. این روال آموزش واژه‌نامه مبتنی بر تکنیک K-SVD^۳ نتایج مناسبی را به منظور حذف نویز تصویر بدست می‌دهد. در این روش، اتم‌های واژه‌نامه براساس بلوک‌های مختلف تصویر نویزی تعیین و با تنظیم پارامتر تنکی مناسب، حذف نویز به کمک بازنمایی تنک تصویر نویزی بر روی واژه‌نامه حاصل صورت می‌پذیرد. در [۱۱] نیز از

پردازش سیگنال تصویر کاربردهای مختلفی در حوزه‌های گوناگون مانند پزشکی، کشاورزی، تعیین هویت در سیستم‌های امنیتی و ... دارد. از این میان بهسازی یا حذف نویز تصویر به علت موارد استفاده مختلف در صنایع دفاعی، شهرسازی، هواشناسی، ارسال داده ماهواره‌ای، نجوم و ... بسیار حائز اهمیت است [۳-۱]. روش‌های پیشنهاد شده تاکنون در زمینه حذف نویز تصویر شامل طراحی فیلترهای ساده و وفقی، استفاده از معادلات مشتقات جزئی، استفاده از تبدیل موجک، روش مبتنی بر مورفولوژی، روش‌های آماری و ... بوده است [۷-۴]. در این مقاله، الگوریتم حذف نویز سیگنال تصویر مبتنی بر تکنیک‌های بازنمایی تنک^۱ و آموزش واژه‌نامه^۲ می‌باشد. یادگیری واژه‌نامه در سال‌های اخیر به منظور بازنمایی تنک سیگنال، یکی از موضوعات مورد علاقه پژوهشگران در زمینه پردازش تصویر بوده است [۱۳-۸]. در [۸]، یک روش حذف نویز تصویر براساس بازنمایی تنک چند مقیاسی ارائه شده است که در آن بازنمایی تنک در مقیاس‌های مختلف با توجه به

تنک مدنظر می‌توان در گام اول آن استفاده نمود. روش بکارگرفته شده به منظور بازنمایی تنک بر اساس تعداد پارمترهای لازم برای تنظیم، دقت و چگونگی تنظیم نرخ تنکی برای هر کاربرد متفاوت می‌باشد. همانطور که انتظار می‌رود با افزایش داده آموزش، سرعت همگرایی الگوریتم نیز کاهش می‌یابد. تنظیم مناسب نرخ تنکی در این شرایط، می‌تواند خطای بازسازی در محدوده قابل قبول را بدست دهد. تاکنون روش‌های یادگیری واژه‌نامه متعددی ارائه شده است که در این بین می‌توان از الگوریتم‌های مبتنی بر درست‌نمایی بیشینه، روش جهت‌های بهینه^۹، روش ماکزیمم احتمال پسیه نام برد [۱۷، ۱۸].

۳- روش پیشنهادی

در این بخش الگوریتم بهسازی تصویر پیشنهادی، ارائه شده و هر بخش در آن مورد بررسی دقیق قرار می‌گیرد.

۳-۱- بازنمایی تنک در روش پیشنهادی

برای پردازش مناسب‌تر سیگنال می‌توان آن را با استفاده از یک ساختار مشخص که در اصطلاح به آن واژه‌نامه گفته می‌شود به صورتی بازنمایی نمود که تعداد عناصر غیرصفر آن نسبت به سیگنال اولیه بسیار کمتر باشد. به این عمل، کدگذاری تنک آن سیگنال گفته می‌شود که عموماً گام ابتدایی در الگوریتم‌های یادگیری واژه‌نامه می‌باشد. هدف از بازنمایی تنک، مدل‌کردن فریم‌های ورودی به عنوان ترکیب خطی از تعداد کمی از بردارهای پایه است. مرحله بعد در الگوریتم‌های یادگیری واژه‌نامه، تطبیق بیشتر اتم‌های واژه‌نامه با داده‌های ورودی است. بیشتر روش‌های طراحی واژه‌نامه، شامل این دو مرحله می‌باشد. ضرایب در مرحله اول برای یک واژه‌نامه داده شده محاسبه می‌گردد و سپس واژه‌نامه با فرض ضرایب تنک ثابت، به روز می‌شود. عمده تفاوت میان الگوریتم‌های مختلف در مورد روش محاسبه ضرایب تنک و در رویکرد آموزش واژه‌نامه خلاصه می‌شود. گام کدگذاری تنک در روش پیشنهادی در این بخش و آموزش واژه‌نامه در بخش بعدی این مقاله مورد بررسی قرار می‌گیرد. در روش پیشنهادی، از الگوریتم بازنمایی تنک^{۱۰} LARC استفاده شده که تعمیمی از الگوریتم^{۱۱} LARS با شرط توقف براساس مقدار همدوسی مانده است [۲۰، ۱۹]. این الگوریتم به منظور بازنمایی تنک سیگنال گفتار ارائه شده و برای اولین بار در این مقاله در راستای حذف نویز داده تصویر بکار گرفته می‌شود [۱۹]. بازنمایی تنک با تکیه بر این تکنیک می‌تواند به صورت زیر بیان گردد:

$$x^* = LRAC(D, x, T, Coh) \quad (3)$$

T ، در این رابطه نرخ تنکی یا کاردینالیته است زیرا الگوریتم بازنمایی تنک LARC بر طبق مقدار تنکی متغیر انجام می‌گیرد. همچنین، پارامتر Coh بیانگر میزان همدوسی مانده است. در نظر گرفتن معیار همدوسی مانده برای خاتمه دادن به الگوریتم کدگذاری تنک، قابلیت مصالحه میان انحراف منبع و اعوجاج منبع را بدست می‌دهد. انحراف منبع زمانی رخ می‌دهد که بازنمایی تنک با نرخ تنکی پایین انجام یا بازنمایی بسیار تنک باشد. بنابراین تعداد اتم‌های مورد نیاز برای نمایش کافی بسیار متنوع بوده و داده نمی‌تواند به درستی بر اتم‌های واژه‌نامه کد

الگوریتم K-SVD به منظور آموزش واژه‌نامه استفاده می‌گردد و ضرایب حاصل از بازنمایی تنک به کمک الگوریتم تعقیب طراحی شده به حذف نویز سیگنال تصویر ورودی منجر می‌شود. طراحی الگوریتم مناسب برای یادگیری اتم‌های واژه‌نامه برای داده‌های تنک و ارائه راهکار برای چگونگی حذف نویز از تصاویر سه بعدی پزشکی در [۱۲] بررسی شده است. یادگیری واژه‌نامه در این مرجع طی دو مرحله کدگذاری تنک و بروزسانی واژه‌نامه انجام می‌شود. در گام ابتدایی این روش، یک الگوریتم موثر به منظور بازنمایی تنک با نام تعقیب چند خوشه‌ای ارائه شده است که در کنار روال آموزش واژه‌نامه در نظر گرفته شده، نتایج مطلوبی را در حذف نویز تصویر بدست می‌دهد. الگوریتم K-SVD به عنوان یک روش بسیار موفق، برای آموزش دیکشنری فراکامل به منظور بازنمایی سیگنال تنک با خطای بازسازی کم و نرخ همگرایی مناسب در کنار هر الگوریتم کدگذاری تنک مدنظر می‌تواند مورد استفاده قرار گیرد [۱۳]. در این مقاله از نسخه اصلاح شده این روش با هدف کاهش همدوسی^۹ اتم‌ها در جهت کاهش نویز سیگنال تصویر بهره گرفته می‌شود. در بخش دوم این مقاله مسئله حذف نویز تصویر مورد بررسی قرار می‌گیرد. سپس در بخش سوم، بخش‌های مختلف روش پیشنهادی معرفی شده و در بخش چهارم نتایج شبیه‌سازی‌های انجام شده ارائه می‌گردد. در ادامه نتیجه‌گیری در مورد روش ارائه شده بیان می‌شود.

۲- تعریف مسئله

سیگنال تصویر در مواجهه با نویز سفید گوسی به صورت زیر قابل مدل شدن خطی می‌باشد:

$$Y(m) = I(m) + \eta(m) \quad (1)$$

که در آن Y ، I و η به ترتیب ماتریس داده سیگنال نویزی، سیگنال تمیز و سیگنال نویز سفید گوسی با مشخصه $\eta(m) \sim N(0, \sigma^2)$ است. تصویر ورودی I به بلوک‌های $I_{m \in M}$ تقسیم می‌شود که M مختصات این بلوک‌ها خواهد بود. تکنیک بازنمایی تنک بیان می‌کند که یک سیگنال $I \in \mathbb{R}^N$ می‌تواند با ترکیب خطی تنکی از اتم‌ها به صورت $I = Dx$ که $D \in \mathbb{R}^{N \times L}$ ، $L > n$ یک واژه‌نامه فراکامل است، نمایش داده شود. واژه‌نامه شامل L اتم در ستون‌ها $\{d_l\}_{l=1}^L$ با نُرم واحد $\|d_{(:,l)}\|_2 = 1, \forall l = 1, \dots, L$ و بردار کدگذار K -تنک^{۱۰} $x \in \mathbb{R}^L$ ، $L \gg K$ شامل ضرایب بازنمایی سیگنال I خواهد بود [۱۴-۱۵]. بنابراین مسئله بازنمایی تنک که شامل بخش‌های خطای بازسازی و قید تنکی است به صورت زیر قابل بیان می‌باشد [۱۴-۱۵]:

$$x^* = \arg \min_x \|I - Dx\|^2 \quad \text{s.t.} \quad \|x\|_0 \leq K \quad (2)$$

تعداد ضرایب غیرصفر در x به صورت $\|x\|_0 = K$ که بیانگر کاردینالیته^۷ سیگنال است تعریف می‌شود. نُرم l_0 در رابطه (۲) به یک مسئله غیرمحدب منتهی می‌گردد که NP-دشوار است و ممکن است موجب گرفتار شدن مسئله بهینه سازی در مینیمم‌های محلی گردد. آزادسازی^۸ این مسئله با جایگزینی نُرم l_0 با نُرم l_1 در [۱۶] گزارش شده است. همانطور که بیان گردید، الگوریتم K-SVD به عنوان یک روش مناسب به منظور آموزش واژه‌نامه فراکامل می‌تواند مورد استفاده قرار می‌گیرد [۱۰]. این روش شامل دو گام بازنمایی تنک و تطبیق اتم‌ها است که به علت انعطاف‌پذیری بودن این روش، از هر الگوریتم کدگذاری

استفاده قرار می‌گیرد، فریم سخت متساوی الزاویه^{۱۳} (ETF) یا فریم گرسمانیان^{۱۴} نامیده می‌شود [۲۲]. اما مسئله اساسی این است که این ماتریس برای هر ابعاد دلخواه از واژه‌نامه وجود ندارد بنابراین حل مسئله به صورت تقریبی انجام خواهد گرفت. یکی از نزدیک‌ترین راه‌حل‌های قابل قبول، پس‌پردازش واژه‌نامه مورد نظر خواهد بود. در [۲۳]، یک روش چرخش و تصویرکردن تکرارشونده^{۱۵} (IPR) معرفی شده است تا به‌طور غیرمستقیم فریم ETF را در دو گام بدست دهد. در گام اول، ماتریس گرام اولیه توسط مجموعه قیودی که قیود ساختاری و طیفی نامیده می‌شوند، تصویر می‌شود. این کار با آستانه‌گذاری ضرایب غیرقطری و محدود کردن تعداد مقادیر ویژه غیرصفر ماتریس گرام G صورت می‌پذیرد. گام بعدی شامل مینیمم‌سازی نُرم باقیمانده در رابطه (۳) با چرخش اتم‌های واژه‌نامه به کمک یک ماتریس متعامد خواهد بود. تصویرکردن ساختاری ماتریس گرام G به فریم سخت متساوی الزاویه K_{μ_0} با مقدار همدوسی تعیین شده μ_0 به این صورت است که در ابتدا عناصر قطری K_{μ_0} به مقدار صفر تنظیم می‌شوند تا نُرم فرینوسی خطای میان ماتریس گرام و ماتریس واحد کاهش یابد $\|G - I\|_F$. سپس عناصر غیرقطری آستانه‌گذاری می‌شوند [۲۳]:

$$K_{i,j} = \text{Thresh}(G_{i,j} |_{i \neq j}, \mu_0) = \begin{cases} g_{i,j} & \text{if } |g_{i,j}| \leq \mu_0 \\ \mu_0 & \text{if } g_{i,j} > \mu_0 \\ -\mu_0 & \text{if } g_{i,j} < -\mu_0 \end{cases} \quad (5)$$

تصویرکردن طیفی با محدود نمودن تعداد مقادیر ویژه ماتریس گرام G صورت می‌پذیرد. این تعداد به‌وسیله حفظ تنها N بزرگترین مقادیر ویژه انجام می‌شود. گام تصویرکردن G با تجزیه مقادیر ویژه^{۱۶} در $G = Q\Lambda Q^T$ به صورت زیر انجام می‌گردد [۲۳]:

$$\hat{\Lambda} = \text{Thresh}(\Lambda |_{i=j}, N) = \begin{cases} \lambda_{i,j} & \text{if } i \leq N, \lambda_{i,j} > 0 \\ 0 & \text{if } i > N \text{ or } \lambda_{i,j} < 0 \end{cases} \quad (6)$$

سپس، ماتریس گرام براساس $\hat{G} = Q\hat{\Lambda}Q^T$ به‌روز شده و رتبه آن به بعد داده مشاهده شده N محدود می‌گردد. اگرچه این گام، ماتریس گرام با حداقل مقدار هموسی بدست می‌دهد اما خطای تقریب داده آموزش افزایش پیدا می‌کند زیرا تطبیق بهینه‌شده که توسط الگوریتم یادگیری K -SVD حاصل شده از دست رفته است. بنابراین برای حل این مسئله، گام چرخش واژه‌نامه با استفاده از یک ماتریس متعامد W به منظور بهبود خطای تجزیه بدون تغییر مقدار همدوسی متقابل در نظر گرفته می‌شود. این ماتریس متعامد $N \times N$ از رابطه بهینه سازی زیر حاصل می‌شود:

$$W^* = \arg \min_w \|Y - WDx\|_F \quad (7)$$

این تکنیک در ابتدا به منظور بازسازی تنک سیگنال موسیقی ارائه گردید و برای اولین بار در مرجع [۲۴] در راستای آموزش واژه‌نامه برای بهسازی سیگنال تصویر مورد استفاده قرار گرفت.

۳-۲-۲-۳- اصلاح اتم‌های واژه‌نامه براساس انرژی ضرایب تنک

شود. همچنین همدوسی متقابل میان اتم‌های واژه‌نامه به صورت زیر بدست می‌آید:

$$\mu(D) = \max_{1 \leq i, j \leq L, i \neq j} |d_i \cdot d_j| \quad (4)$$

از طرف دیگر، انتخاب پارامتر تنکی با نرخ بالا یا کدگذاری بسیار متراکم موجب اعوجاج منبع می‌گردد یعنی تعداد اتم‌ها در واژه‌نامه تصویر برای بازنمایی درست آن کافی نخواهد بود و خطای تقریب افزایش می‌یابد. بنابراین انتخاب مقدار تنکی باید به دقت صورت پذیرد. در روش پیشنهادی از این روش بازنمایی تنک در گام اول الگوریتم یادگیری واژه‌نامه به منظور دستیابی به واژه‌نامه‌های با همدوسی متقابل بالا استفاده می‌شود.

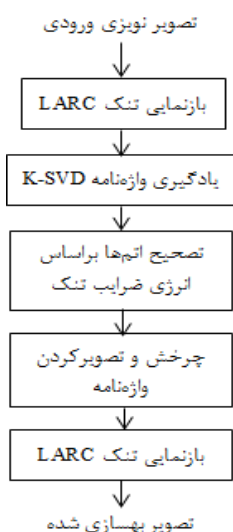
۳-۲- آموزش واژه‌نامه ناهمدوس در روش پیشنهادی

روال یادگیری واژه‌نامه فراکامل مبتنی بر الگوریتم K -SVD، یک روش موثر برای آموزش اتم‌ها براساس مجموعه‌ای از سیگنال‌های نمونه است [۱۰]. هر فریم ورودی در این الگوریتم با ترکیب خطی تنکی از ضرایب K اتم در یک روش مبتنی بر تجزیه مقادیر منفرد، بازنمایی می‌شود. نکته قابل توجه در آموزش واژه‌نامه که می‌تواند خطای تقریب در بازنمایی تنک را تا حد مطلوب کاهش دهد، در نظر گرفتن معیار همدوسی در این روال می‌باشد. این ویژگی از دو جهت قابل بررسی است: (۱) همدوسی اتم‌ها با داده (همدوسی متقابل). (۲) همدوسی بین اتم‌های واژه‌نامه. همدوسی متقابل به نمایش مناسب فریم داده ورودی براساس اتم‌های واژه‌نامه بر می‌گردد که بدیهی است روالی در آموزش واژه‌نامه باید اتخاذ گردد تا اتم‌هایی با حداکثر همدوسی با فریم‌های داده در واژه‌نامه انتخاب و خطای تقریب در بازنمایی تنک کاهش یابد. از طرف دیگر در مورد ویژگی دوم، می‌بایست کمترین همدوسی میان اتم‌های واژه‌نامه وجود داشته باشد تا پایه‌های فضایی چپینش شده، بهترین بازنمایی فریم‌های داده را ایجاد نمایند. این پارامترها هنگامی که سیگنال‌های داده آموزش و نویزهای مورد نظر ساختار یکسان دارند، باید با توجه بیشتری مورد بررسی قرار گیرد زیرا همدوسی بالاتر با کلاس داده و همدوسی متقابل پایین‌تر میان اتم‌ها موجب می‌گردد که خطای تقریب تنک کوچک‌تر و نتایج بهسازی بالاتری حاصل گردد.

۳-۲-۱- اصلاح اتم‌های واژه‌نامه به منظور افزایش همدوسی اتم با داده

بر طبق رابطه (۴)، همدوسی متقابل به عنوان ماکزیمم مقدار مطلق حاصل همبستگی متقابل میان اتم‌ها تعریف می‌گردد [۱۹]. واژه‌نامه‌ها باید بیشترین میزان همدوسی متقابل را داشته باشند تا مانع از اعوجاج در سیگنال بازسازی شده گردند. مسئله یافتن یک واژه‌نامه با همدوسی متقابل کم می‌تواند به مقید نمودن ماتریس گرام^{۱۷} $G = D^T D$ واژه‌نامه D ماکزیمم مقدار مطلق عناصر منتهی گردد. مقدار همدوسی واژه‌نامه D ماکزیمم مقدار مطلق عناصر غیرقطری ماتریس گرام وقتی اتم‌ها نرمالیزه باشند، معرفی می‌گردد [۲۱]. این واژه‌نامه نرمالیزه شده که به منظور بازسازی تنک مناسب مورد

شبه‌سازی‌ها بر روی مجموعه‌ای از تصاویر ۸ بیتی با سطوح خاکستری و ابعاد 256×256 انجام شده است. تکه‌های انتخاب شده از تصویر ورودی به منظور ساخت ماتریس داده Y ، به صورت بلوک‌های 8×8 انتخاب شده است. واژه‌نامه با نرخ ۴ فراکامل می‌باشد و ابعاد 256×64 خواهد داشت. همچنین نرخ تنگی برابر با $T=10$ انتخاب شده که به این معنی است که هر فریم داده ورودی می‌تواند به منظور بازنمایی، حداکثر از ۱۰ اتم در واژه‌نامه استفاده نماید. در شکل ۲ تصاویر حذف نویز شده توسط الگوریتم پیشنهادی، نمایش داده شده است. این تصاویر تحت تاثیر نویز گوسی با انحراف معیار $\sigma = 25$ قرار گرفته‌اند. نتایج نسبت سیگنال به نویز^{۱۷} (SNR) و نسبت پیک سیگنال به نویز^{۱۸} (PSNR) حاصل در جدول ۱ آورده شده است. مقدار نرخ همدوسی متقابل میان اتم-داده و نرخ همدوسی میان اتم‌ها برای روش‌های یادگیری واژه‌نامه مختلف در جدول ۲ گزارش شده است. همانطور که بیان گردید مطلوب است که نرخ همدوسی متقابل میان اتم-داده، بالاترین مقدار ممکن را داشته باشد تا خطای تقریب حداقل مقدار شود. روش یادگیری واژه‌نامه K-SVD با گام بازنمایی تنگ مبتنی بر الگوریتم تعقیب تطابق معامد^{۱۹} (OMP) بدون داشتن معیاری به منظور بررسی پارامتر همدوسی، کمترین مقدار همدوسی متقابل میان اتم-داده را به خود اختصاص داده است که این عملکرد در روش بعدی با اعمال روال بازنمایی تنگ LARC بهبود پیدا می‌نماید. همچنین اگر در ادامه گام تصحیح اتم‌ها که در بخش ۳-۲-۲ معرفی گردید به این روال یادگیری اعمال شود، مقدار همدوسی متقابل کمتر می‌شود از آنجاییکه عملکرد این بلوک کاهش همدوسی میان اتم‌های واژه‌نامه می‌باشد و در انجام این روال، شرایط حصول حداکثر همدوسی اتم-داده نادیده گرفته می‌شود. در ادامه با افزودن گام IPR، از آنجاییکه کاهش خطای تقریب نیز علاوه بر اصلاح مقادیر ویژه ماتریس گرام واژه‌نامه انجام می‌شود، هر دو این پارامترها بهبود پیدا می‌کنند. در مورد نرخ همدوسی میان اتم‌ها نیز از آنجاییکه روش‌های پیشنهادی به منظور کاهش هر چه بیشتر این پارامتر، تصحیح اتم‌ها و تکنیک IPR می‌باشد، در دو سطر اول تغییر محسوس به چشم نمی‌آید. با افزودن این گام‌ها به ترتیب، خواهیم دید که عملکرد مطلوب مورد انتظار حاصل می‌گردد.



شکل ۱: بلوک دیاگرام روش حذف نویز تصویر پیشنهادی.

روال تصحیح اتم‌های واژه‌نامه فراکامل در این بخش به منظور تضمین تطبیق بیشتر میان داده آموزش و واژه‌نامه متناظر پیشنهاد می‌شود. هدف کاهش اطلاعات متقابل و همدوسی احتمالی میان اتم‌های واژه‌نامه است تا در کنار هم قرار گرفتن آن‌ها در روال بهسازی، به خطای کمتری منجر گردد. بنابراین اتم‌های با بیشترین انرژی در بازنمایی تنگ سیگنال تصویر، با اتم‌های با کمترین انرژی بازنمایی تنگ جایگزین می‌شوند. این جایگزینی با این شرط انجام می‌گیرد که اتم‌های انتخاب شده با کمترین انرژی، حداقل خطای بازسازی در بازنمایی تنگ داده را داشته تا همدوسی متقابل با داده مورد انتظار از دست نرود. اگر تغییری در اتم‌های واژه‌نامه‌ها در چندین تکرار متوالی رخ ندهد، روش تصحیح اتم‌ها متوقف خواهد شد. روال پیشنهادی در ادامه آورده شده است:

$$1- \text{بازنمایی تنگ } Y = LARC(D, x)$$

۲- برای هر سطر ماتریس ضرایب تنگ x :

$$l^{max} = \arg \max_{1 \leq l \leq L} \text{norm}(x) \quad l^{min} = \arg \min_{1 \leq l \leq L} \text{norm}(x)$$

$$\text{or} \quad \arg \min_{1 \leq l \leq L} \text{sparsity}(x) \quad \text{or} \quad \arg \max_{1 \leq l \leq L} \text{sparsity}(x)$$

۳- هر ستون $D_{l^{min}}$ یا $D_{l^{max}}$ جایگزین می‌شود اگر $D_{l^{min}}$ نقش اساسی در بازنمایی داده نداشته باشد و برای واژه‌نامه جدید D' داشته باشیم:

$$|Y - D'x|_2 < \varepsilon_1$$

۴- به گام اول برو تا زمانی که تفاوت میان اتم‌های واژه‌نامه در چند تکرار متوالی مشخص از مقادیر تعیین شده ε_2 کمتر باشند:

$$|D - D'|_2 < \varepsilon_2$$

با استفاده از این روال در تصحیح اتم‌ها، یادگیری واژه‌نامه‌های ناهمدوس به دقت تکمیل گردیده و واژه‌نامه‌های جامع با بازنمایی مناسبی بر روی هر بلوک‌های داده حاصل می‌شود.

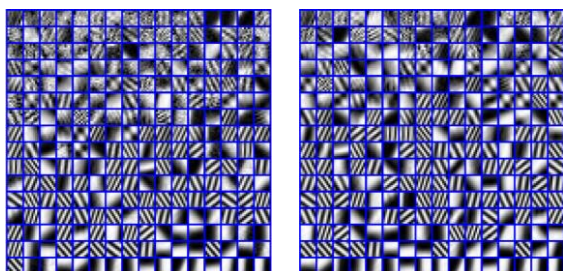
۳-۳- روال بهسازی پیشنهادی

در این مقاله یک روش جدید به منظور حذف نویز تصویر براساس یادگیری واژه‌نامه با حداکثر ناهمدوسی اتم‌ها پیشنهاد می‌گردد. روال یادگیری واژه‌نامه براساس در نظر گرفتن معیار همدوسی به منظور حصول واژه‌نامه‌های فراکامل با اتم‌های ناهمدوس به منظور دستیابی به تصویر حذف نویز شده با دقت بیشتر می‌باشد. بلوک دیاگرام روش پیشنهاد شده در این فصل در شکل ۱ نشان داده شده است. بلوک‌های مختلف این دیاگرام در بخش‌های قبل توضیح داده شده است. در انتها نیز به کمک الگوریتم LARC، ماتریس تصویر حاصل از بلوک‌بندی داده ورودی بر روی اتم‌های واژه‌نامه دارای کمترین میزان همدوسی با یکدیگر و بیشترین میزان همدوسی با داده ورودی، بازنمایی تنگ داده می‌شود. بخش نویزی داده تصویر ورودی در این بازنمایی، به عنوان خطای تقریب با بازسازی ظاهر می‌شود و ضرایب تنگ به آن اختصاص نمی‌یابد. در ادامه از حاصل ضرب ضرایب تنگ جدید در واژه‌نامه ناهمدوس، سیگنال تصویر بهسازی شده حاصل می‌گردد.

۴- شبه‌سازی و تحلیل نتایج

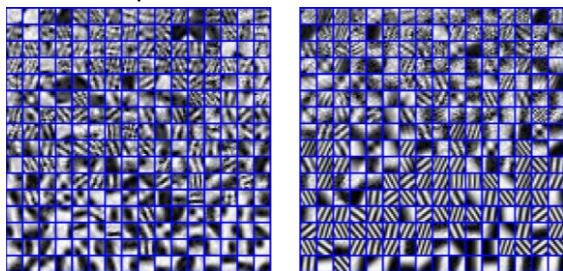
جدول ۲: مقدار نرخ همدوسی متقابل میان اتم-داده و نرخ همدوسی میان اتم‌ها برای روش‌های یادگیری واژه‌نامه مختلف.

نرخ همدوسی میان اتم‌ها	نرخ همدوسی متقابل میان اتم-داده	روش ↓
۰/۶۱	۰/۳۲	K-SVD/OMP
۰/۵۹	۰/۴۶	K-SVD/LARC
۰/۴۸	۰/۴۱	K-SVD/LARC/ تصحیح اتم‌ها
۰/۳۵	۰/۵۴	K-SVD/LARC/ تصحیح اتم‌ها (روش پیشنهادی)



(ب)

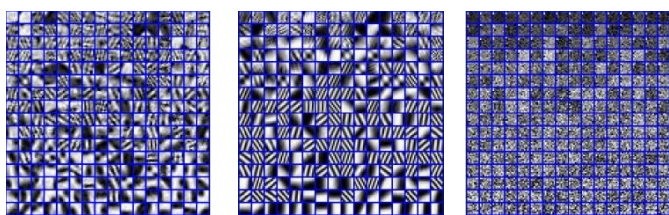
(الف)



(د)

(ج)

شکل ۳: نمایشی از اتم‌های واژه‌نامه فراکامل بعد از اعمال هر مرحله از روال ناهمدوسی برای تصویر چیتا. (الف) K-SVD. (ب) LARC/K-SVD. (ج) اعمال روال تصحیح اتم‌ها. (د) اعمال چرخش و تصویرکردن واژه‌نامه.



(ج)

(ب)

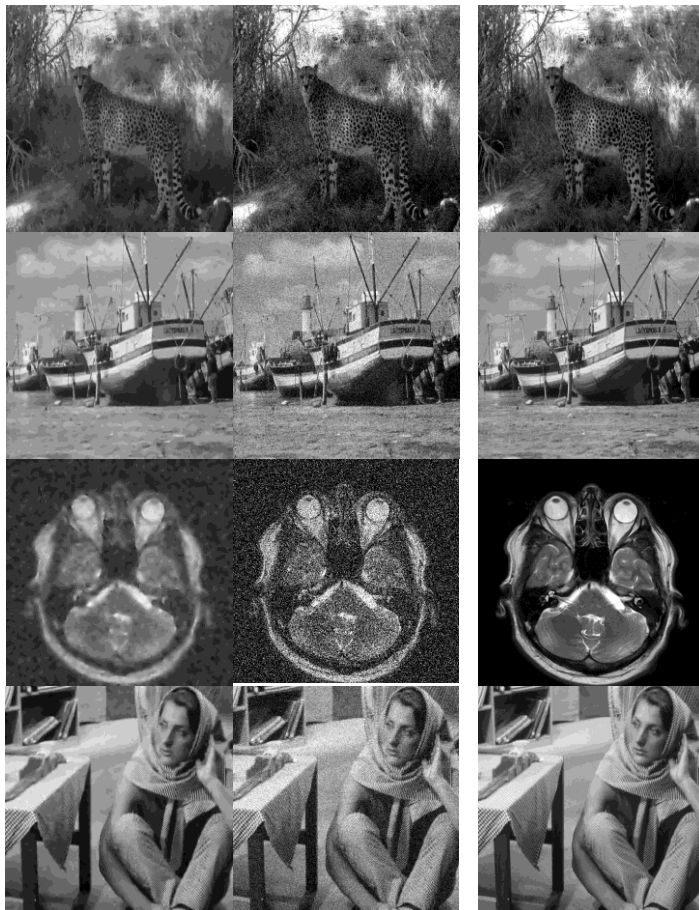
(الف)

شکل ۴: نمایشی از اتم‌های آموزش دیده برای تصویر چیتا توسط الگوریتم‌های: (الف) یادگیری واژه‌نامه مبتنی بر آنالیز مولفه‌های اساسی. (ب) یادگیری واژه‌نامه مبتنی بر روش K-میانگین. (ج) یادگیری واژه‌نامه به روش پیشنهادی.

۵- نتیجه

در این مقاله یک الگوریتم یادگیری واژه‌نامه اصلاح شده بر مبنای ناهمدوسی به منظور بهسازی سیگنال تصویر ارائه شد. دو ویژگی همدوسی متقابل میان اتم‌ها و داده ورودی همچنین ناهمدوسی میان اتم‌های یک واژه‌نامه، با بکارگیری الگوریتم بازنمایی تنک با نرخ تنکی

در شکل ۳ نمایشی از اتم‌های یادگیری شده بعد از اعمال مراحل مختلف روال ناهمدوسی آورده شده است. قابل مشاهده است که با اعمال هر مرحله از بلوک دیاگرام شکل ۱، اتم‌های واژه‌نامه آموزش دیده تغییر خواهند کرد و جزئیات بیشتری را از ساختار تصویر بازنمایی می‌نمایند. در شکل ۴ نیز اتم‌های آموزش دیده توسط الگوریتم‌های یادگیری واژه‌نامه براساس آنالیز مولفه‌های اساسی و یادگیری واژه‌نامه مبتنی بر روش K-میانگین^{۲۰} نشان داده شده که بیان می‌کند اتم‌ها در روش پیشنهادی بیشترین شباهت را با داده تصویر دارند [۱۰].



شکل ۲: نتایج حذف نویز تصویر حاصل از بکارگیری روش پیشنهادی با نویز گوسی $\sigma = ۲۵$. ستون سمت راست: تصاویر تمییز. ستون وسط: تصاویر نویزی. ستون سمت چپ: تصاویر حذف نویز شده.

جدول ۱: مقادیر نسبت سیگنال به نویز حاصل از حذف نویز گوسی تصویر با بکارگیری روش پیشنهادی با نویز گوسین $\sigma = ۲۵$.

PSNR	SNR	روش پیشنهادی	چیتا
۳۰/۷۱	۲۷/۵۶	K-SVD	
۲۹/۸۵	۲۶/۳۵	K-SVD	کشتی
۳۰/۶۵	۲۷/۹۷	روش پیشنهادی	
۳۰/۰۱	۲۷/۱۳	K-SVD	ام آر آی مغز
۳۱/۲۸	۲۸/۵۳	روش پیشنهادی	
۳۰/۷۷	۲۷/۷۶	K-SVD	باربارا
۳۰/۸۹	۲۷/۸۶	روش پیشنهادی	
۲۹/۸۷	۲۸/۰۵	K-SVD	

- [۱۴] Agarwal, A., Anandkumar, A., Jain, P., Netrapalli, P., Tandon, R., "Learning sparsely used overcomplete dictionaries", JMLR: Workshop and Conference Proceedings, Vol. 35, pp. 1-15, 2014.
- [۱۵] Lee, H., Battle, A., Raina, R., Ng, A.Y., "Efficient sparse coding algorithms", Advances in Neural Information Processing Systems, 2006.
- [۱۶] Portilla, J., Mancera, L., "L0-based sparse approximation: Two alternative methods and some applications", Proceedings of the 16th IEEE international conference on Image processing, pp. 3865-3868, 2009.
- [۱۷] Engan, K., Aase, S.O. and Hakon-Husoy, J.H., "Method of optimal directions for frame design", IEEE International Conference on Acoustics, Speech, and Signal Processing., Vol. 5, pp. 2443-2446, 1999.
- [۱۸] Delgado, K., Murray, J.F., Rao, B.D., Engan, K., Lee, T., Sejnowski, T.J., "Dictionary learning algorithms for sparse representation", Neural Computation., Vol.15, pp. 349-396.
- [۱۹] Sigg, C.D., Dikk, T., Buhmann, J.M., "Speech enhancement using generative dictionary learning", IEEE Transactions on Audio, Speech and Language Processing, Vol. 20, Vol. 6, pp.1698-1712, 2012.
- [۲۰] Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., "Least angle regression", Ann. Stat., Vol. 32, pp. 407-499, 2004.
- [۲۱] Tropp, J., Dhillon, I., Heath, R.J., Strohmer, T., "Designing structural tight frames via an alternating projection method", IEEE Trans. on Information Theory, Vol. 51, No.1, pp.188-209, 2005.
- [۲۲] Sustik, M., Tropp, J., Dhillon, I., Heath, R., "On the existence of equiangular tight frames", Linear Algebra and Its Applications, Vol. 426, No. 2, pp. 619-635, 2007.
- [۲۳] Tropp, J., Dhillon, I., Heath, R.J. and Strohmer, T., "Designing structural tight frames via an alternating projection method", IEEE Trans. on Information Theory, Vol. 51, No. 1, pp. 188-209, 2005.
- [۲۴] Mavaddaty, S., Ahadi, S.M., Seyedin, S., "Modified coherence-based dictionary learning method for speech enhancement", Signal Processing, IET, Vol. 9, No. 7, pp. 1-9, 2015.

زیر نویس‌ها

- 1 Sparse representation
- 2 Dictionary learning
- 3 Principle component analysis (PCA)
- 4 K-Singular value decomposition
- 5 Coherence
- 6 K-sparse
- 7 Cardinality
- 8 Relaxation
- 9 Method of optimal directions(MOD)
- 10 Least angle regression with coherence criterion
- 11 Least-angle regression
- 12 Gram matrix
- 13 Equiangular tight frame
- 14 Grassmannian frame
- 15 Iterative projection and rotation (IPR)
- 16 Eigen value decomposition (EVD)
- 17 Signal to noise ratio
- 18 Peak signal to noise ratio
- 19 Orthogonal matching pursuit
- 20 K-means

متغیر، اصلاح اتم‌ها بر مبنای انرژی ضرایب تنک و چرخش و تصویرکردن واژه‌نامه نهایی به منظور حصول واژه‌نامه با ماتریس گرام یک، در روال الگوریتم بهسازی پیشنهادی در نظر گرفته می‌شود. اعمال این تکنیک‌ها به منظور دستیابی به واژه‌نامه با بالاترین میزان انطباق میان داده و اتم‌های بازنمایی کننده و نیز پایین‌ترین نرخ هم‌دوسی میان اتم‌ها در واژه‌نامه است تا پایه‌های فضایی نسبت به یکدیگر تا حد ممکن مستقل باشند. شبیه‌سازی‌های انجام شده نشان داد که روال یادگیری واژه‌نامه پیشنهادی مبتنی بر معیار هم‌دوسی موجب دستیابی به تصویر حذف نویز شده با دقت بالاتر نسبت به روش پایه در این زمینه گردیده است.

مراجع

- [۱] Naimi, H., Adamou-Mitiche, A.B.H. and Mitiche, L., "Medical image denoising using dual tree complex thresholding wavelet transform and Wiener filter", Journal of King Saud University-Computer and Information Sciences, Vol: 27, No.1, pp. 40-45, 2015.
- [۲] Teodoro A., Almeida, M. and Figueiredo, M., "Single-frame image denoising and inpainting using Gaussian mixtures", International Conference on Pattern Recognition Applications and Methods (ICPRAM), pp. 283-288, 2015.
- [۳] Beckouche, S., Starck, J.L. and Fadili, J., "Astronomical image denoising using dictionary learning", Astronomy & Astrophysics 556, A132, 2013.
- [۴] Zhu, Y., Huang, C., "An Improved Median Filtering Algorithm for Image Noise Reduction", Physics Procedia, No. 25, pp. 609-616, 2012.
- [۵] You, Y.L., Kaveh, M., "Fourth Order Partial Differential Equations for Noise Removal?", IEEE Trans. Image Processing, Vol. 9, No. 10, pp. 1723-1730, 2000.
- [۶] مهدی زاده همت آبادی، امیر، حسینی فاطمی، محمدرضا، "حذف نویز تصاویر با استفاده از الگوریتم بهبود یافته تطبیق بلوک و فیلتر سه بعدی"، اولین کنفرانس ملی مهندسی برق دانشگاه آزاد اسلامی واحد لنگرود، ۱۳۹۳.
- [۷] Huang, C., Zhu, Y., "New Morphological Filtering Algorithm for Image Noise Reduction", Image and Signal Processing 2009. CISP '09. 2nd International Congress on, pp. 1-6, 2009.
- [۸] Gan, T., Lu, W., "Image denoising using multi-scale sparse representation", School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Proceedings of IEEE 17th International Conference on Image Processing, pp. 1165-1168, 2010.
- [۹] Chatterjee, P., Milanfar, P., "Image Denoising using Locally Learned Dictionaries", Department of Electrical Engineering, University of California, Santa Cruz, CA 95064, USA, 2009.
- [۱۰] Aharon, M., Elad, M., Bruckstein, A., "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation", IEEE Trans. Signal Process, Vol. 54, No. 11, pp. 4311-4322, 2006.
- [۱۱] Fen, Z., Kai, X., "A novel image denoising method based on DCT basis and sparse representation", Cross Strait Quad-Regional Radio Science and Wireless Technology Conference, College of information & mechanical engineering Beijing Institute of Graphic Communication, pp.1307-1310, 2011.
- [۱۲] Li, S., Fang, L., Yin, H. "An efficient dictionary learning algorithm and its application to 3-D medical image denoising", IEEE transactions on biomedical engineering, Vol. 59, No. 2, pp. 417-427, 2012.
- [۱۳] Davis, G., Mallat, S. and Zhang, Z. "Adaptive time-frequency decompositions", Optical-Engineering., Vol. 33, pp. 218-391, 1994.



بهبود امنیت لایه فیزیکی با استفاده از تخصیص توان تطبیقی در شبکه‌های بی‌سیم مشارکتی

مهدی محمدی^۱، محمد ترابی^۲

^۱ دانشجوی کارشناسی ارشد دانشکده مهندسی برق، دانشگاه شهید بهشتی، تهران

^۲ عضو هیأت علمی دانشکده مهندسی برق، دانشگاه شهید بهشتی، تهران

چکیده

در این مقاله، امنیت لایه فیزیکی را در شبکه‌های بی‌سیم مشارکتی بررسی می‌کنیم که از روش تقویت و ارسال بهره می‌برند و سپس روشی را برای تخصیص توان تطبیقی بین منبع و رله انتخابی به منظور بهبود امنیت بی‌سیم در برابر حملات شنود پیشنهاد می‌کنیم. به منظور مقایسه، طرح انتخاب رله بهینه با تخصیص توان برابر را نیز بررسی می‌کنیم که در آن توان کل به طور برابر در منبع و رله انتخابی تقسیم می‌شود. ابتدا احتمال قطع محرمانگی روش پیشنهادی را بررسی می‌کنیم و یک عبارت تقریبی برای آن بدست می‌آوریم. نتایج بدست آمده از طریق شبیه سازی برتری روش پیشنهادی و همچنین نزدیکی رابطه تقریبی احتمال قطع محرمانگی بدست آمده به رابطه دقیق را نشان می‌دهند.

کلمات کلیدی

امنیت لایه فیزیکی، تخصیص توان، شنود، رله بهینه، تقویت و ارسال

استفاده از گره‌های متعدد مشارکتی یک جایگزین مناسب برای تکنیک چند ورودی - چند خروجی است [7-9]. مشارکت کاربر نه تنها قابلیت اطمینان و عملکرد انتقال بی‌سیم را بهبود می‌بخشد بلکه همچنین پتانسیل فوق‌العاده‌ای در بهبود امنیت بی‌سیم در برابر حملات شنود دارد. در [2] و [3] نشان داده شده است که در حضور یک شنودگر، ظرفیت محرمانگی به صورت اختلاف بین ظرفیت کانال منبع - مقصد (کانال اصلی) و منبع - شنودگر (کانال شنود) بدست می‌آید. اگر ظرفیت محرمانگی منفی شود، شنودگر موفق به قطع سیگنال خواهد شد و یک رخداد قطع اتفاق می‌افتد.

در [10] نویسندگان، ظرفیت محرمانگی ارتباطات بی‌سیم را در حضور یک شنودگر و یک گره رله با پروتکل‌های رمزگشایی و ارسال، تقویت و ارسال و فشرده‌سازی و ارسال مطالعه و با یکدیگر مقایسه کرده‌اند.

در [11]، یک طرح انتخاب رله بهینه با در نظر گرفتن اطلاعات کانال‌های اصلی و شنود پیشنهاد شد و نشان داده شد که عملکرد محرمانگی بهتری از انتخاب رله مرسوم و روش‌های ترکیبی دارد. در [12] استفاده از نویز مشارکتی پیشنهاد شد و از منظر نرخ محرمانگی قابل دستیابی آنالیز شد. در [13] یک طرح ارسال نویز پیشنهاد شد که

۱- مقدمه

با توجه به طبیعت انتشار رسانه بی‌سیم، محرمانگی ارتباطات بی‌سیم به شدت در برابر حملات شنود آسیب‌پذیر است. امروزه امنیت لایه فیزیکی به عنوان یک روش کارآمد در برابر شنود توجهات قابل ملاحظه‌ای را به خود جذب کرده است. برخلاف تکنیک‌های رمزنگاری مرسوم که مبتنی بر کلیدهای محرمانه هستند، امنیت لایه فیزیکی مشخصه‌های فیزیکی کانال را برای جلوگیری از قطع سیگنال انتقالی منبع به مقصد مورد نظر توسط شنودگر، بکار می‌گیرد [1]. برای اولین بار وینر در [2] از این موضوع برای یک کانال شنود بدون حافظه گسسته بهره برد. بعدها این نتایج در [3] به کانال شنود گوسی تعمیم داده شد جایی که از ظرفیت محرمانگی به عنوان ماکزیمم نرخ محرمانگی قابل دستیابی برای ارزیابی عملکرد امنیت لایه فیزیکی استفاده شد. در سال‌های اخیر تکنیک چند ورودی - چند خروجی با تجهیز فرستنده و گیرنده به آنتن‌های متعدد، به عنوان یک روش کارآمد برای افزایش امنیت لایه فیزیکی شناخته شده است [4-6]. هر چند که با توجه به محدودیت توان باتری و اندازه ترمینال ممکن است که پیاده‌سازی آن در برخی موارد سخت یا غیرممکن باشد. کارهای اخیر نشان می‌دهند که

رله‌های مشارکتی منتشر می‌کند که از میان آن‌ها فقط بهترین رله برای ارسال سیگنال دریافتی به مقصد انتخاب خواهد شد در حالیکه شنودگر انتقال از رله انتخابی به مقصد را شنود می‌کند و در تلاش برای ترجمه سیگنال منبع می‌باشد. با استناد به متون امنیت لایه فیزیکی موجود، فرض می‌کنیم که شنودگر همه چیز را در رابطه با انتقال سیگنال از منبع به مقصد شامل طرح رمزگذاری در منبع، پروتکل ارسال در رله‌ها و روش رمزگشایی در مقصد را می‌داند و فقط سیگنال منبع محرمانه است.

۲-۲- انتخاب رله بهینه

همانطور که در [11] بررسی شده رله بهینه، رله‌ای است که ظرفیت محرمانگی انتقال را ماکزیمم کند. بنابراین ملاک انتخاب رله بهینه را می‌توان به صورت زیر نوشت:

$$\text{Optimal Relay} = \arg \max_{i \in \mathcal{R}} C_i$$

$$\arg \max_{i \in \mathcal{R}} \left(\frac{1 + \frac{|h_{si}|^2 |h_{id}|^2 P}{2(|h_{si}|^2 + |h_{id}|^2) \sigma_n^2}}{1 + \frac{|h_{si}|^2 |h_{ie}|^2 P}{2(|h_{si}|^2 + |h_{ie}|^2) \sigma_n^2}} \right) \quad (1)$$

که \mathcal{R} مجموعه‌ای از M رله را نشان می‌دهد. چیزی که می‌توان از معادله (1) به آن پی برد این است که در انتخاب رله نه تنها اطلاعات حالت کانال لینک‌های اصلی یعنی $|h_{si}|^2$ و $|h_{id}|^2$ بلکه اطلاعات حالت کانال لینک شنود یعنی $|h_{ie}|^2$ نیز به حساب آمده است. با توجه به ملاک انتخاب رله بهینه می‌توان از الگوریتم انتخاب رله متمرکز بهره برد، بدین صورت که گره منبع یک جدول شامل M رله و اطلاعات حالت کانال مربوط به آن‌ها $(|h_{si}|^2, |h_{id}|^2, |h_{ie}|^2)$ را ایجاد می‌کند. در این روش، رله بهینه می‌تواند به سادگی با یک نگاه به جدول و با استفاده از ملاک پیشنهادی در معادله (1) مشخص شود.

۳- تخصیص توان تطبیقی

در روش پیشنهادی که در [11] آمده، برای سادگی در محاسبات تخصیص توان برابر در نظر گرفته شده است یعنی توان منبع و رله انتخابی هر دو برابر با $\frac{P}{2}$ است در حالی‌که ما در این مقاله پیشنهاد می‌کنیم که ابتدا رله بهینه را با ملاک معادله (1) انتخاب کنیم اما در هنگام ارسال پیام به جای تخصیص توان برابر از تخصیص توان تطبیقی استفاده کنیم، بدین صورت که توان منبع را P_s و توان رله انتخابی را نیز برای یک مقایسه منصفانه با طرح تخصیص توان برابر، برابر $P - P_s$ در نظر می‌گیریم. حال با توجه به توان‌های اختصاصی به محاسبه پارامتری ظرفیت محرمانگی متغیر با P_s می‌پردازیم و سپس مقدار P_s منجر به ماکزیمم ظرفیت محرمانگی را محاسبه می‌کنیم. با توجه به این که در روش پیشنهادی، منبع سیگنال خود را با توان P_s انتقال می‌دهد، سیگنال دریافتی در رله R_i را می‌توان به صورت زیر نوشت:

$$r_i = \sqrt{P_s} h_{si} S + n_i \quad (2)$$

که h_{si} ضریب محو شونندگی لینک منبع - رله i را نشان می‌دهد و $n_i \sim \mathcal{CN}(0, \sigma_n^2)$ نویز گوسی جمع شونده را نشان می‌دهد. با توجه به ملاک انتخاب رله بهینه می‌دانیم که رله i به عنوان رله انتخابی برای ارسال سیگنال دریافتی خود به مقصد انتخاب شده است. رله i با ضرب

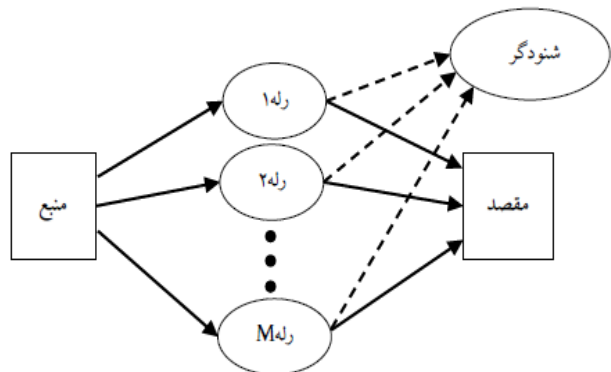
در آن رله با ارسال کدهای مستقل از پیام منبع تلاش می‌کند تا شنودگر را سردرگم کند. در [14] انتخاب اخلاک‌گر و تخصیص توان توأم در یک استراتژی اخلاک مشارکتی پیشنهاد شده است.

در این مقاله یک شبکه بی‌سیم مشارکتی با گره‌های متعدد را در حضور یک شنودگر در نظر می‌گیریم و یک روش تخصیص توان تطبیقی (Adaptive Power Allocation-APA) را نیز علاوه بر انتخاب رله بهینه که در [11] آورده شده را پیشنهاد می‌کنیم. در اینجا فرض می‌کنیم که اطلاعات حالت کانال لینک شنود نیز در دسترس است و تمام رله‌ها را با پروتکل تقویت و ارسال در نظر گرفته‌ایم. برای مقایسه با تخصیص توان تطبیقی برای رله بهینه انتخابی، تخصیص توان برابر (Equal Power Allocation-EPA) را نیز که در [11] مطالعه شده را مورد بررسی قرار داده‌ایم. باقی مقاله بدین صورت می‌باشد که در بخش ۲ مدل سیستم و نحوه انتخاب رله بهینه را مطرح می‌کنیم. در بخش ۳ روش تخصیص توان تطبیقی برای رله انتخابی را ارائه داده و روش تخصیص توان برابر را نیز به عنوان مقایسه با روش پیشنهادی بررسی می‌کنیم. در بخش ۴ یک عبارت تقریبی برای احتمال قطع محرمانگی روش پیشنهادی بدست می‌آوریم. در بخش ۵ نتایج عددی و شبیه‌سازی، برتری روش تخصیص توان تطبیقی را نسبت به روش تخصیص توان برابر به نمایش خواهند گذاشت.

۲- مدل سیستم و نحوه انتخاب رله بهینه

۲-۱- مدل سیستم

یک شبکه بی‌سیم مشارکتی شامل یک منبع، یک مقصد و M رله را در حضور یک شنودگر همان‌طور که در شکل (1) نشان داده شده، در نظر بگیرید. همه رله‌ها مجهز به فقط یک آنتن هستند و ارتباط مستقیمی بین منبع با مقصد مورد نظر و شنودگر وجود ندارد، یعنی مقصد و شنودگر هر دو خارج از ناحیه پوشش منبع هستند و ارتباط فقط از طریق رله‌ها ممکن می‌شود. لینک‌های شنود و اصلی هر دو به صورت کانال‌های رایلی فیدینگ مدل شده‌اند و نویز حرارتی دریافتی در هر گره به صورت یک متغیر تصادفی گوسی مختلط با میانگین صفر و واریانس σ_n^2 $(\mathcal{CN}(0, \sigma_n^2))$ مدل می‌شود. M رله با پروتکل تقویت و ارسال را به صورت $\mathcal{R} = \{R_i | i = 1, 2, \dots, M\}$ مشخص می‌کنیم. در این مقاله، فقط از رله بهینه برای ارسال سیگنال منبع به مقصد مورد نظر استفاده می‌کنیم. بدین صورت که ابتدا منبع سیگنال خود را به



شکل ۱: مدل سیستم

چون مدل استفاده شده رایلی فیدینگ است، می‌توان به این موضوع پی برد که $|h_{id}|^2$ ، $|h_{ie}|^2$ و $|h_{si}|^2$ توزیع‌های نمایی زیر را دارند.

$$f_{|h_{id}|^2}(x) = \frac{1}{\sigma_{id}^2} e^{-\frac{x}{\sigma_{id}^2}} \quad (11)$$

$$f_{|h_{ie}|^2}(x) = \frac{1}{\sigma_{ie}^2} e^{-\frac{x}{\sigma_{ie}^2}} \quad (12)$$

$$f_{|h_{si}|^2}(x) = \frac{1}{\sigma_{si}^2} e^{-\frac{x}{\sigma_{si}^2}} \quad (13)$$

کجه $\sigma_{si}^2 = E(|h_{si}|^2)$ و $\sigma_{ie}^2 = E(|h_{ie}|^2)$ ، $\sigma_{id}^2 = E(|h_{id}|^2)$ هستند. از آن‌جا که بدست آوردن یک عبارت بسته دقیق با توجه به پیچیدگی آنالیز آماری برای جمله اول احتمال قطع محرمانگی بالا یک چالش است، از یک روش تقریبی مشابه آنچه در [15] آمده، استفاده می‌کنیم. با استفاده از این واقعیت که $\frac{xy}{x+y+1} \leq \min\{x, y\}$ همواره برقرار است، می‌توانیم جمله اول معادله (10) را به صورت زیر تقریب بزنیم:

$$I = Pr \left[\frac{1 + \min\{(P - P_s)|h_{id}|^2, P_s|h_{si}|^2\}/\sigma_n^2}{1 + \min\{(P - P_s)|h_{ie}|^2, P_s|h_{si}|^2\}/\sigma_n^2} < 2^{2R_s} \right] \approx Pr \left[\frac{1 + \frac{(P - P_s)|h_{id}|^2|h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{id}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2}}{1 + \frac{(P - P_s)|h_{ie}|^2|h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{ie}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2}} < 2^{2R_s} | P_s \right] \quad (14)$$

با استفاده از قانون کل می‌توانیم معادله (14) را به ۴ معادله به صورت زیر تجزیه کنیم:

$$I = Pr [P_s |h_{si}|^2 < (P - P_s) |h_{id}|^2, P_s |h_{si}|^2 < (P - P_s) |h_{ie}|^2] + Pr \left[\frac{1 + \frac{P_s |h_{si}|^2}{\sigma_n^2}}{1 + \frac{(P - P_s) |h_{ie}|^2}{\sigma_n^2}} < 2^{2R_s}, P_s |h_{si}|^2 < (P - P_s) |h_{id}|^2, P_s |h_{si}|^2 > (P - P_s) |h_{ie}|^2 \right] + Pr \left[\frac{1 + \frac{(P - P_s) |h_{id}|^2}{\sigma_n^2}}{1 + \frac{P_s |h_{si}|^2}{\sigma_n^2}} < 2^{2R_s}, P_s |h_{si}|^2 > (P - P_s) |h_{id}|^2, P_s |h_{si}|^2 < (P - P_s) |h_{ie}|^2 \right] + Pr \left[\frac{1 + \frac{(P - P_s) |h_{id}|^2}{\sigma_n^2}}{1 + \frac{(P - P_s) |h_{ie}|^2}{\sigma_n^2}} < 2^{2R_s}, P_s |h_{si}|^2 > (P - P_s) |h_{id}|^2, P_s |h_{si}|^2 < (P - P_s) |h_{ie}|^2 \right] \quad (15)$$

حال می‌توانیم به صورت جداگانه به محاسبه این ۴ جمله پردازیم.

جمله اول را می‌توانیم به صورت زیر بازنویسی کنیم:

$$I_1 = Pr [P_s |h_{si}|^2 < \min\{(P - P_s) |h_{id}|^2, (P - P_s) |h_{ie}|^2\}] \quad (16)$$

حال اگر $W_k = \min\{(P - P_s) |h_{id}|^2, (P - P_s) |h_{ie}|^2\}$ را در نظر بگیریم می‌توانیم I_1 را بدین صورت محاسبه کنیم:

$$I_1 = Pr [P_s |h_{si}|^2 < W_k] = \int_0^\infty f_{W_k}(x) \int_0^x f_{P_s |h_{si}|^2}(y) dy dx = 1 - \left(\frac{1}{(P - P_s) \sigma_{id}^2} + \frac{1}{(P - P_s) \sigma_{ie}^2} \right) \left(\frac{1}{(P - P_s) \sigma_{id}^2} + \frac{1}{(P - P_s) \sigma_{ie}^2} + \frac{1}{P_s \sigma_{si}^2} \right)^{-1} \quad (17)$$

جمله دوم معادله (15) را نیز بدین صورت مرتب کرده و حل می‌کنیم:

r_i به $h_{si}^* r_i$ آشکارسازی همدوس را انجام داده و سپس مقدار حاصل را در عامل $\frac{1}{\sqrt{P_s |h_{si}|^2}}$ ضرب می‌کند و به مقصد ارسال می‌کند. بعد از انجام این مراحل در R_i ، سیگنال دریافتی در مقصد و شنودگر را می‌توانیم بدین صورت بنویسیم:

$$r_d = \sqrt{P - P_s} h_{id} \frac{h_{si}^* r_i}{|h_{si}|^2 \sqrt{P_s}} + n_d = \sqrt{P - P_s} h_{id} S + \sqrt{P - P_s} \frac{h_{id} h_{si}^*}{|h_{si}|^2 \sqrt{P_s}} n_i + n_d \quad (3)$$

$$r_e = \sqrt{P - P_s} h_{ie} \frac{h_{si}^* r_i}{|h_{si}|^2 \sqrt{P_s}} + n_e = \sqrt{P - P_s} h_{ie} S + \sqrt{P - P_s} \frac{h_{ie} h_{si}^*}{|h_{si}|^2 \sqrt{P_s}} n_i + n_e \quad (4)$$

مطابق با معادلات (3) و (4) ظرفیت انتقال رله AF از R_i به مقصد و شنودگر بدین صورت بدست می‌آید:

$$C_{id} = \frac{1}{2} \log_2 \left(1 + \frac{(P - P_s) |h_{id}|^2 |h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{id}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2} \right) \quad (5)$$

$$C_{ie} = \frac{1}{2} \log_2 \left(1 + \frac{(P - P_s) |h_{ie}|^2 |h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{ie}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2} \right) \quad (6)$$

با ترکیب معادلات (5) و (6) ظرفیت محرمانگی انتقال رله AF به صورت زیر بدست می‌آید:

$$C_i = C_{id} - C_{ie} = \frac{1}{2} \log_2 \left(\frac{1 + \frac{(P - P_s) |h_{id}|^2 |h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{id}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2}}{1 + \frac{(P - P_s) |h_{ie}|^2 |h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{ie}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2}} \right) \quad (7)$$

با ساده کردن عبارت فوق به عبارت زیر می‌رسیم:

$$C_i = \frac{1}{2} \log_2 \left(\frac{-a P_s^3 + b P_s^2 + c P_s + d}{-e P_s^3 + f P_s^2 + g P_s + h} \right) \quad (8)$$

ضرایب معادله (8) را می‌توان به سادگی از روی معادله (7) بدست آورد. حال که C_i را به ازای یک P_s مجهول محاسبه کردیم، می‌توانیم P_s بهینه که منجر به C_i ماکزیمم می‌شود را نیز محاسبه کنیم.

برای ماکزیمم شدن C_i در معادله (8) کافی است از عبارت داخل لگاریتم برحسب P_s مشتق گرفته و برابر صفر قرار دهیم:

$$AP_s^4 + BP_s^3 + CP_s^2 + DP_s + E = 0 \quad (9)$$

که ضرایب این معادله به آسانی و با مشتق گیری و برابر صفر قرار دادن عبارت داخل لگاریتم معادله (8) بدست می‌آید. معادله ۴ جمله‌ای بالا دارای چهار جواب است که جواب مورد نظر ما، جوابی است که حقیقی بوده و در $(0, P)$ قرار داشته باشد.

۴ - احتمال قطع محرمانگی

یک رخداد قطع زمانی اتفاق می‌افتد که ظرفیت محرمانگی از یک مقدار از پیش تعیین شده‌ای کمتر شود. بنابراین می‌توانیم احتمال قطع محرمانگی روش پیشنهادی را به صورت زیر بازگو کنیم:

$$SOP = Pr [C_i < R_0] = Pr \left[\frac{1 + \frac{(P - P_s) |h_{id}|^2 |h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{id}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2}}{1 + \frac{(P - P_s) |h_{ie}|^2 |h_{si}|^2 P_s}{[(P - P_s) \cdot |h_{ie}|^2 + P_s \cdot |h_{si}|^2] \sigma_n^2}} < 2^{2R_s} | P_s \right] \quad (10)$$

$$\begin{aligned}
& \int_{\infty}^{\infty} f_{P_S|h_{Si}|^2}(x) \int_0^x f_{(P-P_S)|h_{ie}|^2}(z) \int_0^x f_{(P-P_S)|h_{id}|^2}(y) dy dz dx + \\
& \int_a^{\infty} f_{P_S|h_{Si}|^2}(x) \int_{\frac{x-a}{b}}^x f_{(P-P_S)|h_{ie}|^2}(z) \int_0^x f_{(P-P_S)|h_{id}|^2}(y) dy dz dx = \\
& 1 - e^{-\frac{a}{P_S\sigma_{Si}^2}} + \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \left[e^{-a\left(\frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)} - 1 \right] + \\
& \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \left[e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)} - 1 \right] - \\
& \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \\
& \cdot \left[e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)} - 1 \right] + e^{-\frac{a}{P_S\sigma_{Si}^2}} - \\
& \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{b(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} e^{-\frac{a}{P_S\sigma_{Si}^2}} + \\
& \frac{1}{(P-P_S)\sigma_{ie}^2 P_S\sigma_{Si}^2} \left(\frac{b}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} \right)^{-1} \\
& \cdot \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{b(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)} - \\
& \frac{1}{(P-P_S)\sigma_{ie}^2} \left(\frac{b}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} \right)^{-1} e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)} + \\
& \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{b(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} e^{-\frac{a}{P_S\sigma_{Si}^2}} - \\
& \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{b(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)} \\
& - \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} e^{-a\left(\frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)} + \\
& \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \\
& \cdot e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)}
\end{aligned} \tag{23}$$

حال که I_1, I_2, I_3 و I_4 را محاسبه کردیم، می‌توانیم جمله اول معادله (10) را با در نظر گرفتن M رله به صورت زیر بنویسیم:

$$\begin{aligned}
I &= \prod_{k=1}^M \left[2 - \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \left\{ \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \right. \right. \right. \\
& \left. \left. \frac{1}{(P-P_S)\sigma_{ie}^2} \right) - \frac{1}{(P-P_S)\sigma_{ie}^2 P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \right\} - \\
& \frac{e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)}}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \left\{ \frac{1}{(P-P_S)\sigma_{ie}^2} \left(\frac{b}{(P-P_S)\sigma_{id}^2} + \right. \right. \\
& \left. \left. \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{b}{P_S\sigma_{Si}^2} \right)^{-1} - 1 \right\} + \frac{e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)}}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \right. \\
& \left. \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \left\{ \frac{1}{(P-P_S)\sigma_{ie}^2} \left(\frac{b}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} \right)^{-1} - 1 \right\} - \\
& \frac{1}{P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} - \frac{e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)}}{(P-P_S)\sigma_{ie}^2} \left(\frac{b}{(P-P_S)\sigma_{id}^2} + \right. \\
& \left. \frac{1}{(P-P_S)\sigma_{ie}^2} \right)^{-1} \left. \right]
\end{aligned} \tag{24}$$

از آنجایی که برای تقریب احتمال قطع محرمانگی داشتیم:

$$SOP \approx I \cdot f(P_S) \tag{25}$$

$$\begin{aligned}
I_2 &= \Pr\{P_S|h_{Si}|^2 < a + b(P-P_S)|h_{ie}|^2, \\
& P_S|h_{Si}|^2 < (P-P_S)|h_{id}|^2, P_S|h_{Si}|^2 > (P-P_S)|h_{ie}|^2\} = \\
& \int_0^{\infty} f_{(P-P_S)|h_{ie}|^2}(z) \int_z^{\infty} f_{P_S|h_{Si}|^2}(x) \int_x^{\infty} f_{(P-P_S)|h_{id}|^2}(y) dy dx dz \\
&= \frac{1}{(P-P_S)\sigma_{ie}^2 P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \\
&\cdot \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} - \frac{e^{-a\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2}\right)}}{(P-P_S)\sigma_{ie}^2 P_S\sigma_{Si}^2} \\
&\cdot \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \left(\frac{b}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{b}{P_S\sigma_{Si}^2} \right)^{-1}
\end{aligned} \tag{18}$$

که $a = 2^{2R_0} \sigma_n^2 - \sigma_n^2$ و $b = 2^{2R_0}$ است. با مرتب کردن جمله سوم معادله (15) نیز داریم:

$$\begin{aligned}
I_3 &= \Pr\{(P-P_S)|h_{id}|^2 < a + bP_S|h_{Si}|^2, \\
& P_S|h_{Si}|^2 > (P-P_S)|h_{id}|^2, P_S|h_{Si}|^2 < (P-P_S)|h_{ie}|^2\}
\end{aligned} \tag{19}$$

از آنجایی که داریم: $P_S|h_{Si}|^2 > (P-P_S)|h_{id}|^2$ و از آنجایی که کمترین مقدار ممکن b برابر یک است و a همواره مقداری نامنفی است، پس با برقرار بودن عبارت دوم داخل احتمال، عبارت اول نیز برقرار خواهد بود، بنابراین می‌توانیم I_3 را به صورت زیر بنویسیم:

$$\begin{aligned}
I_3 &= \Pr\{P_S|h_{Si}|^2 > (P-P_S)|h_{id}|^2, P_S|h_{Si}|^2 < (P-P_S)|h_{ie}|^2\} = \\
& \int_0^{\infty} f_{(P-P_S)|h_{ie}|^2}(z) \int_0^z f_{P_S|h_{Si}|^2}(x) \int_0^x f_{(P-P_S)|h_{id}|^2}(y) dy dx dz \\
&= 1 - \frac{1}{(P-P_S)\sigma_{ie}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} + \\
& \frac{1}{(P-P_S)\sigma_{ie}^2 P_S\sigma_{Si}^2} \left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} \\
& \cdot \left[\left(\frac{1}{(P-P_S)\sigma_{id}^2} + \frac{1}{(P-P_S)\sigma_{ie}^2} + \frac{1}{P_S\sigma_{Si}^2} \right)^{-1} - (P-P_S)\sigma_{ie}^2 \right]
\end{aligned} \tag{20}$$

برای محاسبه I_4 آن را به چند عبارت تجزیه کرده و از قانون کل استفاده می‌کنیم. یک بار فرض می‌کنیم $P_S|h_{Si}|^2 < a$ باشد از آنجا که $(P-P_S)|h_{ie}|^2$ و b هر دو مقادیر غیرمنفی هستند، پس با برقرار بودن عبارت دوم، عبارت اول نیز برقرار خواهد بود و در این صورت می‌توانیم بنویسیم:

$$\begin{aligned}
I_{41} &= \Pr\{(P-P_S)|h_{id}|^2 < P_S|h_{Si}|^2, \\
& P_S|h_{Si}|^2 > (P-P_S)|h_{ie}|^2\}
\end{aligned} \tag{21}$$

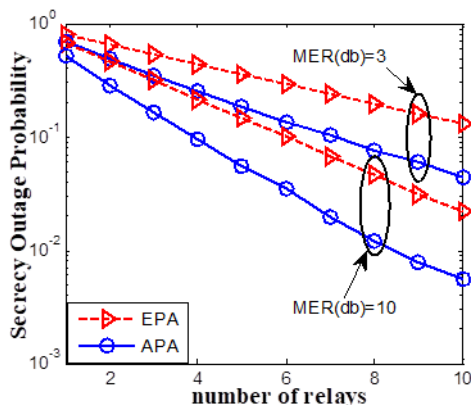
و یک بار هم فرض می‌کنیم که $P_S|h_{Si}|^2 > a$ که در این صورت خواهیم داشت:

$$\begin{aligned}
I_{42} &= \Pr\{(P-P_S)|h_{id}|^2 < \min\{a + b(P-P_S)|h_{ie}|^2, P_S|h_{Si}|^2\}, \\
& P_S|h_{Si}|^2 > (P-P_S)|h_{ie}|^2\} = \\
& \Pr\left\{(P-P_S)|h_{id}|^2 < a + b(P-P_S)|h_{ie}|^2, (P-P_S)|h_{ie}|^2 < \frac{P_S|h_{Si}|^2 - a}{b}\right\} + \\
& \Pr\left\{(P-P_S)|h_{id}|^2 < P_S|h_{Si}|^2, \frac{P_S|h_{Si}|^2 - a}{b} < (P-P_S)|h_{ie}|^2 < P_S|h_{Si}|^2\right\}
\end{aligned} \tag{22}$$

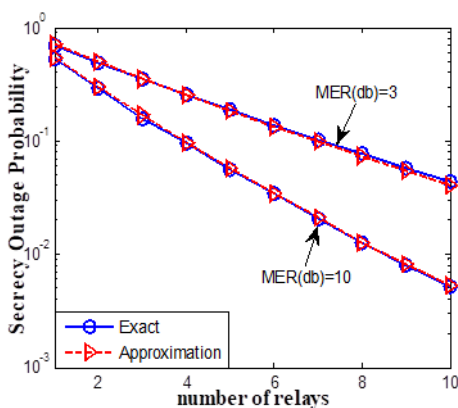
حال می‌توانیم I_4 را به کمک معادلات (21) و (22) بدین صورت بدست آوریم:

$$\begin{aligned}
I_4 &= I_{41} + I_{42} = \\
& \int_0^a f_{P_S|h_{Si}|^2}(x) \int_0^x f_{(P-P_S)|h_{ie}|^2}(z) \int_0^x f_{(P-P_S)|h_{id}|^2}(y) dy dz dx +
\end{aligned}$$

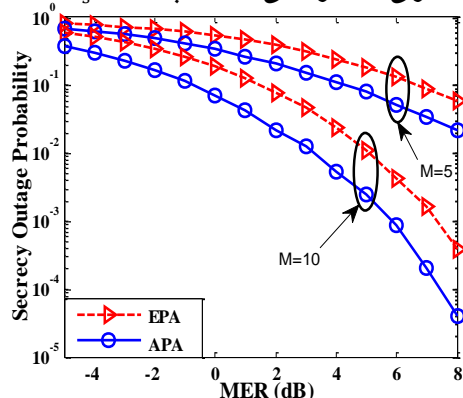
افزایش یابد، ظرفیت محرمانگی نیز افزایش یافته و احتمال قطع محرمانگی کاهش می‌یابد.



شکل ۲: احتمال قطع محرمانگی به ازای تعداد رله‌های مختلف برای تخصیص توان تطبیقی و برابر با $R_s = 0.5$



شکل ۳: عبارت تقریبی و دقیق احتمال قطع محرمانگی روش تخصیص توان تطبیقی به ازای تعداد رله‌های مختلف با $R_s = 0.5$



شکل ۴: احتمال قطع محرمانگی به ازای λ_{de} های مختلف برای طرح تخصیص توان تطبیقی و برابر با $R_s = 0.5$

حال باید به محاسبه $f(P_s)$ بپردازیم، برای این کار کافی است که $(0, P)$ را به چند بخش مساوی تقسیم کرده (مثلاً ۱۰ بخش) و به ازای چندین آزمایش مستقل مقدار P_s را محاسبه کنیم (مثلاً ۱۰۰۰ بار) و با تقسیم تعداد P_s هر بخش بر کل تعداد آزمایش‌ها چگالی احتمال P_s بدست می‌آید. حال مقدار میانی هر بخش را به عنوان مقدار P_s در معادله (25) قرار داده و مقدار I متناظر با آن P_s را از معادله (24) بدست آوریم. اگر تعداد بخش‌های مساوی بین ۰ و P را ۱۰ در نظر بگیریم ($N = 10$)، این مراحل بدین صورت فرمول‌بندی می‌شوند:

$$SOP \approx \sum_{i=1}^N I(P_s(i)) \cdot f(P_s(i)) \quad (23)$$

که در آن $P_s(i) = i \cdot \frac{P}{N} - \frac{P}{2N}$ است.

۵- نتایج عددی و شبیه‌سازی

در این بخش عملکرد روش تخصیص توان تطبیقی را مورد بررسی قرار می‌دهیم و آن را با روش تخصیص توان برابر مقایسه می‌کنیم. در اینجا نشان می‌دهیم که تخصیص توان تطبیقی پیشنهادی عملکرد بهتری نسبت به تخصیص توان برابر دارد و همچنین نشان می‌دهیم که عبارت تقریبی که برای احتمال قطع محرمانگی بدست آوردیم به احتمال قطع محرمانگی دقیق بسیار نزدیک است. در مدل سیستم ما رله‌ها نیمه دوطرفه هستند یعنی نمی‌توانند به طور همزمان دریافت و ارسال را انجام دهند. کانال بین هر دو گره رایلی فیدینگ است و ضرایب کانال بین دو گره توزیع نمایی با پارامترهای σ_{id}^2 و σ_{ie}^2 دارند. $\lambda_{de} = \frac{\sigma_{sd}^2}{\sigma_{se}^2}$ نسبت کانال اصلی به کانال شنود (MER) است. پارامترهای توابع چگالی احتمال ضرایب کانال به صورت $\sigma_{id}^2 = \sigma_{ie}^2 = \alpha_{ie} \sigma_{se}^2$ $\alpha_{si} = \alpha_{id} = \alpha_{ie}$ تعریف می‌شوند. $\sigma_{si}^2 = \alpha_{si} \sigma_{sd}^2$ و $\alpha_{id} \sigma_{sd}^2$ با $\sigma_n^2 = -60 \text{dbm}$ و $P = 1 \text{dbm}$ ، 0.5 استفاده از شبیه‌سازی مونت کارلو و انجام 5×10^4 آزمایش با کانال‌های مستقل و متوسط‌گیری بر روی آن‌ها بدست آمده است.

شکل (2) احتمال قطع محرمانگی را برای روش تخصیص توان تطبیقی پیشنهادی و تخصیص توان برابر به ازای تعداد رله‌های متفاوت و برای دو λ_{de} (MER) متفاوت نشان می‌دهد. همان‌طور که مشاهده می‌کنید احتمال قطع محرمانگی روش تخصیص توان پیشنهادی نسبت به تخصیص توان برابر به طور قابل توجهی کاهش یافته است و احتمال قطع محرمانگی هر دو روش با افزایش تعداد رله‌ها کاهش می‌یابد. همچنین می‌بینیم که با افزایش λ_{de} احتمال قطع محرمانگی کاهش می‌یابد.

شکل (3) احتمال قطع محرمانگی را برای روش تخصیص توان تطبیقی پیشنهادی به ازای تعداد رله‌های متفاوت و از دو روش متفاوت که یکی بر اساس عبارت تقریبی بدست آمده در مقاله و دیگری شبیه‌سازی دقیق می‌باشد را نشان می‌دهد. همان‌طور که مشاهده می‌کنیم عبارت تقریبی بدست آمده بسیار نزدیک به عبارت دقیق است.

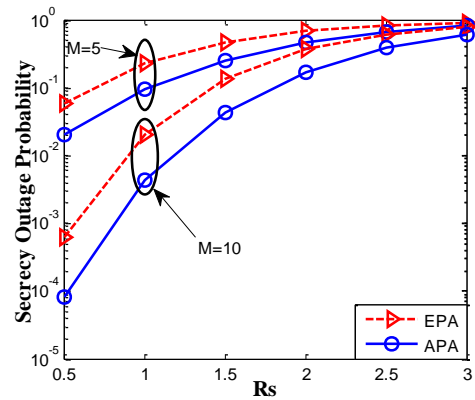
شکل (4) احتمال قطع محرمانگی روش تخصیص توان تطبیقی پیشنهادی و روش تخصیص توان برابر را به ازای مقادیر مختلف λ_{de} به نمایش می‌گذارد. مشاهده می‌شود که با افزایش λ_{de} احتمال قطع محرمانگی کاهش می‌یابد یعنی هرچه نسبت کانال اصلی به کانال شنود

محرمانگی طرح پیشنهادی بدست آوردیم و نشان دادیم که بسیار نزدیک به مقدار دقیق می‌باشد. باید به این نکته توجه داشت که ما در این مقاله فقط به تخصیص توان بین یک رله بهینه و منبع پرداختیم و در آینده می‌خواهیم تخصیص توان و نحوه پرتودهی و ارسال نویز مشارکتی توأم با چندین رله را بررسی کنیم. همچنین محاسبه یک عبارت بسته دقیق برای طرح پیشنهادی یک چالش جذاب در آینده خواهد بود.

مراجع

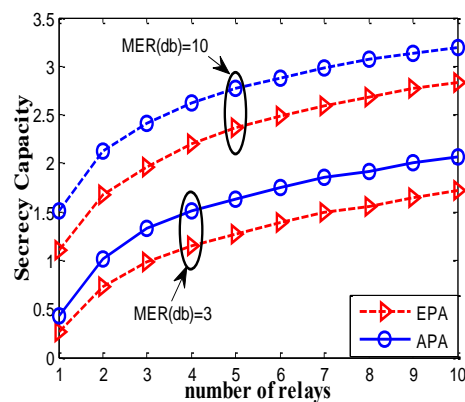
- [1] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. IEEE*, vol. PP, no. 99, pp. 1-39, May 2016.
- [2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355-1387, Oct. 1975.
- [3] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inf Theory*, vol. 24, no. 4, pp. 451-456, Jul. 1978.
- [4] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961-4972, Oct. 2007.
- [5] Ashish Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088-3104, Jul. 2010.
- [6] K. P. Peppas, N. C. Sagias and A. Maras, "Physical layer security for multiple-antenna systems: A unified approach," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 314-328, Jan. 2016.
- [7] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875-1888, Mar. 2010.
- [8] J. Li, A. P. Petropulu, and S. Weber, "On cooperative relaying schemes for wireless physical layer security," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4985-4997, Oct. 2011.
- [9] J. Kim, A. Ikhlef, and R. Schober, "Combined relay selection and cooperative beamforming for physical layer security," *J Commun Netw-s Kor*, vol. 14, no. 4, pp. 364-373, Aug. 2012.
- [10] M. Yuksel and E. Erkip, "Secure communication with a relay helping the wiretapper," in *Proc. 2007 IEEE Information Theory Workshop*, Lake Tahoe, CA, Sep. 2007.
- [11] Y. Zou, X. Wang, and W. Shen, "Optimal relay selection for physical layer security in cooperative wireless networks," *IEEE J Sel Area Comm*, vol. 31, no. 10, pp. 2099-2111, Oct. 2013.
- [12] E. Tekin and A. Yener, "The general Gaussian multiple access and two-way wire-tap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735-2751, Jun. 2008.
- [13] L. Lai and H. E. Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005-4019, Sep. 2008.
- [14] C. L. Wang, T. N. Cho, and F. Liu, "Power allocation and jammer selection of a cooperative jamming strategy for physical-layer security," in *Proc. IEEE VTC*, May 2014, pp. 1-5.
- [15] B. V. Nguyen and K. Kim, "Secrecy outage probability of optimal relay selection for secure AnF cooperative networks," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2086-2089, Dec. 2015.

شکل (5) احتمال قطع محرمانگی روش‌های تخصیص توان تطبیقی و تخصیص توان برابر را به ازای مقادیر مختلف R_s نشان می‌دهد. مشاهده می‌کنیم که با افزایش نرخ محرمانگی هدف، احتمال قطع محرمانگی نیز افزایش می‌یابد.



شکل ۵: احتمال قطع محرمانگی به ازای R_s های مختلف برای طرح تخصیص توان تطبیقی و برابر با $\lambda_{de} = 10$

شکل (6) ظرفیت محرمانگی روش‌های تخصیص توان تطبیقی و تخصیص توان برابر را به ازای مقادیر مختلف تعداد رله‌های کمکی نشان می‌دهد. مشاهده می‌کنیم که با افزایش تعداد رله‌های کمکی، ظرفیت محرمانگی نیز افزایش می‌یابد.



شکل ۶: ظرفیت محرمانگی به ازای تعداد رله‌های مختلف برای طرح تخصیص توان تطبیقی و برابر

۶- نتیجه گیری

در این مقاله یک طرح انتخاب رله را برای بهبود امنیت لایه فیزیکی در شبکه‌های بی‌سیم مشارکتی بررسی کردیم و یک روش تخصیص توان تطبیقی را برای آن پیشنهاد کردیم. به منظور مقایسه ما یک روش تخصیص توان برابر را نیز بررسی کردیم و نشان دادیم که روش پیشنهادی به طور قابل توجهی عملکرد محرمانگی سیستم را بهبود می‌بخشد. همچنین یک عبارت قطع تقریبی برای احتمال قطع



تخصیص عادلانه منابع در سیستم‌های چندپخش مبتنی بر MIMO-OFDM با استفاده از الگوریتم ژنتیک

علیرضا غلامرضائی^۱، حمید فرخی^۲

^۱ دانشجوی دکترای مخابرات، دانشکده برق و کامپیوتر، دانشگاه بیرجند، بیرجند

alireza.gholamrezaee@birjand.ac.ir

^۲ دانشیار گروه مخابرات، دانشکده برق و کامپیوتر، دانشگاه بیرجند، بیرجند

h.farrokhi@birjand.ac.ir

چکیده

تمرکز اصلی در این مقاله، تخصیص توان و زیرحامل در یک سیستم چندپخش مبتنی بر MIMO-OFDM و بمنظور افزایش ظرفیت کلی سیستم است، ضمن آنکه موضوع تامین سطح مشخص از QoS برای کاربران در گروه‌های چندپخش نیز مطرح می‌گردد. برای رسیدن به این هدف، یک الگوریتم تخصیص منابع مناسب با پیچیدگی کم و با حصول اطمینان از تخصیص سطح مطلوبی از منابع برای تمام کاربران در گروه‌های چندپخش و برقراری عدالت معرفی می‌شود. بدین منظور یک الگوریتم ژنتیک (GA) جهت تخصیص زیرحامل‌ها معرفی می‌گردد و برای تخصیص بهینه توان از روش واترفیلینگ استفاده خواهد شد. نتایج شبیه‌سازی نشان می‌دهد که روش ارائه شده، باعث افزایش ظرفیت کلی سیستم نسبت به روش‌های پیشین می‌شود.

کلمات کلیدی

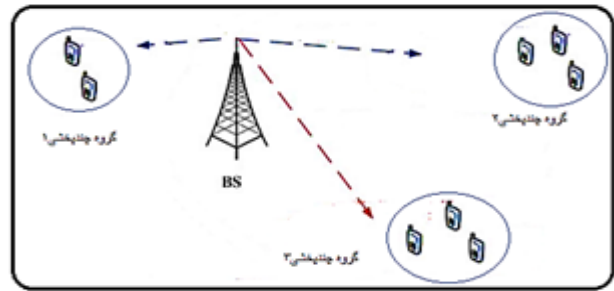
تخصیص منابع، چندپخش، سیستم‌های MIMO-OFDM، عدالت، الگوریتم ژنتیک

یک کاربر به یک گروه ارسال می‌شود. این موضوع در شکل (۱) نشان داده شده است (البته ممکن است کاربران در یک مکان نباشند). این مزیت فرایند چندپخش، می‌تواند بهره‌وری پهنای باند را افزایش دهد و توان ارسال را بطور قابل ملاحظه‌ای نسبت به فرایند تک پخش، کاهش دهد [3]. با توجه به حضور چندین کاربر با موقعیت مکانی مختلف در گروه‌های مختلف سیستم چندپخش، یکی از مهمترین چالش‌ها در سیستم‌های چندپخش تعیین نرخ ارسال به گروه‌های چندپخش است. یکی از رایج‌ترین روش‌های تعیین نرخ ارسال به گروه‌های چندپخش، ارسال با نرخ است که بدترین کاربر حاضر در یک گروه (کاربری که بدترین کیفیت کانال را دارد) نیز بتواند دیتای ارسال شده را درک کند [4]. با انتخاب این روش می‌توان اطمینان داشت که تمام کاربران یک گروه، می‌توانند دیتاهای ارسال شده را بدون مشکل دریافت کنند.

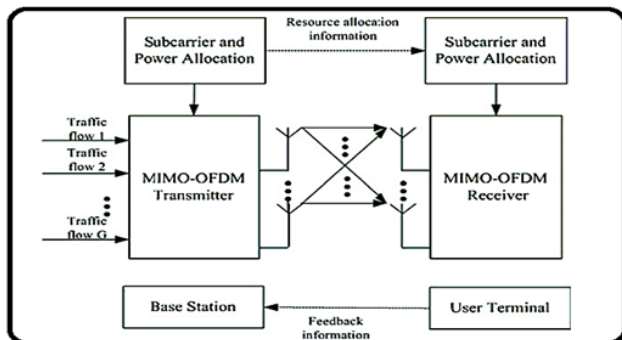
۱- مقدمه

باتوجه به لزوم پشتیبانی شبکه‌های بی‌سیم نسل آینده از سرویس‌های گوناگون مولتی‌مدیا با درجه کیفی^۱ متفاوت، شبکه‌های MIMO-OFDM^۲، بخاطر بهره‌وری طیفی بالا و انعطاف‌پذیری در مدیریت منابع، یک تکنولوژی امیدبخش برای این شبکه‌ها محسوب می‌شوند. یکی دیگر از سرویس‌ها در سیستم مخابرات بی‌سیم، ارسال چندپخش^۳ است که امروزه بسیاری از خدمات چندرسانه‌ای نظیر ویدئو کنفرانس توسط آن پیاده‌سازی شده است و برای سیستم مخابرات بی‌سیم جهانی در نسل سوم مخابرات سیار نیز تعریف گردیده است [1,2]. برخلاف ارسال تک‌پخش^۴، که در آن در هر اسلات زمانی، دیتاها تنها به یک کاربر تحویل داده می‌شود، در ارسال چندپخش زمانیکه چندین کاربر محتوای یکسانی درخواست کنند، سیستم چندپخش به آن کاربران اجازه تشکیل گروه و به اشتراک گذاشتن منابع تخصیصی را می‌دهد و هر زیرحامل به جای ارسال به

در این مقاله فرض شده است که BS اطلاعات حالت کانال^{۱۲} (CSI) تمام کاربران داخل سیستم را دقیقاً می‌داند، بعلاوه فرض می‌شود سیگنال‌های ارسالی یک فیدینگ آهسته^{۱۳} را تجربه می‌کنند، یعنی ضرایب کانال در طول فرایند تخصیص منبع ثابت است. همچنین با در نظر گرفتن تعداد Nt آنتن در BS و تعداد Nr آنتن برای هر کاربر بی‌سیم، می‌توان ماتریس کانال کاربر k ام بر روی زیر حامل n ام که یک ماتریس با بعد Nt*Nr است، را بصورت زیر نوشت:



شکل ۱: مثالی از یک سیستم چندپخش‌ی با چندگروه



شکل ۲: بلوک دیاگرام یک سیستم مبتنی بر MIMO-OFDM

$$\mathbf{H}_{kn} = \begin{pmatrix} h_{11}^{kn} & h_{12}^{kn} & \dots & h_{1N_r}^{kn} \\ h_{21}^{kn} & h_{22}^{kn} & \dots & h_{2N_r}^{kn} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_t,1}^{kn} & h_{N_t,2}^{kn} & \dots & h_{N_t,N_r}^{kn} \end{pmatrix} \quad (1)$$

که در رابطه (۱)، h_{rt}^{kn} گین کانال t امین آنتن فرستنده به r امین آنتن گیرنده کاربر k ام و بر روی زیرحامل n ام می‌باشد. همچنین مقادیر ویژه $\mathbf{H}_{kn}\mathbf{H}_{kn}^T$ بصورت $\{\lambda_{kn}^i\}_{i=1}^M$ فرض شده است که در آن $M = \min(N_r, N_t)$ می‌باشد. همچنین فرض شده است که در این رابطه $\sum_{k=1}^K \sum_{n=1}^{N_t} p_{kn} = P_t$ ، که در این رابطه p_{kn} توان تخصیص داده شده به کاربر k ام و بر روی زیرحامل n ام است و P_t کل توان ارسالی در BS است.

با فرض استفاده از مدولاسیون QAM مرتبه M، می‌توان BER^{۱۴} را از رابطه زیر تقریب زد [12]:

$$BER_{kn}^i \approx \frac{1}{5} \exp \left[\frac{-1.5 p_{kn} \lambda_{kn}^i}{N_0 (2^{b_{kn}} - 1)} \right] \quad (2)$$

$$b_{kn}^i = \log_2 \left(1 + \frac{p_{kn} \lambda_{kn}^i}{N_0 \tau_{kn}^i} \right) \quad (3)$$

که در رابطه (۳)، $\tau_{kn}^i = -\ln(5BER_{kn}^i)/1.5$ ، برای سادگی فرض می‌کنیم، $BER_{kn}^i = BER$ و $\tau_{kn}^i = \tau$. با توجه به اینکه هر کاربر در هر گروه فیدینگ متفاوتی را تجربه می‌کند، یک روش معمول، ارسال با نرخ است که توسط بدترین کاربر از لحاظ وضعیت کانال تعیین شده است. این روش تضمین می‌کند که تمام کاربران درون یک

مطالعات زیادی در زمینه تخصیص منابع صورت گرفته است که اکثر آنها بر روی سیستم‌های با ارسال تک‌پخش‌ی کار کرده‌اند و مواردی که به تحقیق بر روی سیستم‌های با ارسال چندپخش‌ی پرداخته‌اند نیز بر سیستم‌های مبتنی بر OFDM تمرکز داشته‌اند [5-7]. در [8] یک الگوریتم به منظور افزایش ظرفیت سیستم‌های چندپخش‌ی مبتنی بر MIMO-OFDM ارائه شده است، که در آن هر زیرحامل^{۱۵} به گروهی که بهترین موقعیت کانال و تعداد بیشتری کاربر دارد، اختصاص می‌یابد و توان را نیز بطور بهینه به زیرحامل‌ها تخصیص می‌دهد، اما در این مقاله بحث عدالت^{۱۶} در توزیع منابع مطرح نشده است (موضوع عدالت که در اکثر مقالات مورد توجه قرار نگرفته، نقش مهمی را در دنیای تجاری واقعی ایفا می‌کند) و این احتمال وجود دارد گروهی که از نظر موقعیت کانال در وضعیت خوبی به سر نمی‌برد، بطور کل از دریافت منابع محروم شود. در [9] نیز به تخصیص منابع در سیستم‌های چندپخش‌ی مبتنی بر MIMO-OFDM پرداخته شده است که در آن تخصیص کمترین زیرحامل درخواستی از طرف هر گروه نیز تامین و تضمین می‌شود، اما در آن مقاله توان بصورت بهینه اختصاص نمی‌یابد که این امر موجب افت بهره‌وری سیستم می‌شود. همچنین اکثر مطالعات در این حوزه از روش‌های کلاسیک برای حل مسائل بهینه‌سازی بهره‌گرفته‌اند. در [10, 11] تخصیص زیرحامل با استفاده از روش الگوریتم ژنتیک (GA)^{۱۷} تخصیص پیدا کرده است، ولی سیستم مورد بحث در آن‌ها سیستم OFDM می‌باشد و همچنین توان در آن‌ها بصورت بهینه تخصیص داده نشده است. در روش ارائه شده در این مقاله زیرحامل‌ها با استفاده از الگوریتم ژنتیک و توان نیز با روش واترفیلینگ^{۱۸} تخصیص می‌یابند به گونه‌ای که موجب افزایش بهره‌وری سیستم می‌شود و علاوه بر این، بحث عدالت در تخصیص زیرحامل‌ها نیز مدنظر قرار خواهد گرفت. ادامه مقاله بصورت زیر سازمان‌دهی شده است.

در فصل ۲ مدل سیستم مورد استفاده و فرمول‌بندی مسئله نشان داده خواهد شد، فصل ۳ به بیان الگوریتم پیشنهادی می‌پردازد. فصل ۴ شبیه‌سازی الگوریتم‌ها و نتایج را نشان می‌دهد. نتیجه‌گیری در فصل ۵ بیان می‌شود و در انتها مراجع مورد استفاده، آورده شده است.

۲- مدل سیستم

بلوک دیاگرام سیستم چندپخش‌ی بی‌سیم مبتنی بر MIMO-OFDM مورد استفاده در این مقاله در شکل (۲) نشان داده شده است. در اینجا یک BS^{۱۱} در نظر گرفته می‌شود که به K کاربر بی‌سیم و بر روی N زیرحامل سرویس می‌دهد.

زیرحامل‌ها با استفاده از الگوریتم ژنتیک بطور بهینه تخصیص می‌یابند، در مرحله بعد با توجه به زیرحامل‌های تخصیص داده شده در مرحله اول، توان بصورت بهینه تخصیص داده می‌شود. در این مقاله در مرحله دوم و برای تخصیص بهینه توان از روش واترفیلینگ استفاده می‌شود.

۳- الگوریتم پیشنهادی

همانطور که گفته شده در این مقاله از یک روش زیربهینه برای تخصیص منابع و بمنظور افزایش کلی ظرفیت استفاده می‌شود. بدین صورت که مسئله به دو مرحله تقسیم می‌شود، مرحله اول تخصیص زیرحامل با فرض تخصیص یکسان توان و در مرحله بعد، تخصیص توان با در نظر گرفتن زیرحامل‌های تخصیصی مرحله اول صورت می‌گیرد. مرحله تخصیص زیرحامل با استفاده از الگوریتم ژنتیک انجام می‌پذیرد، بدین منظور در ادامه به معرفی این الگوریتم پرداخته می‌شود و سپس روش استفاده از آن در تخصیص زیرحامل توضیح داده می‌شود.

۳-۱- الگوریتم ژنتیک

الگوریتم‌های فراابتکاری یکی از انواع الگوریتم‌های بهینه‌سازی تصادفی (احتمالی) هستند که دارای راهکارهای برون رفت از بهینه محلی می‌باشند و برای طیف گسترده‌ای از مسائل کاربرد دارند. از الگوریتم‌های شناخته شده فراابتکاری می‌توان به الگوریتم ژنتیک اشاره کرد که در این مقاله از آن استفاده شده است.

الگوریتم ژنتیک یک تکنیک جستجو در علم رایانه برای یافتن راه‌حل تقریبی در مسائل بهینه‌سازی و جستجو است. الگوریتم ژنتیک نوع خاصی از الگوریتم‌های تکامل است که از تکنیک‌های زیست‌شناسی فرگشتی مانند وراثت و جهش استفاده می‌کند. این الگوریتم برای اولین بار توسط جان هالند^{۱۷} معرفی شد. به طور کلی، الگوریتم‌های ژنتیکی از اجزاء زیر تشکیل می‌شوند:

کروموزوم^{۱۸}:

در الگوریتم ژنتیک، هر کروموزوم نشان دهنده یک نقطه در فضای جستجو و یک راه‌حل ممکن برای مسئله مورد نظر است. خود کروموزوم‌ها (راه‌حل‌ها) از تعداد ثابتی ژن^{۱۹} (متغیر) تشکیل می‌شوند.

جمعیت^{۲۰}:

مجموعه‌ای از کروموزوم‌ها یک جمعیت را تشکیل می‌دهند. با تاثیر عملگرهای ژنتیکی بر روی هر جمعیت، جمعیت جدیدی با همان تعداد کروموزوم تشکیل می‌شود.

تابع برازندگی^{۲۱}:

به منظور حل هر مسئله با استفاده از الگوریتم ژنتیک، ابتدا باید یک تابع برازندگی برای آن مسئله ابداع شود. برای هر کروموزوم، این تابع عددی غیر منفی را برمی‌گرداند که نشان دهنده شایستگی یا توانایی فردی آن کروموزوم است.

عملگرهای ژنتیکی:

در الگوریتم ژنتیک، در طی مرحله تولید مثل^{۲۲} از عملگرهای ژنتیکی استفاده می‌شود. با تاثیر این عملگرها بر روی یک جمعیت، نسل^{۲۳}

گروه می‌تواند از خدمات ارسالی استفاده کند. بنابراین BS، دیتا را از طریق زیرحامل n به گروه gام با نرخ زیر ارسال می‌کند:

$$r_{gn} = \arg \min_{k \in k_g} \frac{B}{N} \sum_{i=1}^M \log_2 \left(1 + \frac{P_{kn} \lambda_{kn}^i}{N_0 \tau} \right) \quad (4)$$

که در رابطه فوق kg مجموعه کاربر گروه g و B پهنای باند کلی سیستم را نشان می‌دهد. از اینرو می‌توان مجموع نرخ دیتای گروه g روی زیرحامل n را با استفاده از رابطه (۴)، بصورت زیر نوشت:

$$R_{gn} = |k_g| \cdot r_{gn} \quad (5)$$

که منظور از $|k_g|$ ، تعداد کاربرگروه g می‌باشد.

با توجه به فرضیات گفته شده، می‌توان مسئله بهینه‌سازی مورد نظر در این مقاله را بصورت زیر نوشت:

$$\max_{x_{gn} P_{kn}} \sum_{g=1}^G \sum_{n=1}^N x_{gn} R_{gn} \quad (6)$$

تحت شرایط:

$$\sum_{g=1}^G x_{gn} = 1; \quad \forall n. \quad (7)$$

$$x_{gn} \in \{0,1\}. \quad \forall g,n. \quad (8)$$

$$\sum_{n=1}^N x_{gn} \geq Q_g. \quad \forall g. \quad (9)$$

$$\sum_{n=1}^N \max_k P_{kn} \leq P_t \quad (10)$$

$$P_{kn} \geq 0. \quad \forall k,n. \quad (11)$$

در روابط فوق متغیر باینری x_{gn} تخصیص زیرحامل n به گروه g را نشان می‌دهد و در صورتیکه زیرحامل به گروه تخصیص یابد مقدار آن ۱ و در غیر اینصورت مقدارش برابر با ۰ است. قیود (۷) و (۸) بیان می‌کند که هر زیرحامل تنها به یک گروه می‌تواند اختصاص یابد و با رعایت این قیود تداخل داخل سلولی بوجود نمی‌آید. قید (۹) اختصاص حداقل زیرحامل درخواستی توسط هر گروه را تضمین می‌کند و با توجه به این قید می‌توان اطمینان داشت که هیچ گروهی از داشتن زیرحامل محروم نمی‌ماند. البته باید توجه داشت که مجموع زیرحامل‌های درخواستی توسط گروه‌ها نباید از کل زیرحامل‌های موجود بیشتر باشد. قیود (۱۰) و (۱۱) نیز در مورد تخصیص توان بکار می‌رود و تضمین می‌کند که مجموع توان اختصاصی بر روی زیرحامل‌ها از کل توان موجود در سیستم بیشتر نشود و همینطور توان منفی در روابط استفاده نشود. برای حل بهینه این مسئله باید توان زیرحامل بطور همزمان تخصیص پیدا کنند، اما با توجه به اینکه در این مسئله متغیرهای پیوسته و باینری بطور همزمان درگیر هستند و چون قیود مسئله محدب و خطی نیستند، این مسئله به یک مسئله NP-hard مبدل شده که حل بهینه^{۱۵} آن بسیار پیچیده است. از اینرو برای حل این مسئله از الگوریتم زیربهینه^{۱۶} بهره گرفته شده است.

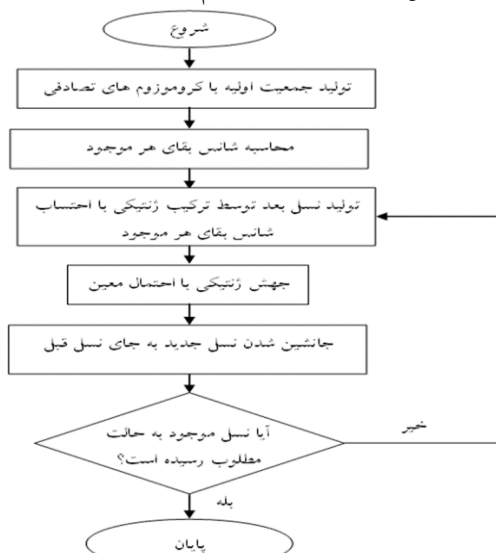
برای الگوریتم زیر بهینه مورد استفاده در این مقاله، مسئله مورد نظر به دو قسمت تقسیم می‌شود و سپس حل می‌گردد. برای این منظور در مرحله اول با فرض تخصیص توان یکسان بر روی زیرحامل‌ها،

و با PC نمایش داده می‌شود. در صورتی که بر روی یک زوج کروموزوم عمل آمیزش صورت نگیرد T فرزندان با تکرار نمودن والدین تولید می‌شوند.

عملگر جهش:

پس از اتمام عمل آمیزش، عملگر جهش بر روی کروموزوم‌ها اثر داده می‌شود. این عملگر یک ژن از یک کروموزوم را به طور تصادفی انتخاب نموده و سپس محتوای آن ژن را تغییر می‌دهد. اگر ژن از جنس اعداد دودویی باشد، آن را به وارونش تبدیل می‌کند و چنانچه متعلق به یک مجموعه باشد، مقدار یا عنصر دیگری از آن مجموعه را به جای آن ژن قرار می‌دهد.

احتمال انجام عمل جهش بر روی هر کروموزوم را نرخ جهش μ و با Pm نمایش می‌دهند. معمولاً این عدد را بسیار کوچک در نظر می‌گیرند. پس از اتمام عمل جهش، کروموزوم‌های تولید شده به عنوان نسل جدید شناخته شده و برای دور بعد اجرای الگوریتم ارسال می‌شوند. در شکل (۳) فلوچارت الگوریتم ژنتیک نشان داده شده است.



شکل ۳: فلوچارت الگوریتم ژنتیک

۳-۲- الگوریتم تخصیص منابع

با توجه با ماهیت گسسته الگوریتم ژنتیک برای تخصیص زیر حامل در این مقاله از این الگوریتم استفاده می‌شود. برای تولید جمعیت اولیه، کروموزوم بصورت نشان داده شده در شکل (۴) تعریف می‌گردد:

زیرحامل	۱	۲	۳	...	N
کروموزوم	۳	۲	G	...	۳

شکل ۴: نمونه ای از کروموزوم بکار رفته در تخصیص زیرحامل

همانطور که در شکل (۴) مشخص است، هر کروموزوم از N تعداد زیر حامل‌ها) ژن تشکیل شده است. برای مقادری اولیه کروموزوم‌ها اعداد صحیح بین ۱ تا G (تعداد گروه‌ها) بطور تصادفی داخل هر ژن قرار می‌گیرند که این امر نشان دهنده آن است که هر زیر حامل به

بعدي آن جمعیت تولید می‌شود. عملگرهای انتخاب^{۲۴}، آمیزش^{۲۵} و جهش^{۲۶} معمولاً بیشترین کاربرد را در الگوریتم‌های ژنتیک دارند. حال هر یک از عملگرهای فوق به صورت جداگانه معرفی می‌گردد. عملگر انتخاب:

این عملگر از بین کروموزوم‌های موجود در یک جمعیت، تعدادی کروموزوم را برای تولید مثل انتخاب می‌کند. کروموزوم‌های برانزنده‌تر شانس بیشتری دارند تا برای تولید مثل انتخاب شوند. روش‌های انتخاب متعددی برای استفاده در الگوریتم‌های ژنتیکی پیشنهاد شده‌اند که در ادامه برخی از این روش‌ها معرفی می‌شوند.

نمونه‌برداری به روش چرخ رولت:

در این روش، به هر فرد قطعه‌ای^{۲۷} از یک چرخ رولت مدور اختصاص داده می‌شود. اندازه این قطعه متناسب با برانزندگی آن فرد است. چرخ M بار چرخانده می‌شود که M تعداد افراد در جمعیت است. در هر چرخش، فرد زیر نشانگر چرخ انتخاب می‌شود و در مخزن والدین نسل بعد قرار می‌گیرد. این روش می‌تواند به صورت زیر پیاده‌سازی شود:

۱- نرخ انتظار کل افراد جمعیت را جمع کنید و حاصل آن را T بنامید.
۲- مراحل زیر را M بار تکرار کنید:
یک عدد تصادفی r بین ۰ و T انتخاب کنید.

در میان افراد جمعیت برگردید و نرخ‌های انتظار آنها را با هم جمع کنید تا این که مجموع بزرگتر یا مساوی r شود. فردی که نرخ انتظارش باعث بیشتر شدن جمع از این حد می‌شود، به عنوان فرد برگزیده انتخاب می‌شود.

انتخاب نخبگان^{۲۸}:

روش انتخاب نخبگان که اولین بار توسط دی. جانگ^{۲۹} مطرح شده است، به عنوان الحاقی به خیلی از روش‌های انتخاب محسوب می‌شود که GA را وادار می‌کند تا بهترین افراد را در هر نسل نگه دارد. چنین افرادی در صورتی که برای تولید مثل انتخاب نشوند یا توسط آمیزش و جهش خراب شوند، ممکن است از بین بروند. بسیاری از محققین دریافته‌اند که انتخاب نخبگان به میزان قابل ملاحظه‌ای کارایی GA را افزایش می‌دهد.

عملگر آمیزش:

عملگر آمیزش بر روی یک زوج کروموزوم از نسل مولد عمل کرده و یک زوج کروموزوم جدید تولید می‌کند. عملگرهای آمیزش متعددی از قبیل، آمیزش تک نقطه‌ای^{۳۰} و آمیزش دو نقطه‌ای^{۳۱} وجود دارد. در آمیزش تک نقطه‌ای، یک موقعیت تصادفی بین دو ژن در نظر گرفته می‌شود. سپس تمامی ژن‌های طرف راست یا طرف چپ این موقعیت در کروموزوم‌های والد با یکدیگر جابجا می‌شوند تا کروموزوم‌های جدید بدست آیند. در آمیزش دو نقطه‌ای، دو موقعیت به صورت تصادفی انتخاب می‌شود و تمامی ژن‌های بین این دو موقعیت در کروموزوم‌های والد با یکدیگر جابجا می‌شوند. لازم به ذکر است که آمیزش معمولاً بر روی همه زوج کروموزوم‌های انتخاب شده برای جفت‌گیری به کار برده نمی‌شود. معمولاً احتمال آمیزش برای هر زوج کروموزوم بین ۰/۶ تا ۰/۹۵ در نظر گرفته می‌شود که به این عدد نرخ آمیزش^{۳۲} گفته می‌شود

زیرحامل‌های تخصیصی در مرحله قبل، یک الگوریتم تخصیص توان کارآمد ارائه می‌شود. در این مرحله بهره‌ی کانال ضعیف‌ترین کاربر مورد استفاده قرار می‌گیرد. روش واترفیلینگ را برای بدست آوردن بهینه اعمال می‌کنیم. توزیع توان روی زیرحامل‌ها بصورت زیر می‌باشد:

$$p_n^* = \max(0, p_n) \quad (12)$$

که منظور از p_n توان تخصیص داده شده به هر آنتن زیرحامل n می‌باشد که از رابطه‌ی زیر بدست می‌آید:

$$\sum_{i=1}^{M_{k_n, n}} \frac{\lambda_{k_n, n}^{(i)}}{\lambda_{k_n, n}^{(i)} p_n + N_0} + \alpha = 0, n = 1, 2, \dots, N. \quad (13)$$

که در رابطه (۱۳) k_n ، کاربر مورد نظر بر روی زیرحامل n می‌باشد، α سطح واترفیلینگ می‌باشد که شرط $\sum_{n=1}^N M_{k_n, n} p_n^* = P_t$ در طول فرایند الگوریتم باید برقرار باشد. الگوریتم واترفیلینگ مورد استفاده در این مقاله، یک روش تکرارشونده می‌باشد که با استفاده از آن می‌توان توزیع بهینه‌ی توان را برای رسیدن به حداکثر ظرفیت پیدا کرد.

۴- مدل‌سازی و نتیجه شبیه سازی

در این بخش نتایج شبیه‌سازی برای نشان دادن عملکرد الگوریتم پیشنهادی، ارائه شده است. مدل سیستم و پارامترهای استفاده شده در این شبیه سازی بصورت زیر می‌باشد:

در اینجا یک سیستم MIMO-OFDM با یک BS و تعداد آنتن‌های فرستنده و گیرنده ۲ در نظر گرفته می‌شود. سه گروه و تعداد کاربران در هر گروه برابر ۴ انتخاب شده است. پهنای باند ۱ MHz و AWGN، ۱ قرار داده شده است. همچنین کانال frequency-selective فرض شده که شامل ۳ مسیر مستقل Rayleigh می‌باشد. تعداد زیرحامل‌ها ۳۲، حداکثر شیفت داپلر ۶۰ Hz و $BER=10^{-3}$ فرض می‌گردد. برای الگوریتم ژنتیک نیز حداکثر تکرار ۱۵۰ و تعداد جمعیت کروموزوم‌ها ۱۰۰ قرار داده شده است. پارامترهای بکار رفته در جدول (۱) آورده شده است.

جدول ۱: مقادیر بکار رفته در شبیه سازی

مقدار	پارامتر
1MHz	B
32	N
3	G
4	K_g
8	Q_g
2	N_f
2	N_t
3	تعداد مسیرها

کدام گروه تخصیص یافته است. بعنوان مثال در شکل (۴)، زیر حامل اول به گروه سوم، زیر حامل دوم به گروه دوم، زیرحامل سوم به گروه G ام و ... تخصیص پیدا کرده است که با این روش مقداردهی، قید اول برقرار می‌گردد یعنی امکان تداخل رفع می‌گردد. پس از تعریف کروموزوم به فرم شکل (۴)، می‌توان مراحل تخصیص زیر حامل را بصورت زیر نوشت (با فرض تخصیص یکسان توان بر روی هر زیر حامل):

۱- تعیین ورودی: در ابتدا موارد زیر به عنوان ورودی به الگوریتم داده می‌شود:

ضرایب ماتریس بهره کانال (H_{km})، توان کل (P_t)، پهنای باند (B)، تعداد گروه‌ها (G)، تعداد کاربران داخل هر گروه ($|k_g|$)، تعداد زیر حامل‌ها (N)، تعداد زیرحامل مورد نیاز هر گروه (Q_g)، تعداد جمعیت اولیه (N_p) و حداکثر تعداد تکرار الگوریتم ژنتیک (L_{Max})

۲- مقدار دهی اولیه: در این مرحله با مقدار دهی تصادفی N_p کروموزوم با اعداد بین ۱ تا G، جمعیت اولیه تولید می‌گردد.

نکته: در این مرحله بعد از مقدار دهی، تعداد زیرحامل تخصیص داده شده به هر زیر حامل شمارش می‌شود، اگر تعداد زیر حامل‌های تخصیص داده شده به هر گروه از تعداد زیرحامل درخواستی توسط هر گروه کمتر باشد، کروموزوم‌ها مجدداً مقدار دهی می‌شوند. با این عمل، قید (۹) ارضا می‌گردد.

۳- محاسبه تابع برازندگی: در این مرحله با قرار دادن مقادیر هر کروموزوم (که نشان دهنده زیر حامل تخصیص داده شده به هر گروه می‌باشد) در رابطه (۶)، مقدار تابع برازندگی برای هر کروموزوم محاسبه می‌گردد.

۴- مرتب سازی: در این مرحله با توجه به مقدار تابع برازندگی هر کروموزوم، کروموزوم‌ها مرتب می‌شوند. کروموزومی که بیشترین مقدار تابع برازندگی را دارد در ابتدا و به همین ترتیب کروموزومی که مقدار تابع برازندگی کمتری دارد، در ادامه قرار می‌گیرند.

۵- تولید جمعیت جدید: در این مقاله برای تولید جمعیت جدید از عملگرهای انتخاب نخبگان، آمیزش، جهش و انتخاب استفاده شده است.

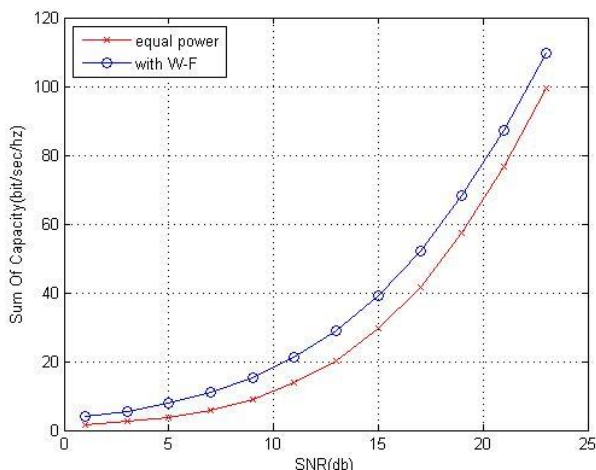
نکته: پس از استفاده از عملگرهای جهش و آمیزش، مجدداً تعداد زیرحامل‌های تخصیص داده شده به هر گروه شمارش می‌گردد و اگر تعداد زیرحامل‌های تخصیص داده شده از تعداد زیرحامل‌ها درخواستی توسط هر گروه کمتر باشد، عملگر مورد نظر مجدداً اجرا می‌گردد.

۶- افزایش شمارنده: به شمارنده یک واحد اضافه می‌گردد ($L=L+1$)

۷- بررسی شرط توقف: اگر مقدار L برابر با L_{Max} نباشد، پرش به مرحله ۳ و ادامه الگوریتم و اگر مقدار L و L_{Max} برابر باشند انتخاب کروموزومی که بیشترین مقدار تابع برازندگی را دارد به عنوان جواب نهایی.

با اجرای الگوریتم توضیح داده شده تمام قیود مورد نظر ارضا می‌گردند.

الگوریتم تخصیص زیرحامل در مرحله اول بهینه نیست، چراکه در آن توزیع یکسان توان اعمال شده است. در این مرحله با توجه به



شکل ۷: مقایسه تخصیص توان با روش واترفیلینگ و تخصیص یکسان توان

همانطور که در شکل (۷) مشخص است، استفاده از روش واترفیلینگ برای تخصیص توان موجب افزایش ظرفیت کلی سیستم می‌گردد.

۵- نتیجه گیری

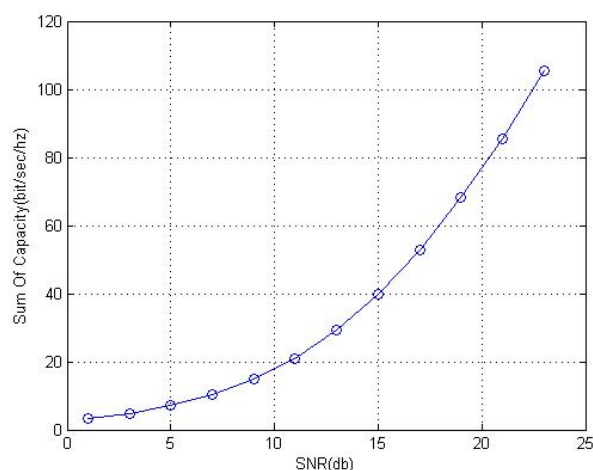
در این مقاله یک روش پیشنهادی در زمینه تخصیص زیرحامل و توان، برای بهبود و افزایش ظرفیت سیستم‌های چندپخش مبتنی بر MIMO-OFDM ارائه شد. در این روش با استفاده از الگوریتم ژنتیک به تخصیص زیرحامل پرداخته شد و با استفاده از روش واترفیلینگ، توان بصورت بهینه تخصیص پیدا کرد، همچنین بحث عدالت در تخصیص منابع نیز در نظر گرفته شد. همانطور که نتایج شبیه سازی نشان داد روش ارائه شده باعث افزایش ظرفیت کلی سیستم گردید.

مراجع

- [1] C. Jie "Mobile TV - a great opportunity for WiMAX," Communicate, no. 41, pp. 34-36, Jun. 2009
- [2] K. Bakano glu, W. Mingquan, L. Hang, and M. Saurabh, "Adaptive resource allocation in multicast OFDMA systems," in Proc. IEEE Wireless Communications and Networking Conference (WCNC'10), Apr., pp. 1-6, 2010
- [3] U. Varshney "Multicast support in mobile commerce applications," Computer, vol. 35, no. 2, pp. 115-117, Feb. 2002.
- [4] J. Liu, W. Chen, Z. Cao, and K. Letaief "Dynamic power and sub-carrier allocation for OFDMA-based wireless multicast systems," in Proc. IEEE International Conference on Communications (ICC'08), 2008.
- [5] Ngo, D. T., Tellambura, C., & Nguyen, H. H. . "Efficient resource allocation for OFDM multicastsystems With spectrum-sharing control". IEEE Transactions on Vehicular Technology, 58(9), 4878-4889. 2009.
- [6] Liu, J., Chen, W., Cao, Z., & Letaief K. B. . "Dynamic power and subcarrier allocation for OFDMA-based wireless multicast systems". In Proceedings of IEEE ICC '08, Beijing, 19-23 May 2008 (pp. 2607-2611). 2008.
- [7] Papoutsis, V. D., Fraimis, I. G., & Kotsopoulos, S. A. . "Fairness-aware resource allocation for the SISO downlink over frequency-selective channels". In Proceedings of IEEE WCNC 2010, NSW, April 18-21 (pp. 1-6). , 2010.

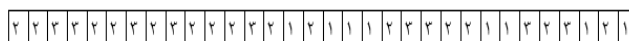
چگالی طیفی توان نویز	1
حداکثر شیفیت داپلر	60Hz
BER	10^{-3}
P_t	1W
زمان نمونه برداری	1e-4s
L_{Max}	150
N_p	100

شکل (۵) ظرفیت کلی سیستم را برای مقادیر داده شده در جدول (۱) نشان می‌دهد.



شکل ۵: ظرفیت کلی سیستم بر حسب SNR با فرض مقادیر داده شده در جدول (۱)

کروموزوم نهایی نیز در شکل (۶) نشان داده شده است و همانطور که مشخص است، ۸ زیرحامل به گروه یک، ۱۵ زیرحامل به گروه دو و ۹ زیرحامل به گروه سه تخصیص پیدا کرده است و این بدان معناست که تعداد زیرحامل در خواستی توسط هر گروه برآورده شده است.



شکل ۶: کروموزوم نهایی

شکل (۷) مقایسه تخصیص بهینه توان با استفاده از روش واترفیلینگ و تخصیص یکسان توان را نشان می‌دهد.

- [8] Xu, J., Lee, S. J., Kang, W. S., & Seo, J. S. . "Adaptive resource allocation algorithm for MIMO-OFDM based wireless multicast systems".*IEEE Transactions on Broadcasting*, 56(1), 98-102.2010.
- [9] Fraimis, Ioannis G., and Stavros A. Kotsopoulos"Fair radio resource allocation for MIMO OFDM-based multicast systems." *Wireless Personal Communications* Springer 1-15.2012.
- [10] D. T. Ngo, C. Tellambura, and H. H. Nguyen, "Efficient resource allocation for OFDMA multicast systems with spectrum-sharing control," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4878-4889, Nov. 2009.
- [11] Nitin Sharma and A. S. Madhukumar"Genetic Algorithm Aided Proportional Fair Resource Allocation in Multicast OFDM Systems",*IEEE Trans On Broadcasting*,2015.
- [12] Chung, S. T., & Goldsmith, A. "Degrees of freedom in adaptive modulation: A unified view.*IEEE Transactions on Communications*, 49(9), 1561-1571,2001.

زیر نویس‌ها

- 1 Quality of Service
- 2 Multiple-Input and Multiple-Output
- 3 Orthogonal Frequency-Division Multiplexing
- 4 Multicasting
- 5 Unicast
- 6 Rate
- 7 Sub-Carrier
- 1 Fairness
- 2 Genetic Algorithm
- 3 Water-Filling
- 4 Base Station
- 5 Channel State Information
- 6 Slow Fading
- 7 Bit Error Rate
- 1 Optimal
- 2 Sub-Optimal
- 3 John L. Holland
- 4 Chromosome
- 5 Gene
- 6 Population
- 7 Fitness Function
- 8 Reproduction
- 9 Generation
- 10 Selection
- 11 Crossover
- 12 Mutation
- 13 Slice
- 14 Elitism
- 15 De Jong
- 1 One-point Crossover
- 2 Two-point Crossover
- 3 Crossover Rate
- 4 Mutation Rate



شناسایی نواحی برجسته تصویر با تلفیق ویژگی‌های رنگ، بافت و موقعیت مکانی پیکسل‌ها

سکینه اسدی امیری^۱

^۱ استادیار، دانشکده فنی و مهندسی، گروه کامپیوتر، دانشگاه مازندران، بابلسر

s.asadi@umz.ac.ir

چکیده

تمام نواحی تصویر به یک اندازه، توجه بیننده را جلب نمی‌کنند، بلکه برخی از نواحی تصویر چشم‌گیرتر بوده و بیشتر مورد توجه مخاطب قرار می‌گیرد. به این نواحی، نواحی چشم‌گیر، نواحی برجسته یا نواحی مورد علاقه می‌گویند. در این مقاله، یک روش جدید برای شناسایی نواحی برجسته تصویر ارائه شده است. اغلب ناحیه برجسته از نظر نوع بافت و رنگ با سایر نواحی تصویر متفاوت است. علاوه بر این، نواحی میانی تصویر بیشتر مورد توجه مخاطبان قرار می‌گیرد. از این رو، در روش پیشنهادی برای شناسایی نواحی برجسته تصویر از ویژگی‌های سطح پایین همچون ویژگی بافت و رنگ و ویژگی سطح بالا همچون موقعیت مکانی پیکسل‌ها استفاده شده است. در این روش، برای تمایز نوع بافت از ویژگی موجک و برای تمایز رنگ از فیلتر Log-Gabor استفاده شده است. روش پیشنهادی روی پایگاه داده MSRA-1000 ارزیابی شده و نتایج آن با چند مقاله اخیر مقایسه شده است. نتایج نشان‌دهنده برتری روش پیشنهادی نسبت به مقالات اخیر است.

کلمات کلیدی

نواحی برجسته تصویر، ویژگی رنگ، بافت تصویر، سیستم بینایی انسان.

واقع ضریب برجستگی، وزن تأثیرگذاری هر پیکسل را در کیفیت تصویر مدل می‌کند. هر چه ضریب برجستگی یک پیکسل بیشتر باشد، خرابی آن پیکسل تأثیر بیشتری بر افت کیفیت تصویر از نظر سیستم بینایی انسان خواهد داشت. شکل ۱ یک مثال از نقشه برجستگی تصویر را نشان می‌دهد. نواحی روشن‌تر در نقشه برجستگی بیانگر نواحی مورد توجه تصویر در سیستم بینایی انسان است.



ب) نقشه برجستگی



الف) تصویر ورودی

شکل ۱: مثالی از نقشه برجستگی تصویر.

۱- مقدمه

برخی از نواحی تصویر بیشتر مورد توجه مخاطب قرار می‌گیرد که به آن نواحی برجسته تصویر می‌گویند. میزان جذب نواحی تصویر به این بستگی دارد که چطور از محیط اطرافش متمایز است. این تفاوت می‌تواند توسط ویژگی‌های سطح پایین و ویژگی‌های سطح بالا مشخص شود [۱]. ویژگی‌های رنگ، روشنایی و بافت نمونه‌هایی از ویژگی‌های سطح پایین، و ویژگی‌های موقعیت مکانی پیکسل‌ها و عمق نمونه‌هایی از ویژگی‌های سطح بالا هستند. هر چه تفاوت یک شی از محیط اطرافش بیشتر باشد، آن شی در تصویر توجه انسان را بیشتر به خود جلب می‌کند. برای به دست آوردن نواحی مورد توجه در تصویر می‌توان از نقشه برجستگی تصویر استفاده کرد.

نقشه برجستگی، تصویری در اندازه تصویر ورودی است که ضریب برجستگی هر پیکسل از تصویر ورودی را نشان می‌دهد. در

نواحی فرکانس میانی تصویر اهمیت بیشتری نسبت به سایر فرکانس‌های تصویر دارد. در این روش از فیلتر Log-Gabor [۱۳] برای استخراج فرکانس‌های میانی تصویر استفاده شده است. این فیلتر میان‌گذر، یک دنباله وسیع^۲ در فرکانس‌های بالا دارد که قابلیت خوبی در کد کردن تصاویر طبیعی نسبت به سایر فیلترهای میان‌گذر دارد [۱۳] و [۱۴]. معیار دوم این است که نواحی مرکزی تصویر بیشتر مورد توجه سیستم بینایی انسان قرار می‌گیرند. معیار سوم حاکی از این است که رنگ‌های گرم (مانند قرمز و زرد) نسبت به رنگ‌های سرد (مانند سبز و آبی) برای انسان جذاب‌تر هستند [۱۵]. این روش بر روی سه کانال رنگ Lab انجام می‌شود و سه نقشه برجستگی برای تصویر ورودی به دست می‌آید. در پایان، با میانگین‌گیری از سه نقشه برجستگی حاصل، یک نقشه برجستگی نهایی برای تصویر ورودی به دست می‌آید.

در [۱۶] پیکسل‌های تصویر با استفاده از k -means به تعدادی خوشه تقسیم می‌شوند و میانگین سطح خاکستری پیکسل‌های هر خوشه، به عنوان مقدار تمام پیکسل‌های موجود در آن خوشه لحاظ می‌شود. در واقع در این روش، با تقسیم تصویر به تعداد زیادی خوشه، جزئیات کم اهمیت تصویر حذف می‌شوند و تنها ساختار اصلی تصویر حفظ می‌گردد. در این روش، از تمایز رنگ خوشه، برای تعیین نواحی برجسته تصویر استفاده شده است. در واقع نواحی برجسته تصویر، در یک ناحیه متمرکز است و برخلاف نواحی پس‌زمینه در کل تصویر پخش نمی‌باشد.

در [۱۷] از دو مدل رنگ Lab و RGB برای شناسایی نواحی برجسته تصویر استفاده شده است. روش مطرح شده در این مقاله بر روی هر یک از کانال‌های رنگ به صورت مجزا انجام می‌شود. ابتدا هر کانال رنگ به بلوک‌های کوچک غیرهمپوشان تقسیم می‌شود. سپس عدم شباهت هر بلوک با بلوک‌های همسایه‌اش محاسبه می‌شود و این مقدار جایگزین کل پیکسل‌های بلوک می‌شود. بدین گونه برای هر کانال رنگ، یک تصویر بر اساس میزان عدم شباهت هر بلوک نسبت به بلوک‌های همسایه‌اش به دست می‌آید. همچنین عدم شباهت هر بلوک نسبت به تمامی بلوک‌های موجود در آن کانال رنگ نیز محاسبه می‌شود. از این‌رو برای هر کانال رنگ، دو تصویر ایجاد می‌شود که با ترکیب این تصاویر، نقشه برجستگی نهایی به دست می‌آید. این روش زمان محاسبات زیادی نیاز دارد.

در [۱۸] با استفاده از ویژگی‌های استخراج شده از تبدیل موجک تصویر، نواحی برجسته تصویر شناسایی می‌شوند. در این روش، ابتدا تصویر با فیلتر گوسی هموار می‌شود تا جزئیات کم اهمیت تصویر حذف شود. سپس تبدیل موجک بر روی تصویر اعمال می‌شود و زیرباند تقریب آن سطح موجک حذف می‌گردد، سپس با فرآیند معکوس تبدیل موجک، تصویر در حوزه مکان به دست می‌آید. تصویر حاصل بیانگر ویژگی‌های لبه و بافت تصویر است. فرآیند تبدیل موجک تا نه سطح بر روی تصویر انجام می‌شود تا ویژگی‌های لبه و بافت تصویر در مقیاس‌های مختلف به دست آیند. سرانجام این نه تصویر حاصل که هر یک بیانگر ویژگی‌های لبه و بافت تصویر در مقیاس‌های مختلف هستند، با هم ترکیب شده تا یک تصویر ویژگی به دست آید.

استخراج نقشه برجستگی تصویر در کاربردهای زیادی همچون شناسایی اشیا [۲-۳]، معیار ارزیابی کیفیت تصویر [۴-۵]، فشرده‌سازی تصویر [۶-۷] و قطعه‌بندی تصویر [۸-۹] مفید است.

اگر چه مقالات زیادی برای شناسایی نواحی برجسته تصویر مطرح شده‌اند، ولی این موضوع همچنان به عنوان یک موضوع فعال است. برخی از روش‌ها با استفاده از یک یا چند ویژگی سطح پایین نقشه برجستگی تصویر را استخراج نمودند. در مقابل روش‌هایی مطرح شدند که با ترکیب ویژگی‌های سطح پایین و سطح بالا عملکرد بهتری در شناسایی نقشه برجستگی تصویر داشتند.

در این مقاله، یک روش جدید برای شناسایی نواحی برجسته تصویر ارائه شده است. همان‌طور که اشاره شد، نواحی برجسته نسبت به سایر نواحی تصویر بر اساس ویژگی‌هایی متمایز است. در روش پیشنهادی از سه ویژگی بافت، رنگ و موقعیت مکانی برای شناسایی نواحی برجسته تصویر استفاده شده است. در واقع ناحیه برجسته از نظر نوع بافت و رنگ ممکن است با سایر نواحی تصویر متفاوت باشد. در روش پیشنهادی برای تمایز نوع بافت از ویژگی موجک استفاده شده است. برای تمایز رنگ نیز از فیلتر Log-Gabor استفاده شده است. علاوه بر این دو ویژگی، ناحیه برجسته از نظر موقعیت مکانی نیز بررسی شده است.

در بخش بعدی، کارهای پیشین در زمینه شناسایی نواحی برجسته تصویر مرور می‌شود. در بخش سوم، روش پیشنهادی برای شناسایی نواحی برجسته تصویر ارائه شده است. نتایج این مقاله در بخش چهارم و نتیجه‌گیری در بخش پنجم مطرح شده‌اند.

۲- مروری بر کارهای پیشین

در ادامه چند مقاله اخیر در رابطه با شناسایی نواحی برجسته تصویر مرور می‌شود. در [۱۰] با استفاده از دامنه تبدیل فوریه، نواحی برجسته تصویر شناسایی می‌شوند. دامنه تبدیل فوریه هر بلوک از تصویر، بیانگر اطلاعات رنگ و بافت آن بلوک از تصویر است. در این روش، ابتدا تصویر به بلوک‌های کوچکی تقسیم می‌شود، سپس دامنه تبدیل فوریه هر بلوک محاسبه می‌گردد و اختلاف مقادیر دامنه تبدیل فوریه هر بلوک با سایر بلوک‌های تصویر محاسبه می‌شود، و مقدار حاصل در بلوک مربوطه از نقشه برجستگی درج می‌شود. هر چه میزان اختلاف، بزرگتر باشد، بیانگر کمیاب بودن آن بلوک نسبت به سایر بلوک‌های تصویر است. قابل ذکر است، در این روش یک وزنی به مقدار تفاضل هر بلوک نسبت به سایر بلوک‌ها اختصاص می‌یابد که مقدار این وزن با توجه به سیستم بینایی انسان به دست می‌آید. این روش علاوه بر این که به زمان محاسبات زیادی نیاز دارد، از دقت خوبی نیز برخوردار نمی‌باشد.

در [۱۱] از فیلتر میان‌گذر گوسی برای استخراج نواحی برجسته تصویر استفاده شده است. زیرا فرکانس‌های میانی تصویر اهمیت بیشتری نسبت به سایر فرکانس‌های تصویر برای سیستم بینایی انسان دارند. در [۱۲] برای شناسایی نواحی برجسته تصویر از سه معیار استفاده شده است. معیار اول این است که برای سیستم بینایی انسان،

² Extended tail

¹ Rarity

- شناسایی ناحیه برجسته با ترکیب ویژگی‌ها
- افزایش اندازه نقشه برجستگی برای رسیدن به اندازه تصویر اولیه
- هموارسازی نقشه برجستگی

۳-۱- تبدیل تصویر از فضای رنگ RGB به Lab

همان‌طور که اشاره شد، ناحیه برجسته نسبت به سایر نواحی تصویر در معیارهایی همچون رنگ، بافت و عمق متفاوت است. برای تشخیص این تمایز از مدل‌های رنگی استفاده می‌شود که به سیستم بینایی انسان نزدیک‌تر باشد. در روش پیشنهادی از مدل رنگ Lab استفاده شده است. همان‌طور که اشاره شد، نقشه برجستگی، تصویری در اندازه تصویر ورودی است. برای افزایش سرعت روش پیشنهادی، اندازه تصویر ورودی را کاهش داده و نقشه برجستگی برای تصویر حاصل محاسبه می‌شود. برای نمونه، تصاویر در اندازه 400×300 به اندازه 256×256 تبدیل شدند.

۳-۲- نقشه برجستگی مبتنی بر بافت

همان‌طور که اشاره شد، ناحیه برجسته نسبت به سایر نواحی تصویر در نوع بافت ممکن است متفاوت باشد. در روش پیشنهادی برای تعیین ویژگی‌های بافت تصویر از تبدیل موجک استفاده شده است. ابتدا فیلتر وینر با اندازه پنجره 3×3 به تصویر ورودی اعمال می‌شود تا جزئیات کم اهمیت تصویر حذف شود. برای تعیین ویژگی‌های بافت مبتنی بر موجک از روش مشابه ویژگی‌های محلی مطرح شده در [۱۸] استفاده شده است. در این روش با استفاده از سطوح مختلف موجک، ویژگی‌های بافت تصویر استخراج می‌شود. ابتدا تصویر با یک سطح تبدیل موجک تجزیه می‌شود و زیرباند تقریب این سطح حذف می‌شود، با معکوس تبدیل موجک از سه زیرباند باقیمانده، ویژگی‌های بافت تصویر در سطح یک به دست می‌آید. به طور کلی، برای تعیین ویژگی بافت در سطح l از تبدیل موجک، تصویر تا l سطح با تبدیل موجک تجزیه می‌شود و زیرباند تقریب بالاترین سطح، حذف می‌شود و با معکوس تبدیل موجک از این سطوح، ویژگی‌های بافت تصویر در آن سطح استخراج می‌شود. در این روش، فرآیند استخراج ویژگی‌های بافت تا نه سطح موجک انجام شده است. بدین گونه نه تصویر ویژگی مرتبط با بافت برای هر کانال رنگ به دست می‌آید که با مجموع نه تصویر، یک تصویر ویژگی برای هر کانال دست می‌آید. با جمع این سه تصویر نیز یک تصویر ویژگی بافت مبتنی بر موجک برای کل تصویر به دست می‌آید. شکل ۲ ویژگی بافت مبتنی بر تبدیل موجک را برای یک تصویر نمونه نشان می‌دهد.



الف) تصویر ورودی

ب) ویژگی بافت

شکل ۲: نقشه برجستگی مبتنی بر بافت.

این تصویر بیانگر ویژگی‌های محلی بافت و لبه تصویر است. قابل ذکر است، روش پیشنهادی در این مقاله برای هر سه کانال در مدل رنگ Lab به صورت مجزا انجام می‌شود. در این مقاله علاوه بر ویژگی محلی مبتنی بر موجک، با استفاده از اطلاعات آماری توزیع ویژگی‌های محلی (با استفاده از نه تصویر ویژگی مذکور)، نقشه برجستگی سراسری نیز برای تصویر محاسبه می‌شود. در پایان با ضرب دو تصویر ویژگی‌های سراسری و محلی، نقشه برجستگی نهایی برای تصویر به دست می‌آید.

در [۱۹] از سه مدل رنگ RGB، YCbCr و Lab برای شناسایی نواحی برجسته تصویر استفاده شده است. در این روش، ابتدا تصویر به بلوک‌های کوچکی تقسیم می‌شود، سپس PCA به هر بلوک از تصویر اعمال می‌شود تا ویژگی‌های موثر برای هر بلوک به دست آید. در این مقاله از اختلاف یک بلوک نسبت به بلوک‌های همسایه آن، برای شناسایی نواحی برجسته تصویر استفاده شده است. در این روش، از تابع گوسی برای ضریب وزنی هر بلوک همسایه استفاده شده است. یعنی بلوک‌های همسایه نزدیک‌تر از اهمیت بیشتری برخوردار هستند. همچنین در این روش، اختلاف هر بلوک نسبت به سایر بلوک‌های تصویر نیز محاسبه شده است تا تفاوت هر بلوک نسبت به کل تصویر نیز مورد سنجش قرار گیرد. سرانجام از ضرب دو تصویر اختلاف سراسری و محلی، نقشه برجستگی نهایی برای تصویر به دست می‌آید. روش [۲۰] از هر دو ویژگی اختلاف رنگ و بافت برای شناسایی نواحی برجسته تصویر استفاده نموده است. در این روش ابتدا تصویر را با استفاده از روش خوشه‌بندی مطرح شده در [۱۶] به تعداد زیادی خوشه تقسیم می‌کند تا جزئیات کم اهمیت تصویر حذف شوند و تنها ساختار اصلی تصویر حفظ شود. خوشه‌بندی مطرح شده در [۱۶]، نسخه بهبودیافته خوشه‌بندی k-means است. پس از بخش‌بندی تصویر، برای شناسایی نواحی برجسته با توجه به تمایز رنگ، از روش مطرح شده در [۱۶] استفاده شده است. برای استخراج ویژگی بافت نیز از الگوی دودویی محلی (lbp) استفاده شده است.

در [۲۰] برای ترکیب دو ویژگی رنگ و بافت از نگاشت عمقی تصویر استفاده شده است. نگاشت عمقی، تصویری در اندازه تصویر ورودی است که بیانگر نقطه تمرکز تصویربردار در حین اخذ تصویر است. در این مقاله، با تخمین میزان تاری پیکسل‌های تصویر، نگاشت عمقی تصویر با یک تابع گوسی مدل شده است.

۳- روش پیشنهادی

در این مقاله، یک روش جدید برای شناسایی نواحی برجسته تصویر با تلفیق دو ویژگی سطح پایین رنگ و بافت و ویژگی سطح بالای موقعیت مکانی پیکسل‌ها ارائه شده است. مراحل روش پیشنهادی به صورت خلاصه در زیر بیان شده است:

- کاهش اندازه تصویر
- تبدیل تصویر از فضای رنگ RGB به Lab
- شناسایی ناحیه برجسته با توجه به ویژگی بافت
- شناسایی ناحیه برجسته با توجه به ویژگی رنگ
- شناسایی ناحیه برجسته با توجه به موقعیت مکانی پیکسل‌ها

۳-۳- نقشه برجستگی مبتنی بر رنگ

علاوه بر ویژگی بافت، ممکن است ناحیه برجسته از نظر رنگ نیز نسبت به سایر نواحی تصویر متمایز باشد. قابل ذکر است چشم انسان به فرکانس‌های میانی تصویر حساسیت بیشتری نسبت به سایر فرکانس‌های تصویر دارد. از این‌رو فرکانس‌های میانی هر یک از کانال‌های رنگ با یک فیلتر میان‌گذر، استخراج می‌شوند. برای استخراج فرکانس‌های میانی تصویر از فیلتر Log-Gabor [۱۲] استفاده شده است. سرانجام نقشه برجستگی مبتنی بر رنگ، از مجذور حاصل از مجموع مربع این سه تصویر به دست می‌آید [۱۲]. شکل ۳، نمونه‌ای از نقشه برجستگی مبتنی بر رنگ را نشان می‌دهد.

۳-۴- نقشه برجستگی مبتنی بر موقعیت مکانی پیکسل‌ها

اغلب نواحی برجسته در نواحی مرکزی تصویر قرار دارند. این فرض را می‌توان با یک تابع گوسی در مرکز تصویر مطابق رابطه زیر مدل کرد [۱۲]:

$$S = \exp\left(-\frac{\|x - c\|_2^2}{\sigma^2}\right) \quad (1)$$

در رابطه بالا، x مختصات هر پیکسل از تصویر، c مرکز تابع گوسی و σ^2 واریانس تابع گوسی است. نگاهت گوسی حاصل از رابطه (۱) برای تصویری در اندازه 256×256 با مختصات مرکز تصویر و با واریانس ۱۱۰ در شکل ۴ الف نشان داده شده است. نواحی روشن در این نگاهت بیانگر تخمینی از مکان ناحیه برجسته در تصویر است، در حالی که نواحی تیره در این نگاهت احتمال اندکی را برای وجود ناحیه برجسته در این نواحی از تصویر برآورد می‌کند.



ب) ویژگی رنگ

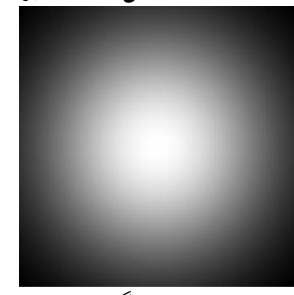


الف) تصویر ورودی

شکل ۳: نقشه برجستگی مبتنی بر رنگ.



ب) نقشه برجستگی نهایی



الف) نقشه برجستگی مبتنی بر موقعیت مکانی

شکل ۴: نقشه برجستگی مبتنی بر موقعیت مکانی پیکسل‌ها و نقشه برجستگی نهایی.

۳-۵- نقشه برجستگی با ترکیب ویژگی‌ها

تاکنون سه ماتریس نقشه برجستگی برای تصویر برآورد شده است. یک نقشه برجستگی براساس تمایز رنگ (G)، نقشه برجستگی دوم براساس تمایز بافت (T) و نقشه برجستگی سوم براساس تخمین مکان ناحیه برجسته با تابع گوسی (G) است. با استفاده از این سه تصویر، نقشه برجستگی نهایی براساس رابطه زیر برای تصویر به دست می‌آید:

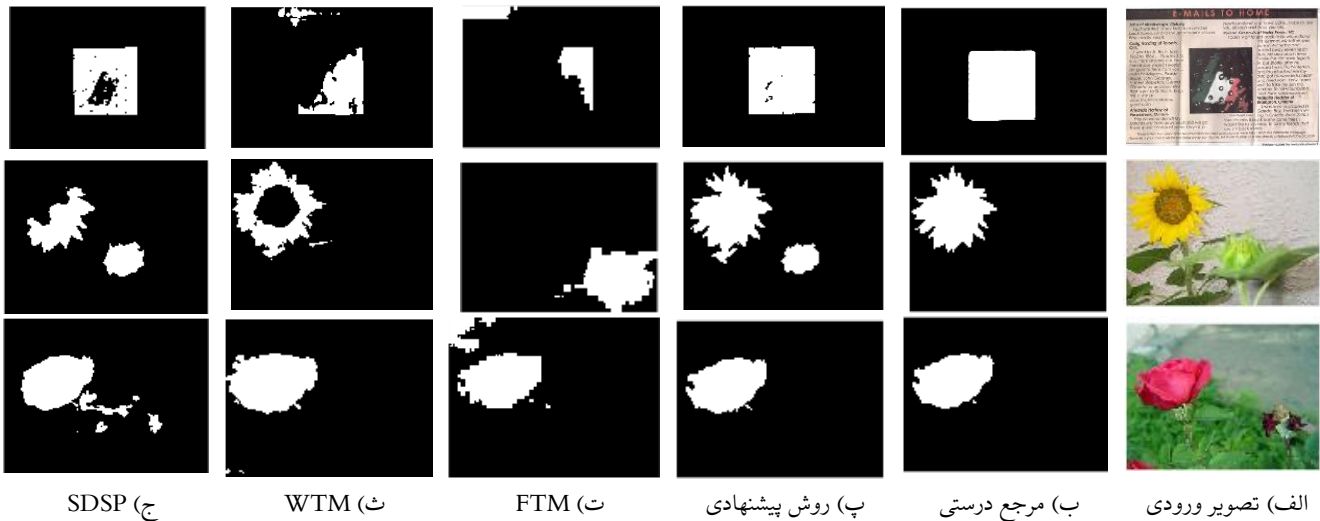
$$Sal = C \times T \times G \quad (2)$$

همان‌طور که اشاره شد، برای بهبود سرعت روش پیشنهادی، در ابتدا اندازه تصویر ورودی کاهش یافت و نقشه برجستگی برای این تصویر حاصل محاسبه شد. در این مرحله، اندازه نقشه برجستگی به اندازه تصویر ورودی افزایش می‌یابد تا نواحی برجسته متناظر با هر ناحیه از تصویر ورودی تعیین گردد. برای حذف نویز و هموار نمودن تصویر نیز از فیلتر وینر در اندازه پنجره 3×3 در این مرحله استفاده شده است. شکل ۴ ب) نقشه برجستگی نهایی را نشان می‌دهد.

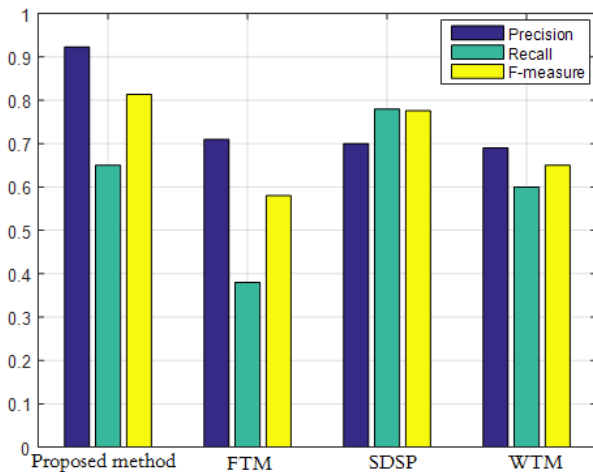
برای تبدیل نقشه برجستگی خاکستری به تصویر دودویی از مقدار آستانه‌ای استفاده می‌شود. اغلب مقدار آستانه را دو برابر مقدار میانگین نقشه برجستگی در نظر می‌گیرند. در نقشه برجستگی دودویی ممکن است، برخی از پیکسل‌ها به اشتباه جز ناحیه برجسته یا پس‌زمینه شناسایی شوند. در روش پیشنهادی، از مفهوم نواحی متصل برای حذف این نواحی اشتباه استفاده شده است. در تصویر دودویی، پیکسل‌های همسایه با مقادیر یکسان، یک ناحیه متصل را تشکیل می‌دهند. برای در نظر گرفتن همسایه‌های مجاور یک پیکسل، از هشت همسایه آن پیکسل استفاده شده است. بدین طریق، تمامی پیکسل‌های نقشه برجستگی دودویی در چند ناحیه قرار می‌گیرند. برخی از این ناحیه‌ها، نویز هستند. برای شناسایی نواحی نویز، نیاز است تعداد پیکسل‌های موجود در هر ناحیه شمارش شود. ناحیه‌هایی که تعداد پیکسل‌های آن‌ها از مقدار آستانه‌ای کمتر باشد، نویز تلقی می‌شوند، و رنگ این پیکسل‌ها معکوس می‌شود. در روش پیشنهادی، مقدار آستانه وابسته به اندازه تصویر (تعداد پیکسل‌های تصویر)، به مقدار $\frac{1}{45}$ تعداد کل پیکسل‌های تصویر در نظر گرفته شده است. این مقدار آستانه، به صورت تجربی به دست آمده است.

۴- نتایج روش پیشنهادی

برای مقایسه روش پیشنهادی با سایر روش‌های مطرح شده برای نقشه برجستگی تصویر از پایگاه‌داده MSRA استفاده شده است. این پایگاه‌داده شامل ۱۰۰۰ تصویر رنگی در اندازه 400×300 (یا 300×400) و ۱۰۰۰ تصویر دودویی مرجع درستی است. این تصاویر با استفاده از نظرات نه‌کاربر به دست آمده است. مرجع درستی، یک تصویر دودویی است که نواحی سفید آن، بیانگر نواحی برجسته تصویر است. روش پیشنهادی با روش‌های WTM [۱۸]، FTM [۱۰] و SDSP [۱۲] بر روی تصاویر موجود در پایگاه‌داده MSRA مقایسه شده است. در شکل ۵ نتیجه مقایسه کیفی روش پیشنهادی با مراجع مذکور بر روی چند تصویر از پایگاه‌داده MSRA نشان داده شده است.



شکل ۵: مقایسه کیفی روش پیشنهادی با چند روش متداول در شناسایی نواحی برجسته تصویر.



شکل ۶: مقایسه کمی روش پیشنهادی با چند روش متداول در شناسایی نواحی برجسته تصویر.

جدول ۱: مقایسه زمانی روش پیشنهادی با چند روش متداول در شناسایی نواحی برجسته تصویر (زمان برحسب ثانیه است).

روش پیشنهادی	FTM	SDSP	WTM
۳۷۰۵۱۰	۱۴۱۱۸۲	۰٫۵۳۳۳	۴٫۸۰۳۰

۵- نتیجه

در این مقاله، روش جدیدی برای شناسایی نواحی برجسته تصویر با تلفیق ویژگی‌های سطح پایین رنگ و بافت و ویژگی‌های سطح بالای موقعیت مکانی پیکسل‌ها پیشنهاد شده است. اگر چه هر یک از ویژگی‌های سطح بالا و ویژگی‌های سطح پایین به تنهایی قادر به شناسایی نواحی برجسته تصویر هستند، تلفیق این دو نوع ویژگی منجر به بهبود قابل توجهی در شناسایی نواحی برجسته تصویر می‌شود. روش پیشنهادی بر روی ۱۰۰۰ تصویر از پایگاه داده MSRA مورد ارزیابی قرار گرفت. نتایج کیفی و کمی نشان‌دهنده برتری روش پیشنهادی نسبت به روش‌های شناسایی نقشه برجستگی متداول است.

همان‌طور که از نتایج مشخص است روش پیشنهادی با دقت خوبی نواحی برجسته تصویر را استخراج نموده است، و نقشه برجستگی حاصل، به تصویر مرجع درستی نزدیک است. علاوه بر مقایسه کیفی روش پیشنهادی با سایر مراجع، از معیارهای کمی دقت^۱، فراخوانی^۲، معیار- F^3 نیز برای مقایسه استفاده شده است. این سه معیار به ترتیب در روابط زیر معرفی شده‌اند:

$$P = \frac{\sum_{x=1}^M \sum_{y=1}^N (t(x,y) \times s(x,y))}{\sum_{x=1}^M \sum_{y=1}^N s(x,y)} \quad (3)$$

$$R = \frac{\sum_{x=1}^M \sum_{y=1}^N (t(x,y) \times s(x,y))}{\sum_{x=1}^M \sum_{y=1}^N t(x,y)} \quad (4)$$

$$F = \frac{(1 + \alpha) \times P \times R}{\alpha \times P + R} \quad (5)$$

در روابط بالا، t تصویر مرجع درستی، s نقشه برجستگی دودویی، M و N اندازه سطر و ستون تصویر است. پارامتر α اهمیت دقت را به فراخوانی تعیین می‌کند. در اکثر مقالات مقدار این پارامتر ۰٫۳ در نظر گرفته شده است. مقایسه روش پیشنهادی با روش‌های مذکور توسط سه معیار دقت، فراخوانی و معیار-F در شکل ۶ نشان داده شده است. همان‌طور که از نتایج مشخص است، معیارهای کمی نیز حاکی از برتری روش پیشنهادی نسبت به سایر مراجع است. البته مقدار معیار فراخوانی در روش SDSP از روش پیشنهادی ما بیشتر است، در حالی که مقدار معیار دقت و معیار-F برای روش پیشنهادی ما بیشتر از روش SDSP است.

روش پیشنهادی به همراه سه روش FTM، SDSP و WTM به کمک کامپیوتر پنتیوم با پردازنده ۳ مگاهرتز هفت هسته‌ای و RAM ۸ گیگا بایت با نرم افزار Matlab ۲۰۱۴ اجرا شده‌اند. نتایج مقایسه زمانی برحسب ثانیه در جدول ۱ درج شده است. این نتایج، با توجه به میانگین زمانی هر یک از روش‌ها بر روی ۱۰۰۰ تصویر پایگاه‌داده MSRA به دست آمده است. همان‌طور که از نتایج مشخص است، زمان اجرای روش پیشنهادی کمتر از دو روش FTM و WTM است. ولی زمان اجرای روش SDSP، کمتر از روش پیشنهادی ما است، در حالی که دقت روش ما برتر از این روش است.

مراجع

international conference on image processing, pp. 171-175, 2013.

[13] Field D. J., "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America A*, Vol. 4, pp. 2379-2394, 1987.

[14] Kovese P., "Image features from phase congruency", *Videre: Journal of Computer Vision Research*, Vol. 1, pp. 1-26, 1999.

[15] Chen X., Wu Y., "A unified approach to salient object detection via low rank matrix recovery", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 853860, 2012.

[16] Perazzi F., Krahenbuhl P., Pritch Y., Hornung A., "Saliency filters: Contrast based filtering for salient region detection", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 733-740, 2012.

[17] Borji A., Itti L., "Exploiting Local and Global Patch Rarities for Saliency Detection", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 478-485, 2012.

[18] Imamoglu N., Lin W., Fang Y., "A saliency detection model using low level features based on wavelet transform", *IEEE Transactions Multimedia*, Vol. 15, No. 1, pp. 96-105, 2013.

[19] Zhang Y. Y., Liu X. Y., Wang H. J., "Saliency detection via two-directional 2DPCA analysis of image patches", *Optik*, Vol. 125, pp. 7222-7226, 2014.

[20] Tian H., Fang Y., Zhao Y., Lin W., Ni R., Zhu Zh., "Salient Region Detection by Fusing Bottom-Up and Top-Down Features Extracted From a Single Image", *IEEE transactions on image processing*, Vol. 23, No. 10, pp. 4389-4398, 2014.

زیر نویس‌ها

¹ Precision

² Recall

³ F-measure

[1] Imamoglu N., Lin W., Fang Y., "A saliency detection model using low level features based on wavelet transform", *IEEE Transactions on Multimedia*, Vol. 15, No. 1, pp. 96-105, 2013.

[2] Zhang J., Sclaroff S., Lin Zh., Shen X., Price B., Mech R., "Unconstrained Salient Object Detection via Proposal Subset Optimization", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5733-5742, 2016.

[3] Walther D., Koch Ch., "Modeling attention to salient proto-objects. *Neural Networks*", Vol. 19, No. 9, pp. 1395-1407, 2006.

[4] Zhang L., Shen Y., Li H., "VSI: A Visual Saliency-Induced Index for Perceptual Image Quality Assessment", *IEEE Transactions on Image Processing*, Vol. 23, No. 10, pp. 4270-4281, 2014.

[5] Zhang W., Borji A., Wang Zh., Le Callet P., Liu H., "The Application of Visual Saliency Models in Objective Image Quality Assessment: A Statistical Evaluation", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 27, No. 6, pp. 1266-1278, 2016.

[6] Guo Ch., Zhang L., "A Novel Multiresolution Spatiotemporal Saliency Detection Model and Its Applications in Image and Video Compression", *IEEE Transactions on Image Processing*, Vol. 19, No. 1, pp. 185-198, 2010.

[7] Yu S. X., Lysin D. A., "Image Compression Based on Visual Saliency at Individual Scales", *International Symposium on Visual Computing*, pp. 157-166, 2009.

[8] Li Z., Liu G., Zhang D., Xu Y., "Robust single-object image segmentation based on salient transition region", *Pattern Recognition*, Vol. 52, pp. 317-331, 2016.

[9] Zou W., Liu Z., Kpalma K., Ronsin J., Zhao Y., Komodakis N., "Unsupervised Joint Salient Region Detection and Object Segmentation", *IEEE Transactions on Image Processing*, Vol. 24, No. 11, pp. 3858-3873, 2015.

[10] Fang Y., Lin W., Lee B.S., Lau Ch. T., Chen Zh., Lin Ch. W., "Bottom-Up Saliency Detection Model Based on Human Visual Sensitivity and Amplitude Spectrum", *IEEE Transactions on Multimedia*, Vol. 14, No. 1, pp. 187-194, 2012.

[11] Achanta R., Hemami S., Estrada F., Susstrunk S., "Frequency-tuned salient region detection," *CVPR'09*, pp.1597-1604, 2009.

[12] Zhang L., Gu Zh., Li H., "SDSP: a novel saliency detection method by combining simple priors", *IEEE*



ردیابی بلادرنگ هدف با استفاده از پردازش چند هسته‌ای و دو ویژگی مبتنی بر رنگ و مبتنی بر بافت

مهدی مقیمی^۱، حسین خسروی^۲

^۱ دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود
MehdiMoghimi@Chmail.ir

^۲ دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود
HosseinKhosravi@shahroodut.ac.ir

چکیده

در این مقاله روشی بر پایه الگوریتم مونت کارلو و بهره‌گیری از پردازش چند هسته‌ای برای ردیابی بلادرنگ هدف در دنباله‌ای از تصاویر (ویدئو) ارائه می‌شود. در این روش ابتدا دو گروه ویژگی که شامل ویژگی مبتنی بر رنگ و ویژگی مبتنی بر شکل-بافت هستند از هدف استخراج می‌شود، سپس با استفاده از ویژگی مبتنی بر رنگ به علت سرعت پردازش بالا به ردیابی شیء می‌پردازیم. در صورت از دست رفتن هدف، با برگشت به فریم ماقبل، و استفاده از ویژگی شکل-بافت، که دقت بهتری را در تشخیص هدف ارائه می‌دهد، به ردیابی هدف می‌پردازیم؛ پس از اطمینان از یافتن شیء، مجدداً با ویژگی مبتنی بر رنگ ردیابی هدف را ادامه می‌دهیم؛ به منظور بهبود سرعت الگوریتم، پردازش را روی چند هسته‌ی واحد پردازنده مرکزی توزیع می‌کنیم. نتایج بدست آمده، نشان می‌دهد روش پیشنهادی در بسیاری از چالش‌ها مانند مواجهه با شیء هم رنگ هدف نسبت به فیلتر ذره‌ای، که بر اساس مونت کارلو است، موفق‌تر عمل می‌کند؛ همچنین از لحاظ سرعت پردازش، به عنوان نمونه برای حالتی که از بیست هزار ذره برای ردیابی استفاده شده باشد، تا ۶۴ برابر سریع‌تر از فیلتر ذره‌ای عمل می‌کند.

کلمات کلیدی

ردیابی بلادرنگ، مونت کارلو، پردازش چند هسته‌ای، فیلتر ذره‌ای، انتخاب ویژگی

۱- مقدمه

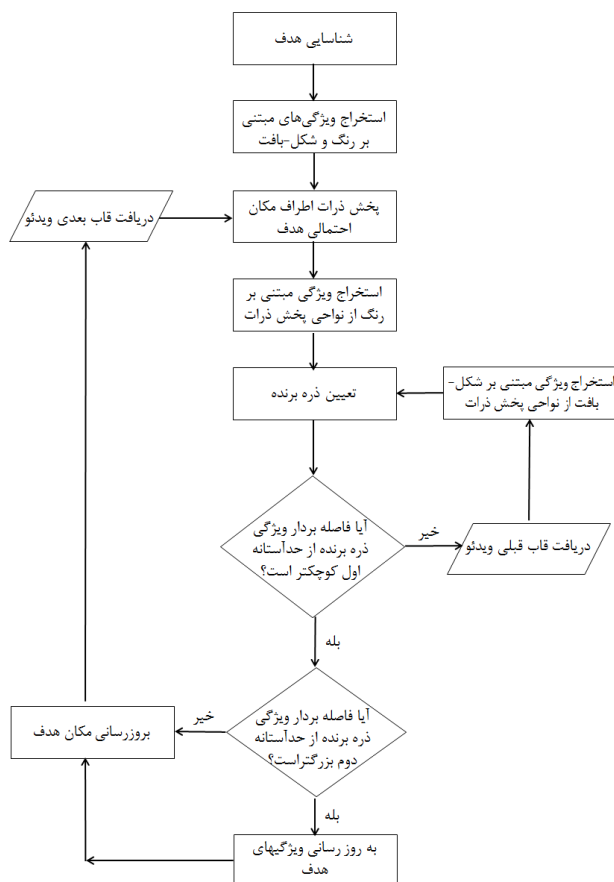
منظور از ردیابی اشیاء، نمایش تغییرات موقعیت یک شیء و دنبال کردن آن در یک دنباله از تصاویر ویدئویی است. تشخیص حرکت یکی از مسائل مهم در زمینه پژوهش‌ها و برنامه‌های مرتبط با حوزه بینایی ماشین است که کاربردهای فراوانی دارد. اگرچه سابقه موضوع ردیابی اشیاء به مسائل نظامی برمی‌گردد ولی امروزه به دلیل کاربردهای بسیار گسترده آن در زمینه‌های مختلفی مانند کنترل ترافیک، تشخیص حرکات غیر معمول، سیستم‌های نظارتی، رباتیک و... این موضوع مورد توجه ویژه‌ای قرار گرفته است [1].

از جمله مسائلی که همواره عملکرد فرآیندهای ردیابی را با مشکل مواجه ساخته است، تعامل آن با روش‌های تشخیص هدف، ظاهر متغیر اهداف و همچنین ردیابی همزمان چندین هدف است [2]. بنابراین برای دستیابی به نتیجه‌ای مطلوب و برطرف ساختن مسائل مطرح شده نیازمند استفاده از فرآیندهایی با دقت بالا هستیم که عمدتاً زمان بر می‌باشند. این در حالی است که اکثر سیستم‌های کاربردی در این حوزه نیازمند پردازش‌های بلادرنگ هستند. باید توجه داشت که انجام عملیات‌های پردازشی پیچیده در پردازش‌های بلادرنگ، صرفاً با انجام عملیات نرم افزاری قابل حصول نیست و لذا استفاده از قابلیت‌های سیستم‌های سخت افزاری و پردازش موازی برای حصول نتیجه مطلوب ضروری می‌نماید [3,4].

شدن هدف است و به یک قاب قبل برمی‌گردیم و هدف را با ویژگی مبتنی بر شکل-بافت ردیابی می‌کنیم. پس از پیدا کردن مجدد هدف، ردیابی را با ویژگی مبتنی بر رنگ ادامه می‌دهیم. در صورتی که پس از طی این مرحله بازم هدف پیدا نشد، احتمالاً انسداد رخ داده است. بنابراین تا زمانی که هدف مجدداً پیدا شود محل پخش ذرات تغییر نمی‌کند و حوالی مکان قبلی هدف ذرات را پخش می‌کنیم.

به منظور کم شدن خطای ردیابی و تعامل روش با تغییر شکل هدف در طول این فرآیند باید بردار ویژگی هدف بروزرسانی شود. بدین منظور حدآستانه دوم که از حد آستانه اول کمتر است را معیار تصمیم‌گیری قرار می‌دهیم. اگر فاصله بردار ویژگی ذره برنده از آستانه دوم هم کمتر بود نشانگر این است که هدف تغییرات چندانی نداشته و ویژگی نیاز به بروزرسانی ندارد، چراکه فاصله بردار ویژگی از بردار ویژگی هدف بسیار اندک است. اما اگر فاصله بردار ویژگی ذره برنده بین حد آستانه اول و دوم قرار داشت بیانگر این نکته است که ذره برنده همان هدف است اما دچار تغییراتی شده و باید بردار ویژگی مبنای آن مورد مقایسه است بروزرسانی شود. بنابراین بردار ویژگی هدف را تغییر می‌دهیم و در قاب‌های بعد، از بردار ویژگی این ذره به عنوان مرجع مقایسه استفاده می‌کنیم.

روند توضیح داده شده برای هر قاب تکرار می‌شود و موقعیت جدید هدف را مشخص می‌کند. روند بیان شده در شکل ۱ قابل مشاهده است.



شکل ۱: روند کلی روش پیشنهادی

در سال‌های اخیر پژوهش‌های متعددی در زمینه ردیابی اشیا با استفاده از ویژگی‌های مبتنی بر رنگ و ویژگی‌های مبتنی بر شکل-بافت صورت گرفته است. با توجه به پیچیدگی تصاویر و مشخصات ویدئو مورد پردازش، می‌توان از این ویژگی‌ها به تنهایی و یا به صورت ترکیبی استفاده کرد. در [5,6] برای ردیابی هدف، ویژگی رنگ به تنهایی مورد استفاده قرار گرفته است؛ در حالیکه در پژوهش [7] با استفاده از ترکیب ویژگی‌های رنگ و شکل-بافت ردیابی هدف انجام شده است. قابل ذکر است که بلادرنگ بودن فرآیند نیز از جمله مسائل مهم در ردیابی هدف است. یکی از راهکارهای کاهش زمان پردازش، استفاده از فرآیندهایی با قابلیت موازی سازی است، همانطور که در [8,9] با بهره‌گیری از پردازش موازی و واحد پردازنده گرافیک، ردیابی بلادرنگ هدف انجام شده است.

با توجه به مسائل مطرح شده، نیاز به یک فرآیند ردیابی بلادرنگ اشیا در حالی که از دقت و پایداری کافی برخوردار باشد، احساس می‌شود. در این مقاله روشی بر پایه الگوریتم مونت کارلو و بهره‌گیری از پردازش چند هسته‌ای برای ردیابی بلادرنگ هدف در دنباله‌ای از تصاویر ویدئویی ارائه می‌شود. در این روش از ویژگی مبتنی بر رنگ به عنوان ویژگی اصلی هدف استفاده می‌شود؛ چراکه استخراج ویژگی‌های مبتنی بر رنگ عموماً ساده بوده و توان پردازشی کمی نیاز دارند. اگر هدف از دست رفت ویژگی انتخاب شده برای ردیابی در آن لحظه نامناسب بوده و بنابراین به قاب ماقبل برمی‌گردیم و با ویژگی دیگر یعنی ویژگی مبتنی بر شکل-بافت به ردیابی هدف می‌پردازیم و پس از اطمینان از پیدا شدن هدف، مجدداً ردیابی را با ویژگی مبتنی بر رنگ ادامه می‌دهیم. در نهایت با توجه به اینکه روش پیشنهادی قابلیت موازی سازی را دارد، با استفاده از هسته‌های متعدد CPU سرعت پردازش را بهبود می‌بخشیم.

۲- روش پیشنهادی

در این روش پس از شناسایی هدف در اولین قاب، به منظور ردیابی هدف در قاب‌های بعدی، ویژگی مبتنی بر رنگ و ویژگی مبتنی بر شکل-بافت را از هدف استخراج می‌کنیم.

در مرحله بعد، با فرض بالا بودن نرخ قاب ویدئو باندازه کافی، بطوریکه جابجایی هدف در دو قاب متوالی زیاد نباشد، در قاب بعدی ویدئو، ذراتی را در حوالی مکان قبلی هدف، به صورت یکنواخت پخش می‌کنیم؛ این ذرات مرکز حدودی است که احتمالاً هدف در آنجا خواهد بود. در مرحله بعد، از این نواحی، ویژگی مبتنی بر رنگ را استخراج می‌کنیم. بنابراین به تعداد ذرات پخش شده در اطراف هدف بردار ویژگی خواهیم داشت.

فاصله این بردارهای ویژگی را از بردار ویژگی هدف محاسبه می‌کنیم. ذره برنده، ذره‌ای است که بردار ویژگی آن با بردار ویژگی هدف کمترین فاصله را داشته باشد.

نهایتاً برای تعیین مکان جدید هدف، دو حد آستانه تعریف می‌کنیم. در صورتی که فاصله بردار ویژگی ذره برنده و بردار ویژگی هدف از حد آستانه اول کمتر بود، ذره برنده به عنوان موقعیت جدید هدف انتخاب می‌شود و اگر از آن حد آستانه بیشتر بود به معنی گم

رابطه (۱) استفاده می‌کنیم که در آن P و Q بیانگر بردار ویژگی هدف و ذره، و p_i و q_i ها بیانگر مقادیر این بردارها هستند.

$$d(P, Q) = \sqrt{\sum_{i=1}^K (p_i - q_i)^2} \quad (1)$$

۳-۲-۳- ردیابی و تخمین مکان هدف

پس از انتخاب شدن ذره برنده، باید مشخص شود که این ذره بیانگر مکان هدف است و یا هدف را گم کرده‌ایم. برای این منظور دو حد آستانه تعریف می‌کنیم.

حد آستانه اول را برای تشخیص اینکه ذره، هدف است یا هدف گم شده به کار می‌بریم و حد آستانه دوم را که کوچکتر از حد آستانه اول است، به منظور معیاری برای بروزرسانی بردار ویژگی هدف استفاده می‌کنیم.

اگر فاصله بردار ویژگی ذره برنده و بردار ویژگی هدف، از حد آستانه اول بیشتر باشد یعنی هدف گم شده. در این شرایط به یک قاب قبل برمی‌گردیم و هدف را با ویژگی مبتنی بر شکل-بافت ردیابی می‌کنیم و پس از پیدا شدن هدف ردیابی آن را با ویژگی مبتنی بر رنگ ادامه می‌دهیم.

اگر فاصله بردار ویژگی ذره برنده و بردار ویژگی هدف، کمتر از حد آستانه اول باشد یعنی ذره برنده بیانگر مکان جدید هدف است؛ و اگر این فاصله از حد آستانه دوم هم کمتر باشد بدین معنی است که ویژگی استخراج شده هدف تغییر چندانی نداشته و نیاز به بروزرسانی ندارد. اما اگر فاصله بین دو حد آستانه باشد، نشان می‌دهد تغییرات ویژگی هدف زیاد بوده و باید اصلاح شود. در این حالت از ذره برنده ویژگی‌های مبتنی بر رنگ و شکل-بافت را استخراج می‌کنیم و در پردازش‌های بعدی به عنوان مرجع مقایسه (بردار ویژگی هدف) استفاده می‌کنیم.

این دو حد آستانه بصورت تجربی انتخاب می‌شوند. تعیین حد آستانه اول در رهگیری صحیح هدف و تعداد از دست رفتن هدف موثر خواهد بود؛ و حد آستانه دوم در تعداد بروزرسانی بردار ویژگی هدف موثر خواهد بود؛ که اگر مقدار آن زیاد در نظر گرفته شود (نزدیک به حد آستانه اول) باعث کاهش تعداد بروزرسانی‌ها در طی پردازش می‌شود و منجر به افزایش احتمال از دست رفتن شیء می‌گردد. اگر مقدار این حد آستانه، کم انتخاب شود منجر به بروزرسانی بیش از حد می‌شود که باعث کند شدن فرآیند و کاهش سرعت پردازش خواهد شد.

۴-۲-۴- استفاده از پردازش چند هسته‌ای

در روش‌های مبتنی بر مونت کارلو مانند فیلتر ذره‌ای هرچه تعداد ذرات پخش شده در اطراف هدف بیشتر باشد، احتمال یافتن هدف و تخمین دقیق‌تر از موقعیت هدف افزایش می‌یابد، اما باید توجه داشت که افزایش تعداد این ذرات حجم پردازش‌ها و در نتیجه زمان پردازش هر قاب را افزایش می‌دهد و حتی ممکن است این افزایش زمان منجر به خارج شدن پردازش از حالت بلادرنگ شود [8].

نهایتاً به منظور افزایش سرعت پردازش، یک هسته CPU را برای طی روند اصلی فرآیند، و بقیه هسته‌های CPU را برای استخراج ویژگی از ذرات و محاسبه فاصله بردار ویژگی ذرات از بردار ویژگی هدف به کار می‌گیریم.

۲-۱- استخراج ویژگی

به منظور استخراج ویژگی از هدف و نواحی احتمال وجود هدف، از ویژگی‌های رنگی و شکل-بافت استفاده می‌کنیم. استفاده از ویژگی‌های رنگ و شکل-بافت برای ردیابی اشیا و صورت در سال‌های اخیر بسیار مورد توجه بوده است [10-13]. ویژگی‌های مربوط به رنگ، حساسیت کمتری نسبت به چرخش و تغییر اندازه هدف دارند و به سادگی نیز قابل محاسبه هستند و از لحاظ سرعت پردازش، برای کارهای بلادرنگ مناسب می‌باشند [14]. ویژگی‌های مرتبط با رنگ نسبت به تغییرات ناگهانی نور بسیار تاثیر پذیر بوده و هنگام دخالت یک شیء با رنگ مشابه هدف، منجر به شکست ردیابی می‌شوند. عموماً هدف و اشیا هم‌رنگ هدف، با استفاده از ویژگی‌های مبتنی بر شکل-بافت قابل تمایز هستند و با توجه به این مسئله، این نوع ویژگی‌ها، زمانی که توصیفگرهای رنگی ضعیف عمل می‌کنند، می‌توانند توصیفگر مناسبی باشند [14]. اما باید توجه داشت که این ویژگی‌ها نسبت به ویژگی‌های رنگ محاسبات پیچیده‌تری دارند و لذا زمان پردازش بیشتری را به خود اختصاص می‌دهند.

یکی از ویژگی‌های مبتنی بر رنگ هیستوگرام رنگی است. انتخاب ویژگی رنگ از فضای RGB و HSV متداول‌ترین بوده و در ویدئوهای مورد آزمایش ما فضای رنگی HSV با توجه به پس زمینه و هدف، بهتر از RGB عمل می‌کند. در این مقاله ویژگی مبتنی بر رنگ، هیستوگرام رنگی نرمال شده در فضای رنگی HSV با ۳۲ سطح است. برای ویژگی مبتنی بر شکل-بافت از HOG استفاده می‌کنیم [15]. بدین صورت که در ابتدا ناحیه مورد نظر را به ناحیه‌ای به عرض ۶۴ و طول ۱۲۸ تغییر اندازه می‌دهیم و سپس با HOG از آن استخراج ویژگی می‌کنیم.

بنابراین بردار ویژگی مبتنی بر رنگ برای سه کانال رنگی، یک بردار ۹۶ تایی و بردار ویژگی مبتنی بر شکل-بافت مورد استفاده در این روش، یک بردار ۳۷۸۰ تایی خواهند بود.

۲-۲- پخش ذرات و انتخاب ذره برنده

برای ردیابی هدف در قاب‌ها، در محدوده مکان هدف در قاب قبلی تعدادی ذره را بصورت یکنواخت پخش می‌کنیم. این ذرات مرکز حدودی است که احتمالاً هدف در آنجا خواهد بود. در مرحله بعد از این نواحی استخراج ویژگی می‌کنیم، بنابراین به تعداد ذرات پخش شده در اطراف هدف بردار ویژگی خواهیم داشت.

فاصله بردار ویژگی هریک از ذرات را از بردار ویژگی هدف محاسبه می‌کنیم و ذره‌ای که بردار ویژگی آن با بردار ویژگی هدف کمترین فاصله را داشته باشد را به عنوان ذره برنده انتخاب می‌کنیم.

برای این مقایسه می‌توان از فاصله‌های باتاچاریا، ماهالانویس، اقلیدسی و ... استفاده کرد [11]. ما در این مقاله از فاصله اقلیدسی با

به منظور سنجش دقت ردیابی هدف در روش پیشنهادی، به صورت دستی مکان هدف را در هر قاب از ویدئو مشخص کردیم؛ با توجه به نتایج حاصل از ردیابی هدف توسط روش پیشنهاد شده و مکان واقعی هدف، میزان دقت و مقاومت تعقیب هدف را مورد ارزیابی قرار دادیم. حدود قاب ۳۵۰-ام از ویدئو برای هدف انسداد داریم که فیلتر ذره‌ای تک ویژگی، در هر دو حالت مبتنی بر رنگ و مبتنی بر شکل-بافت، در ردیابی هدف ناموفق عمل کرده و هدف از دست می‌رود؛ در حالیکه روش پیشنهادی هدف را گم نکرده و به درستی ردیابی را ادامه می‌دهد. همچنین در حدود قاب ۱۶۶-ام و ۴۸۸-ام، هدف با شیء مشابه از لحاظ رنگی مواجه می‌شود که فیلتر ذره‌ای مبتنی بر ویژگی رنگ با مشکل مواجه می‌شود. در حالی که روش پیشنهادی در این چالش نیز موفق عمل می‌کند.

قابل ذکر است، زمانی که هدف از صحنه خارج شود و یا دچار انسداد به صورت طولانی مدت شود، روش پیشنهادی دچار مشکل می‌شود، در حالی که فیلتر ذره‌ای در این چالش بهتر عمل می‌کند؛ زیرا فیلتر ذره‌ای با توجه به مدل دینامیکی هدف، مکان بعدی هدف را پیش‌بینی می‌کند و ذرات را حول آن ناحیه متمرکز می‌کند؛ اما در روش پیشنهادی ذرات به صورت تصادفی و یکنواخت در حوالی مکان هدف پخش می‌شوند؛ که البته این مشکل با افزایش تعداد ذرات در روش پیشنهادی تا حدودی قابل رفع است.

شکل ۲ نتایج تجربی حاصل از آزمایش دقت روش پیشنهادی در قاب‌های چالش برانگیز را نشان می‌دهد.

در نهایت بررسی نتایج بدست آمده، مربوط به بررسی دقت روش پیشنهادی نشان می‌دهد که روش پیشنهادی به دلیل عدم توانایی در پیش‌بینی مکان هدف در آینده، در مواجه با چالشی نظیر انسداد طولانی مدت هدف، ضعیف عمل می‌کند. بنابراین روش پیشنهادی برای ردیابی هدف در ویدئوهای نسبتاً ساده که انسداد و شباهت رنگی، به دفعات کم و کوتاه مدت اتفاق می‌افتند بسیار مناسب بوده و از عملکرد مناسبی برخوردار است.

قابل ذکر است روش‌های مقایسه شده، با استفاده از زبان برنامه نویسی ++C و بهره‌گیری از نرم افزار Visual Studio 2015 و کتابخانه OpenCV2.4.8 پیاده‌سازی شده‌اند؛ و آزمایش‌ها بر روی رایانه‌ای با مشخصات ذکر شده در جدول (۲) انجام شده است.

جدول ۲: مشخصات رایانه مورد استفاده

Processor	Intel® Core™ i7-2670QM The number of CPU cores: 4 The number of threads: 8
RAM	8.0 GB
System type	64 bit
Operating system	Windows 7 Home Premium

برای رفع این مشکل پیشنهاد می‌شود از خاصیت پردازش چند هسته‌ای CPU استفاده شود. بدین منظور یک هسته را به عنوان پردازش کننده اصلی و طی کننده روند کلی فرآیند در نظر می‌گیریم و با برنامه نویسی چند ریسمانی، بقیه هسته‌ها را برای پخش کردن ذرات اطراف هدف و استخراج ویژگی و نهایتاً محاسبه فاصله بردارهای ویژگی بکار می‌گیریم.

۲-۵- نتایج تجربی

به منظور بررسی سرعت و کارایی روش پیشنهادی، روش پیشنهاد شده را بر روی دیتابیس PETS2010 [16] اعمال کردیم و نتایج به دست آمده را با روش فیلتر ذره‌ای که از یک ویژگی برای ردیابی هدف استفاده می‌کنند، مقایسه کردیم.

جدول (۱) و شکل ۲، نتایج تجربی به کارگیری روش پیشنهادی را نمایش می‌دهند. قابل ذکر است که آزمایش‌ها با بردار ویژگی رنگی با اندازه ۹۲ و بردار ویژگی HOG با اندازه ۳۷۸۰، انجام شده است.

در جدول (۱)، زمان متوسط پردازش هر قاب، در روش‌های مختلف بیان شده است. به علت پخش ذرات به صورت اتفاقی در اطراف هدف، در هر بار آزمایش مقداری تفاوت در زمان پردازش وجود خواهد داشت. به همین دلیل اعداد ذکر شده در جدول (۱)، متوسط زمان پردازش هر قاب از ویدئوها در ۱۰ بار تکرار آزمایش است.

جدول ۱: مقایسه زمانی پردازش هر قاب

زمان متوسط پردازش هر قاب (میلی ثانیه)				
روش پیشنهادی سازی شده	روش پیشنهادی بدون موازی سازی	فیلتر ذره‌ای با ویژگی رنگی	فیلتر ذره‌ای با ویژگی HOG	تعداد ذرات
۱۰	۱۳	۲۵	۳۵	۴۰۰
۱۳	۱۹	۴۳	۷۳	۱۰۰۰
۲۲	۳۲	۱۰۰	۱۷۶	۴۰۰۰
۳۶	۸۰	۳۴۶	۴۶۵	۱۰۰۰۰
۶۱	۱۷۷	۳۷۶۳	۳۹۳۲	۲۰۰۰۰

نتایج بیان شده در جدول (۱) نشان می‌دهد، از نظر زمان پردازش هر قاب، روش پیشنهادی بسیار سریع‌تر از فیلتر ذره‌ای عمل می‌کند. این افزایش سرعت، هنگام ردیابی با تعداد زیاد ذرات، نمود بیشتری پیدا می‌کند؛ بطوریکه با بیست هزار ذره به منظور ردیابی، روش پیشنهادی در حالت موازی سازی نشده، نسبت به فیلتر ذره‌ای با HOG، پردازش هر قاب را با سرعتی بیش از ۲۰ برابر انجام می‌دهد، و در حالت موازی سازی شده این افزایش سرعت به بیش از ۶۴ برابر می‌رسد.



قاب ۲۵۸



قاب ۲۴۸



قاب ۱۶۳



قاب ۴۹۷



قاب ۴۸۷

شکل ۲: نتایج تجربی، حاصل از آزمایش دقت روش پیشنهادی در قاب‌های چالش برانگیز ناشی از انسداد و تشابه ویژگی رنگی

and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing, 2015, pp. 276-282.

- [6] T. Liu, R. R. Varior, and G. Wang, "Visual tracking using learned color features," in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 1976-1980.
- [7] D. Ding, Z. Jiang, and C. Liu, "Object tracking algorithm based on particle filter with color and texture feature," in 2016 35th Chinese Control Conference (CCC), 2016, pp. 4031-4036.
- [8] C. Choi and H. I. Christensen, "RGB-D object tracking: A particle filter approach on GPU," in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013, pp. 1084-1091.
- [9] J. Steinbring and U. D. Hanebeck, "GPU-accelerated progressive Gaussian filtering with applications to extended object tracking," in 2015 18th International Conference on Information Fusion (Fusion), 2015, pp. 1038-1045.
- [10] A. Rezik, A. Ben-Hamadou, and W. Mahdi, "Face pose tracking under arbitrary illumination changes," in 2014 International Conference on Computer Vision Theory and Applications (VISAPP), 2014, vol. 3, pp. 570-575.
- [11] A. M. Nagy, A. Ahmed, and H. H. Zayed, "Particle filter based on joint color texture histogram for object tracking," in International Image Processing, Applications and Systems Conference, 2014, pp. 1-6.
- [12] S. Heo and N. I. Cho, "Level set tracking using shape prior and global-local color model," in 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 2015, pp. 469-472.
- [13] D. P. Chau, F. Bremond, and M. Thonnat, "A multi-feature tracking algorithm enabling adaptation to context variations," in 4th International Conference on Imaging for Crime Detection and Prevention (ICDP 2011), 2011, pp. 1-6.
- [14] W. Mingming, Z. Weining, and Y. Yang, "Particle filter-based target tracking algorithm with adaptive multi-feature fusion," in 2011 International Conference on Multimedia Technology, 2011, pp. 3595-3598.
- [15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, pp. 886-893, 2005.
- [16] <http://www.cvg.rdg.ac.uk/WINTERPETS09/a.htmlA>, retrieved at 2017-01-04.

۳- نتیجه

در این مقاله روشی برای ردیابی هدف در دنباله‌ای از تصاویر به صورت بلادرنگ مطرح شد. این روش براساس روش مونت کارلو بوده و براساس شرایط هدف به صورت تطبیقی، از یکی از ویژگی‌های مبتنی بر رنگ و یا مبتنی بر شکل-بافت برای ردیابی هدف استفاده می‌کند. روش پیشنهادی مطرح شده قابلیت موازی سازی دارد که با بهره‌گیری از پردازش چند هسته‌ای بر روی CPU عملکرد آن را بهبود بخشیدیم.

در نهایت نتایج حاصل از ردیابی شیء توسط فرآیند مطرح شده با روش فیلتر ذره‌ای که از یک ویژگی برای ردیابی هدف استفاده می‌کنند مورد مقایسه قرار گرفت. نتایج بدست آمده نشان می‌دهد سرعت پردازش روش پیشنهادی نسبت به فیلتر ذره‌ای بیشتر بوده و در پردازش‌های بلادرنگ سریع‌تر عمل می‌کند. همچنین روش ارائه شده، در مقابل چالش‌هایی نظیر انسداد و مواجهه با اشیا مشابه رنگی با هدف، عملکرد مناسبی دارد.

مراجع

- [1] J. Baskaran and R. Subban, "Compressive object tracking: A review and analysis," in 2014 IEEE International Conference on Computational Intelligence and Computing Research, 2014, pp. 1-7.
- [2] W. Żorski and P. Skłodowski, "Object tracking and recognition using massively parallel processing with CUDA," in 2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR), 2015, pp. 977-982.
- [3] M. Rofouei, M. Moazeni, and M. Sarrafzadeh, "Fast GPU-based space-time correlation for activity recognition in video sequences," in 2008 IEEE/ACM/IFIP Workshop on Embedded Systems for Real-Time Multimedia, 2008, pp. 33-38.
- [4] I. Ishii, H. Ichida, and T. Takaki, "GPU-based face tracking at 500 fps," in 2011 18th IEEE International Conference on Image Processing, 2011, pp. 557-560.
- [5] B. Zhang, Y. Xu, and X. Yang, "Online Pedestrian Tracking Using Ensemble Color Feature," in 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing



بهبود سامانه ردیابی تصویری هدف مبتنی بر ویژگی‌های استخراج شده از شبکه عصبی عمیق

احسان عبداللهی^۱، محمد تقی صادقی^۲

^۱ دانشجوی کارشناسی ارشد، دانشکده مهندسی برق، دانشگاه یزد، یزد
eaca89@gmail.com

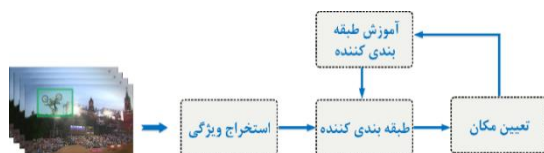
^۲ دانشیار، دانشکده مهندسی برق، دانشگاه یزد، یزد
m.sadeghi@yazd.ac.ir

چکیده

ردیابی تصویری هدف یکی از چالش برانگیزترین زمینه‌ها در حوزه بینایی ماشین است. هدف از ردیابی، تعیین موقعیت هدف در هر لحظه است. با این حال ردیابی با مشکلات مهمی مانند تغییرات روشنایی، انسداد، کلاتر، تغییر شکل و چرخش هدف روبه‌رو است. در دهه‌های گذشته، الگوریتم‌های بسیاری برای غلبه بر این مشکلات ارائه شده‌اند. اخیراً استفاده از شبکه‌های عصبی عمیق برای استخراج ویژگی‌های تصویری مورد توجه قرار گرفته است. این پژوهش بر مبنای یک ردیاب که از شبکه عصبی عمیق به عنوان توصیف‌کننده هدف و از فیلتر همبستگی به عنوان تشخیص دهنده هدف استفاده می‌کند، انجام شده است. در این ردیاب، یک ترکیب خاص از لایه‌های شبکه عصبی عمیق برای توصیف هدف به کار برده شده است. به منظور بهبود عملکرد ردیاب، سایر ترکیب‌های ممکن از لایه‌ها و استفاده از دو شبکه عمیق دیگر بررسی شده است. نتایج حاصل نشان داده که ترکیب ارائه شده در روش مینا بهترین ترکیب نیست و هم‌چنین، افزایش عمیق شبکه، لزوماً باعث تولید ردیاب بهتری نخواهد شد. جهت انتخاب ویژگی از معیار فیشر استفاده شده است تا ویژگی‌های حاصل از شبکه عمیق با نظارت انتخاب شوند.

کلمات کلیدی

ردیابی تصویری هدف، شبکه عصبی عمیق، فیلتر همبستگی، انتخاب ویژگی، نرخ یادگیری وفقی



شکل ۱ بلوک دیاگرام سامانه ردیابی

این ویژگی‌ها یک مدل نمایشی از هدف می‌سازند. به وسیله این مدل، طبقه‌بندی‌کننده آموزش می‌بیند. حال در فریم بعدی به مرکز مشخص شده در مرحله قبل پنجره‌ای بریده شده و ویژگی‌های آن استخراج می‌شود. سپس به وسیله طبقه‌بندی‌کننده، هدف از پس‌زمینه تمیز داده شده و مکان جدید هدف استخراج می‌شود. اکنون در مرکز جدید ویژگی‌ها استخراج شده و طبقه‌بندی‌کننده با ویژگی‌های جدید آموزش می‌بیند. این روند برای سایر فریم‌ها ادامه پیدا می‌کند.

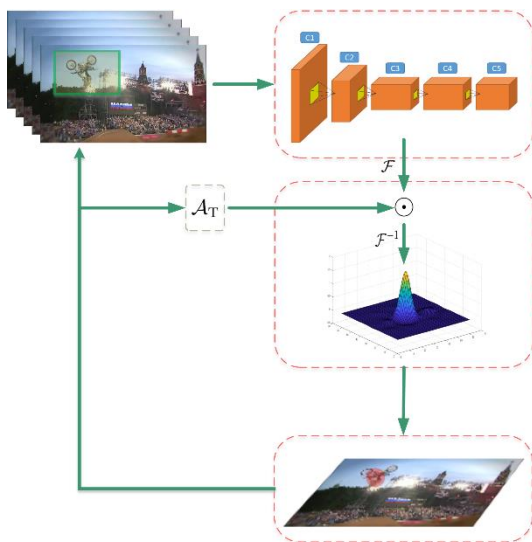
۱- مقدمه

ردیابی تصویری هدف یکی از مسائل پایه‌ای در بینایی ماشین است. یک سناریوی معمول در ردیابی هدف، دنبال کردن یک هدف ناشناخته که در فریم اول با یک مستطیل محصور مشخص شده است می‌باشد. علی‌رغم پیشرفت‌های ردیابی هدف در دهه‌های اخیر، به دلیل تغییرات ظاهری شدید ناشی از انسداد، تغییر فرم شیء، حرکت‌های ناگهانی، تغییرات روشنایی و کلاتر پس‌زمینه، ردیابی هدف هنوز هم یک مساله‌ی چالش برانگیز است [۱]. بلوک دیاگرام سناریوی ردیابی به صورت شکل ۱ است. بر اساس این بلوک دیاگرام، در فریم اول مکان هدف به ردیاب داده می‌شود. سپس پنجره‌ای به مرکز هدف بریده شده و با استفاده از بلوک استخراج ویژگی، ویژگی‌های آن استخراج می‌شود.

در این پژوهش می‌خواهیم روش ارائه شده را از سه جهت مورد بررسی قرار دهیم. اول، استفاده از تک تک لایه‌های شبکه عمیق به عنوان توصیف کننده هدف را به صورت جداگانه بررسی می‌کنیم. دوم، ترکیب‌های مختلف از لایه‌های شبکه عمیق را بررسی می‌کنیم. سوم، تاثیر استفاده از معیار انتخاب ویژگی فیشر بر روی ویژگی‌های شبکه عمیق را بررسی می‌کنیم. در بخش ۳ به معرفی این موارد می‌پردازیم.

۲- الگوریتم روش مبنا

در این بخش، بلوک دیاگرام روش اصلی که CF2 نام دارد، شرح داده می‌شود. در فریم ورودی و به مرکز هدف، با دو برابر ابعاد هدف پنجره‌ای برش داده شده و به عنوان ورودی به شبکه عمیق VGG-19 اعمال می‌شود [۷]. از خروجی لایه‌های سوم، چهارم و پنجم این شبکه به عنوان توصیف کننده هدف استفاده می‌شود. هر کدام از این توصیف‌ها به صورت جداگانه به یک فیلتر همبستگی اعمال شده که در خروجی یک شبه گوسی به دست خواهد آمد. این ۳ شبه گوسی با هم جمع شد و محل قله در شبه گوسی نهایی، مکان هدف خواهد بود. با استفاده از مکان جدید هدف، فیلتر به روزرسانی شده تا با تغییرات ظاهری هدف انطباق پیدا کند. همان‌طور که در شکل ۲ نشان داده شده است. اعمال ویژگی‌های حاصل از شبکه عمیق به فیلتر همبستگی A_T در حوزه فرکانس انجام می‌شود. برای سادگی، در این شکل فقط یک شبه گوسی نشان داده شده است [۵]. در بخش‌های ۱-۲ و ۲-۲ به معرفی شبکه عصبی عمیق و فیلتر همبستگی خواهیم پرداخت.



شکل ۲ بلوک دیاگرام روش مبنا

۲-۱- شبکه عصبی کانولوشن

شبکه‌های عصبی کانولوشن تا حد بسیار زیادی شبیه شبکه‌های عصبی مصنوعی معمول هستند. این نوع شبکه‌ها متشکل از نورون‌هایی با وزن‌ها و بایاس‌های قابل یادگیری (تنظیم) هستند. هر نورون تعدادی ورودی دریافت کرده و سپس حاصل ضرب وزن‌ها در ورودی‌ها را محاسبه کرده و در انتها با استفاده از یک تابع تبدیل (فعال‌ساز) غیرخطی نتیجه‌ای را ارائه می‌دهد [۸]. در این شبکه‌ها ابتدا یک تصویر

اخیرا، ویژگی‌های مبتنی بر شبکه‌های عصبی کانولوشن (CNN) جدیدترین نتایج را در طیف وسیعی از کارهای مربوط به تشخیص اهداف ارائه داده‌اند [۲]. بنابراین، به دست آوردن درکی صحیح جهت استفاده از ویژگی‌های سلسله مراتبی قوی در CNN‌ها برای ردیابی مقاوم دارای اهمیت است. معمولاً ردیاب‌های مبتنی بر یادگیری عمیق، نمونه‌های آموزشی مثبت و منفی در اطراف مکان تخمین زده شده‌ی هدف را برای آموزش یک طبقه بندی کننده به دست می‌آورند؛ که این نمونه‌ها بر اساس ویژگی‌های استخراج شده از یک CNN هستند. اما در این روش دو اشکال عمده وجود دارد. اشکال اول مربوط به استفاده از شبکه عمیق به عنوان یک طبقه بندی کننده‌ی زمان واقعی در الگوریتم‌های تشخیص هدف می‌باشد که فقط از خروجی لایه‌ی آخر برای نمایش هدف استفاده می‌کند. در برخی از کارهای تشخیص هدف، استفاده از ویژگی‌های لایه آخر بسیار موثر است، چرا که این ویژگی‌ها دارای اطلاعات معنایی جهت دسته بندی بوده و هم‌چنین نسبت به تغییرات آزاردهنده مانند تغییرات ظاهری مقاوم هستند. با این حال، هدف اصلی الگوریتم‌های ردیابی هدف، تعیین مکان دقیق هدف است. در نتیجه، استفاده‌ی تنها از ویژگی‌های لایه آخر نمی‌تواند نمایش مناسبی از هدف به دست آورد [۳].

اشکال دوم مربوط به استخراج نمونه‌های آموزشی است. برای آموزش یک طبقه بندی کننده‌ی قوی نیاز به تعداد زیادی از نمونه‌های مثبت و منفی بوده که در بحث ردیابی هدف وجود ندارد. به علاوه در معین کردن مرز تصمیم‌گیری برای تشخیص نمونه‌های مثبت و منفی، همیشه ابهام وجود دارد. چرا که همبستگی این نمونه‌ها به دلیل نمونه برداری در اطراف هدف و هم‌پوشانی زیاد پنجره‌ها، بسیار زیاد است [۴]. برای رفع این مشکلات می‌توان از دو راهکار به صورت توأم استفاده کرد:

۱. استفاده از ویژگی‌های لایه‌های مختلف CNN به جای استفاده تنها از لایه‌ی آخر جهت نمایش هدف در بلوک استخراج کننده‌ی ویژگی.
۲. آموزش فیلتر همبستگی و فقی در خروجی هر لایه بدون نیاز به نمونه برداری جهت طبقه بندی کننده.

اگرچه لایه‌های آخر CNN برای دریافت اطلاعات معنایی موثرتر هستند؛ اما برای دریافت اطلاعات جزئیات مکانی مانند موقعیت هدف، کافی نیستند. از طرف دیگر، لایه‌های اولیه برای تعیین مکان هدف دقت بهتری دارند ولی دربرگیرنده اطلاعات معنایی چندانی نیستند. بنابراین به نظر می‌رسد که استفاده از چندین لایه‌ی CNN برای ردیابی هدف موثر باشد؛ چرا که ویژگی‌های معنایی نسبت به تغییرات ظاهری شدید مقاوم بوده و جزئیات مکانی برای تعیین موقعیت دقیق لازم هستند [۵].

هم‌چنین یک فیلتر همبستگی با ویژگی‌های استخراج شده از هر لایه آموزش می‌بیند و با استفاده از نگاشت پاسخ همبستگی چند لایه‌ای موقعیت هدف به دست می‌آید. تمام نمونه‌های جابه‌جا شده از ویژگی‌ها به عنوان نمونه‌ی آموزشی در نظر گرفته شده و به وسیله یک تابع گوسی به هر کدام یک برجسب با مقدار پیوسته بین صفر تا یک اختصاص داده می‌شود. در نتیجه ابهام نمونه برداری برای آموزش یک طبقه بندی کننده کاهش پیدا خواهد کرد [۶].

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda} \quad (۲)$$

در رابطه (۲)، \hat{k}^{xx} کرنل همبستگی بین نمونه‌های آموزشی بوده علامت \wedge بیان‌گر حوزه فوریه است. λ عبارت تنظیم‌کننده و \hat{y} خروجی مطلوب است. جهت محاسبه کرنل بین دو بردار می‌توان از (۳) استفاده کرد.

$$(۲-۲) \quad (۳)$$

در رابطه (۳)، c نشان دهنده تعداد کانال‌های بردار ورودی و F^{-1} تبدیل فوریه معکوس است. در ضمن برای محاسبه خروجی مطلوب از رابطه (۴) استفاده می‌شود.

$$y(m, n) = \exp\left(-\frac{1}{\sigma^2}\left(m - \frac{M}{2}\right)^2 + \left(n - \frac{N}{2}\right)^2\right) \quad (۴)$$

در رابطه (۴)، M و N ابعاد تصویر ورودی هستند. هم‌چنین σ انحراف استاندارد تابع گوسی است. در بخش ۳، روش‌هایی را برای بهبود عملکرد الگوریتم ارائه خواهیم داد [۶].

۳- روش‌های پیشنهادی

در الگوریتم CF2، علی‌رغم کارایی مشکلاتی وجود دارند. از جمله این‌که، ویژگی‌های خروجی از شبکه VGG بدون هیچ‌گونه پیش‌پردازش به فیلتر همبستگی اعمال شده است. از آن‌جا که این ویژگی‌ها عمومی هستند، امکان دارد در یک سناریوی خاص، برخی از آن‌ها عملکرد مناسبی نداشته و باعث افت عملکرد سامانه شوند. پس شاید اعمال روشی برای حذف این ویژگی‌ها بتواند مفید باشد. هم‌چنین برای بهبود عملکرد الگوریتم، ترکیب‌های دیگری از لایه‌های شبکه VGG و استفاده از سایر شبکه‌های عمیق مورد بررسی قرار می‌گیرد.

۳-۱- پیشنهاد اول: بررسی لایه‌های شبکه عمیق

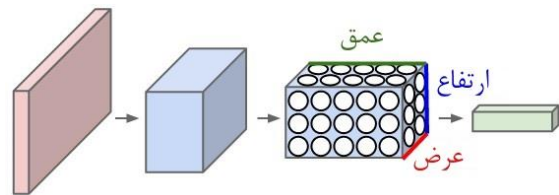
جهت بررسی دقیق‌تر شبکه VGG، ابتدا ویژگی‌های هر کدام از لایه‌ها به صورت جداگانه به ردیاب اعمال شده و کارایی آن از نظر دقت بررسی می‌شود. جهت افزایش دقت ردیاب، سایر ترکیب‌های ممکن از لایه‌ها بررسی شده و بهترین ترکیب به عنوان استخراج‌کننده ویژگی انتخاب می‌شود.

۳-۲- پیشنهاد دوم: استفاده از سایر شبکه‌ها

جهت تأمل بیشتر روی بخش استخراج ویژگی، دو شبکه دیگر با نام Alex Net و VGG-16 مورد بررسی قرار می‌گیرند [۲]. شبکه Alex دارای ۲۱ و شبکه VGG-16 دارای ۳۷ لایه است. از این نظر می‌توان رابطه بین عمق شبکه (معنایی‌تر شدن ویژگی‌ها) و رزولوشن مکانی در ردیابی را بررسی کرد.

به شبکه داده می‌شود؛ تصویر از یک سری لایه‌های کانولوشنی، غیر خطی، پولینگ و تمام متصل عبور کرده و در نهایت خروجی تولید می‌شود. خروجی می‌تواند به صورت یک عدد تنها یا احتمالی که طبقه تصویر را توصیف می‌کند باشد.

شبکه‌های عصبی کانولوشن از این واقعیت که ورودی شامل تصاویر است استفاده کرده و معماری شبکه را به روشی معقول محدود می‌نمایند. به طور خاص، برخلاف یک شبکه عصبی معمولی، لایه‌های یک شبکه عصبی کانولوشن شامل نورون‌هایی است که در سه بعد عرض، ارتفاع و عمق قرار گرفته‌اند. توده فعال‌سازی به یک توده سه بعدی حاوی مقادیر عددی گفته می‌شود که به عنوان ورودی به تابع فعال‌سازی ارسال می‌شوند. برای همین به آن توده فعال‌سازی نیز گفته می‌شود.



شکل ۳ شبکه عصبی کانولوشن

همان‌طور که در شکل ۳ قابل مشاهده است، در هر لایه از یک شبکه عصبی کانولوشن، نورون‌ها در سه بعد مرتب شده‌اند. هر لایه‌ی این شبکه ورودی را در قالب یک توده سه بعدی دریافت کرده و به یک توده سه بعدی خروجی از مقادیر فعال‌سازی نورون‌ها تبدیل می‌کند. در این مثال لایه ورودی قرمز رنگ حاوی تصویر است. بنابراین، عرض و ارتفاع آن ابعاد تصویر خواهند بود و عمق آن هم برابر با ۳ است. یک شبکه از چند لایه تشکیل می‌شود و هر لایه شیوه کار ساده‌ای دارد؛ که در آن یک توده سه بعدی ورودی دریافت کرده و آن را با استفاده از توابعی مشتق پذیر که ممکن است دارای پارامتر یا بدون پارامتر باشند به یک توده سه بعدی خروجی تبدیل می‌کند [۸].

۲-۲- فیلتر همبستگی

اخیراً ردیاب‌های مبتنی بر فیلتر همبستگی، جزء الگوریتم‌هایی هستند که محققان بیشترین توجه را به آن‌ها داشته‌اند. بر اساس تئوری کانولوشن، به دلیل امکان محاسبه همبستگی به صورت ضرب دو سیگنال در حوزه فرکانس، محاسبات آن بسیار بهینه است. بر این اساس، ردیاب‌های مبتنی بر فیلتر همبستگی، حجم محاسبات کمی دارند [۳]. مطابق رابطه (۱)، برای به دست آوردن مکان هدف باید در حوزه فرکانس ضرب نقطه‌ای بین کرنل خودهمبستگی و فیلتر همبستگی صورت پذیرد.

$$\hat{f}(z) = \hat{k}^{xz} \odot \hat{\alpha} \quad (۱)$$

در رابطه (۱)، α فیلتر همبستگی، k^{xz} کرنل همبستگی بین نمونه‌های آموزشی و تکه‌های کاندیدا برای مکان هدف و $\hat{f}(z)$ پاسخ فیلتر همبستگی است. هم‌چنین \wedge نشان‌دهنده حوزه فرکانس است. برای محاسبه فیلتر همبستگی باید از رابطه (۲) استفاده کرد.

۳-۳- پیشنهاد سوم: انتخاب ویژگی

شاید اعمال ویژگی‌های شبکه‌ی VGG به طبقه‌بندی‌کننده بدون هیچ‌گونه نظارتی، باعث ورود ویژگی‌های نامناسب به سامانه شده و فرآیند ردیابی را دچار اختلال کند. بنابراین به نظر می‌رسد که استفاده از معیاری برای انتخاب ویژگی‌های استخراج شده بتواند مفید باشد. از آن‌جا که شبکه عمیق نقش استخراج‌کننده ویژگی را بر عهده دارد، باید بعد از مرحله استخراج ویژگی، فرآیند انتخاب صورت پذیرد.

هر نگاشت پاسخ در شبکه عمیق حاصل اعمال یک فیلتر خاص بر تصویر ورودی یا نگاشت پاسخ موجود در لایه‌ی قبل است. بنابراین هر نگاشت پاسخ به عنوان یک ویژگی در نظر گرفته می‌شود.

یکی از معیارهای انتخاب ویژگی که بر اساس معیار فاصله رتبه‌بندی را انجام می‌دهد معیار فیشر است. به وسیله این روش، زیر مجموعه‌ای از ویژگی‌ها به دست خواهند آمد؛ به طوری که فاصله نمونه‌های متعلق به یک طبقه کمینه و فاصله‌ی نمونه‌های متعلق به سایر طبقه‌ها بیشینه باشد. محاسبه این معیار برای تک تک ویژگی‌ها به صورت رابطه (۴) می‌باشد [۹].

$$F_i = \frac{\sum_{j=1}^c P_j (\mu_{ji} - \mu_i)^2}{\sum_{j=1}^c P_j (\sigma_{ji})^2} \quad (5)$$

در (۵) P_j احتمال تعلق نمونه آموزشی به طبقه j ام بوده و μ_{ji} و σ_{ji} به ترتیب بردار میانگین و انحراف استاندارد برای ویژگی i ام در آن طبقه هستند. μ_i میانگین ویژگی i ام در داده‌های کلیه طبقات می‌باشد. تعداد طبقات نیز با c مشخص شده است. برای محاسبه P_j از رابطه‌ی $P_j = \frac{m_j}{m}$ استفاده می‌شود که در آن m_j تعداد نمونه‌های طبقه j ام و m تعداد کل نمونه‌ها است.

۴- نتایج

اکنون نتایج اجرای این روش‌ها و تاثیر آن‌ها روی عملکرد کلی سامانه بررسی شده و ضمن مقایسه با الگوریتم اصلی (CF2) سعی می‌شود تحلیلی دقیق از علت به دست آمدن نتایج ارائه گردد.

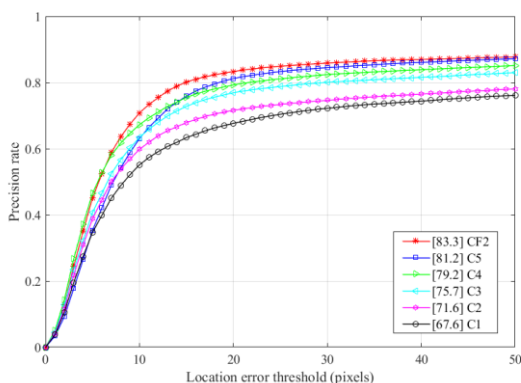
لازم به ذکر است که برای انجام شبیه‌سازی‌های این پژوهش از نرم افزار Matlab 2017 a و جعبه ابزار MatconvNet روی سامانه‌ای با مشخصات سخت‌افزاری CPU Intel i7-6700 3.4GHz*8 و RAM 16GB استفاده شده است. هم‌چنین محاسبات انجام شده توسط CNN در انتشار رو به جلو به GeForce GTX 970 GPU محول شده است.

برای آغاز به کار ردیاب، مرکز اولین فریم در شروع فرآیند به ردیاب داده شده و تا پایان یافتن فریم‌ها، ردیاب بدون دسترسی به مراکز واقعی، مراکز هدف را تخمین می‌زند. سپس با استفاده از معیار ارزیابی و مقایسه با مراکز واقعی، دقت ردیاب به دست می‌آید.

معیار ارزیابی که برای مقایسه‌ی روش‌های مختلف استفاده شده، فاصله اقلیدسی مرکز هدف تخمین زده شده از مرکز هدف واقعی است. بدین منظور حد آستانه‌ای برابر با ۵۰ پیکسل در نظر گرفته شده و به ازای هر کدام از مقادیر بین صفر تا ۵۰، تعداد فریم‌هایی که فاصله مراکز آن‌ها با مراکز واقعی کمتر از مقدار مورد نظر است حساب می‌شوند. اما مقایسه روش‌ها، معمولاً بر اساس تعداد فریم‌هایی که فاصله مرکز آن‌ها با مرکز واقعی کمتر از ۲۰ پیکسل است انجام می‌شود.

روش‌های پیشنهادی روی پایگاه داده‌ای به نام OTB-100 با ۱۰۰ ویدیوی چالشی آزمایش شده است. این ویدیوها در تعداد فریم‌های متفاوتی ارائه شده‌اند. به منظور بررسی روش‌ها در شرایط دشوار و پیچیده‌ی ردیابی، در این پایگاه داده انواع حالت‌های چالشی در ردیابی هدف در نظر گرفته شده است. این شرایط شامل انسدادهای جزئی و کامل، کلاتر پس‌زمینه، تغییرات روشنایی، تغییرات شدید یا جزئی شکل ظاهری هدف، حرکات ناگهانی یا سریع هدف و وجود چندین هدف مشابه در صحنه می‌باشند.

ابتدا تاثیر پنج لایه از VGG-19 روی خروجی ردیاب، به صورت جداگانه بررسی می‌شود. همان‌طور که در شکل مشخص است، بهترین عملکرد در حالتی که از ۳ لایه‌ی شبکه VGG-19 استفاده شود به دست می‌آید (روش CF2 که با رنگ قرمز مشخص شده است). دلیل این امر استفاده از لایه‌های معنایی سطح بالا برای مشخص کردن هدف و جزئیات لایه‌های سطح پایین برای تعیین مکان دقیق هدف است. پس از آن در حالتی که فقط از لایه‌ی C5 استفاده شود عملکرد ردیاب دارای دقت ۸۱/۲ درصد است (دقت شود که دقت گزارش شده برای مقدار ۲۰ پیکسل از محور افقی است). بعد از آن لایه C4 دارای دقت ۷۹/۲ درصد است. علت کمتر بودن دقت این لایه نسبت به لایه‌ی پنج، عمق کمتر ویژگی‌های استخراج شده توسط این لایه است. رتبه‌های بعدی به ترتیب مربوط به لایه‌های سوم، دوم و اول شبکه VGG است. لایه‌های اولیه شبکه VGG معمولاً ویژگی‌های عمومی را استخراج می‌کنند و از آن‌جا که این ویژگی‌ها بین بخش‌های مختلف تصویر مشترک هستند نمی‌توانند باعث تولید یک مدل خاص برای هدف شده و امکان تشخیص اشتباه هدف افزایش پیدا می‌کند. بنابراین با پیش‌روی در عمق شبکه و بررسی خروجی لایه‌ها به صورت جداگانه، دقت ردیاب افزایش پیدا خواهد کرد. اما ترکیب مناسبی از لایه‌های اولیه و لایه‌های معنایی رشد قابل قبولی را در پی دارد.

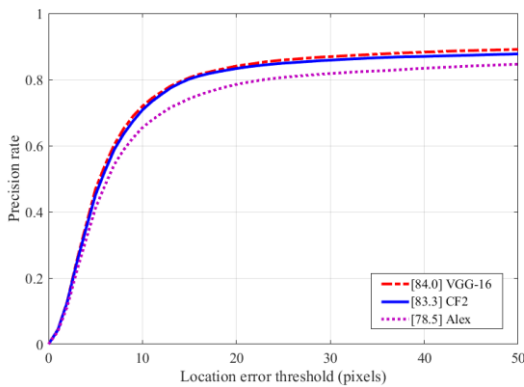


شکل ۴ عملکرد ردیاب با استفاده از لایه‌های مختلف

در الگوریتم CF2 ترکیب مناسبی از لایه‌های پنج، چهار و سه استفاده شده و نتایج قابل قبولی ارائه شده است. اکنون سایر ترکیب‌های لایه‌ها که می‌تواند در فرآیند ردیابی عملکرد مثبتی داشته باشد بررسی می‌گردد. در این الگوریتم خروجی‌های لایه‌های تمام متصل استفاده نشده است. دلیل این امر رزولوشن مکانی بسیار پایین این لایه‌ها می‌باشد. اما اولین لایه‌ی تمام متصل رزولوشن مکانی قابل قبولی دارد. بنابراین به نظر می‌رسد که مناسب است تا از ویژگی‌های استخراج شده

شبکه به اندازه کافی مقاوم نبوده و سبب افت دقت شبکه می‌شود. در این حالت دقت ردیاب برابر با ۷۸/۵ است که نسبت به الگوریتم CF2 حدوداً ۵ درصد کمتر است. اما از نظر سرعت، استفاده از شبکه Alex باعث رشد چشمگیری در سرعت ردیاب شده و آن را حدوداً سه برابر افزایش می‌دهد.

شبکه دیگر VGG-16 است که در فصل چهار در مورد آن توضیح داده شد. این شبکه نسبت به VGG-19 سه لایه‌ی کانولوشنی کمتر دارد. پس از این لحاظ تعداد پارامترهای کمتری داشته و سرعت کار آن بیشتر است. با توجه به کمتر بودن عمق شبکه، انتظار می‌رود که دقت کمتری به دست آید. اما خروجی‌های شکل ۷ نشان می‌دهد که دقت این شبکه ۰/۷ درصد بیشتر از CF2 است. علت این امر را می‌توان بدین شکل توضیح داد که افزایش عمق شبکه لزوماً باعث دقت عملکرد ردیاب نخواهد شد؛ چرا که باعث افت رزولوشن مکانی می‌شود.



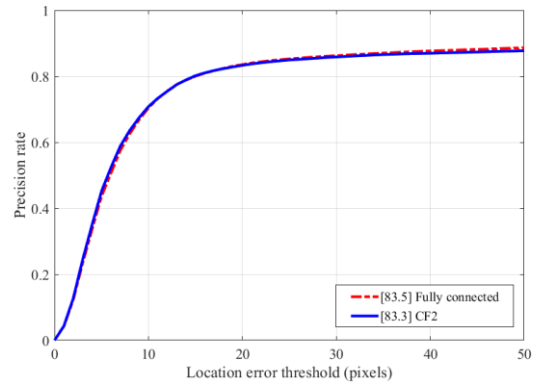
شکل ۷ مقایسه سایر شبکه‌ها با الگوریتم CF2

همان‌طور که در فصل گذشته اشاره شد، برای انتخاب ویژگی از معیار فیشر استفاده شده است. جهت بهره‌گیری از این معیار نیاز است تا رابطه مورد استفاده با شرایط مسئله تطبیق داده شود. در فرآیند ردیابی، هر لحظه یک فریم پردازش می‌شود و در این فریم باید طبقه‌ی هدف و پس زمینه از هم تفکیک گردند. در نتیجه با یک مسئله دو طبقه روبه‌رو هستیم. اما تعداد نمونه‌ها در هر فریم دو می‌باشد که یکی مربوط به طبقه هدف و دیگری مربوط به طبقه پس‌زمینه است. این یک مشکل در استفاده از معیار فیشر است؛ چرا که تعداد نمونه‌ها بسیار ناچیز است. بنابراین، جهت رفع این مشکل دو راه حل ارائه می‌شود:

۱. محاسبه معیار فیشر روی سی فریم اول و اعمال آن از فریم سی‌ام به بعد،
۲. محاسبه معیار فیشر روی فریم اول تمام ویدیوها قبل از مرحله تست.

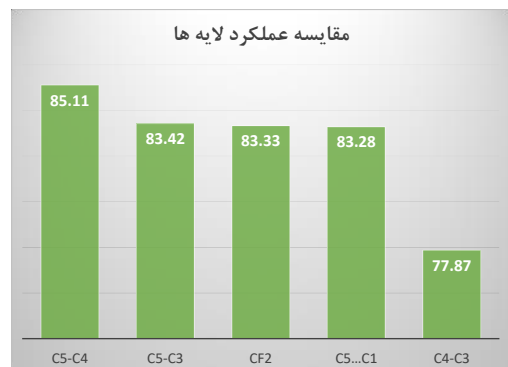
ویژگی‌ها در هر لایه که حاصل اعمال فیلترهای کانولوشنی به نگاشت پاسخ‌اند به صورت ماتریسی بوده و تعداد آن‌ها در لایه پنجم و چهارم ۵۱۲ عدد و در لایه سوم ۲۵۶ عدد است. پس باید میانگین و انحراف استاندارد هر ویژگی روی تعداد نمونه‌ها به دست آید. حال می‌توان با استفاده از معیار فیشر برای هر ویژگی یک امتیاز حساب کرد. انتظار می‌رود که استفاده از این امتیاز، باعث حفظ ویژگی‌های مطلوب و حذف شدن ویژگی‌های نامناسب شود. پس از محاسبه معیار فیشر، با توجه به بازه‌ی گسترده‌ی مقادیر به دست آمده، ابتدا امتیازها به بزرگترین

از این لایه در کنار لایه‌های قبلی برای ساختن مدل ظاهری هدف استفاده نمود. نتیجه استفاده از این لایه در شکل ۵ مشخص شده است. در آستانه بیست پیکسل، روش پیشنهاد شده به میزان ۰/۲ درصد دقت بیشتری دارد و در آستانه‌های بالاتر، این اختلاف بیشتر خواهد شد.



شکل ۵ مقایسه روش لایه‌های تمام متصل و الگوریتم CF2

ترکیب‌های دیگر لایه‌ها شامل استفاده از پنج لایه به صورت هم‌زمان، لایه پنج و چهار، لایه چهار و سه و لایه‌ی پنج و سه است. با توجه به شکل ۶ بالاترین دقت با استفاده از لایه‌های پنج و چهار به دست می‌آید. از آنجایی که لایه پنج معنایی بوده و لایه چهار علاوه بر معنایی بودن، رزولوشن قابل قبولی دارد و مرز اشیاء در آن قابل تشخیص است؛ نیازی به استفاده از لایه سه نبوده و بدین ترتیب سرعت محاسبات بالاتر رفته و ردیابی مقاوم‌تری انجام می‌شود. هم‌چنین در حالتی که از لایه پنج و سه استفاده شود دقت به دست آمده کمی بیشتر از الگوریتم CF2 است. بنابراین دو حالت فوق نتایج بهتری از نظر دقت و سرعت نسبت به روش اصلی به دست آورده‌اند. اما استفاده از پنج لایه به صورت هم‌زمان، علاوه بر اینکه سرعت محاسبات را کاهش می‌دهد؛ دقت روش اصلی را نیز کمی پایین آورده و روش درستی نمی‌باشد. به دلیل آن‌که ویژگی‌های استخراج شده توسط لایه‌های سه و چهار به اندازه کافی عمیق نیست؛ روی کیفیت ردیابی اثر گذاشته و نتیجه ضعیفی در پی خواهد داشت.



شکل ۶ ترکیب لایه‌های مختلف

جهت بررسی قدرت ویژگی‌های استخراج شده توسط سایر شبکه‌های عصبی، شبکه Alex و VGG-16 مورد مطالعه قرار گرفته است. در این شبکه‌ها نیز، سه لایه از شبکه استفاده شده است. با توجه به کمتر بودن عمق شبکه Alex نسبت به شبکه VGG، ویژگی‌های این

باعث بهبود عملکرد ردیاب شد. اما استفاده هم‌زمان از لایه‌های چهار و سه نتیجه مطلوبی ندارد.

در مرحله آخر تاثیر استفاده از معیار انتخاب ویژگی فیشر را بررسی کردیم. اما به دلیل تعداد نمونه‌های آموزشی کم که از مشکلات همیشگی موجود در ردیابی است؛ این معیار نتیجه در حد انتظار از خود به نمایش نگذاشت. در واقع این معیار ویژگی‌هایی را کنار گذاشت که حضور یا عدم حضور آن‌ها در فرآیند انتخاب ویژگی، نقش مهمی ندارد.

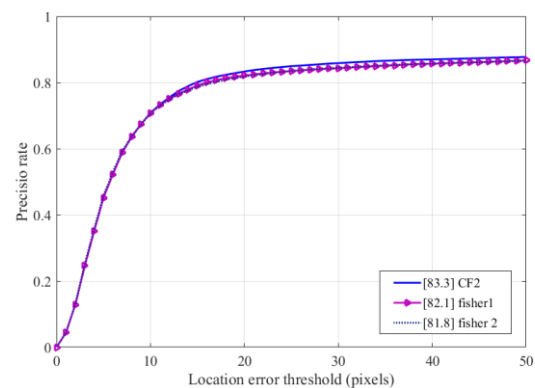
سپاسگزاری

نویسندگان این مقاله بر خود لازم می‌دانند تا از زحمات جناب آقای سعید سعیدی در راستای انجام این پژوهش، نهایت تشکر و قدردانی را به عمل آورند.

مراجع

- [1] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *Acm computing surveys (CSUR)*, vol. 38, p. 13, 2006.
- [2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [3] S. Jeong, G. Kim, and S. Lee, "Effective Visual Tracking Using Multi-Block and Scale Space Based on Kernelized Correlation Filters," *Sensors*, vol. 17, p. 433, 2017.
- [4] B. Babenko, M.-H. Yang, and S. Belongie, "Robust object tracking with online multiple instance learning," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, pp. 1619-1632, 2011.
- [5] C. Ma, J.-B. Huang, X. Yang, and M.-H. Yang, "Hierarchical Convolutional Features for Visual Tracking," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 3074-3082.
- [6] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, pp. 583-596, 2015.
- [7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436-444, 2015.
- [9] C. C. Aggarwal, *Data classification: algorithms and applications*: CRC Press, 2014.

مقدار نرمالیزه شده و سپس ویژگی‌های دارای امتیاز کمتر از ۰/۵ حذف می‌شوند. در ادامه مقایسه‌ای بین روش فیشر و روش CF2 انجام می‌شود.



شکل ۸ مقایسه معیار فیشر و الگوریتم CF2

همان‌طور که در شکل ۸ مشاهده می‌شود در آستانه‌ی ۲۰ پیکسل، روش‌های پیشنهادی یک و دو برای محاسبه معیار فیشر به ترتیب دقت‌های ۸۲/۱ و ۸۱/۸ درصد را به دست آورده‌اند؛ که نسبت به روش اصلی حدود یک درصد افت دارد. در فرآیند ردیابی، از آنجایی که تعداد نمونه‌ها برای معیارهای انتخاب ویژگی کم است نمی‌توان انتظار تاثیر محسوس روی افزایش دقت ردیاب داشت. عملکرد این معیار را می‌توان از نقطه نظر سرعت نیز تحلیل کرد. از آنجایی که حذف تعدادی از ویژگی‌ها روی عملکرد کلی سامانه تاثیر محسوس نگذاشته، در واقع ویژگی‌هایی حذف شده‌اند که حاوی اطلاعات مفیدی نبوده‌اند؛ پس قاعدتا باید سرعت انجام پردازش‌ها افزایش پیدا کند؛ اما به دلیل آن‌که معیار فیشر روی خروجی‌های شبکه عمل کرده و باعث حذف ویژگی‌های نامطلوب می‌شود، نمی‌تواند روی افزایش سرعت ردیاب تاثیر محسوس بگذارد؛ چرا که حدود ۴۵٪ از زمان پردازش هر فریم، برای استخراج ویژگی از شبکه عمیق سپری شده و نمی‌توان با حذف لایه‌های موجود در شبکه باعث کاهش پارامترهای شبکه و افزایش سرعت ردیاب شد.

۵- نتیجه

در این پژوهش به بررسی یک الگوریتم ردیابی هدف مبتنی بر شبکه عصبی عمیق و فیلتر همبستگی پرداختیم (CF2). ابتدا تاثیر هر کدام از لایه‌های شبکه عمیق را بر ردیابی بررسی کردیم. نتیجه به دست آمده نشان داد که هرچه در عمق شبکه پیش برویم، ویژگی‌های استخراج شده توسط شبکه باعث افزایش دقت ردیاب خواهند شد. در واقع ویژگی‌های لایه آخر معنایی‌تر بوده و نسبت به تغییرات ظاهری مقاوم هستند. اما ویژگی‌های اولیه حاوی جزئیات مکانی بیشتری هستند و نسبت به تغییرات ظاهری هدف مقاومت کمتری از خود نشان می‌دهند.

سپس ترکیب لایه‌ها را بررسی کردیم و نتیجه به دست آمده نشان داد که ترکیب لایه چهار و پنج بهترین ردیاب را نتیجه خواهد داد؛ این موضوع نشان می‌دهد که ما توانستیم یک ترکیب بهتر از لایه‌ها را نسبت به روش مبنا پیدا کنیم. در ضمن استفاده از لایه‌ی تمام متصل (که در روش مبنا به دلیل رزولوشن مکانی پایین کنار گذاشته شده بود) هم



روشی سریع و کارآمد برای حذف سایه خودروهای متحرک بمنظور تخمین محدوده دقیق خودرو در تصاویر دریافتی از دوربین جاده‌ای

رسول عسگریان دهکردی^۱، حسین خسروی^۲

^۱ دانشجوی دکتری الکترونیک، دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، شاهرود
r_asgarian_dehkordi@yahoo.com

^۲ استادیار، دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود، شاهرود
Hosseinkhosravi@gmail.com

چکیده

استفاده از دوربین‌های کنار جاده‌ای بمنظور آنالیز ترافیک و بررسی رفتار خودروها در سال‌های اخیر بشدت مورد توجه قرار گرفته است. یکی از چالش‌های رایج در این زمینه تابش شدید نور خورشید و ایجاد سایه خودروها است که شناسایی دقیق خودرو را با مشکل روبرو می‌کند. در این مقاله، الگوریتمی کارا بمنظور شناسایی و حذف اثر سایه خودروهای متحرک برای ویدیوی دریافتی از دوربین‌های کنار جاده‌ای ارائه داده‌ایم. در روش پیشنهادی ابتدا با مدل کردن جاده بعنوان پس‌زمینه، محدوده خودرو و سایه آن بعنوان پیش‌زمینه شناسایی می‌شود. سپس با بدست آوردن مرز پیش‌زمینه و پس‌زمینه و همچنین لبه‌یابی از پیش-زمینه و پس‌زمینه، با استفاده از روشی سریع اثر لبه‌های پس‌زمینه و لبه‌های سایه از لبه‌های پیش‌زمینه حذف و محدوده تقریبی خودرو بدون سایه بدست می‌آید. در ادامه برای افزایش دقت و بهبود پاسخ در نورپردازی‌های متفاوت و ساعات مختلف روز، روشی پیشنهاد داده‌ایم که پس از شناسایی محدوده تقریبی بدون سایه برای چند خودروی متوالی، ناحیه و راستای سایه را در هر زمان تعیین و محدوده دقیق هر خودرو را بدست می‌آورد. برای آزمایش روش پیشنهادی از چندین ویدیوی کنار جاده‌ای که از زوایای مختلف و در نورپردازی‌های متفاوت تهیه شده‌اند استفاده کردیم، نتایج روش پیشنهادی بیانگر عملکرد بسیار مناسب این روش است.

کلمات کلیدی

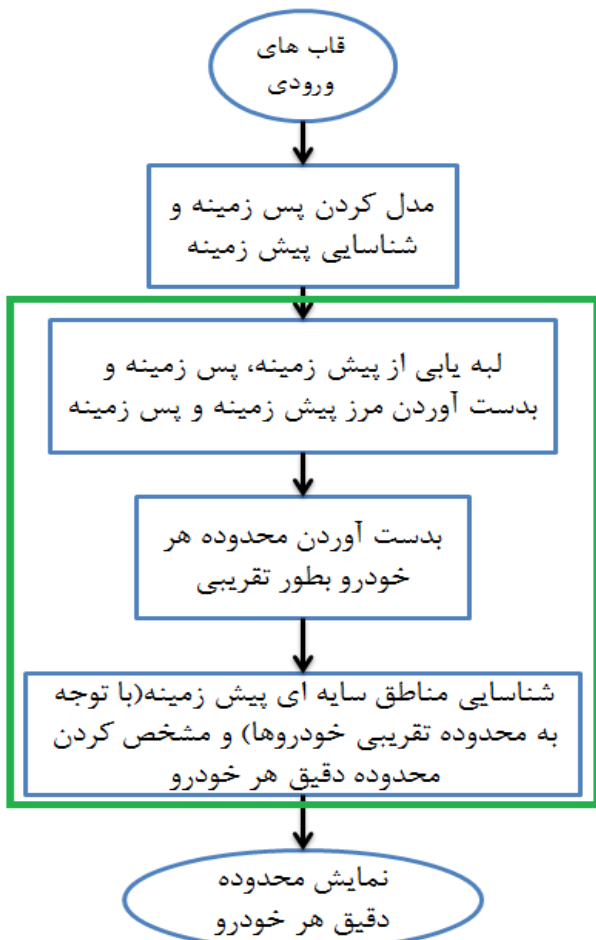
تشخیص ناحیه و جهت سایه، دوربین کنار جاده‌ای، حذف سایه، لبه‌یابی، محدوده خودرو

۱- مقدمه

است. غالباً برای شناسایی خودروهای متحرک از روش‌های مدل کردن پس‌زمینه و شناسایی پیش‌زمینه استفاده می‌شود. در صورت بروز سایه نرم معمولاً این روش‌ها فقط خودرو را بعنوان پیش‌زمینه در نظر می‌گیرند و مشکلی پیش نمی‌آید. اما در صورت بروز سایه سنگین این روش‌ها مقاومت مناسبی ندارند و سایه خودرو را نیز بعنوان پیش‌زمینه و بخشی از خودرو در نظر می‌گیرند. در نظر گرفتن سایه بعنوان بخشی از خودرو، مشکلات فراوانی را برای سامانه‌های تشخیص نوع خودرو، سرعت خودرو و تردد شمار فراهم می‌کند [1].

بسیاری از روش‌ها هستند که از اطلاعات فضاهای مختلف رنگی برای شناسایی و حذف سایه استفاده می‌کنند [2-4]. این روش‌ها سعی در شناسایی محدوده رنگ مربوط به نقاط دارای سایه و نقاط بدون سایه دارند. ژانگ و همکاران [5] یک توزیع آماری Chi-square برای

در حال حاضر در بسیاری از جاده‌ها از ابزارهایی مانند رادار و لیزر برای بررسی ترافیک استفاده می‌شود که اطلاعات حاصل از این روش‌ها محدود است (اکثراً فقط سرعت خودرو و یا نوع خودرو را شناسایی می‌کنند)؛ همچنین هزینه سخت‌افزاری بالایی دارند. یکی از زمینه‌های مورد علاقه در بین ماشین و پردازش تصویر در سال‌های اخیر، استفاده از ویدیوی دوربین‌های جاده‌ای برای آنالیز ترافیک و بدست آوردن پارامترهای مربوط به وسایل نقلیه می‌باشد. استفاده از ویدیوی دوربین‌های نظارتی چالش‌های گوناگونی دارد، یکی از مهمترین این چالش‌ها بروز سایه خودروها در اثر تابش نور خورشید



شکل ۱: مراحل الگوریتم پیشنهادی

۲-۱- مدل کردن پس زمینه و شناسایی پیش زمینه

با توجه به متحرک بودن خودروها، می توان برای شناسایی محدوده خودروها از روش های آشکارسازی پیش زمینه استفاده کرد. تاکنون روش های متعددی برای مدل کردن پس زمینه و شناسایی پیش زمینه ارائه شده است. در روش پیشنهادی انتخاب روش مناسب برای شناسایی پیش زمینه از اهمیت ویژه ای برخوردار است و تاثیر بسزایی در افزایش کارایی و بهبود پاسخ آن دارد. مدل کردن پس زمینه و شناسایی پیش زمینه با چالش های فراوانی از جمله نویز فراوان، تغییر نور محیط و ... روبرو است. بر همین اساس و با توجه به انجام آزمایش های متعدد بر روی روش های مختلف؛ استفاده از روش ارائه شده در [12] را برای مدل کردن پس زمینه و شناسایی پیش زمینه پیشنهاد می کنیم. این روش که یک روش وفقی بهبود یافته بر اساس مدل ترکیبی گوسی (IGMM^۲) است، سرعت بالایی دارد و مقاومت خوبی در برابر سایه نرم و نویز از خود نشان می دهد، در ضمن قابلیت بروزرسانی در طول زمان را دارد.

تصویر سطر اول شکل 2 بخشی از یک فریم ورودی و تصویر سطر دوم این شکل، تصویر خروجی روش استفاده شده برای تقطیع پس زمینه و پیش زمینه را نشان می دهد. همانطور که مشخص است

سایه در نظر گرفته و پیکسل های هر جسم متحرک را به دو کلاس سایه و غیر سایه تقسیم بندی می کنند. مارتل و همکارانش [6] روش GSM^۱ را برای شناسایی سایه ارائه داده اند. در این روش ابتدا با استفاده از روش GMM اجسام متحرک شناسایی می شوند سپس با استفاده از مدل گوسی که برای سایه بدست آورده اند پیکسل های سایه ای را شناسایی می کنند، همچنین برای افزایش کارایی، مدل گوسی سایه را در طول زمان بروز می کنند. برخی روش ها هستند که از اطلاعات بافت برای شناسایی سایه استفاده می کنند. این روش ها فرض را بر این می گیرند که بافت پس زمینه، پیش زمینه و سایه کاملاً از هم تفاوت دارند [7,8]. لیون و همکاران [9] از ویژگی های گابور برای مشخص کردن سایه استفاده کرده است.

لبه بعنوان یک ویژگی مقاوم می تواند در شرایط نورپردازی گوناگون مورد استفاده قرار گیرد. از اینرو در صورت بروز سایه می تواند قابل اتکا باشد [10]. در [11] با استفاده از اطلاعات لبه ای خودروها و پس زمینه، همچنین استفاده از اطلاعات رنگ مربوط به خودروها و پس زمینه، مناطق سایه ای مشخص و با حذف این مناطق محدوده دقیق خودروها بدست آمده است.

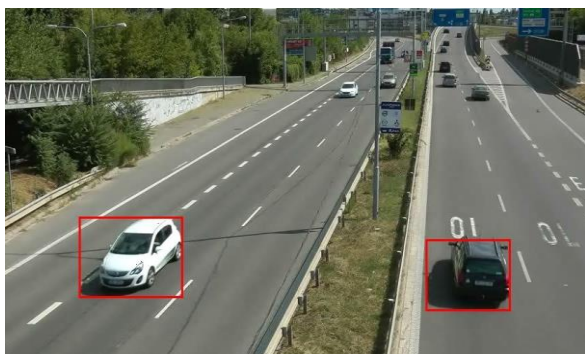
در روش پیشنهادی هدف اصلی این است که بتوانیم محدوده دقیق خودروها را در شرایط نورپردازی گوناگون بدست آوریم. همچنین تمایل داریم تا حد امکان پیچیدگی محاسباتی نداشته باشیم تا الگوریتم کند نباشد. استخراج بافت و استفاده از مدل های آماری برای رنگ و ... محاسبات بالایی دارد و منجر به کندی الگوریتم می گردد. همانطور که اشاره کردیم ویژگی لبه یک ویژگی مقاوم در برابر تغییرات نورپردازی است، علاوه بر این استخراج و استفاده از لبه زیاد زمانبر نیست. بر همین اساس در روش پیشنهادی با استفاده از اطلاعات لبه ای پیش زمینه و پس زمینه، ابتدا محدوده تقریبی خودروها را بدست می آوریم سپس با توجه به محدود تقریبی خودروها، جهت و ناحیه سایه را در هر زمان شناسایی و در نهایت محدوده دقیق هر خودرو را تعیین می کنیم. در ادامه به بیان روش پیشنهادی و نتایج آن می پردازیم.

۲- روش پیشنهادی

شکل شماره ۱ مراحل الگوریتم پیشنهادی را نمایش می دهد. کادر سبز رنگ در این شکل بخش های فرایند مربوط به حذف سایه را نشان می دهد. همانطور که مشخص است در روش پیشنهادی ابتدا پس زمینه مدل می شود و محدوده پیش زمینه بدست می آید. سپس مراحل حذف سایه شروع می شود. به اینصورت که ابتدا لبه های پیش زمینه و پس زمینه آشکار می شود و مرز پیش زمینه و پس زمینه نیز مشخص می گردد. در ادامه با حذف لبه های مرزی و پس زمینه از لبه های پیش زمینه، محدوده تقریبی خودرو بدون سایه با توجه به بقیه لبه های پیش زمینه تعیین می شود. برای افزایش دقت در تعیین محدوده خودرو و همچنین افزایش مقاومت الگوریتم در ساعات مختلف روز (چرخش سایه)، با شناسایی نواحی سایه ای در هر زمان با استفاده از مدل تقریبی چند خودرو، نهایتاً محدوده دقیق هر خودرو تعیین می گردد. در ادامه هر کدام از مراحل روش پیشنهادی بطور کامل تشریح می شود.

² Improved Adaptive Gaussian Mixture Model

¹ Gaussian Mixture Shadow Model



شکل ۲: پاسخ روش [12] در حضور سایه

هدف اصلی در الگوریتم پیشنهادی ارائه روشی است که بتواند محدوده خودروها را با حذف اثر سایه بطور دقیق تعیین کند. بر همین اساس در مراحل بعدی الگوریتم پیشنهادی روشی ارائه می‌دهیم که با استفاده از اطلاعات لبه‌ای پس‌زمینه و پیش‌زمینه، سایه را از پیش‌زمینه حذف می‌کند و محدوده دقیق هر خودرو را مشخص می‌کند.

۲-۲- حذف سایه

برای حذف سایه در روش پیشنهادی لبه‌های پیش‌زمینه، پس‌زمینه و مرز پیش‌زمینه و پس‌زمینه را بدست می‌آوریم. سپس لبه‌های پیش‌زمینه که مشترک با مرز و لبه‌های پس‌زمینه نیستند را شناسایی می‌کنیم تا محدوده تقریبی هر خودرو مشخص شود. در ادامه با تعیین محدوده و جهت سایه، دقت روش را افزایش می‌دهیم.

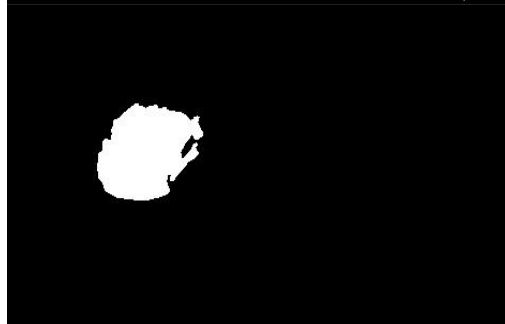
۲-۲-۱- لبه‌یابی از پیش‌زمینه، پس‌زمینه و بدست‌آوردن

مرز پیش‌زمینه و پس‌زمینه

تصویر سطر اول شکل ۳ تصویر باینری محدوده تشخیص داده شده بعنوان پیش‌زمینه در قاب ورودی (تصویر mask)؛ تصویر سطر دوم شکل ۳ محدوده تشخیص داده شده بعنوان پیش‌زمینه (foreground) در قاب ورودی و تصویر سطر سوم محدوده معادل خودرو در تصویر پس‌زمینه (background) را نشان می‌دهد.



پیش‌زمینه به رنگ سفید، پس‌زمینه به رنگ مشکی و نواحی اطراف خودرو که دارای سایه و یا کمی تغییر رنگ هستند به رنگ خاکستری درآمده‌اند؛ با حذف نواحی خاکستری از تصویر سطر دوم و استفاده از عملگرهای مورفولوژی تصویر سطر سوم حاصل می‌شود. با استفاده از این تصویر می‌توان محدوده دقیق خودرو را بدست‌آورد، محدوده خودرو در شکل سطر چهارم نشان داده شده‌است.



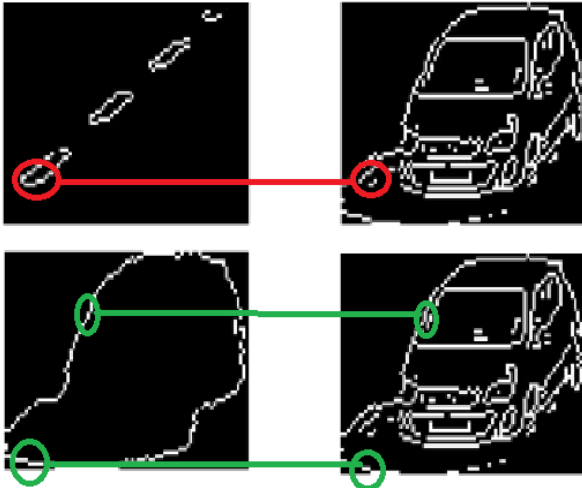
شکل ۲: بدست‌آوردن محدوده خودرو

با توجه به تصاویر شکل ۲ می‌توان نتیجه گرفت که روش IGMM در مقابل سایه نرم مقاومت مناسبی دارد، اما یکی از مشکلات اصلی روش‌های شناسایی پیش‌زمینه، پاسخ نامناسب این روش‌ها در برابر سایه سنگین است. تصویر شکل ۳ پاسخ روش IGMM را در برابر سایه سنگین نشان می‌دهد، همانطور که مشخص است این روش قادر نیست خودرو را به تنهایی شناسایی کند و سایه را نیز همراه خودرو بعنوان پیش‌زمینه در نظر می‌گیرد.

در صورتی که به تصویر f_edge دقت کنیم متوجه می‌شویم که این تصویر علاوه بر اطلاعات لبه‌ای خودرو شامل بخش‌هایی از تصاویر $contour$ و b_edge نیز می‌باشد. این موضوع در تصاویر شکل ۵ نمایش داده شده است.



شکل ۳: سطر اول محدوده باینری یک خودرو و سایه آن بعنوان پیش‌زمینه، سطر دوم تصویر محدوده پیش‌زمینه در قاب ورودی، سطر سوم محدوده معادل با پیش‌زمینه در تصویر پس‌زمینه



شکل ۵: نمایش اشتراک بین تصویر f_edge با تصاویر b_edge و $contour$

پس از شناسایی پیش‌زمینه و محدوده معادل آن در پس‌زمینه، نوبت لبه‌یابی از تصاویر حاصله می‌رسد. برای لبه‌یابی از تصاویر از عملگر $sobel$ استفاده می‌کنیم. پس از لبه‌یابی از تصاویر $mask$ ، تصاویر حاصله را با $contour$ نشان می‌دهیم، این تصویر در حقیقت بیانگر مرز بین پس‌زمینه و پیش‌زمینه است. همچنین پس از لبه‌یابی از تصاویر $background$ ، تصاویر بدست‌آمده را در تصویر $mask$ ضرب می‌کنیم و با b_edge نشان می‌دهیم. در مورد تصویر $foreground$ نیز تصویر حاصله را f_edge نام‌گذاری می‌کنیم.

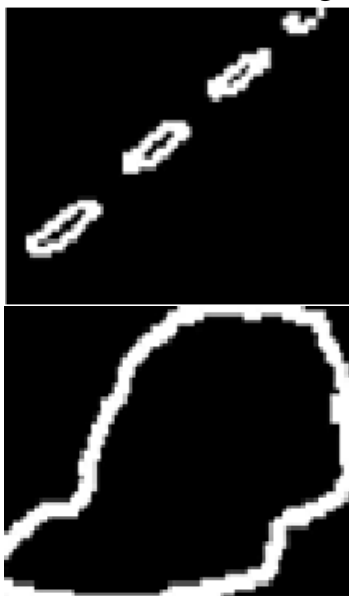
شکل ۴ تصاویر $contour$ ، f_edge و b_edge را نشان می‌دهد.



شکل ۴: نمایش تصاویر $contour$ ، f_edge و b_edge

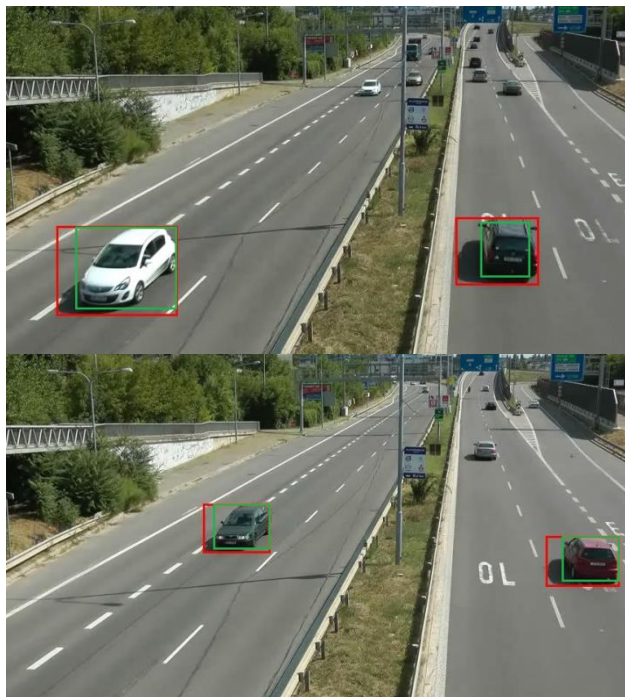
۲-۲-۲- بدست‌آوردن محدوده هر خودرو بطور تقریبی

همانطور که اشاره شد تصاویر $contour$ و b_edge دارای اشتراک‌هایی با تصویر f_edge می‌باشند، بنابراین در صورتی که بتوانیم این اشتراک‌ها را از تصویر f_edge حذف کنیم می‌توانیم محدوده تقریبی خودرو را بدون اثر سایه بدست‌آوریم. برای حذف اشتراک‌های بین تصاویر از روشی مشابه روش ارائه شده در [11] استفاده می‌کنیم. بدین منظور ابتدا تصاویر $contour$ و b_edge را با استفاده از عملیات مورفولوژی به اندازه یک پنجره 2×2 گسترش می‌دهیم. شکل ۶، نتیجه گسترش این تصاویر را نشان می‌دهد.



شکل ۶: گسترش لبه‌های پس‌زمینه و مرز

روشی جدید پیشنهاد می‌دهیم که محاسبات چندانی ندارد و منجر می‌شود محدودده دقیق خودرو مشخص شود.



شکل ۸: نمایش محدودده پیش‌زمینه (کادر قرمز) و محدودده تقریبی خودرو (کادر سبز)

۲-۲-۳- شناسایی محدودده دقیق خودرو

تصویر سمت راست شکل ۹، تصویر خودرو و سایه آن درون جاده، تصویر وسط، لبه‌های مربوط به پیش‌زمینه (خودرو و سایه) و تصویر سمت چپ بخشی از لبه‌های خودرو را نمایش می‌دهد. لبه‌های مرزی تصویر و لبه‌های مشترک با پس‌زمینه در تصویر دوم به رنگ قرمز مشخص شده‌اند، در حقیقت با حذف لبه‌ها از تصویر وسط، تصویر سمت چپ حاصل می‌شود.



شکل ۹: تصویر ستون سمت راست: خودرو درون جاده، تصویر ستون وسط: لبه‌های پیش‌زمینه، تصویر ستون سمت چپ: بخشی از لبه‌های خودرو

با دقت در تصاویر متوجه می‌شویم که حذف کردن اثر پس‌زمینه و لبه‌های مرزی هر چند باعث از بین رفتن محدودده سایه می‌شود اما لبه‌های مرزی خودرو که در مناطق بدون سایه هستند را هم حذف می‌کند و باعث می‌شود محدودده مشخص شده برای خودرو کمتر از محدودده واقعی آن باشد. در صورتی که بتوانیم تشخیص دهیم سایه در کدام ناحیه خودرو قرار گرفته، می‌توانیم از حذف لبه‌های مرزی در نواحی

پس از گسترش تصاویر، برای بدست آوردن محدودده تقریبی خودرو ابتدا تصویر f_edge را بمنظور از بین بردن نویزهای احتمالی به فیلتر median اعمال می‌کنیم و سپس مطابق روابط ۱ و ۲ عمل می‌کنیم:

$$I(x, y) = \max(f_edge(x, y), b_edge(x, y) - b_edge(x, y)) \quad (1)$$

$$car_edge(x, y) = \max(I_edge(x, y), contour(x, y) - contour(x, y)) \quad (2)$$

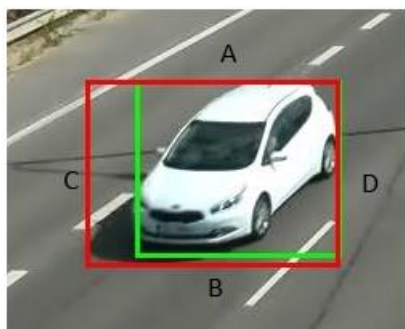
عملگر \max در روابط بالا بیانگر انتخاب بیشترین مقدار برای یک پیکسل است، بعنوان مثال تصویر حاصله از رابطه $\max(f_edge, b_edge)$ تصویری است که در پیکسل‌هایی که تصویر f_edge یا تصویر b_edge برابر با یک هستند مقدار یک را دارد (با توجه به باینری بودن تصاویر). بنابراین تصویر I ، تصویری است که فقط شامل لبه‌های پیش‌زمینه است و لبه‌های پس‌زمینه از آن حذف گردیده، و تصویر car_edge تصویری است که فقط شامل بخشی از لبه‌های خودرو است و لبه‌های پس‌زمینه، سایه و $contour$ در آن وجود ندارد. تصویر سمت راست شکل ۷ تصویر f_edge و تصویر سمت چپ این شکل تصویر car_edge را نشان می‌دهد. همانطور که مشخص است در تصویر car_edge تاثیر لبه‌های سایه و پس‌زمینه از بین رفته، و این تصویر شامل اکثر لبه‌های خودرو است. بنابراین می‌تواند برای بدست آوردن محدودده تقریبی خودرو استفاده گردد. با توجه به اینکه با استفاده از روابط ۱ و ۲ می‌توانیم اثر لبه‌های اضافی بر روی محدودده خودرو را از بین ببریم، علت اصلی گسترش $contour$ و b_edge این است که اطمینان حاصل کنیم همه لبه‌های اضافی از محدودده خودرو حذف شده باشند.



شکل ۷: تصویر سمت راست تصویر پیش‌زمینه (f_edge)، تصویر سمت چپ تصویر (car_edge)

تصاویر شکل ۸ محدودده پیش‌زمینه و محدودده بدست آمده برای چند خودرو را تا این مرحله نمایش می‌دهد. در صورتی که به تصاویر دقت کنیم متوجه می‌شویم که توانسته‌ایم سایه خودرو را حذف کنیم اما متأسفانه برای برخی از خودروها، بخش‌هایی از خودرو که در نواحی بدون سایه خودرو بوده حذف شده که این موضوع منجر به کاهش دقت الگوریتم می‌گردد. در [3,11] برای افزایش دقت در حذف سایه و بدست آوردن محدودده دقیق خودرو از ویژگی رنگ در کنار ویژگی لبه استفاده شده، استفاده از محدودده رنگ برای شناسایی محدودده‌های متعلق به خودرو و پس‌زمینه نیازمند استفاده از روش‌های آماری است که این خود پیچیدگی را افزایش می‌دهد و از سرعت الگوریتم می‌کاهد. بر همین اساس برای افزایش دقت الگوریتم در مرحله بعد

آوریم. سپس برای تعیین اینکه کدام اضلاع نباید جابجا شوند محدوده تقریبی سایه برای n خودروی قبل از این خودرو را در نظر می‌گیریم و هر کدام از اضلاع که برای کمتر از ۲۰ درصد از این n خودرو حداکثر ۲ پیکسل جابجایی داشته‌اند را بعنوان اضلاع بدون نیاز به جابجایی در نظر می‌گیریم. تصویر شکل ۱۲ محدوده تقریبی خودرو (کادر سبز) و محدوده پیش‌زمینه (کادر قرمز) را نمایش می‌دهد. همانطور که مشخص است اضلاع A و D که در مرز بدون سایه قرار داشته، جابجا نشده ولی اضلاع B و C جابجا شده‌اند.



شکل ۱۲: نمایش جابجایی اضلاع مرز سایه (B,C) در محدوده

تقریبی خودروها، کادر قرمز محدوده تشخیص داده شده بعنوان پیش-

زمینه و کادر سبز محدوده تقریبی بدست آمده برای خودرو

با آزمایشاتی که انجام دادیم مشخص شد بررسی محدوده تقریبی سایه برای حدود ۱۰ خودرو قبل ($n=10$) بمنظور شناسایی اضلاع بدون جابجایی کافی است (لازم به ذکر است تردد خودروها در جاده‌های دارای دوربین نظارتی بالاست بنابراین تردد این ۱۰ خودرو در یک محدوده زمانی کوتاه صورت می‌گیرد که در این محدوده زمانی می‌توان منطقه و جهت سایه را ثابت در نظر گرفت). بنابراین برای هر خودرو کفایت محدوده تقریبی خودرو در ۱۰ خودروی قبل از آنرا در جاده تحلیل کنیم. پس از شناسایی اضلاع با جابجایی و بدون جابجایی می‌توانیم محدوده دقیق هر خودرو را تعیین کنیم. همچنین ممکن است شرایطی پیش آید که همه اضلاع بعنوان اضلاع بدون جابجایی شناخته شوند که این حالت زمانی پیش می‌آید که سایه‌ای وجود نداشته باشد (مثل اوایل صبح، زمانی که آسمان ابری است و...).

شکل ۱۳ محدوده دقیق بدست آمده برای چند خودرو را نمایش می‌دهد.



شکل ۱۳: نمایش محدوده دقیق خودروها

۳- بررسی نتایج

برای بررسی نتایج روش پیشنهادی از مجموعه BrnoCompSpeed استفاده کردیم. این مجموعه شامل چندین ویدیوی کنار جاده‌ای است که از زاویه‌های مختلف و در شرایط نورپردازی متفاوت تهیه شده‌اند. بدلیل در دسترس نبودن پیاده‌سازی و پاسخ روش‌های دیگر امکان مقایسه منصفانه با سایر روش‌ها، که چندان هم

بدون سایه جلوگیری کنیم و در نتیجه محدوده دقیقتری از خودرو بدست آوریم.

با آزمایشات و مشاهداتی که با استفاده از چندین ویدیوی کنار جاده‌ای انجام دادیم مشخص شد موقعیت دوربین کنار جاده‌ای بنحوی است که با توجه به زاویه تابش خورشید برای یک محدوده زمانی مشخص، موقعیت سایه برای تمامی خودروها یکسان است. علاوه بر این سایه نمی‌تواند همزمان جلو و عقب خودرو و همزمان سمت راست و چپ راست خودرو قرار گیرد. تصاویر شکل ۱۰ چند خودرو و سایه آنرا برای یک جاده در یک محدوده زمانی مشخص نشان می‌دهد. در صورتی که اضلاع کادر در برگرفته پیش‌زمینه را مطابق شکل در نظر بگیریم، برای تمامی خودروها ضلع A و D منطبق بر مرز خودرو و ضلع B و C منطبق بر مرز سایه هستند. بنابراین در فرایند حذف سایه و تعیین محدوده دقیق خودرو ضلع A و ضلع D نباید جابجایی داشته باشد.



شکل ۱۰: نمایش خودروهای درون جاده برای یک محدوده زمانی،

ضلع‌های A، B، C، D: اضلاع کادر محدوده خودرو

تصویر شکل ۱۱ محدوده تقریبی بدست آمده برای خودروها در مرحله ۲-۲-۲ روش پیشنهادی را نمایش می‌دهد. در صورتی که دقت کنیم متوجه می‌شویم ضلع A و D برای اکثر خودروها جابجا نشده ولی اضلاع B و C برای همه خودروها از مرز سایه جابجا شده و بر مرز خودرو منطبق شده‌است.



شکل ۱۱: نمایش محدوده تقریبی خودروها

با توجه به مشابه بودن جهت سایه برای همه خودروها در یک زمان، عدم قرارگیری سایه بطور همزمان در جلو و عقب یا چپ و راست خودرو و همچنین در دست بودن محدوده تقریبی خودروها، می‌توانیم با بدست آوردن کادر تقریبی خودرو برای چندین خودرو محدوده سایه را بدست آوریم (جهت سایه را شناسایی کنیم) و اضلاعی از کادر که طی فرایند حذف سایه نباید جابجا شوند را تشخیص دهیم. در حقیقت اضلاعی از کادر که بیشترین جابجایی را داشته‌اند در سمتی از خودرو بوده‌اند که سایه و پس‌زمینه وجود داشته و با حذف لبه‌های مربوط به سایه و پس‌زمینه از لبه‌های پیش‌زمینه و بدست آوردن محدوده تقریبی خودرو، مشاهده می‌شود که برای اکثر خودروها (بیش از ۸۰ درصد) این اضلاع بیشتر از بقیه اضلاع جابجا شده‌اند.

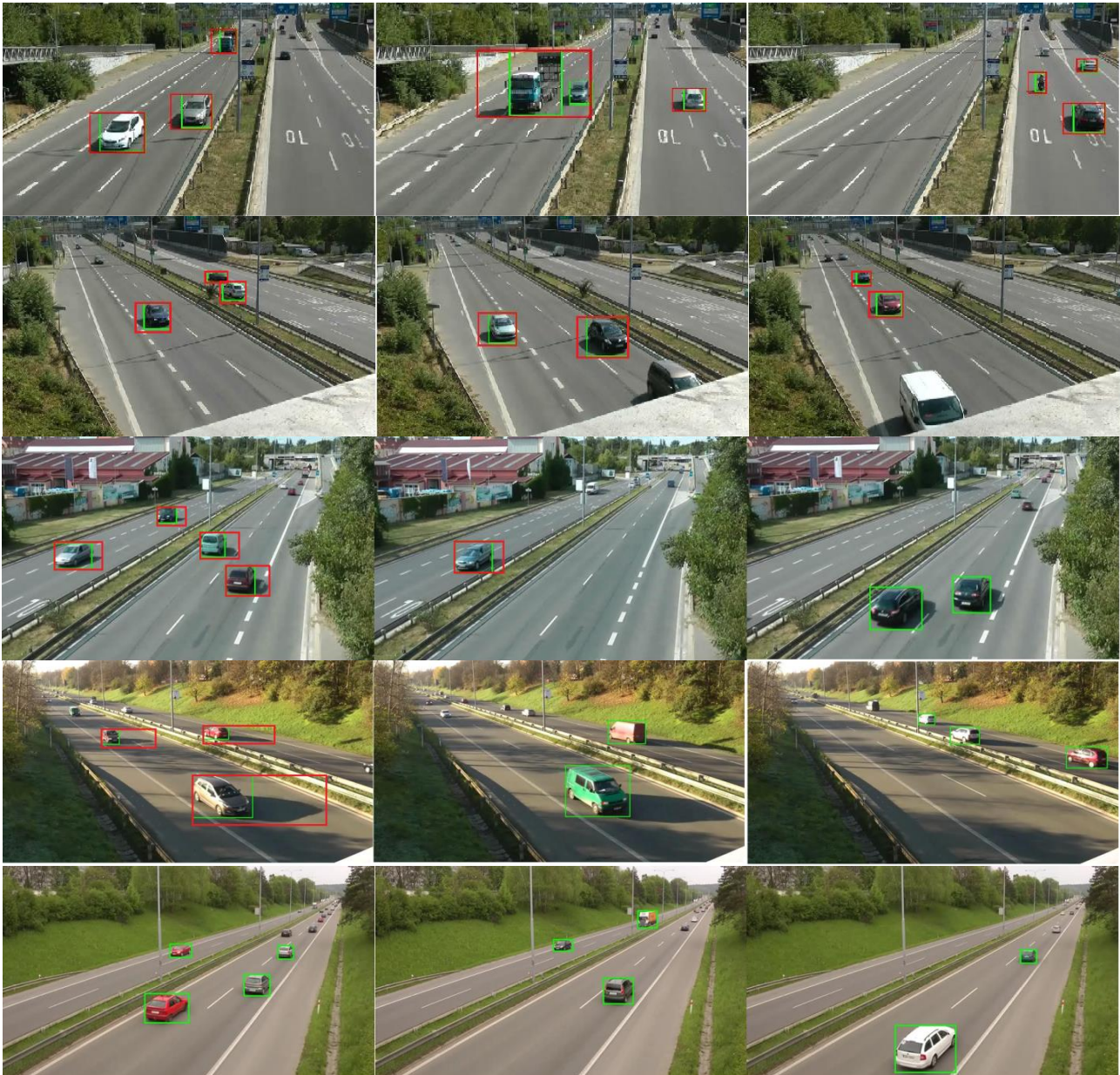
بنابراین به این ترتیب عمل می‌کنیم که پس از بدست آوردن پیش-زمینه با توجه به بخش ۲-۲-۲ محدوده تقریبی خودرو را بدست می-

همانطور که مشخص است روش پیشنهادی در این حالت نیز پاسخ خوبی داشته‌است.

اندازه قاب‌های تمام ویدئوهای مورد آزمایش برابر است با 480×854 ؛ پیاده‌سازی روش پیشنهادی در یک رایانه با 16 GB RAM و CPU 2/6 GHZ صورت گرفته و میانگین سرعت روش پیشنهادی در نرم افزار متلب 2016b برابر با ۷ قاب در ثانیه بدست-آمد.

زیاد نیستند، وجود ندارد. همچنین بدلیل در دست نبودن ground truth فقط بررسی کیفی نتایج صورت گرفته‌است. در تصاویر شکل ۱۴ نتایج روش پیشنهادی برای قاب‌هایی از ویدئوهایی مورد آزمایش نشان داده شده‌است. در این تصاویر، محدوده بدست‌آمده برای خودروها در روش پیشنهادی با کادر سبز رنگ مشخص شده‌است.

همچنین برای تعدادی از فریم‌ها، محدوده شناسایی شده بعنوان پیش‌زمینه با کادر قرمز رنگ مشخص شده‌است. مقایسه کادرهای قرمز و سبز، بیانگر پاسخ خوب روش پیشنهادی است، همانطور که در تصاویر پیداست روش پیشنهادی توانسته با دقت بالایی محدوده خودرو را بدست‌آورد. سطر آخر تصاویر شکل ۱۴ نتایج روش پیشنهادی را برای حالتی نشان می‌دهد که سایه‌ای وجود ندارد،



شکل ۱۴: نمایش نتایج روش پیشنهادی

۴- نتیجه‌گیری

امروزه در اکثر جاده‌ها برای نظارت و کنترل ترافیک از دوربین‌های نظارتی استفاده می‌شود. یکی از مشکلات اصلی در زمینه استفاده از ویدیوی دوربین‌های کنار جاده‌ای ایجاد سایه خودروهای متحرک در ساعات مختلف روز است که شناسایی دقیق خودرو را دشوار می‌سازد. در این مقاله بمنظور ایجاد مقاومت در صورت بروز سایه و شناسایی محدوده دقیق خودرو الگوریتمی ارائه دادیم که پس از شناسایی و لبه‌یابی از پیش‌زمینه و پس‌زمینه و همچنین بدست‌آوردن مرز پیش‌زمینه و پس‌زمینه، تاثیر لبه‌های پس‌زمینه و مرزی را از لبه‌های پیش‌زمینه حذف می‌کند تا تاثیر سایه بطور کامل از لبه‌های پیش‌زمینه حذف گردد. حذف لبه‌های مرزی از لبه‌های پیش‌زمینه منجر می‌شود بخشی از لبه‌های مرز خودرو که در مناطق بدون سایه است نیز از بین برود و در نتیجه محدوده خودرو کوچکتر شود. برای حل این مشکل در مرحله بعد با شناسایی محدوده سایه برای خودروهای متحرک، مرزهای بدون سایه خودروها در هر زمان مشخص می‌شود و با حفظ این مرزها، محدوده دقیق هر خودرو تعیین می‌گردد.

بررسی پاسخ روش پیشنهادی بر روی چندین ویدیو بیانگر پاسخ خوب این روش در شرایط مختلف نورپردازی محیط است. ساختار الگوریتم به گونه‌ای است که پیچیدگی زیادی ندارد و از سرعت مناسبی برخوردار است، علاوه بر این می‌توان در آینده با پیاده‌سازی در OpenCV و یا بر روی GPU، امکان استفاده بلادرنگ را فراهم آورد.

مراجع

[۱۰] معلم، پیمان و رسول عسگریان دهکردی "ردیابی تصویری سریع، مطمئن و مقاوم نسبت به انسداد با کمک یک مدل تقسیم‌شده مبتنی بر لبه"، نشریه مهندسی برق و مهندسی کامپیوتر ایران، ۱-ب، ۳۷-۴۵، بهار ۱۳۹۶.

- [11] Lin, CT., Yang, CT., Shou, YW. et al. "An Efficient and Robust Moving Shadow Removal Algorithm and Its Applications in ITS" in *EURASIP Journal on Advances in Signal Processing*, 2010.
- [12] Zivkovic, Z., "Improved adaptive Gaussian mixture model for background subtraction," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, Vol.2, pp. 28-31, 2004.

- [1] Sochor, J., Juránek R., Špaňhel J., Maršik L. and Široký, A., "BrmoCompSpeed: Review of Traffic Camera Calibration and Comprehensive Dataset for Monocular Speed Measurement", In *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, 2017.
- [2] Cucchiara, R., Grana, C., Piccardi, M., and Prati, A., "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1337-1342, 2003.
- [3] Yang, M.-T., Lo, K.-H., Chiang, C.-C., and Tai, W.-K., "Moving cast shadow detection by exploiting multiple cues," *IET Image Processing*, vol. 2, no. 2, pp. 95-104, 2008.
- [4] Cavallaro, A., Salvador, E., and Ebrahimi, T., "Shadow-aware object-based video processing," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 152, no. 4, pp. 398-406, 2005.
- [5] Zhang, W., Fang, X. Z. and Yang, X. K., "Moving cast shadows detection using ratio edge," *IEEE Transactions on Multimedia*, vol. 9, no. 6, pp. 1202-1214, 2007.
- [6] Martel-Brisson, N. and Zaccarin, A., "Learning and removing cast shadows through a multidistribution approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1133-1146, 2007.
- [7] Joshi, A. J. and Papanikolopoulos, N. P., "Learning to detect moving shadows in dynamic environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 2055-2063, 2008.
- [8] Joshi, A. J. and Papanikolopoulos, N. P., "Learning of moving cast shadows for dynamic environments," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '08)*, pp. 987-992, May 2008.
- [9] Leone, A. C., and Buccolieri, F., "A texture-based approach for shadow detection," in *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS '05)*, pp. 371-376, September 2005.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

تعیین میزان بهینه تبادل برق در نیروگاه‌های تلمبه ذخیره‌ای شبکه الکتریکی هوشمند با استفاده از الگوریتم کرم شب تاب

محسن خنجری^۱، محمد حسینی ابرده^۲

^۱ دانشجوی دکتر، دانشکده برق و کامپیوتر، انشگاه آزاد اسلامی، واحد شاهرود، شاهرود
mohsenkhanjari1@gmail.com

^۲ استادیار، دانشکده برق و کامپیوتر، انشگاه آزاد اسلامی، واحد شاهرود، شاهرود
mohamad.hosseini1@gmail.com

چکیده

الگوی مصرف برق بسیار نامتعادل و غیریکنواخت ایران در ساعات مختلف روز موجب وارد شدن فشار بیش از حد به مجموعه نیروگاهی می‌گردد. شبکه الکتریکی ایران به عنوان یک شبکه رو به رشد به سمت هوشمند سازی می‌رود. در این مقاله مساله اختصاص تولید بهینه به نیروگاه تلمبه ذخیره‌ای، به عنوان پرکاربردترین نوع نیروگاه کنترل کننده در یک شبکه هوشمند، با هدف پیک‌سایبی مورد بررسی قرار گرفته است. بر اساس شرایط بومی ایران و یک سناریو نمونه، الگوریتم کرم شب تاب برای تعیین تولید بهینه یک نیروگاه تلمبه ذخیره‌ای پیشنهاد شده است. یک سناریو عملی با در نظر گرفتن انواع مختلف نیروگاه‌ها و لحاظ کردن شرایط آب و هوایی ایران در نظر شده است. روش پیشنهادی بر روی مدلی از یک شبکه برق واقعی و با در نظر گرفتن احتمال خطا در پیش بینی بار ساعتی پیاده سازی شده و تاثیر آن بر پیک سایی نسبت به حالت بدون وجود نیروگاه نشان داده شد.

کلمات کلیدی

اصلاح الگوی مصرف، الگوریتم کرم شب تاب، شبکه الکتریکی هوشمند، پیک سایی، نیروگاه تلمبه ذخیره‌ای

ذخیره‌ای می‌توان به صورت غیر مستقیم انرژی برق را ذخیره نمود. این نیروگاهها در حقیقت با استفاده غیر مستقیم از اضافه تولید شبکه در ساعات کم مصرف و تزریق آن به شبکه در مواقع پیک، موجب بهینه‌سازی مصرف و کاهش اختلاف بین ساعات کم مصرف و ساعات پر مصرف در شبکه می‌شود. در ایران نیروگاه تلمبه ذخیره‌ای سیاه بیشه با ظرفیت ۱۰۰۰ مگاوات شامل ۴ واحد ۲۵۰ مگاواتی مورد بهره برداری قرار گرفته است. مدیریت نیروگاه‌های تلمبه ذخیره‌ای و توجیه پذیری اقتصادی آنها خصوصا در کشور ما که می‌توان از آنها به عنوان سد و ذخیره کننده آب نیز استفاده کرد بسیار حائز اهمیت می‌باشد. با این دیدگاه در این مقاله سعی شده بر اساس شرایط بومی ایران و یک سناریو نمونه، یک الگوریتم هوشمند برای تعیین تولید بهینه یک نیروگاه تلمبه ذخیره‌ای پیشنهاد گردد. از جمله پارامترهای دخیل برای تعیین تولید

۱- مقدمه

در سالهای اخیر با توجه به اینکه منابع انرژی تجدید ناپذیر روزی به اتمام می‌رسند و آلودگیهای زیست محیطی زیادی نیز هنگام تولید انرژی از خود بر جای می‌گذارند، منابع انرژی پاک و تجدیدپذیر در شبکه های الکتریکی هوشمند مورد توجه ویژه قرار گرفته‌اند. در سال ۲۰۱۳ طبق آمار سازمانهای جهانی حدود ۲۲٪ انرژی مصرفی جهان از راه انرژی‌های تجدیدپذیر به دست می‌آمد که عمدتا انرژی‌های پاک محسوب می‌گردند [۱]. در بین حالت‌های مختلف انرژی‌های تجدیدپذیر می‌توان به نیروگاه‌های تلمبه ذخیره‌ای اشاره کرد. اگرچه ذخیره کردن الکتریسیته به عنوان یکی از انواع انرژی به صورت مستقیم عملی نیست اما با استفاده از روشهایی مانند نیروگاه‌های تلمبه

مدرن باید مدل بازار برق اصلاح شود [۶]. بدیهی است افزایش تعداد نیروگاههای دوستدار محیط زیست و ذخیره‌ساز در یک شبکه ضمن تغییر فاحش در میزان حجم سرمایه‌گذاری لازم و زمان برگشت سرمایه، در قیمت برق تغییرات عمده و اثر گذاری از خود برجای می‌گذارد [۶]. تحلیل جامع از نیروگاه ذخیره‌ساز در ترکیب با یک نیروگاه مبتنی بر انرژی‌های تجدیدپذیر با هدف آنالیز اقتصادی از نظر مدل ارائه شده و دیدگاه ترکیبی آنها بسیار جالب محسوب می‌شود [۷]. ماکزیمم ظرفیت بهینه یک نیروگاه تلمبه ذخیره‌ای در یک شبکه نمونه از نظر عملی بسیار مهم است [۸]. روش استفاده شده برای یافتن این ظرفیت یک روش جستجوی خطی و قیود و تابع هدف مورد بحث نیز در [۸] خطی بوده است. ترکیب یک نیروگاه بادی با یک نیروگاه تلمبه ذخیره‌ای در یک مقطع زمانی کوتاه با هدف پیک‌سایی و بهبود کیفیت توان در [۹] مورد بررسی قرار گرفته است. در نهایت روش پیشنهادی در یک شبکه با چند سناریو مختلف پیاده سازی شده و کارایی روش پیشنهادی در این حالت نشان داده شده است.

پارامترهای دخیل در نیروگاههای تلمبه ذخیره‌ای و نحوه کارکرد اقتصادی آنها نیز یکی از حوزه‌های مهم مدنظر در طراحی است [۱۰]. در انتهای این مرجع یک طراحی نمونه از نیروگاه تلمبه ذخیره‌ای با دیدگاه بهبود کیفیت توان برای یک کشور افریقایی صورت گرفته و کارایی آن در بهبود پارامترهای شبکه نشان داده شده است. جبران سازی شکاف ولتاژ شبکه شامل نیروگاههای تجدیدپذیر با استفاده از نیروگاههای تلمبه ذخیره‌ای یکی از ویژگیهای این نیروگاهها است [۱۱]. نیروگاههای تلمبه ذخیره‌ای با مدیریت صحیح می‌تواند در این موارد برای از بین بردن این شکاف ولتاژی بسیار پر فایده باشد. در این مقاله شرایط مختلف احتمالی شبکه و نحوه مدیریت نیروگاه تلمبه ذخیره‌ای با هدف کنترل و حذف شکاف ولتاژ در حداقل زمان مورد بحث و بررسی قرار گرفته است. میزان ذخیره یا رزرو انرژی در نیروگاه تلمبه ذخیره‌ای نیز باید هنگام کاربرد عملی مد نظر قرار گیرد [۱۲]. این کار با هدف محاسبه انرژی لازم هنگام راه اندازی و ورود به شبکه تخمین زده انجام می‌شود. در نهایت مدل مورد بحث پیاده سازی و از نظر شرایط به صورت کامل با تحلیل غیر خطی و در نظر گرفتن تمام جزئیات به صورت عملی بررسی شده است [۱۲].

در یک سناریو خاص نیروگاه بادی با هدف هماهنگ کردن عملکرد آن با یک نیروگاه تلمبه ذخیره‌ای در هنگام راه اندازی مد نظر قرار گرفته است [۱۳]. تابع هدف در نظر گرفته شده در این مرجع حداقل نوسانات شبکه و حداکثر بودن کیفیت توان بوده است. زمان بندی استفاده بهینه از نیروگاه زمین گرمایی در ترکیب آن با نیروگاه تلمبه ذخیره‌ای و شبکه سناریو دیگری است که میتواند مد نظر قرار گیرد [۱۴]. روش استفاده شده در [۱۴] ازدحام ذرات بوده است. نقش نیروگاههای تلمبه ذخیره‌ای در مجموعه شبکه برق در کشور جمهوری چک از نظر بهبود کیفیت توان و کنترل پایداری شبکه در [۱۵] نشان داده شده است. انتخاب ظرفیت بهینه یک نیروگاه بادی که با نیروگاه تلمبه ذخیره‌ای به صورت همزمان عمل می‌کند با در نظر گرفتن قیود مربوط به نیروگاه بادی، منحنی بار منطقه، میزان وزش باد در منطقه و جزئیات مربوط به جریان آب به صورت کامل در [۱۶] انجام شده است.

بهینه در این نیروگاههای تلمبه ذخیره‌ای می‌توان به شرایط شبکه، بهبود کیفیت توان، عوامل اقتصادی هنگام احداث، شرایط آبی منطقه، شرایط مصرف برق و آب در منطقه و در نهایت امتیازات به دست آمده بر اساس ذخیره نیرو و آب در نیروگاه اشاره نمود. بر اساس سناریو در نظر گرفته شده در این مقاله شبکه الکتریکی هوشمند مد نظر شامل نیروگاه آبی متعارف، نیروگاه خورشیدی یا فوتو ولتاییک، نیروگاه بادی، نیروگاه فسیلی و سیکل ترکیبی خواهد بود که نیروگاه تلمبه ذخیره‌ای در این میان در ترکیب با این نیروگاهها باید بتواند بیشترین کارایی بر اساس سرمایه‌گذاری انجام شده را از خود نشان داده و از نظر اقتصادی توجیه پذیری لازم را داشته باشد.

۲- نیروگاه تلمبه ذخیره‌ای

نیروگاه تلمبه ذخیره‌ای از مجموعه اقداماتی است که برای تغییر الگوی مصرف برق به منظور بهبود قابلیت اطمینان شبکه و جلوگیری از جهش قیمت‌ها، خصوصاً در ساعات پیک بار شبکه اجرا می‌شوند و بدلیل اهمیت موضوع کارهای زیادی در این حوزه در سالهای اخیر انجام شده است.

یک تحلیل جامع با در نظر گرفتن شرایط محیطی اقتصادی برای مشخص کردن ماکزیمم ظرفیت ممکن برای یک نیروگاه تلمبه ذخیره‌ای در یک شبکه برق نمونه در [۲] انجام شده است. در نهایت یک تابع هدف چند وجهی برای بهینه‌سازی بدست آمده و سعی شده با بهینه انتخاب کردن ظرفیت نیروگاه ماکزیمم مقدار برای تابع سود محاسبه شود. مساله به صورت حالتی دو هدفه و چند هدفه با استفاده از روش جستجوی خطی و قرار دادن تعدادی قید خطی عمل بهینه‌سازی صورت گرفته است. نتایج به دست آمده در انتهای مقاله با شرایط عملی مقایسه و بهینه بودن آنها نشان داده شده است. به نیروگاه تلمبه ذخیره‌ای و تاثیر آنها بر قیمت برق در بازار برق در [۳] پرداخته شده است. روش استفاده شده برای بهینه‌سازی در این مقاله یک روش عددی غیر هوشمند بوده و توانسته براساس سناریو طراحی شده در بازار برق تا حد نسبتاً مناسبی در ساعتهای اوج مصرف، قیمت برق را در حدود ۱۸٪ کاهش دهند. زمانبندی هفتگی نیروگاه تلمبه ذخیره‌ای و نیروگاههای ذخیره‌ساز مبتنی بر باتری برای حفظ حاشیه اطمینان شبکه بسیار مهم است [۴]. روش استفاده شده برای بهینه‌سازی این مورد در این مقاله جستجوی ممنوعه بوده است. این روش قابلیت مدل کردن قیود و شرایط غیرخطی و عادی را به صورت کامل داشته، اگر چه تابع هدف در این پیاده سازی یک تابع هدف خطی در نظر گرفته شده است.

تاثیر قیمت برق منفی در برخی ساعات بر کارایی تجهیزات ذخیره‌ساز انرژی حائز اهمیت است [۵]. قیمت برق منفی در مواردی رخ می‌دهد که میزان حداقل تولید نیروگاهها برای باقی ماندن در خط تولید در مجموع از نیاز شبکه در آن ساعت بیشتر است. طبق پیاده سازی انجام گرفته نیروگاهها با کارایی و راندمان پایین در این حالت توصیه نشده و فقط تجهیزات ذخیره‌ساز با هزینه نگهداری کم و میزان ذخیره‌سازی مناسب می‌تواند توجیه اقتصادی داشته باشد. براساس توسعه نیروگاههای تجدیدپذیر شامل نیروگاههای آبی، بادی و فوتوولتاییک و نیروگاههای ذخیره‌ساز و تغییر ساختار در یک شبکه برق

۳-۱- منابع تولید توان

واحدهای تولیدی در شبکه در نظر گرفته شده از نظر قابلیت برنامه ریزی اقتصادی به دو دسته تقسیم می‌شوند.

دسته اول واحدهایی هستند که از نظر تولید کاملاً قابل برنامه‌ریزی بوده و به راحتی می‌توان توان خروجی آنها را کم و زیاد کرد. منابع تولید مبتنی بر سوخت‌های فسیلی از این گروه هستند. دسته دوم شامل گروهی می‌شود که قابلیت کنترل خروجی آنها کمتر از گروه اول و تا حدی تحت تاثیر وضعیت ورودی است. نیروگاه‌های آبی و خورشیدی از این دسته هستند. این نیروگاه‌ها خود به دو دسته تقسیم می‌شوند.

(الف) منابع تجدیدپذیر: از آنجا که توان خروجی منابع تجدیدپذیر مانند سلولهای فتوولتائیک، نیروگاه‌های آبی و حتی سدها دوره‌ای بوده و تحت تاثیر وضعیت ورودی است لذا قابلیت برنامه ریزی در آنها چندان دارای مانور نبوده و همواره هدف استخراج حداکثر توان تولیدی از این منابع و به صورت همزمان رعایت قوانین و قیود مربوط به شبکه است.

(ب) منابع پشتیبان مانند نیروگاه تلمبه ذخیره‌ای: این نیروگاه‌ها خود مولد نیستند و از قابلیت‌های شبکه و ذخیره آنها برای تولید کمک می‌گیرند. به صورت خاص هدف در این پروژه استفاده از قابلیت نیروگاه تلمبه ذخیره‌ای برای پیک‌سایبی است.

بنابراین مفروضات مربوط به نیروگاه‌ها و شبکه را می‌توان به صورت زیر بیان نمود:

(I) انرژی قطعی که شبکه می‌تواند به شما تحویل دهد تا بتوانید نیروگاه تلمبه ذخیره‌ای خود را با آن شارژ کنید به عنوان انرژی تضمین شده شبکه مشخص و معلوم است.

(II) تمامی شبیه‌سازی‌ها با توجه به آنکه نیروگاه‌های فتوولتائیک و بادی بر اساس پیش‌بینی‌های روزانه تنظیم می‌شود در یک دوره زمانی ۲۴ ساعته پیش‌بینی شده و برای هر ۲۴ ساعت آینده مجدداً باید نحوه تحویل توان از شبکه و تزریق مجدد آن به شبکه توسط نیروگاه تلمبه ذخیره‌ای را محاسبه و تنظیم نمود [۱۶ و ۱۷]

(III) انرژی ذخیره شده در لحظه شروع شبیه‌سازی در نیروگاه تلمبه ذخیره‌ای برابر عدد صفر فرض می‌شود.

(IV) انرژی ذخیره شده در نیروگاه تلمبه ذخیره‌ای از اضافه تولید در ساعت کم بار شبکه دریافت و الزاماً در پایان دوره زمانی ۲۴ ساعته در ساعات پیک مصرف به صورت کامل به شبکه تحویل داده شده و نیروگاه در ۲۴ ساعت بعدی از حالت صفر ذخیره دوباره کار خود را آغاز می‌کند.

(V) راندمان نیروگاه تلمبه ذخیره‌ای در پروسه دریافت توان از شبکه و بازگرداندن آن به شبکه برابر عدد ۶۷٪ در نظر گرفته شده که در کاربردهای عملی متداول می‌باشد [۸].

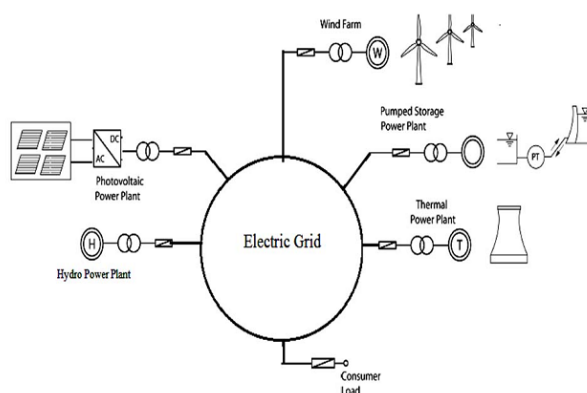
(VI) در شبکه پیاده‌سازی شده منحنی مصرف پیش‌بینی شده برای ۲۴ ساعت آینده در دسترس محسوب می‌شود. به دلیل آنکه الگوریتم‌های پیش‌بینی مصرف در سیستم‌های مدرن و امروزی دقت فوق العاده مناسبی یافته‌اند، چنین فرضی بیهوده به نظر نرسیده و کاملاً منطقی است.

بر اساس مرور انجام شده می‌توان دید نیروگاه‌های تلمبه ذخیره‌ای در شبکه‌های برق مدرن رو به گسترش بوده و به دلیل تنوع عملکرد، شاخص‌های بسیار خوب و گسترده‌ای برای تحقیق محسوب می‌گردند. با توجه به پتانسیل بالای کشور ما در پیاده‌سازی چنین الگوهای، بررسی دقیق‌تر آنها بر اساس جغرافیای مناطق و بومی‌سازی طرح‌های توجیحی می‌تواند به گسترش این گونه نیروگاه‌ها در مناطق مختلف کمک کرده و از این نظر وابستگی صنعت برق به نیروگاه‌های سوخت فسیلی را کاهش دهد. در ادامه ساختار مقاله به شرح زیر است. در بخش بعد ضمن ارائه مقدمه‌ای بر الگوریتم و روش مقاله، ساختار شبکه، مورد بحث و بررسی قرار خواهد گرفت. بخش چهارم به تئوری روش پیشنهادی و جزئیات آن می‌پردازد. بخش پنجم شبیه‌سازی و نتایج مربوط را آورده است. بخش انتهایی نیز به جمع‌بندی و نتیجه‌گیری اختصاص داده شده است.

۳- ساختار شبکه

برای تبیین دقیق روش پیشنهادی ابتدا باید شبکه مد نظر و قیود موجود در آن به صورت کامل توضیح داده شود و سپس با معرفی تابع هزینه و روش کرم شب تاب بهینه‌سازی صورت گیرد. متغیر مد نظر به عنوان پارامتر در اختیار طراح میزان توان تزریقی به نیروگاه تلمبه ذخیره‌ای در ساعات کم مصرف و توان خروجی آن در ساعات پر مصرف است. در این بخش ساختار شبکه و قیود آن را توضیح می‌دهیم و در بخش ۴ به الگوریتم و روش پیشنهادی می‌پردازیم. شکل ۱ ساختار کلی شبکه در نظر گرفته شده را نشان می‌دهد.

همانطور که در شکل ۱ می‌توان دید شبکه مورد بحث شامل انواع مختلفی از نیروگاه‌ها مانند نیروگاه‌های بادی، خورشیدی، سدهای آبی به عنوان نیروگاه‌های تولید پراکنده، نیروگاه‌های فسیلی و سیکل ترکیبی به عنوان نیروگاه‌های متمرکز و نیروگاه تلمبه ذخیره‌ای به عنوان گزینه مد نظر ما بوده است. برای تک تک این وضعیت‌ها و نیروگاه‌ها باید قیود مختلف و شرایط عملی لحاظ شود و پس از آنکه شرایط تک‌تک آنها به صورت کامل در ساختار شبکه مورد بحث مشخص گردید، سناریو مناسب تری در شبکه پیاده‌سازی گردد.



شکل ۱: شبکه در نظر گرفته شده و انواع نیروگاه‌های موجود در آن

۲-۳- پارامترهای نیروگاهی شبکه

برای پیاده سازی یک نمونه از مدل ۳۰ باس IEEE در نظر گرفته شده که برای هر واحد تولیدی کف و سقف تولید مجاز مشخص شده است. در جدول ۱، سقف و کف تولید مجاز واحدهای تولیدی آورده شده است.

با فرض این که تقاضا تابعی اکیداً نزولی از قیمت است، تابع تقاضا در هر بازه زمانی در یک روز مشخص مورد مطالعه می‌تواند بصورت زیر نوشته شود:

$$P_{load}(p) = P_{load}^0(p) + \alpha p \quad (۱)$$

جدول ۱: واحدهای تولیدی، سقف و کف تولید مجاز هر یک

شماره نیروگاه	مینیم تولید	ماکزیم تولید	شماره نیروگاه	مینیم تولید	ماکزیم تولید
1	0	2050	11	0	300
2	0	1398	12	0	356
3	0	988	13	0	300
4	0	726	14	0	0۱۵
5	0	946	15	0	162
6	0	1056	16	0	136
7	0	1056	17	0	66
8	0	1026	18	0	44
9	0	846	19	0	34
10	0	696			

که $P_{load}^0(p)$ تقاضای پیش بینی شده برای آن بازه زمانی و α الاستیسیته تقاضا است. در شرایط بار استاندارد سی باس IEEE در یک بازه زمانی طبق [۱۸] می‌توان فرض کرد.

$$P_{load}(p) = 189.2 - 0.5p \quad (۲)$$

که می‌توان رابطه فوق را برحسب قیمت بصورت زیر نوشت

$$P_{load}(p) = \sum_{i=1}^{Ng} P_{gi} - P_{loss} \quad (۳)$$

Ng تعداد کل تولید کننده‌ها و P_{loss} اتلاف سیستم انتقال را نمایش می‌دهد.

۴- مقدمات شبیه‌سازی

۴-۱- فرض‌های اولیه

• نیروگاههای ۱ تا ۱۰ نیروگاههای سوخت فسیلی و سیکل ترکیبی هستند و محدودیت تامین سوخت در آنها وجود ندارد.
• نیروگاههای ۱۱ و ۱۴ و ۱۷ فوتولتاییک محسوب می‌شود. در این نیروگاهها به صورت ساده از ۷ صبح تا ۷ شب تولید برق داریم. با توجه به آمار روزهای ابری در ایران فرض می‌شود در ۸۲٪ روزها هوای کاملاً صاف داریم. منحنی تولید بر اساس ساعت در روزهای صاف به صورت تابع دوزنقه و به فرم زیر فرض شده است:

$$P = \begin{cases} \frac{P_{max}(t-7)}{3} & 7 < t < 10 \\ P_{max} & 10 < t < 16 \\ \frac{P_{max}(19-t)}{3} & 16 < t < 19 \\ 0 & other \end{cases} \quad (۵)$$

• نیروگاههای ۱۲ و ۱۵ و ۱۸ برق آبی محسوب می‌شود. به صورت ساده در فصل بهار ۹۰٪ ظرفیت پیش‌بینی شده آبی سد و توان نامی در اختیار است. این عدد در فصل تابستان ۵۰٪، در پاییز ۷۰٪ و در زمستان ۹۰٪ خواهد بود. همواره ۱۰٪ ذخیره آب سد به صورت رزرو باقی می‌ماند.

• نیروگاههای ۱۳ و ۱۶ و ۱۹ بادی محسوب می‌شود. در این نیروگاهها با توجه به الگوی وزش باد در منطقه یک تابع چگالی احتمال گسسته به صورت زیر تعریف می‌شود.

جدول ۲: درصد توان خروجی و احتمال رخداد باد در روز

توان خروجی	احتمال رخداد باد
۹۰٪	۰٫۳
۷۵٪	۰٫۳
۵۰٪	۰٫۲
۰	۰٫۲

در نهایت بر اساس اطلاعات موجود از الگوی مصرف برق یک بازار نمونه (PJM در سال ۲۰۰۶) [۱۹] و با تابع هزینه حداکثر پیک سایبی با استفاده از الگوریتم کرم شب تاب میزان توان ذخیره شده در نیروگاه تلمبه ذخیره‌ای محاسبه و پیک های رخ داده قبل و بعد از اعمال روش بررسی می‌شود. به عنوان نتیجه جانبی ظرفیت نیروگاه تلمبه ذخیره‌ای نیز برآورد خواهد شد.

۴-۲- الگوریتم کرم شب تاب

الگوریتم کرم شب تاب یک الگوریتم بهینه سازی است که در سال ۲۰۰۸ توسط یانگ بر اساس رفتار نوری کرم های شب تاب طراحی شده است. این الگوریتم با قرار دادن جمعیتی n عضوی از کرم های شب تاب در نقاط مختلف فضای جستجوی مسئله بهینه سازی به صورت تصادفی آغاز می‌شود. در ابتدا کرم ها مقدار یکسانی از لوسیفیرین (روشنایی) به اندازه ی I در اختیار دارند. هر تکرار الگوریتم شامل یک فاز به روز کردن روشنایی، جذابیت و یک فاز به روز کردن مکان کرم‌ها می‌باشد. در خلال فاز حرکت، هر کرم به صورت احتمالاتی به سمت یکی از همسایگانش که جذابیت بالاتری دارد حرکت می‌کند. به این ترتیب کرم ها به سمت همسایگان با جذابیت بیشتر حرکت می‌کنند. روابط کلی الگوریتم کرم شب تاب به صورت زیر می‌باشد.

$$I = I_0 e^{-\gamma r} \quad (۶)$$

$$\beta = \beta_0 e^{-\gamma r^2}$$

$$r_{ij} = \|x_i - x_j\|$$

$$x_{new_i} = x_{old_i} + \beta_0 e^{-\gamma r_{ij}^2} (x_j - x_{old_i}) + \alpha (rand - \frac{1}{2})$$

I روشنایی، I_0 روشنایی اولیه، β جذابیت، β_0 جذابیت اولیه، γ ضریب جذب، r فاصله اقلیدسی دو کرم، x مکان کرم، rand عددی تصادفی بین صفر و یک، α صفر یا یک می‌باشد. شبه کد کامل الگوریتم کرم شب تاب در [۲۰] آورده شده است.

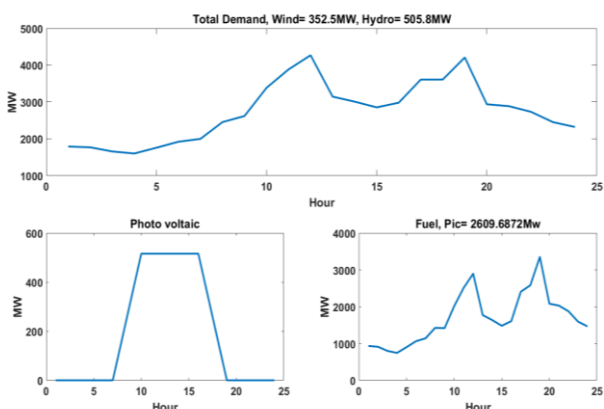
۵- شبیه‌سازی عددی و نتایج

در شبیه‌سازی صورت گرفته اطلاعات مربوط به هر روز به صورت جداگانه مورد پردازش قرار گرفت چراکه فرض این بود نیروگاه تلمبه ذخیره‌ای در هر روز به صورت مستقل از روزهای قبل و بعد مورد استفاده قرار خواهد گرفت [۸].

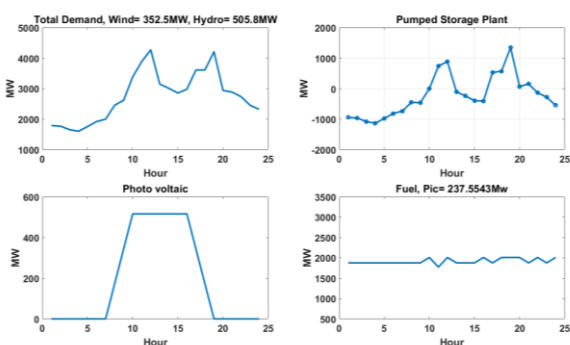
پس از آنکه اطلاعات پیش بینی شده بار و تقاضای هر روز مشخص می‌شود، ابتدا برای قیمت یک نوسان ۶٪ به صورت تصادفی اعمال می‌گردد. این میزان نوسان در پیش بینی قیمت کاملاً طبیعی بوده و در تمامی بازارهای برق دنیا که به صورت خصوصی اداره می‌گردند متعارف است. بدیهی است این میزان برآحتی در یک بازار انحصاری مانند بازار برق فعلی ما می‌تواند صفر منظور گردد. پس از اعمال این نوسان قیمت بر اساس رابطه (۲) مقدار تقاضا بروزسانی می‌گردد. پس از اتمام این دو مرحله، مقدار مصرف نهایی در ۲۴ ساعت روز، به زیربرنامه محاسبه عادی تولید نیروگاهها اعمال می‌گردد. این زیربرنامه کلیه قیود موجود در بخش ۴-۱ را به وضعیت جوی، آفتابی و بازه زمانی روز اعمال کرده و در نهایت تولید مفید هر یک از نیروگاههای تولید پراکنده آبی، خورشیدی و بادی در هر ساعت روز را بدست می‌آورد. این اعداد با مصرف ساعتی مقایسه و در نهایت میزان تولید نیروگاههای فسیلی، یا اضافه تولید در دسترس در هر ساعت به تفکیک محاسبه می‌شوند. در مرحله نهایی تولید نیروگاههای فسیلی به تفکیک ساعت با هدف پیک سابی به زیربرنامه بهینه ساز داده می‌شود. همراه این اعداد مجموع نیروگاهی در دسترس تلمبه ذخیره‌ای به همراه راندمان آن به این قسمت به صورت ورودی اعمال می‌گردد. این زیربرنامه باید به صورت هوشمند و با در نظر گرفتن راندمان و اتلاف در ساعات کم بار تولید را افزایش داده و توان ذخیره شده را در ساعات پر بار مجدداً به شبکه تزریق و از تولید نیروگاههای فسیلی بکاهد. در زیربرنامه بهینه سازی بر اساس الگوریتم کرم شب تاب موارد زیر در نظر گرفته شده است:

تعداد کرم اولیه ۵۱۲ کرم بوده است. مجموعاً الگوریتم یک بردار ۲۴ تایی عدد را تشکیل می‌دهد. این بردار ۲۴ تایی شامل اعداد مثبت و منفی است. اعداد منفی توان دریافتی از شبکه برای شارژ نیروگاه در ساعات کم بار و اعداد مثبت برای ورود به شبکه در ساعات پر بار در نظر گرفته شده اند. مجموع اعداد منفی برابر ماکزیمم توان ورودی نیروگاه تلمبه ذخیره‌ای و مجموع توان خروجی برابر خروجی آن با احتساب راندمان می‌باشد. در صورتی که مجموع نوسانات شبکه در یک روز خاص طوری باشد که نیازی به تزریق توان خروجی ماکزیمم نیروگاه تلمبه ذخیره‌ای به شبکه نباشد و با توان کمتری پیک شبکه حذف شود، در آن صورت برنامه به صورت تطبیقی توان ورودی نیروگاه را کاهش می‌دهد. در نهایت با اعمال موارد فوق تولید نیروگاههای فسیلی متوازن شده و پیک مصرف روزانه تا حد قابل ملاحظه‌ای حذف شده است. الگوریتم کرم شب تاب پیشنهادی با حالت شبکه بدون نیروگاه و بهینه سازی توسط الگوریتم ژنتیک مقایسه و پاسخ در جدول ۳ آورده شده است. همانطور که مشخص است با اضافه شدن نیروگاه تلمبه ذخیره‌ای که میتوانست در یک روز مجموعاً ۵۰۰۰ مگاوات به شبکه تزریق کند، اختلاف موجود بین ساعات پرمصرف و کم مصرف شبکه

به صورت بسیار مشهودی کاهش یافته است. درصد محاسبه شده کاهش پیک بر اساس مجموع تولید کل نیروگاههای فسیلی در یک ساعت یعنی ۱۰۷۸۸ مگاوات محاسبه شده است. ماکزیمم اختلاف بین ساعات پرمصرف و کم مصرف شبکه در یک سال کامل در بدترین حالت بدون اضافه شدن نیروگاه تلمبه ذخیره‌ای ۹۸٪ بوده است که این عدد با اصلاح سیستم به ۲۹٪ کاهش پیدا کرده است. کاهش میانگین پیک به ۳،۳٪ نیز از جمله امتیازات بسیار مشهود روش پیشنهادی است. برای نشان دادن کارایی بدست آمده این استدلال کافی است که با اضافه شدن تنها ۵۰۰۰ مگاوات در طی یک روز کاری پیک موجود در شبکه به صورت میانگین تا ۱۳ برابر کاهش یافته که بسیار قابل ملاحظه است. در شکل‌های ۲ و ۳ برای یک روز بهاری تقاضای کل شبکه و مجموع تولیدات نیروگاههای مختلف به تفکیک قبل و بعد از اعمال روش پیشنهادی آورده شده تا قدرت روش پیشنهادی در حذف پیک و متعادل کردن بار در ساعات مختلف نشان داده شود. برای شرایط یکسان با شکل ۲ و تخصیص بهینه نیروگاه تلمبه ذخیره‌ای همانطور که در شکل ۲ می‌توان دید اختلاف ساعات پرمصرف و کم مصرف در حالت معمولی ۲۶۰۹ مگاوات بوده که با اعمال روش پیشنهادی و به کمک نیروگاه تلمبه ذخیره‌ای در شکل ۳ به ۲۳۷ مگاوات کاهش یافته است. این تغییر مشهود، پتانسیل بسیار بالای این نیروگاه در افزایش پایداری شبکه و متعادل کردن سیستم به کمک روش پیشنهادی را نشان می‌دهد.



شکل ۲: تقاضای بار و تولید انواع نیروگاههای مختلف موجود در شبکه برای یک روز بهاری بدون حضور نیروگاه تلمبه ذخیره‌ای



شکل ۳: تقاضای بار و تولید انواع نیروگاههای مختلف موجود در شبکه برای یک روز بهاری پس از بهینه سازی

جدول ۳: مقایسه نتایج بدست آمده از روش پیشنهادی، وضعیت فعلی شبکه و الگوریتم ژنتیک

انحراف معیار تغییرات مصرف	میانگین بیک%	میانگین بیک (مگا وات)	مینیمم اختلاف %	مینیمم اختلاف (مگا وات)	ماکزیمم اختلاف %	ماکزیمم اختلاف (مگاوات)	
2057	44%	4742	98%	1024	9.50%	10597	بدون نیروگاه تلمبه ذخیره ای
994	9.30%	1011	53.20%	432	4%	5741	الگوریتم ژنتیک
355	3.30%	359	29%	6	0.06%	3088	الگوریتم کرم شب تاب

efficient management of renewable energy sources," 2015 IEEE Students Conference on Engineering and Systems (SCES), Allahabad, 2015, pp. 1-5.

- [9] J. Zou, S. Rahman and X. Lai, "Mitigation of wind output curtailment by coordinating with pumped storage and increasing transmission capacity," 2015 IEEE Power & Energy Society General Meeting, Denver, CO, 2015, pp. 1-5.
- [10] S. A. O. Ilupeju, F. L. Inambao, N. M. Mutombo, T. Nalubega and T. Ustun, "Sustainable energy generation from pumped hydropower," 2015 International Conference on the Domestic Use of Energy (DUE), Cape Town, 2015, pp. 177-185.
- [11] N. Aghanoori and M. A. S. Masoum, "Voltage sag compensation in renewable plant using hydro-pump storage," 2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Brisbane, QLD, 2015, pp. 1-5.
- [12] R. R. Singh and T. R. Chelliah, "Energy saving start-up strategy of pumped storage power plant equipped with doubly-fed asynchronous machine," 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, 2016, pp. 1-6.
- [13] S. Li, J. Wen, J. Li, X. Han and Fang Jiakun, "Coordinated optimization of wind generation and pumped-storage plant by robust unit commitment," 2016 IEEE International Energy Conference (ENERGYCON), Leuven, 2016, pp. 1-6.
- [14] R. S. Patwal and N. Narang, "Heuristic optimization technique for hydrothermal scheduling considering pumped storage unit," 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, 2016, pp. 1-5.
- [15] M. Vitek, T. Králik and J. Tůma, "The role of pumped storage power plants in the power system operation," 2016 17th International Scientific Conference on Electric Power Engineering (EPE), Prague, 2016, pp. 1-5.
- [16] Huiling Li, Chao Zheng, Sizhuo Lv, Shuanbao Liu and Chao Huo, "Research on optimal capacity of wind power based on coordination with pumped storage power," 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi'an, 2016, pp. 1214-1218.
- [17] J. C. Smith, M. R. Milligan, E. A. DeMeo, B. Parsons, "Utility Wind Integration and Operating Impact State of the Art", IEEE Transactions on Power Systems, Volume 22, No. 3, pp. 900 – 908, 2007.
- [18] J. Contreras, M. Klusch, and J. B. Krawczyk, "Numerical solutions to Nash-Cournot equilibria in coupled constraint electricity markets," IEEE Trans. Power Syst., vol. 19, no. 1, pp. 195-206, Feb. 2004.
- [19] <http://www.pjm.com/markets-and-operations/ops-analysis/historical-load-data.aspx>
- [20] Yang, X. S. (2009, October). Firefly algorithms for multimodal optimization. In International symposium on stochastic algorithms (pp. 169-178). Springer, Berlin, Heidelberg.

۶- نتیجه گیری

با توجه به الگویی مصرف بسیار نامتعادل در ساعات مختلف در شبکه برق ایران در این مقاله یک اختصاص تولید بهینه با الگوریتم کرم شب تاب به نیروگاه تلمبه ذخیره‌ای با هدف بیک سایه صورت گرفته است. نتایج بدست آمده نشان داد با اضافه کردن نیروگاه تلمبه ذخیره‌ای به شبکه علاوه بر اصلاح الگویی مصرف، نیروگاه‌های فسیلی به صورت کاملاً متعادل در ساعات مختلف روز فعالیت کرده و الگویی تولید آنها دچار تغییرات زیاد نمی‌شود.

الگوریتم پیشنهادی ضمن لحاظ کردن همه قیودهای ممکن؛ برای هر وضعیت مصرفی در شبکه کاملاً آماده بوده و مشکلی از نظر تخصیص بهینه مقادیر دریافت و تزریق توان به شبکه نداشت. در نهایت این روش بر روی اطلاعات یک شبکه واقعی و با در نظر گرفتن احتمال خطا در پیش بینی بار ساعتی پیاده سازی شد و ماکزیمم بیک نسبت به حالت معمولی ۶۹٪ و نسبت به روش الگوریتم ژنتیک با تابع هزینه یکسان ۲۴٪ کاهش داشت.

مراجع

- [1] Institute for Energy Research's Legacy Society. <http://instituteforenergyresearch.org/topics/encyclopedia/renewabl-e-energy/>
- [2] M. Khatibi and M. Jazaeri, "An analysis for increasing the penetration of renewable energies by optimal sizing of pumped-storage power plants," 2008 IEEE Canada Electric Power Conference, Vancouver, BC, 2008, pp. 1-5.
- [3] D. Connolly, H. Lund, P. Finn, B.V. Mathiesen, M. Leahy, Practical operation strategies for pumped hydroelectric energy storage (PHES) utilising electricity price arbitrage, Energy Policy, Volume 39, Issue 7, 2011, Pages 4189-4196, ISSN 0301-4215,
- [4] J. Kumano and A. Yokoyama, "Optimal weekly operation scheduling on pumped storage hydro power plant and storage battery considering reserve margin with a large penetration of renewable energy," 2014 International Conference on Power System Technology, Chengdu, 2014, pp. 1120-1126.
- [5] Barbour, Edward, Grant Wilson, Peter Hall, and Jonathan Radcliffe. "Can negative electricity prices encourage inefficient electrical energy storage devices?" International Journal of Environmental Studies 71, no. 6 (2014): 862-876.
- [6] C. Spieker, J. Teuwsen, V. Liebenau, S. C. Müller and C. Rehtanz, "European electricity market simulation for future scenarios with high renewable energy production," 2015 IEEE Eindhoven PowerTech, Eindhoven, 2015, pp. 1-6.
- [7] B. Di Pietra and D. A. Sbordone, "Analysis of a energy storage system integrated with renewable energy plants and heat pump for residential application," 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC), Rome, 2015, pp. 410-414.
- [8] Banshwar, N. K. Sharma, Y. R. Sood and R. Srivastava, "Determination of optimal capacity of pumped storage plant by



دسته‌بندی قطبی پاسخ‌ها در یک سیستم پرسش و پاسخ تعاملی

مولود آیت^۱، محمدمهدی حسینی^۲، مرتضی زاهدی^۳

^۱ دانشگاه صنعتی شاهرود، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، molood.ayat@gmail.com

^۲ دانشگاه صنعتی شاهرود، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، hosseini_mm@yahoo.com

^۳ دانشگاه صنعتی شاهرود، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، zahedi@ganjineh.co.ir

چکیده

دسته‌بندی قطبی یکی از شاخه‌های مشهور در پردازش متن است که متن‌های ورودی را با توجه به محتوای درونی آن‌ها، به سه سطح مثبت، منفی و خنثی دسته‌بندی می‌کند. در این مقاله، یک روش پیشنهادی برای دسته‌بندی قطبی پاسخ‌ها در یک سیستم پرسش و پاسخ تعاملی پیشنهاد می‌شود که به بررسی تأثیر مکالمه به دو صورت وابستگی زمانی و ساختاری می‌پردازد. برای این منظور از روش‌های دسته‌بندی با ناظر ماشین بردار پشتیبان، بیزین ساده، درخت تصمیم و الگوریتم bagging استفاده می‌گردد. برای کاهش ابعاد مجموعه ویژگی‌ها و سرعت بیشتر الگوریتم‌ها، از روش‌های کاهش ویژگی استفاده می‌شود. نتایج دسته‌بندی‌ها نشان می‌دهند که استفاده از وابستگی ساختاری موجب بهبود نتایج می‌شود و دقت با استفاده از ماشین بردار پشتیبان با هسته خطی و ویژگی پیشنهادی به ۸۲٫۶۲٪ می‌رسد در صورتی که دقت همین دسته‌بندی بدون در نظر گرفتن گفته‌های قبلی، ۷۷٫۵۴٪ می‌شود و با استفاده از روش کاهش ویژگی مربع کای، دقت برابر ۷۶٫۲۹٪ می‌گردد.

کلمات کلیدی

دسته‌بندی قطبی، مکالمه، استخراج ویژگی، ماشین بردار پشتیبان، سیستم پرسش و پاسخ تعاملی

۱ - مقدمه

از روش‌های پیشرفته پردازش متن، یادگیری ماشین، بازیابی اطلاعات و پردازش زبان طبیعی برای پردازش محتوای متن تولید شده توسط کاربر استفاده می‌کند [3]. آنالیز احساس می‌تواند برای پردازش دیدگاه‌های مربوط به محصولات، نگرش‌های سیاسی، افکار عمومی و خلاصه‌سازی متن استفاده شود [1].

معمولاً آنالیز احساس به صورت نادرست به دسته‌بندی قطبی^۲ تعبیر می‌شود. دسته‌بندی قطبی یکی از وظایف آنالیز احساسی است که احساس مثبت، منفی و خنثی (که به قطب یا بار معروف هستند) را از متون مختلف استخراج می‌کند [4]. تاکنون دسته‌بندی قطبی بر روی متون مختلف، بدون در نظر گرفتن سابقه پیشین متن انجام شده است. بسیاری از عقیده‌ها در پاسخ به دیدگاه‌های گذشته بیان شده‌اند و حالتی مکالمه گونه دارند. در بعضی از این نوع عقیده‌ها بدون در نظر گرفته گذشته عقیده، نمی‌توان نوع دسته آن را مشخص کرد.

تحقیقات، نشان داده است که مهمترین مشکل موتورهای جستجو کنونی عدم تعامل آن با کاربر است. هنگامی که شخصی نیازمند اطلاعاتی است، باید آن را از انبوه منابعی که این موتورها برای

امروزه تعامل میان کاربران و اینترنت از طریق وب افزایش یافته است. در گذشته کاربران اینترنتی تنها قادر به دریافت اطلاعات بودند و نمی‌توانستند دیدگاه‌های خود را در مورد وقایع مختلف بیان کنند. با گسترش منابع اینترنتی نظیر سایت‌های تجاری، وبلاگ‌ها، انجمن‌ها، شبکه‌های اجتماعی و سیستم‌های پرسش و پاسخ، کاربران آزادانه می‌توانند عقیده خود را در مورد یک موضوع خاص بیان کنند. بنابراین حجم اطلاعات بدون ساختار یا نیمه ساختار بیش از پیش افزایش یافته است.

سازماندهی و پردازش حجم عظیم این نوع متن، چالش بزرگی برای فناوری اطلاعات است. زیرا افراد باید بتوانند اطلاعات مفید را مطابق نیاز خود، از اطلاعات بی‌ارزش استخراج کنند [1]. آنالیز احساس^۱ محدوده جدیدی از تحقیق است که در چند سال اخیر، به عنوان شاخه‌ای از مدیریت اطلاعات شناخته شده است [2]. این شاخه

بردار پشتیبان، بیزین ساده، بیشترین آنتروپی و شبکه‌های عصبی برای این منظور انتخاب شدند و ماشین بردار پشتیبان دقت ۹۳٪ را بر روی کلمه‌ها بدست می‌آورد.

مرجع [8-10] دسته‌بندی قطبی را بر روی زبان‌های مختلف انجام داده‌اند. مرجع [8] داده‌های مسابقه‌های بین‌المللی TASS'15 (اسپانیایی)، SemEval'15-16 (انگلیسی) و SENTIPOLC'14 (ایتالیایی) را به عنوان مجموعه داده در نظر می‌گیرد. این مرجع، ویژگی‌های شکلک‌ها، کلمه‌های نفی، n-گرم‌های مبتنی بر کلمه و کاراکتر را بر روی این مجموعه داده‌ها آزمایش می‌کند و ماشین بردار پشتیبان (هسته خطی) را به عنوان روش دسته‌بندی انتخاب می‌کند. دقت برای TASS'15 ۶۳٫۷٪، برای SemEval'15 ۶۲٫۹٪، برای SemEval'16 ۵۳٫۴٪ و برای SENTIPOLC'14 ۶۱٪ بدست آمد.

مرجع [9] از ۳۰۰ مگابایت توییت‌های مربوط به چهار زبان انگلیسی، فرانسوی، آلمانی و ایتالیایی استفاده می‌کند. این مرجع با استفاده از یادگیری عمیق^{۱۵} و skip-gram با پنجره‌ای به اندازه ۵، این توییت‌ها را دسته‌بندی می‌کند. دقت برای توییت‌های انگلیسی ۶۳٫۴۹٪، فرانسوی ۶۴٫۷۹٪، آلمانی ۶۵٫۰۹٪، ایتالیایی ۶۷٫۷۹٪ بدست آمده است.

مرجع [10] از ۵۰۰۰۰ مجموعه دیدگاه فیلم (انگلیسی)، ۱۶۴۴۸ دیدگاه کتاب (عربی) و ۸۸۶ دیدگاه محصول (چینی) استفاده می‌کند. n-گرم‌ها را بر روی هسته‌های خطی^{۱۶} آزمایش می‌کند. دقت برای زبان انگلیسی ۹۲٫۶٪، برای عربی ۸۶٫۵٪ و برای چینی ۹۴٫۲٪ بدست آمد.

مرجع [11] عقیده‌های مربوط به عملکرد هتل‌ها را به عنوان مجموعه داده در نظر گرفت. از کیسه‌ای از کلمه‌ها به عنوان مجموعه ویژگی استفاده کرد و با استفاده از دامنه آنتولوژی فازی^{۱۷} و ماشین بردار پشتیبان دقت ۸۵٫۴٪ برای اتساق، ۸۲٫۷٪ برای رستوران، ۸۱٫۴٪ برای مکان، ۸۲٫۴٪ برای خدمات، ۸۲٫۹٪ برای کارکنان و ۸۱٫۴٪ برای استخر را بدست آورد.

مرجع [12] از ۵۰۰ دیدگاه مربوط به مدل‌های مختلف تلفن همراه در آمازون استفاده می‌کند. روش‌های اطلاعات متقابل نقطه به نقطه^{۱۸}، اطلاعات متقابل^{۱۹}، مربع کای و نرخ شباهت لگاریتمی را با یکدیگر مقایسه می‌کند. روش‌های دسته‌بندی بیزین ساده، درخت تصمیم، چندلایه پرسپترون، الگوریتم bagging را بررسی می‌کند. درخت تصمیم بهترین عملکرد را با دقت ۸۵٫۵٪ بدست آورد.

مرجع [13] کیسه‌ای از کلمه‌ها را به عنوان ویژگی بر روی مجموعه داده‌های ترکی Hepsiburada، Beyazperde، Antoloji اعمال می‌کند. برای دسته‌بندی از ماشین بردار پشتیبان، الگوریتم bagging (ماشین بردار پشتیبان)، بیزین ساده و الگوریتم رای گیری استفاده می‌شود. الگوریتم رای گیری در هر سه مجموعه بهترین دقت را داشت که برای پایگاه‌داده‌های ذکر شده به ترتیب برابرند با ۸۶٫۱۳٪، ۸۳٫۶۸٪ و ۷۹٫۹۶٪.

مرجع [] از مجموعه‌ای از دیدگاه‌های مربوط به یک سایت آرایشی با نام MakeupAlley استفاده می‌کند. این مجموعه حاوی ۱۱۰۰ دیدگاه مثبت و ۱۱۰۰ دیدگاه منفی است. این مرجع ویژگی‌های پایه

او به نمایش می‌گذارند، پیدا کند. در صورت اضافه شدن تعامل، کاربر سریع‌تر می‌تواند اطلاعات مورد نیاز خود را بازیابی کند. همچنین سیستم می‌تواند ابهام‌هایی را که در درخواست کاربر رخ می‌دهند را با پرسش سوال‌هایی برطرف کند [8]. سیستم‌های پرسش و پاسخ تعاملی و سیستم‌های گفتگو^{۲۰} نمونه سیستم‌هایی هستند که با استفاده از مکالمه، اطلاعات مورد نیاز کاربر را فراهم می‌کنند. به طور کلی سیستم‌های گفتگو را می‌توان به دو نوع هدفمند و بدون هدف دسته‌بندی کرد [فاطی]. در سیستم‌های گفتگو هدفمند، کاربر با هدف یافتن اطلاعات خاصی با سیستم گفتگو می‌کند. سیستم‌های پرسش و پاسخ تعاملی نمونه‌ای از این دسته هستند. سیستم‌های گفتگو بدون هدف، اطلاعات خاصی را برای کاربر فراهم نمی‌کنند. اصولاً این نوع سیستم‌ها تا حدودی بدون انسجام عمل می‌کنند. چت بات‌ها نمونه‌ای از این نوع سیستم‌ها هستند [خودم].

تمرکز اصلی این تحقیق بر روی دسته‌بندی قطبی پاسخ‌های یک سیستم پرسش و پاسخ تعاملی است. این مقاله نقطه شروعی برای در نظر گرفتن سابقه پاسخ‌ها در دسته‌بندی قطبی است. یکی از مزایای استفاده از تاریخچه این است که می‌توان پاسخ را پیش‌بینی کرد و در صورت دسته‌بندی صحیح، می‌توان از گفته‌های نامربوط بعدی جلوگیری کرد. این کار موجب کوتاه شدن طول مکالمه و جلب رضایت کاربر می‌شود [14].

تاکنون روش‌های مختلفی برای دسته‌بندی قطبی به کار رفته‌اند. مرجع [1] دسته‌بندی قطبی را بر روی دیدگاه‌های خودرو انجام داده و از دو مجموعه داده COAE2008s (۲۷۳۹ دیدگاه) و اسناد خودرو (۱۰۰۶ دیدگاه) استفاده کرده است. بهترین دقت توسط ماشین بردار پشتیبان، ۸۶٫۶۱٪ برای COAE2008s و ۸۲٫۸۰٪ برای اسناد خودرو بدست آمده است. این مرجع از نسبت جداساز فیشر به عنوان روشی برای انتخاب ویژگی‌های مناسب از فرکانس کلمه استفاده کرده است.

مرجع [3,5,6] از دیدگاه‌های فیلم استفاده می‌کنند. این مجموعه داده حاوی ۱۰۰۰ نظر مثبت و ۱۰۰۰ نظر منفی است. مرجع [3] روش‌های کاهش ویژگی بهره اطلاعات، نرخ بهره^{۲۱}، مربع کای^{۲۲} و باور F^{۲۳} را بر روی دسته‌بندی‌های k-نزدیکترین همسایه^{۲۴}، بیزین ساده^{۲۵}، وینو^{۲۶}، بیشترین آنتروپی^{۲۷}، درخت تصمیم^{۲۸}، ادابوست^{۲۹} و ماشین بردار پشتیبان^{۳۰} آزمایش می‌کند. بهترین دقت دسته‌بندی توسط بیزین ساده و نرخ بهره بدست آمد که این مقدار برابر ۹۰٫۹٪ است. مرجع [5] علاوه بر این دیدگاه‌ها از مجموعه‌ای دیگر با عنوان BLOGS06 استفاده می‌کند. این مرجع، مدل‌های مختلف فرکانس سند و فرکانس معکوس سند را بر روی ماشین بردار پشتیبان آزمایش می‌کند. بهترین دقت برای BLOGS06 ۸۴٫۸۰٪ و برای دیدگاه‌های فیلم ۹۶٫۴٪ است. مرجع [6] n-گرم‌های ۲، ۱ و ۳ را توسط بیزین ساده، بیشترین آنتروپی، گرادیان نزولی تصادفی^{۳۱} و ماشین بردار پشتیبان آزمایش می‌کند. بهترین دقت توسط کلمه‌ها و با استفاده از ماشین بردار پشتیبان بدست آمد که این مقدار برابر ۸۸٫۴۸٪ است.

مرجع [7] نیز از n-گرم‌ها به عنوان مجموعه ویژگی استفاده می‌کند و بر روی دو مجموعه توییت انتخابات سال ۲۰۱۲ آمریکا و ۲۰۱۳ هند، آزمایش‌ها را انجام می‌دهد. روش‌های دسته‌بندی ماشین

پیش‌پردازش استفاده شده شامل سه مرحله می‌باشد. در مرحله اول تمام علائم نگارشی، حذف می‌گردند؛ در مرحله دوم کلیه حروف بزرگ را به کوچک تبدیل می‌شوند و در مرحله سوم تنها یک تکرار از حروف تکراری مجاور در کلمه‌ها نمایش داده می‌شود.

۲-۲- استخراج ویژگی

انتخاب بهترین ویژگی‌ها در تعداد مناسب یک موضوع مطرح در طبقه‌بندی قطبی مبتنی بر یادگیری ماشین است. مطالعه‌های مختلف سعی در حل مسائل مربوط به استخراج ویژگی‌های پیچیده و مقایسه روش‌های انتخاب ویژگی دارد [1,16]. بیشتر تحقیق‌های موجود بر روی ویژگی‌های ساده، از جمله فرکانس کلمه‌ها [1,3,5,7,11]، حضور کلمه‌ها [3,5,6]، n -گرم مبتنی بر حروف [8]، n -گرم‌های مبتنی بر کلمه [4,6,7,8,17]، یا ترکیبی از ویژگی‌های فوق متمرکز شده‌اند. این ویژگی‌ها نسبت به نوع کلاس مستقل هستند. در این تحقیق تاثیر ویژگی‌های زیر در دسته‌بندی، آزمایش می‌گردد.

فرکانس عبارت: فرکانس عبارت تعداد تکرار یک عبارت را در یک متن محاسبه می‌کند. این ویژگی به طور گسترده در پردازش متن استفاده می‌گردد. عبارت به صورت‌های مختلفی مانند n -گرم‌ها وجود دارد. n -گرم‌ها ($n=1$) و کیسه‌ای از کلمه‌ها^{۲۳}، به کلمه‌ها نسبت داده می‌شوند.

حضور عبارت: حضور عبارت در واقع یک بردار دودویی از عبارت‌ها است که حضور و عدم حضور آن‌ها را در متن نمایش می‌دهد. مرجع [16] با بررسی ویژگی‌های فرکانس و حضور عبارت بر روی توییت‌ها به این نتیجه رسید که حضور عبارت‌ها برای دسته‌بندی قطبی موثرتر هستند.

Tf-Idf: این ویژگی یکی از توابع وزن دهی در بازیابی اطلاعات است که فرمول اصلی آن توسط رابطه (۱) محاسبه می‌شود:

$$Tf - Idf(t) = tf(t) \times \log \frac{N}{df(t)} \quad (1)$$

که در آن $tf(t)$ فرکانس عبارت t را بدست می‌آورد؛ N تعداد کل متن‌ها است و $df(t)$ تعداد متن‌هایی که حاوی عبارت t هستند. **ویژگی احتمالی:** یک عبارت ممکن است در هر سه دسته مثبت، منفی و خنثی پدیدار شود. هیچ یک از ویژگی‌های ذکر شده تاثیر کلاس را بر روی عبارت در نظر نمی‌گیرند و تنها بر روی عبارت تمرکز دارند. ویژگی احتمالی میزان مفید بودن یک عبارت را در یک دسته خاص محاسبه می‌کند. رابطه (۲) نحوه محاسبه این احتمال را نشان می‌دهد.

$$prob(t) = \frac{\max(tf_{pos}(t), tf_{neg}(t), tf_{neu}(t))}{tf_{pos}(t) + tf_{neg}(t) + tf_{neu}(t)} \quad (2)$$

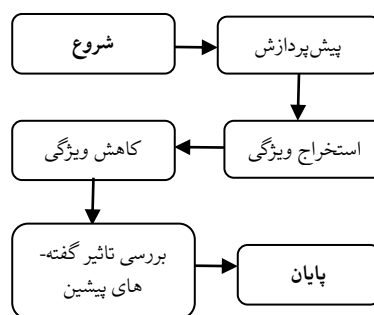
که در آن $prob(t)$ احتمال رخداد و وزن عبارت t را نشان می‌دهد و $tf_i(t)$ فرکانس عبارت t را در دسته i محاسبه می‌کند. به طور کلی این ویژگی بیشترین احتمال رخداد یک عبارت در یک دسته را به عنوان وزن در نظر می‌گیرد. عبارتی که در هر سه دسته به صورت

(کلمه‌ها و کلمه‌های نفی)، ویژگی‌های واژگانی عمومی، ویژگی‌های واژگانی محدود به دامنه و ویژگی‌های الگوی معنایی را توسط ماشین بردار پشتیبان مقایسه می‌کند. بهترین نتیجه توسط ویژگی‌های الگوی معنایی با معیار $f - 82.31\%$ بدست آمد.

ساختار مقاله به شرح زیر است:

در بخش دوم روش پیشنهادی در چند مرحله توصیف می‌گردد؛ در بخش سوم آزمایش‌ها و نتایج بررسی می‌شوند و در پایان در بخش چهارم نتیجه‌گیری انجام می‌شود.

۲- روش پیشنهادی



شکل ۱: مراحل روش پیشنهادی

۲-۱- پیش پردازش

پاسخ‌هایی که کاربران به صورت مکتوب ارائه می‌کنند، معمولاً مانند هر داده خام دیگر حاوی ناخالصی‌هایی است. این ناخالصی‌ها به صورت غلط‌های املایی، حالت‌های اختصار و نامتداول هستند. روش‌های مختلفی برای پیش‌پردازش و برطرف کردن این ناخالصی‌ها وجود دارد. ریشه‌یابی^{۲۰} و حذف کلمه‌های اضافی^{۲۱} نمونه‌ای از این روش‌ها هستند.

یک کلمه ممکن است به شکل‌های مختلف ظاهر شود. ریشه‌یابی کلمه‌ها را به حالت پایه آنها تبدیل می‌کند. بعضی از کلمه‌ها تکرار زیادی در متن‌ها دارند ولی با حذف آن‌ها معنی متن تغییر نمی‌کند. از جمله کلمه‌های با، از، در، و، یا و ... را می‌توان به عنوان نمونه‌هایی از کلمه‌های اضافی در متن‌های فارسی اشاره کرد. امروزه کلمه‌های اضافی برای زبان‌های مختلف به صورت آماده وجود دارد. اما این مجموعه، حاوی کلمه‌های نفی مانند نه، نیست، نمی‌شود و ... هستند و نمی‌توان از آن‌ها برای دسته‌بندی قطبی استفاده کرد.

مرجع [15] شش روش متداول برای استخراج کلمه‌های اضافی را بر روی دسته‌بندی قطبی آزمایش کرده است. این مجموعه کلمه‌ها، از متن‌های کوتاه و خلوت^{۲۲} توییت‌ها استخراج شده‌اند و در کل به این نتیجه رسیده‌اند که حذف کلمه‌های اضافی در متن‌های خلوت و کوتاه موجب کاهش دقت می‌شود. در این مقاله از روش‌های پیش‌پردازش ذکر شده استفاده نمی‌شود. مزیت این کار این است که در آینده می‌توان این کار را بر روی مکالمه‌های زبان دیگر استفاده کرد.

۳- آزمایش‌ها و نتایج

پایگاه داده مورد استفاده شامل ۱۴۴۳ مکالمه انگلیسی است که از یک سیستم پرسش و پاسخ تعاملی استخراج شدند. این مکالمه‌ها در مجموع شامل ۶۹۳۰ پاسخ مثبت، ۲۳۱۹ پاسخ منفی و ۱۱۲۲ پاسخ خنثی هستند. برای توازن در تعداد داده، ۱۱۲۲ پاسخ به صورت تصادفی از هر دسته انتخاب شدند. از بین این پاسخ‌ها برای هر دسته ۷۴۸ پاسخ برای آموزش و ۳۷۴ پاسخ برای آزمون در نظر گرفته شده‌اند. این پاسخ‌ها بگونه‌ای هستند که در یک مکالمه هر دو برچسب آموزش و آزمون وجود نداشته باشد. زیرا در صورت تداخل، هنگامی که تاثیر سابقه مکالمه بررسی می‌شود، از داده‌هایی استفاده می‌گردد که قبلاً آموزش دیده‌اند. در کل ۹۳۷ مکالمه برای آموزش و ۵۰۷ مکالمه برای آزمون استفاده شدند.

جدول (۱) و (۲) نتایج دسته‌بندی‌های مختلف را بر روی ویژگی‌های ذکر شده در بخش (۲-۲) نشان می‌دهد. این ویژگی‌ها بر روی n-گرم‌های ۱ و ۲ آزمایش شده‌اند. اندازه مجموعه ویژگی‌های مربوط به کلمه‌ها برابر ۲۲۰۵ و برای n-گرم‌ها (n=۲) برابر ۹۵۸۲ عبارت است. این ابعاد بزرگ سرعت روش‌های ماشین بردار پشتیبان و درخت تصمیم را به شدت کاهش می‌دهد.

جدول ۱: بررسی درصد دقت وزندهی‌های مختلف کلمه‌ها بر روی داده‌های آزمون

فرکانس	حضور	Tf-Idf	احتمالی
۷۵,۹۳	۷۵,۲۲	۷۵,۲۲	۷۶,۰۲
۷۵,۹۳	۷۵,۲۲	۶۷,۰۲	۷۷,۵۴
۶۴,۹۷	۶۶,۱۳	۶۶,۱۳	۶۴,۹۷
۷۰,۶۷	۷۰,۴۹	۷۰,۴۹	۷۰,۶۷
۷۴,۴۲	۷۶,۳۸	۷۶,۳۸	۷۴,۵۰

جدول ۲: بررسی درصد دقت وزندهی‌های مختلف n-گرم (n=۲) بر روی داده‌های آزمون

فرکانس	حضور	Tf-Idf	احتمالی
۷۲,۳۷	۷۲,۳۷	۷۲,۳۷	۷۲,۳۷
۷۲,۷۲	۷۲,۳۷	۶۵,۴۱	۷۲,۹۰
۶۳,۰۱	۶۱,۹۴	۶۱,۹۴	۶۳,۰۱
۵۶,۶۸	۵۶,۶۸	۵۶,۶۸	۵۶,۶۸
۷۱,۶۳	۵۷,۴۸	۵۷,۴۸	۵۷,۴۸

با بررسی جدول‌های (۱) و (۲)، مشاهده می‌شود که ویژگی احتمالی تقریباً در بیشتر روش‌های کلاس‌بندی عملکردی بهتری نسبت به سایر ویژگی‌ها داشته است. در هر دو جدول بهترین دقت را ماشین بردار پشتیبان با استفاده از ویژگی احتمالی بدست آورد. در جدول (۲) بیزین ساده و الگوریتم کیسه‌ای با استفاده از ویژگی‌های حضور کلمه و tf-idf نتایج بهتری نسبت به ویژگی احتمالی داشتند. نتایج جدول (۲) تا حدودی به هم نزدیک هستند. ویژگی فرکانس n-گرم‌ها (n=۲) نسبت به سایر ویژگی‌ها تقریباً عملکرد بهتری داشته است. به خصوص در الگوریتم کیسه‌ای نسبت به سایر ویژگی‌ها ۱۴٪ بهبود داشته است. مشابه [11,13]، با مقایسه جدول (۱) و (۲) این نتیجه حاصل می‌

یکنواخت تکرار شده ارزش کمتری نسبت به عبارتی که در یک دسته خاص به تعداد زیاد تکرار شده است، دارد.

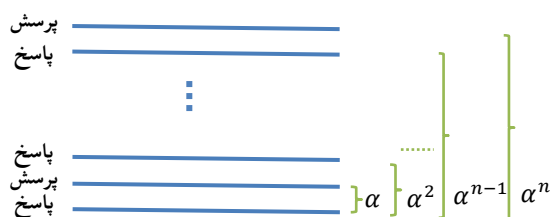
۲-۳- روش‌های کاهش ویژگی

یک مسئله مهم در دسته‌بندی قطبی، ابعاد ویژگی‌ها است که مانع بزرگی را در اجرای بسیاری از الگوریتم‌های پیشرفته بوجود می‌آورد. هدف از روش‌های کاهش ویژگی، کاهش مجموعه ویژگی اصلی است که با حذف برخی از ویژگی‌های غیرمرتبط، موجب بهبود دقت دسته‌بندی و کاهش زمان اجرای الگوریتم‌های یادگیری استفاده می‌شوند. در برخی از تحقیق‌ها، روش‌های مختلفی نظیر بهره‌بردار اطلاعات^{۲۴} [3]، آزمون شباهت لگاریتمی^{۲۵} [2,12]، نسبت جداسازی فیشر^{۲۶} [1,3] و مربع کای [3,12] برای استخراج تعداد ویژگی مناسب به کار گرفته شده‌اند. در این تحقیق از بهره‌بردار اطلاعات، نرخ بهره، مربع کای برای کاهش ویژگی استفاده می‌شود.

۲-۴- بررسی تاثیر گفته‌های پیشین

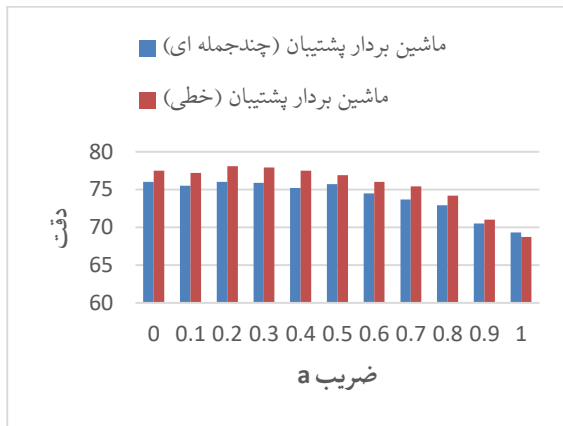
بررسی تاثیر حافظه در مکالمه‌ها برای شناسایی اهداف اجتماعی و تعاملات انسانی در بسیاری از برنامه‌های کاربردی از جمله خلاصه‌سازی [18]، سیستم‌های گفتگو [19-21] و بازی‌های بیست سوالی [7] مهم است. در این بخش گذشته پاسخ‌ها به دو صورت وابستگی زمانی و ساختاری بررسی می‌گردد.

تاثیر وابستگی زمانی: در این روش با استفاده از یک ضریب، گفته‌های قبلی مکالمه تاثیر داده می‌شوند. گفته‌هایی که به پاسخ نزدیک‌ترند ارزش بیشتری خواهند داشت. بنابراین با استفاده از این ضریب به آنها ارزش بیشتری داده می‌شود. شکل (۲) مثالی از نحوه استفاده گفته‌های قبلی را در پاسخ فعلی نمایش می‌دهد.



شکل ۲: نحوه استفاده از ضریب α

تاثیر وابستگی ساختاری: در این روش با استفاده از نوع گفتگوهای قبلی تاثیر مکالمه بررسی می‌گردد. به طور کلی گفته‌های قبلی از نوع پاسخ مثبت، منفی، خنثی و سوال هستند. مرجع [21] از این روش برای دسته‌بندی عواطف در مکالمه استفاده کرده است. این مرجع اندازه تاریخچه را ثابت در نظر می‌گیرد. به عنوان مثال دو گفتگوی قبلی را در دسته‌بندی تاثیر می‌دهد. در این مقاله اندازه تاریخچه متفاوت است. زیرا پاسخ‌ها در هر مکانی ممکن است وجود داشته باشند.



شکل ۳: بررسی تاثیر ضریب‌های مختلف a بر روی دقت دسته‌بندی جدول (۴) نتیجه دسته‌بندی قطبی را با استفاده از وابستگی ساختاری نشان می‌دهد.

جدول ۴: بررسی درصد دقت روش‌های دسته‌بندی با استفاده از ساختار گفته‌های قبلی

احتمالی	Tf-Idf	حضور	فرکانس	
۸۱,۱۰	۸۱,۱۰	۸۱,۱۹	۸۱,۱۹	ماشین بردار پشتیبان (چندجمله‌ای)
۸۲,۶۲	۶۹,۶۰	۸۰,۶۵	۷۱,۵۶	ماشین بردار پشتیبان (خطی)
۷۳,۱۷	۷۳,۱۷	۷۳,۱۷	۷۳,۱۷	بیزین ساده
۷۶,۰۲	۷۶,۰۲	۷۶,۰۲	۷۶,۰۲	درخت تصمیم
۷۸,۶۰	۷۷,۸۹	۷۷,۸۹	۷۷,۸۹	الگوریتم bagging (درخت تصمیم)

همانطور که در جدول (۴) مشاهده می‌شود نتایج دسته‌بندی‌های یکسان بر روی مجموعه ویژگی‌های مختلف، به هم نزدیک هستند. تنها در ماشین بردار پشتیبان با هسته خطی مقادارها با هم متفاوت است. بهترین نتیجه توسط همین دسته‌بندی برای ویژگی احتمالی به دست آمد. در کل نتایج نسبت به دسته‌بندی پاسخ‌ها بدون در نظر گرفتن گذشته، بهبود یافته است.

۴- نتیجه‌گیری

این مقاله دسته‌بندی قطبی را بر روی پاسخ‌های یک سیستم پرسش و پاسخ تعاملی بررسی می‌کند. برای این منظور از ویژگی‌های مشهور آماری استفاده می‌شود و همچنین ویژگی جدیدی با نام ویژگی احتمالی پیشنهاد می‌شود که با توجه به قرارگیری عبارت‌ها در دسته‌های مثبت، منفی و خنثی آنها را وزن‌دهی می‌کند. تمام ویژگی‌های مذکور توسط روش‌های دسته‌بندی باناظر مورد آزمایش قرار می‌گیرند. برای کاهش تعداد ویژگی‌ها، روش‌های استخراج ویژگی به کار می‌روند. با بررسی گذشته پاسخ به دو صورت وابستگی زمانی و وابستگی ساختاری می‌توان به این نتیجه رسید که با استفاده از نوع پاسخ‌های پیشین می‌توان پاسخ فعلی را پیش‌بینی کرد. به عبارت دیگر، در صورتی که بسیاری از پاسخ‌های گذشته مربوط به یک دسته خاص باشند، پیش‌بینی می‌شود پاسخ فعلی نیز مربوط به همان دسته است.

مراجع

- [1] Wang, Suge, Deyu Li, Xiaolei Song, Yingjie Wei, and Hongxia Li. "A feature selection method based on improved fisher's

شود که ویژگی‌های مربوط به کلمه‌ها دقت بهتری نسبت به n -گرم-ها ($n=2$) در همه کلاس‌بندی‌ها داشته‌اند.

در جدول (۳) سه روش مشهور استخراج ویژگی بهره اطلاعات، نرخ بهره و مربع کای بر روی ویژگی احتمالی بررسی شدند. این روش‌ها به هر ویژگی یک امتیاز می‌دهند. بنابراین با استفاده از این روش‌ها علاوه بر کاهش تعداد ویژگی‌ها، می‌توان عبارت‌های مهمی را که در دسته‌بندی قطبی موثر هستند را پیدا کرد. روش‌های کاهش ویژگی مبتنی بر تجزیه بردارهای ویژه^{۲۷} مانند تجزیه و تحلیل مؤلفه اصلی^{۲۸} تنها بر روی کاهش ویژگی تمرکز می‌کنند و عبارت‌های مهم را پیدا نمی‌کنند. بنابراین در این تحقیق از این روش‌ها استفاده نشده است. با توجه به اینکه n -گرم‌ها ($n=2$) در دسته‌بندی قطبی پاسخ‌ها به خوبی کلمه‌ها عمل نکردند، از بررسی روش‌های استخراج ویژگی بر روی آنها صرف نظر گردید.

جدول ۳: بررسی درصد دقت روش‌های دسته‌بندی با استفاده از روش‌های استخراج ویژگی بر روی مجموعه ویژگی احتمالی

مربع کای	نرخ بهره	بهره اطلاعات	
۷۶,۲۹	۷۶,۱۱	۷۶,۱۱	ماشین بردار پشتیبان (چندجمله‌ای)
۷۵,۴۹	۷۵,۵۷	۷۵,۵۷	ماشین بردار پشتیبان (خطی)
۶۷,۶۴	۶۷,۷۳	۶۷,۷۳	بیزین ساده
۷۳,۳۵	۷۳,۳۵	۷۳,۳۵	درخت تصمیم
۷۵,۳۱	۷۴,۰۶	۷۴,۰۶	الگوریتم bagging (درخت تصمیم)

با مقایسه جدول (۳) و نتایج دسته‌بندی ویژگی احتمالی در جدول (۱) مشاهده می‌شود که نتایج به صورت جزئی بهبود یافتند. به غیر از ماشین بردار پشتیبان با هسته خطی که حدود ۲٪ کاهش یافت. به طور کلی مربع کای نسبت به دو روش دیگر دقت بالاتری داشته است. تعداد کلمه‌های کل در مجموعه آموزش برابر ۲۲۰۵ است که با اعمال روش‌های استخراج ویژگی به ۱۷۳ کلمه کاهش یافت. دو روش بهره اطلاعات و نرخ بهره، نتایج یکسانی را تولید می‌کنند زیرا کلمه‌های مشابه با ترتیب متفاوتی را استخراج می‌کنند. این سه روش به کلمه‌ها رتبه‌ای بین صفر و یک می‌دهند و آنها را به ترتیب نزولی مرتب می‌کنند. بعد از ترتیب ۱۷۳ در هر سه روش، رتبه باقی کلمه‌ها صفر می‌شود. بنابراین از مجموعه ویژگی حذف شدند.

به علت دقت بالای ویژگی احتمالی در دسته‌بندی قطبی پاسخ‌ها، از این ویژگی برای بررسی وابستگی زمانی پاسخ‌ها استفاده شد. ضریب مورد نظر باید مقداری بین صفر و یک داشته باشد تا بتواند تاثیر گفته‌های با فاصله بیشتر را در پاسخ فعلی کم کند. همچنین بتواند گفته‌های با فاصله کم را بیشتر تاثیر دهد. نمودار شکل (۳) تاثیر ضریب‌های مختلف را بر روی دقت مکالمه نشان می‌دهد. همانطور که مشاهده می‌شود، استفاده از محتوای گفتگو قبلی تاثیر قابل توجهی در بهبود دقت دسته‌بندی ندارد.

- techniques." In Computational and Business Intelligence (ISCB), 2013 International Symposium on, pp. 178-181. IEEE, 2013.
- [19] Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. "Thumbs up?: sentiment classification using machine learning techniques." In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pp. 79-86. Association for Computational Linguistics, 2002.
- [20] Gamon, Michael. "Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis." In *Proceedings of the 20th international conference on Computational Linguistics*, p. 841. Association for Computational Linguistics, 2002.
- [21] Murray, Gabriel, Giuseppe Carenini, and Raymond Ng. "Generating and validating abstracts of meeting conversations: a user study." In *Proceedings of the 6th International Natural Language Generation Conference*, pp. 105-113. Association for Computational Linguistics, 2010.
- [22] Tavafi, Maryam, Yashar Mehdad, Shafiq R. Joty, Giuseppe Carenini, and Raymond T. Ng. "Dialogue Act Recognition in Synchronous and Asynchronous Conversations." In *SIGDIAL Conference*, pp. 117-121. 2013.
- [23] Kim, Su Nam, Lawrence Cavedon, and Timothy Baldwin. "Classifying dialogue acts in one-on-one live chats." In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 862-871. Association for Computational Linguistics, 2010.
- [24] Herzig, Jonathan, Guy Feigenblat, Michal Shmueli-Scheuer, David Konopnicki, Anat Rafraeli, Daniel Altman, and David Spivak. "Classifying Emotions in Customer Support Dialogues in Social Media." In *SIGDIAL Conference*, pp. 64-73. 2016.
- discriminant ratio for text sentiment classification." *Expert Systems with Applications* 38, no. 7 (2011): 8696-8702.
- [2] Vinodhini, G., and R. M. Chandrasekaran. "Sentiment analysis and opinion mining: a survey." *International Journal of Advanced Research in Computer Science and Software Engineering* 2, no. 6 (2012): 282-292.
- [3] Sharma, Anuj, and Shubhamoy Dey. "A comparative study of feature selection and machine learning techniques for sentiment analysis." In *Proceedings of the 2012 ACM research in applied computation symposium*, pp. 1-7. ACM, 2012.
- [4] Pozzi, Federico Alberto, Elisabetta Fersini, Enza Messina, and Bing Liu. *Sentiment Analysis in Social Networks*. Morgan Kaufmann, 2016.
- [5] Tellez, Eric S., Sabino Miranda-Jiménez, Mario Graff, Daniela Moctezuma, Ranyart R. Suárez, and Oscar S. Siordia. "A Simple Approach to Multilingual Polarity Classification in Twitter." *Pattern Recognition Letters* (2017).
- [6] Braun, Daniel, Adrian Hernandez-Mendez, Florian Matthes, and Manfred Langen. "Evaluating Natural Language Understanding Services for Conversational Question Answering Systems." In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pp. 174-185. 2017.
- [7] Konstantinova, Natalia, and Constantine Orasan. "Interactive question answering." *Emerging Applications of Natural Language Processing: Concepts and New Research* (2012): 149-169.
- [8] Kamal, Ahmad, and Muhammad Abulais. "Statistical features identification for sentiment analysis using machine learning techniques." In *Computational and Business Intelligence (ISCB)*, 2013 International Symposium on, pp. 178-181. IEEE, 2013.
- [9] Paltoglou, Georgios, and Mike Thelwall. "A study of information retrieval weighting schemes for sentiment analysis." In *Proceedings of the 48th annual meeting of the association for computational linguistics*, pp. 1386-1395. Association for Computational Linguistics, 2010.
- [10] Tripathy, Abinash, Ankit Agrawal, and Santanu Kumar Rath. "Classification of sentiment reviews using n-gram machine learning approach." *Expert Systems with Applications* 57 (2016): 117-126. Zięba, Maciej, Jakub M. Tomczak, and Krzysztof Brzostowski. "Selecting right questions with Restricted Boltzmann Machines." In *Progress in Systems Engineering*, pp. 227-232. Springer, Cham, 2015.
- [11] Anjaria, Malhar, and Ram Mohana Reddy Guddeti. "Influence factor based opinion mining of Twitter data using supervised learning." In *Communication Systems and Networks (COMSNETS)*, 2014 Sixth International Conference on, pp. 1-8. IEEE, 2014.
- [12] Deriu, Jan, Aurelien Lucchi, Valeria De Luca, Aliaksei Severyn, Simon Müller, Mark Cieliebak, Thomas Hofmann, and Martin Jaggi. "Leveraging Large Amounts of Weakly Supervised Data for Multi-Language Sentiment Classification." In *Proceedings of the 26th International Conference on World Wide Web*, pp. 1045-1052. International World Wide Web Conferences Steering Committee, 2017.
- [13] Popescu, Marius, Cristian Grozea, and Radu Tudor Ionescu. "HASKER: An efficient algorithm for string kernels. Application to polarity classification in various languages." *Procedia Computer Science* 112 (2017): 1755-1763.
- [14] Ali, Farman, Kyung-Sup Kwak, and Yong-Gi Kim. "Opinion mining based on fuzzy domain ontology and Support Vector Machine: A proposal to automate online review classification." *Applied Soft Computing* 47 (2016): 235-250.
- [15] Tripathy, Abinash, Ankit Agrawal, and Santanu Kumar Rath. "Classification of sentiment reviews using n-gram machine learning approach." *Expert Systems with Applications* 57 (2016): 117-126.
- [16] Gamon, Michael. "Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis." In *Proceedings of the 20th international conference on Computational Linguistics*, p. 841. Association for Computational Linguistics, 2002.
- [17] Tan, Sang-Sang, and Jin-Cheon Na. "Mining Semantic Patterns for Sentiment Analysis of Product Reviews." In *International Conference on Theory and Practice of Digital Libraries*, pp. 382-393. Springer, Cham, 2017.
- [18] Kamal, Ahmad, and Muhammad Abulais. "Statistical features identification for sentiment analysis using machine learning

زیر نویس‌ها

¹ Sentiment Analysis

² Polarity classification

³ Dialogue systems

⁴ Gain Ratio(GR)

⁵ CHI statistics

⁶ Relief-F

⁷ K-nearest neighbor (KNN)

⁸ Naïve Bayes (NB)

⁹ Winnow

¹⁰ Maximum Entropy (ME)

¹¹ Decision Tree

¹² Adaboost

¹³ support vector machine (SVM)

¹⁴ Stochastic Gradient Descent

¹⁵ Deep Learning

¹⁶ String Kernels

¹⁷ Fuzzy Domain Ontology (FDO)

¹⁸ Pointwise Mutual Information(PMI)

¹⁹ Mutual Information(MI)

²⁰ Stemming

²¹ StopWords

²² Sparse

²³ Bag of words(BWO)

²⁴ Information Gain(IG)

²⁵ Log likelihood tests

²⁶ Fisher's discriminant ratio

²⁷ eigenvalue decomposition

²⁸ Principal component analysis (PCA)



ارزیابی امنیت لایه فیزیکی در شبکه‌های مخابرات بی سیم با استفاده از معیارهای ظرفیت کانال امن و احتمال قطع امن

پریناز کاوندی^۱، محمد ترابی^۲

^۱ دانشجوی کارشناسی ارشد مهندسی برق، دانشگاه شهید بهشتی، تهران
pa.kavandi@mail.sbu.ac.ir

^۲ عضو هیات علمی دانشکده مهندسی برق، دانشگاه شهید بهشتی، تهران
m_torabi@sbu.ac.ir

چکیده

در این مقاله به بررسی و تحلیل کارایی یک سیستم مخابراتی بی سیم از لحاظ امنیت لایه فیزیکی شبکه مخابراتی خواهیم پرداخت. بدین منظور معیارهای ظرفیت امن و احتمال قطع کانال محوشدگی رایلی در یک سیستم مخابراتی ارسال مستقیم در نظر گرفته می‌شود. با استفاده از روابط ریاضی مربوط به ارزیابی کارایی و امنیت ارسال و دریافت داده در سیستم مورد نظر می‌پردازیم. همچنین با استفاده از شبیه‌سازی‌های کامپیوتری مونت-کارلو صحت و دقت روابط ریاضی و تحلیل‌های انجام شده را بررسی می‌کنیم. سپس برای بهبود سیستم مخابراتی ارسال مستقیم، سیستم مخابراتی مشارکتی رله‌ای را در نظر می‌گیریم. در ادامه یک سیستم مخابراتی مشارکتی رله‌ای تقویت و ارسال را در نظر گرفته و روابط احتمال ظرفیت امن مثبت و احتمال قطع سیستم مورد نظر را بدست می‌آوریم. با استفاده از نتایج شبیه‌سازی کامپیوتری، مشاهده می‌شود که سیستم مخابراتی مشارکتی رله‌ای تقویت و ارسال قابلیت ایجاد امنیت لایه فیزیکی بهتری را نسبت به سیستم مخابراتی متداول بدون استفاده از رله دارد.

کلمات کلیدی

امنیت لایه فیزیکی، ظرفیت امن، احتمال قطع

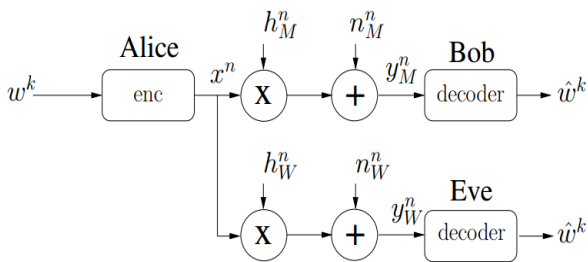
معنی که امنیت آن‌ها هرچند دراز مدت باشد ولی نهایتاً محدود بوده و به اتمام می‌رسد. علاوه بر مشکلاتی که در مدیریت کلید و پیچیدگی منابع محاسباتی برای تولید و توزیع کلیدهای محرمانه وجود دارد، لایه دیگری را نیز در طراحی شبکه‌ها می‌افزاید.

روشی دیگر در برقراری امنیت در یک شبکه مخابراتی، به کارگیری تکنیک‌های امنیت لایه فیزیکی [7-11] می‌باشد. هدف اصلی این تکنیک‌ها، افزایش ظرفیت مسیرهای اصلی ارسال اطلاعات و همزمان کاهش کیفیت مسیرهای شنودی است. امنیت مبتنی بر لایه فیزیکی امنیت بدون قید و شرط یا نظریه اطلاعات است که از امنیت محاسباتی در رمزنگاری مطلوب‌تر می‌باشد.

امنیت نظریه اطلاعات که بر اساس نظریه محرمانگی کامل Shannon (perfect secrecy) [1] ساخته شده، به وسیله Wyner [2] و بعدها توسط Csizsar و Korner [3] مطرح شد، کسانی که در مقاله-

۱- مقدمه

با توجه به ماهیت پخشی کانال سیستم‌های مخابراتی بی سیم، کاربرانی که در محدوده پوشش اطلاعات هستند، قادر به شنود اطلاعات می‌باشند. بنابراین مسائل مربوط به امنیت در شبکه‌های بی سیم اهمیت فزاینده‌ای دارد. روش‌های مرسوم تضمین امنیت پیام‌ها در سامانه‌های مخابراتی شامل استفاده از الگوریتم‌های رمزنگاری است که در لایه‌های بالاتر شبکه انجام می‌شود. میزان امنیت و کارایی این روش‌ها را میزان پیچیدگی محاسباتی با محاسبه تعداد عملیات لازم برای بدست آوردن کلید به کار برده شده در روش رمزنگاری تعیین می‌کند. این روش‌ها اگرچه امنیت را تضمین می‌کنند ولی وابسته به زمان می‌باشند به این



شکل ۱: مدل سیستم اول

که $h_M(i)$ ضریب محوشدگی مختلط متغیر با زمان و $n_M(i)$ نویز گوسی مختلط متقارن چرخشی میانگین صفر است. ضریب $h_M(i)$ ، همچنین اشاره به اطلاعات وضعیت کانال (CSI) دارد، که از خروجی کانال مستقل است و با توجه به توزیع احتمال $p(h_M)$ که برای محوشدگی رایلی، گوسی مختلط میانگین صفر است؛ فرض می‌شود که حاصل، مستقل با توزیع یکسان (i.i.d.) است. اگر محوشدگی را ایستادگوسوسی فرض کنیم ضرایب محوشدگی برای تمام کانال‌های کاربردی (یا برای تمامی زمان‌ها) ثابت است.

$$h_M(i) = h_M, \forall i \quad (2)$$

شخص سوم Eve که دشمن نامیده می‌شود، با مشاهده خروجی کانال، قادر به شنود سیگنال‌های فرستاده شده توسط Alice است.

$$y_W(i) = h_W(i)x(i) + n_W(i) \quad (3)$$

در یک کانال محوشدگی رایلی مستقل با ضرایب محوشدگی ایستادگوسی $h_W(i) = h_W, \forall i$ و $n_W(i)$ نویز گوسی مختلط متقارن چرخشی میانگین صفر، کانال دارای توان محدود است.

$$\frac{1}{n} \sum_{i=1}^n E[|X(i)|^2] \leq P \quad (4)$$

که E نشان دهنده امید ریاضی است و P با متوسط توان سیگنال ارسالی برابر است. علاوه بر این، توان نویز کانال اصلی و کانال شنودگر به ترتیب با N_M و N_W نشان داده می‌شود. لحظه‌ای گیرنده Bob برابر است با

$$\gamma_M(i) = P|h_M(i)|^2/N_M = P|h_M|^2/N_M = \gamma_M \quad (5)$$

و SNR متوسط گیرنده Bob مطابق زیر است.

$$\bar{\gamma}_M(i) = PE[|h_M(i)|^2]/N_M = PE[|h_M|^2]/N_M = \bar{\gamma}_M \quad (6)$$

همچنین SNR لحظه‌ای در گیرنده شنودگر برابر است با

$$\gamma_W(i) = P|h_W(i)|^2/N_W = P|h_W|^2/N_W = \gamma_W \quad (7)$$

و SNR متوسط گیرنده شنودگر مطابق زیر است.

$$\bar{\gamma}_W(i) = PE[|h_W(i)|^2]/N_W = PE[|h_W|^2]/N_W = \bar{\gamma}_W \quad (8)$$

همچنین در نظر گرفتن تابع چگالی احتمال γ_M و γ_W کاربردی خواهد بود. به دلیل اینکه ضرایب محوشدگی کانال (h)، متغیرهای تصادفی

های ابتدایی ثابت کردند کدهای کانالی وجود دارند که توان ضمانت درجه‌ی تعیین شده‌ای از محرمانگی داده و ارسال خطاها را دارد. در کانال شنودگر پیشنهاد شده توسط Wyner، دو کاربر مجاز به وسیله یک کانال اصلی ارتباط دارند و شنودگر به نسخه‌ی نزول یافته اطلاعاتی که به گیرنده می‌رسد دسترسی دارد. در مرجع [4] نشان داده شد که اگر کانال اصلی و کانال شنودگر، کانال نویز گوسی سفید جمع‌شونده (AWGN) باشند و کانال شنودگر ظرفیت کم‌تری نسبت به کانال اصلی داشته باشد، ظرفیت امن (بیشینه نرخ ارسال داده که شنودگر نتواند هیچ اطلاعاتی را شنود کند) برابر با تفاوت میان ظرفیت دو کانال است. بنابراین ارتباط محرمانه امکان پذیر نیست مگر اینکه کانال گوسی اصلی نسبت سیگنال به نویز (SNR) بهتری نسبت به کانال گوسی شنودگر داشته باشد.

در این مقاله به بررسی کارایی یک سیستم مخابراتی بی‌سیم از لحاظ امنیت لایه فیزیکی شبکه مخابراتی می‌پردازیم. بدین منظور معیارهای ظرفیت امن و احتمال قطع کانال محوشدگی رایلی را در یک سیستم مخابراتی ارسال مستقیم در نظر می‌گیریم. با استفاده از روابط ریاضی به ارزیابی کارایی و امنیت ارسال و دریافت داده در سیستم مورد نظر می‌پردازیم. همچنین با استفاده از شبیه‌سازی‌های کامپیوتری مونت-کارلو صحت و دقت روابط ریاضی و تحلیل‌های انجام شده بررسی می‌شود. سپس برای بهبود سیستم مخابراتی ارسال مستقیم، سیستم مخابراتی مشارکتی رله‌ای را در نظر می‌گیریم. نهایتاً با توجه به روابط احتمال ظرفیت امن مثبت و احتمال قطع کانال محوشدگی رایلی در سیستم مخابراتی مشارکتی رله‌ای، به شبیه‌سازی کامپیوتری احتمال قطع و تحلیل سیستم مورد نظر می‌پردازیم.

ساختار این مقاله به صورت زیر است. در قسمت ۲ به مدل سیستم مربوط به شبکه مخابراتی ارسال مستقیم پرداخته می‌شود. در قسمت ۳ به تحلیل عملکرد این سیستم از لحاظ ظرفیت امن و احتمال قطع پرداخته می‌شود و همچنین نتایج عددی و شبیه‌سازی این سیستم بررسی می‌شود. در قسمت ۴ مدل سیستم مربوط به شبکه مخابراتی مشارکتی رله‌ای تقویت و ارسال (AF-Relaying) و نتایج شبیه‌سازی این سیستم بررسی می‌شود و نهایتاً در قسمت ۵ نتیجه‌گیری ارائه خواهد شد.

۲- مدل سیستم اول (ارسال مستقیم)

مسئله‌ای که در شکل ۱ نشان داده شده را در نظر بگیرید. یک کاربر مجاز به نام Alice می‌خواهد پیام w را به کاربر دیگری به نام Bob ارسال کند. بلوک پیام w^k به کدواژه $x^n = [x(1), \dots, x(i), \dots, x(n)]$ تبدیل شده تا در طول یک کانال محوشدگی رایلی زمان گسسته (کانال اصلی) با خروجی ارسال می‌شود [12].

$$y_M(i) = h_M(i)x(i) + n_M(i) \quad (1)$$

$$C_s = C_M - C_W \quad (۱۱)$$

که ظرفیت کانال اصلی از رابطه‌ی زیر به دست می‌آید.

$$C_M = \frac{1}{2} \log_2 \left(1 + \frac{P}{N_M} \right) \quad (۱۲)$$

و ظرفیت کانال شنودگر از رابطه‌ی زیر به دست می‌آید.

$$C_W = \frac{1}{2} \log_2 \left(1 + \frac{P}{N_W} \right) \quad (۱۳)$$

فرض کنید دو کانال اصلی و شنود کانال نویز گوسی سفید جمع-شونده مختلط باشند. نمونه‌های دریافتی و ارسال، مختلط هستند و نویز جمع‌شونده هر دو کانال نیز گوسی مختلط متقارن چرخشی میانگین صفر است. توان ورودی مختلط X طبق رابطه‌ی $\frac{1}{n} \sum_{i=1}^n E[|X(i)|^2] \leq P$ محدود است. به دلیل اینکه یک بار به کار بردن کانال نویز گوسی سفید جمع‌شونده مختلط معادل دو بار به کار بردن کانال نویز گوسی سفید جمع‌شونده مقدار حقیقی [5] در نظر گرفته می‌شود؛ بنابراین ظرفیت امن کانال شنود مختلط در هر بعد مختلط مطابق زیر خواهد بود.

$$C_s = \log_2 \left(1 + \frac{P}{N_M} \right) - \log_2 \left(1 + \frac{P}{N_W} \right) \quad (۱۴)$$

درگام نهایی، ضرایب محوشدگی مختلط برای کانال اصلی و کانال شنود، در نظر گرفته می‌شوند. در مورد کانال ایستای گوسی h_M و h_W تصادفی هستند اما همیشه مقدار ثابتی دارند، می‌توان کانال اصلی (با محوشدگی) را به عنوان یک کانال نویز گوسی سفید جمع‌شونده مختلط با SNR و ظرفیتی به صورت زیر در نظر گرفت.

$$\gamma_M = P |h_M|^2 / N_M \quad (۱۵)$$

$$C_M = \log_2 \left(1 + |h_M|^2 \frac{P}{N_M} \right) \quad (۱۶)$$

به طور مشابه، ظرفیت کانال شنودگر به صورت زیر می‌باشد.

$$C_W = \log_2 \left(1 + |h_W|^2 \frac{P}{N_W} \right) \quad (۱۷)$$

و SNR این کانال نیز به فرم زیر است.

$$\gamma_W = P |h_W|^2 / N_W \quad (۱۸)$$

بنابراین با توجه به رابطه‌ی (۱۱) و غیر منفی بودن ظرفیت کانال، می‌توان ظرفیت امن را برای شرایط محوشدگی ایستای گوسی به صورت زیر نوشت:

$$C_s = \begin{cases} \log_2(1 + \gamma_M) - \log_2(1 + \gamma_W) & \text{if } \gamma_M > \gamma_W \\ 0 & \text{if } \gamma_M \leq \gamma_W \end{cases} \quad (۱۹)$$

وجود ظرفیت امن بین Alice و Bob را در نظر می‌گیریم. همانطور که توضیح داده شد، برای موارد خاص محوشدگی، کانال اصلی (از Alice به Bob) و کانال شنود (از Alice به Eve) می‌تواند به صورت

گوسی مختلط میانگین صفر [5] و SNR لحظه‌ای $\gamma \alpha |h|^2$ است، بنابراین γ بصورت نمایی توزیع شده است.

$$f(\gamma_M) = \frac{1}{\gamma_W} \exp\left(-\frac{\gamma_M}{\gamma_W}\right), \quad \gamma_M > 0 \quad (۹)$$

و

$$f(\gamma_W) = \frac{1}{\gamma_W} \exp\left(-\frac{\gamma_W}{\gamma_W}\right). \quad \gamma_W > 0 \quad (۱۰)$$

نرخ ارسال بین Alice و Bob برابر با $R = H(w^k)/n$ است و احتمال خطا به صورت $P_e^k = \text{Prob}(w^k \neq \hat{w}^k)$ تعریف می‌شود که \hat{w}^k اشاره به تخمین Bob از پیام فرستاده شده دارد. $\text{Prob}(\cdot)$ احتمال متغیر مورد نظر است. ما نه تنها علاقه به افزایش نرخ ارسال بین Alice و Bob داریم بلکه می‌خواهیم عدم اطمینان Eve راجع به حدود ω بیشتر شود.

همچنین نرخ ابهام به صورت $\Delta = H(w^k | Y_W^n) / H(w^k)$ است [6].

(R', d') قابل دسترسی است اگر برای تمامی $\epsilon > 0$ یک جفت کدگذار (encoder) و کدگشا (decoder) به طوری که $R \geq R' - \epsilon$ ، $d' - \epsilon \leq P_e^k \leq \epsilon$ و $\Delta \geq d' - \epsilon$ باشد، وجود داشته باشد. هدف، مشخص کردن ظرفیت امن C_s برحسب بیشینه نرخ ارسال R در $\Delta = 1$ می‌باشد.

فرض شده که Alice و Bob، اطلاعات وضعیت کانال درست کانال اصلی را دارند، اما هیچ دانشی راجع به اطلاعات وضعیت کانال شنود ندارند. در عوض Eve، اطلاعات وضعیت کانال شنود را دارد.

۳- بررسی کارایی و عملکرد سیستم اول

در این قسمت به بررسی کارایی و عملکرد سیستم اول می‌پردازیم. بدین منظور ابتدا ظرفیت امن را بررسی کرده و سپس احتمال قطع این سیستم را مورد ارزیابی قرار خواهیم داد.

۳-۱- ظرفیت امن کانال محوشدگی رایلی

در این قسمت ظرفیت امن یک کانال محوشدگی رایلی شبه ساکن تعیین می‌شود. در ابتدا ظرفیت امن بر حسب ضریب محوشدگی محاسبه می‌شود. سپس درباره وجود ظرفیت امن (کاملاً مثبت) به صورت کلی بحث می‌شود و در قسمت بعدی یک دیدگاه نهایی برای مشخص کردن احتمال قطع ارائه می‌دهیم.

ابتدا ظرفیت امن یک جفت کانال محوشدگی ایستای گوسی با نویز مختلط و ضرایب محوشدگی مختلط را به دست می‌آوریم.

به این منظور، نتایج [4] را برای کانال گوسی مقدار حقیقی شنود بازخوانی می‌کنیم، جایی که Alice و Bob در طول یک کانال با نویز گوسی سفید جمع‌شونده مقدار حقیقی استاندارد به همراه توان نویز $N_M > N_W$ ارتباط برقرار می‌کنند و شنود Eve به وسیله N_W که $N_M > N_W$ است ناهنجار می‌شود. گیرنده Eve، SNR کمتری نسبت به Bob دارد. توان نیز با توجه به $\frac{1}{n} \sum_{i=1}^n E[|X(i)|^2] \leq P$ محدود است. برای این مثال: ظرفیت امن به وسیله رابطه‌ی زیر تعریف می‌شود.

$$\begin{aligned}
&= \int_0^\infty \int_{\gamma_W}^{2^{R_s}(1+\gamma_W)-1} f(\gamma_M, \gamma_W | \gamma_M > \gamma_W) d\gamma_W d\gamma_M \\
&= \int_0^\infty \int_{\gamma_W}^{2^{R_s}(1+\gamma_W)-1} \frac{f(\gamma_M) f(\gamma_W)}{\text{Prob}(\gamma_M > \gamma_W)} d\gamma_W d\gamma_M \\
&= 1 - \frac{\bar{\gamma}_M + \bar{\gamma}_W}{\bar{\gamma}_M + 2^{R_s} \bar{\gamma}_W} \exp\left(-\frac{2^{R_s}-1}{\bar{\gamma}_M}\right) \quad (25)
\end{aligned}$$

و از آنجا که $R_s > 0$

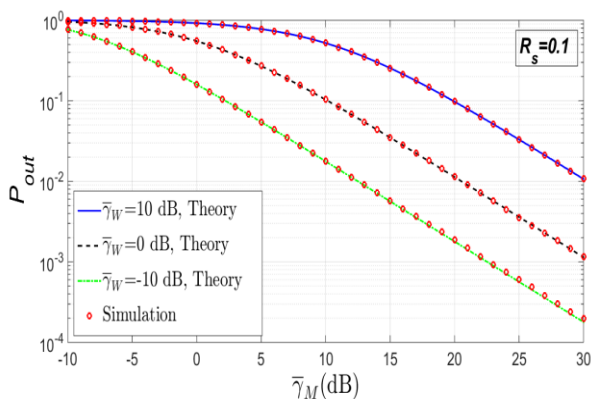
$$\text{Prob}(C_s < R_s | \gamma_M \leq \gamma_W) = 1 \quad (26)$$

با ترکیب دو معادله‌ی قبل رابطه زیر بدست می‌آید [12]

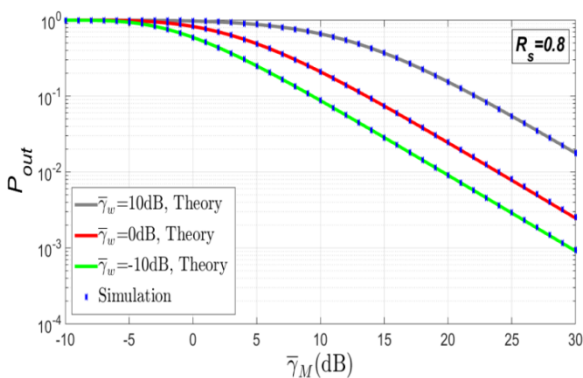
$$P_{out}(R_s) = 1 - \frac{\bar{\gamma}_M}{\bar{\gamma}_M + 2^{R_s} \bar{\gamma}_W} \exp\left(-\frac{2^{R_s}-1}{\bar{\gamma}_M}\right) \quad (27)$$

۳-۳- نتایج تئوری و شبیه‌سازی سیستم اول

در این قسمت به بررسی کارایی و عملکرد سیستم اول با استفاده از نتایج عددی و شبیه‌سازی مونت کارلو می‌پردازیم. بدین منظور نتایج ظرفیت امن را بررسی کرده و همچنین نتایج احتمال قطع سیستم مورد نظر را مورد ارزیابی قرار خواهیم داد.



شکل ۲: نمودار P_{out} بر حسب $\bar{\gamma}_M$ برای $R_s = 0.1$



شکل ۳: نمودار P_{out} بر حسب $\bar{\gamma}_M$ برای $R_s = 0.8$

کانال نویز گوسی سفید جمع‌شونده مختلط با SNR های به ترتیب γ_M و γ_W در نظر گرفته شود. علاوه بر این از رابطه‌ی (۱۹) نتیجه می‌شود ظرفیت امن مثبت است زمانی که $\gamma_M > \gamma_W$ باشد و زمانی که $\gamma_M \leq \gamma_W$ باشد صفر است. با توجه به استقلال کانال اصلی و کانال شنود و علم به اینکه متغیرهای تصادفی γ_M و γ_W به صورت نمایی به ترتیب با توابع چگالی احتمال روابط (۹) و (۱۰) توزیع شده‌اند، بنابراین احتمال وجود یک ظرفیت امن غیرصفر را می‌توان به صورت زیر نوشت [12]

$$\text{Prob}(C_s > 0) = \text{Prob}(\gamma_M > \gamma_W)$$

$$\begin{aligned}
&= \int_0^\infty \int_0^{\gamma_M} f(\gamma_M, \gamma_W) d\gamma_W d\gamma_M \\
&= \int_0^\infty \int_0^{\gamma_M} f(\gamma_M) f(\gamma_W) d\gamma_W d\gamma_M \\
&= \frac{\bar{\gamma}_M}{\bar{\gamma}_M + \bar{\gamma}_W} \quad (20)
\end{aligned}$$

$\text{Prob}(\cdot)$ احتمال متغیر مورد نظر می‌باشد.

۳-۲- احتمال قطع کانال محوشدگی رایلی

در این قسمت یک دیدگاه نهایی برای مشخص کردن احتمال قطع ارائه می‌دهیم.

$$P_{out}(R_s) = \text{Prob}(C_s < R_s) \quad (21)$$

طبق تعریف احتمال قطع، برابر است با احتمال این که ظرفیت امن لحظه‌ای کمتر از نرخ امن مورد نظر $R_s > 0$ باشد. تعریف احتمال قطع، اهمیت کاربردی دارد. بنا بر قاعده احتمال کل داریم:

$$\begin{aligned}
P_{out}(R_s) &= \text{Prob}(C_s < R_s | \gamma_M > \gamma_W) \text{Prob}(\gamma_M > \gamma_W) \\
&+ \text{Prob}(C_s < R_s | \gamma_M \leq \gamma_W) \text{Prob}(\gamma_M \leq \gamma_W) \quad (22)
\end{aligned}$$

با توجه به رابطه‌ی (۲۰) می‌دانیم که

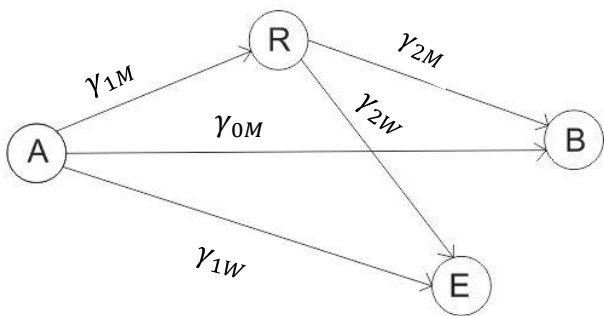
$$\text{Prob}(\gamma_M > \gamma_W) = \frac{\bar{\gamma}_M}{\bar{\gamma}_M + \bar{\gamma}_W} \quad (23)$$

در نتیجه داریم:

$$\text{Prob}(\gamma_M \leq \gamma_W) = 1 - \text{Prob}(\gamma_M > \gamma_W) = \frac{\bar{\gamma}_W}{\bar{\gamma}_M + \bar{\gamma}_W} \quad (24)$$

از طرفی داریم:

$$\begin{aligned}
&\text{Prob}(C_s < R_s | \gamma_M > \gamma_W) \\
&= \text{Prob}(\log_2(1 + \gamma_M) - \log_2(1 + \gamma_W) < R_s | \gamma_M > \gamma_W)
\end{aligned}$$



شکل ۶: مدل سیستم مشارکتی رله ای

رابطه سیگنال به نویز در گیرنده اصلی B (Bob) که سیستم مشارکتی رله ای است به صورت زیر بیان می شود [13]:

$$\gamma_M = \gamma_{0M} + \frac{\gamma_{1M}\gamma_{2M}}{1+\gamma_{1M}+\gamma_{2M}} \quad (28)$$

هم چنین رابطه سیگنال به نویز در گیرنده فرعی E (Eve) به صورت زیر بیان می شود:

$$\gamma_W = \gamma_{1W} + \frac{\gamma_{1M}\gamma_{2W}}{1+\gamma_{1M}+\gamma_{2W}} \quad (29)$$

در نهایت با توجه به رابطه (۱۹) و (۲۸) و (۲۹) خواهیم داشت:

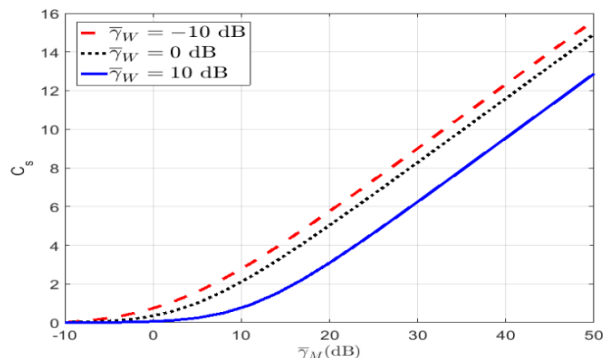
$$C_s = \begin{cases} 0.5 * [\log_2 \left(1 + \gamma_{0M} + \frac{\gamma_{1M}\gamma_{2M}}{1+\gamma_{1M}+\gamma_{2M}} \right) - \log_2 \left(1 + \gamma_{1W} + \frac{\gamma_{1M}\gamma_{2W}}{1+\gamma_{1M}+\gamma_{2W}} \right)] & \text{if } \gamma_M > \gamma_W \\ 0 & \text{if } \gamma_M \leq \gamma_W \end{cases} \quad (30)$$

که ضریب ۰,۵ به این دلیل است که با استفاده از رله، ارسال در دو مرحله انجام می شود. بنابراین خواهیم داشت:

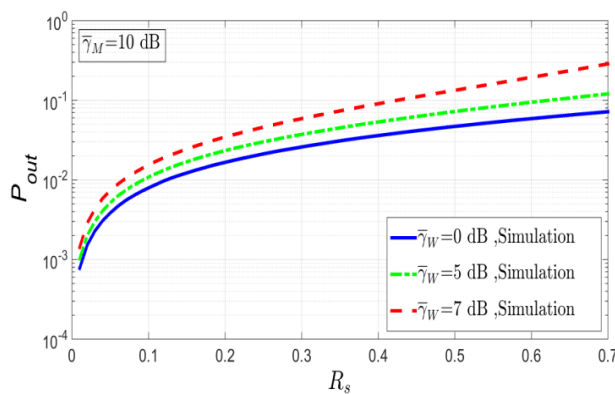
$$P(C_s > 0) = Prob \left(0.5 \log_2 \left(\frac{1+\gamma_{0M} + \frac{\gamma_{1M}\gamma_{2M}}{1+\gamma_{1M}+\gamma_{2M}}}{1+\gamma_{1W} + \frac{\gamma_{1M}\gamma_{2W}}{1+\gamma_{1M}+\gamma_{2W}}} \right) > 0 \right) \quad (31)$$

$$P_{out}(R_s) = Prob \left(0.5 \log_2 \left(\frac{1+\gamma_{0M} + \frac{\gamma_{1M}\gamma_{2M}}{1+\gamma_{1M}+\gamma_{2M}}}{1+\gamma_{1W} + \frac{\gamma_{1M}\gamma_{2W}}{1+\gamma_{1M}+\gamma_{2W}}} \right) \leq R_s \right) \quad (32)$$

از نمودارهای شکل ۲ و ۳ نتیجه می شود با افزایش احتمال قطع کاهش می یابد. همچنین با کاهش $\bar{\gamma}_W$ احتمال قطع کاهش می یابد. از مقایسه دو نمودار نیز نتیجه می شود هرچه نرخ امن عدد کوچکتری باشد احتمال قطع کاهش می یابد. نتایج عددی به دست آمده برای اعتباردهی با شبیه سازی مونت کارلو مقایسه شده و کاملاً برهم منطبق بودند که نشانگر دقت و صحت نتایج است.



شکل ۴: نمودار C_s بر حسب $\bar{\gamma}_M$



شکل ۵: نمودار P_{out} بر حسب R_s برای $\bar{\gamma}_M = 10 \text{ dB}$

از نمودار شکل ۴ نتیجه می شود با افزایش $\bar{\gamma}_M$ ظرفیت امن افزایش می یابد. همچنین با کاهش $\bar{\gamma}_W$ نیز ظرفیت امن افزایش می یابد. از شکل ۵ نتیجه می شود هرچه نرخ امن بیشتر باشد احتمال قطع افزایش می یابد. همچنین نتیجه می شود که با کاهش $\bar{\gamma}_W$ احتمال قطع کاهش می یابد.

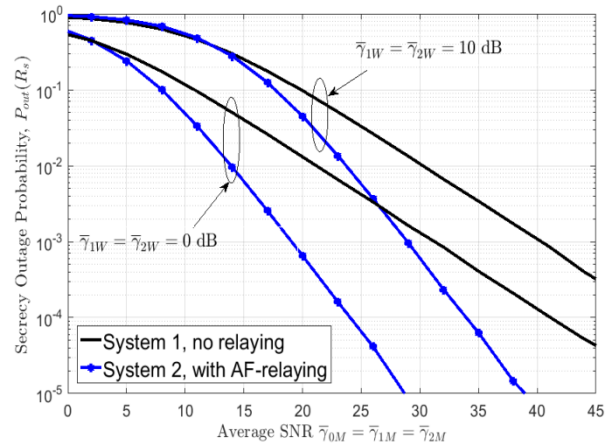
۴- مدل سیستم دوم (مشارکتی رله ای تقویت و ارسال)

در این قسمت ابتدا روابط ظرفیت امن و احتمال ظرفیت امن مثبت و احتمال قطع در کانال های محوشدگی رایلی در یک سیستم مخابراتی با یک رله ای مشارکتی بیان می شود. سپس نتایج شبیه سازی احتمال قطع سیستم مخابراتی با یک رله ای مشارکتی و سیستم مخابراتی ارسال مستقیم در یک نمودار آورده می شود.

- [5] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
- [6] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," *Lecture Notes in Computer Science*, vol. 1807, pp. 351+, 2000.
- [7] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd Edition, Prentice Hall, 2001.
- [8] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733-742, 1993.
- [9] I. Alfred O. Hero, "Secure space-time communication," *IEEE Trans. On Inform. Theory*, vol. 49, no. 12, pp. 3235-3249, December 2003
- [10] A. Thangaraj, S. Dihidar, A. Calderbank, S. W. McLaughlin, and J. Merolla, "Capacity achieving codes for the wiretap channel with applications to quantum key distribution," *CoRR*, vol. cs.IT/0411003, 2004.
- [11] M. A. Thangaraj, S. W. McLaughlin, and J. Merolla, "LDPC-based Gaussian key reconciliation," in *Proc. of the IEEE International Workshop on Information Theory*, Punta del Este, Uruguay, March 2006.
- [12] J. Barros and M. R. D. Rodrigues, "Secrecy Capacity of Wireless Channels," *IEEE Int'l. Symp. Info. Theory*, pp. 356-60, 2006.
- [13] M. Torabi, W. Ajib and D. Haccoun, "Performance Analysis of Amplify-and-Forward Cooperative Networks with Relay Selection Over Rayleigh Fading Channels," *IEEE Vehicular Technology Conference*, Barcelona, Spain, April 2009.

۱-۴ - نتایج شبیه‌سازی سیستم دوم، سیستم مشارکتی رله‌ای تقویت و ارسال

در این قسمت به بررسی عملکرد سیستم دوم با استفاده از شبیه‌سازی مونت کارلو احتمال قطع می‌پردازیم. در شکل های ۷ و ۸ مشاهده شد که در مقایسه با سیستم مخابراتی متداول بدون استفاده از رله، یک سیستم مخابراتی مشارکتی رله‌ای تقویت و ارسال می‌تواند میزان احتمال قطع امن کمتری ایجاد کند.



شکل ۸: نمودار P_{out} بر حسب $\bar{\gamma}_{1M}$ برای $R_s = 0.1$

۵- نتیجه‌گیری

در این مقاله به بررسی کارایی سیستم مخابراتی بی‌سیم از لحاظ امنیت لایه فیزیکی پرداختیم. برای تحلیل سیستم معیارهای ظرفیت امن و احتمال قطع در کانال محوشدگی رایلی در یک سیستم مخابراتی ارسال مستقیم را در نظر گرفتیم و برای آنها روابط ریاضی بدست آوردیم. با استفاده از روابط ریاضی به ارزیابی امنیت ارسال و دریافت داده در سیستم مورد نظر پرداختیم. همچنین با استفاده از شبیه‌سازی‌های کامپیوتری صحت روابط ریاضی بدست آمده را بررسی کردیم. همچنین مشاهده کردیم که با افزایش $\bar{\gamma}_M$ و کاهش $\bar{\gamma}_W$ ظرفیت امن افزایش و احتمال قطع کاهش می‌یابد. در ادامه سیستم مخابراتی مشارکتی رله‌ای تقویت و ارسال را در نظر گرفته و به شبیه‌سازی کامپیوتری احتمال قطع سیستم مورد نظر پرداختیم. مشاهده شد که سیستم مخابراتی مشارکتی رله‌ای تقویت و ارسال قابلیت تامین امنیت لایه فیزیکی بهتری نسبت به سیستم مخابراتی متداول بدون استفاده از رله را دارد.

مراجع

- [1] C.E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. Journal*, vol. 29, pp. 656-715, 1949.
- [2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. Journal*, vol. 54, pp. 1355-1387, 1975.
- [3] I. Csiszar and J. Kerner, "Broadcast channels with confidential messages," *IEEE Trans. on Inform. Theory*, vol. 24, no. 3, pp. 339-348, 1978.
- [4] S. K. Leung-Yan-Cheong and M. E. Hellman, "The gaussian wiretap channel," *IEEE Trans. on Inform. Theory*, vol. 24, no. 4, pp. 451456, July 1978.



تعیین هویت بلادرنگ چهره با استفاده از ویژگی‌های مکان فرکانس چهره در مجموعه‌های بزرگ

خشایار نوریان^۱، حسین خسروی^۲

^۱ دانشجوی رشته هوش مصنوعی و ریاتیک، دانشکده برق و ریاتیک دانشگاه صنعتی شاهرود، شاهرود
k.nourian@shahroodut.ac.ir

^۲ استادیار دانشکده برق و ریاتیک دانشگاه صنعتی شاهرود، شاهرود
HosseinKhosravi@shahroodut.ac.ir

چکیده

بازشناسی چهره یکی از روش‌های مرسوم در تعیین هویت افراد است که کاربردهای زیادی از جمله کنترل و ثبت تردد در سیستم‌های حضور و غیاب، ورود به دستگاه‌های شخصی و شناسایی مجرمین دارد. هدف در بازشناسی چهره، شناسایی هویت فرد از یک تصویر با استفاده از پایگاه داده است. در این مقاله، به ارائه روشی سریع مبتنی بر استخراج ویژگی می‌پردازیم. بعد از آشکارسازی چهره و ویژگی‌های محلی استخراج می‌شود. بر اساس ویژگی محلی، پیش‌پردازش انجام و پارامترهای آن تنظیم می‌شود تا بهترین عملکرد را داشته باشد، سپس با چسباندن ویژگی‌های بهینه، یک بردار ویژگی ساخته می‌شود. با استفاده از کلاسه بند نزدیک‌ترین همسایگی و معیار فاصله منهن به کلاسه‌بندی می‌پردازیم. برای ارزیابی روش پیشنهادی از پایگاه داده FERET که شامل تصاویر حدود ۱۰۰۰ نفر است و برای هر فرد یک نمونه آموزش وجود دارد استفاده کردیم. تمرکز در این مقاله بر چالش‌های تعداد کم نمونه‌های آموزش، تعداد بالای افراد و بازشناسی بلادرنگ می‌باشد. نتایج تجربی روی این پایگاه داده نرخ بازشناسی ۹۵/۹۵٪ را روی مجموعه FB، با زمان ۱۵۸ میلی ثانیه برای هر تصویر مجموعه آزمون دارد.

کلمات کلیدی

بازشناسی چهره، استخراج ویژگی، پایگاه داده حجیم.

پوشیده‌شدن بخشی از چهره، تغییر موقعیت سر، تغییر سن، کم بودن تعداد نمونه‌های آموزش به ازای هر فرد، تعداد زیاد افراد موجود در پایگاه داده و شناسایی بلادرنگ اشاره کرد [۱، ۲].

هر سیستم خودکار بازشناسی چهره ۳ بخش اساسی دارد که در شکل مشاهده می‌شود. اولین بخش، آشکارسازی چهره است. در این بخش، پس از شناسایی و تعیین محل صورت، تصویر چهره از تصویر ورودی جدا می‌شود. در بخش بعدی، اطلاعات مهم و متمایز کننده چهره از ناحیه چهره استخراج می‌شود. در آخرین مرحله، با مقایسه بین بردارهای ویژگی بدست آمده از تصویر ورودی و بردارهای ویژگی ذخیره شده در پایگاه داده با استفاده از کلاسه بند هویت فرد تعیین می‌شود. [۱، ۲].

۱- مقدمه

بازشناسی چهره یکی از روش‌های زیست‌شناختی برای تعیین هویت افراد است. در مسئله بازشناسی چهره، ورودی یک تصویر از چهره شناخته نشده است و هدف تعیین هویت این چهره با استفاده از پایگاه داده است. میزان شباهت بین تصویر داده شده و تصاویر موجود در پایگاه داده محاسبه می‌شود و هویت تصویری که بیشترین شباهت را با تصویر ورودی داشت هویت تصویر مورد آزمایش می‌باشد [۱]. از کاربردهای بازشناسی چهره می‌توان به کنترل و ثبت تردد در سیستم‌های حضور غیاب، کنترل نامحسوس و ایجاد امنیت در بانک‌ها و فرودگاه‌ها، ورود به دستگاه‌های شخصی و شناسایی مجرمین اشاره کرد.

در بازشناسی چهره چالش‌های زیادی وجود دارد که از این چالش‌ها می‌توان به تغییرات روشنایی، تغییرات در حالت چهره،

۲- کارهای مرتبط

تایگمان و همکاران [۴] با ارائه مدل شبکه عصبی عمیق به نام DeepFace به بازشناسی چهره پرداخته‌اند. این شبکه از ۹ لایه کانولوشن تشکیل شده است. برای آموزش این شبکه پایگاه داده‌ای متشکل از ۴ میلیون تصویر از ۴۰۰۰ نفر ساخته شد. این شبکه عصبی روی پایگاه داده LFW نرخ بازشناسی ۹۷,۳۵ درصد را دارد.

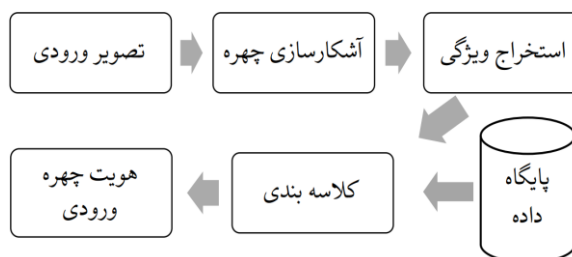
اسکروف و همکاران [۵] روشی به نام FaceNet را معرفی کردند. این روش با استفاده از ساختار شبکه عصبی کانولوشن، نداشت مستقیمی از تصویر به فضای اقلیدسی فشرده انجام می‌دهد که در این فضا فاصله بین دو تصویر همان معیار مشابهت نهایی است. شبکه عصبی کانولوشن معرفی شده در این روش ۶ لایه کانولوشن و حدود ۴,۲ میلیون پارامتر دارد. این روش روی پایگاه داده LFW به نرخ بازشناسی ۹۹,۶۳ درصد رسیده است.

رایت و همکاران [۶] با استفاده از بازنمایی تنک محاسبه شده توسط کمینه سازی L1 به عمل بازشناسی پرداخته‌اند. روش معرفی شده با انجام تنک سازی روی داده‌های ورودی و استفاده از رگرسیون نسبت به تغییرات روشنایی و وجود انسداد عملکرد قابل قبولی داشته است. نرخ بازشناسی این روش روی پایگاه داده AR، ۹۴,۷ درصد می‌باشد. با ۳۰ درصد انسداد نسبت به نبود انسداد ۱,۵ درصد کاهش در نرخ بازشناسی را داشته است.

آهون و همکاران [۷] با استفاده از ویژگی الگوی باینری محلی (LBP)، توصیفگری را برای بازشناسی چهره معرفی کرده‌اند. روش استخراج این توصیفگر به این صورت است که پس از اعمال اپراتور LBP روی ناحیه چهره، تصویر بدست آمده را بلوک بندی می‌کند. از هر بلوک، هیستوگرام الگوهای باینری محاسبه می‌شود و با چسباندن بردارهای ویژگی بلوک‌ها، توصیفگر نهایی ساخته می‌شود. این روش روی پایگاه داده FERET، نرخ بازشناسی ۸۱٪ را دارد.

بخشی و همکاران [۸] روشی برای استخراج یک ویژگی مکان-فرکانسی برای بازشناسی چهره معرفی کرده‌اند. در این روش با اعمال پنجره‌های هم مرکز روی تصویر و استخراج ضرایب بخصوصی از تبدیل فوری و گرفتن میانگین روی مؤلفه‌های فرکانسی به ساخت بردار ویژگی می‌پردازد. نرخ بازشناسی این روش روی زیرمجموعه FB از پایگاه داده FERET، ۹۰,۴ درصد می‌باشد.

وارون و همکاران [۹] بر اساس تبدیل هاف استخراج ویژگی برای بازشناسی را انجام می‌دهند. این روش پس از آشکارسازی چهره و اعمال پیش‌پردازش لبه‌ها را با الگوریتم سوپل پیدا می‌کند. لبه‌ها با عملگر گسترش مورفولوژی پهن تر می‌شود. با بلوک بندی و گرفتن تبدیل هاف از هر بلوک به دنبال پیدا کردن خطوط می‌باشد. نقاطی از فضای هاف که تعداد رای آنها از حد مشخصی بیشتر بود به عنوان ویژگی در نظر گرفته می‌شود. سپس با چسباندن بردارهای ویژگی بلوک‌ها به هم بردار ویژگی نهایی ساخته می‌شود. این روش روی ۳۵ فرد از پایگاه داده FERET دارای نرخ بازشناسی ۸۰,۸۱ درصد می‌باشد.



شکل ۱ پیکربندی کلی سیستم بازشناسی چهره

روش‌های موجود در بازشناسی چهره را می‌توان به ۳ دسته تقسیم بندی کرد [۱-۳].

روش‌های تطبیق سراسری

روش‌های مبتنی بر ویژگی

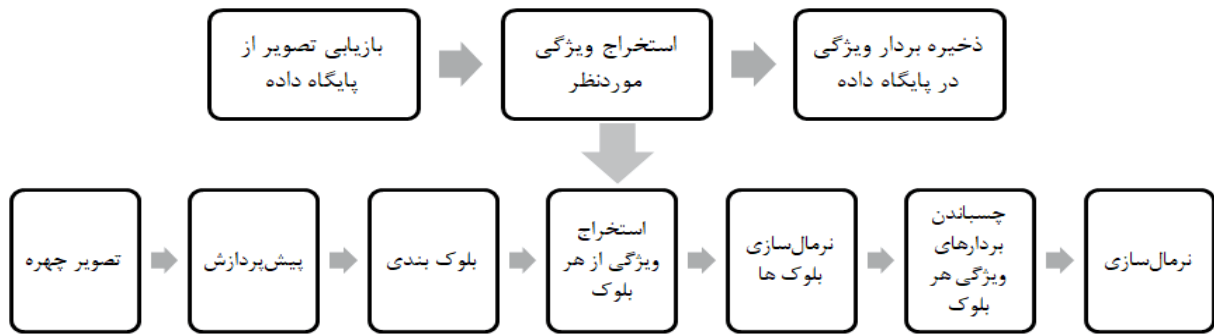
روش‌های ترکیبی

در روش‌های تطبیق سراسری، کل ناحیه چهره مستقیماً برای بازشناسی در نظر گرفته می‌شود سپس با استفاده از روش‌های آماری (تحلیل مولفه اصلی، تحلیل جداساز خطی، تحلیل مولفه مستقل) کلاسه بندی صورت می‌گیرد. این روش‌ها چند عیب عمده دارند، اگر اندازه ناحیه چهره 100×100 پیکسل باشد، حدود ۱۰۰۰۰ پیکسل به عنوان ورودی برای بازشناسی چهره استفاده می‌شود که از نظر محاسباتی و زمانی پیچیدگی بالایی دارند. همچنین تصویر گرفته شده ممکن است تحت تغییرات محل قرار گیری دوربین یا سر، تغییرات حالت چهره و ... باشد پس برای ساختن یک سیستم بازشناسی مقاوم به تعداد زیادی از نمونه‌های آموزشی برای هر کلاس نیاز دارد.

در روش‌های مبتنی بر ویژگی با استخراج ویژگی‌های محلی از ناحیه تصویر علاوه بر کاهش بعد، جمع آوری اطلاعات بهتر صورت می‌گیرد. ویژگی‌های آماری (هیستوگرام، میانگین و انحراف معیار)، ضرایب تبدیل (تبدیل فوری، تبدیل کسینوسی گسسته)، ویژگی‌های جبری و ویژگی‌های ساختاری مثال‌هایی از ویژگی‌های محلی مورد استفاده می‌باشند. خروجی استخراج ویژگی از هر تصویر یک بردار ویژگی است.

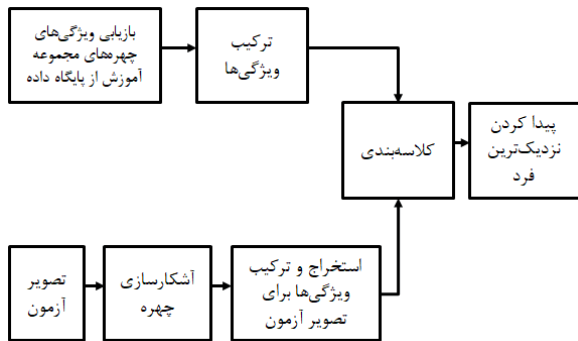
روش‌های ترکیبی رویکردی مثل روش‌های تطبیق سراسری دارند با این تفاوت که بجای استفاده از کل ناحیه چهره به عنوان ورودی از بردار ویژگی استخراج شده از ناحیه تصویر استفاده می‌کنند. این روش‌ها سعی در استفاده از مزیت‌ها در حین کاهش معایب هر دو روش دارند.

در این مقاله با ارائه روش مبتنی بر استخراج ویژگی با تاکید بر چالش‌های تعداد کم نمونه‌های آموزشی، تعداد زیاد افراد موجود در پایگاه داده و بلادرنگ بودن به بازشناسی چهره می‌پردازیم. در بخش ۲ کارهای مرتبط با بازشناسی چهره را مطرح می‌کنیم. در بخش ۳ روش پیشنهادی مطرح می‌شود و در بخش ۴ به گزارش نتیجه پیشنهادی می‌پردازیم. در بخش ۵ به بیان نتیجه گیری و پیشنهاداتی برای کارهای آینده می‌پردازیم.



شکل ۲ فلوجارت استخراج ویژگی پیشنهادی

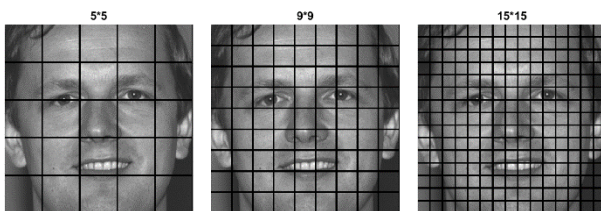
در نهایت با استفاده از کلاسه‌بند نزدیک‌ترین همسایگی، نزدیک‌ترین تصویر از مجموعه آموزش برای تصویر آزمون پیدا می‌کنیم. در شکل فرآیند مرحله آزمون روش پیشنهادی دیده می‌شود.



شکل ۳ فرآیند مرحله آزمون روش پیشنهادی

۳-۱- استخراج ویژگی

در این مرحله پس از بازیابی تصویر چهره از پایگاه داده رابطه‌ای پس از اعمال پیش‌پردازش، تصویر مطابق شکل بلوک بندی می‌شود. از هر بلوک ویژگی موردنظر استخراج و نرمال‌سازی L1 برای هر بلوک انجام می‌شود. سپس با چسباندن بردار ویژگی بلوک‌ها به هم بردار ویژگی نهایی ساخته می‌شود. بردار ویژگی نهایی نیز بعد از نرمال‌سازی در پایگاه داده ذخیره می‌شود.



شکل ۴ نحوه بلوک بندی برای استخراج ویژگی

۳-۱-۱- استخراج ویژگی HOG

نحوه استخراج این ویژگی بصورت زیر است.

- i. از تصویر ورودی گرادیان مرکزی افقی و عمودی گرفته می‌شود.
- ii. تصویر ورودی به $m \times m$ بلوک تقسیم‌بندی می‌شود.
- iii. برای هر پیکسل در هر بلوک مقدار گرادیان و زاویه گرادیان محاسبه می‌شود.

چن و همکاران [۱۰] با بررسی ویژگی‌های با ابعاد بالا در حدود ۱۰۰ هزار بعد به اهمیت تعداد بالای ویژگی پرداخته‌اند. سپس با ارائه یک رگرسیون تنک چرخان، از نظر محاسباتی پردازش را امکان پذیر کرده‌اند. با استفاده از این روش نرخ بازشناسی را بین ۴ تا ۶ درصد نسبت به حالت کم بعد بهبود داده‌اند.

دینگ و همکاران [۱۱] با استخراج الگوهای چندجهتی چندسطحی متقابل از تصویر چهره به بازشناسی چهره پرداخته‌اند. الگوریتم پیشنهادی نسبت به تغییرات شدت روشنایی، تغییر زاویه چهره و تغییر حالت چهره مقاوم گزارش شده است. روش کار این الگوریتم به این صورت است که در ابتدا مشتق اول گوسی از تصویر چهره گرفته می‌شود و تبدیل کسینوسی گسسته بر نتیجه آن اعمال می‌شود. سپس با استفاده از بلوک بندی در دو مرحله ویژگی همسایه‌های اطراف هر پیکسل در جهت‌های ۰، ۴۵، ۹۰، ۱۳۵، ۱۸۰ درجه استخراج می‌شود و برای کلاسه‌بندی از ترکیب دو کلاسه‌بند SVM و KNN استفاده شده است. نرخ بازشناسی این روش روی پایگاه داده FERET، ۹۲٫۸ درصد و روی پایگاه داده CAS-PEAL، ۶۸٫۰۴ درصد می‌باشد.

۳- روش پیشنهادی

با توجه به پیکربندی معرفی شده در شکل اولین مرحله در یک سیستم خودکار بازشناسی چهره آشکار سازی چهره می‌باشد. در روش پیشنهادی از الگوریتم آشکار ساز Viola-jones [۱۲] استفاده شده است. پس از جداسازی ناحیه چهره، چهره بدست آمده به اندازه ۳۶۰×۳۶۰ تغییر اندازه داده می‌شود و سپس آدرس تصویر در پایگاه داده رابطه‌ای ذخیره می‌شود.

دومین مرحله، استخراج ویژگی می‌باشد که در شکل فلوجارت استخراج ویژگی مشاهده می‌شود. سه ویژگی استخراج شده در این مقاله عبارتند از هیستوگرام جهت‌دار گرادیان (HOG)^۱، الگوهای باینری محلی (LBP)^۲ و تبدیل فوریه زمان کوتاه (STFT)^۳. ویژگی HOG اطلاعات گرادیان، ویژگی LBP ساختارهای محلی هر سلول و ویژگی STFT اطلاعات فرکانسی را دارد.

برای هر ویژگی با تغییر پارامترها به دنبال یافتن بهترین عملکرد با استفاده از هر ویژگی بصورت مستقل می‌باشیم. سپس با ترکیب ویژگی‌های بهینه که از بهم چسباندن بردارهای ویژگی بدست می‌آید بهترین ترکیب ویژگی‌ها را بررسی می‌کنیم.

۳-۲- کلاسه‌بندی

بعد از استخراج ویژگی از نمونه‌های آموزشی، هر نمونه یک بردار ویژگی d بعدی می‌شود. با توجه به این‌که کلاسه‌بندی تنها به کمک یک نمونه آموزشی است از کلاسه‌بند نزدیک‌ترین همسایگی استفاده می‌شود. در این کلاسه‌بند با داشتن بردارهای ویژگی مجموعه آموزش و بردار ویژگی نمونه آزمون نزدیک‌ترین کلاس را در بین مجموعه آموزش انتخاب می‌شود. برای انتخاب نزدیک‌ترین نمونه در مجموعه آموزش با معیار فاصله مناسب نیاز داریم. در این مقاله از معیارهای فاصله زیر استفاده می‌کنیم.

- فاصله اقلیدسی:

$$D(a, b) = \left(\sum_{i=1}^d (a_i - b_i)^2 \right)^{\frac{1}{2}} \quad (1)$$

- آماره کای دو:

$$D(a, b) = \sum_{i=1}^d \frac{(a_i - b_i)^2}{a_i + b_i} \quad (2)$$

- فاصله کسینوسی:

$$D(a, b) = 1 - \frac{\sum_{i=1}^d a_i b_i}{\sum_{i=1}^d a_i^2 \sum_{i=1}^d b_i^2} \quad (3)$$

- فاصله منهن:

$$D(a, b) = \sum_{i=1}^d |a_i - b_i| \quad (4)$$

که در این فاصله‌ها a و b ، دو بردار ویژگی d بعدی هستند.

۴- نتایج تجربی

برای ارزیابی از پایگاه داده FERET استفاده می‌کنیم که شامل ۹۹۰ نفر و ۱۳ موقعیت چهره به ازای هر فرد می‌باشد. در این مقاله فقط از تصاویر از روبرو استفاده شده است، این تصاویر شامل تغییرات اندازه، روشنایی، حالت چهره، پوشیده شدن بخشی از چهره توسط عینک، حالت مو و سالخوردگی است. این تصاویر در ۴ زیرمجموعه FA، FB، Dup1 و Dup2 توزیع شده‌اند.

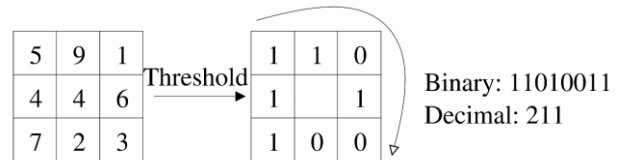
- FA: این تصاویر برای مجموعه آموزش استفاده شده‌اند که به ازای هر نفر یک تصویر وجود دارد.
- FB: بعد از گرفتن تصاویر FA گرفته شده‌اند، تصاویر موجود در این مجموعه فقط تغییرات چهره دارند و به ازای هر فرد یک تصویر در این مجموعه داریم.
- Dup1: تصاویر در این مجموعه تأثیر افزایش سن، داشتن یا نداشتن عینک، تغییر اندازه، تغییر حالت مو و تغییرات روشنایی را بررسی می‌کند. ۷۳۴ تصویر از ۲۴۹ نفر در این مجموعه است.
- Dup2: این تصاویر زیرمجموعه از Dup1 هستند که تصویر گرفته شده بیشتر از ۵۴۰ روز فاصله با تصویر FA دارند. ۲۲۸ تصویر از ۷۵ نفر در این مجموعه قرار دارد.

- iv. زاویه ۰ تا ۳۶۰ درجه به k دسته تقسیم می‌شود. بر اساس زاویه گرادین دسته موردنظر انتخاب شده و به مقدار آن دسته به اندازه مقدار گرادین افزوده می‌شود.
- v. با انجام عمل ۴ برای همه پیکسل‌های بلوک، هیستوگرام جهت‌دار برای سلول ساخته می‌شود. سپس این بردار به‌گونه‌ای نرمال می‌شود که نرم آن یک باشد.
- vi. با چسباندن بردارهای سلول‌های مختلف به هم بردار ویژگی ساخته می‌شود.

۳-۱-۲- استخراج ویژگی LBP

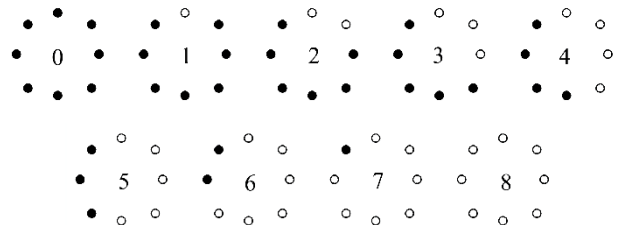
نحوه استخراج این ویژگی بصورت زیر است.

- i. ابتدا روی تصویر چهره ورودی اپراتور LBP اعمال می‌شود. شکل ۵ چگونگی اعمال این اپراتور را روی یک پنجره ۳*۳ نشان می‌دهد.



شکل ۵ چگونگی اعمال اپراتور LBP

- ii. تصویر LBP به $m*m$ بلوک تقسیم می‌شود.
- iii. در هر بلوک هیستوگرام الگوهای یک‌ریخت تشکیل می‌شود. در شکل ۶ این الگوها مشاهده می‌شوند.



شکل ۶ الگوهای یکنواخت برای اپراتور LBP دایره‌ای ۸ همسایگی با شعاع ۱

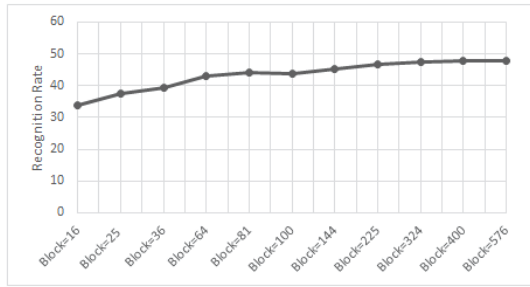
- iv. هیستوگرام هر بلوک نرمال شده و با چسباندن این هیستوگرام‌ها بردار ویژگی نهایی ساخته می‌شود.

۳-۱-۳- استخراج ویژگی STFT

نحوه استخراج این ویژگی بصورت زیر است.

- i. تصویر ورودی به $m*m$ بلوک تقسیم‌بندی می‌شود.
- ii. برای هر ناحیه تبدیل فوریه گرفته می‌شود بطوریکه مولفه‌های کم فرکانس در مرکز باشد. سپس اندازه تبدیل فوریه محاسبه می‌شود.
- iii. با گذاشتن پنجره‌ای مربعی به مرکز اندازه تبدیل فوریه برای هر بلوک، k المان مرکزی را به عنوان بردار ویژگی آن بلوک انتخاب می‌کنیم.
- iv. با چسباندن بردارهای بلوک‌های مختلف به هم بردار ویژگی نهایی ساخته می‌شود.

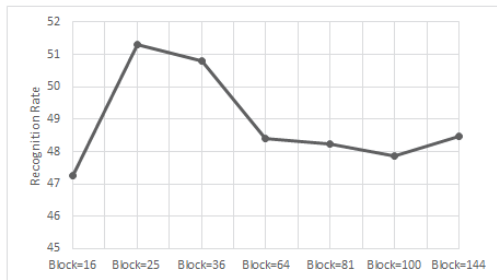
در شکل ۸ تاثیر تغییرات تعداد بلوک‌ها بر نرخ بازشناسی مشاهده می‌شود. با توجه به این شکل تعداد ۳۲۴ بلوک نرخ بازشناسی مطلوبی دارد.



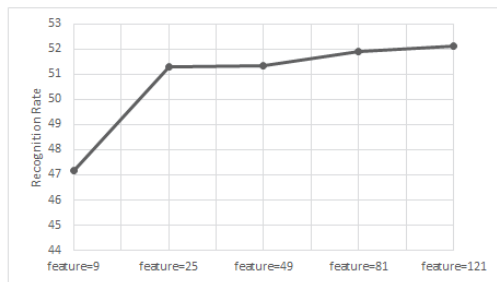
شکل ۸ تاثیر تغییرات تعداد بلوک‌های LBP

۴-۳- تاثیر تغییر پارامترهای STFT

برای ویژگی STFT دو پارامتر تعداد بلوک‌ها و تعداد ویژگی به ازای هر بلوک را داریم. در ابتدا تعداد ویژگی‌ها به ازای هر سلول را ثابت و برابر ۲۵ در نظر می‌گیریم و تغییرات تعداد بلوک‌ها را بررسی می‌کنیم سپس تعداد بلوک‌ها را ثابت و برابر ۲۵ در نظر می‌گیریم و تاثیر تغییرات تعداد ویژگی به ازای هر بلوک را بررسی می‌کنیم.



الف



ب

شکل ۹ تاثیر تغییرات پارامترهای STFT بر نرخ بازشناسی

با توجه به شکل ۹ تعداد بلوک‌ها ۲۵ و تعداد ویژگی‌های فرکانسی به ازای هر بلوک را ۲۵ و ۸۱ ویژگی بر اساس نرخ بازشناسی انتخاب می‌کنیم.

۴-۴- تاثیر معیار فاصله و نرمالسازی

بردارهای ویژگی را در دو حالت با نرمالسازی و بدون نرمالسازی را برای ویژگی‌ها استخراج می‌کنیم. ویژگی HOG با ۱۴۴ بلوک و ۹ دسته هیستوگرام، ویژگی LBP با ۳۲۴ بلوک و ۱۰ ویژگی برای هر بلوک و ویژگی STFT با ۲۵ سلول و ۲۵ ویژگی فرکانسی مورد آزمایش قرار گرفتند.

ارزیابی به این شکل انجام می‌شود که از مجموعه FA به عنوان نمونه‌های آموزش، استخراج ویژگی صورت می‌گیرد و روی ۳ زیرمجموعه FB، Dup1 و Dup2 بصورت جداگانه آزمایش می‌شود. نرخ بازشناسی P برای هر مجموعه بصورت زیر محاسبه می‌شود.

$$p_j = \frac{C_j}{N_j}; j = FB, Dup1, Dup2 \quad (5)$$

که C_j تعداد نمونه‌های با کلاسه‌بندی صحیح در مجموعه Z و N_j تعداد کل نمونه‌های مجموعه Z است.

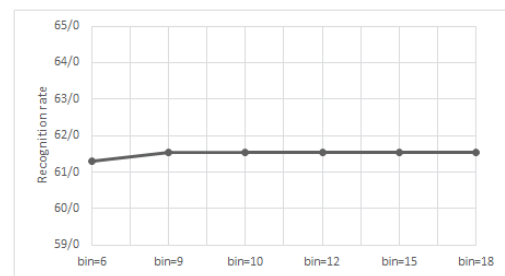
عملکرد میانگین P بصورت زیر محاسبه می‌شود.

$$P = \frac{p_{FB} + p_{Dup1} + p_{Dup2}}{3} \quad (6)$$

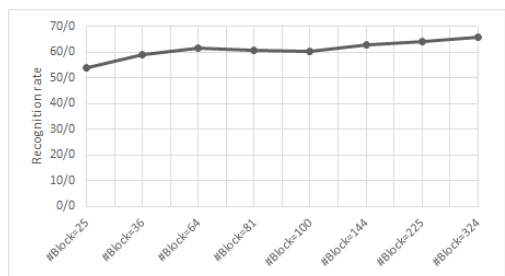
۴-۱- تاثیر تغییر پارامترهای HOG

با توجه به الگوریتم استخراج این ویژگی دو پارامتر تعداد بلوک‌ها و تعداد دسته هیستوگرام را دارد. در ابتدا تعداد بلوک‌ها را ۶۴ در نظر گرفته و تاثیر تغییرات تعداد دسته‌ها را ارزیابی می‌کنیم، سپس تعداد دسته‌های هیستوگرام را ۹ در نظر می‌گیریم و تاثیر تعداد بلوک‌ها را ارزیابی می‌کنیم.

در شکل ۷-الف تاثیر تعداد دسته‌های هیستوگرام و در شکل ۷-ب تاثیر تعداد بلوک‌ها مشاهده می‌شود.



الف



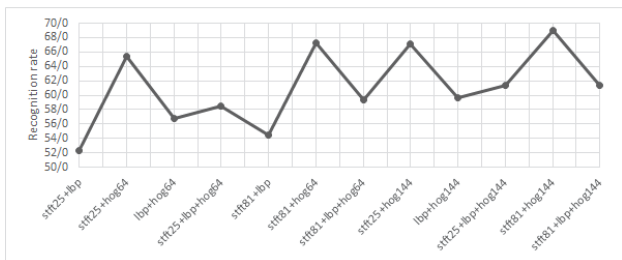
ب

شکل ۷ تاثیر تغییرات پارامترهای HOG بر نرخ بازشناسی

تعداد دسته‌ها ۹ و تعداد بلوک‌ها ۶۴ و ۱۴۴ در نظر می‌گیریم. تعداد بیشتر دسته‌ها تاثیر روی نرخ بازشناسی ندارد. با افزایش تعداد بلوک‌ها اطلاعات جزئی‌تری از اطلاعات گرادیان بدست می‌آید ولی باعث افزایش طول بردار ویژگی می‌شود.

۴-۲- تاثیر تغییر پارامترهای LBP

با توجه به نحوه استخراج ویژگی تعداد ویژگی هر بلوک ثابت و برابر ۱۰ می‌باشد. با تغییر تعداد بلوک‌ها بدنبال بیشترین نرخ بازشناسی هستیم.



شکل ۱۲ نرخ بازشناسی ویژگی‌های ترکیبی

با توجه به این شکل ترکیب دو ویژگی HOG و STFT توانسته نتیجه بهتری داشته باشد.

۴-۶- ویژگی ترکیبی نهایی و مقایسه

در جدول عملکرد الگوریتم پیشنهادی مشاهده می‌شود. که با استفاده از ترکیب ویژگی HOG با ۱۴۴ بلوک و ۹ دسته هیستوگرام و ویژگی STFT با ۸۱ بلوک و ۲۵ ویژگی فرکانسی بهترین عملکرد بدست آمده است.

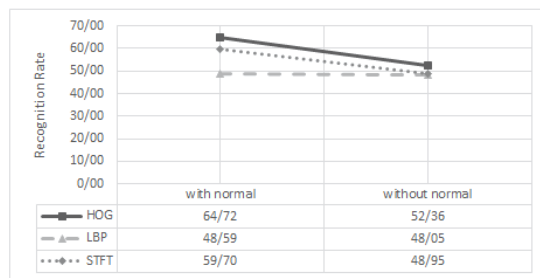
جدول ۲ عملکرد الگوریتم پیشنهادی روی پایگاه داده FERET

میانگین	Dup2	Dup1	FB	طول بردار ویژگی	ویژگی ترکیبی
%۶۷,۱۹	%۴۸,۶۸	%۵۶,۹۵	%۹۵,۹۵	۲۶۰۱	STFT81 + HOG64
%۶۸,۹۴	%۵۳,۳۱	%۵۷,۳۶	%۹۵,۹۵	۳۳۲۱	STFT81 + HOG144

در جدول ۳ مقایسه روش پیشنهادی با سایر روش‌ها از نظر دقت بازشناسی، زمان ارزیابی برای هر تصویر، تعداد کلاس‌ها و تعداد نمونه‌های آموزشی آمده است.

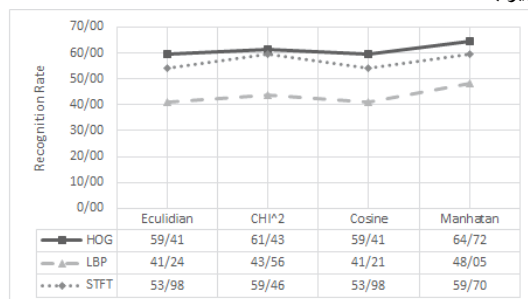
جدول ۳ مقایسه روش پیشنهادی با سایر روش‌ها

زمان آزمون (ms)	نرخ بازشناسی	تعداد نمونه‌های آموزش	تعداد نفرات	پایگاه داده	روش
۸۷۵	FB: %۹۷ Dup1: %۶۶ Dup2: %۶۴	۱	۱۱۹۶	FERET	[۷]
۸۶۰	FB: %۹۰,۴	۱	۹۹۰	از FB FERET	[۸]
۹۰	%۸۰,۸۱	۲۰	۳۵	۳۵ کلاس FERET	[۹]
۱۵۸	FB: %۹۵,۹۵ Dup1: %۵۷,۳۶ Dup2: %۵۳,۳۱	۱	۹۹۰	FERET	روش پیشنهادی



شکل ۱۰ تاثیر نرمالسازی بر نرخ بازشناسی

در شکل ۱۰ مشاهده می‌شود برای ۲ ویژگی HOG و STFT بدون نرمالسازی افت قابل ملاحظه‌ای در نرخ بازشناسی نسبت به حالت با نرمالسازی داریم. در ویژگی LBP افت در نرخ بازشناسی نسبت کم‌تری دارد. چون از ترکیب این ویژگی‌ها استفاده می‌شود باید هر ۳ ویژگی نرمال شوند و سپس با هم ترکیب شوند. همچنین با ۴ معیار فاصله متفاوت آزمایش را تکرار کردیم که نتایج زیر بدست آمد.



شکل ۱۱ تاثیر معیار فاصله بر نرخ بازشناسی

همانطور که در شکل ۱۱ مشاهده می‌شود معیار فاصله منتهی به بیشترین نرخ بازشناسی را برای هر کدام از ویژگی‌ها داشته است.

۴-۵- ترکیب ویژگی‌ها

با تغییر پارامترهای هر ویژگی، ویژگی بهینه را بدست آوردیم در این بخش با ترکیب این ویژگی‌ها بدنبال دستیابی به عملکرد بهتری هستیم. در جدول جزئیات این ویژگی‌ها مشاهده می‌شود.

جدول ۱ جزئیات و عملکرد ویژگی‌های بهینه

ویژگی	HOG	LBP	STFT	تعداد بلوک
تعداد بلوک	۶۴	۳۲۴	۲۵	۲۵
تعداد ویژگی در هر بلوک	۹	۱۰	۲۵	۸۱
طول بردار	۵۷۶	۳۲۴۰	۶۲۵	۲۰۲۵
نام	HOG64	HOG144	STFT25	STFT81
FB	۹۴,۳۳	۹۴,۸۳	۹۳,۷۲	۹۴,۴۳
Dup1	۵۰,۶۸	۵۳,۲۷	۴۸,۰۹	۵۰,۴۱
Dup2	۴۱,۶۷	۴۶,۰۵	۳۷,۲۸	۴۲,۵۴

تمام حالت‌های مختلف ترکیب از ۵ ویژگی فوق را مورد ارزیابی قرار می‌دهیم تا بهترین ترکیب را مشخص کنیم. در شکل ۱۲ نتیجه این آزمایش مشاهده می‌شود.

۵- نتیجه‌گیری

زیرنویس‌ها

- ¹ Histogram of Oriented Gradient
² Local Binary Pattern
³ Short Time Fourier Transform

در این مقاله روشی برای بازشناسی چهره با استفاده از تصاویر از روبروی چهره ارائه شد. در این روش ابتدا از تصاویر آموزش سه ویژگی HOG، LBP و STFT استخراج شد. با تغییر پارامترهای هر ویژگی به دنبال ویژگی بهینه هستیم. از کلاسه‌بند نزدیک ترین همسایگی با معیار فاصله منتهی برای کلاسه‌بندی استفاده کردیم. با امتحان کردن حالات مختلف ترکیب این ویژگی‌های بهینه، به بهترین عملکرد رسیدیم. روش پیشنهادی از نظر سرعت بازشناسی، سریع ترین روش موجود می باشد. که با زمان ۱۵۸ میلی ثانیه عمل بازشناسی را انجام می‌دهد.

برای بهبود عملکرد الگوریتم پیشنهادی میتوان از ویژگی‌های دیگر مثل تبدیل موجک، فیلتر گابور و یا تبدیل کسینوسی گسسته استفاده کرد. همچنین می‌توان بجای به هم چسباندن ویژگی‌ها در ترکیب ویژگی‌ها، برای هر ویژگی یک کلاسه بند در نظر گرفت و با رای گیری نتایج کلاسه‌بندها، نتیجه نهایی تعیین گردد.

مراجع

- [1] Hassaballah, M. and S. Aly. "Face recognition: challenges, achievements and future directions." IET Computer Vision 9(4): 614-626, 2015.
- [2] Zhao, W., R. Chellappa, P. J. Phillips and A. Rosenfeld. "Face recognition: A literature survey." ACM computing surveys (CSUR) 35(4): 399-458, 2003.
- [3] Jafri, R. and H. R. Arabnia. "A survey of face recognition techniques." Jips 5(2): 41-68, 2009.
- [4] Taigman, Y., M. Yang, M. A. Ranzato and L. Wolf. "DeepFace: Closing the Gap to Human-Level Performance in Face Verification". Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2014.
- [5] Schroff, F., D. Kalenichenko and J. Philbin. "Facenet: A unified embedding for face recognition and clustering". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [6] Wright, J., A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma. "Robust face recognition via sparse representation." IEEE transactions on pattern analysis and machine intelligence 31(2): 210-227, 2009.
- [7] Ahonen, T., A. Hadid and M. Pietikainen. "Face description with local binary patterns: Application to face recognition." IEEE transactions on pattern analysis and machine intelligence 28(12): 2037-2041, 2006.
- [۸] بخشی. م. (۱۳۹۵). "استخراج ویژگی از تصاویر چهره جهت بازیابی تصویر از پایگاه‌داده". کارشناسی ارشد پایان نامه، دانشگاه شاهرود.
- [9] Varun, R., Y. V. Kini, K. Manikantan and S. Ramachandran. "Face Recognition using Hough Transform based Feature Extraction." Procedia Computer Science 46: 1491-1500, 2015.
- [10] Chen, D., X. Cao, F. Wen and J. Sun. "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013.
- [11] Ding, C., J. Choi, D. Tao and L. S. Davis. "Multi-directional multi-level dual-cross patterns for robust face recognition." IEEE transactions on pattern analysis and machine intelligence 38(3): 518-531, 2016.
- [12] Viola, P. and M. J. Jones. "Robust real-time face detection." International journal of computer vision 57(2): 137-154, 2004.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

طراحی و بکارگیری فیلتر وینر در جداسازی سیگنال‌های EMI ناشی از مبدل‌های DC/DC افزایشدهنده در حالت موازی

جواد برسلانی^۱، علی دستفان^۲، جواد قالیبافان^۳

^۱ دانشجوی دکتری مهندسی برق قدرت، دانشکده مهندسی برق، دانشگاه صنعتی شاهرود، شاهرود
j.borsalani@shahroodut.ac.ir

^۲ دانشیار گروه قدرت، دانشکده مهندسی برق، دانشگاه صنعتی شاهرود، شاهرود
dastfan@shahroodut.ac.ir

^۳ استادیار گروه الکترونیک و مخابرات، دانشکده مهندسی برق، دانشگاه صنعتی شاهرود، شاهرود
jghalibafan@shahroodut.ac.ir

چکیده

این مقاله نحوه‌ی بکارگیری یک روش بخصوص پردازش سیگنال به منظور تخمین و جداسازی منابع بوجود آورنده تداخلات الکترومغناطیس (EMI) در یک سیستم شامل بیش از یک مبدل الکترونیک قدرت را بیان می‌دارد. این روش این امکان را فراهم می‌دارد که فقط با انجام تعداد محدودی اندازه‌گیری و محاسبه توابع تبدیل بین منابع نویز و خروجی مورد نظر، هم EMI کلی حاصل را تخمین زده و هم تاثیر هر یک از منابع را تعیین کرد. این روش بر پایه تئوری فیلتر وینر است که نحوه طراحی و بکارگیری آن بیان شده و عملکرد آن نیز در یک سیستم نمونه شامل دو مبدل DC/DC افزایشدهنده در حالت موازی، با استفاده از شبیه سازی اعتبار سنجی شده است. نتایج شبیه سازی نشان دهنده رفتار مناسب روش ارائه شده در تخمین میزان EMI کلی در سیستم است.

کلمات کلیدی

تداخلات الکترومغناطیس، فیلتر وینر، جداسازی منابع نویز، تخمین EMI، مبدل‌های الکترونیک قدرت موازی

گرفتن تمهیداتی جهت کاهش میزان EMI تولیدی، قادر به تامین خواست‌های استاندارد نخواهند بود [5]. به منظور کاهش EMI تولیدی یک مبدل، دستیابی به مدل فرکانس بالای سیستم در باند فرکانس مورد نظر (150KHz تا 30MHz) با اهمیت خواهد بود لذا در مراجع به مدلسازی مبدل‌های الکترونیک قدرت از دیدگاه EMC پرداخته شده است [7-5]. این مدل‌ها شامل منبع نویز (کلید نیمه هادی) و مسیر انتشار آن (مسیرهای قدرت در سیستم و نیز امپدانس‌های پراکندگی) می‌شوند. سطح EMI تولیدی به شکل هندسی سیستم، بسته‌بندی، PCBها، پراکندگی‌های اجزای پسیو و مقادیر ولتاژ

۱- مقدمه

امروزه با گسترش بکارگیری مبدل‌های الکترونیک قدرت در صنعت، لزوم کاهش اندازه و افزایش چگالی توان آن‌ها ضروری شده است. جهت افزایش چگالی توان مبدل‌ها لازم است فرکانس کلیدزنی در مبدل را افزایش داد [4-1] که خود سبب افزایش تداخلات الکترومغناطیس (EMI^۱) ناشی از مبدل خواهد شد. استانداردهای حوزه سازگاری الکترومغناطیس (EMC^۲) تعیین کننده سطوح مجاز تداخلات تولیدی توسط مبدل‌ها هستند. مبدل‌های الکترونیک قدرت اغلب بدون در نظر

منبع ثابت V_{dc} و مبدل (دو مبدل افزایشدهنده در حالت موازی) قرار می‌گیرد [3]. یک LISN استاندارد، متشکل از عناصر خطی است که با ایجاد امپدانس ثابت و معلوم در باند فرکانس EMI هدایتی، مسیر نویز ناشی از مبدل الکترونیک قدرت را از طریق یک مقاومت مرجع $50\ \Omega$ می‌بندد و از انتشار آن به منبع قدرت جلوگیری می‌کند. در واقع افت ولتاژ بوجود آمده روی مقاومت مرجع $50\ \Omega$ در LISN، میزان نویز ناشی از مبدل الکترونیک قدرت را نشان خواهد داد. مدار داخلی LISN مورد استفاده در شکل ۲ آمده است.

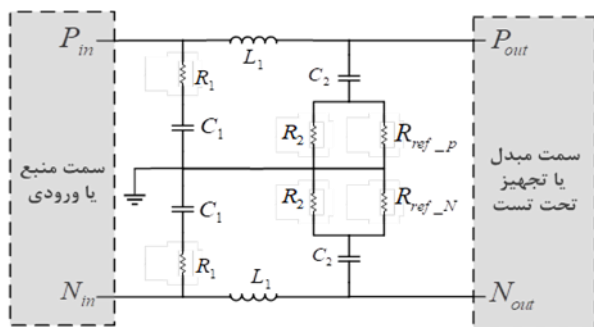
سطح تداخلات اندازه‌گیری شده در LISN و نیز شکل آن، نه تنها به تغییرات ولتاژ (dv/dt) و تغییرات جریان (di/dt) در فرایند کلید زنی در سوئیچ‌های قدرت وابسته بوده بلکه به مسیر انتشار آن نیز بستگی دارد. این مسیر هم از طریق مسیرهای قدرت موجود در مدار (شکل ۱) و هم از طریق پراکندگی‌ها در مدار حاصل می‌گردد. به بیان دیگر در مدلسازی مسیر انتشار EMI، پراکندگی‌های موجود شامل خازن‌ها و سلف‌های پراکندگی سوئیچ‌های قدرت، پراکندگی‌های عناصر پسیو (مقاومت، سلف، خازن و...) قابل صرف نظر کردن نیستند و لازم است به طور دقیق مشخص شوند. شکل ۱ سیستم تحت بررسی شامل دو مبدل DC به DC افزایشدهنده در ورودی با یکدیگر موازی شده‌اند را نشان می‌دهد. هدف یافتن سهم هر یک از مبدل‌ها در سیگنال کلی نویز بوجود آمده در LISN می‌باشد. این کار با طراحی و بکارگیری فیلتر وینر انجام پذیر است که در بخش بعدی معرفی می‌گردد. در بکارگیری این روش فرض‌های زیر در نظر گرفته شده است:

- تداخلات در هر یک از مبدل‌های شکل ۱ فقط ناشی از تغییر وضعیت کلیدهای S1 و S2 بوده؛ اما سطح و شکل آن به مدار متصل به سوئیچ و پراکندگی‌ها دارد.
- کلیدهای قدرت در دو مبدل به طور همزمان تغییر وضعیت نمی‌دهند یا به بیان دیگر دو مبدل مستقل از یکدیگر کنترل شده و به صورت کلی، سوئیچ زنی هر یک از مبدل‌ها دارای اختلاف زمانی است.

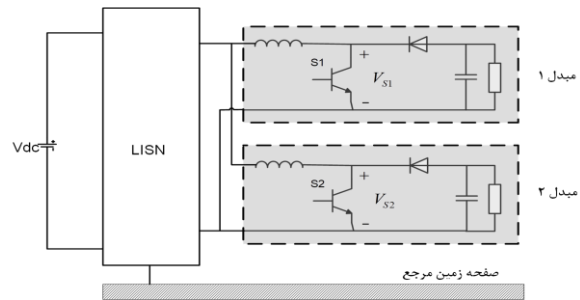
با در نظر گرفتن فرض‌های بالا، سیستم تحت بررسی از دیدگاه EMI، تبدیل به یک سیستم LTI دو ورودی (ولتاژهای کلیدها) و یک خروجی (ولتاژ مقاومت مرجع $50\ \Omega$ در LISN) شده است.

۳- پیاده سازی فیلتر وینر در تخمین منابع نویز

روش کلی پیاده سازی فیلتر وینر در تخمین EMI در شکل ۳ نمایش داده شده است. همان گونه که در شکل ۳ نشان داده شده، طراحی فیلتر



شکل ۲: مدار داخلی LISN مورد استفاده جهت اندازه‌گیری نویز



شکل ۱: بلوک دیاگرام سیستم کلی تحت بررسی شامل دو مبدل DC/DC افزایشدهنده در حالت موازی متصل به LISN جهت اندازه‌گیری میزان EMI تولیدی

(dv/dt) و جریان (di/dt) و... در سیستم وابسته است که سبب شده تا مدل پیشنهاد شده دارای پیچیدگی شود [8].

زمانی که سیستم تحت بررسی شامل بیش از یک مبدل شود، مسئله مدلسازی EMI، پیچیده‌تر نیز خواهد شد [9]. در این سیستم‌ها بیش از یک منبع نویز موجود است که ممکن است عملکرد همزمان مبدل‌ها سبب تجاوز میزان EMI کلی سیستم از سطوح استاندارد شود. شکل ۱ نمونه‌ای از یک سیستم شامل دو مبدل الکترونیک قدرت متصل به یک منبع را نشان می‌دهد که به طور نوعی مستقل از یکدیگر کنترل می‌شوند. با بهره‌گیری از نرم افزارهای تحلیل حوزه زمان مانند MATLAB/Simulink [10] و با داشتن مدل مداری فشرده هریک از مبدل‌ها می‌توان به میزان کلی EMI سیستم در LISN دست یافت. طیف سیگنال ولتاژ حاصل از کلیدزنی در هر مبدل را می‌توان به عنوان دو منبع نویز (ورودی) و طیف سیگنال ولتاژ LISN را به عنوان خروجی یک سیستم LTI^۱ در نظر گرفت لذا با استفاده از تکنیک فیلتر وینر [11] می‌توان تاثیر هر یک از منابع در EMI کلی را تعیین کرد. با دانستن این تاثیر و متعاقباً تفکیک نویز حاصل از هر یک از مبدل‌ها، می‌توان نقاط بحرانی تشدید نویز را شناسایی و حذف کرد. لذا در این مقاله با اندازه‌گیری سیگنال‌های ولتاژ سوئیچ‌های هر یک از مبدل‌ها و نیز ولتاژ EMI کلی در LISN، فیلتر وینر جهت جداسازی تاثیر هر یک از منابع ارائه شده است.

در بخش دوم این مقاله، سازوکار کلی تولید EMI در مبدل‌های افزایشدهنده موازی و نیز نحوه اندازه‌گیری آن بیان شده است. در بخش سوم ابتدا تئوری نحوه یافتن فیلتر وینر در جداسازی منابع و سپس نحوه بکارگیری آن در سیستم تحت بررسی ارائه شده است. سیستم شبیه‌سازی شده و نیز نتایج حاصل از بکارگیری فیلتر ارائه شده در بخش چهارم بیان شده و نهایتاً جمع بندی و نتیجه‌گیری این مقاله در قسمت پنجم آمده است.

۲- تداخل در مبدل‌های افزایشدهنده موازی

در مبدل‌های الکترونیک قدرت منابع اصلی تداخلات الکترومغناطیس، خاموش و روشن شدن سریع کلیدهای نیمه هادی است که در اثر آن ولتاژها و جریان‌های فرکانس بالا در مدار بوجود آمده و از طریق مسیرهای قدرت و نیز پراکندگی‌ها در مدار منتشر می‌گردد. جهت اندازه‌گیری این تداخلات از شبکه پایداریساز امپدانس خط (LISN) استاندارد همانند شکل ۱ استفاده می‌شود که میان منبع قدرت (در اینجا

$$E_r(j\omega) = P(j\omega) - P_v(j\omega) \quad (2)$$

$$= P(j\omega) - H(j\omega)V(j\omega)$$

هدف از طراحی فیلتر حداقل سازی خطا در رابطه (۲) خواهد بود. لذا روش کاهش مربعات برای رابطه (۲) به صورت (۳) حاصل می‌گردد.

$$E\left[|E_r(j\omega)|^2\right] = E\left[(P(j\omega) - H(j\omega)V(j\omega)) \times (P(j\omega) - H(j\omega)V(j\omega))^*\right] \quad (3)$$

که در آن $E[\cdot]$ نشان دهنده عملگر امید ریاضی و * عمل مزدوج گیری اعداد مختلط است.

هدف یافتن مقدار بهینه برای پاسخ فرکانسی H جهت کمینه سازی (۳) است لذا کافی است از رابطه (۳) مشتق گرفته و برابر صفر قرار داده شود [10]. لذا با استفاده از روش‌های مشتق گیری اعداد مختلط رابطه (۴) حاصل می‌گردد.

$$\frac{\partial E\left[|E_r(j\omega)|^2\right]}{\partial H(j\omega)} = 2H(j\omega)S_{vv}(j\omega) - 2S_{vp}(j\omega) \quad (4)$$

که در آن $S_{vv}(j\omega) = E\left[|V(j\omega)|^2\right]$ طیف توان سیگنال v و $S_{vp}(j\omega) = E\left[P(j\omega)V^*(j\omega)\right]$ همبستگی متقابل میان سیگنال‌های p و v است.

با صفر قرار دادن (۴) مقدار بهینه پاسخ فرکانسی برای فیلتر وینر به صورت زیر حاصل می‌گردد:

$$H_o(j\omega) = \frac{S_{vp}(j\omega)}{S_{vv}(j\omega)} \quad (5)$$

در مراجع، تحلیل مشابهی جهت یافتن پاسخ ضربه بهینه برای فیلتر وینر (h) در حوزه زمان ارائه شده است [10] که در آن پاسخ ضربه فیلتر از توابع خودهمبستگی سیگنال v (R_{vv}) و همبستگی متقابل سیگنال‌های p و v (r_{vp}) در حوزه زمان به صورت (۶) حاصل می‌گردد:

$$h = R_{vv}^{-1}r_{vp} \quad (6)$$

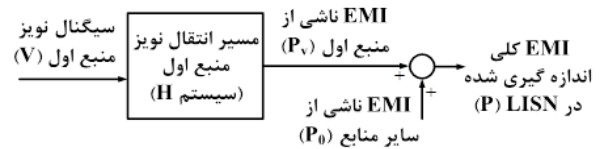
روابط (۵) و (۶) بیان از فیلتر وینر را یکی در حوزه فرکانس و دیگری در حوزه زمان دارند. در این مقاله فرایند طراحی فیلتر وینر جهت تخمین منبع نویز در حوزه فرکانس زمان انجام شده است.

۳-۲- بکارگیری فیلتر در جداسازی منابع نویز

همانگونه که پیشتر گفته شد، سیستم تحت بررسی دارای دو ورودی و یک خروجی است. لذا باید به ازای هر یک از ورودی‌ها به خروجی، یک فیلتر وینر مجزا با رابطه (۶) طراحی کرد پس سیگنال کلی EMI بازسازی شده حاصل از خروجی فیلترهای وینر طراحی شده به صورت زیر بدست خواهد آمد.

$$P' = P_{v1} + P_{v2} = h_1V_{s1} + h_2V_{s2} \quad (7)$$

که در آن P' سیگنال کلی EMI بازسازی شده، P_{v1} و P_{v2} سیگنال‌های تفکیک شده حاصل از هر یک از منابع h_1 و h_2 به ترتیب فیلتر وینر طراحی شده برای مبدل اول و دوم می‌باشد. جهت اعتبار سنجی فیلترهای طراحی شده، کافی است سیگنال بازسازی شده P' را با سیگنال کلی EMI اندازه گرفته شده مقایسه کرد.



شکل ۳: بلوک دیاگرام نحوه انتشار نویز در سیستم تحت بررسی و چگونگی پیاده سازی فیلتر وینر

وینر دارای دو هدف است: اول این که سیستم H را شناسایی کرده و از طرفی دیگر با انجام فقط دو اندازه گیری، میزان سیگنال p_v را تخمین بزنند. این اندازه گیری‌ها شامل سیگنال‌های زیر است:

- منبع اول نویز v که در این مقاله ولتاژ یکی از دو کلید نیمه هادی در شکل ۱ است.
- میزان کلی نویز اندازه گرفته شده p که شامل مجموع اغتشاش ناشی از دیگر منبع نویز p_0 و اغتشاش ناشی از منبع اول p_v در این مقاله مقدار p همان افت ولتاژ اندازه گرفته شده در امپدانس‌های LISN در شکل ۱ خواهد بود.

در طراحی و به کارگیری فیلتر وینر، دو فرض اساسی موجود است [12] که عبارتند از:

- ۱- سیستم H یک سیستم خطی و تغییر ناپذیر نسبت به زمان (LTI) است بدین معنی که می‌توان آن را به طور کامل توسط پاسخ فرکانسی یا پاسخ ضربه بیان کرد.
 - ۲- میزان نویز ناشی از منبع دیگر (p_0) در سیستم کلی، با سیگنال‌های اندازه گرفته شده p و v جمع شونده و ناهمبسته است.
- در رابطه با دو مبدل افزاینده که به صورت موازی با یکدیگر عمل می‌کنند که در این مقاله مورد بررسی قرار گرفته شده، هر دو فرض بالا مورد قبول است چراکه اولاً همانگونه که در شکل ۱ نشان داده شده است کلیه عناصر مدار در باند فرکانس EMI (150KHz تا 30MHz)، خطی فرض شده و با مدل المان‌های فشرده مدل‌سازی شده‌اند لذا سیستم کلی حاصل (که شامل زیر سیستم H نیز خواهد شد) خطی بوده که در نتیجه سیستم H هم خطی است. در ثانی دو مبدل افزاینده در نظر گرفته شده به صورت مستقل از یکدیگر عمل کرده و کلیدزنی آن‌ها هم به صورت همزمان انجام نمی‌گیرد؛ پس فرض ناهمبسته بودن سیگنال‌های نویز ناشی از هر یک از مبدل‌ها هم قابل در نظر گرفتن است.

۳-۱- طراحی فیلتر وینر

با نظر گرفتن فرض‌های یاد شده، می‌توان کلیه سیگنال‌ها را با تبدیل فوریه آن جایگزین کرد. لذا با توجه به شکل ۳ داریم:

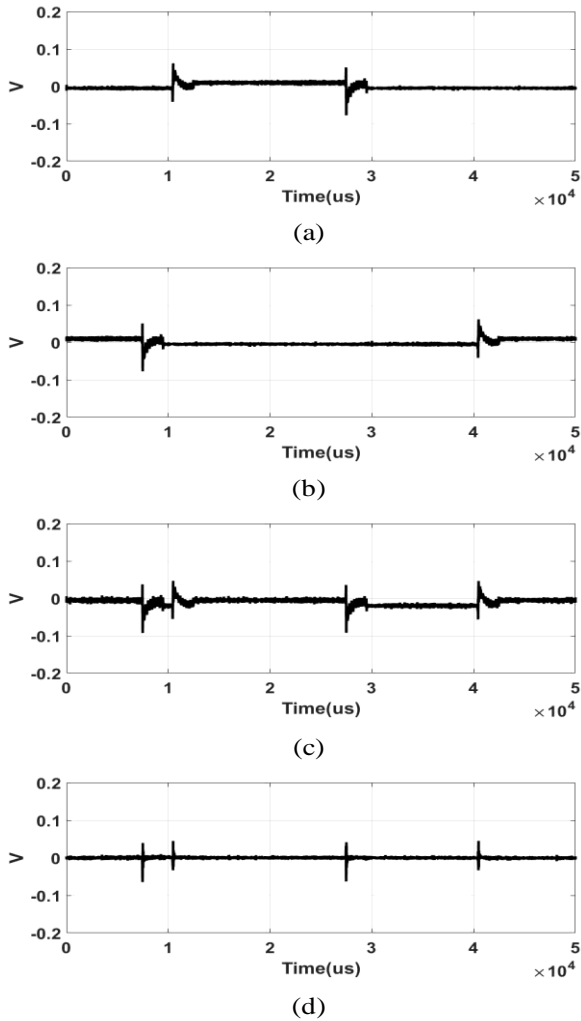
$$p_v(j\omega) = H(j\omega)V(j\omega) \quad (8)$$

که در آن به ترتیب $V(j\omega)$ و $H(j\omega)$ تبدیل فوریه سیگنال‌های v و پاسخ فرکانسی سیستم H هستند.

می‌توان خطای تخمین $E_r(j\omega)$ را به صورت اختلاف میان میزان کلی نویز اندازه گیری شده و نویز ناشی از منبع اول به صورت (۲) در نظر گرفت:

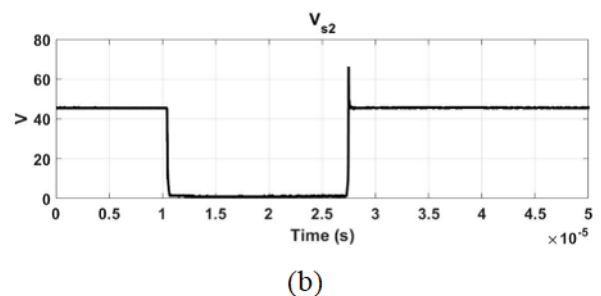
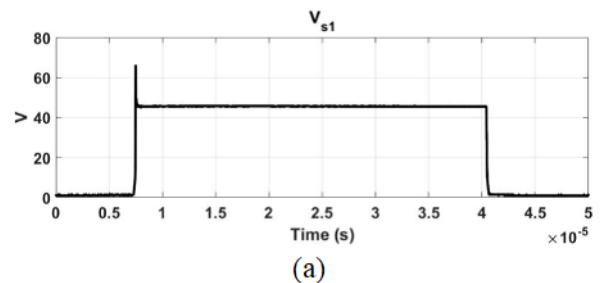
۴- شبیه سازی و تحلیل نتایج شبیه سازی

در این بخش به اعتبار سنجی فیلتر وینر طراحی شده توسط شبیه سازی در نرم افزار MATLAB/Simulink پراخته شده است. همانگونه که پیشتر اشاره شد، برای انجام شبیه سازی، لازم است مدل کاملی از سیستم تحت بررسی موجود باشد. در این مقاله از مدل ارائه شده در [13] جهت شبیه سازی هر یک از مبدل‌های شکل ۱ استفاده شده است. همچنین LISN نیز همانند شکل ۲ توسط المان‌های مداری فشرده مدلسازی شده است. پله‌های زمانی شبیه سازی جهت مشاهده رفتار گذرایی تغییرات ولتاژ و جریان برابر 1ps انتخاب شده که به خوبی بتواند حالت گذرای خاموش و روشن شدن کلیدهای نیمه هادی را مدل کند. زمان کلی شبیه سازی برابر 200us قرار داده شده تا هم حالت گذرایی موجود در سیستم از بین رفته و هم در سیگنال‌های حاصل، کمترین فرکانس نویز مورد نظر (150KHz) قابل اندازه‌گیری باشد و همچنین شکل موج‌های حاصل از شبیه سازی جهت استفاده در طراحی فیلتر وینر، با فرکانس نمونه‌گیری 1GHz اندازه‌گیری شده‌اند. هر یک از مبدل‌های سیستم شکل ۱ به صورت مستقل از یکدیگر و با فرکانس ثابت و برابر کنترل شده‌اند. فرکانس کلیدزنی برابر 20KHz انتخاب شده و زمان و وظیفه کلیدها هم به طور یکسان برابر 70% در نظر گرفته شده است. شکل ۴، شکل موج ولتاژ کلیدهای هر یک از مبدل‌ها را نشان می‌دهد. همانگونه که از این شکل مشخص است به منظور جلوگیری از برهم نهی و نیز کاهش همبستگی میان این دو سیگنال، لحظه روشن شدن کلیدها نسبت به یکدیگر دارای تاخیر زمانی است که این کار سبب خواهد شد شروط لازم جهت طراحی فیلتر وینر که در بخش سوم آن‌ها اشاره شد، برآورده شود.



شکل ۵: شکل موج خروجی فیلتر وینر طراحی شده برای (a) مبدل اول (b) مبدل دوم به همراه (c) شکل موج سیگنال کلی EMI بازسازی شده از فیلتر و نیز (d) شکل موج سیگنال EMI کلی اندازه‌گیری شده در حوزه زمان

در شکل ۵ (a) شکل موج حوزه زمان خروجی فیلتر وینر حاصل از مبدل اول نشان داده شده است. مقایسه این شکل موج و شکل موج ولتاژ کلیدزنی در مبدل اول (شکل ۴ (a)) به خوبی وجود اثر کلیدزنی را مشخص می‌کند. تحلیل مشابهی را می‌توان در رابطه با مبدل دوم (شکل ۵ (b)) نیز بیان کرد. شکل ۵ (c)، شکل موج EMI کلی بازسازی شده حاصل از مجموع شکل موج‌های (a) و (b) را نشان می‌دهد و در شکل ۵ (d) EMI کلی اندازه‌گیری شده در LISN ارائه شده است. مقایسه این دو شکل موج، صحت فیلترهای وینر طراحی شده را نشان می‌دهد. میزان EMI شبیه سازی شده حاصل از تبدیل فوریه ولتاژ اندازه‌گرفته شده در LISN و نیز میزان EMI تخمین زده شده حاصل از بکارگیری فیلتر وینر را می‌توان با به کارگیری تبدیل فوریه از سیگنال‌های حوزه زمان بدست آورد [14] که در شکل ۶ نشان داده شده است.



شکل ۴: شکل موج ولتاژهای در نظر گرفته شده برای کلیدهای نیمه هادی مبدل‌های DC/DC افزاینده (a) مبدل اول (b) مبدل دوم

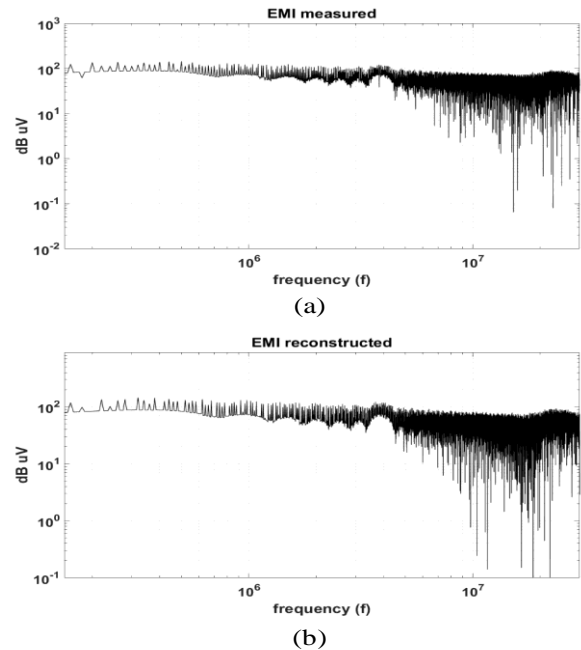
شکل ۵ نتایج حاصل از به کارگیری فیلتر وینر طراحی شده را نشان می‌دهد.

مراجع

- [1] Bondarenko, N.; Li Zhai; Bingjie Xu; Guanghua Li; Makharashvili, T.; Loken, D.; Berger, P.; Van Doren, T.P.; Beetner, D.G., "A Measurement-Based Model of the Electromagnetic Emissions from a Power Inverter," in Power Electronics, IEEE Transactions on, vol.30, no.10, pp.5522-5531, Oct. 2015.
- [2] Xun Gong; Josifovic, I.; Ferreira, J.A., "Modeling and Reduction of Conducted EMI of Inverters with SiC JFETs on Insulated Metal Substrate," in Power Electronics, IEEE Transactions on, vol.28, no.7, pp.3138-3146, July 2013
- [3] M. Rucinski, P. Musznicki and P. J. Chrzan, "Electromagnetic interference frequencies prediction model of flyback converter for snubber design," in IET Power Electronics, vol. 8, no. 6, pp. 994-999, 6 2015. doi: 10.1049/iet-pel.2014.0250
- [4] Hedayati, M.H.; John, V., "Filter Configuration and PWM Method for Single-Phase Inverters with Reduced Conducted EMI Noise," in Industry Applications, IEEE Transactions on, vol.51, no.4, pp.3236-3243, July-Aug. 2015
- [5] Toure, B.; Schanen, J.-L.; Gerbaud, L.; Meynard, T.; Roudet, J.; Ruelland, R., "EMC Modeling of Drives for Aircraft Applications: Modeling Process, EMI Filter Optimization, and Technological Choice," in Power Electronics, IEEE Transactions on, vol.28, no.3, pp.1145-1156, March 2013
- [6] Amarir, S.; Al-Haddad, K., "A Modeling Technique to Analyze the Impact of Inverter Supply Voltage and Cable Length on Industrial Motor-Drives," in Power Electronics, IEEE Transactions on, vol.23, no.2, pp.753-762, March 2008.
- [7] Bishnoi, H.; Baisden, A.C.; Mattavelli, P.; Boroyevich, D., "Analysis of EMI Terminal Modeling of Switched Power Converters," in Power Electronics, IEEE Transactions on, vol.27, no.9, pp.3924-3933, Sept. 2012.
- [8] Koyama, Y.; Tanaka, M.; Akagi, H., "Modeling and Analysis for Simulation of Common-Mode Noises Produced by an Inverter-Driven Air Conditioner," in Industry Applications, IEEE Transactions on, vol.47, no.5, pp.2166-2174, Sept.-Oct. 2011.
- [9] Musznicki, P.; Schanen, J.-L.; Granjon, P.; Chrzan, P.J., "The Wiener Filter Applied to EMI Decomposition," in Power Electronics, IEEE Transactions on, vol.23, no.6, pp.3088-3093, Nov. 2008.
- [10] "Matlab: Simulink," The MathWorks, Nattick, MA, 1994.
- [11] S.-V. Vaseghi, Advance Digital Signal Processing and Noise Reduction. New York: Wiley, 2000.
- [12] P. Musznicki, J. L. Schanen, P. Granjon and P. J. Chrzan, "EMI Estimation for DC/AC Hard Switching Converter Using Wiener Filter," 2006 12th International Power Electronics and Motion Control Conference, Portoroz, 2006, pp. 473-478.
- [13] N. Poon, B. M. Pong, C. Liu, and C. K. Tse, "Essential-coupling-path models for non-contact EMI in switching power converters using lumped circuit elements," IEEE transactions on power electronics, vol. 18, pp. 686-695, 2003.
- [14] Labrousse, D.; Revol, B.; Gautier, C.; Costa, F., "Fast Reconstitution Method (FRM) to Compute the Broadband Spectrum of Common Mode Conducted Disturbances," in Electromagnetic Compatibility, IEEE Transactions on, vol.55, no.2, pp.248-256, April 2013.

زیر نویس‌ها

- 1 Electromagnetic Interference
- 2 Electromagnetic Compatibility
- 3 Line Impedance Stabilization Network
- 4 Linear Time Invariant
- 5 Wiener Filter
- 6 Duty Cycle



شکل ۶: میزان EMI شبیه سازی شده (a) سیگنال اصلی (b) سیگنال بازسازی شده

محور افقی نمودارهای شکل ۶ فرکانس و محور عمودی، دامنه سیگنال در حوزه فرکانس را نمایش می‌دهد که طبق استاندارد بر حسب dBuV در نظر گرفته شده است. شکل ۶ (a) میزان EMI را برای سیگنال اصلی نشان می‌دهد به بیان دیگری این نمودار طیف حوزه فرکانس سیگنال حوزه زمان نشان داده شده در شکل ۵ (d) است و نیز به طور مشابه، شکل ۶ (b) میزان EMI را برای سیگنال بازسازی شده از طریق بهره‌گیری از فیلتر وینر را نشان می‌دهد که طیف حوزه فرکانس سیگنال شکل ۵ (c) است. جهت بررسی سیستم از نقطه نظر EMC، طیف حوزه فرکانس EMI مورد استفاده قرار می‌گیرد [14]، لذا در این مقاله مطلوب آن است که سیگنال‌های شکل ۶ (a) و شکل ۶ (b) مشابه هم باشند. به بیان دیگر استفاده از فیلتر وینر برای شناسایی منابع نویز موجود در سیستم بر میزان EMI کلی اندازه گرفته شده تأثیری نگذارد. مقایسه این دو سیگنال، شباهت دو طیف را بدست می‌دهد.

۵- نتیجه

در این مقاله، به بررسی کاربرد فیلتر وینر به عنوان ابزار پردازش دیجیتالی سیگنال در جداسازی منابع نویز میدل‌های الکترونیک قدرت پرداخته شده است. سازوکار بوجود آمدن EMI در یک سیستم نمونه تحت بررسی شامل دو میدل DC/DC در حالت موازی بیان شده و سپس نحوه طراحی فیلتر وینر مورد نظر و نیز چگونگی پیاده‌سازی آن ارائه شده است. شبیه‌سازی سیستم تحت بررسی به کمک نرم افزار MATLAB/Simulink و توسط مدلسازی المان فشرده در حوزه زمان صورت پذیرفته است. نتایج شبیه سازی نشان دهنده توانایی قابل قبول فیلتر ارائه شده در تخمین میزان EMI کلی سیستم، چه در حوزه زمان و چه در حوزه فرکانس می‌باشد.



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

تشخیص افعال روزمره با استفاده از تلفن همراه و یادگیری عمیق

محمود معلم^۱، علی اکبر پویان^۲

^۱ دانشجوی دکتری، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود،

moallem@shahroodut.ac.ir

^۲ استادیار، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی شاهرود، شاهرود،

a.pouyan@shahroodut.ac.ir

چکیده

کشف و بازشناسی خودکار افعال روزمره برای تحقق اینترنت اشیا و ارائه خدمات مبتنی بر مکان بسیار مهم است. با پیشرفت کمی و کیفی حسگرهای موجود در تلفن‌های همراه هوشمند - حتی غیرهوشمند - این دستگاه‌ها به دستگاه‌های پایش مداوم و ارزان‌قیمتی تبدیل شده‌اند که می‌توانند به شکل برخط، حرکت‌های مالکان خود را ثبت کنند؛ اما تشخیص حرکت انجام‌شده از روی این داده‌های خام مستلزم وجود روش‌های طبقه‌بندی قدرتمند است. تا پیش از ظهور یادگیری عمیق تشخیص این افعال معمولاً بر اساس ویژگی‌های دست‌ساز انجام می‌شد و زمان‌بر بودن استخراج این ویژگی‌های عملاً امکان تشخیص بر خط را منتفی می‌کرد. ما در این مقاله سه ساختار عمیق را بر روی داده‌های خام یک دادگان مشهور تشخیص افعال به نام ActiTrack بکار گرفته‌ایم. نتایج حاصل، به‌ویژه در مورد ساختار LSTM ثابت می‌کند که این ساختارها می‌توانند امکان تشخیص بر خط افعال را با دقت بالایی فراهم کنند.

کلمات کلیدی

تشخیص افعال، حسگرهای پوشیدنی، یادگیری عمیق، LSTM

کمیت و کیفیت حسگرهای موجود، رشد چشمگیر و پیشرفت قابل توجهی داشته‌اند. در ساختار یک تلفن همراه امروزی حسگرهای متعددی مثل حسگر مکان، مجاورت، مغناطیس، نور، صدا، شتاب زاویه‌ای و شتاب خطی تعبیه شده است. مورد آخر، یعنی شتاب سنج خطی، تقریباً در همه تلفن‌های همراه هوشمند وجود دارد. وجود این حسگرها در تلفن‌های همراه هوشمند و امکان یکپارچه‌سازی این تلفن‌ها با اینترنت اشیا - و شبکه جهانی اینترنت به‌عنوان بستر محتوایی - باعث پیدایش مقولات مؤثر و مهمی مثل خدمات حال آگاه^۱ و رایانش فراگیر^۲ شد که بسیاری از محققین آن‌ها را موج سوم رایانش می‌دانند [2].

یکی از ملزومات و شاخه‌های فرعی محاسبات حال آگاه بازشناسی خودکار افعال انسانی^۳ است. در این شاخه حسگرهای متفاوت داده‌های مربوط به احوال یک یا چند نفر را جمع‌آوری می‌کنند تا با تحلیل و آکاوی این داده‌ها، حال و وضع و فعل آن فرد یا افراد به‌طور خودکار و

۱ - مقدمه

جهان اطراف ما امروزه مملو از انواع حسگرهایی است که به‌طور دائم و پیوسته داده‌های محیطی را جمع‌آوری می‌کنند تا از تجمیع و تحلیل این داده‌ها، اطلاعات و دانشی فراهم شود که در نهایت برای پالایش، بهبود و ارتقای همان محیط و همین جهان مورد استفاده قرار می‌گیرد. این حسگرها معمولاً برای ما مزاحمتی فراهم نمی‌کنند و حتی غال با از دید ما پنهان هستند، اما وجود دارند و در کنار هم شبکه گسترده‌ای را تشکیل می‌دهند که از شبکه جهان‌گستر وب بسیار بزرگ‌تر است. شبکه‌ای که آن را اینترنت اشیا^۴ [1] می‌نامند.

یکی از اجزای بسیار رایج و پرکاربرد این شبکه، تلفن‌های همراه هوشمند هستند. این تلفن‌ها در دهه گذشته در جمیع جهات، از جمله

استخراج دستی ویژگی‌ها علیرغم رواجی که در تجزیه و تحلیل سیگنال دارد از دو محدودیت رنج می‌برد [11]:

- مرحله استخراج ویژگی بیش از اینکه به اطلاعات موجود در سیگنال ربط داشته باشد به دانش زمینه‌ای محقق‌ای که سیگنال را بررسی می‌کند مربوط است، به همین خاطر به دلیل تنوع حرکات غیرمعمول در افراد مختلف ممکن است برخی خصائص این حرکات از قلم بیافتد یا به‌طور کامل مورد بررسی قرار نگیرد.
- استخراج ویژگی‌های دست‌ساز معمولاً هزینه‌های محاسباتی و زمانی را به فرآیند پردازش تحمیلی کند و در نتیجه قابلیت‌های این فرآیند را برای به‌کارگیری در سناریوهای زمان واقعی محدود می‌نماید.

پیشنهاد ما برای حل مشکلات مذکور استفاده از ویژگی‌های ذاتی داده‌هاست که به‌طور خودکار و در طول فرآیند آموزش از داخل خود داده‌های خام استخراج می‌شود. بدین منظور ما از روش یادگیری عمیق و ابزار LSTM، برای یادگیری ویژگی‌های اصلی و مهم در تشخیص و شناسایی الگوی افعال روزمره استفاده می‌کنیم.

۳- کارهای انجام شده

تا پیش از ظهور یادگیری عمیق، مطالعات متعددی در زمینه‌ی بازشناسی افعال با استفاده از گوشی تلفن همراه انجام شده است. تقریباً تمامی مطالعات یادشده یک فرآیند چهار مرحله‌ای را طی کرده‌اند:

۱. جمع‌آوری اطلاعات
۲. پیش‌پردازش اطلاعات خام
۳. استخراج ویژگی‌ها
۴. طبقه‌بندی

در اغلب این مطالعات جمع‌آوری داده‌ها از طریق حسگر شتاب سنج موجود در یک یا چند تلفن همراه هوشمند صورت پذیرفته است [12]. البته گاهی ژيروسکوپ [13] و بندرت مگنومتر [14] نیز مورد استفاده قرار گرفته‌اند. مرحله پیش‌پردازش هم معمولاً شامل حذف نویز از طریق اعمال پالایه‌های بالا یا پایین‌گذر و تقسیم داده‌ها به پنجره‌های همپوشان است. این پنجره‌ها غالباً هم طول هستند و طول آن‌ها بر مبنای نرخ نمونه‌برداری حسگرها و نوع افعال مورد نظر انتخاب می‌شود. البته برخی از محققین طول پنجره‌ها و میزان همپوشانی آن‌ها را مورد توجه قرار داده و تأثیر این عوامل بر دقت و جامعیت بازشناسی افعال را ارزیابی کرده‌اند [15].

استخراج ویژگی‌ها مهم‌ترین بخش فرآیند چهارگانه تشخیص افعال را تشکیل می‌دهد. انتخاب ویژگی‌های مناسبی که از درون داده‌های خام استحصال شده و معیارهای طبقه‌بندی، مثل دقت، صحت و جامعیت را افزایش دهد یکی از بزرگ‌ترین چالش‌هایی است که محققین با آن‌ها روبرو بوده‌اند. بیشتر تحقیقات موجود در حوزه بازشناسی افعال انسانی بر مبنای ویژگی‌های دست‌ساز بنا شده است. این ویژگی‌های مکاشفه‌ای هستند، یعنی انتخاب آن‌ها بر اساس دانش، تجربه و گاه علاقه و سلیقه محقق صورت می‌پذیرد. رایج‌ترین انواع ویژگی‌های که تاکنون در مقوله کشف افعال به کار گرفته شده‌اند شامل یک یا چند مورد از موارد زیر می‌باشند:

از طریق یک سیستم محاسباتی تخمین زده شود. حسگرهای مورد استفاده در فرآیند کشف افعال می‌توانند از نوع حسگرهای بصری، پوشیدنی، متراکم و یا ترکیبی از آن‌ها باشد. همچنین افعال مورد نظر ممکن است از نوع ساده (مثل راه رفتن، نشستن، ...) یا پیچیده (مثل غذا پختن یا نظافت منزل) باشند. به دلیل شیوع حسگرهای پوشیدنی و ارائه آن‌ها در قالب تلفن همراه، ساعت، مچ‌بند، کمربند و حتی دکمه هوشمند، ما در مقاله حاضر تنها این نوع از حسگرها را مورد توجه قرار می‌دهیم. گرچه روش ارائه شده برای سایر انواع حسگرها نیز قابل استفاده هست. همچنین برای ساده‌تر شدن موضوع، افعال، جملگی از نوع ساده در نظر گرفته شده‌اند.

کشف و بازشناسی خودکار افعال انسانی با استفاده از تلفن‌های هوشمند در چند سال گذشته توجه فراوانی را به خود جلب کرده است. بخش مهمی از این توجه به دلیل کاربردهای بدیع، مفید و ذی‌قیمتی است که از آن حاصل می‌شود. به‌عنوان بخشی از کاربردهای یادشده می‌توان به نظارت بر رفتار سالخوردهگان [3]، پایش سلامتی بیماران و کودکان [5]، [4]، ردیابی مکان باهدف ارائه خدمات خاص [6]، تبلیغات حال آگاه [7] و ارزیابی کیفیت خواب و تمرین ورزشکاران [8] اشاره کرد. همه این کارها صرف نظر از تفاوت‌هایی که باهم دارند، مستلزم پردازش سیگنال پیوسته یا گسسته‌ای هستند که از طریق نمونه‌برداری یا تغییر وضعیت حسگرها در قالب سری‌های زمانی یا سلسله‌ها بروز و ظهور پیدا می‌کنند.

در ادامه این مقاله، ما در بخش ۲ به بیان مسئله می‌پردازیم. در بخش ۳ به مرور تاریخچه کشف افعال و سابقه به‌کارگیری یادگیری عمیق در این موضوع خواهیم پرداخت. در بخش ۴ راه‌حل پیشنهادی و ساختار شبکه LSTM مورد استفاده معرفی می‌شود. بخش ۵ شامل توصیف دادگان و بیان تجربیات انجام شده است. مقاله حاضر در بخش ۶ و با نتیجه‌گیری به پایان خواهد رسید.

۲- بیان مسئله

بازشناسی افعال انسانی با شتاب سنج، مثل بسیاری از کاربردهای دیگر پردازش سیگنال بر اساس استخراج ویژگی‌های دست‌ساز از داده‌های سیگنال شتاب سنج صورت می‌پذیرد. تا به حال طیف گسترده‌ای از روش‌های استخراج ویژگی در حوزه پردازش سیگنال مورد استفاده قرار گرفته است. این روش‌ها به‌طور کلی دو نوع اصلی از ویژگی‌ها را استخراج می‌کنند [9].

ویژگی‌های دامنه زمان

ویژگی‌های دامنه فرکانس

ویژگی‌های دامنه زمان معمولاً شامل خصائص آماری مثل میانگین، انحراف از معیار، عبور از صفر، انرژی و همبستگی هستند که از پنجره‌های همپوشان سیگنال استخراج می‌شوند. در ویژگی‌های دامنه فرکانس نیز تبدیل فوریه گسسته از بقیه رایج‌تر است، این ویژگی برای برآورد توان باندهای فرکانسی مختلف مورد استفاده قرار می‌گیرد. بعلاوه منبع [4] پیشنهاد می‌کند که برای طبقه‌بندی دقیق‌تر حرکات غیرارادی، تبدیل Stockwell [10] نیز در مرحله استخراج ویژگی‌ها از سیگنال شتاب سنج سه محوره بکار گرفته شود.

استفاده می‌کنند. این شبکه‌ها که در بسیاری از کاربردهای طبقه‌بندی نتایج خوبی را تولید کرده‌اند در مواجهه با داده‌های مسلسل مثل سری‌های زمانی و رشته‌ها از یک نقطه ضعف بزرگ رنج می‌برند. شبکه‌های کاملاً متصل واحدهای ورودی خود را مستقل در نظر می‌گیرند و در نتیجه از کشف و درک وابستگی‌های محلی و سلسله مراتبی زمانی و مکانی آن‌ها عاجز هستند. استخراج این روابط به شبکه‌های خاص‌تری مثل شبکه‌های کانولوشنی^{۱۲} یا شبکه‌های بازگشتی^{۱۳} نیاز دارد.

در میان ساختارهای یادگیری عمیق، شبکه‌های کانولوشنی بیشترین کارکرد و تأثیر را در حوزه محاسبات فراگیر، به‌طور عام و شاخه کشف افعال، به‌طور خاص برجای گذاشته‌اند [33]. این شبکه‌ها معمولاً از لایه‌های پنهان متعددی برخوردار هستند که در آن‌ها پالایه‌های کانولوشنی متفاوتی پیاده‌سازی می‌شود. پالایه‌های یادشده ویژگی‌ها ذاتی داده‌ها را استخراج می‌کنند و در ترکیب با لایه‌های واکشی^{۱۴} و لایه‌های کاملاً متصل یک ساختار قدرتمند برای طبقه‌بندی با ناظر ایجاد می‌نمایند. مراجعی مثل [34] و [35] و [36] انواع شبکه‌های کانولوشنی را برای بازشناسی و طبقه‌بندی افعال انسانی بکار گرفته‌اند. برخی از محققان سعی کرده‌اند با افزودن لایه‌های کانولوشنی یا سرکشی جدید یا تغییر لایه‌های موجود مدل‌های بدیعی ایجاد کنند که به بهای پیچیدگی محاسباتی بیشتر نتایج دقیق‌تری را تولید نماید [34]. برخی دیگر نیز کوشیده‌اند تا با ارائه شبکه‌های ساده‌تر و کوچک‌تر، امکان تعبیه و یادگیری این مدل‌ها بر روی دستگاه‌های همراه و قابل حمل، و در نتیجه بازشناسی برخط افعال را فراهم کنند [37].

چنانکه گفته شد در فرایند کشف افعال، داده‌ها به پنجره‌هایی تقسیم می‌شوند که معمولاً از طول یکسانی برخوردار بوده و دارای هم‌پوشانی هستند. متأسفانه در بسیاری از مدل‌های رایج، این پنجره‌ها به‌عنوان بسته‌های مستقل تلقی شده و وابستگی‌های زمانی آن‌ها نادیده گرفته می‌شود. در سال‌های اخیر بسیاری از محققین برای رفع این نقیصه به شبکه‌های بازگشتی به‌ویژه حافظه طولانی کوتاه‌مدت یا LSTM^{۱۵} روی آورده‌اند. شبکه‌های عصبی بازگشتی در واقع سامانه‌های پویایی هستند که وضعیت درونی خود را در خلال گام‌های یک فرآیند، مثلاً طبقه‌بندی، حفظ می‌کنند و به همین دلیل از نوعی حافظه برخوردار هستند.

حافظه طولانی کوتاه‌مدت یا LSTM تاکنون در زمینه‌های متعددی مثل پردازش گفتار [38]، تشخیص متون دست‌نویس [39]، زیرنویس کردن تصاویر [40] و به‌ویژه کاربردهای پردازش زبان طبیعی [41] مثل ترجمه ماشینی [42] با موفقیت بکار گرفته شده است. از آنجاکه این شبکه‌ها توانایی زیادی در مدل‌سازی و تشخیص وابستگی‌های (درازمدت) زمانی دارند، انتظار می‌رود که توفیق آن‌ها در تحلیل و طبقه‌بندی سیگنال‌های شتاب سنج نیز تکرار شود.

۴- راه‌حل پیشنهادی

در سال ۱۹۹۷ سپ هواختر و هوشیدر [43] سعی کردند با ابداع مفهوم درگاه‌های ورودی و خروجی مشکل اختفای گرادیان در شبکه‌های بازگشتی را حل کنند. تلاش ایشان به ارائه نوعی از شبکه‌های بازگشتی به نام حافظه طولانی کوتاه‌مدت یا LSTM منجر گردید. این ساختار که

۱. خصوصیات آماری سیگنال خام مثل میانگین، میانسه، واریانس و ... [16], [13]
۲. تبدیلات پایه مثل تبدیل فوریه [17]، موجک^{۱۶} [18] و تبدیل کسینوسی گسسته^{۱۷} [19]
۳. مدل خود هم‌بسته^{۱۸} [20]
۴. نمایش سیگنال به شکل رشته‌ای از نمادها [21]

در آخرین گام فرایند چهارگانه بازشناسی افعال، ویژگی‌های فوق‌الذکر به یک طبقه بند تحویل می‌شوند تا حدود و خطوط ذاتی آن‌ها را فرا بگیرد و در نتیجه در مواجهه با داده‌های جدید با دقت و جامعیت بالایی تخمین بزند که این داده‌ها مربوط به کدام یک از افعال هستند. تقریباً تمام روش‌های طبقه‌بندی در ادبیات کشف افعال مورداستفاده قرار گرفته‌اند. در میان مهم‌ترین آن‌ها می‌توان به شیوه‌های نزدیک‌ترین همسایگی یا kNN [22] ماشین بردار پشتیبان یا SVM [23]، شبکه‌های عصبی مصنوعی یا ANN [16]، مدل مخفی مارکوف یا HMM [24] و درخت تصمیم [25] اشاره کرد. برخی از محققین برای افزایش توان مدل پیشنهادی خود از تجمیع طبقه بندها^{۱۹} استفاده کرده‌اند. از جمله [26] و [27] سه روش درخت تصمیم، شبکه عصبی چندلایه، و رگرسیون لجستیک را با یکدیگر جمع نموده‌اند. نتایج به‌دست‌آمده نشان می‌دهد که تجمیع طبقه بندها می‌تواند در بازشناسی افعال انسانی بسیار خوب عمل کند، بطوریکه خروجی مدل حاصل از خروجی هر یک از اجزای آن به‌طور معتنا به‌ای بهتر باشد.

یادگیری عمیق از زمان ظهور خود در سال ۲۰۰۶ تاکنون شاخه‌های مختلفی از هوش مصنوعی را تحت تأثیر قرار داده و به‌ویژه در بیشتر کاربردهای یادگیری ماشینی نتایج را بهبود بخشیده است. طبیعی است که حوزه کشف افعال هم از این موج بی‌نصیب نمانده است. یکی از دلایل جذابیت یادگیری عمیق، عرضه مدل‌های پیچیده‌ای است که ظرفیت بیشتری را برای یادگیری ماشینی‌ها فراهم می‌کنند؛ اما شاید اصلی‌ترین دلیل استقبال از این شیوه، تأثیر آن در حذف یا دستکم کاهش تأثیر مهندسی ویژگی‌های و افزایش امکان استفاده از داده‌های خام می‌باشد. در واقع مدل‌های پیچیده یادگیری عمیق با گذر داده‌های خام از پالایه لایه‌های غیرخطی متعدد و متفاوت، ویژگی‌های ذاتی و اصلی این داده‌ها را به‌طور خودکار استخراج می‌کنند و عملاً نیاز به استخراج دستی ویژگی‌های دست‌ساز را منتفی می‌نمایند.

در نخستین مطالعه‌ای که در مورد به‌کارگیری یادگیری عمیق برای کشف افعال انسانی انجام پذیرفته، ماشینی بولتزمن محدود^{۲۰} یا RBM در جهت استخراج خودکار ویژگی‌ها مورداستفاده قرار گرفته است [28]. این مطالعه روش‌های دیگری مثل PCA و FFT را با RBM مقایسه کرده و RBM در مورد هر چهار دادگان مورداستفاده از سایر روش‌ها پیشی گرفته است. پس از آن، مطالعات متعدد دیگری مثل [29] و [30] و [31] روش‌های مختلف و متفاوت یادگیری عمیق را در جهت استخراج خودکار ویژگی‌ها از داده‌های خام حسگرهای بی‌سیم مورداستفاده قرار داده و به نتایج قابل قبولی دست پیدا کردند. همچنین مرجع [32] با ترکیب یک شبکه عصبی عمیق از نوع RBM و مدل مخفی مارکوف روشی را ارائه کرده که نتایج آن از نتایج همه‌کاره‌ای پیشین بهتر است. تحقیقات یادشده علیرغم تفاوت‌ها در یک نکته شباهت دارند: همه آن‌ها از شبکه‌های عصبی عمیق کاملاً متصل

فرمول‌های زیر نشان می‌دهند که سلول‌های حافظه در هر گام زمانی چگونه بهنگام می‌شوند. در این فرمول‌ها f_t, i_t و o_t به ترتیب معرف درگاه‌های ورودی، فراموشی و خروجی هستند. W و b پارامترهای مدل، یعنی وزن‌های اتصالات و بایاس‌ها می‌باشند. σ و \tanh توابع تانژانت هیپربولیک و سیگمید و عملگر ضرب عنصری را نشان می‌دهند.

ابتدا درگاه ورودی که در واقع یک لایه سیگمیدی است در مورد میزان تأثیر ورودی گام فعلی در وضعیت سلول تصمیم‌گیری می‌کند:

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} * C_{t-1} + b_i) \quad (2)$$

سپس \hat{C} به‌عنوان مقدار احتمالی و کاندید لحظه t تشکیل می‌شود

$$\hat{C} = \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \quad (3)$$

در قدم بعدی درگاه فراموشی میزان تأثیر وضعیت جاری در تشکیل وضعیت آتی سلول را ارزیابی می‌نماید:

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} * C_{t-1} + b_f) \quad (4)$$

اکنون میانگین وزن‌دار وضعیت موجود و وضعیت کاندید، تکلیف وضعیت سلول در گام بعدی را مشخص می‌کنند. چنانکه گفته شد، مقادیر درگاه‌های فراموشی و ورودی به ترتیب به‌عنوان وزن‌های این دو مقدار در نظر گرفته خواهند شد:

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \quad (5)$$

در نهایت درگاه خروجی، ورودی فعلی، خروجی بلوک قبل و وضعیت بهنگام شده را از پالایه یک‌لایه سیگمیدی عبور می‌دهد تا مقدار خروجی بلوک جاری به سمت بلوک مجاور بعدی را مشخص کند:

$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} * C_t + b_o) \quad (6)$$

$$h_t = o_t \odot \tanh(C_t) \quad (7)$$

۵- تجربه‌ها

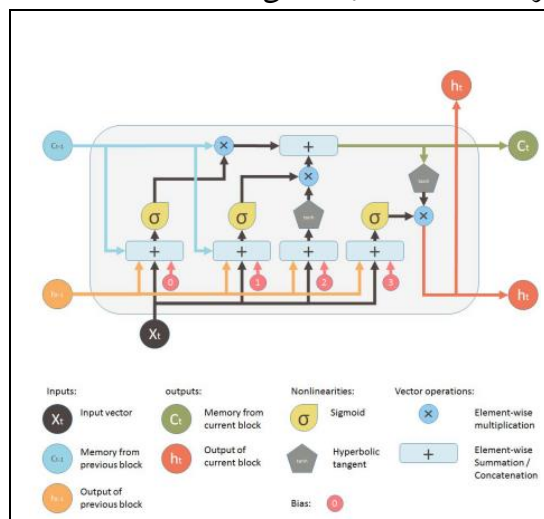
۵-۱- دادگان

ما برای انجام این مطالعه از یکی از دادگان رایج در حوزه کشف افعال استفاده کرده‌ایم. این دادگان که شامل افعال ساده یک نفره است، در آزمایشگاه داده‌کسای شبکه‌های حسگر بی‌سیم دانشگاه فورد هام ایجاد شده است و به دادگان ActiTracker یا WISDM شهرت دارد [16]. در فرایند ایجاد دادگان مذکور، ۳۶ داوطلب یک تلفن هوشمند مبتنی بر سیستم عامل اندروید را در جیب جلویی شلوار خود گذاشته و ۶ کار روزمره را به‌دفعات انجام داده‌اند. این افعال شامل راه رفتن، تکاپو^{۱۷}، بالا رفتن از پله، پایین آمدن از پله، نشستن و برخاستن است. در خلال انجام این افعال، حسگر شتاب سنج تلفن همراه با نرخ HZ ۲۰ (یک نمونه در هر ۵۰ ms) شتاب خطی کاربران را در طول سه محور X, Y, Z ثبت می‌کرده است. دادگان نهایی شامل ۱۰۹۸۲۰۷ رکورد است. توزیع افعال این دادگان در شکل ۲ نشان داده شده است.

ما این داده‌ها را به پنجره‌های لغزان و همپوشانی به طول یک ثانیه تقسیم کرده‌ایم. با توجه به نرخ نمونه‌برداری ۲۰ هرتزی داده‌ها، هر پنجره حاوی ۲۰۰ نمونه است. همچنین هر پنجره از پنجره قبلی خود

بعدها با اضافه شدن درگاه فراموشی کامل‌تر شد [44] در حل بسیاری از مسائل، از جمله طبقه‌بندی و کشف ناهنجاری مورد استفاده قرار گرفته است. شبکه‌های LSTM در واقع به‌جای نورون از واحدهای بزرگ‌تر و پیچیده‌تری به نام بلوک حافظه تشکیل شده‌اند. اصلی‌ترین مفهوم در مورد هر بلوک حافظه، وضعیت بلوک است که برآیند پیچیده و غیرخطی از ورودی بلوک، خروجی بلوک‌های مجاور،

وضعیت بلوک در گام‌های پیشین و وضعیت بلوک قبلی می‌باشد. هر بلوک حافظه از یک سلول حافظه و تعدادی درگاه تشکیل شده است. سلول حافظه وضعیت جاری بلوک را نگهداری می‌کند و درگاه‌ها شارش اطلاعات را به سمت داخل و خارج مدیریت می‌نمایند. شکل ۱ ساختار یک بلوک حافظه LSTM را نشان می‌دهد.



شکل ۱ - ساختار یک بلوک LSTM

چنانکه از شکل برمیآید، هر درگاه در واقع یک لایه شبکه عصبی با تابع فعال‌سازی (معمولاً) سیگمید است که خروجی آن در انتها به یک عملگر ضرب هامادار یا ضرب عنصری ختم می‌شود. در معماری هوشیدر [43] درگاه‌ها، ورودی جاری بلوک (x_t) و خروجی بلوک قبلی (h_{t-1}) را به‌عنوان ورودی دریافت می‌نمایند و با عبور آن‌ها از پالایه یک‌لایه سیگمیدی، عددی را در مقام خروجی تولید می‌کنند که در فاصله ۰ تا ۱ قرار دارد و مقدار آن معرف میزان تأثیرگذاری درگاه می‌باشد. درگاه ورودی تعیین می‌کند که کدام داده‌ها باید در سلول حافظه بلوک ذخیره شوند، درگاه فراموشی تصمیم می‌گیرد که این داده‌ها تا چه زمانی باید در بلوک باقی بمانند و درگاه خروجی مشخص می‌نماید که کدام داده‌ها باید از بلوک خارج بشوند. لازم به ذکر است که نویسندگان مرجع [44] با ملاحظه کاستی‌های موجود در بلوک‌های LSTM زمان خود، وضعیت جاری بلوک را نیز به مجموعه ورودی‌های هر درگاه اضافه کردند و با نام‌گذاری اتصالات مذکور به اتصالات آفتابگردان رایج‌ترین نوع بلوک‌های LSTM، یعنی LSTM آفتابگردانی^{۱۸} را پدید آوردند. بلوک‌هایی که ما از آن‌ها استفاده کرده‌ایم از این نوع هستند.

یک بلوک LSTM را می‌توان به‌عنوان یک تابع F در نظر گرفت که با دریافت ورودی جاری (x_t)، وضعیت موجود (C_{t-1}) و خروجی بلوک قبلی (h_{t-1})، خروجی بلوک جاری (h_t) را تولید می‌کند:

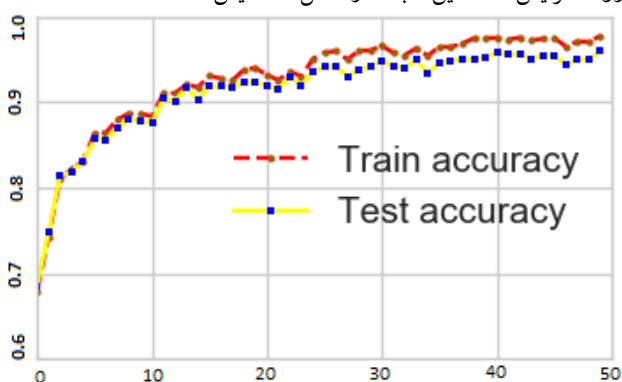
$$h_t = F(x_t, h_{t-1}, C_{t-1}) \quad (1)$$

تا برای لایه‌های طبقه بند قابل استفاده باشد. ما این ساختار را CNN نامیده‌ایم.

(ج) یک شبکه عصبی عمیق که از ترکیب یک لایه ورودی، دو لایه LSTM و یک لایه عادی تشکیل شده است:

LSTM (32) -> LSTM (32) -> Dense (6)

این شبکه که ساختار اصلی پیشنهادی مقاله حاضر است، داده‌ها را به صورت پنجره‌هایی با ابعاد $3 * 1 * 200$ دریافت می‌کند و آن‌ها را از دو لایه LSTM می‌گذراند. هر یک از این دو لایه شامل ۳۲ بلوک LSTM هستند و توسط این بلوک‌ها، ویژگی‌های ذاتی و درونی پنجره‌های مذکور را استخراج می‌کنند. توانایی عمده این شبکه‌ها بیشتر در کشف مجاورت‌ها، شباهت‌ها و وابستگی‌های زمانی بروز می‌نماید. این ویژگی‌ها پس از استخراج به یک لایه softmax تحویل داده می‌شود تا کار طبقه‌بندی و افراز آن‌ها به یکی از ۶ کلاس مربوطه را انجام بدهد. روند افزایش دقت این شبکه در شکل ۳ نمایش داده شده است.

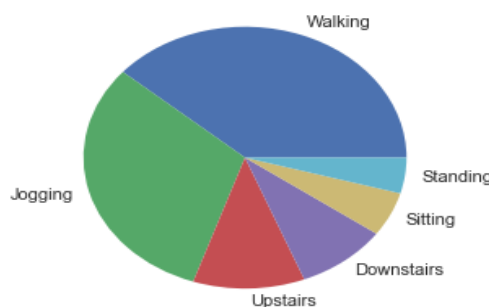


شکل ۳- روند افزایش دقت در شبکه LSTM

صرف‌نظر از تفاوت‌های ساختاری این شبکه‌ها، پارامترهای فرایندهای آموزش و آزمایش آن‌ها در تجربیات ما یکسان بوده است. هر یک از این شبکه‌ها، ۱۲ بار برای طبقه‌بندی افعال ۶ گانه مورد مطالعه بکار گرفته شده‌اند. در هر اجرا داده‌های ورودی به دو بخش ۷۰ و ۳۰ درصدی تقسیم شده که بخش اول مجموعه آموزشی و بخش دوم مجموعه آزمایشی اجرا را تشکیل داده‌اند. داده‌ها در بسته‌های ۱۰۰۰ تایی تحویل شبکه‌ها شده‌اند (Batch size = 1000) و هر شبکه ۵۰ بار تحت آموزش قرار گرفته است (Epochs = 50). نهایتاً در همه موارد نرخ یادگیری برابر ۰/۰۲۵ بوده و از تابع آنتروپی متقابل دودویی به عنوان تابع خطای شبکه استفاده شده است.

علاوه بر ساختارهای سه‌گانه فوق‌الذکر، ما به منظور مقایسه نتایج به دست آمده، نتایج مرجع [16] را عیناً نقل کرده‌ایم. البته نویسندگان منبع مذکور که دادگان مقاله حاضر از کار ایشان اخذ شده است، برای کشف و طبقه‌بندی افعال از داده‌های خام استفاده نکرده‌اند. آن‌ها این داده‌ها را به پنجره‌های غیر همپوشانی به طول ۱۰ ثانیه (شامل ۲۰۰ نمونه) تقسیم نموده و از هر پنجره ۴۳ ویژگی آماری را استخراج کرده‌اند. سپس سه روش طبقه‌بندی درخت تصمیم (J48)، رگرسیون لجستیک و شبکه عصبی چندلایه را بر روی این مجموعه داده جدید اعمال کرده‌اند. نتایج کار ایشان به همراه نتایج روش‌های عمیق مقاله حاضر در جدول ۱ نمایش داده شده است.

۴۰ نمونه جلوتر است، بنابراین هر یک از پنجره‌ها، با پنجره‌های مجاور خود ۸۰٪ همپوشانی دارد.



شکل ۲- توزیع کلاس‌های افعال در دادگان Actitrack

۲-۵- پیاده‌سازی

ما در مطالعه حاضر، سه ساختار عمیق را پیاده‌سازی کرده و با یکدیگر مقایسه نمودیم:

(الف) یک شبکه عصبی پیش‌خور عادی عمیق^{۱۸} که از ۵ لایه تشکیل شده است. برای تغذیه این شبکه، پنجره‌های ورودی به شکل تخت درآمده‌اند. می‌دانیم که هر پنجره در واقع حاوی ۲۰۰ نمونه از مقادیر x و y است؛ بنابراین هر پنجره را می‌توان به عنوان یک ماتریس دو بعدی $3 * 200$ در نظر گرفت. ما در حالت اخیر، این ماتریس دو بعدی را به شکل سطری امتداد داده و آن را به یک بردار ۶۰۰ عنصری تبدیل کرده‌ایم. سپس این بردارها را به یک شبکه عصبی عمیق تحویل داده‌ایم که ساختار آن به صورت (۶, ۸, ۱۲۰۰, ۱۲۰۰, ۱۲۰۰, ۶۰۰) است. تابع فعال‌سازی این شبکه در لایه‌های دوم و سوم ReLU، در لایه چهارم سیگمید و در لایه پنجم softmax است. برای اینکه شبکه به دام بیش‌برازش^{۱۹} در نیفتد، ما از روش Dropout با نرخ ۰/۵ بهره‌جسته‌ایم. ما این ساختار را DNN می‌نامیم.

Input (600) -> Dense (1200) -> Dense (1200) -> Dense (8) -> Dense (6)

(ب) یک شبکه کانولوشنی که شامل دو لایه کانولوشن، دو لایه نمونه‌برداری و دو لایه عادی است. این لایه‌ها در آرایشی به شکل زیر چیده شده‌اند:

Conv () -> MaxPool () -> Conv () -> MaxPool () -> Flatten () -> Dense (8) -> Dense (6)

داده‌های ورودی در قالب پنجره‌های یک‌ثانیه‌ای ($3 * 200$) در اختیار شبکه قرار می‌گیرند و بخش کانولوشن، شامل چهار لایه اول ویژگی‌های اصلی این پنجره‌ها را استخراج می‌کند. با توجه به شیوه عملکرد شبکه‌های کانولوشنی، این ویژگی‌ها بیشتر از سنخ مجاورت‌ها و قرابت‌های مکانی هستند. سپس خصایص یادشده برای طبقه‌بندی در اختیار دو لایه آخر که از نوع شبکه‌های عصبی عادی هستند، قرار می‌گیرد. لایه تخت‌سازی که در میان این دو بخش قرار گرفته، داده‌های دو بعدی که از بخش اول خارج می‌شوند را به شکل یک بعدی درمی‌آورد

- [7] M. Baldauf, S. Dustdar, and F. Rosenberg, "A survey on context-aware systems," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 2, no. 4, p. 263, 2007.

جدول ۱ - مقایسه دقت روش‌های مختلف طبقه‌بندی						
	% Of Records Correctly Predicted					
	J48 [16]	Logistic Regression [16]	MLP [16]	LSTM	CNN	Dense
Walking	89.9	93.6	91.7	98.0	89.0	74.0
Jogging	96.5	98.0	98.3	99.0	92.0	94.9
Upstairs	59.3	27.5	61.5	85.0	55.0	36.0
Downstairs	55.5	12.3	44.3	85.0	48.0	50.0
Sitting	95.7	92.2	95.0	99.0	87.0	100
Standing	93.3	87.0	91.9	96.0	75.0	87.0
Overall	85.1	78.1	91.7	96.0	80.0	76.0

- [8] C. Ladha, N. Y. Hammerla, P. Olivier, and T. Plötz, "ClimbAX," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13*, 2013, p. 235.
- [9] R. San-Segundo, J. M. Montero, R. Barra-Chicote, F. Fernández, and J. M. Pardo, "Feature extraction from smartphone inertial signals for human activity segmentation," *Signal Processing*, vol. 120, pp. 359–372, 2016.
- [10] R. G. Stockwell, L. Mansinha, and R. Lowe, *Localisation of the complex spectrum: The S Transform*, vol. 44, 1996.
- [11] H. P. Martinez, Y. Bengio, and G. Yannakakis, "Learning deep physiological models of affect," *IEEE Comput. Intell. Mag.*, vol. 8, no. 2, pp. 20–33, May 2013.
- [12] "Physical activity recognition by smartphones, a survey," *Biocybern. Biomed. Eng.*, vol. 37, no. 3, pp. 388–400, Jan. 2017.
- [13] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A Public Domain Dataset for Human Activity Recognition Using Smartphones."
- [14] X. Yun, J. Calusdian, E. R. Bachmann, and R. B. McGhee, "Estimation of Human Foot Motion During Normal Walking Using Inertial and Magnetic Sensor Measurements," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 7, pp. 2059–2072, Jul. 2012.
- [15] B. Yamansavaşçılar and M. A. Güvensan, "Activity Recognition on Smartphones: Efficient Sampling Rates and Window Sizes," *2016 IEEE Int. Conf. Pervasive Comput. Commun. Work. PerCom Work. 2016*, no. PerMoby, 2016.
- [16] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explor. Newsl.*, vol. 12, no. 2, p. 74, 2011.
- [17] T. Huynh and B. Schiele, "Analyzing features for activity recognition," in *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence innovative context-aware services: usages and technologies - sOc-EUSAI '05*, 2005, p. 159.
- [18] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 871–879, Mar. 2009.
- [19] Z. He and L. Jin, "Activity recognition from acceleration data based on discrete cosine transform and SVM," in *2009 IEEE International Conference on Systems, Man and Cybernetics*, 2009, pp. 5041–5044.
- [20] A. M. Khan, Young-Koo Lee, S. Y. Lee, and Tae-Seong Kim, "A Triaxial Accelerometer-Based Physical-Activity Recognition via Augmented-Signal Features and a Hierarchical Recognizer," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 5, pp. 1166–1172, Sep. 2010.
- [21] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, "A symbolic representation of time series, with implications for streaming algorithms," in *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery - DMKD '03*, 2003, p. 2.
- [22] W. Wu, S. Dasgupta, E. E. Ramirez, C. Peterson, and G. J. Norman, "Classification Accuracies of Physical Activities Using Smartphone Motion Sensors," *J. Med. Internet Res.*, vol. 14, no. 5, p. e130, Oct. 2012.

۶- نتیجه

در این مقاله ما پیشنهاد کردیم که برای کشف و بازشناسی افعال انسانی از روی داده‌های حسگرهای تلفن‌های همراه - به‌ویژه شتاب سنسور - از یادگیری عمیق استفاده شود. پیشنهاد مشخص این مقاله به‌کارگیری LSTM در کاربرد یادشده بود. نتایج تجربی به‌دست‌آمده نشان می‌دهد که یادگیری عمیق روش در کشف افعال انسانی بسیار خوب عمل می‌کند و در میان ساختارهای مورد استفاده LSTM بالاترین دقت را دارد. طبیعی است که برای مطالعه قابلیت و ثبات روش پیشنهادشده، به داده‌های بیشتری نیاز است.

به‌کارگیری LSTM و سایر ساختارهای عمیق بر روی داده‌هایی که شامل افعال بیشتری هستند، تنظیم پارامترهای این ساختارها در جهت کارایی بیشتر و ارائه راه‌حل برای کاهش منابع مورد نیاز و زمان اجرای ساختارهای عمیق، بخصوص LSTM می‌تواند موضوع مطالعات آینده باشد.

مراجع

- [1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Futur. Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013.
- [2] K. J. Lyytinen, Y. Yoo, U. Varshney, M. Ackerman, and G. Davis, "Communications of the Association for Information Systems Surfing the Next Wave: Design and Implementation Challenges of Ubiquitous Computing," *Commun. Assoc. Inf. Syst.*, vol. 13, 2004.
- [3] M. Á. Álvarez de la Concepción, L. M. Soria Morillo, J. A. Álvarez García, and L. González-Abriol, "Mobile activity recognition and fall detection system for elderly people using Ameva algorithm," *Pervasive Mob. Comput.*, vol. 34, pp. 3–13, 2017.
- [4] M. S. Goodwin, M. Haghghi, Q. Tang, M. Akcakaya, D. Erdogmus, and S. Intille, "Moving Towards a Real-time System for Automatically Recognizing Stereotypical Motor Movements in Individuals on the Autism Spectrum Using Wireless Accelerometry," *Proc. 2014 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput.*, pp. 861–872, 2014.
- [5] F. Casamassima, A. Ferrari, B. Milosevic, P. Ginis, E. Farella, and L. Rocchi, "A Wearable System for Gait Training in Subjects with Parkinson's Disease," *Sensors*, vol. 14, no. 4, pp. 6229–6246, Mar. 2014.
- [6] U. Maurer, A. Rowe, A. Smailagic, and D. P. Siewiorek, "eWatch: A Wearable Sensor and Notification Platform," in *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)*, pp. 142–145.

زیرنویس‌ها

- 1 Internet of Things
- 2 Context aware
- 3 Pervasive Computing
- 4 Human Activity Recognition
- 5 Low Pass and High Pass Filter
- 6 Heuristic
- 7 Wavelet
- 8 Discrete Cosine Transformation
- 9 Autoregressive Model
- 10 Ensemble of Classifiers
- 11 Restricted Boltzmann Machine
- 12 Convolutional Neural Network
- 13 Recurrent Neural Network
- 14 Pooling Layer
- 15 Long Short Term Memory
- 16 Peephole LSTM
- 17 Jogging
- 18 Deep Feed forward Neural Network
- 19 Overfitting

- [23] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine," Springer, Berlin, Heidelberg, 2012, pp. 216–223.
- [24] T. Westeyn, K. Vadas, X. Bian, T. Starner, and G. D. Abowd, "Recognizing mimicked autistic self-stimulatory behaviors using HMMs," in *Proceedings - International Symposium on Wearable Computers, ISWC*, 2005, vol. 2005, pp. 164–167.
- [25] L. Fan, Z. Wang, and H. Wang, "Human Activity Recognition Model Based on Decision Tree," in *2013 International Conference on Advanced Cloud and Big Data*, 2013, pp. 64–68.
- [26] C. Catal, S. Tufekci, E. Pirmit, and G. Kocabag, "On the use of ensemble of classifiers for accelerometer-based activity recognition," *Appl. Soft Comput. J.*, vol. 37, pp. 1018–1022, 2015.
- [27] M. N. S. Zainudin, M. N. Sulaiman, N. Mustapha, and T. Perumal, "Activity recognition based on accelerometer sensor using combinational classifiers," in *2015 IEEE Conference on Open Systems (ICOS)*, 2015, pp. 68–73.
- [28] T. Plötz, N. Y. Hammerla, and P. Olivier, "Feature learning for activity recognition in ubiquitous computing," in *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, 2011, vol. 22, no. 1, p. 1729.
- [29] "Using unlabeled data in a sparse-coding framework for human activity recognition," *Pervasive Mob. Comput.*, vol. 15, pp. 242–262, Dec. 2014.
- [30] C. Vollmer, H.-M. Gross, and J. P. Eggert, "Learning Features for Activity Recognition with Shift-Invariant Sparse Coding," Springer, Berlin, Heidelberg, 2013, pp. 367–374.
- [31] Y. Li, D. Shi, B. Ding, and D. Liu, "Unsupervised Feature Learning for Human Activity Recognition Using Smartphone Sensors," Springer, Cham, 2014, pp. 99–107.
- [32] L. Zhang, X. Wu, and D. Luo, "Human activity recognition with HMM-DNN model," in *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*, 2015, pp. 192–197.
- [33] N. Y. Hammerla, S. Halloran, and T. Ploetz, "Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables," *Ijcai*, pp. 1533–1540, 2016.
- [34] M. Zeng *et al.*, "Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors," *Proc. 6th Int. Conf. Mob. Comput. Appl. Serv.*, vol. 6, 2014.
- [35] J. Bo Yang, M. Nhut Nguyen, P. Phyo San, X. Li Li, and S. Krishnaswamy, "Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition," *Proc. IJCAI*, no. Ijcai, pp. 3995–4001, 2015.
- [36] Y. Chen and Y. Xue, "A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer," *2015 IEEE Int. Conf. Syst. Man, Cybern.*, pp. 1488–1492, 2015.
- [37] D. i, C. Wong, B. Lo, and G.-Z. Yang, "Deep Learning for Human Activity Recognition: A Resource Efficient Implementation on Low-Power Devices."
- [38] F. Eyben, M. Wollmer, B. Schuller, and A. Graves, "From speech to letters - using a novel neural network architecture for grapheme based ASR," in *2009 IEEE Workshop on Automatic Speech Recognition & Understanding*, 2009, pp. 376–380.
- [39] P. Doetsch, M. Kozielski, and H. Ney, "Fast and Robust Training of Recurrent Neural Networks for Offline Handwriting Recognition," in *2014 14th International Conference on Frontiers in Handwriting Recognition*, 2014, pp. 279–284.
- [40] L. Gao, Z. Guo, H. Zhang, X. Xu, and H. T. Shen, "Video Captioning With Attention-Based LSTM and Semantic Consistency," *IEEE Transactions on Multimedia*, vol. 19, no. 9, pp. 2045–2055, 2017.
- [41] M. Sundermeyer, H. Ney, and R. Schlüter, "From Feedforward to Recurrent LSTM Neural Networks for Language Modeling," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 3, pp. 517–529, 2015.
- [42] S. P. Singh, A. Kumar, H. Darbari, L. Singh, A. Rastogi, and S. Jain, "Machine translation using deep learning: An overview," *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, pp. 162–167, 2017.
- [43] S. Hochreiter and J. Urgan Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [44] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to Forget: Continual Prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.



خوشه‌بندی هوشمند شبکه‌های حسگر بی‌سیم با رویکرد الگوریتم فراابتکاری

سید مجتبی حسینی راد

استادیار، دانشکده فنی و مهندسی، گروه مهندسی کامپیوتر - دانشگاه پیام نور شاهرود- شاهرود

hosseinirad@se.pnu.ac.ir

چکیده

شبکه‌های حسگر بیسیم از بهم پیوستن صدها و شاید هزاران سنسور بیسیم کوچک که در یک محیط غیرقابل دسترس برای نظارت و کنترل پدیده‌های مختلف فیزیکی ایجاد می‌شوند. با توجه به محدودیت منابع سنسورهای بیسیم، ذخیره انرژی از مهمترین چالش‌های طراحی و ایجاد یک توپولوژی با بازدهی مطلوب می‌باشد. به دلیل عدم بکارگیری هرگونه مکانیسم خواب در الگوریتم‌های خوشه‌بندی (توزیع شده و متمرکز)، می‌توان با این بکارگیری این الگوریتم‌ها انرژی مصرفی شبکه را در معماری سلسله مراتبی به خوبی مدیریت کرد. در این مقاله با بهره‌گیری از مدل لایه‌ای و بکارگیری الگوریتم‌های ترکیبی فراابتکاری - PSO، GA، یک الگوریتم خوشه‌بندی با زمان‌بند خواب موثر به جهت مدیریت انرژی ارائه شده که قابل پیاده‌سازی در تمامی پروتکل‌های خوشه‌بندی است. نتایج تحقیق نشان می‌دهد که روش پیشنهادی با ارائه خوشه‌بندی مطلوب از کمینه حسگرهای فعال، بیشینه میزان باتری باقیمانده، حداکثر پوشش و حداقل حسگر فعال بدون سرخوشه در هر دوره زمانی توانسته است طول عمر شبکه را تا ۳ برابر افزایش دهد.

کلمات کلیدی

شبکه‌های حسگر بیسیم، زمان‌بند، مسیریابی، ذخیره انرژی.

امروزه علاوه بر محدودیت انرژی، کنترل توپولوژی، پوشش شبکه، تلفیق و تجمیع داده‌های جمع‌آوری شده، بهنگام‌سازی زمانی، مکان‌یابی و مسیریابی بهینه و امن از جمله چالش‌های مهم در طراحی شبکه‌های حسگر می‌باشند [۳]. به‌منظور ارتباط با محیط پیرامون و گره‌های همسایه، سنسورها به یک باتری کوچک، ثابت با توان کاری کم تجهیز شده‌اند. به دلیل دسترسی بسیار دشوار، این باتری‌های کوچک، غیرقابل تعویض و یا شارژ مجدد هستند و یک گره تا زمانی قادر است وظایف مختلف از قبیل جمع‌آوری، پردازش، تلفیق و ارسال داده‌ها را انجام دهد که منابع انرژی در اختیار داشته و با اتمام این منابع به نوعی سنسور می‌میرد [۴].

یکی از توپولوژی‌های موثر و دارای بازده بالای انرژی در شبکه‌های حسگر، توپولوژی سلسله‌مراتبی به جهت خوشه‌بندی گره‌های سنسور بوده و الگوریتم‌ها و روش‌های گوناگونی برای خوشه‌بندی در شبکه‌های حسگر بیسیم معرفی شده‌اند. در طراحی و ایجاد یک الگوریتم خوشه‌بندی، اگر علاوه بر میزان انرژی مصرفی، سایر پارامترهای مهم شبکه از قبیل پوشش شبکه، انتخاب گره‌ها با بیشترین میزان باتری باقیمانده، کمترین میزان افزونگی در داده‌های جمع‌آوری شده و غیره با

۱- مقدمه

شبکه‌های حسگر بی‌سیم متشکل از تعداد زیادی گره‌های حسگر ارزان قیمت پخش شده در یک محیط غیرقابل دسترس با هدف نظارت و کنترل پدیده‌های مختلف فیزیکی از قبیل دما، رطوبت، صدا، تصویر، فشار و غیره و جمع‌آوری اطلاعاتی در مورد محیط پیرامون سنسورها پراکنده شده‌اند، ایجاد می‌شوند. به دلیل استفاده این شبکه‌ها در مکان‌های غیرقابل دسترس، غالباً مکان قرارگیری این سنسورها از قبل معین و مشخص نبوده و به صورت تصادفی در محیط کاربردی پخش می‌شوند تا با رصد پدیده‌های مختلف و گردآوری داده از وضعیت آنان، به صورت مسقیم و یا از طریق گره‌های واسط به گره مقصد که ایستگاه اصلی نامیده می‌شود، ارسال نمایند [۱]. شبکه‌های حسگر بی‌سیم در کاربردهای متعددی نظارت ترافیکی، نظارت بر شرایط محیطی، نظارت بر سلامت ساختارها یا تجهیزات یک سیستم، اتوماسیون صنعتی، روبات‌ها، حفاظت نظارتی-مراقبتی به‌کار گرفته می‌شوند [۲].

مشخص شده است، ولی در بیشتر الگوریتم‌های خوشه بندی به علت احتمالی و یا تصادفی بودن انتخاب سرخوشه‌ها در شکل‌گیری آنان، تعداد خوشه‌هایی کاملاً متفاوت در هر بازه زمانی از طول عمر شبکه ایجاد می‌گردند [۸]. سرخوشه با زمان‌بند فعالیت درون خوشه‌ی (فعال-خواب) می‌تواند با تغییر وضعیت کاری گره‌ها به حالت خواب یا کم توان میزان انرژی مصرفی را کمینه کند. به همین صورت با به‌کارگیری چرخشی دوره‌ای گره‌ها می‌توان دوره زمانی خاصی را برای ارسال و دریافت داده‌ها تعیین کرد. لذا خوشه‌بندی با الگوریتم زمان‌بند خواب سبب اجتناب از افزونگی داده در محیط کاری، جلوگیری از تصادم در دسترسی کانال و ارسال دوباره داده‌ها می‌شود [۹].

انتخاب سرخوشه‌ها و شکل‌گیری خوشه‌ها، دو گام اصلی در فرآیند خوشه‌بندی است که در انجام دو مرحله چالش‌های متفاوتی وجود دارد. در مرحله ایجاد خوشه‌ها، سایر سنسورهای با توجه به شرایط مسئله برای اتصال به یک سرخوشه از میان سرخوشه‌های در دسترس تصمیم‌گیری می‌کنند. با توجه به مصرف بالای انرژی ارتباطی مستقیم میان گره‌های فرستنده و گیرنده و تناسب مستقیم آن با فاصله اقلیدسی میان گره‌ای، برای کاهش هزینه ارتباطی، باید سرخوشه‌ها در فاصله‌ای بهینه از ایستگاه کاری مکان‌یابی شوند [۱۰].

تجمع و انتقال داده‌ها از دیگر وظایف سرخوشه‌هاست که در صورت تکرار متوالی فعالیت سرخوشه‌ی یک گره باعث می‌شود انرژی گره به سرعت تخلیه شود. لذا برای انتخاب دوباره سرخوشه‌ها در هر دوره زمانی، ضروری به نظر می‌رسد که میزان انرژی باقیمانده و تعداد دوره‌های فعالیت سرخوشه‌ی گره‌ها در نظر گرفته شوند. تعداد مسیرهای داده‌ای طی شده توسط یک گره تا رسیدن به سرخوشه و اندازه خوشه که گره حسگر بر اساس برای اتصال به خوشه تصمیم‌گیری کند از مهمترین چالش‌ها در انتخاب سرخوشه هستند [۱۱].

در برخی از رویکردها، مفهوم روش سلسله مراتبی و یا خوشه‌بندی چند سطحی در برابر خوشه‌بندی تک سطحی مطرح شده که هدف آن پرکنندگی متوازن، توزیع و مصرف بهینه‌تر انرژی است. با توجه به اهمیت کارایی ارتباط برون خوشه‌ی در شبکه‌های بزرگ بهره‌جستن از خوشه بندی چند سطحی مورد توجه بیشتری قرار گرفته است. قابلیت اطمینان و کیفیت سرویس در شبکه‌های حسگر بی‌سیم نیز دارای مفاهیم متفاوتی است. سرویس‌های ارائه شده توسط شبکه‌های حسگر نامشخص هستند. بنابراین علاوه بر نرخ ارسال بسته، مقدار و کیفیت اطلاعات ارسال به گره‌ها معیارهای مهم ارزیابی می‌باشند [۱۲].

در محیط‌های همگن که همه گره‌ها قابلیت‌های یکسانی دارند، فرآیند شکل‌دهی خوشه و انتخاب سرخوشه به صورت توزیع شده، صحیح‌ترین تکنیک در راستای کسب انعطاف پذیری بیشتر و زمان‌های سریع‌تر اجرا یا همگرایی مستقل از تعداد گره‌ها می‌باشد. در مقابل، تعداد کمی از رویکردها از تکنیک‌های متمرکز شده یا ترکیبی استفاده می‌کنند که در آنها یک یا چند هماهنگ‌کننده یا ایستگاه اصلی، مسئول تفکیک کردن تمام شبکه به صورت منفصل و کنترل عضویت در خوشه‌ها هستند. در رویه شکل‌دهی خوشه پویا، هرگاه شامل انتخاب مجدد سرخوشه‌ها به صورت منظم دوره‌ای یا رخدادگرا باشد یا شامل رویه‌های سازماندهی مجدد خوشه‌ها باشد [۱۳].

ماهیتی متضاد در خوشه‌بندی گره‌ها در نظر گرفته شوند، طراحی به یک مشکل ان-پی سخت تبدیل شده که برای یافتن یک پاسخ نزدیک به مطلوب بهینه، استفاده از الگوریتم‌های فراابتکاری تکاملی پیشنهاد می‌شود. در این مقاله با توجه به با مزایای و معایب الگوریتم‌ها، از الگوریتم ترکیبی بهینه‌سازی ذرات-ژنتیک بهره گرفته شده تا میزان تنوع در جمعیت پاسخ الگوریتم به صورت مطلوبی افزایش یابد. همگرایی سریع، ارائه جستجویی قدرتمند و مناسب در فضای پاسخ گسترده با پیچیدگی محاسباتی پایین از مزیت‌های مهم این الگوریتم ترکیبی است [۵].

در این مقاله الگوریتم‌های مختلف تکاملی برای طراحی یک الگوریتم زمان‌بند خواب موثر مبتنی بر انرژی و افزونگی داده با هدف انتخاب کمینه تعداد سنسورهای فعال با بالاترین میزان انرژی باقیمانده و افزونگی در داده‌های جمع آوری شده و بیشترین میزان پوشش شبکه بکار گرفته شده است. سپس در لایه دوم روش پیشنهادی با بهره‌گیری از الگوریتم فراابتکاری ترکیبی سنسورهای فعال منتخب خوشه‌بندی شده و به دلیل نوآوری الگوریتم پیشنهادی و عدم وجود هرگونه روش مشابه و به جهت بررسی صحت آن، این روش با الگوریتم خوشه بندی WEEC مقایسه شده است.

نتایج حاصله نشان می‌دهد که با استفاده از الگوریتم‌های یک الگوریتم زمان‌بند خواب مبتنی بر معماری چندلایه‌ای و بهره‌گیری از آن توپولوژی سلسله مراتبی می‌توان عملکرد پروتکل‌های خوشه بندی را به صورت مطلوبی بهبود بخشید. افزایش مطلوب طول عمر، کاهش میزان افزونگی در داده‌های جمع آوری شده، داشتن پوشش حداکثری با حداقل تعداد سنسور فعال از نتایج این تحقیق می‌باشد.

۲- مروری بر ادبیات

در توپولوژی سلسله مراتبی گره‌های سنسور شبکه در دسته‌هایی که خوشه نامیده می‌شوند، گروه بندی می‌شوند خوشه‌بندی در شبکه‌های حسگر، می‌تواند به‌طور قابل توجهی میزان بازدهی انرژی، طول عمر و مقیاس پذیری شبکه را افزایش داده و روشی موثر جهت کاهش انرژی مصرفی درون یک خوشه با جمع و تلفیق داده‌های گردآوری شده در جهت کاهش تعداد پیام‌های ارسالی به ایستگاه اصلی می‌باشد. لذا استفاده از روش‌های خوشه‌بندی در کاربردهایی که نیاز به جمع‌آوری داده‌ها دارند، بسیار سودمند هستند [۶].

خوشه‌بندی قادر است با محدودسازی تعاملات درون خوشه‌ای و کاهش میزان افزونگی داده‌های انتقالی میان سنسورها پهنای باند ارتباطی را به خوبی حفظ کند. بیشتر الگوریتم‌های پیشنهادی خوشه بندی، انتقال داده‌ها را در مسیر داده‌ای تک‌گام پیشنهاد می‌کنند که ممکن است با افزایش چگالی سنسورها، باعث ترافیک و ازدحام در گره‌چاهک و یا ایستگاه اصلی شود. این ترافیک ممکن است باعث تاخیر در ارتباطات و پیگیری ناموفق رخدادها گردد. از این رو برای افزایش مقیاس پذیری، توازن بار و بهره‌برداری مناسب از منابع را ضروری می‌نماید [۷].

تعداد خوشه‌ها، چگونگی پراکندگی، تعداد اعضا و ارتباطات درون خوشه‌ای از پارامترهای مهم در بررسی و دسته‌بندی خوشه‌ها هستند. در بعضی از روش‌های پیشنهادی، تعداد سرخوشه‌ها از قبل معین و

روشها و عملیات محاسباتی و تکرار پاسخ‌های تصادفی ایجاد شده برای جستجو و یافتن راه حل‌هایی بهینه در فضای پاسخ ممکن استفاده می‌کنند [۱۹]. بیشتر الگوریتم‌های تکاملی برای حل مسائل پیوسته ارائه شده‌اند و تنها از الگوریتم ژنتیک (GA)، الگوریتم بهینه‌سازی ذرات (PSO) و رقابت استعماری (ICA) می‌توان در حل مسائل گسسته بهره جست.

الگوریتم بهینه‌سازی ذرات یک الگوریتم ذاتا پیوسته، بسیار ساده و دارای قدرت محاسباتی قوی است که در ابتدا برای یافتن پاسخ و حل مسائل پیوسته معرفی گردید. این الگوریتم در الگوریتم‌های فراابتکاری دسته بندی شده و می‌توان با انجام برخی تغییرات از آن برای حل مسائل گسسته و یافتن پاسخ‌های نزدیک مطلوب بهره جست. در این الگوریتم فرض شده است که تعدادی از موجودات زنده که ذره نامیده و در فضای جستجو پخش شده و هر ذره مقدار تابع هدف را در موقعیتی از فضا که در آن قرار گرفته است محاسبه کرده و سپس با استفاده از ترکیب اطلاعات محل فعلی اش و بهترین محلی که قبلاً در آن بوده و همچنین اطلاعات یک یا چند ذره از بهترین ذرات موجود در جمع، جهتی را برای حرکت انتخاب کند. پس از انجام حرکت جمعی، یک مرحله از الگوریتم به پایان رسیده و این مراحل تا بدست آمدن جواب مطلوب تکرار می‌شوند. بهره‌مندی از حافظه، همکاری و اشتراک گذاری اطلاعات بین ذرات، سرعت همگرایی بالا، انعطاف‌پذیری بهتر در برابر مشکل بهینه محلی، راحتی پیاده‌سازی و اجرا و همگرایی زودرس از مزایای استفاده از الگوریتم PSO و کاهش تنوع جمعیت از مهمترین معایب آن می‌باشد.

الگوریتم‌های تکاملی از دو مرحله اصلی جستجو و استخراج تشکیل شده‌اند. بهره‌جستن از یک مکانیسم جستجوی قدرتمند می‌تواند پاسخ‌های ممکن بیشتری را در فضای جستجو بررسی و ارزیابی نماید. همگرایی سریع الگوریتم نیز در کارایی آن بسیار با اهمیت است. همان‌گونه که بیان شد یکی از مهمترین معایب الگوریتم PSO کاهش تنوع الگوریتم می‌باشد که فضای جستجو را کاهش می‌دهد، در صورتی که سایر الگوریتم‌ها مانند الگوریتم ژنتیک یکی از قدرتمندترین الگوریتم‌هایی تکاملی است که می‌تواند با ایجاد جمعیت‌های متنوع جدید پاسخ‌های بسیاری را در فضای جستجو بیابد. در این تحقیق سعی شده است تا با بهره‌گرفتن از مکانیسم‌های جستجو (چرخ رولت)، ترکیب یکنواخت و جهش الگوریتم ژنتیک در الگوریتم PSO و ایجاد یک الگوریتم فراابتکاری ترکیبی به نام PSO-GA، با رفع مشکل کاهش جمعیت، کارایی الگوریتم را به شکل مطلوب افزایش یابد. لذا بعد از اتمام هر دوره اجرای الگوریتم PSO، این مراحل ترکیب یکنواخت و جهش الگوریتم ژنتیک را اعمال کرده و سپس شرایط خاتمه را بررسی می‌شود.

۳- الگوریتم پیشنهادی

در این مقاله برای طراحی یک پروتکل خوشه‌بندی m لایه‌ای، در مرحله اول طراحی یک الگوریتم زمان‌بند متمرکز بر اساس پارامترهای مهم یک شبکه حسگر مبتنی بر معماری لایه‌ای پیشنهاد می‌شود تا در شبکه‌های حسگر با نوع پراکندگی تصادفی و چگالی بالا بتوان با حداقل تعداد

بیشتر الگوریتم‌های شناخته شده خوشه‌بندی، بر اساس معیارهای شکل دهی خوشه و پارامترهای استفاده شده جهت انتخاب سرخوشه می‌توانند به دو دسته اصلی الگوریتم‌های خوشه‌بندی یا احتمالاتی (تصادفی، ترکیبی) و غیر احتمالاتی (قطعی) تقسیم شوند. در الگوریتم‌های خوشه‌بندی احتمالاتی، جهت تعیین سرخوشه‌های آغازین، یک مقدار احتمالی به هر گره تخصیص داده می‌شود. ممکن است معیارهای ثانوی دیگری هم چه در خلال فرآیند انتخاب سرخوشه مثل انرژی باقی مانده و چه در طول فرآیند شکل دهی خوشه مثل نزدیکی یا هزینه ارتباط و همگی در جهت رسیدن به مصرف انرژی بهتر و طول عمر شبکه وجود داشته باشند [۱۴].

الگوریتم‌های خوشه‌بندی احتمالاتی، فرآیند بازدهی انرژی بالاتر، معمولاً باعث سریع شدن زمان‌های اجرا یا همگرایی و کاهش حجم پیام‌های تبادل می‌گردند. در الگوریتم‌های خوشه‌بندی غیراحتمالی، اصولاً معیارهای قطعی ویژه بیشتری برای انتخاب سرخوشه و شکل دهی خوشه مورد توجه قرار گرفته‌اند؛ این معیارها اساساً مبتنی بر نزدیکی و مجاورت گره‌ها مثل اتصال و درجه و اطلاعات دریافت شده از گره‌های نزدیکتر می‌باشند. در الگوریتم‌ها، یک کلاس نسبتاً جدید و کاملاً چالش برانگیز به نام پروتکل‌های الهام گرفته از زیست‌شناسی معرفی شده که بر پایه هوش گرومی بوده و تا به امروز احتمالاً امیدبخش‌ترین رویکرد قابل انتخاب برای خوشه‌بندی شبکه‌های حسگر به شمار می‌رود [۱۵].

برای کاهش میزان انرژی مصرفی گره‌های سنسور الگوریتم‌های زمان‌بندی خواب مختلفی پیشنهاد شده است. با توجه به میزان مصرف انرژی یکسان در حالات ارسال و بیکار در مدت زمان عدم وجود داده، خاموش کردن فرستنده رادیویی گره بسیار مطلوب خواهد بود. یک الگوریتم زمان‌بند خواب تلاش می‌کند تا با تغییر حالت کاری و یا حتی خاموش کردن قسمتهای مختلف سنسور در یک دوره کاری و تمرکز بر مدیریت کارایی فرستنده و گیرنده رادیویی، میزان انرژی مصرفی گره را کمینه کند تا گره سنسور بتواند با حداقل میزان انرژی به فعالیت ادامه دهد [۱۶]. به بیان دیگر این الگوریتم‌ها با به کار بستن تکنیکها و مکانیسم‌های مختلف، فعالیت و توان قسمتهای مختلف گره را کنترل می‌کنند. بنابراین طراحی و بکارگیری الگوریتم‌های زمان‌بند خواب موثر برای کاهش اتلاف انرژی مصرفی و کنترل انرژی مصرفی برای تبادل داده‌ها میان گره‌های سنسور بسیار ضروری و اجتناب ناپذیر است [۱۷].

فقط تعداد بسیار اندکی از توپولوژیهای خوشه بندی متمرکز که با کمک الگوریتم‌های فراابتکاری طراحی و ایجاد شده‌اند، حالت خواب را در حالات کاری سنسور تعریف کرده‌اند. همچنین در این روشها هیچ پارامتر خاصی برای ارزیابی و سنجش کارایی حالت خواب و تاثیر آن بر توپولوژی و شبکه تعریف نشده است. به همین صورت در نتایج این تحقیقات، هیچ جدول و یا نموداری برای ارزیابی حالت خواب و تاثیرات آن بر روی سایر پارامترهای شبکه از جمله پوشش، اتصال و افزونگی داده و یا سایر پارامترها ارائه نشده است [۱۸].

در طراحی شبکه‌های حسگر با الگوریتم فراابتکاری، غالباً یک مدل شبکه به صورت یک مسئله بهینه‌سازی چند هدفه پیشنهاد می‌شود. الگوریتم‌های تکاملی زیرمجموعه‌ای از محاسبات تکاملی بوده که از

باید متذکر شد که پارامتر Π_1 مهمترین و تاثیرگذارترین پارامتر در تابع هزینه برای یافتن یک چیدمان سنسورهای فعال بهینه می‌باشد. پوشش محیط کاری یکی دیگر از پارامترهای بسیار مهم در شبکه‌های حسگر بی‌سیم بوده و همواره سعی می‌شود پوشش حداکثری با توجه به شرایط شبکه ایجاد شود.

$$\Pi_1 = \frac{\sum_{i=1}^m \sum_{j=1}^n W_{ij} \cdot \Omega_{ij}}{\sum_{k=1}^{m \times n} EP_k} \quad (3)$$

همچنین اگر میزان پوشش از یک حد آستانه کمتر شود دیگر وجود شبکه حسگر فاقد هرگونه معنی بوده و شبکه متوقف و از بین می‌رود. در این تحقیق نیز برای تصمیم‌گیری اتمام عملکرد شبکه با توجه به میزان پوشش آن صورت می‌پذیرد. برای به دست آوردن میزان درصد عدم پوشش شبکه (Π_2) کافی است تا تعداد درایه‌های ماتریس افزونگی داده (Ω) که برابر با صفر می‌باشد، را محاسبه کنیم (معادله ۴).

$$\Pi_2 = \frac{(\sum_{i=1}^m \sum_{j=1}^n \Omega_{ij}) \times 100}{m \times n} \quad \text{where } r_{ij} < 1 \quad (4)$$

پارامتر بسیار مهم دیگر میزان باتری باقیمانده در گره‌های سنسور می‌باشد. از دیگر اهداف الگوریتم پیشنهادی فعالیت گره‌هایی با انرژی باقیمانده بیشتر و خواب گره‌هایی با انرژی باقیمانده کمتر می‌باشد. اگر فرض کنیم تعداد n گره زنده در شبکه وجود داشته باشد، الگوریتم پیشنهادی باید تعداد k گره را از n گره به‌صورتی انتخاب کند که علاوه بر کاهش میزان افزونگی داده، گره‌های فعال با حداکثر میزان باتری باقیمانده نسبت به گره‌های خواب انتخاب شوند. اگر پارامتر α میزان متوسط باتری باقیمانده گره‌های فعال و پارامتر β میزان متوسط باتری باقیمانده گره‌های خواب را نمایش دهند، نسبت این دو پارامتر (Π_3) در تابع هزینه همواره باید مقداری کوچکتر از یک باشد که در رابطه (۵) بیان داده است.

$$\Pi_3 = \frac{\frac{1}{n} \times \sum_{i=1}^n \beta_i}{\frac{1}{k} \times \sum_{i=1}^k \alpha_i} \quad (5)$$

لذا با توجه به پارامترهای تعریف شده Π_1 تا Π_3 ، تابع هزینه در رابطه (۶) تعریف می‌شود.

$$\text{Cost Function} = \min \left\{ \sum_{i=1}^3 W_i \times \Pi_i \right\} \quad (6)$$

که در آن پارامترهای تعریف مسئله و W_i وزنهای مختلف این پارامترهاست که بر اساس آزمون صحیح و خطا به دست می‌آید (جدول ۱). در مرحله دوم برای خوشه‌بندی مجموعه سنسورهای فعال، باید پارامترهای Ω_1 تا Ω_7 در ادامه تعریف می‌شوند. یکی از پارامترهای مهم در خوشه‌بندی میزان اتصال شبکه و تعداد گره‌های معمولی است که در دامنه رادیویی خود نتوانسته‌اند هیچ سرخوشه‌ای را بیابند. اگر یک سنسور فعال در شعاع رادیویی خود سرخوشه‌ای را نیابد، گره خارج از دامنه نامیده شده و با S_{out} نمایش داده می‌شود.

جدول (۱): اوزان مختلف پارامترهای Π_1 تا Π_3 .

Algorithm	Π_1	Π_2	Π_3
GA	10^3	10	10^3

سنسور فعال دارای مقدار بیشینه باتری باقیمانده به پوشش و اتصالات حداکثری و کمینه میزان افزونگی داده رسید. به بیان دیگر ارائه روش لایه‌ای برای طراحی توپولوژی سلسله مراتبی در شبکه‌های حسگر و پیشنهاد یک الگوریتم زمان‌بند-خوشه بندی ترکیبی مطلوب با بهره‌گیری از الگوریتم فراابتکاری ترکیبی حرکت ذرات-ژنتیک به جهت خوشه‌بندی هوشمند از اهداف این پژوهش می‌باشد. پس از اتمام ایجاد و خوشه‌بندی سنسورهای فعال، در آخرین مرحله سنسورها با استفاده از روش TDMA اطلاعات جمع آوری شده از محیط را برای گره چاهک تخصیص یافته در لایه بالاتر ارسال می‌کنند.

در حل مسائل با استفاده از الگوریتم‌های تکاملی یافتن پارامترهای مهم و موثر مسئله برای ایجاد یک تابع ارزیابی مطلوب به جهت سنجش و ارزیابی پاسخ‌های ممکن ایجاد شده از مهمترین و پیچیده‌ترین مرحله می‌باشد. از این رو پارامترهای موثر مختلفی برای انتخاب سنسورهای فعال در لایه اول به جهت ایجاد یک الگوریتم زمان‌بند خواب با راندمان بالا و در لایه دوم پارامترهای موثر برای خوشه‌بندی حسگرهای فعال منتخب تعیین و تشریح می‌شود.

در روش پیشنهادی، برای محاسبه افزونگی داده‌ها ابتدا محیط کاربردی به صورت مشبک و بر اساس چگالی حسگری سنسورها تقسیم بندی می‌شود و فرض است که در هر تقاطع یک رویداد قابل گزارش رخ داده است. با فرض ابعاد محیط کاربردی $m \times n$ (مترمربع) و چگالی پراکنندگی شبکه $1 \text{ sen}/m^2$ (شعاع حسگری دایره‌ای با شعاع $1m$)، آنگاه ماتریس مورد نظر نیز دارای $m \times n$ آرایه خواهد بود. با تشکیل ماتریسی متناسب برای این رخدادها به جهت محاسبه تعداد گزارشات آنها می‌توان ماتریس افزونگی داده‌ها (Ω) و سپس میزان متوسط افزونگی داده‌ها را محاسبه کرد که در معادله (۱) نمایش داده شده است.

$$\Omega = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \vdots & \vdots & \dots & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mm} \end{bmatrix}_{m \times m} \quad (1)$$

که قطر اصلی این ماتریس برابر صفر و درایه r_{ij} بیانگر مجموع گزارشات یک رویداد توسط گره‌های مجاور می‌باشد. مقادیر درایه‌های این ماتریس باید $1 \leq r_{ij} \leq 2$ باشد. در صورتی که مقدار درایه از یک کمتر باشد آن نقطه از محیط کاری پوشش داده نشده و برای مقادیر بزرگتر از دو دارای افزونگی داده خواهیم بود. به همین صورت برای هر درایه r_{ij} ، وزن w_{ij} متناسب با میزان برهم افتادگی سنسورها نیز محاسبه و پس از ایجاد ماتریس وزن W در نظر گرفته می‌شود. با توجه به دایره‌ای بودن شعاع حسگری، میزان برهم افتادگی حسگرها از رابطه (۲) محاسبه شده و در ماتریس Ω اعمال می‌گردد.

$$W_{ij} = R_{sen} - (2R_{sen}^2 \cos^{-1}(\frac{d_{ij}}{2R_{sen}}) - \frac{1}{2}d \sqrt{4R_{sen}^2 - d_{ij}^2}) \quad (2)$$

که در آن d_{ij} میزان فاصله اقلیدسی دو گره و R_{sen} بیانگر شعاع حسگری سنسورها می‌باشد. بنابراین میزان متوسط افزونگی داده (Π_1) را از می‌توان پس از اعمال وزن هر درایه ماتریس، ضرب اسکالر دو ماتریس افزونگی داده (Ω) و وزن (W)، از رابطه (۳) محاسبه حاصل می‌شود. که EP_k بیانگر مجموع تعداد رخدادها در هر تقاطع می‌باشد.

سرخوشه‌ها حداکثر اعضا سرخوشه‌ها، پارامتر Ω_6 تعریف شده است تا با جلوگیری از بهینه‌های محلی جلوگیری از خوشه‌های غیریکنواخت و خوشه‌هایی همگن با تعداد اعضای نزدیک به حداکثر تشکیل شده و از که از رابطه (۱۲) محاسبه می‌شود.

$$\Omega_6 = \sum_{j=1}^k (N_{CL-members})_j \quad \text{where } (N_{CL-members})_j < \frac{N_{CH-max}}{2} \quad (12)$$

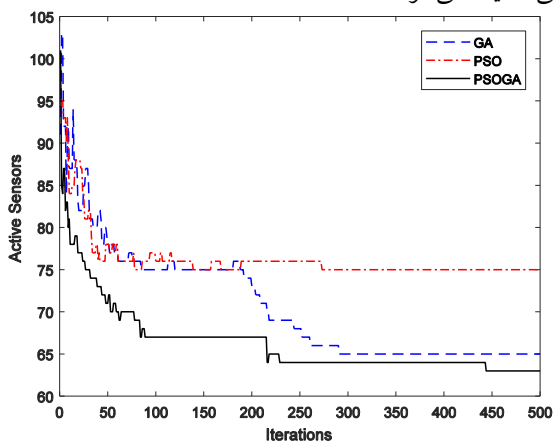
بنابراین تابع هزینه خوشه بندی متشکل از پارامترهای Ω_1 تا Ω_6 در رابطه (۱۳) بیان شده است.

$$CCost \text{ Funtion} = \min \left\{ \sum_{i=1}^6 w_i \Omega_i \right\} \quad (13)$$

که مقادیر اوزان w_1 تا w_6 با انجام تکرارهای متعدد و به صورت صحیح و خطا بدست می‌آید که در این تحقیق $w_1=10^5, w_2=10^5, w_3=10^5, w_4=10^2, w_5=10^3, w_6=10^3$ می‌باشد.

۴- ارزیابی و بررسی نتایج

با در نظر گرفتن شعاع حسگری ۱ متری برای سنسورها، میزان چگالی واحد پراکندگی، ۱ سنسور در هر مترمربع محاسبه و چگالی پراکندگی شبکه با توجه به تعداد سنسور استفاده شده در محیط و چگالی واحد شبکه معین می‌شود. لذا با توجه به شعاع حسگری، تعداد ۱۵۰ گره سنسور ثابت و همگن با توانایی‌های یکسان و برابر و به صورت تصادفی یکنواخت در یک محیط کاری غیرقابل دسترس با ابعاد 10×10 مترمربع پراکنده شده است. ایستگاه کاری در فاصله ای دور از شبکه قرار داشته که قادر است با تمامی گره‌ها تبادل اطلاعات نماید، در حالی که گره‌ها فقط در وضعیت فعالیت سرخوشه‌ای قادرند داده‌ها را به ایستگاه کاری ارسال نمایند. برای ارسال و دریافت، سنسورها از دو مدل کانال رادیویی، مدل کانال فضای آزاد و مسیره‌های داده چندگانه متقارن براساس فاصله بین انتقال دهنده و دریافت کننده بهره گرفته‌اند. پس از گسترانیدن سنسورها، گره‌ها مقادری اولیه می‌شوند. در ابتدا به منظور ارزیابی عملکرد الگوریتم زمان‌بند خواب، به هر گره یک مقدار باتری تصادفی بین صفر تا ۰/۵ ژول تخصیص داده می‌شود. برای ایجاد لایه اول و انتخاب گره‌های فعال در گام نخست ابتدا روش پیشنهادی با الگوریتم‌های GA، PSO، و PSOGA، پیاده سازی شده و نتایج آن بررسی مقایسه می‌شود.



شکل ۱: تعداد سنسورهای فعال الگوریتم‌های مختلف فراابتکاری.

لذا تعداد گره‌های خارج از دامنه Ω_1 بیانگر مجموع تعداد سنسورهای خارج از دامنه است که از رابطه (۷) محاسبه می‌شود.

$$\Omega_1 = \begin{cases} \sum_{j=1}^k (SN_{out})_j & \text{if } k > 0 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

تعداد سرخوشه‌ها در هر بازه زمانی از طول عمر باید متناسب با تعداد سنسورهای فعال بوده و اگر چنانچه سنسورها بیشتر از نیاز به وضعیت سرخوشه‌ای تغییر وضعیت دهند، برخی از سرخوشه‌ها بدون هرگونه عضو شده و باعث اتلاف بیشتر انرژی خواهند شد. پارامتر Ω_2 تعداد سرخوشه‌های بدون عضو را نمایش می‌دهد و از رابطه (۸) محاسبه می‌شود.

$$\Omega_2 = \begin{cases} \sum_{j=1}^k (CL_{mem-less})_k & \text{if } k > 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

در مدل پیشنهادی فرض شده است که با توجه به محدودیت‌های منابع سنسورها هر گره قادر است تنها به تعداد محدودی سنسور معمولی خدمات سرخوشه‌ای ارائه نماید. اگر چنانچه تعداد اعضا از میزان حداکثری بیشتر شود، آن سرخوشه دارای سربار اعضا شده و ممکن است نتواند به تمامی اعضا خدمات ارائه داده و باعث ایجاد تاخیر و یا حتی از دست رفتن داده‌ها شود.

$$\Omega_3 = \begin{cases} \sum_{j=1}^k (CL_{overloaded})_j & \text{if } k > 0 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

یکی دیگر از پارامترهای بسیار مهم در انتخاب سرخوشه‌ها میزان انرژی باقیمانده گره‌هاست. انتخاب گره‌هایی با میزان انرژی بالاتر سبب می‌شود تا از بروز خطاهای ناشی از فقدان انرژی در سرخوشه‌ها در طول یک بازه جلوگیری شده و سرویس‌های مسیریابی با قابلیت اعتماد بالاتری ارائه شوند. برای دست یافتن به مهم، پارامتر Ω_4 میزان متوسط باتری باقیمانده سنسورهای معمولی به سنسورهای سرخوشه را محاسبه کرده و همیشه باید این میزان از عدد یک کمتر باشد که در رابطه (۱۰) نمایش داده است.

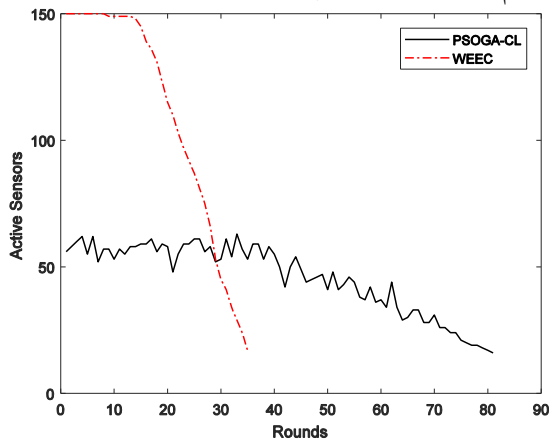
$$\Omega_4 = \frac{\frac{1}{m} \times \sum_{j=1}^m (SN_{Residual-Battery})_j}{\frac{1}{n} \times \sum_{i=1}^n (CH_{Residual-Battery})_i} \quad (10)$$

در روند شکل گیری خوشه‌ها خوشه‌هایی با پراکندگی یکنواخت اعضا مطلوب است. پارامتر Ω_5 میزان واریانس پراکندگی را در سرخوشه‌ها محاسبه کرده و تلاش می‌شود تا این مقدار به صفر میل کند که از رابطه (۱۱) محاسبه می‌شود.

$$\Omega_5 = \frac{1}{k} \times \left(\sum_{j=1}^k \left| N_{CL-members} - \left[\frac{N_{SN}}{N_{CH}} \right] \right| \right)_j \quad (11)$$

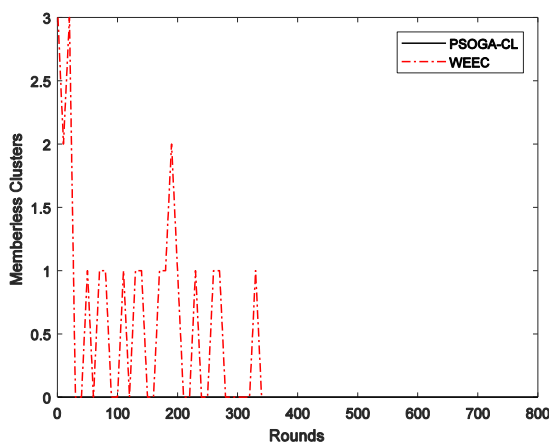
به منظور بهره‌گیری از توان حداکثری سرخوشه‌ها، خوشه‌هایی با حداکثر تعداد اعضا سبب ذخیره بیشتر مصرف انرژی در سنسورها خواهد شد. به منظور شکل متوسط تعداد اعضا سنسورهای فعال به

میزان باتری باقیمانده و پوشش شبکه و حداقل میزان افزونگی داده و برهم افتادگی حسگری است. در ادامه این مقاله با توجه به راندمان بسیار مطلوب، برای طراحی زمان‌بندی خواب و خوشه‌بندی از الگوریتم بهینه‌سازی ذرات-ژنتیک استفاده می‌شود. در گام دوم تحقیق، به توپولوژی مورد نظر یک لایه خوشه‌بندی اضافه شده و توپولوژی حاصل با الگوریتم WEEC مقایسه می‌شود.



شکل ۴: تعداد سنسورهای فعال در طول عمر شبکه.

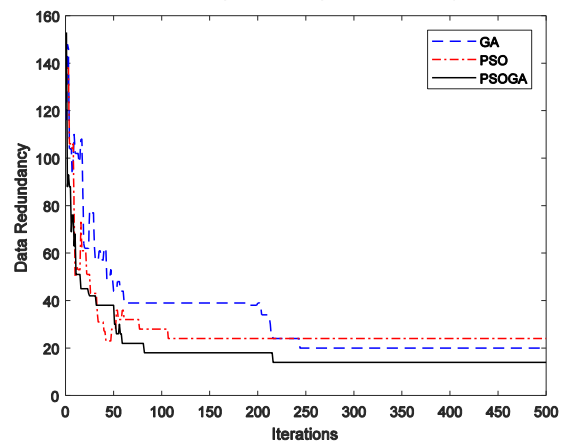
در شکل (۴) تعداد سنسورهای فعال در طول عمر شبکه نمایش داده شده است. الگوریتم خوشه‌بندی پیشنهادی PSOGA-CL با بهره‌گیری از الگوریتم زمان‌بندی به خوبی توانسته است در مقایسه با الگوریتم WEEC با تعداد $\frac{1}{3}$ سنسور زنده در هر دوره فعالیت کرده و به خوبی به انتظارات خواسته شده پاسخ دهد. یکی دیگر از پارامترهای مهم در خوشه‌بندی، تعداد سرخوشه‌ها در هر بازه زمانی می‌باشد. تعداد زیاد سرخوشه‌ها باعث اتلاف انرژی و تعداد کم آنان باعث آسیب رساندن به اتصال و پوشش و افزایش زمان تاخیر انتقال خواهد شد. همانطور که مشخص است تعداد سرخوشه‌ها در روش پیشنهادی نسبت به الگوریتم WEEC دارای روندی متعادل‌تر و کمتری می‌باشد که باعث افزایش کارایی شبکه خواهد شد. در الگوریتم WEEC تعداد زیاد سرخوشه‌ها باعث ایجاد خوشه‌هایی بدون عضو در طول عمر شبکه شده است در حالی که در روش پیشنهادی هیچ خوشه بدون عضو وجود ندارد شکل (۵).



شکل ۵: تعداد خوشه‌های بدون عضو در طول عمر شبکه.

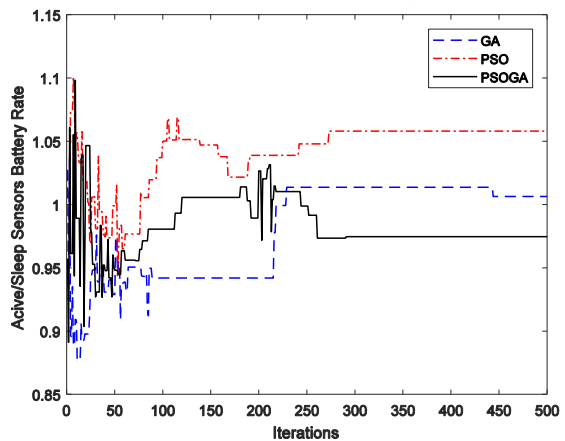
در شبکه‌های تصادفی دوردست، شبکه‌هایی با تعداد بسیار زیاد سنسور با پراکندگی تصادفی ایجاد می‌شوند. لذا عدم استفاده از یک الگوریتم

هرچه تعداد گره‌های کمتری در هر دوره زمانی در وضعیت فعال انجام وظیفه نمایند، باعث می‌شود تا علاوه بر ذخیره بیشتر انرژی، بازدهی شبکه نیز افزایش یابد. در شکل (۱) تعداد حسگرهای فعال انتخاب شده بعد از ۱۵۰ تکرار نمایش داده شده است. الگوریتم GA بیشترین تعداد و الگوریتم PSOGA کمترین تعداد سنسورهای فعال با پوشش شبکه حداکثری را ارائه داده‌اند. با کاهش میزان افزونگی داده، تعداد پیام‌های مبادلاتی شبکه به صورت مطلوبی کاهش یافته و علاوه بر ذخیره بیشتر انرژی سبب کاهش میزان ترافیک شبکه نیز می‌شود. در شکل (۲) میزان افزونگی داده با استفاده از الگوریتم PSOGA بعد از گذشت ۱۰۰ تکرار به میزان مطلوبی کاهش یافته و تقریباً همگرا شده و در نهایت کمترین میزان افزونگی را ارائه می‌دهد.



شکل ۲: میزان افزونگی داده الگوریتم‌های مختلف فراابتکاری.

همچنین میزان باتری در دسترس سنسورها در عملکرد آتی گره‌ها بسیار با اهمیت بوده و خطاهای ناشی از فقدان انرژی لازم در روند دوره زمانی جاری را کمینه می‌کند. در شکل (۳) میزان متوسط انرژی باقیمانده در سنسورهای خواب نسبت به سنسورهای فعال نمایش داده شده و همواره باید این میزان از یک کمتر باشد.



شکل ۳: میزان نسبت باتری باقیمانده سنسورهای خواب به فعال.

لذا الگوریتم PSOGA کمترین میزان نسبت باتری سنسورهای خواب به فعال را نسبت به سایر الگوریتم‌ها ارائه نموده است. میزان درصد عدم پوشش شبکه در تمامی الگوریتم‌ها برابر صفر می‌باشد. لذا می‌توان گفت الگوریتم زمان‌بند خواب با کمک الگوریتم فراابتکاری ترکیبی بهینه‌سازی ذرات-ژنتیک به خوبی توانسته است با حداقل تعداد سنسور مجموعه‌ای کمینه از حسگرهای فعال را ارائه نماید که دارای بیشترین

Sensor Networks”, in IEEE/ACM Transaction on Networking, To Appear.

[6] E. R. Hruschka, J. G. B. Ricardo, A. A. Freitas, "A Survey of Evolutionary Algorithms for Clustering", In: IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 39 issue 2, March 2009, DOI: <http://dx.doi.org/10.1109/TSMCC.2008.2007252>.

[7] Robert M. Curry, J. Cole Smith, "A survey of optimization algorithms for wireless sensor network lifetime maximization", In: Journal of Computers & Industrial Engineering, Vol. 101, November 2016, pp. 145–166, DOI: <https://doi.org/10.1016/j.cie.2016.08.028>.

[8] J. Roselina, P. Lathab, S. Benittac, "Maximizing the wireless sensor networks lifetime through energy efficient connected coverage", In Journal of Ad Hoc Networks, Vol. 62, pp. 1–10, July 2017, DOI: <https://doi.org/10.1016/j.adhoc.2017.04.001>.

[9] A. A. Abbasia, M. Younisb, "A survey on clustering algorithms for wireless sensor networks", Computer Communications, vol. 30, issues 14–15, 15 October 2007, pp. 2826–2841, DOI: <http://dx.doi.org/10.1016/j.comcom.2007.05.024>.

[10] N. Behboudi, A. Abhari, "A Weighted Energy Efficient Clustering (WEEC) for Wireless Sensor Networks", Seventh International Conference on Mobile Ad-hoc and Sensor Networks, 2011, DOI: <http://dx.doi.org/10.1109/MSN.2011.13>.

[11] S. Kousalya Devi, C. Kavitha, "Survey on Sleep Scheduling Methods in Wireless Sensor Networks", In International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Vol. 3 Issue 1, Jan. 2014.

[12] K. P. Ferentinos, T. A. Tsiligiridis, "Adaptive design optimization of wireless sensor networks using genetic algorithms", In: Journal of Computer Networks, Vol. 51, Issue 4, 14 March 2007, pp. 1031–1051, DOI: <https://doi.org/10.1016/j.comnet.2006.06.013>.

[13] Pooja, S. Singh, "Improved O-LEACH protocol: A clustering based approach in wireless microsensor network", In: 10th International Conference on Intelligent Systems and Control (ISCO), 2016, DOI: <http://dx.doi.org/10.1109/ISCO.2016.7727015>.

[14] I. Demirkol, C. Ersoy, F. Alagoz, "MAC protocols for wireless sensor networks: a survey", In Journal of IEEE Communications, Vol. 44, Issue 4, April 2006, DOI: <http://dx.doi.org/10.1109/MCOM.2006.1632658>.

[15] Ajit R. Pagar, Prof. D. C. Mehrete, "A Survey on Energy Efficient Sleep Scheduling in Wireless Sensor Network", In: International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 5, Issue 1, Jan. 2015.

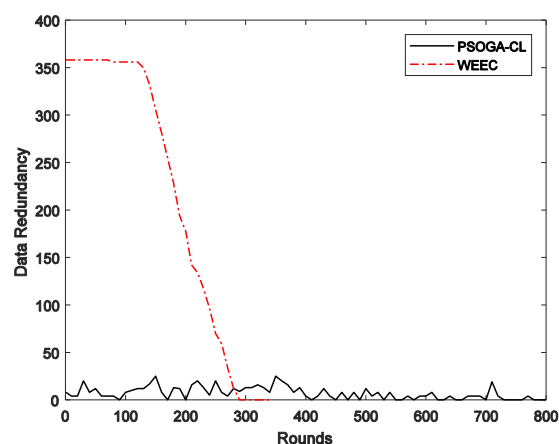
[16] S. Hussain, A. W. Martin, and O. Islam, "Genetic algorithm for hierarchical wireless sensor networks", In: Journal of Networks, vol. 2, no. 5, pp. 87–97, 2007, DOI: <http://dx.doi.org/10.4304/jnw.2.5.87-97>.

[17] R. L. Haupt, S. E. Haupt, Practical genetic algorithms. John Wiley & Sons, 2004, DOI: <http://dx.doi.org/10.1002/0471671746>.

[18] Jenn-Long Liu, Chinya V. Ravishankar, "LEACH-GA: Genetic Algorithm-Based Energy-Efficient Adaptive Clustering Protocol for Wireless Sensor Networks", In International Journal of Machine Learning and Computing, Vol.1, No. 1, April 2011.

[19] S. M. Hosseinirad, M. Alimohammadi, S. K. Basu, and A. A. Pouyan, "LEACH Routing Algorithm Optimization through Imperialist Approach," Int. J. Eng. A Basics, vol. 27, no. 1, p. 39, 2013, DOI: <http://dx.doi.org/10.5829/idosi.ije.2014.27.01a.06>.

زمان‌بند باعث افزونگی بسیار زیاد داده خواهد شد. الگوریتم خوشه‌بندی پیشنهادی به دلیل استفاده از الگوریتم زمان‌بند کارآمد دارای میزان داده افزونه بسیار کمتری نسبت به الگوریتم WEEC می‌باشد شکل (۶). با توجه به نتایج حاصله الگوریتم پیشنهادی توانسته است با انتخاب بهینه مجموعه‌ای از حسگرهای فعال در هر بازه زمانی و خوشه‌بندی یکنواخت و موثر طول عمر شبکه را به خوبی افزایش دهد.



شکل ۶: میزان افزونگی داده بر طول عمر شبکه.

۵- نتیجه گیری

استفاده از توپولوژی سلسله‌مراتبی در شبکه‌های حسگر بی‌سیم می‌تواند بازدهی شبکه را به نحو مطلوب افزایش دهد. ایجاد یک روش خوشه‌بندی با یک زمان‌بند خواب متمرکز از اهداف این مقاله بوده است. نتایج این تحقیق نشان می‌دهد که استفاده از یک زمان‌بند خواب موثر در شبکه‌های حسگر بی‌سیم اجتناب ناپذیر است. افزایش طول عمر شبکه، قابلیت اطمینان و پوشش و کاهش میزان افزونگی داده، مصرفی صرفی از مزایای استفاده از الگوریتم خوشه‌بندی پیشنهادی می‌باشد.

منابع

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," Computer networks, vol. 38, no. 4, pp. 393–422, 2002, DOI: [http://dx.doi.org/10.1016/S1389-1286\(01\)00302-4](http://dx.doi.org/10.1016/S1389-1286(01)00302-4).
- [2] T. Arampatzis, J. Lygeros, and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks," in Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation, 2005, pp. 719–724, DOI: <http://dx.doi.org/10.1109/2005.1467103>.
- [3] Robert M. Curry, J. Cole Smith, "A survey of optimization algorithms for wireless sensor network lifetime maximization", In: Journal of Computers & Industrial Engineering, Vol. 101, November 2016, pp. 145–166, DOI: <https://doi.org/10.1016/j.cie.2016.08.028>.
- [4] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," Ad hoc networks, vol. 7, no. 3, pp. 537–568, 2009, DOI: <http://dx.doi.org/10.1016/j.adhoc.2008.06.003>.
- [5] W. Ye, J. Heidemann, and D. Estrin, "Medium Access Control with Coordinated, Adaptive Sleeping for Wireless



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

اعضای کمیته راهبردی

رئیس دانشگاه صنعتی شاهرود	پروفسور محمد مهدی فاتح
معاون پژوهشی دانشگاه صنعتی شاهرود	دکتر علی دستفان
رئیس قطب علمی سامانه‌های پردازش دیجیتال	پروفسور حسن غفوری فرد
دبیر کمیته دائمی کنفرانس	پروفسور کریم فائز
دبیر کنفرانس سوم	پروفسور حمید حسن پور
دبیر علمی کنفرانس	دکتر علیرضا احمدی فرد
دبیر کمیته مالی	دکتر علی سلیمانی
دبیر کمیته روابط عمومی	دکتر حسین مروی
دبیر کمیته پایگاه داده ها و مقالات	دکتر هدی مشایخی
دبیر کمیته روابط بین الملل	دکتر محسن رضوانی
دبیر کمیته تدارکات و امور اجرایی	دکتر مرتضی زاهدی
دبیر کمیته پشتیبانی و ارتباط با صنعت	دکتر علی اکبر پویان
دبیر کمیته انتشارات	دکتر امیدرضا معروضی
دبیر کمیته تشریفات	دکتر منصور فاتح
دبیر کمیته انفورماتیک	مهندس علیرضا تجری



قطب علمی سامانه‌های پردازش دیجیتال

کنفرانس پردازش سیگنال و سیستم‌های هوشمند

۲۹ الی ۳۰ آذرماه ۱۳۹۶



دانشگاه صنعتی شاهرود

اعضای کمیته علمی

دانشگاه تربیت مدرس	پروفسور احسان الله کبیر
دانشگاه صنعتی امیرکبیر	پروفسور کریم فائز
دانشگاه صنعتی امیرکبیر	پروفسور سیداحمد معتمدی
دانشگاه صنعتی امیرکبیر	پروفسور محمدحسن مرادی
دانشگاه صنعتی امیرکبیر	پروفسور حمیدرضا امین‌داور
دانشگاه صنعتی شاهرود	پروفسور حمید حسن پور
دانشگاه صنعتی امیرکبیر	دکتر ابوالقاسم راعی
دانشگاه صنعتی امیرکبیر	دکتر سیدمحمد احدی
دانشگاه صنعتی امیرکبیر	دکتر سیدعلی سیدصالحی
دانشگاه صنعتی امیرکبیر	دکتر حسن آقایی‌نیا
دانشگاه صنعتی امیرکبیر	دکتر حسن طاهری
دانشگاه صنعتی امیرکبیر	دکتر سعید شریفیان
دانشگاه صنعتی امیرکبیر	دکتر مجید شالچیان
دانشگاه شهید بهشتی	دکتر فرح ترکمنی‌آذر
دانشگاه گیلان	دکتر اسدالله شاه‌بهرامی
دانشگاه یزد	دکتر محمد قاسم‌زاده
دانشگاه صنعتی شاهرود	دکتر علیرضا احمدی‌فر
دانشگاه گنکوردیا	دکتر نعمت‌الله شیری
دانشگاه صنعتی شاهرود	دکتر وحید ابوالقاسمی
دانشگاه صنعتی شاهرود	دکتر علیرضا الفی
دانشگاه صنعتی شاهرود	دکتر حیدر طوسی‌ان
دانشگاه صنعتی شاهرود	دکتر امیدرضا معروضی
دانشگاه صنعتی شاهرود	دکتر حسین مروی

دانشگاه صنعتی شاهرود	دکتر علی سلیمانی
دانشگاه صنعتی شاهرود	دکتر هادی گرایلو
دانشگاه صنعتی شاهرود	دکتر حسین قلی زاده نرم
دانشگاه صنعتی امیرکبیر	دکتر حمیدرضا حبیبیان
پژوهشگاه ارتباطات و فناوری اطلاعات (مرکز تحقیقات مخابرات ایران)	دکتر کامبیز بدیع
دانشگاه سمنان	دکتر فرزین یغمایی
دانشگاه صنعتی خواجه نصیرالدین طوسی	دکتر یوسف درمانی
دانشگاه بیرجند	دکتر حسن فرسی
دانشگاه شهید باهنر کرمان	پروفسور حسین نظام آبادی پور

Dr. Shahaboddin Shamshirband	University of Malaya, Malaysia
Prof. Lauro Cássio Martins de Paula	Federal University of Goiás, Brazil
Dr. J. Arturo Olvera Lopez	Autonomous University of Puebla
Prof. Jinde Cao	Southeast University
Dr. Pasquale De Meo	University of Messina
Dr. Ali Gholipour	Harvard Medical School
Prof. Yudong Zhang	Nanjing Normal University
Dr. Christophe Feltus	Luxembourg Institute of Science and Technology
Dr. Alina Campan	Northern Kentucky University
Prof. Noor Azina Ismail	University of Malaya
Prof. Firuz Zare	University of Queensland
Dr. Mehregan Mahdavi	UNSW Sydney – Australia's Global University
DR. Amin Zehtabian	Free University of Berlin

فهرست الفبایی نویسندگان

نویسنده	صفحه
ا	
ابوطالبی، حمیدرضا	۶۱
احمدی فرد، علیرضا	۱۰۵
اسدی امیری، سکینه	۱۳۱
اصدقی، فائزه	۸۵
آیت، مولود	۱۶۲
ب	
برسلانی، جواد	۱۸۱
پ	
پویان، علی‌اکبر	۱۸۶, ۴۸, ۵
ت	
ترابی، محمد	۱۶۸, ۱۱۸
ح	
حسن‌پور، حمید	۵۴
حسینی، زینب	۳۲
حسینی ابرده، محمد	۱۵۶
حسینی راد، سیدمجتبی	۱۹۳
حسینی ورکی، مهدیه سادات	۹۱
حسینی، محمد مهدی	۱۶۲, ۷۳
خ	
خسروی، حسین	۱۷۴, ۱۴۸, ۱۳۷
خنجری، محسن	۱۵۶
د	
دستفان، علی	۱۸۱
ر	
راعی، ابوالقاسم-اسدالله	۷۹
رشمه کریم، مریم	۳۷
رضوانی، محسن	۲۰
ز	
زارعی، مانی	۹۸
زاهدی، مرتضی	۱۶۲, ۷۳
س	
سادات، الهه سادات	۲۶

۱۳	سلطانی محبوب، امیر
۴۳	سلطانی، آزاده
۴۳	سلطانی، محمود
۸۵	سلیمانی ایوری، علی
	ص
۴۸	صادقی زاده، مهدی
۷۹	صادقی، حمید
۱۴۲	صادقی، محمد تقی
۳۲	صمدیانی، نجمه
	ظ
۱۳	ظهیری، سید حمید
	ع
۱۴۲	عبداللهی، احسان
۱۴۸	عسگریان دهکردی، رسول
	غ
۱۲۴	غلامرضائی، علیرضا
۱۰۵	غلامی، سوسن
	ف
۲۰	فاتح، منصور
۹۱، ۲۶	فائز، کریم
۱۲۴	فرخی، حمید
۵	فردوسی، سعیده
	ق
۱۸۱	قالیبافان، جواد
۲۰	قیصری، پرنیا
	ک
۶۷	کارگران، علی
۱۶۸	کاوندی، پریناز
۹۸	کرماجانی، حمیدرضا
	گ
۶۷	گرایلو، هادی
۵۴	گندمی، محمدرضا
	م
۱۱۸	محمدی، مهدی
۴۸، ۳۷	معروضی، امیدرضا
۱۸۶	معلم، محمود

- ۱۳۷.....مقیمی، مهدی
- ۹۸.....ملکی نودهی، آرش
- ۱۱۲.....مودتی، سمیرا
- ۵.....میرزائی، فائزه
- ن
- ۱۷۴.....نوریان، خشایار
- و
- ۶۱.....ولی‌نژاد طرقلبه، حمید

