

نشریه علمی - پژوهشی
انجمن مهندسان کنترل و ابزار دقیق ایران - دانشگاه صنعتی خواجه نصیرالدین طوسی
جلد ۵، شماره ۱، بهار ۱۳۹۰

فهرست مقالات

- ۱ کنترل پیش بین مقید متمرکز سیستم مدیریت زنجیره تامین کاملاً جفت شده از دیدگاه وابستگی اطلاعاتی رده‌ها
محمد میران بیگی، علی اکبر جلالی
- ۱۴ یک الگوریتم یادگیری ترکیبی جدید جهت طراحی شبکه های موجک فازی برای تقریب توابع، شناسایی و کنترل بهنگام سیستم های غیر خطی
مریم شهریاری کاهکشی، مریم ذکری
- ۲۷ کاربرد مفهوم ظرفیت جذب انرژی در تحلیل پایداری سیستم های چند عاملی
کریم رحمانی، احمد افشار، علی اکبر جمشیدی فر
- ۳۸ تقسیم بهینه توان اکتیو و راکتیو بین چند منبع تولید پراکنده در یک شبکه مستقل با شبکه عصبی مصنوعی
علی اصغر قدیمی، حسن رستگار، فرزاد رضوی
- ۵۰ تقریب تابع ارزش عمل با استفاده از شبکه توابع پایه شعاعی برای یادگیری تقویتی
ولی درهمی، امید محرابی
- ۶۴ حل مساله می نیمم های محلی در مسیریابی به روش میادین پتانسیل مجازی و براساس دیدگاه حرکت مایعات
بیژن معاونی، داوود قنبری گل
- ۷۷ تعیین موقعیت بازیکنان فوتبال و استخراج همزمان مسیر آنها در مدل واقعی زمین با استفاده از روش گراف در تصاویر ویدیویی
مهرتاش منافی فرد، حمید عبادی، حمید ابریشمی مقدم



نشریه علمی - پژوهشی، انجمن مهندسان کنترل و ابزار دقیق ایران - دانشگاه صنعتی خواجه نصیرالدین طوسی،

جلد ۵، شماره ۱، بهار ۱۳۹۰

پست الکترونیکی: control@isice.ir

صاحب امتیاز: انجمن مهندسان کنترل و ابزار دقیق ایران

مدیر مسئول: پروفسور ایرج گودرزنیا

سردبیر: پروفسور علی خاکی صدیق - تلفن: ۸۴۰۶۲۳۱۷ - پست الکترونیکی: sedigh@kntu.ac.ir

آدرس محل کار: خیابان دکتر شریعتی، پل سیدخندان، دانشکده برق دانشگاه صنعتی خواجه نصیرالدین طوسی

سمت: استاد دانشگاه

شورای سردبیری: پروفسور علی خاکی صدیق، دکتر حمید خالوزاده، دکتر علیرضا فاتحی

دبیر اجرایی: دکتر حمید خالوزاده

هیأت تحریریه:

پروفسور علی خاکی صدیق (استاد) - پروفسور ایرج گودرزنیا (استاد) - دکتر حمید خالوزاده (دانشیار) - پروفسور پرویز جبه دار مارالانی (استاد) - پروفسور

علی غفاری (استاد) - دکتر حمیدرضا مومنی (دانشیار) - پروفسور سید کمال الدین نیکروش (استاد) - پروفسور مسعود شفیعی (استاد) - پروفسور بهزاد مشیری

(استاد)

هیأت مشاوران:

دکتر حمیدرضا مومنی، پروفسور بهزاد مشیری، پروفسور مسعود شفیعی، پروفسور علی خاکی صدیق، پروفسور پرویز جبه دار مارالانی، پروفسور علی غفاری،

دکتر حمید خالوزاده، دکتر حمیدرضا تقی راد، دکتر کیوان مسروری، دکتر محمدتقی بطحایی، دکتر محمدتقی بهشتی، دکتر فرزاد جعفر کاظمی، دکتر رویا

امجدی فرد، دکتر سید علی اکبر موسویان، دکتر محمد تشنه لب، پروفسور محمد حایری، دکتر سید علی اکبر صفوی، پروفسور حسین سیفی، دکتر احد

کاظمی، دکتر علیرضا فاتحی، دکتر محمدرضا اکبرزاده توتونچی، دکتر مسعود علی اکبر گلکار، دکتر ناصر پریز، دکتر مهرداد جوادی، دکتر جعفر حیرانی

نوبری، پروفسور فرامرز حسین بابایی، دکتر بیژن معاونی، دکتر مهدی علیاری شوره دلی، دکتر محمد عاروان، دکتر محمد توکلی بینا.

هیأت مدیره انجمن مهندسان کنترل و ابزار دقیق:

مهندس عباس شعری مقدم، پروفسور مسعود شفیعی، دکتر حمیدرضا مومنی، دکتر حمید خالوزاده، دکتر مهرداد جوادی، دکتر داود کریم زادگان، مهندس علی

کیانی.

تهران، خیابان انقلاب، میدان فردوسی، خیابان عباس موسوی (فرصت)، پلاک ۷۱، طبقه دوم، اتاق ۲۴۱

فاکس: ۸۸۳۲۴۹۷۹

تلفن: ۸۸۱۳۰۰۲

صندوق پستی: ۱۵۸۱۵-۳۵۹۵

www.isice.ir

نشریه علمی- پژوهشی (ISC) کنترل وابسته به "انجمن مهندسان کنترل و ابزار دقیق ایران" و "دانشگاه صنعتی خواجه نصیرالدین طوسی" تنها نشریه تخصصی در زمینه مهندسی کنترل و ابزار دقیق است.

این مجله که بصورت فصلنامه منتشر می شود و تا کنون ۱۰ شماره آن به چاپ رسیده، با استقبال فراوان جامعه علمی دانشگاهی و مهندسان شاغل در صنایع گوناگون روبرو شده است و هم اکنون از مجلات منظم و پربار تخصصی در سطح علمی محسوب می گردد. تعداد مقالات دریافتی جهت بررسی و چاپ از دانشگاهها و مراکز علمی- پژوهشی سرتاسر کشور به خوبی نمایانگر این مدعا است. جایگاه این مجله در زمینه های مختلف علمی و فن آوری داخل کشور به خوبی توسط محققان و اندیشمندان این حوزه ها شناخته شده است و بدیهی است که چاپ به هنگام و با کیفیت بالای این مجله تخصصی می تواند گام بلندی در راستای توسعه پایدار علمی کشور باشد.

در یازدهمین جشنواره تجلیل از پژوهشگران و فناوران برتر که در تاریخ هفتم دی ماه ۱۳۸۹ برگزار گردید، مجله کنترل به عنوان نشریه علمی برتر در زمینه فنی و مهندسی معرفی شد. خداوند علیم و حکیم را برای دست یابی به این رتبه علمی با وجود تعداد زیادی مجله ی علمی و پژوهشی در حوزه فنی و مهندسی در داخل کشور، شکر گزاریم. بدیهی است که چاپ به هنگام و با کیفیت بالای این مجله تخصصی می تواند گام بلندی در راستای توسعه پایدار علمی کشور باشد.

از آنجا که این نشریه برای داوری و چاپ مقالات هزینه ای را از ارسال کنندگان مقالات دریافت نمی نماید، لذا برای حفظ کیفیت و روند رو به رشد آن نیاز به حمایت مالی دارد که در این خصوص موسسات و شرکتهای زیر از جمله حامیان این شماره مجله بوده اند که در اینجا از ایشان تشکر و قدردانی بعمل می آید.

یارگران این شماره

- شرکت پتروشیمی امیر کبیر
- شرکت پلیمر آریاساسول
- شرکت پتروشیمی لاله
- شرکت جابون
- شرکت پتروشیمی پردیس
- شرکت سپهر انرژی
- شرکت فیرم کو
- شرکت مهندسی سنس

فهرست مقالات

- ۱ کنترل پیش‌بین مقید متمرکز سیستم مدیریت زنجیره تامین کاملاً جفت شده از دیدگاه وابستگی اطلاعاتی رده‌ها
محمد میران‌بیگی، علی‌اکبر جلالی
- ۱۴ یک الگوریتم یادگیری ترکیبی جدید جهت طراحی شبکه‌های موجک فازی برای تقریب توابع، شناسایی و کنترل بهنگام سیستم‌های غیرخطی
مریم شهریاری کاهکشی، مریم ذکری
- ۲۷ کاربرد مفهوم ظرفیت جذب انرژی در تحلیل پایداری سیستم‌های چند عاملی
کریم رحمانی، احمد افشار، علی‌اکبر جمشیدی فر
- ۳۸ تقسیم بهینه توان اکتیو و راکتیو بین چند منبع تولید پراکنده در یک شبکه مستقل با شبکه عصبی مصنوعی
علی اصغر قدیمی، حسن رستگار، فرزاد رضوی
- ۵۰ تقریب تابع ارزش عمل با استفاده از شبکه توابع پایه شعاعی برای یادگیری تقویتی
ولی درهمی، امید محرابی
- ۶۴ حل مساله می نیمم‌های محلی در مسیریابی به روش میداین پتانسیل مجازی و براساس دیدگاه حرکت مایعات
بیژن معاونی، داوود قنبری گل
- ۷۷ تعیین موقعیت بازیکنان فوتبال و استخراج همزمان مسیر آنها در مدل واقعی زمین با استفاده از روش گراف در تصاویر ویدیویی
مهرتاش منافی فرد، حمید عبادی، حمید ابریشمی مقدم

مجله کنترل، مجله ای علمی - پژوهشی است که در برگیرنده تازه ترین نتایج تحقیقات نظری و کاربردی در علوم مختلف مرتبط با مهندسی کنترل و ابزار دقیق می باشد. مقالات ارسالی به مجله کنترل می بایست به زبان فارسی و دارای چکیده انگلیسی باشند. از میان مباحث مورد نظر این مجله میتوان به موارد زیر اشاره نمود:

- (۱) مدل سازی، شناسایی، شبیه سازی و بهینه سازی سیستمها.
- (۲) تحلیل و طراحی سیستم های کنترل پیشرفته همچون سیستم های کنترل خطی و غیرخطی، سیستمهای کنترل تطبیقی، کنترل مقاوم و کنترل بهینه، سیستمهای کنترل هوشمند، سیستمهای کنترل تصادفی، سیستمهای کنترل گسسته پیشامد و ترکیبی، سیستمهای ابعاد وسیع.
- (۳) مکاترونیک و رباتیک.
- (۴) ابزار دقیق و سیستمهای ترکیب داده و اطلاعات سنسوری.
- (۵) اتوماسیون صنعتی همچون سیستمهای کنترل گسترده، رابط انسان - ماشین، سیستمهای ایمنی و تشخیص خطا، سیستمهای زمان حقیقی و سیستمهای کنترل سوپروایزری.

کاربردهای مورد علاقه این مجله، وسیع بوده و می تواند در برگیرنده موارد زیر باشد:

- (۱) سیستمهای هدایت و ناوبری.
- (۲) فرآیندهای صنعتی شامل فرآیندهای شیمیایی و بیوتکنولوژی.
- (۳) فرآیندهای استخراج و فراوری موارد معدنی.
- (۴) سیستمهای حمل و نقل و خودروهای خودکار.
- (۵) تولید و توزیع نیروی برق.
- (۶) مهندسی محیط زیست و هواشناسی.
- (۷) مهندسی تکنولوژی تولید.
- (۸) سیستمهای اقتصادی و مالی.
- (۹) سیستمهای اطلاعاتی، مخابراتی و شبکه های صنعتی.
- (۱۰) مهندسی پزشکی.
- (۱۱) سیستمهای آموزش هوشمند.

از کلیه پژوهشگران و کارشناسان فعال در زمینه های مرتبط با مهندسی کنترل و ابزار دقیق دعوت بعمل می آید تا مقالات و نتایج آخرین دستاوردهای علمی و پژوهشی خود را به این مجله ارسال نمایند. خواهشمند است مقالات خود را به صورت الکترونیکی به آدرس control@isice.ir ارسال فرمایید. برای کسب اطلاعات بیشتر و دریافت نحوه تهیه و ارسال مقالات می توانید به سایت مجله با آدرس www.isice.ir مراجعه نمایید.

شیوه تدوین

متن مقالات شامل چکیده، بدنه مقاله، مراجع و زیرنویسها باید با فونت ۱۲ B Zar و با فاصله double میان خطوط، در صفحات A4 یک ستونی و تحت نرم‌افزار Word تهیه گردد.

آدرس نویسندگان

آدرس پستی کامل همه نویسندگان همراه با شماره تلفن و دورنگار (فکس) و نشانی پیام‌نگار (email) نویسنده عهده‌دار مکاتبات در برکه مستقلى چاپ و به همراه مقاله ارسال گردد.

چکیده

هر مقاله باید شامل، عنوان (فارسی و انگلیسی)، چکیده (فارسی و انگلیسی) مقاله در حداکثر ۲۰۰ واژه، کلیدواژه (فارسی و انگلیسی) در حداکثر ۵ واژه باشد.

تصاویر و عکسها

در هنگام ارسال مقاله جهت داوری نیازی به ارسال اصل تصاویر و عکسها نمی‌باشد، ولی رونوشت ارسالی باید واضح باشد. پس از تایید مقاله، ارسال اصل تصاویر و عکسها جهت چاپ مقاله ضروری می‌باشد.

مراجع

به کلیه مراجع باید در متن ارجاع داده شده باشد. مراجع باید با شماره مشخص گردند و جزئیات آنها به شرح زیر در پایان مقاله به ترتیب حروف الفبای نویسندگان ظاهر گردد:

مقالات: [شماره مرجع] نام خانوادگی و علامت اختصاری اول نام، سال انتشار یا تاریخ برگزاری، "عنوان مقاله"، نام کامل نشریه یا کنفرانس، شماره مجله یا شماره جلد، شماره صفحات.

کتابها: [شماره مرجع] نام خانوادگی و نام کامل همه نویسندگان، عنوان کتاب، نام مترجم (در صورت وجود)، نام کامل ناشر، سال انتشار.

واحدها: کلیه مقالات باید از واحد استاندارد SI (متریک) در تمام بخشهای مقاله استفاده نمایند. در کنار واحد SI می‌توان از واحد انگلیسی در داخل پرانتز نیز استفاده نمود.

طول مقالات

حداکثر حجم مقالات در هنگام چاپ ۱۵ صفحه می‌باشد که معادل حدود ۷۵۰۰ واژه می‌باشد. برای چاپ صفحات بیشتر و یا رنگی لازم است هزینه‌ای معادل ۲۵۰,۰۰۰ ریال (۲۵ دلار آمریکا) برای هر صفحه به حساب مجله واریز گردد.

فرایند ارسال مقاله

مقالات قابل چاپ در مجله شامل مقالات کامل پژوهشی، مقالات کوتاه و یادداشتهای پژوهشی می‌باشد. مقالات ارسالی نباید در هیچ مجله داخلی و یا خارجی چاپ شده باشد و یا در حال داوری باشد.

- لازم است نسخه الکترونیکی مقاله به شکل pdf و word جهت داوری به نشانی مجله control@isice.ir ارسال شود.
- مقالات جهت داوری به داوران متخصص ارسال میگردد. در پایان تایید یا رد هر مقاله توسط هیئت تحریریه مجله انجام خواهد پذیرفت. سردبیر مجله نتیجه داوری را برای نویسنده عهده‌دار مکاتبات ارسال خواهد نمود.
- در صورتی که نیاز به تصحیح مقاله باشد، تصحیحات باید منحصرأ محدود به موارد ذکر شده باشد. در سایر موارد نویسنده لازم است سردبیر را در جریان هر گونه تغییر و یا تصحیح دیگری قرار دهد. در هر صورت مسئولیت صحت و سقم مطالب بر عهده نویسنده خواهد بود.
- در صورتی که مقاله جهت چاپ پذیرفته شود، یک نسخه از مجله همراه با ۵ نسخه از مقاله به هر یک از نویسندگان اهدا خواهد گردید.

حق کپی: در صورت تایید مقاله، نویسندگان لازم است فرم انتقال حق انتشار آن به "انجمن مهندسان کنترل و ابزار دقیق ایران" را تکمیل و به همراه اصل مقاله ارسال نمایند. نویسندگان لازم است موافقت کتبی دارندگان حق کپی بخشهایی از مقاله که از مراجع و منابع دیگر نسخه‌برداری شده است را دریافت و به دفتر مجله ارسال نمایند.

بدینوسیله از کلیه اساتید، پژوهشگران و کارشناسان مهندسی کنترل و ابزار دقیق جهت ارائه مقالات خود در این نشریه دعوت به عمل می‌آورد. خواهشمند است مقالات خود را به صورت الکترونیکی به آدرس: control@isice.ir ارسال فرمایید. برای کسب اطلاعات بیشتر میتوانید به سایت: <http://www.isice.ir> مراجعه نمایید.

کنترل پیش‌بین مقید متمرکز سیستم مدیریت زنجیره تامین کاملاً جفت شده از دیدگاه وابستگی اطلاعاتی رده‌ها

محمد میران بیگی^۱، علی اکبر جلالی^۲

^۱ دانشجوی دکتری مهندسی کنترل، دانشکده مهندسی برق و کامپیوتر، دانشگاه تهران m.miran@ut.ac.ir
^۲ دانشیار و عضو هیئت علمی، گروه کنترل، دانشکده برق دانشگاه علم و صنعت ایران drjalali@iust.ac.ir

(تاریخ دریافت مقاله ۱۳۸۸/۱۲/۱۶، تاریخ پذیرش مقاله ۱۳۹۰/۱/۱۵)

چکیده: سیستم مدیریت زنجیره تامین، یک شبکه از واحدها یا رده‌های تامین کالا و تولید و انبار و توزیع و خرده‌فروش همراه با تاخیرهای زمانی بزرگ می‌باشد. هدف سیستم کنترلی، کارکردن زنجیره تامین در نقطه کاری بهینه یعنی برآوردن تقاضا، علیرغم تنوع تقاضای مشتری است. در این پژوهش، از یک مدل سیستم مدیریت زنجیره تامین جفت شده‌ی اطلاعاتی مبتنی بر "نظریه بازی آشامیدنی"^۱ زمان پیوسته استفاده شد و به یک مدل سیستم مدیریت زنجیره تامین زمان گسسته‌ی چندطبقه- چند رده، شامل واحدهای تامین کننده و تولید کننده و انبار و توزیع کننده و خرده‌فروشی بسط داده شد. در این مدل دو نوع جریان دینامیکی در طول زنجیره تامین در نظر گرفته می‌شوند: جریان کالا از سمت تامین کننده به سوی مصرف کننده نهایی، و جریان برگشتی اطلاعات برای آگاه ساختن رده‌ها از تقاضای مشتری. با توجه به کارایی کنترل کننده پیش‌بین مبتنی بر مدل در کار با این گونه سیستم‌ها که هم تاخیرهای زمانی بزرگ و هم یک ساختار دینامیکی متغیر (منتج از اغتشاش تقاضای مشتری) دارند، یک کنترل کننده پیش‌بین مقید متمرکز بکار گرفته شده است. همچنین برای بالا بردن مقاومت سیستم در مقابل تغییرات تقاضا، یک بخش توقیف حرکت به تابع هزینه اضافه شده است.

کلمات کلیدی: سیستم مدیریت زنجیره تامین، تقاضا، سیستم مدیریت زنجیره تامین جفت شده‌ی اطلاعاتی، کنترل کننده پیش‌بین مبتنی بر مدل، بخش توقیف حرکت.

Centralized Constrained Predictive Control on Information Coupled Supply Chain Management System

Mohammad Miranbeigi, Ali Akbar Jalali

Abstract: Supply chain management system (SCM) is a network of suppliers and manufacturers and warehouses and distributors and retailers, with large delay times. The control system aims at operating the supply chain at the optimal point despite the influence of demand changes. In this paper, an information coupled supply chain management system model based on "Beer Game Theory"¹ was used and was developed to an supply chain management consist of supply, manufacture, warehouse, distribution, retailer units. Then a centralized constrained model predictive controller applied on that. Also a move suppression term added to cost function, which increased system robustness toward changes on demands.

Keywords: Supply Chain Management System, Demand, Information Coupled Supply Chain Management System, Model Predictive Controller, Move Suppression Term.

¹ Beer Game Theory

۱- مقدمه

توصیف می‌شوند. در برخی مدل‌ها مانند مدل مورد نظر مقاله‌ی حاضر، این جریان‌ها بین تمامی گره‌های متوالی زنجیره تامین وجود دارد و این باعث ایجاد چرخه‌هایی از وابستگی اطلاعات میان گره‌ها می‌شود. در برخی مدل‌ها نیز تنها جریان مواد بین گره‌های زنجیره دیده می‌شود و هیچ‌گونه چرخه‌ای وجود ندارد.

کنترل پیش‌بین (MPC) از جهت تشکیل مدلی صریح از سیستم به منظور محاسبات کنترلی، بررسی رفتار سیستم در یک افق زمانی آینده و در نظر گرفتن محدودیت‌های ورودی‌ها، حالت‌ها و خروجی‌های سیستم در محاسبات کنترلی، بکارگیری مدل‌های اغتشاش برای در نظر گرفتن نامعینی تقاضا و پاسخ مناسب به تغییرات تقاضا، یک روش مناسب برای کنترل این سیستم‌ها به نظر می‌رسد.

سیستم‌های مدیریت زنجیره تامین غالباً توسط کنترل‌کننده‌های کلاسیک PID کنترل شده‌اند که برخی از آنها در [۱] تا [۴] دیده می‌شوند. استفاده گسترده از کنترل‌کننده‌های پیش‌بین مبتنی بر مدل، غالباً به خاطر اعمال کارآمد قیدها و شدت‌های غیرخطی در سیستم‌های دینامیکی چند متغیره است [۵] و [۶]. کاپسوتیس^۱ برای اولین بار MPC را برای کنترل یک مسئله مدیریت موجودی تک محصولی بکاربرد [۷]. در این مقاله، یک بخش جریمه در مقابل انحرافات از یک مسیر مرجع موجودی در طول زمان‌های تحویل تولید در نظر گرفته شد. زافستاس^۲ نیز در حل مسئله کنترل بهینه‌ی تولید/موجودی، با استفاده از MPC، انحراف از مقادیر موجودی مطلوب را جریمه کرد [۸]. در این کار، دینامیک‌های فروش با یک معادله تراز موجودی جفت می‌شوند و مسئله کنترل بهینه به صورت یک MPC که در آن متغیرهای کنترل شامل تلاش تبلیغاتی و سطوح تولید هستند، در نظر گرفته می‌شوند. در پژوهش بعدی، MPC برای مدیریت یک شبکه توزیع - تولید چند رده‌ای - چند محصولی با زمان‌های تحویل و بدون سفارش‌های برگشتی بکار گرفته شد [۹]. وانگ و همکاران او^۳ تقاضا را به عنوان اغتشاش بار و به عنوان یک سیگنال تصادفی بدست آمده از نویز سفید یکپارچه در نظر گرفتند و یک MPC مبتنی بر تخمین برای افزایش عملکرد مقاومت سیستم در برابر تغییرپذیری تقاضا و پیش‌بینی‌های غلط، پیشنهاد کردند [۱۰]. در یک مقاله ارائه شده در سال‌های اخیر نیز یک MPC برای بهینه‌سازی سیستم‌های تولید و توزیع شامل یک مدل

امروزه در بازار رقابتی موجود، بنگاه‌های اقتصادی و تولیدی علاوه بر پرداختن به سازمان و منابع داخلی، خود را به مدیریت و نظارت بر منابع و ارکان مرتبط خارج از سازمان نیازمند یافته‌اند. علت این امر در واقع دستیابی به مزیت یا مزایای رقابتی با هدف کسب سهم بیشتری از بازار است. مدیریت زنجیره تامین پدیده‌ای است که این کار را به طریقی انجام می‌دهد که مشتریان بتوانند خدمت قابل اطمینان و سریع را با محصولات با کیفیت در حداقل هزینه دریافت کنند. در سراسر دنیا از آمریکا گرفته تا آسیا، از اروپا تا آفریقا، مقتضیات یک اقتصاد جهانی که هر روز رقابتی‌تر می‌شود، همه را واداشته تا درباره زنجیره تامین و مدیریت هرچه بهتر آن، صحبت کنند. مسئله کلیدی در یک زنجیره تامین، مدیریت و کنترل هماهنگ تمامی این فعالیت‌ها است. مفهوم کلی مدیریت زنجیره تامین برآوردن تقاضای مشتری برای دریافت خدمات کارآمدتر و مناسب‌تر و نیاز به کاهش هزینه‌های نگهداری کالاها موجود در انبار، از مواد اولیه تا تحویل محصول تمام شده به مشتری است. در کل، زنجیره‌های تامین، سیستم‌های دینامیکی پیچیده‌ای هستند که بوسیله تقاضاهای مشتری تحریک می‌شوند.

در اغلب زنجیره‌های تامین، هدف ما تنظیم متغیرهای تصمیم برای پیشینه‌کردن رضایت مشتری با کمترین هزینه عملیاتی است. زیرسیستم‌های نمونه شامل واحدهای فروشنده‌های مواد (تامین‌کننده) در رده تامین، واحدهای تولیدکننده در رده تولید، واحدهای ذخیره‌کننده کالا در رده انبار، واحدهای توزیع‌کننده در رده توزیع و واحدهای فروشنده در رده خرده‌فروش برای محصولات مختلف، گره‌های سیستم بوده و با یکدیگر در ارتباط هستند. در این سیستم هر گره بوسیله یک مجموعه از گره‌های بالادستی تغذیه شده و یک مجموعه از گره‌های پایین دستی را تغذیه می‌کنند. هر کدام از این گره‌ها یک سطح موجودی کالای تنظیم‌شده دارند که در کنترل این سیستم باید برآورده شوند و در برابر هرگونه دستور مشتری، این سطح موجودی کالا باید حفظ شود.

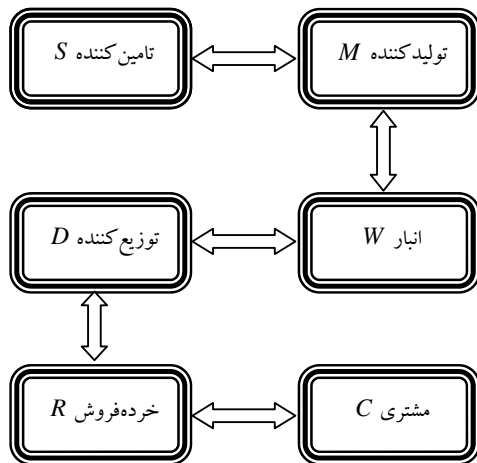
در مدل‌سازی دینامیکی سیستم‌های مدیریت زنجیره تامین عموماً دو نوع جریان وجود دارد. بین رده‌های متوالی دو نوع جریان فرآیند یعنی جریان اطلاعات همانند یک سفارش نیاز کالا و جریان مواد یعنی انتقال حقیقی کالاها وجود دارند. در حقیقت زنجیره‌های تامین معمولاً بوسیله جریان مستقیم مواد اولیه و کالاها و جریان برگشتی اطلاعات

¹ Kapsiotis² Tzafestas³ Wang

همچنین پیشنهادهای برای زمینه‌های باز پژوهشی آینده در این مقوله، ارائه می‌شوند.

۲- مدل دینامیکی سیستم مدیریت زنجیره تامین و توسعه آن

یک سیستم مدیریت زنجیره تامین، شامل تامین‌کننده‌ها، کارخانجات، انبارها و توزیع‌کننده‌ها و خرده‌فروش‌ها می‌باشد که عملیات تامین مواد خام، انتقال مواد خام داخل محصولات میانی و نهایی، و توزیع محصولات نهایی به مشتریان را با توجه به تقاضای مشتری و پیش‌بینی آن، کنترل سطوح موجودی و جریان‌های اطلاعاتی انجام می‌دهد. هنگامی که بتوان یک رویه معین را به عنوان قسمتی از فرایند تولید-توزیع تمیز داد، این را یک طبقه یا یک رده در زنجیره تامین می‌نامند (که این تشخیص‌ها معمولاً بوسیله فاکتورهای جغرافیایی انجام می‌شوند). زنجیره‌های تامین معمولاً بوسیله جریان مستقیم مواد اولیه و کالاها و جریان برگشتی اطلاعات توصیف می‌شوند. همان‌طور که در شکل ۱ دیده می‌شود، در اکثر طرح‌های اولیه، جریان کالا از یک رده به رده بعدی منتقل می‌شود تا کالا به مصرف‌کننده نهایی برسد.



شکل ۱: سیستم مدیریت زنجیره تامین

یک زنجیره تامین نوعاً به وسیله یک جریان رو به جلو از مواد و جریان رو به عقب از اطلاعات تعریف می‌شود. در این سیستم هر گره بوسیله یک مجموعه از گره‌های بالادستی تغذیه شده و یک مجموعه از گره‌های پایین دستی را تغذیه می‌کنند. هر کدام از این گره‌ها یک سطح موجودی کالای تنظیم‌شده دارند که در کنترل این سیستم باید

زمان‌بندی شده ساده شده برای تابع تولید پیشنهاد شده است. روش کنترلی پیشنهادی فقط نوع قطعی تقاضا را که نیاز برای یک رویه کنترل موجودی کالا را کاهش می‌دهد، ملاحظه کرده است [۱۱] و [۱۲].

در این مقاله، یک مدل وابسته‌ی اطلاعاتی گسسته زمان مبتنی بر نظریه بازی آشامیدنی [۱۳] استفاده و به یک سیستم مدیریت زنجیره تامین ابعاد گسترده بسط داده می‌شود. براساس این نظریه، بین تمامی گره‌های رده‌های متوالی دو نوع جریان فرآیند یعنی جریان اطلاعات همانند یک سفارش نیاز کالا و جریان مواد یعنی انتقال حقیقی کالاها وجود دارند. طبق آنچه گفته شد، از آنجا که روش MPC یک روش بهینه‌سازی مقید برخط^۱ می‌باشد و یک مسئله با ورودی‌ها و خروجی‌های ابعاد بالا، به راحتی توسط آن قابل حل است، مدل ابعاد گسترده تاخیردار سیستم مدیریت زنجیره تامین حاضر به صورت مقید و با ملاحظه‌ی برخط تغییرات تقاضای مشتری در طول زمان با ساختار بهینه‌سازی پیش‌بین، کنترل می‌شود. همچنین اثر شلاق چرمی در طول شبیه‌سازی‌ها مورد بررسی قرار می‌گیرد. اثر شلاق چرمی از جمله معمول‌ترین پویایی‌های زنجیره‌های تامین است که به خاطر آن، تغییرات کوچک تقاضای محصول از سوی مشتری در جلوی زنجیره تامین، هر چه به عقب زنجیره تامین حرکت می‌کنیم به نوسانات بیشتر و بیشتر در تقاضا برای شرکت‌ها تبدیل می‌شود [۱۴]. همچنین از یک بخش توقیف حرکت برای مقابله با اثر تغییرات ناگهانی تقاضا روی بیشینه ورودی کنترلی استفاده می‌شود. یک بخش توقیف حرکت که سرعت تغییر در کمیت‌های منتقل شده از طریق زنجیره (تغییر سرعت تغییر متغیرهای کنترل) را جریمه می‌کند، مقاومت سیستم کنترلی را افزایش می‌دهد.

از اینرو در بخش دوم این مقاله یک سیستم مدیریت زنجیره تامین دارای چرخه‌های اطلاعاتی و یک مدل دینامیکی از آن توسعه داده می‌شود. در بخش سوم، صورت مسئله کنترلی تشریح شده، دلایل استفاده از روش کنترل پیش‌بین بار دیگر تبیین شده و تابعی معیار بهینه‌سازی برای مدل سیستم مدیریت زنجیره تامین موردنظر طراحی می‌شود. در بخش چهارم نیز شبیه‌سازی‌های سیستم دینامیکی مدیریت زنجیره تامین جفت‌شده‌ی اطلاعاتی مقید تحت تنوع تقاضاهای مشتری نشان داده می‌شوند و روی اثر بخش توقیف حرکت بحث می‌شود. در بخش پنجم، نتایج بکارگیری روش کنترلی مذکور و نوآوری‌های آن و

¹ Online

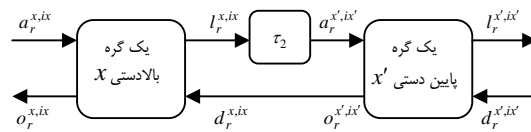
$$0 \leq (s^{x,ix}(t), o_u^{x,ix}(t), b^{x,ix}(t)) \leq s_{\max}^x \quad (2)$$

$$0 \leq o_r^{x,ix}(t) \leq o_{r\max}^{x,ix}$$

$$l_r^{x,ix}(t) = d_r^{x,ix}(t - \tau_1) + \frac{b^{x,ix}(t)}{t_b} \quad (3)$$

که در آن $s^{x,ix}(t)$ موجودی گره i رده x ، $o_u^{x,ix}(t)$ سفارش‌های انجام نشده گره i رده x و $b^{x,ix}(t)$ سفارش‌های انباشته شده گره i رده x از دوره‌های زمانی قبل، متغیرهای حالت این سیستم هستند. $s_{\max}^{x,ix}$ نیز یک مجموعه سه عضوی ثابت حداکثر برای متغیرهای حالت می‌باشد. در معادله نرخ انتقال (۳) سفارش‌ها بر مبنای ارضای تقاضای مشتری به علاوه تسویه سفارش‌های انباشته شده انجام می‌شوند. در این معادله یک تاخیر خط لوله‌ای τ_1 به عنوان زمان لازم برای پردازش سفارش و یک تاخیر مرتبه اول t_b به عنوان زمان جریان تسویه سفارش‌های برگشتی می‌باشند. بعد از جانشانی معادله (۳) در (۱) دیده می‌شود که متغیر حالت $b^{x,ix}(t)$ کنترل‌ناپذیر است، اما اگر نرخ تقاضا به یک مقدار ماندگار همگرا شود $b^{x,ix}(t)$ به صفر همگرا خواهد شد. پس $b^{x,ix}(t)$ پایدارپذیر است و می‌توان این سیستم را کنترل کرد. هدف کنترل برای هر رده میل کردن $s^{x,ix}(t)$ به موجودی مطلوب ثابت $s_d^{x,ix}$ و همگرا شدن $o_u^{x,ix}(t)$ به سفارش اجرا نشده مطلوب $o_{ud}^{x,ix}(t)$ می‌باشد که $o_{ud}^{x,ix}(t) = \tau_2 \cdot d_{rss}^{x,ix}$ (میانگین تقاضای حالت ماندگار مشتری‌هاست) است.

دو گره مجاور از دوره نیز به صورت شکل ۳ مدل می‌شوند و روابط موجود میان پارامترهای دینامیکی آن‌ها نیز مطابق دسته روابط ۴ می‌باشد. گره بالادستی x و گره پایین دستی x' در نظر گرفته شده‌اند.



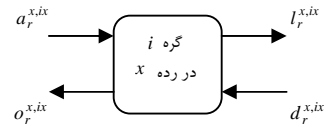
شکل ۳: مدل دو گره مجاور از سیستم مدیریت زنجیره تامین

$$a_r^{x',ix'}(t) = l_r^{x,ix}(t - \tau_2), d_r^{x',ix'}(t) = o_r^{x,ix}(t) \quad (4)$$

τ_2 در این روابط نماینده تاخیر در انتقال کالا از گره بالادستی به گره پایین دستی است. به دلیل ابعاد بزرگ مسئله مورد نظر این مقاله، نمودار بلوکی شکل ۳ به صورت ساده شده‌ی شکل ۴ نمایش داده

برآورده شوند و در برابر هرگونه دستور مشتری، این سطح موجودی کالا باید حفظ شود [۱۴]. این سطح موجودی کالا تحت عنوان سطح موجودی ایمنی شناخته می‌شود به این معنی که برای برآوردن هر میزان تقاضای ناگهانی مشتری باید این سطح از موجودی در گره‌های زنجیره تامین به خصوص انبارها برقرار باشد. البته به سبب هزینه‌های نگهداری موجودی، در پژوهش‌های جاری به دنبال روش‌هایی هستند که نیازی به انباشت زیاد کالا در گره‌های زنجیره ندارند و در نهایت درصدد رساندن سطوح موجودی کالا به مقدار صفر و برآوردن آتی تقاضای مشتری هستند.

در این مقاله از یک مدل زمان پیوسته سیستم مدیریت زنجیره تامین مبتنی بر نظریه بازی آشامیدنی در [۱۳] استفاده شده است. بر اساس این مدل دونوع جریان در دینامیک مدل باید لحاظ شود: جریان کالا و جریان اطلاعات. طبق این مدل‌سازی مدل یک گره از سیستم مدیریت زنجیره تامین در ساختار وابستگی اطلاعاتی گره‌ها (رده‌ها) به صورت شکل ۲ نشان داده می‌شود.



شکل ۲: مدل یک گره از سیستم مدیریت زنجیره تامین در ساختار وابستگی اطلاعاتی گره‌ها (رده‌ها)

گره‌های تغذیه شونده، گره‌های پایین دستی برای گره‌های تغذیه کننده هستند و گره‌های تغذیه کننده نیز به همین ترتیب، نسبت به گره‌های تغذیه شونده، بالادستی به شمار می‌آیند. در این شکل، اگر $x \in \{S, M, W, D, R\}$ باشد $a_r^{x,ix}$ نماینده نرخ اکتساب گره i رده x ، تعداد کالاهایی که در روز از رده بالادستی کسب می‌شوند و $o_r^{x,ix}$ نرخ سفارش گره i رده x ، $l_r^{x,ix}$ نرخ انتقال گره i رده x ، تعداد کالاهایی که در روز به رده پایین دستی انتقال داده می‌شوند و $d_r^{x,ix}$ نرخ انتقال گره i رده x ، تعداد کالاهایی که در روز به وسیله رده پایین دستی سفارش داده می‌شوند، می‌باشند. در نهایت مدل دینامیکی زمان پیوسته و قیده‌ای حالت این سیستم مدیریت زنجیره تامین برای مقادیر $t \geq 0$ به صورت زیر در می‌آید:

$$\begin{cases} \dot{s}^{x,ix}(t) = a_r^{x,ix}(t) - l_r^{x,ix}(t) \\ \dot{o}_u^{x,ix}(t) = o_r^{x,i}(t) - a_r^{x,ix}(t) \\ \dot{b}^{x,ix}(t) = d_r^{x,ix}(t) - l_r^{x,ix}(t) \end{cases} \quad (1)$$

¹ Pipeline delay

حال اگر روابط ۴ و ۵ در روابط ۲ و ۳ جانشانی شوند، روابط زیر برای مدل دینامیکی ساده شده‌ی سیستم مدیریت زنجیره تامین حاصل می‌شوند.

$$\dot{s}^{S,iS}(t) = o_r^{S,iS}(t - \tau_2) - \sum_{iM} o_r^{M,iM}(t - \tau_1) - \frac{b^{S,iS}(t)}{t_b} \quad (۶)$$

$$\dot{o}_u^{S,iS}(t) = o_r^{S,iS}(t) - o_r^{S,iS}(t - \tau_2)$$

$$\dot{b}^{S,iS}(t) = \sum_{iM} o_r^{M,iM}(t - \tau_2) - \sum_{iM} o_r^{M,iM}(t - \tau_1) - \frac{b^{S,iS}(t)}{t_b}$$

$$\dot{s}^{M,iM}(t) = o_r^{M,iM}(t - \tau_1 - \tau_2) - \sum_{iW} o_r^{W,iW}(t - \tau_1) + \sum_{iS} \frac{b^{S,iS}(t - \tau_2)}{t_b} - \frac{b^{M,iM}(t)}{t_b} \quad (۷)$$

$$\dot{o}_u^{M,iM}(t) = o_r^{M,iM}(t) - o_r^{M,iM}(t - \tau_1 - \tau_2) - \sum_{iS} \frac{b^{S,iS}(t - \tau_2)}{t_b}$$

$$\dot{b}^{M,iM}(t) = \sum_{iW} o_r^{W,iW}(t) - \sum_{iW} o_r^{W,iW}(t - \tau_1) - \frac{b^{M,iM}(t)}{t_b}$$

$$\dot{s}^{W,iW}(t) = o_r^{W,iW}(t - \tau_1 - \tau_2) - \sum_{iD} o_r^{D,iD}(t - \tau_1) + \sum_{iM} \frac{b^{M,iM}(t - \tau_2)}{t_b} - \frac{b^{W,iW}(t)}{t_b} \quad (۸)$$

$$\dot{o}_u^{W,iW}(t) = o_r^{W,iW}(t) - o_r^{W,iW}(t - \tau_1 - \tau_2) - \sum_{iM} \frac{b^{M,iM}(t - \tau_2)}{t_b}$$

$$\dot{b}^{W,iW}(t) = \sum_{iD} o_r^{D,iD}(t) - \sum_{iD} o_r^{D,iD}(t - \tau_1) - \frac{b^{W,iW}(t)}{t_b}$$

$$\dot{s}^{D,iD}(t) = o_r^{D,iD}(t - \tau_1 - \tau_2) - \sum_{iR} o_r^{R,iR}(t - \tau_1) + \sum_{iW} \frac{b^{W,iW}(t - \tau_2)}{t_b} - \frac{b^{D,iD}(t)}{t_b} \quad (۹)$$

$$\dot{o}_u^{D,iD}(t) = o_r^{D,iD}(t) - o_r^{D,iD}(t - \tau_1 - \tau_2) - \sum_{iW} \frac{b^{W,iW}(t - \tau_2)}{t_b}$$

$$\dot{b}^{D,iD}(t) = \sum_{iR} o_r^{R,iR}(t) - \sum_{iR} o_r^{R,iR}(t - \tau_1) - \frac{b^{D,iD}(t)}{t_b}$$

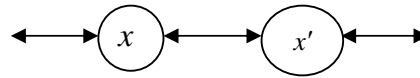
$$\dot{s}^{R,iR}(t) = o_r^{R,iR}(t - \tau_1 - \tau_2) - d_r^{R,iR}(t - \tau_1) + \sum_{iD} \frac{b^{D,iD}(t - \tau_2)}{t_b} - \frac{b^{R,iR}(t)}{t_b} \quad (۱۰)$$

$$\dot{o}_u^{R,iR}(t) = o_r^{R,iR}(t) - o_r^{R,iR}(t - \tau_1 - \tau_2) - \sum_{iD} \frac{b^{D,iD}(t - \tau_2)}{t_b}$$

$$\dot{b}^{R,iR}(t) = d_r^{R,iR}(t) - d_r^{R,iR}(t - \tau_1) - \frac{b^{R,iR}(t)}{t_b}$$

این روابط تمامی گره‌های شکل ۵ را در رده‌های مختلف مدل می‌کنند. در این معادلات به ازای هر گره از $x \in \{S, M, W, D, R\}$ متغیر حالت $s^{x,iX}(t)$ و $o_u^{x,iX}(t)$ و $b^{x,iX}(t)$ یک متغیر کنترل $o_r^{x,iX}(t)$ و برای هر گره خرده‌فروش، یک ورودی اغتشاش اندازه‌گیری شده $d_r^{R,iR}(t)$ در نظر گرفته می‌شوند. بر اساس این معادلات دیده می‌شود که یک وابستگی اطلاعاتی رده‌های مجاور را کاملاً جفت می‌کند. به طور مثال رابطه (۶) برای تامین‌کننده‌ها به $o_r^{M,iM}(t - \tau_1)$ تولیدکننده‌ها و رابطه (۷) برای تولیدکننده‌ها به $b^{S,iS}(t - \tau_2)$ تامین‌کننده‌ها نیاز دارند. این روند وابستگی برای تمامی گره‌های مجاور دیده می‌شود. در قسمت بعد

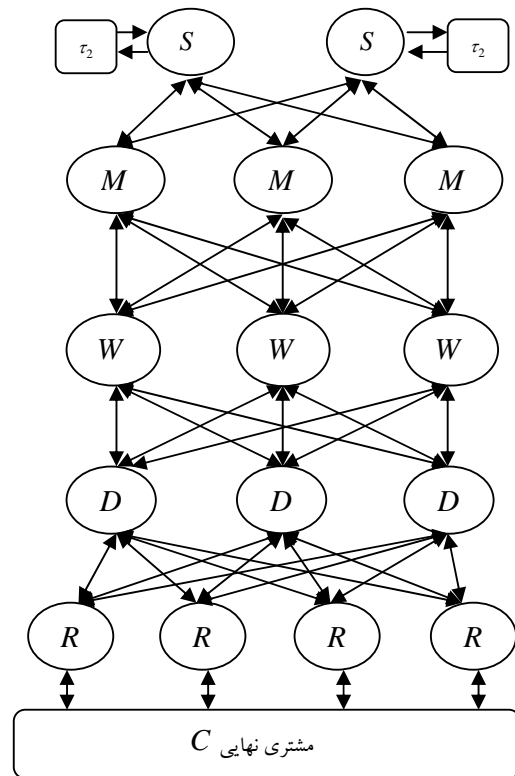
می‌شود. در این شکل پیکان دوطرفه نمایانگر جریان تاخیردار انتقال کالا و برگشت اطلاعات از سوی مشتری است.



شکل ۴: مدل جایگزین شکل ۳

سیستم مدیریت زنجیره تامین مدنظر، تک محصولی بوده و دارای ۲ گره تامین، ۳ گره تولید، ۳ گره انبار، ۳ گره توزیع و ۴ گره خرده‌فروش در ۵ رده است. این سیستم در شکل ۵ به صورت عمودی و از بالا به پایین نشان داده شده است. تنها یک استثنا در ابتدای زنجیره وجود دارد که جریان اطلاعات آمده از تامین‌کننده از طریق تولید به کالا تبدیل می‌شود و سپس جریان به تامین‌کننده بر می‌گردد. برای سادگی این تبدیل نیز با یک تاخیر و رابطه‌ی مدل می‌شود.

$$a_r^{S,iS}(t) = o_r^{S,iS}(t - \tau_2) \quad (۵)$$



شکل ۵: سیستم مدیریت زنجیره تامین ۵ رده‌ای مدنظر

بنابراین زنجیره تامین به وسیله تقاضای مشتری یعنی تعداد دستگاه‌ها و کالاهای فروخته شده در روز تحریک می‌شود و سپس یک سری جریان‌های اطلاعات و کالا در سیستم به راه می‌افتند.

$W_{ou}^{x,ix}$ مربوط به هزینه‌های سفارش‌های برآورده نشده و $W_u^{x,ix}$ مربوط به هزینه‌های کنترل و سفارش دادن، هستند. همچنین $s_d^{x,ix}$ سطح موجودی مطلوب ثابت و $o_{ud}^{x,ix}$ سفارش اجراشده مطلوب در رده X می‌باشند. همچنین چون متغیرهای $b^{x,ix}(t_k)$ کنترل‌ناپذیرند، در مسئله بهینه‌سازی وارد نمی‌شوند. همان‌گونه که ذکر شد، اعمال یک بخش توقیف حرکت که سرعت تغییر متغیرهای ورودی کنترلی را جریمه می‌کند، مقاومت سیستم کنترلی را در برابر تغییرات تقاضا، اثر شلاق چرمی و نامعینی‌های سیستم افزایش می‌دهد. از اینرو یک بخش توقیف حرکت نیز به این تابعی معیار اضافه شده است که در آن فاکتور $W_{ou}^{x,ix}$ مربوط به هزینه‌های بخش توقیف حرکت استفاده در رده X است.

۴- مطالعه موردی و نتایج شبیه‌سازی

از آنجا که مدل‌سازی حاضر بر اساس فیزیک یک مسئله مدیریت زنجیره تامین واقعی همراه با اطلاعات تقاضا در همه‌ی رده‌هاست و برخلاف مدل‌سازی‌های تجربی یک مدل مبتنی بر اصول اساسی سیستم به شمار می‌آید، روابط ۶ الی ۱۰ هر سیستم مدیریت زنجیره تامین با هر تعداد رده و هر تعداد گره را دربر می‌گیرند. در مورد مطالعه‌ی این مقاله، سیستم مدیریت زنجیره تامین شکل ۵ در نظر گرفته شده است. در این مورد سعی بر اینست که مدل حاضر به یک زنجیره تامین با ابعاد واقعی نزدیک شود و همچنین انتقال کالا بین همه‌ی گره‌ها در نظر گرفته شده است. در طول این زنجیره تنها یک نوع کالا پخش می‌شود. مدل‌های دینامیکی که مدل‌های دینامیکی زمان پیوسته هستند با استفاده از امکانات نرم‌افزار مطلب^۱ با زمان نمونه‌برداری کنترل‌کننده پیش‌بین به مدل‌های زمان گسسته تبدیل می‌شوند. تاخیرهای زمانی موجود نیز با تقریب زنده مرتبه چهار پیدا^۲ تقریب زده شده و جایگزین می‌شوند. نرخ‌های داده‌شده در این مسئله بر حسب روز هستند و زمان نمونه‌برداری (واحد زمان) مدل سیستم و کنترل‌کننده پیش‌بین برابر با ۰.۲ روز (۴.۸ ساعت) است. مقادیر نقاط تنظیم موجودی، ظرفیت‌های ذخیره ماکزیمم در هر گره و فاکتورهای وزنی نیز در جدول ۱ آمده‌اند.

به تشریح مسئله کنترلی و دلایل استفاده از کنترل‌کننده پیش‌بین برای حل این مسئله پرداخته شده و در نهایت، کنترل‌کننده پیش‌بین مقتضی طراحی می‌شود.

۳- روش تحقیق و طراحی کنترل‌کننده

پیش‌بین مقید

با توجه به مدل‌سازی قسمت قبل، سیستم مدیریت زنجیره تامین مدنظر مقاله، یک سیستم چند ورودی-چند خروجی ابعاد بالا همراه با اغتشاش تقاضای مشتری است که (در صورت تغییر) دینامیک مسئله را در هر لحظه از زمان دچار تغییر می‌کند. همچنین یک سری محدودیت‌های ورودی و خروجی همانند نامنفی بودن تعداد واحد کالا و حداکثر ظرفیت موجودی هر گره زنجیره وجود دارند. کنترل پیش‌بین از جهت در نظر گرفتن محدودیت‌های ورودی‌ها، حالت‌ها و خروجی‌های سیستم در محاسبات کنترلی، و پاسخ مناسب به تغییرات تقاضا، یک روش مناسب برای کنترل این سیستم‌ها به نظر می‌رسد [۱۵]. بطور کلی اهداف اصلی روش کنترل زنجیره تامین بدین صورت می‌باشد: بیشینه کردن رضایت مشتری و، حداقل کردن هزینه‌های عملیاتی زنجیره تامین [۱۶]. در ادامه باید این معادلات (۶ الی ۱۰) بر اساس زمان نمونه‌برداری کنترل‌کننده پیش‌بین (روز) استفاده شده به صورت زمان گسسته بازنویسی شوند. پس در حالت کلی با یک تابعی معیار کاملاً مربعی و با توجه به محدودیت‌های رابطه ۲ و دینامیک‌های روابط ۶ الی ۱۰ به صورت زمان گسسته، مسئله‌ی زیر حاصل می‌شود.

(۱۱)

$$J_{total} = \sum_{x \in \{S, M, W, D, R\}} \sum_{ix} \left(\sum_{j=1}^p (W_s^{x,ix} [s^{x,ix}(t_{k+j}) - s_d^{x,ix}]^2 + W_{ou}^{x,ix} [o_u^{x,ix}(t_{k+j}) - o_{ud}^{x,ix}]^2) + \sum_{j=0}^{c-1} (W_u^{x,ix} [o_r^{x,ix}(t_{k+j})]^2 + W_{ou}^{x,ix} [o_r^{x,ix}(t_{k+j}) - o_r^{x,ix}(t_{k+j-1})]^2) \right),$$

$$\text{subject to } \begin{cases} \text{Equations 6 to 10} \\ 0 \leq (s^{x,ix}(t_k), o_u^{x,ix}(t_k), b^{x,ix}(t_k)) \leq s_{max}^{x,ix} \\ 0 \leq o_r^{x,ix}(t_k) \leq o_{rmax}^{x,ix} \end{cases}$$

در اینجا نقاط تنظیم موجودی کالا پارامترهای زمان ثابت هستند و t_k به عنوان نمونه‌های زمان گسسته، فاکتورهای وزنی ثابت نامنفی $W_s^{x,ix}$ مربوط به هزینه‌های ذخیره موجودی کالا و وسایل موجودی کالا،

¹ Matlab² Pade

جدول ۱ - داده‌های زنجیره تامین

خرده فروش	توزیع کننده	انبار	تولید کننده	تامین کننده	رده داده
۲۵۰	۲۵۰	۳۵۰	۳۵۰	۴۰۰	بیشینه موجودی مجاز
۱۵۰	۱۵۰	۲۰۰	۲۰۰	۲۵۰	بیشینه سفارش برآورده نشده مجاز
۸۰					بیشینه سفارش انباشته شده مجاز
۸۰	۱۰۰	۱۸۰	۲۵۰	۳۲۰	نقطه تنظیم موجودی محصول
مقدار حالت ماندگار تقاضا + مقدار اولیه سفارش برآورده نشده					نقطه تنظیم سفارش برآورده نشده
[۰.۸, ۰.۱, ۰.۱, ۰.۱]					$W_{ou}^{s,i,x}, W_u^{s,i,x}, W_{\delta u}^{s,i,x}$
۸۰	۱۰۰	۱۸۰	۲۵۰	۳۲۰	سطح موجودی اولیه
۳۰					مقدار اولیه سفارش برآورده نشده
۳۰					مقدار اولیه سفارش انباشته شده

به شکل ۵ نرخ سفارش برای هر رده پایین دستی برابر با نرخ تقاضای رده بالادستی است. تغییرات تقاضای محصول از سوی مشتری در جلوی زنجیره تامین، هر چه به عقب زنجیره تامین حرکت می‌کنیم به نوسانات بیشتر و بیشتر در تقاضا برای رده‌ها تبدیل می‌شود (اثر شلاق چرمی).

در ادامه برای بررسی بهتر اثر تغییرات تقاضای مشتری، از یک نرخ تقاضای مشتری پالسی با دامنه‌های متفاوتی از واحد کالا در دوره‌های زمانی مختلف استفاده می‌شود. شکل‌های ۸ و ۹، انحراف از سطوح موجودی در کل زنجیره و ورودی‌های سیستم را بدون استفاده از بخش توقیف حرکت و شکل‌های ۱۰ و ۱۱، انحراف از این سطوح موجودی و ورودی‌های سیستم را همراه با استفاده از بخش توقیف حرکت نشان می‌دهند. با توجه به شکل ورودی رده‌های مختلف دیده می‌شود که از رده‌های پایین دستی به سمت رده‌های بالادستی میزان بیشینه ورودی کنترلی بیشتر می‌شود، اما در نهایت همگی به مقدار حالت ماندگار خواهند رسید (اثر شلاق چرمی). در این شبیه‌سازی‌ها به وضوح، استفاده از بخش توقیف حرکت، اثر شلاق چرمی و اثر بد نامعینی‌های سیستم را کاهش می‌دهد و همچنین با پیک و ورودی کمتری روبرو خواهیم بود.

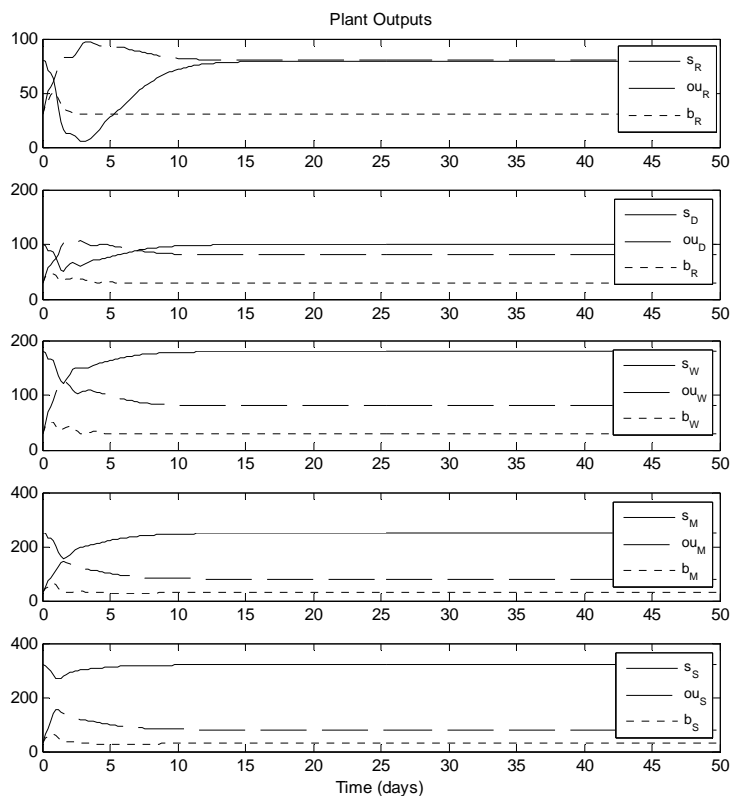
در نهایت برای درک بهتر عملکرد مناسب روش کنترلی روی سیستم مدیریت زنجیره تامین، نرخ تقاضای مشتری به صورت تصادفی بین ۱ تا ۱۰۰ واحد کالا در نظر گرفته می‌شود. در این بخش نیز کنترل کننده یکبار بدون بخش توقیف حرکت و یکبار همراه با این بخش در نظر گرفته می‌شود. با توجه به شکل‌های ۸ و ۹ و ۱۰ و ۱۱، قبل از شبیه‌سازی‌های این بخش به نظر می‌رسد استفاده از بخش توقیف حرکت در برابر تغییرات ناگهانی تقاضا و دستیابی به یک مقاومت نسبی در برابر این تغییرات ضروری است. شکل‌های ۱۲ و ۱۳، انحراف از سطوح موجودی در کل زنجیره و ورودی‌های سیستم را بدون استفاده از بخش توقیف حرکت و شکل‌های ۱۴ و ۱۵، انحراف از این سطوح موجودی و ورودی‌های سیستم را همراه با استفاده از بخش توقیف حرکت نشان می‌دهند. همانطور که انتظار می‌رفت، جریمه نکردن تغییرات ورودی منجر به نوسانات شدید در نمودار خروجی‌ها و ورودی‌های سیستم می‌شود. اما با استفاده از بخش توقیف حرکت، سطوح موجودی در حوالی نقاط تنظیم موجودی، کنترل می‌شوند. همچنین اثر بد شلاق چرمی در طول زنجیره به شدت کاهش می‌یابد.

باید توجه داشت که در تمام شبیه‌سازی‌های انجام شده، محدودیت‌های ورودی و خروجی در نظر گرفته شده‌اند. حداقل ورودی‌ها و خروجی‌ها، صفر در نظر گرفته شده و حداکثر آن‌ها در

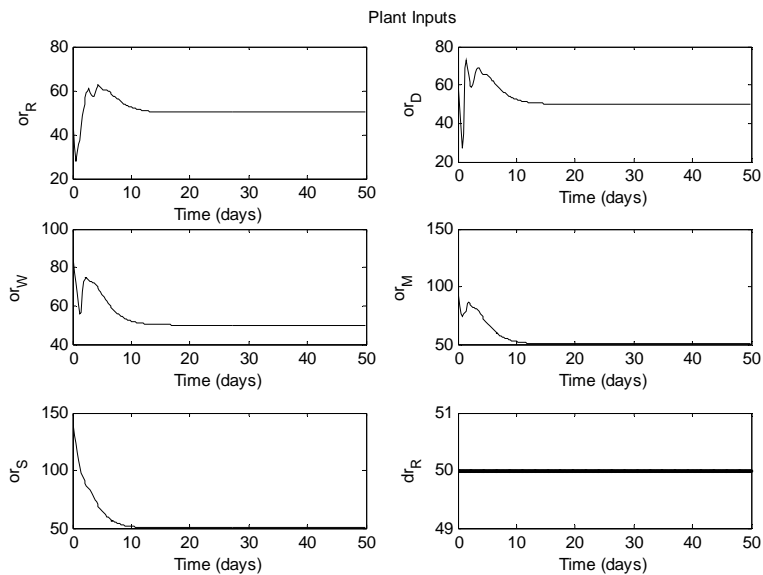
افق پیش‌بینی برابر با ۲۰ واحد زمانی (۴ روز) و افق کنترل برابر با ۸ واحد زمانی (۱.۶ روز) در نظر گرفته می‌شوند. تمامی متغیرهای حالت به عنوان خروجی‌های سیستم مدیریت زنجیره تامین هم در نظر گرفته می‌شوند. در نهایت ۴۵ متغیر خروجی و ۳۸ متغیر ورودی و ۴ متغیر اغتشاشی اندازه‌گیری شده در این مدل وجود دارد. مقادیر حالت‌ماندگار سفارش‌های انباشته شده به علت پایدار بودن در نقطه تعادل صفر و شکل نمایی، خود به خود پس از حالت گذار صفر می‌شوند. تاخیرهای زمانی t_b و t_1 و t_2 ، نیز به ترتیب معادل ۰.۵ و ۱ و ۱ واحد زمانی در نظر گرفته می‌شوند. پاسخ دینامیکی زنجیره برای یک مقدار ثابت ۵۰ واحد کالا در تقاضای مشتری محاسبه شد. شکل‌های ۶ و ۷، انحراف از سطوح موجودی (خروجی‌های سیستم) برای رده‌های مختلف زنجیره و ورودی‌های سیستم (ورودی‌های کنترلی و اغتشاش اندازه‌گیری شدهی تقاضا) را در ۵۰ روز بدون استفاده از بخش توقیف حرکت نشان می‌دهند. از آنجا که مقدار تقاضای مشتری ثابت است، پاسخ همراه با توقیف حرکت برای تقاضای ثابت خیلی متفاوت از شکل‌های ۶ و ۷ نمی‌شود (ضمناً با اعمال بخش توقیف حرکت، پیچیدگی محاسباتی بالاتر می‌رود).

اثر شلاق چرمی بوضوح در این شکل‌ها قابل مشاهده است. با توجه

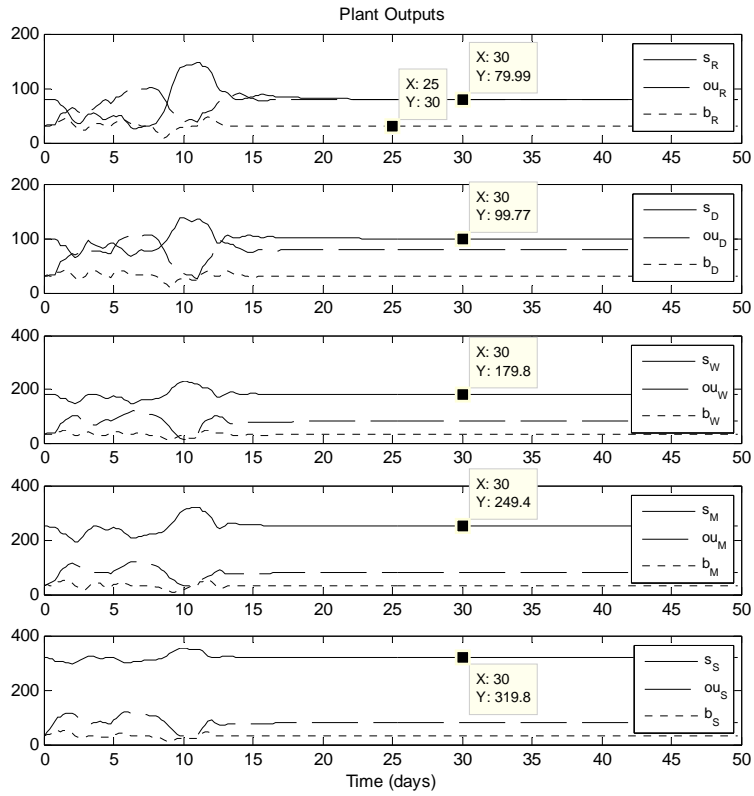
جدول ۱ تعیین شده‌اند.



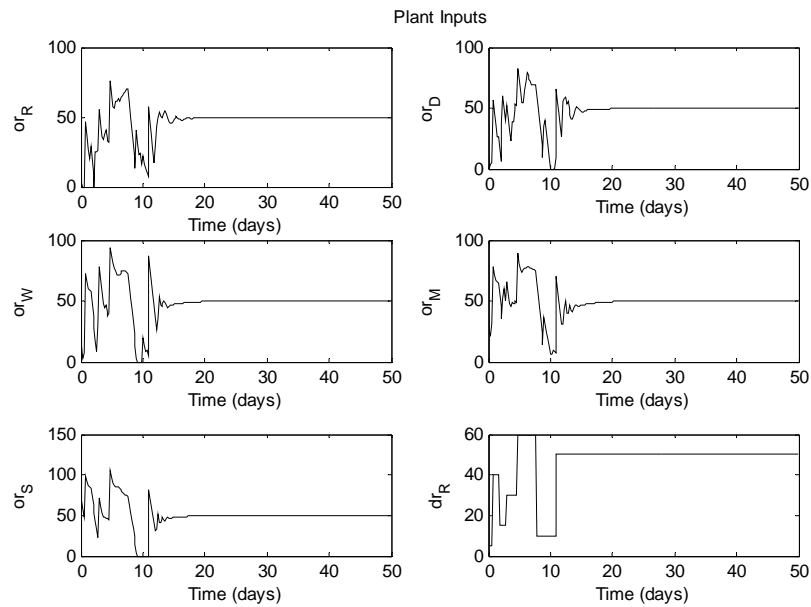
شکل ۶- خروجی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای ثابت ۵۰ واحد کالا در روز



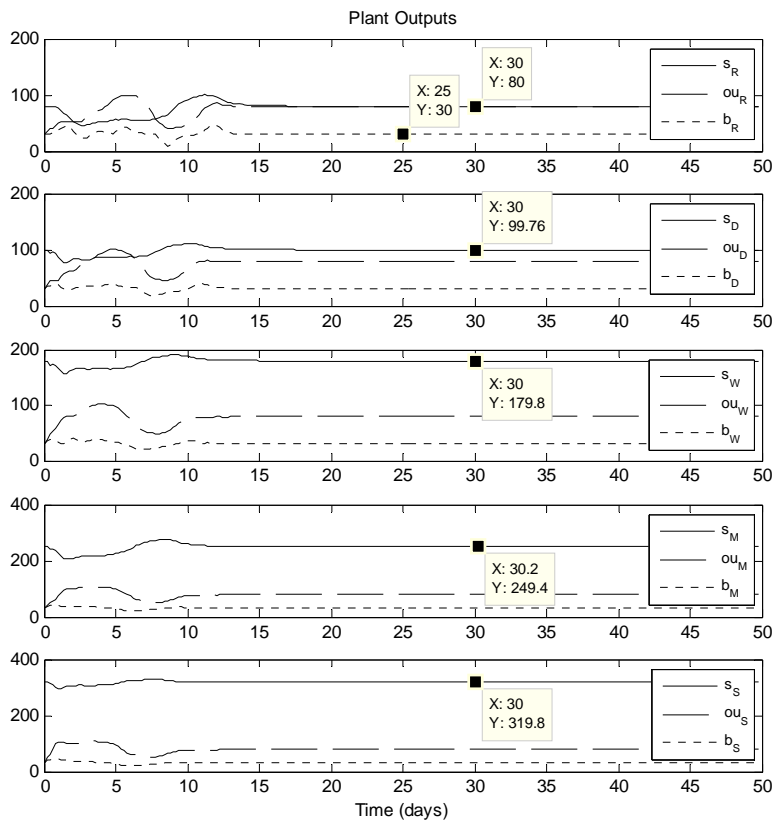
شکل ۷- ورودی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای ثابت ۵۰ واحد کالا در روز



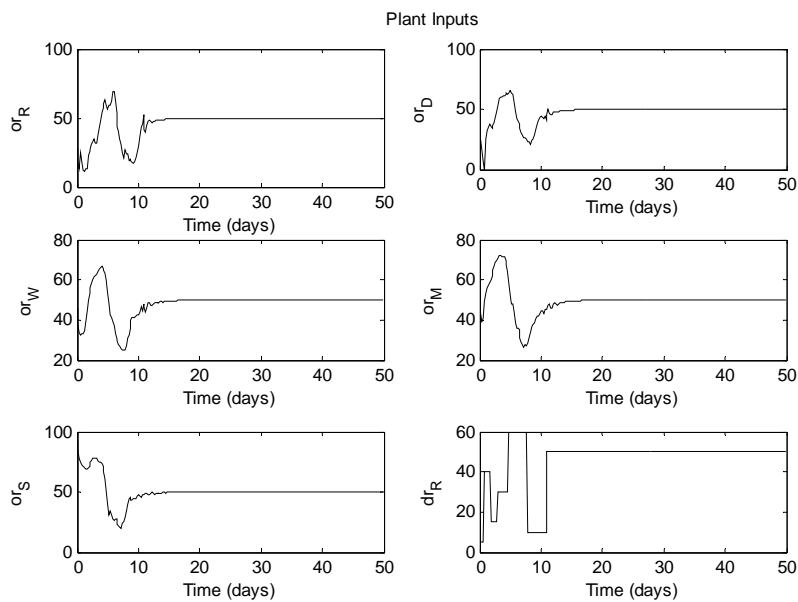
شکل ۸- خروجی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای پالسی بدون توقف حرکت



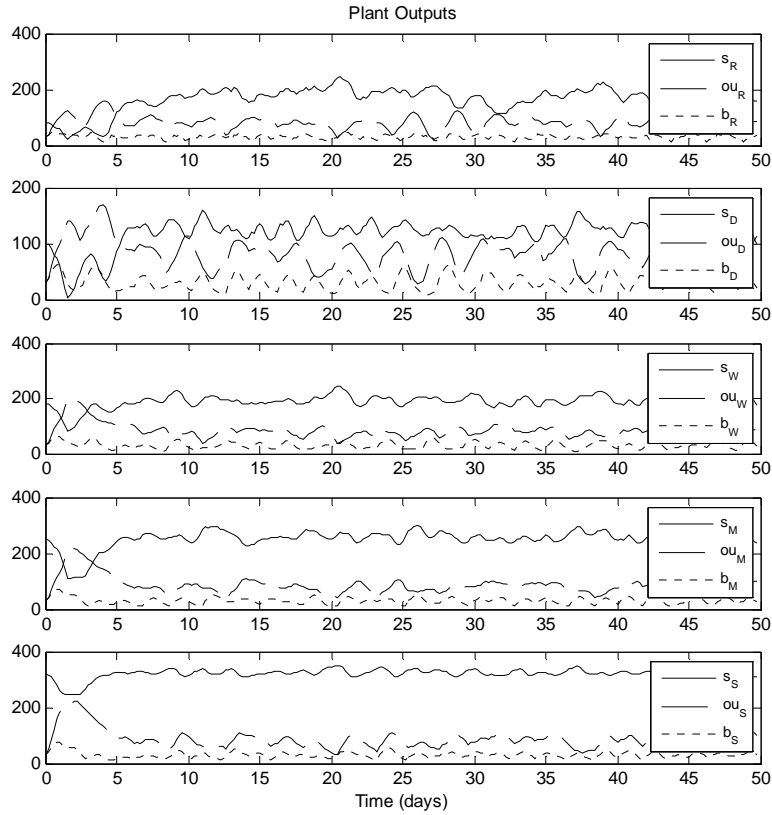
شکل ۹- ورودی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای پالسی بدون توقف حرکت



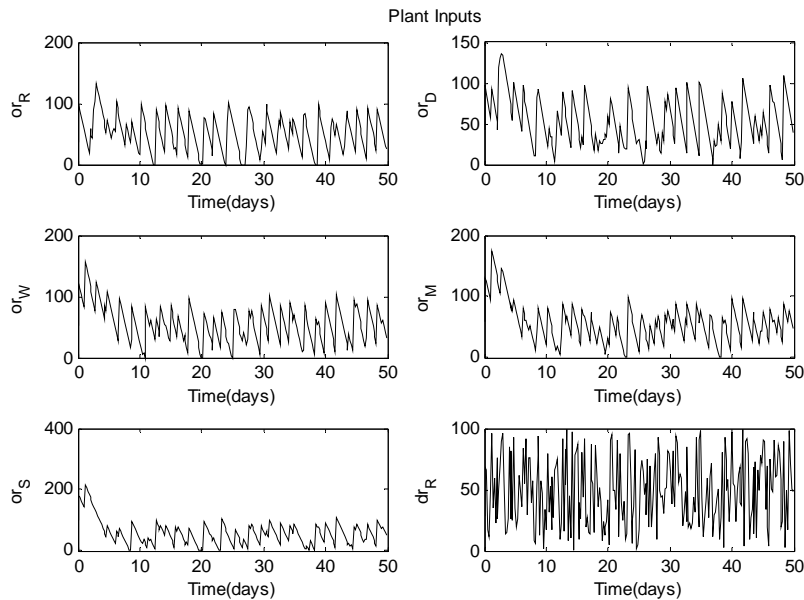
شکل ۱۰- خروجی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای پالسی همراه با توقیف حرکت



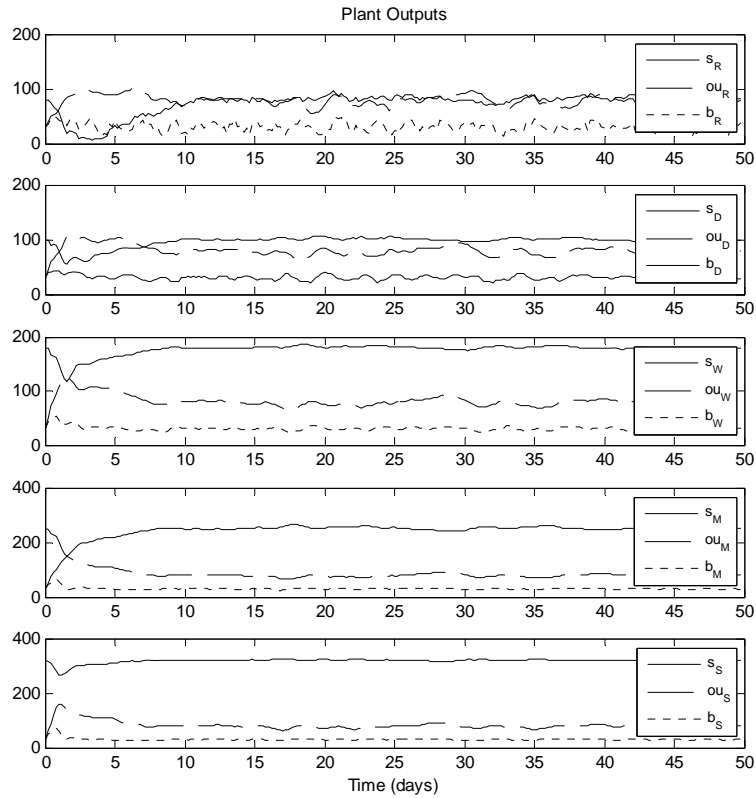
شکل ۱۱- ورودی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای پالسی همراه با توقیف حرکت



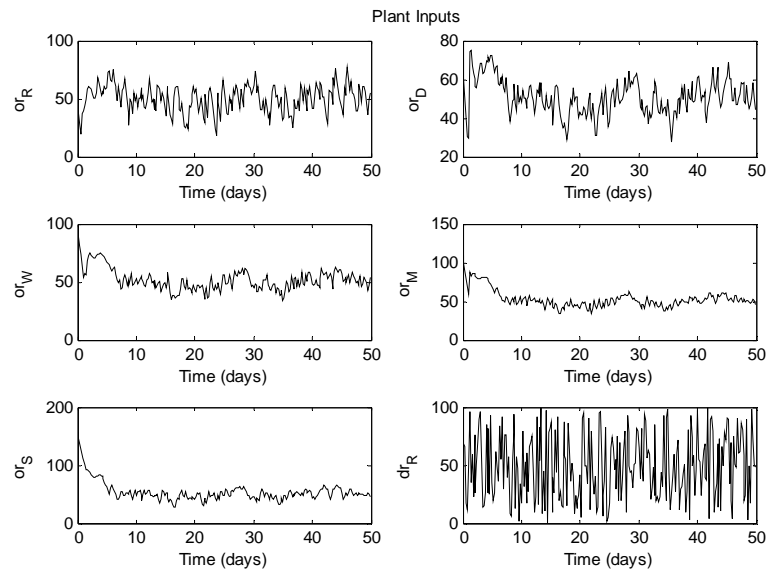
شکل ۱۲ - خروجی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای تصادفی بدون توقیف حرکت



شکل ۱۳ - خروجی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای تصادفی بدون توقیف حرکت



شکل ۱۴ - خروجی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای تصادفی همراه با توقیف حرکت



شکل ۱۵ - ورودی‌های سیستم مدیریت زنجیره تامین در برابر نرخ تقاضای تصادفی همراه با توقیف حرکت

- [4]. R. Grubbstrom, T. Huynh, "multi-stage capacity constrained production inventory systems in discrete-time with non-zero lead times using MRP theory," *International Journal of Production Economics*, vol. 101, pp. 53–62, 2006.
- [5]. S. Keerthi, E. Gilbert, "Optimal, infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations," *Journal of Optimization Theory and Application*, vol. 57, pp. 265–93, 1998.
- [6]. P. S. Agachi, Z. K. Nagy, M. V. Cristea, A. Imre-Lucaci, *Model Based Control*, WILEY-VCH Verlag GmbH & Co, 2009.
- [7]. G. Kapsiotis, S. Tzafestas, "Decision making for inventory/production planning using model-based predictive control," *Parallel and distributed computing in engineering systems*. Amsterdam: Elsevier, pp. 551–556, 1992.
- [8]. S. Tzafestas, G. Kapsiotis, "Model-based predictive control for generalized production planning problems," *Computers in Industry*, vol. 34, pp. 201–210, 1997.
- [9]. P. Lopez, B. Ydstie Grossmann, "A model predictive control strategy for supply chain management," *Computers & Chemical Engineering*, vol. 27, pp. 1201–18, 2003.
- [10]. W. Wang, R. Rivera, "A novel model predictive control algorithm for supply chain management in semiconductor manufacturing," *Proceedings of the American control conference*, vol. 1, pp. 208–213, 2005.
- [11]. I. Yildirim, B. Tan, F. Karaesmen, "A multiperiod stochastic production planning and sourcing problem with service level constraints," *OR Spektrum*, vol. 27, pp. 471–89, 2005.
- [12]. S. Chopra, P. Meindl, *Supply Chain Management Strategy, Planning and Operations*, Pearson Prentice Hall Press, New Jersey, pp. 58–79, 2004.
- [13]. W. Dunbar, S. Desa, "Model predictive control for dynamic supply chain management," *Proceedings of the International Workshop on Assessment and Future Directions of NMPC*, Freudenstadt-Lauterbad, Germany, August, 2005.
- [14]. H. Sarimveis, P. Patrinos, D. Tarantilis, T. Kiranoudis, "Dynamic modeling and control of supply chain systems: A review," *Computers & Operations Research*, vol. 35, pp. 3530 – 3561, 2008.
- [15]. E. F. Camacho, C. Bordons, *Model Predictive Control*. Springer, 2004.
- [16]. J. D. Stermann, *Business Dynamics Systems Thinking and Modeling in A Complex World*, McGraw Hill Press, pp. 113–128, New York, 2000.

۸- نتیجه‌گیری

روش کنترل پیش‌بین از جهت در نظر گرفتن محدودیت‌های ورودی‌ها، حالت‌ها و خروجی‌های سیستم در محاسبات کنترلی، و پاسخ مناسب به تغییرات تقاضا، یک روش مناسب برای کنترل سیستم‌های مدیریت زنجیره تامین است. شبیه‌سازی‌های انجام شده در این مقاله، این امر را با هدف کنترلی حفظ سطوح موجودی برای پیشینه کردن رضایت مشتری با کمترین هزینه عملیاتی به خوبی نشان می‌دهند. در این مقاله یک سیستم مدیریت زنجیره تامین جفت شده‌ی اطلاعاتی، برای مدل کردن یک زنجیره تامین واقعی به ابعاد وسیعی در رده‌ها و گره‌ها گسترش داده شد. بدین ترتیب این سیستم پنج رده‌ای و کنترل پیش‌بین روی آن قابل تعمیم به هر مدل از زنجیره تامین است. این سیستم یکبار همراه با تقاضای ثابت مشتری و بار دیگر با نرخ تقاضای متغیر پالسی و در نهایت با نرخ تقاضای متغیر تصادفی در نظر گرفته شد، تا کارایی اعمال کنترل پیش‌بین روی این سیستم‌ها نشان داده شود. همچنین در این مقاله اثر شلاق چرمی و همچنین اثر اعمال بخش توقیف حرکت بررسی شد. جریمه نکردن تغییرات ورودی منجر به نوسانات شدید در نمودار خروجی‌ها و ورودی‌های سیستم می‌شود. اما با استفاده از بخش توقیف حرکت، سطوح موجودی در حوالی نقاط تنظیم موجودی، کنترل می‌شوند. همچنین اثر بد شلاق چرمی در طول زنجیره به شدت کاهش می‌یابد.

مراجع

- [1]. H. Wiendahl, J. Breithaupt, "Automatic production control applying control theory," *International Journal of Production Economics*, vol. 63, pp. 33–46, 2000.
- [2]. J. Dejonckheere, SM. Disney, M. R. Lambrecht, D. R. Towill, "The impact of information enrichment on the bullwhip effect in supply chains: a control engineering perspective," *European Journal of Operational Research*, vol. 153, pp. 727–50, 2009.
- [3]. D. R. Towill, G. N. Evans, P. Cheema, "Analysis and design of an adaptive minimum reasonable inventory control system," *Production Planning & Control*, vol. 8, pp. 545–557, 1997.

یک الگوریتم یادگیری ترکیبی جهت طراحی شبکه‌های موجک فازی برای تقریب توابع، شناسایی و کنترل بهنگام سیستم‌های غیرخطی

مریم شهریاری کاهکشی^۱، مریم ذکری^۲

^۱ دانشجوی دکتری مهندسی برق، گروه کنترل، دانشگاه صنعتی اصفهان، m.shahriarikahkeshi@ec.iut.ac.ir

^۲ استادیار، دانشکده مهندسی برق و کامپیوتر، گروه کنترل، دانشگاه صنعتی اصفهان، mzekri@cc.iut.ac.ir

(تاریخ دریافت مقاله ۱۳۸۹/۱۱/۱۹، تاریخ پذیرش مقاله ۱۳۹۰/۱/۲۶)

چکیده: در این مقاله، یک الگوریتم یادگیری ترکیبی جدید برای طراحی شبکه‌های موجک فازی به منظور تقریب توابع، کنترل و شناسایی سیستم‌های غیرخطی ارائه شده است. الگوریتم پیشنهادی مبتنی بر الگوریتم حداقل مربعات متعامد (OLS)، الگوریتم جهش قورباغه‌های بهم آمیخته (SFL) و روش حداقل مربعات بازگشتی (RLS) می‌باشد. الگوریتم حداقل مربعات متعامد جهت تعیین ابعاد شبکه، غربال کردن موجکها به منظور انتخاب موجکهای مؤثر و تعیین تعداد قوانین فازی استفاده می‌شود. بدین ترتیب با انتخاب موجکهای مؤثر بر اساس داده‌های آموزشی شبکه موجک فازی ساخته شده و مقادیر اولیه پارامترهای شبکه تعیین می‌شوند. سپس پارامترهای خطی و غیرخطی شبکه به ترتیب با استفاده از روش حداقل مربعات بازگشتی و الگوریتم جهش قورباغه‌ها تنظیم می‌گردند. برای نشان دادن قابلیت و توانایی روش پیشنهادی، نتایج شبیه‌سازی در طی چند مثال برای تقریب تابع، کنترل و شناسایی سیستم‌های غیرخطی آورده شده است. همچنین، نتایج بدست آمده از این روش با نتایج روش‌های گزارش شده در مقالات دیگر نیز مقایسه شده است. نتایج شبیه‌سازی نشان می‌دهند که روش پیشنهادی مقاله، ضمن استفاده از تعداد قانون فازی کمتر، صحت تقریب و شاخص عملکرد مدل را بهبود بخشیده و از عملکرد بهتری برای سیستم‌های مورد مطالعه، در مقایسه با سایر روشها، برخوردار است.

کلمات کلیدی: شبکه‌های موجک فازی، الگوریتم جهش قورباغه‌ها، تقریب توابع، شناسایی و کنترل سیستم‌های غیرخطی.

A Hybrid Learning Algorithm for Fuzzy Wavelet Networks Design for Functions Approximation, Online Identification and Control of Nonlinear Systems

Maryam Shahriari Kahkeshi, Maryam Zekri

Abstract: In this paper, a hybrid learning algorithm is presented for fuzzy wavelet networks (FWNs) design for functions approximation, online identification and control of nonlinear systems. The proposed algorithm is based on orthogonal least square (OLS) algorithm, Shuffled Frog Leaping (SFL) algorithm and recursive least square method (RLS). The OLS algorithm is used for determine network dimensions, number of fuzzy rules and wavelets in each fuzzy rule and for purifying wavelets in each sub-WNN. So, after selection of important wavelets based on training data set, FWN structure is constructed and initial values of the network parameters are determined. Then linear and nonlinear parameters of the network are tuned based on recursive least square method and SFL algorithm, respectively. In order to show the capabilities and effectiveness of the proposed method, simulation results are presented for some example: function approximation, online identification and control of nonlinear systems. Also, the results obtained by the proposed approach are compared with the previous approaches reported in the literature. Simulation results show that the proposed method improves model approximation accuracy and performance index by using less number of fuzzy rules compare to other methods for study systems.

Keywords: Fuzzy Wavelet Networks, Shuffled Frog Leaping Algorithm, Functions Approximation, Identification and Control of Nonlinear System.

۱- مقدمه

در سال های اخیر، تئوری موجک به طور گسترده مورد مطالعه قرار گرفته و در حوزه های مختلفی از علوم و مهندسی اعمال شده است [۱]-[۴]. موجکها دارای ویژگی بسیار مهم مکان یابی زمان فرکانس بوده و قادرند به آسانی رفتار محلی (فرکانس بالا) و رفتار سراسری (فرکانس پایین) هر تابع را نمایش دهند. اما ساخت و ذخیره موجکهای پایه با ابعاد بزرگ هزینه زیادی می طلبد. لذا برای مواجهه با مسائل دارای ابعاد بزرگ، لازم است الگوریتم هایی توسعه داده شوند که اجرای آنها به ابعاد ورودی حساسیت کمتری داشته باشد.

از طرفی، شبکه های عصبی به عنوان ابزاری توانمند برای حل مسائل با ابعاد بزرگ معرفی شده و مطالعات بسیاری بر روی توانایی این شبکه ها برای تقریب توابع غیرخطی گزارش شده است [۵]-[۷]. اگرچه شبکه های عصبی دارای خواص مهمی از جمله قابلیت یادگیری، تعمیم پذیری و محاسبات موازی می باشند، اما فاقد روش مناسبی برای تعیین ساختار شبکه و تنظیم مقادیر اولیه پارامترهای شبکه هستند. به منظور حل مشکل مطرح شده، ژانگ و بنویست در سال ۱۹۹۲، با ادغام خواص مفید شبکه های عصبی و خواص مکان یابی زمان- فرکانس موجکها، شبکه های عصبی موجک (WNN) را بعنوان یک ابزاری قوی و کارآمد ارائه نمودند [۸]. در شبکه های عصبی موجک، توابع موجک بعنوان تابع فعالیت، به جای توابع سیگموئید در لایه مخفی استفاده می- شوند. شبکه های موجک در مسائل مختلفی از قبیل تقریب توابع، شناسایی و کنترل سیستم های غیرخطی، پردازش سیگنال و طبقه بندی بکار رفته اند [۴]، [۹]-[۱۲].

پارامتر مقیاس، در شبکه های موجک، از نظر فیزیکی معرف دقت می باشد، لذا نقش مهمی در تقریب زنی یک تابع ایفا می کند. براساس این دیدگاه، دنیل و همکارانش، با ادغام تئوری آنالیز با دقت چندگانه تئوری موجک (MRA) و مدل فازی تاکاگی- سوگنو، شبکه های موجک فازی را برای تقریب توابع غیرخطی دلخواه ارائه نمودند [۱۳]. هدف از ادغام مدل فازی و شبکه های عصبی موجک، بهبود صحت تقریب زنی توابع با استفاده از نسخه های بسط و شیفته یافته یک موجک مادر، بدون افزایش تعداد پایه های موجک می باشد. از آن پس مدل های متفاوتی برای شبکه های موجک فازی به منظور شناسایی سیستم، پیش بینی سری های زمانی، مسائل کنترل و تقریب توابع مطرح شده است [۱۳]-[۱۹].

در [۱۳]، [۱۴] و [۱۹]، هر قانون فازی توسط یک زیرشبکه عصبی موجک که شامل موجکهای تک مقیاسی با پارامتر مقیاس یکسان است، نمایش داده شده است و از الگوریتم حداقل مربعات متعامد برای انتخاب موجکهای مهم استفاده شده است. شبکه موجک فازی به دست آمده در [۱۳]، برای تقریب تابع و در [۱۴] و [۱۹]، به منظور کنترل سیستم های غیرخطی بکار برده شده است. مدل ارائه شده در [۱۵]، تبدیل موجک گسسته را با سیستم های فازی تاکاگی- سوگنو ترکیب می کند، این مدل متشکل از یک مجموعه قوانین فازی است که بخش آنگاه قوانین، بسط سری های توابع موجک هستند، این مدل برای مدلسازی سیستم استفاده شده است.

نویسندگان مقاله در [۱۶]، یک شبکه موجک فازی مبتنی بر سیستم فازی تاکاگی- سوگنو- کانگ پیشنهاد کرده اند که بخش مؤخر قوانین فازی با توابع پایه موجک که قابلیت مکان یابی زمان- فرکانسی دارا می باشند، جایگزین شده است. همچنین از مدل پیشنهادی برای پیشگویی و شناسایی سیستم های دینامیکی غیرخطی استفاده نموده اند.

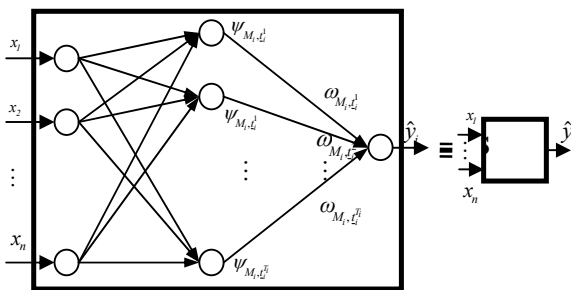
در مدل ارائه شده در [۱۷]، ورودی ها وارد بلوک تبدیل موجک گسسته شده و سپس خروجی هر بلوک فازی سازی شده و به عنوان ورودی وارد یک شبکه عصبی می شود، مدل پیشنهادی برای مسائل شناسایی سیستم و کنترل به کار رفته است. در [۱۸] شبکه موجک فازی برای تقریب توابع و شناسایی سیستم استفاده شده است که در آن یک روند خود-تنظیم مبتنی بر الگوریتم ژنتیک برای تعیین پارامترهای مهم شبکه موجک فازی مانند توابع عضویت، وزنها، پارامتر شیفته و مقیاس پیشنهاد شده است.

در این مقاله یک الگوریتم یادگیری ترکیبی برای طراحی شبکه های موجک فازی به منظور تقریب توابع، شناسایی و کنترل سیستم های غیرخطی ارائه شده است. این الگوریتم، ترکیبی از الگوریتم حداقل مربعات متعامد، الگوریتم جهش قورباغه ها و حداقل مربعات بازگشتی را به منظور غربال کردن موجکها، تعیین ابعاد شبکه، تعداد قوانین فازی و آموزش پارامترهای شبکه بکار می برد. مقداردهی اولیه مناسب پارامترهای شبکه، نقش بسزایی در سرعت همگرایی پاسخ شبکه ایفا می کند. به این ترتیب ابتدا با استفاده از الگوریتم حداقل مربعات متعامد، موجکهای مهم و مؤثر از میان موجکهای کاندید انتخاب شده و با توجه به پارامتر مقیاس موجکها، ابعاد شبکه، تعداد قوانین فازی، تعداد زیر شبکه های عصبی موجک و پارامتر مقیاس مربوط به هر زیرشبکه تعیین می شوند. به این ترتیب مقادیر اولیه پارامترهای شبکه شامل پارامتر شیفته و مقیاس موجکها، وزنها، زیرشبکه های عصبی موجک و

i ام و \hat{y}_i خروجی محلی برای قانون i ام می باشد که برابر مجموع خطی یک مجموعه محدود از موجک های $\psi_{M_i, t^k}^{(k)}(\underline{x})$ با پارامتر مقیاس یکسان $M_i \in \mathbb{Z}$ است. در نهایت نمایانگر مجموعه فازی است که توسط تابع عضویت گوسی (۲) توصیف می شود. $A_j^i(x_j)$ درجه عضویت x_j در A_j^i می باشد.

$$A_j^i(x_j) = e^{-\frac{(x_j - p_{j1}^i)^2}{p_{j2}^i}} \quad (2)$$

بطوریکه $p_{j1}^i, p_{j2}^i \in \mathbb{R}$ می باشند. پارامترهای p_{j1}^i و p_{j2}^i به ترتیب مرکز و عرض تابع عضویت را نشان می دهند. هر شبکه موجک فازی شامل C زیرشبکه عصبی موجک می باشد. هر زیرشبکه عصبی موجک یک ساختار رگرسیون غیرخطی است که نگاهی ورودی- خروجی را توسط نسخه های بسط و شیفت داده شده توابع موجک نمایش می دهد [۱۴]. ساختار یک شبکه عصبی موجک در شکل (۱) نشان داده شده است.



شکل ۱: ساختار شبکه عصبی موجک [۱۴].

خروجی شبکه عصبی موجک طبق رابطه (۳) محاسبه می شود:

$$\hat{y}_i = \sum_{k=1}^{T_i} w_{M_i, t^k} \psi_{M_i, t^k}^{(k)}(\underline{x}) \quad (3)$$

در رابطه فوق $\psi_{M_i, t^k}^{(k)}(\underline{x})$ تابع فعالیت موجک مربوط به زیرشبکه عصبی موجک i ام می باشد که توسط ضرب تانسور توابع موجک یک بعدی بیان می شوند:

$$\begin{aligned} \psi_{M_i, t^k}^k &= 2^{M_i/2} \psi^{(k)}(2^{M_i} \underline{x} - t^k) \\ &= \prod_{j=1}^n 2^{M_i/2} \psi^{(k)}(2^{M_i} x_j - t_j^k) \end{aligned} \quad (4)$$

بطوریکه $t_j^k \in \mathbb{R}$ و $M_i \in \mathbb{Z}$ به ترتیب پارامترهای شیفت و مقیاس توابع موجک می باشند. همچنین بردار ورودی \underline{x} و بردار w_{M_i, t^k} وزن بین لایه مخفی و خروجی زیرشبکه عصبی موجک i ام می باشد. پارامتر

پارامترهای مربوط به توابع عضویت فازی تعیین می شوند. سپس با استفاده روش حداقل مربعات بازگشتی، پارامترهای خطی شبکه که شامل ضرایب وزن مربوط به هر زیرشبکه عصبی موجک می باشند، تنظیم می شوند. همچنین پارامترهای غیرخطی شبکه شامل، پارامترهای شیفت توابع موجک و پارامترهای مربوط به توابع عضویت فازی با در نظر گرفتن مربعات خطای بین خروجی شبکه موجک فازی و خروجی مطلوب به عنوان تابع هدف، با استفاده از الگوریتم شهودی جهش قورباغه ها بروز رسانی می شوند. در انتها عملکرد الگوریتم پیشنهادی برای مسائل تقریب تابع، کنترل و شناسایی سیستم های غیرخطی بررسی می شود. مقایسه نتایج بدست آمده از روش پیشنهادی با کارهای قبلی انجام شده، مؤید ویژگیهای بارز روش پیشنهادی در کاهش تعداد قوانین فازی، کاهش تعداد پارامترهای شبکه، بهبود شاخص عملکرد شبکه، صحت تقریب زنی بالا، افزایش سرعت همگرایی و کاهش حجم محاسبات می باشد.

ادامه مقاله به صورت زیر سازمان یافته است: برای فراهم نمودن یک زمینه مناسب، بخش دوم به اختصار به توضیح ساختار شبکه های موجک فازی و پس از آن، بخش سوم به معرفی الگوریتم جهش قورباغه های بهم آمیخته پرداخته است. بخش چهارم به معرفی الگوریتم یادگیری ترکیبی پیشنهادی اختصاص یافته است. در بخش پنجم مثال ها و سیستم های مورد مطالعه معرفی شده و نتایج حاصل از شبیه سازی ها و مطالعات عددی بر روی این سیستم های نمونه آورده شده است. سرانجام، بخش ششم به نتیجه گیری نهایی مقاله اختصاص یافته است.

۲- ساختار شبکه های موجک فازی

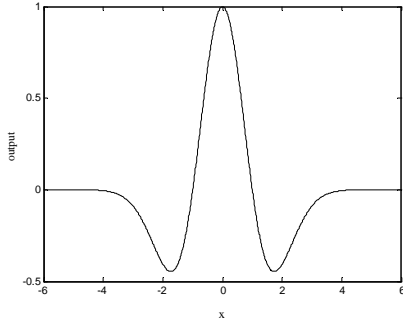
هر شبکه موجک فازی، برای تقریب یک تابع دلخواه را می توان توسط مجموعه ای از قوانین فازی به صورت رابطه (۱) توصیف کرد:

R^i : If x_1 is A_1^i and x_2 is A_2^i and ... and x_q is A_q^i ,

$$\text{Then } \hat{y}_i = \sum_{k=1}^{T_i} w_{M_i, t^k} \Psi_{M_i, t^k}^{(k)}(\underline{x})$$

$$\text{and } w_{M_i}^k \in \mathbb{R}, M_i \in \mathbb{Z}, t^k \in \mathbb{R}^q, x \in \mathbb{R}^q \quad (1)$$

که R^i قانون فازی i ام ($1 \leq i \leq C$) و C تعداد قوانین فازی است. x_j متغیر ورودی j ام ($1 \leq j \leq n$) و n ابعاد ورودی می باشد. همچنین، $t^k = [t_1^k, t_2^k, \dots, t_n^k]$ که مقدار شیفت برای موجک k ام است و $k = 1, 2, \dots, S$ که S برابر تعداد کل موجکهای انتخاب شده می باشد. همچنین T_i تعداد کل موجکها برای قانون فازی



شکل ۳: تابع موجک کلاسه مکزیک

همچنین پارامتر مقیاس برای موجک کلاسه مکزیک در گستره ۴- تا ۴ انتخاب شده است. ابتدا بر اساس داده های آموزشی، مقادیر پارامتر مقیاس M_i برای تابع موجک انتخاب شده، تعیین می شود. داده های آموزشی به صورت مجموعه (۸) تعریف می شوند:

$$\left\{ \left(\underline{x}_l^d, y_l^d \right), 1 \leq l \leq L, \underline{x}_l^d \in R^n, y_l^d \in R \right\} \quad (۸)$$

که L تعداد نمونه های آموزشی می باشد. هدف، ساخت و آموزش شبکه موجک فازی نشان داده شده در شکل (۲) براساس مجموعه داده های آموزشی (۸) می باشد، به گونه ای که خطای بین خروجی شبکه موجک فازی و y_l^d حداقل شود.

۳- الگوریتم جهش قورباغه های بهم آمیخته

الگوریتم جستجوی جهش قورباغه های بهم آمیخته (SFL) یا بطور خلاصه الگوریتم جهش قورباغه ها، یک الگوریتم جستجوی فرا ابتکاری جدید مبتنی بر جمعیت اولیه از خانواده الگوریتم های ممیتیک می باشد که از تکامل طبیعی گروهی از قورباغه ها زمانی که بدنبال محل با بیشترین ذخیره غذایی در دسترس می گردند، الهام گرفته شده است [۲۰].

الگوریتم SFL نخستین بار توسط یوسف و لنزی در سال ۲۰۰۳ برای تعیین اندازه بهینه لوله های جدید در توسعه شبکه لوله کشی مورد استفاده قرار گرفت [۲۱]. از آن پس SFL به عنوان یک روش بهینه سازی کارآمد مورد توجه محققین در زمینه های مختلف قرار گرفت [۲۲].

روند کار الگوریتم SFL

پروسه این الگوریتم با تولید جمعیت اولیه شروع می شود. ابتدا جمعیت اولیه ای شامل N قورباغه $P = \{X_1, X_2, \dots, X_N\}$ (قورباغه ها در الگوریتم SFL معادل کروموزوم ها در الگوریتم ژنتیک (GA) می باشند و نمایانگر یک راه حل شدنی در مسئله بهینه سازی هستند) بصورت تصادفی و در فضای ممکن Ω تولید می شود. در یک مسأله

شیفت مرکز تابع موجک را تعیین می کند، در حالیکه پارامتر مقیاس پهنای موجک را کنترل می نماید.

مطابق رابطه (۴) در هر قانون فازی یا زیر شبکه عصبی موجک، موجک های تک مقیاسی هستند و دارای پارامتر مقیاس یکسان در تمام ابعاد ورودی می باشند. به این ترتیب با استفاده از مکانیزم استنتاج فازی TSK خروجی شبکه موجک فازی (\hat{y}) توسط رابطه (۵) بیان می شود.

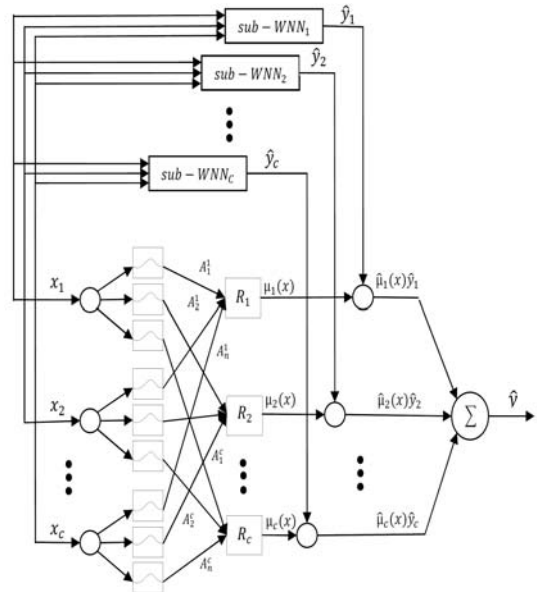
$$\hat{y}(x) = \frac{\sum_{i=1}^c \mu_i(x) \hat{y}_i}{\sum_{i=1}^c \mu_i(x)} = \sum_{i=1}^c \hat{\mu}_i(x) \hat{y}_i \quad (۵)$$

که داریم:

$$\hat{\mu}_i(x) = \mu_i(x) / \sum_{i=1}^c \mu_i(x), \quad \mu_i(x) = \prod_{j=1}^q A_j^i(x) \quad (۶)$$

بطوریکه $0 \leq \hat{\mu}_i \leq 1$ و $\sum_{i=1}^c \hat{\mu}_i = 1$ می باشد. $\hat{\mu}_i$ درجه شراکت

هر شبکه عصبی موجک با سطح دقت M_i را در خروجی کل شبکه تعیین می کند. شکل (۲)، ساختار شبکه های موجک فازی را نمایش می دهد.



شکل ۲: ساختار شبکه های موجک فازی

در این مقاله از نسخه های شیفت و بسط یافته تابع موجک کلاسه مکزیک که توسط رابطه (۷) بدست می آیند، استفاده است. شکل (۳)، تابع موجک کلاسه مکزیک را نشان می دهد.

$$\psi \left(\frac{x-t_j}{M_i} \right) = \left(1 - \left(\frac{x-t_j}{M_i} \right)^2 \right) \exp \left(-\frac{1}{2} \left(\frac{x-t_j}{M_i} \right)^2 \right) \quad (۷)$$

در روابطه فوق، r یک عدد تصادفی یکنواخت بین صفر و یک است و D_{max} حداکثر مقدار تغییرات مجاز در موقعیت قورباغه در یک پرش می باشد.

چنانچه این تغییر موقعیت، جوابی با شایستگی بهتر تولید کرد، این جواب جایگزین X_w می گردد. در غیر اینصورت، محاسبات انجام شده با استفاده از روابط (۱۰) و (۱۱) و با جایگزینی بهینه فرامحلی X_g به جای بهینه محلی X_b تکرار می شوند. در صورتی که باز هم بهبودی در جواب حاصل نگردد، X_w حذف شده و یک راه حل جدید بصورت تصادفی جایگزین آن می گردد. این سیر تکاملی برای تعداد گام های تکاملی ممتیک (تعداد تکرارهای جستجوی محلی)، که از قبل مشخص شده است، در هر مپلکس تکرار می شود [۲۰]-[۲۲].

پس از اتمام فرآیند جستجوی محلی در مپلکس ها، تمامی اعضای جمعیت به هم آمیخته شده و بر اساس ارزش شایستگی آنها مجدداً به صورت نزولی مرتب می شوند. سپس دوباره به چند زیر مجموعه تقسیم شده و روند گفته شده تکرار می شود. تکامل جمعیت در مپلکس ها (فرآیند جستجوی محلی) و ترکیب دوباره کل جمعیت تا جایی ادامه پیدا می کند که شرط توقف الگوریتم (اتمام تعداد تکرارها یا رسیدن به یک درصد خطای از پیش تعیین شده) برآورده گردد. در این صورت، الگوریتم SFL خاتمه یافته و جواب با بهترین مقدار شایستگی بعنوان بهترین جواب پیدا شده گزارش می شود. به منظور درک بهتر، فلوجارت الگوریتم SFL در شکل (۵) نشان داده شده است.

۴- الگوریتم یادگیری ترکیبی پیشنهادی

الگوریتم یادگیری ارائه شده برای شبکه موجک فازی دارای ۲ مرحله است، مرحله اول شامل تعیین مقادیر اولیه پارامترهای شبکه و مرحله دوم شامل بروزرسانی پارامترهای شبکه می باشد.

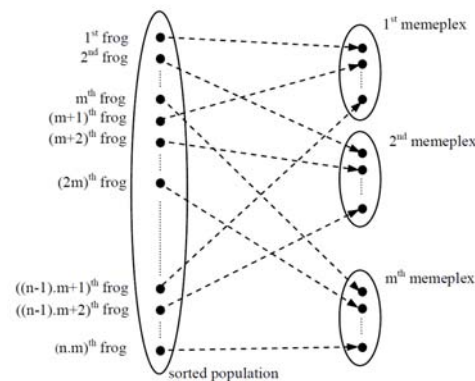
۴-۱- روش مقدار دهی اولیه به پارامترهای شبکه

در این مقاله موجکهای اولیه با تشکیل یک توری موجک بدست می آیند [۲۴]. اما هنگامی که موجکهای اولیه با استفاده از تشکیل توری موجک انتخاب می شوند، دارای افزونگی می باشند و این افزونگی منجر به افزایش تعداد نرونهای موجک می شود. به همین دلیل از الگوریتم حداقل مربعات متعامد برای انتخاب موجکهای مؤثر و پرنفوذ استفاده می شود [۱۹ و ۲۵]. پارامترهای آزاد در شبکه موجک

فازی که بایستی تحت آموزش قرار گیرند شامل P_{jr}^i ، l_j^k و $w_{Mi,t,k}$

S بعدی (S تعداد متغیرها)، موقعیت قورباغه (پاسخ) i ام در فضای جستجو به عنوان یک راه حل قابل قبول در مسئله بهینه سازی در نظر گرفته می شود و آنرا بصورت بردار $X_i = [x_{i1}, x_{i2}, \dots, x_{iS}]^T$ نشان می دهند (X_i موقعیت قورباغه i ام می باشد و $i = 1, \dots, N$).

در مرحله بعدی، با استفاده از تابع برازندگی تعریف شده، هر یک از جواب های مسئله ارزیابی می گردند. تابع برازندگی با توجه به مسئله تعریف می گردد. در ادامه راه حل ها، با توجه به مقادیر شایستگی شان، به صورت نزولی مرتب می گردند. سپس کل جمعیت به m بخش مساوی تقسیم می شود که به هر کدام از این زیربخش ها یک مپلکس گفته می شود. در هر مپلکس n راه حل مسئله قرار می گیرد ($n = N/m$) به گونه ای که اولین راه حل (راه حل با بالاترین مقدار شایستگی) در مپلکس اول قرار می گیرد، دومین راه حل در مپلکس دوم، m امین راه حل در مپلکس m ام و $(m+1)$ امین راه حل مجدداً در مپلکس اول قرار می گیرد؛ این روند تا توزیع تمامی راه حل ها ادامه می یابد. در شکل (۴) روند توزیع قورباغه ها یا راه حل ها در مپلکس ها نشان داده شده است.



شکل ۴: روند شکل گیری مپلکس ها [۲۳]

سپس، در هر کدام از مپلکس ها راه حل های با بدترین و بهترین میزان شایستگی مشخص شده و به ترتیب با X_w و X_b نشان داده می شوند. همچنین راه حل با بهترین مقدار شایستگی در میان کل جمعیت نیز مشخص می گردد (X_g). در ادامه، در طی فرآیند تکامل مپلکس ها، در هر مپلکس یا زیر گروه، موقعیت بدترین راه حل (X_w) به سمت موقعیت بهترین راه حل (X_b) بروز رسانی می شود. موقعیت جدید راه حل بدتر با استفاده از قانون پرش قورباغه بصورت زیر بدست می آید.

$$\text{Position change (D)} = r \times (X_b - X_w) \quad (10)$$

$$X_w(\text{new}) = X_w + D, (\|D\| < D_{max}) \quad (11)$$

$$P_{j2}^i = \sqrt{\frac{2 \sum_{l=1}^L \bar{\mu}_{il} (x_{jl}^d - P_{j1}^i)^2}{\sum_{l=1}^L \bar{\mu}_{il}}} \quad (14)$$

مقادیر اولیه P_{j1}^i و P_{j2}^i بنحوی انتخاب شده اند که به ترتیب میانگین وزنی و انحراف استاندارد داده های ورودی باشند.

۴-۲- روش بروز رسانی پارامترها

پس از ساخت شبکه موجک فازی و تعیین تعداد زیر شبکه های عصبی موجک، تعداد قوانین فازی و تعیین مقادیر اولیه پارامترهای شبکه مطابق با بخش (۴-۱)، بایستی پارامترهای شبکه تحت آموزش قرار گیرند. در این مرحله، هدف از آموزش شبکه، حداقل نمودن میانگین مربعات خطا بعنوان تابع هزینه می باشد، که به صورت زیر تعریف می شود:

$$E = \frac{1}{L} \sum_{l=1}^L (y_l^d - \hat{y}_l)^2 \quad (15)$$

در رابطه (۱۵)، \hat{y}_l و y_l^d به ترتیب خروجی شبکه موجک فازی و خروجی مطلوب می باشند. همچنین L تعداد داده های آموزشی و E میانگین مربعات خطا پس از اعمال تمام ورودیها به شبکه می باشد. در این مقاله به منظور افزایش سرعت همگرایی و رسیدن به پاسخ، از دو الگوریتم یادگیری کارآمد برای آموزش پارامترهای شبکه استفاده شده است: روش حداقل مربعات بازگشتی برای آموزش ضرایب وزن مربوط به هر زیر شبکه عصبی موجک و الگوریتم جهش قورباغه ها برای آموزش پارامترهای غیرخطی شبکه شامل، پارامترهای شیفیت توابع موجک و پارامترهای مربوط به توابع عضویت فازی.

قوانین بروز رسانی پارامترهای خطی با استفاده از الگوریتم حداقل مربعات بازگشتی به صورت روابط (۱۶) و (۱۷) می باشد. جزئیات این الگوریتم در [۲۶] آمده است.

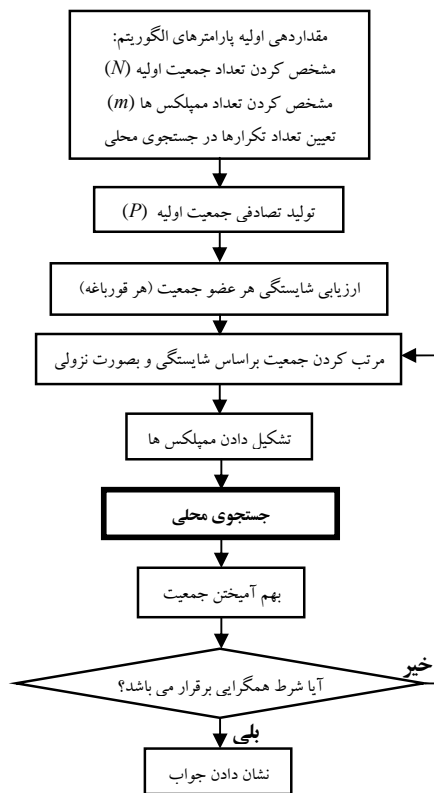
$$P_{n+1} = P_n - \frac{P_n a_{n+1} a_{n+1}^T P_n}{1 + a_{n+1}^T P_n a_{n+1}} \quad (16)$$

$$w_{n+1} = w_n + P_{n+1} a_{n+1} (y_{n+1} - a_{n+1}^T w_n) \quad (17)$$

همانطور که قبلا بیان شد، در آموزش پارامترهای شبکه، هدف حداقل نمودن میانگین مربعات خطا بعنوان تابع هزینه می باشد، که به صورت رابطه (۱۵) تعریف می شود. بنابراین این مسئله را می توان به صورت یک مسئله بهینه سازی با تابع هدف زیر در نظر گرفت:

$$fitness = \frac{1}{\sum_{l=1}^L (y_l^d - \hat{y}_l)^2 + C} \quad (18)$$

می باشند که به ترتیب $r = 1, 2$ ، $j = 1, 2, \dots, n$ ، $i = 1, 2, \dots, c$ و $k = 1, 2, \dots, S$ هستند.



شکل ۵: فلوجارت الگوریتم [۲۲].

مقادیر t_j^k و وزن های اولیه $w_{M_{i,l}^k}$ در مرحله انتخاب موجکها و با استفاده از الگوریتم OLS بدست می آیند اما پارامترهای P_{jr}^i نامشخص هستند و بایستی مقداردهی اولیه شوند.

پارامترهای آزاد توابع عضویت هستند که درجه شراکت هر زیر شبکه عصبی موجک را با دقت خاصی تعیین می کنند. با محاسبه $\hat{y}_l^d(x_{jl}^d)$ در رابطه (۳) درجه شراکت تقریبی $\bar{\mu}_{il}$ بصورت ذیل محاسبه می شود.

$$\bar{\mu}_{il} = \frac{|\hat{y}_l^d(x_{jl}^d)|}{\sum_{i=1}^c |\hat{y}_l^d(x_{jl}^d)|}, \quad i = 1, 2, \dots, c, \quad l = 1, 2, \dots, L \quad (12)$$

بدین ترتیب مقادیر اولیه P_{jr}^i مطابق با روابط زیر مقداردهی اولیه می شوند.

$$P_{j1}^i = \frac{\sum_{l=1}^L \bar{\mu}_{il} x_{jl}^d}{\sum_{l=1}^L \bar{\mu}_{il}} \quad (13)$$

همانطور که مشخص است، تابع به طور قطعه ای پیوسته و قابل تجزیه و تحلیل است. در این مثال، ۲۰۰ داده آموزشی به طور یکنواخت در فاصله [10, -10] انتخاب شده است.

از آنجایی که انتخاب پارامترهای الگوریتم های SFL تأثیر قابل توجهی بر کیفیت جواب بدست آمده دارد، قبل از حل مسئله آزمایش های متعددی انجام گرفته و با توجه به نتایج به دست آمده، مقادیر مناسب برای تعداد مپلکس ها، تعداد تکرارهای محلی و مقدار Dmax به ترتیب برابر با ۱۰، ۲۰ و inf در نظر گرفته شده است. همچنین تعداد اعضای جمعیت ۲۰۰ عضو و نیز تعداد تکرارها که شرط توقف الگوریتم است ۴۰۰۰ تکرار در نظر گرفته شده است.

در [۱۸] از یک روند خود-تنظیم مبتنی بر الگوریتم ژنتیک استفاده شده است. تعداد اعضای جمعیت و تعداد تکرار الگوریتم ژنتیک با الگوریتم SFL یکسان در نظر گرفته شده تا بتوان نتایج را در شرایطی برابر با یکدیگر مقایسه کرد. سایر پارامترهای الگوریتم ژنتیک مانند نرخ جهش و نرخ همبری به ترتیب برابر ۰/۱ و ۰/۹ در نظر گرفته شده اند. شکل (۶) تقریب تابع $f(x)$ با استفاده از شبکه موجک فازی بر اساس الگوریتم پیشنهادی و روشهای اعمال شده را نشان می دهد. همچنین به منظور ارزیابی کارایی روش پیشنهادی در مقایسه با سایر روشهای ارائه شده در مقالات، از شاخص عملکرد شبکه تعریف شده در [۱۸]، استفاده می کنیم.

$$J = \frac{\sum_{i=1}^N (y_i^d - \hat{y}_i)^2}{\sum_{i=1}^N (y_i^d - \bar{y})^2} \quad (18)$$

که \hat{y}_i و y_i^d به ترتیب خروجی شبکه و خروجی مطلوب می باشند و \bar{y} طبق رابطه (۱۹) محاسبه می شود:

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i^d \quad (19)$$

لازم به ذکر است که شاخص عملکرد (۱۸) با استفاده از ۲۰۰ داده تست که بصورت یکنواخت نمونه برداری شده اند، محاسبه شده است. مقایسه میان عملکرد شبکه پیشنهادی با دیگر شبکه های مبتنی بر موجک [۸، ۱۳، ۱۸، ۲۷] در جدول (۱) آورده شده است. نتایج ارائه شده، مؤید قابلیت روش پیشنهادی جهت بهبود شاخص عملکرد شبکه، کاهش تعداد قوانین فازی و همچنین کاهش تعداد پارامترهای شبکه می باشد.

در این مقاله به دلیل سرعت همگرایی زیاد و قابلیت بالای الگوریتم SFL در یافتن پاسخ های زیر بهینه، از این الگوریتم برای حل مسئله بهینه سازی فوق استفاده شده است. از طرفی پارامترهای غیر خطی که بایستی تحت آموزش قرار بگیرند عبارتند از: t^k ، p_{j1}^i و p_{j2}^i . در نتیجه هر بردار راه حل یا هر قورباغه را بصورت زیر تعریف می کنیم:

$$frog = X = [p_{j1}^i \quad p_{j2}^i \quad t^k] \quad (17)$$

در هر بار تکرار الگوریتم، هر بردار راه حل مطابق با رابطه (۱۷) با استفاده از تابع هدف (۱۵) ارزیابی شده و بروز رسانی می شود. باید به این نکته توجه داشت که در الگوریتم پیشنهادی بجای تولید تصادفی جمعیت اولیه برای قورباغه ها، مقادیر اولیه ای که توسط الگوریتم OLS بدست آمده اند، به عنوان جمعیت اولیه الگوریتم SFL بکار می روند. همچنین بازه تغییرات بردار (۱۷) نسبت به این مقادیر اولیه تعیین می شود. از آنجا که الگوریتم RLS دارای قابلیت خوب و سرعت همگرایی بالا در آموزش پارامترهای خطی می باشد، لذا استفاده از آن ضمن کاهش حجم محاسبات و پیچیدگی روش (به سبب کاهش طول بردار جواب رابطه (۱۷) و تسهیل مسئله بهینه سازی)، موجب تسریع سرعت همگرایی الگوریتم پیشنهادی می شود.

۵- نتایج شبیه سازی

در این بخش، به منظور ارزیابی کارایی الگوریتم پیشنهادی برای یادگیری شبکه های موجک فازی، عملکرد این روش برای تقریب تابع، شناسایی و کنترل سیستم غیر خطی مورد بررسی قرار گرفته و نتایج آن با نتایج روش های گزارش شده در مقالات دیگر، مقایسه شده است. کلیه شبیه سازی ها در محیط برنامه نویسی نرم افزار MATLAB و با استفاده از یک کامپیوتر با پردازشگر Core2Duo 2.53MHz و حافظه 4GB انجام شده است.

مثال ۱- تقریب تابع قطعه ای پیوسته

در این بخش عملکرد الگوریتم پیشنهادی برای تقریب تابع غیرخطی و قطعه ای هموار $f(x)$ بررسی شده است. همچنین به منظور انجام یک مقایسه، روشهای ارائه شده در [۱۳] و [۱۸] نیز به مسئله اعمال شده است. تابع $f(x)$ به صورت زیر تعریف می شود:

$$f(x) = \begin{cases} -2.186x - 12.864 & -10 \leq x < -2 \\ 4.246x & -2 \leq x < 0 \\ 10e^{-0.05x-5} \sin[(0.03x+0.7)x] & 0 \leq x \leq 10 \end{cases} \quad (17)$$

مثال ۲- کنترل سیستم های غیر خطی

در این قسمت از دو مثال برای نشان دادن قابلیت روش پیشنهادی در کنترل سیستم های غیر خطی و مقاوم بودن آن در مقابل عدم قطعیت در پارامترهای سیستم و اغتشاش وارد بر سیستم، استفاده شده است.

الف- در این مثال، الگوریتم پیشنهادی به منظور طراحی یک کنترل کننده برای یک سیستم غیر خطی استفاده شده در مقالات متفاوت [۱۸] و [۲۸-۳۰]، بکار برده شده است. همچنین به منظور انجام یک مقایسه، روش مبتنی بر الگوریتم ژنتیک ارائه شده در [۱۸] نیز به مسئله اعمال و نتایج حاصل از آن با روش پیشنهادی مقایسه شده است.

سیستم غیرخطی (۲۰) را در نظر بگیرید که در آن تابع غیرخطی $f(x)$ نامعلوم می باشد. هدف، طراحی کنترل کننده شبکه موجک فازی به منظور پایدار نمودن سیستم (۲۰) در مبدأ می باشد.

$$\dot{x}(t) = \frac{1 - e^{-x(t)}}{1 + e^{-x(t)}} + u(t) \quad (20)$$

همانطور که آشکار است، سیستم فوق در در عدم حضور سیگنال کنترل ناپایدار است. با فرض $u = 0$ و برای $x > 0$ ، $\dot{x} = \frac{1 - e^{-x}}{1 + e^{-x}} > 0$ و به ازای $x < 0$ ، $\dot{x} = \frac{1 - e^{-x}}{1 + e^{-x}} < 0$ می باشد. با استفاده از روش خطی سازی با فیدبک، کنترل کننده مناسب برای پایدار کردن این سیستم در مبدأ، به صورت زیر تعریف می شود [۱۸]:

$$u(t) = -\frac{1 - e^{-x(t)}}{1 + e^{-x(t)}} - ax(t) \quad (21)$$

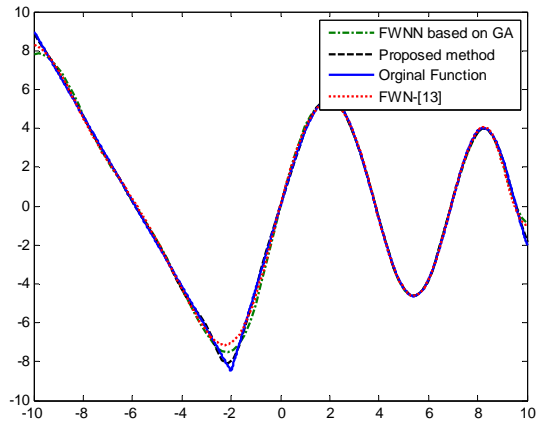
در رابطه فوق، a ثابت وزنی است که سرعت نسبی همگرایی پاسخ سیستم را کنترل می نماید. اگر مدل شبکه موجک فازی برای تابع غیرخطی و نامشخص $f(x)$ ، با $\hat{f}(x)$ نشان داده شود، کنترل کننده طراحی شده، به صورت (۲۲) می باشد:

$$u(t) = -\hat{f}(x) - ax(t) \quad (22)$$

در این مثال جمعیت اولیه و تعداد تکرارها برای هر دو الگوریتم SFL و GA به ترتیب برابر ۱۰۰ و ۲۰۰۰ در نظر گرفته شده است. شکل (۸)، پاسخ سیستم و سیگنال کنترل را به ازای $a = 3$ برای دو روش نشان می دهد. همچنین نمودار همگرایی دو روش در شکل (۹) نشان داده شده است. همانطور که از شکل های (۸) و (۹) پیداست، هر دو روش به خوبی عمل می کنند اما روش پیشنهادی مقاله از نظر دقت رسیدن به جواب و سرعت همگرایی بهتر عمل نموده است.

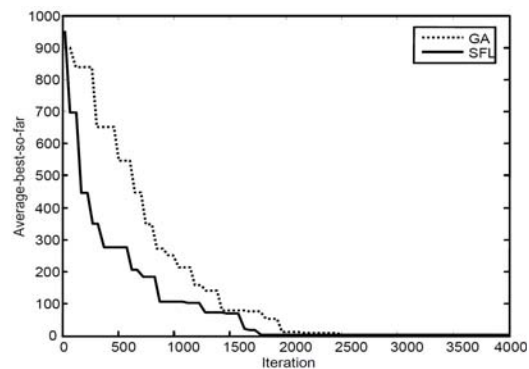
جدول ۱: مقایسه عملکرد الگوریتم پیشنهادی با سایر شبکه های مبتنی بر موجک

روش	J	تعداد قوانین فازی	تعداد پارامترهای شبکه	زمان اجرا
الگوریتم پیشنهادی	۰/۰۰۵۸	۴	۱۶	۷۶۴ s
FWN+GA [۱۸]	۰/۰۲۱۵	۴	۲۰	۹۸۷ s
FWN [۱۳]	۰/۰۳۱۲	۵	۲۴	۷۷۵ s
WARX NN [۲۷]	۰/۰۴۸۰	-	۲۳	-
WNN [۸]	۰/۰۵۰۶	-	۲۲	-



شکل ۶: تقریب تابع $f(x)$

به منظور مقایسه عملکرد و نحوه همگرایی الگوریتم ها، مشخصه همگرایی الگوریتم ها در یافتن مینیمم هزینه، در شکل (۷) رسم شده است. این مشخصه برای میانگین ۵۰ مرتبه اجرای مستقل الگوریتم ها ترسیم شده است. همانطور که از مشخصه های همگرایی پیداست، روش پیشنهادی بدلیل استفاده از الگوریتم های SFL و RLS و نیز تولید جمعیت اولیه مناسب و استفاده از OLS، روند همگرایی بهتر و سریع تری نسبت به روش مبتنی بر GA داشته است.



شکل ۷: مشخصه همگرایی الگوریتم های مبتنی بر SFL و GA

(۲) مشاهده می‌شود، تعداد قوانین فازی با استفاده از روش پیشنهادی کاهش یافته اما در عین حال ضمن بهبود مشخصات پاسخ گذرای سیستم، سرعت همگرایی پاسخ نیز افزایش یافته است. همچنین در مقایسه با [۲۸]، نوسانات پاسخ حذف شده، مشخصات حالت گذرا بهبود یافته و دامنه سیگنال کنترل نیز کاهش یافته است.

جدول ۲: مقایسه عملکرد الگوریتم پیشنهادی با سایر روش های ارائه شده

روش	تعداد قوانین فازی	دامنه سیگنال کنترل	زمان پایداری	زمان اجرا
الگوریتم پیشنهادی	۲	$-2 \leq u \leq 0$	3.3 s (a=1)	۶۵۵ S
الگوریتم پیشنهادی	۲	$-1.5 \leq u \leq 0$	1.7 s (a=3)	۶۴۵ S
FWN+GA[۱۸]	۶	-	3.2 s (a=3)	۸۷۳ S
IAFSMC[۲۸]	۲	$-7 \leq u \leq 5$	2.2 s (a=3)	-
SAFC[۳۰]	۶	-	14 s (a=3)	-

ب - در این مثال، از الگوریتم پیشنهادی جهت طراحی یک کنترل کننده شبکه موجک فازی تطبیقی [۱۹] استفاده شده و اثر تغییر پارامترهای سیستم و نیز اغتشاش وارد به سیستم مورد بررسی قرار گرفته است. هدف کنترل عبارت است از طراحی یک کنترل کننده موجک فازی تطبیقی بطوریکه خروجی سیستم، خروجی مطلوب، $y_d(t)$ ، را دنبال نماید.

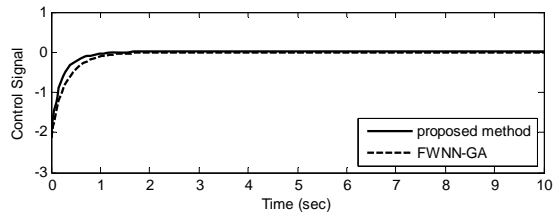
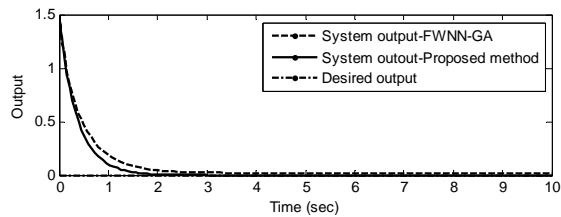
سیستم غیرخطی (۲۳) را در نظر بگیرید، که در آن تابع غیرخطی $f(x)$ نامعلوم می‌باشد.

$$\begin{cases} \dot{x}^{(n)} = f(x) + u \\ y = x \end{cases} \quad (23)$$

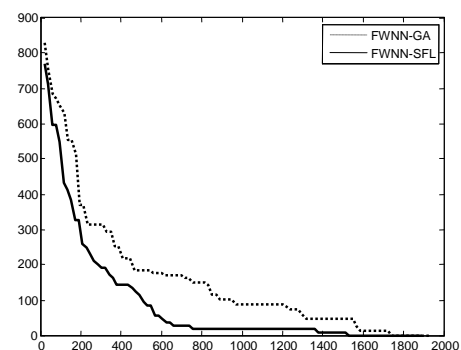
اگر تقریب شبکه موجک فازی برای تابع غیرخطی و نامشخص $f(x)$ ، با $\hat{f}(x|k)$ نشان داده شود، کنترل کننده طراحی شده، به صورت زیر می‌باشد [۳۱].

$$u(t) = -\hat{f}(x|k) + y_d^{(n)} + c^T e \quad (24)$$

که در آن e خطای ردیابی و برابر با $e = y_d - y = y_d - x$ بردار خطا به صورت $e = (e, \dot{e}, \dots, e^{(n-1)})^T$ تعریف می‌شود. همچنین بردار ضرایب $c = (c_1, c_2, \dots, c_n)^T$ به گونه ای انتخاب می شود که ریشه های چندجمله‌ای $s^n + c_1 s^{n-1} + \dots + c_n = 0$ در نیم صفحه باز سمت چپ صفحه مختلط گیرند [۳۱]. ضرایب k در $\hat{f}(x|k)$ ، پارامترهای تطبیقی هستند که به صورت بهنگام تنظیم می‌شوند. اکنون با توجه به مطالب فوق، روش پیشنهادی مقاله را در مدل سازی دینامیک های سیستم و طراحی کنترل کننده برای یک سیستم سرو مکانیزم بکار می‌بریم.

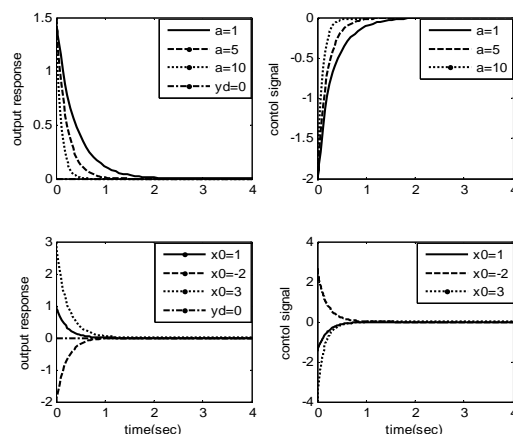


شکل ۸: پاسخ سیستم و سیگنال کنترل به ازای $a = 3$



شکل ۹: مشخصه همگرایی الگوریتم های مبتنی بر SFL و GA

همچنین، شکل (۱۱) پاسخ سیستم و سیگنال کنترل به ازای مقادیر متفاوت a و شرایط اولیه متفاوت را نشان می‌دهد. همانطور که از شکل (۱۱) مشاهده می‌شود، کنترل کننده پیشنهادی قادر است به ازای شرایط اولیه متفاوت با دامنه سیگنال کنترل قابل قبول، سیستم را کنترل نماید.



شکل ۹: پاسخ سیستم و سیگنال کنترل به ازای مقادیر متفاوت a و x_0

مقایسه نتایج حاصل از روش پیشنهادی و کارهای قبلی انجام شده در سایر مقالات، در جدول (۲)، ارائه شده است. همانطور که از جدول

سیستم سرومکانیزم غیرخطی که توسط معادله دیفرانسیل مرتبه دوم (۲۵) مدل می شود را در نظر بگیرید.

$$m\ddot{q} + l\dot{q} + \Delta f(q) = \tau + d \quad (25)$$

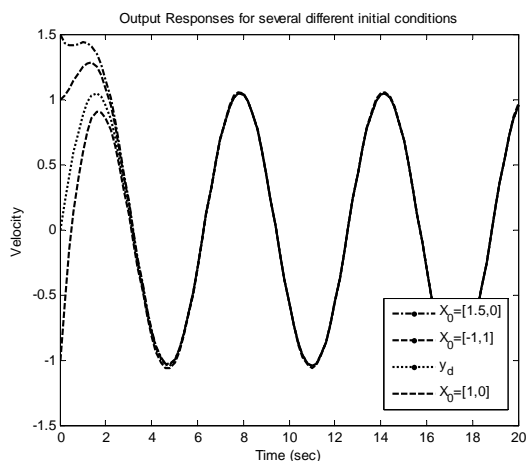
پارامترهای مدل که بیانگر مشخصات فیزیکی سیستم می باشد به صورت زیر توصیف می شوند. \dot{q} : سرعت، q : موقعیت، $\Delta f(q)$: ترم غیرخطی وابسته به q ، m و l : جرم و میرایی سیستم، τ : گشتاور و d ، اغتشاش را نمایش می دهند. با فرض $x_1 = q$ و $x_2 = \dot{q}$ ، موقعیت و سرعت سیستم، به عنوان متغیر حالت برای سیستم در نظر گرفته شده و فرض می شود که قابل اندازه گیری باشند. بنابراین معادلات دینامیکی سرومکانیزم در فضای حالت، می تواند به صورت (۲۶) توصیف شوند.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= f(x) + u + d(t) \\ y &= x \end{aligned} \quad (26)$$

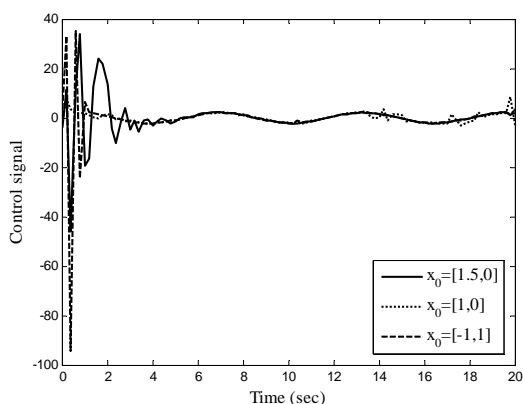
که در رابطه فوق داریم: $f(x) = -lx_2 - \Delta f(x_1)$ و $u = \tau$ ، $\Delta f(x_1) = 0.4 \sin(x_1)$ هدف کنترل در این مثال، ردیابی مسیر مطلوب $y_d(t) = (\pi/3) \sin(t)$ توسط خروجی سیستم می باشد. پارامترهای سیستم به صورت $m = 1 \text{ kg}$ و $l = 1$ و $d(t)$ به عنوان اغتشاش خارجی وارد بر سیستم در نظر گرفته شده است. در این مثال جمعیت اولیه و تعداد تکرارها برای هر دو الگوریتم SFL و GA به ترتیب برابر ۲۰۰ و ۱۰۰۰ در نظر گرفته شده است.

شکلهای (۱۰) و (۱۱) به ترتیب، پاسخ خروجی سیستم در ردیابی مسیر مطلوب $y_d(t)$ و سیگنال کنترل مربوط به آن را در عدم حضور اغتشاش و تغییرات زمانی در پارامترهای سیستم به ازای چندین شرایط اولیه متفاوت $x_0 = [1.5, 0]$ ، $x_0 = [-1, 1]$ ، $x_0 = [1, 0]$ نشان می دهند. همانطور که از نتایج شبیه سازی مشاهده می شود، به ازای شرایط اولیه متفاوت، خروجی سیستم قادر به ردیابی مسیر مطلوب می باشد و سریعاً به سمت مسیر مطلوب همگرا می شود. از طرفی دامنه سیگنال کنترل نیز در هر حالت قابل قبول و مناسب می باشد.

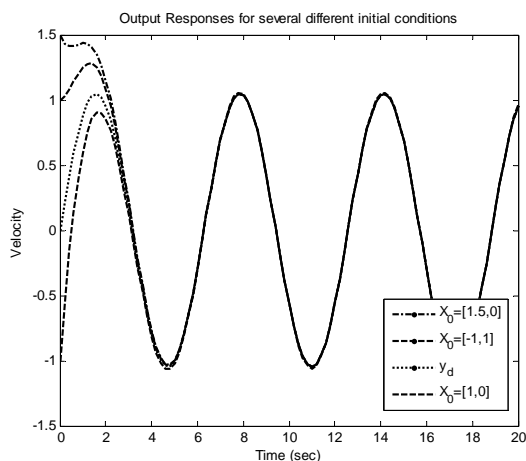
حال اگر اغتشاش خارجی وارد بر سیستم به صورت $d(t) = \cos(2\pi t)$ و فاکتور میرایی سیستم به صورت $l = 1 + 0.3 \sin(t)$ تغییر کند، پاسخ سیستم در ردیابی مسیر مطلوب در حضور اغتشاش و نامعینی و سیگنال کنترل مربوط به آن به ازای چندین شرایط اولیه متفاوت بصورت نشان داده شده در شکل های (۱۲) و (۱۳) بدست می آید.



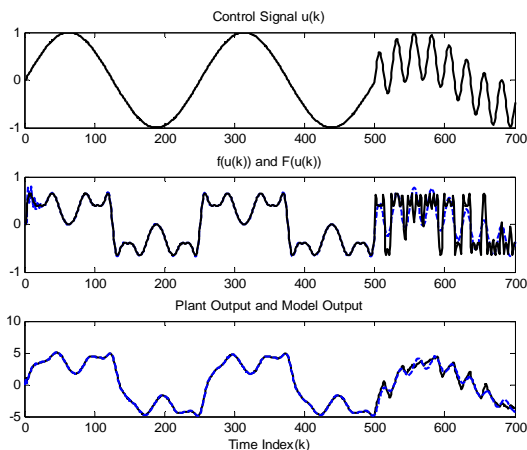
شکل ۱۰: پاسخ خروجی در عدم حضور اغتشاش و نامعینی.



شکل ۱۱: سیگنال کنترل در عدم حضور اغتشاش و نامعینی.

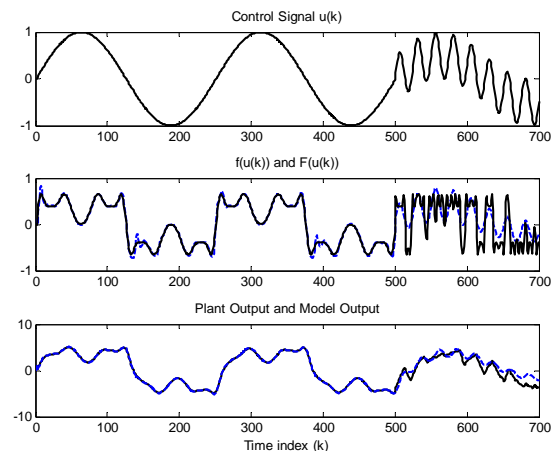


شکل ۱۲: پاسخ خروجی به در حضور اغتشاش و نامعینی.



شکل ۱۴: الف) سیگنال کنترل، ب) $F(u(k))$ به صورت خط چین و $f(u(k))$ با خط پر، ج) خروجی مدل به صورت خط چین و خروجی پلنت با خط پر.

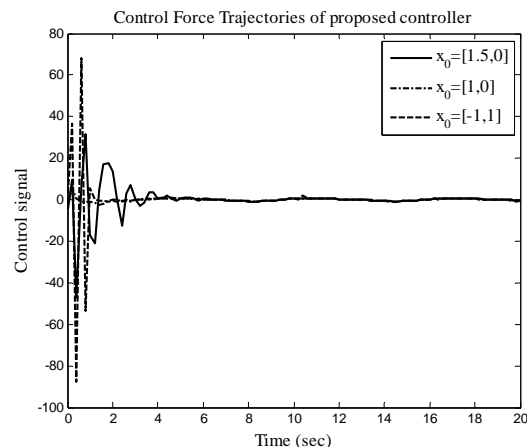
به منظور مقایسه روش ارائه شده با روش های قبلی اعمال شده، مثال مذکور با استفاده از روش مبتنی بر الگوریتم ژنتیک معرفی شده در [۱۸] نیز پیاده سازی شده است. شکل (۱۵) خروجی سیستم و سیگنال کنترلی را برای سیستم فوق نشان می دهد.



شکل ۱۵: الف) سیگنال کنترل، ب) $F(u(k))$ به صورت خط چین و $f(u(k))$ با خط پر، ج) خروجی مدل به صورت خط چین و خروجی پلنت با خط پر.

همانطور که از شکل های (۱۴) و (۱۵) مشاهده می شود، صحت و دقت روش پیشنهادی مقاله در مقایسه با روش مبتنی بر الگوریتم ژنتیک بیشتر است.

به منظور ارزیابی بیشتر، نتایج حاصل از روش های اعمال شده و نیز نتایج حاصل از شبکه انفیس و شبکه عصبی چند لایه گزارش شده در [۲۶] و [۳۲] در جدول (۳) آورده شده است. همانطور که از نتایج



شکل ۱۳: سیگنال کنترل در حضور اغتشاش و نامعینی.

همانطور که مشاهده می گردد، کنترل کننده در مقابل تغییرات زمانی در پارامترهای سیستم و اغتشاش خارجی وارد بر سیستم مقاوم بوده و پارامترهای تطبیقی در حضور اغتشاش و نامعینی به نحوی تغییر می یابند که خروجی سیستم با مشخصات گذرای قابل قبول سریعاً به سمت مسیر مطلوب همگرا شود.

مثال ۳- شناسایی بهنگام در سیستم های کنترل

در این مثال، شبکه موجک فازی برای شناسایی بخش غیرخطی در یک سیستم کنترلی به کار رفته است. پلنت مورد بررسی توسط معادله تفاضلی زیر توصیف می شود:

$$y(k+1) = .3y(k) + .6y(k-1) + f(u(k)) \quad (27)$$

به منظور شناسایی پلنت، یک مدل سری- موازی با معادله تفاضلی (۲۸) بکار می رود.

$$\hat{y}(k+1) = .3\hat{y}(k) + .6\hat{y}(k-1) + F(u(k)) \quad (28)$$

در رابطه (۲۸)، $F(u(k))$ تابعی است که با یک شبکه موجک فازی شامل ۴ قانون فازی (۴ تابع عضویت) جایگزین شده و پارامترهای آن در هر گام زمانی به صورت بهنگام پروژراسانی می شوند. در مرحله آموزش ورودی پلنت و مدل، یک سیگنال سینوسی به صورت $u(k) = \sin(2\pi k / 250)$ در بازه $k = 1$ تا $k = 250$ و یک سیگنال بصورت $u(k) = .5 \sin(2\pi k / 250) + .5 \sin(2\pi k / 25)$ در بازه $k = 250$ تا $k = 500$ می باشد.

خروجی سیستم و سیگنال کنترلی در شکل (۱۴) نشان داده شده اند همانطور که شکل (۱۴) نشان می دهد، خروجی مدل قادر است خروجی پلنت را بلافاصله بعد از توقف آموزش شبکه در $k = 250$ و حتی بعد از تغییر $u(k)$ به صورت در لحظه $k = 500$ دنبال نماید.

- Approximation", *IEEE Trans, Systems, Man, and Cybernetics, Part B*.
- [6] Narendra, K.S., Parthasarathy, K., 1990, "Identification and Control of Dynamical Systems Using Neural Networks", *IEEE Trans, Neural Networks*.
- [7] Rios-Patron, E., Braatz, R.D., 1997, "Identification and Control of Dynamical Systems Using Neural Networks", *IEEE Trans, Neural Networks*.
- [8] Zhang, Q., Benveniste, A., 1992, "Wavelet Networks," *IEEE Trans, Neural Networks*.
- [9] Faa-Jeng, L., Shen, P.H., Kung, Y.S., 2005, "Adaptive Wavelet Neural Network Control for Linear Synchronous Motor Servo Drive", *IEEE Trans, Magnetics*.
- [10] Jun, Z., Zhang, J., Gilbert G., Walter, Y. M., Lee W.N.W., 1995, "Wavelet Neural Networks for Function Learning", *IEEE Trans, Signal Processing*.
- [11] Shi, D., Chen, F., Ng, G.S., Gao, J., 2006, "The Construction of Wavelet Network for Speech Signal Processing", *Neural Computing & Applications*.
- [12] Galvao, R. K. H., Yoneyama, T., 2004, "A Competitive Wavelet Network for Signal Clustering", *IEEE Trans, Systems, Man, and Cybernetics, Part B*.
- [13] Ho, D.W.C., Zhang, P.A., Xu, J., 2001, "Fuzzy Wavelet Networks for Function Learning", *IEEE Trans, Fuzzy Systems*.
- [14] Zekri, M., Sadri, S., Sheikholeslam, F., 2008, "Adaptive Fuzzy Wavelet Network Control Design for Nonlinear Systems", *Fuzzy Sets and Systems*.
- [15] Karatepe, E., Alci, M., 2005, "A New Approach to Fuzzy Wavelet System Modeling", *International Journal of Approximate Reasoning*.
- [16] Yilmaz, S., Oysal, Y., 2020, "Fuzzy Wavelet Neural Network Models for Prediction and Identification of Dynamical Systems", *IEEE Trans, Neural Networks*.
- [17] Srivastava, S., Singh, M., Hanmandlu, M., Jha, A. N., 2005, "New Fuzzy Wavelet Neural Networks for System Identification and Control", *Appl. Soft Comput.*
- [18] Tzeng, S.T., 2010, "Design of Fuzzy Wavelet Neural Networks Using the GA Approach for Function Approximation and System Identification", *Fuzzy Sets System*.
- [۱۹] ذکری م، طراحی شبکه موجک فازی تطبیقی برای کنترل سیستم های غیرخطی، رساله دکتری مهندسی برق کنترل، دانشگاه صنعتی اصفهان، اصفهان، ۱۳۸۶.

ارائه شده در جدول مشاهده می شود، تعداد پارامترهای شبکه با استفاده از الگوریتم پیشنهادی به طور قابل ملاحظه ای کاهش یافته است.

جدول ۳: مقایسه عملکرد الگوریتم پیشنهادی با سایر روش های ارائه شده

روش	تعداد پارامترها	تعداد توابع عضویت	زمان اجرا
الگوریتم پیشنهادی	۱۶	۴	۱۶۵ s
FWN+GA [۱۸]	۲۰	۴	۲۸۳ s
MLP [۳۲]	۴۰	-	-
ANFIS [۲۶]	۳۵	۷	-

۶- نتیجه گیری

در این مقاله، یک الگوریتم ترکیبی برای شبکه های موجک فازی به منظور حل مسائل تقریب توابع، شناسایی بهنگام و کنترل تطبیقی سیستم های غیرخطی ارائه شد. الگوریتم ترکیبی پیشنهادی مبتنی بر الگوریتم حداقل مربعات متعامد، الگوریتم جهش قورباغه ها و روش حداقل مربعات بازگشتی می باشد. این الگوریتم ضمن انتخاب موجک های مهم و مؤثر و تعیین تعداد قوانین فازی و تعداد زیر شبکه های عصبی موجک برای ساخت شبکه موجک فازی، مقادیر اولیه پارامترهای شبکه را تعیین کرده و آنها را تحت آموزش قرار می دهد. الگوریتم پیشنهادی برای تقریب تابع، شناسایی و کنترل سیستم غیرخطی بکار رفته و نتایج آن با سایر کارهای گزارش شده مقایسه گردید. نتایج بدست آمده، نشان می دهد که سرعت همگرایی با استفاده از شبکه پیشنهادی افزایش یافته و ضمن بهبود شاخص عملکرد شبکه در مقایسه با سایر روشها، تعداد قوانین فازی و تعداد پارامترهای شبکه و نیز حجم محاسبات و پیچیدگی روش کاهش یافته است.

مراجع

- [1] Chun-Fei, H., Chih-Min, L., Tsu-Tian, L., 2006, "Wavelet Adaptive Backstepping Control for a Class of Nonlinear Systems", *IEEE Transactions on Neural Networks*.
- [2] Jian-Xin, X., Ying, T., 2007, "Nonlinear Adaptive Wavelet Control Using Constructive Wavelet Networks", *IEEE Trans, Neural Networks*.
- [3] Sadabadi, M.S., Shafiee, M., Karrari, M., 2008, "System Identification of Two-dimensional Continuous-time Systems Using Wavelets as Modulating Functions", *ISA Transactions*.
- [4] Billings, S.A., Hua-Liang, W., 2005, "A New Class of Wavelet Networks for Nonlinear System Identification", *IEEE Trans, Neural Networks*.
- [5] Shioh-Shung, Y., Ching-Shioh, T., 1996, "An Orthogonal Neural Network for Function

- [27] Chen, J., Bruns, D.D., 1995, "WaveARX Neural Network Development for System Identification Using a Systematic Design Synthesis", *Industrial & Engineering Chemistry Research*.
- [28] Wang, J., Rad, A.B., Chan, P.T., 2001, "Indirect Adaptive Fuzzy Sliding Mode Control: Part I: fuzzy Switching", *Fuzzy Sets and Systems*.
- [29] Shahriari-Kahkeshi, M., Zekri, M., Sheikholeslam, F., 2010, " Application of Fuzzy Wavelet Neural Network to Adaptive Control of Nonlinear Systems", *4th International Conference on Intelligent Information Technology Application*.
- [30] Wang, L. X., 1993, "Stable Adaptive Fuzzy Control of Nonlinear Systems", *IEEE Trans, Fuzzy Systems*.
- [۳۱] شهریاری کاهکشی م.، طراحی شبکه موجک فازی تطبیقی به روش غیر مستقیم برای کنترل سیستم های غیرخطی، پایان نامه کارشناسی ارشد مهندسی برق کنترل، دانشگاه صنعتی اصفهان، اصفهان، ۱۳۸۹.
- [32] Narendra, K.S., Parthasarathy, K., 1990, "Identification and Control of Daynactical Systems using Neural Networks ", *IEEE Trans, Neural Networkem*.
- [20] Li, Q., 2009, "Shuffled Frog Leaping Algori thm Based Optimal Reactive Power Flow", *International symposium on computer network and multimedia technology*.
- [21] Eusuff, M.M., Lansey, K., Pasha, F., 2003, "Shuffled Frog Leaping Algorithm: a memetic meta heuristic for discrete optimization", *Engineering Optimization*.
- [22] Bijami, E., Askari, J., Farsangi, M.M., 2010, "Power System Stabilizer Design by Using Shuffled Frog Leaping", *The 6th international conference on technical and physical problems of power engineering*.
- [23] Huynh, T.H., 2008, "A Modified Shuffled Frog Leaping Algorithm for Optimal Tuning of Multivariable PID controllers", *IEEE Internatioanal Conference on Industrial Technology*.
- [24] Qinghua, Z., 1997, "Using Wavelet Network in Nonparametric Estimation," *IEEE Trans, Neural Networks*.
- [25] Chen, S., Cowan, C.F.N, Grant, P.M., 1991, "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks", *IEEE Trans, Neural Networks*.
- [26] Jyh, C.T.S., Jang, R., Mizutani, E., 1997, "*Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*", Prentice-Hall, Englewood Cliffs.

کاربرد مفهوم ظرفیت جذب انرژی در تحلیل پایداری سیستم های چند عاملی

کریم رحمانی^۱، احمد افشار^۲، علی اکبر جمشیدی فر^۳

^۱ استادیار، پژوهشکده برق و فن آوری اطلاعات، سازمان پژوهش های علمی و صنعتی ایران، k_rahmani2003@yahoo.com

^۲ دانشیار، دانشکده مهندسی برق، دانشگاه صنعتی امیرکبیر، aafshar@aut.ac.ir

^۳ استادیار، پژوهشکده برق و فن آوری اطلاعات، سازمان پژوهش های علمی و صنعتی ایران، gamshidifar@irost.org

(تاریخ دریافت مقاله ۱۳۹۰/۱/۱۷، تاریخ پذیرش مقاله ۱۳۹۰/۲/۳)

چکیده: امروزه تحلیل پایداری^۱ یک موضوع حیاتی در سیستم های واقعی از جمله سیستم های چند عاملی^۲ می باشد. در این مقاله مفهوم جدید ظرفیت جذب انرژی^۳ برای تحلیل پایداری سیستم های چند عاملی توسعه داده شده است. ظرفیت جذب انرژی در هر نقطه تعادل یک سیستم، ماکزیمم انرژی جذب شده در سیستم به شرط پایدار بودن نقطه تعادل تعریف شده است. در این مقاله عامل^۴ یک موجود هایبرید^۵ شامل تعدادی مود در نظر گرفته شده است که دارای قابلیت سیر تکامل پیوسته درون هر مود و همچنین انتقال بین مودها می باشد. سیستم چند عاملی نیز متشکل از چندین عامل بوده که هر یک فعالیت خود را از یک شرایط اولیه خاص آغاز نموده و با همکاری دیگر عامل ها به دنبال دستیابی به یک هدف مشترک از پیش تعریف شده می باشند. در این مقاله، قابلیت مفهوم ظرفیت جذب انرژی در تحلیل پایداری عامل و سیستم های چند عاملی نشان داده شده است. در خاتمه، کار آرایشی روش ارائه شده با انجام پیاده سازی روش پیشنهادی بر روی دو پروسه مختلف مورد بررسی قرار گرفته است.

کلمات کلیدی: پایداری، ظرفیت جذب انرژی، عامل، سیستم های چند عاملی، موجود هایبرید.

Application of Energy Absorption Capacity Concept for Stability Analysis of Multi-Agent Systems

Karim Rahmani, Ahmad Afshar, Ali Akbar Jamshidifar

Abstract: The issue of stability analysis of multi-agent systems (MASs) is focal point in this paper. In this paper, new concept of “Energy Absorption Capacity (EAC)” is extended to analyze the stability of MAS. EAC is defined for every equilibrium point of a system and is the maximum absorbed energy by the system preserving the stability of that point. An agent is considered a hybrid entity compromising a number of modes with continuous evolution within the modes and transition between them. A MAS is a group of agents that start from an initial state and cooperate with each other to achieve a pre-determined goal. In this work, it is shown that the EAC concept has good potential for stability analysis of both agent and multi-agent (MAS) systems. Finally, two examples are presented to support the proposed approach.

Keywords: Stability, Energy Absorption Capacity (EAC), Absolute EAC, Instantaneous EAC, Hybrid Agent, Multi-Agent Systems.

¹ Stability

² Multi-Agent System

³ Energy Absorption Capacity

⁴ Agent

⁵ Hybrid Entity

۱- مقدمه

تکنولوژی های مبتنی بر سیستم عامل در زمینه های کاربردی در مراحل مختلف خطوط تولید در لایه های محلی و جهانی مخصوصا در توسعه الگوریتم های کنترلی توزیع شده خط تولید کاربرد داشته است [۱] و [۲]. همچنین در زمینه سیستم های چند رباتی همکار، سیستم های حمل و نقل شهری و وسایل نقلیه هوشمند کاربرد سیستم های چند عاملی ملاحظه شده است [۳]. در سیستم چند عاملی، دانشمندان کامپیوتر به خاصیت های عملی سیستم از قبیل هماهنگی، منطق و استدلال و یادگیری علاقه مند بوده ولیکن برای مهندسين کنترل که بیشتر علاقه به پیاده سازی این سیستم ها دارند، پایداری سیستم جایگاه ویژه ای دارد.

پایداری یک خاصیت مهم در سیستم می باشد و در صورت عدم تحقق آن، سیستم ناپایدار و بدون استفاده می گردد. در مطالعه و بررسی سیستم های دینامیکی، کلاسه های مختلفی از مسائل پایداری مطرح می باشند. در این راستا روش های گوناگونی جهت مقابله با این مسئله مطرح شده اند. از جمله روش های شناخته شده، روش لیپانوف^۶ در مطالعه پایداری نقاط تعادل می باشد. روش دیگر مطرح شده، روش ورودی-خروجی^۷ بوده که در آن تفصیل پایداری بر آن اساس می باشد که اعمال ورودی محدود شده به سیستم منجر به خروجی محدود شده در سیستم گردد [۴] و [۵]. روش دیگر تحلیل پایداری سیستم های فیزیکی، قضیه انفعال^۸ می باشد که بر پایه ورودی و خروجی سیستم و انرژی آن استوار است [۶]. در زمینه تحلیل پایداری سیستم های چند عاملی، تحقیقات متعددی در مقالات گزارش شده است [۷] و [۸] و [۹] و [۱۰] و [۱۱]. از جمله این تحقیقات، در نظر گرفتن سیستم های چند عاملی در قالب حالت های مجزا و بکارگیری روش های آماری در تحلیل پایداری آن ها می باشد. تمرکز محققین دیگر در طبقه بندی دیگری از سیستم های چند عاملی از قبیل گروه و یا به عبارتی حرکت دسته جمعی می باشد، که مدل سازی دینامیکی عامل ها انجام شده و با استفاده از روش های موجود از قبیل لیپانوف، تحلیل پایداری سیستم را انجام داده اند.

در این مقاله، مفهوم ظرفیت جذب انرژی جهت بنا نهادن شرایط پایداری یک عامل (موجود های پدید) و پایداری سیستم های چند عاملی بسط داده شده است. روش پیشنهادی بطور خاص مناسب سیستم هایی

می باشد که در مدلسازی آن ها از روش های شی گرا^۹ و روش مدلسازی بلوک تابع^{۱۰} [۱۲] و [۱۳] استفاده شده است. مفهوم جدید پیشنهادی همخوانی لازم را با تحقیقات گذشته محققین در توسعه روش های مدلسازی و کنترل سیستم های توزیع شده [۱۴] و [۱۵] و همچنین تحلیل پایداری آن ها دارد [۱۶] و [۱۷]. نقش انرژی در مدلسازی فیزیکی سیستم ها در گذشته بنا نهاده شده است و چندین قالب کاری شناخته شده در مدلسازی توسعه داده شده است [۱۸] و [۱۹]. بر اساس بررسی های انجام شده، تنها تحقیقات مربوطه در زمینه تحلیل پایداری سیستم های چند عاملی، در زمینه رباتیک می باشد که در آن حداقل انرژی پتانسیل مورد نیاز جهت واژگون کردن ربات در اطراف تکیه گاه چند ضلعی آن^{۱۱} حاشیه انرژی پایداری^{۱۲} تعریف شده است. این حاشیه انرژی مناسب مطالعات پایداری سیستم های استاتیک بوده و توسعه حاشیه انرژی پایداری که منجر به تعریف حاشیه پایداری دینامیکی^{۱۲} شده است جهت معرفی راهکاری مناسب برای مطالعه پایداری سیستم های دینامیکی (حرکت ربات) می باشد [۲۰] و [۲۱].

بسط مفهوم ظرفیت جذب انرژی پیشنهادی در این مقاله در راستای معیار های توسعه یافته جهت تحلیل پایداری سیستم های چند عاملی بوده و در بر گیرنده حالت های پیوسته درون هر مود و سوئیچینگ بین مود ها می باشد. مقاله پیشنهادی متشکل از بخش های مختلف می باشد. در بخش ۲، مفهوم ظرفیت جذب انرژی در توسعه معیار پایداری سیستم های دینامیکی ارائه شده و در بخش ۳ بسط آن برای تحلیل پایداری سیستم های تک عامل و چند عاملی انجام شده است. در بخش ۴ مقایسه بین روش لیپانوف و روش جذب انرژی آورده شده است. نهایتا در بخش ۵ دو مثال کاربردی جهت نشان دادن عملکرد روش پیشنهادی و نتایج پیاده سازی ارائه شده و در خاتمه (بخش ۶) نتیجه گیری و پیشنهادات در ارتباط با موضوع تحقیق در آینده مطرح شده است.

۲- کاربرد ظرفیت جذب انرژی در توسعه معیارهای پایداری برای حالت های پیوسته

از نقطه نظر لیپانوف، نقطه تعادل یک سیستم پایدار است اگر تمام حالات شروع شده از نقاط مجاور آن نقطه تعادل در همسایگی آن

⁹ Object-Oriented Methodology

¹⁰ Function Block Modeling

¹¹ Edge Stability Margin

¹² Dynamic Stability Margin

⁶ Lyapunov

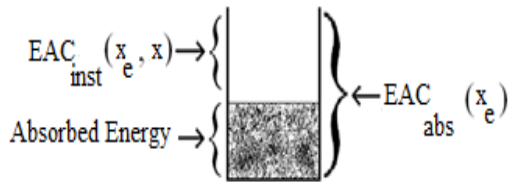
⁷ Input-Output

⁸ Passivity Theorem

تعریف شده است. گام بعدی توسعه معیار پایداری با شرایط مرزی دیگر در یک سیستم می باشد. جهت بسط مفهوم ظرفیت جذب انرژی به شرایط فوق عبارت جدید ظرفیت جذب انرژی مطلق تعریف شده است. **تعریف ۴:** فرض کنید سیستم (۱) در شرایط اولیه x_0 باشد و یا به عبارت دیگر انرژی اولیه $E(x_0)$ را جذب کرده باشد. ظرفیت جذب انرژی لحظه ای آن در حالت اولیه x_0 به صورت اختلاف ظرفیت جذب انرژی مطلق سیستم و انرژی جذب شده آن تعریف می شود.

$$EAC_{inst}(x_e, x_0) = EAC_{abs}(x_e) - E(x_0) \quad (۳)$$

بعنوان یک مثال، ظرفیت جذب انرژی لحظه ای یک فنر خطی برابر $\frac{1}{2}kx^2$ $EAC - \frac{1}{2}x^2$ می باشد که در آن عبارت دوم نشان دهنده انرژی پتانسیل ذخیره شده در تغییر طول فنر می باشد. جهت درک بهتر مفاهیم ظرفیت جذب انرژی لحظه ای و مطلق یک سیستم و ارتباط بین آن ها، شکل شماتیکی (۱) پیشنهاد شده است.



شکل ۱: نمایش شماتیک ظرفیت جذب انرژی مطلق و آنی

در ادامه دو قضیه زیر در خصوص ارتباط بین ظرفیت جذب انرژی و پایداری سیستم بیان می شود.

قضیه ۱:

فرض کنید $x_e = 0$ یک نقطه تعادل سیستم (۱) باشد. x_e پایدار است اگر و فقط اگر ظرفیت جذب انرژی مطلق سیستم در آن نقطه تعادل بزرگتر از صفر باشد. بعبارت دیگر

$$EAC_{abs}(x_e) > 0$$

اثبات قضیه فوق در مرجع [۱۷] ارائه شده است.

قضیه ۲:

نقطه تعادل $x_e = 0$ در سیستم (۱) پایدار مجانبی است اگر و فقط اگر پایدار بوده و

باقی بمانند و در غیر اینصورت سیستم ناپایدار می باشد [۴]. همچنین سیستم در آن نقطه تعادل پایدار مجانبی می باشد به شرط آنکه تمام حالات شروع شده از نقاط مجاور آن نقطه تعادل در مجاورت آن باقی مانده و با افزایش زمان به سمت آن نقطه تعادل میل کنند. تعریف دیگر از پایداری بر اساس قضیه انفعال ارائه شده است که در آن یک رابطه بین پایداری و انرژی سیستم بیان شده است [۶].

در این بخش بر اساس معیار های لیاپانوف و انفعال و مفهوم جدیدی بنام ظرفیت جذب انرژی، معیار مناسبی جهت تحلیل پایداری نقاط تعادل سیستم ارائه شده است. بر اساس مفهوم ظرفیت جذب انرژی، هر نقطه تعادل یک سیستم پایدار است اگر ظرفیت جذب انرژی آن مثبت بوده و ناپایدار است اگر ظرفیت جذب انرژی آن منفی باشد. همچنین پایدار مجانبی است اگر علاوه بر پایدار بودن، ظرفیت جذب انرژی لحظه ای با افزایش زمان به ظرفیت جذب انرژی مطلق آن میل کند. در ادامه تعاریف زیر برای توسعه معیار ها بیان شده اند. شرح و تفصیل این تعاریف در مرجع [۱۶] آورده شده است.

تعریف ۱: ظرفیت جذب انرژی مطلق سیستم^{۱۳}، حداکثر میزان انرژی قابل جذب توسط سیستم است بگونه ای که به ازای بیش از آن سیستم به وضعیت غیر قابل برگشت برود.

تعریف ۲: برای یک المان ذخیره کننده انرژی، ظرفیت جذب انرژی مطلق در نقطه تعادل آن بصورت زیر تعریف می شود.

$$EAC_{abs}(x) = E(x) \Big|_{x_{EAC}}$$

که در آن x متغیری است که المان در آن راستا می تواند انرژی ذخیره نماید، $E(x)$ انرژی ذخیره شده در المان و در راستای x و x_{EAC} نشان دهنده حداکثر مقدار x است که المان به ازای بیش از آن به وضعیت غیر قابل برگشت می رود.

تعریف ۳: سیستم دینامیکی زیر با نقطه تعادل x_e را در نظر بگیرید.

$$\dot{x} = f(x), f(x_e) = 0 \quad (۱)$$

ظرفیت جذب انرژی مطلق به صورت زیر تعریف شده است

$$EAC_{abs}(x_e) = \min_i E(x_i) \Big|_{x_i \in EAC} \quad (۲)$$

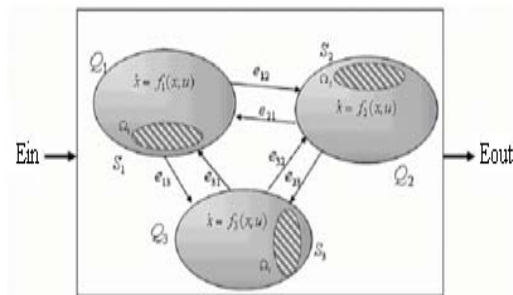
که در آن $E(x_i)$ انرژی ذخیره شده در x_i است. تا کنون ظرفیت جذب انرژی مطلق سیستم در ارتباط با نقاط تعادل

¹³ Absolute Energy Absorption Capacity

تفاوت درون هر مود یک عامل نشان داده شده و حالت های مستقل همان مود های یک سیستم می باشند. عامل با متغیر $A(Q, X)$ نشان داده شده که در آن Q نشان دهنده حالت های مستقل (یا مود ها) و X نشان دهنده حالت های پیوسته می باشد. درون هر مود، عامل دارای سیر تکاملی پیوسته می باشد. شکل (۲)، ساختار داخلی یک عامل هایبرید را به تصویر می کشاند. عامل فوق متشکل از سه مود مستقل Q_1, Q_2, Q_3 می باشد. دینامیک داخلی هر عامل درون هر مود توسط معادله دیفرانسیلی زیرین تشریح شده است.

$$\dot{x} = f_i(x, u)$$

حالت پیوسته X متعلق به فضای حالت S_i می باشد. مناطق سایه دار Ω_i درون هر فضای حالت نشان دهنده مناطق نامن می باشند و حوادث عامل ها e_{ij} که باعث انتقال بین مود ها Q_i می گردند در اثر دریافت ورودی های خارجی E_{in} و یا توسط سیر تکاملی حالت ها در راستای دستیابی به مقادیر از پیش تعریف شده می باشد. بعضی از این حوادث و یا توالی خاص خود، مولد خروجی های خارجی عامل می شوند (E_{out}).



شکل ۲: ساختار داخلی موجود هایبرید

روشی که عامل در نحوه برخورد با حوادث و تغییر مودها اتخاذ می نماید مربوط به خواص استقلال و انعطاف پذیری آن می باشد. استقلال عامل ها بدلیل داشتن نوعی کنترل در پذیرفتن و یا رد کردن حوادث دریافت شده می باشد که منجر به تغییر و یا عدم تغییر مود می گردد. انعطاف پذیری یک عامل مربوط به چگونگی نحوه برخورد با حوادث ورودی و زمان و روش تولید شدن حوادث خروجی می باشد. این موضوع می تواند در ارتباط با دیگر خواص عامل ها از قبیل هوش و یادگیری بسط پیدا کند.

ب - پایداری یک عامل

برای بحث در مورد پایداری یک سیستم عامل، یک عامل متشکل از ۳ مود طبق شکل (۲) در نظر گرفته شده است. در واقع عامل هایبرید

$$EAC(x, x) \rightarrow EAC(x) \text{ as } t \rightarrow \infty$$

inst eq abs eq

اثبات قضیه فوق در مرجع [۱۷] ارائه شده است.

۳- بسط ظرفیت جذب انرژی به تحلیل

پایداری عامل و سیستم های چند عاملی

در این بخش، مفهوم ظرفیت جذب انرژی جهت تحلیل پایداری

سیستم های چند عاملی توسعه داده شده است. در ابتدا به منظور نشان دادن قابلیت مفهوم ظرفیت جذب انرژی در تحلیل پایداری عامل و سیستم های چند عاملی، یک مثال کیفی اجتماعی ارائه شده است.

فردی را در نظر بگیرید که در یک جامعه زندگی کرده و نحوه تعامل او با دیگران از طریق تعامل اقتصادی می باشد. زندگی عادی او در جامعه را می توان بعنوان یک نقطه تعادل تصور نمود. اعتبار اقتصادی این فرد میزان سرمایه وی و درآمدی است که می تواند کسب نماید. از طرف دیگر زندگی در اجتماع هزینه های مختلف اقتصادی را به وی تحمیل می کند. بطور معمول بخشی از سرمایه و درآمد این فرد صرف هزینه های جاری زندگی شده و بخش دیگر صرف هزینه های پیش بینی نشده می شود. واضح است که هر چقدر میزان سرمایه و درآمد یک فرد بیشتر باشد و هزینه های معمول زندگی بخش کمتری از این سرمایه و درآمد وی را در برگیرد، ثبات اقتصادی این فرد بیشتر است. از دیدگاه مفهوم ظرفیت جذب انرژی، اعتبار اقتصادی این فرد همان ظرفیت جذب انرژی وی و هزینه های معمول پیش بینی نشده زندگی، انرژی است که به وی تحمیل می شود. بدیهی است هر چقدر ظرفیت جذب انرژی این فرد بالاتر باشد، امکان تحمل هزینه های اضافی بیشتری را در زندگی خود دارد و در صورتی که هزینه زندگی بیش از توان اقتصادی وی باشد، عملاً از لحاظ اقتصادی ورشکسته می شود که می توان به ناپایداری اقتصادی وی تعبیر نمود.

در این بخش، نخست یک سیستم عامل هایبرید تشریح شده و سپس مفهوم ظرفیت جذب انرژی برای تحلیل پایداری عامل و سیستم های چند عاملی توسعه داده می شود.

الف - موجود هایبرید

در این تحقیق، عامل یک موجود هایبرید در نظر گرفته شده است که شامل تعدادی مود مستقل با تکامل پیوسته درون هر مود بوده و ارائه دهنده ارزش های پیوسته و رویداد های مستقل با رفتار دینامیکی می باشد. حالت های ارزش پیوسته توسط معادلات دیفرانسیلی و معادلات

یک سیستم سوئیچینگ می باشد، در جایی که دو مفهوم در بحث پایداری بایستی در نظر گرفت. اول، تحلیل پایداری حالت های پیوسته درون هر مود و دوم تحقیقات بر روی اثرات سوئیچینگ بین مود ها [۲۶] مد نظر می باشد.

پایداری حالت های پیوسته درون هر مود می تواند بطور مستقیم توسط روش ظرفیت جذب انرژی (بخش ۲) بررسی گردد. فرض کنید x_e نقطه تعادل هر مود باشد. بنابر این

$$x_e \in S \setminus \Omega$$

$$f_i(x_e, 0) = 0, \quad \forall t \geq t_e$$

که در آن t_e نشان دهنده زمان رسیدن به x_e می باشد. دینامیک حالت های پیوسته عامل درون هر مود توسط رابطه زیر توصیف می شود.

$$\dot{x} = f_i(x, u)$$

با توجه به قضایای ۱ و ۲ بخش ۲، سیستم پایدار است اگر و فقط اگر:

$$EAC_{abs}(x_e) > 0$$

پایدار مجانبی است اگر و فقط اگر:

1 - Stable

$$2 - EAC_{inst}(x_{eq}, x) \rightarrow EAC_{abs}(x_{eq}) \text{ as } t \rightarrow \infty$$

در صورتی که بخشی از فضای حالت درون هر مود بعنوان منطقه ناامن Ω شناخته شود، حالت های پیوسته عامل نباید وارد این مناطق نا امن شود و در غیر اینصورت ناپایدار خواهد شد. این بدان معنی است که انرژی جذب شده توسط سیستم پیش از ظرفیت انرژی جذب شده لحظه ای در منطقه نا امن می باشد. هر چند ممکن است برخی مودها دارای مناطق نا امن درون خود نباشند، بنابراین عامل تا زمانی که درون آن مود ها است، پایدار می باشد.

شق دیگر پایداری عامل، مربوط به سوئیچینگ بین مود های گسسته آن می باشد. روش رایج برای تحلیل سیستم های سوئیچینگ با فرکانس سوئیچ دلخواه و البته محدود، روش توابع لیاپانوف مشترک (چندگانه)^{۱۴} می باشد [۲۶] و [۲۷].

با توجه به اینکه فرکانس سوئیچینگ بین مود ها محدود است، بنابراین می توان یک $\tau > 0$ یافت که زمان بین دو سوئیچ متوالی از آن کوچکتر نباشد. در بازه زمانی

$$t \in [t_i, t_{i+1})$$

که

$$t_{i+1} - t_i \geq \tau$$

در مود Q_i قرار گرفته باشد.

با توجه به مطالب بیان شده در بخش دوم، عامل در این مود تا زمانیکه ظرفیت جذب انرژی لحظه ای آن مثبت باشد، پایدار بوده و در غیر اینصورت ناپایدار می گردد. ظرفیت جذب انرژی لحظه ای x_e در مود Q_i در زمان t_i را با $EAC_{inst}(x_e, x(t_i))$ نشان می دهیم. همچنین تصور کنید t_j زمانی می باشد که در آن عامل مجدداً به مود Q_i سوئیچ می نماید. جهت تضمین پایداری عامل، ظرفیت جذب انرژی لحظه ای x_e در مود Q_i و در زمان t_j که برابر $EAC_{inst}(x_e, x(t_j))$ است، نبایستی کمتر از $EAC_{inst}(x_e, x(t_i))$ باشد. به عبارت دیگر میزان انرژی جذب شده عامل هر بار که وارد مود Q_i می شود نباید بیشتر از مقدار آن در زمان های قبلی ورود به این مود باشد.

بعلاوه، برخی انتقالات خاص بین مود ها ممکن است منجر به ناپایداری در سیستم گردد. این انتقالات می تواند در مرحله طراحی مشخص و به عنوان انتقال غیر مجاز مشخص گردند. برای مثال یک موتور القائی را با سه مود: ۱- آماده بکار، ۲- شروع و ۳- عملکرد نرمال در نظر بگیریم. برای سیستم فوق، در صورتیکه انتقال از مود آماده بکار به مود عملکرد نرمال به طور مستقیم انجام شود، ممکن است به موتور خسارت وارد نموده و بنابراین ممنوع می باشد. بنابر این عملکرد صحیح یک موتور القائی از مود آماده به کار آغاز شده و سپس مدتی در مود شروع باقی مانده و نهایتاً وارد مود عملکرد نرمال می گردد. در مود عملکرد نرمال، مقادیر غیر مجاز سرعت و جریان موتور که همان متغیر های حالت می باشند، را می توان مناطق نا امن معرفی نمود که سیستم باید از ورود به آن نواحی خودداری نماید. در نتیجه یک عامل پایدار است اگر شرایط زیر برقرار باشند:

۱- ظرفیت جذب انرژی مطلق عامل درون هر مود مثبت باشد.

۲ - حالت های سیستم از تکامل به مناطق نا امن $x \in S_i \setminus \Omega_i$ اجتناب کنند.

۳- ظرفیت جذب انرژی لحظه ای عامل در مود Q_i نبایستی در زمان های متناوب ورود به این مود کاهش یابد.

$$EAC_{inst}(x_e, x(t_j)) \geq EAC_{inst}(x_e, x(t_i)) \quad (۴)$$

¹⁴ Common (Multiple) Lyapunov Functions

که انرژی اولیه آن کمتر از ظرفیت جذب انرژی مجموعه هدف باشد یعنی

$$\Xi(\Theta_0) < EAC_{abs}(\Theta_d)$$

آنگاه با افزایش تعداد مود های میانی که سیستم از آن ها گذشته است،

$$EAC_{inst}(\Theta_d, \Theta_k) \rightarrow EAC_{abs}(\Theta_d) \text{ as } k \rightarrow \infty$$

دسترسی به مجموعه هدف از طریق مود های میانی و از طریق انتقال برخی رخدادهای بین عامل ها میسر می باشد. برای این منظور لازم است که عامل ها پایدار درونی باشند تا پاسخ به رخدادهای مختلف آن ها را ناپایدار ننمایند. در نتیجه، شرایط پایدار بودن سیستم چند عاملی به شرح زیر می باشد.

۱- هر عامل یک سیستم درونی پایدار باشد.

$$\Xi(\Theta_0) < EAC_{abs}(\Theta_d)$$

۲- برای هر شرایط اولیه Θ_0 که انرژی جذب شده آن کمتر از ظرفیت جذب انرژی مجموعه هدف سیستم چند عاملی باشد،

$$\Xi(\Theta_0) < EAC_{abs}(\Theta_d)$$

و رابطه زیرین بایستی برقرار باشد.

$$EAC_{inst}(\Theta_d, \Theta_k) \rightarrow EAC_{abs}(\Theta_d)$$

as $k \rightarrow \infty$

(۵)

۴- مقایسه روش لیاپانوف و روش ظرفیت جذب انرژی

روش های لیاپانوف و ظرفیت جذب انرژی هر دو بر مبنای انرژی هستند. بنابراین شباهتهای زیادی با هم دارند. هر چند که تفاوت هایی نیز دارند. هر دو روی پایداری نقطه تعادل سیستم تمرکز می کنند. هر دو بر مبنای یک تابع انرژی برای تحلیل پایداری عمل می کنند که تعیین این تابع انرژی برای برخی سیستم ها ممکن است سخت ترین مرحله تحلیل باشد.

ممکن است ادعا شود که تعیین ظرفیت جذب انرژی مناسب یک سیستم یک مرحله چالش برانگیز آن است. ولی این ادعا در صورتی درست است که فقط توصیف ریاضی سیستم در دسترس باشد. ولی اگر اطلاعات فیزیکی سیستم نیز معلوم باشد، ظرفیت جذب انرژی سیستم را می توان راحتتر محاسبه نمود. این کار با استفاده از تجزیه سیستم به المان های اصلی تشکیل دهنده آن که در واقع المان های ساده ذخیره کننده انرژی و تلفاتی هستند، انجام خواهد شد. ظرفیت جذب انرژی این المان ها براحتی از روی اطلاعات دینا شیت آن ها و یا توسط آزمایشات ساده قابل تعیین خواهد بود.

چنانکه t_1 و t_j زمان های ورود اول و دوم عامل به مود Q_j می باشند.

۴- عامل بایستی از بعضی انتقالات مستقیم اجتناب بنماید.

برای عامل های بدون دینامیک مانند عامل های نرم افزاری، شرط اول فوق قابل صرف نظر می باشد. از طرف دیگر، برای عامل هایی که فقط یک مود دارند، فقط شرط اول بایستی بررسی گردد.

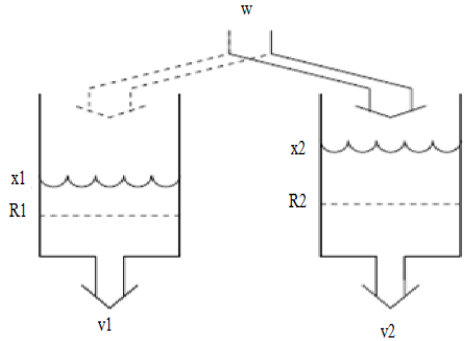
ج- پایداری سیستم های چند عاملی

یک سیستم چند عاملی مطلوب معمولاً از یک شرایط اولیه Θ_0 شروع نموده و سعی می کند از طریق تعداد محدودی مود های میانی Θ_k به هدف Θ_d دست یابد. اگر مجموعه هدف یک مجموعه نامتغیر^{۱۵} تعریف شود، هر انتقال مود درون آن برای همیشه درون آن باقی می ماند. بنابراین پایداری یک سیستم چند عاملی منسوب را می توان به پایداری مجانبی مجموعه هدف Θ_d مرتبط نمود. فرض کنید که $\rho(\Theta_0, \Theta_d)$ نشان دهنده فاصله بین مجموعه اولیه Θ_0 و مجموعه هدف Θ_d باشد. این فاصله را می توان به اختلاف انرژی بین آن دو مجموعه متناظر نمود. این اختلاف انرژی برای هر دو مود متوالی به فرم

$$\rho(\Theta_{k-1}, \Theta_k) = \Xi(\Theta_{k-1}) - \Xi(\Theta_k)$$

قابل بیان است که در آن $\Xi(\Theta_k)$ انرژی سیستم در مود میانی Θ_k می باشد. بنابراین $\rho(\Theta_k, \Theta_d)$ نشان دهنده اختلاف انرژی جذب شده سیستم در مود Θ_k نسبت به مجموعه اهداف Θ_d می باشد. بر اساس قضیه لیاپانوف، سیستم چند عاملی پایدار است و یا به طور معادل، مجموعه هدف آن پایدار مجانبی می باشد اگر $\rho(\Theta_k, \Theta_d)$ در هر k کاهش یابد و به ازای $k \rightarrow \infty$ مقدار $\rho(\Theta_k, \Theta_d)$ به سمت صفر میل کند. از نقطه نظر ظرفیت جذب انرژی، هر گونه انحراف از نقطه تعادل بدلیل جذب انرژی توسط سیستم می باشد. فرض کنید که سیستم با جذب انرژی $E_0 = \Xi(\Theta_0)$ وارد شرایط اولیه Θ_0 شده است. همچنین فرض کنید $EAC_{abs}(\Theta_d)$ ظرفیت جذب انرژی مطلق مجموعه اهداف بوده و $EAC_{inst}(\Theta_d, \Theta_k) = EAC_{abs}(\Theta_d) - \Xi(\Theta_k)$ ظرفیت جذب انرژی لحظه ای در هر مود میانی Θ_k می باشد. مجموعه هدف بطور مجانبی پایدار می باشد اگر برای هر شرایط اولیه Θ_0

¹⁵ Invariant Set



شکل ۳: سیستم ذخیره و توزیع آب

شکل (۴) نشان دهنده ساختار داخلی این سیستم می باشد، که در آن عامل دارای دو مد مستقل q_1 , q_2 بوده و دینامیک داخلی مود q_1 ، q_2 به ترتیب توسط روابط زیر نشان داده شده است.

$$dx_1/dt = W - V_1 \quad (۶)$$

$$dx_2/dt = -V_2$$

$$x_2 \geq R$$

$$dx_1/dt = -V_1 \quad (۷)$$

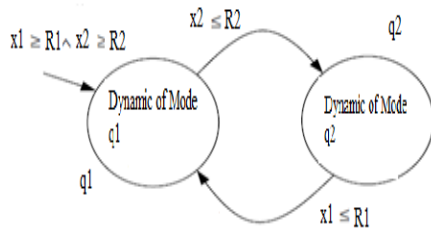
$$dx_2/dt = W - V_2$$

$$x_2 \geq R$$

هدف این سیستم نگه داشتن سطح آب (x_1 , x_2) متغیرهای حالت سیستم و نشان دهنده سطوح آب در مخازن ۱ و ۲) در مخازن ۱ و ۲ بالای سطوح مرجع R_1 , R_2 می باشد. هدف کنترل، تضمین قیود زیر می باشد.

$$x_1 > R_1 \text{ and } x_2 > R_2$$

از طرف دیگر، رخداد های ($x_2 \leq R_2$, $x_1 \leq R_1$) که منجر به انتقال از یک مود به مود دیگر می شود، حاصل تکامل حالت های پیوسته در رسیدن به مقادیر از پیش تعریف شده R_1 , R_2 می باشد.



شکل ۴: ساختار داخلی سیستم ذخیره و توزیع آب

برای تحلیل پایداری سیستم ذخیره و آبرسانی، یک عامل هایبرید با دو مود در شکل (۴) نشان داده شده است. در واقع این عامل هایبرید یک سیستم سوئیچ می باشد که در آن دو مفهوم پایداری مطرح می باشند. (۱) تحلیل پایداری حالت های پیوسته درون هر مود و (۲) تحقیق

تحلیل پایداری در روش لیپانوف بر اساس یک تابع شبه انرژی (کاندیدای لیپانوف) انجام می شود. در صورت یافتن این تابع می توان پایداری را تحلیل نمود و در غیر اینصورت نمی توان در این خصوص اظهار نظر کرد. علاوه بر این توابع لیپانوف مختلف، تخمین های مختلفی از نواحی پایداری ارائه می دهند. به این دلیل است که روش لیپانوف یک شرط کافی برای پایداری و ناپایداری ارائه می دهد.

در روش ظرفیت جذب انرژی، تحلیل پایداری بر اساس مقدار حد بالای تابع انرژی و نه خود تابع انجام می شود. یک سیستم تا زمانیکه انرژی جذب شده آن کمتر از این مقدار باشد پایدار باقی می ماند و بر عکس. بنابراین می توان ادعا نمود که روش ظرفیت جذب انرژی نسبت به روش لیپانوف کمتر محافظه کارانه است.

همچنین ظرفیت جذب انرژی معیار ملموس تری از پایداری ارائه می دهد چرا که پایداری را به یک مقدار اسکالر ربط می دهد. با کمک آن می توان میزان پایداری در سیستم را با هم مقایسه نمود. هر چقدر ظرفیت جذب انرژی بیشتر باشد، پایداری بیشتر است و بر عکس. بنابر این از این روش می توان برای مقایسه پایدار کننده های مختلف برای یک سیستم مشخص نیز استفاده نمود.

۵- پیاده سازی

کار آرای روش پیشنهادی بر روی دو پروسه مختلف مورد بررسی قرار گرفته است. پروسه اول، یک سیستم ذخیره و توزیع آب متشکل از دو مخزن می باشد که در آن سیستم اصلی نشان دهنده یک عامل هایبرید می باشد و هدف دستیابی به اهداف از پیش تعریف شده است. پروسه دوم یک سیستم چند عاملی تعقیب - گریز می باشد که در آن عملیات از یک مجموعه شرایط اولیه آغاز شده و از طریق مد های میانی محدود به هدف نهایی، که دستگیری عامل فراری است، دست می یابد.

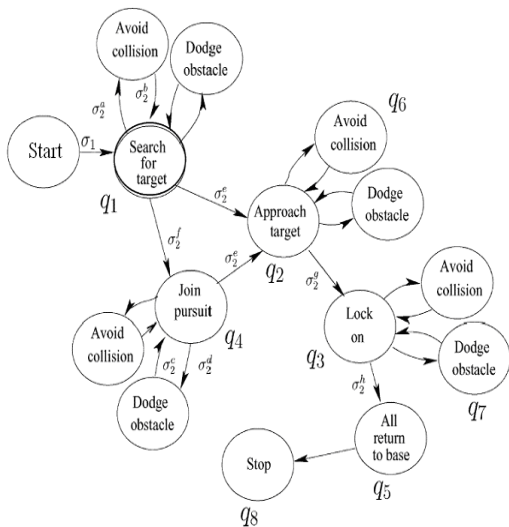
الف- سیستم ذخیره و توزیع آب

شکل (۳) نشان دهنده دو مخزن آب می باشد که هر دو دارای خروجی آب با تغییرات حجمی در واحد زمان و مستقل از یکدیگر می باشند [۲۹]. آب ورودی به سیستم با تغییرات حجمی در واحد زمان از طریق یک شیر در هر لحظه از زمان می تواند به یکی از دو مخزن هدایت گردد.

در ارتباط با اثرات سوئیچ بین مود ها. طبق تعریف ۱ این سیستم را پایدار گوئیم اگر به وضعیت برگشت پذیر نرود. وضعیت برگشت پذیر در این سیستم وضعیتی است که آب از مخازن سر ریز شده و بیرون بریزد که در این صورت سیستم عملاً ناپایدار شده است. جهت تحلیل پایداری هر مخزن، ظرفیت جذب انرژی مطلق مخزن ۱ و ۲ در نقاط تعادل باید تعریف شود. در اینجا ظرفیت جذب انرژی سیستم معادل حداکثر حجم آب است که می شود داخل مخزن ریخت بدون آنکه سرریز شود.

در آن $EAC_{abs i} = A_i H_i$ که در آن A_i سطح مقطع و H_i ارتفاع مخزن A_i می باشد. اگر سطح آب در هر کدام از مخازن از ارتفاع ماکزیموم مخزن تجاوز نماید، آب از مخزن سر ریز شده و سیستم ناپایدار می گردد. بنابراین برای هر مود می توان یک ظرفیت جذب انرژی بر اساس ظرفیت حجمی آن تعیین نمود. در ارتباط با شق دوم پایداری باید فرکانس سوئیچ مورد نیاز بر اساس دینامیک سیستم را بدست آورد و آنرا با مشخصات فیزیکی سوئیچ ω مقایسه نمود. اگر سوئیچ ω از نظر مکانیکی توانایی سوئیچ کردن با آن فرکانس را نداشت، پایداری سیستم از این نظر دچار اختلال خواهد شد.

طرح منطقی دستگیری هدف، شامل ۸ مود نشان داده شده در شکل (۵) می باشد. فهرست حوادث در جدول شماره (۱) داده شده است. مود پیش فرض کنترلی برای، جستجو برای هدف^{۱۹} می باشد که در آن محیط اطراف توسط سنسورها بطور تصادفی جاروب شده تا هدف پیدا گردد و یا آنکه بر اثر رخداد جدیدی، خودرو به مود "به سمت هدف رفتن"^{۲۰} و یا مود "تعقیب تیمی"^{۲۱} سوئیچ نماید. در این مود، خودرو جستجوی خود را رها کرده و به سمت هدف می رود و یا به طرف مختصات هم تیمی خود می رود. به هدف نزدیک شده و بر روی هدف قفل می نماید و بتدریج هدف دستگیر می شود.



شکل ۵: تعیین توالی مود های یک تیم "خودروی زمینی بدون سرنشین"

جدول ۱: فهرست حوادث برای تیم "خودروی زمینی بدون سرنشین"

حوادث	تفسیر
σ_2^a	دیگر خودرو ها قابل رویت بوده و برخورد امکانپذیر میباشد
σ_2^b	چیزی قابل رویت نمی باشد
σ_2^c	مانعی دیگر دیده نمیشود
σ_2^d	هدف در زاویه دید می باشد
σ_2^e	هدف توسط عضو تیم بدست آورده شده است
σ_2^g	هدف به اندازه کافی نزدیک برای دستگیری می باشد
σ_2^h	هدف دستگیر شده است

ب - سیستم چند عاملی تعقیب^{۱۶} و گریز^{۱۷}
 در این مثال، دو خودروی زمینی بدون سرنشین^{۱۸} با همکاری یکدیگر سعی در دستگیری یک خودروی فراری (خودروی زمینی بدون سرنشین) را دارند [۲۱]. هر عامل (خودروی تعقیب کننده) دارای سنسور های بینایی جهت تشخیص خودروی در حال فرار و سنسور های مناسب جهت تشخیص موانع و اندازه گیری موقعیت آن ها نسبت به مرجع می باشد. دینامیک زمان هر خودرو توسط روابط زیر توصیف می شود.

$$\begin{aligned} \dot{X} &= v \cos(\theta) \\ \dot{Y} &= v \sin(\theta) \\ \dot{\theta} &= \omega \end{aligned} \quad (8)$$

جائیکه X و Y مرکز جرم و θ جهت خودرو می باشد. سیستم (۸) همچنین می تواند در فرم کلی زیر نشان داده شود.

$$\begin{aligned} \dot{x} &= f_{UGV}(x, u) \\ u &= [v, \omega]^T \text{ و } x = [X, Y, \theta]^T \end{aligned}$$

جائیکه بردار ورودی برای اهداف کنترلی می باشد که در آن

¹⁹ Search for Target

²⁰ Approach Target

²¹ Join Pursuit

¹⁶ Pursuit

¹⁷ Evasion

¹⁸ Unmanned Ground Vehicle

در عمل، اگر عامل ها (خودروی زمینی بدون سرنشین) واژگون شوند، ناپایدار می گردند. اگر سرعت های انتقالی و زاویه ای خودروها از حد مجاز v_{EAC} و ω_{EAC} تجاوز نماید، واژگون می شوند. این مقادیر مرز های مناطق غیر مجاز درون هر مود را نشان می دهند. ظرفیت جذب انرژی مطلق هر یک از عامل ها بر اساس انرژی جنبشی آن ها در v_{EAC} و ω_{EAC} داده شده است.

$$EAC_{abs}(x_e) = \left(\frac{1}{2}mv^2 + \frac{1}{2}I\omega^2 \right) \Big|_{v=v_{EAC}}^{\omega=\omega_{EAC}} \quad (9)$$

چنانکه m, I نشان دهنده جرم و اینرسی می باشند. بدلیل آنکه هیچگونه محدودیت های فیزیکی در این مثال ذکر نشده است، بنابراین $EAC_{abs}(x_e)$ درون هر مود مثبت بوده و انرژی جذب شده عامل ها کمتر از $EAC_{abs}(x_e)$ برای تمام شرایط مجاز می باشد. نهایتاً، سوئیچ بین مود ها بگونه ای طراحی شده است که انتقال ممنوع انجام نشود. بنابراین هر یک از عامل ها (خودروها) داخلی پایدار می باشند.

تیم چند عاملی وقتی ناپایدار می گردد که هدف نهایی (دستگیری عامل فراری به دلایل هماهنگی ضعیف و یا تصادف با دیگر عامل ها و یا موانع ویا موارد مشابه) قابل دسترسی نباشد. تیم به هدف Θ_d (دستگیری عامل فراری) در زمان های محدود در صورتی نزدیک می شود که از شرایط اولیه Θ_0 با قید

$$\Xi(\Theta_0) < EAC_{abs}(\Theta_d)$$

آغاز به کار کرده باشد. مقدار انرژی $EAC_{abs}(\Theta_d)$ مرتبط به پارامتر هایی از قبیل هماهنگی سیستم های چند عاملی، تاخیر در ارتباطات و توانمندی عامل ها در تشخیص هدف، موانع، یکدیگر و ظرفیت انرژی ذخیره شده مطلق هر عامل می باشد. سیستم چند عاملی بطوری طراحی شده است که رابطه

$$EAC_{abs}(\Theta_d) > 0$$

برقرار می باشد. بعلاوه، سیستم چند عاملی با گذشت زمان به هدف نزدیک می شود، به عبارت دیگر

$$EAC_{inst}(\Theta_d, \Theta_k) \rightarrow EAC_{abs}(\Theta_d) \text{ as } k \rightarrow \infty$$

برقرار می باشد. این به آن معنی می باشد که مجموعه هدف Θ_d بطور مجانبی پایدار و یا بطور هم ارز سیستم چند عاملی پایدار می

فرامین اصلی کنترل مربوط به هر مود، مجموعه ای از حرکات چرخشی و رو به جلو برای جستجوی محیط اطراف تولید می کند. کنترل اولیه مودها در جدول (۲) داده شده است. در مود q_1 ، q_1 یک سرعت ثابت و θ_i جهت خودروی i ام می باشد. در مود q_2 ، q_2 ، ρ_1 فاصله اقلیدسی^{۲۲} بین خودروی i ام و خودروی در حال فرار می باشد. در مود q_3 ، θ_p ، θ_i جهت خودروی i ام و هدف نهایی می باشد. در مود q_4 ، θ_1 ، θ_2 جهت اولین و دومین خودروی زمینی بدون سرنشین می باشند. در مود q_5 ، q_5 ، ρ_3 فاصله اقلیدسی بین خودروی i ام و نقطه شروع عملیات می باشد. تابع ω در مودهای ششم و هفتم یک میدان برداری دافع^{۲۳} با رابطه مربع معکوس فاصله بین خودرو و هدف تولید می کند. در این تابع

$$\tilde{\theta} = \text{atan2}(Y_{ob} - Y_i, X_{ob} - X_i)$$

که در آن (X_i, Y_i) مختصات خودرو و (X_{ob}, Y_{ob}) موقعیت تخمین زده شده مانع (هدف عدم برخورد با مانع) می باشد. در این روش، خودروی زمینی بدون سرنشین، سعی بر عدم برخورد با دیگر خودروها (UGV)، هدف و موانع موجود در محیط خواهد بود.

جدول ۲: فرامین اصلی کنترل خودروی بدون سرنشین

مود	برچسب	فرامین اصلی کنترل
q_1	جستجو برای هدف	$v = \text{fwd}, \omega = \text{rnd } \theta_i$
q_2	نزدیک شدن به هدف	$v = K_1 \rho_1, \omega = 0$
q_3	قفل کردن بر روی هدف	$v = K_1 \rho_1 \cos(\theta_p - \theta_i),$ $\omega = K_3(\theta_p - \theta_i) + K_4(\dot{\theta}_p - \dot{\theta}_i)$
q_4	ملحق شدن به تعقیب	$v = K_5 \rho_2 \cos(\theta_1 - \theta_2),$ $\omega = K_6(\theta_1 - \theta_2) + K_7(\dot{\theta}_1 - \dot{\theta}_2)$
q_5	برگشت به پایگاه	$v = K_8 \rho_3, \omega = K_9(\theta_i - \theta_b)$
$q_{6,7}$	دوری از تصادف و فرار از موانع	$v = 0, \omega = f_{AC}(d)(\pi - \tilde{\theta} - \theta_i)$
q_8	توقف	$v = 0, \omega = 0$

²² Distance Euclidean

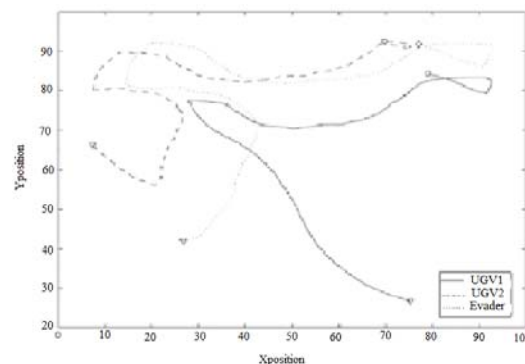
²³ Repulsive Vector Field

از پیش تعریف شده می باشند. مفهوم ظرفیت جذب انرژی مقدار ماکزیموم انرژی می باشد که هر سیستم در هر نقطه تعادل خود میتواند جذب کند قبل از اینکه به وضعیت برگشت ناپذیر برود. در این تحقیق، هدف اصلی در بسط مفهوم ظرفیت جذب انرژی در توسعه معیارهای تحلیل پایداری سیستم های چند عاملی در بر گیرنده تکامل حالت های پیوسته درون هر مود و سوئیچینگ بین مود ها می باشد. در ادامه مقایسه ای بین روش پیشنهادی و روش لیاپانوف انجام شده و شباهت ها و تفاوت ها بیان شده است. نتایج شبیه سازی نشان دهنده تحلیل مناسبتر و ملموس تر روش پیشنهادی در مقایسه با روش های معمول می باشد. دیدگاه های آتی نویسندگان در توسعه و استفاده از روش پیشنهادی در تحلیل پایداری گروه^{۲۴} و سیستم های چند عاملی دیگر می باشد.

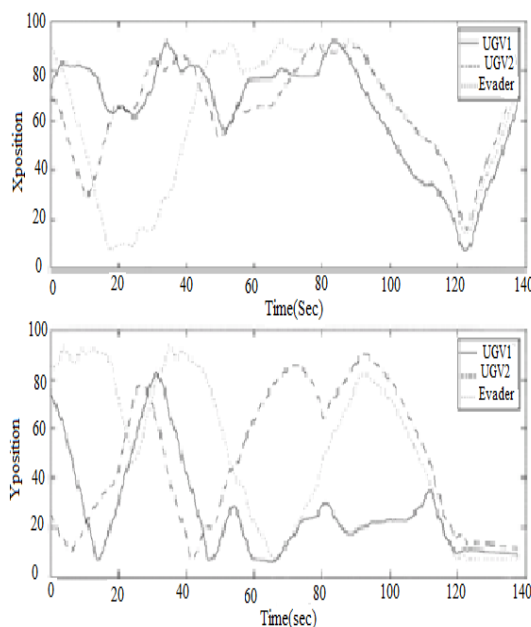
مراجع

- [1] Xiaohua, L. and J. L. Feng Wang, "Simulation system of Production Scheduling Multi-Agent-Based," World Congress on Computer Science and Inf. Eng. (2009).
- [2] Leitão, P., "Agent-based Distributed Manufacturing Control: A State-of-the-art Survey," International Journal of Engineering Applications of Artificial Intelligence, vol. 22, no. 7, pp. 979-991, Elsevier, (2009).
- [3] H. Hattori, Y. and Y. Nakajima, and T. Ishida, "Modeling Behaviors for Multi-Agent Traffic Simulation," In Proceeding of Autonomous Agents and Multiagent Systems, 6th Workshop on Agents in Traffic and Transportation, Toronto, Canada, (2010).
- [4] Khalil, H. K., "Nonlinear Systems," Prentice Hall, 3th edn., (2001).
- [5] Krstic, M. and I. Kanellakopoulos, and P. Kokotovic, "Nonlinear and Adaptive Control Design," John Wiley & Sons, (1995).
- [6] Willems, J., "Dissipative dynamical systems, part i: General theory," Arch. Rat. Mech. An. 45, (1972).
- [7] Chli, M. and P. D. Wilde, and R. Ribeiro, "Stability of Multi-Agent Systems," Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, pp 551-556, Piscataway, NJ, (2003).
- [8] Moreau, L., "Stability of Multiagent Systems With Time-Dependent Communication Links," IEEE Trans. Autom. Control, vol. 50, no. 2, pp. 169-182, Feb. (2005).

باشد. شبیه سازی های انجام شده نیز نشان دهنده دستگیری عامل فراری در زمان محدود می باشند. نتایج شبیه سازی در شکل های (۶) و (۷) نشان داده شده اند. با توجه به این نتایج، عامل فراری در مدت زمان ۱۲۰ ثانیه دستگیر شده است. بنابراین با توجه به موارد اشاره شده در بخش سوم در خصوص پایداری عامل و سیستم چند عاملی، می توان نتیجه گرفت که این سیستم پایدار است.



شکل ۶: مختصات حرکتی خودروهای تعقیب گر و فراری



شکل ۷: مسیر حرکت خودروهای تعقیب گر و فراری

۶ - جمع بندی و دیدگاه های آتی

مفهوم ظرفیت جذب انرژی در تحلیل پایداری عامل و سیستم های چند عاملی گسترش داده شده است، جاییکه عامل یک موجود هاپرید در نظر گرفته شده است و سیستم های چند عاملی، گروهی از عامل ها می باشند که با همکاری یکدیگر بدنبال رسیدن به یک هدف

²⁴ Swarm

- Simulation”, Ph.D. Thesis, University of Twente, The Netherlands, (2002).
- [20] Garcia, E. and J. Estremera and P. Gonzalez de Santos, “A comparative study of stability margins for walking machines,” *Robotica*, volume 20, pp 595–606, (2002).
- [21] Garcia, E. and P. Gonzalez de Santos, “An improved energy stability margin for walking machines subject to dynamic effects,” *Robotica*, volume 23, pp. 13–20, (2005).
- [22] Fregene, K. and D. C. Kennedy, and D. W. L. Wang, “Toward a Systems- and Control-Oriented Agent Framework,” *IEEE Trans. on System, Man, and Cybernetics—Part B: Cybernetics*, Vol. 35, No. 5, pp 999-1012, Oct. (Oct., 2005).
- [23] Van Der Schaft, A. “Port-Hamiltonian Systems: an Introductory Survey,” in *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, (2006).
- [24] Siljak, D. D. “Decentralized control complex systems,” Academic Press INC., (1991).
- [25] Ortega, R. and A. Van Der Schaft, and B. Maschke & G. Escobar, “Energy-shaping of Port-Controlled Hamiltonian Systems by Interconnection,” in *proceedings of IEEE Conference on Decision and Control*, Phoenix, Arizona, (1999).
- [26] Liberzon, D. and A. S. Morse, “Basic Problems in stability and Design of Switched Systems,” *IEEE Control Systems Magazine*, vol.19, no.5, pp.59-70, (1999).
- [27] Branicky, M. S. “Multiple Lyapunov Functions and other Analysis Tools for Switched and Hybrid Systems,” *IEEE Trans. Automat. Contr.*, vol. 43, no. 4, pp. 475–482, (1998).
- [28] Guenther, O. and T. Hogg, and B. A. Huberman, “Learning in Multiagent Control of Smart Matter,” in *proceeding of AAAI97 Workshop on Multiagent Learning*, (1997).
- [29] Bak, T. and R. Izadi-Zamanabadi, “Lecture Notes-hybrid systems,” Department of Control Engineering, Aalborg University, Aalborg, Denmark, (Oct. 27, 2004).
- [9] Gazi, V. and K. M. Passino, “Stability analysis of swarms,” *IEEE Trans. Autom. Control*, vol. 48, No. 4, pp. 692-697, Apr. (2003).
- [10] Olfati-Saber, R. and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. (2004).
- [11] Lee, L. C. and H. S. Nwana and D. T. Ndumu, and P. De Wilde, “The stability, scalability and performance of multi-agent systems,” *BT Technol J*, Vol 16, No 3, pp 94-103, July (1998).
- [12] <http://www.modelica.org>
- [13] Lewis, R., “ Modeling control systems using IEC61499: applying function blocks to distributed systems,” *Institution of Electrical Engineers, U.K.*, (2001).
- [14] Jamshidifar, A. A. and A. Afshar, A. Salari, “FBLab, an IEC61499 based software tool for modeling of distributed control systems,” *proceedings of the 2003 IEEE International Conference on Control Application*, Istanbul, Turkey, pp. (551-556), June (2003).
- [15] Jamshidifar, A. A. and A. Afshar, “Application of Object-oriented Approach to Large-Scale System, a Case Study,” *Proceedings of the 2004 IFAC International Conference on Large-Scale Systems*, Osaka, Japan, pp. 537-541, July (2004).
- [16] Jamshidifar, A. A. and A. Afshar, “Stability Analysis of Large-Scale Agent-Based Systems,” *Proceedings of the 2006 IEEE International Conference on Information, Control and Innovative Computing*, pp. 382-385, Beijing, China, (2006).
- [17] Jamshidifar, A. A. and A. Afshar, “Energy Absorption Capacity: A New Concept for Stability Analysis of Nonlinear Dynamic Systems,” *Asian Journal of Control*, Vol.11, no.1, pp.1-9, (2008).
- [18] Wellstead, P. E., “Introduction to Physical Pystem Modelling,” Academic press LTD., (1979).
- [19] Golo, G., “Interconnection Structures in Port-Based Modelling: Tools for Analysis and

تقسیم بهینه توان اکتیو و راکتیو بین چند منبع تولید پراکنده در یک شبکه مستقل با شبکه عصبی مصنوعی

علی اصغر قدیمی^۱، حسن رستگار^۲، فرزاد رضوی^۳

^۱استادیار، دانشگاه اراک، دانشکده فنی و مهندسی، گروه مهندسی برق، a-ghadimi@araku.ac.ir

^۲دانشیار، دانشگاه صنعتی امیرکبیر، دانشکده مهندسی برق، rastegar@aut.ac.ir

^۳استادیار، دانشگاه تفرش، دانشکده مهندسی برق، razavi.farzad@taut.ac.ir

(تاریخ دریافت مقاله ۱۳۸۹/۱۲/۲۵، تاریخ پذیرش مقاله ۱۳۹۰/۲/۱۱)

چکیده - در این مقاله به تقسیم بهینه توان اکتیو و راکتیو بین منابع تولید پراکنده تغذیه کننده یک شبکه مستقل با استفاده از شبکه عصبی مصنوعی پرداخته شده است. منابع تولید پراکنده تغذیه کننده شبکه مستقل توسط اینورتر به شبکه متصل شده و یک سیستم مدیریت هوشمند و on-line با استفاده از شبکه های عصبی مصنوعی، مقادیر بهینه میزان تولید واحدها را ارائه می دهد. کنترل کننده محلی موجود در کنار هر سیستم تحقق میزان توان تولیدی ارائه شده به آن را تضمین می کند. شبکه عصبی با داده هایی که از حل مسئله پخش بار بهینه با استفاده از الگوریتم ژنتیک بدست آمده بگونه ای آموزش داده شده که میزان توانهای تولیدی واحدها را به منظور کمینه سازی تلفات سیستم و با در نظر گرفتن قیود ولتاژ شین ها و مقادیر نامی واحدها تعیین نماید. نتایج شبیه سازی بر روی یک سیستم توزیع نمونه نشان می دهد که روش ارائه شده قادر به برآورده نمودن نیازهای سیستم و مقادیر نامی واحدهای تولید پراکنده در حالت های مختلف کارکرد بوده و کنترل کننده ارائه شده سیستم را بصورت on-line در حالت بهینه مدیریت و پیوستگی تغذیه بارها را تضمین می نماید.

کلمات کلیدی: تولید پراکنده (DG)، MicroGrid، کنترل، تخمین حالت، الگوریتم ژنتیک، شبکه عصبی مصنوعی، بهینه سازی

Optimal Active and Reactive Power Sharing between several Distributed Generation Units in a Stand-Alone MicroGrid Using Artificial Neural Network

Ali Asghar Ghadimi, Farzad Razavi, Hassan Rastegar

Abstract: This paper is concerned with the generation control in small stand-alone MicroGrids consisting of inverter interfaced Distributed Generation (DG) units. An intelligent and on-line Microgrid Management System (MGMS) using Artificial Neural Network (ANN) controller is used in this study and it determines the amount of power produced from generation units in a stand-alone MicroGrid. The ANN trained with a data generated from a Genetic Algorithm (GA) solved optimal power flow problem, which defines generation unit's power in order to have a minimum power loss in the system, considering normal buses voltage and rating of generation units. Simulation results in a typical distribution system consisting two DG units show that the proposed method can meet the requirements of the system and DG units rating in stand-alone operations. In this way the system is managed in an on-line, optimal, and reliable situation that guarantee the continuity of power to loads in stand-alone mode of a MicroGrid.

Keywords: MicroGrid, State Estimation, Genetic Algorithm, Artificial Neural Network, Load Sharing, Distributed Generation

۱- مقدمه

تمرکز نموده اند و کنترل توان حقیقی چندین منبع تولید پراکنده مد نظر قرار نگرفته است.

به منظور افزایش سطح نفوذ منابع تولید پراکنده، ساده نمودن طراحی و کارکرد انعطاف پذیر، سیستمهای متصل شده با تعداد زیادی واحد تولید پراکنده نیاز به سیستمهای کنترل on-line و هوشمند دارند. این وسایل باید رفتار سیستم را پس از هر گونه تغییر یا اغتشاش زیر نظر داشته و بر اساس آن، سیگنالهای کنترلی بهینه و مناسب و یا عملیات پیشگیرانه لازم را (شبیه کارکرد سیستم AGC در شبکه های قدرت کنونی) ارائه دهند. بدین منظور در این مقاله ابتدا با استفاده از الگوریتم ژنتیک با توجه به تغییرات و سطوح مختلف بار در شین های سیستم، سیگنال های کنترلی مناسب و بهینه به صورت off-line بدست آمده است. با توجه به اینکه بهینه سازی پروسه ای زمان بر می باشد، لذا یک شبکه عصبی مصنوعی^۳ (ANN) بگونه ای طراحی شده است تا عملکرد بهینه سیستم را به صورت on-line تضمین نماید شبکه عصبی با اطلاعات بدست آمده از بهینه سازی توسط الگوریتم ژنتیک آموزش داده شده و سرانجام شبکه عصبی مصنوعی سیستم را در حالت بهینه، دقیق و به صورت on-line مدیریت خواهد کرد.

بر اساس اهداف اشاره شده، در بخش ۲ سیستم مدیریت MicroGrid معرفی و همه اجزای آن تشریح شده است. در بخش ۳ نحوه بهینه سازی با الگوریتم ژنتیک شرح داده خواهد شد. بخش ۴ طراحی یک شبکه عصبی جهت تعیین مقادیر تولید توان اکتیو و راکتیو هر عنصر DG به صورت on-line انجام می شود. نتایج شبیه سازی بر روی یک MicroGrid نمونه در بخش ۵ ارائه شده است. سرانجام بخش ۶ نتایج مقاله را ارائه می دهد.

۲- تشریح سیستم مدیریت MicroGrid و اجزای آن

مدیریت فعال^۴ نوعی از کنترل مرکزی برای شبکه های توزیع است و بعنوان روشی برای افزایش سطح نفوذ منابع تولید پراکنده مورد استفاده قرار گرفته است [۱۸]-[۲۱]. شبیه سیستم های قدرت سنتی این سیستم مدیریت می تواند برای کنترل ولتاژ و فرکانس و همچنین توان اکتیو و راکتیو یک سیستم شامل چندین MicroGrid بکار رود. شکل (۱) این ساختار را نشان می دهد [۲۱] که سه سطح کنترلی برای آن پیشنهاد می گردد:

- سیستم مدیریت توزیع^۵ (DMS) و اپراتور بازار^۶ (MO) که در سطح ولتاژ متوسط نصب می شوند.

امروزه مسائل زیست محیطی، بازار آزاد برق، نیاز به کاهش آلودگی CO₂، تغییرات آب و هوایی و همچنین قیمت و محدودیتهای سوختهای فسیلی باعث رغبت شدید در اتصال منابع کوچک نظیر سلولهای خورشیدی، پیلهای سوختی، میکروتوربین ها، ژنراتورهای بادی کوچک و تجهیزات ذخیره ساز انرژی در شبکه های فشار ضعیف شده است [۱]-[۳]. این شیوه جدید تولید انرژی الکتریکی باعث گسترش شبکه های کوچک شامل منابع تولید پراکنده^۱ و سایر تجهیزات تحت عنوان MicroGrid شده که پیش بینی می شود نقش بسیار مهمی در سیستمهای قدرت آینده داشته باشد [۴] و [۵].

مسئله اصلی در یک MicroGrid مستقل که در اثر قطع اتفاقی یا برنامه ریزی شده شبکه ایجاد می گردد، هماهنگی منابع تولید موجود برای تقسیم توانهای اکتیو و راکتیو خروجی آنها و کنترل ولتاژ و فرکانس است.

کنترل واحدهای تولید پراکنده در حالت مستقل به طرق مختلفی قابل انجام است. مرجع [۶]، یک کنترل کننده مود لغزشی^۲ برای اینورترهای سه فاز چهار سیمه ارائه نموده است. مرجع [۷] با استفاده از یک تکنیک PWM جدید روشی برای تولید ولتاژ سینوسی با وجود بارهای غیر خطی در سیستم ارائه نموده است. مرجع [۸] یک استراتژی کنترل برای هر دو حالت متصل به شبکه و جدا از شبکه ارائه نموده است که دارای خاصیت تشخیص جزیره ای شدن داخلی می باشد. همچنین مرجع [۹] یک استراتژی کنترل جدید برای کنترل یک واحد تولید پراکنده متصل شده با اینورتر ارائه داده است. کلیه این مقالات بر روی کنترل اینورترها در حالت مستقل و تنها یک اینورتر و تعدادی بار تمرکز نموده اند.

در برخی دیگر از مقالات [۱۰]-[۱۴]، مفهوم دروپ توان اکتیو-فرکانس و توان راکتیو-ولتاژ شبیه سیستمهای قدرت کنونی برای کنترل توان منابع تولید متصل شده توسط اینورتر مورد استفاده قرار گرفته است. مزیت این روش در استفاده از متغیرهای محلی برای کنترل واحدهای تولیدی می باشد. اما با توجه به تغییر فرکانس بعد از هر گونه اغتشاش، یک حلقه کنترل ثانویه برای کنترل دقیق و پایدار فرکانس لازم است. بعلاوه تقسیم بار بین واحدهای تولید به منظور کارکرد بهینه در شبکه و برآورده کردن همه قیود سیستم با این روش امکان پذیر نمی باشد. سیستمی با یک اینورتر منبع ولتاژ بعنوان واحد اصلی و مابقی بعنوان واحد کمکی در [۱۵] ارائه شده است. تحقیقات حاضر نظیر [۱۶] و [۱۷] از کنترل مرکزی برای کنترل بهینه تپ چنجرها یا جبران سازهای استاتیکی توان راکتیو برای تنظیم ولتاژ در سیستم توزیع استفاده نموده اند و تنها بر روی حالت متصل به شبکه و کنترل ولتاژ

³ Artificial Neural Network

⁴ Active Management

⁵ Distribution Management System

⁶ Market Operator

¹ Distributed Generation

² Sliding Mode

می گیرد تا مقدار توان تولیدی واحدهای تولید پراکنده و فرمانهای کنترلی برای سایر تجهیزات قابل کنترل را محاسبه نماید. هدف نهایی این مقاله ارائه استراتژی کنترلی برای MicroGrid جدا از شبکه^۳ به منظور کارکرد نرمال و بهینه سیستم است. بنابراین هر یک از تجهیزات سیستم باید کنترل کننده ای داشته و هماهنگی بین همه اجزا و مدیریت سیستم با MGMS انجام می شود. زیر بخشهای زیرین بر روی تشریح و مدل سازی اجزای مختلف MGMS در حالت کارکرد جزیره ای می پردازد.

۱-۲- واحد تخمین حالت^۴

امروزه با توجه به توسعه اتوماسیون در شبکه های توزیع، سیستم های SCADA^۵ در بسیاری از شبکه های توزیع نصب شده اند که می توانند دامنه ولتاژ، توان اکتیو، توان راکتیو و سایر مقادیر در یک خط یا یک شین از شبکه را اندازه گیری نموده و به مرکز کنترل انتقال دهند [۲۵] و [۲۶]. در سیستمهای مدرن توزیع، تخمین حالت نقش حساسی در تخمین حالت زمان واقعی سیستم (که نمی تواند از اندازه گیریهای محدود و گاهی کم اطمینان در سیستم توزیع بدست آیند) دارد. تخمین حالت سیستم توزیع می تواند تلفات توان، ولتاژ شین ها، مقدار تولید، توان بارها، توان خطوط توزیع و غیره را محاسبه نماید [۲۷].

بخاطر محدود بودن تعداد اندازه گیریهای واقعی در سیستم توزیع با تعداد بسیار زیاد گره ها، تخمین حالت در سیستم توزیع مشکلتر است. بنابراین اندازه گیریهای مجازی برای تخمین حالت در سیستم توزیع مورد نیاز می باشد که می تواند تخمین و پیش بینی بار در نقاط مختلف با استفاده از اطلاعات گذشته و روشهای مخصوص باشد. مدل سازی بار می تواند تخمین پروفیل بار مشتریان را بصورت زمان واقعی انجام دهد و بعنوان اندازه گیری مجازی در تخمین حالت مورد استفاده قرار گیرد [۲۶] و [۲۷]. در این مقاله از الگوریتم معمول بکار گرفته شده در سیستم های قدرت سنتی، بر اساس فرمولاسیون حداقل مربعات وزن دار شده [۲۶] و [۲۷]، مطابق رابطه (۱) استفاده شده است.

$$\text{Min } J(x) = \sum_{i=1}^M \frac{[z_i - h_i(x)]^2}{\sigma_i^2} \quad (1)$$

که:

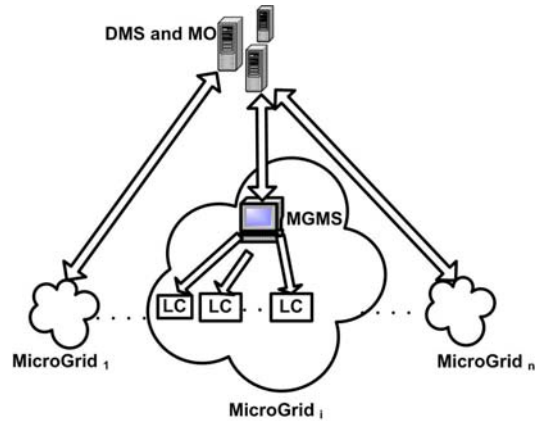
x : بردار متغیرهای حالت (دامنه و فاز ولتاژ شین ها)

M : تعداد اندازه گیری های واقعی و اندازه گیری های مجازی

z_i : مقدار اندازه گیری شده i ام

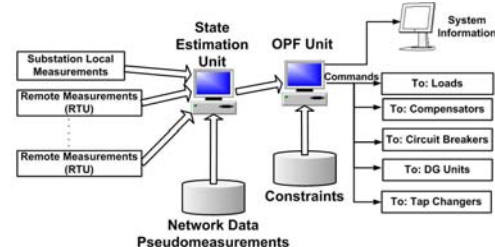
h_i : عبارت اندازه گیری ها بصورت تابعی از متغیرهای حالت

- سیستم مدیریت MicroGrid^۱ (MGMS) که در هر یک از MicroGrid ها وجود دارد.
- کنترل محلی^۲ (LC) که در برخی بخشهای کنترل پذیر سیستم توزیع نصب می شوند.



شکل (۱): مدیریت فعال یک سیستم توزیع شامل چند MicroGrid

در این ساختار وسایل مخابراتی ارتباطی قابل اطمینان را بین همه سطوح کنترلی فراهم می نماید. با پیشرفتهای اخیر در زمینه فناوری مخابرات، استفاده از این ساختار کاربردی است و اتصال آنها به اندازه کافی قابل اطمینان است. همچنین پروژه های مختلفی در سراسر جهان وجود دارد که مسائل فنی و عملی بکارگیری مدیریت فعال در شبکه های توزیع را بررسی می کنند و این سیستم توسط شرکتهای مختلف گسترش یافته و در برخی کشورهای اروپایی مثل آلمان و اسپانیا بصورت آزمایشی بکار گرفته شده است [۲۱]-[۲۴]. به نظر می رسد در آینده ای نزدیک این سیستم در همه شبکه های توزیع جا باز خواهد نمود. شکل (۲) یک ساختار ممکن برای MGMS را نشان می دهد [۲۱].



شکل (۲): ساختار سیستم مدیریت MicroGrid (MGMS)

در طراحی MicroGrid باید توجه ویژه ای به سیستم مدیریت آن (MGMS) شود. این سیستم اطلاعات نیازهای حرارتی و الکتریکی محلی، نیازهای کیفیت توان، قیمت برق و سوخت و غیره را بکار

³ Stand-Alone

⁴ State Estimation

⁵ Supervisory Control And Data Acquisition

¹ MicroGrid Management System

² Local Control

این مقاله تنها بر روی واحدهای تولید پراکنده که از طریق اینورتر به MicroGrid متصل شده اند تمرکز کرده و سایر تجهیزات در نظر گرفته نمی شوند. همچنین نوع و دینامیک تولید کننده که می تواند پیل سوختی، توربین بادی، سلول خورشیدی و ... باشد نیز در عملکرد سیستم مؤثر می باشد که در این مقاله تولید کننده به همراه مبدل DC-DC و ذخیره ساز، بصورت یک منبع DC ثابت در نظر گرفته می شوند. لذا روشی که در این مقاله مورد بحث قرار می گیرد برای انواع منابع تولید پراکنده صادق خواهد بود. با این فرضیات کنترل اینورتر اتصال، عامل اصلی در اجرای فرمانها می باشد و در این قسمت مورد بررسی قرار می گیرد.

توان اکتیو و راکتیو تغذیه شده توسط یک منبع DG متصل شده با اینورتر می تواند بصورت مستقل با پارامترهای اینورتر کنترل شوند. روشهای کنترلی PQ (که در آن اینورتر مقدار مرجع توانهای اکتیو و راکتیو را به سیستم تزریق می کند) و VF (که در آن اینورتر وظیفه تولید ولتاژ با دامنه و فرکانس مورد نظر را به عهده دارد) طرحهای مشترکی هستند که برای کنترل واحدهای تولید پراکنده متصل شده با اینورتر مورد استفاده قرار می گیرند [۱۲]-[۱۵] و [۲۹]. در این مطالعه آنچنانکه در بخشهای بعد تشریح خواهد شد از هر دو نوع کنترل استفاده خواهد شد.

۳- سیستم مدیریت MicroGrid مبتنی بر شبکه عصبی (NN-MGMS)

کارکرد بهینه یک MicroGrid بصورت on-line بخاطر طبیعت متغیر بارها و واحدهای تولید پراکنده و همچنین زمان محاسباتی زیاد بسادگی قابل بکارگیری نمی باشند [۳۰]. مفهوم شبکه های عصبی بخاطر توانایی محاسباتی بالا، پیچیدگی کم، راحتی بکارگیری و زمان کم مورد نیاز برای پاسخ می تواند این محدودیتها را از بین ببرد. از سوی دیگر شبکه های عصبی نیاز به آموزش مناسب دارند و لازم است که داده های مناسب لازم برای آموزش آن ایجاد شوند. با وجود این پروسه آموزش شبکه عصبی یک روال off-line است که بصورت دوره ای برای تست عملکرد آن و یا وقتی که تغییرات اساسی در مشخصات MicroGrid (نظیر بارها یا منابع تولید و یا ساختار شبکه) اتفاق می افتد انجام می گیرد. در این مطالعه طراحی سیستم مدیریت MicroGrid با استفاده از شبکه عصبی بر اساس قدمهای زیرین انجام گردیده است.

۳-۱- استراتژی کنترل MicroGrid مستقل

در یک سیستم توزیع مستقل دو مسئله مهم وجود دارد. اولین مسئله وجود منابع تولید بدون اینرسی و با پاسخ دهی کند است که این مسئله باید با نصب ذخیره سازهای انرژی در سمت DC اینورترها مرتفع گردد تا تغذیه بار بصورت سریع انجام شود. مسئله دوم عدم وجود مرجع ولتاژ

واریانس توزیع اندازه گیری i

تابع h خطی نیست و لازم است که روش نیوتن تکرار شونده برای حل آن استفاده شود. بصورت واضح تر گرادینان تابع J محاسبه شده و با استفاده از روش نیوتون مساوی صفر قرار داده می شود. این عمل موجب ایجاد معادله تکرار شونده (۲) که باید حل شود می گردد:

$$\begin{aligned} x_{k+1} &= x_k + G(x_k)H^T(x_k)W[z - h(x_k)] \\ G(x_k) &= [H^T(x_k)WH(x_k)]^{-1} \\ H(x) &= \frac{\partial h(x)}{\partial x} \end{aligned} \quad (2)$$

واحد تخمین حالت از اندازه گیریهای واقعی سیستم، بعضی اندازه گیریهای مجازی و ساختار شبکه نظیر وضعیت کلیدها استفاده می کند و سپس حالت فعلی سیستم را محاسبه می کند. خروجی تخمینگر حالت مقادیر ولتاژ شین ها، مقادیر بارها، تلفات خطوط، مقادیر تولید و توان خطوط می باشد.

۲-۲- واحد پخش بار بهینه^۱

هدف این واحد بدست آوردن مقادیر مرجع توانی است که باید توسط هر یک از واحدهای تولید پراکنده به منظور کارکرد بهینه و با در نظر گرفتن همه قیود به شبکه تزریق گردد. برای رسیدن به این هدف رابطه ریاضی (۳) باید حل شود تا مقادیر بهینه بدست آیند:

$$\begin{aligned} \text{Minimize } P_{Loss} &= f(|v|, \delta) \\ \text{Subject to:} \\ \text{Load Flow Equation (it mean } \sum_{i=1}^{ngen} P_{G_i} &= \sum_{i=1}^{nload} P_{D_i} + \text{Losses)} \\ V_i^{\min} \leq V_i &\leq V_i^{\max} \quad \text{for } i = 1 \text{ to no. of Buses} \\ P_{G_i}^{\min} \leq P_{G_i} &\leq P_{G_i}^{\max} \quad \text{for } i = 1 \text{ to no. of DG units} \\ Q_{G_i}^{\min} \leq Q_{G_i} &\leq Q_{G_i}^{\max} \quad \text{for } i = 1 \text{ to no. of DG units} \\ S_{Line-i} &< S_{Max-i} \quad \text{for } i = 1 \text{ to no. of Lines} \end{aligned} \quad (3)$$

در این رابطه تلفات توان، بصورت تابعی از دامنه و زاویه فاز ولتاژها می باشد که در ضمیمه آمده است.

۲-۳- کنترل اینورترهای اتصال منابع تولید پراکنده

همانگونه که ساختار شکل (۲) نشان می دهد، در یک MicroGrid عوامل کنترلی مختلفی نظیر مقادیر توان اکتیو و راکتیو واحدهای تولید پراکنده، موقعیت تپ ترانسفورماتورها، وضعیت کلیدهای قدرت، مقادیر جبران سازهای توان راکتیو و همچنین فرمانهایی برای بارزدایی وجود دارند که هر یک می توانند در عملکرد سیستم نقش داشته باشند.

¹ Optimal Power Flow (OPF)

حالت سیستم را تخمین زده و مقادیر زده و ولتاژها و همچنین ولتاژ شین ها و جریان خطوط را محاسبه می نماید.

قدم بعدی تقسیم توان بین سایر واحدها با در نظر گرفتن مقادیر نامی آنها و هدایت سیستم به کارکرد بهینه و نرمال است. این کار با حل معادله پخش بار بهینه (توسط مقادیر تخمین زده شده و رابطه (۳)) می باشد. این کار چون زمانبر است توسط واحد ANN که با استفاده از داده های بدست آمده توسط الگوریتم ژنتیک آموزش داده شده است بصورت زمان واقعی انجام می شود. واحد ANN به این تغییرات پاسخ داده و فرمانهای بهینه را صادر می نماید و ارتباط های مخابراتی این مقادیر جدید مرجع توان را به منابع تولید پراکنده انتقال می دهند.

واضح است که واحد اصلی باید قادر به تولید توان اضافی در مدت زمان تغییر بار تا رسیدن مقادیر مرجع جدید به سایر منابع تولید پراکنده باشد. از این رو ظرفیت وسیله ذخیره ساز باید با توجه به حداکثر تغییرات بار، تأخیر کنترل کننده ها و سرعت تغییر خروجی واحدهای تولید باشد. در مراجع [۳۱] و [۳۲] نحوه محاسبه میزان ذخیره ساز مناسب و نحوه طراحی آن شرح داده شده است.

۲-۳- تعیین ورودی ها و خروجی های شبکه

عصبی

ورودیهای شبکه عصبی باید بگونه ای انتخاب شوند که بتوانند رفتار سیستم را توصیف نمایند و تا حد امکان کم باشند تا از تعداد پارامتر زیاد شبکه عصبی جلوگیری شود. همچنین ورودیهای شبکه عصبی باید شامل پارامترهای قابل اندازه گیری و قابل کنترل سیستم باشند. در این مقاله ورودیهای شبکه عصبی توانهای اکتیو و راکتیو بارها که از واحد تخمین حالت می آیند و ساختار شبکه (یعنی اینکه بعضی از خطوط برای تعمیرات خارج از سرویس باشند) هستند.

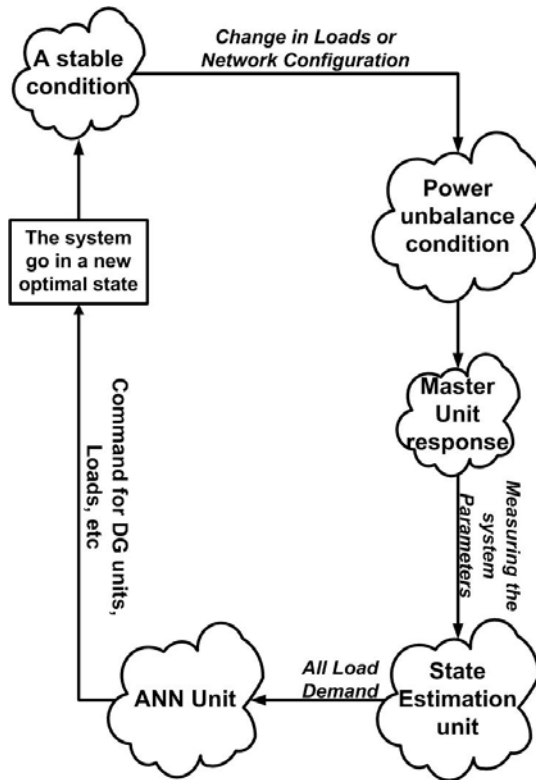
خروجی شبکه عصبی مقادیر مرجع توانهای اکتیو و راکتیو منابع تولید پراکنده غیر اصلی می باشد. زیرا همانگونه که در استراتژی کنترلی ذکر شد، واحد اصلی وظیفه جبران اختلافات توان بین مقدار تولید و مصرف که می تواند ناشی از خطاها و ... باشد را بعهده دارد و توسط کنترل کننده های محلی تنظیم لازمه انجام می پذیرد.

۳-۳- تولید داده های آموزش

تولید داده های آموزش ایجاد نمونه هایی از داده هاست که نشان دهنده رفتار سیستم در هنگام تغییر حالت آن است. در این مطالعه این کار بصورت زیر انجام می شود:

- در نظر گرفتن مقادیر متفاوت برای بارها. در این مطالعه برای کاهش تعداد داده ها، سه سطح بار (حداقل، متوسط و حداکثر) برای هر شین در نظر گرفته می شود و ترکیب بارها بصورت یکنواخت از بین این سطوح بار انتخاب می شود.
- در نظر گرفتن ساختارهای مختلف برای شبکه توزیع.
- حل روش بهینه سازی بر اساس الگوریتم ژنتیک برای بدست

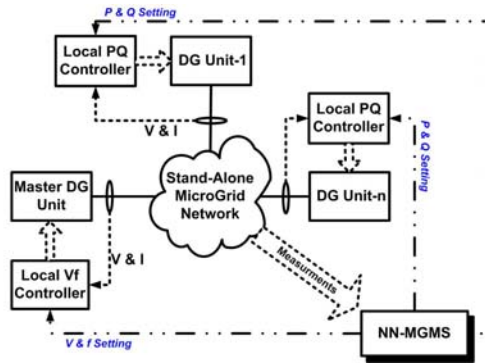
و فرکانس است که یک یا چند منبع تولید پراکنده باید بعنوان مرجع فرکانس و ولتاژ عمل نمایند. این چنین واحد بعنوان واحد اصلی^۱ در این مقاله نامیده می شود و با نصب مقدار مناسب ذخیره ساز در سمت DC اینورتر آن پاسخگویی سریع به تغییرات توان و ولتاژ پایدار را بعهده دارد. سایر منابع تولید کننده پراکنده خود را با این مرجع هماهنگ نموده و در حالت کنترل توان ثابت (مود PQ) کار می کنند و در تنظیم پایدار بار- تولید شرکت می نمایند. بر اساس استراتژی اشاره شده، روش ارائه شده برای کنترل تقسیم بار بین واحدهای تولید پراکنده در این مقاله مطابق فلوچارت شکل (۳) است.



شکل (۳): دیاگرام عملکرد سیستم مدیریت

در ابتدا فرض می شود که همه واحدها در یک سیستم مستقل زیر مقدار نامی خود کار می کنند تا اینکه بتوانند تغییرات توان را داشته باشند. بعد از هر گونه تغییر در سیستم، ابتدا واحد اصلی پاسخ خواهد داد و توان جذب یا تزریق می نماید تا اینکه نیاز بار بر طرف شود و در حالیکه مابقی واحدها به تغذیه همان مقادیر قبلی برنامه ریزی شده ادامه می دهند. در فاصله های زمانی ثابت، واحدهای اندازه گیری پارامترهای سیستم را اندازه گیری نموده و آنها را به MGMS می فرستند. واحد تخمین حالت با استفاده از این اندازه گیری ها و اندازه گیریهای مجازی

¹ Master Unit



شکل (۵): نمای کلی استراتژی کنترلی ارائه شده

۴- سیستم مورد مطالعه و شبیه سازی

استراتژی کنترل ارائه شده بر روی سیستم توزیع نمونه که در شکل (۶) نشان داده شده است آزمایش شده است. سیستم شامل چهار فیدر است که توسط یک ترانسفورماتور ۲۰/۰.۴ کیلو ولت از شبکه قدرت تغذیه می شود و اطلاعات شبکه در ضمیمه ۲ آمده است. باید توجه شود که بارها بصورت توان ثابت با مقادیر اولیه اشاره شده در ضمیمه مدل شده اند. قسمت مهم شبکه شامل ۹ شین، ۸ بار در دو فیدر شعاعی می باشد. وقتی شبکه سراسری وصل است، دو واحد تولید پراکنده (DG 1) و (DG 2) قسمتی از انرژی مورد نیاز را جهت کاهش برق مصرفی مشتریان تولید می نمایند. هنگامیکه خطا در شبکه رخ دهد یا اینکه سیستم از شبکه قدرت از طریق کلید اصلی (Main Switch) بصورت از پیش برنامه ریزی شده قطع شود، سیستم بصورت مستقل در می آید. در این مطالعه تنها بر روی حالت جزیره ای شبکه تمرکز کرده و گذر از حالت اتصال شبکه به جزیره ای در نظر گرفته نمی شود.

برای ساده سازی دینامیک واحدهای تولید پراکنده در نظر گرفته نشده و ولتاژ DC ثابت در همه اینورها در نظر گرفته می شود. همچنین DG 1 با ذخیره ساز مناسب واحد اصلی است و مانند یک ژنراتور سنکرون برای تولید مرجع ولتاژ و فرکانس مورد استفاده قرار می گیرد. واحد DG 2 با استفاده از PLL خود را با واحد اصلی همزمان نموده و در حالت کنترل توان ثابت (PQ) کار می کند و مقادیر مرجع توان آن از طریق فرمانهای سیستم مدیریت MicroGrid تغییر می نماید. داده های واحدهای تولید پراکنده نیز در ضمیمه آمده است.

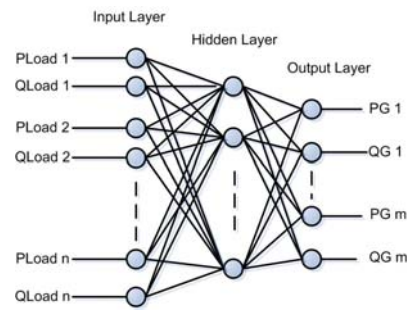
برای تخمین حالت، سه اندازه گیر ولتاژ در شینهای ۱، ۲ و ۶ و دو اندازه گیر توان اکتیو و راکتیو در شینهای شامل منابع تولید پراکنده (شین های ۱ و ۲) نصب شده اند. همچنین پیش بینی بارهای برخی شینها بعنوان داده های اندازه گیری مجازی برای تخمین حالت سیستم مورد استفاده قرار گرفته است.

آوردن مقادیر بهینه برای هر یک از منابع تولید پراکنده به منظور کمینه سازی تلفات توان و برآورده سازی قیود. بعد از این قدمها ورودی و خروجی سیستم کنترل مشخص شده و می تواند برای آموزش شبکه عصبی مورد استفاده قرار گیرند.

۳-۴ طرح شبکه عصبی

در این مطالعه شبکه عصبی feed-forward با یک لایه مخفی مورد استفاده قرار گرفته است و توجه زیادی به کاهش تعداد ورودیها شده است. لازم بذکر است که تعداد نرونها مخفی باید با دقت انتخاب شده تا اینکه قابلیت شبکه عصبی بالا باشد (قابلیت تولید پاسخهای مناسب برای ورودیهایی که آموزش دیده نمی شوند).

بر اساس ورودیها و خروجیهای شرح داده شده در بخش ۴-۲ ساختار شبکه عصبی مورد استفاده را می توان بصورت شکل (۴) نشان داد. لایه ورودی شامل توانهای اکتیو و راکتیو بارها و لایه خروجی توانهای تولیدی واحدهای تولید پراکنده به جز واحد اصلی می باشد.



شکل (۴): ساختار شبکه عصبی مورد استفاده

۳-۵ آموزش شبکه عصبی

جعبه ابزار شبکه عصبی نرم افزار MATLAB جهت ایجاد، آموزش و ارزیابی شبکه عصبی مورد استفاده قرار گرفته است. الگوریتم BP برای آموزش شبکه عصبی با استفاده از پارامترهای آموزش مناسب که با استفاده از محاسبات پخش بار بهینه بدست آمده اند مورد استفاده قرار گرفته است. بعد از مرحله آموزش، عملکرد شبکه عصبی با بقیه داده هایی که از حل معادلات با الگوریتم ژنتیک برای یک مجموعه جدید بارها بدست آمده، ارزیابی می شود.

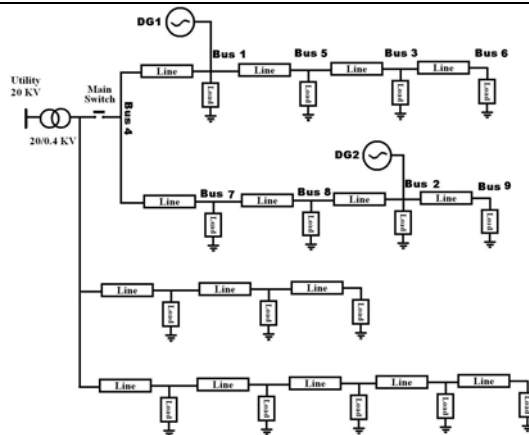
۳-۶ بکارگیری شبکه عصبی در کنترل کننده on-Line

همه قدمهای قبلی باید بصورت off-Line انجام گیرد. محصول نهایی این دستورالعمل باید در یک فضای on-line در سیستم مدیریت MicroGrid بکار گرفته شود. شکل زیرین نمای کلی سیستم MicroGrid مستقل به همراه استراتژی کنترلی را نمایش می دهد.

جدول ۱ تعدادی از این داده ها که توسط الگوریتم ژنتیک بدست آمده است را نشان می دهد.

سیس تعدادی از داده های بدست آمده جهت آموزش شبکه عصبی مورد استفاده قرار می گیرند. همانگونه که قبلاً بیان شد از آنجائیکه سیستم ۸ شین بار دارد، شبکه عصبی ۱۶ ورودی (توانهای اکتیو و راکتیو هر شین) و همچنین ۲ خروجی (توانهای اکتیو و راکتیو DG 2) دارد. ساختارهای متفاوت برای شبکه عصبی آزمایش شده تا بهترین عملکرد بدست آید. بهترین دقت از یک شبکه سه لایه FF با چهار نرون در لایه مخفی با تابع tan-sigmoid و ۲ نرون با تابع تبدیل خطی در لایه خروجی بدست آمده است.

جدول ۲ نتایج شبیه سازی شبکه عصبی و الگوریتم ژنتیک را برای برخی از حالت‌هایی که در آموزش مورد استفاده قرار نگرفته است نشان می دهد. همانگونه که ملاحظه می گردد نتایج بسیار نزدیک هم بوده و شبکه عصبی می تواند برای تعیین سیگنال کنترلی بصورت on-line مورد استفاده قرار گیرد.



شکل (۶): دیاگرام تک خطی سیستم مورد مطالعه

همانگونه که در بخش قبلی بیان شد، داده های آموزش با تغییر بارها بین سه سطح حداقل، متوسط و حداکثر بار در هر شین ایجاد شده و الگوریتم ژنتیک سطوح تولید مناسب برای واحد DG 2 را معین می کند. این سه سطح به اندازه ۶۰٪، ۱۰۰٪ و ۱۲۰٪ مقادیر داده شده در ضمیمه در نظر گرفته شده اند.

در این مسئله الگوریتم ژنتیک برای بهینه سازی مقادیر توان اکتیو و راکتیو واحدهای تولید پراکنده در سیستم انجام می شود و هدف این است که تلفات توان اکتیو (تابع هدف) حداقل شده و قیود سیستم نیز ارضاء شوند. لذا در اینجا هر کروموزوم بصورت شکل (۷) شامل توانهای اکتیو و راکتیو منابع DG انتخاب می شود [۲۸]:

P DG 1	Q DG 1	...	P DG m	Q DG m
--------	--------	-----	--------	--------

شکل (۷): ساختار کروموزومهای مورد استفاده

با توجه به اینکه روش کد گذاری ارائه شده دودویی بوده، هر یک از مقادیر توان با توجه به حداقل و حداکثر آنها نرمالیزه و بصورت رشته هایی از صفر و یک تبدیل می شوند. سپس بعد از اتمام بهینه سازی با توجه به مقدار نامی توان هر یک از منابع مقدار واقعی متغیرها بدست می آید.

در این مقاله عملگر انتخاب بر اساس روش چرخ رولت^۱ و بر اساس میزان برازش هر کروموزوم انجام می گردد. برازندگی هر کروموزوم معکوس تابع هدف آن می باشد که در رابطه (۳) مشخص شده است. از عملگر تقاطع دو نقطه ای^۲ با احتمال جهش^۳ برابر با ۰/۲٪ برای هر کروموزوم استفاده شده است. این عملگر دو نقطه ای جهت جلوگیری از انتخاب جوابهای محلی بهینه می باشد.

¹ Roulette Wheel

² Two cut-point crossover

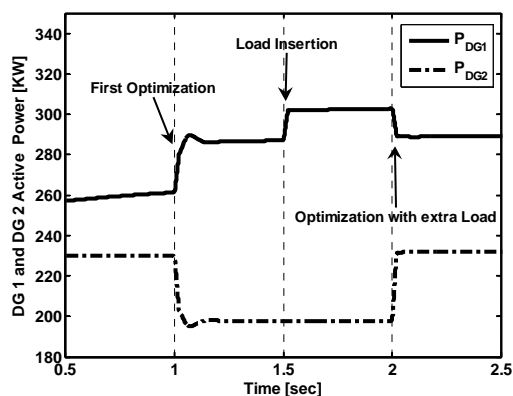
³ Mutation Probability

جدول ۱: تعدادی از داده های بدست آمده از الگوریتم ژنتیک

Set	Load Level for each Bus								PG1 (kW)	PG2 (kW)	QG1 (kVar)	QG2 (kVar)	PLoss (kW)
	1	2	3	5	6	7	8	9					
1	0.60	1.00	1.20	1.20	1.00	0.60	1.00	1.20	260.47	229.91	101.86	84.65	14.29
2	1.20	1.20	1.00	0.60	1.20	1.20	1.00	0.60	275.27	182.20	86.11	77.47	9.33
3	1.20	0.60	1.20	0.60	1.00	0.60	0.60	0.60	236.12	148.62	120.18	16.00	7.51
4	0.60	1.00	0.60	1.00	1.00	1.20	1.00	1.20	223.18	255.24	113.36	66.46	12.57
5	1.00	0.60	1.00	1.20	1.00	1.00	1.00	1.20	294.73	222.13	101.89	92.14	13.97
6	0.60	0.60	0.60	1.20	0.60	1.00	1.20	1.20	224.38	234.56	77.49	101.60	10.78
7	0.60	0.60	1.00	1.00	0.60	0.60	1.20	1.20	209.79	226.51	67.79	104.58	10.45
8	1.00	1.20	0.60	1.00	0.60	1.20	1.20	1.00	244.28	236.12	79.13	103.55	8.66
9	1.20	1.20	0.60	0.60	0.60	1.00	1.00	0.60	219.27	180.37	67.81	79.01	4.65
10	1.00	1.00	0.60	1.20	0.60	0.60	1.20	0.60	254.43	174.39	86.16	75.59	6.28

جدول ۲: مقایسه نتایج شبکه آموزش داده شده و الگوریتم ژنتیک برای برخی از داده ها

Set	Load Level for each Bus								PG2(kW) NN	PG2(kW) GA	QG2(kVar) NN	QG2(kVar) GA
	1	2	3	5	6	7	8	9				
1	1.00	1.00	0.60	1.00	1.20	0.60	1.00	0.60	145.36	144.59	62.31	61.55
2	0.60	1.00	1.00	1.00	1.00	0.60	1.20	1.00	208.03	208.05	97.04	92.92
3	1.00	1.20	0.60	1.00	1.00	0.60	0.60	1.00	192.78	193.56	71.73	73.17
4	0.60	1.20	1.20	1.00	1.00	1.20	1.00	1.20	252.76	255.18	95.10	90.30
5	0.60	1.00	1.00	1.00	1.00	0.60	1.20	1.00	208.03	208.05	97.04	92.92



شکل (۸): توان اکتیو واحدهای DG 1 و DG 2

پروسه شبیه سازی در چهار حالت مختلف انجام شده و در زیر بخشهای زیرین شرح داده خواهند شد.

۴-۱- راه اندازی سیستم

ابتدا فرض می شود که سیستم با مقادیر مرجع پخش بار شروع به کار کند. مقادیر مرجع توان DG 2 در ۲۳۰ کیلووات و ۶۰ کیلووات تنظیم شده اند. همانگونه که از شکل (۸) و شکل (۹) دیده می شود در این حالت کنترل کننده ها خروجی DG 2 را در مقادیر مطلوب تنظیم نموده اند و همانطور که انتظار می رود مابقی توان توسط واحد اصلی - DG 1 - تأمین می گردد.

کل سیستم را به کار در حداقل تلفات توان و ولتاژ نرمال (بین ۰.۹۵ تا ۱.۰۵ پرینت) و مقدار نامی واحدهای تولید هدایت می کند.

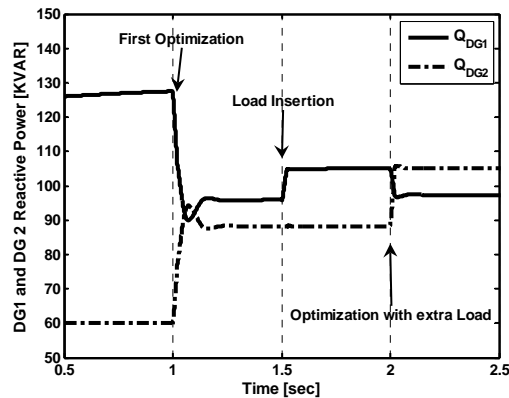
۳-۴- اضافه شدن بار

در سومین حالت شبیه سازی یک اغتشاش مثل اضافه شدن یک بار به سیستم اعمال می شود. در زمان ۱.۵ ثانیه یک بار ۴۰ کیلووات و ۲۰ کیلووار در شین ۹ اضافه می شود و همانگونه که در شکل (۸) و شکل (۹) نشان داده شده است توان اضافه شده توسط واحد اصلی DG 1 تأمین شده و واحد دیگر در مقادیر قبلی مرجع کار می کند. در این حالت همانگونه که شکل (۱۰) نشان می دهد ولتاژ بعضی شین ها از حداقل (۰.۹۵) تجاوز نموده و همچنین تلفات توان به ۰.۳۲ پرینت افزایش می یابد.

۴-۴- بهینه سازی سیستم با بار اضافی

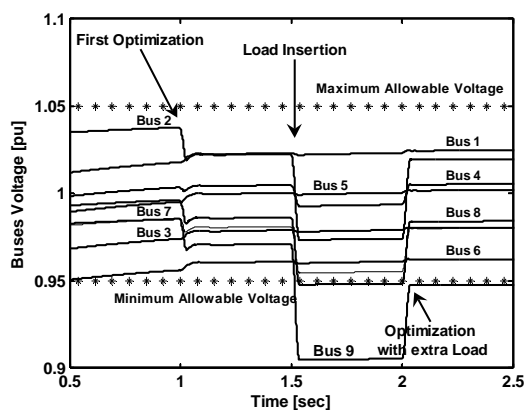
مجدداً در زمان $t=2$ ثانیه سیستم مدیریت MicroGrid پارامترهای سیستم را اندازه گیری کرده و حالت سیستم را تخمین می زند و لذا مقادیر توان مصرفی معلوم می شوند. واحد شبکه عصبی اطلاعات بدست آمده از واحد تخمین حالت را مورد استفاده قرار داده و مقادیر مرجع توان برای واحدهای تولید پراکنده تعیین می کند تا اینکه تلفات توان را کاهش دهد و همچنین سیستم را به داشتن ولتاژ نرمال هدایت نماید. نتایج کنترل کننده شبکه عصبی برای توان مرجع DG 2 در این حالت ۲۳۱ کیلووات و ۱۰۵ کیلووار است. بعد از انتقال این سیگنالها به DG 2، سیستم این توانها را تولید نموده و همانگونه که شکل (۱۰) نشان می دهد ولتاژ شین ها به مقدار نرمال برمی گردد و همچنین تلفات توان به ۰.۲۸ پرینت رسیده و واضح است که به اندازه ۱۰٪ کاهش یافته است.

همانگونه که از نتایج شبیه سازی پیداست، روش ارائه شده برای کنترل منابع تولید پراکنده مؤثر است و می تواند به نیازهای شبکه در هنگام جزیره ای شدن از شبکه سراسری پاسخ دهد و این طرح می تواند پیوستگی توان بارها را تضمین نماید. هر تغییر در سیستم با استفاده از اندازه گیری های واقعی، سابقه سیستم و واحد تخمین حالت معین می شود. بر اساس مقادیر تخمین زده شده توسط شبکه عصبی فرمانهای مناسب تولید شده و فرستادن این فرمانها به کنترل کننده های محلی، منابع تولید پراکنده را به حالت بهینه هدایت نموده و سیستم در حالت بهینه و نرمال کار خواهد کرد. لازم به ذکر است که اضافه کردن استراتژی بارزدایی (یعنی حذف بار اضافی در صورت کافی نبودن توان تولیدی منابع جهت تأمین بارها به ترتیب اولویت بارهای مهم و غیرمهم) و حفاظت (منظور حفاظت سیستم در برابر خطاها و چگونگی هماهنگی تجهیزات حفاظتی) به روش تقسیم بار ارائه شده می تواند باعث عملکرد کامل سیستم MicroGrid را در یک شبکه جزیره ای شود.



شکل (۹): توان راکتیو واحدهای DG 1 و DG 2

در این حالت مطابق شکل (۱۰) ولتاژها در مقادیر نرمال هستند و کل تلفات توان سیستم ۰.۲۵۶ پرینت است که حدود ۴٪ کل بار سیستم می باشد.



شکل (۱۰): ولتاژ شین های سیستم

۲-۴- بهینه سازی سیستم

در زمان $t=1$ ثانیه واحد تخمین حالا پارامترهای سیستم و مقادیر بار را محاسبه می کند. نتایج تخمین و مقادیر واقعی نشان دهنده دقت خوب واحد تخمین حالت است و مقدار بار و همچنین ولتاژ شین ها بعد از تخمین معلوم خواهند شد. واحد ANN در این حالت مقدار بهینه DG 2 را به اندازه ۱۹۷ کیلووات و ۸۸ کیلووار محاسبه نموده و به واحد می فرستد. از شکل (۸) و شکل (۹) واضح است که کنترل کننده محلی فرمانها را اجرا نموده و توان را در مقادیر دلخواه تنظیم می نماید. با این مقادیر بهینه، تلفات سیستم به ۰.۲۱۲۰ پرینت تغییر یافته که مشخص است که به اندازه ۲۴٪ کاهش یافته است و همچنین مطابق شکل (۱۰) ولتاژها در مقادیر نرمال باقی مانده اند.

همانگونه که در هنگام آموزش شبکه عصبی انتظار می رفت، مقادیر محاسبه شده و ارسال شده مرجع جدید برای واحدهای تولید پراکنده

۵- نتیجه گیری

این مقاله یک استراتژی مدیریت و کنترل برای کارکرد قابل اطمینان و بهینه یک MicroGrid در حالت مستقل ارائه نمود. طرح ارائه شده از تجهیزات مخابراتی برای ارسال اطلاعات به کنترل کننده مرکزی استفاده می کند. مقادیر بهینه تولید برای هر یک از واحدهای تولید پراکنده با استفاده از روش بهینه سازی الگوریتم ژنتیک برای چندین حالت بار و ساختار شبکه بدست آمده و شبکه عصبی بصورت Off-line با این اطلاعات آموزش داده می شود. سرانجام شبکه عصبی در سیستم کنترل on-line برای تولید فرمانهای کنترلی سیستم بلافاصله بعد از هر تغییر مورد استفاده قرار می گیرد.

همانگونه که نتایج شبیه سازی روی یک MicroGrid نمونه در این مقاله نشان می دهد، وقتی که در قسمتی از شبکه توزیع شامل چندین منبع تولید پراکنده، شبکه قدرت وجود ندارد طرح ارائه شده می تواند سیستم را در این حالت کنترل نموده تا واحدهای تولید پراکنده موجود بارها را تأمین نموده و پیوستگی توان بارها را تضمین نماید.

۶- مراجع

- [7] Fang Gao, and M. Reza Irvani, "A Control Strategy for a Distributed Generation Unit in Grid-Connected and Autonomous Modes of Operation", IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 23, NO. 2, APRIL 2008, Pages: 850-859
- [8] Houshang Karimi, Hassan Nikkhajoei, and Reza Irvani, "Control of an Electronically-Coupled Distributed Resource Unit Subsequent to an Islanding Event", IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 23, NO. 1, JANUARY 2008, Pages: 493-501
- [9] Charles K. Sao, Peter W. Lehn, "Autonomous Load Sharing of Voltage Source Converters", IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 20, NO. 2, APRIL 2005, Pages: 1009-1016
- [10] Yunwei Li, D. Mahinda Vilathgamuwa, and Poh Chiang Loh, "Design, Analysis, and Real-Time Testing of a Controller for Multibus Microgrid System", IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 19, NO. 5, SEPTEMBER 2004, Pages: 1195-1204
- [11] J. A. Peças Lopes, C. L. Moreira, and A. G. Madureira "Defining Control Strategies for MicroGrids Islanded Operation", IEEE Transactions on Power Systems, May 2006, vol. 21, no. 2, Pages: 916-924
- [12] F. Katiraei, and M. R. Irvani, "Power Management Strategies for a Microgrid With Multiple Distributed Generation Units", IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 21, NO. 4, NOVEMBER 2006, Pages: 1821-1831
- [13] F. Katiraei, M. R. Irvani, and P. W. Lehn, "Micro-Grid Autonomous Operation During and Subsequent to Islanding Process", IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 20, NO. 1, JANUARY 2005, Pages: 248-257
- [14] A. Engler, "Control of parallel operating battery inverters", In proceeding of the Photovoltaic Hybrid Power Systems Conference, 2000, Aix-en-Provence
- [15] Nick Hiscock, Terence G. Hazel, and Jonathan Hiscock, "Voltage Regulation at Sites With Distributed Generation", IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS, VOL. 44, NO. 2, MARCH/APRIL 2008, Pages: 445-454
- [17] Ali Asghar Ghadimi, Hassan Rastegar, "Control of a Stand-Alone MicroGrid with Active Management Scheme," Journal of
- [1] A.A. Ghadimi, A. M. Daryani, H. Rastegar, "Detailed Modeling and analysis of a full bridge PWM DC-DC converter", In proceedings of the Australasian Universities Power Engineering Conference (AUPEC06), December 2006, Melbourne, Australia
- [2] R. H. Lasseter and P. Piagi, "Microgrid: A conceptual solution", In proceeding of the Power Electronics Specialists Conference, Aachen, Germany, June 2004, vol. 6, pages: 4285-4290
- [3] M. N. Marwali, J. W. Jung, and A. Keyhani, "Control of Distributed Generation Systems, Part II: Load Sharing Control", IEEE Transaction on Power Electronics, VOL. 19, NO. 6, NOVEMBER 2004, Pages: 1551-1561
- [4] P. Piagi, R. Lasseter, "Autonomous control of Microgrids", in proceeding of the IEEE PES General Meeting, 18-22 June 2006
- [5] Min Dai, Mohammad Nanda Marwali, Jin-Woo Jung, and Ali Keyhani, "A Three-Phase Four-Wire Inverter Control Technique for a Single Distributed Generation Unit in Island Mode", IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 23, NO. 1, JANUARY 2008, Pages: 322-331
- [6] Hiren Patel and Vivek Agarwal, "Control of a Stand-Alone Inverter-Based Distributed Generation Source for Voltage Regulation and Harmonic Compensation", IEEE TRANSACTIONS ON POWER DELIVERY, VOL. 23, NO. 2, APRIL 2008, Pages: 1113-

- Islanding Operation”, in proceeding of the IEEE PES Power Systems Conference and Exposition, 2004.
- [29] G. Celli, F. Pilo, G. Pisano, G. G. Soma, “Optimal Participation of a Microgrid to the Energy Market with an Intelligent EMS”, in proceeding of the IPEC 2005 Conference, Singapore, 2005.
- [30] J. W. Jung, and A. Keyhani, “Modeling and Control of Fuel Cell Based Distributed Generation Systems in a standalone Ac Power System”, Journal of Iranian Association of Electrical and Electronics Engineers, Vol. 2, No. 1, 2005, Pages: 10-23
- [31] M. Uzunoglu and M. S. Alam, “Dynamic Modeling, Design, and Simulation of a Combined PEM Fuel Cell and Ultracapacitor System for Stand-Alone Residential Applications”, IEEE TRANSACTIONS ON ENERGY CONVERSION, VOL. 21, NO. 3, SEPTEMBER 2006, Pages: 767-775
- Electrical Engineering, Vol. 8, No. 3, 2008, Pages: 53–60
- [18] S. N. Liew and G. Strbac, “Maximizing penetration of wind generation in existing distribution networks,” Proc. Inst. Elect. Eng., Gen., Transm., Distrib., vol. 149, No. 3, May 2002, Pages: 256–262
- [19] A. Shafiu, T. Bopp, I. Chilvers, and G. Strbac, “Active management and protection of distribution networks with distributed generation”, in proceeding of the IEEE PES General Meeting, Jun. 6–10, 2004, Vol. 1, Pages:1098–1103
- [20] Aris L. Dimeas, Nikos D. Hatziaargyriou, “Operation of a Multiagent System for Microgrid Control”, IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 20, NO. 3, AUGUST 2005, Pages:1447-1455
- [21] B. Buchholz, U. Schluecking, “Energy Management in Distribution Grids: European Cases”, in proceeding of the IEEE Power Engineering Society General Meeting, June 2006
- [22] P. Bresesti, A. Cerretti, “SDNO: Smart Distribution Network Operation Project”, in proceeding of the IEEE Power Engineering Society General Meeting, June 2007
- [23] K. Okuyama, T. Kato, Y. Suzuoki, T. Funabashi, K. Wu, Y. Yokomizu, T. Okamoto, “Improvement of reliability of power distribution system by information exchange between distributed generators-sharing of all DGs information”, in proceeding of the IEEE Power Engineering Society Summer Meeting, 2001
- [24] A. Shafiu, N. Jenkins, and G. Strbac, “Measurement location for state estimation of distribution networks with generation”, IEE Proc.-Gener. Transm. Distrib., Vol. 152, No. 2, March 2005, Pages: 240-246
- [25] Haibin Wang, Noel. N. Schulz, “A Load Modeling Algorithm for Distribution System State Estimation”, IEEE/PES Transmission and Distribution Conference and Exposition, 2001
- [26] Inigo Cobelo, Ahmed Shafiu, Nick Jenkins, Goran Strbac, “State estimation of networks with distributed generation”, EUROPEAN TRANSACTIONS ON ELECTRICAL POWER, Vol 17, 2007, Pages: 21–36
- [27] R. L. Haupt and S. E. Haupt, “Practical Genetic Algorithms”, New York, Wiley, 2004
- [28] A. Bertani, C. Bossi, F. Fornari, S. Massucco, S. Spelta, F. Tivegna, “A Microturbine Generation System for Grid Connected and

ضمیمه

ض ۱: تلفات سیستم توزیع

در یک سیستم n باسه تلفات توان می تواند بصورت رابطه (۱) تعریف شود:

$$P_L + jQ_L = \sum_{i=1}^n S_i = \sum_{i=1}^n V_i I_i^* =$$

$$[V_1 \quad V_2 \quad \dots \quad V_n] \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix}^* = V^T I^* \quad (۱)$$

که V و I بردار ولتاژ و جریان شینها می باشند و رابطه ساده ای با ماتریس سیستم دارند و جایگزینی آن در رابطه قبلی نتیجه می دهد:

$$V = Z_{Bus} I$$

$$\text{Therefore: } P_L + jQ_L = (Z_{Bus} I)^T I^* = I^T Z_{Bus} I^* \quad (۲)$$

با نوشتن جریان شین ها و Z_{Bus} بصورت قسمتهای حقیقی و موهومی:

$$Z_{Bus} = R + jX = \begin{bmatrix} r_{11} & \dots & r_{1n} \\ \vdots & & \vdots \\ r_{n1} & \dots & r_{nn} \end{bmatrix} + j \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nn} \end{bmatrix} \quad (۳)$$

$$I = I_a + jI_r = \begin{bmatrix} I_{a1} \\ \vdots \\ I_{an} \end{bmatrix} + j \begin{bmatrix} I_{r1} \\ \vdots \\ I_{rn} \end{bmatrix}$$

ض ۲: پارامترهای سیستم نمونه

همه خطها شبیه هم هستند با اطلاعات زیر:

Cable Resistance: $R=0.164$ ohms/kmCable Inductance: $L=0.26$ mH/km

اطلاعات خطوط توزیع

From Bus	To Bus	Length (m)
4	1	500
1	5	100
5	3	200
3	6	250
4	7	650
7	8	500
8	2	400
2	9	400

اطلاعات بارها

Bus No.	P (kW)	Q (kVar)
1	80	25
2	36	12
3	25	12
5	90	33
6	60	12
7	36	12
8	56	32
9	96	42

اطلاعات سیستم های تولید پراکنده

DG	Nominal Power (kVA)	Control Mode	P-Gain	I-Gain
DG 1	400	Vf	1.1	3.23
DG 2	280	PQ	6	0.12

با جایگزینی قسمت های حقیقی و موهومی جریانها و ماتریس Zbus در رابطه (۱) نتیجه می دهد:

$$P_L + jQ_L = (I_a + jI_r)^T (R + jX)(I_a - jI_r)$$

$$so:$$

$$P_L = I_a^T R I_a + I_r^T R I_r = \sum_{j=1}^n \sum_{k=1}^n r_{jk} (I_a_j I_a_k + I_r_j I_r_k)$$

به منظور داشتن رابطه تلفات توان بصورت تابعی از توان تزریقی شین ها می توان از رابطه بین توان و جریان بصورت زیر استفاده نمود:

$$P_i + jQ_i = V_i I_i^* = |V_i| (\cos \delta_i + j \sin \delta_i) (I_a_i - jI_r_i)$$

که δ_i زاویه ولتاژ شین i ام نسبت به شین مرجع است. قسمت حقیقی و موهومی جریان از این رابطه قابل بدست آوردن است.

$$I_a_i = \frac{1}{|V_i|} (P_i \cos \delta_i + Q_i \sin \delta_i)$$

$$I_r_i = \frac{1}{|V_i|} (P_i \sin \delta_i - Q_i \cos \delta_i)$$

بنابراین با جایگزینی این روابط در رابطه (۵) تلفات توان از رابطه (۷) بدست می آید:

$$P_L = \sum_{j=1}^n \sum_{k=1}^n [\alpha_{jk} (P_j P_k + Q_j Q_k) + \beta_{jk} (Q_j P_k - P_j Q_k)]$$

where:

$$\alpha_{jk} = \frac{r_{jk}}{|V_j| |V_k|} \cos(\delta_j - \delta_k)$$

$$\beta_{jk} = \frac{r_{jk}}{|V_j| |V_k|} \sin(\delta_j - \delta_k)$$

تقریب تابع ارزش عمل با استفاده از شبکه توابع پایه شعاعی برای یادگیری تقویتی

ولی درهمی^۱، امید محرابی^۲

^۱استادیار دانشکده مهندسی برق و کامپیوتر، دانشگاه یزد، vderhami@yazduni.ac.ir

^۲فارغ‌التحصیل کارشناسی ارشد مهندسی برق- کنترل، omidmehrabi62@yahoo.com، کنترل،

(تاریخ دریافت مقاله ۱۳۸۹/۱۰/۲۶، تاریخ پذیرش مقاله ۱۳۹۰/۲/۲۱)

چکیده: مشکل تنگنای ابعاد، یکی از چالش‌هایی است که کاربرد الگوریتم‌های یادگیری تقویتی گسسته را در مورد مسائل کنترلی واقعی که دارای فضای حالت و عمل بزرگ و یا پیوسته می‌باشند محدود نموده است. ترکیب روش‌های آموزشی گسسته با تقریب زنده‌های تابعی برای حل این مشکل چندی است مورد توجه محققان قرار گرفته است. در همین راستا در این مقاله یک الگوریتم جدید یادگیری تقویتی عصبی (NRL) بر مبنای معماری نقاد- تنها معرفی می‌گردد. الگوریتم مذکور از ترکیب الگوریتم یادگیری سارسا با شبکه عصبی RBF به عنوان یک تقریب زنده‌ی تابعی حاصل شده است و ما آن را "یادگیری سارسای عصبی" (NSL) می‌نامیم. ورودی‌های شبکه جفت حالت و عمل‌های مسئله و خروجی آن تابع ارزش عمل تقریب زده شده می‌باشد. وزن‌های شبکه به صورت بر خط با توجه به الگوریتم ارائه شده تنظیم می‌گردند. به عنوان یک شرط لازم همگرایی ما همچنین وجود نقاط ایستای منطبق بر نقاط ثابت الگوریتم "تکرار تقریب ارزش عمل" برای NSL را اثبات می‌نماییم. نتایج شبیه سازی ارائه شده در مورد مسائل خودرو در کوهستان و آکروبات حاکی از عملکرد مناسب تر روش ارائه شده از لحاظ سرعت آموزش و کیفیت عملکرد می‌باشد.

کلمات کلیدی: یادگیری تقویتی عصبی، معماری نقاد-تنها، شبکه عصبی RBF، یادگیری سارسا، نقاط ایستا

Action Value Function Approximation Based on Radial Basis Function Network for Reinforcement Learning

Vali Derhami, Omid Mehrabi

Abstract: One of the challenges encountered in the application of classical reinforcement learning methods to real-control problems is the curse of dimensionality. In order to overcome this difficulty, hybrid algorithms that combine reinforcement learning with various function approximators have attracted many research interests. In this paper, a novel Neural Reinforcement Learning (NRL) scheme which is based on Sarsa learning and Radial Basis Function (RBF) network is proposed. The RBF network is used to approximate the Action Value Function (AVF) on-line. The inputs of RBF network are state-action pairs of system and its outputs are corresponding approximated AVF. As the necessary condition for the convergence of NSL to the optimal task performance, the existence of stationary points for NSL which coincide with the fixed points of Approximate Action Value Iteration (AAVI) are proved. The validity of the proposed algorithm is tested through simulation examples: mountain car control task, and acrobat problem. Overall results demonstrate that our algorithm can effectively improve convergence speed and the efficiency of experience exploitation.

Keywords: Neural reinforcement learning, Critic-only architecture, RBF neural network, Sarsa, stationary points.

۱- مقدمه

طراحی کنترلگر بهینه برای سیستم‌هایی با دینامیک پیچیده، جایی که داشتن یک تعامل هوشمندانه با یک محیط پویایی که دارای عدم قطعیت^۱ و نایقینی^۲ نیز هست، همواره به صورت کارا و مؤثر توسط روش‌های کنترل کلاسیک قابل انجام نیست. با توجه به اینکه عموماً نمی‌توان مقدار خروجی مطلوب کنترل کننده در این نوع سیستم‌ها را تعیین نمود، استفاده از روش‌های آموزش بدون ناظر برای طراحی و تنظیم پارامترهای کنترلگر ارجحیت دارد [۱]. یادگیری تقویتی^۳ (RL)، یک روش قوی مدرن برای آموزش روی خط استراتژی‌های کنترل از طریق تعامل با محیط است. در این روش سیستم تلاش می‌کند تا تقابلات خود با یک محیط پویا را از طریق سعی و خطا بهینه نماید. ایده بنیادی یادگیری تقویتی بر این اصل استوار است که اگر عملی منجر به بهبود عملکرد گردد میل به انجام آن عمل تشدید یا تقویت می‌گردد [۲]. این روش تنها با استفاده از یک معیار اسکالر راندمان، که سیگنال تقویت یا پاداش نامیده می‌شود، بدون نیاز به سرپرست قادر به آموزش عامل‌ها در محیط‌های پیچیده، ناقطعی و تصادفی^۴ می‌باشد و کنترلگر تنها با توجه به نقد عملکرد کنترل که توسط سیگنال تقویتی بیان می‌گردد، به جای یک خروجی مطلوب داده شده توسط سرپرست، آموزش می‌بیند. قابلیت‌های مذکور باضافه قدرت کاوش^۵ بالا در جهت یافتن پاسخ بهینه و عملکرد در زمان واقعی، آن را تبدیل به یک استراتژی کارآمد برای آموزش کنترلگرهای هوشمند نموده است.

در یادگیری تقویتی گسسته مقدار ارزش حالت (یا جفت حالت-عمل) در جدول ارزش^۶ ذخیره شده و در هر قدم که آن حالت (یا جفت حالت-عمل) ملاقات شوند به روز رسانی انجام می‌گیرد [۲]. به کارگیری الگوریتم‌های یادگیری تقویتی گسسته در مسائل کنترل با دو چالش عمده روبروست. (۱) با توجه به اینکه حتی مسائل کنترلی ساده نیز عمدتاً دارای فضای حالت و عمل بزرگ و یا پیوسته می‌باشند و از آنجا که تعداد پارامترهای قابل تنظیم در یادگیری تقویتی گسسته، رابطه مستقیمی با بعداصلی^۷ فضای متغیرهای حالت و عمل مسأله دارد، در نتیجه در مسائل کنترل با مشکل تنگنای ابعاد^۸ مواجه هستیم. (۲) در فرایند آموزش ممکن است بسیاری از حالت‌ها (جفت‌های حالت و عمل) ملاقات نشوند و در نتیجه مقادیر آنها در جدول ارزش به روز رسانی نگردد. واضح است که در این صورت عامل نمی‌تواند عملکرد مناسبی را در مواجهه با این حالت‌ها پس از آموزش داشته باشد. در حالی که اصولاً ما از یک الگوریتم یادگیر انتظار داریم که قدرت استدلال و

تعمیم^۹ در مورد موارد ملاقات نشده را نیز دارا باشد. هر دو مشکل نامبرده شده به رویکرد استفاده از تقریب زنده‌های تابع برای تقریب تابع ارزش^{۱۰} منجر شده است [۴، ۵]. بر این اساس محققین با ترکیب الگوریتم‌های یادگیری تقویتی گسسته با تقریب زنده‌های تابعی^{۱۱}، الگوریتم‌های یادگیری تقویتی پیوسته راه ارائه داده‌اند. کاربردهای گسترده و عملکرد مطلوب شبکه‌های عصبی در کنترل و مسائل پیچیده و نیز مزایایی چون کنترل سیستم‌های غیرخطی با دقت دلخواه، مقاوم بودن و تحمل خطا، قابل یادگیر بودن (یعنی توانایی تنظیم وزن‌های شبکه)، قابلیت تعمیم، سرعت بالا به دلیل پردازش‌های موازی و قابلیت تقریب توابع باعث شده است تا محققین با ترکیب شبکه‌های عصبی به عنوان تقریب زنده با روش‌های یادگیری تقویتی، الگوریتم‌های یادگیری تقویتی عصبی^{۱۲} (NRL) را ارائه دهند [۶، ۷]. شبکه‌های MLP و RBF^{۱۳} از جمله پرکاربردترین شبکه‌های عصبی به کارگرفته شده بدین منظور بوده‌اند [۸-۱۰].

دومعماری معروف در زمینه الگوریتم‌های یادگیری تقویتی پیوسته عبارتند از: معماری عملگر-نقاد^{۱۴} و معماری نقاد-تنها^{۱۵}. قالب کلی اکثر پژوهش‌های انجام شده در این زمینه دارای معماری عملگر-نقاد است. این معماری دارای دو بخش عملگر و نقاد بوده و در آن بخش نقاد برای تقریب تابع ارزش و بخش عملگر برای تولید عمل استفاده می‌گردد [۹-۱۴]. با وجود کاربرد گسترده الگوریتم‌های عملگر-نقاد عصبی در بسیاری از مسائل توسط محققین، در مطالعات قبلی نشان داده شده است که الگوریتم‌های مذکور دارای فقدان کاوش مناسب می‌باشند. البته در تعدادی از مقالات [۱۲، ۱۵] برای افزایش درجه کاوش سیستم، خروجی شبکه را با یک مقدار تصادفی بر مبنای تابع توزیع چگالی احتمال گوسی با میانگین صفر جمع کرده و به فرآیند اعمال می‌کنند. علیرغم بهبودهای صورت گرفته بخاطر لحاظ کردن کاوش در کنترلگر، روش دارای مشکلات زیر می‌باشد: مقدار کاوش، تقریباً در همه وضعیت‌های سیستم یکسان، متقارن، حول یک عمل دارای ماکزیمم ارزش تخمینی و متناسب با مقدار متوسط پاداش‌ها است.

درمقابل معماری عملگر-نقاد، روش نقاد-تنها فقط دارای، یک بخش نقاد است که برای تقریب تابع ارزش عمل^{۱۶} استفاده می‌شود و عمل‌نهایی با توجه به مقادیر ارزش تقریب زده شده با استفاده از یکی از روش‌های انتخاب عمل همچون شبه حریصانه^{۱۷} و یا بیشینه نرم^{۱۸} تولید

⁹. Generalization

¹⁰. Value function

¹¹. Function approximators

¹². Neural Reinforcement Learning (NRL)

¹³. Radial Basis Function

¹⁴. Actor-critic

¹⁵. Critic-only

¹⁶. Action Value Function

¹⁷. ϵ -Greedy

¹⁸. Softmax

¹. Nondeterministic

². Uncertainty

³. Reinforcement learning (RL)

⁴. Stochastic

⁵. Exploration

⁶. Lookup table

⁷. Cardinality

⁸. Curses of dimensionality

ارزش عمل. بعلاوه، با توجه به آنکه در فرمول پیشینه نرم مقدار ارزش عمل مورد نیاز است، در مقاله‌ی مذکور مقدار ارزش عمل در هر قدم زمانی، از مقدار تابع ارزش با فرض دانستن مدل محیط محاسبه شده است، در حالی که در مسائل یادگیری تقویتی معمولاً مدل محیط، ناشناخته فرض می‌گردد.

در مقالات [۲۲، ۲۳] از ترکیب روش سارسا^۶ (که یک روش وابسته به سیاست است) با تقریب زنده‌های تابعی خطی (ما این ترکیب را سارسای خطی می‌نامیم) برای تقریب تابع ارزش عمل استفاده شده است. در [۱۶، ۲۲] همگرایی پارامترهای وزن سارسای خطی به یک ناحیه در صورت استفاده از یک سیاست ایستا^۷ در همه رویدادها، اثبات شده است. در [۲۳] همگرایی الگوریتم ارائه شده به پاسخ یکتا به شرط ایستا بودن سیاست‌ها در هر رویداد اثبات گردیده است. الگوریتم مذکور دارای دو ضعف عمده می‌باشد: ۱- نتایج قضیه، تضمینی در خصوص کیفیت سیاست نهایی که الگوریتم به آن همگرا می‌شود ارائه نمی‌دهد، ۲- با توجه به آنکه پس از تعیین هر سیاست جدید باید آموزش تا همگرایی وزن‌ها ادامه یابد، سرعت آموزش کند می‌باشد. در حالیکه در مسائل کنترل با آموزش روی خط، مطلوب آن است که در هر قدم زمانی سیاست انتخاب عمل به روز رسانی گردد. توجه شود که هیچ‌گونه پیاده سازی و شبیه سازی برای سه الگوریتم اشاره شده در بالا [۲۱-۲۳] ارائه نشده و در واقع دو چالش عمده برای پیاده سازی این الگوریتم‌ها وجود دارد: انتخاب تابع تقریب زنده تابعی خطی و نحوه انتخاب عمل، به نحوی که این دو بتوانند شرایط بیان شده در لم‌ها و قضایای بیان شده در مقالات مذکور را ارضاء نمایند.

در مقاله حاضر سعی داریم با ارائه یک ساختار تقریب زنده عصبی جدید و ترکیب آن با روش یادگیری تقویتی سارسا، یک الگوریتم جدید یادگیری تقویتی پیوسته که ما آن را یادگیری سارسای عصبی^۸ (NSL) می‌نامیم مورد مطالعه و بررسی قرار دهیم. به عنوان شرط لازم همگرایی برای الگوریتم NSL در این مقاله ما همچنین وجود نقاط ایستای الگوریتم ارائه شده را که منطبق بر نقاط ثابت^۹ الگوریتم تکرار تقریب ارزش عمل^{۱۰} است ثابت می‌نماییم. براساس تحقیقات ما تاکنون هیچ الگوریتم یادگیری تقویتی عصبی بر مبنای معماری نقاد- تنها با هدف تعیین بهترین عمل در هر حالت (استفاده در کنترل) که دارای تحلیل ریاضی باشد ارائه نشده است. لذا نتایج ارائه شده در این مقاله بعنوان اولین NRL با معماری نقاد- تنها است که دارای تحلیل ریاضی هم می‌باشد.

ساختار این نوشتار به این ترتیب صورت گرفته است. در بخش بعد مفاهیم پایه‌ای یادگیری تقویتی به طور مختصر توضیح داده می‌شوند.

می‌گردد. این نحوه‌ی انتخاب عمل درجه کاوش بالا را باعث می‌شود [۱۶، ۱۷].

در مقاله [۱۸] از یک شبکه عصبی بازگشتی برای تقریب تابع ارزش عمل در روش یادگیری Q استفاده شده است. ورودی‌های شبکه زوج های حالت و عمل مسأله و خروجی شبکه تابع ارزش عمل تقریب زده شده می‌باشد. در [۱۹] از ترکیب یک شبکه MLP با یک لایه از توابع گوسی یک تقریب زنده نیمه محلی^۱ معرفی شده است. ورودی‌های شبکه در این تقریب زنده بردار حالت مسأله می‌باشد. در [۲۰] از یک شبکه عصبی RBF چند ورودی چند خروجی به عنوان تقریب زنده تابع ارزش عمل استفاده شده است در این روش ورودی‌های شبکه بردار حالت مسأله بوده و خروجی‌های شبکه توابع ارزش عمل تقریب زده شده به ازای هر عمل گسسته می‌باشد به عبارت دیگر به تعداد عمل‌های گسسته از شبکه خروجی گرفته می‌شود. هیچ تحلیل ریاضی برای روش های بالا ارائه نشده است و بررسی‌های ما نشان داد که ساختار تقریب زنده و روش انتخاب عمل در الگوریتم‌های فوق به نوعی است که نمی‌تواند فرض‌های مربوط به لم‌ها و قضایای در این زمینه را برآورده کنند.

همچنین در سال‌های اخیر پژوهش‌هایی در مورد حل مسائل کنترل بهینه به صورت بر خط با استفاده از یادگیری تقویتی با فرض داشتن مدل کامل (جزئی) از دینامیک سیستم انجام شده است [۲۴-۲۷]. الگوریتم‌های ارائه شده در مقالات مذکور بر مبنای روش‌های برنامه ریزی پویا^۲ می‌باشند. روش‌های برنامه ریزی پویا برای بدست آوردن سیاست بهینه^۳ با فرض داشتن مدل کاملی از محیط به کار می‌روند. روش‌های مذکور اگرچه از لحاظ ریاضی قابل توجه‌اند ولی به خاطر احتیاج به مدل کاملی از محیط و همچنین حجم محاسبات گسترده، از اهمیت کمی در یادگیری تقویتی برخوردارند. در روش‌های یادگیری تقویتی مطلوب آن است که همان نتایج حاصل شده از روش‌های مذکور با محاسبات کمتر و بدون فرض داشتن یک مدل کامل از محیط حاصل گردد.

همانگونه که از مباحث بالا دریافت می‌شود عدم وجود تحلیل ریاضی یکی از مشکلات اکثر الگوریتم‌های یادگیری تقویتی پیوسته‌ای است که تاکنون توسط محققین ارائه شده‌اند. با این وجود در تعدادی از مقالات تحلیل‌های ریاضی مثبتی در مورد ترکیب الگوریتم‌های یادگیری تقویتی وابسته به سیاست با تقریب زنده‌های تابعی خطی موجود است. در [۲۱] وجود نقاط ایستا^۴ در روش تفاضل موقتی^۵ (TD) خطی بهره برنده از سیاست انتخاب عمل پیشینه نرم اثبات شده است. البته این اثبات تنها مربوط به تقریب تابع ارزش حالت است نه تابع

6. Sarsa learning

7. Stationary

8. Episodes

9. Neural Sarsa Learning (NSL)

10. Fixed points

11. Approximate Action Value Iteration

1. Semi-localised

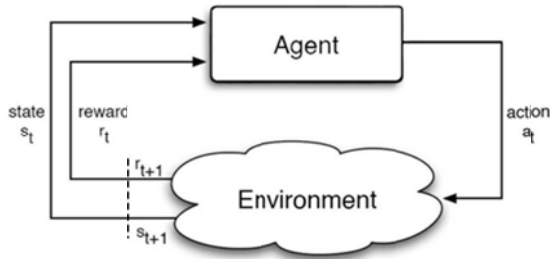
2. Dynamic programming

3. Optimal policy

4. Stationary points

5. Temporal Difference (TD)

اجرا انتخاب می‌کند سیاست می‌نامند. سیاست یا استراتژی عامل، $[0,1] \rightarrow \psi: \pi$ تابع احتمالی است که احتمال انتخاب شدن هر عمل



شکل ۱: چارچوب یادگیری تقویتی

را در هر حالت و با توجه به گام زمانی می‌دهد. به طور مثال $p = \pi_t(s, a)$ می‌گوید اگر عامل در زمان t در حالت S قرار گرفته باشد یا احتمال p عمل a را انتخاب می‌کند [۲].

در یادگیری تقویتی هدف عامل در قالب سیگنال پاداشی که از محیط دریافت می‌کند بیان می‌شود. در هر مرحله زمانی این پاداش به صورت عددی ساده بیان می‌شود $r_t \in R$. در بیانی ساده، هدف عامل بیشینه کردن مجموع این پاداش‌ها در بلندمدت^۳ است و این لزوماً به معنای بیشینه کردن پاداش در هر مرحله نیست.

$$E \left\{ \sum_{t=0}^{\infty} \gamma^t r_t \right\} \quad (1)$$

که در آن E بیانگر امید ریاضی و $\gamma \in [0,1]$ نرخ تنزیل می‌باشد که اهمیت نسبی پاداش‌های بلندمدت را نسبت به پاداش‌های کوتاه مدت تنظیم می‌کند. امید ریاضی کل پاداشی که عامل با شروع از حالت $S_t = S$ تحت سیاست π به دست می‌آورد ارزش یک حالت تحت سیاست π ، $V^\pi(s)$ ، می‌نامند [۲۸، ۲]:

$$V^\pi(s) = E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s \right\} \quad (2)$$

که در آن $E_\pi\{\cdot\}$ نشان دهنده امید ریاضی است در صورتی که از سیاست π پیروی شود. به طور مشابه تابع دیگری به نام $Q^\pi(s, a)$ تعریف می‌شود که بیانگر امید ریاضی کل پاداشی است که عامل با شروع از حالت $S_t = S$ ، انجام عمل $a_t = a$ و سپس در پیش گرفتن سیاست π بدست می‌آورد [۲]:

$$Q^\pi(s, a) = E_\pi \left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+1+k} \mid s_t = s, a_t = a \right\} \quad (3)$$

یادگیری سارسای عصبی در بخش سوم توضیح داده شده است. در بخش چهارم وجود نقاط ایستای منطبق بر نقاط ثابت الگوریتم تکرار تقریب ارزش عمل برای الگوریتم NSL ارائه شده در بخش سوم اثبات می‌گردد. در بخش چهارم مسائلی که برای ارزیابی عملی ایده‌ها مورد شبیه سازی قرار گرفته تشریح گشته و مختصری از نتایج شبیه سازی ارائه می‌گردد. نهایتاً بخش پنجم به طرح نتایج مهم، نتیجه‌گیری مطالب و پیشنهادات اختصاص دارد.

۲- مدل تصمیم گیری مارکف و یادگیری تقویتی

چارچوب ریاضی مرسوم محیط استفاده شده در اکثر مسائل RL "مدل تصمیم گیری مارکف" (MDP) می‌باشد [۲۷]. در یک سیستم مارکف حالت بعدی محیط و پاداش دریافتی تنها به عمل و حالت قبلی عامل در محیط بستگی دارد. یک MDP چندتایی (S, A, ψ, P, R) است. جایی که $S = \{s_1, s_2, \dots, s_n\}$ یک مجموعه محدود حالت‌های گسسته محیط، $A = \{a_1, a_2, \dots, a_m\}$ مجموعه محدود عمل‌های گسسته عامل، $\psi \subseteq S \times A$ مجموعه زوج-های حالت و عمل قابل قبول و $P: \psi \times S \rightarrow [0,1]$ تابع احتمال گذر از حالت S_t به حالت بعدی $S_{t+1} = S'$ با انجام عمل a_t است به طوری که $\forall (s, a) \in \psi \sum_{s' \in S} P(s_t, a_t, s') = 1$ و $R: \psi \rightarrow R$ تابع پاداش دریافتی توسط عامل می‌باشد. در هر مرحله زمانی $t = 0, 1, 2, \dots$ ، عامل حالت جدیدی از محیط را دریافت می‌کند $S_t \in S$ و بر مبنای این حالت، عمل خود $a_t \in A(S_t)$ را بر اساس سیاست اتخاذ شده انجام می‌دهد. یک گام بعد یعنی در $t + 1$ ، محیط یک پاداش عددی $r_{t+1} = r(S_t, a_t)$ بر حسب عمل عامل به وی می‌دهد و با احتمال $P(S_t, a_t, S')$ به حالت جدید S_{t+1} می‌رود (شکل ۱).

در یادگیری تقویتی عامل همواره می‌بایست مصالحه‌ایی بین دو هدف مخالف برقرار کند. از یک طرف کلیه‌ی جفت‌های حالت-عمل مسئله باید به حد کافی برای بدست آوردن دانش مورد کاوش قرار گیرند و از طرف دیگر تجربه‌های بدست آمده باید در انتخاب عمل بکار گرفته شوند. برای آموزش مؤثر، عمل‌ها باید بگونه‌ای انتخاب شوند که محیط بطور مناسب آزموده شده و از جریمه‌ها نیز حتی الامکان اجتناب گردد. انجام کامل این دو کار بطور همزمان ممکن نیست. به منظور برقراری تعادل بین کاوش و بهره‌برداری از تجربیات^۴ در حین آموزش عامل می‌بایست مکانیزمی برای انتخاب عمل خود داشته باشد تا بدین وسیله تعیین نماید که در هر گام زمانی مناسب‌ترین عمل کدام است. قانونی که عامل با توجه به آن در هر حالت، عملی را برای

^۱. Markovian Decision Process (MDP)

^۲. Exploitation

^۳. Long term reward

$$\sum_{t=0}^{\infty} \alpha_t = \infty, \quad \sum_{t=0}^{\infty} \alpha_t^2 < \infty \quad (۶)$$

۳- یادگیری سارسای عصبی

یکی از قدرتمندترین شبکه‌های عصبی مورد استفاده در مسائل تخمین تابع، شبکه عصبی RBF است. از مزایای این شبکه می‌توان به سرعت همگرایی بالا، ساختار ساده و در عین حال همگرایی به جواب بهینه دقیق اشاره نمود. این شبکه، الگوی ورودی n بعدی را با استفاده از گره‌های لایه پنهان به یک الگوی خروجی m بعدی نگاشت می‌کند. بنابراین از این شبکه‌ها می‌توان در هر مسأله‌ای که حل آن نیاز به یک نگاشت از فضای ورودی به خروجی باشد استفاده کرد. شبکه عصبی RBF نرمالیزه شده^۴ (NRBF) به وسیله نرمالیزه کردن توابع گوسی با جمع جبری خروجی کل توابع گوسی حاصل می‌شود. استفاده از شبکه NRBF در یادگیری تقویتی به واسطه حذف تأثیرات محدوده‌های توابع گوسی و عملکرد بهتر در تقریب توابع، مناسب‌تر است. در این مقاله یک ساختار تقریب زنده عصبی جدید مطابق با شکل (۲) که در آن از یک شبکه عصبی NRBF چهار لایه با یک نرون در لایه خروجی برای تقریب تابع ارزش عمل، پیشنهاد شده است.

لایه ورودی شبکه شامل $n + 1$ نرون می‌باشد. $x(t) = (s_t, a_t)^T \in R^{n+1}$ که متشکل از جفت حالت و عمل‌های مسأله است. که در آن $S_t = (s_1, s_2, \dots, s_n)^T \in R^n$ بردار حالت‌های مسأله می‌باشد و همچنین $a_t \in A$ که $A = [a_1, a_2, \dots, a_M]$ بردار عمل‌های مسأله است.

لایه ورودی، بردار ورودی $x_t \in R^{n+1}$ را به هر یک از گره‌های لایه پنهان ارسال می‌کند. در شبکه‌ی RBF لایه پنهان از نرون‌ها با توابع فعالساز شعاعی تشکیل شده است. متداول‌ترین تابع فعالساز در این نوع از شبکه‌ها تابع انتقال گوسی است.

به ازای هر متغیر ورودی $x_i = (x_1, x_2, \dots, x_{n+1})$ تابع پایه حالت و عمل گوسی به فرم زیر تعریف می‌گردد:

$$\varphi_{ip}(x_i) = \exp\left(-\frac{\|x_i - c_{ip}\|^2}{\sigma_{ip}^2}\right), p = 1, 2, \dots, k_i \quad (۷)$$

که در آن c_{ip} را مرکز^۵ گوسی می‌نامیم و σ_{ip} نیز مشخص کننده میزان گستردگی^۶ گوسی است که در آن مقدار تابع به طور مشخص از صفر متفاوت است. $\|\cdot\|$ نشان دهنده نرم اقلیدسی^۱ است.

تابع $Q^\pi(s, a)$ را تابع ارزش عمل برای سیاست π می‌نامند. در ادامه روش خطای تفاضل موقتی و یک تعمیم معروف آن که برای تخمین توابع ارزش عمل به کار می‌رود می‌آید.

۲-۱ یادگیری تفاضل موقتی

روش‌های TD به عنوان یکی از روش‌های پایه‌ای حل مسأله یادگیری تقویتی، ایده‌ای است که به عنوان هسته مرکزی در یادگیری تقویتی شناخته می‌شود. ساده‌ترین روش تفاضل موقتی که آن را TD(0) می‌نامند، برای تخمین تابع ارزش حالت به صورت زیر به کار می‌رود [۲]:

$$V(s_t) \leftarrow V(s_t) + \alpha_t [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)] \quad (۴)$$

که α نرخ آموزش، s_t بردار حالت، γ نرخ تنزیل^۱ و r مقدار پاداش آنی است. در این رابطه عبارت داخل کروشه را خطای تفاضل موقتی می‌نامند.

۲-۱-۱ یادگیری سارسا

یک قدم مناسب جهت حصول به کاوش بالاتر، هنگام بکارگیری یادگیری تقویتی در کنترل، تخمین تابع ارزش عمل بجای تابع ارزش حالت می‌باشد. روش سارسا مقادیر ارزش عمل را تحت سیاست جاری تخمین می‌زند. فرمول به روز رسانی مقادیر ارزش عمل در روش سارسا به صورت زیر است [۲]:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha_t [r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (۵)$$

در روش سارسا گذر از یک جفت حالت-عمل به جفت حالت-عمل بعدی لحاظ می‌گردد. قاعده به روز رسانی فوق، از همه عناصر چند تایی $(s_t, a_t, r_{t+1}, s_{t+1}, a_{t+1})$ استفاده می‌کند، به همین دلیل با توجه به حروف این چند تایی روش مذکور Sarsa نامیده شده است.

مقادیر ارزش در هر دو الگوریتم شرح داده شده در بالا، تحت فرض‌های زیر به مقدار بهینه همگرا می‌شوند [۳۲-۳۰]:

فرض ۱: محیط مسأله MDP غیر نوسانی^۲، کاهش ناپذیر^۳ و با سیگنال‌های تقویت محدود است.

فرض ۲: نرخ آموزش مثبت، کاهشی و شرایط زیر را، ارضا می‌کند.

^۴ Normalized Radial Basis Function (NRBF)

^۵ Center

^۶ Width

^۱ Discount factor

^۲ Aperiodic

^۳ Irreducible

$$c_{i1} = x_{i_{min}} \quad (۱۴)$$

$$c_{ip} = c_{i(p-1)} + \Delta c_i, \quad \text{for } p = 2, 3, \dots, k_i \quad (۱۵)$$

که Δc_i فاصله بین دو مرکز مجاور می‌باشد و بر طبق رابطه زیر تعیین می‌گردد:

$$\Delta c_i = \frac{x_{i_{max}} - x_{i_{min}}}{k_i - 1} \quad (۱۶)$$

با این توصیف، در هر مسأله از $h = k_1 \times k_2 \times \dots \times k_{n+1}$ تابع پایه حالت و عمل بر روی فضای $n + 1$ بعدی فضای حالت و عمل برای تقریب تابع ارزش عمل استفاده می‌گردد.

پارامترهای گسترده‌گی گوسی نیز با توجه به نوع مسأله توسط طراح تنظیم می‌گردند. هدف آموزش، به روز رسانی روی خط مقادیر وزن شبکه عصبی W به گونه‌ای است که بهترین تقریب از تابع ارزش عمل صورت پذیرد.

مقدار خطای تفاضل موقتی ارزش عمل بین دو حالت متوالی به صورت زیر تقریب زده می‌شود:

$$\Delta \hat{Q}_t(s_t, a_t) = r_{t+1} + \gamma \hat{Q}_t(s_{t+1}, a_{t+1}) - \hat{Q}_t(s_t, a_t) \quad (۱۷)$$

در این مقاله از روش گرادیان نزولی برای تنظیم بردار وزن در جهت کاهش مربع خطای تابع ارزش عمل تقریب زده شده $E(t) = \frac{1}{2} (\Delta \hat{Q}_t)^2$ استفاده شده است. پارامترهای وزن شبکه عصبی بر طبق قاعده زنجیری به صورت زیر به روز رسانی می‌گردند:

$$\begin{aligned} \Delta w_j(t+1) &= -\eta_t \frac{\partial E(t)}{\partial w_j(t)} \\ &= \eta_t \Delta \hat{Q}_t \frac{\partial Q(s_t, a_t)}{\partial w_j(t)} \\ &= \eta_t \left(r_{t+1} + \gamma \hat{Q}_t(s_{t+1}, a_{t+1}) - \hat{Q}_t(s_t, a_t) \right) \frac{\partial Q(s_t, a_t)}{\partial w_j(t)} \end{aligned} \quad (۱۸)$$

که در آن η_t نرخ یادگیری وزن‌های شبکه عصبی در بازه‌ی $[0, 1]$ در زمان t می‌باشد. از رابطه (۱۱) به صورت زیر محاسبه می‌گردد:

$$\frac{\partial Q(s_t, a_t)}{\partial w_j(t)} = \phi_j(x_t) \quad (۱۹)$$

برای یک الگوی ورودی x خروجی j -امین گره $\varphi_j: \mathcal{R}^{n+1} \rightarrow \mathcal{R}$ در لایه‌ی پنهان برابر است با:

$$\varphi_j(x_1, x_2, \dots, x_{n+1}) =$$

$$\prod_{i=1}^{n+1} \varphi_{ip} = \exp\left(-\sum_{i=1}^{n+1} \frac{\|x_i - c_{ip}\|^2}{\sigma_{ip}^2}\right), j = 1, 2, \dots, h \quad (۸)$$

نرمالیزه سازی در لایه‌ی سوم به صورت زیر انجام می‌گیرد:

$$\phi_j(x) = \frac{\varphi_j(x)}{\sum_{k=1}^h \varphi_k(x)}, j = 1, 2, \dots, h \quad (۹)$$

خروجی شبکه $\hat{Q}: \mathcal{R}^{n+1} \rightarrow \mathcal{R}$ با استفاده از حاصلجمع وزندار خطی پاسخ‌های لایه‌ی پنهان در گره‌ی خروجی بدست می‌آید:

$$\hat{Q}_t(s_t, a_t) = \sum_{j=1}^h \phi_j(x) w_j(t) \quad (۱۰)$$

که w_1, w_2, \dots, w_h پارامترهای وزن شبکه عصبی و $\phi_1, \phi_2, \dots, \phi_h$ توابع پایه ارزش عمل نرمالیزه شده روی فضای حالت-عمل می‌باشند [۲۳].

رابطه (۱۰) را می‌توان به فرم زیر به صورت برداری نوشت:

$$\hat{Q}_t(s_t, a_t) = \phi^T(s, a) w_t \quad (۱۱)$$

که $\phi^T(s, a) = [\phi_1(s, a), \dots, \phi_h(s, a)]$ ترانهاده ماتریس $\phi(s, a)$ می‌باشد. ماتریس فضای حالت-عمل Φ با ابعاد $|S| \times |A| \times |A|$ را به صورت زیر تعریف می‌کنیم:

$$\Phi = \begin{bmatrix} \phi_1(s, a) & \dots & \phi_h(s, a) \\ \vdots & & \vdots \\ \phi^T(s_N, a_M) \end{bmatrix} \quad (۱۲)$$

که S مجموعه حالت‌ها، A مجموعه عمل‌ها، $|S| = N$ و $|A| = M$ است. لذا داریم:

$$\hat{Q}_t = \Phi w_t \quad (۱۳)$$

بردار مراکز $c_r = [c_{r1}, c_{r2}, \dots, c_{rk_r}]$ برای k_i تابع پایه گوسی تعریف شده برای ورودی x_i به صورت زیر تنظیم می‌گردند:

^۱. Euclidean

خاصیت مهم از لحاظ همگرایی در یادگیری تقویتی پیوسته می باشد [۲۳].

۲- مقدار احتمال انتخاب عمل در فرمول پیشنهادی نرم به همه مقادیر ارزش وابسته است، این ویژگی امکان کاوش مناسب را فراهم می کند [۲].

۳- با تعیین مقدار ضریب دما می توان میزان کاوش و بهره برداری از تجربیات را کنترل کرد.

۶- محاسبه مقدار ارزش عمل تقریب زده شده جدید $\hat{Q}_{t+1}(s_{t+1}, a)$ برای کلیه عملها $a \in A$ با استفاده از فرمول (۱۱).

۷- اعمال عمل انتخاب شده به محیط.

۸- $t \leftarrow t + 1$ و بازگشت به قدم اول.

۴- تحلیل ریاضی الگوریتم NSL

مهمترین نتایج این مقاله در این بخش ارائه می گردد که در آن ما وجود نقاط ایستای منطبق بر نقاط ثابت الگوریتم تکرار تقریب ارزش عمل را برای NSL اثبات می کنیم.

۴-۱- اثبات وجود نقاط ایستا برای الگوریتم NSL

همچنانکه نشان داده خواهد شد، NSL یک پیاده سازی از الگوریتم سارسای خطی بهره برنده از سیاست پیشنهادی نرم می باشد. لذا در ابتدا شرایطی را که تحت آن وجود نقاط ایستا برای روش سارسای خطی وقتی که انتخاب عمل در آن بر اساس فرمول پیشنهادی نرم انجام می پذیرد، را بیان نموده و سپس نشان می دهیم که این شرایط برای NSL نیز برقرار است.

در الگوریتم سارسای خطی از K تابع پایه برای تقریب تابع ارزش عمل $Q(s, a)$ در قدم t به فرم زیر استفاده می شود [۱۶]:

$$\hat{Q}_i(W) = \sum_{k=1}^k w_k f_k(i) = W^T F(i) \quad \forall i \quad (22)$$

$$\in (s, a)$$

که W بردار وزنهای قابل تنظیم و $F_i = (f_1(i), \dots, f_k(i))$ توابع پایه حالت- عمل روی فضای حالت- عمل مسأله می باشند. ماتریس قطری D_π با ابعاد $|S| \times |S|$ شامل المانهای احتمال حالت ماندگار هر

حالت- عمل تحت سیاست π را در نظر بگیرید. با استفاده از روش

با توجه به این نکته، فرمول به روز رسانی پارمترهای وزن شبکه عصبی به صورت زیر در می آید:

$$\Delta w_j(t+1) = [\eta_t \times \Delta \hat{Q}_t(s_t, a_t) \times \phi_j(x_t)] \quad (20)$$

لازم به ذکر است که در رابطه (۱۷) مقدار $\hat{Q}_t(s_{t+1}, a_{t+1})$ وابسته به سیاستی است که عامل با توجه به آن در حالت بعدی، انتخاب عمل می نماید به همین دلیل الگوریتم NSL پیشنهادی یک الگوریتم "وابسته به سیاست" است.

رویه اجرای یادگیری بر اساس الگوریتم NSL به صورت زیر است:

۱- مشاهده حالت s_{t+1} و دریافت سیگنال تقویتی r_{t+1} .

۲- توقف آموزش، اگر الگوریتم به تعداد دفعات لازم موفقیت (رسیدن به هدف) و یا ماکزیمم تعداد رویدادها رسیده باشد.

۳- محاسبه مقدار ارزش عمل تقریب زده شده $\hat{Q}_t(s_{t+1}, a)$ برای کل عمل های ممکن $a \in A$ با استفاده از شبکه عصبی RBF.

۴- محاسبه $\Delta \hat{Q}$ و به روز رسانی مقادیر وزن های شبکه عصبی با استفاده از فرمول های (۱۷) و (۲۰).

۵- انتخاب عمل با استفاده از یکی از روش های مرسوم انتخاب عمل.

در این مقاله با توجه به آنکه می خواهیم از نتایج و فرض های ارائه شده در این بخش در تحلیل NSL استفاده کنیم، برای انتخاب عمل از رابطه پیشنهادی نرم به فرم زیر استفاده شده است:

$$\text{prob}(a_t = a_k) = \frac{\exp(\hat{Q}(s_t, a_k)/T)}{\sum_{b \in A} \exp(\hat{Q}(s_t, b)/T)} \quad (21)$$

که در آن $T > 0$ ضریب دما^۲ نامیده می شود. معمولاً در ابتدای آموزش مقدار ضریب دما بزرگ و در حین آموزش هر چه به سمت جلو می رویم مقدار ضریب دما کاهش می یابد تا از تجربیات قبلی بیشتر استفاده گردد. گذشت زمان در قالب تعداد تلاش های انجام شده و تعداد مراحل به روز شدن جدول دانش عامل و یا وزن های شبکه نشان داده می شود.

روش پیشنهادی نرم دارای ویژگی های مطلوب زیر است:

۱- تابع توزیع احتمال عمل در فرمول پیشنهادی نرم از توزیع بولتزمن که یک توزیع پیوسته است، تبعیت می کند. این ویژگی یک

^۱. On-policy

^۲. Temperature factor

به عبارت دیگر باید ثابت شود که: الف) ستون‌های ماتریس Φ مستقل خطی هستند. (۲) تابع احتمال انتخاب عمل در NSL برابر با فرمول پیشینه نرم سنتی می‌باشد.

الف) برای یک مجموعه گسسته محدود از حالت و عمل می‌توان ماتریس حالت و عمل Φ تعریف شده در رابطه (۱۲) را، برای NSL به صورت زیر بازنویسی کرد:

$$\Phi = \begin{bmatrix} \phi_1(s_1, a_1) & \cdots & \phi_h(s_1, a_1) \\ \phi_1(s_1, a_2) & \cdots & \phi_h(s_1, a_2) \\ \vdots & & \vdots \\ \phi_1(s_1, a_M) & \cdots & \phi_h(s_1, a_M) \\ \vdots & & \vdots \\ \phi_1(s_N, a_1) & \cdots & \phi_h(s_N, a_1) \\ \phi_1(s_N, a_2) & \cdots & \phi_h(s_N, a_2) \\ \vdots & & \vdots \\ \phi_1(s_N, a_M) & \cdots & \phi_h(s_N, a_M) \end{bmatrix} = [\phi_1 | \dots | \phi_h] \quad (25)$$

است. RBF-امین گره i خروجی نرمالیزه شده $\phi_i(s_j, a_k)$ که متمایز IRBF از آنجا که بردار مراکز و پارامترهای گسترده‌گی توابع دارای مرتبه Φ هستند، به سادگی می‌توان استدلال نمود که ماتریس NSL کامل یا به عبارت دیگر توابع پایه ارزش عمل در الگوریتم مستقل خطی هستند.

ب) از آنجا که استراتژی انتخاب عمل ما در این مقاله پیشینه نرم انتخاب شده بود، بنابراین فرض ۴ نیز برقرار است. □

۵- شبیه سازی

به منظور ارزیابی عملکرد استراتژی آموزشی ارائه شده و همچنین مقایسه آن با روش ارائه شده در مقاله [۲۰]، از این دو روش در حل مسائل حالت پیوسته خودرو در کوهستان و آکروبات استفاده می‌کنیم.

۵-۱- مسأله خودرو در کوهستان

مسأله خودرو در کوهستان^۱ یکی از مسائل معروف در زمینه یادگیری تقویتی است، که در مقالات مختلف [۱۴، ۳۵] و همچنین مسابقات علمی مرتبط در این زمینه [۳۶] جهت ارزیابی الگوریتم‌های یادگیری تقویتی جدید از آن استفاده شده است. اتومبیلی را در سراسیمه یک جاده کوهستانی مانند شکل (۳) در نظر بگیرید. هدف راننده ماشین به بالای تپه می‌باشد، اما از آنجا که شتاب گرانش زمین قویتر از نیروی محرک موتور اتومبیل است، قادر نیست به سادگی از سطح شیب دار

گرایان نزولی برای تنظیم بردار وزن در جهت کاهش مجموع مربعات خطای $\|Q - \hat{Q}\|_{D_\pi}^2$ فرمول به روز رسانی بردار وزن پس از اجرای عمل a_t در حالت S_t برای روش سارسای خطی به صورت زیر حاصل می‌شود [۳۳]:

$$w_{t+1} = w_t + \alpha_t \phi^T(s_t, a_t) [r_{t+1} + \gamma \phi^T(s_{t+1}, a_{t+1}) w_t - \phi^T(s_t, a_t) w_t] \quad (23)$$

که r_{t+1} پاداش آنی بعد از اجرای عمل a_t در حالت S_{t+1} بعدی، و a_{t+1} عمل بعدی می‌باشد.

لازم به ذکر است که برای تحلیل NSL از روابط و ویژگی‌های یک نسخه از الگوریتم "تکرار تقریب ارزش عمل" استفاده شده است. یک نسخه از الگوریتم مذکور در [۱۶] ارائه گردیده و وجود حداقل یک نقطه ثابت برای آن اثبات شده است.

لم: الگوریتم سارسای خطی (۲۳) تحت فرض‌های ۱ تا ۴ دارای نقاط ایستای منطبق بر نقاط ثابت روش تکرار تقریب ارزش عمل می‌باشد.

فرض ۳: توابع پایه حالت-عمل $\{F_i | i = 1, 2, \dots, k\}$ در رابطه (۲۲) مستقل خطی هستند.

فرض ۴: سیاست انتخاب عمل در هر حالت با استفاده از فرمول پیشینه نرم سنتی به صورت زیر به دست می‌آید [۲، ۳۴]:

$$\pi_w^\beta(s, a) = \frac{\exp[\beta \cdot \hat{Q}(s_t, a_t)]}{\sum_{b \in A} \exp[\beta \cdot \hat{Q}(s_t, b)]} \quad (24)$$

که در آن β ضریب کاوش می‌باشد و به صورت عکس ضریب دما تعریف می‌شود. از آنجا که عمل‌ها بر طبق تابع ارزش عمل تقریب زده شده، حاصل می‌گردند احتمال انتخاب عمل به β وابسته است؛ به همین دلیل β و w در اندیکس π در روابط آورده شده‌اند.

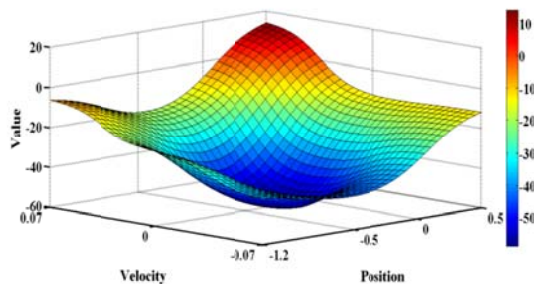
قضیه ۱: الگوریتم یادگیری سارسای عصبی (NSL) ارائه شده، تحت فرض‌های ۱ و ۲ دارای نقاط ایستای منطبق بر نقاط ثابت روش تکرار تقریب ارزش عمل می‌باشد.

اثبات: کافی است نشان دهیم که NSL، یک پیاده سازی از الگوریتم سارسای خطی (۲۳) است. به سادگی می‌توان نشان داد که با جایگزینی رابطه (۱۱) در رابطه (۱۷) و (۲۰)، رابطه (۲۳) حاصل می‌شود. بنابراین قانون به روز رسانی وزن‌های الگوریتم NSL معادل با الگوریتم سارسای خطی است.

اکنون با توجه به لم ۱، اگر ما ثابت کنیم که شرایط ذکر شده در فرض‌های ۳ و ۴ برای NSL برقرار است، آنگاه نتایج لم ۱ هم برای NSL صادق است و قضیه اثبات شده است.

¹. Mountain car problem

مشخص است مقدار ارزش در نواحی نزدیک به هدف و نواحی سمت چپ مسیر حرکت ماشین بیشتر از قسمت های دیگر است.

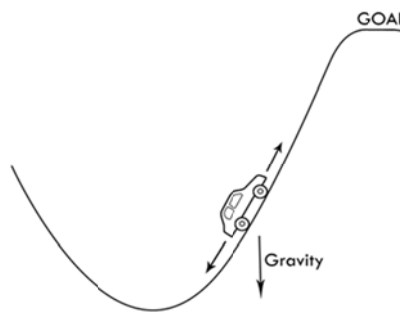


شکل ۴: ارزش عمل با بالاترین ارزش در NSL

۵-۱-۲- جزئیات آموزش

در این مسأله ورودی های شبکه عصبی $x = [s, a]^T$ در $[x, v, a]^T$ متشکل از موقعیت و سرعت عامل $s \in [x, v]$ باضافه تک تک عمل ها $a \in A = [-1, 0, +1]$ در آن حالت می باشد. در این مسأله از $3 \times 3 \times 3$ تابع پایه حالت و عمل برای تقریب تابع ارزش عمل در روش NSL استفاده شده است. برای اینکه ارزیابی دقیق و عادلانه ای در مقایسه بین دو روش داشته باشیم تمامی شرایط در هر دو روش یکسان در نظر گرفته شده اند بنابراین در شبیه سازی روش ارائه شده در مقاله [۲۰] نیز ۳ تابع پایه گوسی بر روی هر یک از متغیرهای حالت ورودی تعریف شده است. در هر دو الگوریتم $\lambda = 0.99$ و همچنین نرخ نمونه برداری ۰.۰۲ ثانیه در نظر گرفته شده است. نتایج هر آزمایش متوسط ارزیابی انجام شده در ۳۰ اجرا است. در ابتدای هر اجرا مقادیر وزن های شبکه عصبی با مقدار اولیه صفر مقداردهی می شوند. هر اجرا شامل دو بخش آموزش و تست می باشد. اگر تعداد رویدادها (منظور از یک رویداد شروع از نقطه آغاز و رسیدن به هدف می باشد) از ۲۰۰ بیشتر شود یا تعداد دفعات متوالی که عامل به هدف رسیده به ۴۰ برسد بخش آموزش پایان می یابد. شماره رویداد در پایان بخش آموزش به عنوان "معیار زمان آموزش" (LDI)^۲ در نظر گرفته می شود. هر رویداد از یک نقطه تصادفی آغاز می شود و زمانی پایان می پذیرد، که یا عامل به هدف برسد و یا تعداد قدم های انجام شده توسط عامل در یک رویداد از ۷۰۰ بیشتر شود. تعداد قدم های انجام شده در رسیدن به هدف در یک رویداد به عنوان معیار "تعداد قدم زمانی رسیدن به هدف" (LE)^۳ در نظر گرفته می شود. همچنین درصد موفقیت عامل در رسیدن به هدف به عنوان "معیار نرخ موفقیت"^۴ لحاظ می گردد. بخش تست شامل ۴۰ رویداد با شروع از یک نقطه تصادفی می باشد.

جاده حتی با تمام نیرو بالا رود. تنها راه حل ممکن در این شرایط بازگشت به عقب و گرفتن شتاب لازم جهت غلبه بر شتاب گرانش و عبور از شیب جاده می باشد. متغیرهای ورودی، موقعیت اتومبیل (x) و سرعت آن (v) می باشند. هدف از یادگیری، تنظیم روی خط وزن های یک کنترلر عصبی به گونه ای است که بتواند نیروی کنترلی لازم (F) جهت راندن اتومبیل به بالای تپه از هر موقعیت و سرعت اولیه را در حداقل زمان فراهم نماید.



شکل ۳: مسأله خودرو در کوهستان

۵-۱-۱- دینامیک سیستم

معادلات حرکت اتومبیل به صورت زیر است [۲، ۳۵]:

$$v_{t+\tau} = \min(0.07, \max(-0.07, v_t + 0.001 \times F_t + g \times \cos(3x_t)))$$

$$x_{t+\tau} = \min(0.5, \max(-1.2, x_t + v_{t+1}))$$

(۲۶)

که در آن $F \in \{-1, 0, 1\}$ ، کل عمل های ممکن قابل انجام در این مسأله می باشد. محدوده تغییرات موقعیت ماشین (x) و سرعت آن (v) به صورت زیر در نظر گرفته می شوند:

$$\{(x, v \in R^2) | -1.2 \leq x \leq 0.5, -0.07 \leq v \leq 0.07\}$$

(۲۷)

هنگامی که موقعیت اتومبیل به نقطه $x_t = -1.2$ برسد، سرعت ماشین صفر شده و رسیدن به موقعیت $x = 0.5$ هدف عامل بوده و به عنوان موفقیت در نظر گرفته می شود.

سیگنال تقویتی در این مسأله به صورت زیر تعریف می گردد:

$$r(x, v) = \begin{cases} +1 & x = 0.5 \\ -1 & \text{otherwise} \end{cases}$$

(۲۸)

شکل (۴) منحنی ارزش عمل با بالاترین ارزش $\max_a(Q(s, a))$ در مسأله خودرو در کوهستان را نشان می دهد.^۱ همانطور که از شکل

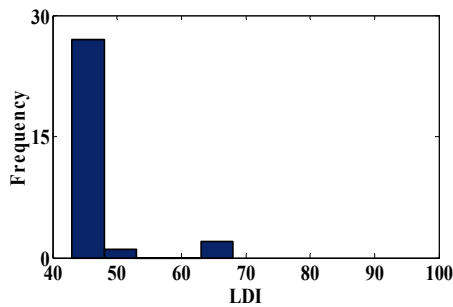
^۱ این شکل توسط الگوریتم NSL تولید شده است.

^۲ Learning Duration Index
^۳ Length of Episode
^۴ Success Rate

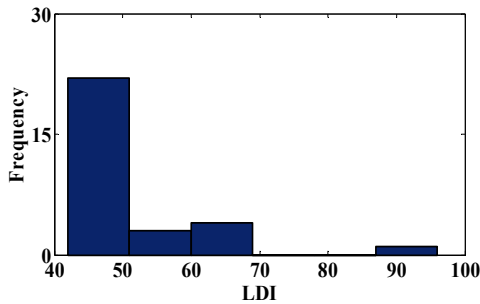
جدول ۱: نتایج شبیه سازی برای مقادیر متفاوت نرخ آموزش اولیه و ضریب دمای اولیه

Initial Parameters	Method	Avg. LDI	Success Rate	LE	Success Rate (Test)	LE (Test)
$T_0 = 1.0$ $\eta_0 = 0.2$	NSL method's [20]	47.83 51.73	89.75 90.85	147.34 143.74	98.25 97.33	67.42 77.12
$T_0 = 0.1$ $\eta_0 = 0.2$	NSL method's [20]	44.26 45.16	93.07 92.62	128.11 130.55	98.33 97.11	68.66 76.58
$T_0 = 0.01$ $\eta_0 = 0.1$	NSL method's [20]	50.30 51.14	89.86 89.44	165.28 164.41	82.93 81.92	170.32 175.41
$T_0 = 5.0$ $\eta_0 = 0.2$	NSL method's [20]	53.50 53.06	85.29 84.55	180.87 184.52	99.16 97.83	64.05 71.03

LDI, Learning duration index; LE, Length of Episode; NSL, Neural Sarsa Learning.



شکل ۵: هیستوگرام LDI ها در NSL



شکل ۶: هیستوگرام LDI ها در [20]

الگوریتم در شکل‌های (۵) و (۶) برای حالت $\eta_0 = 0.2$ و $T_0 = 1.0$ نشان داده شده است. همانگونه که از شکل‌ها مشخص است هیستوگرام‌های رسم شده نیز پراکندگی کمتر LDI ها در NSL را نشان می‌دهد که مؤید عملکرد بهتر روش NSL می‌باشد. نتایج ارائه شده در جدول (۱)، حاکی از وابستگی کیفیت و زمان آموزش در هر دو الگوریتم به مقدار اولیه ضریب دما است. هر چه ضریب دما کاهش می‌یابد سرعت آموزش سریع‌تر می‌شود. اما این کاهش ضریب دما، کیفیت آموزش را نیز کاهش می‌دهد. اگر ضریب دما خیلی کوچک شود، سیاست استفاده شده به سیاست حریصانه همگرا می‌شود. در این حالت زمان آموزش برای الگوریتم‌ها کوتاه شده و سیستم به سرعت به یکی از نقاط اکسترمم همگرا می‌گردد، لیکن به دلیل نبود

نرخ یادگیری و ضریب دما مطابق با روابط زیر در هر مرحله آموزش به صورت زیر کاهش می‌یابد [۱۶]:

$$\eta_t = \begin{cases} \eta_{t-1}/1.001 & \text{if } t = 4k \\ \eta_{t-1} & \text{otherwise} \end{cases} \quad (29)$$

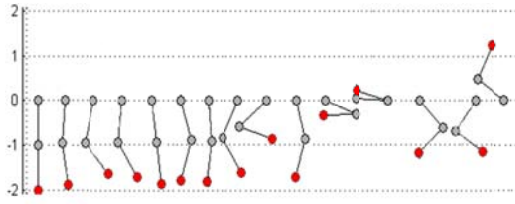
$$T_t = \begin{cases} T_{t-1} - (0.99)^t \times 0.4 \times T_t & \text{if } t = 4k, t < 25 \\ T_{t-1} - (0.99)^t \times 0.2 \times T_t & \text{if } t = 4k, t \geq 25 \\ T_{t-1} & \text{otherwise} \end{cases} \quad (30)$$

که $t \geq 1$ شماره رویداد می‌باشد. همچنانکه از روابط (۲۹) و (۳۰) مشخص است مقدار ضریب دما و نرخ آموزش در طول هر رویداد ثابت است و در شروع هر رویداد طبق روابط مذکور مقدار آن‌ها تعیین می‌گردد.

به منظور افزایش درجه کاوش و جلوگیری از گیر افتادن در مینیمم‌های محلی، در صورتی که الگوریتم نتواند بعد از ۷۰ رویداد در بخش آموزش همگرا گردد ضریب آموزش روی مقدار اولیه و ضریب دما روی ۲۵٪ مقدار اولیه تنظیم می‌گردد.

۵-۱-۳- نتایج شبیه سازی

جدول (۱) نتایج ۴ سری آزمایش را بر روی هر دو روش برای مقدارهای اولیه متفاوت نرخ یادگیری و ضریب دما نشان می‌دهد. ستون‌های ۳ تا ۵ به ترتیب بیانگر معیار زمان آموزش، تعداد قدم‌های انجام شده در رسیدن به هدف و نرخ موفقیت، در بخش آموزش می‌باشند. ستون‌های ۶ و ۷ نیز به ترتیب بیانگر نرخ موفقیت و تعداد قدم‌های انجام شده در رسیدن به هدف در بخش تست هستند. همچنانکه نتایج شبیه سازی نشان می‌دهد، زمان آموزش و کیفیت عملکرد در مجموع در روش NSL بهتر از روش ارائه شده در [۲۰] می‌باشد. در حقیقت ساختار تقریب زنده در NSL به نوعی است که تقریب مناسب‌تری از ارزش عمل صورت می‌پذیرد. هیستوگرام LDI ها برای هر دو



شکل ۸: نمایی از مجموعه حرکات آکروبات با نرخ نمونه برداری ۰.۴ ثانیه

جدول ۲: پارامترهای مسأله آکروبات شبیه سازی شده

Parameter	Description	Value
l_1, l_2	Link lengths	1.0
lc_1, lc_2	Joint to mass center	0.5
m_1, m_2	Link masses	1.0
I_1, I_2	Link inertias	1.0
τ	Torque range	$\{-1, 0, 1\}$
g	Gravitational force	9.8
t_i	Integration time step	0.05
t_c	Control time step	0.2

۵-۲-۱- دینامیک سیستم

معادلات دینامیکی حاکم بر سیستم به شکل زیر می باشند [۳۸، ۶]:

$$\ddot{\theta}_1 = -d_1^{-1}(d_2\ddot{\theta}_2 + \phi_1) \quad (۳۱)$$

$$\ddot{\theta}_2 = \left(m_2lc_2^2 + I_2 - \frac{d_2^2}{d_1}\right)^{-1} \left(\tau + \frac{d_2}{d_1}\dot{\phi}_1 - \phi_2\right) \quad (۳۲)$$

که در آن

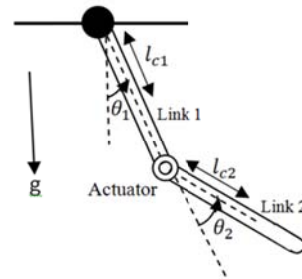
$$d_1 = m_1lc_1^2 + m_2(l_1^2lc_2^2 + 2l_1l_2\cos\theta_2) + I_1 + I_2 \quad (۳۳)$$

$$d_2 = m_2(lc_2^2 + l_1lc_2\cos\theta_2) + I_2 \quad (۳۴)$$

$$\phi_1 = -m_2l_1lc_2\dot{\theta}_2^2 \sin\theta_2 - 2m_2l_1lc_2\dot{\theta}_2\dot{\theta}_1 \sin\theta_2 + (m_1lc_1 + m_2l_1)g\cos(\theta_1 - \pi/2) + \dot{\phi}_2 \quad (۳۵)$$

$$\phi_2 = m_2lc_2g\cos(\theta_1 + \theta_2 - \pi/2) \quad (۳۶)$$

شایان ذکر است معادلات فوق تنها جهت شبیه سازی سیستم مورد استفاده قرار گرفته اند و از دید عامل (بخش کنترل) دینامیک سیستم ناشناخته می باشد.



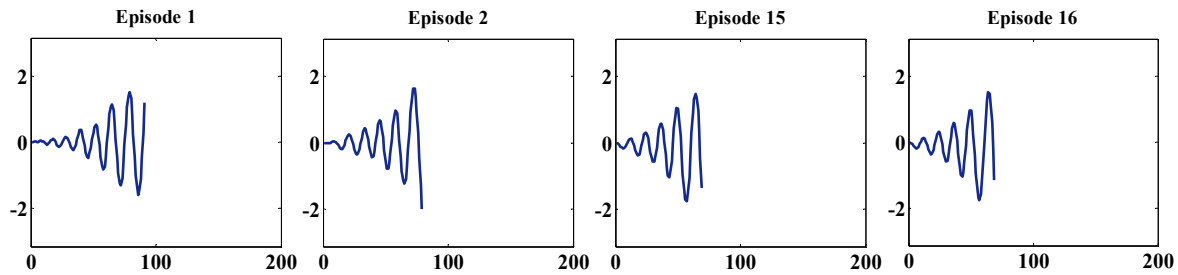
شکل ۷: مسأله آکروبات

کاوش کافی، کیفیت آموزش افت قابل ملاحظه ای می نماید. مقادیر بزرگ ضریب دما کاوش بالاتری را به سیستم می دهد، اما اگر این ضریب خیلی بزرگ باشد به علت فضای بزرگ مسأله، عامل دچار سردرگمی می شود و زمان آموزش و کیفیت آموزش افت می نمایند. از این رو یک مقدار میانی برای ضریب دما جواب مناسب تری داده است. باید توجه شود که در فرمول پیشنهادی، درجه کاوش به ضریب دما و اختلاف مقادیر ارزش عمل وابسته است. اگر اختلاف مقادیر ارزش عمل کوچک باشد، آنگاه انتخاب یک ضریب دمای متوسط باعث کاوش بالا می شود. همچنین اگر اختلاف مقادیر ارزش عمل بزرگ باشد، یک ضریب دمای نسبتاً کوچک می تواند باعث استفاده از تجربیات بالایی گردد. با توجه به نکات فوق، مقدار اولیه ضریب دما باید به گونه ای انتخاب شود که اولاً در شروع آموزش که اختلاف بین مقادیر ارزش عمل کوچک است، میزان کاوش نسبتاً زیاد باشد، ثانیاً مقدار ضریب دما پس از حدود ۲۵ رویداد به حد کافی کوچک شده باشد به گونه ای که اگر اختلاف بین مقادیر ارزش عمل زیاد شده است، انتخاب عمل به سمت سیاست حریصانه نزدیک گردد.

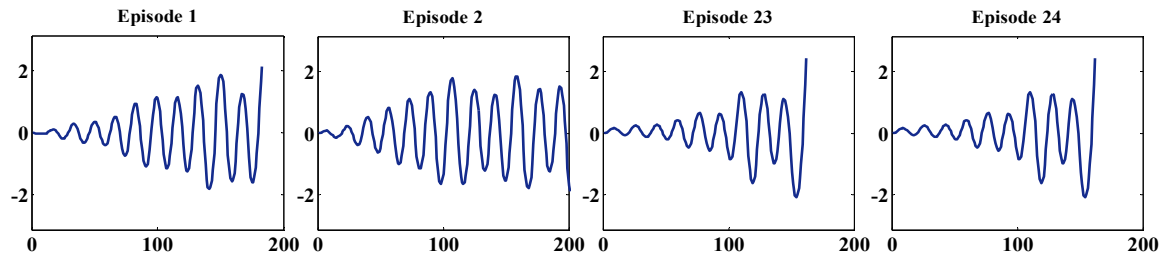
۵-۲- مسأله آکروبات

آکروبات^۱ در واقع یک روبات صفحه ای دو لینکی است (شکل ۷)، که از نظر فیزیکی رفتار یک ژیمناستیک کار که با خم و راست نمودن کمر خود سعی می کند تا انرژی لازم را برای بالا کشیدن بدن خود به بالای میله ای که از آن آویزان شده را فراهم آورد، مدل می کند [۳۷، ۳۸]. در این مسأله هدف بالا آوردن نوک لینک دوم (پای ژیمناستیک کار) به اندازه یک واحد بالای نقطه ی آویز می باشد. برای این منظور ربات باید حالت تاب خوردن گرفته تا بتواند انرژی لازم را بدین منظور به دست آورد. متغیرهای حالت مسأله، شامل زاویای دو لینک و مشتقات آنها می باشد. در این مسأله سرعت زاویه ای لینک اول به بازه ی $\dot{\theta}_1 \in [-4\pi, 4\pi]$ و سرعت زاویه ای لینک دوم به بازه ی $\dot{\theta}_2 \in [-9\pi, 9\pi]$ محدود شده اند [۶]. شکل (۸) نمایی از مجموعه حرکات آکروبات برای رسیدن به هدف را با نرخ نمونه برداری ۰.۴ ثانیه نشان می دهد.

^۱. Acrobot



شکل ۹: تغییرات زاویه‌ای لینک اول (θ_1) در روش NSL



شکل ۱۰: تغییرات زاویه‌ای لینک اول (θ_1) در روش [۲۰]

سیاست نهایی که در نهایت روش NSL بدان همگرا شده است به مراتب بهتر بوده و تعداد قدم‌های زمانی بسیار کمتری در رسیدن به هدف انجام شده است.

۶- نتیجه گیری و کارهای آینده

در این مقاله یک الگوریتم جدید یادگیری تقویتی که از ترکیب روش‌های یادگیری تقویتی با شبکه‌های عصبی حاصل شده و در واقع تعمیمی از روش گسسته سارسا است معرفی شد. الگوریتم مذکور که NSL نامیده شد از لحاظ معماری دارای ساختار نقاد - تنها بوده و به منظور تنظیم وزن‌های یک شبکه عصبی به صورت بر خط به کار می‌رود. شبکه عصبی مورد نظر یک شبکه RBF است که به عنوان تقریب زنده‌ی تابع ارزش عمل به کار گرفته شد. از نتایج شبیه سازی انجام شده در مورد مسائل خودرو در کوهستان و آکروبات می‌توان نتیجه گرفت که به کارگیری معماری نقاد-تنها به خاطر پتانسیل کاوش بالاتر شانس رسیدن به جواب بهینه را زیاده‌تر می‌نماید. هر چند نتایج شبیه سازی در مقایسه روش NSL با روش ارائه شده در مقاله [۲۰] در مورد مسئله خودرو در کوهستان به هم نزدیک بود. نتایج مربوط به مسئله آکروبات که مسئله‌ای بسیار سخت‌تر از مسئله خودرو در کوهستان است کارآیی بالاتر روش ارائه شده توسط ما را مشخص نمود.

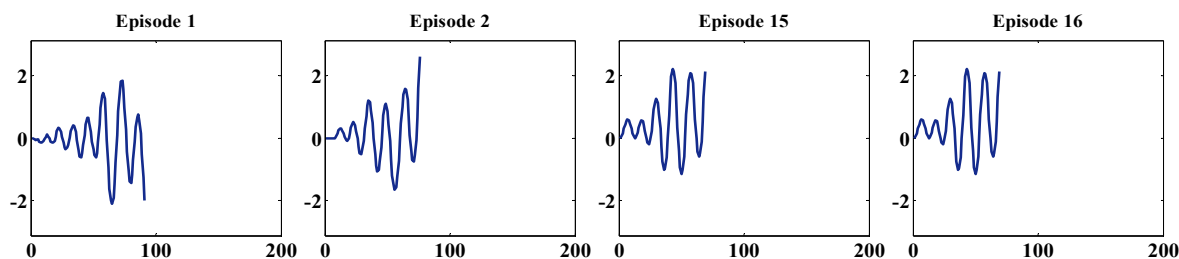
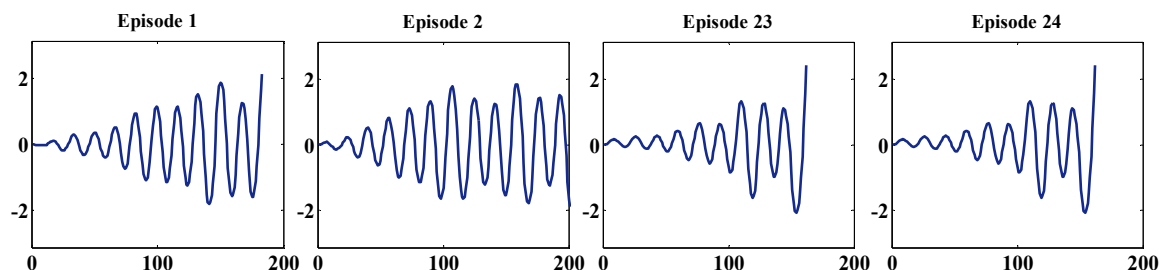
با استفاده از نتایج مربوط به وجود نقاط ثابت برای روش تکرار تقریب ارزش عمل و همچنین تعریف تابع پایه حالت- عمل در الگوریتم NSL به نحوی که شرایط سارسای خطی بیان شده را ارضا نماید، وجود نقاط ایستای

۵-۲-۲- جزئیات آموزش

ورودی‌های شبکه عصبی در این مسأله در روش NSL مشکل از زوایای دو لینک و مشتقات آنها باضافه‌ی تک تک عمل‌ها در آن حالت می‌باشد $x = [s, a]^T = [\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2, a]^T$. در این مسأله $3 \times 3 \times 3 \times 3 \times 3$ تابع پایه‌ی حالت و عمل برای تقریب تابع ارزش عمل در روش NSL استفاده شده است در روش [۲۰] نیز برای هر متغیر پیوسته‌ی حالت ورودی ۳ تابع پایه حالت گوسی در نظر گرفته شده است. پارامترهای سیگنال تقویتی، نرخ آموزش، ضریب دما مانند مثال قبل است. برای مشاهده تأثیر کنترل اعمالی، کنترل‌گر و مکانیسم آموزش در هر ۰.۲ ثانیه تحریک می‌شوند. تعریف کلیه پارامترهای استفاده شده و همچنین مقادیری که در شبیه سازی برای آنها در نظر گرفته شده در جدول (۲) آمده است. برای مقایسه دو الگوریتم تغییرات زوایای دو لینک در یک اجرا که شامل حداکثر ۵۰ رویداد با شروع از نقطه $(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) = (0, 0, 0, 0)$ می‌باشد بررسی شده است. حداکثر تعداد قدم‌ها در یک رویداد در بخش آموزش ۲۰۰ لحاظ شده است. شکل‌های (۹) و (۱۰) تغییرات زوایای لینک اول (θ_1) و شکل‌های (۱۱) و (۱۲) تغییرات زاویه‌ای لینک دوم را در هر دو روش برای حالتی که $\eta_0 = 0.2$ ، $T_0 = 0.1$ و $\lambda = 0.9$ نشان می‌دهند.

۵-۲-۳- نتایج شبیه سازی

همانگونه که از شکل‌ها مشخص است روش NSL هم از جهت سرعت همگرایی و هم از نظر سیاست نهایی که بدان همگرا شده است به مراتب بهتر از روش ارائه شده در [۲۰] عمل کرده است. در روش NSL به طور متوسط الگوریتم پس از ۱۵ رویداد همگرا شده است در حالی که این معیار در روش [۲۰] در حدود ۲۳ رویداد می‌باشد همچنین

شکل ۱۱: تغییرات زاویه‌ای لینک دوم (θ_2) در روش NSLشکل ۱۲: تغییرات زاویه‌ای لینک دوم (θ_2) در روش [γ^*]

- [2] Sutton, Richard S., and Barto, Andrew.G. "Reinforcement Learning: An Introduction", Cambridge, MA: MIT Press, 1998.
- [3] Sutton, R.S., Barto, A.G., and Williams, R.J., 1992, "Reinforcement learning is direct adaptive optimal control", IEEE control systems magazine, pp.19-22.
- [4] Wiering, M.A., 2004, "Convergence and divergence in standard and averaging reinforcement learning", Proceedings of European Conference on Machine Learning, Italy, pp.477 – 488.
- [5] Tsitsiklis, J. N., and Van Roy, B., 1997, "An analysis of temporal-difference learning with function approximation", IEEE Transactions on Automatic Control, Vol. 42, No. 5, pp. 674–690.
- [6] Barto, A.G., Sutton, R.S., and Anderson, C.W., 1983, "Neuronlike adaptive elements that can solve difficult learning control problems", IEEE Transactions on Systems, Man, and Cybernetics, Vol. 13, No.5.
- [7] Potocnil, P., and Grabec, I., 2000, "Adaptive self-tuning neurocontrol", Mathematics and Computers in Simulation, Vol.51, pp.201-207.
- [8] Haykin, S., 1994, "Neural Networks, A Comprehensive Foundation", New York: Macmillan.
- [9] Guili, C., Wang, M., Hung, Z.J., and Zhang, Z.F., 2009, "An actor-critic reinforcement learning algorithm based on adaptive RBF network", Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, pp.984-988.
- [10] Si, J., and Wang, Y.T., 2001, "Online learning control by association and reinforcement", IEEE Transactions on Neural Networks, Vol. 12, No.2, pp. 264-276.
- [11] Konda, V.R., and Tsitsiklis, J.N., 2003, "On Actor-Critic Algorithms", SIAM Journal on Control and Optimization, 42(4):1143–1166.

منطبق بر نقاط ثابت روش تکرار تقریب ارزش عمل برای NSL بعنوان اولین کار تحلیلی در این زمینه اثبات شد.

توجه شود که اثبات وجود نقاط ایستا لزوماً همگرایی روش را نتیجه نمی‌دهد، بلکه این یک شرط لازم برای همگرایی است. به طور کلی این نتیجه از چند جهت مهم می‌باشد: ۱- بر طبق تحقیقات و جستجوهای که ما انجام دادیم، نتایج نظری ارائه شده برای NSL در این مقاله اولین نتایج مثبت ریاضی برای الگوریتم‌های NRL با معماری نقاد-تنها است، وقتی که سیاست در هر قدم زمانی به روز رسانی می‌شود، می‌باشد. ۲- NSL یک پیاده سازی سارسای خطی با استفاده از شبکه‌های عصبی است که در این پیاده سازی ما با ارائه یک ساختار تقریب زننده عصبی جدید توانستیم راهکاری برای رفع چالش، تعریف توابع پایه حالت- عمل ارائه دهیم. ۳- اثبات وجود نقاط ایستای منطبق بر نقاط ثابت روش تکرار تقریب ارزش عمل حاکی از امید بالا برای دست یافتن به راندمان مناسب می‌باشد. ۴- در الگوریتم NSL سیاست انتخاب عمل در هر قدم زمانی به روز رسانی می‌گردد که منجر به سرعت بالای آموزش در این روش می‌شود.

در آینده تلاش خواهیم نمود تا الگوریتمی عمل پیوسته بر مبنای ایده‌ی ارائه شده برای حل مسائلی که در آنها پیوستگی عمل مزیت بارزی نسبت به عمل گسسته دارد ارائه نماییم.

مراجع

- [1] Gullapallia, V., 1991, "A comparison of supervised and reinforcement learning methods on a reinforcement learning task", Proceedings of IEEE International Conference on Intelligent control, USA, pp. 394-399.

- [24] Vamvoudakis, K. G., and Lewis, F. L., 2010, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem", *Automatica*, vol. 46, pp. 878-888.
- [25] Vrabie, D., and Lewis, F. L., 2008, "Adaptive optimal control algorithm for continuous-time nonlinear systems based on policy iteration", *IEEE Proc. CDC08*, pp. 73-79.
- [26] D. Vrabie, Pastravanu, O., Abu-Khalaf, M., and Lewis, F.L., 2009, "Adaptive optimal control for continuous-time linear systems based on policy iteration", *Automatica*, vol. 45, no. 2, pp. 477-484.
- [27] Al-Tamimi, A. Lewis, F.L., 2008, "Discrete-Time Nonlinear HJB Solution Using Approximate Dynamic Programming: Convergence Proof", *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 38, no. 4.
- [28] Kalyanasundaram, S., Chong, E. K. P., and Shroff, N. B., 2004, "Markov decision processes with uncertain transition rates: sensitivity and max hyphen min control", *Asian J. Control*, Vol. 6, No. 2, pp.253-269.
- [29] Kaelbling, L. P., Littman, M. L., and Moore, A. W., 1996, "Reinforcement learning: a survey", *Journal of Artificial Intelligence Research*, No. 4, pp. 237- 285.
- [30] Watkins, C., and Dayan, P., 1992, "Q-Learning", *Machine Learning*, Vol. 8, pp. 279-292.
- [31] Singh, S., Jaakkola, T., Littman, M. L., and Szepesvari, C., 2000, "Convergence results for single-step on-policy reinforcement learning algorithms", *Machine Learning*, Vol. 39, pp. 287-308.
- [32] Tsitsiklis, J. N., 1994, "Asynchronous stochastic approximation and Q-learning", *Machine Learning*, Vol. 16, pp.85-202.
- [33] Baird, L. C., 1995, "Residual algorithms: Reinforcement learning with function approximation", *Proceedings of 12th International Conference on Machine Learning*, California, pp. 30-37.
- [34] E. Alpaydin., 2004, "Introduction to Machine Learning", the MIT Press, Cambridge, Massachusetts.
- [35] Derhami, V., Majd, V.J., and Nili Ahmadabadi, M., 2010, "Exploration and Exploitation Balance Management in Fuzzy Reinforcement Learning", *Fuzzy Sets and Systems*, Elsevier, Vol. 161, No. 4, pp. 578-595.
- [36] Second Annual Reinforcement Learning Competition <http://rl-competition.org>.
- [37] Spong, M. W., 1995, "The swing up control problem for the acrobat", *IEEE Control Systems Magazine*, Vol. 15, pp. 49-55.
- [38] Xin Xu, Dewen Hu, and Xicheng Lu, 2007, "Kernel-Based Least Squares Policy Iteration for Reinforcement Learning", *IEEE Transactions on Neural Networks*, Vol. 18, No. 4, pp. 973-992.
- [12] Potocnil, P., and Grabec, I., 2000, "Adaptive self-tuning neurocontrol", *Mathematics and Computers in Simulation*, Vol.51, pp.201-207.
- [13] Hwang, K.S., Tan, S.W., and Tsai, M. C., 2003, "Reinforcement Learning to Adaptive Control of Nonlinear Systems", *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, Vol.33, No.3, pp.514-521.
- [14] Bhatnagar, S., Sutton, R.S., Ghavamzadeh, M., and Lee, M., 2008, "Incremental natural actor-critic algorithms", *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, pp. 105-112.
- [15] Hwang, K.S., Tan, S.W., and Tsai, M. C., 2003, "Reinforcement Learning to Adaptive Control of Nonlinear Systems", *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, Vol.33, No.3, pp.514-521.
- [16] Derhami, V., Majd, V.J., and Nili Ahmadabadi, M., 2008, "Fuzzy Sarsa learning and the proof of existence of its stationary points", *Asian journal of control*, John Wiley InterScience, Vol. 10, No. 5, pp.535-549.
- [17] Jouffe, L., 1998, "Fuzzy inference system learning by reinforcement methods", *IEEE Transactions on Systems, Man and Cybernetics-part C*, Vol. 28, No. 3, pp. 338-355.
- [18] Sun, W., Wang, X., and Yuhu, C., 2008, "Reinforcement Learning Method for Continuous State Space Based on Dynamic Neural Network", *proceedings of the 7th World Congress on Intelligent Control and Automation*, June 25 - 27, Chongqing, China.
- [19] Cetina, V.U., 2008, "Multilayer Perceptrons with Radial Basis Functions as Value Functions in Reinforcement Learning", *European symposium on Artificial Neural Networks- Advances in Computational Intelligence and Learning*, Bruges, Belgium.
- [20] FRÄMLING, K., 2008, "Light-weight reinforcement learning with function approximation for real-life control tasks", In: *Proceedings of 5th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, Funchal, Madeira, Portugal, pp. 127-134.
- [21] De Farsis, D.P., and Van Roy, B., 2000, "On the existence of fixed points for approximate value iteration and temporal-difference learning", *Journal of Optimal Theory and Application*, Vol. 105, No. 3, pp. 25-36.
- [22] Gordon, G. J., 2000, "Reinforcement Learning with function approximation converges to a region", *Proc. 8th Int. Conf. Neural Inf. Process. Syst.*, Colorado, pp. 1040-1046.
- [23] Perkins, T.J., and Precup, D., 2000, "A convergent form of approximate policy iteration", *Proc. 9th Int. Conf. Neural Inf. Process. Syst.*, Singapore, pp. 1595-1602.

حل مساله می نیمم های محلی در مسیر یابی به روش میادین پتانسیل مجازی و بر اساس دیدگاه حرکت مایعات

بیژن معاونی^۱، داوود قنبری گل^۲

^۱استادیار دانشکده مهندسی راه آهن، دانشگاه علم و صنعت ایران، تهران، ایران. b_moaveni@iust.ac.ir

^۲دانشجوی کارشناسی ارشد برق-کنترل، دانشگاه آزاد واحد علوم و تحقیقات، تهران، ایران. Davoodgol_34@yahoo.com

(تاریخ دریافت مقاله ۱۳۸۹/۱۱/۱۳، تاریخ پذیرش مقاله ۱۳۹۰/۳/۱)

چکیده: یکی از مهمترین مسائل در سیستم های چند عامله متحرک هوشمند، مسیریابی و عدم برخورد آنها با موانع در طی مسیر است. به همین دلیل در این زمینه تحقیقات گسترده ای صورت گرفته و راه کارهای متنوعی نیز ارائه شده است. هر کدام از این راه کارها، مولود دیدگاه خاص ارائه دهندگان آن نسبت به تغییر نوع و تعریف مسأله بوده است. که این امر موجب شده بسته به شرایط داخلی مسأله، مانند نوع موانع، قابلیت های عامل هوشمند و ... این راه کارها نقاط قوت و ضعفی نسبت به یکدیگر داشته باشند. روش توابع پتانسیل یکی از روش های مطرح برای حل این مساله است. نقطه ضعف عمده این روش قرار گرفتن در مینیمم های محلی است که در سال های اخیر راه حل های مختلفی برای آن پیشنهاد گشته است. در این مقاله نیز راه حل جدیدی برای حل مسأله می نیمم های محلی در مسیریابی عامل های متحرک مبتنی بر روش توابع پتانسیل مجازی و با اجتناب از برخورد با موانع ارائه می گردد. دیدگاه ارائه شده در این مقاله بر اساس تحلیل و فرموله نمودن مسیر یابی مایعات شکل گرفته و همان اندازه که جریان مایعات در مسیریابی و پشت سر گذاشتن انواع موانع توانمند هستند، راه کار پیشنهادی نیز توانمند می باشد و رسیدن به نقطه هدف و عدم برخورد با موانع را در طول مسیر تضمین می نماید. در این مقاله همچنین، الگوریتم اصلاح شده ای، به منظور استفاده بلادرنگ، از روش پیشنهادی ارائه گردیده است. شبیه سازی هایی نیز به منظور مشخص شدن قابلیت ها و توانمندی های روش های پیشنهادی در این مقاله ارائه شده اند.

کلمات کلیدی: عامل متحرک، مسیریابی، عدم برخورد، توابع پتانسیل، می نیمم های محلی.

Solving the Local Minimum Problem in Path Planning Algorithm based on the Virtual Potential Field and the Principles of Liquid Movement

Bijan Moaveni, Davoud Ghanbari Gol

Abstract: Path planning and obstacle avoidance are important problems in intelligent mobile agent systems. In this paper, a novel path planning algorithm and a new obstacle avoidance approach to solve the local minimum problem of mobile agents are presented. This path planning algorithm is introduced based on the principles of liquid movement. In addition to, by modifying the potential function methodology for this algorithm, a new obstacle avoidance approach is presented. This new algorithm, by finding the collision free path, guarantees the global convergence to the target when the target is reachable. Also, in this paper, an online version of the algorithm is presented. Simulation results are given to show the effectiveness of the methodology.

Keywords: Mobile Agent, Path Planning, Obstacle Avoidance, Potential Functions.

۱- مقدمه

تحقیقات عمده‌ای برای حل مسأله، مسیریابی عاملهای متحرک به منظور رسیدن به نقطه هدف بدون برخورد با موانع صورت گرفته است و راه کارهای مختلفی نیز ارائه شده است. از جمله راه کارهای پر کاربرد در این زمینه، تحلیل و ارائه راه حل مسأله بر اساس میدان های پتانسیل مجازی است، که در آن یک میدان پتانسیل دافعه برای موانع و یک میدان پتانسیل جاذبه برای نقطه هدف در نظر می گیرند و سپس توسط یک الگوریتم می نیم یابی، حرکت به سوی می نیم مطلق که همان نقطه هدف است را دنبال می نمایند [1-3] [6] [16-22]. الگوریتم A^* یکی از روش های مطرح در این حوزه بوده است که به نوعی مدل گسسته ای از روش میادین پتانسیل است و در یافتن کوتاهترین مسیر کاربرد دارد [23] [26]. از جمله دیگر روشهای مهم و مطرح در حرکت به سمت هدف با اجتناب از برخورد با موانع می توان به روش ارائه شده در [4] اشاره نمود. در این روش از روش کنترل حلقه بسته مبتنی بر فیدبک خطا، که یک روش پایه ای در مهندسی کنترل می باشد، استفاده شده است که براساس آن ورودی های محرک عامل، به گونه ای اعمال می شوند که به کمترین خطای نهایی، که همان فاصله عامل از هدف است، برسند. روش دیگری که در مقالات [5] و [7] پیشنهاد شده مبتنی بر برنامه ریزی خطی و غیرخطی قیود حاصل از موانع و مکانیسم حرکتی عامل برای رسیدن به هدف در کمترین زمان یا کمترین جابجایی بوده است. روش متداول دیگر در حل مسأله مسیریابی، که بر اساس نوعی پردازش تصویری از ورودی سنسورهای تشخیص دهنده موانع است، در مقالات [8-10] و [12-13] بیان شده است. در این روش با پردازش تصاویر دریافتی از سنسورها ی نصب شده بر روی عامل، فضاها ی خالی و قابل عبور بین موانع تشخیص داده شده و بر این اساس حرکت به سمت هدف مسیر دهی می شود. در برخی تحقیقات از جمله [24-25]، از محاسبات دینامیک سیالات (CFD) که یک فرمولاسیون متداول از دینامیک سیالات می باشد و هدف اصلی آن پیش بینی دقیق جریان سیالات، انتقال حرارت و عکس العملهای شیمیایی در سیستم های پیچیده است، برای مسیریابی عامل های متحرک استفاده شده است. روشهای مختلف دیگری نیز به صورت خاص وجود دارد که صرفاً برای شرایط خاصی پیشنهاد شده اند [14-15]، که یک نمونه از آنها که در [11] ارائه شده است، برای حل مساله عدم برخورد با موانع بزرگ مناسب است. لازم به ذکر است که با توجه به اهمیت زیاد مساله مسیریابی عامل های متحرک، روشها و الگوریتم های متنوعی نیز در این زمینه در داخل کشور ارائه شده

است [26-28]. ولیکن در این مقاله بیشتر به معرفی روشهای متداول و/یا مرتبط با مطالب پیشنهادی در مقاله حاضر خواهیم پرداخت که با توجه با حجم بالای تحقیقات در این موضوع منطقی به نظر می رسد. پر واضح است که هر یک از روشهای اشاره شده در بالا دارای نقاط قوت و وضعی هستند ولیکن از آنجایی که روش ارائه شده در این مقاله قرار است به نوعی اصلاح شده روش های مبتنی بر میادین پتانسیلی باشد، به نقاط ضعف این دسته از روش ها که در [1] نیز به آنها اشاره شده است، خواهیم پرداخت. مهمترین نقاط ضعف روش های مبتنی بر توابع پتانسیل عبارتند از:

- ۱- متوقف شدن عامل در می نیم های محلی.
 - ۲- ناتوانی در مسیر یابی در مواقعی که موانع در نزدیکی هدف باشند.
 - ۳- حرکت نوسانی عامل در مسیر های باریک مابین موانع.
- البته تحقیقاتی نیز جهت حل این مشکلات صورت گرفته و راه حل هایی نیز پیشنهاد شده است [1]. به طور ویژه در زمینه رفع مشکل توقف در می نیم های محلی روش میادین پتانسیل، تحقیقات فراوانی حتی در سالهای اخیر انجام گرفته است [19-22]، که یکی از نتایج اصلی تحقیقات صورت گرفته در این مقاله نیز ارائه روش نوین دیگری در این زمینه با استفاده از دیدگاه حرکت مایعات می باشد.
- در مقاله حاضر سعی بر این است تا با ارائه یک دیدگاه نو به مسأله مسیریابی و استفاده از راه کار میادین پتانسیل مجازی، الگوریتم جدیدی برای حل مسأله می نیم های محلی در مسیریابی، بدون برخورد با موانع معرفی گردد، که بر اساس آن روش نوینی نیز برای خروج از نقاط می نیم محلی ارائه خواهد شد. این دیدگاه نو، بر اساس تحلیل مکانیسم حرکت مایعات ایجاد شده است، که ناشی از تمایل مایعات برای رفتن به سطوح پایینتر به علت نیروی جاذبه می باشد. مایعات توانایی بالایی در پشت سرگذاشتن انواع موانع موجود در مسیر حرکت خود به سمت سطوح پایینتر دارند، که براساس تجربیات طبیعی در صورت مکفی بودن حجم مایع دست یافتن به می نیم مطلق قطعی است. با توجه به این توانایی مایعات، تحلیل حرکت مایعات می تواند در بدست آوردن یک الگوریتم مسیریابی توانمند بسیار مفید باشد. در ادامه ابتدا مکانیسم حرکت و مسیر یابی مایعات مورد بررسی قرار گرفته و فرموله می شود و سپس بر اساس آن الگوریتم مسیر یابی مدنظر در دو بخش ارائه می گردد. در قسمت شبیه سازی، نتایج حاصل از این الگوریتم در شرایط مختلف مورد بررسی قرار خواهد گرفت و توانمندیهای این روش مشخص می گردد.

۲- مدل سازی حرکت مایعات

جهت مدل سازی و تحلیل دقیق حرکت و مسیریابی مایعات، مایع را به جای یک مجموعه به هم پیوسته، به صورت یک سری قطرات یا ذرات به اندازه کافی کوچک و در کنار هم در نظر خواهیم گرفت. بر این اساس، حرکت مایعات را می توان اینگونه تحلیل نمود که ابتدا اولین قطره از مایع طبق قاعده حرکت به سمت سطوح پایتتر پیرامونی شروع به حرکت می نماید و اگر در هیچ می نیم محلی، که ممکن است توسط موانع ایجاد شده باشد، قرار نگیرد تا رسیدن به پایتترین سطح یا همان نقطه هدف به حرکت خود ادامه خواهد داد. اما اگر در مسیر حرکت به یک می نیم محلی برخورد کند، در آنجا متوقف شده و باقی می ماند. حضور این قطره در می نیم محلی خود باعث ایجاد یک اختلاف سطح در آن نقطه می شود، به همین دلیل شرایط محیط برای حرکت قطرات بعدی تغییر خواهد کرد. این تغییر محیط می تواند با تداوم به دام افتادن قطرات در آن نقطه ادامه یابد تا جایی که قطرات بعدی بتوانند طبق قاعده حرکت مشابه، از کنار این محل عبور نمایند و به حرکت خود به سمت سطوح پایتتر ادامه دهند. بعد از آنکه اولین قطره توانست خود را به پایتترین نقطه برساند، از این پس بقیه قطرات مایع طبق مسیر حرکت این قطره حرکت خواهند نمود. با کمی دقت متوجه می شویم که قطرات مایع در نقاط می نیم محلی خود نقش مانع را بازی می کنند و با حضور خود مسیر حرکت قطرات بعدی به سمت می نیم مطلق را هموارتر می نمایند.

با الگو برداری از این تحلیل می توان یک الگوریتم نو برای حل مسأله مسیریابی کلیه عاملهای متحرک ارائه نمود. قدم اول ایجاد یک محیط مجازی مناسب جهت حرکت عامل، مشابه حرکت یک مایع، به سمت هدف است که این محیط با تعریف میداین پتانسیل مجازی مناسب برای موانع و نقطه هدف بدست می آید. لذا برای داشتن یک محیط مجازی مناسب توسط میداین پتانسیل، لازم است دو شرط زیر برآورده گردند:

۱- سطح پتانسیل نقطه هدف پایتترین سطح باشد.

۲- سطح پتانسیل موانع به اندازه کافی بزرگتر از نقاط پیرامونشان باشد.

۲-۱- تعریف میداین پتانسیل

همانگونه که در بخش قبل اشاره شد، لازم است جهت حل مسأله توابع پتانسیل مناسبی تعریف گردند. در اینجا نمونه ساده ولی راه گشا از توابع پتانسیل که بتواند چنین میدانی را ایجاد نماید ارائه می گردد. در

معادله (۱) میداین پتانسیل دافعه برای موانع معرفی می گردند که از توابع پیشنهادی مقالات [1] و [3] الگو برداری شده اند.

$$P_O = \frac{K_O}{r_o^2} \quad (1)$$

که

$$r_o = \begin{cases} R & R \square . \\ . & R < . \end{cases} \quad (2)$$

و

$$R_O = \sqrt{(O_X - X)^2 + (O_Y - Y)^2} - O_r - A_r \quad (3)$$

در معادله (۳)، A_r شعاع عامل، O_r شعاع مانع، زوج (O_x, O_y) موقعیت مانع، زوج (X, Y) موقعیت عامل و K_O ضریب تناسبی پتانسیل مانع می باشند. لازم به ذکر است که در این تابع پتانسیل سطح مقطع موانع و عامل را با تقریب دایره فرض شده است. پر واضح است که این فرض محدود کننده نیست، چرا که به موانع کوچک با سطح مقطع غیر دایره، به سادگی می توان یک دایره محیط نمود و یا اینکه موانع بزرگ را با کنار هم قرار دادن موانع دایره ای کوچکی جایگزین کرد. لذا، با توجه به سطح مقطع عامل که دایره ای فرض شده است، حداقل فاصله ای که عامل متحرک می تواند با مانع داشته باشد به اندازه شعاع عامل است. در اینجا مشابه شکل ۱، برای خروج شعاع عامل از محاسبات آتی، این شعاع را به شعاع موانع افزوده ایم تا با این شرایط فاصله بین عامل و موانع به طور خودکار در محاسبات دخیل باشد. با این عمل، شعاع موانع به اندازه شعاع عامل بیشتر شده و شعاع عامل متحرک صفر در نظر گرفته می شود. لذا معادله (۳) را می توان به صورت (۴) بازنویسی نمود. که صفر بودن شعاع عامل به معنی نقطه ای بودن و نقطه ای عمل کردن عامل در طی حرکت در مسیر است. این امر باعث تسهیل محاسبات بدون از دست دادن قابلیت پیاده سازی عملی الگوریتم خواهد شد.

$$R_O = \sqrt{(O_X - X)^2 + (O_Y - Y)^2} - O'_r \quad (4)$$

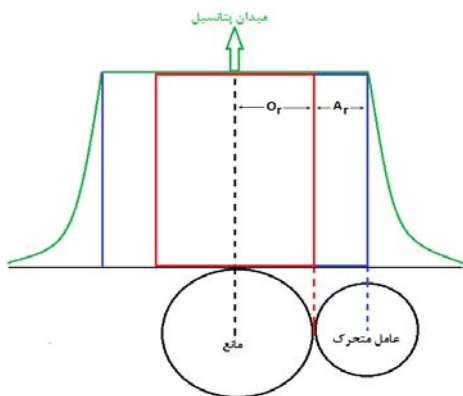
$$O'_r = O_r + A_r$$

از طرفی میداین پتانسیل هدف به صورت (۵) قابل بیان خواهد بود.

$$P_t = K_t r_t \quad (5)$$

$$r_t = \sqrt{(T_x - X)^2 + (T_y - Y)^2}$$

که در آن، زوج (T_x, T_y) موقعیت هدف، زوج (X, Y) موقعیت عامل، K_t ضریب تناسبی پتانسیل هدف و P_t میداین پتانسیل هدف در هر نقطه از محیط می باشد. با این تعاریف میدان کل حاصل از موانع و هدف (\bar{P}) در هر نقطه از محیط، به صورت مجموع جبری مقدار تابع

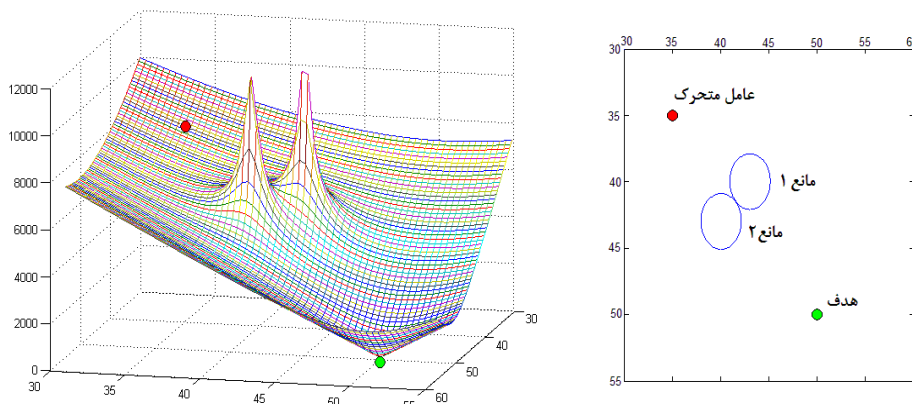


شکل ۱: میدان پتانسیل موانع (منحنی سبز رنگ) توان میدان پتانسیل را به صورت تابعی از فاصله از مانع بیان میکند.

پتانسیل حاصل از هدف و مقدار تابع پتانسیل هر مانع در معادله (۶) قابل بیان خواهد بود.

$$\bar{P} = P_r + \sum_{j=1}^N P_{o,j} \quad (6)$$

که N تعداد موانع در محیط و $P_{o,j}$ مقدار تابع پتانسیل مانع j ام می باشد. در بخش سمت چپ شکل ۲ یک فضای سه بعدی برای دو مانع فرضی و یک هدف، که موقعیت آنها در محیط واقعی در قسمت راست همین شکل دیده می شود، تهیه شده است که می تواند دید مناسبی از شرایط محیط مجازی در اختیار قرار دهد. در این شکل به خوبی وجود شرایط دوگانه گفته شده برای میدان پتانسیل مناسب، قابل مشاهده است



شکل ۲: نمایش سه بعدی میدان پتانسیل محیط برای دو مانع و یک هدف (عامل متحرک با رنگ قرمز و هدف با رنگ سبز مشخص است).

در این روابط $Step$ ، گام یا دامنه حرکتی عامل در هر مرحله از الگوریتم است که حداقل آن با محدودیتهای موجود در طراحی و ساخت مکانیسم سخت افزاری و نرم افزاری عامل محدود می شود و حد اکثر آن نیز باید به گونه ای باشد که باعث عبور عامل از روی موانع در فضای مجازی نگردد. لذا:

$$O'_r = (A_r + O_r) \ll Step \ll$$

در رابطه (۷)، از پارامتر $m = 0, 1, \dots, A_{nsd}$ به منظور تعیین زاویه یا جهت حرکت عامل استفاده می شود. همچنین A_{nsd} نشان دهنده تعداد زاویه های ممکن برای جهت گیری عامل است. حداقل این پارامتر نیز توسط محدودیتهای ذاتی مکانیسم حرکتی عامل محدود شده است. بدیهی است برای اینکه عامل قابلیت مسیر یابی را داشته باشد این پارامتر حداقل باید $3 (A_{nsd} > 3)$ باشد. در شکل ۳ زوایای قابل انتخاب عامل، برای مثالی با $A_{nsd} = 12$ مشاهده می شود.

۲-۲- دینامیک حرکتی عامل

پیش از بیان الگوریتم مسیر یابی و پس از ارائه معادلات مربوط به میادین پتانسیل، لازم است نحوه حرکت عامل در محیط نیز فرموله و تعریف گردد. در این مقاله سعی شده است یک دینامیک ساده برای حرکت عامل معرفی گردد که با کمترین پارامتر بتواند برای گستره زیادی از انواع عامل های متحرک استفاده شود. به این منظور موقعیت جدید عامل که با زوج (X_{new}, Y_{new}) نمایش داده می شود بر اساس موقعیت قبلی عامل (X_{old}, Y_{old}) دامنه حرکت و جهت حرکت به صورت (۷) قابل بیان خواهد بود.

$$Y_{new} = Y_{old} + \left(Step \cdot \cos \left(m \frac{2\pi}{A_{nsd}} \right) \right) \quad (7)$$

$$X_{new} = X_{old} + \left(Step \cdot \sin \left(m \frac{2\pi}{A_{nsd}} \right) \right)$$

$$P = \bar{P} + \sum_{i=1}^M P_{v,i} \quad (10)$$

که در آن M تعداد موانع مجازی موجود در محیط، $P_{v,i}$ مقدار تابع پتانسیل مانع مجازی "ام" و \bar{P} تابع پتانسیل موانع واقعی و هدف است که در (۶) معرفی شد. البته لازم به ذکر است که در تعریف تابع پتانسیل موانع مجازی، لزومی ندارد که توابع پتانسیل شبیه توابع موانع واقعی تعریف شوند، بلکه می توان بسته به شرایط حاکم بر مسیریابی متفاوت بوده و فقط دارا بودن شرایط گفته شده برای توابع پتانسیل، الزامی است. برای مثال، هرگاه تراکم موانع زیاد نباشد، جهت بالا بردن سرعت مسیریابی می توان از توابع پتانسیل با ضرایب تناسبی بزرگتر برای موانع مجازی استفاده نمود و/یا اینکه برای بهبود مسیریابی در محیطی با موانع متراکم، جهت داشتن مسیرهای کوتاهتر، می توان از توابعی با ضرایب تناسب کوچکتر برای موانع مجازی سود برد.

۳- مسیر یابی به سمت هدف

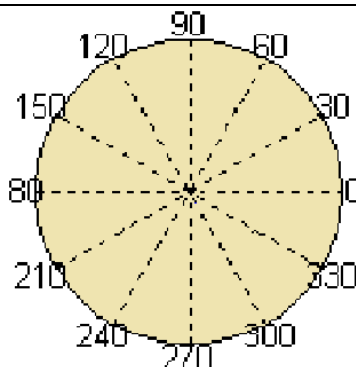
با توجه به اساس تعریف میدانی پتانسیل، حرکت عامل در هر گام می بایست بر اساس گرادیان نزولی باشد. به این معنی که عامل لازم است در جهتی گام بردارد که به سطح پتانسیل پایینتری نسبت به هر جهت انتخابی دیگر برسد. این اصول حرکتی برای یک عامل متحرک با دینامیک معرفی شده در بخش پیشین (بخش ۲-۲) که در آن تعداد زوایای قابل انتخاب (A_{nsd}) برای تعیین جهت گام برداشتن محدود می باشد، الگوریتم ساده ای خواهد داشت. بدین صورت که، ابتدا با استفاده از معادلات (۷) و قرار دادن مقدار m از ۰ تا A_{nsd} ، مختصات کلیه نقاطی که عامل می تواند با برداشتن یک گام برسد بدست می آید. حال با محاسبه سطح پتانسیل حاصل از هدف و کلیه موانع واقعی و مجازی موجود در این نقاط و بدست آوردن می نیم این سطوح پتانسیل، جهت مناسب بر اساس گرادیان نزولی را که همان زاویه مربوط به این سطح پتانسیل می نیم است محاسبه می گردد. البته باید توجه داشت که اگر این سطح پتانسیل می نیم از سطح پتانسیل محل فعلی عامل بزرگتر باشد، آنگاه محل فعلی عامل یک می نیم محلی است.

۳-۱- الگوریتم مسیریابی

حال بر اساس توابع پتانسیل معرفی شده در بالا می توان الگوریتمی را برای مساله مسیریابی بدون برخورد با موانع به صورت زیر بیان نمود، که در شکل ۴ نیز فلوجارت کامل آن رسم شده است:

۱- تعریف میدان پتانسیل مناسب برای هدف (۵)، موانع واقعی (۱)

و موانع مجازی (۸).



شکل ۳: زوایای قابل انتخاب (بر حسب درجه) برای عامل متحرک با

$$A_{nsd} = 12$$

۳-۲- موانع مجازی

در بخش پیشین نحوه تعریف محیط مجازی و توابع پتانسیل مربوط به نقطه هدف و موانع واقعی و همچنین معادلات مربوط به دینامیک حرکتی عامل ارائه گردید. حال در محیط پتانسیلی تعریف شده، می توان بر اساس منطق حرکتی مایعات که حرکت به سمت سطوح با پتانسیل پایینتر است حرکت را آغاز نمود و هر جا که عامل با می نیم محلی برخورد نماید یک مانع مجازی با یک تابع پتانسیل معین قرار داده و حرکت را از نو آغاز نمود، تا جایی که دیگر عامل در هیچ می نیم محلی قرار نگرفته و به نقطه هدف که می نیم مطلق است برسد و در نهایت مسیری که منجر به رسیدن به نقطه هدف می شود یک مسیر مناسب برای حرکت عامل در محیط واقعی است. به این منظور لازم است تابع پتانسیل موانع مجازی نیز تعریف و به فرمولاسیون پیشین اضافه گردد. لذا، به طریق مشابه تابع پتانسیل موانع مجازی را می توان به صورت (۸) در نظر گرفت.

$$P_v = \frac{K_v}{r_v^2} \quad (8)$$

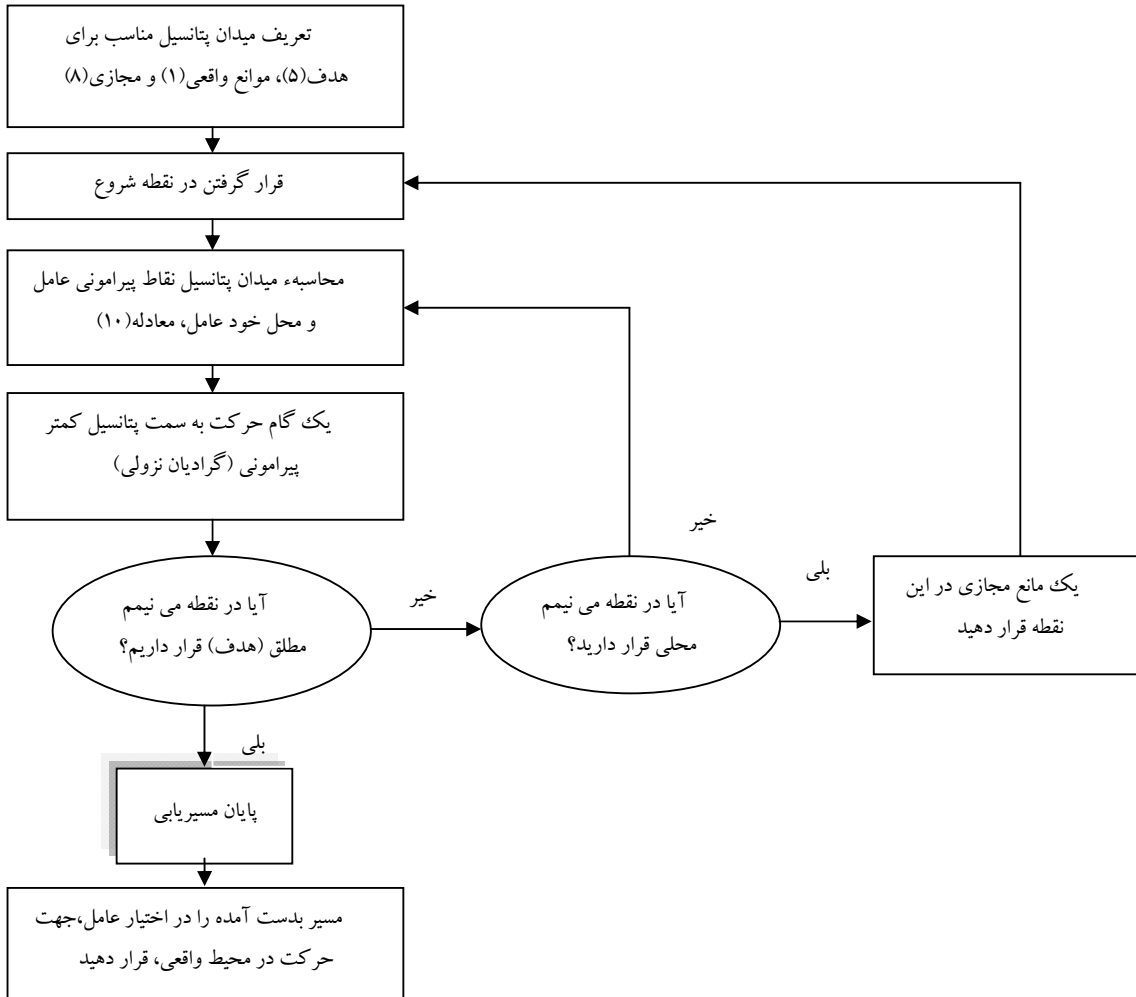
$$r_v = \begin{cases} R_v & R_v \geq . \\ . & R_v < . \end{cases}$$

$$R_v = \sqrt{(V_x - X)^2 + (V_y - Y)^2} - V_r' \quad (9)$$

$$V_r' = V_r + A_r$$

A_r شعاع عامل، V_r شعاع مانع مجازی، زوج (V_x, V_y) موقعیت مانع، (X, Y) موقعیت عامل و K_v ضریب تناسبی پتانسیل مانع مجازی می باشد. در نتیجه با لحاظ نمودن موانع مجازی، میدان پتانسیل کل در هر نقطه از محیط و با استفاده از (۶) و (۸) به صورت (۱۰) در خواهد آمد.

- ۲- قرار گرفتن در نقطه شروع.
 ۳- محاسبه میدان پتانسیل، (۱۰)، نقاط پیرامونی عامل و محل خود عامل.
 ۴- یک گام حرکت به سمت پتانسیل کمتر پیرامونی (گرادیان نزولی).
 ۵- اگر در می نیم محلی قرار گرفتید یک مانع مجازی در آن نقطه قرار داده و به مرحله ۲ بازگردید.
 ۶- اگر به نقطه هدف رسیدید الگوریتم را متوقف ساخته و مسیر بدست آمده را در اختیار عامل قرار دهید. در غیر این صورت به مرحله ۳ بروید.



شکل ۴: فلوجارت الگوریتم مسیریابی

۲-۳- بررسی همگرایی الگوریتم

پیش از پرداختن به بررسی همگرایی الگوریتم، در اینجا به بیان ریاضی عملکرد الگوریتم خواهیم پرداخت. لذا با تعاریف ارائه شده در (۱۱)، Q نشان دهنده موقعیت عامل در محیط است و نقطه شروع حرکت $Q(1)$ خواهد بود.

$$Q(k+1) = \begin{bmatrix} X_{new} \\ Y_{new} \end{bmatrix}$$

(۱۱)

$$Q(k) = \begin{bmatrix} X_{old} \\ Y_{old} \end{bmatrix}$$

برای شروع الگوریتم ابتدا توسط رابطه (۱۲) موقعیت های مجاور نقطه کنونی عامل متحرک محاسبه می گردد.

۱-۱- در یک نقطه می نیم محلی قرار داریم. بر اساس اینکه عملکرد اصلی الگوریتم در نقاط می نیم محلی افزایش مقدار پتانسیل آن طبق بخش دوم رابطه (۱۳) است، لذا این امر به طور قطع، باعث خروج عامل از می نیم محلی می شود، و در نتیجه Q_{∞} نمی تواند یک می نیم محلی باشد.

۲-۱- هیچ مسیری برای به سمت هدف وجود ندارد. این به معنی در حصر کامل بودن عامل متحرک و/یا نقطه هدف در میان موانع است که مخالف فرض ۲ قضیه می باشد.

۳-۱- فاصله مابین موانع به گونه ایست که عامل متحرک قادر به عبور از میان آنها نیست. این حالت نیز مخالف فرض ۱ قضیه می باشد.
۲- اگر باشد $P(Q_{\infty}) \leq P(Q_d)$ که به معنی پایتیر بودن پتانسیل یک نقطه در محیط از پتانسیل نقطه هدف است و این خلاف شرط اصلی قضیه می باشد که طبق آن پتانسیل نقطه هدف می بایست کمترین مقدار را دارا باشد.

لذا با توجه به اینکه کلیه حالات ممکن خلاف فرضها و شروط اصلی قضیه می باشد، درستی قضیه و همگرایی به نقطه هدف اثبات می شود.

۴- شبیه سازی

در این بخش به شبیه سازی الگوریتم ارائه شد در بخش قبل به منظور مسیریابی عامل متحرک، حل مساله می نیم های محلی و عدم برخورد با موانع خواهیم پرداخت. به منظور آماده نمودن مساله برای شبیه سازی ابتدا لازم است یک میدان پتانسیل مناسب طبق شرایط گفته شده در بخش ۲ انتخاب گردد. که در این شبیه سازی از توابع معرفی شده در بخش ۲ استفاده شده است. همچنین دینامیک حرکتی عامل نیز به صورت معادلات (۷) در نظر گرفته شده است.

مثال ۱:

در این قسمت نتایج شبیه سازی برای وضعیت های مختلف موانع و هدف به صورت مرحله ای نمایش داده شده و با نتایج حاصل از شبیه سازی الگوریتم پتانسیلی پیشنهادی در [۳] به عنوان یک روش پایه، مقایسه گردیده است. این نتایج در شرایط زیر و در چهار حالت مختلف صورت گرفته است.

شرایط فرضی شبیه سازی الگوریتم جدید:

$$\begin{aligned} Step &= 10, A_{nsd} = 16, K_O = 1, K_t = 35, \\ O_r &= 1.5, A_r = 0.03, K_v = 60 \end{aligned}$$

$$Q(k+1) = Q(k) + step \cdot \begin{bmatrix} \cos\left(m \frac{2\pi}{A_{nsd}}\right) \\ \sin\left(m \frac{2\pi}{A_{nsd}}\right) \end{bmatrix} \quad (12)$$

سپس توسط روابط و شرطهای قید شده در (۱۳) حرکت به سمت نقطه هدف یا Q_d را ادامه می دهیم.

$$\begin{aligned} &[\exists P(Q(k+1)) < P(Q(k))] \\ &\Rightarrow m_c = \{m | \min P(Q(k+1))\} \\ &\Rightarrow Q(k+1) = Q(k) + step \cdot \begin{bmatrix} \cos\left(m_c \frac{2\pi}{A_{nsd}}\right) \\ \sin\left(m_c \frac{2\pi}{A_{nsd}}\right) \end{bmatrix} \\ &\Rightarrow k = k+1 \Rightarrow \text{تکرار الگوریتم} \quad (13) \end{aligned}$$

$$\begin{aligned} &[\forall P(Q(k+1)) < P(Q(k))] \& [Q(k) \neq Q_d] \\ &\Rightarrow P(Q(k)) = P(Q(k)) + P_v \\ &\Rightarrow k = 1 \Rightarrow \text{تکرار الگوریتم} \end{aligned}$$

$$\begin{aligned} &[\forall P(Q(k+1)) < P(Q(k))] \& [Q(k) = Q_d] \\ &\Rightarrow \text{پایان الگوریتم} \end{aligned}$$

که در روابط (۱۳)، $P(Q(k))$ بیانگر مقدار پتانسیل محیط در موقعیت $Q(k)$ و P_v اندازه پتانسیل موانع مجازی می باشد.

قضیه: عامل متحرک در فضای پتانسیل مجازی که مقدار می نیم مطلق آن مربوط به نقطه هدف است، با استفاده از الگوریتم نشان داده شده در شکل (۴)، همواره به نقطه هدف دست پیدا می کند اگر:

۱- شعاع فرضی عامل متحرک A_r به اندازه کافی کوچک باشد به طوری که بتواند از فواصل مابین موانع عبور نماید.

۲- عامل متحرک و نقطه هدف به طور جداگانه در حصر کامل موانع نباشند.

۳- به تعداد کافی بتوانیم الگوریتم را تکرار نماییم.

اثبات: برای اثبات قضیه از برهان خلف استفاده می شود. با توجه به فرض سوم امکان تکرار الگوریتم بصورت نامحدود وجود داشته باشد ولیکن پس از بی نهایت بار هنوز عامل به نقطه هدف نرسیده باشد که این امر به صورت معادله (۱۴) قابل بازنویسی است.

$$\lim_{n \rightarrow \infty} Q(k) = Q_{\infty} \neq Q_d \quad (14)$$

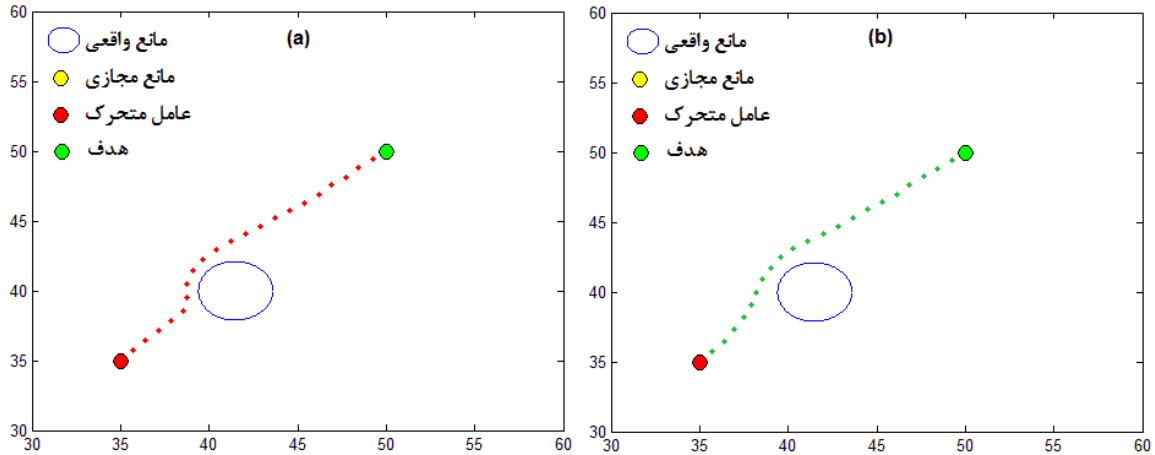
لذا، ممکن است یکی از دو مورد زیر رخ داده باشد:

۱- اگر $P(Q_{\infty}) > P(Q_d)$ باشد آنگاه:

نتیجه شبیه سازی برای چنین شرایطی را نشان می دهد. همانگونه که مشخص است هردو الگوریتم توانسته اند به راحتی مسیر رسیدن به هدف را تقریباً مشابه هم طی نمایند.

شرایط فرضی برای اجرای شبیه سازی الگوریتم پیشنهادی در [۳] عبارت است از اینکه بهره های موقعیت و سرعت برابر مقدار واحد انتخاب شده است.

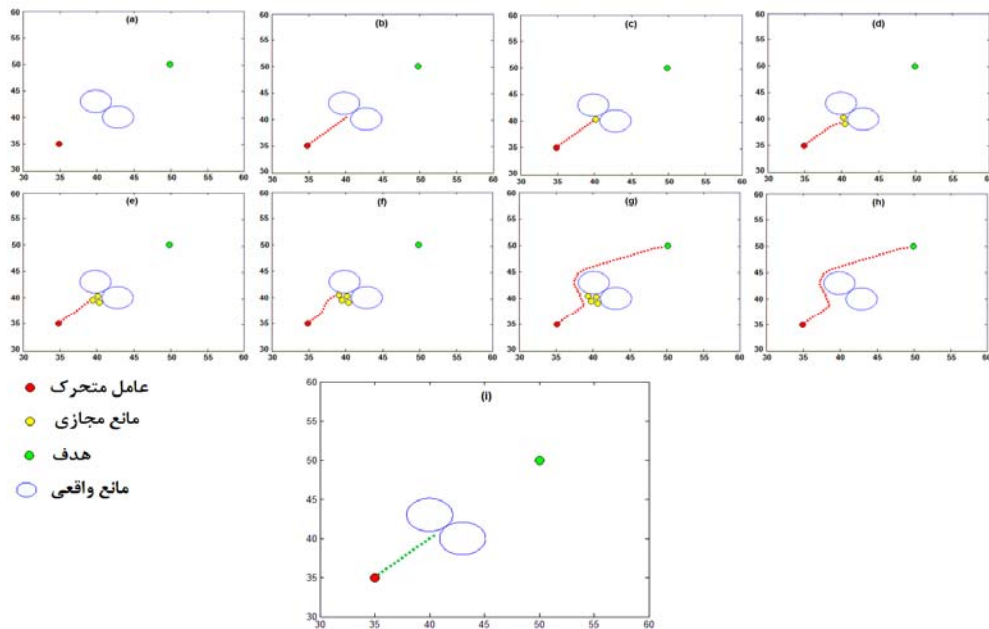
حالت ۱: برای شروع، عملکرد الگوریتم پیشنهادی را در شرایط عدم وجود می نیم های محلی مورد بررسی قرار می دهیم. شکل ۵



شکل ۵: مسیریابی بدون می نیم محلی. (a) الگوریتم جدید، (b) الگوریتم پیشنهادی در [3]

بعد از قرار دادن ۴ مانع مجازی (از شکل c-۶ تا f-۶) توانسته است از می نیم محلی فاصله گرفته و خود را به هدف برساند. در صورتیکه الگوریتم پیشنهادی [3]، همانطور که انتظار می رفت در می نیم محلی متوقف شده است (شکل i-۶).

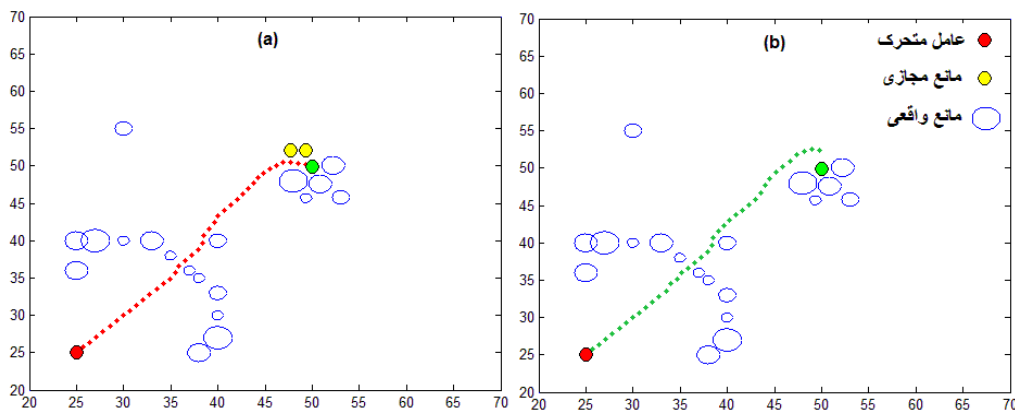
حالت ۲: در این مرحله توانمندی الگوریتم با حضور می نیم های محلی مورد بررسی قرار می گیرد. در شکل ۶، مراحل مختلف عملکرد الگوریتم برای دو مانع که نمایش سه بعدی میدان پتانسیل آنها در شکل ۲ آمده بود، به نمایش در آمده است. در این شبیه سازی عامل



شکل ۶: مسیریابی با حضور می نیم محلی. (a-h) الگوریتم جدید، (i) الگوریتم پیشنهادی در [3]

نزدیکی زیاد باعث به وجود آمدن می نیم های محلی در نزدیکی هدف می گردد و در نهایت عامل در نزدیکی هدف بدون آنکه مانعی مقابلش باشد متوقف می گردد. با توجه به نتایج شبیه سازی این اشکال در الگوریتم جدید با قرار دادن موانع مجازی، که باعث به هم خوردن محل می نیم های محلی می گردد، رفع گردیده است.

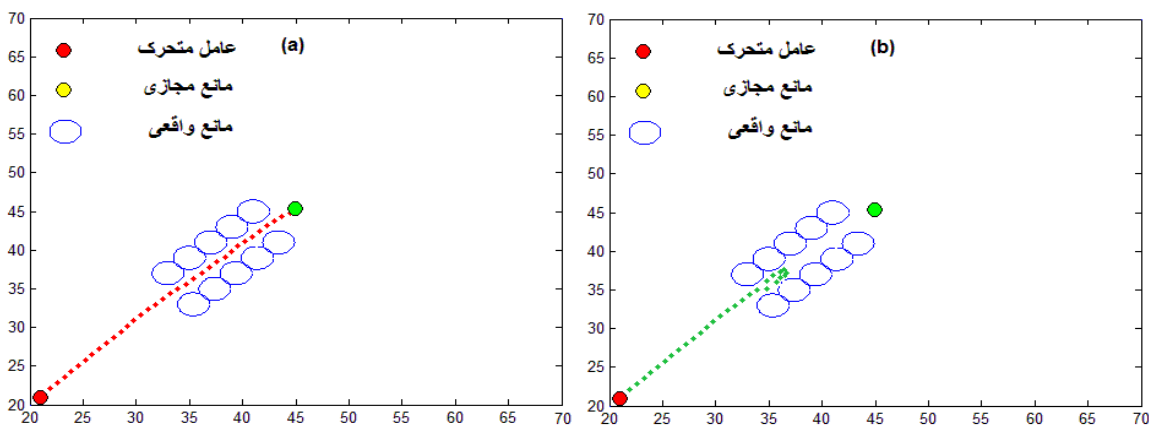
حالت ۳: در شکل ۷ عملکرد الگوریتم با حضور موانع مختلف در نزدیکی هدف و همچنین در مسیر حرکت عامل، مورد بررسی قرار گرفته است. که با توجه به مسیر تعیین شده، توانایی عبور عامل از ما بین موانع حتی در نقاط نزدیک به هدف، کاملا مشهود است. در اینجا نیز نتایج شبیه سازی الگوریتم پیشنهادی در [3] نشان دهنده متوقف شدن عامل به علت نزدیکی زیاد موانع به نقطه هدف می باشد که این



شکل ۷. مسیریابی با حضور موانع در نزدیکی هدف. (a) الگوریتم جدید، (b) الگوریتم پیشنهادی در [3]

از محاسبه گرادیان میدان پتانسیل دچار فرافهش شده و باعث ایجاد نوسان در حرکت عامل متحرک می شود که ممکن است مانند نتایج شبیه سازی در این قسمت، این نوسانات وارد یک حلقه بسته شده و باعث توقف حرکت عامل متحرک شود.

حالت ۴: توانمندی مسیریابی الگوریتم در مسیرهای باریک و با حضور هدف محصور با موانع نیز در شکل ۸ بررسی شده است. در الگوریتم جدید، عامل با عبور از مسیر باریک ما بین موانع توانسته خود را به نقطه هدف برساند (a-8) در صورتیکه در الگوریتم پیشنهادی در [3] به علت فاصله کم بین موانع و مسیر حرکت باریک، نیروی حاصل



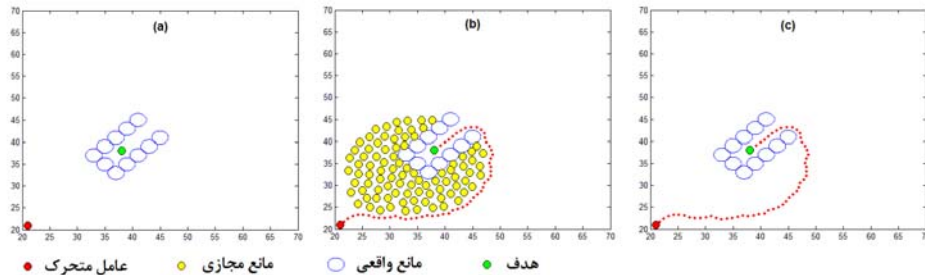
شکل ۸. مسیریابی در مسیرهای باریک با حضور هدف محصور در موانع. (a) الگوریتم جدید، (b) الگوریتم پیشنهادی در [3]

محلی (شکل ۶)، مسیر یابی در شرایط نزدیک بودن موانع به هدف (شکل ۷) و نوسانات حرکتی عامل در مسیرهای باریک (شکل ۸) مورد بررسی قرار گرفت. توجه به نتایج شبیه سازی ها نشان دهنده این واقعیت

در این شبیه سازی ها (در ۴ حالت مختلف) توانمندی روش پیشنهادی در این مقاله برای حل سه مشکل اصلی روش های مبتنی بر توابع پتانسیل اشاره شده در ابتدای مقاله یعنی، گریز از می نیم های

دادن توانمندی مسیریابی این الگوریتم در شرایط پیچیده تر، نشان دهنده شباهت زیاد عملکرد این الگوریتم با رفتار مایعات در برخورد با موانع است.

است که این روش به خوبی توانسته است راه کار مناسبی برای حل مساله مسیریابی حرکت عامل به سمت هدف بدون برخورد و همچنین حل مساله می نیم محلی به موانع ارائه دهد. در شکل ۹ نتیجه یک شبیه سازی دیگر را برای الگوریتم جدید می بینید که در عین نشان



شکل ۹: شباهت مسیریابی الگوریتم جدید با رفتار مایعات در برخورد با موانع و مسیریابی.

۵- اگر در می نیم محلی قرار گرفتید یک مانع مجازی در آن نقطه قرارداده و به مرحله ۳ برگردید.
۶- اگر عامل به نقطه هدف رسید الگوریتم را متوقف کنید. در غیر این صورت به مرحله ۳ بروید.

سرعت محاسبات در این الگوریتم خیلی بالاتر از الگوریتم اصلی می باشد ولی با تغییر در الگوریتم اصلی، اصول مسیریابی آن، که بر مبنای حرکت مایعات بوده نیز تغییر کرده است. پس با توجه به توانایی بالای مایعات در مسیریابی، الگوریتم اصلاح شده به همان نسبت که از اصول حرکتی مایعات فاصله گرفته، از توانایی مسیریابی کمتری برخوردار شده است و در شرایطی که در می نیم های محلی محاصره شده با موانع قرار می گیرد دچار رزنانس شده و متوقف می شود و/یا بعد از طی مسیر طولانی رفت و برگشتی بین موانع، موفق به خروج از می نیم محلی می شود.

مثال ۲:

به منظور بررسی نحوه عملکرد الگوریتم مسیریابی بلادرنگ، در این مثال و در دو بخش، نحوه مسیریابی الگوریتم اولیه و اصلاح شده مقایسه می گردند.

حالت ۱: در شکل ۱۱ نتایج مسیریابی الگوریتم اصلی (شکل ۱۱-ب) و الگوریتم اصلاح شده (شکل ۱۱-ا) که در شرایط کاملا یکسان بدست آمده است نشان داده می شود. الگوریتم اصلی با قرار دادن ۱۰ مانع مجازی و طی ۲۸۸ مرحله محاسباتی، عامل توانسته بدون ورود به فضای مابین موانع، از آنها عبور کرده و در ۳۵ گام حرکتی به نقطه هدف برسد. در صورتیکه در الگوریتم اصلاح شده، عامل پس از ورود به فضای مابین موانع و طی مسیر طولانی رفت و

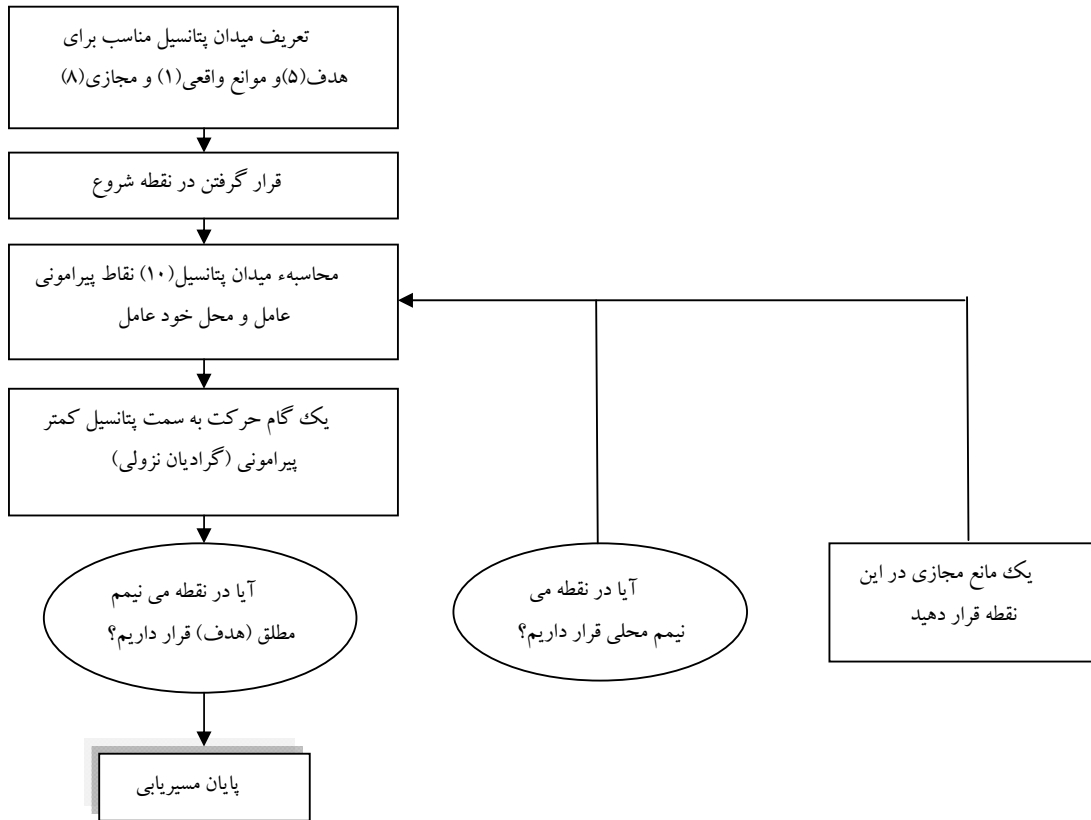
۵- الگوریتم اصلاح شده برای مسیریابی بلادرنگ عامل متحرک

در الگوریتم ارائه شده در بخش سوم این مقاله مشهود است که ابتدا می بایست در یک محیط مجازی مسیریابی انجام شود و سپس عامل شروع به حرکت نماید که این امر در محیط هایی با موانع متحرک می بایست برای هر گام حرکت عامل تکرار گردد که باعث افزایش زمان محاسبات شده و عملکرد حرکتی عامل را از حالت بلادرنگ خارج می نماید. به منظور کاهش زمان مسیریابی در این الگوریتم می توان از تغییر توابع پتانسیل و تغییر پارامترهای مربوط به مکانیسم حرکت عامل متحرک استفاده کرد. روش دیگر برای تسریع مسیریابی، اصلاح الگوریتم ارائه شده می باشد که در آن تلاش می شود که از تکرار برخی محاسبات جلوگیری شود. در الگوریتم اصلاحی ارائه شده سعی می شود تا بخش های تکراری موجود در الگوریتم اصلی، یعنی بازگشت به نقطه شروع بعد از برخورد با می نیم محلی، حذف گردد. بر این اساس الگوریتم اصلاحی به صورت زیر عمل خواهد کرد که فلوچارت مربوطه در شکل ۱۰ نیز ترسیم شده است:

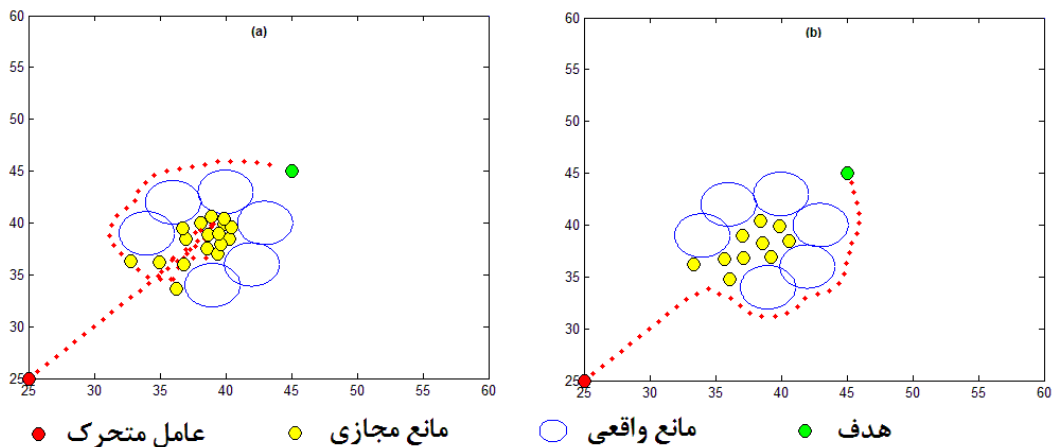
- ۱- تعریف میدان پتانسیل مناسب برای هدف (۵)، موانع واقعی (۱) و موانع مجازی (۸).
- ۲- قرار گرفتن در نقطه شروع.
- ۳- محاسبه میدان پتانسیل (۱۰)، نقاط پیرامونی عامل و محل خود عامل.
- ۴- یک گام حرکت به سمت پتانسیل کمتر پیرامونی (گردانان نزولی).

مسیری را که در عمل می بایست عامل متحرک اول طی نماید بسیار پیچیده و طولانیست و همانطور که گفته شد این پیچیدگی ممکن است در برخی شرایط باعث ناموفقیت در مسیریابی شود.

برگشتی، و همچنین قرار دادن ۱۶ مانع مجازی و انجام ۸۱ مرحله محاسباتی و گام حرکتی، موفق به مسیر یابی شده است. سرعت مسیریابی الگوریتم اصلاحی خیلی بیشتر از الگوریتم اصلی است ولیکن



شکل ۱۰: فلوجارت الگوریتم اصلاح شده.



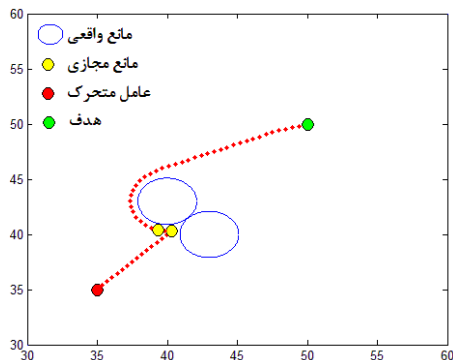
شکل ۱۱: مقایسه عملکرد الگوریتم اصلی (a) و الگوریتم اصلاح شده (b).

بودن الگوریتم اصلاح شده کاملاً محسوس است، به طوریکه عامل در الگوریتم اصلی این مسیریابی در ۱۱۰ مرحله محاسباتی و با قرار دادن ۴

در شکل ۱۲، نتیجه شبیه سازی با الگوریتم جدید برای حالت دوم مثال ۱ نشان داده شده است. با مقایسه نتایج این دو شبیه سازی سریع

افزایش توانایی مسیریابی را در پی داشته باشد. لازم به ذکر است که توانمندی الگوریتم اصلاح شده می تواند با تغییر پارامترهای الگوریتم مانند تعریف توابع پتانسیل و غیره، بسته به شرایط مسأله، شکل و موقعیت موانع، ارتقاء پیدا نماید.

مانع مجازی در ۴۷ گام حرکتی و در الگوریتم اصلاح شده در ۵۳ مرحله محاسباتی و با قرار دادن فقط دو مانع مجازی در ۳۳ گام حرکتی موفق به رسیدن به هدف می شود. لذا، در الگوریتم اصلاح شده، تغییرات صورت گرفته در الگوریتم بسته به نوع مساله می تواند کاهش یا



شکل ۱۲: مسیریابی الگوریتم اصلاح شده.

- [3] Khatib, O., 1986, "Real-Time Obstacle Avoidance for Manipulators and Mobile Robots". International Journal on Robotics Research, 5, 90-98.
- [4] Secchi, H., Mut V., Carelli, R., 1994, "Impedance Control for Mobile Robots with Sensorial Feedback". National Symposium of Automatic Control – AADECA, Argentina, 19-24.
- [5] Perdereau, V., Passi, C., Drouin, M, 2002, "Real-time Control of Redundant Robotic Manipulators for Mobile Obstacle Avoidance" Robotics and Autonomous Systems, 41, 41-59.
- [6] Nguyen, B. Q., Chuang, Y. L., Tung, D., Hsieh, Ch., Jin, Zh., Shi, L., Marthaler, D., Bertozzi, A., Murray, R. M., June 2005, "Virtual Attractive-Repulsive Potentials for Cooperative Control of Second Order Dynamic Vehicles on the Caltech MVWT". Proceeding of American Control Conference, Portland, OR, USA, 8-10.
- [7] Yang, T., Liu, Zh., Chen, H., Pei, R., 2007, "Robust Tracking Control of Mobile Robot Formation with Obstacle Avoidance". Journal of Control Science and Engineering, 2007, 1-10.
- [8] LIU, Sh. C., TAN, D. L., LIU, G. J., 2007 "Formation Control of Mobile Robots with Active Obstacle Avoidance". Acta Automatica Sincia, 33, 529-535.
- [9] Ulrich, I., Borenstein, J., May 1998, "Reliable Obstacle Avoidance for Fast Mobile Robots" . Proceedings of the IEEE International Conference on Robotics and Automation, Belgium, 16-21.
- [10] Fernández, J.L., Sanz, R., Benayas, J.A., Diéguez. A.R., 2004 "Improving collision avoidance for mobile robots in partially known environments: the

۶- جمع بندی

در این مقاله الگوریتم جدیدی بر اساس حرکت مایعات برای حل مسأله مسیر یابی عامل های متحرک ارائه گردید و تواناییهای آن جهت غلبه بر مشکلات متداول مسیریابی مورد بررسی قرار گرفت، که نتایج آن، گویای قدرت بالای این الگوریتم در مسیریابی می باشد. الگوریتم های ارائه شده در این مقاله لزوماً مسیر بهینه و/یا سریعترین را نسبت به سایر روشها ارائه نمی دهند ولیکن یافتن مسیر بهینه را تضمین می نماید که این امر با ارائه قضیه ای اثبات گردید. البته با تغییر پارامترهای دخیل در الگوریتم می توان بسته به شرایط و نوع مسأله به مسیرهای مناسب تری دست یافت. در این مقاله یک دیدگاه جدید در زمینه مسیریابی و خروج از می نیم های محلی برای عامل های متحرک که می تواند کاربردهای عملی و تئوری متنوعی در بر داشته باشد ارائه گردید و توانمندی های آن مورد بحث و بررسی قرار گرفت.

۷- مراجع

- [1] Yongjie, Y., Zhang, Y., February 2009, "Collision Avoidance Planning in Multi-robot based on Improved Artificial Potential Field and Rules". Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics Bangkok, Thailand, 22-25.
- [2] Kim, J. O., Pradeep, K. K., June 1992, "Real-Time Obstacle Avoidance Using Harmonic Potential Functions". IEEE Transactions on Robotics and Automation, 8, 338-349.

- agent states". *Robotics and Autonomous Systems*, 56, 1050-1060.
- [21] Zou, X.Y., Zhu, J., 2003 "Virtual local target method for avoiding local minimum in potential field based robot navigation". *Journal of Zhejiang University Science*, 4, 264-269.
- [22] Park, M. G., Lee., M. Ch., 2003, "A New Technique to Escape Local Minimum in Artificial Potential Field Based Path Planning". *KSME International Journal*, 17, 1876-1885.
- [23] Qu, Y.H., Pan, Q., Yan, J.G., Nov. 2005, "Flight Path Planning of UAV Based on Heuristically Search and Genetic Algorithms", *Proceeding of the 31st Annual Conference of IEEE Industrial Electronics Society*, 6-10.
- [24] Pac, M. R., 2007 "A Fluid Dynamics Framework for Control of Mobile Robot Networks". Master's Thesis, Middle East Technical University.
- [25] Gingras, D., Dupuis, E., Payre, G. de Lafontaine, J., May 2010, "Path planning based on fluid mechanics for mobile robots using unstructured terrain models", *Proceeding of the IEEE International Conference on Robotics and Automation, Anchorage, AK*, 3-7.
- [26] مریم خراشادیزاده. مسیریابی برای حرکت رباتهای متحرک با استفاده از شبکه های خود سازنده. پایان نامه کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر، ۱۳۸۳.
- [27] عزیزالله سیمین فر. مسیریابی و طرح ریزی حرکت برای ربات متحرک. پایان نامه کارشناسی ارشد، دانشکده مهندسی برق، دانشگاه صنعتی شریف، ۱۳۷۷.
- [28] علیرضا خیرخواه. مسیریابی رباتهای متحرک با استفاده از الگوریتم کولونی مورچه ها. پایان نامه کارشناسی ارشد، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، ۱۳۸۰.
- beam curvature method". *Robotics and Autonomous Systems*, 46, 205-219.
- [11] Cai, Ch., Yang, Ch., Zhu, Q., Liang, Y., Aug. 2007 "Collision Avoidance in Multi-Robot Systems". *Proceedings of the IEEE International Conference on Mechatronics and Automation August, Harbin* 5-8.
- [12] Watanabe, Y., Calisey, A.J., Johnson, E.N., August 2007, "Vision-Based Obstacle Avoidance for UAVs", *AIAA Guidance, Navigation and Control Conference and Exhibition, Hilton Head, South Carolina, USA*, 20-23.
- [13] Ku, Ch.H., Tsai, W.H., 2001 "Obstacle Avoidance in Person Following for Vision-Based Autonomous Land Vehicle Guidance Using Vehicle Location Estimation and Quadratic Pattern Classifier", *IEEE Transaction on Industrial Electronics*, 48, 205-215.
- [14] Aude, E. P. L., Lopes, E. P., Menezes, L. F. S., Mendes, F. S., Silveira, J. T. C., Silva, F. A. B., Serdeira, H., Martins, M. F., October 2003 "Real-Time Obstacle Avoidance performed by an Autonomous Vehicle throughout a Smooth Trajectory using an Electronic Stick" *Proceedings of the 2003 IEEE. Conference on Intelligent Robots and Systems, Las Vegas, Nevada, USA*, 27-31.
- [15] Pradalier, C., Hermosillo, J., Koike, C., Braillon, C., Bessihre, P., Laugier, C., April 2004 "An Autonomous Car-Like Robot Navigating Safely Among Pedestrians", *Proceeding of the IEEE International Conference on Robotics & Automation, New Orleans, LA, USA, April 26-May 1*.
- [16] Zhang X.B., Fang Y.Ch., Ma B.J., 2008 "A PFM-based Global Convergence Visual Servo Path Planner". *Acta Automation Sinica*, 34, 1250-1256.
- [17] Shimoda, Sh., Kuroda, Y., Iagnemma, K., April 2005, "Potential Field Navigation of High Speed Unmanned Ground Vehicles on Uneven Terrain". *Proceedings of the IEEE International Conference on Robotics and Automation, Barcelona, Spain*, 18-22.
- [18] Ge, S. S., Cui, Y. J., 2000, "New Potential Functions for Mobile Robot Path Planning". *IEEE Transaction on Robotics and Automation*, 16, 615-620.
- [19] Tang, L., Dian, S., Gu, G., Zhou, K., Wang, S., Feng, X., July 2010 "A Novel Potential Field Method for Obstacle Avoidance and Path Planning of Mobile Robot". In *Proceeding of, 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), Chengdu*, 9-11.
- [20] Mabrouk, M.H., McInnes., C.R. 2008 "Solving the potential field local minimum problem using internal

تعیین موقعیت بازیکنان فوتبال و استخراج همزمان مسیر آنها در مدل واقعی زمین با استفاده از روش گراف در تصاویر ویدیویی

مهرداد منافی فرد^۱، حمید عبادی^۲، حمید ابریشمی مقدم^۳

^۱ دانشجوی کارشناسی ارشد مهندسی نقشه برداری، گروه فتوگرامتری، دانشگاه صنعتی خواجه نصیرالدین طوسی، mehrtash64@yahoo.com

^۲ دانشیار، دانشکده مهندسی نقشه برداری، گروه فتوگرامتری، دانشگاه صنعتی خواجه نصیرالدین طوسی، ebadi@kntu.ac.ir

^۳ دانشیار، دانشکده مهندسی برق و کامپیوتر، گروه مهندسی پزشکی، دانشگاه صنعتی خواجه نصیرالدین طوسی، moghadam@eedt.kntu.ac.ir

(تاریخ دریافت مقاله ۱۳۹۰/۱/۱۲، تاریخ پذیرش مقاله ۱۳۹۰/۳/۲۷)

چکیده: در تجزیه و تحلیل مسابقات فوتبال، تعیین دقیق موقعیت بازیکنان از موارد کلیدی است که اساس آنالیزهای بعدی نظیر ردیابی بازیکنان قرار می گیرد. استخراج مسیر بازیکنان، اطلاعات مفیدی جهت ارزیابی نقاط قوت و ضعف تیم و بازیکنان و تنظیم استراتژی تیمی کارا به منظور بهبود عملکرد تیم در اختیار سرمربی و متخصصین قرار می دهد. از آنجا که با حرکت دوربین، بیننده فریم های قبلی را از دست می دهد، با مرتبط سازی تصاویر به هم و تعیین مسیر بازیکنان روی موزاییک تصویری ایجاد شده می توان مسیر پیوسته بازیکنان را برای بیننده به تصویر کشید. هدف اصلی این مقاله، انتقال موقعیت بازیکنان به فضای مدل زمین و استخراج همزمان مسیر حرکت تمام بازیکنان در فضای تصویر و مدل زمین است. برای این منظور، موقعیت بازیکنان با دو روش مدل گوسی ترکیبی و آدابوست استخراج شده و بعد از حذف نواحی اضافی نظیر دروازه و خطوط، جداسازی بازیکنان همپوشان، تعیین تیم بازیکنان و انتقال موقعیتها به سیستم مختصات مشترک، ردیابی در دو فضای تصویر و مدل واقعی زمین انجام می شود. در نهایت، دقت دو روش آدابوست و مدل گوسی ترکیبی در تعیین موقعیت بازیکنان روی ۷۸ فریم از ۶ ویدئوی مختلف و دقت ردیابی ارزیابی و مقایسه می شود.

کلمات کلیدی: تصاویر ویدئویی، ردیابی، مدل گوسی ترکیبی، آدابوست، بازیکنان فوتبال.

Player Localization and Tracking in Field Model Space using Graph Representation in Football Broadcast Videos

Mehrtash Manafifard, Hamid Ebadi, Hamid Abrishami Moghaddam

Abstract: Precise player localization is the key step for improved analysis such as player tracking in soccer broadcast videos. Extracting player trajectories provides some essential information for coaches and sport experts to determine weaknesses and strengths of the players and the team and to evaluate overall strategy of the game. As far as previous scenes are missed by camera motion, continuous player trajectory could be depicted by trajectory extraction on constructed mosaic. The goal of this paper is to transform player position to the field model and extract all trajectories in both image and model spaces. Therefore, players are detected using Gaussian Mixture Model and Gentle Adaboost. After removing extra regions such as goal post and lines, isolating occluded players, player labeling and transforming player positions to the same coordinate system, player tracking is carried out. Finally, the proposed player detection and tracking methods are applied for 78 frames taken from six soccer sequences.

Keywords: Video images, Track, GMM, Adaboost, Football players.

۱- مقدمه

با پیشرفت علم و فناوری و کاربردی شدن علوم در سالهای اخیر، بسیاری از کشورهای پیشرفته سعی در به کارگیری علم در آنالیز حرکات ورزشکاران خود به منظور آموزش اصولی افراد از سنین کم و بهبود عملکرد ورزشکاران حرفه ای در مسابقات مهم دارند. آنالیز ویدئوهای ورزشی از کارآمدترین روشها برای این منظور است. از آنجا که فوتبال از پر طرفدارترین ورزشهای دنیاست و برنامه های تلویزیونی مختلف در سراسر دنیا با حضور مربیان و داوران برای تجزیه و تحلیل وقایع اتفاق افتاده و نقاط ضعف و قوت تیمها و داور برگزار می شود، جستجوی روشهای اتوماتیک و دقیق برای آنالیز وقایع در ویدئوهای مسابقات فوتبال از ابزارهای کارآمد برای متخصصین و طرفداران این ورزش پر طرفدار محسوب می شود.

در سالهای اخیر تجزیه و تحلیل ویدئوهای ورزشی مختلف نظیر شنا، شیرجه، والیبال، گلف و فوتبال مورد توجه محققین قرار گرفته است. در مقاله [19] سه هدف اصلی (خلاصه سازی محتوای ویدئو، اضافه کردن اطلاعات به ویدئوی فوتبال و آنالیزهای سطح بالا) برای تحلیل ویدئوهای فوتبال مطرح شده و مروری به تحقیقات انجام گرفته در زمینه این سه هدف صورت گرفته است. خلاصه سازی، بخشهای مهم ویدئوی فوتبال را استخراج می کند. این مرحله به صورت غیر آتی انجام می گیرد و اغلب نیاز به دقت بالایی ندارد. اما در اضافه کردن اطلاعات (فواصل، خط آفساید و غیره) دقت متوسطی مورد نیاز است و اغلب آشکارسازی بازیکنها، کالیبراسیون دوربین، استخراج خطوط از مراحل مورد نیاز می باشند. به منظور آنالیزهای سطح بالا که ردیابی بازیکنها در این سطح طبقه بندی شده است اطلاعات دقیقتر نظیر اطلاعات تاکتیکی و نقاط قوت و ضعف تیمها و غیره از ویدئو استخراج می شود و نیاز به دقتهای بالا مخصوصاً در تعیین موقعیت بازیکنها و توپ است. اغلب تحقیقات صورت گرفته جهت تحلیلهای سطح بالا در زمینه ردیابی بازیکنها و توپ و ایجاد سیستمهای تشخیص آتی وقایع انجام گرفته اند. مقاله [18] به بحث کالیبراسیون دوربین در فریمهای وسط زمین که خطوط کافی و بافت کافی وجود ندارد پرداخته است و با فرض ثابت بودن موقعیت دوربین و با تنها دو نقطه، دوربین را کالیبره کرده است. روش به کار برده شده برای آشکارسازی زمین در این مقاله نیاز به تعریف یکسری پارامترها دارد. نویسندگان مقاله [20] ردیابی سه بعدی توپ و حل مشکل همپوشانی که از مشکلات رایج در ردیابی توپ است را در تصاویر حاصل از چند دوربین ثابت مطرح کرده اند که این مسئله کاربرد روش را محدود کرده است. به علاوه، مقالات زیادی بحث آشکارسازی بازیکنها و زمین را مطرح کرده اند که اغلب هیستوگرام ها در فضای HSV [23] و یا مدل گوسی ترکیبی [7,16] به کار برده شده است. به طور کلی روش مدل گوسی ترکیبی به دلیل اینکه نیاز به تعریف حدود آستانه خاصی ندارد روش بهتری است. در مقاله [21] از k-means برای آشکارسازی بازیکنها استفاده شده و تاثیر دقت تشخیص پاها در

مرحله ردیابی بازیکن مورد بررسی قرار گرفته است. بر اساس نتایج ارائه شده توسط نویسندگان مقاله ذکر شده، از دست دادن زانو به پایین خطیابی بین ۳-۲ متر و از دست دادن کل پا خطیابی بیش از ۵ متر در تعیین موقعیت بازیکن ایجاد کرده است. در این مقاله هم مشابه مقالات دیگر خطا برحسب متر برای موقعیت همه بازیکنها (چه پا از دست برود و چه نرود) ارائه نشده است تا بتوانیم مقایسه ای بین دقت روش به کار برده در این تحقیق که سعی در جبران پاها کرده است با مقالات دیگر داشته باشیم و این مسئله از دیدگاه ما ضعف روشهای موجود در جبران پای بازیکنهاست که منجر به خطاهای زیاد شده است. روشهای قدیمی تر [4] از رنگ لباس بازیکنها یا اطلاعات اندازه و لبه برای آشکارسازی بازیکنها استفاده کرده اند. مسئله ردیابی در ویدئوهای مسابقات فوتبال به دلیل حرکات سریع و ناگهانی بازیکنها و دوربین، کیفیت پایین تصاویر ناشی از نویز و حرکات دوربین که گاه منجر به از دست دادن بازیکن می شود و تعداد زیاد بازیکن ها که در جهات مختلف در حرکتند و در بسیاری موارد با هم همپوشانی دارند با دشواری های خاص خود روبروست. در تحقیقات قبلی روش های مختلفی به منظور ردیابی بازیکنها در ویدئوهای ورزشی مختلف به کار برده شده اند. در مقالات [1-4]، از روش تناظریابی قالب ها^۱ برای ردیابی بازیکنها استفاده شده است. بعضی محققین از فیلتر کالمن به همراه مشخصه های کلیدی نظیر ویژگی های ظاهری یا نمایش گرافی برای ردیابی استفاده کرده اند [5,6]. در ویدئوی فوتبال به دلیل حرکات غیر خطی و تغییر مسیر ناگهانی بازیکنها، فیلتر کالمن به تنهایی جوابگوی مسئله ردیابی نبوده و از فیلتر کالمن اغلب برای ردیابی توپ استفاده شده است. در مواردی روشهایی نظیر MHT^۲ [7]، Mean-shift [8] و TSV^۳ [9] استفاده شده اند. MHT در ردیابی چند شیء (ورزشکاران، افراد پیاده و غیره) که تعداد و محل اهداف در طول زمان متغیر است و برای مواقعی که مدل حرکتی شیء غیر قابل پیش بینی است طراحی شده است. برای این منظور مسیرهای محتمل مختلف به صورت درختی در نظر گرفته شده و محتمل ترین مسیر انتخاب گردیده است. اما در روش Mean-shift بر اساس هیستوگرام رنگ شیء در فریم قبل یک تصویر بر اساس احتمال تعلق پیکسل به شیء ایجاد و پیک (محتملترین موقعیت شیء) جستجو می شود. این روش با تکرار و بر اساس معیار تشابه عمل می کند و در صورتی که رنگ شیء تغییر نکند روشی مقاوم است. روش فیلتر ذره ای^۴ نیز از روشهای متداول ردیابی است و توسط محققین مختلف در ترکیب با روشهای دیگر به کار رفته است [10-13]. اخیراً در تحقیق دیگری [14] از روش متروپولیس هستینگ^۵ به همراه روش تبرید تدریجی^۶ برای ردیابی بازیکنهای فوتبال استفاده شده است که این روش

¹ Template matching² Multiple hypothesis tracking³ Temporal spatio velocity⁴ Particle Filter⁵ Metropolis Hasting(MH)⁶ Simulated Annealing

گیری ها بتوانند در محیط واقعی انجام شوند. اختلاف مدل مذکور با زمین واقعی تنها در یک عنصر مقیاس است که با ضرب فواصل روی مدل در مقیاس، فاصله واقعی روی زمین به دست می آید. همچنین به منظور افزایش دقت تعیین موقعیت بازیکن، دو روش مدل گوسی ترکیبی^۱ و آدابوست^۲ پیاده سازی و مقایسه شده اند. در مواردی که بازیکن از دست رفته است از درونبایی برای مشخص کردن محل بازیکن از دست رفته استفاده شده است. به علاوه، اکثر تشخیص های غلط طی مرحله ردیابی حذف شده اند و مواردی که الگوریتم قادر به تشخیص برجسب بازیکن نبوده است با تعریف همسایگی ها بازیکنها برجسب دهی شده اند. همچنین یکی از نکات متمایز روش پیشنهادی حل مشکل همپوشانی بازیکنها قبل از ردیابی با گراف است که منجر به محدود کردن همسایگی ها و ردیابی دقیقتر شده است. روش پیشنهادی برای حل مشکل همپوشانی روی تعداد زیاد بازیکنهای همپوشان آزمایش شده و قادر است تقریباً همه همپوشانی ها را تشخیص دهد، از هم تفکیک کند و دو مرحله تعیین تیم بازیکنها و برجسب دهی را همزمان انجام دهد. این کار در مقالات دیگر یا در دو مرحله جدا انجام شده است [14,9,8,5] یا با توجه به روش مورد استفاده نظیر تطبیق قالبها [4,2] مرحله برجسب دهی حذف شده است. همچنین در تحقیقات قبلی، الگوریتمهای آشکارسازی بازیکنها با معیارهایی که توانایی آشکارسازی بازیکنها را بررسی می کنند ارزیابی شده اند و توجهی به دقت تعیین موقعیت و قدرت آشکارسازی پای بازیکن توسط الگوریتم که در بسیاری فریمها ممکن است آشکارسازی نشود نشده است [۱۵،۱۴،۱۰،۹،۸،۴،۲]. ولی در اینجا دقت تعیین موقعیت برحسب متر در سیستم واقعی برای دو روش پیشنهادی ارزیابی شده است. علاوه بر این، روش آدابوست که اغلب برای آشکارسازی صورت به کار برده شده برای تعیین موقعیت بازیکنها در فریمهای ویدئویی آزمایش شده است.

در ادامه و در بخش بعدی مقاله به آشکارسازی بازیکنها با دو روش آدابوست و مدل گوسی ترکیبی پرداخته شده است. در بخش سوم این مقاله ایجاد موزایک تصویری و مرتبط سازی تصاویر به مدل توضیح داده شده است. در بخش چهارم برجسب دهی و حل مشکل همپوشانی بازیکنها مطرح شده است. در بخش پنجم پیاده سازی ردیابی بازیکنها به روش گراف شرح داده شده است. در بخش ششم نتایج ارزیابی شده اند و در پایان نتایج و پیشنهادات مطرح شده اند.

۲- آشکارسازی بازیکنان در تصاویر ویدئویی

در اینجا بازیکنها از پس زمینه به دو روش مدل گوسی ترکیبی و آدابوست تفکیک شده اند و روشی که بازیکنها را با دقت بیشتر تعیین موقعیت کند در قسمت های بعد برای ردیابی بازیکنها استفاده شده است.

با پیچیدگیهای محاسباتی و تعریف توابع توزیع پیشنهادی در الگوریتم متروپولیس هستینگ روبروست. روش گراف نیز از روشهای ردیابی بازیکنهاست که برای ردیابی بازیکن و توپ مورد استفاده قرار می گیرد. این روش در بخش پنجم تشریح خواهد شد. در تحقیق [5] در شرایط دوربین ثابت و در مقاله [15] در شرایط تصویر برداری با چند دوربین این روش به کار برده شده است. روش گرافی نیز اغلب به صورت ترکیبی با روش های دیگر به کار برده می شود.

مطالعات نویسندگان این مقاله در تحقیقات قبلی نشان می دهد که اغلب مقالات ردیابی بازیکنهای فوتبال را در فضای تصویر انجام داده اند و کمتر به مشکل همپوشانی بین بازیکنها در تصویر برداری با تک دوربین متحرک که از مشکلات رایج در مسابقات فوتبال است توجه جدی شده است. آنالیز ویدئوهای اخذ شده از تک دوربین متحرک با مشکل حل همپوشانی ها و تاری تصویر روبروست و از آنجا که استفاده از چند دوربین همزمان سازی شده همیشه مقدور نبوده است و بسیاری تصاویر موجود از تک دوربین (دوربینی که جریان اصلی بازی را دنبال می کند) اخذ شده اند، اهمیت ضرورت تحقیق برای استخراج اطلاعات از این تصاویر اجتناب ناپذیر است. مقالات [15,9,2,1] از معدود مقالاتی هستند که ردیابی را در فضای مدل انجام داده اند که مقاله [15] ردیابی را با چند دوربین و مقالات [9,2,1] ردیابی را با تک دوربین مطرح کرده اند. تمام مقالات مطالعه شده نظیر [15,10,5,2,1] در تعیین موقعیت دقیق بازیکنهای همپوشان در فضای مدل که محل بازیکنها برخلاف فضای تصویر برهم منطبق نیست با مشکل مواجه می شوند. در مقاله [2] مراحل پیاده سازی شده در این مقاله (آشکارسازی زمین، آشکارسازی بازیکنها، تعیین تیم، حل مسئله همپوشانی و ردیابی) به روشهای کاملاً متفاوتی پیاده سازی شده اند. روش آشکارسازی زمین در مقاله ذکر شده نیاز به تعیین یکسری پارامترها دارد و وقتی بازیکن آشکارسازی نشده است این مرحله با پارامترهای متفاوت انجام گرفته است. همچنین در مقاله [2] موقعیت تمام بازیکنهای همپوشان یکسان در نظر گرفته شده اند. در نظر گرفتن موقعیت یکسان برای بازیکنهای همپوشان با توجه به وجود همپوشانیهای متعدد در یک فریم که گاه بین چند بازیکن اتفاق می افتد، منجر به خطای موقعیت در حد چندین متر در فضای زمین می شود (دقت تعیین موقعیت بازیکنها در مقاله [2] همانند اکثر مقالات برحسب متر بررسی نشده است). از آنجا که جداسازی بازیکنهای همپوشان در آنالیزهای بعدی می تواند بسیار مهم و کارا باشد، در روش پیشنهادی در مقاله ما به جداسازی بازیکنهای همپوشان توجه جدی شده است و این نکته، امتیاز برجسته ای برای این تحقیق محسوب می شود. مقایسه نتایج آشکارسازی در روش پیشنهادی با نتایج [2] و [14] در بخش ارزیابی نتایج آورده شده است.

در این مقاله محل پای بازیکن در فضای تصویر و مدل تخمین زده شده است و هدف اصلی ردیابی بازیکنها در سیستم مختصات تصویر و مدل است تا تصویر فیزیکی واقعی تری برای بیننده ایجاد شود و اندازه

¹ Gaussian mixture model(GMM)

² Gentle Adaboost

<p>۱- جمع آوری نمونه های آموزشی مثبت و منفی X_i و برچسب دهی به آنها y_i</p> $(x_1, y_1), \dots, (x_m, y_m)$ $x_i \in X, y_i \in Y = \{-1, 1\}$
<p>۲- وزندهی اولیه به نمونه های جمع آوری شده</p> $D_1(i) = \frac{1}{m}, i = 1, \dots, m$ <p>که m تعداد نمونه هاست.</p>
<p>۳- به ازای $t = 1 : T$</p> <p>الف) برای هر ویژگی j (ویژگی h_j) یک طبقه بندی کننده h_j مشخص می شود. سپس طبقه بندی کننده ضعیف (h_j) که خطا را با توجه به وزن نمونه ها مینیمم کند انتخاب می گردد.</p> $h_j(x) = \begin{cases} 1, & \text{if } P_j f_j(x) < P_j \theta_j \\ 0 & \text{otherwise} \end{cases}$ $\varepsilon_j = \sum_{i=1}^m D_t(i) [y_i \neq h_j(x_i)]$ $h_t = \arg \min_{h_j \in H} \varepsilon_j$ $h_t : X \rightarrow \{-1, +1\}$ <p>P_j مشخص کننده جهت نامساوی، θ_j حد آستانه برای ویژگی کاندید و $f_j(x)$ مقدار ویژگی در پنجره مورد نظر هستند.</p> <p>ب) اگر خطا از 0.5 کمتر بود وزنها را به روز می کنیم:</p> $\alpha_t = \frac{1}{2} \ln \frac{1 - \varepsilon_t}{\varepsilon_t}$ $D_{t+1}(i) = \frac{D_t(i) e^{-\alpha_t y_i h_t(x_i)}}{Z_t}$ <p>Z_t ضریب نرمالیزاسیون است.</p>
<p>۴- طبقه بندی کننده نهایی</p> $H(x) = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right)$

شکل ۱: الگوریتم آدابوست [17]



شکل ۲: ویژگیهای هار مورد استفاده برای تشخیص بازیکنها پیشنهاد شده توسط

Viola

۲-۱ آشکارسازی بازیکنان با استفاده از روش طبقه بندی آدابوست و ساختار آبخاری^۱

روشهای آدابوست و ساختار آبخاری به منظور آشکارسازی صورت در تصاویر پیشنهاد شدند [17] و نتایج قابل قبولی در مقالات مختلف برای آشکارسازی صورت با این روش ارائه شده اند. در اینجا این روش روی فریمهای ویدئوی فوتبال برای آشکارسازی بازیکنها به کار برده می شود. ایراد آدابوست وقت گیر بودن فرآیند آموزش الگوریتم و نیاز به تعداد نمونه های آموزشی زیاد است. همچنین آدابوست در آموزش دیدن کند است ولی در آشکارسازی سریع عمل می کند و کارایی آن به نمونه های آموزشی و طبقه بندی کننده های ضعیف^۲ بستگی دارد. طبقه بندی کننده های ضعیف مختلف با انتخاب ویژگیهای هار مختلف با حدود آستانه و سائزهای مختلف تعریف می شوند. با این روش تعداد زیاد طبقه بندی کننده ضعیف قابل تعریفند. در هر تکرار، خطا به ازای طبقه بندی کننده های ضعیف محاسبه می شود و طبقه بندی کننده ضعیفی که کمترین خطا را دارد انتخاب می گردد. اختلاف قسمت سیاه و سفید ویژگیهای هار با f_j در الگوریتم شکل ۱ نشان داده شده است. همانطور که در الگوریتم شکل ۱ مشاهده می شود، اساس روش آدابوست این است که در هر دور نمونه ها مجدد وزندهی شوند و نمونه هایی که در دور قبل توسط طبقه بندی کننده ضعیف انتخابی درست تشخیص داده شده بودند وزن کمتر و نمونه هایی که اشتباه تشخیص داده شده بودند وزن بیشتر بگیرند. علاوه بر اینها ساختار آبخاری به منظور کم کردن حجم و زمان محاسبات و آشکارسازی های نادرست پیشنهاد شد که هر مرحله آن شامل تعدادی از طبقه بندی کننده های ضعیفی است که با آدابوست آموزش می بینند. در هر مرحله از ساختار آبخاری نمونه هایی که منفی تشخیص داده شدند حذف و نمونه هایی که مثبت تشخیص داده شدند به مرحله بعدی فرستاده می شوند. نمونه هایی که از تمام مراحل ساختار آبخاری بگذرند به عنوان شیء مورد نظر طبقه بندی می شوند. در نهایت طبقه بندی کننده آموزش دیده کل تصویر ورودی را اسکن می کند تا بازیکنها را آشکار کند.

به منظور افزایش دقت و کاهش تعداد نمونه های آموزشی مورد نیاز ابتدا زمین بازی در فضای رنگ YCbCr و با تشخیص پوش محدب^۳ زمین از تماشاگرها جدا شده است. سپس بازیکنها با آدابوست و ساختار آبخاری آشکارسازی شده اند. در نهایت عملیات مورفولوژی برای اتصال نواحی بازیکن که از هم جدا تشخیص داده شده اند استفاده شده است. در شکل ۱ الگوریتم مورد استفاده و در شکل ۲ ویژگیهای هار مورد استفاده و در اشکال ۳ تا ۶ نتایج حاصل از اعمال طبقه بندی کننده روی یکی از فریم های ورودی نشان داده شده است.

¹ Cascade of classifiers

² Weak classifier

³ Convex hull

⁴ Haar feature

احتمالات موخر مربوط به تعلق هر پیکسل به هر کلاس محاسبه و در مرحله M میانگین و کواریانس کلاسها با توجه به احتمالات موخر به روز می شوند. این روند تکرار و به این ترتیب تابع درستنمایی بیشینه می گردد. بعد از اتمام تکرارها هر پیکسل به کلاسی که بیشترین احتمال موخر را دارد منتسب می شود. الگوریتم با مقادیر اولیه (μ_k^0, Σ_k^0) شروع شده است که این مقادیر اولیه با روش k-means به دست آمده اند. الگوریتم EM شامل مراحل زیر است:

مرحله E: در این مرحله وزن عضویت داده $x(i)$ به خوشه k ام با داشتن پارامترها محاسبه شده است:

$$w_{ik} = p(C=k | x(i), \theta) = \frac{p_k(x(i) | \theta_k) \alpha_k}{\sum_{m=1}^k p_m(x(i) | \theta_m) \alpha_m} \quad (2)$$

مرحله M: در این مرحله با داشتن $(\mu_k^t, \Sigma_k^t, \alpha_k^t)$ پارامترها برای تکرار t+1 به صورت زیر به روز شده اند:

$$\alpha_k^{new} = \frac{1}{N} \sum_{i=1}^N w_{ik}$$

$$\mu_k^{new} = \frac{1}{\sum_{i=1}^N w_{ik}} \sum_{i=1}^N w_{ik} \cdot x(i) \quad (3)$$

$$\Sigma_k^{new} = \frac{1}{\sum_{i=1}^N w_{ik}} \sum_{i=1}^N w_{ik} \cdot (x(i) - \mu_k^{new})(x(i) - \mu_k^{new})^t$$

اگر میزان تغییر احتمال تعلق پیکسلها به خوشه ها از یک حد آستانه کوچکتر بود الگوریتم متوقف می شود و در غیر این صورت به مرحله E بر می گردد.

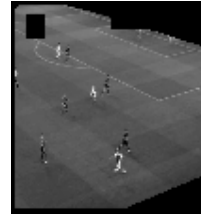
در شکل زیر با استفاده از عملیات مورفولوژی روی نتایج و تشخیص پوسته محدب زمین، نواحی پس زمینه حذف شده اند. نتایج حاصل از پیاده سازی روش ذکر شده در شکلهای ۷ و ۸ نشان داده شده است.



شکل ۸: بازیکنها



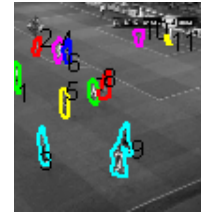
شکل ۷: جدا کردن بازیکنها با مدل گوسی ترکیبی



شکل ۴: حذف پس زمینه و آرهما



شکل ۳: تصویر اولیه



شکل ۶: بازیکنها (رنگها و شماره ها برای نمایش بهتر نواحی مربوط به بازیکنهای مختلف استفاده شده اند)



شکل ۵: جدا کردن بازیکنها با آدابوست

۲-۲ آشکارسازی بازیکنها با استفاده از روش مدل گوسی ترکیبی

در این قسمت از روش EM^۱ به همراه مدل گوسی ترکیبی برای خوشه بندی مشاهدات استفاده شده است [16]. در این روش فرض بر این است که مشاهدات متعلق به یک تابع پخش متشکل از ترکیب توابع گوسی با پارامترهای مختلف هستند و احتمال تعلق پیکسل X به چمن به صورت ترکیب K تابع گوسی مدل شده است (معادله ۱).

$$p_k(x | \theta_k) = \frac{1}{2\pi^{\frac{d}{2}} \left| \Sigma_k \right|^{\frac{1}{2}}} e^{-\frac{1}{2}(x - \mu_k)^t \Sigma_k^{-1} (x - \mu_k)}$$

$$p(x | \theta) = \sum_{k=1}^K \alpha_k p_k(x | \theta_k) \quad (1)$$

$$\theta = \{\alpha_1, \dots, \alpha_k, \theta_1, \dots, \theta_k\}$$

$$\theta_k = \{\mu_k, \Sigma_k\}$$

که هر تابع گوسی p_k با میانگین μ_k و کواریانس Σ_k بردار ویژگیهای پیکسل X با \underline{x} و وزن هر کدام از توابع گوسی با α_k نشان داده شده اند که مجموع وزنها یک است. هر گوسی میانگین و ماتریس کواریانس مربوط به خود را دارد که باید تخمین زده شود. بنابراین از روش EM برای تخمین پارامترهای مدل گوسی ترکیبی (میانگین، کواریانس و وزنها) استفاده می شود. در EM عضویت هر پیکسل به هر خوشه با احتمال موخر^۲ تعریف می گردد و به تعداد کلاسها احتمال موخر داریم. به عبارت دیگر هر کلاس با میانگین و کواریانس و تابع احتمال^۳ هر کلاس به صورت تابع گوسی تعریف می شود. در مرحله E

¹ Expectation Maximization (EM)

² Posterior probability

³ Likelihood function

۳- ایجاد موزاییک تصویری و مرتبطسازی

تصاویر با مدل

در این مرحله برای مرتبط سازی تصاویر به مدل و ایجاد موزاییک تصویری از انتقال افکنشی دو بعدی استفاده شده است. مختصات نقاط کنترل مشترک در فضای تصویر و مدل به عنوان ورودی معرفی شده اند و ماتریسهای انتقال بین تصاویر با یکدیگر و تصاویر با مدل به عنوان خروجی تعریف شده اند. در رابطه ۴ مختصات نقاط در فضای تصویر با X, Y و مختصات نقاط روی مدل با X, Y نشان داده شده است و پارامترهای a, b, c با ۴ نقطه مشترک برای ایجاد ارتباط بین دو فضا تعیین شدند. همچنین انتقال را می توان بین تصاویر انجام داد. به علاوه، موزاییک ایجاد شده حاصل از مرتبط سازی تصاویر به یکدیگر در مرحله ردیابی بازیکنها استفاده می شود. در شکلهای ۹ و ۱۰ نتایج مرتبط سازی تصاویر و مدل نشان داده شده است. همچنین بعد از برقراری ارتباط بین فضای تصویر و مدل، خطوط و دروازه می توانند با یکسری نقاط کلیدی روی مدل مشخص شوند و به فضای تصویر انتقال یابند. در این مرحله به منظور تفکیک بازیکنها، پیکسلهای مربوط به این نواحی با توجه به ویژگی های خاص ایندو نظیر رنگ و ضخامت شناسایی و حذف شدند.

$$X = \frac{a_0x + a_1y + a_2}{c_0x + c_1y + 1} \quad (4)$$

$$Y = \frac{b_0x + b_1y + b_2}{c_0x + c_1y + 1}$$



شکل ۹: مرتبط سازی تصویر و مدل



شکل ۱۰: مرتبط سازی دو فریم تصویری

۴- حل همزمان برچسب دهی و جداسازی

بازیکنان همپوشان

در این بخش، جداسازی و تعیین موقعیت بازیکنهایی که با هم همپوشانی دارند به عنوان سخت ترین مرحله در تعیین موقعیت بازیکن ها مورد توجه قرار گرفته است. در روش پیشنهادی از اطلاعات رنگ لباس بازیکن ها برای جدا کردن و برچسب دهی تمام بازیکن ها استفاده شده است. لذا شناسایی فضاهای رنگ و حدود آستانه مناسب برای عناصر این فضاها که رنگ لباس بازیکن ها را از هم متمایز کند در دقت این مرحله بسیار تاثیر گذار است. از آنجا که رنگ های متمایزی برای لباس بسیاری تیمها به کار برده شده است مثلا لباس تیم برزیل یا هلند یا تیم ملی ایران، اگر یک بار فضای رنگ مناسب و حدود آستانه مناسب شناسایی شود قابل استفاده برای تمامی بازی های این تیم ها خواهد بود. پس از آن، نواحی همپوشان بر اساس رنگهای لباس تیم های مختلف موجود در ناحیه مشخص شده و اندازه مستطیل در برگیرنده بازیکن^۱ شناسایی شده اند. این مستطیلها در شکل ۱۶ نشان داده شده است. اگر ناحیه به عنوان ناحیه همپوشان مشخص نشده باشد، بر اساس رنگ موجود در مستطیل بازیکن برچسب داده شده است و در غیر این صورت دو بازیکن غیر هم تیمی بر اساس رنگ لباس و اندازه نزدیکترین مستطیل همسایه تفکیک شده اند. در صورتی که دو بازیکن هم تیمی پوشش داشته باشند تنها معیار جداسازی آنها اندازه مستطیل های همسایه است. به علاوه، در این روش دقت آشکارسازی بازیکن در مواقعی که پاهای بازیکن از دست رفته یا نواحی اضافی در مستطیل بازیکن قرار گرفته اند، با بررسی مستطیل هایی که اندازه بزرگتر یا کوچکتر از حد معمول نسبت به مستطیل های دیگر یا مستطیل های همسایه دارند و اصلاح اندازه آنها به کمک رنگ، افزایش داده شده است.

۵- ردیابی بازیکنان به روش گراف

به منظور ردیابی بازیکنها، بعد از انتقال موقعیتهای به دست آمده بازیکنها به روش مدل گوسی ترکیبی به یک سیستم مختصات (تصویر اول یا مدل)، مراحل زیر انجام می گیرد:

الف) تعریف همسایگی: همسایگی بین نقاط با شروط زیر تعریف می شود:

- دو بازیکن با فاصله کمتر از حد مجاز (ماکسیمم سرعت مجاز X اختلاف زمانی فریم دو بازیکن) و اختلاف زمانی معادل اختلاف زمانی بین دو فریم متوالی به عنوان همسایه تعریف شده اند. برای تعریف دقیقتر همسایگی از آنجا که حرکت بازیکن در جهت افقی و رو به دروازه بسیار سریعتر از حرکت بازیکن در جهت قائم یعنی در امتداد عرض زمین است، سرعت مجاز در جهت افقی و قائم به طور جداگانه

¹ Bounding box

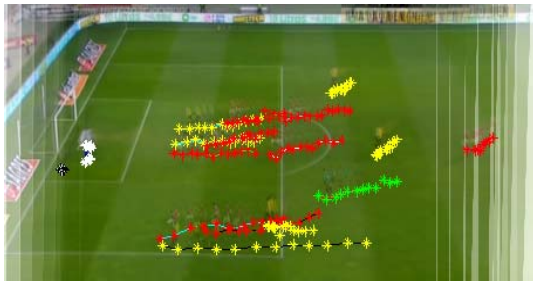
۳- فاصله بین آخر یک مسیر و عنصر اول مسیر دیگر از (ماکسیم سرعت مجاز \times اختلاف زمانی دو نقطه) کوچکتر و اختلاف زمانی دو نقطه از یک بزرگتر باشد.

۴- تعداد فریم هایی که بازیکن آشکارسازی نشده است از ۵ کمتر باشد.

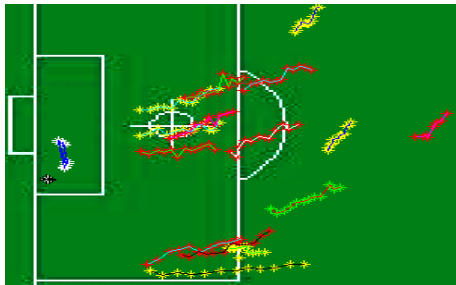


شکل ۱۱: درونبایی برای تعیین موقعیت بازیکن از دست رفته بین دو مسیر روی موزاییک ایجاد شده

ب) ردیابی بر اساس مینیمم فاصله: در اینجا از نمایش گراف برای ردیابی بازیکنها در فریمهای ویدئو استفاده شده است. پس از تعریف همسایگی به طریق ذکر شده، از بازیکن واقع در فریمی که کمترین زمان را دارد، ایجاد مسیر بازیکن مذکور شروع شده است و نزدیکترین بچه بازیکن به عنوان نقطه بعدی مسیر انتخاب می شود. به همین ترتیب مسیر بازیکن ادامه داده شده است تا بچه ای برای ادامه مسیر بازیکن یافت نشود. این کار ادامه داده می شود تا تمام مسیرهای ممکن ایجاد شوند. بعد از ایجاد تمام مسیرهای ممکن سیری که هیچ کدام از بازیکنهای آن برجسب ندارند به عنوان تشخیص های اشتباه متوالی حذف می شود. در شکل ۱۲ مسیر بازیکنها در سیستم مختصات تصویری و در شکل ۱۳ در سیستم مختصات مدل نشان داده شده اند. داور با نقاط سبز، دروازه بان با نقاط سفید، تیم ۱ با نقاط قرمز و تیم ۲ با نقاط زرد نشان داده شده اند. نقاط سیاه نقاطی هستند که برجسب معین نگرفتند و به عنوان تشخیص های اشتباه حذف می گردند. شکستگی های ناگهانی ایجاد شده در مسیر بازیکنها ناشی از خطای تعیین موقعیت بازیکن است که اغلب از دست رفتن پای بازیکن منجر به این خطا شده است. خالی بودن ناحیه جلوی دروازه بان از چمن و پیوستن این ناحیه به پاهای دروازه بان علت خطای زیادی است که در مسیر دروازه بان مشاهده می شود.



شکل ۱۲: ردیابی در فضای تصویر



شکل ۱۳: ردیابی در فضای مدل

تعریف می شود. با این کار بسیاری همسایگی های غلط حذف می گردند.

• بعد از انتقال نقاط به سیستم مختصات مدل، فاصله دو نقطه همسایه روی مدل باید کمتر از حداکثر فاصله مجاز در فضای مدل باشد. از آنجا که بعضی بازیکنهای هم تیمی که نزدیک هم هستند در یکی از دو فضا همسایه اند و در فضای دیگر همسایه نیستند، از هر دو فضای تصویر و مدل برای تعریف دقیق تر همسایگی و کاهش تشخیص های غلط استفاده می کنیم. بنابراین نقاطی که جابجایی آنها در هر دو فضا در حد مجاز تعریف شده برای آن فضا بود با هم همسایه اند.

• دو بازیکن برجسب یکسان گرفته باشند (برجسبها: داور، تیم یک، تیم دو، دروازه بان).

• در نمایش گرافی، بازیکن در فریم قبل والد بازیکن و بازیکن در فریم بعد بچه بازیکن در فریم حاضر نامیده می شود. به منظور حذف همسایگی بازیکن با تشخیص های اشتباه، اگر بازیکنی بیش از یک والد (بچه) داشت که یکی از آنها برجسب مشخصی در مرحله برجسب دهی نگرفته بود، همسایگی بازیکن با بازیکن بدون برجسب حذف می شود.

• اگر ناحیه بیش از یک بچه (والد) دارد که یکی از بچه ها (والدها) یک والد (بچه) و بقیه بیش از یک والد (بچه) دارند، همسایگی بین ناحیه و نواحی دارای بیش از یک والد (بچه) حذف می شود. با این کار یکسری همسایگی های غلط حذف می گردند.

• اگر بازیکن هیچ همسایه ای نداشت، از لیست کاندیدهای ردیابی حذف می شود.

تعریف همسایگی ها به هدایت ردیابی مخصوصا وقتی دو بازیکن نزدیک هم در حرکتند و تعداد انتخابها زیاد است کمک کرده است.

گاه اتفاق می افتد که به دلایلی نظیر تاری بازیکن در چند فریم یا ریز بودن بازیکن به دلیل دوری از دوربین و کیفیت بد تصویر یا همپوشانی های حل نشده، بازیکن آشکارسازی نشده است. همانطور که در تصویر چپ شکل ۱۱ نشان داده شده است بازیکن در دو فریم به دلیل حرکت دوربین و تاری از دست رفته است. اگر شرایط زیر برقرار باشد مسیر بازیکن به دلیل از دست رفتن بازیکن به دو قسمت تقسیم شده است که در این گونه مواقع، درونبایی خطی برای تعیین موقعیت بازیکن های از دست رفته بین دو مسیر به کار برده می شود.

۱- برجسب دو مسیر یکسان باشد.

۲- زمان فریم های تشکیل دهنده دو مسیر کاملا با هم متفاوت باشند.

۶- ارزیابی نتایج

آدایوست (ستون راست) و مدل گوسی ترکیبی (ستون چپ) و جداسازی بازیکنهای همپوشان و جداسازی بازیکنهای همپوشان

میانگین دقت دو روش روی ۷۸ فریم از ۶ ویدیوی مختلف در جدول ۱ آورده شده است. خطاها نسبت به نقطه وسط دو پای بازیکن سنجیده شده اند و از آنجا که تمام فضای بین دو پا که هنگام دیدن این فاصله از ۱ متر هم بیشتر است می تواند به عنوان موقعیت بازیکن در نظر گرفته شود، نتایج حاصل از پیاده سازی روشها با توجه به خطاهای ذکر شده در جدول ۱ قابل اعتماد هستند. معیار های ارزیابی دقت تعیین موقعیت بازیکن به صورت زیر تعریف شده اند که تعداد کل بازیکنها در فریمهای ویدئوهای مختلف با N و اختلاف مقدار محاسباتی موقعیت بازیکن با مقدار داده کنترلی با ΔX و ΔY نشان داده شده است.

$$RMSE = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\Delta X_i^2 + \Delta Y_i^2)}$$

$$RMSEX = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\Delta X_i^2)} \quad (5)$$

$$RMSEY = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\Delta Y_i^2)}$$

جدول ۱: خطای تعیین موقعیت بازیکن

خطا	RMSEX		RMSEY		RMSE	
	مدل گوسی ترکیبی	مدل آدایوست	مدل گوسی ترکیبی	مدل آدایوست	مدل گوسی ترکیبی	مدل آدایوست
روش	0.37	0.46	0.65	0.74	0.75	0.87

از دو معیار فراخوان^۱ و دقت^۲ برای ارزیابی کارایی روشهای آشکارسازی بازیکنها استفاده شده است. برای یک طبقه بندی کننده خوب باید هر دو معیار بالا باشند.

$$recall = \frac{TP}{TP + FN} \quad (6)$$

$$precision = \frac{TP}{TP + FP}$$

جدول ۲: تعریف پارامترها

	نمونه منفی	نمونه مثبت
نمونه مثبت تشخیص داده شده	FP	TP
نمونه منفی تشخیص داده شده	TN	FN

در جدول زیر مقایسه ای بین نتایج به دست آمده در آشکارسازی بازیکنها در این مقاله و مقالات [2,14] نشان داده شده است.

به منظور آشکارسازی بازیکنها به روش آدایوست، ۷۰۳ نمونه مثبت و ۷۱۲ نمونه منفی جمع آوری شد که از این تعداد دو سوم برای آموزش و یک سوم برای ارزیابی استفاده شدند. ساختار آبخاری شامل سه مرحله است که در مرحله اول ۶۴۷، در مرحله دوم ۲۶ و در مرحله سوم ۵ نمونه منفی تشخیص داده و حذف شدند. کارایی روش آدایوست را می توان با آموزش بهتر الگوریتم افزایش داد. در روش توسعه داده شده در این تحقیق اکثراً آشکارسازی خطوط و بازیکنها در آدایوست با مشکل مواجه است و خطوط به عنوان بازیکن تشخیص داده می شوند که منجر به تشخیص های اشتباه زیاد شده است. در مرحله آموزش ۹۰.۱٪ از نمونه های منفی مثبت تشخیص داده شدند. این روش به دلیل سختی آموزش آن پیشنهاد نمی شود. در آشکارسازی بازیکنها به روش مدل گوسی ترکیبی چند تکه شدن بازیکن کمتر دیده می شود و پیوستگی بیشتر پاها به بدن نسبت به روش آدایوست قابل مشاهده است. این روش برای آشکارسازی بازیکنها مناسب تر، دقیقتر و راحتتر به نظر می رسد. تشخیص های غلط شامل تکه های خطوط باقیمانده در تصویر و نویزهاست که در مقایسه با آدایوست بسیار کمتر است و با آشکارسازی خطوط و حذف نواحی کوچک حذف می شوند. مزیتی که در این تحقیق روش آدایوست به مدل گوسی ترکیبی دارد این است که ویژگیهای موجک هار در آدایوست استفاده می شود که به رنگ سبز حساس نیست ولی عناصر رنگ در مدل گوسی ترکیبی استفاده می شود که به رنگ لباس سبز که همرنگ چمن است حساس است. در ۷۸ فریم، ۸۱ مورد همپوشانی کلی مشاهده شد که روش ذکر شده تقریباً تمامی موارد را به درستی آشکار نمود. نتایج حاصل از تشخیص و حل مشکل همپوشانی بازیکنها با استفاده از دو روش فوق روی فریمهای ویدئوهای مختلف در شکل ۱۴ نشان داده شده است:



شکل ۱۴: مستطیلهای نشان دهنده بازیکنها بعد از مرحله آشکارسازی با دو روش

¹ Recall² Precision

در کارهای آینده به استخراج اطلاعات سه بعدی از ویدئوی فوتبال، ردیابی توپ و آشکارسازی اتوماتیک خطوط به منظور مرتبط سازی اتوماتیک فریم ها پرداخته خواهد شد.

مراجع

- [1] H.S.- Yoon, et al. A Soccer Image Sequence Mosaicing and Analysis Method Using Line and Advertisement Board Detection. ETRI Journal, 2002, vol. 24, no. 6, pp. 443-454.
- [2] S.H.- Khatoonabadi, M. Rahmati. Automatic soccer players tracking in goal scenes by camera motion elimination. Image and Vision Computing 27 (2009) 469-479.
- [3] Han, H., et al. A Unified and Efficient Framework for Court-Net Sports Video Analysis Using 3-D Camera Modeling. IEEE Transactions on Circuits and Systems for Video Technology, 2008. pp. 1628 - 1638.
- [4] Utsumi, O., et al. An object detection method for describing soccer games from video. In: IEEE International Conference on Multimedia and Expo (ICME), 2002, Vol. 1, pp. 26-29.
- [5] P. Figueroa, et al. Tracking soccer players using the graph representation. In Proceedings of the 17th International Conference on Pattern Recognition (ICPR), 2004, pp. 787-790.
- [6] G. Liu, et al. A novel approach for tracking high speed skaters in sports using a panning camera, Pattern recognition, 42(11)(2009) 2922-2935.
- [7] Beetz, M., et al. Visually tracking football games based on tv broadcasts. IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, 2007, Hyderabad, India.
- [8] T.S.- Chiang, et al. An Improved Mean Shift Algorithm Based Tracking System for Soccer Game Analysis. Proceedings of APSIPA Annual Summit and Conference, 2009, Sapporo, Japan, pp. 4-7.
- [9] K. Sato, J. K. Aggarwal. Tracking soccer players using broadcast tv images. In IEEE internat. Conf. on Advanced Video and Signal Based Surveillance(AVSS), 2005.
- [10] Liu, Y., et al. Extracting 3D information from broadcast soccer video. Image and Vision Computing 24 (2006) 1146-1162
- [11] Bilmes, J.A. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models, 1998. TR-97-021, Berkeley.
- [12] Y. Zhang, H. Lu, Ch. Xu. Collaborate ball and player trajectory extraction in broadcast soccer video. In: Proc. IEEE Internat. Conf. on Pattern Recognition (ICPR), 2008, pp. 1-4.

جدول ۳: ارزیابی کارایی روش های آشکارسازی بازیکن و مقایسه کارایی

روشهای پیشنهادی و روش مراجع [2,14]

معیار	دقت			
	مقاله [14]	مقاله [2]	مدل گوسی ترکیبی	آدابوست
میانگین (%)	88.65	92.38	۹۴	۸۸

معیار	فراخوان			
	مقاله [14]	مقاله [2]	مدل گوسی ترکیبی	آدابوست
میانگین (%)	۹۲.۱۹	۸۸.۸۲	۹۸	۹۹

خطای تعیین موقعیت یا خطای برچسب دهی بازیکنها در مراحل قبلی منجر به خطاهای تعیین مسیر بازیکنها شده است. مثلا بازیکن به دو تکه تقسیم شده و یک مسیر اضافی ایجاد شده است یا ممکن است بازیکنی درست برچسب دهی نشده یا آشکارسازی نشده باشد. با اینکه در روش پیشنهاد شده خطاهای ذکر شده با توجه به همسایگی ها و با به کارگیری درون یابی تا حدی جبران گردیده اند، ولی در بعضی مواقع این مسیرهای اضافی ایجاد شده اند. همچنین مسیرهای اضافی که متعلق به بازیکنها نبوده اند با توجه به نداشتن برچسب حذف شده اند. نتایج تعیین مسیر بازیکنها در جدول ۴ نشان داده شده است.

جدول ۴: دقت ردیابی به روش گراف روی نتایج مدل گوسی ترکیبی

شماره ویدئو	درصد مسیرهای درست	تعداد مسیرهای اضافی	تعداد کل مسیرها	تعداد فریم ها
	مدل و تصویر	مدل و تصویر	مدل و تصویر	
۰	۹۰	۲	۱۰	۱۵
۱	۱۰۰	۱	۱۲	۱۵
۲	۱۰۰	۰	۱۸	۵
۳	۱۰۰	۱	۱۱	۱۵
۴	۹۰	۱	۱۰	۴
۵	۱۰۰	۱	۱۴	۱۲
۶	۱۰۰	۰	۱۶	۱۲

۷- نتیجه گیری

تعیین موقعیت بازیکنها اساس آنالیزهای بعدی به منظور استخراج اتوماتیک اطلاعات از ویدئوی مسابقات فوتبال است. هر قدر این مرحله دقیقتر انجام شود، دقت مراحل بعدی بهبود می یابد. علاوه بر این همپوشانی بازیکنها از مشکلات رایجی است که کمتر به آن پرداخته شده است. در این تحقیق به مشکل همپوشانی بازیکنها پرداخته شد و بازیکنها با دقت قابل قبول تعیین موقعیت و ردیابی شدند. همچنین دقت تعیین موقعیت بازیکنها با دو روش آدابوست و مدل گوسی ترکیبی در سیستم متریک بررسی گردید و مسیر بازیکنها با نمایش بهتری نسبت به کارهای قبلی روی موزاییک تصویری حاصل از مرتبط سازی تصاویر به یکدیگر و مدل زمین بازی ایجاد شده در محیط اتوکد برحسب استانداردهای فیفا نمایش داده شد.

- [19] T. D’Orazio, M. Leo, “A review of vision-based systems for soccer video analysis”, *Pattern Recognition*, Vol. 43, pp. 2911–2926, 2010.
- [20] Jinchang Ren, James Orwell, Graeme A. Jones, Ming Xu, “Tracking the soccer ball using multiple fixed cameras”, *Computer Vision and Image Understanding*, Vol. 113, pp. 633–642, 2009.
- [21] Hanjoo Kim, Siwook Nam, and Jaihie Kim. Player segmentation evaluation for trajectory estimation in soccer games. *Image and Vision Computing 2003*, New Zealandv.
- [22] A. Yamada, Y. Shirai, J. Miura. Tracking players and a ball in video image sequence and estimating camera parameters for 3D interpretation of soccer games. in: *Proceedings of the International conference on Recognition*, 2002, pp. 303–306.
- [23] Ilkoo Ahn, Youngwoo Kim and Changick Kim. Customizing Ground Color to Deliver Better Viewing Experience of Soccer Video. *ETRI Journal*, vol. 30, no. 1, pp. 101-112, Feb. 2008.
- [13] Mauthner, TH., et al. Visual tracking of athletes in beach volleyball using a single camera. *International journal of computer science in sports*, 2007, 6(2), 21-35
- [14] J. Liu et al. Automatic player detection, labeling and tracking in broadcast soccer video. *Pattern Recognition 30(2009)* 103-113.
- [15] P.J. Figueroa et al. Tracking soccer players aiming their kinematical motion analysis. *Computer Vision and Image Understanding*, 2006. Vol. 101, No.2 pp. 122–135.
- [16] S. Jiang, et al. A new method to segment playfield and its applications in match analysis in sports video, In *Proceedings of the 12th annual ACM international conference on Multimedia*, 2004. New York, USA, pp. 292-295.
- [17] Viola, P., Jones, M. Rapid object detection using a boosted cascade of simple features. In: *Proc. IEEE Internat. Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2001. Kauai, Hawaii, vol.1, pp.511.
- [18] Xinbo Gao, Zhenxing Niu, Dacheng Tao, Xuelong Li, “Non-goal scene analysis for soccer video”, *Neurocomputing*, Vol. 74, pp. 540–548, 2011.



K.N. Toosi Univ. of Tech.

Journal of Control (ISSN 2008-8345)



A Joint Publication of the Iranian Society of Instrumentation and Control Engineers and the K.N. Toosi University of Technology, Vol. 5, No. 1, Spring 2011.

Publisher: **Iranian Society of Instrumentation and Control Engineers**

Managing Director: Prof. Iraj Goodarznia

Editor-in-Chief: Prof. Ali Khaki-Sedigh

Tel: 84062317

Email: sedigh@kntu.ac.ir

Assistant Editor: Dr. Hamid Khaloozadeh, Dr. Alireza Fatehi

Executive Director: Dr. Hamid Khaloozadeh

Editorial Board:

Prof. A. Khaki-Sedigh, Prof. I. Goodarznia, Dr. H. Khaloozadeh (Associate Prof.), Prof. P. Jabedar-Maralani, Prof. A. Ghafari, Dr. H.R. Momeni (Associate Prof), Prof. S.K. Nikraves, Prof. M. Shafiee, Prof. B. Moshiri.

Advisory Board:

Dr. H.R. Momeni, Prof. B. Moshiri, Prof. M. Shafiee, Prof. A. Khaki-Sedigh, Prof. P. Jabedar-Maralani, Prof. A. Ghaffari, Dr. H. Khaloozadeh, Dr. H.R. Taghirad, Dr. K. Masroori, Dr. M.T. Bathaei, Dr. M.T. Hamidi-Beheshti, Dr. F. Jafarkazemi, Dr. R. Amjadifard, Dr. S.A. Mousavian, Dr. M. Teshnelab, Prof. M. Haeri, Dr. S.A. Safavi, Prof. H. Seifi, Dr. A. Kazemi, Dr. A. Fatehi, Dr. M.R. Akbarzadeh-Toutounchi, Dr. M. Golkar, Dr. N. Pariz, Dr. M. Javadi, Dr. J. Heirani-Nobari, Prof. F. Hossein-Babaei, Dr. B. Moaveni, Dr. M. Aliari-Shourehdeli, Dr. M. Arvan, Dr. M. Tavakoli-Bina.

The ISICE Board of Director:

Abbas. Sheri-Moghadam, Prof. Masoud Shafiee., Dr. Hamid Reza Momeni, Dr. Hamid Khaloozadeh, Dr. Mehrdad Javadi, Dr. Davod Karimzadegan, Ali Kiani.

Address: Room 241, 2nd floor, No.71, Mousavi Ave. Ferdowsi Sq. Enghelab St. Tehran, Iran.

P.O. Box: 15815-3595

Tel: (+9821) 88813002

Fax: (+9821) 88324979

<http://www.isice.ir>



A Joint Publication of the Iranian Society of Instrument and Control Engineers and the K.N. Toosi University of Technology
Vol. 5, No. 1, Spring 2011

Contents

- Centralized Constrained Predictive Control on Information Coupled Supply Chain Management System** 1
Mohammad Miranbeigi, Ali Akbar Jalali
- A Hybrid Learning Algorithm for Fuzzy Wavelet Networks Design for Functions Approximation, Online Identification and Control of Nonlinear Systems** 14
Maryam Shahriari Kahkeshi, Maryam Zekri
- Application of Energy Absorption Capacity Concept for Stability Analysis of Multi-Agent Systems** 27
Karim Rahmani, Ahmad Afshar, Ali Akbar Jamshidifar
- Optimal Active and Reactive Power Sharing between several Distributed Generation Units in a Stand-Alone MicroGrid Using Artificial Neural Network** 38
Ali Asghar Ghadimi, Farzad Razavi, Hassan Rastegar
- Action Value Function Approximation Based on Radial Basis Function Network for Reinforcement Learning** 50
Vali Derhami, Omid Mehrabi
- Solving the Local Minimum Problem in Path Planning Algorithm based on the Virtual Potential Field and the Principles of Liquid Movement** 64
Bijan Moaveni, Davoud Ghanbari Gol
- Player Localization and Tracking in Field Model Space using Graph Representation in Football Broadcast Videos** 77
Mehrtash Manafifard, Hamid Ebadi, Hamid Abrishami Moghaddam