

Robust Modeling Using a Combination of Simulated Annealing and Reinforcement Learning for Route Optimization

Hami Talebi¹, Mehran khalaj², Davood Jafari³, Parisa Mousavi Ahranjani⁴, Amir Hossein Kamali Dolatabadi⁵.

1. PhD student, Department of Industrial Engineering, Parand and Robat Karim Branch, Islamic Azad University, Parand, Iran. Hami_talebi@yahoo.com
2. Assistant Professor, Department of Industrial Engineering, Parand and Robat Karim Branch, Islamic Azad University, Parand, Iran. *Corresponding Author, mekhalaj1@gmail.com
3. Assistant Professor, Department of Industrial Engineering, Parand and Robat Karim Branch, Islamic Azad University, Parand, Iran. djafari5071@yahoo.com
4. Assistant Professor, Department of Industrial Engineering, Parand and Robat Karim Branch, Islamic Azad University, Parand, Iran. P_mousavi_a@yahoo.com
5. Assistant Professor, Department of Industrial Engineering, Parand and Robat Karim Branch, Islamic Azad University, Parand, Iran. amir.kamali2002@gmail.com

Abstract

Introduction: In today's dynamic and highly competitive business environment, optimizing supply chain operations is a critical necessity. This study proposes a robust optimization framework that synergistically integrates simulated annealing with reinforcement learning to enhance routing efficiency within a complex supply chain network.

Method: The problem is formulated as a mixed-integer linear programming (MILP) model whose objective function minimizes total transportation costs, vehicle utilization expenses, and time-window penalties. Real-world uncertainties in customer demand are represented as uncertain parameters and addressed through a robust optimization approach. To obtain practical, scalable solutions, a metaheuristic based on simulated annealing is employed and further augmented with reinforcement learning strategies to improve algorithmic effectiveness.

Results: A real-world case study from the logistics sector demonstrates the practical application and effectiveness of the proposed model and solution approach. The integrated simulated annealing–reinforcement learning algorithm yields efficient routing plans that reduce overall costs and enhance resource utilization under demand uncertainty.

Discussion: A comprehensive sensitivity analysis provides actionable intelligence on the model's adaptability and performance across various operational scenarios. The findings offer supply chain decision-makers insights into how cost, vehicle usage, and service levels respond to changes in demand and system parameters, underscoring the framework's practical value for robust, data-driven routing decisions.

Keywords: Supply Chain, Robust Optimization, Simulated Annealing, Reinforcement Learning, Vehicle Routing Problem

مدل سازی استوار با استفاده از ترکیب شبیه سازی تبرید و یادگیری تقویتی جهت بهینه سازی مسیرها

دوره پنجم، تابستان ۱۴۰۳
شماره دوم، صص: ۱۹۳-۲۰۹

تاریخ دریافت: ۱۴۰۳/۰۱/۲۸
تاریخ پذیرش: ۱۴۰۳/۰۳/۰۸

حامی طالبی^۱، مهران خلیج^{۲*}، داود جعفری^۳، پریسا موسوی اهرنجانی^۴، امیرحسین کمالی^۵

۱. دانشجوی دکتری، گروه مهندسی صنایع، واحد پرند و رباط کریم، دانشگاه آزاد اسلامی، پرند، ایران. Hami_talebi@yahoo.com
۲. استادیار، گروه مهندسی صنایع، واحد پرند و رباط کریم، دانشگاه آزاد اسلامی، پرند، ایران. (نویسنده مسئول) mekhalaj1@gmail.com
۳. استادیار، گروه مهندسی صنایع، واحد پرند و رباط کریم، دانشگاه آزاد اسلامی، پرند، ایران. djafari5071@yahoo.com
۴. استادیار، گروه مهندسی صنایع، واحد پرند و رباط کریم، دانشگاه آزاد اسلامی، پرند، ایران. P_mousavi_a@yahoo.com
۵. استادیار، گروه مهندسی صنایع، واحد پرند و رباط کریم، دانشگاه آزاد اسلامی، پرند، ایران. amir.kamali2002@gmail.com

چکیده: در محیط پویا و رقابتی کسب و کارهای امروزی، بهینه سازی عملیات زنجیره تأمین نقشی حیاتی ایفا می کند. این مقاله یک چارچوب بهینه سازی استوار ارائه می دهد که با ترکیب الگوریتم تبرید شبیه سازی شده و یادگیری تقویتی، به بهبود کارایی مسیرها در شبکه ای پیچیده از زنجیره تأمین می پردازد. مدل ریاضی مورد استفاده، یک مدل برنامه ریزی خطی عدد صحیح است که هدف آن به حداقل رساندن مجموع هزینه های حمل و نقل، بهره برداری از وسایل نقلیه و پنجره های زمانی است. با در نظر گرفتن عدم قطعیت تقاضا به عنوان یک واقعیت دنیای واقعی، از رویکرد بهینه سازی استوار برتسیماس برای مدیریت این عدم قطعیت استفاده شده است. برای حل مدل در مقیاس های واقعی، الگوریتم فراابتکاری تبرید شبیه سازی شده به کار گرفته شده و برای ارتقای عملکرد، با رویکرد یادگیری تقویتی ترکیب شده است. یک مطالعه موردی واقعی در حوزه لجستیک، کاربردپذیری و اثربخشی مدل و الگوریتم پیشنهادی را نشان می دهد. همچنین، تحلیل حساسیت انجام شده، بینش های عملیاتی ارزشمندی درباره انطباق پذیری و عملکرد مدل در شرایط مختلف ارائه می کند که می تواند راهگشای تصمیم گیری در مدیریت زنجیره تأمین باشد.

واژه های کلیدی: زنجیره تأمین، بهینه سازی استوار، شبیه سازی تبرید، یادگیری تقویتی، مسیریابی وسایل نقلیه.

۱. مقدمه

در فضای پویای کسب‌وکار امروز، بهینه‌سازی عملیات زنجیره تأمین برای سازمان‌هایی که در تلاش برای حفظ مزیت رقابتی هستند، امری حیاتی است. تعامل پیچیده میان تأمین‌کنندگان ثابت، مراکز توزیع و مشتریان، نیازمند تخصیص استراتژیک منابع برای پاسخگویی کارآمد به تقاضا است. این چالش چندوجهی شامل تخصیص مراکز توزیع به تأمین‌کنندگان، تخصیص مشتریان به مراکز توزیع و ایجاد مسیرهای بهینه برای خودروهایی است که در این شبکه پیچیده تردد می‌کنند. محور اصلی این تلاش بهینه‌سازی، تعادل تخصیص منابع با در نظر گرفتن عواملی چون الگوهای متغیر تقاضا، ظرفیت‌های مراکز توزیع و محدودیت‌های خودروها است.

پیچیدگی این مسئله با پهنه‌های زمانی نرم ارائه‌شده توسط مشتریان افزایش می‌یابد که بُعد زمانی را وارد می‌کند و نیازمند زمان‌بندی دقیق برای برآورده کردن انتظارات تحویل است. انحراف از بازه‌های زمانی مشخص‌شده موجب جریمه می‌شود که نیاز به تعادل دقیق میان رضایت مشتری و بهره‌وری منابع را برجسته می‌سازد.

برای نمایش کاربرد عملی چارچوب پیشنهادی، یک مطالعه موردی برگرفته از سناریویی واقعی در بخش لجستیک ارائه شده است. در این مورد، یک پلتفرم تجارت الکترونیک بزرگ به دنبال افزایش کارایی عملیات زنجیره تأمین خود، شامل شبکه‌ای از تأمین‌کنندگان ثابت، مراکز توزیع متعدد و پایگاه متنوعی از مشتریان است. چالش‌های موجود شامل الگوهای متغیر تقاضا، ظرفیت‌های محدود خودروها و مراکز توزیع، و نیاز به رعایت چارچوب‌های زمانی دقیق تحویل است.

این مقاله یک چارچوب جامع برای پرداختن به این پیچیدگی‌ها پیشنهاد می‌دهد که به دنبال بهینه‌سازی عملیات زنجیره تأمین در عین حداکثرسازی بهره‌وری منابع و برآورده کردن انتظارات تحویل است. اساس این رویکرد بر یک چارچوب بهینه‌سازی استوار بنا شده است که برای در نظر گرفتن عدم قطعیت‌های موجود در تقاضا طراحی شده است. با ادغام تهرید شبیه‌سازی شده و یادگیری تقویتی، روش ارائه‌شده ما به طور کارآمد فضای راه‌حل را کاوش می‌کند و راه‌حل‌های بهینه یا نزدیک به بهینه برای نمونه‌های پیچیده و بزرگ مقیاس ارائه می‌دهد.

برای ارزیابی اثربخشی رویکرد پیشنهادی، آزمایش‌های دقیق روی مجموعه‌ای از مسائل آزمایشی انجام می‌شود. مجموعه داده‌های سفارشی که از پژوهش‌های پیشین الهام گرفته‌اند، اساس ارزیابی را تشکیل می‌دهند و پیچیدگی‌های دنیای واقعی را منعکس می‌کنند. عملکرد الگوریتم‌های ما با یک حل‌کننده بهینه مقایسه می‌شود که مزایای ملموس گنجاندن استراتژی‌های مبتنی بر یادگیری را نشان می‌دهد.

علاوه بر این، تحلیل حساسیت برای ارزیابی انطباق‌پذیری و عملکرد چارچوب بهینه‌سازی ما تحت شرایط متغیر انجام می‌شود. این تحلیل بینش‌های ارزشمندی در باره چگونگی تأثیر تغییرات در پارامترهای خاص بر استواری رویکرد ما ارائه می‌دهد و اطلاعات قابل اجرا برای تصمیم‌گیری در مدیریت زنجیره تأمین فراهم می‌کند.

در بخش‌های بعدی، فرمول‌بندی استوار مسئله را ارائه می‌دهیم، روش‌شناسی راه‌حل ارائه‌شده را به تفصیل شرح می‌دهیم و نتایج عددی را ارائه می‌کنیم که در یک ارزیابی جامع از چارچوب پیشنهادی ما به اوج می‌رسد. از طریق این پژوهش، هدف ما ارائه یک رویکرد کاربردی و مؤثر برای بهینه‌سازی زنجیره تأمین در دنیای واقعی است که سازمان‌ها را قادر می‌سازد تا با چابکی و کارایی پیچیدگی‌های عملیات زنجیره تأمین مدرن را مدیریت کنند.

ادامه این مقاله به شرح زیر است: بخش ۲ مرور ادبیات مقالات مرتبط را ارائه می‌دهد. بخش ۳ تعریف مسئله و فرمول‌بندی‌های ریاضی را ارائه می‌کند. در ادامه بخش ۳ فرمول‌بندی‌های بهینه‌سازی استوار ارائه شده است. روش حل پیشنهادی در بخش ۴ شرح داده شده است. همچنین در بخش ۴ شامل نتایج عددی است. مطالعه موردی در ادامه بخش ۴ شرح داده شده است و در نهایت در بخش ۵ به نتیجه‌گیری می‌پردازیم.

۲. مرور ادبیات

۲.۱. مقدمه

در این بخش، به مرور ادبیات مرتبط با زنجیره تأمین، لجستیک، مسائل مسیریابی، حوزه بهینه‌سازی مقاوم، و پژوهش‌های مرتبط با روش‌های حل مبتنی بر یادگیری می‌پردازیم. این بررسی با هدف ایجاد درکی جامع از روندهای جاری، چالش‌ها و رویکردهای نوآورانه در این حوزه‌ها انجام شده است.

۲.۲. زنجیره تأمین، لجستیک و مسائل مسیریابی

ادغام مسیریابی در مسائل زنجیره تأمین یکی از زیرشاخه‌های کلیدی در زمینه مسائل بهینه‌سازی زنجیره تأمین است که توجه بسیاری از پژوهشگران را به خود جلب کرده است. این ادغام، امکان طراحی سیستم‌های کارآمدتر و مقاوم‌تر را فراهم می‌کند. به عنوان مثال، جاوید و آزاد [۱] به مسئله‌ای پیچیده شامل تصمیم‌گیری‌های همزمان در مکان‌یابی، تخصیص، ظرفیت، موجودی و مسیریابی در یک سیستم زنجیره تأمین تصادفی پرداختند. آن‌ها از ترکیب روش‌های جستجوی ممنوعه و شبیه‌سازی تهرید برای حل مسئله خود استفاده کردند. این مطالعه نشان داد که استفاده از روش‌های ترکیبی می‌تواند عملکرد بهتری نسبت به رویکردهای سنتی داشته باشد.

در مطالعه‌ای دیگر، لی و همکاران [۲] به طراحی شبکه زنجیره تأمین پرداختند که شامل تصمیم‌گیری‌های مکان‌یابی تأسیسات، تخصیص تأسیسات، و مسیریابی بود. آن‌ها کاربردهای عملی این مسئله را برجسته کرده و یک الگوریتم ابتکاری برای حل ارائه کردند. این مطالعه نمونه‌ای از نحوه به کارگیری روش‌های نوآورانه در حل مسائل پیچیده زنجیره تأمین است.

از سوی دیگر، شیمید و همکاران [۳] به بررسی گسترش های مسئله کلاسیک مسیریابی وسایل نقلیه در مدیریت زنجیره تأمین پرداختند. گسترش هایی که مورد توجه قرار گرفت شامل اندازه گیری دسته ها، زمان بندی، بسته بندی، گروه بندی، موجودی و مسائل بین مدلی بود. این پژوهش نشان داد که چگونه ترکیب این عوامل می تواند تأثیر چشمگیری بر کارایی زنجیره تأمین داشته باشد.

۱.۲.۲. مسیریابی وسایل نقلیه در زنجیره تأمین

مسائل مسیریابی وسایل نقلیه در زنجیره تأمین برای انواع مختلف محصولات بررسی شده است. به عنوان نمونه، عواد و همکاران [۴] به مرور مطالعات اخیر در توزیع محصولات حساس به دما در زنجیره تأمین سرد پرداختند. این مطالعه نشان داد که توزیع محصولات حساس به دما نیازمند روش هایی است که بتواند تغییرات دمای محیط را مدیریت کند.

در زمینه زنجیره تأمین غذایی، موسوی و بزرگامیری [۵] یک مدل مکان یابی هاب-زمان بندی وسایل نقلیه برای محصولات فاسدشدنی ارائه دادند. آن ها از الگوریتم ژنتیک مرتب سازی غیرمغلوب- (NSGA-II) برای حل مدل چندهدفه خود استفاده کردند. این الگوریتم نه تنها کارایی مدل را بهبود بخشید، بلکه توانست زمان محاسباتی را نیز کاهش دهد.

علاوه بر این، کائو و همکاران [۶] به بررسی مسیریابی وسایل نقلیه در زنجیره تأمین بیومس پرداختند. این مطالعه نشان داد که طراحی مناسب مسیرها می تواند تأثیرات مثبت اقتصادی و زیست محیطی داشته باشد.

توانا و همکاران [۷] یک مدل مکان یابی-موجودی-مسیریابی برای زنجیره تأمین سبزی با انتشار کم کربن ارائه دادند. این مدل به طور خاص بر کاهش انتشار گازهای گلخانه ای تمرکز داشت و عدم قطعیت را نیز در مدل خود لحاظ کرد. این پژوهش نمونه ای از تلاش های پژوهشگران برای ایجاد تعادل بین ملاحظات اقتصادی و زیست محیطی است.

۲.۲.۲. مسیریابی وسایل نقلیه در زنجیره تأمین

پرداختن به پنجره زمانی در تصمیم گیری های مسیریابی یکی از جنبه های پیچیده در طراحی سیستم های زنجیره تأمین است. گوویندان و همکاران [۸] مسئله مکان یابی-مسیریابی دو سطحی با پنجره زمانی را در طراحی شبکه زنجیره تأمین مواد غذایی فاسدشدنی معرفی کردند. این مطالعه نشان داد که چگونه می توان از محدودیت های زمانی به عنوان ابزاری برای بهبود کارایی سیستم استفاده کرد.

یاسینوفسکیا و همکاران [۹] نیز به بررسی پنجره های زمانی و برداشت و تحویل همزمان در زنجیره تأمین با اقلام بازگشت پذیر پرداختند. این مطالعه نشان داد که در نظر گرفتن این عوامل می تواند تأثیر به سزایی در کاهش هزینه ها و بهبود سطح خدمات داشته باشد.

۳.۲.۲. نوآوری های اخیر در بهینه سازی زنجیره تأمین و لجستیک

در سال های اخیر، پژوهشگران توجه بیشتری به مسائل نوظهور در حوزه زنجیره تأمین و لجستیک داشته اند. به عنوان مثال، دای و همکاران [۱۰] با توجه به تقاضای فزاینده برای مراقبت های بهداشتی خانگی، مسئله مکان یابی-مسیریابی را با ناوگانی ترکیبی از وسایل نقلیه برقی و معمولی و همچنین ایستگاه های تعویض باتری مورد بررسی قرار دادند. استفاده آن ها از یک الگوریتم شبیه سازی تبرد رقابتی، عملکرد بهتری نسبت به روش های موجود نشان داد.

دو و همکاران [۱۱] به چالش های شرکت های اکسپرس در مدیریت تحویل های ناشی از خریدهای آنلاین رو به افزایش پرداختند. آن ها از طریق مسئله مسیریابی مکان یابی مشترک دو سطحی چند انباری توانستند راه حل هایی نوآورانه ارائه دهند.

ذبیحیان بی شه و همکاران [۱۲] نیز یک مدل برنامه ریزی مختلط عدد صحیح چندهدفه برای مسئله مکان یابی-مسیریابی پایدار زباله های خطرناک ارائه دادند. تمرکز این پژوهش بر کاهش اثرات زیست محیطی مدیریت زباله های خطرناک بود که به طور مستقیم با ملاحظات پایداری مرتبط است.

این مطالعات نشان دهنده اهمیت استفاده از مدل های نوآورانه و روش های پیشرفته در حل مسائل لجستیکی پیچیده است. چنین روش هایی می توانند به ایجاد سیستم هایی کارآمدتر و پایدارتر کمک کنند و راه حل هایی برای چالش های متعدد در حوزه زنجیره تأمین و لجستیک ارائه دهند.

۳.۲.۲. بهینه سازی استوار

حوزه بهینه سازی استوار طی سال های اخیر رشد قابل توجهی داشته است و مدل های بسیاری را معرفی کرده که به طور چشمگیری به پیشرفت این زمینه کمک کرده اند. این پیشرفت ها، از جمله در کارهای مولوی و همکاران [۱۳]، بن تال و همکاران [۱۴]، و برتسیماس و سیم [۱۵] مشاهده می شود. این مدل ها نه تنها از محبوبیت گسترده ای برخوردارند بلکه با کارایی بالا در حل چالش های پیچیده بهینه سازی، جایگاه خود را به عنوان مشارکت های برجسته در این حوزه تثبیت کرده اند.

۱.۳.۲. کاربردهای بهینه سازی استوار

بهینه سازی استوار کاربردهای گسترده ای در زمینه های مختلف یافته است که در مقالات و مطالعات متعدد منعکس شده است. به عنوان نمونه، پیشوایی و همکاران [۱۶] مدلی برای طراحی شبکه زنجیره تأمین حلقه بسته تحت عدم قطعیت داده ها ارائه کردند. این مدل به طور خاص به مدیریت داده های نامطمئن پرداخته و کاربردهای عملی متعددی دارد. به همین ترتیب، رهبری و همکاران [۱۷] دو مدل استوار برای مدیریت عدم قطعیت های مرتبط با زمان سفر و تازگی محصولات

در مسئله مسیریابی و زمان‌بندی خودروها با داکینگ متقاطع ارائه دادند.

حبیب‌زاده بوکانی و همکاران [۱۸] در مسائل مکان‌یابی هاب‌های تک‌اختصاصی و چنداختصاصی با ظرفیت، پارامترهای هزینه ثابت راه‌اندازی و ظرفیت هر هاب را به عنوان پارامترهای نامطمئن در نظر گرفتند. آن‌ها با استفاده از بهینه‌سازی استوار، این عدم قطعیت‌ها را مدیریت کردند. وارس و همکاران [۱۹] مسئله زمان‌بندی تولید تحت عدم قطعیت را در کارخانه ابره‌بری بررسی کردند و با استفاده از شبیه‌سازی مونت کارلو، رفتار راه‌حل‌های استوار را تحلیل نمودند.

گو و همکاران [۲۰] به اهمیت بهینه‌سازی استوار در مدیریت نوسانات قیمت اجاره در لجستیک شهری اشاره کردند، به‌ویژه در مکان‌یابی میکروهاب‌ها و چالش‌های مسیریابی خودروها. همچنین، فایز و همکاران [۲۱] چارچوب محاسباتی استواری را برای مسئله مسیریابی دوسطحی خودروها ارائه دادند. این مدل که از وسایل نقلیه خودکار برای ارائه کمک‌های اضطراری استفاده می‌کند، از یک مدل بهینه‌سازی استوار دومرحله‌ای برای مدیریت عدم قطعیت تقاضا بهره برد و تکنیک‌های تولید ستون و محدودیت را در خود جای داده‌است.

۲.۳.۲. بهینه‌سازی استوار داده‌محور

در سال‌های اخیر، بهینه‌سازی استوار به حوزه‌های داده‌محور نیز گسترش یافته‌است [۲۲]. این روش‌ها از داده‌ها برای تعریف مجموعه‌های عدم قطعیت استفاده می‌کنند و باعث انعطاف‌پذیری بیشتری در تصمیم‌گیری می‌شوند. شانگ و همکاران [۲۳] مجموعه‌ای نوآورانه از عدم قطعیت‌ها را با استفاده از یادگیری هسته‌ای برای بهینه‌سازی استوار داده‌محور ارائه دادند. این روش به‌طور گسترده توسط محققان مختلف مورد استفاده قرار گرفته‌است.

موسوی و بزرگی‌امیری [۲۴] از بهینه‌سازی استوار داده‌محور مبتنی بر هسته در مسئله مکان‌یابی-مسیریابی هاب با عدم قطعیت در زمان سفر استفاده کردند. همچنین، محسنی و همکاران [۲۵] چارچوب مدیریت انرژی ترانکشی دوسطحی داده‌محور را ارائه دادند. آن‌ها از روش بهینه‌سازی استوار داده‌محور با استفاده از مجموعه عدم قطعیتی که با تخمین چگالی هسته استوار ساخته شده بود، بهره‌بردند.

در نهایت، لی و همکاران [۲۶] مدلی برای بهینه‌سازی استوار داده‌محور دومرحله‌ای ارائه کردند که برای مسئله تولید نیروگاه با انتشار کربن نزدیک به صفر طراحی شده‌است. این مدل توانست تأثیر عدم قطعیت در تولید انرژی را کاهش دهد و نقش مهمی در کاهش آثار زیست‌محیطی ایفا کند.

۴.۲. روش‌های مبتنی بر یادگیری

برای افزایش کارایی روش‌های حل، به‌ویژه روش‌های مبتنی بر فراابتکاری، پژوهشگران اخیراً تکنیک‌های مبتنی بر یادگیری را با روش‌های کلاسیک ترکیب کرده‌اند. این پیشرفت‌ها عمدتاً از حوزه یادگیری ماشین الهام گرفته‌اند، به‌ویژه اصول ریشه‌گرفته از یادگیری

تقویتی. کاریمی مامقان و همکاران [۲۷] یک فراکاوش مبتنی بر یادگیری بر اساس NSGA-II، روش خوشه‌بندی k-Means، و یک الگوریتم جستجوی محلی تکرارشونده برای مسئله مکان‌یابی هاب تحت تراکم ارائه کردند. چنگ و همکاران [۲۸] یک فراکاوش مبتنی بر جنگل تصادفی را در مسئله زمان‌بندی موازی بررسی کردند.

سیدعباسی و همکاران [۲۹] از یادگیری تقویتی برای ترکیب سه نسخه از الگوریتم گرگ خاکستری برای حل مسائل بهینه‌سازی جهانی استفاده کردند. کین و همکاران [۳۰] یک فراابتکاری مبتنی بر یادگیری تقویتی برای مسئله مسیریابی وسایل نقلیه ارائه کردند. همچنین، دی سانتیاگو جونیر و همکاران [۳۱] یک ابر-فراکاوش مبتنی بر یادگیری تقویتی برای حل مسائل چندهدفه پیشنهاد کردند.

در حوزه ادغام تکنیک‌های یادگیری تقویتی، روش q-learning به عنوان رویکردی کلیدی شناخته شده‌است که توسط بسیاری از پژوهشگران مورد استفاده قرار گرفته‌است. گولچوک و اوزسویدان [۳۲] یک فراابتکاری مبتنی بر q-learning و فراابتکاری پیشنهاد کردند. شی و لی [۳۳] یک روش حل مبتنی بر q-learning برای مسئله زمان‌بندی جریان هیبریدی دو مرحله‌ای توزیع‌شده معرفی کردند. نی و همکاران [۳۴] یک الگوریتم کلونی مصنوعی زنبور عسل مبتنی بر q-learning را برای مسئله برنامه‌ریزی مسیر وسیله نقلیه بدون سرنشین مورد بررسی قرار دادند.

اخیراً، روند پیشرفته‌ای در میان پژوهشگران شامل بررسی عمیق‌تر یادگیری تقویتی عمیق برای ارتقاء کارایی روش‌های حل شکل گرفته‌است. ژانگ و همکاران [۳۵] یک رویکرد یادگیری تقویتی عمیق مبتنی بر متا یادگیری برای حل مسائل پیچیده بهینه‌سازی چندهدفه معرفی کردند که با آموزش یک مدل متا به منظور استخراج زیرمدل‌ها برای مسائل مختلف، به همگرایی سریع‌تر به جبهه پارتو دست یافتند. کالستاد و همکاران [۳۶] یک چارچوب فراابتکاری مبتنی بر یادگیری تقویتی عمیق پیشنهاد کردند که انتخاب هیوریستیک را در مقایسه با روش‌های Adaptive Large Neighborhood Search و Uniform Random Selection در حل مسائل بهینه‌سازی ترکیبی بهبود می‌بخشد.

در مورد مقالات اخیر SA، به‌ویژه در ادغام‌های مبتنی بر یادگیری، لین و همکاران [۳۷] یک الگوریتم SA مبتنی بر یادگیری (LSA) برای مسئله چیدمان تسهیلات با نواحی نابرابر معرفی کردند. الگوریتم آن‌ها از نمایش راه‌حل جدید، یک تابع جریمه بهبودیافته، و مجموعه متنوعی از عملگرهای همسایگی بهره‌مند بود که همگی توسط یک کنترل‌کننده مبتنی بر یادگیری تقویتی برای کاوش کارآمد هدایت می‌شدند. الگوریتم ترکیبی کوسانوگلو و همکاران [۳۸] الگوریتم شبیه‌سازی تبرید را با شبکه Q دوگانه عمیق ترکیب کرد که SA از راه‌حل‌های DRL به عنوان حالت‌های اولیه استفاده می‌کند و کاوش در برنامه‌ریزی نگهداری پیچیده را افزایش می‌دهد.

۵.۲. تحلیل شکاف

این مقاله در تقاطع چندین حوزه پویا در بهینه سازی زنجیره تأمین، بهینه سازی استوار و روش های مبتنی بر یادگیری قرار دارد. با این حال، علی رغم کاوش جامع در این زمینه ها، برخی شکاف ها در چشم انداز فعلی همچنان وجود دارد:

در حالی که تکنیک های بهینه سازی استوار به طور مؤثر برای مقابله با عدم قطعیت ها در زنجیره های تأمین به کار می روند، شکافی قابل توجه در زمینه رسیدگی دقیق به الگوهای تقاضای نامطمئن همچنان باقی است.

الگوریتم شبیه سازی تبرید تقویت شده از طریق تکنیک های یادگیری تقویتی، رویکردی نوآورانه در این مطالعه به شمار می رود. با این حال، عمق کاوش در مورد سازگاری و مقیاس پذیری این روش اصلاح شده SA محدود باقی مانده است. با وجود اذعان به گنجاندن یادگیری تقویتی، تجزیه و تحلیل جامع تری از کارایی آن در سناریوهای مختلف و پیچیده زنجیره تأمین می تواند درک عمیق تری از پتانسیل و محدودیت های آن ارائه دهد.

نوآوری های این تحقیق به شرح موارد زیر هستند:

- به کارگیری بهینه سازی استوار بر اساس روش برتسیماس در مسئله زنجیره تأمین با در نظر گرفتن هم زمان مسیریابی و تقاضای غیرقطعی
- ارائه الگوریتم فراابتکاری مبتنی بر یادگیری نوین برای حل مسئله
- ارائه دو روش مبتنی بر یادگیری و یک روش مبتنی بر یادگیری تقویتی به منظور افزایش کارایی الگوریتم فراابتکاری شبیه سازی تبرید
- انجام مطالعه موردی با استفاده از داده های یکی از بزرگ ترین فروشگاه های آنلاین ایران جهت ارزیابی عملکرد مدل در محیط واقعی

۳. تعریف مسئله و مدل سازی ریاضی

۱.۳. بیان مسئله

چالش اصلی مورد بررسی در این تحقیق حول بهینه سازی عملیات زنجیره تأمین با تأمین کنندگان ثابت، مراکز توزیع و مشتریان می چرخد. این مسئله چندوجهی شامل تخصیص مراکز توزیع به تأمین کنندگان، تخصیص مشتریان به مراکز توزیع و ایجاد مسیرهای کارآمد برای وسایل نقلیه بین مراکز توزیع و مشتریان و همچنین بین خود مشتریان است. در حالی که مسیرهای بین تأمین کنندگان و مراکز توزیع از پیش تعیین شده است، تصمیم گیری اصلی در تعیین تخصیص بهینه منابع برای تأمین تقاضا در هر نقطه از زنجیره تأمین است.

علاوه بر این، مسئله نیازمند بررسی دقیق تخصیص وسایل نقلیه به هر مرکز توزیع است، با در نظر گرفتن ظرفیت های محدود هم وسایل

نقلیه و هم مراکز توزیع. این امر مستلزم ایجاد تعادل ظریف در اطمینان از استفاده بهینه منابع بدون اضافه بار کردن هر گره خاص است. پیچیدگی ها ناشی از تعامل عوامل مختلف مانند الگوهای متغیر تقاضای مشتریان، ظرفیت های متفاوت مراکز توزیع و محدودیت های وسایل نقلیه از نظر حجم است.

در افزودن به پیچیدگی مسئله، مشتریان پنجره های زمانی نرم را ارائه می دهند که بر نامهریزی تحویل را پیچیده تر می کند. تعیین زمان های بهینه تحویل باید ملاحظات زودرسی و تأخیر را در نظر بگیرد، با جریمه هایی که با انحراف از چارچوب های زمانی مشخص شده مرتبط هستند. این بُعد زمانی لایه اضافی پیچیدگی را به فرآیند بهینه سازی اضافه می کند، زیرا نیازمند یک عمل متعادل سازی دقیق بین برآورده کردن انتظارات مشتری و استفاده کارآمد از منابع موجود است.

در شکل ۱، نمایش گرافیکی مسئله ارائه شده است که شبکه ای از تأمین کنندگان ثابت، مراکز توزیع و مشتریان و همچنین مسیرهای مختلف متصل کننده آن ها را نشان می دهد. این کمک بصری یک نمای کلی مختصر از روابط پیچیده لجستیکی که نیاز به بهینه سازی دارند را ارائه می دهد. هدف اصلی این مطالعه توسعه یک چارچوب جامع است که به این پیچیدگی ها می پردازد. از طریق این تلاش، ما به دنبال دستیابی به عملیات زنجیره تأمین ساده شده، مقرون به صرفه و مشتری محور هستیم که استفاده از منابع را حداکثر می کند در حالی که انتظارات تحویل را برآورده می سازد.

جدول ۱: مقایسه مقالات مطرح در حوزه ادبیات موضوع با تحقیق حاضر

مقاله	سال انتشار	مسیریابی	عدم قطعیت	بهینه سازی استوار	محدودیت پنجره زمانی	یادگیری تقویتی	روش حل مبتنی بر یادگیری
[۱]	۲۰۱۰	✓	✓				
[۲]	۲۰۱۰	✓					
[16]	۲۰۱۱		✓	✓			
[3]	۲۰۱۳	✓					
[8]	۲۰۱۴	✓	✓		✓		
[18]	۲۰۱۶	✓	✓	✓			
[9]	۲۰۱۷	✓	✓		✓		
[17]	۲۰۱۹	✓	✓	✓			
[4]	۲۰۲۱	✓	✓				
[6]	۲۰۲۱	✓					
[7]	۲۰۲۱	✓	✓				
[33]	۲۰۲۲					✓	✓
[36]	۲۰۲۳					✓	✓
[39]	۲۰۲۴					✓	✓
تحقیق حاضر		✓	✓	✓	✓	✓	✓

$$\text{Min} \sum_{u=1}^S \sum_{k=1}^C TC_{uk} \theta_{uk} + VE \sum_{k=1}^C v_k + \sum_{i=1}^N at_i + \sum_{i=1}^N e_i + \sum_{i=1}^N l_i \quad (1)$$

محدودیت‌ها:

$$\sum_{k=1}^C y_{ik} = 1 \quad \forall i = 1, 2, \dots, N \quad (2)$$

$$2x_{ijk} \leq y_{jk} + y_{ik} \quad \begin{matrix} \forall i, j \\ = 1, 2, \dots, N, k \\ = 1, 2, \dots, C \end{matrix} \quad (3)$$

$$\sum_{\substack{l=0 \\ i \neq j \\ N+1}}^N x_{ijk} = y_{jk} \quad \begin{matrix} \forall j \\ = 1, 2, \dots, N, k \\ = 1, 2, \dots, C \end{matrix} \quad (4)$$

$$\sum_{\substack{j=1 \\ i \neq j \\ N}}^N x_{ijk} = y_{ik} \quad \begin{matrix} \forall i \\ = 1, 2, \dots, N, k \\ = 1, 2, \dots, C \end{matrix} \quad (5)$$

$$\sum_{j=1}^N x_{0,jk} = v_k \quad \forall k = 1, 2, \dots, C \quad (6)$$

$$\sum_{i=1}^N x_{i,N+1,k} = v_k \quad \forall k = 1, 2, \dots, C \quad (7)$$

$$at_i \geq TC_{ki} x_{0,ik} \quad \begin{matrix} \forall i \\ = 1, 2, \dots, N, k \\ = 1, 2, \dots, C \end{matrix} \quad (8)$$

$$at_j + M(1 - x_{ijk}) \geq TC_{ij} + at_i \quad \begin{matrix} \forall i, j \\ = 1, 2, \dots, N, k \\ = 1, 2, \dots, C, i \\ \neq j \end{matrix} \quad (9)$$

$$load_i \geq \tilde{d}_i x_{0,ik} \quad \begin{matrix} \forall i \\ = 1, 2, \dots, N, k \\ = 1, 2, \dots, C \end{matrix} \quad (10)$$

$$load_j + M(1 - x_{ijk}) \geq \tilde{d}_i + load_i \quad \begin{matrix} \forall i, j \\ = 1, 2, \dots, N, k \\ = 1, 2, \dots, C, i \\ \neq j \end{matrix} \quad (11)$$

$$load_i \leq VC \quad \forall i = 1, 2, \dots, N \quad (12)$$

$$at_i \leq B_i + l_i \quad \forall i = 1, 2, \dots, N \quad (13)$$

$$at_i + e_i \geq A_i \quad \forall i = 1, 2, \dots, N \quad (14)$$

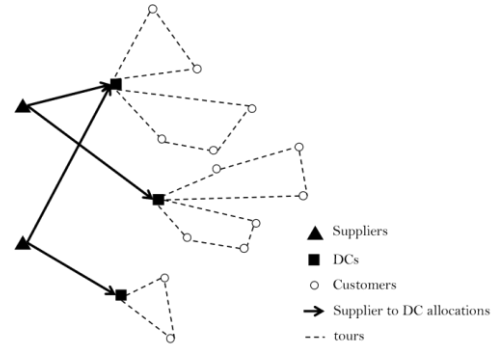
$$\sum_{i=1}^N y_{ik} \tilde{d}_i \leq \beta_k \quad \forall k = 1, 2, \dots, C \quad (15)$$

$$\sum_{u=1}^S \theta_{uk} = \beta_k \quad \forall k = 1, 2, \dots, C \quad (16)$$

$$\beta_k \leq CC_k \quad \forall k = 1, 2, \dots, C \quad (17)$$

$$\begin{matrix} at_i, load_i, e_i, l_i, \beta_k, \theta_{uk} \geq 0 \\ v_k \in \mathbb{Z} \\ x_{ijk}, y_{ik} \in \{0, 1\} \end{matrix} \quad \begin{matrix} \forall i, j \\ = 0, 1, 2, \dots, N \\ + 1, k \\ = 1, 2, \dots, C \end{matrix} \quad (18)$$

تابع هدف (۱) شامل پنج بخش است. بخش اول هزینه‌های کل حمل و نقل بین تأمین‌کنندگان و مراکز توزیع را محاسبه می‌کند. بخش دوم هزینه استفاده از وسایل نقلیه در مراکز توزیع را محاسبه می‌کند. بخش سوم هزینه حمل و نقل بین مشتریان را محاسبه می‌کند. بخش چهارم هزینه زودرسی در تحویل‌ها و بخش پنجم هزینه تأخیر در تحویل‌ها را محاسبه می‌کند.



شکل ۱: نمایش گرافیکی مسئله

۲.۳. فرموله‌سازی مسئله

اندیس‌ها:

$i, j \in \{0, 1, 2, \dots, N + 1\}$	اندیس مشتریان
$k \in \{1, 2, \dots, C\}$	اندیس مراکز توزیع
$u \in \{1, 2, \dots, S\}$	اندیس تأمین‌کنندگان

پارامترها:

VE	هزینه استفاده از وسایل نقلیه
VC	ظرفیت وسایل نقلیه
d_i	تقاضای مشتری i
A_i	زمان شروع پنجره زمانی مشتری i
B_i	زمان پایان پنجره زمانی مشتری i
CC_k	ظرفیت مرکز توزیع k
M	یک عدد بزرگ
TC_{ij}	هزینه حمل و نقل بین مشتری i و j
TC_{ki}	هزینه حمل و نقل بین مرکز توزیع k و مشتری i
TC_{uk}	هزینه حمل و نقل بین تأمین‌کننده u و مرکز توزیع k

متغیرهای مثبت:

at_i	زمان رسیدن سفارش مشتری i
$load_i$	میزان بار وسیله نقلیه در زمان ارسال سفارش مشتری i
l_i	دیر کرد رسیدن سفارش مشتری i
e_i	زود کرد رسیدن سفارش مشتری i
β_k	مجموع محصولات تجمیع شده در مرکز توزیع k
θ_{uk}	جریان محصولات از تأمین‌کننده u به مرکز توزیع k

متغیرهای عدد صحیح:

v_k	تعداد وسیله نقلیه تخصیص داده شده به مرکز توزیع k
-------	--

متغیرهای صفر و یک:

x_{ijk}	مساوی با یک اگر سفارش مشتری i قبل از مشتری j رسیده باشد (سفارش هر دو مشتری در مرکز توزیع k است)، در غیر این صورت مساوی با صفر
y_{ik}	مساوی با یک اگر مشتری i به مرکز توزیع k تخصیص یافته باشد، در غیر این صورت مساوی با صفر

تابع هدف:

که η_{ij} در آن در بازه -1 و 1 قرار می گیرد. η_{ij} در واقع یک متغیر بازه است که برای کنترل میزان تاثیر عدم قطعیت درون محدودیت ها است. در ادامه مسئله (۲۰) را می تواند به صورت زیر نوشت:

$$\begin{aligned} \min cx \\ s.t. \sum_j \bar{a}_{ij}x_j + \max_{\eta_{ij}}(\bar{a}_{ij}\eta_{ij}x_j) \leq b_i, \forall i \\ \sum_j \eta_{ij} \leq \Gamma_i, \forall i \\ 0 \leq \eta_{ij} \leq 1, \forall j \in J_i \\ x_j \geq 0, \forall \end{aligned} \quad (22)$$

در مسئله (۲۲) یک پارامتر Γ_i به عنوان محدودیت بودجه تعریف شده است که تجمع انحراف محدود شود. پارامتر Γ_i آستانه ای برای عدم قطعیت تجمعی مشخص می کند و انحراف کلی در تقاضا را محدود می سازد تا مقاومت مدل افزایش یابد. این بدین معناست که فرض می کنیم تغییرات پارامتر نامعین نمی تواند از آستانه تعریف شده Γ_i فراتر رود. برای تابع حفاظتی در رابطه (۲۲)، داریم:

$$\begin{aligned} \max_{\eta_{ij}}(\bar{a}_{ij}\eta_{ij}x_j) \\ s.t. \sum_j \eta_{ij} \leq \Gamma_i, \forall i \\ 0 \leq \eta_{ij} \leq 1, \forall j \in J_i \\ x_j \geq 0, \forall j \end{aligned} \quad (23)$$

مسئله دوگان (۲۳) به صورت زیر نوشته می شود:

$$\begin{aligned} \min \Gamma_i z_i + \sum_{j \in J_i} p_{ij} \\ s.t. \Gamma_i + p_{ij} \geq \bar{a}_{ij}x_j, \forall i, j \in J_i \\ p_{ij} \geq 0, \forall i, j \in J_i \\ z_i \geq 0, \forall i, x_j \geq 0, \forall j \in J_i \end{aligned} \quad (24)$$

در ادامه معادل استوار مسئله (۲۲) به صورت زیر خواهد بود:

$$\begin{aligned} \min cx \\ s.t. \sum_j \bar{a}_{ij}x_j + \Gamma_i z_i + \sum_{j \in J_i} p_{ij} \leq b_i, \forall i \\ \Gamma_i + p_{ij} \geq \bar{a}_{ij}x_j, \forall i, j \in J_i \\ p_{ij} \geq 0, \forall i, j \in J_i \\ z_i \geq 0, \forall i, x_j \geq 0, \forall j \in J_i \end{aligned} \quad (25)$$

در مسئله مورد نظر در این تحقیق تقاضای شبکه زنجیره تامین در بازه $[\bar{d}_i - \bar{d}_i, \bar{d}_i + \bar{d}_i]$ نوسان می کند که در آن بیانگر تقاضای پایه مشتری i است و به عنوان یک نقطه مرجع برای ارزیابی نوسانات تقاضا به کار می رود و $\bar{d}_i$ بیانگر بیشترین مقدار انحراف تقاضا برای مشتری i است. در نتیجه عبارت $\phi_i = \frac{\bar{d}_i - \bar{d}_i}{\bar{d}_i}$ را تعریف می کنیم که بیانگر انحراف معیار از تقاضای مشاهده شده است.

بر اساس روابط بالا، فرمولاسیون استوار محدودیت های (۱۰)، (۱۱) و (۱۵) به صورت زیر تعریف می شوند:

$$\bar{d}_i x_{0,ik} + \Gamma_i z'_i + p'_i \leq load_i \quad \forall i = 1, 2, \dots, N, k = 1, 2, \dots, C \quad (26)$$

$$\Gamma_i + p'_i \geq \bar{d}_i x_{0,ik} \quad \forall i = 1, 2, \dots, N, k = 1, 2, \dots, C \quad (27)$$

$$\begin{aligned} \bar{d}_i + \Gamma_i z''_i + p''_i + load_i &\leq load_j & \forall i, j &= 1, 2, \dots, N, k = 1, 2, \dots, C, i \neq j \\ &+ M(1 - x_{ijk}) & & \end{aligned} \quad (28)$$

$$\Gamma_i + p''_i \geq \bar{d}_i \quad \forall i = 1, 2, \dots, N \quad (29)$$

معادله (۲) اطمینان می دهد که هر مشتری فقط به یک مرکز توزیع تخصیص داده شود. معادله (۳) مشخص می کند که هر وسیله نقلیه درون مرکز توزیع تنها در صورتی به مشتریان i و j خدمت رسانی خواهد کرد که این مشتریان به مرکز توزیع k تخصیص داده شده باشند. معادلات (۴) و (۵) محدودیت های کلاسیک مسئله مسیریابی و سیله نقلیه هستند. این معادلات بررسی می کنند که آیا هر وسیله نقلیه فقط یک بار وارد و خارج مشتری تخصیص داده شده می شود. معادلات (۶) و (۷) تأیید می کنند که شروع و پایان مسیرهای وسایل نقلیه در مراکز توزیع به تعداد وسایل نقلیه در آن مراکز توزیع بستگی دارد. معادلات (۸) و (۹) زمان رسیدن و وسایل نقلیه به مشتریان را محاسبه می کنند. معادلات (۱۰) و (۱۱) بار وسایل نقلیه را هنگام خدمت رسانی به مشتریان محاسبه می کنند. معادله (۱۲) اطمینان می دهد که بار وسایل نقلیه از ظرفیت آن ها تجاوز نکند. معادلات (۱۳) و (۱۴) زودرسی و تأخیر سفارش ها را محاسبه می کنند. معادله (۱۵) مقدار کل محصولات تجمیع شده در هر مرکز توزیع را محاسبه می کند. معادله (۱۶) تعداد محصولات از هر تأمین کننده به مراکز توزیع را محاسبه می کند. معادله (۱۷) اطمینان می دهد که تعداد محصولات در هر مرکز توزیع از ظرفیت آن ها تجاوز نکند. در نهایت، معادله (۱۸) نوع متغیرها را در مدل پیشنهادی نشان می دهد.

۳.۳. فرموله سازی استوار

در این مقاله ما برای پرداختن به عدم قطعیت های موجود در تقاضا، از رویکرد بهینه سازی استوار استفاده شده است. هدف از این بهینه سازی استوار، طراحی مسیریابی است که نوسانات تقاضا را در نظر می گیرد. فرمول بندی بهینه سازی استوار از کار برتسیماس و سیم [15] الهام گرفته شده است. فرض کنید مسئله بهینه سازی به شرح زیر است:

$$\begin{aligned} \min cx \\ s.t. \sum_j \bar{a}_{ij}x_j \leq b_i, \forall i, \forall \bar{a}_{ij} \in J_i \\ x \in X, x_j \end{aligned} \quad (19)$$

یا:

$$\begin{aligned} \min cx \\ s.t. \max_{\bar{a}_{ij} \in J_i} \left(\sum_j \bar{a}_{ij}x_j \right) \leq b_i, \forall i \\ x \in X, x_j \geq 0 \end{aligned} \quad (20)$$

پارامتر $\bar{a}_{ij}$ تحت شرایط عدم قطعیت قرار دارد که در بازه $[\bar{a}_{ij} - \bar{a}_{ij}, \bar{a}_{ij} + \bar{a}_{ij}]$ قرار می گیرد که در آن مقدار اسمی پارامتر است و $\bar{a}_{ij}$ بیشترین مقدار انحراف از مقدار اسمی است. همچنین J_i بیانگر مجموعه عدم قطعیت پارمترها برای محدودیت i است. در ادامه داریم:

$$\eta_{ij} = \frac{\bar{a}_{ij} - \bar{a}_{ij}}{\bar{a}_{ij}} \quad (21)$$

ترکیبی، از جمله VRP ، به دلیل توانایی آن در کاوش کارآمد فضای راه حل و فرار از بهینه‌های محلی مناسب است [40].

به منظور انتخاب آگاهانه ساختار همسایگی برای تولید راه حل جدید، سه رویکرد یادگیری تقویتی به کار گرفته می‌شود. این رویکردها برای یادگیری و ارزیابی ارزش و کارایی هر ساختار همسایگی موجود طراحی شده‌اند. با بهره‌گیری از اصول یادگیری تقویتی، الگوریتم به طور پویا استراتژی اکتشاف خود را تطبیق می‌دهد و بر حرکت‌هایی تمرکز می‌کند که احتمال بیشتری برای دستیابی به راه حل‌های بهتر دارند. این ماهیت تطبیقی، توانایی الگوریتم را در پیمایش مؤثر فضای راه حل پیچیده نمونه‌های VRP در مقیاس بزرگ افزایش می‌دهد و در نهایت به دستیابی به راه حل‌های با کیفیت بالا کمک می‌کند. از طریق ادغام الگوریتم شبیه‌سازی تبرید با یادگیری تقویتی، رویکرد ما به طور هم‌افزایی نقاط قوت هر دو تکنیک را ترکیب می‌کند و ابزاری قوی و همه‌جانبه برای پرداختن به چالش‌های بهینه‌سازی زنجیره تأمین در دنیای واقعی ارائه می‌دهد.

نمایش راه حل: نمایش راه حل در مدل پیشنهادی این تحقیق از تعدادی بردار استفاده می‌کند که به طور جمعی تخصیص‌ها و تصمیمات مسیریابی را در برمی‌گیرد. هر بردار مربوط به مسیرهای یک مرکز توزیع خاص است و به طور مؤثر هم تخصیص مشتریان به هر مرکز توزیع و هم مسیرهای بعدی برای خدمت‌رسانی به این مشتریان تخصیص یافته را نشان می‌دهد. در هر بردار، وجود مقادیر صفر به عنوان جداکننده عمل می‌کند و مرزهای بین وسایل نقلیه مختلف در ناوگان مرکز توزیع را مشخص می‌کند. مسئله تخصیص از تأمین‌کنندگان به مراکز توزیع به عنوان یک مرحله بهینه‌سازی ثانویه در نظر گرفته می‌شود که به راحتی به عنوان یک مسئله برنامه ریزی خطی قابل حل است. علاوه بر این، محدودیت‌های اضافی مانند محدودیت‌های ظرفیت به عنوان متغیرهای نقض در تابع هدف گنجانده شده‌اند. شکل ۲ نمایش یک راه حل را نشان می‌دهد. در این مثال که دارای دو مرکز توزیع است، گره‌های ۲، ۳، ۵ و ۷ به مرکز توزیع ۱ تخصیص داده شده‌اند. اولین وسیله نقلیه در این مرکز توزیع گره‌های ۲ و ۳ را خدمت‌رسانی می‌کند و وسیله نقلیه دوم ابتدا گره ۵ و سپس گره ۷ را خدمت‌رسانی می‌کند.

مسیرهای مرکز توزیع اول	2	3	0	5	7
مسیرهای مرکز توزیع دوم	6	0	1	4	

شکل ۲: مثالی از نمایش راه حل

در این تحقیق، ما هفت ساختار جستجوی همسایگی متمایز را برای پیمایش و بهره‌برداری نظام‌مند از فضاها راه حل طراحی کرده‌ایم. این ساختارهای جستجوی سفارشی نقش محوری در افزایش کارایی و اثربخشی رویکرد بهینه‌سازی ما ایفا می‌کنند:

$$\sum_{i=1}^N y_{ik} \bar{d}_i + \sum_{i=1}^N (\Gamma_i z_i''' + p_i''') \leq \beta_k \quad \forall k = 1, 2, \dots, C \quad (30)$$

$$\sum_{i=1}^N (\Gamma_i + p_i''') \geq \sum_{i=1}^N y_{ik} \bar{d}_i \quad \forall k = 1, 2, \dots, C \quad (31)$$

$$z_i', z_i'', z_i''', p_i', p_i'', p_i''' \geq 0 \quad \forall i = 1, 2, \dots, N \quad (32)$$

بر اساس محدودیت‌های فوق، مدل برنامه‌ریزی خط عدد صحیح مختلط استوار به صورت زیر خواهد بود:

$$\text{Min} \sum_{u=1}^S \sum_{k=1}^C TC_{uk} \theta_{uk} + VE \sum_{k=1}^S v_k + \sum_{i=1}^N at_i + \sum_{i=1}^N e_i + \sum_{i=1}^N l_i \quad \text{s.to.} \quad (2)-(9), (12)-(14), (16)-(18), (26)-(32)$$

۴. روش حل

مدل پیشنهادی این تحقیق بر اساس چارچوب مسئله مسیریابی وسایل نقلیه (VRP) طراحی شده است که به عنوان یک مدل ریاضی کلاسیک، امکان بهینه‌سازی مسیرهای توزیع را فراهم می‌آورد. انتخاب VRP به عنوان چارچوب اصلی تحقیق، به دلیل انطباق آن با نیازهای بهینه‌سازی توزیع در زنجیره‌های تأمین واقعی صورت گرفته است؛ زنجیره‌هایی که اغلب شامل چندین تأمین‌کننده، مراکز توزیع و مشتریان متنوع می‌باشند. VRP به طور گسترده در بهینه‌سازی مسیرها کاربرد داشته و پایه ریاضی قدرتمندی برای مدیریت محدودیت‌های مرتبط با ظرفیت، تقاضا و مسیریابی در شبکه‌های لجستیکی ارائه می‌دهد.

با این حال، لازم به ذکر است که VRP به عنوان یک مسئله NP -سخت شناخته می‌شود، به این معنا که یافتن یک راه حل بهینه دقیق برای نمونه‌های بزرگ، از نظر محاسباتی غیرممکن است. این چالش موجب می‌شود که کاربرد مستقیم VRP در سناریوهای واقعی محدود شود و به بررسی روش‌های جایگزین نیاز باشد که بتوانند نمونه‌های بزرگ را در بازه‌های زمانی منطقی مدیریت کنند.

برای رفع این محدودیت، در این پژوهش از الگوریتم شبیه‌سازی تبرید مبتنی بر یادگیری (LSA) استفاده شده است. این روش فراابتکاری، ترکیبی از اصول شبیه‌سازی تبرید (SA) و قابلیت‌های تطبیق‌پذیری یادگیری است. با ادغام این تکنیک‌ها، هدف رویکرد پیشنهادی ما، پیمایش کارآمد فضای جستجو و ارائه راه حل‌های بهینه یا نزدیک بهینه برای نمونه‌های پیچیده و بزرگ VRP می‌باشد. بدین ترتیب، این پژوهش ابزاری مؤثر و عملی برای بهینه‌سازی توزیع در زنجیره‌های تأمین واقعی ارائه می‌دهد.

۱.۴. شبیه‌سازی تبرید

این بخش الگوریتم شبیه‌سازی تبرید را که برای حل مدل پیشنهادی توسعه یافته است، ارائه می‌دهد. الگوریتم شبیه‌سازی تبرید یک تکنیک بهینه‌سازی فراابتکاری قدرتمند است که از فرآیند تبرید در متالورژی الهام گرفته شده است. این روش به‌ویژه برای حل مسائل بهینه‌سازی

درج درون مسیری: این ساختار جستجوی هم سایگی بر عملیات درون مسیری تمرکز دارد. شامل درج یک مشتری در مسیر موجود یک وسیله نقلیه است.

درج بین مسیری: ساختار درج بین مسیری دامنه را به عملیات بین مسیری گسترش می دهد. شامل انتقال یک مشتری از مسیر یک وسیله نقلیه به وسیله نقلیه دیگر است.

تعویض درون مسیری: ساختار تعویض درون مسیری، مشابه ساختار درج درون مسیری، بر عملیات درون مسیری تمرکز دارد. با این حال، در این مورد، عملیات شامل تعویض موقعیت دو مشتری در همان مسیر یک وسیله نقلیه است.

تعویض بین مسیری: عویض بین مسیری دامنه را به عملیات بین مسیری گسترش می دهد. شامل تبادل مشتریان بین مسیرهای دو وسیله نقلیه مختلف است.

معکوس سازی درون مسیری: ساختار معکوس سازی درون مسیری به عملیات درون مسیری می پردازد. بر معکوس کردن بخش هایی از مسیر در برنامه سفر یک وسیله نقلیه تمرکز دارد.

افزودن و سیله نقلیه: افزودن وسیله نقلیه بعد متفاوتی را به فرآیند بهینه سازی معرفی می کند. شامل افزودن یک وسیله نقلیه کاملاً جدید به ناوگان است.

حذف وسیله نقلیه: در مقابل، حذف وسیله نقلیه حذف یک وسیله نقلیه از ناوگان را در نظر می گیرد. این عملیات تأثیر کوچک سازی ناوگان بر راه حل کلی را ارزیابی می کند.

۲.۴. شبیه سازی تیرید مبتنی بر یادگیری

در الگوریتم های فرا ابتکاری، انتخاب راهبردی ساختارهای جستجوی همسایگی (SS) برای تولید راه حل های جدید یک ملاحظه حیاتی است. اغلب، SS ها به صورت تصادفی انتخاب می شوند، بدون آنکه نشانه روشنی از مناسب بودن آنها وجود داشته باشد. با این حال، رویکرد آگاهانه تر شامل ارزیابی عملکرد هر SS و انتخاب بر اساس این ارزیابی است. به عنوان مثال، وقتی یک SS بدون هیچ بهبودی به کار گرفته می شود، نشان می دهد که جستجوی همسایگی انتخاب شده ممکن است برای وضعیت فعلی مناسب نباشد و بنابراین مستلزم نوعی جریمه است. برعکس، اگر یک SS منجر به بهبود قابل توجهی شود، نشان دهنده کارآمدی آن است و پاداش متناسب را شایسته می شود. این اصل در چارچوب یادگیری قرار می گیرد، جایی که مدل با پیگیری جریمه ها و پاداش های مرتبط، درک خود را از ارزش هر اقدام تطبیق و پالایش می کند.

تابع ارزش برای هر اقدام، با $Q(a)$ برای همه a در مجموعه A نشان داده می شود. علاوه بر این، $SS(A)$ نشان دهنده فراوانی انتخاب اقدام A که مربوط به یک ساختار جستجوی همسایگی خاص است. همچنین، موقعیتی را در نظر بگیرید که پس از انجام یک اقدام، سیستم پاداشی دریافت می کند، و α نرخ یادگیری را نشان می دهد. در این زمینه،

فرمول بندی برای هر یک از سه نوع یادگیری پی شنهادی را می توان به طور مختصر به شرح زیر ارائه کرد:

- روش تطبیقی: این روش یک قاعده به روزرسانی پویای اساسی است که تابع ارزش هر اقدام را با میانگین گیری از پاداش های واقعی دریافت شده تخمین می زند و استفاده کارآمد از منابع حافظه را فراهم می کند. این رویکرد توابع ارزش را به صورت تکراری به روز می کند و به الگوریتم اجازه می دهد تا تخمین های خود را در طول زمان تطبیق دهد و تغییرات محیط را به طور مؤثر درک کند.

$$Q(a) = Q(a) + \frac{R - Q(a)}{SS(a)}$$

- رویکرد تطبیق با محیط پویا: این رویکرد برای پرداختن به محیط های پویا طراحی شده است، جایی که پاداش های مرتبط با اقدامات ممکن است در طول زمان تغییر کنند. در حالی که روش رویکرد تطبیقی یک محیط ثابت با توزیع احتمال ثابت پاداش ها را فرض می کند، این فرض می تواند به راحتی نقض شود.

$$Q(a) = Q(a) + \alpha[R - Q(a)]$$

- یادگیری Q: این رویکرد یک الگوریتم یادگیری تقویتی خارج از سیاست است که برای تشخیص راه حل های بهینه یا نزدیک به بهینه در مسائل فرآیند تصمیم گیری مارکوف طراحی شده است. برای کاربرد مؤثر، نیازمند تعیین حالت ها، اقدامات و پاداش ها است.

با این حال، تعریف حالت ها در یادگیری Q ویژگی متمایزی دارد. این امر بر اساس کمی سازی تعداد مواردی است که راه حل فعلی بهبودی نشان نمی دهد. به طور خاص، اگر UI_t تعداد دفعاتی باشد که راه حل فعلی بهبود نیافته است، متغیر حالت به فواصل مشخصی گسسته سازی می شود که هر کدام با یک حالت متفاوت مطابقت دارد. این حالت ها به عنوان بازتابی از فرآیند بهینه سازی عمل می کنند و تصمیم گیری الگوریتم را هدایت می کنند.

$$s_t = \begin{cases} 1, & 0 \leq UI_t < \zeta_1 \\ 2, & \zeta_1 \leq UI_t < \zeta_2 \\ 3, & \zeta_2 \leq UI_t < \zeta_3 \\ 4, & \zeta_3 \leq UI_t < \zeta_4 \\ 5, & \zeta_4 \leq UI_t \end{cases}$$

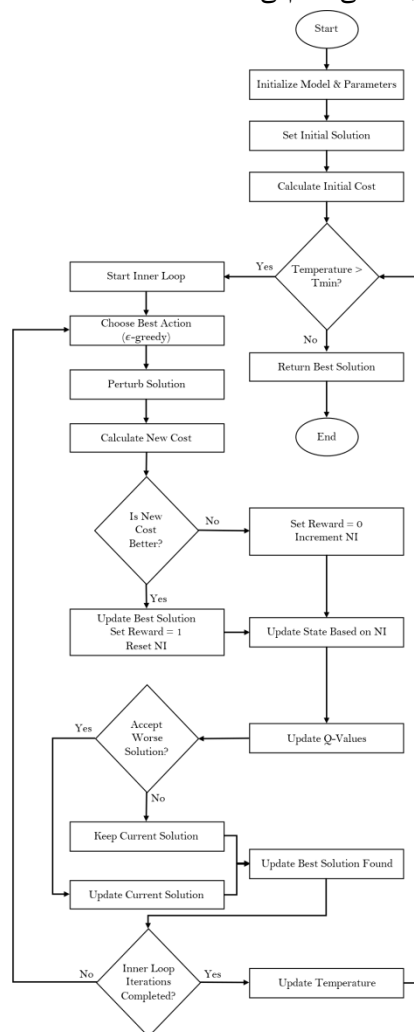
در این محیط پویا، هنگامی که حالت سیستم s_t است و اقدام a_t انتخاب می شود و پاداش متناظر r_t دریافت می شود، تابع ارزش اقدام با استفاده از یک قاعده یادگیری تفاوت زمانی به روز می شود. این به روزرسانی هم پاداش فوری (r_t) و هم ارزش تخمینی حالت بعدی (s_{t+1}) را در نظر می گیرد که توسط ضریب تخفیف $\gamma \in (0,1)$ وزن دهی شده اند. قاعده به روزرسانی اطمینان حاصل می کند که الگوریتم می آموزد تا پیامدهای بلندمدت اقدامات خود را پیش بینی و ارزیابی کند، که به پالایش کلی تخمین های ارزش اقدام کمک می کند.

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha[r_t + \gamma \max_{b \in A} Q(s_{t+1}, b) - Q(s_t, a_t)]$$

انتخاب ساختارهای جستجوی همسایگی توسط استراتژی اپسیلون-حریه صانه کنترل می شود. این رویکرد تعادلی بین اکتشاف و بهره‌برداری ایجاد می کند و به الگوریتم اجازه می دهد گاهی اوقات یک ساختار تصادفی را انتخاب کند (اکتشاف)، در حالی که عمدتاً انتخاب ساختارهایی را ترجیح می دهد که در تکرارهای قبلی امیدوارکننده تر نشان داده اند (بهره‌برداری).

$$SS(a_t) = \begin{cases} \text{Random Selection,} & \text{rand}(0,1) < \epsilon \\ \text{argmax}_{s_t \in S} (Q(s_t, a_t)), & \text{Otherwise} \end{cases}$$

در شکل ۳ نمودار جریان الگوریتم شبیه سازی تبرید مبتنی بر یادگیری را نشان می دهد که اصول یادگیری تقویتی را با بهینه سازی شبیه سازی تبرید سنتی ادغام می کند.



شکل ۲: فلوچارت الگوریتم شبیه سازی تبرید مبتنی بر یادگیری

۵. محاسبات عددی

۱.۵. ارزیابی کارایی الگوریتم

مدل پیشنهادی برای ارزیابی جامع عملکرد الگوریتم های پیشنهادی، ضروری است که این روش ها تحت آزمایش های دقیق و متنوع بر روی مجموعه ای از مسائل تست قرار گیرند. با توجه به این که هیچ مسئله

تست استاندارد ی که به طور خاص برای مدل پیشنهادی ما طراحی شده باشد وجود ندارد، مجموعه داده ها با دقت و به صورت تصادفی تولید شده اند و از داده های به کار رفته در پژوهش های پیشین، به ویژه تحقیق انجام شده توسط کوردو و همکاران [41]، الهام گرفته شده اند.

مجموعه داده استفاده شده از [41] شامل موقعیت مشتریان و مراکز توزیع، تقاضاها و پنجره های زمانی است. این مجموعه داده های سفارشی تولید شده، پایه و اساس ارزیابی ما را تشکیل داده و پیچیدگی ها و ظرایف دنیای واقعی را منعکس می کنند.

برای اطمینان از یکپارچگی تنظیمات آزمایش، سایر پارامترهای ضروری نیز به صورت سیستماتیک تولید شده اند. این پارامترها به صورت زیر تعیین شده اند:

$$\begin{aligned} VE &= 2000 \\ VC &= 70 \\ CC_k &\sim U\left(\frac{\sum_i d_i}{N}, \frac{1.2 \times \sum_i d_i}{N}\right) \end{aligned}$$

شایان ذکر است که تمام الگوریتم های پیشنهادی با دقت بالا به زبان برنامه نویسی جولیا کدنویسی و اجرا شده اند. محاسبات روی سیستمی با پردازنده Intel Core i5 با فرکانس ۲ گیگاهرتز و ۱۶ گیگابایت حافظه انجام گرفته است. برای اعتبار سنجی کیفیت نتایج به دست آمده، از حل کننده Gurobi به منظور یافتن راه حل های بهینه استفاده شده است.

اندازه مسائل تست بر اساس تعداد تأمین کنندگان، مراکز توزیع و مشتریان موجود در هر سناریو تعیین می شود. در نتایج محاسباتی این پژوهش، اندازه هر مسئله تست به صورت یک سه تایی (S, C, N) نمایش داده شده است، که در آن C ، N و به ترتیب نشان دهنده تعداد تأمین کنندگان، مراکز توزیع و مشتریان در نمونه مورد نظر هستند.

در این مطالعه، مسائل کوچک شامل سناریوهایی با حداکثر ۲۵ مشتری (نمونه های شماره ۱ تا ۷) تعریف شده اند، در حالی که مسائل متوسط شامل پیچیدگی هایی با کمتر از ۵۰ مشتری (نمونه های شماره ۸ تا ۱۵) هستند.

مسائل تست به عنوان پایه ای برای مقایسه روش های مختلف حل مسئله عمل می کنند. به دلیل ماهیت تصادفی الگوریتم های فراابتکاری در انتخاب راه حل و حرکات، هر نمونه مسئله به منظور اطمینان از پایداری در فرآیند ارزیابی، تحت ۲۰ تکرار قرار می گیرد. روش های مورد ارزیابی شامل موارد زیر هستند:

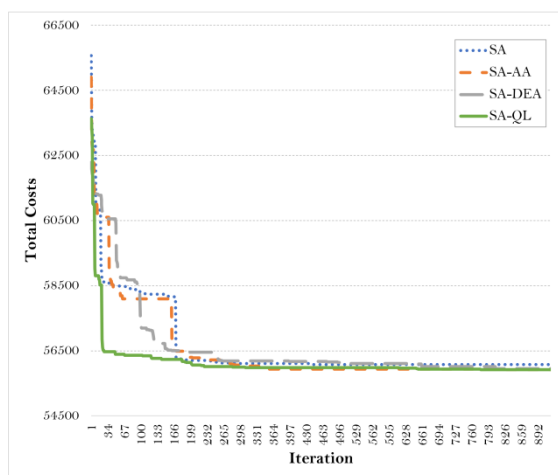
- حل کننده بهینه (Optimal Solver)
- شبیه سازی تبرید بدون رویکرد یادگیری (SA)
- شبیه سازی تبرید با روش یادگیری "رویکرد تطبیقی-SA" (AA)
- شبیه سازی تبرید با روش یادگیری "تطبیق با محیط پویا" (SA-DEA)
- شبیه سازی تبرید با روش یادگیری (SA-QL) Q-Learning

این مقایسه به منظور ارزیابی کارایی و اثربخشی روش های پیشنهادی در مواجهه با چالش های مسائل مختلف انجام می پذیرد.

15	(3,5,50)	61835.63	61835.63	62000.93	63020.92	61877.97
16	(3,5,55)	64158.91	67578.70	67863.68	68236.92	67505.03
17	(4,5,55)	68119.43	70190.46	70523.58	70289.65	70130.58
18	(4,5,60)	70415.14	72450.16	72442.59	72533.00	72404.20
19	(4,6,65)	74142.22	75645.87	76089.52	76145.30	75477.01
20	(4,6,70)	69431.45	80276.88	80933.69	81500.40	80124.97
21	(5,6,75)	82243.31	85211.33	86495.53	87496.04	84942.89
22	(5,5,80)	-	95863.90	96541.05	96646.18	94925.55
23	(5,5,85)	-	98804.82	98706.36	100492.69	98367.77
24	(5,6,90)	-	105367.21	106198.51	106882.37	104668.88
25	(5,7,100)	-	131268.97	131398.47	131919.72	130913.92

جدول ۳: مقایسه زمان محاسبات (بر حسب ثانیه) در تمامی روش‌های حل

#	Size	Optimal	SA	SA-AA	SA-DEA	SA-QL
1	(2,2,10)	0.05	0.31	0.38	0.41	0.37
2	(2,3,10)	0.05	0.41	0.54	0.39	0.46
3	(2,2,15)	0.15	0.42	0.43	0.32	0.45
4	(2,3,15)	0.15	0.43	0.44	0.41	0.47
5	(2,4,15)	0.17	0.44	0.49	0.46	0.67
6	(3,4,20)	4.33	0.54	0.53	0.43	0.56
7	(3,4,25)	0.46	0.61	0.56	0.40	0.50
8	(3,4,30)	8.21	0.54	0.68	0.42	0.63
9	(3,5,30)	9.59	0.52	0.55	0.41	0.53
10	(3,5,35)	10.16	0.51	0.56	0.42	0.59
11	(4,5,35)	8.44	0.50	0.56	0.86	0.56
12	(3,5,40)	161.62	0.51	0.67	0.54	0.60
13	(3,5,45)	194.81	0.57	0.64	0.59	0.67
14	(4,5,45)	209.60	0.59	0.65	0.50	0.59
15	(3,5,50)	558.84	0.53	0.61	0.50	0.61
16	(3,5,55)	801.91	0.59	0.65	0.59	0.67
17	(4,5,55)	850.51	0.61	0.57	0.49	0.54
18	(4,5,60)	1614.77	0.55	0.58	0.47	0.55
19	(4,6,65)	1992.21	0.59	0.77	0.60	0.68
20	(4,6,70)	2801.59	0.65	0.70	0.66	0.72
21	(5,6,75)	4980.10	0.72	0.72	0.62	0.75
22	(5,5,80)	-	0.66	0.73	0.59	0.70
23	(5,5,85)	-	0.66	0.72	0.61	0.72
24	(5,6,90)	-	0.73	0.81	0.66	0.81
25	(5,7,100)	-	0.76	0.84	0.70	0.85



شکل ۳: مقایسه همگرایی الگوریتم‌ها

۲.۵. تحلیل حساسیت روی مدل استوار

در این بخش، ما یک تحلیل حساسیت انجام می‌دهیم تا ارزیابی کنیم که چگونه تغییرات در پارامترهای خاص بر پایداری چارچوب بهینه‌سازی پیشنهادی ما تأثیر می‌گذارند. این تحلیل بینش‌های حیاتی در مورد قابلیت انطباق و عملکرد سیستم تحت شرایط متغیر ارائه می‌دهد.

جدول ۲ ارزش‌یابی جامعی از عملکرد الگوریتم‌ها ارائه می‌دهد و اثربخشی آن‌ها را بر اساس مقادیر هدف بررسی می‌کند. شایان توجه است که ادغام روش‌های یادگیری، بهبود چشمگیری در کارایی کلی الگوریتم شبیه‌سازی تبرید (SA) به همراه داشته و اهمیت به‌کارگیری این تکنیک‌های تطبیقی را برجسته می‌کند.

بررسی دقیق‌تر نتایج عددی نشان می‌دهد که در نمونه‌های مسائل با مقیاس کوچک و متوسط، الگوریتم‌های شبیه‌سازی تبرید مبتنی بر یادگیری، نتایجی تقریباً مشابه ارائه می‌دهند. این امر نشان‌دهنده آن است که در سناریوهایی با پیچیدگی‌های نسبتاً محدود، انتخاب استراتژی یادگیری ممکن است تأثیر کمتری بر راه حل نهایی داشته باشد.

با این حال، با افزایش مقیاس مسائل و ورود به نمونه‌های بزرگ‌تر، مزیت منحصربه‌فرد روش *Q-Learning* آشکار می‌شود. این روش با عملکرد برتر خود در چنین چالش‌های بزرگ مقیاس، به‌عنوان گزینه‌ای برتر معرفی می‌گردد و برتری خود را در مواجهه با این دسته از مسائل نشان می‌دهد.

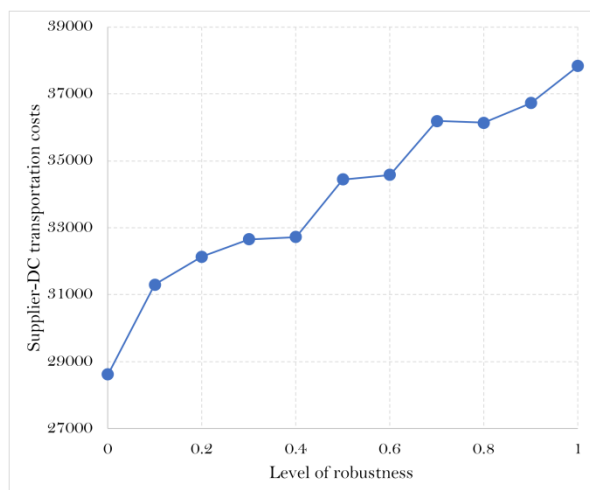
رفتار همگرایی الگوریتم‌های پیشنهادی در شکل ۳ نشان داده شده است. این نمودار کاهش مقدار تابع هدف را در طول تکرارها نمایش می‌دهد و بیانگر یک روند نزولی ثابت است که نشان می‌دهد الگوریتم‌ها به تدریج به سمت راه حل‌های بهینه حرکت می‌کنند. روندهای همگرایی مشاهده شده در تمامی الگوریتم‌ها نشان‌دهنده اثربخشی آن‌ها در حل مسئله است، هرچند هر الگوریتم ویژگی‌های منحصربه‌فردی در رفتار همگرایی از خود نشان می‌دهد.

با توجه به جدول ۳، عملکرد الگوریتم‌ها این بار از منظر کارایی محاسباتی مورد بررسی قرار گرفته است. این معیار کلیدی به ما امکان می‌دهد نه تنها کیفیت راه حل‌ها، بلکه سرعت دستیابی به آن‌ها را نیز ارزیابی کنیم. شایان ذکر است که استفاده از استراتژی‌های مبتنی بر یادگیری ممکن است منجر به افزایش جزئی در زمان محاسبات شود، اما این سرمایه‌گذاری اغلب با بهبودهای چشمگیر در کیفیت راه حل‌ها جبران می‌شود.

این یافته‌ها به‌طور کلی بر تطبیق پذیری و قابلیت اطمینان الگوریتم‌های شبیه‌سازی تبرید مبتنی بر یادگیری در طیف وسیعی از اندازه‌های مسائل تأکید می‌کنند. پایداری آن‌ها در برابر تکرارهای متعدد، که در قالب ۲۰ تکرار برای هر نمونه مسئله نشان داده شده است، اعتبار و کارایی این روش‌ها را در مواجهه با چالش‌های پیچیده بهینه‌سازی زنجیره تأمین تقویت می‌کند.

جدول ۲: مقایسه مقادیر تابع هدف در تمامی روش‌های حل

#	Size	Optimal	SA	SA-AA	SA-DEA	SA-QL
1	(2,2,10)	14194.16	14194.16	14194.16	14194.16	14194.16
2	(2,3,10)	15311.26	15337.83	15311.26	15337.83	15311.26
3	(2,2,15)	19646.81	19761.54	19679.82	19646.81	19679.82
4	(2,3,15)	24656.54	24656.54	24709.69	24717.02	24656.54
5	(2,4,15)	24533.81	24533.81	24618.69	24533.81	24696.49
6	(3,4,20)	30869.42	30869.42	31058.97	30989.09	31413.60
7	(3,4,25)	33733.78	33733.78	35746.38	35916.44	35753.81
8	(3,4,30)	40444.32	40748.29	40555.50	40708.44	40752.94
9	(3,5,30)	40134.33	41249.13	40777.00	40522.75	40871.95
10	(3,5,35)	46956.22	47166.14	47878.05	47813.64	47784.27
11	(4,5,35)	42661.51	43650.75	44268.55	44038.84	44026.69
12	(3,5,40)	51013.40	51633.60	51159.94	51272.83	51565.69
13	(3,5,45)	55212.11	55851.47	56013.09	56235.08	55790.47
14	(4,5,45)	53011.56	55686.67	55990.27	55928.23	55550.69



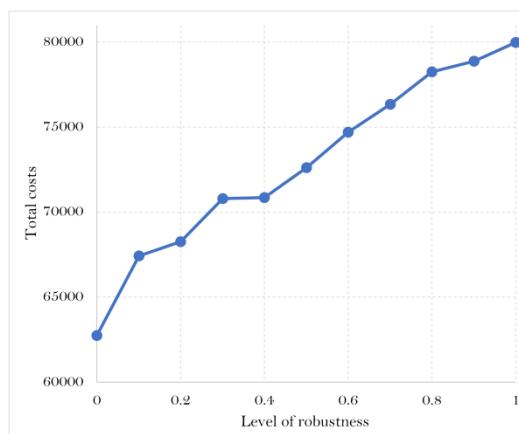
شکل ۶: تغییرات هزینه‌های حمل و نقل از تامین کنندگان به مراکز توزیع به ازای تغییرات درجه عدم قطعیت

شکل ۶ به تحلیل حساسیت چارچوب بهینه‌سازی نسبت به هزینه‌های حمل و نقل بین تأمین کنندگان و مراکز توزیع می‌پردازد. در سطوح پایین استواری (۰٫۱ تا ۰٫۴)، انحراف قابل توجهی مشاهده نمی‌شود. با این حال، هنگامی که از این بازه عبور می‌کنیم، یک روند افزایشی چشمگیر و مداوم به‌وضوح نمایان می‌شود. این نشان می‌دهد که در سطوح پایین استواری، سیستم ممکن است نسبت به تغییرات در هزینه‌های حمل و نقل حساسیت کمتری داشته باشد. اما با افزایش استواری، تأثیر این هزینه‌ها بیشتر نمایان می‌شود. این بینش می‌تواند در تصمیم‌گیری‌ها در زمینه انتخاب تأمین کنندگان و استراتژی‌های مدیریت هزینه‌های حمل و نقل راهگشا باشد.

۳.۵. مطالعه موردی

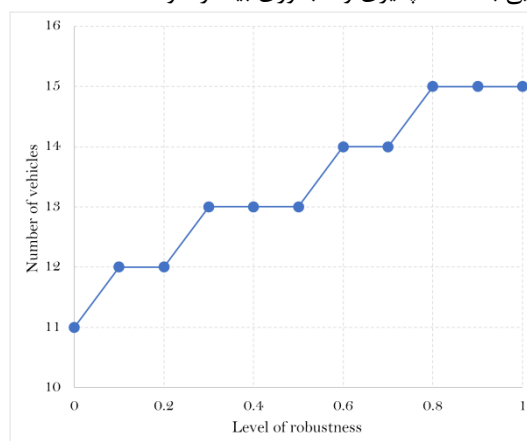
در این بخش، یک مطالعه موردی جامع در خصوص مسئله پیشنهادی ارائه خواهد شد که یک کاربرد دنیای واقعی را نمایش می‌دهد و جنبه‌ها و راه‌حل‌های بالقوه آن را توضیح می‌دهد.

مطالعه موردی با استفاده از داده‌های واقعی به دست آمده از یکی از بزرگ‌ترین فروشگاه‌های آنلاین تجاری در ایران انجام شده است، که به دلیل دسترسی گسترده به بازار و پایگاه مصرف کنندگان خود شناخته شده است. مجموعه داده‌های استفاده شده در این مطالعه به استان تهران اختصاص داشتند و نمونه‌ای نمایانگر از دینامیک‌های بازار و رفتارهای مصرف کنندگان در این منطقه را ارائه می‌دهند. این مجموعه داده‌ها شامل مختصات جغرافیایی تأمین کنندگان، مراکز توزیع (DCs) و مشتریان، همراه با اطلاعات تقاضای اساسی شامل پنجره‌های زمانی خاص مشتریان و تقاضاهای آن‌ها می‌باشد. علاوه بر این، مجموعه داده‌ها شامل اطلاعات حیاتی درباره ظرفیت وسایل نقلیه و مراکز توزیع است که برای تحلیل دینامیک‌های لجستیکی و زنجیره تأمین در استان تهران ضروری است.



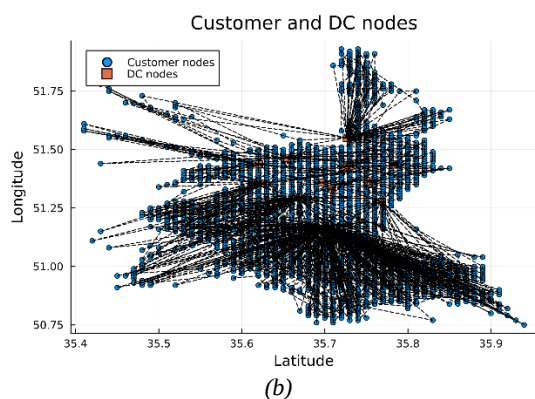
شکل ۴: تغییرات هزینه به ازای تغییرات درجه عدم قطعیت

شکل ۴ تأثیر تغییرات در هزینه‌های کل بر استواری رویکرد بهینه‌سازی ما را نمایش می‌دهد. به‌وضوح یک روند افزایشی در سطوح استواری مشاهده می‌شود که با افزایش هزینه‌های کل همراه است. این امر نشان می‌دهد که در چارچوب این تحلیل، هزینه‌های بالاتر با سطوح بالاتری از استواری ارتباط دارند. این ممکن است به این معنا باشد که در این سناریو، سرمایه‌گذاری در منابع با هزینه بالاتر می‌تواند منجر به راه‌حلی با انعطاف‌پذیری و تاب‌آوری بیشتر شود.



شکل ۵: تغییرات تعداد وسیله نقلیه در شبکه به‌ازای تغییرات درجه عدم قطعیت

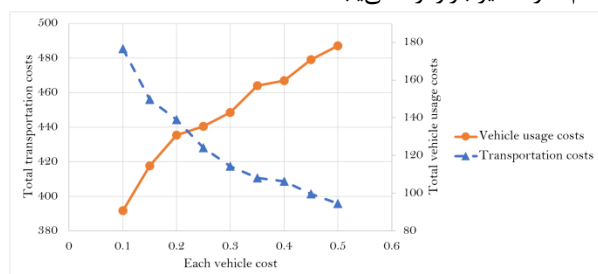
شکل ۵ حساسیت چارچوب بهینه‌سازی را به تغییرات در تعداد وسایل نقلیه در شبکه نشان می‌دهد. ما یک روند افزایشی در سطوح استواری مشاهده می‌کنیم که با افزایش تعداد وسایل نقلیه همراه است. با این حال، نکته مهم این است که اگرچه یک روند صعودی وجود دارد، شیب این روند نسبتاً ملایم است. این نشان می‌دهد که در بازه مورد بررسی، تأثیر تعداد وسایل نقلیه بر استواری قابل توجه است اما به اندازه‌ای قوی نیست که به‌طور کامل تعیین کننده باشد. این ممکن است به این معنا باشد که عوامل دیگری نیز نقش‌های مهمی در تعیین اندازه بهینه ناوگان دارند.



شکل ۸: مسیر وسایل نقلیه در مطالعه موردی برای حالت a: بدون عدم قطعیت و b: بیشترین میزان عدم قطعیت

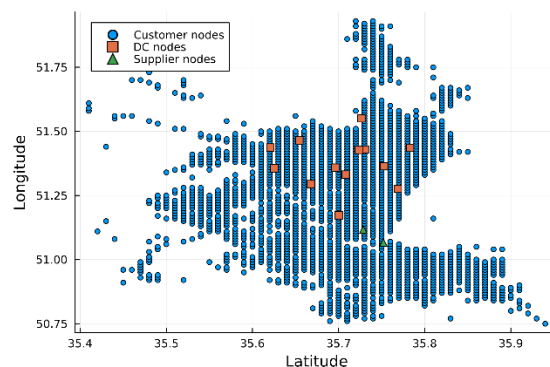
در تحلیل عدم قطعیت مطالعه موردی، اجرای سطح کامل پایداری تغییر قابل توجهی در پیکربندی شبکه نشان می دهد. استراتژی بهینه سازی که برای سازگاری با عدم اطمینانها در داده های تقاضا طراحی شده است، موجب ادغام قابل توجهی در مسیرها می شود و آنها را به سمت یک مرکز توزیع متمرکز هدایت می کند. این متمرکزسازی در بهره برداری از کارایی شبکه لجستیکی نقش حیاتی دارد. با هدایت تحویل ها از طریق مجموعه ای متمرکز از مراکز توزیع (DCs)، سیستم در برابر نوسانات تقاضا تاب آوری بیشتری پیدا می کند. به طور کلی، این رویکرد نه تنها عملیات لجستیکی را ساده سازی می کند، بلکه شبکه را در برابر اختلالات یا نوسانات تقاضای مشتریان مقاوم تر می سازد.

علاوه بر این، اعمال سطح کامل عدم قطعیت باعث توزیع گسترده تری از تقاضا میان چندین مرکز توزیع می شود. این پراکندگی نقش مهمی در تقویت انطباق پذیری و تاب آوری شبکه ایفا می کند. با غیرمتمرکز کردن نقاط تحویل در میان مجموعه ای وسیع تر از مراکز توزیع، سیستم رویکردی متعادل تر و پاسخ دهنده تر به تقاضای مشتریان فراهم می کند. این استراتژی توزیع خطر مرتبط با بارگذاری بیش از حد مراکز توزیع خاص را کاهش می دهد و تخصیص منابع را به طور عادلانه تر و کارآمدتر تضمین می کند. در نتیجه، شبکه قادر به مدیریت بهتر الگوهای مختلف تقاضا می شود و عملکرد کلی و پاسخ دهی آن به چشم انداز متغیر بازار ارتقا می یابد.



شکل ۹: تغییرات در هزینه های کل حمل و نقل و هزینه های کل وسایل نقلیه مطابق با هزینه هر وسیله نقلیه

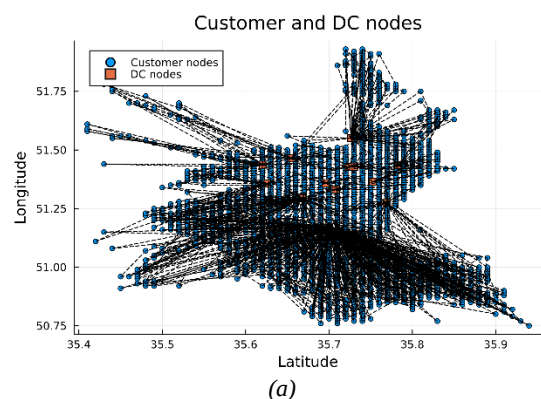
این مجموعه داده ها شامل مجموعاً ۱۸۸۷ گره مشتری، ۱۳ گره مرکز توزیع و ۲ گره تأمین کننده است که یک شبکه جامع از ذینفعان کلیدی در چارچوب لجستیکی را نشان می دهد. چیدمان فضایی و توزیع این گره ها در شکل ۷ به طور بصری نمایش داده شده است و محل های استراتژیک آنها در استان تهران را به تصویر می کشد. تقاضاهای مشتریان در چهار پنجره زمانی مختلف شامل ۹-۱۲ صبح، ۱۲-۳ عصر، ۳-۶ عصر و ۶-۹ شب تقسیم بندی شده اند که الگوهای زمانی متنوع تقاضای مشتریان را منعکس می کند.



شکل ۷: توزیع گره ها در شبکه لجستیکی استان تهران، شامل گره های مشتری، مراکز توزیع و تأمین کنندگان برای مطالعه موردی

مطالعه موردی با استفاده از روش حل پیشنهادی که در ابتدای این بخش توضیح داده شده است، حل می شود و قابلیت و کارایی آن در مواجهه با چالش های پیچیده لجستیکی موجود در چشم انداز بازار متنوع تهران را نشان می دهد.

برای این مطالعه موردی، مجموعه داده با سطوح مختلفی از استواری (robustness) مورد بررسی قرار می گیرد که به طور استراتژیک برای مدیریت عدم اطمینان های ذاتی در داده های تقاضا طراحی شده اند. شکل ۸ نتایج این مطالعه موردی را به طور بصری نمایش می دهد و راه حل های مسیریابی وسایل نقلیه را در دو سناریو متمایز نشان می دهد: یکی بدون سطح صفر پایداری و دیگری با استفاده از سطح کامل پایداری. این تصاویر به وضوح تفاوت های آشکار در استراتژی های بهینه سازی مسیرها را در مواجهه با درجات مختلف عدم اطمینان در مجموعه داده ها نشان می دهند.



یکی دیگر از جنبه‌های حیاتی که در این مطالعه موردی تحلیل شده است، تأثیر هزینه وسیله نقلیه (VE) بر چارچوب لجستیکی است. این پارامتر به‌طور قابل توجهی بر تعداد وسایل نقلیه مورد استفاده در شبکه و همچنین مدت زمان‌های حمل و نقل تجمعی تأثیر می‌گذارد. تصویر در شکل ۹ رابطه مستقیم بین هزینه وسیله نقلیه و هزینه‌های حمل و نقل را روشن می‌کند. با تغییرات در پارامتر هزینه وسیله نقلیه، هزینه‌های کل حمل و نقل و هزینه‌های کلی وسایل نقلیه تغییرات متناسبی را تجربه می‌کنند. این نمایش بصری حساسیت سیستم لجستیکی به تغییرات در هزینه‌های وسیله نقلیه را برجسته می‌کند و توضیح می‌دهد که چگونه تنظیمات در این پارامتر می‌تواند مستقیماً بر جنبه‌های اقتصادی شبکه تأثیر بگذارد و در نتیجه بر تعداد وسایل نقلیه مورد استفاده و هزینه‌های کل حمل و نقل تأثیر بگذارد.

روند قابل مشاهده یک رابطه جذاب بین افزایش هزینه وسیله نقلیه (VE) و تأثیرات آن بر شبکه لجستیکی را نشان می‌دهد. با افزایش VE ، که نشان‌دهنده استفاده از وسایل نقلیه گران‌تر است، هزینه‌های کلی وسایل نقلیه به‌طور قابل پیش‌بینی افزایش می‌یابد. همزمان، یک اثر جریان‌کننده جالب به‌وجود می‌آید، به‌طوری‌که هزینه‌های کل حمل و نقل کاهش می‌یابد. این کاهش در هزینه‌های حمل و نقل به‌وجود می‌آید و وسایل نقلیه گران‌تر مربوط می‌شود که با وجود افزایش هزینه‌های کلی وسایل نقلیه، قادر به کاهش زمان‌های سفر به دلیل قابلیت‌ها یا کارایی‌های بهبود یافته خود هستند.

با این حال، در میان این همبستگی مستقیم، یک مشاهده حیاتی به‌وجود می‌آید: یک مقدار بهینه برای هزینه وسیله نقلیه (VE) وجود دارد که به‌طور قابل توجهی به نفع شبکه است. پس از عبور از یک آستانه خاص، افزایش تدریجی هزینه وسیله نقلیه دیگر به‌طور متناسب هزینه‌های حمل و نقل را کاهش نمی‌دهد و منجر به بازدهی کاهش می‌شود. این نقطه تغییر نشان‌دهنده مقدار بهینه برای VE در چارچوب لجستیکی است، که تعادلی را نشان می‌دهد که در آن معامله میان هزینه‌های بالاتر وسیله نقلیه و کاهش هزینه‌های حمل و نقل به تعادل می‌رسد و در نتیجه کارایی و اثربخشی هزینه شبکه بهینه می‌شود.

۶. نتیجه‌گیری

در این پژوهش، چارچوبی جامع برای بهینه‌سازی عملیات زنجیره تأمین شامل تأمین‌کنندگان ثابت، مراکز توزیع و مشتریان ارائه شده است. چالش‌های مورد بررسی شامل تخصیص منابع، مسیریابی خودروها و برآورده کردن انتظارات مشتریان در پنجره‌های زمانی مشخص بود. در رویکرد ما تکنیک‌های بهینه‌سازی استوار با تبرید شبیه‌سازی شده، ادغام شده و از انطباق‌پذیری یادگیری تقویتی برای کاوش فضای پیچیده راه‌حل بهره‌می‌برد.

از طریق آزمایش و ارزیابی گسترده، اثربخشی چارچوب پیشنهادی خود را در طیفی از اندازه‌های مسئله و سناریوها نشان داده‌ایم. نتایج، مزایای قابل توجهی را در کاهش استراتژی‌های مبتنی بر یادگیری را

برجسته می‌کند که به‌ویژه در نمونه‌های بزرگ مقیاس مشهود است، جایی که روش یادگیری Q عملکرد برتری را نشان داد.

علاوه بر این، تحلیل حساسیت ما بینش‌های ارزشمندی درباره انطباق‌پذیری سیستم با تغییرات در پارامترهای کلیدی ارائه کرد. درک چگونگی تأثیر تغییرات در هزینه‌ها، تعداد خودروها و هزینه‌های حمل و نقل بر استواری سیستم، اطلاعات قابل اجرا برای تصمیم‌گیرندگان در مدیریت زنجیره تأمین فراهم می‌کند.

مطالعه موردی، برگرفته از یک سناریوی لجستیک واقعی، کاربرد عملی چارچوب ما را به نمایش گذاشت. پیاده‌سازی موفقیت‌آمیز، مزایای ملموسی را از نظر افزایش کارایی، بهره‌وری منابع و برآورده کردن انتظارات تحویل نشان داد.

پژوهش ما یک ابزار کاربردی و مؤثر برای بهینه‌سازی زنجیره تأمین در دنیای واقعی ارائه می‌دهد. چارچوب ما با پرداختن به پیچیدگی‌های ذاتی در عملیات زنجیره تأمین مدرن، سازمان‌ها را قادر می‌سازد تا با چابکی و کارایی با چالش‌ها روبرو شوند. ادغام تکنیک‌های بهینه‌سازی استوار با تبرید شبیه‌سازی شده و یادگیری تقویتی، پتانسیل روش‌شناسی‌های پیشرفته را برای متحول کردن شیوه‌های مدیریت زنجیره تأمین نشان می‌دهد.

همچنان که حوزه مدیریت زنجیره تأمین به تکامل خود ادامه می‌دهد، مسیرهای امیدبخش متعددی برای پژوهش‌های آینده وجود دارند. اولاً، بررسی ادغام فناوری‌های نوظهور مانند بلاک‌چین و هوش مصنوعی می‌تواند کارایی و شفافیت عملیات زنجیره تأمین را بیشتر افزایش دهد. علاوه بر این، گسترش چارچوب برای در نظر گرفتن بهینه‌سازی چندهدفه، گنجاندن معیارهای پایداری زیست‌محیطی و پرداختن به الگوهای تقاضای پویا در زمان واقعی، حوزه‌هایی هستند که نیازمند بررسی بیشتر هستند.

References

- [1] A. A. Javid and N. Azad, "Incorporating location, routing and inventory decisions in supply chain network design," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 5, pp. 582–597, 2010.
- [2] J.-H. Lee, I.-K. Moon, and J.-H. Park, "Multi-level supply chain network design with routing," *International Journal of production research*, vol. 48, no. 13, pp. 3957–3976, 2010.
- [3] V. Schmid, K. F. Doerner, and G. Laporte, "Rich routing problems arising in supply chain management," *European Journal of Operational Research*, vol. 224, no. 3, pp. 435–448, 2013.
- [4] M. Awad, M. Ndiaye, and A. Osman, "Vehicle routing in cold food supply chain logistics: A literature review," *The International Journal of Logistics Management*, vol. 32, no. 2, pp. 592–617, 2021.
- [5] M. Musavi and A. Bozorgi-Amiri, "A multi-objective sustainable hub location-scheduling problem for perishable food supply chain," *Computers & Industrial Engineering*, vol. 113, pp. 766–778, 2017.

- [22] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Mathematical Programming*, vol. 167, pp. 235–292, 2018.
- [23] C. Shang, X. Huang, and F. You, "Data-driven robust optimization based on kernel learning," *Computers & Chemical Engineering*, vol. 106, pp. 464–479, 2017.
- [24] M. Musavi and A. Bozorgi-Amiri, "Data-driven robust optimization for hub location-routing problem under uncertain environment," *Journal of Industrial and Systems Engineering*, 2023.
- [25] S. Mohseni, M. S. Pishvae, and R. Dashti, "Privacy-preserving energy trading management in networked microgrids via data-driven robust optimization assisted by machine learning," *Sustainable Energy, Grids and Networks*, vol. 34, p. 101011, 2023.
- [26] Y. Li, Y. Sun, J. Liu, C. Liu, and F. Zhang, "A data driven robust optimization model for scheduling near-zero carbon emission power plant considering the wind power output uncertainties and electricity-carbon market," *Energy*, p. 128053, 2023.
- [27] M. Karimi-Mamaghan, M. Mohammadi, A. Pirayesh, A. M. Karimi-Mamaghan, and H. Irani, "Hub-and-spoke network design under congestion: A learning based metaheuristic," *Transportation research part e: logistics and transportation review*, vol. 142, p. 102069, 2020.
- [28] C.-Y. Cheng, P. Pourhejazy, K.-C. Ying, S.-F. Li, and C.-W. Chang, "Learning-based metaheuristic for scheduling unrelated parallel machines with uncertain setup times," *IEEE Access*, vol. 8, pp. 74065–74082, 2020.
- [29] A. Seyyedabbasi, R. Aliyev, F. Kiani, M. U. Gulle, H. Basyildiz, and M. A. Shah, "Hybrid algorithms based on combining reinforcement learning and metaheuristic methods to solve global optimization problems," *Knowledge-Based Systems*, vol. 223, p. 107044, 2021.
- [30] W. Qin, Z. Zhuang, Z. Huang, and H. Huang, "A novel reinforcement learning-based hyper-heuristic for heterogeneous vehicle routing problem," *Computers & Industrial Engineering*, vol. 156, p. 107252, 2021.
- [31] V. A. de Santiago Junior, E. Özcan, and V. R. de Carvalho, "Hyper-heuristics based on reinforcement learning, balanced heuristic selection and group decision acceptance," *Applied Soft Computing*, vol. 97, p. 106760, 2020.
- [32] İ. Gölcük and F. B. Ozsoydan, "Q-learning and hyper-heuristic based algorithm recommendation for changing environments," *Engineering Applications of Artificial Intelligence*, vol. 102, p. 104284, 2021.
- [33] B. Xi and D. Lei, "Q-learning-based teaching-learning optimization for distributed two-stage hybrid flow shop scheduling with fuzzy processing time," *Complex System Modeling and Simulation*, vol. 2, no. 2, pp. 113–129, 2022.
- [34] X. Ni, W. Hu, Q. Fan, Y. Cui, and C. Qi, "A Q-learning based multi-strategy integrated artificial bee colony algorithm with application in unmanned vehicle path planning," *Expert Systems with Applications*, vol. 236, p. 121303, 2024.
- [35] Z. Zhang, Z. Wu, H. Zhang, and J. Wang, "Meta-learning-based deep reinforcement learning for multiobjective optimization problems," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.
- [36] J. Kallestad, R. Hasibi, A. Hemmati, and K. Sörensen, "A general deep reinforcement learning hyperheuristic framework for solving combinatorial optimization problems," *European Journal of Operational Research*, vol. 309, no. 1, pp. 446–468, 2023.
- [37] J. Lin, A. Shen, L. Wu, and Y. Zhong, "Learning-based simulated annealing algorithm for unequal area facility
- [6] J. X. Cao, Z. Zhang, and Y. Zhou, "A location-routing problem for biomass supply chains," *Computers & Industrial Engineering*, vol. 152, p. 107017, 2021.
- [7] M. Tavana, H. Tohidi, M. Alimohammadi, and R. Lesansalmasi, "A location-inventory-routing model for green supply chains with low-carbon emissions under uncertainty," *Environmental Science and Pollution Research*, vol. 28, pp. 50636–50648, 2021.
- [8] K. Govindan, A. Jafarian, R. Khodaverdi, and K. Devika, "Two-echelon multiple-vehicle location-routing problem with time windows for optimization of sustainable supply chain network of perishable food," *International journal of production economics*, vol. 152, pp. 9–28, 2014.
- [9] G. Iassinovskaia, S. Limbourg, and F. Riane, "The inventory-routing problem of returnable transport items with time windows and simultaneous pickup and delivery in closed-loop supply chains," *International Journal of Production Economics*, vol. 183, pp. 570–582, 2017.
- [10] Z. Dai, Z. Zhang, and M. Chen, "The home health care location-routing problem with a mixed fleet and battery swapping stations using a competitive simulated annealing algorithm," *Expert Systems with Applications*, vol. 228, p. 120374, 2023.
- [11] J. Du, X. Wang, X. Wu, F. Zhou, and L. Zhou, "Multi-objective optimization for two-echelon joint delivery location routing problem considering carbon emission under online shopping," *Transportation Letters*, vol. 15, no. 8, pp. 907–925, 2023.
- [12] A. Zabihiyan-Bisheh, H. R. Vandchali, V. Kayvanfar, and F. Werner, "A sustainable multi-objective model for the hazardous waste location-routing problem: A real case study," *Sustainable Operations and Computers*, vol. 5, pp. 1–14, 2024.
- [13] J. M. Mulvey, R. J. Vanderbei, and S. A. Zenios, "Robust optimization of large-scale systems," *Operations research*, vol. 43, no. 2, pp. 264–281, 1995.
- [14] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*, vol. 28. Princeton university press, 2009.
- [15] D. Bertsimas and M. Sim, "The price of robustness," *Operations research*, vol. 52, no. 1, pp. 35–53, 2004.
- [16] M. S. Pishvae, M. Rabbani, and S. A. Torabi, "A robust optimization approach to closed-loop supply chain network design under uncertainty," *Applied mathematical modelling*, vol. 35, no. 2, pp. 637–649, 2011.
- [17] A. Rahbari, M. M. Nasiri, F. Werner, M. Musavi, and F. Jolai, "The vehicle routing and scheduling problem with cross-docking for perishable products under uncertainty: Two robust bi-objective models," *Applied Mathematical Modelling*, vol. 70, pp. 605–625, 2019.
- [18] F. Habibzadeh Boukani, B. Farhang Moghaddam, and M. S. Pishvae, "Robust optimization approach to capacitated single and multiple allocation hub location problems," *Computational and Applied Mathematics*, vol. 35, pp. 45–60, 2016.
- [19] M. Varas, S. Maturana, R. Pascual, I. Vargas, and J. Vera, "Scheduling production for a sawmill: A robust optimization approach," *International Journal of Production Economics*, vol. 150, pp. 37–51, 2014.
- [20] F. Guo, Z. Wang, Z. Huang, and X. Ma, "Robust optimization of microhub network and mixed service strategy for a multidepot location-routing problem," *Computers & Industrial Engineering*, vol. 190, p. 110070, 2024.
- [21] T. I. Faiz, C. Vogiatzis, J. Liu, and M. Noor-E-Alam, "A robust optimization framework for two-echelon vehicle and UAV routing for post-disaster humanitarian logistics operations," *Networks*, vol. 84, no. 2, pp. 200–219, 2024.

layout problem,” *Soft Computing*, vol. 28, no. 6, pp. 5667–5682, 2024.

[38] F. Kosanoglu, M. Atmis, and H. H. Turan, “A deep reinforcement learning assisted simulated annealing algorithm for a maintenance planning problem,” *Annals of Operations Research*, vol. 339, no. 1, pp. 79–110, 2024.

[39] Z. Zhang, Z. Shao, W. Shao, J. Chen, and D. Pi, “MRLM: A meta-reinforcement learning-based metaheuristic for hybrid flow-shop scheduling problem with learning and forgetting effects,” *Swarm and Evolutionary Computation*, vol. 85, p. 101479, 2024.

[40] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi, “Optimization by simulated annealing,” *science*, vol. 220, no. 4598, pp. 671–680, 1983.

[41] J.-F. Cordeau, M. Gendreau, and G. Laporte, “A tabu search heuristic for periodic and multi-depot vehicle routing problems,” *Networks: An International Journal*, vol. 30, no. 2, pp. 105–119, 1997.

[42] Borzou, A. and M. Taherkhani, *The effect of internal and external green supply chain management activities on company performance*. Intelligent Multimedia Processing and Communication Systems (IMPCS), 2024. 5(4): p. 59-66.

[43] Saadati, Z. and M. Zarei, *Solving the Multi-Objective Problem of IoT Service Placement in Fog Computing Using Reinforcement Learning Approaches*. Intelligent Multimedia Processing and Communication Systems (IMPCS), 2023. 4(3): p. 29-41.