

بهینه‌سازی تخصیص و زمان‌بندی شاول و کامیون در معادن روباز با استفاده از شبیه‌سازی مبتنی بر الگوریتم‌های یادگیری تقویتی

عرفان مزینانی

کارشناس ارشد، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران

erfan.mazinani@modares.ac.ir

نسیم نهاوندی^۱

استاد، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران، ایران

n_nahavandi@modares.ac.ir

چکیده

با توجه به پیچیدگی ذاتی سیستم‌های حمل‌ونقل در معادن روباز، روش‌های سنتی برنامه‌ریزی در بسیاری از موارد توان پاسخ‌گویی به پویایی و عدم قطعیت موجود در این فرایندها را ندارند. در این پژوهش، با هدف ارتقای بهره‌وری عملیاتی، یک سامانه شبیه‌سازی مبتنی بر الگوریتم‌های یادگیری تقویتی توسعه داده شده است که به بهینه‌سازی هم‌زمان تخصیص و زمان‌بندی شاول و کامیون در معادن روباز می‌پردازد. در این سامانه، یک عامل هوشمند با استفاده از الگوریتم‌های یادگیری تقویتی آموزش می‌بیند تا تصمیم‌هایی کارآمد در زمینه‌ی مسیر حرکت وسایل نقلیه، زمان‌بندی عملیات بارگیری و تخلیه، و تخصیص بهینه‌ی منابع اتخاذ کند. نتایج حاصل از پیاده‌سازی مدل نشان می‌دهد که استفاده از الگوریتم Q-learning موجب یادگیری تخصیص بهینه‌ی کامیون‌ها به شاول‌ها شده و نسبت به دو مدل مقایسه‌ای ساده‌تر عملکرد بهتری از خود نشان می‌دهد. این الگوریتم از طریق به‌روزرسانی تدریجی مقادیر Q و یادگیری بر مبنای تجربیات پیشین، سیاستی بهینه برای تصمیم‌گیری می‌آموزد که در نهایت به افزایش تعداد بارگیری‌های موفق و بهبود قابل توجه بهره‌وری کل سیستم منجر می‌شود.

کلمات کلیدی: شبیه‌سازی، الگوریتم‌های یادگیری تقویتی، تخصیص و زمان‌بندی شاول، تخصیص و زمان‌بندی کامیون، معادن روباز.

^۱ نویسنده مسئول

۱. بیان مسئله

فرآیند معدن کاری به طور کلی شامل سه مرحله اصلی اکتشاف، استخراج و فرآوری مواد معدنی است که هر یک از این مراحل به زیرفرآیندهای متعددی تقسیم می شود. در میان این مراحل، حمل و نقل مواد معدنی به عنوان پیوند میان استخراج و فرآوری، از اهمیت ویژه ای برخوردار است؛ چراکه حدود نیمی از کل هزینه های عملیاتی در معادن روباز به این بخش اختصاص دارد.

با در نظر گرفتن هزینه های بالای تجهیزات معدنی، بهینه سازی سیستم حمل و نقل نقشی اساسی در کاهش هزینه ها و افزایش بهره وری دارد. به طور میانگین، قیمت هر شاول در حدود ۲۰۰ تا ۳۰۰ هزار دلار و بهای هر کامیون با ظرفیت ۱۰۰ تن حدود ۵۰ تا ۶۰ هزار دلار است. این ارقام نشان می دهد که نحوه تخصیص کامیون ها به شاول ها و زمان بندی مناسب عملیات بارگیری و حمل، از مهم ترین و در عین حال پیچیده ترین فرآیندهای عملیاتی در معادن روباز محسوب می شود. بهینه سازی این فرآیندها می تواند سهم قابل توجهی در کاهش هزینه های معدن کاری و بهبود کارایی سیستم حمل و نقل داشته باشد [۱].

در معادن روباز، پس از خرد کردن سنگ ها، مرحله حمل و نقل مواد به عنوان یکی از مهم ترین و پرهزینه ترین بخش ها محسوب می شود. حدود نیمی از هزینه های کل استخراج به این مرحله اختصاص می یابد. به همین دلیل، بهبود عملکرد

عملیات حمل و نقل، تأثیر مستقیمی بر کاهش هزینه های کلی معدن خواهد داشت [۲].

از دیدگاه برنامه ریزی تولید، مسئله تخصیص و زمان بندی تجهیزات حمل و نقل، یکی از مهم ترین و پیچیده ترین مسائل بهینه سازی منابع محسوب می شود. هدف این نوع بهینه سازی، همانند سایر حوزه های صنعتی، کاهش هزینه و زمان عملیات است. در معادن روباز، این هدف به طور خاص شامل کمینه سازی هزینه ی حمل کامیون ها و بهینه سازی زمان بندی شاول ها برای بارگیری کامیون ها است. در چنین محیطی، تصمیم گیری درباره ی نحوه ی حرکت کامیون ها و استفاده ی مؤثر از تجهیزات، فرآیندی چندبعدی و چالش برانگیز است. مدیران معدن ناگزیرند میان اهداف متضادی مانند افزایش نرخ تولید، کاهش هزینه های عملیاتی، کوتاه کردن زمان چرخه ی حمل و نقل و حفظ کیفیت مواد معدنی تعادل برقرار کنند. افزون بر این، محدودیت هایی همچون ظرفیت کامیون ها، توان بارگیری شاول ها و ظرفیت نقاط تخلیه نیز تصمیم گیری را پیچیده تر می سازد.

یادگیری تقویتی یکی از شاخه های مهم و پرکاربرد یادگیری ماشین است که به طور خاص به مسئله یادگیری از طریق تعامل با محیط می پردازد. در این نوع یادگیری، عامل یا همان سیستم یادگیرنده تلاش می کند تا با انجام اعمال مختلف در محیط و دریافت پاداش ها یا تنبیه ها، بهینه ترین راهکار را برای

رسیدن به هدف مشخص پیدا کند. تفاوت اصلی یادگیری تقویتی با سایر انواع یادگیری ماشین در این است که این روش بر اساس تجربه و تعامل مداوم با محیط عمل می‌کند. در یادگیری تقویتی، عامل با محیطی در تعامل است که می‌تواند تغییرات را بر اساس اعمال عامل تجربه کند. این محیط ممکن است بسیار ساده یا پیچیده باشد و می‌تواند شامل عوامل دیگری باشد که هم‌زمان با عامل در حال یادگیری در آن فعالیت می‌کنند. هدف نهایی عامل این است که سیاست یا راهبردی را بیاموزد که با استفاده از آن بتواند بیشترین پاداش ممکن را در بلندمدت به دست آورد [۳].

با توجه به محدودیت‌های ذکر شده، هدف پژوهش حاضر شبیه‌سازی مدل تخصیص و زمان‌بندی شاول/کامیون در معادن روباز و سپس حل آن توسط روش‌های مبتنی بر یادگیری ماشین و الگوریتم یادگیری تقویتی که بتواند بازخورد مناسب را به سامانه بدهد و همچنین پیاده‌سازی مدل شبیه‌سازی شده بر روی یک معدن روباز واقعی است. ساختار این پژوهش در ادامه به صورت زیر است. در بخش دوم مروری بر پژوهش‌های پیشین انجام می‌شود. در بخش سوم شبیه‌سازی ارائه خواهد شد. در بخش چهارم مدل ریاضی به دست آمده از بخش سوم توسط نرم‌افزار پایتون روی یک سیستم مطالعاتی استاندارد پیاده‌سازی و تحلیل حساسیت و مقایسه صورت می‌گیرد و در بخش آخر نتایج پژوهش ارائه و پیشنهاداتی برای تحقیقات آینده ارائه خواهد شد.

۲. مروری بر پیشینه پژوهش

مسئله‌ی زمان‌بندی در عملیات باربری معادن روباز همواره از چالش‌های اصلی در مدیریت بهره‌برداری معادن بوده است. پژوهشگران متعددی در سال‌های اخیر با استفاده از روش‌های شبیه‌سازی و بهینه‌سازی، تلاش کرده‌اند راهکارهایی برای بهبود تخصیص تجهیزات حمل‌ونقل و افزایش بهره‌وری سیستم ارائه دهند. در ادامه، مروری بر برخی از مطالعات انجام‌شده در این حوزه ارائه می‌شود. در پایان این بخش نیز، با جمع‌بندی نتایج پژوهش‌های پیشین، جایگاه نوآوری و ضرورت انجام تحقیق حاضر به صورت شفاف تبیین خواهد شد.

اسناتفیت‌تری و همکاران [۴]، مدلی بهینه‌سازی برای تعیین نحوه‌ی تخصیص کامیون‌ها در معادن روباز توسعه دادند. هدف این مدل، حداقل‌سازی هزینه‌ی کل سیستم شامل دو بخش اصلی بود:

۱- هزینه‌های سرمایه‌گذاری اولیه، که به قیمت خرید کامیون‌ها در پروژه‌ی استخراج اولیه با توجه به تعداد کامیون‌های مورد نیاز در هر مسیر مربوط می‌شود.

۲- هزینه‌های عملیاتی حمل‌ونقل، شامل هزینه‌ی ثابت نیروی انسانی (اپراتورها) و هزینه‌های متغیر ناشی از مصرف سوخت کامیون‌ها و بیل‌های مکانیکی.

نتایج این مطالعه نشان داد که انتخاب نوع و ترکیب بهینه‌ی کامیون‌ها با مشخصات فنی متفاوت، نقش قابل‌توجهی در حفظ پایداری عملیات حمل‌ونقل مواد و کاهش هزینه‌های کل معدن دارد. همچنین، در نظر گرفتن تفاوت‌های عملکردی کامیون‌ها (از نظر ظرفیت، مصرف سوخت و دوام) می‌تواند موجب بهبود کارایی اقتصادی سیستم حمل‌ونقل شود.

یوپاده‌ی و عسگری‌نسب [۵]، یک ابزار بهینه‌سازی به عنوان یک تصمیم‌گیرنده تخصیص پویای شاول برای مدل‌های شبیه‌سازی عملکرد معدن ارائه داده‌اند. اهداف ابزار بهینه‌سازی فراهم آوردن تخصیص شاول به سینه کارهای معدن به منظور به حداکثر رساندن تولید، برآورده شدن درجه مطلوب و تناژ در سنگ شکن‌ها و به حداقل رساندن حرکات شاول است. در این مقاله توسعه و اجرای ابزار بهینه‌سازی با مطالعه موردی معدن سنگ آهن ارائه شده است. رویکردی برای تولید برنامه‌های کوتاه مدت از طریق عملیات تولید شبیه‌سازی شده و بهینه‌سازی تصمیم‌گیری در مورد تخصیص شاول آن را عملی‌تر کرده و با تجزیه و تحلیل سناریو، فرصت‌هایی را برای برنامه‌ریزی پیشگیرانه فراهم می‌کند.

هدف پژوهش ماکاروا و همکاران [۶]، کاهش زمان بیکاری کامیون‌ها و بیل‌های معدن در یک معدن زغال سنگ روباز است. یک الگوریتم اکتشافی برای تصمیم‌گیری‌های اعزام در شرایط تخصیص پویای کامیون‌ها توسعه داده شده است. پارامترهای اولویت برای انتخاب بیل پس از تخلیه پایان

کامیون معرفی شده است. همچنین الگوریتمی برای جستجوی پارامترهای اولویت بهینه برای برآورده کردن معیار کارایی مورد نیاز توسعه داده شده است. این الگوریتم مبتنی بر یک مدل شبیه‌سازی سیستم بیل-کامیون است. رویکرد پیشنهادی از نظر گروهی از بیل‌ها با نقطه تخلیه مشترک در معادن زغال سنگ روباز مختلف قابل اجرا است. اهمیت این کار در این واقعیت نهفته است که مدل پیشنهادی عوامل تصادفی مرتبط با مدت زمان بارگیری و تخلیه، حرکت کامیون، تعمیر بیل‌ها و کامیون‌های حمل بار و همچنین مدت زمان بین تعمیرات را در نظر می‌گیرد.

فایز فادین [۷]، در پژوهش خود، با بهره‌گیری از یک رویکرد نوآورانه مبتنی بر الگوریتم‌های فراابتکاری و مدل شبیه‌سازی-بهینه‌سازی، به مسئله بهینه‌سازی تخصیص کامیون در معادن روباز پرداخته است. با استفاده از داده‌های واقعی، این پژوهش به دنبال یافتن پاسخ به این سوال بوده که چگونه می‌توان با تخصیص بهینه کامیون‌ها، به حداکثر تولید و حداقل هزینه در صنعت معدنکاری دست یافت. نتایج این پژوهش نشان می‌دهد که این روش می‌تواند به عنوان یک ابزار قدرتمند برای بهبود بهره‌وری در عملیات معدنکاری مورد استفاده قرار گیرد.

بخت‌آور و محمودی [۸]، یک رویکرد بهینه‌سازی قوی مبتنی بر سناریو را برای حل مشکل تخصیص کامیون-بیل توسعه دادند. برای این منظور، مشکل تخصیص کامیون-بیل در دو فاز با استفاده از مفاهیم رویکرد بهینه‌سازی قوی، تجزیه و تحلیل شبکه و مسئله کوتاه‌ترین مسیر و برنامه‌نویسی عدد صحیح باینری تحت عدم قطعیت فرموله شد. عدم قطعیت در خروجی بیل و ظرفیت سنگ شکن از فاز اول و تعداد کامیون‌های موجود از فاز دوم بر اساس رویکرد بهینه‌سازی قوی در نظر گرفته شد. این رویکرد تخصیص کامیون-بیل در تمام معادن روباز که در آن کامیون‌هایی با ظرفیت‌های مختلف استفاده می‌شود و مسیرهای متفاوتی بین نقاط بارگیری و تخلیه وجود دارد، قابل اجرا است.

هاشمی و ستاروند [۹]، با استفاده از یک روش شبیه‌سازی، سیستم‌های مدیریتی مختلف تجهیزات معدن روباز را در حالات مختلف از جمله راه‌حل‌های بدون اعزام، اعزام و ترکیبی از دو حالت، برای معدن مس سونگون مدل‌سازی کردند. یک

مدل شبیه‌سازی اعزام با تابع هدف کمینه‌سازی زمان انتظار کامیون‌ها، که از قبل وجود داشته را توسعه دادند و مدل قبلی را که با تخصیص ثابت شبیه‌سازی شده بود را به صورت تخصیص پویا شبیه‌سازی کردند و در نهایت واحد کنترل مخلوط درجه سنگ معدن را در مدل وارد کردند. مدل توسعه یافته توانایی در نظر گرفتن ویژگی‌های دقیق تجهیزات بارگیری و حمل بار را دارد. نتایج حاصل از این شبیه‌سازی نشان می‌دهد که زمان انتظار کامیون‌ها به صورت قابل توجهی کاهش یافته است.

محتشم و همکاران [۱۰]، در پژوهش خود یک مدل برنامه‌ریزی هدف محدود تصادفی را بر اساس چهار هدف مهم برای تخمین اثرات عدم قطعیت بر کارایی سیستم‌های بیل کامیون ارائه کردند. مدل پیشنهادی با استفاده از ۱۱ سناریو برنامه‌ریزی و سطوح مختلف اطمینان برای تولید لودر برای تعیین بهترین تخصیص کامیون‌ها در یک معدن مس روباز پیاده‌سازی شده است. نتایج نشان داد که مدل می‌تواند کیفیت و کمیت مواد مورد نیاز برای دستیابی به اهداف برنامه تولید کوتاه‌مدت معدن را در تمام سطوح مختلف اطمینان، حتی در بالاترین سطح ریسک، کنترل کند.

پژوهش یائو و همکاران [۱۱]، یک سیستم نوآورانه پیشنهاد می‌کند که تصمیمات ترکیب سنگ معدن را به صورت پویا و بر اساس تخصیص کامیون در معادن روباز بهینه‌سازی می‌نماید. برخلاف روش‌های سنتی، این سیستم علاوه بر در نظر گرفتن مسیرهای بهینه برای کامیون‌ها، اثرات ترکیب سنگ‌های مختلف بر کیفیت نهایی و سودآوری معدن را نیز پیش‌بینی می‌کند. رویکرد پیشنهادی بر اهمیت تنظیم پویا بر اساس تغییرات کیفیت سنگ معدن، در دسترس بودن کامیون‌ها، و محدودیت‌های مربوط به ترکیب‌ها تأکید دارد. نتایج حاکی از آن است که ترکیب بهینه پویا می‌تواند نه تنها بهره‌وری کل معدن را افزایش داده، بلکه هزینه‌ها را نیز کاهش دهد. این سیستم، رویکردی جامع برای مدیریت کامیون‌ها و بهبود کارایی عملیات معدنی ارائه می‌دهد.

پژوهش پاتهان و همکاران [۱۲]، مدلی را معرفی می‌کند که پارامترهای ژئوتکنیکی و عملکرد تجهیزات حفاری را برای تخصیص بهینه کامیون‌ها در معادن ادغام می‌کند. با استفاده از شاخص کل زمان بارگیری (TLT) و زمان سفر کامیون‌ها،

مقدار بهینه کامیون برای هر شاول محاسبه شده است. مطالعه موردی در یک معدن روباز زغال سنگ در پاکستان، بهبود قابل توجهی در کارایی عملیاتی و کاهش هزینه‌ها نشان داده است. این رویکرد با استفاده از داده‌های بلادرنگ و شبیه‌سازی پیشرفته، بهره‌وری عملیاتی و پایداری اقتصادی معادن را افزایش می‌دهد.

پژوهش ژانگ و همکاران [۱۳]، یک مدل برنامه‌ریزی عدد صحیح مختلط برای بهینه‌سازی هم‌زمان مسیرها و سرعت کامیون‌های خودران به منظور کاهش مصرف انرژی ارائه می‌دهد. الگوریتم پیشنهادی شامل دو بخش است: یک مدل تخصیص جریان بهینه و یک الگوریتم جستجوی ممنوعه برای تخصیص کامیون‌ها. همچنین یک سیستم زمان‌بندی بلادرنگ برای محیط‌های پویا و تصادفی معدنکاری توسعه داده شده است. آزمایش‌ها در یک معدن زغال سنگ در چین نشان داد که این مدل بهره‌وری را افزایش داده و نیازهای محاسباتی زمان‌بندی بلادرنگ را تأمین می‌کند.

یعقوبی و همکاران [۱۴]، یکی از چالش‌های اساسی در سیستم‌های حمل‌ونقل صنعتی، ماهیت پویا و غیرقطعی فرآیندهای عملیاتی است. این پژوهش با ارائه مدل‌های توسعه‌یافته از مسئله‌ی VRP و معرفی رویکرد GVRP نشان می‌دهد که در محیط‌های واقعی، عوامل مختلفی نظیر تغییرات زمان سفر، محدودیت‌های سوخت، و شرایط عملیاتی متغیر، باعث پیچیده‌تر شدن فرآیند تصمیم‌گیری می‌شوند. آن‌ها تأکید می‌کنند که در چنین شرایطی، تصمیم‌گیری‌های بلادرنگ و مبتنی بر داده می‌تواند به شکل محسوسی بهره‌وری سیستم حمل‌ونقل را افزایش دهد. اهمیت این موضوع در مسئله تخصیص کامیون-شاول نیز مشاهده می‌شود، زیرا خرابی ماشین‌آلات، تأخیرهای عملیاتی و تغییرات لحظه‌ای در ظرفیت تجهیزات، نیازمند به‌کارگیری رویکردهایی است که توان سازگاری با عدم قطعیت را داشته باشند.

در پژوهش آقادی [۱۵]، نشان داده شده است که در سیستم‌های عملیاتی پیچیده، ترکیب محدودیت‌های زمانی، ترتیب انجام وظایف و ظرفیت تجهیزات، نقش تعیین‌کننده‌ای در عملکرد کلی سیستم دارد. این پژوهش با ارائه مدل‌های

دقیق زمان‌بندی نشان می‌دهد که کوچک‌ترین تأخیر یا اختلال می‌تواند اثر زنجیره‌ای بر کل فرآیند ایجاد کند و بهره‌وری را کاهش دهد. در چنین شرایطی، وجود الگوریتم‌های هوشمند برای تعیین برنامه کارآمد و سازگار با شرایط واقعی ضروری است. این بینش‌ها برای مسئله تخصیص کامیون-شاول ارزش ویژه‌ای دارد، زیرا خرابی ماشین‌آلات، تغییرات در نرخ استخراج، و نوسان ظرفیت حمل، نیازمند مدل‌هایی هستند که بتوانند با انعطاف‌پذیری بالا، زمان‌بندی و سیاست تخصیص بهینه را ارائه دهند.

با مرور مطالعات پیشین در حوزه‌ی تخصیص و زمان‌بندی تجهیزات حمل‌ونقل در معادن روباز، می‌توان دریافت که بیشتر پژوهش‌ها بر بهینه‌سازی مبتنی بر شبیه‌سازی یا الگوریتم‌های فراابتکاری متمرکز بوده‌اند. با این حال، در اغلب این مطالعات، شرایط واقعی عملیات معدنکاری از جمله احتمال خرابی تجهیزات، نوسانات عملکردی کامیون‌ها و شاول‌ها، و محدودیت‌های عملیاتی محیط معدن به‌صورت کامل لحاظ نشده است.

در پژوهش حاضر، تفاوت اصلی رویکرد پیشنهادی با تحقیقات پیشین در در نظر گرفتن عوامل واقعی و پویا مانند خرابی کامیون‌ها و اثر آن‌ها بر بهره‌وری کل سیستم نهفته است. در این روش، با بهره‌گیری از نرم‌افزارهای آماری، نرخ خرابی و زمان تعمیر کامیون‌ها بر اساس داده‌های واقعی استخراج و به مدل شبیه‌سازی وارد شده است. سپس با استفاده از الگوریتم یادگیری تقویتی، مدل قادر است احتمالات خرابی و تأثیر آن بر زمان‌بندی و تخصیص تجهیزات را پیش‌بینی و بهینه‌سازی کند.

افزون بر این، استفاده از رویکرد پیشنهادی موجب کاهش چشمگیر زمان حل مسئله نسبت به روش‌های سنتی می‌شود، در حالی که دقت و واقع‌گرایی مدل نیز به‌طور محسوسی افزایش می‌یابد. این ویژگی‌ها، مدل حاضر را به ابزاری کارآمد

برای مدیریت هوشمند عملیات حمل و نقل در معادن روباز تبدیل می‌کند.

۳. مدل‌سازی

مدل معدن به عنوان یک محیط شبیه‌سازی توسط کتابخانه SimPy پیاده‌سازی شده است. در این مدل، تعدادی شاول و کامیون با نرخ‌های بارگیری و حمل و نقل مشخصی وجود دارند. هر شاول با یک نرخ مشخص مواد را از معدن بارگیری می‌کند و کامیون‌ها نیز با نرخ‌های مشخص به معدن می‌آیند تا مواد را حمل کنند. در این فرایند، احتمال خرابی کامیون‌ها تعریف شده است که می‌تواند باعث ایجاد وقفه در عملیات آن‌ها شود.

در کلاس MineEnv، محیط شبیه‌سازی معدن با تعیین تعداد شاول‌ها و کامیون‌ها، نرخ‌های بارگیری و حمل و نقل، و احتمال خرابی، تنظیم می‌شود. کامیون‌ها در یک صف قرار می‌گیرند تا برای بارگیری آماده باشند و به نوبت توسط شاول‌ها بارگیری می‌شوند. هر شاول در این محیط به صورت یک فرایند عمل می‌کند که مواد را بارگیری کرده و سپس کامیون‌ها را به حمل مواد تخصیص می‌دهد.

الگوریتم Q-Learning برای تصمیم‌گیری در تخصیص کامیون‌ها به شاول‌ها استفاده می‌شود. در این الگوریتم، جدولی به نام Q-table برای ذخیره ارزش‌های Q مربوط به هر حالت و اقدام نگهداری می‌شود. در هر گام زمانی، الگوریتم با توجه به مقدار ϵ (احتمال انتخاب تصادفی)، یا یک کامیون را به صورت تصادفی برای بارگیری انتخاب می‌کند یا بر اساس ارزش‌های Q-table، کامیونی را انتخاب می‌کند که بیشترین پاداش مورد انتظار را دارد. این فرایند با یادگیری پیوسته و به‌روزرسانی ارزش‌های Q در طی اپیزودهای مختلف، به الگوریتم اجازه می‌دهد تا سیاست بهینه‌ای برای تخصیص کامیون‌ها به شاول‌ها بیابد.

مدل ارائه شده ترکیبی از شبیه‌سازی و یادگیری تقویتی را برای مدیریت بهینه تخصیص شاول‌ها و کامیون‌ها در یک محیط معدنی نشان می‌دهد. این کد با استفاده از کتابخانه‌های مختلف پایتون مانند simpy برای شبیه‌سازی رویدادهای گسسته و numpy برای انجام محاسبات عددی کار می‌کند. هدف اصلی این کد بهینه‌سازی تخصیص کامیون‌ها به شاول

ها برای حداکثر کردن بهره‌وری در حمل و نقل مواد معدنی است.

محیط شبیه‌سازی با استفاده از کلاس MineEnv تعریف شده است. در این کلاس، صفی برای کامیون‌ها وجود دارد که هر کامیون پس از بارگیری به آن بازمی‌گردد. هر شاول به طور مداوم تلاش می‌کند کامیون‌ها را بارگیری کند و اگر خراب نشود، این فرایند ادامه پیدا می‌کند. برای بارگیری هر کامیون، زمانی بر اساس نرخ بارگیری شاول و نرخ مسافرت کامیون محاسبه می‌شود. یکی از بخش‌های مهم این کد استفاده از یادگیری تقویتی برای بهبود تخصیص کامیون‌ها به شاول‌ها است. این کار با استفاده از الگوریتم Q-learning انجام می‌شود. در این الگوریتم، جدولی به نام Q-table وجود دارد که مقادیر Q را برای هر جفت شاول / کامیون ذخیره می‌کند. مقادیر Q نشان‌دهنده ارزش یا سودمندی هر جفت شاول / کامیون در طول زمان است و با استفاده از پاداش‌هایی که در طول شبیه‌سازی دریافت می‌شود، به‌روزرسانی می‌شود.

برای هر اپیزود از شبیه‌سازی، یک محیط جدید ایجاد می‌شود و شبیه‌سازی تا زمان مشخصی اجرا می‌شود. در پایان هر اپیزود، تعداد تخصیص‌های موفق شاول‌ها به کامیون‌ها محاسبه و به عنوان پاداش در نظر گرفته می‌شود. این پاداش‌ها برای به‌روزرسانی Q-table استفاده می‌شوند. الگوریتم Q-learning از پارامترهای آلفا (نرخ یادگیری)، گاما (عامل تخفیف) و اپسیلون (نرخ اکتشاف) برای کنترل فرایند یادگیری استفاده می‌کند. پس از آموزش عامل با استفاده از Q-learning، عامل قادر خواهد بود بهینه‌ترین تخصیص‌ها را برای شاول‌ها و کامیون‌ها تعیین کند. این تخصیص‌ها بر اساس مقادیر Q موجود در Q-table انجام می‌شود و عامل می‌تواند بهره‌وری سیستم را به حداکثر برساند.

به‌منظور اطمینان از پایداری و کارایی فرایند یادگیری، سه پارامتر کلیدی الگوریتم یعنی نرخ یادگیری (α)، ضریب تخفیف (γ) و نرخ اکتشاف (ϵ) به‌صورت دقیق تنظیم و آزمون شدند:

نرخ یادگیری (α)

این پارامتر نشان‌دهنده میزان تأثیر تجربیات جدید بر

به روزرسانی مقادیر Q است. مقدار بیش از حد بالا می تواند باعث نوسان در یادگیری شود و مقدار خیلی پایین منجر به همگرایی کند گردد.

در این پژوهش، مقدار اولیه ی $\alpha = 0.1$ در نظر گرفته شد و پس از انجام آزمون های حساسیت در بازه ی $[0.05, 0.3]$ ، مقدار 0.15 به عنوان گزینه ی نهایی انتخاب گردید؛ زیرا این مقدار بهترین توازن را بین سرعت همگرایی و پایداری تصمیم گیری عامل فراهم می کرد.

ضریب تخفیف (γ)

این پارامتر اهمیت پاداش های آینده را نسبت به پاداش های فعلی تعیین می کند. مقدار بالاتر باعث تمرکز عامل بر پاداش های بلندمدت می شود. در این مطالعه، برای تأکید بر بهینه سازی بلندمدت عملکرد سیستم، مقدار $\gamma = 0.9$ انتخاب شد. این مقدار پس از آزمون مقادیر $[0.7, 0.95]$ بیشترین پایداری و نرخ موفقیت را در تخصیص ها نشان داد.

نرخ اکتشاف (ϵ)

برای جلوگیری از گیر افتادن عامل در بهینه های محلی، از سیاست ϵ -greedy استفاده شد که در آن عامل با احتمال ϵ به صورت تصادفی اقدام کرده و با احتمال $1 - \epsilon$ اقدام با بیشترین مقدار Q را انتخاب می کند. مقدار اولیه $\epsilon = 0.2$ در نظر گرفته شد و طی فرایند یادگیری به صورت خطی تا مقدار 0.05 کاهش یافت تا با پیشرفت آموزش، عامل بیشتر بر استراتژی های بهینه تمرکز کند.

کد همچنین شامل چندین مرحله برای نمایش و تحلیل نتایج است. این مراحل شامل نمایش Q -table به صورت نمودار حرارتی، ترسیم تخصیص کامیون ها به شاول ها با استفاده از نمودار پراکنده و نمایش نمودارهای مختلف برای ارزیابی عملکرد سیستم در طول اپیزودها است. این نمودارها شامل تعداد تخصیص های موفق در هر اپیزود، پاداش های تجمعی در هر اپیزود و میانگین پاداش ها در طول زمان هستند. مقادیر تصادفی نمایی در این مدل برای شبیه سازی زمان های بین وقایع استفاده می شوند.

در کد ارائه شده، مقادیر تصادفی نمایی برای شبیه سازی زمان بارگیری کامیون ها توسط شاول ها، زمان سفر کامیون ها و کامیون ها پس از خرابی استفاده می شود. برای تولید این مقادیر از تابع `np.random.exponential` استفاده می شود. این تابع از کتابخانه `numpy` در پایتون برای تولید اعداد تصادفی از توزیع نمایی استفاده می کند. پارامتر این تابع نرخ وقوع رویداد است، که معکوس میانگین زمان بین وقوع دو رویداد متوالی است. در این مدل، نرخ بارگیری شاول ها، نرخ سفر کامیون ها و نرخ تعمیرات کامیون ها به عنوان ورودی به تابع `np.random.exponential` داده می شوند تا زمان های تصادفی متناظر تولید شوند. در اینجا زمان های مربوط به شاول با S و زمان های مربوط به کامیون با T نمایش داده شده است. داریم:

$$P(T \leq t) = 1 - e^{(-\lambda t)}$$

عبارت بالا توزیع تجمعی T را نمایش می دهد.

$$f(t) = \lambda e^{(-\lambda t)}$$

این توزیع نیز تابع چگالی احتمال T را نمایش می دهد.

$$E(T) = \frac{1}{\lambda}$$

$$Var(T) = 1 / \lambda^2$$

$$cov(x, y) = E[xy] - E[x]E[y]$$

عبارت زیر نیز کوواریانس بین دو متغیر را نمایش می دهد که برای نمایش همبستگی نیز استفاده می شود.

$$\rho(x, y) = \frac{cov(x, y)}{\sqrt{var(x) var(y)}}$$

$$P(X = K) = \frac{e^{-\lambda} \lambda^K}{K!}$$

$$S(t) = e^{-\lambda t}$$

$$H(t) = \frac{f(t)}{S(t)}$$

$$\pi(a|s) = \frac{e^{Q(s,a)/T}}{\sum_b e^{Q(s,b)/T}}$$

در اینجا دلتا برای نشان دادن اختلاف بین تابع هزینه این حالت و حالت بعدی استفاده می‌شود:

$$\delta = r + \gamma V(S') - V(S)$$

تابع زیر میزان سود انباشته شده را نمایش می‌دهد:

$$G_t = \sum_{K=0}^{\infty} \gamma^K r_{t+K+1}$$

تابع زیر نیز سود حاصله از لحظه t به بعد را نمایش می‌دهد:

$$R_t = r_t + \gamma r_{(t+1)} + \gamma^2 r_{(t+2)} + \dots$$

معادله سود انباشته شده حاشیه ای:

$$\Delta W = \alpha \delta \nabla_w V(S)$$

معادله نشان دهنده بهترین سیاست:

$$\pi'(s) = \operatorname{argmax}_a Q(s, a)$$

تمامی احتمالات باید ۱ شود:

$$\sum_a \pi(a|s) = 1$$

معادله مجموع سناریوی قبل نسبت به بعدی ها باید ۱ شود:

$$\sum_{s'} P(s'|s, a) = 1$$

معادله تابع بی نظمی (آنتروپی):

$$H(\pi) = -\sum_a \pi(a|s) \log \pi(a|s)$$

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t)$$

$$\delta_t = r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)$$

بهره وری سیاست گذاری:

$$\rho_t = \frac{\pi(a_t|s_t)}{\mu(a_t|s_t)}$$

$$E[e^{j\omega T}] = \int_0^{\infty} e^{j\omega T} \lambda e^{-\lambda t} dt$$

$$\frac{d}{dt} F(t) = f(t)$$

$$M_X(t) = E[e^{tX}]$$

هدف اصلی این شبیه سازی و یادگیری تقویتی، بهینه سازی فرآیند تخصیص منابع در محیطی پویا و پیچیده است. با استفاده از این روش، می توان کارایی و بهره وری سیستم های معدنی را افزایش داد و از منابع موجود به بهترین شکل ممکن استفاده کرد. این کد نشان می دهد که چگونه می توان از تکنیک های یادگیری تقویتی برای حل مسائل پیچیده و بهبود عملکرد سیستم ها استفاده کرد.

مدل یادگیری تقویتی در این کد با استفاده از تجربیات گذشته، سیاست های بهینه ای را برای تخصیص کامیون ها به شاول ها یاد می گیرد. این یادگیری به طور تدریجی انجام می شود و عامل با تکرار اپیزودهای شبیه سازی و به روزرسانی مقادیر Q ، سیاست های خود را بهبود می بخشد. این فرآیند یادگیری به عامل اجازه می دهد تا بهترین تخصیص ها را پیدا کند و بهره وری سیستم را به حداکثر برساند. همچنین مدل بازگشتی یادگیری تقویتی بصورت زیر است که در آن نرخ یادگیری و γ ضریب تخفیف است. همچنین Q و V توابع هزینه و R جایزه است. در اینجا π سیاست مدنظر را نمایش می دهد:

در اینجا تابع Q برای نمایش هزینه مدل تحت حالت و اقدام استفاده می شود:

$$Q(s, a) = Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$$

تابع V نیز مقدار بیشینه Q را برای اقدام های مختلف نمایش می دهد:

$$V(s) = \max_a Q(s, a)$$

همچنین احتمال هر اقدام از رابطه زیر به دست می آید:

میانگین تابع هزینه نیز به صورت زیر است:

$$Q(s, a) = \frac{1}{N(s, a)} \sum_{i=1}^{N(s, a)} G_i$$

$$\delta_t = r_t + \gamma \sum_{s'} P(s' | s, a) V(s') - V(s)$$

$$\delta = r + \gamma Q(s', a') - Q(s, a)$$

$$\delta_t = r_t + \gamma \sum_{s'} P(s' | s, a) V(s')$$

$$\pi_{\theta}(a|s) = \frac{e^{e^T(s, s, a)}}{\sum_b e^{e^T(s, s, b)}}$$

$$\delta_t = r_t + \gamma V(s_{t+1}) - V(s_t)$$

همچنین، استفاده از مقادیر تصادفی نمایی در شبیه‌سازی، به مدل امکان می‌دهد تا تغییرات و عدم قطعیت‌های موجود در محیط‌های واقعی را به خوبی بازتاب دهد. این ویژگی باعث می‌شود که سیاست‌های یادگرفته‌شده توسط عامل در مواجهه با شرایط واقعی نیز کارآمد باشند و بتوانند عملکرد سیستم را بهبود بخشند.

۴. حل مدل

در این بخش شبیه‌سازی ارائه شده در بخش مدل‌سازی بر روی داده‌های یک معدن واقعی روباز اجرا شده است و سپس با سه رویکرد مختلف مسئله حل شده است و در هر کدام پارامترها گزارش و در نهایت مقایسه شده است. برای این منظور ابتدا اطلاعات سیستم مورد مطالعه، ارائه شده است. در ادامه، نتایج عددی ارائه شده است که برای این منظور، صحت-سنجی و آنالیز حساسیت بر روی نتایج نیز انجام شده است. در انتها بحث و نتیجه‌گیری بر روی نتایج انجام شده است.

۱.۴. داده‌های مدل

در معدن روباز مورد مطالعه، شش شاول و سی و هفت کامیون وجود دارد که هر کدام نرخ‌های بارگیری و مسافرت مشخصی دارند. این نرخ‌ها به عنوان پارامترهای ورودی به شبیه‌سازی داده می‌شوند و برای هر شاول و کامیون متفاوت هستند. همچنین، احتمال خرابی برای کامیون وجود دارد که در

شبیه‌سازی در نظر گرفته شده است. وقتی یک کامیون خراب می‌شود، نیاز به زمان برای تعمیر دارد که با نرخ مشخصی انجام می‌شود، برای شبیه‌سازی نرخ خرابی کامیون و همچنین نرخ تعمیر آن‌ها نیز اطلاعات سیستم حمل و نقل یک معدن در اختیار قرار گرفت و با وارد کردن زمان خرابی و مدت زمانی که هر کامیون در تعمیرگاه می‌ماند.

برای شبیه‌سازی نرخ خرابی و تعمیر کامیون‌ها ابتدا از شرکت معدنی مربوطه داده‌های مربوط به خرابی و توقف هر کامیون را جمع‌آوری کردیم، سپس با کمک نرم‌افزار آماری minitab توزیع مناسب برای نرخ خرابی و زمانی که هر کامیون در تعمیرگاه بود را برازش کردیم و در مدل شبیه‌سازی شده در کد پایتون نیز وارد کردیم.

پس از جمع‌آوری داده‌ها، زمان بین خرابی‌ها برای هر کامیون محاسبه گردید. میانگین زمان بین هر توقف ناشی از خرابی حدود ۲۶ ساعت و همچنین میانگین زمانی که هر کامیون در تعمیرگاه می‌ماند ۱۰ ساعت در نظر گرفته شده است.

در جدول ۱ میانگین زمانی که شاول‌ها کامیون را پر می‌کنند بر حسب دقیقه ثبت شده است.

جدول ۱. میانگین زمانی بارگیری کامیون‌ها

Average truck loading time

Time(min)	shovel
۳,۳	Shovel ۱
۳	Shovel ۲
۲,۶	Shovel ۳
۲,۴	Shovel ۴

		۳,۲	Truck ۱۹
--	--	-----	----------

افق شبیه‌سازی ۳۰۰ واحد زمانی در نظر گرفته شده و برای هر ۵ واحد زمانی یک اپیزود در نظر گرفته شد خروجی مدل اعم از نمودار و تخصیص بهینه استخراج می‌شود.

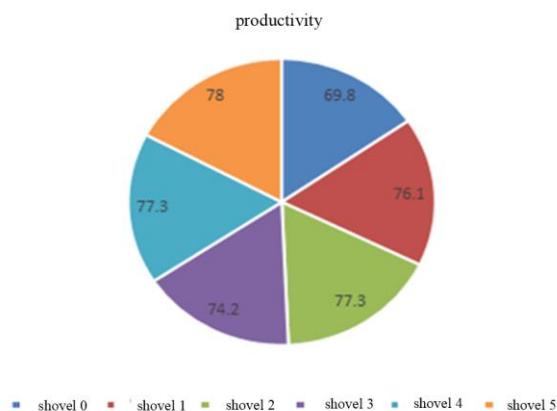
۲.۴. اجرای شبیه‌سازی با در نظر گرفتن سه روش

مختلف تخصیص کامیون به شاول

مدل برای سه روش مختلف تخصیص کامیون به شاول شبیه‌سازی شده است که در زیر توضیح داده می‌شود.

۱.۲.۴. تخصیص به اولین شاول موجود

در این مدل، تخصیص کامیون‌ها به شاول‌ها به صورت ساده و مستقیم انجام می‌شود. هر بار که یک کامیون برای بارگیری آماده می‌شود، به اولین شاول موجود تخصیص داده می‌شود. فرآیند تخصیص در این مدل به این صورت است که ابتدا یک کامیون از صف کامیون‌ها انتخاب می‌شود و سپس به اولین شاول که آماده بارگیری است تخصیص داده می‌شود. اگر کامیون خراب شود، شاول منتظر می‌ماند تا کامیون بعدی اعزام شود و سپس باربری انجام می‌شود. پس از باربری، کامیون برای سفر به محل تخلیه می‌رود و سپس به صف بازمی‌گردد. بهره‌وری شاول‌ها در روش تشکیل صف در نمودار ۱ رسم شده است.



نمودار ۱. بهره‌وری شاول‌ها در روش تشکیل صف

۲,۲	Shovel ۵
۲,۱	Shovel ۶

در جدول ۲ میانگین تعداد سفرهایی که هر کامیون در یک ساعت از پای شاول سفر خود را آغاز می‌کند و دوباره به نقطه ابتدایی برمی‌گردد ثبت شده است.

جدول ۲. میانگین تعداد سفرهای کامیون‌ها

Average number of truck trips

Trip num	truck	Trip num	truck
۲,۱	Truck ۲۰	۲,۱	Truck ۱
۲,۳	Truck ۲۱	۲,۳	Truck ۲
۲,۲	Truck ۲۲	۲,۲	Truck ۳
۲,۵	Truck ۲۳	۲,۵	Truck ۴
۲,۷	Truck ۲۴	۲,۷	Truck ۵
۲,۸	Truck ۲۵	۲,۸	Truck ۶
۲,۹	Truck ۲۶	۲,۹	Truck ۷
۳	Truck ۲۷	۳	Truck ۸
۳,۱	Truck ۲۸	۳,۱	Truck ۹
۳,۲	Truck ۲۹	۳,۲	Truck ۱۰
۲,۱	Truck ۳۰	۲,۲	Truck ۱۱
۲,۳	Truck ۳۱	۲,۳	Truck ۱۲
۲,۲	Truck ۳۲	۲,۵	Truck ۱۳
۲,۵	Truck ۳۳	۲,۷	Truck ۱۴
۲,۷	Truck ۳۴	۲,۸	Truck ۱۵
۲,۸	Truck ۳۵	۲,۹	Truck ۱۶
۲,۹	Truck ۳۶	۳	Truck ۱۷
۳	Truck ۳۷	۳,۱	Truck ۱۸

۳.۲.۴. استفاده از یادگیری تقویتی برای تخصیص کامیون به شاول

در پایان فرایند یادگیری، الگوریتم Q-learning قادر به یادگیری سیاستی است که تخصیص بهینه کامیون‌ها به شاول‌ها را تعیین می‌کند. این تخصیص بهینه به معنای یافتن ترکیب‌هایی است که منجر به بیشترین تعداد بارگیری‌ها و حمل‌ونقل‌های موفق می‌شود. در اینجا، شاول‌ها با نرخ‌های مختلف بارگیری و کامیون‌ها با نرخ‌های مختلف سفر دارند و همچنین احتمال خرابی برای کامیون وجود دارد. شبیه‌سازی با در نظر گرفتن همه این عوامل، تخصیص بهینه را یاد می‌گیرد.

یکی از خروجی‌های اصلی این کد، Q-table یا جدول Q است که یک ماتریس دو بعدی است که مقادیر Q برای هر جفت شاول-کامیون را ذخیره می‌کند. هر مقدار Q نشان دهنده ارزش اقدام تخصیص یک کامیون به یک شاول خاص در حالت فعلی است. مقادیر Q از طریق فرایند یادگیری تقویتی به‌روزرسانی می‌شوند و پس از تعداد زیادی تکرار، این مقادیر نمایانگر سیاست بهینه تخصیص می‌شوند.

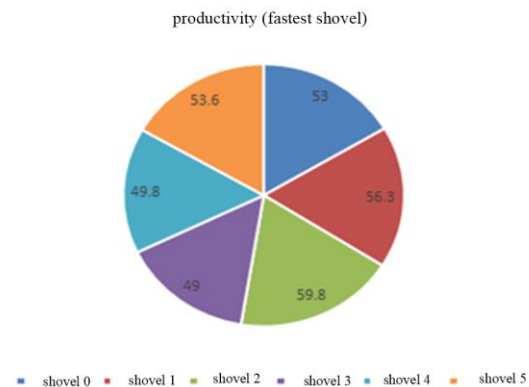
چنانچه در نمودار ۳ نشان داده شده است، نمودار Q-table به صورت یک نقشه حرارتی نمایش داده می‌شود که در آن هر خانه از ماتریس با رنگی نمایش داده می‌شود که نشان‌دهنده مقدار Q برای آن جفت شاول-کامیون است. رنگ‌های گرم‌تر (مثل قرمز) نشان‌دهنده مقادیر Q بالاتر و رنگ‌های سردتر (مثل آبی) نشان‌دهنده مقادیر Q پایین‌تر هستند. این نقشه حرارتی به ما کمک می‌کند تا به سرعت جفت‌های شاول-کامیون با ارزش‌ترین تخصیص‌ها را شناسایی کنیم.

Figure ۱. Productivity of shovels in the queue formation method

تعداد تخصیص‌های موفق در این روش در حدود ۷۰۰ عدد می‌باشد که در افق زمانی ۳۰۰ واحد زمانی رسم شده است، همچنین میانگین تعداد سفرهای کامیون‌ها در این روش در یک ساعت به طور میانگین برابر ۲.۲ سفر می‌باشد.

۲.۲.۴. تخصیص به سریع‌ترین شاول

سناریوی دوم که شبیه‌سازی تخصیص به سریع‌ترین شاول است، تلاش می‌کند تا عملکرد را با تخصیص کامیون‌ها به شاول‌هایی که بالاترین نرخ بارگیری را دارند بهبود ببخشد. در این مدل، هر بار که یک کامیون برای بارگیری آماده می‌شود، به سریع‌ترین شاول موجود تخصیص داده می‌شود. این روش تلاش می‌کند تا با استفاده بهتر از شاول‌هایی که نرخ بارگیری بالاتری دارند، عملکرد کلی سیستم را بهبود دهد. این مدل نیز همچون مدل اول با احتمال خرابی کامیون‌ها روبه‌رو است و این مسئله را در فرآیند تخصیص مد نظر قرار می‌دهد. در نمودار ۲ بهره‌وری شاول‌ها را در این روش را مشاهده می‌کنید.

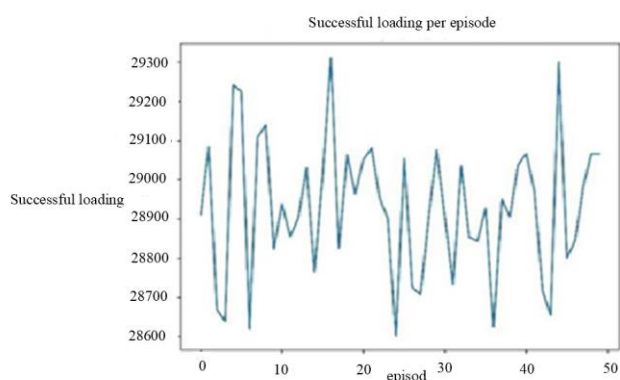


نمودار ۲. بهره‌وری شاول‌ها در روش تخصیص به سریع‌ترین شاول

Figure ۲. Evaluation of shovel productivity under the fastest Allocation Strategy

تخصیص‌های موفق نیز در این روش ۳۵۰ می‌باشد. همچنین میانگین تعداد سفرهای کامیون‌ها در این روش در یک ساعت به طور میانگین برابر ۱.۹ سفر می‌باشد.

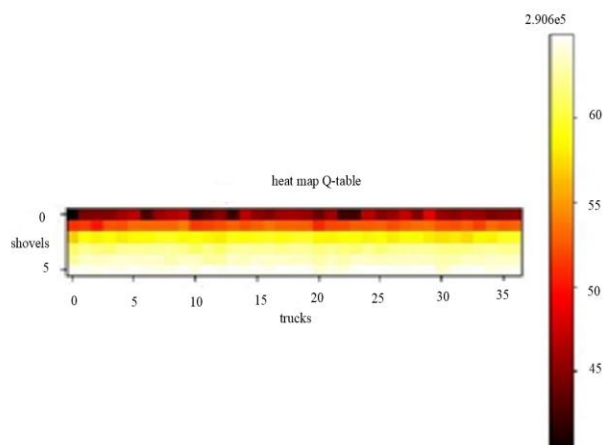
زمان پیشرفت می کند و آیا به سمت بهبود عملکرد می رود یا خیر.



نمودار ۵. تعداد بارگیری های موفق در هر اپیزود

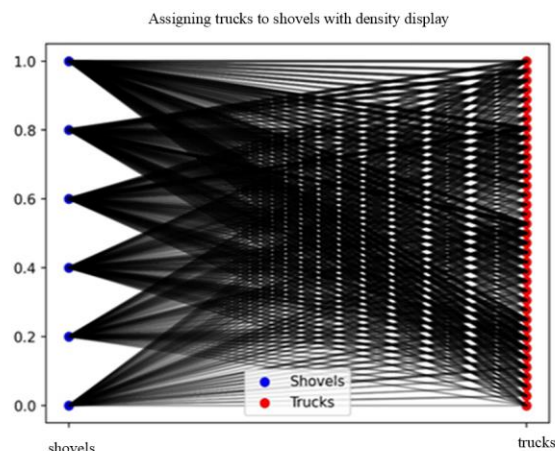
Figure ۵. The number of successful load operations in each episode

یکی دیگر از خروجی های این کد، نموداری است که تخصیص کامیون ها به شاول ها را با استفاده از نمایش چگالی نشان می دهد. در این نمودار، موقعیت های شاول ها و کامیون ها به صورت نقاطی در دو طرف نمودار نمایش داده می شود و خطوطی که این نقاط را به هم متصل می کنند نشان دهنده تخصیص ها هستند. ضخامت و شفافیت این خطوط نشان دهنده تعداد دفعات تخصیص هر جفت شاول-کامیون است. نمودار ۶ نشان می دهد کدام تخصیص ها بیشتر انجام شده اند.



نمودار ۶. نمودار چگالی تخصیص

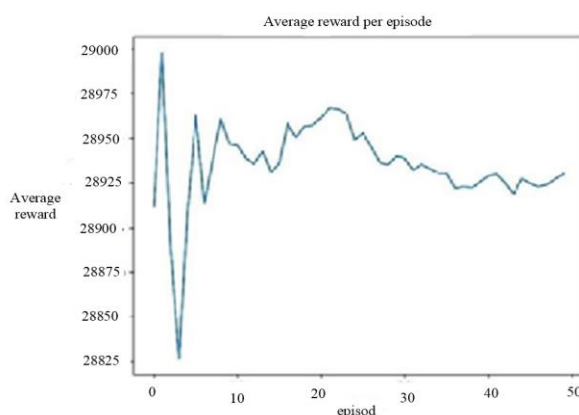
Figure ۶. Density plot of Allocation



نمودار ۳. Q-table حاصل در یادگیری تقویتی

Figure ۳. The resulting Q-table obtained from reinforcement learning

عملکرد الگوریتم از طریق چندین معیار اندازه گیری می شود که شامل تعداد بارگیری های موفق در هر اپیزود و مجموع پاداش های تجمعی در طول زمان است. نمودارهایی که این معیارها را نشان می دهند به تحلیل کارایی الگوریتم کمک می کنند. نمودار ۴ مجموع پاداش های تجمعی در طول اپیزودها را نشان می دهد که نشان دهنده بهبود تدریجی عملکرد الگوریتم در طول زمان است.



نمودار ۴. مجموع پاداش های تجمعی در طول شبیه سازی

Figure ۴. Cumulative sum of Rewards over the Course of the Simulation

نمودار ۵ تعداد بارگیری های موفق در هر اپیزود را نشان می دهد. این نمودار نشان می دهد که چگونه الگوریتم در طول

۳.۴. مقایسه عملکرد مدل‌های شبیه‌سازی با یادگیری تقویتی

برای مقایسه عملکرد سه مدل بررسی شده، از معیارهای مختلفی استفاده می‌شود. یکی از این معیارها تعداد بارگیری‌های موفق در طول زمان شبیه‌سازی است. مدل یادگیری تقویتی با یادگیری از تجربیات گذشته، قادر است تا تعداد بیشتری بارگیری موفق داشته باشد، زیرا از سیاست بهینه‌ای که یاد گرفته است پیروی می‌کند. در مقابل، مدل‌های ساده‌تر که تخصیص به اولین یا سریع‌ترین شاول را انجام می‌دهند، احتمالاً تعداد کمتری بارگیری موفق خواهند داشت زیرا این مدل‌ها بهینه‌سازی پیچیده‌ای انجام نمی‌دهند و فقط به سادگی تخصیص‌ها را انجام می‌دهند. در جدول ۳ مقایسه سه روش در هر قسمت آمده است.

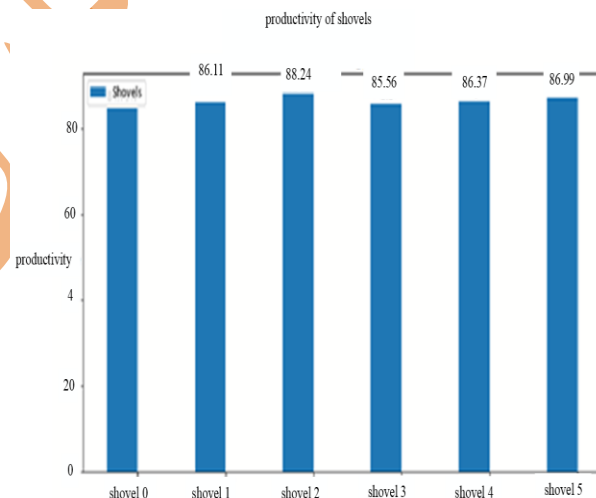
جدول ۳. مقایسه یادگیری تقویتی با دو روش دیگر

Table ۳. Comparing reinforcement learning with two other methods

Reinforcement learning	Allocation to the fastest shovel	Allocation to the first available shovel	method/ Performance criteria
۸۵,۵	۵۳,۵۸	۷۵,۴۵	Average productivity of shovels
۷۲۰	۳۸۰	۶۹۰	Number of successful allocations in the time interval

با توجه به خروجی‌های مختلف این کد، می‌توان نتیجه‌گیری کرد که الگوریتم Q-learning توانسته است با موفقیت تخصیص بهینه را یاد بگیرد. نقشه حرارتی Q-table نشان‌دهنده ارزش‌های تخصیص بهینه است و نمودارهای عملکرد الگوریتم نشان‌دهنده بهبود تدریجی در طول زمان است. همچنین نمودار تخصیص‌ها به صورت بصری نشان می‌دهد که کدام تخصیص‌ها بیشتر مورد استفاده قرار گرفته‌اند.

این مدل بهینه‌سازی تخصیص کامیون‌ها به شاول‌ها در یک محیط معدنی را به صورت کارآمد انجام داده و بهبود عملکرد سیستم را در طول زمان نشان می‌دهد. با استفاده از این مدل می‌توان در محیط‌های واقعی تصمیمات بهتری در مورد تخصیص منابع گرفت و بهره‌وری سیستم را افزایش داد. در نمودار ۷ بهره‌وری شاول‌ها نمایش داده شده است.



نمودار ۷. بهره‌وری شاول‌ها در روش یادگیری تقویتی

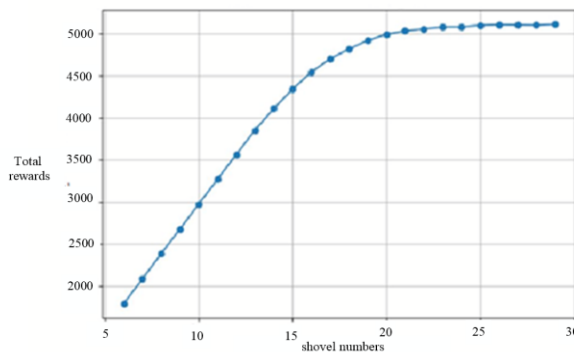
Figure ۷. Productivity of shovels using Reinforcement Learning

باید در نظر داشت تعداد تخصیص‌های موفق در این روش در حدود ۸۰۰ می‌باشد. میانگین تعداد سفرهای کامیون‌ها در این روش ۲.۷ سفر در ساعت می‌باشد. همچنین تخصیص‌های بهینه در روش یادگیری تقویتی در بستر افق شبیه‌سازی نیز بدست می‌آید.

۲,۷	۲,۲	۱,۹	Average number of trips per truck per hour
-----	-----	-----	--

می‌توان نتیجه گرفت که تعداد شاول‌ها پارامتر حساسی در این سیستم است. این اطلاعات به تصمیم‌گیری‌های مدیریتی کمک می‌کند. به عنوان مثال در تعیین تعداد بهینه شاول‌ها و کامیون‌ها برای حداکثرسازی بهره‌وری عملیات معدن‌کاری. به کمک نمودارها می‌توان تحلیل دقیقی از تاثیر هر پارامتر بر سیستم داشت و با استفاده از آن‌ها، تصمیمات بهتری برای بهینه‌سازی عملیات حمل و نقل معدن اتخاذ کرد. در ادامه، به تحلیل هر یک از این نمودارها می‌پردازیم.

۱. تعداد شاول‌ها: نمودار ۸ مربوط به تعداد شاول‌ها نشان می‌دهد که با افزایش تعداد شاول‌ها، تعداد تخصیص‌های موفقیت‌آمیز نیز افزایش می‌یابد. این افزایش به دلیل افزایش ظرفیت بارگیری و کاهش زمان انتظار کامیون‌ها برای بارگیری است. با این حال، با رسیدن به یک نقطه مشخص، نرخ افزایش پاداش کاهش می‌یابد که نشان‌دهنده محدودیت ظرفیت فعلی کامیون‌ها برای حمل مواد است. در این حالت، افزودن شاول‌های بیشتر تأثیر کمتری بر بهره‌وری کلی خواهد داشت و به جای آن، باید به فکر افزایش ظرفیت حمل و نقل یا بهبود نرخ بارگیری باشیم.



نمودار ۸. تحلیل حساسیت تعداد شاول‌ها

Figure ۸. Sensitivity Analysis of the Number of Shoveles

۲. تعداد کامیون‌ها: نمودار ۹ مربوط به تعداد کامیون‌ها نیز نشان‌دهنده رابطه‌ای مشابه با تعداد شاول‌هاست. در این نمودار، با افزایش تعداد کامیون‌ها، میزان تخصیص‌های موفق افزایش می‌یابد زیرا تعداد بیشتری کامیون برای حمل مواد در دسترس است. با این حال، بعد از یک نقطه مشخص، افزایش کامیون‌ها تأثیر قابل توجهی بر پاداش کل ندارد که ممکن

مقایسه عملکرد این سه مدل نشان می‌دهد که مدل یادگیری تقویتی با استفاده از الگوریتم Q-learning قادر است تا تخصیص بهینه کامیون‌ها به شاول‌ها را یاد بگیرد و بهترین عملکرد را نسبت به دو مدل ساده‌تر داشته باشد. این مدل با استفاده از به‌روزرسانی مقادیر Q و یادگیری از تجربیات گذشته، سیاستی بهینه می‌آموزد که منجر به بیشترین تعداد بارگیری‌های موفق می‌شود و بهره‌وری کلی سیستم را افزایش می‌دهد.

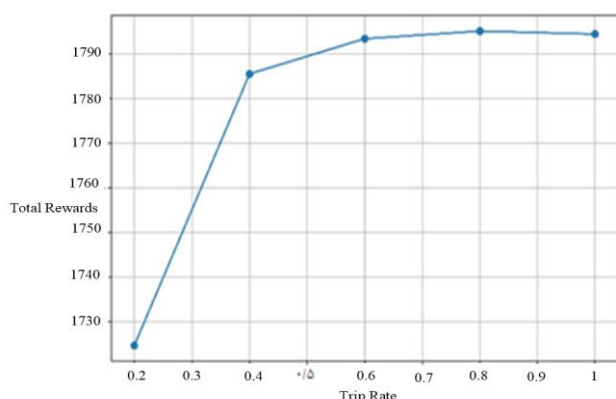
۴.۴. تحلیل حساسیت

تحلیل حساسیت یکی از ابزارهای مهم در مدل‌سازی و شبیه‌سازی سیستم‌ها است که به کمک آن می‌توان فهمید چگونه تغییرات در پارامترهای ورودی بر خروجی‌های سیستم تأثیر می‌گذارند. این فرایند به ویژه در سیستم‌های پیچیده‌ای که با عدم قطعیت‌های زیادی روبه‌رو هستند، بسیار مفید است. هدف از این تحلیل، شناسایی پارامترهایی است که بیشترین تأثیر را بر عملکرد سیستم دارند.

در مدل پژوهش حاضر، بررسی تأثیر تغییرات در پارامترهای مختلفی مانند تعداد شاول‌ها، تعداد کامیون‌ها، نرخ بارگیری شاول‌ها، نرخ حمل و نقل کامیون‌ها، و احتمال خرابی انجام می‌شود. هر یک از این پارامترها به صورت جداگانه تغییر می‌کنند و اثر آن بر روی عملکرد سیستم با استفاده از تعداد موفقیت‌آمیز تخصیص‌ها (تعداد بارگیری‌های انجام شده) به عنوان معیار عملکرد سنجیده می‌شود. به عبارت دیگر، پاداش کل در هر اپیزود به عنوان معیار ارزیابی استفاده می‌شود و میانگین پاداش‌های اپیزودهای مختلف به عنوان نتایج تحلیل حساسیت ثبت می‌شود.

این تحلیل نشان می‌دهد کدام پارامترها بیشترین تأثیر را بر عملکرد سیستم دارند و حساسیت سیستم نسبت به تغییرات در این پارامترها چگونه است. به عنوان مثال، اگر تغییرات در تعداد شاول‌ها تأثیر قابل توجهی بر پاداش کل داشته باشد،

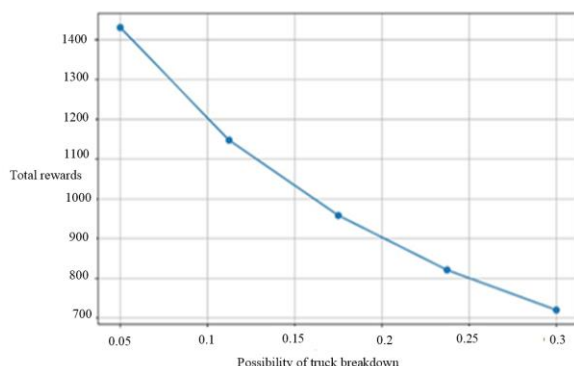
بهبود عملکرد سیستم منجر می‌شود. افزایش نرخ حمل و نقل به معنای کاهش زمان لازم برای رفت و برگشت کامیون‌ها و در نتیجه افزایش تعداد دفعات بارگیری و تخلیه است. با این حال، همان‌طور که در نمودار مشاهده می‌شود، بعد از رسیدن به یک نقطه بهینه، تاثیر افزایش نرخ حمل و نقل کاهش می‌یابد و نیاز به بهبود در دیگر اجزای سیستم (مثل تعداد کامیون‌ها) احساس می‌شود.



نمودار ۱۱. تحلیل حساسیت نرخ حمل و نقل کامیون‌ها

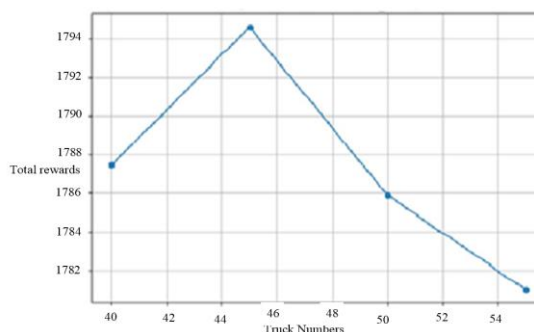
Figure ۱۱. Sensitivity Analysis on Truck Transport Rate

۵. احتمال خرابی کامیون‌ها: نمودار ۱۲ احتمال خرابی کامیون‌ها نیز روند مشابهی با شاول‌ها را نشان می‌دهد. با افزایش احتمال خرابی، تعداد تخصیص‌های موفق کاهش می‌یابد که به دلیل افزایش زمان تعمیر کامیون‌ها و کاهش تعداد کامیون‌های در دسترس برای حمل مواد است. این نمودار نیز بر اهمیت مدیریت و نگهداری کامیون‌ها برای کاهش زمان توقف ناشی از خرابی تأکید دارد.



نمودار ۱۲. تحلیل حساسیت نرخ خرابی کامیون

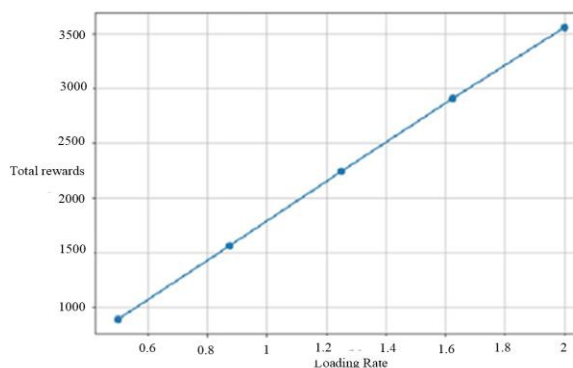
است به دلیل محدودیت ظرفیت بارگیری شاول‌ها یا ترافیک در عملیات حمل و نقل باشد.



نمودار ۹. تحلیل حساسیت تعداد کامیون‌ها

Figure ۹. Sensitivity study of Truck Count in the System

۳. نرخ بارگیری شاول‌ها: نمودار ۱۰ نرخ بارگیری شاول‌ها نشان می‌دهد. چنانچه مشاهده می‌شود، افزایش نرخ بارگیری تاثیر مثبتی بر تعداد تخصیص‌های موفق دارد. این به دلیل کاهش زمان بارگیری و افزایش سرعت عملیات است. این نمودار به وضوح نشان می‌دهد که بهبود نرخ بارگیری می‌تواند به افزایش کارایی عملیات معدن کاری منجر شود. به‌خصوص در شرایطی که ظرفیت کامیون‌ها بهینه است، افزایش نرخ بارگیری می‌تواند به بهبود قابل توجهی در بهره‌وری منجر شود.



نمودار ۱۰. تحلیل حساسیت نرخ بارگیری شاول

Figure ۱۰. Sensitivity analysis of shovel loading rate

۴. نرخ حمل و نقل کامیون‌ها: نمودار ۱۱ مربوط به نرخ حمل و نقل کامیون‌ها نشان‌دهنده این است که افزایش این نرخ به

Figure ۱۲. Sensitivity analysis of truck
breakdown rates

۵. نتیجه‌گیری

بخش قابل توجهی از هزینه‌های عملیاتی در معادن روباز به عملیات حمل‌ونقل مواد اختصاص دارد؛ از این رو، بهینه‌سازی تخصیص و زمان‌بندی کامیون‌ها و شاول‌ها نقش تعیین‌کننده‌ای در کاهش هزینه و افزایش بهره‌وری دارد. هدف اصلی این پژوهش، توسعه و پیاده‌سازی یک مدل ترکیبی شبیه‌سازی-یادگیری تقویتی برای مدیریت هوشمند حمل‌ونقل در محیط‌های پویا و نامطمئن معدنی بود.

نتایج حاصل از اجرای مدل نشان داد که الگوریتم-Q learning توانایی قابل توجهی در یادگیری سیاست‌های بهینه تخصیص کامیون‌ها به شاول‌ها دارد. مدل توسعه‌یافته با یادگیری تدریجی از تجربیات گذشته، موفق شد تعداد بارگیری‌های موفق و بهره‌وری کلی سیستم را به‌طور چشمگیری افزایش دهد.

در مقایسه با پژوهش‌های پیشین نظیر اسناتفیری و همکاران (۲۰۲۱) و یوپادهی و عسگری‌نسب (۲۰۱۷)، نتایج این تحقیق نشان می‌دهد که افزودن فاکتورهای خرابی تجهیزات و یادگیری تطبیقی عامل هوشمند موجب واقع‌گراتر شدن مدل و بهبود عملکرد آن در شرایط عملیاتی شده است. در حالی که بیشتر پژوهش‌های گذشته تنها به بهینه‌سازی استاتیک یا تخصیص ثابت پرداخته‌اند، مدل حاضر با رویکرد یادگیری پویا توانسته است رفتار سیستم را در حضور عدم قطعیت‌های واقعی (مانند نرخ خرابی کامیون‌ها و تغییرات زمانی بارگیری) به‌صورت مؤثر مدل‌سازی کند. بنابراین، یافته‌های این تحقیق در عین هم‌سویی با اصول مطرح‌شده در مطالعات قبلی، نشان‌دهنده‌ی بهبود عملکرد نسبت به مدل‌های سنتی و شبیه‌سازی صرف است.

نتایج تحلیل حساسیت نشان داد که افزایش تعداد شاول‌ها تا حدودی موجب رشد تعداد تخصیص‌های موفق می‌شود، اما پس از رسیدن به نقطه‌ی اشباع، تأثیر آن کاهش می‌یابد. بر همین اساس، می‌توان تعداد بهینه شاول‌ها و کامیون‌ها را برای هر معدن بر مبنای ظرفیت واقعی سیستم استخراج کرد. همچنین مشاهده شد که احتمال خرابی کامیون‌ها تأثیر قابل توجهی بر عملکرد کل سیستم دارد؛ به‌طوری‌که با افزایش نرخ خرابی، میانگین بهره‌وری به‌صورت غیرخطی کاهش می‌یابد. این یافته اهمیت نگهداری پیشگیرانه تجهیزات و استفاده از سامانه‌های هوشمند پایش وضعیت را در مدیریت معادن برجسته می‌کند.

بر اساس نتایج مدل، پیشنهاد می‌شود مدیران معادن از الگوریتم‌های یادگیری تقویتی برای برنامه‌ریزی پویا و بلادرنگ (Real-time) در تخصیص کامیون‌ها استفاده کنند. همچنین استفاده از داده‌های عملیاتی برای به‌روزرسانی مداوم سیاست‌های تصمیم‌گیری می‌تواند بهره‌وری سیستم را در بلندمدت افزایش دهد.

اضافه کردن هرچند مدل پیشنهادی توانسته است جنبه‌های واقع‌گرایانه‌ای همچون خرابی تجهیزات و تعامل پویا میان عامل و محیط را در نظر بگیرد، اما این پژوهش نیز با محدودیت‌هایی روبه‌رو بوده است. از جمله:

محدودیت در اندازه‌ی مطالعه‌ی موردی و دسترسی به داده‌های جامع از تمام دوره‌های عملیاتی معدن. فرضیات ساده‌کننده‌ی مدل در مورد نرخ‌های بارگیری و حمل‌ونقل که در شرایط واقعی ممکن است تغییرپذیرتر باشند.

عدم در نظر گرفتن اثرات شرایط محیطی (مانند وضعیت جوی یا مسیرهای ناهموار) که می‌توانند بر عملکرد کامیون‌ها و زمان‌بندی اثرگذار باشند.

(References) مراجع

١. Namin, F.S., Ghasemzadeh, H. and Aghajari, A.M., ٢٠٢٣. A comprehensive approach to selecting mine transportation system using AHP and FUZZY-TOPSIS. *Decision Making and Analysis*, 1, pp.٢٣-٣٩. <https://doi.org/10.50976/dma.12.23.117323-39>
٢. Alarie, S. and Gamache, M., ٢٠٠٢. Overview of solution strategies used in truck dispatching systems for open pit mines. *International Journal of Surface Mining Reclamation and Environment*, 16(١), pp.٥٩-٧٦. <https://doi.org/10.1076/ijsm.16.1.59.3408>
٣. Souchleris, K., Sidiropoulos, G.K. and Papakostas, G.A., ٢٠٢٣b. Reinforcement learning in game industry review, prospects and challenges. *Applied Sciences*, 13(٤), p.٢٤٤٣. <https://doi.org/10.3390/app13042443>
٤. Isnafitri, M.F., Rosyidi, C.N. and Aisyati, A., ٢٠٢١. A truck allocation optimization model in open pit mining to minimize investment and transportation costs. *IOP Conference Series: Materials Science and Engineering*, 1096(١), p.٠١٢٠٢٤. <https://doi.org/10.1088/1757-899X/1096/1/012024>
٥. Upadhyay, S.P. and Askari-Nasab, H., ٢٠١٧. Dynamic shovel allocation approach to short-term production planning in open-pit mines. *International Journal of Mining, Reclamation and Environment*, 33(١), pp.١-٢٠. <https://doi.org/10.1080/17480930.2017.1310024>
٦. Makarova, I., Khabibullin, R., Belyaev, E., Gabsalikhova, L. and Mukhametdinov, E., ٢٠١٦. Improving the logistical processes in corporate service system. *Transport Problems*, 11(١), pp.٥-١٨. <https://doi.org/10.20808/tp.2016.11.1.1>
٧. Fadin, A.Y.F., ٢٠١٧. Simulation-optimization truck dispatch problem using look-ahead algorithm in open pit mines. *International Journal of GEOMATE*, 13(٣٦). <https://doi.org/10.21660/2017.36.2.843>
٨. Bakhtavar, E. and Mahmoudi, H., ٢٠١٨. Development of a scenario-based robust model for the optimal truck-shovel allocation in open-pit mining. *Computers & Operations Research*, 104, p.١١٥. <https://doi.org/10.1016/j.cor.2018.08.003>
٩. Hashemi, A.S. and Sattarvand, J., ٢٠١٥b. Simulation based investigation of different fleet management paradigms in open pit mines – a case study of Sungun copper mine. *Archives of Mining Sciences*, 60(١), pp.١٩٥-٢٠٨. <https://doi.org/10.1016/amsc-2015-0013>
١٠. Mohtasham, M., Mirzaei-Nasirabad, H. and Alizadeh, B., ٢٠٢١. Optimization of truck-shovel allocation in open-pit mines under uncertainty: a chance-constrained goal

- programming approach. *Mining Technology: Transactions of the Institutions of Mining and Metallurgy*, 13(2), pp.81-100.
<https://doi.org/10.1080/20726668.2021.1916170>.
۱۱. Yao, J., Wang, Z., Chen, H., Hou, W., Zhang, X., Li, X. and Yuan, W., ۲۰۲۳. Open-pit mine truck dispatching system based on dynamic ore blending decisions. *Sustainability*, 15(۴), p.۳۳۹۹.
<https://doi.org/10.3390/su15043399>.
۱۲. Pathan, S.M., Pathan, A.G. and Memon, M.S., ۲۰۲۰. Simulation optimization of shovel-truck system in open-pit mines considering rockmass parameters. *Journal of Advanced Transportation*, ۲۰۲۰(۱), p.۷۹۳۹۰۳۷.
<https://doi.org/10.1155/atr/7939.37>
۱۳. Zhang, X., Guo, A., Ai, Y., Tian, B. and Chen, L., ۲۰۲۲. Real-time scheduling of autonomous mining trucks via flow allocation-accelerated tabu search. *IEEE Transactions on Intelligent Vehicles*, 7(۳), pp.۴۶۶-۴۷۹.
<https://doi.org/10.1109/TIV.2022.3167064>
۱۴. Yaghoubi, A., Bozorgi Amiri, A. & Sadegh Amalnik, M. (۲۰۲۳). A bi-objective programming model for designing a healthcare service network considering location efficiency. *Journal of Industrial Engineering and Management*, 39(۱), pp. ۱۶۹-۱۸۲.
<https://doi.org/10.24200/j75.2022.07193.2187>
۱۵. Aghadavoudi Jelfaei, A. & Alinaghian, M. (۲۰۲۳). Presenting two mathematical models and four heuristic algorithms for the vehicle routing problem considering customers' suggested location-time options. *Journal of Industrial Engineering and Management*, 39(۱), pp. ۷۳-۸۴.
<https://doi.org/10.24200/j75.2022.07180.2221>
۱۶. Mazinani, E. (۲۰۲۰). *Optimizing Shovel-Truck Allocation and Scheduling in Open-Pit Mines Considering Truck Failures*. Master's Thesis, Tarbiat Modares University.[in persian]

Optimization of shovel and truck allocation and scheduling in open pit mines using simulation based on reinforcement learning algorithms

Erfan. Mazinani,

erfan.mazinani@modares.ac.ir

Master of Science, Faculty of Industrial and
Systems Engineering, Tarbiat Modares
University, Tehran, Iran

Nasim. Nahavandi, (Corresponding author),

n_nahavandi@modares.ac.ir

Professor, Faculty of Industrial and Systems
Engineering, Tarbiat Modares University,
Tehran, Iran

significant improvement in the efficiency of the entire system.

Keywords: Simulation, reinforcement learning algorithms, shovel allocation and scheduling, truck allocation and scheduling, open pit mines.

Abstract

Due to the inherent complexity of transportation systems in open-pit mines, traditional planning methods are often unable to address the dynamics and uncertainties inherent in these processes. In this study, a simulation system based on reinforcement learning algorithms has been developed to simultaneously optimize the allocation and scheduling of shovels and trucks in open-pit mines, with the aim of improving operational efficiency. In this system, an intelligent agent is trained using reinforcement learning algorithms to make efficient decisions regarding vehicle routing, loading and unloading scheduling, and optimal resource allocation. The results of the model implementation show that the use of the Q-learning algorithm leads to learning the optimal allocation of trucks to shovels and outperforms two simpler comparative models. By gradually updating Q values and learning from previous experiences, this algorithm learns an optimal decision-making policy, which ultimately leads to an increase in the number of successful downloads and a