

یک رویکرد هوشمند مبتنی بر الگوریتم جامعه مورچگان تقویت‌شده برای تخصیص بهینه وظایف در عملیات امداد و نجات زلزله

ناهید بهرامی^۱، میثم ارگانی^{۱*}، علی درویشی بلورانی^۱، علیرضا وفائی‌نژاد^۲

^۱ دانشگاه تهران، تهران، ایران

nd.bahrami@ut.ac.ir, argany@ut.ac.ir, ali.darvishi@ut.ac.ir

^۲ دانشگاه شهید بهشتی، تهران، ایران

a_vafaei@sbu.ac.ir

(دریافت: مهر ۱۴۰۴، تصویب: آذر ۱۴۰۴)

چکیده

مدیریت مؤثر تیم‌ها، به‌ویژه در شرایط بحرانی، یکی از چالش‌های اساسی و در عین حال حیاتی در عملیات‌های امداد و نجات به شمار می‌رود. در این شرایط، تصمیم‌گیری سریع، هماهنگی مؤثر میان اعضا و تخصیص بهینه فعالیت‌ها، نقش تعیین‌کننده‌ای در موفقیت مأموریت‌ها و کاهش تلفات انسانی دارد. از این رو، شناسایی و به‌کارگیری روش‌های علمی و فناورانه برای ارتقای عملکرد تیم‌های امدادی از اهمیت بالایی برخوردار است. پژوهش حاضر با هدف بررسی و بهبود عملکرد تیم‌های امداد و نجات در زمان وقوع زمین‌لرزه انجام گرفته است. زلزله‌ها به دلیل وسعت خسارات، گستردگی مناطق آسیب‌دیده و محدودیت زمانی در نجات مصدومان، از جمله پیچیده‌ترین و بحرانی‌ترین مخاطرات طبیعی محسوب می‌شوند. هرگونه بهبود در مدیریت تیم‌های عملیاتی و بهینه‌سازی فرآیند تصمیم‌گیری در چنین شرایطی می‌تواند منجر به افزایش سرعت امدادرسانی و در نتیجه نجات جان افراد بیشتری شود. در این پژوهش، برای طراحی و پیاده‌سازی یک مدل بهینه‌سازی عملکرد تیم‌های امدادی، از تلفیق علوم و فنون نوین مکانی با الگوریتم‌های هوش مصنوعی استفاده شده است. بر اساس مطالعات انجام‌شده، الگوریتم هوش جمعی جامعه مورچگان به عنوان یکی از کارآمدترین روش‌های مبتنی بر رفتار جمعی برای حل مسائل تیمی پیچیده مناسب شناسایی شد. با این حال، برای ارتقای توانایی این الگوریتم در انطباق با محیط، از روش یادگیری تقویتی بهره گرفته شد و الگوریتم یادگیری Q-Learning برای ترکیب با مدل جامعه مورچگان انتخاب شد. این ترکیب موجب شد تا فرآیند تخصیص وظایف و فعالیت‌ها به افراد در مدل پیشنهادی، ارتقا و بهبود یابد. برای اجرا و بررسی عملکرد مدل، ابتدا پارامترهای مؤثر در ساختار تیم‌ها و تخصص افراد، وظایف و فعالیت‌ها، و میزان خسارات انسانی و سازه‌ای، شناسایی و در مدل لحاظ شده است. محدوده‌ی مطالعاتی پژوهش در ناحیه چهارم منطقه ۳ شهر تهران انتخاب شد و ۷۰ امدادگر در قالب ۱۴ تیم پنج نفره که همگی مطابق با استانداردها و پروتکل‌های هلال احمر ایران آموزش دیده‌اند، در عملیات شبیه‌سازی‌شده شرکت داده شدند. داده‌های مربوط به موقعیت مکانی، میزان خسارات، تخصص اعضا و فعالیت‌های جست‌وجو و نجات، به عنوان ورودی‌های اصلی مدل مورد استفاده قرار گرفتند. در مدل پیشنهادی، توابع پاداش و هزینه با در نظر گرفتن شاخص‌هایی همچون سرعت حرکت، زمان انجام فعالیت‌ها، فاصله محل فعالیت‌ها و در کل میزان موفقیت در نجات افراد، تعریف و تدوین شدند. نتایج به‌دست‌آمده نشان داد که مدل ترکیبی پیشنهادی توانست به کارایی ۶۹/۴ درصد دست یابد و عملکرد عملیات امداد و نجات را در مقایسه با الگوریتم واحد جامعه مورچگان که به کارایی ۵۰/۱ درصد یافته بود، تا حدود ۱۹ درصد افزایش دهد. به طور کلی، یافته‌های پژوهش حاضر نشان می‌دهد که ترکیب روش‌های هوش جمعی با یادگیری تقویتی می‌تواند چارچوبی نوین برای مدیریت هوشمند تیم‌های امدادی در شرایط بحرانی فراهم سازد. مدل ارائه‌شده نه تنها در عملیات‌های زلزله، بلکه در سایر انواع بحران‌های طبیعی یا انسانی نیز قابلیت توسعه و به‌کارگیری دارد. این رویکرد می‌تواند به عنوان گامی مؤثر در جهت ایجاد سامانه‌های تصمیم‌یار هوشمند برای مدیریت بحران در سطوح شهری و ملی مورد استفاده قرار گیرد.

واژگان کلیدی: هوش مصنوعی، تخصیص وظایف، امداد و نجات، الگوریتم جامعه مورچگان، یادگیری تقویتی، زلزله.

۱- مقدمه

برنامه‌ریزی، زمانبندی و ارتقا عملکرد تیم‌ها در مسائلی که حل آن‌ها تیم‌محور است می‌تواند در راندمان نهایی نقش به‌سزایی ایفا کند. همان‌طور که در سراسر دنیا از روش‌های هوش مصنوعی در حل مسائل مختلف استفاده می‌شود می‌توان از روش‌های ایجاد هوشمندی برای ارتقا عملکرد و کوتاه نمودن زنجیره پاسخ در مواجهه با بحران‌هایی نظیر مخاطرات طبیعی مانند زلزله بهره برد. از طرف دیگر با توجه به گستردگی بحران‌های مخاطرات طبیعی و نیز خطرات جانی مخاطره زمین‌لرزه، موضوع این پژوهش اهمیت دو چندان پیدا می‌کند.

از این‌رو با توجه به اهمیت موضوع و نوآوری مدنظر این پژوهش مطالعه‌ای بر روی پژوهش‌های پیشین و کتب امداد و نجات در سوانح طبیعی و آوارها صورت گرفته، و مصاحبه و گفتگویی نیز با متخصصان امداد و نجات در بحران‌هایی نظیر زلزله انجام شده است. در ادامه برخی از پژوهش‌ها مورد بیان و بررسی قرار می‌گیرد. در سال ۲۰۰۹ وفائی‌نژاد و همکارانشان با توسعه یک مدلی الهام گرفته شده از مناقصات اقتصادی به ارائه روشی نوین و کارآمد جهت تخصیص وظایف افراد امداد رسان در محیطی پویا و با در نظرگیری زمان در جستجو و نجات زلزله پرداخته اند [۱].

همچنین در سال ۲۰۱۰ در پژوهشی وفایی نژاد و همکاران با مدلسازی مکانی- زمانی راهی برای افزایش راندمان انجام فعالیت‌های دسته جمعی انسان‌ها ارائه نمودند. تئوری هاگروستراند را به عنوان مبنای مدلسازی در نظر گرفتند و مدل پیشنهادی آن‌ها باعث حل مسائل چند بعدی شد. برای بررسی کارایی مدل، فعالیت‌های دو گروه زنده‌یابی و آواربرداری جستجو و نجات زلزله شبیه‌سازی و ارزیابی نمودند و نتایج بهبود ۱۸.۸۳ درصدی عملکرد گروه‌ها در صورت استفاده از مدل را نشان داد [۲].

در سال ۲۰۱۴ راسخ و همکاران با استفاده از مدل صف برای عملیات جستجو و نجات زلزله به کمک سامانه اطلاعات جغرافیایی مدلی شبیه‌سازی نمودند که بتواند میان تعداد گروه‌های امدادی در منطقه و توان عملیاتی آن‌ها یک حالت تعادل ایجاد نماید. با استفاده از مدل نرخ بقای افراد زیر آوار، بر اساس نوع سازه، عامل مؤثری برای تحلیل مدل صف شبیه سازی شده فراهم کردند و در نتیجه آن تخصیص تعداد و ترکیب مناسب نیروها در منطقه مورد مطالعه بود [۳].

در پژوهشی در سال ۲۰۱۵ Ole-Christoffer Granmo و همکاران مدلی احتمالاتی برای برنامه‌ریزی تخلیه در وقوع سوانح با در نظر گیری پویایی جمعیت و محیط به هنگام آتش‌سوزی کشتی ارائه کردند. مدل احتمالاتی بر اساس پویایی مخاطره و مدل آتش، روانشناسی جمعیت و زمان طراحی شده است. نهایتاً از مدل DNB برای پیگیری و پیش-بینی حرکت افراد و از مدل ACO برای یافتن راه‌های امن به صورت پویا در پاسخ به خطرات ثانویه هنگام تخلیه استفاده شده و به مدل مکانی- زمانی پویا دست یافته است [۴].

در سال ۲۰۲۰ Singh Patel و همکاران در پژوهشی یک چارچوب یکپارچه برای مدیریت زلزله پیشنهاد شده است که جمع‌آوری داده‌ها از منابع عمومی در دسترس شروع می‌شود، و سپس داده‌ها را پاکسازی، کاوش، ویژگی‌سازی می‌کند و مدل‌سازی با استفاده از مدل‌های CNN مبتنی بر یادگیری عمیق انجام می‌دهد. تحلیل‌ها با تمرکز بر تخریب ساختمان، تأثیر زمین‌لرزه‌ها را بررسی می‌کند و تأثیر آن را اندازه‌گیری و نظارت می‌کند [۵].

در پژوهشی در سال ۲۰۲۰ Li Yan و همکاران یک سیستم نجات مبتنی بر تحرک انسانی در طوفان و سیل به نام MobiRescue پیشنهاد کردند، که هدف آن به حداکثر رساندن تعداد درخواست‌های نجات برآورده شده و به حداقل رساندن تاخیر و تعداد تیم‌های نجات اعزام شده به موقعیت‌ها است. ابتدا دریافتند که شدت تأثیر سانحه در مناطق مختلف متفاوت است و حرکت مردم به طور قابل‌توجهی تحت تأثیر آن است، به این معنی که مسیرهای رانندگی تیم‌های امداد باید به طور سازگار تنظیم شود. سپس، یک روش مبتنی بر ماشین بردار پشتیبانی (SVM) برای پیش‌بینی توزیع درخواست‌های نجات و یک روش اعزام تیم نجات مبتنی بر یادگیری تقویتی (RL) برای دستیابی به اهداف توسعه دادند و برتری MobiRescue را نشان دادند [۶].

در سال ۲۰۲۱ Jalumuru Nalini و P. M. Khilar در یک کار تحقیقاتی، استراتژی جدیدی برای حل زمان‌بندی کار با استفاده از بهینه‌سازی کلونی مورچه‌ها (ACO) با تطبیق یادگیری تقویتی (RL) همراه با تحمل خطا برای مقاوم‌سازی فرآیند زمان‌بندی در برابر خطا و دستیابی به هدف حداقل فاصله زمانی پیشنهاد شده است. الگوریتم پیشنهادی، بهینه‌سازی کلونی مورچه‌های تقویت‌شده (RACO) حدود ۶۰ درصد عملکرد بهتری نسبت به اجرای انفرادی ACO دارد [۷].

در سال ۲۰۲۲ Zhaojia Tang و همکاران در پژوهشی برای دستیابی به هدف برنامه‌ریزی و شناسایی مسیر برای ربات بازرسی و ارزیابی خسارت ساختمان پس از زلزله، یک روش برنامه‌ریزی مسیر گام کوچک توزیع شده با استفاده از یادگیری تقویتی اصلاح شده (MRL) پیشنهاد شده است. MRL جهت را به روز می‌کند و مسیر را برای جلوگیری از اختلالات ناشناخته تنظیم می‌کند. پس از یافتن بهترین زاویه بازرسی، دوربین روی ربات می‌تواند تصویر را به وضوح ثبت کند و در نتیجه قابلیت تشخیص را بهبود بخشد [۸].

در پژوهشی در سال ۲۰۲۳ Xiaoyan Li و همکاران یک روش ارزیابی نوآورانه برای مقادیر ریسک خروجی و احتمال ایمنی هر منطقه قابل عبور در سیل را ارائه دادند. آن‌ها یک الگوریتم برنامه‌ریزی مسیر پیشگام را بر اساس یادگیری تقویتی عمیق توسعه دادند که بهره‌برداری از پاداش تصادفی بهبود یافته و مکانیسم‌های کاوش پاداش ناهمگن را برای عملکرد در یک سناریوی برنامه‌ریزی مسیر نجات با عدم قطعیت و پیچیدگی ترکیب کرده است. با توجه به یافته‌ها، الگوریتم پیشنهادی در همگرایی به مسیر بهینه، نمونه‌برداری کامل و راندمان اجراء بهتر عمل کرده است [۹].

در سال ۲۰۲۴ Rahat Tamzid و همکاران در پژوهشی برای برنامه‌ریزی هوشمند مسیر نجات پس از زلزله با استفاده از اینترنت اشیا، یک محیط شبیه‌سازی شده با ۱۶ نمونه آزمایش شده است و با موفقیت داده‌ها را از دستگاه‌های IoT مستقر در ساختمان‌ها جمع‌آوری و پردازش، و آسیب‌های ساختاری را طبقه‌بندی می‌کند و تعداد افراد آسیب‌دیده را تخمین می‌زند. این رویکرد به طور قابل توجهی زمان مورد نیاز برای شروع عملیات نجات را با ارائه مسیرهای بهینه به امدادگران به سمت مکان‌های فاجعه کاهش می‌دهد، که آثار مثبت بالقوه‌ای برای استراتژی‌های مدیریت بلایا دارد [۱۰].

در مطالعه‌ای که توسط Minal Shahakar و همکاران در سال ۲۰۲۴ سیستم جدیدی که بهینه‌سازی کلونی مورچه‌ها (ACO) و یادگیری تقویتی (RL) را ترکیب می‌کند برای کمک به مشکلاتی که در هنگام تغییر وظایف و نیاز به منابع در سیستم‌های پراکنده و بزرگ پیش می‌آیند. ACO بر اساس نحوه یافتن غذا توسط مورچه‌ها و برای تغییر نحوه توزیع مشاغل بین گره‌های کامپیوتری بر اساس دانش محلی و مسیرهای فرومون استفاده می‌شود. این روش جستجوی سریع راه‌حل‌ها و سازگاری با موقعیت‌های جدید را آسان می‌-

کند و از RL نیز برای یادگیری و تنظیم تغییرات سیستم در طول زمان استفاده می‌شود.

با برنامه‌ریزی تعادل بار به عنوان مجموعه‌ای از تصمیمات، عوامل RL قادر به بهبود قوانین خود هستند تا سیستم بهتر کار کند و منابع به طور موثرتری استفاده شوند. این سیستم رفتار خود را تغییر می‌دهد تا با استفاده از هوش گروهی ACO و انعطاف‌پذیری RL، نسبت به تغییر وظایف و در دسترس بودن منابع واکنش نشان دهد. روش پیشنهادی در بهبود عملکرد سیستم، کاهش زمان واکنش و بهترین استفاده از منابع است و بهتر از روش‌های متعادل کننده بار استاندارد عمل می‌کند [۱۱].

در سال ۲۰۲۵ در پژوهشی Jin Zhang و همکاران از یادگیری تقویت عمیق (DRL) برای رسیدگی به مشکل مسیریابی پویا در تدارکات اضطراری استفاده کردند. این روش از چارچوب کنشگر-انتقادی، همراه با مکانیسم‌های شبکه‌های اشاره‌گر و شبکه‌های عصبی حافظه کوتاه‌مدت برای تعیین مسیر موثر امداد بلایا استفاده، و طرح زمان‌بندی به‌دست‌آمده را با نتایج به‌دست‌آمده از الگوریتم DRL بر اساس مدل تک شبکه‌ای و الگوریتم بهینه‌سازی کلنی مورچه‌ها (ACO) مقایسه می‌کند. نتایج کاهش ۸۰ درصدی در زمان را نسبت به روش سنتی نشان داده است [۱۲].

در نهایت با مطالعات و بررسی‌های انجام شده پارامترهای مهم به منظور برنامه‌ریزی و تخصیص وظایف تیم‌ها شناسایی شد که در قسمت‌های آتی مقاله بیان می‌شود. همچنین استفاده از روش تلفیقی الگوریتم کلونی مورچگان و یادگیری تقویتی Q-learning به عنوان روش پیشنهادی پژوهش انتخاب شده است. بررسی کارایی مدل و مقایسه آن در انتهای بخش پیاده‌سازی آمده است. در این پژوهش مدل ترکیبی انتخاب شده برای برنامه‌ریزی تخصیص وظایف امدادگران در زمین‌لرزه با توجه به پارامترهای اساسی شناسایی شده، نوآوری مدنظر پژوهش است. در نهایت باتوجه به نتایج به دست آمده در این پژوهش، مدل پیشنهادی و نوآوری مورد نظر، برای موضوع امداد و نجات زمین‌لرزه مناسب شناخته شد.

۲- ادبیات پژوهش

در این بخش ساختار تیم‌ها و پارامترهای مهم در امداد رسانی زلزله و فعالیت‌های مهم و مشخصات آن‌ها در

رخداد زمین‌لرزه بررسی و بیان می‌شود. همچنین الگوریتم ACO و یادگیری تقویتی و زیرمجموعه‌های آن که در پژوهش حاضر استفاده می‌شود، پرداخته شده است.

۲-۱- تیم‌ها و پارامترهای مهم

در این پژوهش تیم‌ها مطابق ساختار هلال احمر ایران چیده شده‌اند و طبق بررسی‌های به عمل آمده از مسئولان این سازمان، گروه‌های امداد و نجات در سوانح و مخاطرات طبیعی متشکل از ۵ نفر بوده و تمامی افراد آموزش کامل دیده‌اند تا آمادگی پاسخ به هر رخداد و انجامی تمامی وظایف موجود در اینگونه بحران‌ها را داشته باشند. از این رو، افراد امدادگر در این پژوهش، با تخصص کامل و شرایط جسمانی و تخصصی مشابه در نظر گرفته می‌شوند. در ادامه می‌بایست پارامترهای مهم در این تخصیص وظایف شناسایی شود.

با توجه به اینکه افراد تیم در پژوهش حاضر امدادگران در مخاطره طبیعی زمین‌لرزه است، برای تعیین وظایف و شناسایی پارامترهای مهم در انجام وظایف، مطالعات و بررسی‌هایی در کتب و مقالات مرتبط در زمینه‌های مذکور انجام شد و نیز با مدیران و کارشناسان خبره سازمان هلال احمر گفتگو شد. در بحران‌های مانند زمین‌لرزه اقداماتی که انجام می‌شود در سه بخش قبل از وقوع حادثه، هنگام وقوع و پس از بحران تقسیم می‌شود. ۷۲ ساعت ابتدایی از زمان رخداد زمین‌لرزه، زمان حیاتی برای نجات جان افراد است [۱۳]. از این رو یکی از پارامترهای مهمی که باید در محاسبات لحاظ شود زمان است.

به صورت کلی از منظر صاحب نظران حوزه امداد و نجات در بحران‌هایی مانند زمین‌لرزه که با آوارهای ساختمانی رو به رو هست، امداد رسانی در مراحل کلی جستجو و زنده‌یابی، رهاسازی و تثبیت، و انتقال قرار می‌گیرد [۱۴]. با توجه به گستردگی فعالیت‌های امداد و نجاتی در اینگونه بحران‌ها، در این پژوهش فعالیت‌هایی که قابلیت محاسبه و در نظرگیری زمان حدود در بخش‌های جستجو و نجات و رهاسازی و تثبیت را دارند در نظر گرفته می‌شوند.

این فعالیت‌ها و وظایف به صورت کلی در نظر گرفت می‌شوند و عبارتند از: جستجو، آواربرداری، کمک‌های اولیه و ثبت وضعیت افراد در محل حادثه، قطع خطوط انرژی برای جلوگیری از مخاطرات ثانویه، ارزیابی و نشانه‌گذاری است. اولویت‌بندی این فعالیت‌ها با توجه به مراحل امداد و نجات مشخص شده است و نیز زمان متوسط انجام هر نوع از

فعالیت‌ها با مشورت خبرگان امداد و نجات در هلال احمر صورت پذیرفته است. در ادامه در جدولی (جدول ۱) اولیت-بندی و زمان‌بندی فعالیت‌ها مشخص شده است [۱۵، ۱۶].

جدول ۱- اطلاعات فعالیت‌ها

نوع فعالیت	اولویت‌بندی	متوسط زمان انجام فعالیت (دقیقه)
جستجو افراد	۲	۱۵
آواربرداری سبک	۳	۲۰
آواربرداری سنگین	۶	۳۰۰
کمک‌های اولیه و حمایت روانی	۴	۳۰
ثبت وضعیت افراد	۵	۳
ایمن‌سازی با قطع شریان‌های انرژی	۱	۵
ارزیابی و نشانه‌گذاری	۱	۵

از اطلاعات این جدول در طراحی مدل و اجرای پژوهش استفاده خواهد شد.

۲-۲- ACO

مورچه‌ها چند رفتار جمعی دارند که می‌توان با الهام از آن‌ها به حل مسائل متناسب پرداخت، یکی از رفتارهای معروف مورچه‌ها یافتن غذا است که در بهینه‌یابی بسیار پرکاربرد است و مورد استفاده است. دو رفتار دیگر مورچگان، ساخت قبرستان و مراقبت از کودکان است که در این دو رفتار اجتماعی و عملکرد مورچه‌ها، مفهوم خوشه‌بندی و دسته-بندی را نشان می‌دهد. در این روش‌ها جهت شبیه‌سازی رفتار مورچه‌های واقعی، مورچه‌های مصنوعی تعریف می‌شود و از مطرح‌ترین الگوریتم مورچه‌ها Ant-System، Ant، Min-Max Ant System، System است [۱۷].

در این پژوهش از رفتار جستجوی غذا مورچگان الهام گرفته می‌شود و از سیستم Ant Colony بهره گرفته می‌شود. در این روش این روش پارامتر اثرگذار ماده‌ای به نام فرومون است که مورچه‌های واقعی در حین حرکت بر روی مسیر خود به جای می‌گذارند [۱۸]. در مسیرهای کوتاه‌تر، مورچه‌ها راحت‌تر و سریع‌تر به غذا می‌رسند، در نتیجه در این مسیرها فرومون بیشتری ریخته شده و به دلیل حساسیت مورچه‌ها نسبت به فرومون، بیشتر جذب این مسیرها می‌شوند و در نتیجه مسیرهای مناسب‌تر به این گونه شناسایی می‌شود. رابطه اصلی این روش به صورت زیر (رابطه ۱) تعریف می‌شود.

$$P_{ij}^k(t) = \begin{cases} \frac{\tau_{ij}^\alpha(t) * \eta_{ij}^\beta(t)}{\sum_{j \in N_j^k} \tau_{ij}^\alpha(t) * \eta_{ij}^\beta(t)} & j \in N_j^k \\ 0 & j \notin N_j^k \end{cases} \quad (1)$$

در این رابطه $P_{ij}^k(t)$ احتمال انتقال مورچه k ام در زمان t از موقعیت i به موقعیت j می‌باشد. τ_{ij} چگالی فرومون متناظر حرکت از i به j است. η_{ij} میزان مطلوبیت حرکت از i به j . N_j^k مجموعه موقعیت‌های مجازی است که هنوز توسط مورچه k ام بررسی نشده است [۱۸]. از α جهت تنوع بخشی به پاسخ‌ها و β به عنوان نرخ یادگیری استفاده می‌شود، و در واقع α و β ضرایب کنترل تأثیر فرومون و مطلوبیت هستند. این فرمول تعیین می‌کند که مورچه‌ها چطور بین مسیرهای ممکن یکی را انتخاب کنند. اگر α بزرگ‌تر باشد، تأثیر فرومون بیشتر می‌شود و اگر β بزرگ‌تر باشد، اهمیت مطلوبیت افزایش می‌یابد.

موقعیت‌های قبلی که توسط مورچه بررسی شده در لیستی قرار می‌گیرد تا مجدداً به عنوان پاسخ مناسب انتخاب نشود. در نهایت کل پاسخ‌های انتخاب شده با استفاده از تابع هدف بررسی شده و اگر بهینه‌تر از مجموعه پاسخ قبلی الگوریتم بود، به عنوان پاسخ مناسب جایگزین آخرین پاسخ بهینه می‌شود.

۲-۳- RL

از زیرشاخه‌های اصلی هوش مصنوعی یادگیری ماشین است که یادگیری تقویتی روش از آن است که بر اساس توابعی برای تشویق یا تنبیه در یادگیری، به حل و بهبود حل در مسائل می‌پردازد. روش‌های یادگیری تقویتی به چند دسته اصلی تحت عناوین روش‌های مبتنی بر ارزش^۱، روش‌های مبتنی بر سیاست^۲ و روش‌های ترکیبی^۳ تقسیم می‌شوند که بخشی هم تحت عنوان یادگیری تقویتی عمیق^۴ نیز می‌توان برای آن در نظر گرفت [۱۹، ۲۰].

یادگیری تقویتی مبتنی بر ارزش، بر اساس تابع ارزش $Q(s,a)$ و استخراج سیاست بهینه از آن تعریف می‌شود و عمل می‌کند. الگوریتم‌هایی معروف این روش Q-Learning و SARSA هستند [۲۱]. در روش‌های مبتنی بر سیاست به جای یادگیری تابع ارزش $Q(s,a)$ ، مستقیماً سیاست $\pi_\theta(a|s)$ را با استفاده از گرادینت یاد می‌گیرد. در متد اصلاین روش به نام Policy Gradient امکان یادگیری

سیاست‌های تصادفی و پیوسته را دارد و عملکرد بهتری در محیط‌هایی با فضای عمل پیوسته داراست. واریانس این روش بالاست که نیاز به ترفندهایی مثل baseline برای پایداری را ایجاب می‌کند [۲۲].

در روش ترکیبی هدف ترکیب مزایای روش‌های ارزش-محور و سیاست‌محور است. در مدل اصلی این روش به نام Actor-critic، Actor به‌روزرسانی سیاست π_θ و critic تخمین ارزش Q برای راهنمایی actor را انجام می‌دهد. پایدارتر از روش Policy Gradient است و مناسب برای پیاده‌سازی با شبکه عصبی است اما این روش پیچیدگی پیاده‌سازی دارد و حساس به تنظیمات هاپیر پارامترها است [۲۳]. در روش یادگیری تقویتی عمیق از شبکه عصبی به عنوان تقریب‌گر تابع ارزش و سیاست Q و V استفاده می‌شود. این روش الگوریتم‌هایی مانند DQN و PPO دارد [۲۴].

۳- روش‌شناسی پژوهش

روش پیشنهادی پژوهش حاضر تلفیقی از الگوریتم جامعه مورچگان و یادگیری تقویتی Q-Learning است که در ادامه بخش بیان می‌شود و با توجه به پارامترهای اساسی پژوهش اعم از امدادگران و فعالیت‌ها تابعی تعریف می‌شود که در مدل پیشنهادی ACO-RL استفاده می‌شود.

۳-۱- ACO-RL

در الگوریتم جامعه مورچگان در هر تکرار مورچه‌ها برای انتخاب مسیر و انتقال از موقعیت کنونی به موقعیت بعدی از یک تابع احتمال بر اساس رابطه (۱) استفاده می‌کنند [۱۸]. مقدار فرومون روی یال‌های یک گراف طبق رابطه (۲) به‌روزرسانی می‌شود تا مسیرهای بهینه‌تر مشخص شوند.

$$\tau_{ij}(t+1) = (1-\rho) \cdot \tau_{ij}(t) + \sum_k \Delta \tau_{ij}^k \quad (2)$$

در این رابطه τ_{ij} مقدار فرومون روی یال (i,j) در زمان t است، و ρ نرخ تبخیر فرومون ($0 < \rho < 1$) و $\sum_k \Delta \tau_{ij}^k$ مقدار فرومونی است که توسط مورچه‌های مختلف اضافه می‌شود. این فرمول به الگوریتم کمک می‌کند تا مسیرهای پربازده‌تر را با افزایش مقدار فرومون، و مسیرهای ضعیف‌تر را با تبخیر فرومون شناسایی کند.

^۳ Actor-Critic

^۴ Deep Reinforcement Learning

^۱ Value-Based

^۲ Policy-Based

یکی از مواردی که در اجرای الگوریتم باید مشخص باشد، تعداد مورچه‌ها است که برابر با تعداد امدادگران است. در ابتدای اجرای الگوریتم مورچه‌ها به صورت تصادفی در محیط فرض می‌شوند. برای تنظیم فرومون اولیه از یک مقدار تصادفی در بازه $[0, \tau_0]$ استفاده می‌شود. از یادگیری تقویتی برای بهبود عملکرد الگوریتم ACO که خود به نوعی از ساختار یادگیری تقویتی (فرومون) بهره می‌برد، استفاده می‌شود. یک رویکرد رایج برای این کار، استفاده از تابع Q-Learning است تابع به‌روزرسانی آن به صورت رابطه (۳) تعریف می‌شود.

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (3)$$

$Q(s, a)$ مقدار ارزش (Q-value) مربوط به انجام عمل a در حالت s است. α نرخ یادگیری و r پاداش دریافت‌شده، γ ضریب کاهش ارزش پاداش در حالت بعدی است و طبق بررسی‌های به عمل آمده α برابر با ۰.۱ و γ برابر با ۰.۹۵ لحاظ می‌شود [۱۹، ۲۰]. این فرمول به مورچه‌ها کمک می‌کند تا تصمیم‌گیری بهتری داشته باشند. در واقع، مسیرهایی که پاداش بالاتری دارند، ارزش بیشتری می‌گیرند و به مرور زمان یادگیری بهینه‌ای از مسیرها اتفاق می‌افتد.

۳-۱-۱- ارتباط Q-Learning و فرومون در ACO:

روش به‌روزرسانی فرومون با استفاده از Q-Learning به صورت رابطه ۴ تعریف می‌شود [۲۵].

$$\tau_{(i+1)(j+1)} = (1 - \rho)\tau_{ij} + \alpha Q(s_i, a_i) \quad (4)$$

این رابطه مقدار فرومون را با توجه به ارزش یادگیری Q تنظیم می‌کند. مسیرهایی که Q-value بالاتری دارند، فرومون بیشتری دریافت می‌کنند و در نتیجه احتمال انتخاب آن‌ها توسط سایر مورچه‌ها افزایش می‌یابد. و در واقع ترکیب ACO و یادگیری تقویتی باعث می‌شود که مورچه‌ها از تجربیات قبلی خود یاد بگیرند و مسیرهای بهینه‌تر، سریع‌تر و پایدارتر شناسایی کنند و تنظیم فرومون بهینه‌تر صورت پذیرد. و همچنین پاداش‌ها باعث هدایت بهتر مسیرها می‌شوند و این ترکیب باعث بهبود عملکرد ACO در حل مسئله حاضر می‌شود.

۳-۲- توابع ریاضی

الگوریتم جامعه مورچه‌ها تابع هدف دارد که سعی می‌کند آن را کمینه کند تا به بهترین نتیجه برسد، و در واقع

تابع هزینه نامیده می‌شود. این تابع بر اساس پارامترهای مهم پژوهش شامل امدادگران، ساختار تیم‌ها و فعالیت‌ها تعریف می‌شود. همچنین الگوریتم یادگیری تقویتی نیز تابعی به نام پاداش دارد که با در نظرگیری شرایط لایق پاداش تخصیص و برنامه‌ریزی فعالیت‌ها تعریف می‌شود.

۳-۲-۱- تابع هزینه

با بررسی‌های انجام شده مشخص شد پارامترهای مساحت محل جستجو، میزان خسارت وارده، مدت زمان انجام فعالیت، فاصله نجاتگران تا پلاک ساختمانی، سرعت حرکت نجاتگران در انجام وظیفه محوله بسیار اهمیت دارد. رابطه‌ای احتمالی برای موفقیت امدادگران به صورت تلفیقی از احتمال وجود سوژه در محل و احتمال کاوش تعریف می‌شود [۱۴، ۲۶]. با توجه به شرایط مسئله تابع هزینه احتمال میزان موفق نشدن یک امدادگر در انجام فعالیت محوله می‌شود که باید این میزان توسط مدل به حداقل برسد و مطابق رابطه ۵ تعریف می‌شود.

$$Cost = 1 - (1 - e^{-Area_i * Distance_i * Speed * Time}) \quad (5)$$

در رابطه ۵، $Area_i$ مساحت ناحیه‌ای که فعالیت در آن قرار دارد نسبت به مساحت کل پلاک‌های ساختمانی محدوده مطالعاتی است. $Distance$ فاصله امدادگر تا محل فعالیت، $Time$ زمان انجام هر فعالیت، $Speed$ سرعت حرکت امدادگر است که در این پژوهش به طور میانگین و معمول حدود ۰/۲۳ متر بر ثانیه در نظر گرفته شده است [۲۷].

۳-۲-۲- تابع پاداش

با توجه به اهمیت وجود تابع پاداش در یادگیری تقویتی تصمیم بر آن شد تخصیص‌هایی که رسیدگی به مجروحان بیشتر دارد امتیاز بیشتر و در واقع پاداش بیشتری دریافت کند. در رابطه (۶) تابع پاداش بر اساس پاداش‌دهی هنگام تخصیص وظایف و فعالیت‌ها به امدادگر، تعیین شده است که در آن $MaxInjured_i$ ، تعداد کل مجروحان و $MaxInjured$ تعداد مجروحان در هر سازه مسکونی می‌باشد.

$$Reward = (MaxInjured_i / MaxInjured) \quad (6)$$

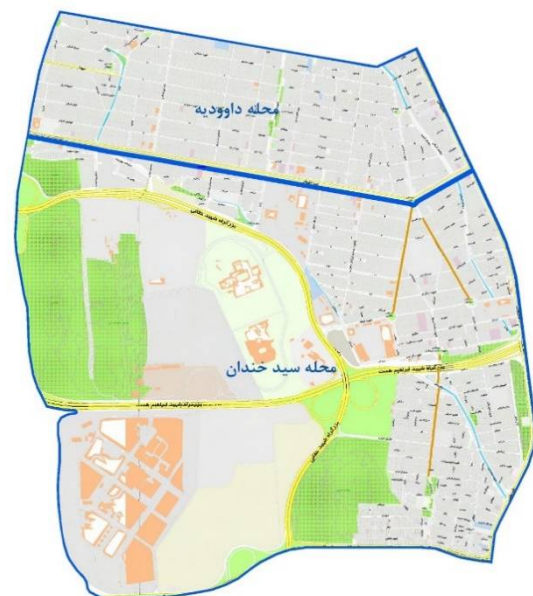
۴- پیاده‌سازی و یافته‌های پژوهش

در این بخش داده‌های ورودی پژوهش وارد نرم‌افزار می‌شود تا اجرای مدل پیشنهادی روی آن پیاده‌سازی شود. ابتدا

محدوده مطالعاتی و داده‌های اولیه که بافت‌های خسارات سازه‌ای و انسانی می‌باشد، مشخص می‌شود. داده‌های دیگری که لازم است وارد مدل پیشنهادی شود جداول مربوط به امدادگران و فعالیت‌ها است. در ادامه نیز به بیان و نمایش پیاده‌سازی مدل پیشنهادی پرداخته می‌شود.

۴-۱- محدوده مطالعاتی

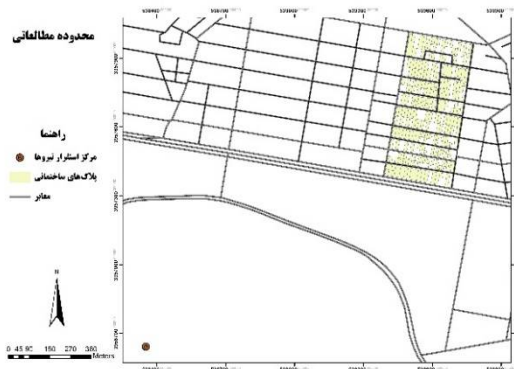
محدوده مطالعاتی بخشی از محدوده میانی منطقه ۳ شهرداری تهران می‌باشد. منطقه ۳ واقع در بخش شمالی شهر، با وسعتی حدود ۳/۲۹۴۵ هکتار، یکی از مناطق مهم و پرجمعیت تهران محسوب می‌شود. این منطقه از شمال به بزرگراه‌های چمران، مدرس و صدر، از شرق به خیابان پاسداران و بخشی از خیابان شریعتی، از جنوب به بزرگراه‌های سلیمانی و همت و از غرب به بزرگراه چمران محدود می‌شود. از نظر موقعیت جغرافیایی این منطقه به عنوان پل ارتباطی بین مناطق شمالی و مرکزی تهران عمل می‌کند و دسترسی‌های مناسبی به سایر نقاط شهر دارد. همچنین دارای بافتی متنوع از بافت‌های مسکونی، تجاری و اداری است و از نظر امکانات رفاهی، تفریحی، فرهنگی و درمانی در سطح بالایی قرار دارد. این منطقه شامل محله‌های مانند ونک، قلهک، دروس، اختیاریه، سید خندان، میرداماد و یوسف آباد است که دسترسی مناسبی به سایر مناطق تهران فراهم است. ترافیک سنگین از نقاط ضعف این منطقه به شمار می‌رود.



شکل ۱- ناحیه ۴ منطقه ۳ شهر تهران (شهرداری تهران)

منطقه ۳ از شمال با منطقه ۱، از شرق با منطقه ۴، از غرب با منطقه ۲ و از جنوب با مناطق ۶ و ۷ هم مرز و همجوار است. این منطقه در حال حاضر دارای ۱۲ محله می‌باشد و برطبق اطلاعات مقدماتی سرشماری عمومی نفوس و مسکن در سال ۱۳۹۵ جمعیت منطقه ۳، برابر ۳۳۰۰۰۴ نفر می‌باشد. ۱۲ محله منطقه ۳ در ۶ ناحیه قرار گرفته است که محدوده مطالعاتی پژوهش حاضر در تقسیم بندی ناحیه چهارم قرار می‌گیرد [۲۹، ۲۸]. ناحیه ۴ شامل دو محله داوودیه و سیدخندان می‌شود که در شکل (۱) نشان داده شده است [۲۹].

محله داوودیه که یکی از محله‌های منطقه ۳ بشمار می‌آید که در ناحیه ۴ تقسیم‌بندی آن قرار گرفته است و آن را به محدوده خیابان میرداماد و میدان مادر می‌شناسند. محدوده میدان مادر فعلی و محسنی سابق از مراکز مهم تجاری و اداری منطقه سه و شهر تهران محسوب می‌شود. این محله از شمال به خیابان شهید دستگردی، از جنوب به اتوبان حقانی، از شرق به خیابان شریعتی و از غرب به اتوبان مدرس منتهی می‌گردد. محدوده مطالعه پژوهش حاضر بخشی از قسمت میانی این محله، در نزدیکی میدان مادر می‌باشد که در شکل (۲) نشان داده شده است.



شکل ۲- محدوده مطالعاتی و محل استقرار نیروها

محدوده مطالعاتی بافتی مسکونی و تجاری دارد، بخش تجاری عموماً محدود به جوار خیابان میرداماد است و عمده سازه‌های دیگر محدوده مسکونی می‌باشد. این بخش از شهر تهران شامل ۲۵۵ پلاک ساختمانی است و اطلاعات خسارات سازه‌ای و انسانی زلزله‌ای ۶ ریشتری برای آن از پیش محاسبه و به عنوان داده ورودی منظور شده است. ۷۰ امدادگر بر اساس مدل پیشنهادی و پاراثرهای شناسایی شده، در آن به انجام وظایف محوله خواهند پرداخت. امدادگران از مرکز مدیریت بحران نزدیک محدوده مطالعاتی اعزام خواهند شد.

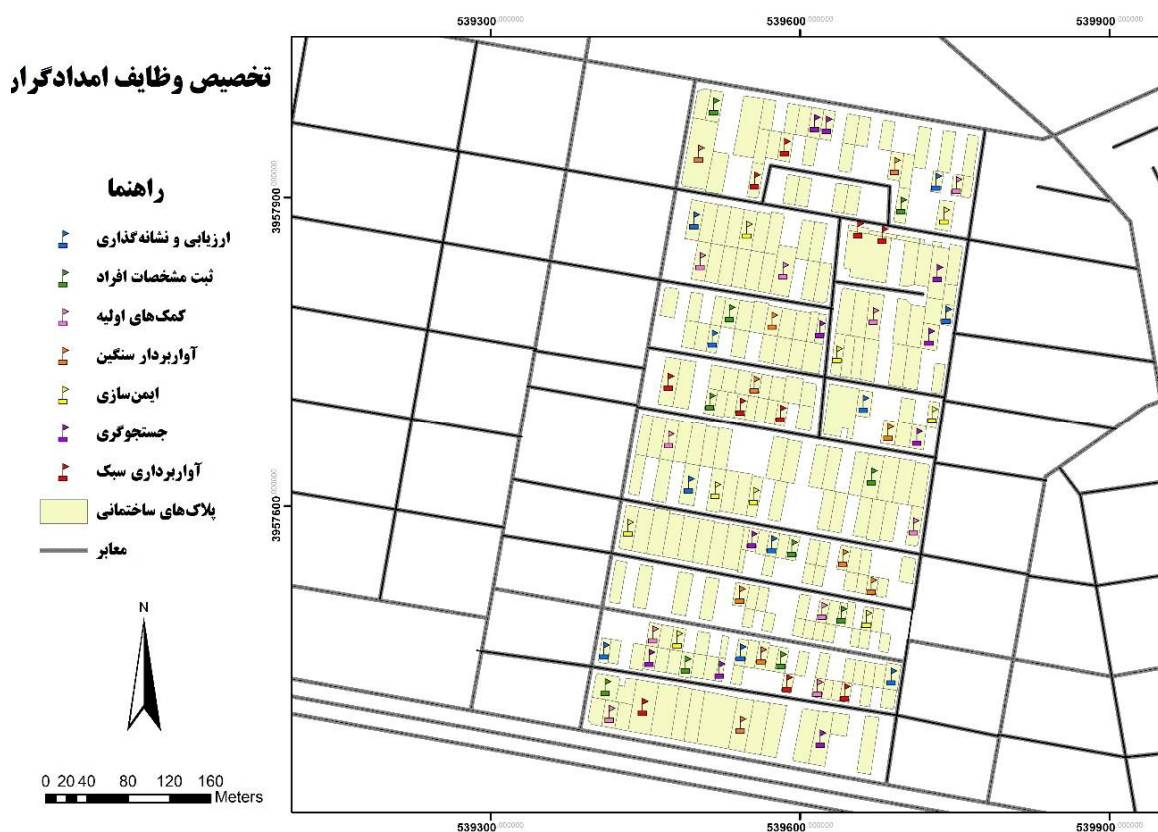
۴-۲- اجرای مدل پیشنهادی

با استفاده از برنامه‌نویسی پایتون مدل پیشنهادی که روابط ریاضی و طراحی آن در بخش‌های قبل بیان شد پیاده‌سازی می‌شود. مدل به این صورت است که الگوریتم تقویتی جامعه مورچگان با الگوریتم Q-Learning که به صورت ضریب به‌روزرسانی فرومون، ترکیب و تلفیق شده و با توابع هزینه و پاداش یادگیری که با استفاده از پارامترهای اثرگذار طراحی شده‌اند، وارد بخش محاسباتی می‌گردند. در ادامه مدل پیشنهادی با داده‌های ورودی مشخصات امدادگران و فعالیت‌ها و میزان خسارات هر سازه اجرا و نتایج محاسبه می‌شود.

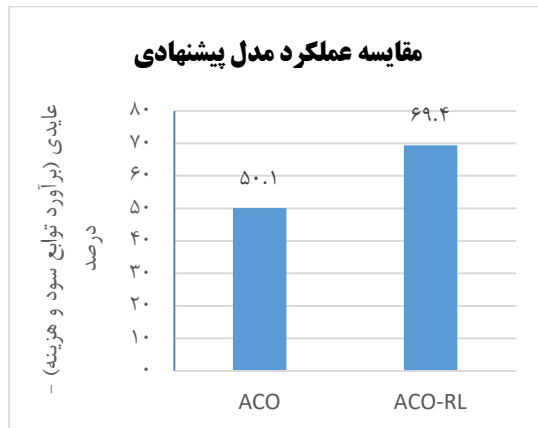
خروجی مدل فعالیت و مکان فعالیت مناسب هر امدادگر را مشخص می‌کند، به گونه‌ای که بهترین چیدمان تیم و بیشترین سود و کمترین هزینه برای تیم به ارمغان آید و در نتیجه بهترین عملکرد تیم ایجاد شود. اینگونه الگوریتم‌ها

تمامی حالات را محاسبه نمی‌کنند و در واقع عملکرد آن‌ها به این صورت است که به شکلی بهینه به پاسخ مناسب برسند. در این راستا عموماً ابتدا چیدمانی به صورت تصادفی ایجاد شده و در ادامه سعی می‌شود با کمک پارامترهای روابط ریاضی تدوین شده در زمانی نسبتاً کوتاه به نتیجه مطلوب دست یابد.

با توجه به منابع معتبر مطالعه شده برای پارامتر نرخ یادگیری (α) میزان ۰.۱ منظور شد و برای پارامتر ضریب کاهش یا تعدیل (γ) ۰.۹۵ لحاظ شد و تعداد تکرار به صورت معمول ۱۰۰ در نظر گرفته شد [۱۷، ۱۸، ۳۰]. حل مسئله مطروحه در این پژوهش با روش‌های دقیق امکان‌پذیر نیست بنابراین می‌بایست از روش‌هایی مانند مدل پیشنهادی که سعی می‌کند در زمانی مناسب به پاسخی بهینه و کارآمد دست یابد، استفاده شود. در شکل (۳) خروجی اجرای مدل پیشنهادی بر روی داده‌های ورودی محاسبه شده و در محیط GIS نمایش داده شده است.



شکل ۳- تخصیص و برنامه‌ریزی انجام‌شده برای امدادگران توسط مدل پیشنهادی



نمودار ۱- مقایسه عملکرد مدل پیشنهادی

پژوهش حاضر قصد داشته است تا فعالیت تیم‌های نجاتگر در رخداد زمین‌لرزه را به شکلی نوین طوری مدل-سازی کند که عملکرد و کارایی نهایی آن‌ها افزایش یابد. جهت نیل به هدف فوق، بررسی‌های گسترده‌ای در ارتباط با مدل‌هایی که می‌توانستند در زمینه‌های یاد شده به کار روند صورت پذیرفت. همچنین از تجربیات افراد خبره و تحقیقات محقق‌های دیگر استفاده شد تا علاوه بر ارائه مدلی جدید در این زمینه، از غیرکاربرد و نامناسب بودن آن جلوگیری شود. بدین منظور مدل‌های مختلفی هوش مصنوعی مطالعه و الگوریتم جامعه مورچگان و یادگیری تقویتی به دلیل قابلیت مدل‌سازی و پیاده‌سازی موضوع مناسب شناخته شد. دلیل این مهم این است که برای مدل نمودن بهتر ساختار تیم‌های عملیاتی نیاز به الگوریتمی بود که هم ساختار تیمی داشته باشد هم سازگاری مناسبی با الگوریتم یادگیری تقویتی داشته باشد. در این راستا الگوریتم جامعه مورچگان به دلیل ساختار تیمی و تقویتی وجود فرومون مناسب ترکیب با یادگیری تقویتی شناخته شد.

همچنین پس از مطالعه و بررسی انواع مدل‌های و ایجاد مدل پیشنهادی و اجرای آن در محدوده مطالعاتی، منجر به حصول نتایج ذیل نیز شد:

- تلفیق الگوریتم جامعه مورچگان و یادگیری تقویتی به نوعی باعث بهبود و تعمیق یادگیری در مسأله حاضر شد که ارزیابی کمی نتایج حاصله نیز گواهی بر آن می‌تواند باشد.
- امکان استفاده از یادگیری تقویتی Q-Learning به دلیل کارایی در محیط‌های گسسته و محدود، در تلفیق با الگوریتم جامعه مورچگان و انجام تخصیص فعالیت‌ها به نجاتگران می‌تواند خوب عمل کند.

همانطور که از شکل (۳) مشخص است و قبل‌تر نیز بیان شد، ۷ نوع فعالیت برای امدادگران در نظر گرفته شده است. همچنین طبق مطالب ذکر شده در بخش‌های پیشین، تیم‌ها پنج نفره هستند و همه امدادگران تمامی آموزش‌ها و تخصص‌های لازم را دیده‌اند. در نتیجه با توجه به تعداد کل نجاتگران که ۷۰ نفر می‌باشد، پس از تخصیص برای هر نوع فعالیت، ۱۰ نجاتگر برای انجام آن فعالیت تخصیص یافته‌اند. در پژوهش حاضر سعی شده تمامی جایگشت‌های تخصیصی افراد درون تیم‌ها در محاسبات در نظر گرفته شود.

با توجه به بررسی نتایج به دست آمده سازه‌های با تخریب بیشتر برای فعالیت آواربرداری اولویت‌دارتر بوده‌اند و سازه با آسیب انسانی بیشتر برای کمک‌های اولیه اولویت‌دارتر بودند. مکان‌هایی که افراد بیشتری در خود جای داده بودند و تخریب شده‌اند نیازمند جستجو و زنده‌یابی هستند. در واقع تخصیص افراد با توجه به اینکه مکان‌های ساختمانی که میزان خسارت بیشتر دارند در اولویت امداد رسانی هستند، انجام می‌شود.

۵- نتیجه‌گیری و پیشنهادات

در راستای بررسی عملکرد مدل پیشنهادی مقایسه‌ای بین استفاده از تک مدل الگوریتم جامعه مورچگان و مدل پیشنهادی تلفیقی از الگوریتم جامعه مورچگان و یادگیری تقویتی انجام شده است. مقایسه عملکرد با استفاده از مقادیر محاسبه شده مدل‌ها برای برآورد توابع سود و هزینه که با عنایت به مدل جهانی احتمال موفقیت در امداد و نجات طراحی شده‌اند، تخصیص بهینه در هر دو حالت انجام شده است. مدل پیشنهادی نسبت به حالتی که تنها از یک متد جامعه مورچگان استفاده شود، بهبود قابل قبولی داشته است.

با توجه به اینکه برآورد عایدی (برآورد توابع سود و هزینه) حاصله از اجرای مدل پیشنهادی حدود ۶۹/۴ درصد و نیز برآورد سود حاصله استفاده تنها از الگوریتم جامعه مورچگان برای پژوهش حاضر ۵۰/۱ درصد محاسبه شده است، در نهایت افزایش ۱۹ درصدی بهره‌وری مدل پیشنهادی حاصل شده است. در نتیجه برآورد حاصله می‌تواند برای عملکرد مناسب تیم‌های امدادی پس از زمین‌لرزه کارا و سودمند باشد و جان انسان‌های زیادی را نجات دهد، از این رو عملکرد مدل خوب ارزیابی می‌شود. مقایسه به صورت نموداری از میزان اثرگذاری مدل پیشنهادی با نمودار ۱ نمایش داده شده است.

داراست و فقط نیازمند ویژه سازی توابع هزینه و پاداش با توجه به پارامترهای مؤثر است. دستاورد دیگر ارائه مدلی حداقل امکان منطبق بر واقعیت و اطلاعات معتبر و داده‌ها به روز است که این موضوع منجر به کاربردی شدن مدل پیشنهادی می‌شود. و همچنین مدل پیشنهادی در یک محیط شهری که احتمال رخداد زمین‌لرزه‌ای مخرب را دارد، پیاده‌سازی شده است که این موضوع کاربردی ساختن مدل پیشنهادی را بیشتر می‌کند.

پیشنهاد می‌شود در پژوهش‌های آتی، کارایی دیگر الگوریتم‌های بهینه‌سازی و هوشمندی جهت مدل‌سازی فرآیند امداد و نجات در زمین‌لرزه بررسی و ارزیابی شود. پیشنهاد دیگر، استفاده و ارزیابی مدل پیشنهادی در امداد و نجات سایر بحران‌ها مانند سیل و آتش‌سوزی است. و پیشنهاد آخر بررسی دیگر روش‌های آماری برای مقایسه و ارزیابی مدل پیشنهادی در موضوع پژوهش است.

• الگوریتم Actor-Critic، Policy Gradient، PPO برای محیط‌های پیوسته مناسب است و مناسب محیط گسسته پژوهش حاضر نمی‌باشد.

• الگوریتم DQN که یک روش از زیرمجموعه یادگیری عمیق در یادگیری تقویتی است به دلیل محیط عملکردی گسسته آن می‌تواند گزینه خوبی باشد اما به دلیل عملکرد مناسب در محیط‌ها با داده انبوه به نظر چندان مقرون به صرفه در مسئله پژوهش حاضر نمی‌باشد اما می‌تواند به عنوان پیشنهاد برای پژوهش‌های آتی به شرط دارا بودن محدوده مطالعاتی وسیع مورد بررسی و ارزیابی قرار گیرد.

در پژوهش انجام شده تلفیق الگوریتم جامعه مورچگان با یادگیری تقویتی برای جستجو و نجات در زمین‌لرزه نوآوری - ای بوده که پیشتر انجام نشده است. همچنین مدل ارائه شده نگرشی کاربردی برای راهبری گروه‌های کاری دیگر را

مراجع

- [۱] Vafaeinezhad, A.R. Alesheikh, A.A. Hamrah, M. Reza Nourjou, R. Shad, R. (2009). "Using GIS to Develop an Efficient Spatio-temporal Task Allocation Algorithm to Human Groups in an Entirely Dynamic Environment Case Study: Earthquake Rescue Teams". Computational Science and Its Applications – ICCSA 2009. Volume 5592 of the series Lecture Notes in Computer Science. pp 66-78. 2009.
- [۲] Vafaeinezhad, A. R, Alesheikh, A., Nouri, J., (2010). "Developing a spatio-temporal model of risk management for earthquake life detection rescue team". International Journal of Environmental Science & Technology. March 2010, Volume 7, Issue 2, pp 243–250.
- [۳] Rasekh, A., Vafaeinezhad, A. R. (2012), "Developing a GIS Based Decision Support System for Resource Allocation in Earthquake Search and Rescue Operation", Computational Science and Its Applications – ICCSA 2012, Volume 7334 of the series Lecture Notes in Computer Science PP. 275-285. Springer
- [۴] Jaziar Radianti, Ole-Christoffer Granmo, Parvaneh Sarshar, Morten Goodwin, Julie Dugdale, Jose J. Gonzalez, 2015: A spatio-temporal probabilistic model of hazard- and crowd dynamics for evacuation planning in disasters. Applied Intelligence. January 2015. Volume 42. Issue 1. pp 3–23. Springer
- [۵] Yashwant Singh Patel, Sourasekhar Banerjee1, Sajal K., 2020, Low-Latency Energy-Efficient Cyber-Physical Disaster System Using Edge Deep Learning. Jan 4, 2020.
- [۶] Li Yan, Shohaib Mahmud, Haiying Shen, Natasha Zhang Foutz and Joshua Anton, 2020, MobiRescue: Reinforcement Learning based Rescue Team Dispatching in a Flooding Disaster, 2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS)|978-1-7281-7002-2/20/\$31.00©2020IEEE|DOI: 10.1109/ICDCS47774.2020.00033
- [۷] Jalumuru Nalini, P. M. Khilar, 2021. Reinforced Ant Colony Optimization for Fault Tolerant Task Allocation in Cloud Environments, Wireless Personal Communications (2021) 121:2441–2459, <https://doi.org/10.1007/s11277-021-08830-4>
- [۸] Zhaojia Tang, Ping Wang, Yong Wang, Changgeng Wang, Yu Han, 2022. Distributed Small-Step Path Planning and Detection Method for Post-earthquake Robot to Inspect and Evaluate Building Damage. Published in Frontiers in Neurobotics 15 August 2022, Engineering, Environmental Science.
- [۹] Xiaoyan Li, Xuedong Liang, Xia Wang, Rong Wang, Lingli Shu, Wentao Xu, 2023, Deep reinforcement learning for optimal rescue path planning in uncertain and complex urban pluvial flood scenarios, Applied Soft Computing 144 (2023) 110543

- [۱۰] Rahat Tamzid, Ehfaz Faisal Mahee, Ashikur Rahman, Tahira Akhtar Tamanna, Muntasir Rahman, Md Mamun Hossain, 2024. Intelligent Route Planning for Post-Earthquake Rescues Using IoT. Engineering, Environmental Science, Computer Science, May 2024.
- [۱۱] Minal Shahakar, S. A. Mahajan, Lalit Patil, 2024, Optimizing System Resources and Adaptive Load Balancing Framework Leveraging ACO and Reinforcement Learning Algorithms, J. Electrical Systems 20-1s (2024): 244 – 256
- [۱۲] Jin Zhang, Hao Xu, Ding Liu and Qi Yu, 2025, Advancing Dynamic Emergency Route Optimization with a Composite Network Deep Reinforcement Learning Model, Systems 2025, 13, 127 <https://doi.org/10.3390/systems13020127>
- [۱۳] N. Bahrami¹, M. Argany^{*,1}, A.D. Bolorani¹, A.R. Vafaeinejad², 2025. A Geo-AI Approach in the Post-Earthquake Rescue. J. RS. GEOINF. RES. 3(1): 187-200, Winter & Spring 2025
- [۱۴] Mezyabadi, Shahryar. 2012. Search and rescue in collapsed buildings (rubble removal). Tehran: Nakhl [In Persian].
- [۱۵] Comprehensive National Rescue and Relief Plan. 2003
- [۱۶] Sharifi Sadeh, Mehrab. 2010. Disaster Assessment. Scientific Quarterly Journal of Rescue and Relief. Year 2. Issue 4 [In Persian].
- [۱۷] N. Bahrami, M. Argany, N. Neysani Samani, A.R. Vafaei Nejad, 2021. Designing a Context-aware Recommender System in the Optimization of the Relief and Rescue by Ant Colony Optimization Algorithm and Geospatial Information System, Journal of Geomatics Science and Technology 11 (2), 153-162.
- [۱۸] N Bahrami, M Argany, NN Samani, AR Vafaeinejad, Designing a context-aware recommender system in the optimization of the relief and rescue. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Volume 42, Pages 171-177.
- [۱۹] Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press. [ISBN: 978-0262039246].
- [۲۰] Watkins, C.J.C.H. and Dayan, P. (1992) Q-Learning. Machine Learning, 8, 279-292.
- [۲۱] RS Sutton, S Singh, D McAllester, (2000). Comparing policy-gradient algorithms. IEEE Transactions on Systems, Man, and Cybernetics
- [۲۲] Konda, V., Tsitsiklis, J. N., (2000). Actor-critic algorithms. Society for Industrial and Applied Mathematics
- [۲۳] Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. Nature. [DOI: 10.1038/nature14236]
- [۲۴] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, Oleg Klimov, (2017). Proximal Policy Optimization Algorithms. arXiv:1707.06347
- [۲۵] Minal Shahakar, S. A. Mahajan, Lalit Patil, 2024. Enhancing Resource Utilization and Load Distribution with ACO and Reinforcement Learning in Dynamic Computing Infrastructures. Panamerican Mathematical Journal ISSN: 1064-9735 Vol 34 No. 1 (2024)
- [۲۶] Kenneth B. Chiacchia, Heather E. Houlahan, (2023). Deriving Objective Probability of Detection for Missing Person Search: Validating Use of Effective Sweep Width and Associated Mathematical Models. Journal of Search & Rescue Volume 6, Issue 1
- [۲۷] N Bahrami, M Argany, M Jelokhani Neyaraki, A Vafaeinezhad, Providing a spatial approach in the rescue and relief management after the earthquake, Environmental Management Hazards 6 (2), 117-129.
- [۲۸] Official website of District 3 of Tehran Municipality. <https://region3.tehran.ir/>
- [۲۹] Tehran City Crisis Prevention and Management Organization, Deputy for Prevention and Risk Reduction. 2008. Rapid Earthquake Damage and Casualty Estimation System in Tehran City [In Persian].
- [۳۰] N. Bahrami, 2019. Using Tabu Search Algorithm and Geospatial Information System for Managing of the Relief and Rescue Teams, Journal of Geomatics Science and Technology 8 (3), 179-18.