

به نام خدا

موضوع

مدل سازی قیمت گذاری رقابتی پویا پارکینگ شهری با رویکرد یادگیری تقویتی
عمیق چندعاملی

نام درس

فرآیند های احتمالی

نام دانشجو

مصطفی کرمی

استاد

دکتر فریبرز جولای

تاریخ

شهریور ماه 1404

چکیده

مدیریت بهینه منابع محدود پارکینگ در محیط‌های شهری متراکم، مستلزم رویکردهایی است که همزمان بتوانند نرخ اشغال مطلوب، سودآوری پایدار و رضایت کاربران را تأمین کنند. این پژوهش یک چارچوب یادگیری تقویتی چندعاملی (MARL) مبتنی بر مدل لاجیت چندگانه (MNL) ارائه می‌کند که رفتار رقابتی اپراتورهای پارکینگ را در یک محیط غیرایستا مدل‌سازی کرده و تخصیص تقاضا را به‌طور درون‌زا بر اساس قیمت و شاخص‌های اشغال پیش‌بینی می‌نماید. محیط شبیه‌سازی به‌صورت فرآیند تصمیم‌گیری مارکوف طراحی شده و تابع پاداش ترکیبی شامل حداکثرسازی سود، نزدیکی به نرخ اشغال هدف و کاهش نرخ رد تقاضا تدوین گردیده است. برای آموزش عوامل، الگوریتم DQN با معماری شبکه عصبی چندلایه و روش تجربه‌تکرار به‌کار گرفته شد. نتایج شبیه‌سازی در سه سناریو (تقاضای پایه، دو برابر شدن تقاضای پایه، و دو برابر شدن حساسیت قیمتی) نشان داد که چارچوب پیشنهادی قادر به یادگیری سیاست‌های کارآمد و تطبیق‌پذیر با شرایط بازار است. در شرایط تقاضای بالا، پارکینگ‌های با ظرفیت بیشتر حداکثر سود را کسب کردند، در حالی که در وضعیت حساسیت قیمتی بالا، اپراتورهای با قیمت پایین توانستند سهم بازار بیشتری به‌دست آورند. مقایسه با مدل‌های تک‌عاملی موجود، برتری چارچوب حاضر را در محیط رقابتی و پویا از نظر سودآوری و کنترل اشغال اثبات کرد. این رویکرد در شهرهایی با زیرساخت هوشمند (حسگرهای اشغال، پرداخت دیجیتال و مدیریت متمرکز) قابلیت پیاده‌سازی دارد و می‌تواند به‌عنوان یک ابزار تصمیم‌یار برای سیاست‌گذاران در جهت توسعه سیستم‌های حمل‌ونقل پایدار مورد استفاده قرار گیرد.

واژگان کلیدی: قیمت‌گذاری پویا، یادگیری تقویتی چندعاملی، پارکینگ شهری، مدل لاجیت چندگانه، محیط غیرایستا

مقدمه

بازار پارکینگ های شهری

پارکینگ های شهری، به ویژه در مناطق پرتراکم، فراتر از یک زیرساخت خدمات عمومی، دارایی های درآمدزایی با ارزش اقتصادی قابل توجه محسوب می شوند. هر جای پارک نمایانگر یک واحد عرضه در بازاری محدود و ارزشمند است که میزان تقاضای آن به عواملی همچون موقعیت مکانی، دسترسی پذیری، و ساختار قیمتی بستگی دارد. برای مالکان و مدیران، بیشینه سازی سود نه تنها مستلزم مدیریت کارای ظرفیت اشغال است، بلکه نیازمند توانایی واکنش به نوسانات بازار و تنظیم بهینه نرخ ها در شرایط متغیر تقاضا و رقابت نیز می باشد.

در بسیاری از شهرها، چندین بهره بردار در فاصله های مکانی نزدیک فعالیت می کنند و این همجواری، بازاری رقابتی و پویا ایجاد می کند که در آن تصمیم های قیمت گذاری یک بازیگر می تواند سهم بازار و سودآوری رقبا را به طور مستقیم تحت تأثیر قرار دهد. در چنین محیطی، سیاست های ایستا و نرخ های ثابت غالباً منجر به استفاده غیربهره ای از ظرفیت – شامل اشباع بیش از حد یا ظرفیت خالی – شده و فرصت های درآمدی قابل توجهی را از بین می برند. بررسی ظرفیت و پتانسیل درآمدی بازار پارکینگ در مقیاس شهری نشان می دهد که مدیریت هوشمند نرخ ها و بهینه سازی اشغال ساعتی می تواند سودآوری و کارایی این بازار را به شکل معناداری ارتقا دهد، به ویژه زمانی که پارکینگ ها به عنوان اجزای یک سامانه تعامل پذیر و نه صرفاً واحدهای منفرد در نظر گرفته شوند.

اهمیت و تاثیر قیمت گذاری

قیمت گذاری، مؤثرترین ابزار مدیریتی برای کنترل تقاضا و بهبود عملکرد مالی پارکینگ های شهری است. تعیین نرخ صرفاً یک اقدام مالیاتی یا حسابداری نیست، بلکه استراتژی ای است که می تواند بر اشغال، نرخ گردش مشتریان، و بازده کل سرمایه گذاری تأثیر بگذارد. نرخ های بیش از حد بالا ممکن است بخشی از تقاضا را به سمت رقبا یا پارک های خیابانی سوق دهد، در حالی که نرخ های بسیار پایین می توانند به ازدحام، کاهش درآمد به ازای هر واحد ظرفیت و اتلاف منابع منجر شوند.

در بازار رقابتی، تغییرات حتی جزئی در قیمت می تواند موجب جابه جایی سریع تقاضا میان پارکینگ ها شود و اثرات قابل توجهی بر جریان درآمد داشته باشد. این پویایی، مدیریت قیمت را به فرایندی حساس و زمان محور بدل می کند که نیازمند پاسخگویی لحظه ای به نوسانات تقاضا و اقدامات رقابتی است. مطالعات تجربی و مدل سازی نشان داده اند که سیاست های قیمت گذاری هوشمند و داده محور، نسبت به سیاست های ثابت، می توانند رشد درآمد چند درصدی – و گاه دورقمی – ایجاد کنند؛ رقمی که در مقیاس بازار شهری به سودهای قابل ملاحظه ای منجر می شود.

قیمت گذاری پویا و ویژگی های مثبت آن و تفاوت با قیمت ثابت

رویکرد قیمت گذاری ثابت، اگرچه از نظر اجرایی ساده و کم هزینه است، اما فاقد انعطاف لازم برای پاسخ به تغییرات سریع تقاضا، شرایط رقابتی یا رویدادهای غیرمنتظره شهری است. این محدودیت باعث می شود در ساعات اوج، ظرفیت به سرعت پر شود و در ساعات کم تقاضا بخشی از ظرفیت بلااستفاده بماند. در مقابل، قیمت گذاری پویا نرخ ها را متناسب با داده های لحظه ای یا دوره ای بازار – شامل میزان اشغال، جریان ورود و خروج، وضعیت رقبا و حتی شرایط آب و هوایی یا رویدادهای خاص – به روزرسانی می کند.

این سازوکار تطبیقی بهره‌برداری بهینه از ظرفیت را تسهیل کرده و امکان حداکثرسازی درآمد را در بازه‌های زمانی متغیر فراهم می‌آورد. در پژوهش حاضر، این انعطاف‌پذیری از طریق پیاده‌سازی یادگیری تقویتی عمیق چندعامله (Multi-Agent Deep Q-Network) محقق شده است، به گونه‌ای که هر پارکینگ به‌مثابه یک عامل مستقل، با در نظر گرفتن تعاملات رقابتی و پاسخ‌های بازار، نرخ بهینه خود را تعیین می‌کند.

ویژگی‌های به خصوص مسئله پارکینگ‌های شهری، مانند ظرفیت و فاصله و تاثیر آن بر تقاضا

بازار پارکینگ شهری دارای ویژگی‌های منحصربه‌فردی است که آن را از دیگر بازارهای خدماتی با ظرفیت محدود متمایز می‌کند:

- محدودیت ظرفیت – ظرفیت هر پارکینگ در کوتاه‌مدت ثابت است؛ بنابراین سیاست قیمت باید علاوه بر جذب مشتری، از اشباع ناگهانی و ازدحام جلوگیری کند.
- موقعیت مکانی – فاصله از مراکز تجاری، اداری یا رویدادهای شهری مستقیماً بر کشش قیمتی تقاضا اثر می‌گذارد. مکان‌های نزدیک‌تر ظرفیت اعمال نرخ‌های بالاتر را دارند، در حالی که مکان‌های دورتر اغلب ناچار به رقابت قیمت یا ارائه خدمات تکمیلی هستند.
- تعاملات رقابتی – قیمت‌گذاری یک پارکینگ نه تنها بر مشتریان خود، بلکه بر تصمیمات و تقاضای رقبا نیز اثرگذار است. این تعاملات، ماهیت مسئله را به یک بازی پویا میان چند عامل تبدیل می‌کند.

این ویژگی‌ها در شبیه‌سازی حاضر به صورت صریح مدلسازی شده‌اند تا نرخ‌گذاری بتواند به‌طور هم‌زمان به شرایط ظرفیت داخلی، موقعیت نسبی و رفتار رقبا واکنش نشان دهد.

اشاره به تحقیقات سابق

ادبیات موجود طیف گسترده‌ای از رویکردها را پوشش که در ادامه به آن می‌پردازیم

- روش‌های سنتی شامل مدل‌های آماری و اقتصادسنجی برای برآورد کشش قیمتی و ارزیابی اثر سیاست‌هایی چون افزایش نرخ یا کاهش عرضه.
- مداخلات ساختاری همچون برنامه‌ریزی مکانی بهینه، نظام‌های رزرو سهمیه‌بندی‌شده، و ترکیب سیاست قیمت‌گذاری با تقویت حمل‌ونقل عمومی.
- رویکردهای نوین مبتنی بر فناوری‌های سنجش لحظه‌ای (حسگرها، سامانه‌های رزرو آنلاین)، نظریه بازی‌ها، الگوریتم‌های بهینه‌سازی، و یادگیری تقویتی چندعامله برای تنظیم خودکار نرخ‌ها با هدف بیشینه‌سازی درآمد و ارتقای رضایت مشتری.

این روند، گذار از مدیریت ایستا به مدیریت تطبیقی و داده‌محور را نشان می‌دهد که به‌ویژه در بازارهای رقابتی و با ظرفیت محدود، مزیت‌های چشمگیری ایجاد می‌کند.

اهداف پژوهش، پرسش‌های پژوهش، نوآوری‌ها و اهمیت کار

در سال‌های اخیر، رشد بی‌وقفه شهرنشینی و افزایش تعداد خودروها فشار روزافزونی بر زیرساخت‌های پارکینگ در بافت شهری وارد کرده است. در حالی که سیاست‌های قیمت‌گذاری پویا به عنوان ابزاری کلیدی برای مدیریت کارآمد ظرفیت و کاهش تراکم ترافیک مطرح شده‌اند، اغلب چارچوب‌های موجود یا به صورت تک‌عاملی عمل کرده‌اند، یا پویایی رفتار رقابتی بهره‌برداران را

نادیده گرفته‌اند. افزون بر این، متغیرهای مکانی و فیزیکی مانند فاصله از مقصد، ظرفیت محدود و حساسیت قیمتی کاربران در بسیاری از رویکردهای پیشین به درستی یکپارچه‌سازی نشده‌اند. این محدودیت‌ها سبب شده است که مدل‌های سنتی نتوانند پاسخ‌گوی پیچیدگی‌های بازار چندبهره‌برداری در شرایط غیرایستا باشند.

این پژوهش چارچوبی مبتنی بر یادگیری تقویتی چندعامله عمیق (MARL-DQN) توسعه داده است که در آن هر پارکینگ به عنوان یک عامل مستقل، در محیط رقابتی با سایر عامل‌ها تعامل می‌کند و ضمن یادگیری مداوم، نرخ‌های خود را به شکلی تنظیم می‌کند که هم سودآوری کلی بهره‌برداران بیشینه شود و هم ظرفیت‌ها به صورت بهینه مورد استفاده قرار گیرد. این چارچوب از مدل انتخاب چندجمله‌ای لاجیت (MNL) برای پیش‌بینی توزیع تقاضا بهره می‌برد و با توجه به وضعیت لحظه‌ای بازار، تصمیمات قیمت‌گذاری را در یک فرآیند یادگیری و سازگاری مداوم اتخاذ می‌کند. در این مسیر، پژوهش حاضر به چند پرسش بنیادی پاسخ می‌دهد: چگونه می‌توان در بازار چندبهره‌برداری، بدون اطلاع کامل از رفتار رقبا، نرخ‌هایی تعیین کرد که سود کل سیستم افزایش یابد؟ محدودیت ظرفیت و فاصله مکانی چه اثری بر سیاست‌های بهینه‌سازی قیمت دارند؟ قیمت‌گذاری پویا چگونه می‌تواند بر توزیع مکانی مشتریان و نرخ اشغال تأثیر بگذارد؟ و در نهایت، الگوهای رفتاری بازار و عامل‌ها در طول زمان چگونه شکل می‌گیرند و تغییر می‌کنند؟ در این پژوهش به مدل‌سازی جامع رقابت میان چند عامل مستقل در بستر پارکینگ شهری، به کارگیری الگوریتم MARL-DQN برای استخراج سیاست‌های نرخ‌گذاری بهینه در محیط غیرایستا و رقابتی، و ادغام ملاحظات فیزیکی و مکانی، شامل ظرفیت و فاصله، مستقیماً در تصمیم‌گیری عامل‌ها پرداخته شده مواردی که در پیشینه پژوهشی کمتر مورد توجه بوده اند. نتایج شبیه‌سازی‌ها نشان می‌دهد که این چارچوب نه تنها توانسته اهداف دوگانه سودآوری و استفاده کارا از ظرفیت را محقق کند، بلکه اثرات مثبتی بر توزیع تقاضا، تعادل بازار و بهبود تجربه کاربری نیز برجای گذاشته است.

مرور ادبیات

اهمیت سیاست گذاری قیمت گذاری و تصمیمات در پارکینگ شهری

از دهه ۱۹۷۰، بحث مدیریت عرضه و قیمت‌گذاری پارکینگ به عنوان بخشی از سیاست‌های مدیریت تقاضای سفر وارد ادبیات حمل‌ونقل شهری شد. نخستین کارها بیشتر بر اثرات مستقیم هزینه پارکینگ بر استفاده از خودرو تمرکز داشتند و شواهد روشنی یافتند که حتی تغییرات نسبتاً کوچک در هزینه می‌تواند انتخاب‌های سفر را تحت تأثیر قرار دهد (Yan et al., 2019). این مطالعات نشان دادند که حذف یا کاهش یارانه پارکینگ (مثلاً پارکینگ رایگان کارمندان) منجر به کاهش چشمگیر استفاده از خودروهای شخصی می‌شود. با گذر زمان، پژوهشگران دریافتند که مسئله پارکینگ تنها به «هزینه» محدود نمی‌شود، بلکه عوامل غیرقیمتی مانند زمان جست‌وجو و زمان پیاده‌روی تا مقصد نیز می‌توانند به اندازه یا حتی بیش از نرخ، بر انتخاب مکان و شیوه سفر اثر بگذارند. این یافته‌ها جایگاه سیاست‌های ترکیبی (قیمت‌گذاری به همراه مدیریت زمان و مکان پارک) را تقویت کرد (Qian & Rajagopal, 2014) در دهه‌های اخیر، با افزایش پیچیدگی تقاضا و محدودیت‌های عرضه، نگاه‌ها به سمت مدل‌سازی یکپارچه انتخاب شیوه سفر و مکان پارک سوق پیدا کرد. این رویکردها نه تنها نقطه‌نظر درآمدزایی برای بهره‌برداران را لحاظ می‌کنند، بلکه تعادل میان کارایی شبکه حمل‌ونقل و رضایت کاربران را نیز مدنظر قرار می‌دهند (P. Wang et al., 2020). در بازارهای واقعاً رقابتی که چندین بهره‌بردار خصوصی و عمومی فعال هستند، اهمیت تصمیمات قیمت‌گذاری دوچندان می‌شود. آثار متقابل قیمت‌گذاری میان رقبا، می‌تواند به نوسانات شدید سهم بازار و درآمد منجر شود و این نیاز به چارچوب‌های

تحلیلی پیشرفته‌تر را روشن کرده است (J. Wang et al., 2024). در چنین شرایطی، دیگر نمی‌توان صرفاً به نرخ ثابت یا تغییرات دوره‌ای محدود اکتفا کرد، بلکه لازم است مکانیزم‌هایی به کار گرفته شوند که پاسخ لحظه‌ای به نوسانات تقاضا و اقدامات رقبا ارائه دهند. همچنین، ورود فناوری‌های حسگر و سیستم‌های اطلاعاتی هوشمند، امکان دسترسی به داده‌های دقیق لحظه‌ای از ظرفیت و اشغال را فراهم کرده و افق‌های جدیدی برای سیاست‌گذاری باز کرده است. اما همان‌طور که مطالعات نشان می‌دهند، اگر این اطلاعات با سیاست‌های قیمتی مناسب تلفیق نشود، اثرگذاری آن بر تغییر رفتار رانندگان محدود خواهد بود (Jioudi et al., 2019; Qian & Rajagopal, 2014). برآیند این روند پژوهشی چند دهه‌ای، گذار از نگاه ایستا و تک‌بعدی به مدیریت پارکینگ، به سمت سیاست‌گذاری پویا، چندعامله و داده‌محور بوده است. این مسیر همان زیرساخت فکری و نظری است که تحقیق حاضر بر روی آن بنا شده و هدف آن ارائه چارچوبی است که بتواند هم منافع اقتصادی بهره‌برداران را بهینه کند و هم با انعطاف‌پذیری بالا به تغییرات سریع محیط شهری پاسخ دهد.

قیمت‌گذاری پویا و الگوریتم‌های آن

مفهوم قیمت‌گذاری پویا ابتدا در صنایع هوایی و هتل‌داری به عنوان روشی برای مدیریت درآمد (Revenue Management) مطرح شد؛ جایی که نرخ‌ها بسته به فصل، اشغال، و زمان رزرو تغییر می‌کردند. این ایده به سرعت به سایر صنایع مانند بزرگراه‌های با عوارض متغیر، خدمات اشتراکی، و نهایتاً پارکینگ شهری تعمیم داده شد (Lei & Ouyang, 2017). در مراحل اولیه، مدل‌های قیمت‌گذاری پویا عمدتاً بر مبنای روش‌های بهینه‌سازی کلاسیک شکل می‌گرفتند. بهره‌برداران از برنامه‌ریزی خطی یا غیرخطی برای تعیین نرخ بر اساس تقاضای پیش‌بینی شده استفاده می‌کردند. این مدل‌ها اغلب فرض می‌کردند که توزیع آماری تقاضا و حساسیت قیمتی شناخته شده است. هرچند، پیچیدگی محیط واقعی – شامل رقابت، نوسانات غیرقابل پیش‌بینی و رفتار غیرخطی کاربران – باعث شد که کارایی این روش‌ها محدود شود (Kalakanti & Rao, 2025). با گسترش قدرت محاسباتی و دسترسی به داده‌های بزرگ، روش‌های داده‌محور وارد صحنه شدند. این رویکردها از مدل‌های آماری و شبیه‌سازی برای تخمین تقاضای آتی و بهینه‌سازی نرخ‌ها استفاده کردند. روش‌هایی نظیر Approximate Dynamic Programming (ADP) توانستند بدون نیاز به پیش‌بینی کامل آینده، به راهکارهای نزدیک به بهینه دست یابند، به‌ویژه در سیستم‌های پیچیده شهری (Lei & Ouyang, 2017). در ادامه، پژوهش‌ها به سمت الگوریتم‌های مبتنی بر بازی‌های ریاضی (Game Theory) حرکت کردند تا بتوانند تعاملات میان بهره‌برداران و مشتریان یا میان چند بهره‌بردار را مدل کنند. مدل‌های Stackelberg Leader-Follower در محیط‌هایی که یک اپراتور پیشرو نرخ‌گذاری می‌کند و دیگران واکنش نشان می‌دهند، بسیار رایج شدند (Lim & Masoud, 2024; J. Wang et al., 2024). این چارچوب‌ها برای شرایط رقابتی و حتی سناریوهای با چندین خدمت همزمان (پارک، بارگیری، شارژ برقی) به کار رفتند. گام بعدی، ورود الگوریتم‌های یادگیری ماشین و یادگیری تقویتی (RL) به حوزه قیمت‌گذاری پویا بود. این الگوریتم‌ها توانستند بدون فرض دانش دقیق از توزیع تقاضا یا مدل سیستم، مستقیماً از داده‌های تجربه و تعامل، سیاست‌های بهینه استخراج کنند (Yin & Han, 2021). این امکان را به بهره‌برداران می‌دهد که به تغییرات محیطی – که ممکن است غیرخطی، تصادفی و چندعامله باشد – به‌طور آنی واکنش نشان دهند. در سال‌های اخیر، ترکیب روش‌های RL با مدل‌سازی چندعامله (Multi-Agent) تحولی چشمگیر ایجاد کرده است، به‌خصوص در محیط‌های رقابتی یا مشارکتی که تصمیمات قیمت‌گذاری یک عامل، بر تقاضا و درآمد دیگران تاثیر مستقیم دارد (Kastius & Schlosser, 2022). افزون بر این، الگوریتم‌های پیوسته‌محور نظیر Deep Deterministic Policy Gradient

(DDPG) برای تنظیم نرخ‌های پیوسته در سناریوهایی مانند شارژ سریع خودروهای برقی یا رزرو پارکینگ به‌کار گرفته شده‌اند (Liu et al., 2021). این مسیر تکاملی – از مدل‌های تحلیلی ساده تا الگوریتم‌های پیشرفته یادگیری تقویتی چندعامله – زمینه‌ساز رویکردهایی است که امروز می‌توانند در زمان واقعی، نرخ‌ها را بر اساس داده‌های لحظه‌ای، تقاضا، و رقابت، هم برای افزایش کارایی سیستم و هم بهبود تجربه کاربر بهینه کنند. این دقیقاً همان بستر نظری و فنی است که تحقیق حاضر از آن بهره می‌گیرد تا یک سامانه قیمت‌گذاری بهینه و پویا برای پارکینگ شهری توسعه دهد.

یادگیری تقویتی در قیمت‌گذاری پویا و موارد کاربرد

مفهوم یادگیری تقویتی (Reinforcement Learning – RL) در ابتدا بیشتر در حوزه‌هایی چون بازی‌های رایانه‌ای و رباتیک مورد استفاده بود؛ جایی که یک عامل از طریق تعامل مکرر با محیط و آزمون و خطا، سیاست بهینه را برای بیشینه‌سازی پاداش بلندمدت خود یاد می‌گرفت. اما با رشد بازارهای داده‌محور و نیاز به تصمیم‌گیری‌های سریع و تطبیقی، پژوهشگران از اواخر دهه ۲۰۱۰ به این نتیجه رسیدند که RL می‌تواند جایگزین یا مکمل مدل‌های سنتی قیمت‌گذاری پویا باشد (Kastius & Schlosser, 2022). اولین موج استفاده از RL در قیمت‌گذاری به حوزه‌هایی چون تجارت الکترونیک و مدیریت موجودی تعلق داشت. در این مطالعات، الگوریتم‌های Q-Learning و Deep Q-Network (DQN) برای تنظیم نرخ‌ها بر اساس بازخورد لحظه‌ای بازار توسعه یافتند. این الگوریتم‌ها قادر بودند بدون داشتن مدل کامل از رفتار تقاضا، از داده‌های تراکنش برای یافتن استراتژی‌های سودآور استفاده کنند (Yin & Han, 2021). به تدریج، این رویکرد به صنایعی انتقال یافت که با محدودیت منابع فیزیکی و شرایط محیطی متغیر روبه‌رو بودند. به عنوان مثال، مدیریت مصرف انرژی در ریزشبکه‌ها با استفاده از DDPG و DQN توانست الگوهای بهینه مصرف و تولید را در محیط‌های ناپایدار یاد بگیرد (Nakabi & Toivanen, 2021). در صنعت خودروی برقی نیز، RL برای تعیین قیمت و زمان‌بندی شارژ به کار گرفته شد تا هم ازدحام ایستگاه‌ها کاهش یابد و هم درآمد بهره‌بردار بیشینه شود (Liu et al., 2021). ورود RL به پارکینگ شهری از دو جهت تقویت شد:

- داده‌های لحظه‌ای که از حسگرها و اپلیکیشن‌های رزرو به دست می‌آید.

- ماهیت پویا و غیرقابل پیش‌بینی تقاضا در شبکه‌های شهری.

در این حوزه، مدل‌های مبتنی بر Q-Learning برای تعیین نرخ‌های ساعتی پارکینگ با در نظر گرفتن عوامل محیطی (مثل آب‌وهوا و تعطیلات) به کار رفتند که توانستند هم درآمد و هم تعادل اشغال را بهبود دهند (Sowmya & Dhabu, 2023). یکی از پیشرفت‌های مهم، استفاده از یادگیری تقویتی چندعامله (Multi-Agent Reinforcement Learning – MARL) بود. در MARL، هر بهره‌بردار یا بخش از سیستم به عنوان یک عامل مستقل عمل می‌کند که تصمیماتش بر دیگران اثرگذار است. به عنوان مثال، در کنترل بلادرنگ شارژ خودروهای برقی، چارچوب *Centralized Training, Decentralized Execution* توانست هماهنگی بین ایستگاه‌های مختلف را بهبود دهد و از تصمیمات ناهم‌سود جلوگیری کند (Li et al., 2025). افزون بر آن، RL در سناریوهای چندهدفه که باید هم درآمد و هم تجربه کاربری بهینه شوند کارآمد است. الگوریتم‌های پیوسته‌محور مثل DDPG یا سیاست‌محور مانند Soft Actor-Critic (SAC) در محیط‌های با فضای اقدام پیوسته عملکرد بالاتری داشتند (Kastius & Schlosser, 2022). به این ترتیب، مسیر پژوهشی RL در قیمت‌گذاری پویا از کاربردهای ساده و تک‌عاملی به سمت چارچوب‌های پیچیده چندعامله، چندهدفه و بلادرنگ حرکت کرده است. این روند همان پایه علمی و فنی‌ای را ایجاد می‌کند که امکان توسعه سامانه‌های قیمت‌گذاری شهری هوشمند، انطباق‌پذیر و رقابتی را فراهم کرده است – هدفی که تحقیق حاضر دنبال می‌کند.

قیمت گذاری پویا در پارکینگ شهری

پژوهش در زمینه قیمت گذاری پارکینگ شهری پیش از دهه ۲۰۱۰ عمدتاً بر روش های ثابت یا تعدیل های دوره ای نرخ ها متمرکز بود. در این رویکردها، نرخ ها بر اساس داده های تاریخی و با فواصل بلندمدت (مثلاً ماهانه یا فصلی) تنظیم می شدند، که فرصت واکنش سریع به نوسانات روزانه یا ساعتی تقاضا از بین می رفت (Yan et al., 2019). ظهور فناوری های شهر هوشمند و سامانه های اطلاع رسانی لحظه ای در اوایل دهه ۲۰۱۰، نقطه عطفی در این مسیر بود. با استفاده از حسگرها، دوربین ها و اپلیکیشن های رزرو، داده های اشغال پارکینگ در زمان واقعی در دسترس قرار گرفتند و امکان آزمون قیمت گذاری پویا با بازه های کوتاه مدت فراهم شد. پروژه هایی مانند Sfpark در سان فرانسیسکو نشان دادند که به کمک بهینه سازی نرخ ها می توان توازن اشغال را بهبود داد، اما این طرح ها عموماً نرخ ها را تنها هر چند هفته یا ماه یکبار به روزرسانی می کردند - کندتر از آنچه تقاضای پرنوسان شهری نیاز داشت (Lei & Ouyang, 2017). مرحله بعدی با ورود مدل سازی پیشرفته و شبیه سازی شهری همراه بود. روش هایی مانند Macroscopic Fundamental Diagram (MFD) برای توصیف تعامل ترافیک و جستجوی پارکینگ در شبکه های چندحالت به کار گرفته شد و قیمت گذاری بازخوردی (Feedback Pricing) توانست بدون پیش بینی کامل آینده، نرخ ها را طوری تنظیم کند که ازدحام و زمان جستجو کاهش یابد (Zheng & Geroliminis, 2016). پس از آن، رویکردهای بازی های ریاضی - مانند مدل های Stackelberg Leader-Follower - وارد صحنه شدند تا تعاملات بین بهره برداران، بخش خصوصی و شهرداری ها را در سناریوهای رقابتی یا مشارکتی مدل کنند. این مدل ها در مدیریت هم زمان پارکینگ عمومی و خصوصی یا فضای حاشیه خیابان موفق عمل کردند (Lim & Masoud, 2024; J. Wang et al., 2024). در سال های اخیر، استفاده از یادگیری تقویتی (RL) و حتی چندعامله (MARL)، انقلابی در قیمت گذاری پارکینگ ایجاد کرده است. این الگوریتم ها می توانند به طور خودکار و بدون مدل کامل از بازار، نرخ های روزانه یا ساعتی را بهینه کنند. نمونه هایی شامل Q-Learning برای تعیین نرخ های پویا بر اساس عوامل محیطی مانند آب و هوا و تعطیلات (Sowmya & Dhabu, 2023) و معماری های چندعامله برای بهینه سازی هم زمان اشغال و تجربه کاربر در اپلیکیشن های رزرو پارکینگ (Jioudi et al., 2019) هستند. در نهایت، همگرایی فناوری های جمع آوری داده در زمان واقعی، الگوریتم های پیشرفته بهینه سازی و یادگیری تقویتی چندعامله، امکان طراحی سامانه های پارکینگ شهری هوشمند را فراهم کرده است که قادرند واکنشی سریع، کارآمد و متعادل به چالش های تقاضای پویا، محدودیت ظرفیت و شرایط رقابتی بازار نشان دهند. این مسیر تکاملی، بستری محکم برای پژوهش حاضر فراهم می آورد که بر توسعه راهبردهای قیمت گذاری در محیط های رقابتی و چندعامله متمرکز است.

یادگیری تقویتی در قیمت گذاری پویا در پارکینگ شهری و انواع الگوریتم ها

کاربرد یادگیری تقویتی (RL) در پارکینگ شهری از زمانی اوج گرفت که محدودیت های روش های تحلیلی سنتی و داده محور در پیش بینی رفتار متغیر رانندگان آشکار شد. سامانه های کلاسیک قیمت گذاری، حتی با استفاده از مدل های حساسیت قیمتی، نمی توانستند به خوبی به شرایط غیرقطعی مانند تغییرات آب و هوا، رویدادهای شهری یا نوسانات ناگهانی تقاضا پاسخ دهند (Qian & Rajagopal, 2014).

نخستین تلاش ها با الگوریتم های مدل محور مانند Q-Learning انجام شد که در آن عامل با آزمون و خطا، نرخ های بهینه را برای ساعات مختلف روز پیدا می کرد. پژوهش Sowmya و Dhabu (2023) نشان داد که Q-Learning، حتی بدون دانش دقیق از مدل تقاضا، می تواند با در نظر گرفتن متغیرهایی مثل فاصله تا مرکز شهر، شرایط جوی و تعطیلات، تعادل بهتری بین

درآمد و اشغال برقرار کند (Sowmya & Dhabu, 2023) در ادامه، ترکیب الگوریتم‌های یادگیری عمیق با RL راه را برای شبکه‌های Q عمیق (DQN) باز کرد که قادر بودند فضای حالت‌های پیچیده‌تری را پردازش کنند. این مسیر، به استفاده از الگوریتم‌های پیوسته‌محور مانند Deep Deterministic Policy Gradient (DDPG) در سناریوهایی که نرخ‌ها باید به‌طور پیوسته تنظیم شوند انجامید، مانند مدیریت ظرفیت شارژ و پارک در یکپارچگی با خدمات خودروهای برقی (Liu et al., 2021) با ظهور یادگیری تقویتی چندعامله (MARL)، امکان مدل‌سازی شرایطی فراهم شد که در آن چند بهره‌بردار یا چند بخش از یک سیستم شهری (مثلاً چند پارکینگ در یک منطقه) به‌صورت همزمان تعامل و رقابت دارند. پلتفرم e-Parking نشان داد که معماری چندعامله می‌تواند با در نظر گرفتن حضور رانندگان با رزرو و بدون رزرو، سیاست قیمتی هوشمند و منصفانه‌تری ایجاد کند (Jioudi et al., 2019) رویکردهای آموزش متمرکز، اجرای غیرمتمرکز (CTDE) در MARL، مانند آنچه در کنترل بلادرنج شارژ خودروهای برقی استفاده شد (Li et al., 2025) قابلیت تعمیم به پارکینگ شهری را دارند و امکان هماهنگی همزمان تصمیمات قیمت‌گذاری میان چند عامل را بدون نیاز به اشتراک کامل داده‌های حساس فراهم می‌کنند. انواع رایج الگوریتم‌های RL برای پارکینگ شامل:

- Q-Learning و SARSA برای محیط‌های گسسته با نرخ‌های دسته‌بندی‌شده.
- DQN برای فضای حالت‌های بزرگ با تصمیمات گسسته اما پیچیده.
- DDPG و SAC برای نرخ‌گذاری پیوسته و سناریوهای با تغییرات ملایم اما مداوم.
- MARL با CTDE یا یادگیری مستقل برای تعاملات چندپارکینگی رقابتی یا مشارکتی.

این مسیر تکاملی، نشان از گذر تحقیق از مدل‌های ساده و تک‌عاملی به الگوریتم‌های عمیق و چندعامله دارد که می‌توانند شرایط واقعی و چندبخشی پارکینگ شهری را بهتر بازسازی کنند. این پیشرفت‌ها مستقیماً با اهداف پژوهش حاضر همسو هستند، که به دنبال توسعه یک سیستم DQN چندعامله رقابتی قابل انطباق با شرایط واقعی و اقتصادی است. مرور پژوهش‌های پیشین نشان می‌دهد که اگرچه استفاده از RL و MARL در قیمت‌گذاری پویا نتایج امیدبخشی داشته، اغلب مدل‌ها یا فرضیات ایستا درباره رفتار رقبا داشته‌اند، یا به‌صورت یک‌عاملی عمل کرده‌اند و تعاملات واقعی میان بهره‌برداران را نادیده گرفته‌اند. همچنین، کمتر مطالعه‌ای توانسته است اثر مشترک ظرفیت فیزیکی، موقعیت مکانی، و حساسیت قیمتی را در یک محیط رقابتی غیرایستا به‌طور همزمان مدل کند. این پژوهش، با طراحی یک چارچوب MARL-DQN که در آن هر پارکینگ به‌عنوان عامل مستقل رفتار می‌کند و تصمیماتش بر بقیه اثرگذار است، می‌کوشد این نیاز را در بستر شبیه‌سازی واقعی‌تر برطرف نماید.

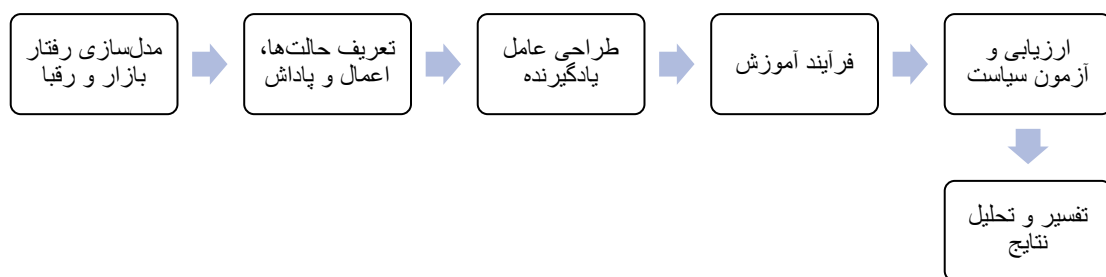
جدول مرور ادبیات

روش شناسی و مدل سازی

روند انجام پروژه

روند انجام این پژوهش با هدف طراحی و ارزیابی یک سامانه هوشمند قیمت گذاری پویای پارکینگ در محیط رقابتی، مبتنی بر یادگیری تقویتی عمیق و ساختار چندعامله (Multi-Agent DQN)، در چند مرحله اصلی پیاده سازی گردید. این مراحل به صورت متوالی و تکرارشونده اجرا شده و در قالب چرخه ای از تعریف مسئله تا ارزیابی مدل، به شکل زیر سازمان دهی شدند

در شکل 1 شماتیک کلی پروژه رسم شده است



شکل 1، شماتیک مراحل اجرای پروژه

روش DQN

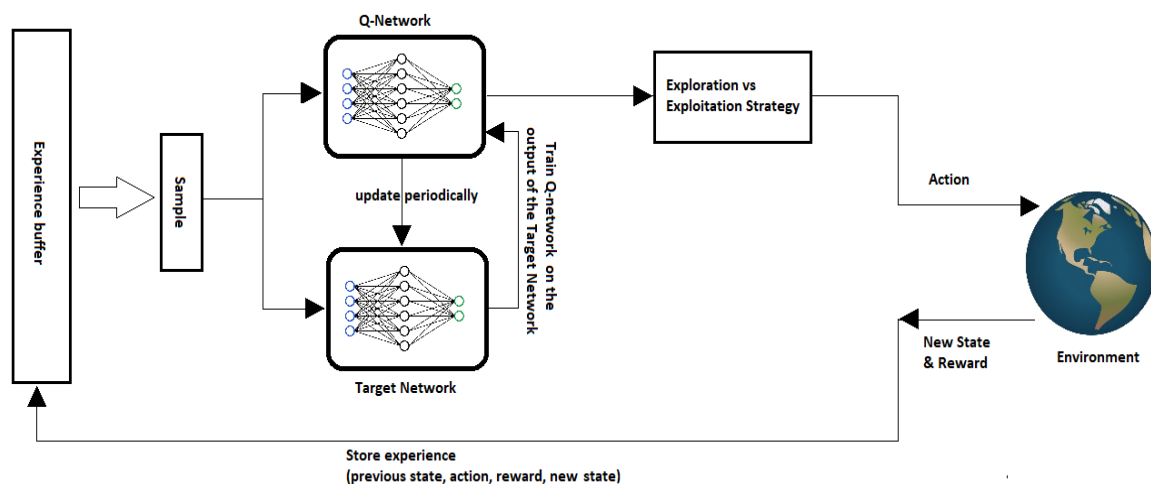
انتخاب الگوریتم شبکه Q عمیق (DQN) در این پژوهش ناشی از نیاز به مدیریت فضای حالت پیوسته و اعمال با ابعاد بزرگ و متغیرهای متعدد تأثیرگذار بر تصمیمات قیمت گذاری است. در محیط رقابتی بین پارکینگ ها، وضعیت سیستم در هر لحظه تابعی از عواملی چون نرخ اشغال، شرایط جوی، ویژگی های روز، و قیمت های رقبا است که مدل سازی تحلیلی دقیق آن پیچیده و بعضاً غیرممکن است. DQN، به عنوان یک روش مدل-آزاد، با استفاده از شبکه های عصبی توانایی تقریب تابع ارزش $Q(s,a)$ را در چنین فضای پیچیده ای فراهم می کند و از طریق مکانیزم های بازپخش تجربه و شبکه هدف ثابت و کارایی آموزش را بهبود می بخشد.

در بخش نظری، چارچوب تصمیم گیری این مدل بر اساس فرآیند تصمیم گیری مارکوفی (MDP) بنا شده است. در یک MDP، فرض می شود که احتمال انتقال به حالت بعدی و پاداش دریافتی تنها به حالت و عمل فعلی وابسته است (خاصیت مارکوفی). این ویژگی امکان مدل سازی پویای محیط را فراهم می کند، به نحوی که وضعیت فعلی پارکینگ (از نظر نرخ اشغال، قیمت، و شرایط محیطی) به همراه اکشن انتخابی (تعیین قیمت) کل اطلاعات لازم را برای پیش بینی گام بعدی و پاداش فراهم می آورد. این ساختار مارکوفی به DQN اجازه می دهد تا از طریق تعامل مداوم با محیط شبیه سازی شده و به روزرسانی تدریجی مقادیر Q، سیاستی بهینه برای قیمت گذاری پویا در شرایط رقابتی استخراج کند.

شبکه Q عمیق (Deep Q-Network) یکی از الگوریتم های کلیدی یادگیری تقویتی عمیق است. با ترکیب Q-Learning سنتی و شبکه های عصبی عمیق، امکان تقریب تابع ارزش-عمل $Q(s,a)$ را در فضاهای حالت و عمل بسیار بزرگ فراهم می سازد. در این روش، بردار حالت ورودی به یک شبکه عصبی چندلایه تغذیه شده و این شبکه برای هر عمل ممکن، برآوردی از ارزش

مورد انتظار (Expected Return) باز می‌گرداند. عامل در هر گام، اکشنی را بر اساس بیشینه‌سازی این خروجی یا با بهره‌گیری از سیاست اکتشافی انتخاب می‌کند. نوآوری‌های کلیدی DQN شامل استفاده از **بازپخش تجربه** (Experience Replay) برای شکستن هم‌بستگی نمونه‌ها و بهره‌گیری از **شبکه هدف** (Target Network) برای پایداری یادگیری است. این رویکرد امکان یادگیری سیاست‌های بهینه را بدون نیاز به مدل صریح محیط، تنها بر اساس تعاملات گذشته و بازخورد پاداش فراهم می‌کند.

در شکل 2 مکانیزم کلی DQN نمایش داده شده است که در آن شبکه عصبی، بردار حالت را دریافت کرده و بردار ارزش‌های Q برای مجموعه اعمال ممکن را پیش‌بینی می‌کند.



شکل 2، دیاگرام روش DQN

مسئله حاضر در تقاطع چند حوزه علمی و نظری قرار می‌گیرد. از منظر **مدیریت شهری و اقتصاد حمل‌ونقل**، سیاست‌گذاری قیمتی یکی از ابزارهای کلیدی برای مدیریت تقاضای پارکینگ، کنترل نرخ اشغال فضا و کاهش ترافیک جستجوی محل پارک است که ریشه در نظریه‌های اقتصاد خرد و رفتار مصرف‌کننده دارد. این مسئله همچنین با حوزه **مدل‌سازی انتخاب گسسته** مرتبط است، جایی که رفتار انتخاب رانندگان بر اساس مطلوبیت نسبی گزینه‌ها و با استفاده از مدل‌هایی نظیر **لاجیت چندجمله‌ای** تبیین می‌شود. در بعد محاسباتی، رویکرد پژوهش در **چارچوب یادگیری تقویتی بدون مدل** و به طور خاص الگوریتم **شبکه Q عمیق (DQN)** جای می‌گیرد که با تقریب تابع ارزش از طریق شبکه عصبی، توانایی پردازش فضای حالت و عمل بزرگ را فراهم می‌سازد. ویژگی چندعامله (Multi-Agent) مدل، آن را به حوزه **سیستم‌های چندعامله و بازی‌های غیرهمکارانه** پیوند می‌دهد، جایی که هر عامل بهینه‌سازی سود خود را در تعاملی پویا با رقبا دنبال می‌کند. پژوهش حاضر با ترکیب این بنیان‌های نظری، جایگاه خود را به عنوان گامی پیشرو در توسعه مدل‌های قیمت‌گذاری پویا برای محیط‌های شهری رقابتی تثبیت می‌کند.

اهداف اصلی مدل

طراحی مدل پیشنهادی بر این اصل استوار است که **قیمت‌گذاری پویا** در محیط رقابتی باید به صورت همزمان دو هدف کلیدی را محقق سازد:

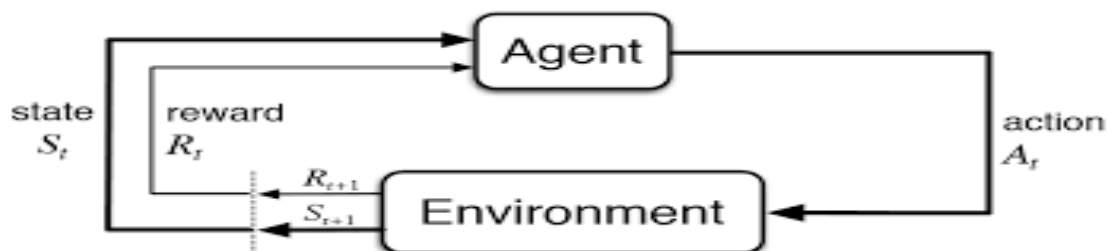
1. **حداکثرسازی سود روزانه هر پارکینگ** از طریق تعیین بهینه قیمت ساعتی که علاوه بر پوشش هزینه‌ها، حاشیه سود مناسبی ایجاد نماید.

2. **متعادل‌سازی نرخ اشغال نسبت به ظرفیت** به منظور جلوگیری از هر دو حالت «کم‌استفاده شدن» و «اشباع بیش‌ازحد» فضاهای پارک.

این اهداف با تغییر پویا قیمت‌ها در بازه‌های زمانی کوتاه و کنترل تقاضا بر مبنای کشش قیمتی تقاضا (PED) و توزیع بهینه آن بین رقبا به کمک مدل انتخاب گسسته چندجمله‌ای/لوجیت (MNL) حاصل می‌شوند؛ به گونه‌ای که قیمت به صورت بلادرنگ متناسب با شرایط بازار (شامل ساعت، روز هفته، وضعیت آب‌وهوا و رفتار رقبا) تعدیل شده و الگوی توزیع تقاضا میان پارکینگ‌ها بهینه گردد.

جدول متغیرها

Q_t^{base}	تقاضای پایه قبل از تعدیل عوامل محیطی
Q_t	کل تقاضای تعدیل یافته
Q_{it}	تقاضای تخصیص یافته به پارکینگ
Q_{target}	تقاضای مطلوب پارکینگ
MS_t	سهم بازار پارکینگ
Rej_{it}	خودروهای رد شده به دلیل پر بودن ظرفیت
S_{it}	خودروهای خدمت‌دهی شده توسط پارکینگ
P_{it}	قیمت واحد (ساعتی) پارکینگ
C_i	هزینه واحد پارکینگ
K_i	ظرفیت کل پارکینگ
O_{it}	نرخ اشغال فعلی
PED_t	کشش قیمتی تقاضا
d_i	میانگین فاصله پارکینگ تا مناطق مهم
q_i	شاخص کیفیت خدمات
CP_{it}	قیمت نسبی پارکینگ به رقا
R_{it}	پاداش عامل
π_{it}	سود خالص پارکینگ
a_{it}	عمل انتخاب‌شده (ضریب قیمت)
W_t	وضعیت آب‌وهوا
D_t	نوع روز
FQ_t	پیش‌بینی تقاضا
OG_{it}	اختلاف اشغال از رقا
RP_{it}	قیمت نسبی رقا



شکل 3 دیاگرام الگوریتم های یادگیری تقویتی

شکل 3 نحوه عملکرد الگوریتم های تقویتی را نشان می دهد برای اجرا و مدل سازی نیاز به تعریف محیط، حالات، اعمال و نوع محاسبه پاداش داریم که در این بخش به تفصیل توضیح داده میشوند.

برای مدل سازی محیط بازار و تمرین عامل ها در این فضا ابتدا تقاضای پایه موجود در بازار تعیین میشود، برای هر بازه زمانی 0 تا 24 ساعت شبانه روز با توجه به افزایش میزان تقاضا در ساعات 14 تا 19 بعد از ظهر لیستی از تقاضای پایه به عنوان مرجع محاسبه کل تقاضا قرار میگیرد

$$Q_t^{base} = [50, 50, 20, 20, 50, 200, 300, 300, 400, 400, 400, 500, 1000, 1200, 1500, 1800, 2000, 1500, 1000, 700, 500, 300, 100, 50]$$

تقاضای کلی بازار بسته به نوع روز و نوع آب و هوا و میانگین قیمت بازار نسبت به قیمت مطلوب تغییر میکند، با محاسبه PED (کشش قیمتی تقاضا) که بسته به ساعت روز و نوع روز و آب و هوا تغییر میکند و ضرب آن در میزان اختلاف قیمت با قیمت مطلوب میزان تغییر تقاضای بازه زمانی از تقاضای پایه محاسبه شده و با قرار دادن در یک تابع توزیع پواسون به عنوان میانگین تقاضای واحد زمانی تقاضای هر بازه زمانی محاسبه میشود

PED: میزان درصد اختلاف تقاضا به ازای هر درصد تغییر قیمت است، این متغیر در آب هوای نامناسب کاهش یافته و در روز های آخر هفته افزایش میابد همچنین در ساعات شلوغی افزایش یافته و در مواقع خلوت کاهش میابد، این متغیر در واقع حساسیت رانندگان به طور کلی به قیمت های بازار در شرایط مختلف میباشد

$$demand\ factor = ((PED * ((\bar{P}_{it} - 1.5\bar{C}_{it})) / 1.5\bar{C}_{it}))$$

$$\lambda = (1 + (demand\ factor)) * Q_t^{base}$$

$$poisson(\lambda) = Q_t$$

مطلوب رانندگان پرداخت 50 درصد سود است. بر اساس کشش تقاضا و تقاضای پایه ورودی تابع توزیع پواسون محاسبه و از این تابع احتمالی مقدار تقاضا در هر بازه زمانی محاسبه میشود

ترک پارکینگ توسط خودرو ها

برای مدل سازی ترک پارکینگ توسر رانندگان فرض شده است نیمی از رانندگان بعد از دو ساعت و ربع آنها بعد از 1 ساعت و باقی مانده بعد از 3 ساعت پارکینگ را ترک میکنند، برای کاهش پیچیدگی مسئله از مدل سازی احتمالی ترک پارکینگ خودداری میکنیم چنانچه هدف اصلی پژوهش مدل سازی قیمت و ورودی پارکینگ میباشد

$$departure_t = (0.25 * S_{t-1}) + (0.5 * S_{t-2}) + (0.25 * S_{t-3}) \quad \forall i = 1, 2, 3$$

نحوه تخصیص تقاضا و سهم بازار

برای محاسبه تقاضای هر پارکینگ در فضای رقابتی از روش (MNL) استفاده میشود در این روش متغیر Utility برای هر پارکینگ بسته به ویژگی های آن محاسبه میشود ویژگی های کیفیت پارکینگ (q_i) فاصله از نقاط شلوغ اطراف (d_i) و قیمت پارکینگ (p_{it}) در محاسبه این مطلوبیت موثر است بازار نسبت به هر یک از این ویژگی ها حساسیت متفاوتی دارند

کیفیت در بازه ای از 0 تا 1 تعریف میشود از کمترین کیفیت تا بالاترین آن

برای هم مقیاس کردن ویژگی ها تا تاثیر مقیاس بر مطلوبیت از بین برود مقادیر قیمت و کیفیت و فاصله با معادله $\min \max$ بی مقیاس میگردد.

$$norm(P)_{it} = \frac{p_{it} - 0}{400 - 0}$$

براعمال تاثیر افزایش قیمت همزمان تمام پارکینگ ها و در نظر گرفتن آن در میزان مطلوبیت از سقف و کف ثابت برای نرمال سازی قیمت استفاده میکنیم

$$norm(q)_i = \frac{q_i - \min(q_i)}{\max(q_i) - \min(q_i)}$$

$$norm(d)_i = \frac{d_i - \min(d_i)}{\max(d_i) - \min(d_i)}$$

معادله محاسبه Utility در ادامه آمده است

$$U_{it} = (\alpha * p_{it}) + (\beta * q_i) + (\gamma * d_i)$$

α : price sensitivity , β : quality sensitivity, γ : distance sensetivity

پس از محاسبه مطلوبیت پارکینگ ها از معادله MNL برای تعیین سهم بازار (MS_{it}) هر رقیب بهره میبریم

$$MS_{it} = \frac{e^{U_{it}}}{\sum_{i=1}^n e^{U_{it}}}$$

برای محاسبه تقاضای هر پارکینگ (Q_{it}) ضرب سهم بازار در تقاضای بازه زمانی استفاده میکنیم.

$$Q_{it} = MS_{it} * Q_t$$

فرضیات:

برای ساده سازی و مدل سازی بهتر فرضیات در نظر گرفته شده در ادامه بیان میشود:

- 1- قیمت در 1 ساعت ثابت است
- 2- تنها دو نوع روز به عنوان طول هفته و آخر هفته در نظر گرفته شده است
- 3- تنها دو نوع آب و هوای مناسب و نامناسب در نظر گرفته شده است
- 4- در ابتدای روز پارکینگ ها خالی هستند
- 5- حساسیت تمام رانندگان به تغییرات قیمت یکسان در نظر گرفته شده است
- 6- همه رقبا از سیاست گذاری قیمت پویا پیروی کرده و اهداف یکسانی را دنبال میکنند، هر سه رقیب عقلانی هستند

زنجیره های مارکوف مسئله

حالت (state)

هر عامل مسئله با مشاهده و سنجش حالت سیستم تصمیم به انتخاب یکی از عمل های موجود میگیرد تا پاداش بیشینه به دست آورد حالت سیستم شامل وضعیت خود پارکینگ، ساعت و روز و وضعیت پارکینگ نسبت به رقبا میباشد برای تعریف حالت 7 متغیر زیر تعریف شده است

- 1- ساعت: یکی از موارد مهم و تاثیر گذار بر تصمیم گیری عامل ساعت روز میباشد عامل بسته به ساعت روز تصمیم میگیرد سیاست های مناسب را برگزیند (H) گسسته از 0 تا 24
- 2- نوع روز: آیا در آخر هفته هستیم و یا طول هفته رفتار مشتریان به نسبت متفاوت است و عامل نیز باید متفاوت عمل کند (D) 1 آخر هفته 0 طول هفته
- 3- نوع آب و هوا: اثر مشابه نوع روز دارد (W) 1 هوای نامناسب 0 هوای مناسب
- 4- شلوغی: میزان شلوغی فعلی پارکینگ تعیین کننده استفاده کم یا زیاد از پارکینگ بوده و برای احقاق اهداف صاحبان پارکینگ از اهمیت بالایی برخوردار است (O) متغیر پیوسته در بازه (0 - 1)
- 5- قیمت نسبی: درصد تفاوت قیمت پارکینگ نسبت به میانگین قیمت بازار متغیر گسسته با ۲ نمایش داده میشود

$$r = \frac{P_{it} - \bar{P}_{it}}{\bar{P}_{it}}$$

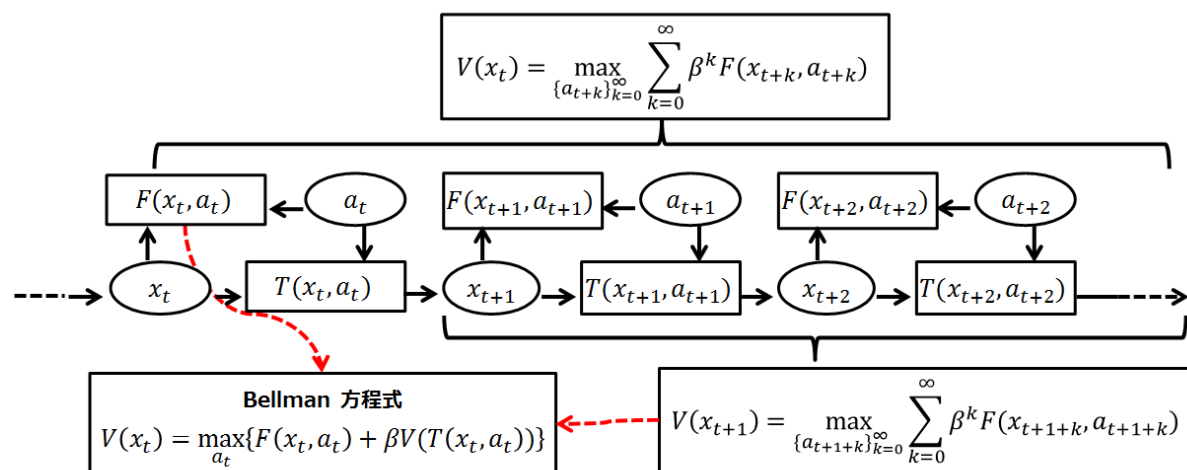
- 6- اختلاف شلوغی نسبت به رقبا متغیر پیوسته با Ogap نمایش داده میشود $Ogap = O_{it} - \bar{O}_{it}$
- 7- پیشبینی تقاضای دو بازه زمانی آینده متغیر پیوسته با FQ نمایش داده میشود $FQ_t = \frac{FQ_{t+1} + FQ_{t+2}}{2}$

حالت سیستم به شکل تابعی از این هفت متغیر ($S(H, D, W, O, r, Ogap, FQ)$) نمایش داده میشود

اعمال (actions)

هر عامل نیاز دارد با توجه به حالت فعلی سیستم عملی را انتخاب کند، در این پروژه در هر بازه زمانی عامل میتواند ضربی از هزینه هر واحد پارکینگ را به عنوان قیمت اعلام کند و این تنها مجموعه عمل عامل است

به عبارتی قیمت ها از ضرب عددی بین (0.8 – 3.5) برابر هزینه پارکینگ تعیین میشود الگوریتم DQN تنها اعمال گسسته را میپذیرد بنابراین این بازه به 20 بخش تقسیم شده و از این 20 ضریب هر کدام میتواند به عنوان تصمیم عامل انتخاب شده و بر اساس این انتخاب و مشاهده وضعیت بعدی پاداش این فعل در این حالت خاص با معادله Belman محاسبه میشود شکل 4 نحوه عملکرد معادله بهینه سازی را نشان میدهد



شکل 4 معادله بلمن و فرآیند زنجیره ای

تابع احتمالی انتقال transition probability distribution

در هر حالت با اتخاذ تصمیم عامل بر تعیین قیمت بازه زمانی بعدی، تقاضا و در نتیجه شلوغی حالت بعد با معادلات بیان شده در بخش تدوین محیط شبیه سازی محاسبه میشود، به علت تغییر و بهبود سیاست های هر سه رقیب در طول بازه شبیه سازی فضای مدل غیر ایستا بوده و توابع احتمالی انتقال به حالت بعد در طول روز تغییر میکند و هر عامل نیاز به یادگیری سیاست بهینه در چنین فضایی را دارد، بنابر این حالت بعدی بسته به اعمال و حالت قبل بوده، اما محیط مدام در حال تغییر است

$$S_{t+1} = P_t(S_{t+1}|S_t, a_t^1, a_t^2, \dots, a_t^n)$$

حالت در زمان $S_t:t$

عمل عامل i در بازه $a_t^i:t$

پاداش: (Reward Function)

در این مدل، پاداش هر عامل به عنوان معیاری برای ارزیابی کیفیت تصمیم گیری در بازه t زمانی جاری تعریف شده و ترکیبی از سه جزء اصلی است: سود اقتصادی، انحراف از هدف اشغال و جریمه برای تقاضای خدمت داده نشده.

برای پاداش باید در نظر داشت برای پارکینگ در هنگام ساعت شلوغی سود از اهمیت بالا برخوردار است و در ساعات عادی تکمیل هر چه بیشتر ظرفیت برای اعمال این موضوع از وزن دهی استفاده شده

$$Reward = X * profit + (1 - X) * arrival\ bonus - rejection\ penalty$$

X وزن دهی اهمیت سود و مشتری از دست رفته است زمان شلوغی از 14 تا 19 بعد از ظهر میباشد

$$X = 0.7 \text{ if peak hour}$$

$$X = 0.1 \text{ if off - peak hour}$$

سود از ضرب تعداد خودرو خدمات دهی شده و میزان حاشیه سود به دست می آید

$$profit = S_{it} * (P_{it} - C_i)$$

میزان خلوتی بیش از حد و یا مازاد شلوغی با $arrival\ bonus$ اندازه گیری میشود در صورتی که پارکینگ بیش از نسبت ظرفیتش به ظرفیت کل از تقاضای بازه زمانی بهره مند گردد به این معناس که قیمت پایین تری ارائه کرده و ظرفیت افزایش قیمت دارد و بر عکس در صورت پایین بودن به این معناس که امکان افزایش قیمت دارد و باید قیمت را زیاد کند البته میزان اهمیت این موضوع در ساعات خلوت اهمیت پیدا میکند که نباید تعداد معدود رانندگان را از دست داد.

$$arrival\ bonus = (Q_{it} - Q_{it}^{target}) * profit$$

$$Q_{it}^{target} = \frac{k_i}{\bar{k}_i} * Q_{it}$$

برای محاسبه میزان خسارات ناشی از عدم خدمات دهی به خودرو ها از معادله زیر استفاده شده است، در صورت رد کردن خودرو به این معناس که ترافیک شهری بیشتر شده و توانایی قیمت گذاری بالا را داشته ایم و این سود از دست رفته است بنابر این ضریب بالایی برای این جریمه قائل میشویم میزان سود واقعی از دست رفته 1.5 برابر به عنوان جریمه

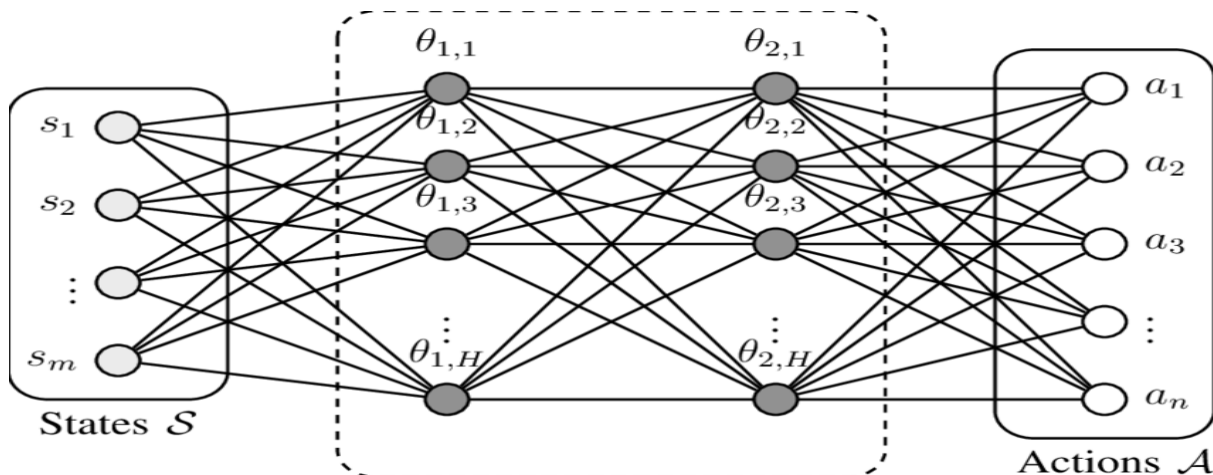
$$rejected\ penalty = rej_{it} * profit * 1.5$$

الگوریتم

باتوجه به ماهیت مسئله، یعنی محیط رقابتی چندعاملی (MAS) با حالت های پیوسته و فضای عمل گسسته، الگوریتم Deep Q-Network (DQN) انتخاب شد. دلیل انتخاب DQN نسبت به الگوریتم های مبتنی بر سیاست (Policy-Based) مانند PPO یا SAC، توانایی آن در تقریب تابع ارزش $Q(s, a)$ در فضاهای حالت بزرگ و پایداری بیشتر در برابر نوسانات محیط غیرایستا است. در محیط حاضر، هر عامل تنها مشاهده نسبی از وضعیت رقبا دارد و نیاز به تابع ارزش تقریبی بر پایه شبکه های عصبی برای هدایت انتخاب عمل بهینه احساس می شود.

معماری شبکه عصبی

معماری شبکه عصبی شامل 7 متغیر حالت $(S, H, D, W, O, r, O_{gap}, FQ)$ به عنوان ورودی و عمل $A(0.8 - 3.5)$ این حالت به عنوان خروجی میباشد، شکل 5 معماری شبکه عصبی DQN را نمایش میدهد



شکل 5 ساختار شبکه عصبی الگوریتم DQN

پارامترهای ساختار شبکه عصبی

- لایه های پنهان: دو لایه با اندازه 128 نورون
- gamma: 0.95
- Learning rate: 0.002 کاهش پس از طی کردن نیمی از اپیزود ها تا کف 0.0005 با نرخ 0.95
- ϵ : 1 کاهش در هر 100 اپیزود با نرخ 0.999 تا کف 0.05
- Buffer size: 100000
- تعداد اپیزود شبیه سازی: 7000
- به روزرسانی تابع هدف بعد از هر 10 اپیزود

روش به روزرسانی بر اساس معادله Bellman و با استفاده از تجربه بازپخش (Experience Replay) انجام می شود:

نرم افزارها و کتابخانه ها

پیاده سازی در زبان Python 3.10 انجام شد. کتابخانه های اصلی:

- PyTorch (برای شبکه عصبی و بهینه سازی)
- NumPy (برای عملیات عددی برداری)
- Pandas (برای مدیریت داده های شبیه سازی)
- Matplotlib / Seaborn (برای تجسم نتایج)

داده های شبیه سازی

3 پارکینگ با مدل ارائه شده شبیه سازی و سیاست بهینه برای هر یک حاصل گردیده است

داده های مسئله:

پارکینگ A با ظرفیت 1000 خودرو و هزینه تمام شده 120 واحد کیفیت 0.7 و فاصله 700 متری میانگین از نقاط مهم اطراف

پارکینگ B با ظرفیت 1500 خودرو و قیمت تمام شده 95 واحد و کیفیت 0.5 و فاصله 1400 متری میانگین از نقاط مهم

پارکینگ C با ظرفیت 1250 خودرو و قیمت تمام شده 70 واحد و کیفیت 0.3 و فاصله 1800 متری میانگین از نقاط مهم

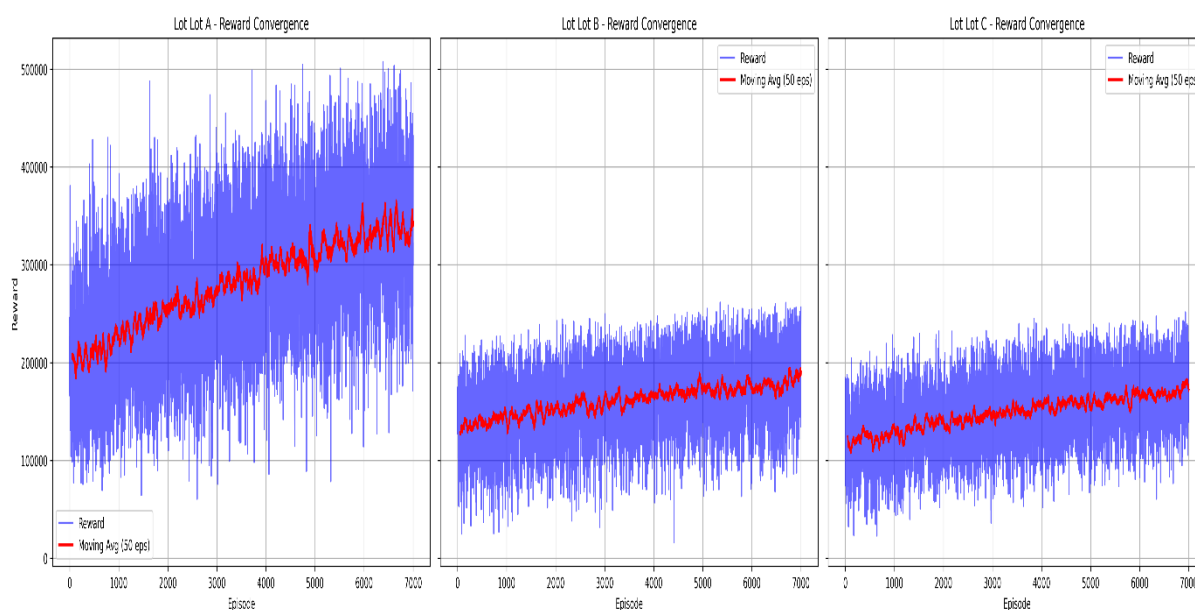
قیمت شروع پارکینگ ها 1.1 هزینه آن ها است

نتایج

در این بخش نمودار ها، سیاست ها و نتایج حاصل از شبیه سازی مسئله با روش ارائه شده در پژوهش بیان شده و مفصلاً به تفسیر و توضیح نتایج میپردازیم.

همگرایی

در سناریوی آموزش عوامل، فرآیند همگرایی پاداش برای سه پارکینگ Lot A، Lot B و Lot C طی حدود ۷۰۰۰ اپیزود بررسی گردید. شکل 6 نمایش همگرایی عامل ها را نمایش میدهد



شکل 6 نمودار های همگرایی عامل ها

نتایج نشان می دهد که با وجود محیط رقابتی و غیرایستا، هر سه عامل موفق به یادگیری سیاست های بهینه و رسیدن به محدوده های پایدار پاداش شده اند.

برای Lot A، روند میانگین متحرک پاداش (۵۰ اپیزود) از ابتدای آموزش تا حدود اپیزود ۵۰۰۰ افزایشی پایدار داشته و در نهایت به محدوده ای بالاتر از ۳۰۰,۰۰۰ واحد پاداش همگرا شده است. دامنه نوسان پاداش در این پارکینگ اگرچه گسترده است، اما نسبت به مقدار میانگین پاداش به مرور متعادل تر شده که نشان دهنده افزایش بهره برداری از سیاست های کارآمد در

کنار حفظ انعطاف‌پذیری برای پاسخ به تغییرات محیطی است. برتری مطلق Lot A در مقایسه با دو پارکینگ دیگر را می‌توان به ظرفیت بیشتر و استراتژی قیمت‌گذاری تهاجمی آن نسبت داد.

Lot B روندی آرام‌تر ولی نسبتاً یکنواخت در رشد پاداش داشته و میانگین آن در پایان آموزش به حدود ۱۸۰'۰۰۰ واحد رسیده است. نوسانات نسبت به Lot A کمتر بوده که حاکی از ثبات بیشتر استراتژی در تعامل با تغییرات بازار است. شکل منحنی و رسیدن زودهنگام به فاز فلات نشان می‌دهد که این عامل سریع‌تر به یک سیاست نسبتاً پایدار دست یافته است.

در Lot C نیز گرچه میانگین نهایی پاداش حدود ۱۵۰'۰۰۰ واحد و از دو پارکینگ دیگر کمتر است، اما روند افزایشی آن نسبت به آغاز آموزش قابل ملاحظه بوده و دامنه نوسان مشابه Lot B است. این امر را می‌توان به ظرفیت پایین‌تر و تمرکز بر جذب بخش حساس به قیمت تقاضا نسبت داد که امکان سودآوری بالا ولی پایدار را محدود می‌سازد.

به طور کلی، نتایج همگرایی تأیید می‌کند که الگوریتم یادگیری تقویتی عمیق در محیط چندعاملی توانسته است در حضور پویایی رقابتی و تغییرات غیردقیق‌پذیر، سیاست‌هایی با توازن بین سودآوری و نرخ اشغال بیاموزد. حفظ بخشی از نوسانات پس از همگرایی نیز طبیعی و ناشی از ماهیت غیریستای محیط و تعاملات استراتژیک میان عوامل رقابتی است.

خروجی‌ها و نتایج مدل

نمودار 1 تغییرات سه پارکینگ (A، B و C) را در طول یک روز نشان می‌دهد و در واقع نمایانگر سیاست قیمتی است که عامل‌های یادگیرنده برای بیشینه‌سازی سود و کنترل نرخ اشغال آموخته‌اند.

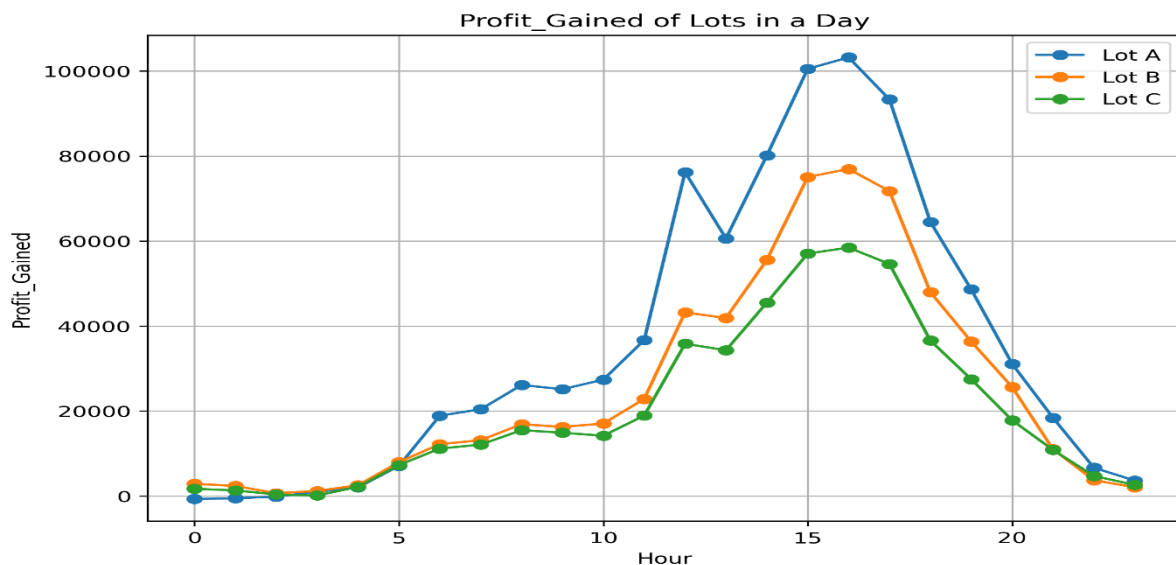


نمودار 1 سیاست قیمت‌گذاری بهینه آموخته شده

در ساعات ابتدایی شبانه‌روز، هر سه پارکینگ با قیمت‌های نسبتاً پایین فعالیت می‌کنند؛ به جز پارکینگ B که قیمت ثابت و بالاتری نسبت به دو رقیب دارد، احتمالاً به دلیل موقعیت یا کیفیتی که کشش قیمتی پایین‌تری ایجاد می‌کند. با نزدیک شدن به ساعات آغاز تقاضای جدی (حدود ساعت 4 تا 6)، هر سه سیاست جهش قیمتی ثبت می‌کنند، اما دامنه این جهش در پارکینگ A بسیار بیشتر است و تا بازه‌های 370–420 واحد در ساعات پیک بعدازظهر باقی می‌ماند. پارکینگ B نیز پس از مدتی افزایش ملایم، از ساعت 12 به بعد وارد سطح بالاتری می‌شود، در حالی که پارکینگ C به‌طور محافظه‌کارانه در میانه نمودار حرکت می‌کند و کمترین نوسان را دارد. این الگوها نشان می‌دهد که پارکینگ A سیاستی تهاجمی و متکی بر جذب حداکثر

درآمد در لحظات اوج اجرا می‌کند، پارکینگ B با کمی تأخیر خود را به موج قیمتی می‌رساند تا بازار را از دست ندهد، و پارکینگ C ثبات را بر نوسان ترجیح داده و شاید به دنبال حفظ سهم ثابت مشتریان وفادار است. در مجموع، این نمودار به روشنی بیانگر سازوکار بهینه‌سازی مبتنی بر یادگیری تقویتی است که در آن قیمت به‌عنوان اهرم اصلی تنظیم تقاضا و تعادل اشغال در طول روز عمل می‌کند.

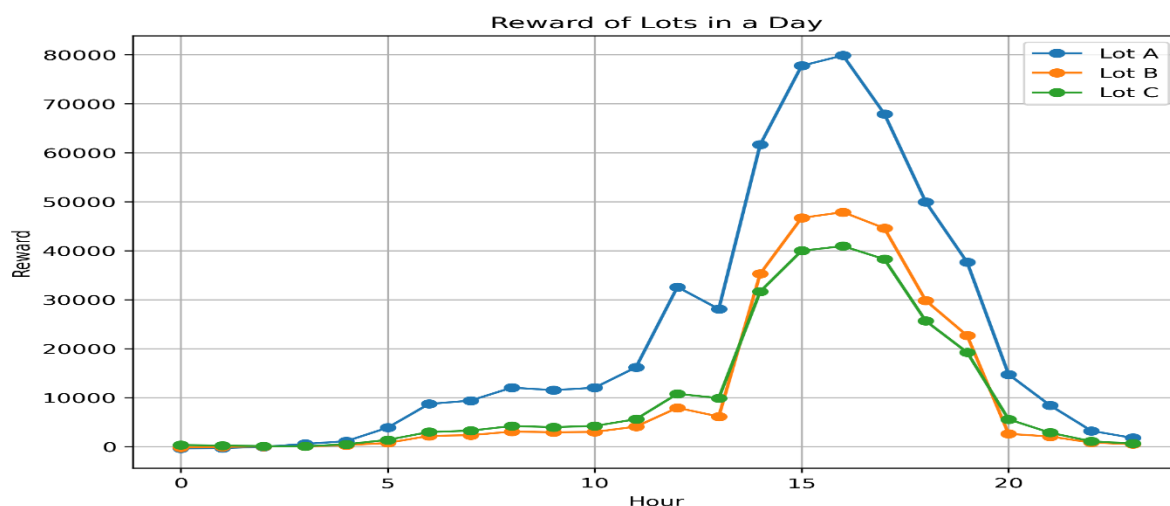
نمودار 2 روند تغییرات سود سه پارکینگ رقیب را در طول 24 ساعت نمایش می‌دهد و بیانگر اثر مستقیم الگوی تقاضا و سیاست‌های قیمت‌گذاری پویا بر عملکرد مالی هر پارکینگ است.



نمودار 2 میزان سود هر پارکینگ

در ساعات اولیه شبانه‌روز (0 تا 5 بامداد) سود هر سه پارکینگ در پایین‌ترین سطح قرار دارد که نتیجه کاهش محسوس تقاضای پایه در این بازه است. با شروع صبح و افزایش تدریجی حجم تقاضا (5 تا 11 صبح)، سود به‌صورت پیوسته رشد می‌کند و تا حوالی ظهر تفاوت معناداری بین رقبا مشاهده نمی‌شود. از ساعت 12 با آغاز بازه پیک، نرخ رشد سود شتاب می‌گیرد و در فاصله 14 تا 17 عصر به بیشینه روزانه می‌رسد؛ در این مرحله پارکینگ A با دستیابی به حدود ۱۰۲ هزار واحد، عملکرد بهتری نسبت به B و C نشان می‌دهد که حاکی از اثربخشی بالاتر سیاست‌های قیمت‌گذاری و مدیریت اشغال آن است. پس از ساعت 17، همگام با کاهش تقاضا، روند نزولی در سود آغاز می‌شود و تا پایان روز به مقادیر نزدیک به صفر در بازه شبانه بازمی‌گردد. این رفتار کلی با مفروضات مدل محیطی (شامل توزیع ساعتی تقاضا، کشش قیمتی و وزن‌دهی بالاتر به ساعات پیک در تابع پاداش) همخوان بوده و بیانگر توانایی مدل در بازتولید واقع‌گرایانه پویایی‌های سود در بازار رقابتی پارکینگ شهری است.

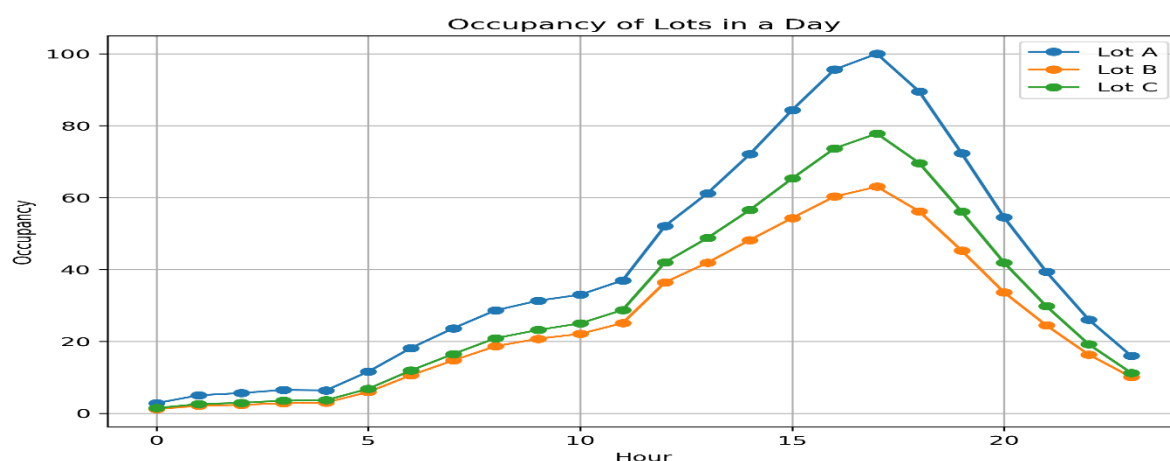
نمودار 3 تغییرات پاداش سه پارکینگ A، B و C را در طول ۲۴ ساعت به تصویر می‌کشد و نشان می‌دهد که عامل‌های یادگیرنده چگونه در پاسخ به شرایط بازار و هدف‌گذاری تابع پاداش عمل کرده‌اند.



نمودار 3 میزان پاداش عامل هر پارکینگ

در ساعات اولیه بامداد (۰ تا ۴)، پاداش هر سه پارکینگ نزدیک صفر و پایدار است که حاصل تقاضای کم و نبود انگیزه برای تغییرات عمده قیمتی است. از حدود ساعت ۵ به بعد، بخصوص برای پارکینگ A، روند افزایشی ملایمی آغاز می‌شود که هم‌راستا با رشد تدریجی تقاضا و بهبود اشغال است. نقطه تحول اصلی حوالی ساعت ۱۲ تا ۱۳ رخ می‌دهد، جایی که هر سه پارکینگ به‌واسطه ورود به بازه پیک و واکنش به سناریوهای با ارزش پاداش بالا، جهش قابل توجهی در پاداش تجربه می‌کنند. در این بازه، تفاوت بین سیاست‌ها به‌خوبی نمایان است: پارکینگ A سیاستی تهاجمی‌تر در قیمت‌گذاری و اشغال دارد و تا سقف نزدیک ۸۰ هزار واحد پاداش در ساعت ۱۵-۱۶ می‌رسد، در حالی که پارکینگ B و C با زمان‌بندی نزدیک اما دامنه کمتر، به ترتیب حدود ۴۹ و ۴۲ هزار واحد را ثبت می‌کنند. پس از ساعت ۱۷، کاهش تقاضای بازار موجب افت سریع پاداش‌ها می‌شود و تا حوالی ساعت ۲۱ به شدت کاهش یافته و در پایان روز تقریباً به همان سطح ابتدایی بازمی‌گردد. الگوی کلی نشان می‌دهد که مدل پاداش — با ترکیب سود مالی، کاهش انحراف از نرخ اشغال هدف، و جریمه برای رد کردن مشتری — به‌خوبی توانسته عامل‌ها را به سمت بیشینه‌سازی کنترل‌شده عملکرد در ساعات با ارزش بیشتر هدایت کند.

نمودار 4 روند تغییر نرخ اشغال سه پارکینگ A، B و C را طی یک شبانه‌روز نمایش می‌دهد و بیانگر تأثیر مشترک سیاست قیمت‌گذاری پویا و الگوی تقاضای ساعتی بر میزان استفاده از فضاهای پارک است.

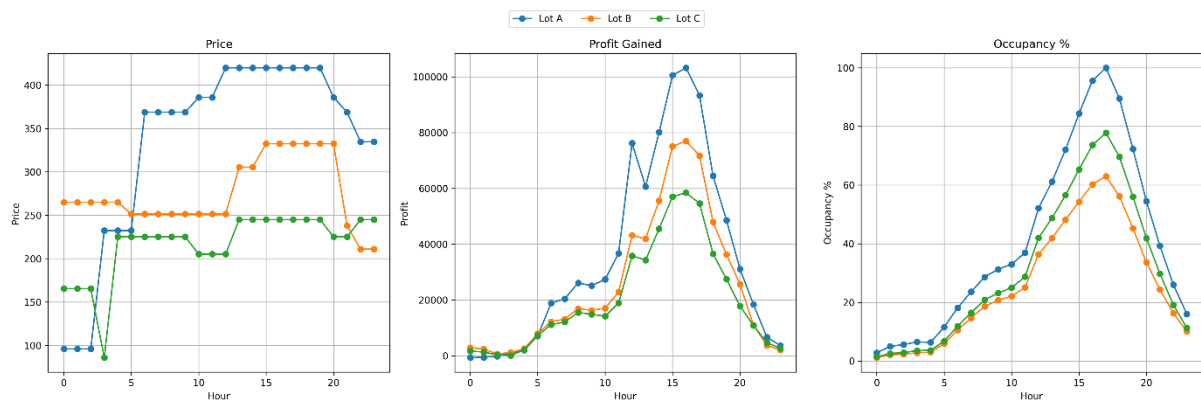


نمودار 4 میزان پر بودن ظرفیت هر پارکینگ

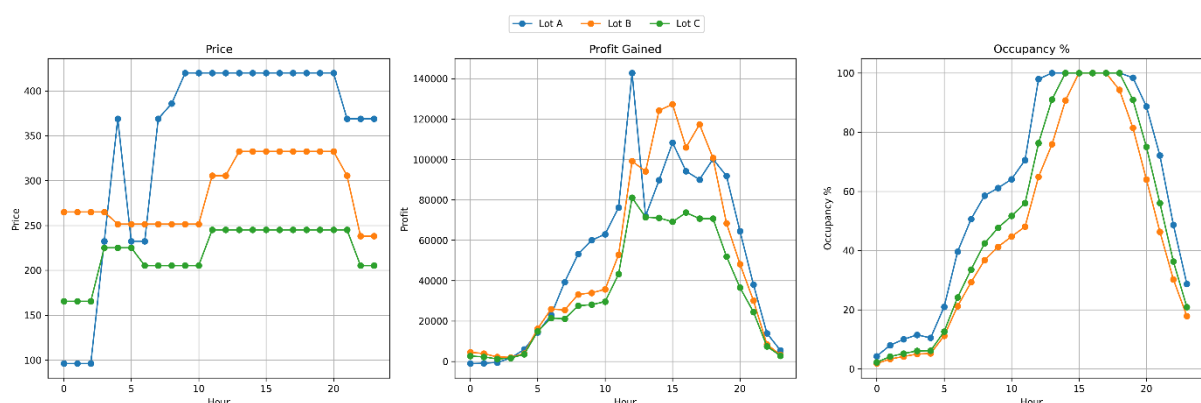
در ابتدای روز (۰ تا ۴ صبح)، اشغال هر سه پارکینگ در سطوح بسیار پایین و پایدار حرکت می‌کند که با حداقل تقاضای شبانه همخوان است. از ساعت ۵، هر سه منحنی به تدریج شیب مثبت پیدا می‌کنند و تا حوالی ظهر با رشد ملایم، وارد محدوده‌های متوسط اشغال می‌شوند. پس از آن، با ورود به بازه پیک تقاضا (حدود ساعت ۱۲ تا ۱۷)، نرخ اشغال با شیبی تند افزایش یافته و برای پارکینگ A به سقف ۱۰۰٪ در ساعت ۱۷ می‌رسد، در حالی که پارکینگ C حدود ۷۵٪ و پارکینگ B نزدیک به ۶۵٪ اشغال را ثبت می‌کند. این تفاوت‌ها نشان می‌دهد که A توانسته با ترکیب قیمت و مدیریت ظرفیت، در جذب تقاضای بیشتر نسبت به رقبای موفق‌تر عمل کند، در حالی که B و C سیاست‌های ملایم‌تری اتخاذ کرده‌اند که احتمالاً برای جلوگیری از ازدحام یا حفظ نرخ اشغال هدف طراحی شده است. پس از ساعت ۱۸، هر سه منحنی به‌طور همزمان روند کاهشی را تجربه می‌کنند و تا نیمه‌شب به سطوح پایین اولیه بازمی‌گردند. این الگو به خوبی با رفتار مدل محیط که در آن انحراف از اشغال بهینه جریمه می‌شود و تقاضا تابعی از زمان، قیمت نسبی، و کشش بازار است، انطباق دارد.

سناریوهای پردازی و تحلیل حساسیت

در نمودارهای ۵ و ۶ گراف قیمت، سود و شلوغ بودن در شرایط عادی و در زمانی که تقاضا دوبرابر شده است را نشان می‌دهد



نمودار ۵ سناریو تقاضای عادی



نمودار ۶ سناریو تقاضای افزایش یافته

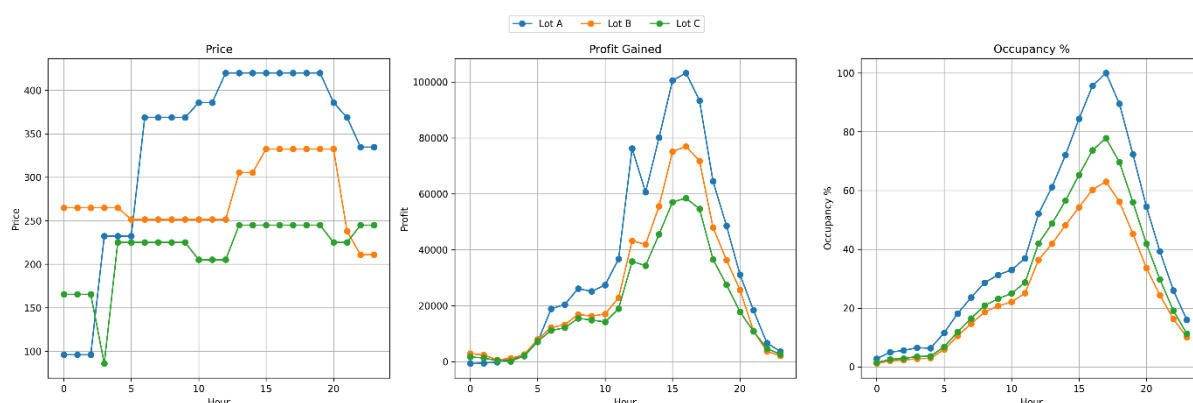
الگوی قیمت‌ها نسبت به حالت تقاضای نرمال تقریباً ثابت مانده است. این امر نشان می‌دهد که عامل‌های یادگیرنده لزوماً نیازی به افزایش قیمت در شرایط وفور تقاضا ندیده‌اند و بیشتر بر مدیریت ظرفیت برای بیشینه‌سازی سود تمرکز کرده‌اند. پارکینگ A همچنان مشی قیمتی تهاجمی با سقف ثابت طولانی‌مدت دارد، در حالی که B و C در سطوح پایین‌تر، ولی پایدار حرکت می‌کنند.

افزایش تقاضا باعث جهش محسوس سود همه پارکینگ‌ها در ساعات پیک شده است. جالب اینجاست که بر خلاف سناریوی نرمال، در اینجا Lot B حتی در برخی ساعات سودی بالاتر از A ثبت می‌کند (تا حدود ۱۲۵ هزار واحد)، که احتمالاً ناشی از پرشدن سریع ظرفیت B با نرخ قیمتی بهینه‌تر و کاهش نرخ «مشتري رد شده» است. همچنین Lot C که معمولاً در جایگاه سوم قرار داشت، در این حالت به سودی نزدیک به Lot A می‌رسد، زیرا تقاضای مازاد بازار تقریباً ظرفیت تمام پارکینگ‌ها را پر کرده است.

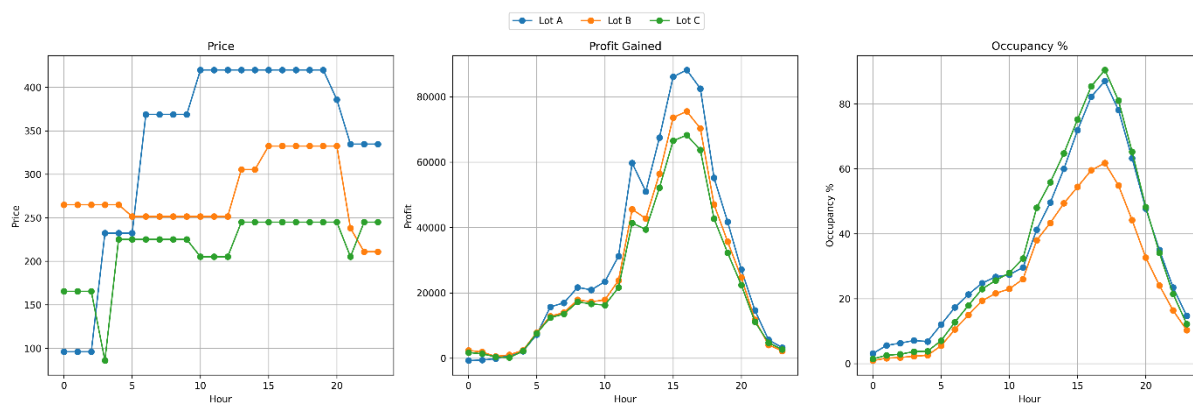
تقریباً همه پارکینگ‌ها در بازه‌ی ۱۳ تا ۱۸ به سقف ۱۰۰٪ اشغال رسیده‌اند. این وضعیت بیانگر آن است که عرضه به‌طور کامل مصرف شده و تفاوت در سود بین بازیگران صرفاً از تفاوت قیمت و توزیع زمانی پرشدن حاصل می‌شود. Lot C به‌رغم ظرفیت کمتر، به دلیل پرشدن کامل، عملکرد نسبی بهتری پیدا کرده است.

در شرایط عرضه محدود و تقاضای افزایشی، ظرفیت فیزیکی و استراتژی قیمتی تعیین‌کننده نهایی سودآوری هستند. این نتایج نشان می‌دهد که در مدل شما، وقتی بازار اشباع می‌شود، مزیت ظرفیت بالا به‌شدت تقویت می‌شود، اما بازیگرانی که قیمتی بهینه بین حداکثرسازی درآمد و جلوگیری از ردشدن تقاضا انتخاب می‌کنند (مثل Lot B در این سناریو)، می‌توانند حتی از دارنده بیشترین ظرفیت نیز پیشی بگیرند.

در نمودارهای ۷ و ۸ به مقایسه حساسیت متوسط و بسیار بالای رانندگان به قیمت می‌پردازیم



نمودار ۷ سناریو حساسیت به قیمت عادی



نمودار ۸ سناریو حساسیت به قیمت افزایش یافته

ساختار قیمتی هر پارکینگ مشابه سناریوهای قبل است، اما اثر آن به دلیل دوبرابر شدن حساسیت تقاضا، بسیار متفاوت مشاهده می‌شود.

Lot A با بالاترین قیمت‌ها (حتی در حالت پایه بیش از ۳۵۰ و در اوج ۴۲۰) به همان شدت قبل در بازار رقابت نمی‌کند، چون مشتریان با حساسیت بالا به سمت گزینه‌های ارزان‌تر تمایل دارند.

Lot C با پایین‌ترین قیمت (حدود ۱۰۰ تا ۲۵۰) در این شرایط بسیار جذاب‌تر شده و موفق به جذب سهم بزرگی از تقاضا شده است. در شرایط حساسیت بالای مشتریان، برتری سود لزوماً با بالاترین قیمت به دست نمی‌آید. Lot B و Lot C به یکدیگر نزدیک شده‌اند و سودشان در ساعاتی از روز تقریباً با Lot A رقابت می‌کند.

سود Lot C حالا با وجود پایین‌تر بودن نرخ متوسط بلیت، در حوالی ساعت ۱۵-۱۷، به سقف تاریخی خود می‌رسد، چون توانسته ظرفیت اشغال بسیار بالایی را برای مدت طولانی حفظ کند.

Lot B که ظرفیت متوسط و قیمت میانه دارد، نرخ اشغال نسبتاً پایدار ولی پایین‌تر از C را تجربه می‌کند.

وقتی کشش قیمتی تقاضا دو برابر می‌شود، رقابت قیمتی اهمیت بیشتری پیدا می‌کند و مزیت ظرفیت بالای Lot A بدون تعدیل قیمت رنگ می‌بازد.

نتیجه گیری و بحث

این پژوهش یک چارچوب یادگیری تقویتی عمیق چندعاملی (MARL) را برای قیمت‌گذاری پویا در بازارهای رقابتی پارکینگ شهری ارائه کرد که در آن از مدل لاجیت چندگانه (MNL) برای تخصیص تقاضا استفاده شده و ماهیت غیرایستای محیط رقابتی ناشی از تغییر سیاست رقبا به‌طور صریح مدل‌سازی شده است. نوآوری اصلی کار در تعریف محیط به‌صورت یک فرایند تصمیم‌گیری مارکوف با عوامل رقابتی، تخصیص تقاضای مبتنی بر نرخ اشغال، و تابع پاداش ترکیبی است که میان حداکثرسازی سود، رعایت نرخ اشغال هدف، و حداقل‌سازی تعداد درخواست‌های رد شده تعادل برقرار می‌کند.

آزمایش‌های شبیه‌سازی در سناریوهای مختلف شدت تقاضا و حساسیت قیمتی نشان دادند که چارچوب پیشنهادی قادر است سیاست‌های قیمت‌گذاری پایداری را بیاموزد که با شرایط بازار سازگار می‌شوند. در شرایط تقاضای پایه، تعدیل تهاجمی ولی دقیق نرخ‌ها به سودآوری بالاتر بدون افت نرخ اشغال ساعات اوج انجامید. دو برابر شدن تقاضای پایه، سودآوری را در همه پارکینگ‌ها افزایش داد و اپراتورهای با ظرفیت بالاتر بیشترین منفعت را بردند. در مقابل، دو برابر شدن حساسیت قیمتی باعث انتقال سهم بازار به سمت اپراتورهای کم‌هزینه شد که نشان‌دهنده واکنش‌پذیری مدل نسبت به کشش قیمتی مشتریان است. تحلیل همگرایی پاداش‌ها تأیید کرد که همه عوامل به استراتژی‌های عملیاتی پایدار رسیده‌اند، هرچند نوسانات ناشی از ماهیت رقابتی و غیرایستای محیط باقی‌مانده است.

مقایسه با پژوهش‌های پیشین، مانند مدل‌های تک‌عاملی مبتنی بر Q-learning برای قیمت‌گذاری پارکینگ (مانند Sowmya & Dhabu, 2023) یا رویکردهای برنامه‌ریزی پویای تقریبی (ADP) همراه با سیستم رزرو (مانند Ouyang, 2017 & Lei)، نشان می‌دهد که رویکرد حاضر سه مزیت شاخص دارد:

1. یادگیری رقابتی چندعاملی — عوامل به صورت هم‌زمان سیاست‌های خود را در یک محیط مشترک تکامل می‌دهند و تعاملات استراتژیک را بهتر بازتاب می‌دهند.

2. یکپارچه‌سازی مکانیزم تخصیص تقاضا — انتخاب مشتریان به‌طور درون‌زا و بر اساس MNL مدل‌سازی شده است و ارتباط قیمت‌گذاری با فرآیند واقعی تصمیم‌گیری سفر برقرار است.

3. انعطاف‌پذیری نسبت به تغییرات سناریو — آموزش در شرایط متنوع بازار، قابلیت تعمیم مدل به شوک‌های ناشی از تغییر شدت تقاضا و کشش قیمتی را بالا می‌برد.

از نظر عملی، این چارچوب در شهرهایی که دارای سیستم سنجش لحظه‌ای نرخ اشغال، پرداخت دیجیتال و سامانه متمرکز مدیریت پارکینگ هستند، قابل اجراست. برای بهره‌برداری واقعی، لازم است سیاست‌های شهری امکان تعدیل پویا در نرخ پارکینگ را فراهم کنند و این فرآیند با شفافیت و پذیرش اجتماعی همراه باشد. نیاز محاسباتی این مدل با زیرساخت‌های ابری و لبه شبکه جدید سازگار بوده و تصمیم‌ها در زمان نزدیک به واقعی (near real-time) به‌روزرسانی می‌شوند.

با این حال، اتکای مطالعه بر داده‌های مصنوعی و ساده‌سازی رفتار رانندگان از محدودیت‌های این رویکرد به شمار می‌رود. اگرچه ساختار MNL اثر قیمت و نرخ اشغال را لحاظ می‌کند، اما تنوع کامل رفتاری و عوامل محیطی پیچیده‌تر (مانند رویدادهای خاص یا تغییرات ناگهانی آب‌وهوا) در آن پوشش داده نشده است. همچنین مقیاس‌پذیری مدل در شبکه‌های بزرگ با ده‌ها یا صدها واحد پارکینگ رقابتی، نیازمند بهینه‌سازی بیشتر الگوریتم و روش‌های آموزش توزیع‌شده است.

برای آینده، استفاده از داده‌های واقعی پارکینگ شهری برای کالیبراسیون، گسترش فضای حالت به منظور پوشش تعاملات حمل‌ونقل چندوجهی، و سنجش اثرات اجتماعی و زیست‌محیطی (مانند کاهش تراکم ترافیک جستجوی پارکینگ و کاهش انتشار CO₂) توصیه می‌شود. چنین توسعه‌هایی، علاوه بر افزایش واقع‌گرایی مدل، مسیر به‌کارگیری آن را به عنوان یک ابزار پشتیبان تصمیم برای شهرداری‌ها و اپراتورهای خصوصی هموار خواهد کرد و پیوندی عملی میان بهینه‌سازی محاسباتی و سیاست‌گذاری حمل‌ونقل پایدار برقرار می‌سازد.

- Jioudi, B., Amari, A., Moutaouakkil, F., & Medromi, H. (2019). e-parking: Multi-agent smart parking platform for dynamic pricing and reservation sharing service. *International Journal of Advanced Computer Science and Applications*, 10(11).
- Kalakanti, A. K., & Rao, S. (2025). Dynamic pricing for electric vehicle charging. *Proceedings of the 40th ACM/SIGAPP Symposium on Applied Computing*, 1609–1611.
- Kastius, A., & Schlosser, R. (2022). Dynamic pricing under competition using reinforcement learning. *Journal of Revenue and Pricing Management*, 21(1), 50–63.
- Lei, C., & Ouyang, Y. (2017). Dynamic pricing and reservation for intelligent urban parking management. *Transportation Research Part C: Emerging Technologies*, 77, 226–244.
- Li, Y., Zhang, Z., & Xing, Q. (2025). Real-time online charging control of electric vehicle charging station based on a multi-agent deep reinforcement learning. *Energy*, 319, 135095.
- Lim, J., & Masoud, N. (2024). Dynamic usage allocation and pricing for curb space operation. *Transportation Science*, 58(6), 1252–1276.
- Liu, D., Wang, W., Wang, L., Jia, H., & Shi, M. (2021). Dynamic pricing strategy of electric vehicle aggregators based on DDPG reinforcement learning algorithm. *IEEE Access*, 9, 21556–21566.
- Nakabi, T. A., & Toivanen, P. (2021). Deep reinforcement learning for energy management in a microgrid with flexible demand. *Sustainable Energy, Grids and Networks*, 25, 100413.
- Qian, Z. S., & Rajagopal, R. (2014). Optimal occupancy-driven parking pricing under demand uncertainties and traveler heterogeneity: A stochastic control approach. *Transportation Research Part B: Methodological*, 67, 144–165.
- Sowmya, K., & Dhabu, M. M. (2023). Model free Reinforcement Learning to determine pricing policy for car parking lots. *Expert Systems with Applications*, 230, 120532.
- Wang, J., Wang, H., Zhang, X., Hu, H., & Peng, W. (2024). Optimizing public parking supply and pricing strategies in a competitive market with shared private parking services. *Transportation Research Part C: Emerging Technologies*, 166, 104774.
- Wang, P., Guan, H., & Liu, P. (2020). Modeling and solving the optimal allocation-pricing of public parking resources problem in urban-scale network. *Transportation Research Part B: Methodological*, 137, 74–98.
- Yan, X., Levine, J., & Marans, R. (2019). The effectiveness of parking policies to reduce parking demand pressure and car use. *Transport Policy*, 73, 41–50.

- Yin, C., & Han, J. (2021). Dynamic pricing model of e-commerce platforms based on deep reinforcement learning. *Computer Modeling in Engineering & Sciences*, 127(1), 291–307.
- Zheng, N., & Geroliminis, N. (2016). Modeling and optimization of multimodal urban networks with limited parking and dynamic pricing. *Transportation Research Part B: Methodological*, 83, 36–58.