

## مدل بهینه‌سازی نرخ ورودی رمپ، سرعت مجاز متغیر و عوارض بزرگراهی

### مبثنی بر یادگیری تقویتی

#### مقاله علمی - پژوهشی

شهریار افندی زاده\*، استاد، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، تهران، ایران

محمدعلی حسن زاده، دانشجوی کارشناسی ارشد، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، تهران، ایران

\* پست الکترونیکی نویسنده مسئول: zargari@iust.ac.ir

دریافت: ۱۴۰۱/۰۹/۱۴ - پذیرش: ۱۴۰۲/۰۴/۲۸

صفحه ۳۰-۱۳

#### چکیده

رشد بدون توقف عبور و مرور شهری به خصوص در بزرگراه‌ها منجر به ایجاد تداخلات ترافیکی و در نتیجه ایجاد صف‌های طولانی، افزایش زمان سفر، افزایش تصادفات و انتشار آلودگی‌های زیست‌محیطی شده است. از آنجایی که توسعه‌ی گسترده‌ی زیرساخت‌ها نمی‌تواند راهکاری دائمی برای حل این معضلات باشد می‌بایست از تسهیلات موجود بهره‌برداری بهینه صورت گیرد. روش‌های متعددی از جمله کنترل نرخ ورودی رمپ، سرعت مجاز متغیر، کنترل تغییر خط، اعمال عوارض و ... تا امروز برای مدیریت بزرگراهی مورد مطالعه قرار گرفته و اثربخش بوده‌اند. یکی از رویکردها برای بهبود عملکرد این کنترل‌ها پیاده‌سازی هم‌زمان و هماهنگ آن‌هاست. در این مقاله به بررسی اثر پیاده‌سازی هماهنگ سه کنترل نرخ ورودی رمپ، سرعت مجاز متغیر و عوارض بزرگراهی بر شرایط ترافیکی شبکه بزرگراهی پرداخته شده است. برای این منظور از روش یادگیری تقویتی و عامل بهینه‌ساز مبدائی سیاست جهت مدیریت مقادیر کنترلی به صورت متغیر استفاده شده است. داده‌های ترافیکی ابتدا به وسیله‌ی مدل‌های حافظه بلندمدت - کوتاه‌مدت پیش‌بینی شده و سپس به شبیه‌سازی نرم‌افزار سومو داده می‌شوند. مدل کنترل‌گر مبتنی بر یادگیری تقویتی سپس بر اساس خروجی‌های شبیه‌سازی، مقادیر نرخ رمپ، سرعت مجاز و عوارض را انتخاب و بر روی شبکه اعمال می‌نماید. نتایج پیاده‌سازی مدل پیشنهادی تأثیر بسزای آن در بهبود پارامترهای ترافیکی همچون میانگین زمان سفر، میانگین سرعت خودروها، میانگین زمان انتظار، طول صف، میانگین تراکم شبکه و تعداد خودروهای رسیده به مقصد را نشان می‌دهد. این مدل میانگین زمان سفر را به میزان ۱۱٫۵ درصد نسبت به شرایط بدون کنترل کاهش داد که مقدار بیشتری نسبت به مدل‌های یک یا دو کنترل‌ه می‌باشد. ضمن اینکه در شرایط رخ دادن تصادف در شبکه نیز شرایط نسبت به وضعیت بدون کنترل به شکل قابل‌ملاحظه‌ای بهتر بوده است. این نتایج اثربخشی رویکرد ادغام کنترل‌ها در بهبود کارایی مدیریت بزرگراهی را نشان می‌دهد.

واژه‌های کلیدی: نرخ ورودی رمپ، سرعت مجاز متغیر، عوارض بزرگراهی، یادگیری تقویتی، شبکه‌ی عصبی

#### ۱- مقدمه

فزاینده‌ی تقاضای جاده‌ای می‌باشند. این افزایش منجر به ایجاد تداخلات ترافیکی شدید و افزایش زمان سفر شده است. چنین تداخلاتی ظرفیت کل سیستم را کاهش داده و از زیرساخت‌های گران بزرگراه دقیقاً در مواقعی که مورد نیاز هستند (ساعت‌های اوج مصرف)، استفاده ناکافی می‌شود

در دهه‌های گذشته عوامل متعددی منجر به افزایش میزان حضور خودروهای شخصی در جاده‌ها و به‌طور کلی رشد استفاده از مدهای حمل‌ونقل جاده‌ای شده است. افزایش کلی جمعیت، توسعه‌ی شهرنشینی، پیشرفت فناوری خودروسازی، کاهش هزینه‌ی نگهداری خودرو و ... همگی از دلایل این رشد

در مقایسه با مدل‌های قیمت‌گذاری تراکم ثابت، مدل‌های قیمت‌گذاری تراکم پویا به دلیل عدم قطعیت، تصادفی بودن و ویژگی‌های متغیر با زمان دستگاه‌های حمل‌ونقل، منطقی‌تر و مؤثرتر هستند (Cheng et al., 2017). کاربرد اعمال عوارض برای مدیریت بزرگراه در مطالعات متعددی مورد بررسی قرار گرفته است و نتایج قابل قبولی نیز داشته است. برای مثال در مطالعه آقای آرانگو (Arango, 2019) به کمک یادگیری تقویتی<sup>۵</sup> و باهدف افزایش جریان کل مسیر از یک الگوریتم شبکه عصبی عمیق دوگانه<sup>۶</sup> برای اخذ عوارض متغیر به ترافیک استفاده شده که بر اساس شبیه‌سازی در سومو افزایش ۷,۷ درصدی جریان و کاهش احتمال رخ دادن صف در شبکه را در پی داشته است؛ بنابراین عوارض بزرگراهی می‌تواند یک مکمل مناسب برای کنترل رمپ و سرعت مجاز متغیر در مدیریت بزرگراهی باشد.

کنترل نرخ ورودی رمپ یک روش مدیریت رمپ ورودی به بزرگراه با استفاده از چراغ‌های ترافیکی برای کنترل تعداد خودروهای وارد شونده به بزرگراه‌ها است. این چراغ‌ها معمولاً با چرخه کوتاه‌تری برنامه‌ریزی شده‌اند تا به یک یا چند وسیله نقلیه (معمولاً دو یا سه) در هر فاز سبز اجازه ورود به مسیر اصلی دهند. نرخ ورود خودروها معمولاً بر اساس حجم ترافیک و سرعت در بزرگراه‌ها و میزان صف موجود در رمپ انتخاب می‌گردد. بهترین رویکرد برای مدیریت رمپ، کنترل رمپ پاسخگو به ترافیک است که در آن نرخ‌های ورودی در پاسخ به شرایط ترافیکی به رمپ‌ها اختصاص می‌یابد (Shaaban, asif khan and Hamila, 2016). همچنین کنترل رمپ را می‌توان به دو صورت محلی و هماهنگ اجرا کرد. روش محلی صرفاً با هدف بهبود عملکرد شبکه در یک ناحیه خاص صورت می‌پذیرد در صورتی که در روش هماهنگ تأثیر کنترل اعمال شده بر نواحی مجاور نیز بررسی شده و کنترل باهدف بهبود عملکرد کلی سیستم تنظیم می‌گردد (Mizuta et al., 2014). مطالعات متعددی برای پیاده‌سازی کنترل رمپ پاسخگو انجام شده است. یکی از پرکاربردترین روش‌های کنترل رمپ مدل آلینا<sup>۷</sup> است که توسط پاپاگرچو و همکاران ارائه شده است (Papageorgiou, Hadj Salem and Blosseville, 1991). این روش که یک الگوریتم کنترل بازخوردی<sup>۸</sup> پاسخگو به ترافیک محلی است سطح اشغال را از آشکارسازهای نصب شده در بزرگراه به‌عنوان ورودی گرفته و

(Carlson et al., 2010). یک روش برای کاهش تداخلات توسعه‌ی زیرساخت‌های ترافیکی می‌باشد. با توجه به افزایش توقف‌ناپذیر خودروها امروزه در بسیاری موارد دیگر امکان کمتری برای ساخت‌وسازهای جدیدتر می‌باشد و باید بین به حداکثر رساندن ارایه عرضه و حفظ بهره‌وری تعادل برقرار شود (Mahajan et al., 2015). ضمن اینکه ایجاد امکانات و زیرساخت‌های جدید خود می‌تواند منجر به استفاده‌ی بیشتر افراد از حمل‌ونقل جاده‌ای و در نتیجه سبب تشدید مشکلات ترافیکی بشود. یکی دیگر از روش‌های کاهش این تداخلات مدیریت و کنترل بزرگراه با هدف استفاده‌ی بهینه‌تر از تسهیلات موجود می‌باشد. با توجه به اینکه بزرگراه‌ها یکی از مهم‌ترین تسهیلات جابجایی در سطح شهر می‌باشند می‌بایست از تدابیر مناسب برای کنترل آن‌ها در ساعات اوج ترافیک بهره برد. نیاز به کنترل‌های ترافیکی منجر به ایجاد روش‌های متعددی برای این امر شده است که شامل کنترل نرخ ورودی رمپ<sup>۹</sup>، سرعت مجاز متغیر<sup>۱۰</sup>، عوارض بزرگراهی، کنترل تغییر خط<sup>۱۱</sup>، مسیریابی لحظه‌ای<sup>۱۲</sup> و ... می‌باشد. از میان این روش‌ها کنترل نرخ ورودی رمپ و سرعت مجاز متغیر بیشترین کارایی را از خود نشان داده‌اند. این دو روش تلاش می‌کنند ترافیک ورودی به بزرگراه را به نحوی مدیریت کنند که تداخلات ترافیکی در ناحیه ورودی کاهش یابد. در صورتی که برخی روش‌ها همچون عوارض بزرگراهی تلاش دارند مسیر خودروها را به مسیر خلوت‌تر تغییر دهند. این تفاوت رویکرد می‌تواند منجر به عملکرد مناسب این سه روش در صورت پیاده‌سازی هماهنگ شده شود. دریافت عوارض یک روش رایج است که ممکن است با هدف ایجاد درآمد برای بازیابی هزینه‌های ساخت‌وساز و نگهداری راه و همچنین مدیریت تراکم بزرگراهی با تغییر ابعاد زمانی و مکانی رفتارهای سفر و تصمیمات مسافران در مورد انتخاب روش سفر یا مسیر بکار گرفته شود (Zhang et al, 2019). این کار می‌تواند به سه روش ثابت، برنامه‌ریزی شده و پویا انجام شود. عوارض پویا نوعی از قیمت‌گذاری تراکم است که در آن عوارض در زمان واقعی به‌عنوان تابعی از شرایط ترافیکی فعلی متفاوت است برخلاف عوارض ثابت که در طول زمان ثابت می‌ماند و عوارض برنامه‌ریزی شده که در آن عوارض بسته زمان روز، روز هفته یا فصل طبق یک برنامه از پیش تعیین شده تغییر می‌کند (Lombardi, Santos and Annaswamy, 2021).

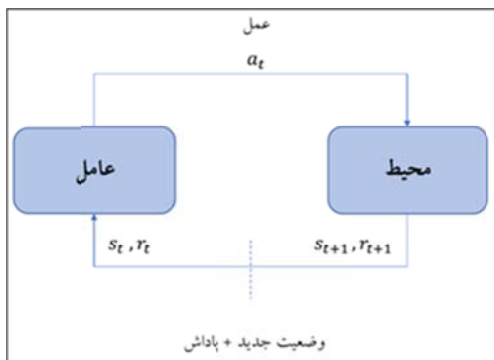
یادگیری تقویتی (Vinitsky et al., 2018) (Dumont and Li et al., (Kusic et al., 2020) (Van Vuuren, 2019) استفاده شده است. این مطالعات اثربخشی سرعت مجاز متغیر را در کاهش زمان سفر، تراکم شبکه و احتمال تصادف نشان می‌دهند. باوجود کارایی کنترل‌های ارایه شده در بسیاری از مطالعات، همچنان نیاز به راهکارهایی برای بهبود عملکرد روش‌های مدیریت بزرگراهی وجود دارد. پیاده‌سازی یکپارچه کنترل‌های ترافیکی مختلف یکی از راهکارهایی است که می‌تواند منجر به افزایش کارایی مدیریت اجرا شده بشود. یکی از نخستین مطالعات با این رویکرد از هیگی و همکاران منتشر شد (Hegy, De Schutter and Hellendoorn, 2005). مدل پیشنهادی این مطالعه که از روش کنترل مدل پیش‌بینی برای ادغام کنترل رمپ و سرعت استفاده می‌کند، جریان خروجی بالاتر و کل زمان سفر کمتری را نسبت به رویکرد کنترل منفرد داشت. روش‌های کنترل پیش‌بینی<sup>۱۱</sup> (Hegy, De Schutter and Hellendoorn, 2005) (Lu et al., 2011) و روش‌های کنترل بهینه (PaPamichail et al., 2008) (Ma et al., 2015) از پرکاربردترین الگوریتم‌های بکار رفته برای ادغام کنترل رمپ و سرعت مجاز متغیر می‌باشد. همچنین الگوریتم‌های بازخورد آبخاری<sup>۱۲</sup> (Carlson et al., 2012) و یادگیری تقویتی (Dumont and Van Vuuren, 2015) نیز برای این موضوع بکار گرفته شده‌اند. نتایج این مطالعات نشان می‌دهد به‌کارگیری کنترل ترکیبی رمپ و سرعت منجر به اثربخشی بیشتر نسبت به کنترل منفرد در کاهش زمان سفر و تداخلات ترافیکی شده است. همچنین طبق (Ma et al., 2015) افزودن کنترل سرعت به کنترل رمپ باعث کاهش طول صف و زمان انتظار شده است. این نتایج پیشنهاد می‌دهند که ترکیب روش‌های کنترلی جدید به‌صورت هماهنگ می‌تواند کارایی مدیریت بزرگراهی را افزایش دهد؛ بنابراین ادغام کنترل نرخ ورودی رمپ، سرعت مجاز متغیر و عوارض بزرگراهی می‌تواند منجر به یک روش مدیریت مفیدتر نسبت به ترکیب‌های ارایه شده در مطالعات پیشین بشود. در این مقاله ارایه‌ی مدل مدیریت بزرگراهی جدیدی متشکل از ترکیب این سه روش کنترلی مدنظر می‌باشد. این مدل به‌وسیله‌ی عامل بهینه‌ساز سیاست مبدائی<sup>۱۴</sup> (که یکی از عامل‌های یادگیری تقویتی می‌باشد) و در یک محیط شبیه‌سازی ترافیکی خردنگر (نرم‌افزار سومو) ساخته شده است.

نرخ ورودی رمپ را به‌عنوان یک متغیر کنترلی محاسبه می‌کند. همچنین پاپاگرچو و همکاران الگوریتم متلاین<sup>۹</sup> را که نوعی توسعه‌ی الگوریتم آلینا برای کنترل هماهنگ رمپ است ارایه نمودند (Papageorgiou, Blosseville and Hadj Salem, 1990). سایر روش‌های ارایه شده برای کنترل رمپ شامل الگوریتم کنترل رمپ هماهنگ شده‌ی اکتشافی<sup>۱۱</sup> ارایه شده توسط پاپامیشیل و همکاران (Papamichail et al., 2010)، الگوریتم منطق فازی-لوجیک<sup>۱۱</sup> توسط قدس و همکاران (Ghods, Kian and Tabibi, 2009) و الگوریتم بهینه‌سازی خطی استاتیک توسط خضرییان و همکاران (Khazraeian, Koohifar and Kalantari, 2017) می‌باشد. یادگیری تقویتی نیز از الگوریتم‌هایی می‌باشد که در کنترل رمپ قابل استفاده است. (Davarynejad et al., 2011) (Fares and Goma, 2014) (Belletti et al., 2017) (Abdelgawad, 2012) (Dumont and Van Vuuren, 2015) (Lu et al., 2011) و همکاران جهت کنترل رمپ با در نظر گرفتن طول صف یکی از نخستین تلاش‌ها برای به‌کارگیری یادگیری تقویتی در کنترل رمپ می‌باشد که عملکرد مناسبی در حفظ جریان نسبتاً برابر با ظرفیت داشته است. کنترل نرخ ورودی رمپ اثربخشی بالایی در بهبود عملکرد بزرگراه دارند اما محدودیت‌هایی نیز وجود دارد. به‌عنوان مثال از آنجاکه فضای رمپ‌ها محدود می‌باشند ممکن است صف‌های ایجاد شده در رمپ پس‌زده و به زیرساخت‌های مجاور سرازیر شوند. یکی دیگر از روش‌های مدیریت بزرگراهی که امکان کاهش طول صف را دارد، سرعت مجاز متغیر می‌باشد (Carlson, Papamichail and Papageorgiou, 2014). سیستم‌های محدودیت سرعت متغیر یکی از روش‌های هوشمند هستند که تغییر پویای محدودیت‌های سرعت را در پاسخ به شرایط ترافیکی و حوادث امکان‌پذیر می‌کنند. تغییرات محدودیت سرعت انتخاب شده با نمایشگرهای دارای قابلیت نمایش پیام متغیر نشان داده می‌شود (Khondaker and Katten, 2015). در این روش با کاهش سرعت مجاز بالادست محل تراکم، جریان ورودی به گلوگاه کاهش یافته و در نتیجه می‌تواند منجر به کاهش طول صف بشود. در مطالعات مختلف برای مدیریت تغییرات سرعت مجاز از روش‌های متعددی همچون الگوریتم ژنتیک-فازی-لوجیک (Chiou, Huang and Lin, 2012) و

## ۲- پیشینه تحقیق

استفاده می‌کند، توزیع احتمال شرطی حالات بعدی از این حالت خاص مستقل می‌ماند. از آنجایی که وضعیت ترافیکی آینده صرفاً بستگی به وضعیت ترافیکی اکنون و همچنین عمل انجام‌شده توسط کنترل‌گر دارد و به حالات قبل وابستگی ندارد، می‌توان فرض تابعیت محیط ترافیکی خود را از این موضوع را فرض نمود (Sewak, 2019).

در این مقاله برای مدیریت مقادیر کنترلی نرخ ورودی رمپ، سرعت مجاز متغیر و عوارض بزرگراهی، از کنترل‌گر مبتنی بر یادگیری تقویتی به کمک عامل بهینه‌ساز مبدائی سیاست استفاده شده است. همچنین برای پیش‌بینی احجام ترافیکی و داد آن به عنوان ورودی به مدل کنترل‌گر، مدل‌های شبکه عصبی سری زمانی حافظه کوتاه‌مدت - بلندمدت بکار رفته‌اند.



شکل ۱. حلقه‌ی یادگیری تقویتی

## ۱-۲- یادگیری تقویتی

یادگیری تقویتی شاخه‌ای از یادگیری ماشین است که در حل مسائل کنترلی با تصمیم‌گیری متوالی تخصص دارد. در این مسائل یک عامل وجود دارد که سعی دارد در ارتباط با محیط در هر مرحله بر اساس وضعیت محیط عمل یا عمل‌هایی را انجام دهد طوری که این عمل‌ها منجر به حداکثر شدن مفهوم معینی از پاداش‌های تجمعی شود. شکل ۱ ساختار عملکرد عامل یادگیری تقویتی را نشان می‌دهد که با دریافت یک وضعیت و انجام یک عمل، محیط را در وضعیت جدیدی قرار می‌دهد. در این روش عامل در یک دوره‌ی طولانی از تکرارها سعی می‌کند با آزمون و خطا یاد بگیرد که در زمان مشاهده‌ی یک وضعیت خاص، انجام چه عملی منجر به بدست آوردن پاداش بهینه می‌شود. این عامل در طول یادگیری به ایجاد یک سیاست بهینه برای انطباق وضعیت به عمل می‌پردازد (Sewak, 2019). در مسئله‌ی ما محیط شبکه‌ی ترافیکی موجود در شبیه‌سازی سومو می‌باشد که هر دو دقیقه یک وضعیت از شبکه را به مدل یادگیری تقویتی می‌دهد. سپس مدل سعی می‌کند با تغییر مقادیر کنترل‌ها شرایط شبکه را بهبود ببخشد. این کار را با تلاش برای افزایش پاداشی که برای آن تعریف‌شده است انجام می‌دهد.

می‌توان محیطی را فرض کرد که در آن رابطه‌ی ۱ برقرار است که در آن  $S$  بیانگر وضعیت و  $a$  عمل کنترل‌گر و  $t$  زمان می‌باشد. این قاعده رابطه‌ای است که وضعیت فعلی و عمل کنترل‌گر را به وضعیت آینده تبدیل می‌کند. اگر در هر زمان براساس وضعیت پاداشی به کنترل‌گر داده شود هدف عامل در یادگیری به حداکثر رساندن این پاداش است (Recht, 2019).

$$S_{t+1} = f(S_t, a_t) \quad (1)$$

$$\text{Maximize } E[\sum_{t=0}^n R_t(S_t, a_t)] \quad (2)$$

مدل یادگیری تقویتی معمولاً به کنترل‌کننده اجازه می‌دهد تا قبل از تصمیم‌گیری در مورد اقدام بعدی، وضعیت را مشاهده کند. این به یک کنترل‌کننده اجازه می‌دهد تا به‌طور مداوم عدم اطمینان را از طریق بازخورد کاهش دهد. از این‌رو، به‌جای بهینه‌سازی بر روی توالی‌های قطعی کنترل‌ها، در عوض روی سیاست‌ها، بهینه‌سازی صورت می‌گیرد. یک خط‌مشی کنترل یا به عبارتی سیاست، تابعی است که یک وضعیت را از محیط می‌گیرد و یک عمل کنترلی جدید را خروجی می‌دهد.

یکی از شروط به‌کارگیری یادگیری تقویتی، تابعیت محیط از فرآیند تصمیم‌گیری مارکوف می‌باشد. برای فرآیندی که چندین حالت را پشت سر گذاشته است و اکنون در یک وضعیت خاص قرار دارد، اگر این فرآیند از ویژگی مارکوف پیروی کند توزیع احتمال شرطی حالت احتمالی بعدی را صرف‌نظر از دنباله‌ای بیان می‌کنیم که برای رسیدن به این وضعیت فعلی طی شده است؛ بنابراین، حتی اگر چندین راه برای رسیدن به یک حالت خاص وجود داشته باشد، صرف‌نظر از اینکه فرآیند برای رسیدن به این حالت از کدام راه خاص

خط‌مشی‌های  $\tau_t$  متغیری از تصمیم‌ها و وضعیت‌های مسئله هستند (Recht, 2019).

$$\tau_t = (u_1, \dots, u_{t-1}, x_0, \dots, x_t) \quad (3)$$

مدل‌های یادگیری تقویتی برای تصمیم‌گیری به تشکیل یک سیاست پرداخته و به‌مرور به بهبود این سیاست می‌پردازند. در حل مسئله یادگیری تقویتی دو روش رایج برای برخورد با مسئله وجود دارد، روش‌های خط‌مشی و روش‌های ارزشی. در روش‌های خط‌مشی، سعی می‌شود مستقیماً خط‌مشی را بهینه نمود، درحالی‌که در روش‌های ارزشی سعی بر آن است، بازده مورد انتظار آینده را ارزیابی نمود و سیاست را از آنجا استخراج کرد. درحال حاضر گروه سوم از روش‌ها، روش‌های عامل-نقاد<sup>۱۵</sup> هستند که سعی دارند دو روش قبلی را به نحو معناداری باهم ترکیب کنند. (Lavet et al., 2018). از آنجایی‌که مدل‌های عملگر نقاد قابلیت انجام چندین عمل گسسته (نه یک عمل و نه چندین عمل پیوسته) را دارا هستند به شکل مناسب‌تری بر روی محیط ما قابل اجرا می‌باشند. لذا در این مقاله از این دسته مدل‌ها استفاده می‌شود. بهینه‌سازی سیاست مبدائی یک روش عملگر-نقاد است که همان‌طور که از نام آن پیداست دو بخش دارد. عملگر با خط‌مشی مطابقت دارد و برای انتخاب عمل و به‌روزرسانی خط‌مشی استفاده می‌شود، درحالی‌که نقاد با تابع مقدار مطابقت دارد و پارامترهای شبکه را برای تابع مقدار مورد استفاده به‌روز می‌کند. شبکه عامل وضعیت را به‌عنوان ورودی دریافت می‌کند و عمل متناسب با آن را انتخاب می‌کند. برای نشان دادن تابع مقدار حالت، شبکه نقاد نیز حالت را به‌عنوان ورودی دریافت می‌کند و یک عدد واحد را که بیانگر ارزش تخمین زده‌شده آن حالت است را خروجی می‌دهد (Schulman et al., 2017).

## ۲-۲- مدل حافظه کوتاه‌مدت - بلندمدت

الگوریتم یادگیری شبکه عصبی مصنوعی، یا شبکه عصبی، یک سیستم یادگیری محاسباتی است که از شبکه‌ای از توابع برای درک و تبدیل داده‌های ورودی از یک فرم به خروجی دلخواه، معمولاً به شکل دیگری، استفاده می‌کند. مفهوم شبکه عصبی مصنوعی از زیست‌شناسی انسان و نحوه عملکرد نورون‌های مغز انسان برای درک ورودی‌های حواس انسان الهام گرفته‌شده است. شبکه‌های عصبی تنها یکی از بسیاری از

ابزارها و رویکردهای مورداستفاده در الگوریتم‌های یادگیری ماشین هستند. خود شبکه عصبی ممکن است به‌عنوان یک قطعه در بسیاری از الگوریتم‌های یادگیری ماشینی مختلف برای پردازش ورودی‌های داده پیچیده در فضایی که رایانه‌ها می‌توانند آن را درک کنند، استفاده شود. این الگوریتم‌ها دارای سه ناحیه اصلی می‌باشند. ناحیه نخست داده‌های ورودی را دریافت می‌کند. این داده‌ها می‌توانند در شکل‌های مختلف همچون اعداد، مجموعه‌ای از اعداد، تصویر، متن و ... باشند. لایه‌های میانی یا لایه‌های پنهان ناحیه‌ای هستند که پردازش بر روی داده‌های ورودی در آن‌ها صورت می‌گیرد. این ناحیه می‌تواند متشکل از یک یا چندلایه باشد و هر لایه می‌تواند تعدادی نورون داشته باشد که هر نورون یک ورودی را دریافت و پس از انجام عملی بر آن (معمولاً عملیاتی ریاضی) خروجی را به دست می‌دهد. معمولاً داده‌ی واردشده به هر نورون در یک وزن ضرب شده و به خروجی می‌رود. ناحیه یا لایه نهایی خروجی می‌باشد که نتیجه‌ی پردازش‌های صورت گرفته بر روی داده ورودی را به دست می‌دهد. (Namini, Tavakoli and Namin, 2018). یکی از انواع شبکه‌های عصبی، شبکه‌های عصبی بازگشتی<sup>۱۶</sup> می‌باشند که برای پیش‌بینی داده‌های سری زمانی مورداستفاده قرار می‌گیرند. این الگوریتم‌ها مجموعه‌ای از داده‌ها را دریافت می‌کنند و این داده‌ها به لایه‌های پنهان می‌روند. در هر نورون توسط تابع فعال‌ساز یک عملیات ریاضی بر روی آن‌ها انجام می‌شود. سپس در لایه خروجی، مقدار پیش‌بینی مدل برای ورودی‌های دریافت شده منتشر می‌شود. ساده‌ترین نوع شبکه‌های عصبی بازگشتی RNN نام دارد. این شبکه‌ها دارای یک حافظه می‌باشند و در این حافظه مقدار خروجی قبلی را ذخیره می‌کنند. سپس در هر مرحله این مقدار در کنار ورودی جدید به‌عنوان ورودی به نورون مدنظر منتقل می‌شود (Yamak, Yujain and Gadosey, 2019). یکی از ایرادات شبکه‌های RNN ضعف آن‌ها در حافظه‌های بلندمدت است. شبکه‌های عصبی حافظه کوتاه‌مدت - بلندمدت این قابلیت را دارند که بخشی از اطلاعات مهم را در حافظه‌ی بلندمدت خود نگه‌دارند. شبکه‌های حافظه کوتاه‌مدت-بلندمدت برخلاف RNN دو ورودی و دو خروجی دارند. بین این ورودی‌ها و خروجی‌ها، یکی از ورودی‌ها مستقیماً به خروجی متصل شده است. این اتصال که از ابتدا تا انتهای پروسه پیش‌بینی ادامه

کنترل عوارض در ایجاد تغییر مکانی و زمانی در تصمیم‌گیری رانندگان برای سفر می‌تواند مکمل مناسبی برای دو روش دیگر باشد. همچنین مشاهده شد که یادگیری تقویتی در بسیاری از مطالعات گذشته به‌عنوان کنترل‌گر استفاده شده است زیرا این مدل‌ها از قبل نیازی اطلاعات راجع به عمل‌های بهینه ندارند و یادگیری از طریق آزمون و خطا اتفاق می‌افتد. در میان انواع مختلف عامل‌های یادگیری تقویتی، عامل بهینه‌ساز مبدائی سیاست قابلیت انجام چندین عمل گسسته برای هر یک از کنترل‌ها را در هر مرحله دارد. لذا برای انتخاب مقادیر کنترلی از مدل یادگیری تقویتی بهینه‌ساز مبدائی سیاست استفاده خواهد شد.

### ۳-۱- ساخت مدل

مدل نهایی ارائه شده در این مقاله شامل دو بخش مدل می‌باشد که در کنار یکدیگر به‌طور هم‌زمان مورد استفاده قرار خواهند گرفت. بخش اول مدل‌های شبکه عصبی حافظه کوتاه‌مدت-بلندمدت هستند که برای تخمین تقاضا در بازه‌ی ده دقیقه‌ی آینده بکار می‌روند. در ابتدا داده‌های تاریخی ترافیکی که ماتریس مبدأ - مقصد برای ۲۴ بازه‌ی ده دقیقه‌ای (در مدت دو ساعت) گذشته است را به این مدل‌ها داده و پیش‌بینی ماتریس مبدأ مقصد برای ده دقیقه‌ی آینده به دست می‌آید. سپس مقادیر پیش‌بینی شده به نرم‌افزار شبیه‌ساز داده می‌شوند تا توسط آن هر خودرو با توجه به مبدأ و مقصد خود به مسیری اختصاص داده شود. اطلاعات تخصیص به‌دست‌آمده در غالب درصد اشغال هر قطعه از نرم‌افزار شبیه‌ساز در بازه‌های دودقیقه‌ای برداشت شده و به‌عنوان ورودی به مدل یادگیری تقویتی داده می‌شوند تا این مدل تغییرات کنترل‌های نرخ ورودی رمپ، سرعت مجاز و عوارض را بر اساس وضعیت شبکه به‌عنوان خروجی به دست بدهد. سپس این مقادیر در شبکه‌ی واقعی و همچنین در شبیه‌سازی اعمال خواهند شد. این رویه به همین نحو و به‌صورت تکرار تا پایان روز ادامه خواهد داشت. در انتها مدل‌های شبکه‌ی عصبی به روش اصلاح خرد به‌روزرسانی می‌شوند. این روند در شکل ۳ بیان شده است. مراحل ساخت این مدل‌ها در ادامه بیان خواهد شد. برای ساخت مدل‌ها ابتدا به داده‌های ترافیکی و همچنین شبکه‌ی بزرگراهی دارای دو مسیر موازی نیاز داریم که در بخش سوم بررسی می‌شوند.

دارد، حافظه‌ی بلندمدت نیز نامیده می‌شود و قابلیت اضافه کردن یا فراموش کردن اطلاعات را نیز دارد. دروازه‌ی فراموشی به همین جهت بکار گرفته می‌شود تا اطلاعات غیرضروری گذشته از حافظه فراموش شود. روابط ریاضی گیت‌های اشاره‌شده در ادامه بیان شده است. شکل ۲ نیز ساختار نورون در شبکه‌های عصبی حافظه کوتاه‌مدت-بلندمدت را بر اساس روابط نشان می‌دهد (Fu, Zhang and Li, 2016).

$$i_t = \sigma(W_i h_{t-1} + W_i h_t) \quad (4)$$

$$f_t = \sigma(W_f h_{t-1} + W_f h_t) \quad (5)$$

$$o_t = \sigma(W_o h_{t-1} + W_o h_t) \quad (6)$$

$$\bar{c} = \tanh(W_c h_{t-1} + W_c h_t) \quad (7)$$

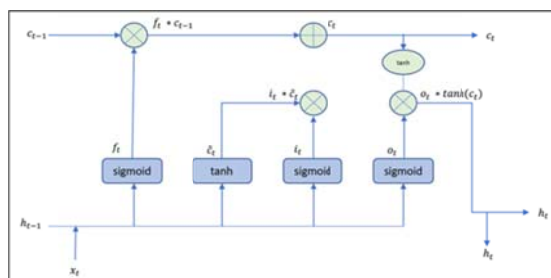
$$c_t = (i_t * \bar{c}) + (f_t * c_{t-1}) \quad (8)$$

$$h_t = o_t * \tanh(c_t) \quad (9)$$

$$X_t = \text{بردار ورودی} \quad (10)$$

$$h_t = \text{بردار خروجی} \quad (11)$$

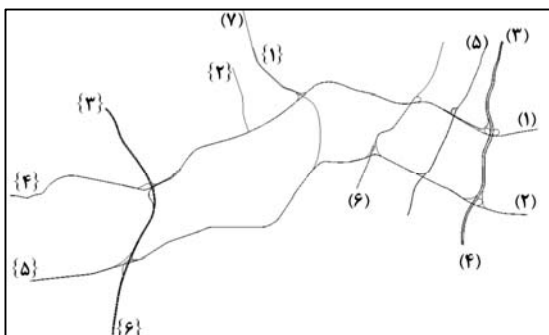
$$W, U \text{ و } f = \text{ماتریس پارامتر و بردار} \quad (12)$$



شکل ۲. ساختار سلول مدل حافظه بلندمدت - کوتاه‌مدت

### ۳- روش تحقیق

کنترل‌های موردنظر این مقاله، سه کنترل نرخ ورودی رمپ، سرعت مجاز و عوارض متغیر می‌باشد. دلیل این انتخاب اثربخشی قابل‌توجه این روش‌ها به‌خصوص کنترل رمپ و سرعت مجاز در مطالعات پیشین می‌باشد. ضمن اینکه توانایی

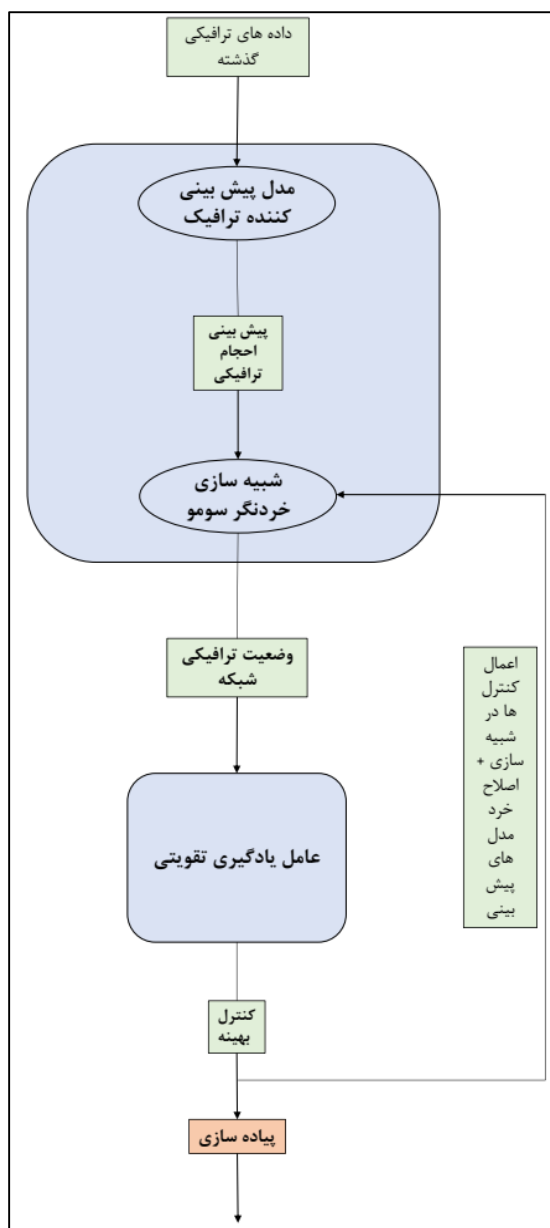


شکل ۴. شبکه‌ی ترافیکی بکار گرفته شده

از آنجایی که هدف اصلی کنترل سرعت مجاز متغیر کاهش نرخ ورود خودرو از مسیر اصلی به نقاط گلوگاهی است بهترین بخش برای پیاده‌سازی کنترل سرعت مجاز بالادست این نقاط می‌باشد. با توجه به این‌که ورودی‌های (ت) و (ث) مشخص شده در شکل ۵ نقاطی از شبکه هستند که پتانسیل بالایی برای تشکیل تداخلات ترافیکی شدید را دارند، محل اجرای کنترل سرعت مجاز متغیر باید به نحوی انتخاب شود که بالادست این دو ورودی باشد. قطعه‌ی (آ) مشخص شده در شکل به طول ۱٫۵ کیلومتر که در بالادست ورودی (ت) قرار دارد محل نخست پیاده‌سازی کنترل سرعت می‌باشد. قطعه‌ی (ب) نیز که خودروهای وارد شونده از رمپ را به مسیر اصلی هدایت می‌کند برای کنترل سرعت مدنظر می‌باشد. اعمال کنترل سرعت در این نقطه می‌تواند منجر به کاهش جریان ورودی به رمپ و در نتیجه کاهش طول صف بشود. قطعه‌ی (پ) نیز به طول ۴۰۰ متر بالادست ورودی (ث) بوده و کنترل سرعت بر روی آن اعمال می‌شود. لازم به ذکر است به علت نزدیکی قطعات (آ) و (پ) و برای جلوگیری از رخداد تصادفات در اثر ترمز گیری خودروها، اختلاف سرعت در این دو بخش به طوری تنظیم می‌گردد که کمتر از ۱۰ کیلومتر بر ساعت اختلاف داشته باشند. تفاوت سرعت بیشتر بین این دو قطعه منجر به ترمز خودروها حین ورود از قطعه‌ی (آ) به (پ) شده و ایمنی مسیر را به خطر می‌اندازد. خودروهای وارد شونده از مبادی ۳ و ۵ در صورت انتخاب به ورود به مسیر اصلی با یکدیگر ترکیب شده و از رمپ (ت) مشخص شده در شکل ۵ وارد این مسیر می‌شوند. ضمن اینکه تمامی

### ۳-۱-۱- داده‌ها و شبکه‌ی ترافیکی

شبکه‌ی مورد استفاده در این مقاله مطابق شکل ۴ دارای دو مسیر موازی با رویکرد شرق به غرب می‌باشد که به هر یک رمپ‌هایی ورود کرده‌اند. این شبکه دارای ۷ مبدأ (نشان داده شده با پرانتز) و ۶ مقصد (نشان داده شده با آکولاد) می‌باشد که در مجموع ۴۲ زوج مبدأ-مقصد را تشکیل می‌دهند. البته به علت مجاورت مبادی ۶ و ۷ به مقصد شماره یک، تردد بین این دو زوج مبدأ-مقصد در نظر گرفته نمی‌شود.

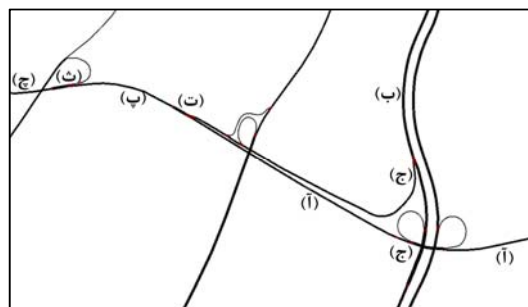


شکل ۳. روند عملکرد مدل‌های پیشنهادی

اعمال کنترل نرخ ورودی خواهند داشت. لذا این دو رمپ برای اعمال کنترل مذکور موردنظر می‌باشند. این دو ورودی با عبور تعداد زیاد خودرو به مسیر اصلی امکان ایجاد تداخل در پایین دست تقاطع ورودی را دارند. بدین منظور چراغ ورودی رمپ در این نقاط قرار گرفته که دارای سه فاز می‌باشد. فاز سبز به نحوی تعیین می‌شود که تعداد محدودی خودرو را به بزرگراه وارد کند. فاز زرد برای جلوگیری از رخداد تصادف ناشی از ترمز گیری شدید خودروها و دارای زمانی ثابت می‌باشد. فاز قرمز نیز مانع ورود خودروها از رمپ به مسیر اصلی جهت روان‌سازی جریان بزرگراه می‌باشد. مدل یادگیری تقویتی در هر مرحله به اعمال تغییر در طول فاز سبز و قرمز می‌پردازد. طول چراغ زرد به‌طور ثابت ۳ ثانیه می‌باشد. اگرچه در برخی از مطالعات برای چراغ ورودی رمپ از فاز زرد صرف‌نظر می‌کنند اما این موضوع می‌تواند منجر به ترمز ناگهانی برخی خودروها و ایجاد تصادف شود لذا این زمان برای فاز زرد در نظر گرفته شد. طول زمان سبز هر چراغ بین ۵ تا ۲۵ و طول فاز قرمز بین ۵ تا ۲۰ ثانیه خواهد بود. در صورت قرار ندادن حداقل و حداکثر برای زمان سبز و قرمز مدل ممکن است تصمیمات خود را به سمتی ببرد که در یک ورودی چراغ سبز آنقدر زیاد یا قرمز آنقدر کم شود که عملاً چراغ غیرفعال شود. همچنین کاهش زیاد زمان سبز یا افزایش زیاد زمان قرمز می‌تواند منجر به عدم امکان ورود خودروهای یک رمپ به مسیر اصلی شده و باعث افزایش طول صف رمپ گردد.

با توجه به اینکه شبیه‌ساز سومو قابلیت اعمال عوارض بر شبکه را ندارد، مقدار عوارض موردنظر باید تبدیل به زمان شده و به زمان سفر اصلی هر قطعه افزوده شود تا نرم‌افزار برای مسیریابی خودروها از این زمان اصلاح‌شده استفاده کند. با فرض میانگین درآمد ۶۰۰۰ دلار برای یک فرد، اگر این فرد ۲۱ روز در ماه و هر روز ۸ ساعت مشغول به کار باشد ارزش هر ثانیه کار این فرد ۱ سنت خواهد بود. بر این مبنای مقدار ۱ سنت برای هر ثانیه برای تبدیل عوارض به زمان استفاده می‌گردد و حداکثر مقدار عوارض در هر یک از نقاط اخذ، ۲۰۰ سنت خواهد بود. عوارض دریافتی در نقاط (ج) دارای مقدار یکسان خواهند بود اما عوارض دریافتی در (چ) متفاوت است. در عمل مدل دو مقدار برای عوارض در هر مرحله انتخاب می‌کند. عمل نخست بر دو عوارضی اول و عمل دوم بر عوارضی سوم اعمال می‌شود. بخش بعدی کار آماده‌سازی داده‌های ترافیکی مبدأ مقصد می‌باشد. برای این کار به دو بخش نیاز است. داده‌ی نخست برای آموزش مدل هلی کنترل‌گر و پیش‌بینی کننده نیاز است و بخش دوم برای استفاده‌ی مدل و بررسی عملکرد آن می‌باشد. از آنجایی که در نظر است توانایی هر دو مدل در احجام ترافیکی متنوعی را

خودروهای وارد شونده از مبادی ۱ تا ۶ امکان ورود به مسیر شماره‌ی دو را نیز دارند. در شکل ۴ مشاهده می‌شود که خودروهای ورودی از مبادی ۱ تا ۵ در صورت انتخاب به ورد به مسیر اصلی در رمپ (ت) با یکدیگر تداخل دارند که این نقطه را به محل اصلی تداخلات احتمالی تبدیل می‌کند. به همین جهت رمپ ورودی این مکان مناسب‌ترین نقطه برای پیاده‌سازی کنترل نرخ رمپ و بزرگراه‌های بالادست آن نیز مناسب برای کنترل سرعت مجاز خواهند بود. همچنین اعمال عوارض در ورودی مسیرهای اصلی به این رمپ می‌تواند منجر به تغییر مسیر خودروها با مسیر پایین شود؛ بنابراین این مکان مناسب‌ترین محل برای اعمال عوارض می‌باشد.



شکل ۵. محل اجرای کنترل‌ها: (ا)، (ب) و (پ) کنترل سرعت، (ت) و (ث) کنترل رمپ، (ج) و (چ) کنترل عوارض

با توجه به شکل ۵ رمپ‌های ورودی (ت) و (ث) نقاطی هستند که با توجه به ورودی‌های متعدد و حجم بالای خودروها بیشترین نیاز را به اعمال کنترل نرخ ورودی خواهند داشت. لذا این دو رمپ برای اعمال کنترل مذکور موردنظر می‌باشند. این دو ورودی با عبور تعداد زیاد خودرو به مسیر اصلی امکان ایجاد تداخل در پایین دست تقاطع ورودی را دارند. بدین منظور چراغ ورودی رمپ در این نقاط قرار گرفته که دارای سه فاز می‌باشد. فاز سبز به نحوی تعیین می‌شود که تعداد محدودی خودرو را به بزرگراه وارد کند. فاز زرد برای جلوگیری از رخداد تصادف ناشی از ترمز گیری شدید خودروها و دارای زمانی ثابت می‌باشد. فاز قرمز نیز مانع ورود خودروها از رمپ به مسیر اصلی جهت روان‌سازی جریان بزرگراه می‌باشد. مدل یادگیری تقویتی در هر مرحله به اعمال تغییر در طول زمان سبز و قرمز می‌پردازد. طبق شکل ۵ رمپ‌های ورودی (ت) و (ث) نقاطی هستند که با توجه به ورودی‌های متعدد و حجم بالای خودروها بیشترین نیاز را به

و پاداش برای مدل می‌باشد. همان‌طور که در توضیحات مربوط به شبکه بیان شد، کنترل‌گر در هر ۲ دقیقه و با توجه به وضعیت سیستم جاده‌ای باید ۲ عمل برای عوارض، ۳ عمل برای سرعت مجاز و ۴ عمل برای چراغ‌های سبز و قرمز ۲ چراغ تعبیه‌شده انجام دهد. عمل تعریف‌شده برای مدل در هرکدام از این موارد به‌صورت کم کردن، زیاد کردن یا تغییر ندادن مقدار کنترلی می‌باشد.

هر واحد تغییر برای عوارض ۱۰ سنت، برای سرعت ۱ متر بر ثانیه و برای زمان هر فاز چراغ ۲ ثانیه می‌باشد. به این صورت که در ابتدا برای هرکدام از این ۹ مقدار کنترلی یک مقدار در شروع روز در نظر گرفته‌شده و هر دو دقیقه مدل می‌تواند هرکدام از این مقادیر را به‌اندازه‌ی یک واحد تعریف‌شده کم یا زیاد کند و یا تغییر ندهد. مقادیر ابتدایی کنترل‌ها در شروع هرروز، حداقل، حداکثر و گام تغییر آن‌ها مطابق جدول ۱ می‌باشد.

میزان سرعت مجاز در دو بخش به علت نزدیکی و افزایش احتمال تصادف در صورت وجود تفاوت زیاد، حداکثر ۳ متر بر ثانیه (۱۰/۸ کیلومتر بر ساعت) تفاوت خواهند داشت.

اگر مدل بعد از پیاده‌سازی مقادیر اولیه تفاوت بیش از ۳ متر بر ثانیه در سرعت مجاز این دو قطعه ببیند، با گام‌های مساوی آن‌ها را به هم نزدیک می‌کند. برای مثال اگر سرعت بخش اول ۲۳ و سرعت بخش دوم ۱۸ کیلومتر بر ثانیه باشد، هرکدام به میزان ۱ متر بر ثانیه به هم نزدیک شده و درنهایت مقادیر ۲۲ و ۱۹ متر بر ثانیه برای این دو بخش اعمال می‌گردد.

مورد سنجش قرار داد، نیاز به داده‌ای می‌باشد که دارای الگوی روزانه، هفتگی و فصلی است. لذا داده‌ی بخش اول را برای یک سال (۳۶۵ روز با فرض هر فصل ۹۱ روز) ایجاد می‌شود. این داده را به نحوی ایجاد می‌شود که در طول روز دارای الگوهای مختلفی باشد. همچنین در نظر است که این داده دارای الگوی هفتگی و فصلی نیز باشد. بخش دوم داده به همین شکل بیست روز خواهد بود که درنهایت آن را برای بررسی عملکرد مدل‌ها مورد استفاده قرار می‌گیرد. همان‌طور که گفته شد شبکه‌ی مورد استفاده، ۴۲ زوج مبدأ- مقصد (۷ مبدأ و ۶ مقصد دارد) است. البته به علت نزدیکی مبادی ۶ و ۷ به مقصد یک فرض می‌شود که ترددی بین این دو زوج مبدأ- مقصد نبوده و مقدار حجم این دو زوج را صفر قرار داده می‌شود؛ اما برای سایر زوج‌های مبدأ مقصد در طول روز الگوی متفاوتی را در نظر گرفته می‌شود. از آنجایی که شبیه‌سازی یک روز ۲۴ ساعته آن‌هم برای بیش از هزار بار آزمون خطای مدل یادگیری تقویتی بسیار زمان‌بر می‌باشد، مدل ارائه‌شده را برای یک بازه‌ی دوساعته از روز آماده می‌شود. همچنین مقادیر مبدأ مقصد به‌صورت تعداد خودرو در بازه‌های ده‌دقیقه‌ای می‌باشد. به عبارتی هرروز ۲ ساعت و دارای ۱۲ بازه متفاوت با ماتریس مبدأ مقصد متفاوت خواهد بود.

### ۳-۱-۲- آموزش مدل کنترل‌گر

پس از آماده‌سازی داده و شبکه‌ی ترافیکی به آموزش مدل یادگیری تقویتی برای تعیین مقادیر کنترلی پرداخته می‌شود. مهم‌ترین گام در ساخت مدل کنترل‌گر، تعریف عمل، وضعیت

جدول ۱. مقادیر اولیه، حداقل، حداکثر و گام تغییر مقادیر کنترلی

عوارض	عوارض	سرعت	سرعت	سرعت	سرعت	فاز سبز	فاز قرمز	فاز سبز	فاز قرمز
بخش یک	بخش یک	بخش یک	بخش دو	بخش سه	بخش سه	چراغ اول	چراغ اول	چراغ دوم	چراغ دوم
صفر	صفر	۲۵ متر بر ثانیه	۲۵ متر بر ثانیه	۲۵ متر بر ثانیه	۲۵ متر بر ثانیه	۱۰ ثانیه	۱۰ ثانیه	۱۰ ثانیه	۱۰ ثانیه
صفر	صفر	۱۵ متر بر ثانیه	۱۵ متر بر ثانیه	۱۵ متر بر ثانیه	۱۵ متر بر ثانیه	۵ ثانیه	۵ ثانیه	۵ ثانیه	۵ ثانیه
۲۰۰ سنت	۲۰۰ سنت	۲۵ متر بر ثانیه	۲۵ متر بر ثانیه	۲۵ متر بر ثانیه	۲۵ متر بر ثانیه	۲۰ ثانیه	۲۰ ثانیه	۲۵ ثانیه	۲۰ ثانیه
۱۰ سنت	۱۰ سنت	۱ متر بر ثانیه	۱ متر بر ثانیه	۱ متر بر ثانیه	۱ متر بر ثانیه	۲ ثانیه	۲ ثانیه	۲ ثانیه	۲ ثانیه

با افزایش پاداش میانگین زمان سفر نیز کاهش یابد اما ممکن است درک ارتباط میان زمان سفر و وضعیت دریافتی مدل برای عامل سخت باشد، این تعریف از پاداش نمی‌تواند کمک زیادی به عملکرد مدل نماید. لذا بهتر است تعریفی از پاداش داشته باشیم که بتواند برای مدل قابل‌درک باشد. از آنجایی که ۲۰ قطعه‌ی انتخاب‌شده نقاطی از شبکه هستند که تداخلات ترافیکی در آن‌ها شکل می‌گیرد، اگر مدل بتواند میانگین سرعت در آن‌ها را افزایش و میانگین سطح اشغال را کاهش دهد می‌تواند منجر به بهبود وضعیت ترافیکی و کاهش زمان سفر گردد. ضمن اینکه تعریف پاداش بر اساس این دو پارامتر می‌تواند به‌نوعی اهداف دیگری جز زمان سفر را نیز مورد اثر قرار دهد. لذا پاداش مدل در هر مرحله مطابق رابطه‌ی ۱۳ تعریف می‌گردد:

$$reward = - \frac{moc}{ms} \quad (13)$$

$moc$  = میانگین سطح اشغال برای بیست قطعه مورد نظر

$ms$  = میانگین سرعت خودروها در بیست قطعه مورد نظر

تمام‌روزهای موجود در داده‌های قبلی پیاده می‌کند. به این ترتیب به‌مرور عامل سیاست خود در انتخاب عمل بهینه بر اساس وضعیت شبکه را بهبود می‌بخشد.

### ۳-۱-۳- آموزش مدل پیش‌بینی کننده ترافیک

مدل یادگیری تقویتی با آزمون و خطا سعی در یادگیری عمل بهینه برای کنترل مسیر می‌نماید. این مدل در هر مرحله اقدامی را بر روی شبکه انجام داده و در نتیجه‌ی آن وضعیت و پاداشی را دریافت می‌کند. از آنجایی که وضعیت باید بیانگر شرایط شبکه و نقاط دارای تداخل باشد، مقدار میانگین سرعت و سطح اشغال ۲۰ قطعه‌ی شبکه که در نقاط گلوگاهی قرار دارند را به‌عنوان وضعیت انتخاب قرار داده می‌شود. برای آگاهی از وضعیت این ۲۰ قطعه در هر مرحله، دو راهکار وجود دارد. راهکار نخست نصب دوربین‌ها یا شناساگرها و شمارش تعداد خودروها و سرعت آن‌ها به‌طور به لحظه است.

بخش دوم تعیین تعریف وضعیت شبکه برای مدل است. از آنجایی که مدل باید تصمیمات خود را بر مبنای میزان شلوغی شبکه در بخش‌های مختلف بگیرد، ۲۰ قطعه مهم از کل شبکه که شامل بالادست رمپ‌ها، پایین‌دست آن‌ها و خود رمپ ورودی می‌باشد را انتخاب نموده و سطح اشغال و میانگین سرعت وسایل نقلیه بر روی آن‌ها را به‌عنوان وضعیت به مدل داده می‌شود. این مقادیر در یک ماتریس ۲ در ۲۰ ذخیره می‌شود. ستون اول شامل مقادیر سطح اشغال به درصد برای ۲۰ قطعه و ستون دوم میانگین سرعت خودروهای روی آن قطعه بر حسب متر بر ثانیه می‌باشد. از آنجایی که مدل برای درک اثر عملکرد خود بر محیط باید بتواند بر اساس وضعیت، پاداش را پیش‌بینی کند، بهترین حالت این است که پاداش بر اساس وضعیت تعریف گردد. برای مثال می‌توان پاداش را برابر با منفی میانگین زمان سفر در دو دقیقه‌ی گذشته تعریف کرد که

بدین ترتیب اگر مدل بتواند پاداش تعریف‌شده را افزایش دهد، می‌توان گفت میانگین سرعت کاهش یافته و یا میانگین سطح اشغال در نقاط گلوگاهی کاهش یافته است. انتظار می‌رود در صورت افزایش این پاداش سایر پارامترهای ترافیکی همچون میانگین زمان سفر نیز بهبود یابند. آموزش مدل به این صورت می‌باشد که برای هر یک از روزها در داده‌ی سال اول، یک شبیه‌سازی ترافیکی ایجاد می‌شود. سپس مدل بر اساس مقادیر ابتدایی و همچنین با توجه به خالی بودن شبکه از خودرو عمل اول خود را انجام خواهد داد. پس از آن به مدت ۱۲۰ ثانیه شبیه‌سازی ادامه خواهد داشت. در این مرحله شبیه‌ساز (محیط) وضعیت موردنظر را بر اساس تعاریف مطرح‌شده در این بخش به عامل بر خواهد گرداند. همچنین مقدار پاداش مدل نیز به شکل بیان‌شده محاسبه شده و به عامل داده خواهد شد. عامل بر اساس مقادیر دریافتی عمل بعدی خود را انتخاب کرده و آن‌ها پیاده‌سازی می‌کند. این روند تا پایان روز و بعد از آن برای

دقت خود برای پیش‌بینی داده‌های آینده می‌نماید. در ادامه به بیان جزئیات این مدل پرداخته خواهد شد. همان‌طور که گفته شد داده‌ی سال نخست برای ۳۶۴ روز ایجاد شده و هر روز (۲ ساعت) به ۱۲ بازه‌ی ده‌دقیقه‌ای تقسیم شده است. در واقع ما ۴۳۶۸ داده برای هر زوج مبدأ مقصد داریم. ۸۰ درصد این داده یعنی ۳۴۹۴ مورد برای آموزش و مابقی برای ارزیابی مدل مورد استفاده قرار گرفته است. مدل به‌گونه‌ای ساخته شده است که با مشاهده‌ی ۲۴ مقدار قبلی، مقدار بعدی را پیش‌بینی کند. برخی از پارامترهای مهم مدل ساخته شده در جدول ۲ مشخص شده است. مدل ساخته شده دارای پنج لایه می‌باشد. لایه‌ی نخست یک‌لایه‌ی LSTM با ۳۲ نورون می‌باشد و برای تابع فعال‌سازی از تابع RELU استفاده شده است. لایه‌ی بعدی دراپ اوت می‌باشد که برای پیش‌گیری از برازش بیش‌ازحد استفاده می‌شود. لایه سوم و چهارم همانند لایه اول و دوم دو لایه‌ی LSTM و دراپ اوت می‌باشد. لایه‌ی نهایی نیز خروجی مدل را می‌دهد.

این روش از آنجایی که نیاز به امکانات زیادی دارد و با توجه به این‌که داده‌های دریافت شده از شناساگرها نیازمند پردازش بوده و بلافاصله تحویل کنترل‌گر نمی‌شوند نمی‌توانند اطلاعات مورد نیاز هر یک از این ابزارها نیز می‌توانند کل سیستم را دچار اختلال کند. روش دوم این است که داده‌های مبدأ-مقصد را به یک نرم‌افزار شبیه‌ساز وارد شوند و اطلاعات ریز شبکه شامل میانگین سرعت و اشغال بخش‌های مختلف در هر مقطع زمانی را از آن گرفته شوند. البته لازمه‌ی این کار این است که داده‌های مبدأ-مقصد را موجود باشند. پیش‌بینی مقادیر ماتریس مبدأ مقصد برای آینده بر مبنای الگوهای استخراج شده از داده‌های پیشین یکی از رویکردهای متداول در بسیاری از مطالعات ترافیکی می‌باشد. مدل‌های شبکه عصبی سری زمانی از مدل‌های بسیار قدرتمند برای این امر است. در این پژوهش ابتدا داده‌ی یک‌ساله‌ی ایجاد شده برای هر زوج مبدأ مقصد به یک مدل شبکه عصبی مختص خود آن زوج داده شده و این مدل با آموزش بر روی الگوهای این داده سعی بر افزایش

جدول ۲. پارامترهای نهایی مورد استفاده در مدل‌های حافظه کوتاه‌مدت - بلندمدت

پارامتر	انتخاب نهایی
تعداد بازگشت به عقب	۱۲
نرخ یادگیری	۰,۰۰۱
تعداد لایه و نورون	۲ لایه ۶۴ نورونی
ساز دسته‌بندی	۳۲
درصد لایه حذفی	۱۵
تابع فعال‌ساز	ReLU
تابع بهینه‌ساز	Adam
تعداد تکرار	۳۰۰

### ۳-۲- استفاده از مدل

می‌شود. به این طریق ادامه می‌شود تا کل مقادیر مبدأ مقصد ۲ ساعت آینده برای ۱۲ بازه‌ی ۱۰ دقیقه‌ای پیش‌بینی شود. پس از این مرحله مقادیر به‌دست‌آمده را بر روی شبکه‌ی بزرگراهی کنترل شده با سه روش مورد نظر در نرم‌افزار سومو بارگذاری شده و شبیه‌سازی آغاز می‌شود. پس از گذشت هر ۲ دقیقه

برای استفاده از مدل در ابتدا ماتریس مبدأ-مقصد برای ۱۲ بازه‌ی قبلی (۲ ساعت قبل) را به مدل شبکه‌ی عصبی وارد می‌شود. این مدل با این مقادیر ماتریس مبدأ مقصد ده دقیقه‌ی آینده را پیش‌بینی می‌کند. سپس به کمک این مقدار و ماتریس ۱۱ بازه‌ی قبلی، ماتریس مبدأ-مقصد بازه‌ی دوم را پیش‌بینی

کوتاه‌تر (در این مقاله روزانه) است. اگر داده‌های مبدأ مقصد هرروز پس از اتمام روز در اختیار سیستم قرار بگیرد می‌توان از طریق روش اصلاح خرد<sup>۱۷</sup> به به‌روزرسانی مدل‌های پیش‌بینی صرفاً با همین داده‌های یک‌روزه پرداخت. این روش به زمان‌بر بودن ساخت مدل‌های پیش‌بینی از ابتدا نبوده و تا حدودی اثر تغییرات الگوی تقاضا در بازه‌ی روزانه نیز قابل در نظر گرفتن خواهد بود. البته لازم است برای پیش‌گیری از برازش بیش‌ازحد مدل‌ها بر روی داده‌ی یک‌روزه و فراموش شدن اطلاعات به‌دست‌آمده از داده‌های پیشین، باید از نرخ یادگیری بسیار پایین‌تر استفاده نمود. لذا شیوه‌ی استفاده از این روش در این تحقیق به این صورت است که پس از اتمام روز و به دست آمدن داده‌های مبدأ مقصد جدید، مدل‌های پیش‌بینی با نرخ یادگیری ۰,۰۰۰۰۱ (۱۰۰ برابر کمتر از نرخ یادگیری روی داده‌ی یک‌ساله) و با ۲۰ تعداد تکرار بر روی داده‌های یک روز آموزش داده می‌شود تا ضرایب مدل با شیب ملایمی به‌روزرسانی گردد.

اطلاعات سرعت و درصد اشغال بیست قطعه مشخص شده از شبیه‌ساز دریافت شده و به مدل یادگیری تقویتی داده می‌شود. این مدل نیز به انتخاب عمل‌کنترلی بهینه پرداخته و آن را بر روش شبکه پیاده می‌نماید. در کاربرد عملی انتظار نیز پس از تعیین عمل‌کنترلی در هر زمان، کنترل‌ها بر شبکه‌ی واقعی پیاده‌سازی می‌شود. یکی از نکاتی که ممکن است در طول زمان موجب کاهش اثربخشی اعمال مدل شود تغییرات در الگوی داده‌های مبدأ مقصد می‌باشد. البته این تغییرات به‌صورت ناگهانی (مثلاً ظرف یک روز یا یک هفته) رخ نمی‌دهد و در بازه‌ی زمانی طولانی‌تری امکان رخ دادن دارد. یکی راه‌حل برای این مشکل این است که مثلاً همراه مدل‌های پیش‌بینی کننده مجدداً از نوساخته شوند. ایراد این موضوع طولانی بودن پروسه‌ی ساخت مدل از ابتدا و همچنین عدم درک تغییرات در بازه‌های کوتاه‌تر مثلاً روزانه و هفتگی می‌باشد. راهکار بهتری که برای این مسئله وجود دارد به‌روزرسانی مدل‌ها صرفاً با داده‌های جدید و در بازه‌های

جدول ۳. پارامترهای استفاده‌شده در بروز رسانی مدل‌های شبکه عصبی

درصد تقسیم داده‌ها به داده آموزش و ارزیابی	۱۰۰ درصد آموزش
نرخ یادگیری مدل	۰,۰۰۱
تعداد تکرار	۲۰
تعداد هر دسته‌ی داده	۲

ضمن اینکه برای بررسی عدم برازش بیش‌ازحد باید مقدار  $R^2$  بر روی داده‌های آموزش و داده‌های ارزیابی مقدار کمی باشد. خروجی مدل‌های (۴۰ مورد برای ۴۰ زوج مبدأ مقصد) استفاده در این مقاله مقادیر  $R^2$  ارزیابی دارای میانگین ۰,۸۸ بوده و در بازه‌ی ۰,۸۰۷ تا ۰,۹۰۵ قرار دارد که نشان‌دهنده‌ی پیش‌بینی مناسبی می‌باشد. ضمن اینکه برای تمامی مدل‌ها بیشترین مقدار تفاوت  $R^2$  در داده‌های آموزش و ارزیابی برابر ۰,۰۲۶ است که کاملاً قابل‌قبول می‌باشد هیچ‌کدام از مدل‌ها دچار برازش بیش‌ازحد نمی‌باشند. مقدار mse برای داده‌های آموزش بین ۰,۰۰۳۸ و ۰,۰۰۵۵ می‌باشد. برای داده‌های ارزیابی نیز این عدد بین ۰,۰۰۳۸ و ۰,۰۰۶۱ می‌باشد. تمامی اعداد به‌دست‌آمده نشانگر عملکرد مناسب مدل در پیش‌بینی داده‌های ترافیکی می‌باشد.

#### ۴- بررسی و مقایسه عملکرد مدل‌ها

در بخش بررسی عملکرد مدل‌ها به دو بخش از بررسی باید پرداخت. بخش نخست بررسی عملکرد مدل‌های پیش‌بینی ترافیکی و بخش دوم مقایسه‌ی عملکرد مدل‌های کنترل‌گر در استراتژی‌های مختلف می‌باشد.

#### ۴-۱- ارزیابی عملکرد مدل‌های پیش‌بینی ترافیکی

برای عملکرد مناسب مدل کنترل‌کننده، پیش‌بینی صحیح مدل‌های شبکه عصبی اهمیت بسیاری دارد. ملاک‌های مورد‌استفاده برای سنجش عملکرد مدل شامل خطای میانگین مربعات و همچنین  $R^2$  می‌باشد. معیار  $R^2$  برای بررسی میزان دقت مدل در انتخاب مقادیر بوده و در مطالعات مهندسی مقادیر  $R^2$  بالای ۰,۷ نشان‌دهنده‌ی عملکرد مناسبی می‌باشد.

#### ۴-۲- سناریوهای کنترلی مختلف و عملکرد آنها

بررسی این موضوع مدل‌های مختلفی شامل یک، دو و یا هر سه کنترل مورد نظر به شرح زیر ساخته شده و عملکرد آنها مقایسه می‌گردد:

در بخش‌های پیشین اشاره شد که یکی از رویکردهای مناسب برای افزایش کارایی مدیریت اعمال شده، ترکیب روش‌های کنترلی به صورت هماهنگ با یکدیگر می‌باشد. برای

۱. سناریو بدون هیچ‌گونه کنترل (N)

۲. سناریو با کنترل هم‌زمان رمپ، سرعت و عوارض (RVT)

۳. سناریو با کنترل رمپ و سرعت (RV)

۴. سناریو با کنترل رمپ و عوارض (RT)

۵. سناریو با کنترل سرعت و عوارض (VT)

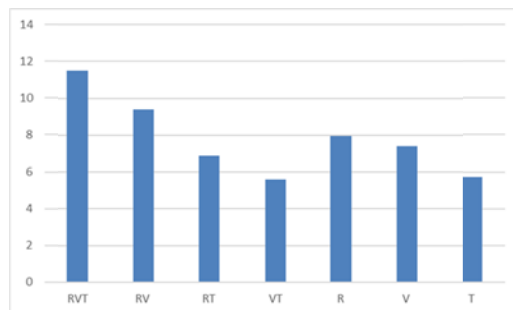
۶. سناریو با فقط کنترل رمپ (R)

۷. سناریو با فقط کنترل سرعت (V)

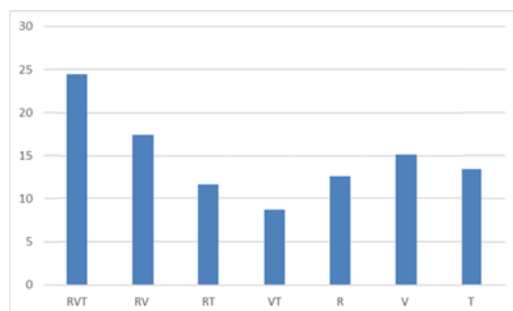
۸. سناریو با فقط کنترل عوارض (T)

از سناریو RVT بیشترین کارایی را داشته است. البته تمامی کنترل‌های اعمال شده در صورت قابلیت بهبود زمان سفر را داشته‌اند. شکل ۷ تغییرات میانگین سرعت خودروها در سناریوهای مختلف را نشان می‌دهد. مشاهده می‌شود که دو سناریو RVT و RV بیشترین کارایی را نسبت به سایر سناریوها داشته‌اند. سناریو RVT با افزایش ۲۴،۴۵ درصدی میانگین سرعت در مقایسه با افزایش ۱۷،۴۹ درصدی آن در سناریو RV، دارای اثر به مراتب بالاتری در بهبود این پارامتر بوده است. همچنین مشاهده می‌شود که تمامی کنترل‌ها توانایی بهبود میانگین سرعت در شبکه را دارند اما دو سناریو RVT و RV بیشترین کارایی را داشته‌اند که با آنچه از مطالعات پیشین برداشت و حدس زده می‌شد مطابقت دارد. درصد تغییرات تمام پارامترهای ترافیکی نسبت به سناریو N در جدول ۴ ارائه شده است.

برای مقایسه هر کدام از این حالت‌ها، مدلی منطبق با سناریو مدنظر ساخته و سپس این مدل را برای ۲۰ روز داده‌های آماده شده برای ارزیابی پیاده‌سازی شده است. این داده‌ها در این بیست روز دارای احجام متفاوت می‌باشند که میزان عملکرد مدل در هر روز بسته به میزان ترافیک متفاوت است. در نهایت میانگین بیست‌روزه تغییرات پارامترهای ترافیکی شامل زمان سفر، سرعت متوسط، زمان انتظار، طول صف، تعداد خودروهای به مقصد رسیده و متوسط تراکم برای هر یک از سناریوها در مقایسه با سناریو N محاسبه شده است. نمودار شکل ۶ میزان تغییرات میانگین زمان سفر در هر یک از سناریوهای مربوطه را نشان می‌دهد. در این شکل مشاهده می‌شود که کاهش میانگین زمان سفر نسبت به سناریو N در سناریوی RVT دارای بیشترین میزان بوده و با کاهش ۱۱،۴۷ درصدی اثربخشی قابل‌ملاحظه‌ای داشته است. کنترل RV نیز که رایج‌ترین ترکیب به کار رفته در مطالعات پیشین می‌باشد پس



شکل ۶. تغییرات میانگین زمان سفر در هر یک از سناریوها



شکل ۷. تغییرات میانگین سرعت در هر یک از سناریوها

جدول ۴. درصد بهبود پارامترهای ترافیکی در هر یک از سناریوها نسبت به سناریو N

سناریو	میانگین زمان سفر (S)	میانگین سرعت (M/S)	میانگین زمان انتظار (S)	میانگین طول صف (VEH)	تعداد خودروهای رسیده به مقصد (VEH)	میانگین تراکم شبکه (VEH/KM)
RVT	۱۱,۴۷	۲۴,۴۵	۳۱,۷۰	۴۷,۶۹	۳,۰۳	۱۶,۷۰
RV	۹,۴۰	۱۷,۴۹	۱۶,۱۹	۳۲,۸۲	۲,۳۶	۱۳,۷۰
RT	۶,۹۱	۱۱,۶۳	۱۴,۰۷	۳۸,۰۰	۱,۷۲	۱۴,۳۰
VT	۵,۵۹	۸,۶۸	۱۲,۳۲	۲۴,۶۹	۱,۲۴	۷,۰۳
R	۷,۹۹	۱۲,۵۵	۶,۶۵	۸,۹۵	۱,۹۵	۱۵,۴۵
V	۷,۴۳	۱۵,۱۴	۲۹,۷۵	۳۰,۵۱	۲,۱۵	۱۴,۵۲
T	۵,۷۳	۱۳,۴۴	۲۰,۰۳	۲۰,۱۵	۱,۴۱	۷,۲۳

مشاهده می‌شود که در بهبود تمامی پارامترهای ترافیکی، سناریو RVT عملکرد بهتری را نسبت به سایر کنترل‌ها از خود نشان داده است. همچنین سناریو RV نیز به‌جز سناریو RVT از تمامی سناریوها عملکرد بهتری را داشت است. این موضوع بیان‌گر این است که ترکیب کنترل رمپ و سرعت توانسته است اثربخشی کنترل منفرد رمپ و سرعت را افزایش دهد. ضمن اینکه افزودن کنترل عوارض بزرگراهی نیز عملکرد کنترل رمپ و سرعت را بهبود بخشیده است. از نکات جالب جدول ۴ عملکرد ضعیف سناریو R در کاهش طول صف و همچنین میانگین زمان انتظار می‌باشد. با توجه به اینکه کنترل نرخ ورودی رمپ به‌تنهایی منجر به افزایش طول صف در رمپ‌های ورودی می‌شود، این موضوع قابل پیش‌بینی بوده و در مطالعات پیشین نیز به این موضوع اشاره شده است. همچنین سناریوهای کنترلی RT و VT در بهبود غالب پارامترها نتایج ضعیف‌تری نسبت به سناریوهای R و V داشته‌اند. می‌توان گفت که ترکیب کنترل عوارض به ترکیب کنترل نرخ ورودی رمپ و سرعت مجاز متغیر نتایج می‌تواند

مناسبی داشته باشد. اما در صورتی که یکی از این دو کنترل وجود نداشته باشند، افزودن کنترل عوارض نیز کارایی قابل توجهی نخواهد داشت.

#### ۴-۳- بررسی تأثیر مدل در مدیریت تصادفات

یکی از مهم‌ترین کاربردهای روش‌های مدیریت ترافیکی، کنترل بزرگراه در شرایط غیرمعمول از جمله رخ دادن تصادفات می‌باشد. برای بررسی عملکرد مدل پیشنهادی در چنین شرایطی یک وضعیت مشابه با تصادف را در پنج روز با احجام مختلف ایجاد نموده و عملکرد مدل در هر یک از این روزها را محاسبه و بررسی می‌شود. بدین منظور در بازه‌ی زمانی بین ثانیه ۳۶۰۰ تا ۳۹۰۰ شبیه‌سازی (۵ دقیقه) یک قطعه‌ی ۵۰ متری در خط سرعت بالادست ورودی دوم را مسدود نموده و اجازه‌ی عبور خودرو را داده نمی‌شود. جدول ۵ نتایج پارامترهای ترافیکی در هر پنج روز، یک‌بار به‌صورت کنترل‌شده و یک‌بار به‌صورت کنترل نشده نشان می‌دهد. مطابق جدول ۵ اجرای سناریو RVT به‌طور میانگین توانسته است

کتر می‌باشد که نشانگر سخت‌تر بودن مدیریت بزرگراه در شرایط تصادف است. البته همچنان مدل توانسته است کارایی مناسبی را در این شرایط بدون نشان دهد.

زمان سفر را ۴٫۸ درصد، سرعت را ۱۱٫۵ درصد، زمان انتظار را ۲۵ درصد، طول صف را ۲۴ درصد، تعداد خودروهای به مقصد رسیده را ۴٫۱ درصد و تراکم شبکه را ۹٫۵ درصد بهبود بخشید. این مقادیر از عملکرد مدل در شرایط بدون تصادف

جدول ۵. تغییرات پارامترهای ترافیکی در دو حالت بودن و نبودن کنترل ترافیکی پیشنهادی

روز	کنترل	زمان سفر ثانیه	سرعت متر بر ثانیه	زمان انتظار ثانیه	طول صف n	خودروهای رسیده	تراکم وسیله بر کیلومتر
۱	N	۹۲۵٫۶۱	۷٫۳۶	۳۲٫۱۱	۲۹۸	۱۱۸۱۰	۴۵٫۸۳
	RVT	۸۸۰٫۳۳	۸٫۲۵	۲۷٫۱۱	۲۳۱	۱۲۴۲۰	۴۱٫۴۵
۲	N	۹۰۱٫۳۴	۷٫۷۲	۲۷٫۵۶	۲۵۶	۱۱۷۱۹	۴۳٫۲۴
	RVT	۸۶۱٫۷۹	۸٫۴۷	۲۴٫۳	۱۹۴	۱۲۲۸۵	۳۹٫۵۶
۳	N	۸۷۹٫۶۷	۸٫۷۳	۲۴٫۳۷	۲۴۲	۱۱۵۵۶	۳۹٫۶۲
	RVT	۸۲۷٫۲۸	۸٫۸۹	۱۱٫۵۲	۱۸۹	۱۲۱۲۲	۳۶٫۳۸
۴	N	۸۱۸٫۹۶	۸	۱۲٫۹۹	۲۱۲	۱۱۴۶۷	۳۷٫۴۵
	RVT	۷۹۰٫۳	۹٫۴۵	۹٫۶۴	۱۶۲	۱۱۷۶۲	۳۳٫۹۷
۵	N	۷۹۳٫۹۴	۸٫۷۱	۱۵٫۵۲	۱۷۶	۱۱۸۱۰	۳۳٫۷۲
	RVT	۷۵۱٫۳۲	۱۰٫۱۵	۱۰٫۹۸	۱۲۰	۱۱۱۳۶	۲۹٫۴۹

### ۵- نتیجه‌گیری

بررسی اثر ترکیب این روش با دو روش قبلی پرداخته شده است و با شبیه‌سازی در نرم‌افزار خرد نگر سومو خروجی‌های پارامترهای ترافیکی موردنظر را در ۸ حالت مختلف دریافت شد. خروجی‌های مقایسه‌ی ۸ سناریوی مختلف را در جداول مربوطه آورده شده است. مشاهده شد که بهترین عملکرد را حالت سه کنترل و پس‌از آن حالت دو کنترل‌ها ترکیب رمپ و سرعت داشتند. بهبود اثر بر زمان سفر از ۹٫۴۱ به ۱۰٫۴۱ درصد کاهش، بهبود عملکرد سرعت از ۱۷٫۵ درصد به ۲۴٫۳ درصد افزایش، افزایش تعداد خودروهای به مقصد رسیده از ۲٫۳۶ درصد به ۳ درصد، بهبود تغییرات چگالی از ۱۳٫۷۱ درصد به ۱۶٫۷۱ درصد کاهش و بهبود درصد اشغال از ۱٫۸۷ درصد به ۱۳٫۲۸ درصد کاهش نشان‌دهنده‌ی عملکرد مناسب‌تر روش سه کنترل نسبت به سناریو ۳ (رمپ و سرعت) می‌باشد. ممکن است ترکیب روش‌های دیگری همچون کنترل تغییر خط، پیشنهاد مسیریابی به لحظه به رانندگان، رانندگی در شانه و ... نیز منجر به بهبود عملکرد مدل نسبت به مدل ارائه‌شده در این مقاله نیز گردد. همچنین استفاده از داده‌های به‌روز و لحظه‌ای

باوجود تأثیر مناسب اجرای منفرد کنترل‌های ترافیکی همچون کنترل رمپ، سرعت مجاز متغیر و عوارض بزرگراهی، به علت عدم رفع کامل تداخلات ترافیکی و بالا بودن میزان تأخیر بزرگراهی نیاز به ترکیب روش‌های کنترلی و پیاده‌سازی هماهنگ آن‌ها رویکردی برای دستیابی به نتایج بهتر می‌باشد. تحقیقات متعددی به بررسی عملکرد ترکیب کنترل‌های رمپ و سرعت پرداخته‌اند. نتیجه‌ی این تحقیقات بهبود عملکرد مدل پیشنهادی نسبت به حالت تک کنترل بوده است اما همچنان می‌توان با استفاده از روش‌های دیگر به نتایج بهتری نیز دست‌یافت. کنترل نرخ رمپ و سرعت مجاز متغیر روش‌هایی هستند که برای مدیریت تقاضای ورودی به یک بزرگراه استفاده می‌شوند. اما غالبا اثر چندانی بر روی انتخاب مسیر رانندگان ندارند. کنترل اعمال عوارض بزرگراهی از طرفی این توانایی را دارد که در شرایطی که مسیر اصلی ترافیک سنگینی دارد و ورود خودروهای جدید می‌تواند وضعیت را بحرانی‌تر کند، افزایش عوارض مسیر اصلی می‌تواند منجر به تغییر مسیر بسیاری از خودروها به مسیرهای جانبی گردد. در این مقاله به

International Journal of Urban Sciences, 21(Sup1), pp.3-18.

-Chiou, Y.C., Huang, Y.F. And Lin, P.C., (2012.), "Optimal Variable Speed-Limit Control under Abnormal Traffic Conditions", Journal of the Chinese Institute of Engineers, 35(3), pp.299-308.

-Davarynejad, M., Hegyi, A., Vrancken, J. And Van Den Berg, J., (2011), "Motorway Ramp-Metering Control with Queuing Consideration Using Q-Learning, in 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), IEEE, pp. 1652-1658.

-Fares, A. and Gomaa, W., (2014), June, "Freeway Ramp-Metering Control Based on Reinforcement Learning", in 11th IEEE International Conference on Control & Automation (ICCA), IEEE., pp. 1226-1231.

-François-Lavet, V., Henderson, P., Islam, R., Bellemare, M.G. and Pineau, J., (2018), "An Introduction to Deep Reinforcement Learning", Arxiv Preprint Arxiv:1811.12560.

-Fu, R., Zhang, Z. And Li, L., (2016), "Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction", in 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), IEEE, pp. 324-328.

-Ghods, A.H., Kian, A.R. And Tabibi, M., (2009), "Adaptive Freeway Ramp Metering and Variable Speed Limit Control: A Genetic-Fuzzy Approach", IEEE, Intelligent Transportation Systems Magazine, 1(1), pp.27-36.

-Hegyi, A., De Schutter, B. And Hellendoorn, H., (2005), "Model Predictive Control for Optimal Coordination of Ramp Metering and Variable Speed Limits", Transportation Research Part C: Emerging Technologies, 13(3), pp.185-209.

-Khazraeian, S., Koohifar, F. And Kalantari, N., (2017), "A Nonlinear Optimal Static Controller for Ramp Control (NOSCO)", in Proceedings of the 96th Annual Meeting of the Transportation Research Board.

-Khondaker, B. And Kattan, L., (2015), "Variable Speed Limit: An Overview", Transportation Letters, 7(5), pp.264-278.

-Kušić, K., Dusparić, I., Guériau, M., Gregurić, M. And Ivanjko, E., (2020), "Extended Variable Speed Limit Control Using Multi-Agent Reinforcement Learning, in 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), IEEE, pp. 1-8.

-Li, Z., Xu, C., Guo, Y., Liu, P. And Pu, Z., (2020), "Reinforcement Learning-Based

مثل استفاده از روش‌های پیش‌بینی، دوربین‌ها و آشکارسازها برای به دست آوردن اطلاعات می‌تواند کمک شایانی به مدل برای تصمیم‌گیری نماید.

#### ۶- پی‌نوشت‌ها

1. Ramp Metering
2. Variable Speed Limit
3. Lane Change Control
4. Online Route Guidance
5. Reinforcement Learning
6. Double Deep Nueral Network
7. ALIENA
8. Feedback Control
9. METALINE
10. Heuristic Ramp-Metering Coordination (HERO)
11. Genetic Fuzzy Logic Control (GFLC)
12. Model Predictive Control
13. Cascade Feedback
14. Proximal Policy Optimization (PPO)
15. Long – Short Term Memory (LSTM)
16. Actor–Critic Method
17. Recurrent Nueral Network
18. Fine Tuning

#### ۷- مراجع

-Arango, M., (2019), "Toll Road With Dynamic Congestion Pricing Using Reinforcement Learning".

-Belletti, F., Haziza, D., Gomes, G. And Bayen, A.M., (2017), "Expert Level Control of Ramp Metering Based on Multi-Task Deep Reinforcement Learning", IEEE Transactions on Intelligent Transportation Systems, 19(4), pp.1198-1207.

-Carlson, R.C., Manolis, D., Papamichail, I. And Papageorgiou, M., (2012), "Integrated Ramp Metering and Mainstream Traffic Flow Control on Freeways Using Variable Speed Limits", Procedia-Social and Behavioral Sciences, 48, pp.1578-1588.

-Carlson, R.C., Papamichail, I. And Papageorgiou, M., (2014), "Integrated Feedback Ramp Metering and Mainstream Traffic Flow Control on Motorways Using Variable Speed Limits", Transportation Research Part C: Emerging Technologies, 46, pp.209-221.

-Carlson, R.C., Papamichail, I., Papageorgiou, M. And Messmer, A., (2010), "Optimal Motorway Traffic Flow Control Involving Variable Speed Limits and Ramp Metering", Transportation Science, 44(2), pp.238-253.

-Cheng, Q., Liu, Z., Liu, F. and Jia, R., (2017), "Urban Dynamic Congestion Pricing: An Overview and Emerging Research Needs",

- Robotics, and Autonomous Systems, 2, pp.253-279.
- Rezaee, K., Abdulhai, B. And Abdelgawad, H., (2012), September, "Application of Reinforcement Learning with Continuous State Space to Ramp Metering in Real-World Conditions", in 2012 15th International IEEE Conference on Intelligent Transportation Systems, IEEE, pp. 1590-1595.
- Schmidt-Dumont, T. And Van Vuuren, J.H., (2015), "Decentralised Reinforcement Learning for Ramp Metering and Variable Speed Limits on Highways", IEEE Transactions on Intelligent Transportation Systems, 14(8), p.1.
- Schmidt-Dumont, T. And Van Vuuren, J.H., (2019), "A Case for the Adoption of Decentralised Reinforcement Learning for The Control of Traffic Flow on South African Highways", Journal of the South African Institution of Civil Engineering, 61(3), pp.7-19.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A. And Klimov, O., (2017), "Proximal Policy Optimization Algorithms", Arxiv Preprint Arxiv:1707.06347.
- Sewak, M., (2019), "Deep Reinforcement Learning", Singapore: Springer Singapore.
- Shaaban, K., Khan, M.A. And Hamila, R., (2016), "Literature Review of Advancements in Adaptive Ramp Metering", Procedia Computer Science, 83, pp.203-211.
- Siami-Namini, S., Tavakoli, N. And Namin, A.S., (2018), "A Comparison of ARIMA And LSTM In Forecasting Time Series", in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, pp. 1394-1401.
- Vinitsky, E., Parvate, K., Kreidieh, A., Wu, C. and Bayen, A., (2018), "Lagrangian Control through Deep-RL: Applications to Bottleneck Decongestion", in 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, pp. 759-765.
- Yamak, P.T., Yujian, L. and Gadosey, P.K., (2019), "A Comparison Between Arima, Lstm, And Gru for Time Series Forecasting", In Proceedings of the 2019 2<sup>nd</sup> International Conference on Algorithms, Computing and Artificial Intelligence pp. 49-55.
- Zhang, Y., Atasoy, B., Akkinepally, A. and Ben-Akiva, M., (2019), "Dynamic Toll Pricing Using Dynamic Traffic Assignment System with Online Calibration", Transportation Research Record, 2673(10), pp.532-546.
- Variable Speed Limits Control to Reduce Crash Risks Near Traffic Oscillations on Freeways", IEEE Intelligent Transportation Systems Magazine, 13(4), pp.64-70.
- Lombardi, C., Picado-Santos, L. and Annaswamy, A.M., (2021), "Model-Based Dynamic Toll Pricing: An Overview", Applied Sciences, 11(11), p.4778.
- Lu, X.Y., Varaiya, P., Horowitz, R., Su, D. and Shladover, S.E., (2011), "Novel Freeway Traffic Control with Variable Speed Limit and Coordinated Ramp Metering", Transportation Research Record, 2229(1), pp.55-65.
- Ma, M.H., Yang, Q.F., Liang, S.D. And Li, Z.L., (2015), "Integrated Variable Speed Limits Control and Ramp Metering for Bottleneck Regions on Freeway", Mathematical Problems in Engineering.
- Mahajan, N., Hegyi, A., Van De Weg, G.S. And Hoogendoorn, S.P., (2015), "Integrated Variable Speed Limit and Ramp Metering Control Against Jam Waves--A COSCAL V2 Based Approach", in 2015 IEEE 18th International Conference on Intelligent Transportation Systems, IEEE, pp. 1156-1162.
- Mizuta, A., Roberts, K., Jacobsen, L., Thompson, N. And Colyar, J., (2014), "Ramp Metering: A Proven, Cost-Effective Operational Strategy: A Primer (No. FHWA-HOP-14-020)", United States, Federal Highway Administration.
- Papageorgiou, M. , Hadj-Salem, H. And Blosseville, J.M., 1991," ALINEA: A Local Feedback Control Law for On-Ramp Metering", Transportation Research Record, No. 1320, Transportation Research Board, Washington, D.C., pp. 58 – 64
- Papageorgiou, M., Blosseville, J. M., And Hadj Salem, H., (1990), "Modeling and Real-Time Control of Traffic Flow on The Southern Part of Boulevard Peripherique in Paris: Part II: Coordinated On-Ramp Metering", Transportation Research. Vol. 24A, No. 5, pp. 361-370.
- Papamichail, I., Kampitaki, K., Papageorgiou, M. And Messmer, A., (2008), "Integrated Ramp Metering and Variable Speed Limit Control of Motorway Traffic Flow", IFAC Proceedings Volumes, 41(2), pp.14084-14089.
- Papamichail, I., M.Papageorgiou, V. Vong, & J. Gaffney, (2010), "Heuristic Ramp-Metering Coordination Strategy Implemented at Monash Freeway, Australia, Research Record", Journal of the Transportation Research Board, No. 2178, 2010, pp. 10–20.
- Recht, B., (2019), "A Tour of Reinforcement Learning: The View from Continuous Control", Annual Review of Control,

# Optimization Model of Ramp Input Rate, Variable Speed Limit and Highway Tolls based on Reinforcement Learning

*Shahriar Afandizadeh, Professor, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran.*

*Mohammad Ali Hassanzadeh, M.Sc., Student, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran.*

*E-mail: zargari@iust.ac.ir*

Received: March 2023- Accepted: August 2023

## **ABSTRACT**

The non-stop growth of urban traffic, especially on highways, has led to traffic jams and, as a result, long queues, increased travel time, increased chances of accidents, and environmental pollution. Since the extensive development of infrastructures cannot be a permanent solution to solve such problems, the existing facilities should be used optimally and appropriately. Several methods, such as ramp metering control, variable speed limit, lane change control, tolls, etc., have been studied and investigated to control highways. One of the approaches to improve the performance of these controls is their simultaneous and coordinated integration and implementation. In this paper, we will investigate the effect of the combination of three concurrent controls of ramp metering, variable speed limit, and highway tolls on the traffic conditions of the highway network. For this purpose, the reinforcement learning method has been used to manage the control values as variables. Traffic data is first predicted by predictive models based on neural networks and then given to the simulation of Sumo software. The controller model then selects the ramp rate, speed limit, and toll values based on the simulation outputs and apply them to the network. The results of implementing the proposed model show the significant impact of the model in improving traffic parameters such as average travel time, the average speed of cars, waiting time, queue length, average network density, and the number of vehicles arriving at their destination. The model reduced the average travel time by 11.5% compared to the condition without controls, which is more than the effect of implementing single controls or a double combination of the mentioned methods. In addition, in the event of an accident in the network, the proposed management can significantly improve the network conditions compared to the situation without control.

**Keywords:** Ramp Metering, Variable Speed Limit, Highway Toll, Reinforcement Learning, Neural Networks