TU Delft

Delft University of Technology

Real-time scheduling and routing of shared autonomous vehicles considering platooning in intermittent segregated lanes and priority at intersections in urban corridors

Wang, Zhimian; An, Kun; Correia, Gonçalo; Ma, Wanjing

DOI
10.1016/j.tre.2024.103546

Publication date
2024

Document Version
Final published version

Published in
Transportation Research Part E: Logistics and Transportation Review

Citation (APA)
Wang, Z., An, K., Correia, G., & Ma, W. (2024). Real-time scheduling and routing of shared autonomous vehicles considering platooning in intermittent segregated lanes and priority at intersections in urban corridors. *Transportation Research Part E: Logistics and Transportation Review*, *186*, Article 103546. https://doi.org/10.1016/j.tre.2024.103546

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Transportation Research Part E 186 (2024) 103546

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

Real-time scheduling and routing of shared autonomous vehicles considering platooning in intermittent segregated lanes and priority at intersections in urban corridors

Zhimian Wang [a], Kun An [a,*], Gonçalo Correia [b], Wanjing Ma [a]

[a] The Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China
[b] Department of Transport & Planning, Delft University of Technology, Mekelweg 8 2628 CD Delft, the Netherlands

ARTICLE INFO

Keywords:
Shared autonomous vehicles
Vehicle platooning
Traffic priority
Vehicle scheduling and routing
Two-stage algorithm

ABSTRACT

Anticipating the forthcoming integration of shared autonomous vehicles (SAVs) into urban networks, the imperative of devising an efficient real-time scheduling and routing strategy for these vehicles becomes evident if one is to maximize their potential in enhancing travel efficiency. In this study, we address the problem of jointly scheduling and routing SAVs across an urban network with the possibility of platooning the vehicles at intersections to reduce their travel time. We argue that this is especially useful in large urban areas. We introduce a novel vehicle scheduling and routing method that allows a specific number of SAVs to converge at the intersections of urban corridors within designated time intervals, facilitating the formation of SAV platoons. Dedicated lanes and signal priority control are activated to ensure that these platoons go through the corridors efficiently. Based on the above concept, we propose a linear integer programming model to minimize the total travel time of SAVs and the delays experienced by the conventional vehicles due to SAV priority, thereby optimizing the overall performance of the road network. For large instances, we develop a two-stage heuristic algorithm to solve it faster. In the first stage, leveraging an evaluation index that manifests the compatibility of each vehicle-to-request combination, we allocate passenger requests to a fleet of SAVs. In the second stage, a customized genetic algorithm is designed to coordinate the paths of various SAVs, thus achieving the desired vehicle platooning effect. The optimization method is tested on a real-world road network in Shanghai, China. The results display a remarkable reduction of 15.76 % in the total travel time of the SAVs that formed platoons. The overall performance of the road network could be improved with the total travel time increase of conventional vehicles significantly smaller than the reduction observed in SAVs' total travel time.

1. Introduction

The increase in the adoption of private cars has exacerbated traffic congestion in numerous metropolitan areas. Public transport is regarded as one of the solutions to mitigate the traffic congestion problem. Shared autonomous vehicle (SAV) services, a future transportation mode combining the benefits of public transportation and the convenience of private cars, offer new options for

* Corresponding author.
E-mail addresses: 2310819@tongji.edu.cn (Z. Wang), kunan@tongji.edu.cn (K. An), g.correia@tudelft.nl (G. Correia), mawanjing@tongji.edu.cn (W. Ma).

https://doi.org/10.1016/j.tre.2024.103546
Received 15 November 2023; Received in revised form 18 April 2024; Accepted 19 April 2024
Available online 2 May 2024
1366-5545/© 2024 Elsevier Ltd. All rights reserved.

travelers. With the development of autonomous vehicle (AV) technology, and the enhancement of shared mobility platforms, SAVs are expected to become an important part of the urban mobility system in the future.

Previous studies have shown that autonomous vehicles (AVs) can increase the road throughput by driving with a much smaller headway compared to human-driven vehicles (Ye et al., 2018; Mohajerpoor et al., 2019). In order to promote the adoption of AVs, various priority strategies have been proposed and tested in previous studies. Chen et al. (2016) proposed dedicated lanes on major roads for AVs, which proved to be an effective way to promote the adoption of AVs. However, the deployment of AV-dedicated lanes requires a relatively high AV penetration rate, otherwise the reduction in the number of lanes will greatly affect the normal operation of the other conventional vehicles, which is likely to cause a series of negative issues, such as vehicle queuing, traffic congestion, and even accidents (Liu et al., 2019).

In this paper, we propose an innovative approach for the allocation of time and space resources at intersections and road segments among SAVs and private conventional vehicles in a mixed-traffic urban network. With carpooling granting SAVs the qualities of public transportation with regards to seat occupancy, SAV platoons, capable of carrying multiple passengers, should in our view be entitled to priority treatment, much like traditional buses. Viegas et al. (2004) considered an Intermittent Bus Lane (IBL) system used for bus priority, where a lane's status changes based on the presence of a bus within its spatial domain: turning into a "BUS lane" when a bus approaches and reverting to a regular lane open to general traffic once the bus departs. Inspired by this concept, we specify that the dedicated SAV lanes and signal priority control on a main road, which are collectively referred to as "priority services" in our study, are activated solely when necessary. These priority services remain dormant until the conditions for SAV platooning are satisfied, necessitating a sufficient number of SAVs concentrated at the upstream intersection within designated time intervals.

We envision a dedicated SAV lane that is reserved on the right-hand side of the road, complete with a harbor-style parking area designed to facilitate the assembly of SAV platoons. The SAVs that arrive earlier can wait at the parking area until the other vehicles arrive, as shown in Fig. 1. In this way, these vehicles will cause no nuisance while waiting. Simultaneously, platooning the vehicles will concentrate the priority given to SAVs only to a specific period of time thus reducing the negative effects on the conventional vehicles which are temporarily not allowed to enter the lane.

Since platooning is not a spontaneous behavior of the vehicles, a systematic and efficient scheduling and routing method is needed to both coordinate the assignment of passengers to the vehicles and the platoons. The SAVs engage in space–time coordination with each other to attain the platooning effect and minimize their total travel time through priority services while adhering to the specified time window constraints for the passengers. The added priority feature including the platoons reflects a significant advancement when compared to the state-of-the-art SAV scheduling problems that typically consider only carpooling and vehicle dispatching.

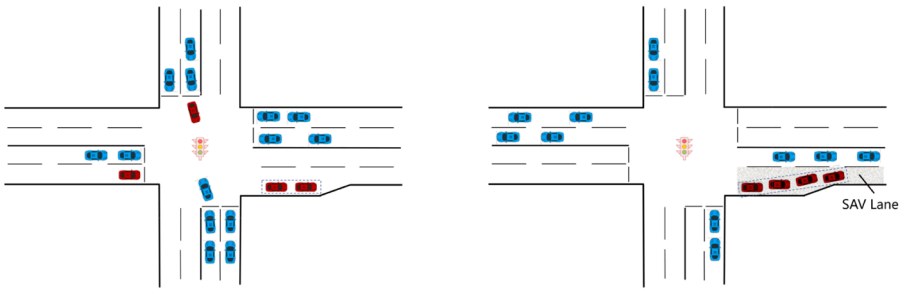
The paper is organized as follows. Firstly, we review the literature regarding AV priority services and SAV scheduling problem in Section 2. Then we formulate a linear integer programming model for SAV scheduling and routing in Section 3. Next, a two-stage heuristic algorithm is proposed to accelerate the solving of the problem in Section 4. Then in Section 5 we report on the tests that we have run on the performance of the proposed model as well as the solution algorithm for the case study of Shanghai, China. Conclusions and future research directions are finally presented in Section 6.

2. Literature review

2.1. AV priority strategies

So-called AV priority strategies are employed to enhance the efficiency of traffic management systems in granting priority to AVs over other vehicles on the road. This performance enhancement primarily focuses on factors such as travel time, reduced delays, and improved safety, all of which are vital for ensuring the seamless and effective operation of AVs.

The deployment of AV-dedicated lanes has been proposed as an effective priority strategy that not only ensures uninterrupted movement for AV operations but also promotes private AV ownership (Chen et al., 2016; Liu et al., 2019; Pourgholamali et al., 2023; Razmi Rad et al., 2020). Chen et al. (2016) studied the impact of AV market penetration on a dedicated lane reservation scheme. Their findings indicated that it is advisable to implement AV-dedicated lanes progressively rather than aggressively before the market penetration reaches a relatively high level, approximately around 20 %. Ghiasi et al. (2017) introduced a capacity analysis model that



(a) Temporary parking before platooning (b) Formation of SAV platoon

Fig. 1. Process of SAV platooning.

centers on a non-uniform and stochastic headway space distribution modeled as a Markov chain. The researchers devised a lane management model for highways, which identifies the ideal number of dedicated connected automated vehicle (CAV) lanes across different demand levels, CAV market penetration rates, and CAV fleet size, with the goal of optimizing the capacity of mixed traffic flows with both CAV and human-driven vehicles (HVs). Seilabi et al. (2023) introduced a robust optimization model for the deployment of CAV-dedicated lanes in a highway network, which aimed to address the inherent uncertainty associated with the forecast of the potential CAV market size. Liu et al. (2019) investigated the impact of exclusive lanes within the context of mixed traffic flow involving both AVs and HVs in a road network. It was found that deploying dedicated AV lanes across the entire network can be impractical, particularly when the AV penetration rate is low, as these lanes might even diminish overall traffic efficiency. As an alternative solution, they introduced a novel concept known as autonomous vehicle/toll (AVT) lanes. These lanes are not exclusive to AVs alone but also permit HVs to access them by paying a toll. The results of the study demonstrated that the settings of AVT lanes can significantly improve the overall performance of the transportation system. Building upon this research, Wang et al. (2021) further explored the optimal toll rates for AVT lanes.

We can observe that several studies have explored the deployment of AV-dedicated lanes to minimize the total travel time for AVs, thereby alleviating traffic congestion in urban road networks. Further, compared to AVs, the significant increase in travel times for HVs raises concerns about social equity, highlighting the need for a more balanced spatiotemporal distribution of resources on a road network for both AVs and HVs. The feasibility of implementing fixed AV dedicated lanes is challenged by relatively low AV market penetration rates. To address this challenge, a promising approach could involve the temporary activation of dedicated lanes, similar to the intermittent dedicated lanes used for bus priority management (Viegas et al., 2004). However, few studies have yet explored the use of such intermittent dedicated lanes for AV platoons.

Transit signal priority (TSP), as another critical priority strategy, is designed to facilitate the movement of specific types of vehicles at signalized intersections by adjusting traffic signal timings (Truong et al., 2018). TSP strategy has been widely explored in the management of conventional buses (Anderson et al., 2020; Zeng et al., 2020; Liang et al., 2023; Xu et al., 2023; Truong et al., 2019). However, instead of focusing solely on TSP strategy, more research effort has been directed in the past years towards signal-free intersection management for AVs, such as the autonomous intersection manager (AIM) protocol, which has been proposed as an alternative to traffic signals for AVs (Dresner and Stone, 2004; Wu et al., 2019; Antonio and Maria-Dolores, 2022; Li et al., 2023). A few studies have also examined configurations where AVs and HVs share traffic intersections. Levin and Boyles (2016) found that a high AV market penetration rate (around 80 %) was necessary for employing AIM as an alternative to signal control. Rey et al. (2019) investigated a novel traffic control policy that accommodates both HVs and AVs sharing the road infrastructure. Assuming the availability of dedicated AV lanes at intersections, they established a dedicated phase during which only AVs were permitted to traverse the intersection, thus coordinating the operation of AVs and HVs.

Therefore, in mixed flow scenarios, where both AVs and HVs are present, AVs can be regulated using the AIM protocol, while signals remain reserved for HVs. Specifically, when AV penetration rates are relatively low, both AVs and HVs should be managed through signal control. In such cases, employing a TSP strategy could serve as an effective solution for alleviating delays for AVs at signaled intersections. In this study, we consider periodically activating TSP management in conjunction with the deployment of intermittent dedicated lanes to provide priority for AV platoons.

2.2. SAV scheduling problem

The emergence of SAVs offers new solutions for urban mobility, presenting a cost-effective and highly flexible mode of transportation for short-haul travelers. However, the inherent spatiotemporal disparities in travel demand require effective vehicle scheduling decisions. To enhance vehicle accessibility and promote carpooling for using capacity in the most efficient way, the optimization of SAV scheduling assumes a vital role, and this area of study has undergone substantial exploration over the past decade (Gurumurthy and Kockelman, 2018; Simonetto et al., 2019; Alazzawi et al., 2018; Cokyasar and Larson, 2020; Seo and Asakura, 2022).

Levin et al. (2019) conducted research on the operational aspects of SAVs, with the impact of traffic congestion taken into consideration. They proposed a linear programming model to address the SAV routing problem, based on the link transmission model. Furthermore, they explored the utilization of SAVs as a first-mile and last-mile solution, which demonstrated SAVs' performance as part of the public transport system. Scheltes et al. (2017) studied a system called Automated Last-Mile Transport (ALMT), which involves a fleet of small, fully automated electric vehicles designed to enhance the last-mile performance of a trip typically taken by train passengers. They developed an agent-based simulation model for ALMT, which included a dispatching algorithm responsible for allocating travel requests among the fleet of available vehicles. The case study results suggested that the performance of ALMT system was comparable to walking, but further enhancements were required to make it competitive with cycling. Lokhandwala et al. (2018) analyzed the merits and drawbacks of shared autonomous taxis, using taxi data from New York City. Their findings demonstrated a substantial 59 % reduction in vehicle fleet size when transitioning from traditional taxis to shared autonomous taxis while maintaining the same service level. Additionally, carpooling with SAVs led to a remarkable increase in occupancy rates from 1.2 to 3, resulting in a significant 55 % reduction in total travel distance. Nevertheless, a side effect was observed, with the reduced fleet size taxis were concentrated in areas of high demand, potentially diminishing service levels in suburban areas. Hasan et al. (2021) investigated the benefits of AVs and carpooling platforms in community-based trip sharing. They developed a column generation program to construct optimized mini routes for serving both inbound and outbound trips. The optimization results showed that their proposed approach reduced daily vehicle usage by approximately 92 % with AVs, concurrently reducing daily vehicle miles driven by 30 %.

In general algorithms designed to address the SAV scheduling problem can be categorized into two types: exact solution algorithms (Liang et al., 2020) and heuristic algorithms (Ropke and Pisinger, 2006; Sun et al., 2018; Tafreshian and Masoud, 2020). Exact solution

algorithms rely on rigorous mathematical formulations and methodically seek the best-feasible solution, often with awareness of an optimality gap. However, these exact algorithms have inherent limitations in real-world applications, as their computational complexity grows exponentially with the problem's scale, resulting in extremely long computation time. In recent years, researchers have increasingly turned to heuristic algorithms to solve practical vehicle routing problems. For instance, Cherkesly et al. (2015) introduced a population-based heuristic algorithm that combines a genetic algorithm with a local search approach to tackle passenger pickup and delivery problems. The results of their case studies demonstrated the algorithm's ability to handle 200 requests within one hour and 300 requests within three hours, with average deviations from known optimal solution values ranging from 0.17 % to 2.84 %. Farhan et al. (2018) addressed the carpooling problem involving SAVs, employing a "cluster first, path second" strategy. They decomposed the original problem into two steps: initially assigning requests to vehicles using a shortest-path algorithm and subsequently constructing each vehicle's path via a tabu search algorithm. Sun et al. (2020) studied the pickup and delivery problems considering time window constraints and time-dependent profit. They improved the Adaptive Large Neighborhood Search (ALNS) algorithm by incorporating an approximation of the objective function with taboo lists and a hybrid stopping criterion. Several removal operators and insertion operators were also designed to adapt to the characteristics of the proposed problem.

Despite abundant literature on SAV scheduling problems and their corresponding solution algorithms, most existing studies have rarely taken into account the influence of SAV platoons on scheduling strategies. SAV platooning necessitate the synchronization of diverse SAVs in carpooling and routing process, resulting in an SAV scheduling problem with extraordinary complexity. Consequently, there arises a pressing need for a comprehensive, unified mathematical model that jointly optimizes both SAV scheduling and platooning.

For bridging that gap, we present in this paper a linear integer programming model aimed at optimizing the scheduling, routing, and platooning strategies for SAVs. This model enables SAVs to collaborate with one another to leverage platooning benefits and reduce their collective travel time through priority services. The solution method, grounded in heuristic algorithms, is tailored to yield high-quality solutions within a relatively short calculation time, rendering it suitable for prospective real-time SAV scheduling systems.

3. Model formulation

In this section, we introduce the problem and explain the basic assumptions and notations to support its formulation. Then we present all constrains and the objective function and explain how each one works in the model.

3.1. Problem description

We consider an SAV travel system designed to facilitate passenger travelling through SAV scheduling and routing as well as platooning in intermittent segregated lanes along urban corridors. Passengers submit their travel requests in the form of reservations. The SAV operator gathers information on both vehicle supply and passenger demand, which serves as a critical input for optimizing vehicle scheduling. SAVs follow the scheduled routes to pick up and drop off passengers and may form platoons on specific link segments if this improves traffic conditions. Signal priority at intersections and lane segregation on these specific link segments are activated appropriately to enhance the travel efficiency of SAV platoons.

The system operates on a traffic network that can be represented by graph $G = (N, A)$ in which N is the set of nodes and A is the set of links. The study horizon $\xi\Gamma$ is partitioned into ξ scheduling periods. Each scheduling period is furtherly divided into $\Gamma = \{t\}$ timesteps, with the platooning of SAVs taking place within the duration of a single time step. In each scheduling period, passenger

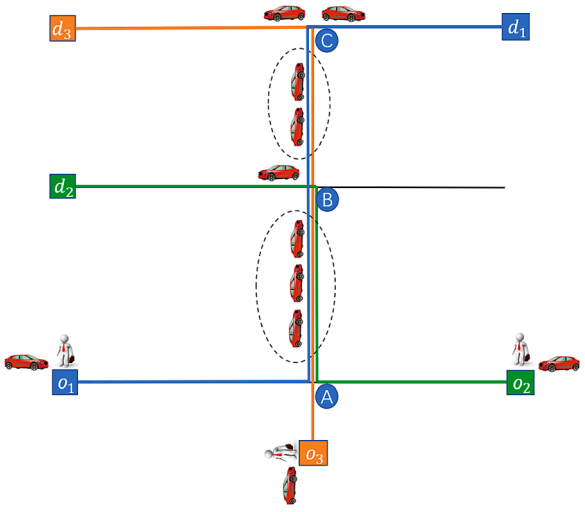


Fig. 2. Route schedule for SAV platooning.

requests $k \in K$ are submitted to the system with their pick-up time window $[dt_k^{min}, dt_k^{max}]$ and arrival time window $[at_k^{min}, at_k^{max}]$. Both the pick-up and drop-off locations are positioned at the nodes of the network. There are $H = \{h\}$ vehicles in the network. Vehicles could be idling at the nodes or driving on the links with or without passengers. At the beginning of the study horizon, the vehicles are positioned on the depot node of the network which is pre-defined. At the start of each scheduling period, we schedule the SAVs to fulfill the users' travel demand that have been reserved within the previous period. This allocation is based on the available vehicle information including the current node in which vehicles are located, current vehicle status (i.e., free or with passengers), and how long for the vehicle to finish the current trip and be free to accept new passengers. An SAV must complete its service task and become free status before it qualifies for dispatch in the subsequent period.

Carpooling services are managed by the SAV operator. Our operational model assumes that SAVs initially pick up passengers who will be sharing the ride and subsequently drop them off one by one. This entire process constitutes a single trip for the SAV. The SAV becomes available for a new trip only after delivering all passengers to their respective destinations. In other words, within a single SAV trip, drop-offs must occur after all passengers have been picked up. We do not consider dynamic ride-sharing scenarios where a vehicle can pick up new passengers while already en route. Therefore, an SAV trip is divided into two segments: the pick-up itinerary with $\eta = 1$ and the drop-off itinerary with $\eta = 2$. For example, an SAV starting from its original node o, fulfills three requests whose OD nodes are (r_1, s_1), (r_2, s_2), (r_3, s_3) with a route that goes through $o \rightarrow r_1 \rightarrow r_3 \rightarrow r_2 \rightarrow s_3 \rightarrow s_2 \rightarrow s_1$. The route of $o \rightarrow r_1 \rightarrow r_3 \rightarrow r_2$ is the pick-up itinerary with index $\eta = 1$, whereas $r_2 \rightarrow s_3 \rightarrow s_2 \rightarrow s_1$ is the drop-off itinerary with index $\eta = 2$. The route such as $o \rightarrow r_1 \rightarrow r_2 \rightarrow s_2 \rightarrow r_3 \rightarrow s_3 \rightarrow s_1$ is not allowed in our model since the pick-up of request r_3 happens after the drop-off of request s_2.

Additionally, we try to coordinate a sufficient number of SAVs to benefit from priority services in the form of platoons. As illustrated in Fig. 2, three SAVs originating from distinct points o_1, o_2, o_3 are scheduled to form a platoon at intersection A. Subsequently, a temporarily dedicated SAV lane connecting A to B, together with the signal priority control at both A and B, is activated to minimize the delays of the SAV platoon at link segments and intersections. The SAV traveling along the green route disassembles from the platoon at intersection B, whereas the SAVs following the blue and orange routes diverge from one another at intersection C. Afterward, These SAVs proceed to their respective destination points, namely d_1, d_2, d_3.

The following is the set of assumptions underlying the formulation of the SAV scheduling and routing problem.

Assumption 1. The SAVs are homogeneous in terms of their capacity and operation speed.

Assumption 2. Requests should be served by the SAVs in their preferred time window. Additional waiting is not allowed. Otherwise, the passenger is considered as non-served.

3.2. Notations

Notations used throughout the paper are summarized in Table 1.

Table 1
Notation table.

Sets and inputs	
$G = (N,A)$	Road network, where N is the node set and A is the link set
$\Gamma = \{t\}$	Time step index within a scheduling period, t marks the instant where the time step begins
$H = \{h\}$	Vehicle index
$K = \{k\}$	Request index
η	Itinerary index, $\eta = 1$ for the pick-up itinerary, $\eta = 2$ for the drop-off itinerary
o_h	Position of vehicle $h \in H$ at the start of the current scheduling period
C_h	Capacity of vehicle $h \in H$
I_h	State of vehicle $h \in H$ at the start of the current scheduling period: equals 1, if vehicle h is free; 0, otherwise
r_k	Pick-up node of request $k \in K$
s_k	Drop-off node of request $k \in K$
$[dt_k^{min}, dt_k^{max}]$	Pick-up time window of request $k \in K$
$[at_k^{min}, at_k^{max}]$	Drop-off time window of request $k \in K$
TT_{ij}	SAV travel time on link (i,j)
l_{ij}^{min}	Minimum number of SAVs required to form a platoon on link $(i,j) \in A$
l_{ij}^{max}	Maximum number of SAVs allowed to join in a platoon on link $(i,j) \in A$
Decision Variables	
f_h	Equals 1, if vehicle h serves any request; 0, otherwise
z_h^k	Equals 1, if vehicle h serves request k; 0, otherwise
χ_k	Equals 1, if request k is rejected; 0, otherwise
$x_{ijt}^{h,\eta}$	Equals 1, if vehicle h enters link (i,j) at the beginning of timestep t in trip itinerary η; 0 otherwise.
$y_h^{k,\eta}$	Equals 1, if vehicle h picks up or delivers request k at the end of itinerary η; 0, otherwise. In other words, for the pick-up itinerary with $\eta = 1$ having $y_h^{k,1} = 1$ indicates that request k is the last request that vehicle h picks up in the current trip. For the drop-off itinerary ($\eta = 2$) $y_h^{k,2} = 1$ indicates that request k is the last request that vehicle h drops off in the current trip.

3.3. Constraints

3.3.1. Vehicle-to-request matching constraints

Constraint (1) imposes that at the start of the scheduling period, only the free SAVs are available to be dispatched.

$$f_h \le I_h \qquad \forall h \in H \tag{1}$$

Constraints (2) and (3) use the decision variable z_h^k to determine the matching between vehicle h and request k. **Constraint (2)** enforces that each request should be served by an available SAV, otherwise it has to be rejected in the current scheduling period (i.e., $\chi_k = 1$). **Constraint (3)** imposes that each SAV utilized in the current period can be matched with multiple requests, within the limits of its maximum capacity.

$$\sum\nolimits_{h\in H} z_h^k = 1 - \chi_k \qquad \forall k \in K \tag{2}$$

$$f_h \le \sum\nolimits_{k\in K} z_h^k \le C_h f_h \qquad \forall h \in H \tag{3}$$

3.3.2. Vehicle routing constraints

Each SAV has two itineraries—one for picking up passengers and the other for dropping them off. The terminal of the pick-up itinerary is the pick-up node of the last passenger who boards the SAV, while the terminal of the drop-off itinerary is the drop-off node of the last passenger who disembarks from the SAV. Nevertheless, the precise locations of these terminals on the network remain uncertain until the vehicle-to-request matching and vehicle routing results are determined. In order to model the terminals in the mathematical program, we add *virtual nodes* to the network, which act as substitute for the uncertain terminal of each itinerary. We connect the virtual nodes to every node on the network via *virtual links*. The travel time on each of these virtual links is equal to 0.

In Fig. 3, we use a hypothetical SAV route as an example to illustrate the concept of virtual nodes and virtual links. The virtual last pick-up node is denoted by d_1 whereas the virtual last drop-off node is denoted by d_2. The SAV first picks up passengers 1, 2, and 3 in sequence and then delivers passengers 2, 3 and 1, one by one. As such, r_3 is the actual terminal of the pick-up itinerary and s_1 is the actual terminal of the drop-off itinerary. For modelling convenience, d_1 and d_2 are added and connected to r_3 and s_1 respectively. The virtual SAV route is $o_h \to r_1 \to r_2 \to r_3 \to d_1 \to r_3 \to s_2 \to s_3 \to s_1 \to d_2$. Please note that d_1 should be connected to the nodes with an undirected link to make sure that the vehicle route is not interrupted. There are three virtual links $r_3 \to d_1$, $d_1 \to r_3$, and $s_1 \to d_2$, which are represented by dashed lines. The actual road links are represented by solid lines.

As previously mentioned, the SAV route consists of two types of itineraries. The links belonging to the pick-up itinerary include $o_h \to r_1 \to r_2 \to r_3 \to d_1$ (represented by gray lines) and the links belonging to the drop-off itinerary include $d_1 \to r_3 \to s_2 \to s_3 \to s_1 \to d_2$ (represented by blue lines). The virtual node d_1 serves as both the end of the pick-up itinerary and the start of the drop-off itinerary, which is a bridge of the two itineraries. The comprehensive diagram of the expanded network, incorporating virtual nodes and links, is presented in Fig. 4. These virtual nodes and virtual links are also included into set N and set A respectively.

In order to ensure the path feasibility considering the virtual links, **Constraints (4) to (5)** are introduced to restrict the opportunity for each SAV to visit the virtual nodes. **Constraint (4)** requires that a dispatched SAV that serves at least one request, should arrive at d_1 at the end of its pick-up itinerary, then it should depart from d_1 at the beginning of its drop-off itinerary, and finally arriving at d_2 at the end of its drop-off itinerary. **Constraint (5)** specifies that the SAV is prohibited from departing from d_1 or visiting d_2 during its pick-up itinerary. **Constraint (6)** mandates that the SAV is not allowed to return to d_1 or depart from d_2 during its drop-off itinerary.

$$\sum\nolimits_{(j,d_1)\in A}\sum\nolimits_{t\in\Gamma} x_{jd_1t}^{h,1} = \sum\nolimits_{(d_1,j)\in A}\sum\nolimits_{t\in\Gamma} x_{d_1jt}^{h,2} = \sum\nolimits_{(j,d_2)\in A}\sum\nolimits_{t\in\Gamma} x_{jd_2t}^{h,2} = f_h \qquad \forall h \in H \tag{4}$$

$$\sum\nolimits_{(d_1,j)\in A}\sum\nolimits_{t\in\Gamma} x_{d_1jt}^{h,1} = \sum\nolimits_{(j,d_2)\in A}\sum\nolimits_{t\in\Gamma} x_{jd_2t}^{h,1} = \sum\nolimits_{(d_2,j)\in A}\sum\nolimits_{t\in\Gamma} x_{d_2jt}^{h,1} = 0 \qquad \forall h \in H \tag{5}$$

$$\sum\nolimits_{(j,d_1)\in A}\sum\nolimits_{t\in\Gamma} x_{jd_1t}^{h,2} = \sum\nolimits_{(d_2,j)\in A}\sum\nolimits_{t\in\Gamma} x_{d_2jt}^{h,2} = 0 \qquad \forall h \in H \tag{6}$$

To ensure the feasibility of SAV routes and that all passengers' pick-up and drop-off nodes are visited as scheduled, a few constraints are given as follows.

Constraint (7) ensures that vehicle h should depart from its initial position o_h, if it has been dispatched to serve at least one request.

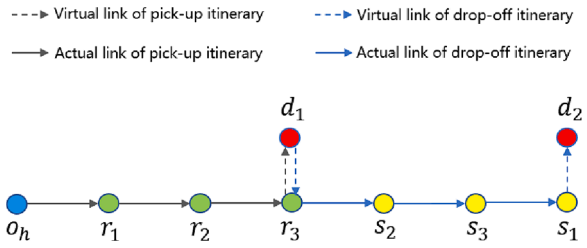


Fig. 3. Diagram for a possible SAV route choice.

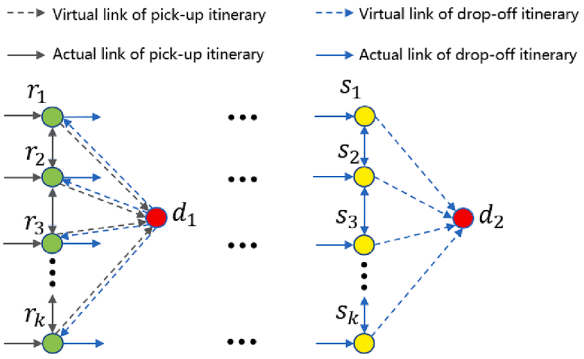


Fig. 4. Diagram for the concept of virtual nodes and virtual links.

$$\sum_{(o_h,j)\in A}\sum_{t\in\Gamma}x^{h,1}_{o_hjt}=f_h \qquad \forall h\in H \tag{7}$$

Constraints (8) to (10) determine the real endpoint (i.e., the last pick-up node) of the pick-up itinerary. **Constraint (8)** imposes that vehicle h can pick up request k as the last passenger during its pick-up itinerary only if vehicle h matches request k. **Constraint (9)** ensures that one and only one request can be the last request that an SAV picks up in the pick-up itinerary. **Constraint (10)** ensures that if request k is the last pick-up request for vehicle h, i.e., $y^{k,1}_h=1$, the vehicle must proceed from pick-up node r_k of request k to enter the virtual node d_1.

$$y^{k,1}_h\le z^k_h \qquad \forall h\in H,\forall k\in K \tag{8}$$

$$\sum_{k\in K}y^{k,1}_h=f_h \qquad \forall h\in H \tag{9}$$

$$\sum_{t\in\Gamma}x^{h,1}_{r_kd_1t}\ge y^{k,1}_h \qquad \forall h\in H,\forall k\in K \tag{10}$$

Constraint (11) is the flow conservation constraint at the virtual node d_1. The SAV must return to the same pick-up node r_k after visiting the virtual node d_1. Since the travel time on virtual links (r_k,d_1) and (d_1,r_k) is zero, the SAV departs and arrives at r_k at exactly the same time t. This is also a connection constraint between the pick-up and drop-off itineraries, where $x^{h,1}_{r_kd_1t}$ with $\eta=1$ specifies the last link in the pick-up itinerary and $x^{h,2}_{d_1r_kt}$ with $\eta=2$ specifies the first link in the drop-off itinerary.

$$x^{h,1}_{r_kd_1t}=x^{h,2}_{d_1r_kt} \qquad \forall h\in H,\forall k\in K,\forall t\in\Gamma \tag{11}$$

Similar to **Constraints (8) to (10)**, **Constraints (12) to (14)** determine the real endpoint (i.e., the last drop-off node) of the drop-off itinerary. **Constraint (12)** imposes that vehicle h can drop off request k as the last passenger during its drop-off itinerary only if vehicle h matches request k. **Constraint (13)** ensures that one and only one request is the last request that an SAV drops off in the drop-off itinerary. **Constraint (14)** ensures that if request k is the last drop-off request for vehicle h, i.e., $y^{k,2}_h=1$, the vehicle must proceed from drop-off node s_k of request k to enter the virtual node d_2.

$$y^{k,2}_h\le z^k_h \qquad \forall h\in H,\forall k\in K \tag{12}$$

$$\sum_{k\in K}y^{k,2}_h=f_h \qquad \forall h\in H \tag{13}$$

$$\sum_{t\in\Gamma}x^{h,2}_{s_kd_2t}\ge y^{k,2}_h \qquad \forall h\in H,\forall k\in K \tag{14}$$

Constraints (15) and (16) ensure that each SAV should arrive and leave the intermediate nodes $j\in N\backslash\{o_h,d_1,d_2\}$on either itinerary the same number of times.

$$\sum_{(i,j)\in A}\sum_{t\in\Gamma}x^{h,1}_{ijt}=\sum_{(j,i)\in A}\sum_{t\in\Gamma}x^{h,1}_{jit} \qquad \forall h\in H,\forall j\in N\backslash\{o_h,d_1,d_2\} \tag{15}$$

$$\sum_{(i,j)\in A}\sum_{t\in\Gamma}x^{h,2}_{ijt}=\sum_{(j,i)\in A}\sum_{t\in\Gamma}x^{h,2}_{jit} \qquad \forall h\in H,\forall j\in N\backslash\{o_h,d_1,d_2\} \tag{16}$$

Constraints (17) and (18) specify that if vehicle h enters link (i,j) at timestep t, it should arrive at node j by timestep $t+TT_{ij}$, where TT_{ij} denotes the total travel time on link (i,j).

$$\sum_{(j,l)\in A}x^{h,1}_{jl\left(t+TT_{ij}\right)}\ge x^{h,1}_{ijt} \qquad \forall h\in H,\forall t\in\Gamma,\forall j\in N\backslash\{o_h,d_1,d_2\},(i,j)\in A \tag{17}$$

$$\sum_{(j,l)\in A} x^{h,2}_{jl\left(t+TT_{ij}\right)} \geq x^{h,2}_{ijt} \qquad \forall h \in H, \forall t \in \Gamma, \forall j \in N\backslash\{o_h, d_1, d_2\}, (i,j) \in A \tag{18}$$

Constraint (19) imposes that the SAV must arrive at the pick-up nodes within the corresponding pick-up time window constraints of the passengers (dt_k^{min}, dt_k^{max}). Similarly, **Constraint (20)** imposes that the SAV must arrive at the drop-off nodes within the corresponding arrival time window constraints of the passengers (at_k^{min}, at_k^{max})

$$\sum_{(j,r_k)\in A} \sum_{t\in\left[dt_k^{min}, dt_k^{max}\right]} x^{h,1}_{jr_kt} \geq z^k_h \qquad \forall h \in H, \forall k \in K \tag{19}$$

$$\sum_{(j,s_k)\in A} \sum_{t\in\left[at_k^{min}, at_k^{max}\right]} x^{h,2}_{js_kt} \geq z^k_h \qquad \forall h \in H, \forall k \in K \tag{20}$$

3.3.3. Vehicle platooning constraints

Vehicles have the potential to form platoons when they converge at the entrance of a link (i,j), provided their arrival times at node i coincide closely in time, specifically falling within the same time step t in this study. To assess whether a platoon can be established, it becomes essential to calculate the number of vehicles entering link (i,j) during time step t. If this number is greater than or equal to l_{ij}^{min}, meaning the minimum number of SAVs required to form a platoon, the SAVs arriving within that particular time step are eligible to receive priority services in the form of platooning.

The SAVs' platooning process entails double constraints involving both time and space, significantly increasing the complexity of the problem. Once again, we employ the convenience of virtual nodes and virtual links to simplify the problem. Consider (p_1, p_2) as a link that is capable of offering priority for SAVs. As shown in Fig. 5, we construct virtual nodes $p_1^{(1)}, \ldots p_1^{(t)} \ldots, p_1^{(m)}$ for the original node p_1, corresponding to time steps 1, 2, 3…t, …m, where m represents the maximum number of time steps in the current scheduling period. This set of virtual nodes is denoted as $\widehat{N}$. Virtual links $\left(p_1, p_1^{(1)}\right), \ldots, \left(p_1, p_1^{(m)}\right)$ are established to connect the upstream node p_1 to these virtual nodes. These virtual links constitute a new subset within the link set A, denoted as $\check{A}$. Simultaneously, virtual links $\left(p_1^{(1)}, p_2\right), \ldots, \left(p_1^{(m)}, p_2\right)$ are constructed to connect the virtual nodes to the downstream node p_2, forming set $\widehat{A}$. In this way, the flow of SAVs traversing through $\left(p_1, p_1^{(t)}\right)$, $\left(p_1^{(t)}, p_2\right)$ represents the number of SAVs visiting link (p_1, p_2) during time step t in the form of a platoon. It is essential to note that since not all road segments are allowed to offer SAV platoon priority we only construct virtual time–space nodes and links for road segments with this function. The travel time along the virtual links in set $\check{A}$ is 0, while the virtual links in set $\widehat{A}$ have a reduced travel time compared to the original link (p_1, p_2).

The vehicle platooning problem can be reformulated as a vehicle routing problem that takes into account the time window constraints and the constraints on the number of SAVs for each virtual link. **Constraint (21)** ensures that the SAVs arriving at the virtual node $i^{(t)} \in \widehat{N}$ at the other timestep $t' \neq t$ (i.e., $t' \in \Gamma\backslash\{t\}$), cannot enter the virtual link $\left(i^{(t)}, j\right)$.

$$x^{h,\eta}_{i^{(t)}jt'} = 0 \qquad \forall h \in H, \forall \eta \in \{1,2\}, \forall t \in \Gamma, \forall \left(i^{(t)}, j\right) \in \widehat{A}, \forall t' \in \Gamma\backslash\{t\} \tag{21}$$

Constraints (22) and (23) restrict the number of platooning vehicles on each virtual link. **Constraint (22)** specifies that if an SAV traverses virtual link $\left(i^{(t)}, j\right)$ (i.e., $x^{h,\eta}_{i^{(t)}jt} = 1$), the total number of SAVs on virtual link $\left(i^{(t)}, j\right)$ should be no less than $l^{min}_{i^{(t)}j}$, the minimum number of SAVs required to form a platoon. The constraint always holds with the right side of the inequality operator equal to 0 (i.e., $x^{h,\eta}_{i^{(t)}jt} = 0$). **Constraint (23)** ensures that the number of SAVs on virtual link $\left(i^{(t)}, j\right)$ cannot exceed $l^{max}_{i^{(t)}j}$, the upper bound of the number of

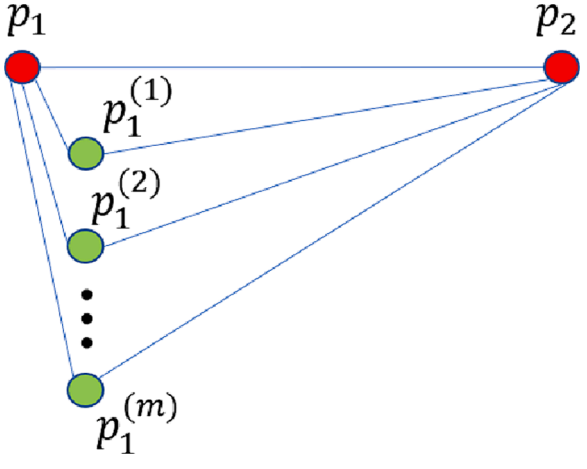


Fig. 5. Conversion of vehicle platooning problem.

SAVs in a single platoon.

$$\sum_{\eta\in\{1,2\}}\sum_{h\in H} x^{h,\eta}_{i^{(t)}jt} \geq l^{min}_{i^{(t)}j} \bullet x^{h,\eta}_{i^{(t)}jt} \qquad \forall h\in H, \forall \eta\in\{1,2\}, \forall t\in\Gamma, \forall\left(i^{(t)},j\right)\in\widehat{A} \tag{22}$$

$$\sum_{\eta\in\{1,2\}}\sum_{h\in H} x^{h,\eta}_{i^{(t)}jt} \leq l^{max}_{i^{(t)}j} \qquad \forall t\in\Gamma, \forall\left(i^{(t)},j\right)\in\widehat{A} \tag{23}$$

3.3.4. Domain of the variables

Constraints (24) to (27) specify variables $f_h, z^k_h, x^{h,\eta}_{ijt}$ and $y^{k,\eta}_h$ to be binary variables.

$$f_h\in\{0,1\} \qquad \forall h\in H \tag{24}$$

$$z^k_h\in\{0,1\} \qquad \forall h\in H, \forall k\in K \tag{25}$$

$$x^{h,\eta}_{ijt}\in\{0,1\} \qquad \forall h\in H, \forall (i,j)\in A, \forall t\in\Gamma, \forall\eta\in\{1,2\} \tag{26}$$

$$y^{k,\eta}_h\in\{0,1\} \qquad \forall h\in H, \forall k\in K, \forall\eta\in\{1,2\} \tag{27}$$

3.4. Travel time computing method

The SAV travel time on link $(i,j)\in A\backslash\check{A}$ consists of two parts: one is the travel time on the link segment (i.e., τ_{ij}), and the other is the dwelling time (or total delays) at the downstream intersection j (i.e., D^0_{ij}), which is generated when SAVs arrive during red time. Besides, the link $(i,j)\in\check{A}$ is merely used to connect virtual node $i\in\widehat{N}$ to the actual node on the network, thereby having an SAV travel time equal to 0. Therefore, the total travel time (i.e., TT_{ij}) is calculated by **Equation (28)**:

$$TT_{ij}=\begin{cases}\tau_{ij}+D^0_{ij} & \forall(i,j)\in A\backslash\check{A}\\ 0 & \forall(i,j)\in\check{A}\end{cases} \tag{28}$$

The travel time of each SAV on a particular link depends on whether or not it has received priority services. We assume that the penetration rate of SAVs is much lower than conventional vehicles, and thus the number of SAVs is unlikely to be sufficient to significantly impact the travel time of a specific link segment. The travel time is primarily determined by the traffic flow of conventional vehicles, which can be predicted or observed in advance. If an SAV does not receive priority services while traversing link (i,j) (i.e., link (i,j) does not belong to set $\widehat{A}$ or set $\check{A}$), its travel time on the link segment can be calculated by **Equation (29)** which is the BPR function:

$$\tau_{ij}=t^0_{ij}\left[1+\alpha_{ij}\left(v_{ij}/c_{ij}\right)^{\beta_{ij}}\right] \qquad \forall(i,j)\in A\backslash\check{A}\backslash\widehat{A} \tag{29}$$

where t^0_{ij} is the free flow travel time on link (i,j), c_{ij} is the capacity of a single lane of link (i,j), v_{ij} is the traffic flow of conventional vehicles on link (i,j), α_{ij}, β_{ij} are positive parameters.

We denote C_j as the signal period and g_j as the duration of the green phase in the priority signal cycle at intersection j. When SAVs arrive during the green phase, their delays are 0. However, for those arriving before or after the green phase, their delays are computed by contrasting their arrival time with the commencement of the subsequent green phase. Consequently, the expected total delays, represented as D^0_{ij}, can be derived by **Equation (30)** (see Appendix A for the proof):

$$D^0_{ij}=\frac{\left(C_j-g_j\right)^2}{2C_j} \qquad \forall(i,j)\in A\backslash\check{A}\backslash\widehat{A} \tag{30}$$

If SAVs form a platoon and travel on the virtual link $(i,j)\in\widehat{A}$, they have exclusive right of way thus having a link travel time equal to the free flow travel time t^0_{ij}, as is shown in **Equation (31)**:

$$\tau_{ij}=t^0_{ij} \qquad \forall(i,j)\in\widehat{A} \tag{31}$$

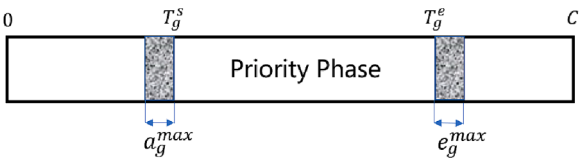


Fig. 6. SAV priority control.

The dwelling time at intersections is also reduced, with two signal priority strategies: green advance (i.e., advance the green time of priority phase) or green extension (i.e., prolong the green time of priority phase), taken to prioritize SAVs. We set a_g^{max} to restrict the maximum time of green advance and e_g^{max} to restrict the maximum time of green extension. The constitution of the SAV priority phase is shown in Fig. 6.

If an SAV receives priority while traversing link (i,j), its expected total delays can be calculated by **Equation (32)** (see Appendix B for the proof):

$$D_{ij}^0 = \frac{\left(C_j - g_j - a_{g_j}^{max} - e_{g_j}^{max}\right)^2}{2C_j} \qquad \forall (i,j) \in \widehat{A} \tag{32}$$

Therefore, the travel time TT_{ij} on any link $(i,j) \in A$ is a parameter with constant value in our model, and it remains unchanged regardless of how the SAV traffic flow is allocated on the road network.

3.5. Objective function

The primary goal of the scheduling method proposed in our study is to reduce the total travel time of SAVs. However, in the context of a relatively low SAV market penetration, it becomes imperative to account for the overall performance of the road network and minimize the negative impacts on traditional vehicles. As a result, the objective function employed in this model comprises three components: the total travel time of SAVs, the additional delays incurred by conventional vehicles due to SAV priority, and the penalty associated with the rejection of requests.

The additional delays stemming from SAV priority can be categorized into two distinct components. One source of the delay arises from the link section $\left(i^{(t)},j\right)$ where dedicated lanes are activated to offer priority to SAVs. In these instances, the reduction in the number of lanes available to conventional vehicles leads to an increase in travel time on the link section. We denote the traffic flow of conventional vehicles on link $(i,j) \in A$ at timestep t as $v_{i^{(t)}j}$. The number of conventional vehicles affected within this time step is $v_{i^{(t)}j}\Delta t$. Then, the delays of conventional vehicles arising from virtual link sections $\left(i^{(t)},j\right) \in \widehat{A}$ are calculated by **Equation (33)**:

$$D_{i^{(t)}j}^1 = B_{i^{(t)}j} v_{i^{(t)}j} \Delta t t_{ij}^0 \alpha_{ij} \left[\left(\frac{v_{i^{(t)}j}}{c_{ij}\left(\delta_{ij}-1\right)} \right)^{\beta_{ij}} - \left(\frac{v_{i^{(t)}j}}{c_{ij}\delta_{ij}} \right)^{\beta_{ij}} \right] \qquad \forall t \in \Gamma, \forall \left(i^{(t)},j\right) \in \widehat{A} \tag{33}$$

where δ_{ij} is the number of lanes on link (i,j), $B_{i^{(t)}j}$ is a binary variable that determines whether virtual link $\left(i^{(t)},j\right)$ is visited where $B_{i^{(t)}j} = 1$, if virtual link $\left(i^{(t)},j\right)$ is visited by SAVs and $B_{i^{(t)}j} = 0$, otherwise. The value of $B_{i^{(t)}j}$ is determined by imposing **Constraints (34) to (36)**:

$$B_{i^{(t)}j} \le \sum\nolimits_{h\in H} \sum\nolimits_{\eta\in\{1,2\}} x_{i^{(t)}jt}^{h,\eta} \qquad \forall t \in \Gamma, \forall \left(i^{(t)},j\right) \in \widehat{A} \tag{34}$$

$$B_{i^{(t)}j} \ge x_{i^{(t)}jt}^{h\eta} \qquad \forall h \in H, \forall \eta \in \{1,2\}, \forall t \in \Gamma, \forall \left(i^{(t)},j\right) \in \widehat{A} \tag{35}$$

$$B_{i^{(t)}j} \in \{0,1\} \qquad \forall t \in \Gamma, \forall \left(i^{(t)},j\right) \in \widehat{A} \tag{36}$$

Another source of delays arises from the downstream intersection of the link $\left(i^{(t)},j\right)$. As the phase providing priority for SAVs extends, the time allocated to other phases at the intersection correspondingly diminishes, leading to increased delays for conventional vehicles. We employ $v_{i^{(t)}j}^1$ and $v_{i^{(t)}j}^2$ to represent the traffic flow of conventional vehicles, which drive through the downstream intersection of link $\left(i^{(t)},j\right)$ with increased delays, respectively caused by green advance and green extension. The increased delays for conventional vehicles, denoted as $D_{i^{(t)}j}^2$, are then determined by **Equation (37)** (see Appendix C for the proof):

$$D_{i^{(t)}j}^2 = B_{i^{(t)}j} \bullet \frac{v_{i^{(t)}j}^1 \left[a_{g_j}^{max3} + 3\left(2C_j + g_j\right) a_{g_j}^{max2} \right] + v_{i^{(t)}j}^2 \left[3\left(5C_j + g_j\right) e_{g_j}^{max2} - e_{g_j}^{max3} \right]}{18C_j} \qquad \forall t \in \Gamma, \forall \left(i^{(t)},j\right) \in \widehat{A} \tag{37}$$

Green advance strategy leads to relatively long delays for a conventional vehicle that arrives precisely within the advanced time segment, which may yield a very unfair system for some individual conventional vehicles. Therefore, we formulate the following proposition to evaluate this impact.

Proposition 1. Suppose an event that a conventional vehicle is affected by the green advance strategy during a signal period at the upstream intersection of link $\left(i^{(t)},j\right) \in \widehat{A}$. $I_{i^{(t)}j}^a$ is a binary variable, which equals 1 if this event occurs, and equals 0, otherwise. $J_{i^{(t)}j}^a$ is a continuous variable, which represents the resulting delays from this event. Then:

The probability of this event to occur is:

$$P\left(I^a_{i^{(t)},j}=1\right)=\int_{T^s_{g_j}-a^{max}_{g_j}}^{T^s_{g_j}}\frac{1}{C_j}\bullet\frac{T^s_{g_j}-x}{C_j}dx=\frac{a^{max2}_{g_j}}{2C_j{}^2}$$

The expected delay that results from the event is:

$$E\left(J^a_{i^{(t)},j}|I^a_{i^{(t)},j}=1\right)=\int_{T^s_{g_j}-a^{max}_{g_j}}^{T^s_{g_j}}\frac{1}{a^{max}_{g_j}}\left[\frac{1}{3}\left(2C_j+g_j\right)+\frac{1}{6}\left(T^s_{g_j}-x\right)\right]dx=\frac{1}{3}\left(2C_j+g_j\right)+\frac{1}{12}a^{max}_{g_j}$$

The maximum delay that results from the event is:

$$max\left\{J^a_{i^{(t)},j}|I^a_{i^{(t)},j}=1\right\}=\frac{1}{3}\sum_p\left(C_j+a^{max}_{g_j}-g^{a,p}_j\right)=a^{max}_{g_j}+\frac{2C_j+T^s_{g_j}}{3}$$

The event and the three indices proposed in Proposition 1 should be taken into account while determining the scheme of the signal priority control for SAV platoons. In summary, the total delays that SAV priority services cause to other conventional vehicles are calculated by **Equation (38)**:

$$D_{total}=\sum\nolimits_{t\in\Gamma}\sum\nolimits_{\left(i^{(t)},j\right)\in\hat{A}}\left(D^1_{i^{(t)}j}+D^2_{i^{(t)}j}\right) \tag{38}$$

A positive parameter γ is introduced to represent the weight of the total delays caused by SAV priority services to conventional vehicles in the objective function, and β is set as a penalty for the rejection of each request. Finally, the model formulated in our study is given as follows where equation (39) is the objective function:

$$min\sum\nolimits_{h\in H}\sum\nolimits_{\eta\in\{1,2\}}\sum\nolimits_{t\in\Gamma}\sum\nolimits_{(i,j)\in A}TT_{ij}x^{h,\eta}_{ijt}+\gamma\sum\nolimits_{t\in\Gamma}\sum\nolimits_{\left(i^{(t)},j\right)\in\hat{A}}\left(D^1_{i^{(t)}j}+D^2_{i^{(t)}j}\right)+\beta\sum\nolimits_{k\in K}\chi_k \tag{39}$$

s.t Constraints (1) − (27) and (33) − (38).

4. Solution method

The real-time SAV scheduling and routing problem is a challenging NP-hard problem characterized by a multitude of variables and constraints incorporated into the model. Commercial solvers often struggle to effectively manage these large-scale problems, leading to difficulties in obtaining high-quality solutions within practical computational time. We propose a two-stage heuristic algorithm to solve the problem efficiently. The core idea of the two-stage algorithm is that we first allocate requests to vehicles and decide the pick-up and drop-off sequence while considering the potential of platooning effects, then endeavor to insert links with priority services into the original paths based on platooning constraints.

4.1. First stage: Vehicle-to-request matching method considering platooning effect

In the first stage, in order to determine the vehicle-to-request matching scheme as well as the pick-up and drop-off sequence, we utilize a reactive anytime optimal method for assigning passenger requests to a fleet of vehicles of varying capacity (Alonso-Mora et al.,2017). The algorithm mainly consists of the following steps:

0) Path feasibility check method

For vehicle h, the optimal travel path σ_h that minimizes the objective function subject to the pick-up and drop-off time constraints is given by the following function:

$$travel(h,R_h)$$

which returns “invalid” if no feasible path exists.

1) Pairwise request-vehicle (RV) graph

In the first step, we try to compute which requests can be pairwise combined and which vehicles can serve which requests individually.

For any two requests k_1 and k_2, let an SAV start from the origin of request k_1 or k_2, and then 4 paths could be generated, with varied sequence selected to pick up and drop off the two requests. If one of those paths could fulfill both requests while satisfying the path feasibility check, they are potentially combined and $e(k_1,k_2)$ is added to the RV-graph.

Likewise, for a vehicle h and a request k, if there is a feasible path that starts at the origin of the vehicle and satisfies the request, the match $e(h,k)$ is added to the RV-graph.

2) Request-trip-vehicle (RTV) graph

The second step is to construct trips (groups of requests) that can be served by a vehicle. For trip T and vehicle h, if the vehicle could pick up and drop off all the requests of the trip in a certain order, while satisfying the path feasibility check, we take that they are potentially matched and add $e(T,h)$ to the RTV-graph. As the trip size increases, it becomes necessary to check trip T only if there is a vehicle h for which every one of its sub-trips T' has an edge $e(T',h)$ in the RTV-graph.

3) Optimal assignment

Based on the RTV-graph, each match of trip and vehicle is comprised of a pick-up itinerary and a drop-off itinerary, which can be further divided into several point-to-point sub-itineraries by pick-up points or drop-off points. When considering the *nth* sub-itinerary of vehicle h, we first evaluate the minimum travel time from O_n (i.e., the origin of the *nth* sub-itinerary) to D_n (i.e., the destination of the *nth* sub-itinerary), denoted as $\tau_t\left(s_h^{O_n}, s_h^{D_n}\right)$.

Then, we insert a "link combination" into this sub-itinerary, which can consist of a single link or a succession of connected links with SAV priority services. We suppose that the upstream and downstream nodes of the link combination are p_{up} and p_{down}, respectively. As shown in Fig. 7, vehicle h departs from O_n, then visits p_{up} and p_{down}, and finally arrives at D_n. The minimum travel time of the whole process is recorded as $\tau_t\left(s_h^{O_n}, s_p^{up}\right) + \tau_t\left(s_p^{up}, s_p^{down}\right) + \tau_t\left(s_p^{down}, s_h^{D_n}\right)$,

where $\tau_t\left(s_p^{up}, s_p^{down}\right)$ represents the total travel time from p_{up} to p_{down} on link combination p, considering priority effect, $\tau_t\left(s_h^{O_n}, s_p^{up}\right)$, and $\tau_t\left(s_p^{down}, s_h^{D_n}\right)$ denotes the minimum travel time from O_n to p_{up} and from p_{down} to D_n, respectively.

By comparing the values of the above two terms, we can determine whether vehicle h reduces its travel time during the *nth* sub-itinerary through priority services on the link combination p. If there is a reduction in travel time, it suggests that this link combination exhibits a close spatial correlation with the sub-itinerary. The greater the reduction in travel time, the higher the likelihood that vehicle h can participate in a platoon on the link combination p during the *nth* sub-itinerary. Conversely, if the travel time is not reduced, this link combination is considered to have a distant spatial correlation with the sub-itinerary, indicating little potential for platooning. Based on this theory, the individual cost of each trip-vehicle match presented in the RTV-graph is established as follows:

$$c_{hT}^i = \sum\nolimits_{n\in\{1,\cdots,N_{hT}\}} \left[\tau_t\left(s_{h,i}^{O_n}, s_{h,i}^{D_n}\right) - \lambda\left(max\left\{\max_{p\in P}\left\{\tau_t\left(s_{h,i}^{O_n}, s_p^{up}\right) + \tau_t\left(s_p^{up}, s_p^{down}\right) + \tau_t\left(s_p^{down}, s_{h,i}^{D_n}\right) - \tau_t\left(s_{h,i}^{O_n}, s_{h,i}^{D_n}\right)\right\}, 0\right\}\right)\right] \tag{40}$$

$$c_{hT} = \min_{i\in R}\left\{c_{hT}^i\right\} \tag{41}$$

where N_{hT} denotes the total number of sub-itineraries of vehicle h in trip T, R is a set of pick-up and drop-off sequences of vehicle h in trip T, which should be confirmed feasible by the path feasibility check. In **Equation (40)**, for each feasible pick-up and drop-off sequence i, in which vehicle h executes trip T, we calculate an evaluation index c_{hT}^i. The index comprises two components: first, the minimum travel time from the origin to the destination without the activation of any priority services, and second, the theoretically maximum reduction in travel time gained from priority services during the whole trip. We use parameter λ to represent the discount of the benefits potentially obtained from priority services. In **Equation (41)**, we determine the minimum cost for vehicle h executing trip T among various pick-up and drop-off sequences.

Based on the proposed individual cost function, we try to find out the optimal assignment of trips to vehicles. First, a greedy solution is computed by assigning trips to vehicles iteratively starting from the matches with the lowest costs and largest number of requests in the trip. It serves as the initial point for the trip-vehicle assignment optimization.

Then, we execute a trip-vehicle assignment optimization as shown in **Function (42)** and **Constraints (43) to (44)**:

$$min\sum\nolimits_{h,T\in RTV} c_{hT}\epsilon_{hT} + \beta\sum\nolimits_{k\in K}\chi_k \tag{42}$$

s.t

$$\sum\nolimits_{T\in\zeta_h}\epsilon_{hT} \le 1 \qquad \forall h \in H \tag{43}$$

$$\sum\nolimits_{T\in\zeta_k}\sum\nolimits_{h\in\zeta_T}\epsilon_{hT} = 1 - \chi_k \qquad \forall k \in K \tag{44}$$

where ϵ_{hT} is a binary variable that decides whether vehicle h is assigned to trip T, ζ_h denotes the set of trips T for which an edge $e(T,h)$ exists in the RTV-graph, ζ_k denotes the set of trips with request k included, and ζ_T denotes the set of vehicles h for which an edge $e(T,h)$ exists in the RTV-graph. **Constraint (43)** restricts each vehicle to be assigned to at most a single trip, and **Constraint (44)** specifies that each request should be assigned to a vehicle, otherwise it has to be rejected.

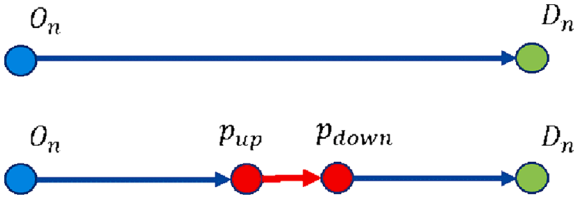


Fig. 7. Insertion of link combination into sub-itinerary.

4.2. Second stage: Vehicle routing and platooning method

In the second stage, we construct the spatiotemporal correlation for SAV platooning with a customized genetic algorithm. Firstly, we establish initial routes for the dispatched SAVs based on the sequence of pick-up and drop-off nodes from multiple matched requests, and the links with priority services are inserted into the pick-up or drop-off itineraries according to the spatial correlations assessed in Step 1. Then we execute a crossover operation involving one of the links found in the original routes and the virtual links derived based on platooning time. This operation generates numerous new routes that enable platooning within varied time intervals. Mutation occurs ahead of the crossover operation with a goal to promote the opportunity of platooning by randomly adding link segments with priority services to the route from which they have been excluded. A checkout process is conducted to ensure that the newly generated routes comply with arrival time and vehicle number constraints, which is essential for meeting the spatiotemporal requirement of platooning. Subsequently, we collect all the feasible routes to form the offspring population. The top three routes in terms of fitness are selected to join the parent population in the subsequent iteration. This sequence of procedures is repeated for multiple iterations until the top three fitness values have remained unchanged for several iterations or the iteration count reaches a preset maximum.

The chromosome structure is illustrated in Fig. 8. Chromosome 0 outlines the spatial feasibility of SAVs to participate in platoons at each link with priority services, while the remaining chromosomes are constructed to describe the platooning features of SAVs within varied time steps. Each chromosome is further subdivided into distinct segments represented by different colors, where each segment corresponds to a specific link intended for platooning. For example, the red segment signifies that vehicle 1 to *N* could potentially form a platoon at link 1, while the yellow segment indicates the possibility of platooning at link 2, etc. The configuration of segments on these chromosomes, combined with the temporal dimension, represents a unique link at a specific time step, essentially embodying the concept of a virtual link.

In each segment, every SAV is assigned two genes to modify its pick-up and drop-off itineraries. Each itinerary can be further subdivided into several sub-itineraries based on pick-up and drop-off points (e.g., 2^+ and 2^- denote the pick-up and drop-off points for Request 2). These sub-itineraries are organized in a predetermined sequence established in Step 1. Since an SAV cannot participate in platoons on a given link segment more than once within a single itinerary, we coded each gene with either "0″ or a positive number "*n*", where "0" signifies that the virtual link is not visited during any sub-itineraries, while "*n*" indicates that the SAV visits the virtual link during the *nth* sub-itinerary.

The specific steps of the vehicle routing and platooning method are introduced as follows:

0) Construction of initial solution

The construction of the initial solution involves determining the initial value of each gene on chromosome 0. This is accomplished by assessing the spatial correlation in Step 1, wherein we identify whether each vehicle can decrease the total travel time by visiting each link during its pick-up and drop-off itinerary respectively.

1) Mutation

There are instances where we observe that the number of SAVs on a particular link is slightly below the lower threshold required to form a platoon. In such cases, we can reintroduce the vehicles that were initially excluded during the initial solution construction process to create a new platoon, which may further reduce the total travel time of the SAVs. Therefore, we implement mutation by changing the code of the specific genes on chromosome 0 from "0″ to "1″, as shown in Fig. 9.

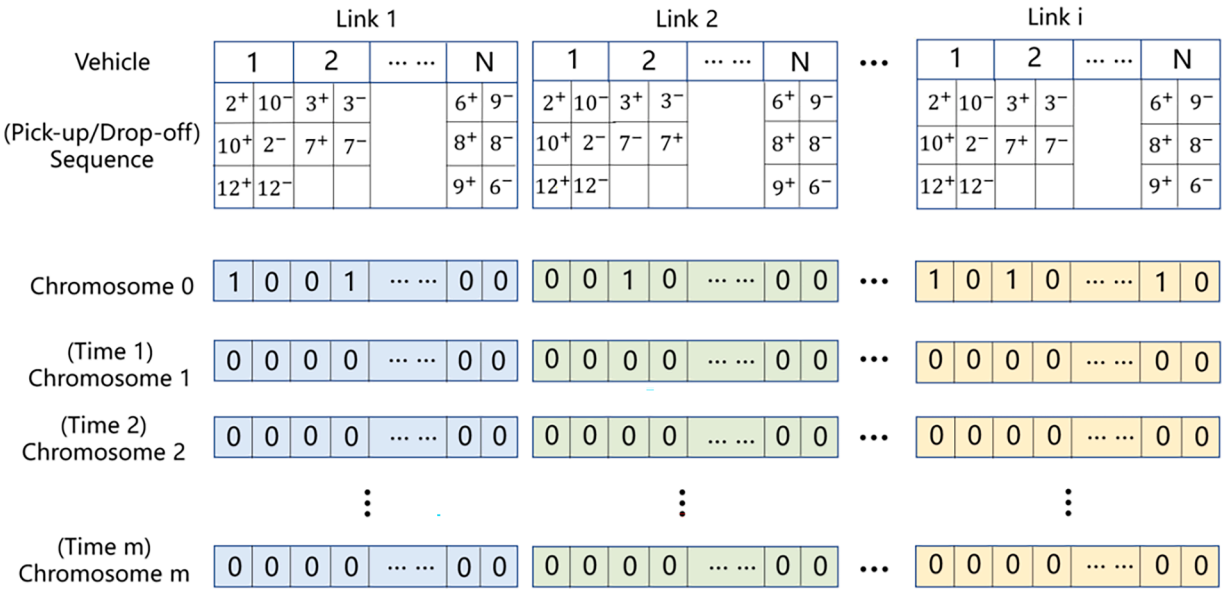


Fig. 8. Diagram for the constitution of initial chromosomes.

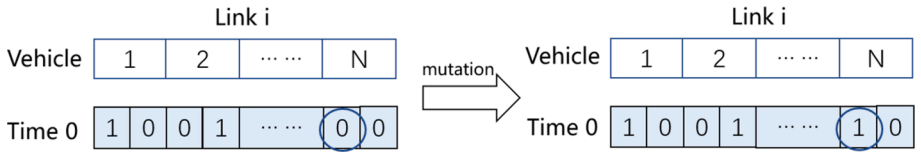


Fig. 9. Diagram for mutation operation.

The mutation probability of the genes varies among segments. It starts at zero in a segment that has not been selected for crossover. Once a segment is initially selected for crossover, the mutation probability of the genes within the segment is updated at the beginning of the subsequent iteration according to the function $P_i^w = min\left\{1, P_i^{w-1}(1-\xi) + \xi\frac{N_i}{(\beta_i+1)N_{min}}\right\}$, where i represents the segment number, w denotes the number of iteration, P_i^w indicates the mutation probability of segment i in iteration w, N_i represents the number of genes which are coded as "1" in segment i of chromosome 0, N_{min} indicates the minimum number of SAVs required for platooning, β_i indicates the number of iterations in which segment i has been selected for crossover, ξ is a positive parameter, ranging from 0 to 1.

2) Crossover

We conduct crossover by swapping a segment of chromosome 0 with the allele segment of another chromosome. In this way, some of the SAVs are thrown into a certain time step, making it convenient to check whether they can reach the upstream nodes of the target link and form a platoon. The crossover operation is shown in Fig. 10.

To enhance the search efficiency, preference should be given to the segments where a significant number of genes are coded as "1" on chromosome 0, because a substantial number of SAVs are likely to traverse the links represented by such segments during their trips, exhibiting higher possibility of forming platoons. Based on this idea, we introduce a parameter referred to as crossover probability, which quantifies the likelihood of each segment being selected for crossover. This crossover probability is updated at the beginning of each iteration according to the function $P_i = \frac{e^{\frac{N_i}{\beta_i+1}}}{\sum_j e^{\frac{N_j}{\beta_j+1}}}$, where i denotes the segment number, P_i indicates the probability that segment i is selected for crossover, N_i represents the number of genes which are coded as "1" in segment i of chromosome 0, β_i indicates the number of iterations in which segment i has been selected for crossover.

3) Time-dependent feasibility check

After the crossover operation, we perform a time-dependent feasibility check to determine whether it is viable to put an SAV within a virtual link for platooning. The departure time for each SAV is constrained within a range $[t_o^{min}, t_o^{max}]$, which is updated at the beginning of each iteration. Initially, the feasible departure time for each SAV spans the entire scheduling period: $[(num-1) \bullet T, num \bullet T]$. As the iterations progress, certain vehicles are designated for platooning on specific virtual links. This requires adjusting the feasible range of departure times to comply with the time window constraints of the corresponding virtual link segments: $[t_p^{min} - \tau_t(s_h^O, s_p^{up}), t_p^{max} - \tau_t(s_h^O, s_p^{up})]$, where t_p^{min} and t_p^{max} represent the lower and upper bounds of the time window constraint of virtual link p, respectively.

Each vehicle that has been thrown into a new virtual link for platooning during the crossover operation must undergo a departure time derivation. This entails considering all potential sub-itineraries for insertion and calculating their respective departure times. If the calculated departure time for all values fall outside the feasible range of time windows, signifying infeasibility regarding the time window constraint, we label the corresponding genes as infeasible (i.e., labeled gray in Fig. 12) and exclude them from the platooning process in subsequent steps. For vehicles with feasible departure times, a further check on their arrival time at the pick-up and drop-off nodes should be conducted as well.

The sub-itinerary for insertion is determined by selecting the earliest departure time that has been confirmed feasible, with the index of the sub-itinerary utilized to replace the initial "1" code on the corresponding gene. The operation after the time-dependent feasibility check is illustrated in Fig. 11.

4) Vehicle-number-based feasibility check

After excluding the genes due to time infeasibility, we count the number of genes that are coded as positive numbers within every segment. If the number is less than the minimum number of vehicles required to form a platoon, all the genes coded as positive numbers within the segment will be labeled as infeasible. Conversely, if the number of nodes exceeds the maximum number of vehicles

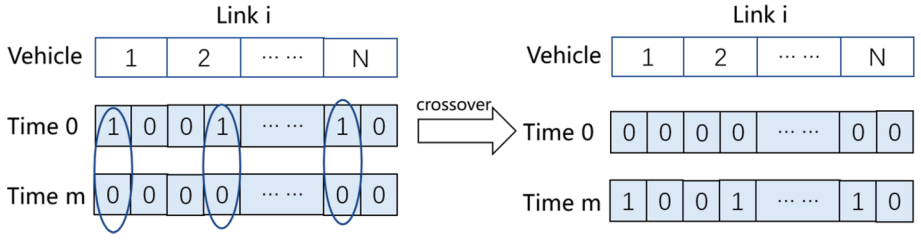


Fig. 10. Diagram for crossover operation.

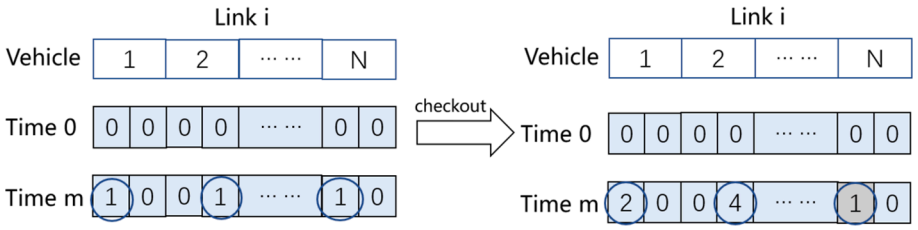


Fig. 11. Diagram for the operation after time-dependent feasibility check.

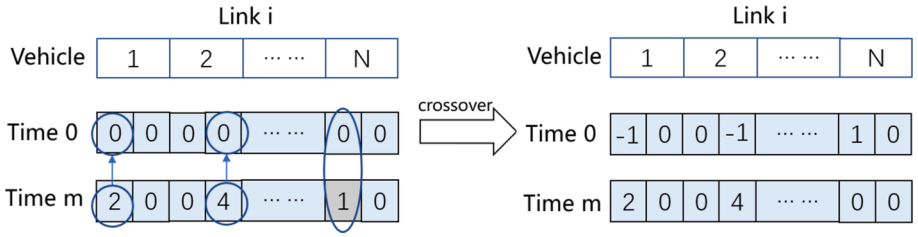


Fig. 12. Diagram for the operation after vehicle-number-based feasibility check.

allowed in a single platoon, some of them must be excluded. We calculate the increase in travel time when each vehicle is excluded from the platoon. The vehicles which lead to the least amount of travel time increase are excluded until the platoon size constraint is fulfilled.

With the vehicle-number-based feasibility checked, another crossover operation is conducted on several pairs of segments, as shown in Fig. 12.

For the genes that have been labeled infeasible, we revert their code back to "1″ and swap them with their allele genes on chromosome 0. As for the nodes that are coded as positive numbers without being labeled as infeasible, we update their allele genes on chromosome 0 from "0″ to "-1″. No further crossover operations are conducted on the nodes coded as "-1″ in the subsequent iterations.

5) Fitness evaluation and selection

In this step, for each feasible offspring obtained through mutation, crossover, and checkout operations, we compute the fitness score based on the objective function. From the offspring population, the individuals with the highest three fitness scores are picked to be part of the parent population in the next iteration.

6) Iteration stopping criteria

The goal of the algorithm is to find the offspring individual with the highest fitness score. Following each iteration, it is vital to update the best fitness and best offspring individual based on the evaluation results of the newly generated offspring. Subsequently, a decision is made on whether to initiate a new iteration in pursuit of an improved solution. The iteration-stopping criteria entail the condition that the top three fitness values have remained unchanged for φ consecutive iterations or the iteration count reaches an

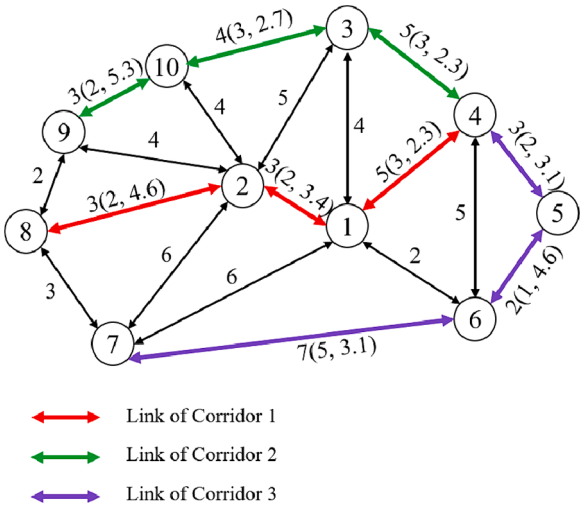


Fig. 13. Synthetic network setting.

upper bound Ω. The optimal solution for SAV scheduling and platooning model is ultimately determined based on the best offspring individual.

5. Evaluations

5.1. Synthetic road network case study

In this section, we preliminarily validate the SAV scheduling, routing, and platooning method using a synthetic network (Wan and Lo., 2009), as displayed in Fig. 13. In this network, we identify three corridors, highlighted in various colors in the figure, which could offer priority services for SAV platoons. The travel times for each link are also indicated in the figure. Specifically, the travel times for SAV platoons and the increased total travel time for conventional vehicles, associated with priority services along the links of these corridors, are denoted within brackets. For example, on link 9–10, the label 3(2, 5.3) indicates that under normal circumstances, an SAV requires 3 min to traverse the link. However, with priority services provided for an SAV platoon, this time is reduced to 2 min, at the expense of introducing increased total travel time of 5.3 min for conventional vehicles. Notably, within the proposed model, the travel times for SAV platoons and the increased total travel time for conventional vehicles are derived from a multitude of parameters tied to real-world link characteristics (such as the number of lanes, signal periods, flow of conventional vehicles, etc.). Thus, for the sake of simplicity in testing on the synthetic network, we contemplate employing constant values.

We choose 20 min as the scheduling period and further divide the scheduling period into 20 timesteps, each of which with one minute duration (i.e., $\Gamma = 30$, $\Delta t = 1$). The minimum number of SAVs required to form a platoon is set to 3 (i.e., $l_{ij}^{min} = 3 \ \forall (i,j) \in \widehat{A}$). The weight of the delays of conventional vehicles in the objective function is $\gamma = 0.5$. We designate an SAV depot at node 4 equipped with 20 SAVs, along with 10 randomly generated requests distributed across the network. Each vehicle can take a maximum of two passengers. Table 2 lists the details of the requests used in the case study.

We designed three cases, each involving different platooning corridors, in order to compare the results of the SAV scheduling, routing, and platooning on the network. Six performance metrics were used: the objective value, number of platoons, average number of SAVs per platoon, reduced total travel time of SAVs, increased total travel time of conventional vehicles, and computational time. We addressed these problems using three methods: Gurobi-only approach, a hybrid approach combining heuristic algorithm with Gurobi (utilizing Gurobi for solving the linear integer programming model based on the vehicle-to-request matching scheme obtained from Stage 1 of the heuristic), and the proposed two-stage heuristic algorithm. The methods were programmed in Windows using Python 3.8, and the processor used is an Intel Core i5-5500 CPU operating at 2.1 GHz, with 16 GB of RAM. The results are presented in Table 3.

The proposed two-stage heuristic algorithm exhibited superior performance in terms of computational efficiency. For Case 1, all three algorithms obtained the optimal solution. However, as the number of platooning corridors increased, Gurobi failed to produce a feasible solution for Cases 2 and 3 within 3 h. The computational times for the proposed two-stage heuristic algorithm were 3.6 s and 5.0 s for Cases 2 and 3, respectively, while those for the hybrid method combining heuristic and Gurobi were 32.1 s and 57.0 s, respectively. Notably, there was a significant increase in computational time observed in the hybrid method, ranging from 17.4 s to 57.0 s, corresponding to the increased number of platooning corridors.

Regarding solution quality, in Cases 1 and 3, both heuristic methods achieved the same globally optimal solution as the exact solver Gurobi. However, in Case 2, the objective function value obtained by the proposed two-stage heuristic method was slightly higher than that obtained by the hybrid method combining heuristic and Gurobi.

We conducted an analysis of the best solutions obtained in the three cases. In Case 1, the number of SAV platoons was minimal, at 1, whereas this number increased to 2 for Cases 2 and 3. The reduction on SAVs' total travel time ranged from 5 to 12 min, and the increased total travel time of conventional vehicles ranged from 3.4 to 8.0 min.

5.2. Real-world case description

The model is further applied to a real-world road network in Yangpu District, center of the city of Shanghai, in China. The district has a total area of 60.61 km^2 and a population of about 131,000. Fig. 14 shows the simplified road network of Yangpu District based on

Table 2
Request information.

Request index	Pickup node	Dropoff node	Pickup time window	Dropoff time window
1	4	8	[7, 17]	$[-\infty, 32]$
2	1	10	[7, 17]	$[-\infty, 29]$
3	5	2	[9, 19]	$[-\infty, 28]$
4	4	9	[9, 19]	$[-\infty, 32]$
5	6	8	[7, 17]	$[-\infty, 30]$
6	1	9	[6, 16]	$[-\infty, 28]$
7	3	8	[9, 19]	$[-\infty, 30]$
8	5	7	[6, 16]	$[-\infty, 30]$
9	4	6	[7, 17]	$[-\infty, 25]$
10	5	3	[8, 18]	$[-\infty, 31]$

Table 3
Performance comparison of different solution algorithms on three SAV platooning corridors cases.

Case		Case 1	Case 2	Case 3
Index of platooning corridors		1	1, 2	1, 2, 3
Objective function value	Gurobi	68.7	—	—
	Stage1 + Gurobi	68.7	68.0	63.9
	Stage1 + Stage2	68.7	68.7	63.9
Number of platoons	Gurobi	1	—	—
	Stage1 + Gurobi	1	2	2
	Stage1 + Stage2	1	1	2
Avg number of SAVs per platoon	Gurobi	3	—	—
	Stage1 + Gurobi	3	3	3
	Stage1 + Stage2	3	3	3
Reduced total travel time of SAVs (min)	Gurobi	5	—	—
	Stage1 + Gurobi	5	8	12
	Stage1 + Stage2	5	5	12
Increased total travel time of conventional vehicles (min)	Gurobi	3.4	—	—
	Stage1 + Gurobi	3.4	8.0	7.7
	Stage1 + Stage2	3.4	3.4	7.7
Computational Time (s)	Gurobi	7190.7	—	—
	Stage1 + Gurobi	17.4	32.1	57.0
	Stage1 + Stage2	1.9	3.6	5.0

Notice: The cases where a feasible solution could not be obtained within 3 h is marked as '—'.

OpenStreetMap.

The road network has 214 links and 67 nodes. Within this network, three major roads, Siping Road (18–26-35–36), Zhoujiazui Road (44–45-46–48), and Quyang Road (12–17-25–34), which are highlighted in red in the figure, may offer priority for SAVs. Siping Road and Zhoujiazui Road have two-way eight lanes, while Quyang Road has two-way six lanes. The number of lanes is sufficient to reserve a temporarily dedicated SAV lane and provide signal priority control for SAV platoons at the intersections. The capacity of each link is 1200veh per hour per direction per lane. The traffic flow of conventional vehicles on each link is 900veh per hour per direction per lane.

We choose *30 min* as the scheduling period and further divide the scheduling period into 30 timesteps, each of which with one minute duration (i.e., $\Gamma = 30, \Delta t = 1$). The minimum number of SAVs required to form a platoon is set to 3 (i.e., $l_{ij}^{min} = 3 \forall (i,j) \in \widehat{A}$). We also constrain the upper bound of SAV number in a single platoon to be 10 (i.e., $l_{ij}^{max} = 10 \forall (i,j) \in \widehat{A}$). The parameters of the BPR function are $\alpha = 1, \beta = 2$. The weight of conventional vehicle delays in the objective function is $\gamma = 0.5$. At every intersection connected to the three major roads that are capable of offering priority for SAVs, the signal period is 60 s and the ratio of green time to cycle is 0.4. Both the maximum time of green advance (i.e.,a_g^{max}) and green extension (i.e., e_g^{max}) are 6 s. Both the traffic flow of conventional vehicles affected by green advance and green extension at these intersections are 600veh per hour. Based on Proposition 1, we assess the impact of green advance strategy on individual conventional vehicle by calculating the proposed three indices: $P\left(I^a_{i^{(t)},j} = 1\right) = 0.005$, $E\left(J^a_{i^{(t)},j} | I^a_{i^{(t)},j} = 1\right) = 48.5$, $max\left\{J^a_{i^{(t)},j} | I^a_{i^{(t)},j} = 1\right\} = 50$. Therefore, a conventional vehicle bears a minimal probability of 0.5 % of being impacted by green advance at an intersection with SAV platoon, with delays not exceeding 50 s.

The initial positions of the SAVs are nodes: 18, 24, 36, and 39, with each node initially accommodating 50 vehicles. We suppose that each vehicle can satisfy a maximum of two requests per trip. We use the Shanghai Taxi data on 30 April 2021 consisting of order ID, order start and end times, and longitudes and latitudes of the trip origins and destinations. There are 25,431 requests in the study area that do a trip of over 3 km. These are taken as the passenger requests in the following experiments. The major OD pairs, specifically those with over 100 requests, are depicted in Fig. 15. Notably, we observe several instances of overlap between the OD pairs and the links that provide priority services.

5.3. Experiments results

Leveraging the road network, model parameters outlined in **Section 5.1**, we conduct several experiments to optimize the SAV scheduling and platooning strategy by employing the two-stage heuristic algorithm proposed in our study.

5.3.1. Demand variation

The number of requests is a key factor that affects the calculation speed. Therefore, in order to test the performance of the heuristic algorithm, we respectively record the calculation time spent on the two stages as the number of requests increases from 30 to 200 per half hour. The requests are randomly generated, drawing upon the distribution of the Shanghai Taxi data. We conducted 50 tests for each experiment with varied number of requests, and took the average values as the results of the experiment.

The results of the experiments, as illustrated in Fig. 16, reveal a gradual increase in the calculation time for both Step 1 and Step 2 as the number of requests grows from 30 to 200. Notably, the stopping criteria applied in our experiments require that the top three fitness values should remain constant for 5 consecutive iterations or until the iteration count reaches a maximum of 30 (i.e., φ=5, Ω =

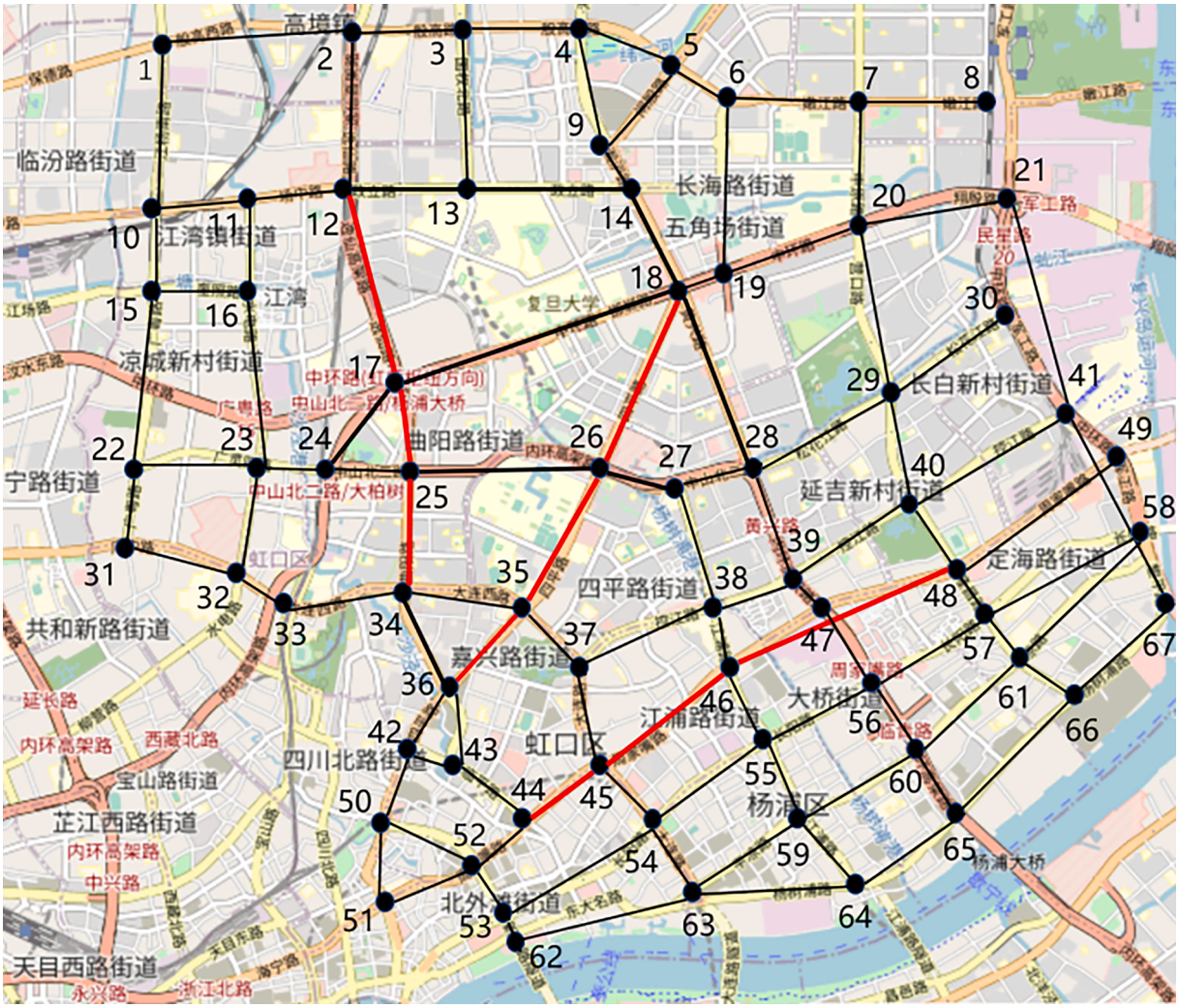


Fig. 14. Road network used in the case study.

30), since we observed that the optimal solution showed only slight improvement beyond 30 iterations. The total calculation time ranges from 10 s to 141 s. The total calculation time is relatively low, taking only 55 s when the number of requests stands at 150. However, both Step 1 and Step 2 experience a doubling of their durations when the number of requests increases from 150 to 200. The solutions along iterations in one of the experiments involving 200 requests are shown in Fig. 17, where the optimal solution is obtained through 23 iterations.

The numerical results of our experiments are presented in Table 4. As the number of requests increases, both the number of SAV platoons and the average number of SAVs per platoon show a gradual increment. In the case of 30 requests, there are only 1.80 platoons, with an average of 3 SAVs per platoon, which barely meets the minimum requirement for platooning. However, as the number of requests increases to 200, the number of platoons rises significantly to 10.20, and the average number of SAVs per platoon reaches 4.33.

We calculate the number of SAVs engaged in platooning on each link and determine their total count. Subsequently, we define the platooning percentage as the ratio of this cumulative count of platooning SAVs to the total number of dispatched SAVs. In this experiment, a noticeable rise in the percentage of platooning is observed with the number of requests, increasing from 33.19 % to 43.78 %.

Both the number of platoons and the average number of SAVs per platoon are closely related to the performance of the road network, while the former shows a more evident impact. The total travel time of SAVs exhibits a notable decline as the number of platoons grows, with the percentage of reduced travel time for the SAVs that formed platoons ranging from 11.74 % to 15.76 %. The total delays of conventional vehicles increase with the number of platoons as well, because each newly generated platoon contributes to an extra delay to the conventional vehicles. However, the total travel time increase of conventional vehicles is much smaller than the reduction observed in SAVs' total travel time, which shows that the overall performance of the road network could be improved by

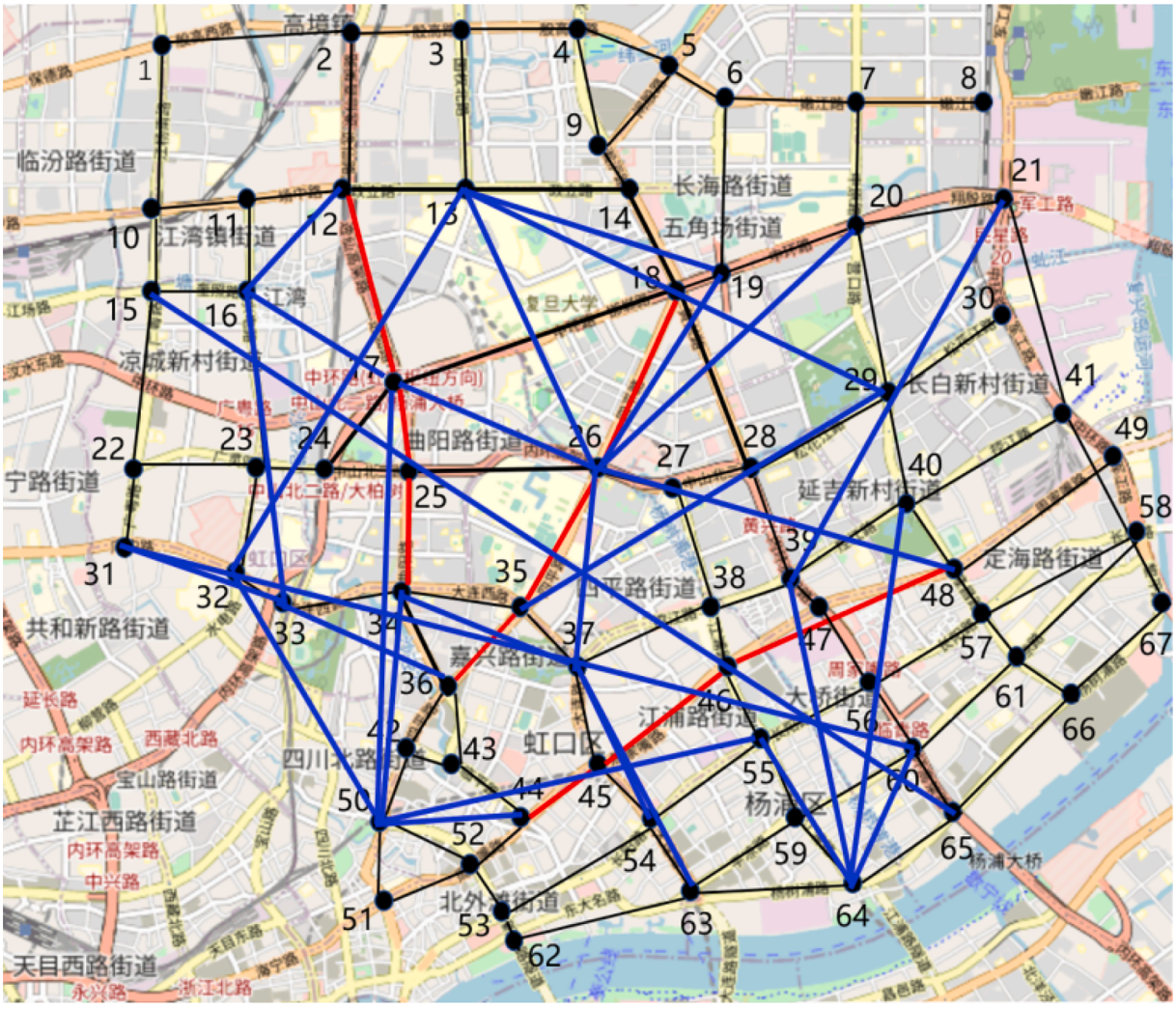


Fig. 15. Major OD pairs in the research area.

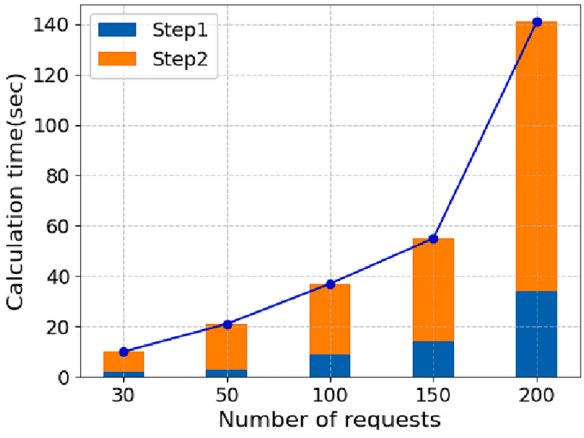


Fig. 16. Calculation time of SAV scheduling experiments.

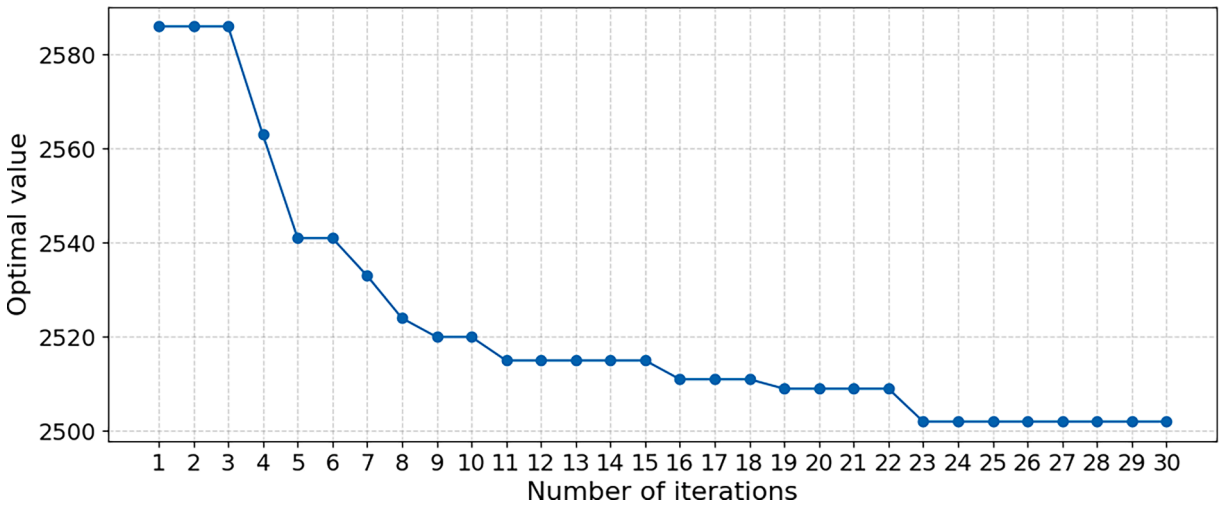


Fig. 17. Solution optimization along iterations.

Table 4
Optimization result of SAV scheduling experiments.

Number of requests	Number of platoons	Avg number of SAVs per platoon	Percentage of platooning	Percentage of reduced travel time of platooning SAVs	Reduced total travel time of SAVs (min)	Increased total travel time of conventional vehicles (min)
30	1.80	3.07	33.19 %	11.74 %	17.72	5.60
50	3.10	3.13	36.50 %	11.92 %	27.63	9.72
100	5.40	3.62	38.72 %	12.68 %	55.12	17.66
150	7.40	4.17	40.70 %	13.84 %	88.46	23.99
200	10.20	4.33	43.78 %	15.76 %	109.80	33.99

using the scheduling and platooning method proposed in our study, if sufficient numbers of SAVs are available to form the platoons.

5.3.2. Sensitivity analysis on vehicle capacity

We use the case of 150 requests as a benchmark and tested the system performance under various vehicle capacities, i.e., the number of passengers shared in a vehicle ride. The numerical results are presented in Table 5. The influence of vehicle capacity on both the usage of SAV fleet and the total travel time of SAVs is illustrated in Fig. 18.

When the capacity of each SAV is constrained to be one passenger (i.e., ridesharing is not allowed), the average number of platoons reaches 18.7 and the platooning percentage exceeds 60 %. However, with the vehicle capacity increased to two passengers, both the number of platoons and the platooning percentage decrease significantly. The reduction is even more pronounced when each SAV can carry up to three passengers. This result manifests the impact of carpooling on diminishing the platooning opportunities. This happens because the requests with similar routes are more likely to be allocated to one vehicle through carpooling rather than first matched to multiple vehicles and then platooned together.

The total travel time of the SAVs, as well as the time saved through priority service, diminishes as the number of shared riders increases. This indicates the predominant role of carpooling in SAV scheduling, which can optimize the total travel time by reducing vehicle usage. Meanwhile, the priority service for SAV platoons serves as an effective supplement to further enhance SAV operational efficiency. The effect of SAV platooning is particularly noticeable in scenarios with low-degree carpooling, or in the case in which passenger requests are more concentrated instead of sparsely distributed.

5.3.3. Sensitivity analysis on the weight of delays

The case with 150 requests and a capacity of 2 passengers per SAV is used as the base line case. We vary the weight of conventional vehicle delays in the objective function to balance the significance of SAVs' total travel time and the delays of conventional vehicles. The results are shown in Fig. 19.

Upon analysis of the optimal solution, it becomes evident that the reduced total travel time of SAVs as well as the increased delays of conventional vehicles remains relatively stable as the weight of delays increases from 0.3 to 0.5. However, when the weight of delays is raised from 0.5 to 1.0, the percentage of platooning drops from 40.70 % to 30.88 %, which leads to a substantial increase in total travel time of SAVs and a corresponding decrease in total delays of conventional vehicles. As the weight of delays increases from 1.0 to 2.0, both the reduction in total travel time of SAVs and the delays of conventional vehicles diminish. Particularly, when the weight of delays reaches 2.0, the reduced total travel time of SAVs and the increased delays of conventional vehicles become marginal, indicating

Table 5
Sensitivity test on vehicle capacity.

Vehicle capacity	Number of platoons	Avg number of SAVs per platoon	Percentage of platooning	Percentage of reduced travel time of platooning SAVs	Reduced total travel time of SAVs (min)	Increased total travel time of conventional vehicles (min)
1	18.70	4.86	60.71 %	16.06 %	242.50	56.59
2	7.40	4.17	40.70 %	13.84 %	88.46	23.99
3	4.30	3.28	27.66 %	12.86 %	40.35	13.50

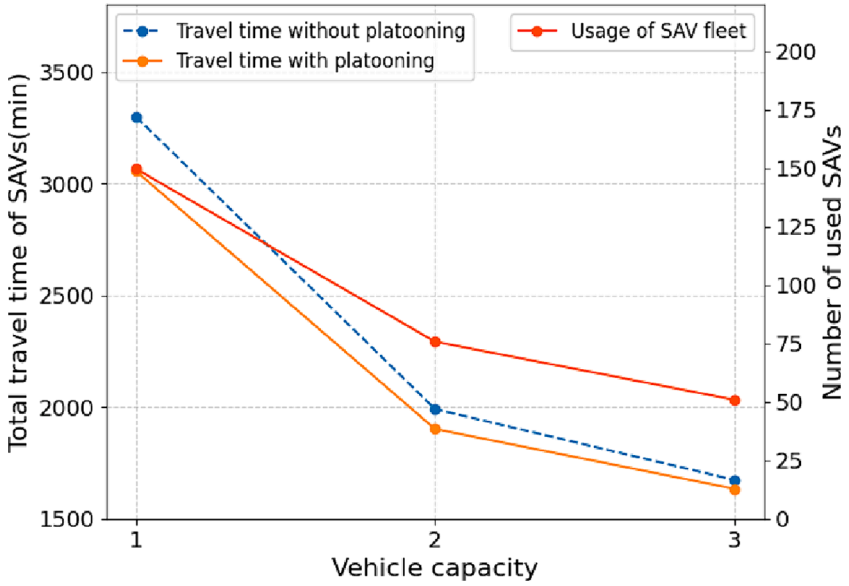
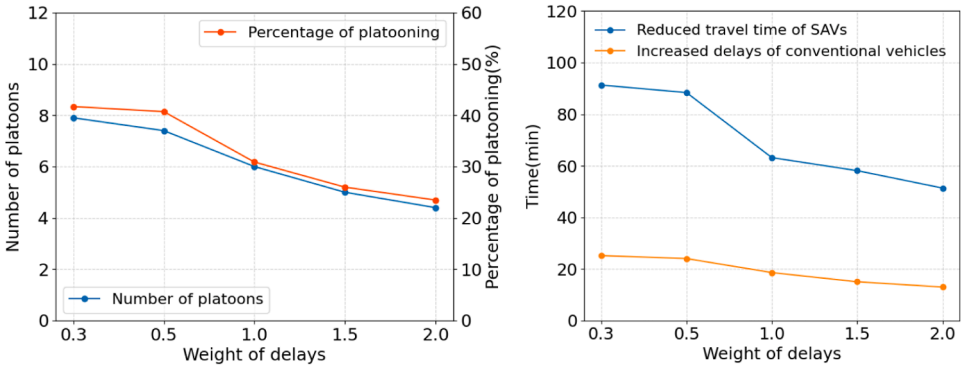


Fig. 18. Usage of SAV fleet and total travel time of SAVs.



(a) Number of platoons and percentage of platooning (b) SAVs' travel time and delays of conventional vehicles

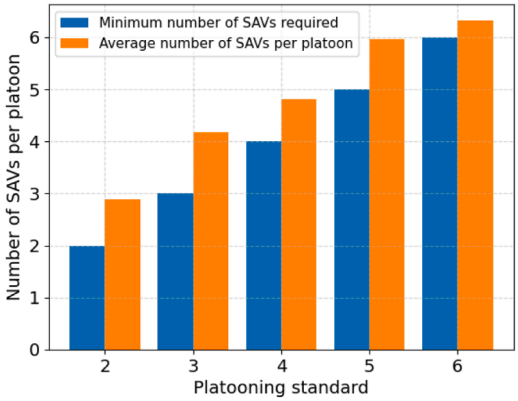
Fig. 19. Sensitivity test on weight of delays.

limited platooning effect. Therefore, the weight of delays is useful to modify the allocation of road resources on certain link segments, thus improving the optimization effect of the SAV scheduling and routing method proposed in our study.

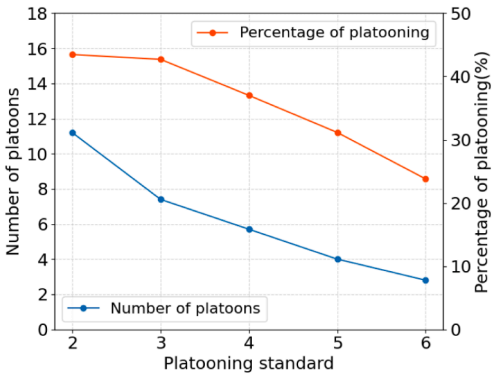
5.3.4. Sensitivity analysis on the minimum number of SAVs required for platooning

We vary the minimum number of SAVs to form a platoon from 2 to 6 based on the baseline case. We find the number of platoons, the average number of SAVs per platoon, and the percentage of platooning are sensitive to the minimum number of SAVs required for platooning, as shown in Fig. 20.

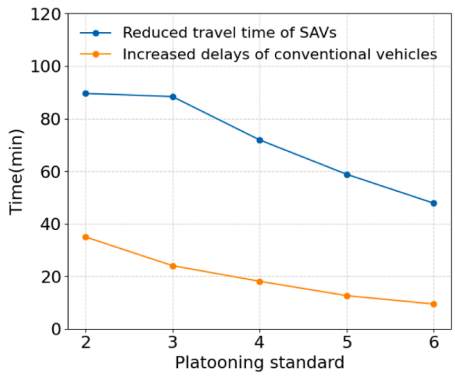
As the minimum number of SAVs required for platooning increases from 2 to 3, the percentage of platooning as well as the reduced total travel time of SAVs did not change much, which indicates that the number of SAVs that receive priority services through



(a) Average number of SAVs per platoon



(b) Number of platoons and percentage of platooning



(c) SAVs' travel time and delays of conventional vehicles

Fig. 20. Sensitivity test on the minimum number of SAVs required for platooning.

platooning is almost kept at the same level. However, with more SAVs required in a platoon, some SAVs are obliged to join another platoon, making the average number of SAVs per platoon grow to a value that narrowly exceeds the updated standard, and giving rise to a significant decline in the number of platoons. The delays of conventional vehicles are reduced while ensuring the efficiency of SAVs. Therefore, we suggest setting the number at 3.

As the minimum number increases from 3 to 6, there is a significant decrease in both the number of platoons and the percentage of platooning. Besides, we also observe a notable increase in the travel time of SAVs and a reduction in the delays of conventional vehicles with higher vehicle number standard. Therefore, it is concluded that adjusting the minimum number of SAVs required for platooning can mitigate the negative impact on conventional vehicles. However, an excessively high standard may considerably prevent the formation of SAV platoons, resulting in limited improvements in SAV operational efficiency.

6. Conclusion

With the expansion of metropolitan areas, as well as the increasing sophistication of intelligent transportation systems, SAVs are poised to become a crucial element of urban public transit in the future. In this study, we introduce a vehicle scheduling and routing method that allows SAVs to operate with priority on designated trunk roads in the form of platooning. An SAV scheduling, routing, and platooning model is developed to minimize the total travel time of SAVs and the delays experienced by conventional vehicles due to SAV priority, thereby optimizing the overall performance of the road network. The problem is formulated as a linear integer programming model which is very hard to solve due to the number of variables and constraints. In order to address the model complexity, we propose a two-stage heuristic algorithm. In the first stage, we allocate passenger requests to a fleet of SAVs, leveraging an evaluation index that manifests the compatibility of each vehicle-to-request combination based on spatial correlations, while considering the potential reduction in travel time gained from priority services. In the second stage, we establish the optimal vehicle routing solution, using a customized genetic algorithm that coordinates the paths of various SAVs in both time and space, thereby achieving the desired vehicle platooning effect.

A case study was presented to illustrate the effectiveness of the algorithm, the benefits of platooning for SAVs, and its impact on the overall performance of the road network. Our results demonstrate that the solution algorithm can compute the optimization result within one minute for scenarios involving 150 requests on a road network with 214 links and 67 nodes, indicating potential

applicability to real-time scheduling systems. In our experiments, the SAVs that formed platoons saw their total travel time reduced by up to 15.76 %, while keeping the total travel time increase of conventional vehicles significantly smaller than the improvements in SAV travel time. Upon examining the impact of vehicle capacity, we have determined that the priority service offered to SAV platoons can serve as a valuable complement to the carpooling approach. Furthermore, the weight of delays in the objective function and the minimum number of SAVs required for platooning emerge as effective control parameters for ensuring the overall performance of the road network while maintaining SAV operational efficiency.

The presented scheduling and routing approach for SAV platooning can be extended to other transportation modes, such as freight trucks and modular autonomous vehicles (MAVs). MAVs represent an innovative concept enabling dynamic docking and undocking of modules with vehicles of varying sizes during the journey (Chen et al., 2021; Pei et al., 2021; Tian et al., 2022). The scheduling, routing, and platooning method can be employed in the MAV transit system, where the operation of multiple modular units shares similarities with the SAV platooning process considered in this study. Additionally, the strategy of intermittent dedicated lanes can be further applied in traffic management, facilitating the operation of specific vehicle types while maintaining overall transportation system equity. However, the extensive management costs for intermittent activation of SAV dedicated lanes and signal priority control may impede widespread adoption of the proposed method. The utilization of such high-cost services may not always be profitable for SAV companies. Future research could address the SAV platooning problem on a larger scale. Additionally, the increased delays of conventional vehicles due to SAV priority services could be checked through a simulation-based method.

CRediT authorship contribution statement

Zhimian Wang: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. **Kun An:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Gonçalo Correia:** Writing – review & editing, Methodology. **Wanjing Ma:** Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is sponsored by the National Natural Science Foundation of China (72361137005, 72101186, 52325210). The authors declare no competing financial interests in this paper.

Appendix A

Among the SAVs that do not receive priority services, those arriving within green phase almost have no delays at the intersection, while the delays of those arriving before or after green phase are calculated by comparing their arrival time to the start time of the following green phase. Therefore, the expected total delays D_{ij}^0 given in **Equation (30)** can be derived as follows:

The probability density function of each SAV's arrival time within one signal period is:

$$f(x) = \begin{cases} \dfrac{1}{C_j} & 0 \le x \le C_j \\ 0 & else \end{cases}$$

The relationship between the dwelling time and the arrival time of each SAV is:

$$D(x) = \begin{cases} T_{g_j}^s - x & 0 \le x < T_{g_j}^s \\ 0 & T_{g_j}^s \le x \le T_{g_j}^e \\ C_j + T_{g_j}^s - x & T_{g_j}^e < x \le C_j \end{cases}$$

The expected total delays D_{ij}^0 are calculated as follows:

$$D_{ij}^0 = \int_0^{C_j} f(x)D(x)dx = \int_0^{T_{g_j}^s} \frac{1}{C_j}\left(T_{g_j}^s - x\right)dx + \int_{T_{g_j}^e}^{C_j} \frac{1}{C_j}\left(C_j + T_{g_j}^s - x\right)dx = \frac{\left(C_j - g_j\right)^2}{2C_j}$$

Appendix B

The maximized green phase, which adopts the maximum time of green advance and green extension, is considered for calculating the delays of SAVs with priority control. For an SAV platoon that arrives before the original green phase, we need to locate the lead car

of the platoon and associate it with the maximized green phase. If the lead car of the SAV platoon arrives within the maximized green phase, the delays of each SAV are equal to 0. Otherwise, the delays vary according to the position of the platoon. For an SAV platoon that arrives after the original green phase, we only need to analyze each SAV of the platoon respectively. The delays are equal to 0, if its arrival time is within the maximized green phase. Otherwise, it needs to disassemble from the platoon and stay at the intersection until the next green phase. Therefore, the expected total delays $D^0_{i^{(t)}j}$ given in **Equation (32)** can be derived as follows:

The probability density function of each SAV's arrival time within one signal period is:

$$f(x)=\begin{cases}\dfrac{1}{C_j} & 0\le x\le C_j\\ 0 & else\end{cases}$$

The relationship between the dwelling time and the arrival time of the SAV platoon is:

$$D(x)=\begin{cases}T^s_{g_j}-a^{max}_{g_j}-x & 0\le x<T^s_{g_j}-a^{max}_{g_j}\\ 0 & T^s_{g_j}-a^{max}_{g_j}\le x\le T^e_{g_j}+e^{max}_{g_j}\\ C_j+T^s_{g_j}-x & T^e_{g_j}+e^{max}_{g_j}<x\le C_j\end{cases}$$

The expected total delays D^0_{ij} are calculated as follows:

$$D^0_{ij}=\int_0^{C_j}f(x)D(x)dx=\int_0^{T^s_{g_j}-a^{max}_{g_j}}\frac{1}{C_j}\left(T^s_{g_j}-a^{max}_{g_j}-x\right)dx+\int_{T^e_{g_j}+e^{max}_{g_j}}^{C_j}\frac{1}{C_j}\left(C_j+T^s_{g_j}-x\right)dx=\frac{\left(C_j-g_j-a^{max}_{g_j}-e^{max}_{g_j}\right)^2}{2C_j}$$

Appendix C

The increased delays at intersections for conventional vehicles, denoted as $D^2_{i^{(t)}j}$ in **Equation (37)**, can be derived as follows:

The delays of the conventional vehicles affected by green advance in phase p is:

$$D^{a,p}_{i^{(t)}j}=\int_{T^s_{g_j}-\frac{a_{g_j}}{3}}^{T^s_{g_j}}\left(C_j+T^s_{g_j}-g^{a,p}_j-x\right)v^1_{i^{(t)}j}dx=v^1_{i^{(t)}j}\left(C_j-g^{a,p}_j\right)\frac{a_{g_j}}{3}+\frac{1}{2}v^1_{i^{(t)}j}\left(\frac{a_{g_j}}{3}\right)^2$$

The total delays of the conventional vehicles affected by green advance is:

$$D^a_{i^{(t)}j}=\sum_p D^{a,p}_{i^{(t)}j}=\frac{1}{3}v^1_{i^{(t)}j}\left(2C_j+g_j\right)a_{g_j}+\frac{1}{6}v^1_{i^{(t)}j}{a_{g_j}}^2$$

The delays of the conventional vehicles affected by green extension in phase p is:

$$D^{e,p}_{i^{(t)}j}=\int_{T^e_{g_j}+g^{e,p}_j-C_j}^{T^e_{g_j}}\frac{e_{g_j}}{3}v^2_{i^{(t)}j}dx+\int_{T^e_{g_j}}^{T^e_{g_j}+e_{g_j}}\left(T^e_{g_j}+\frac{e_{g_j}}{3}-x\right)v^2_{i^{(t)}j}dx=v^2_{i^{(t)}j}\left(2C_j-g^{e,p}_j\right)\frac{e_{g_j}}{3}-\frac{1}{2}v^2_{i^{(t)}j}\left(\frac{e_{g_j}}{3}\right)^2$$

The total delays of the conventional vehicles affected by green extension is:

$$D^e_{i^{(t)}j}=\sum_p D^{e,p}_{i^{(t)}j}=\frac{1}{3}v^2_{i^{(t)}j}\left(5C_j+g_j\right)e_{g_j}-\frac{1}{6}v^2_{i^{(t)}j}{e_{g_j}}^2$$

The total delays of conventional vehicles at intersection are the sum of the above two terms:

$$D_{i^{(t)}j}\left(a_{g_j},e_{g_j}\right)=D^a_{i^{(t)}j}+D^e_{i^{(t)}j}=\frac{1}{3}v^1_{i^{(t)}j}\left(2C_j+g_j\right)a_{g_j}+\frac{1}{6}v^1_{i^{(t)}j}{a_{g_j}}^2+\frac{1}{3}v^2_{i^{(t)}j}\left(5C_j+g_j\right)e_{g_j}-\frac{1}{6}v^2_{i^{(t)}j}{e_{g_j}}^2$$

The probability density function of the arrival time of the head car of the SAV platoon within one signal period is:

$$f(x)=\begin{cases}\dfrac{1}{C_j} & 0\le x\le C_j\\ 0 & else\end{cases}$$

The time of green advance a_{g_j} and green extension e_{g_j} vary according to the arrival time of the SAV platoon. Therefore, the increase in delays of conventional vehicles at intersection is calculated as the expected value across all of these scenarios:

$$D^2_{i^{(t)}j} = \int_{T^s_{gj}-a^{max}_{gj}}^{T^s_{gj}} \frac{1}{C_j}\left[\frac{1}{3}v^1_{i^{(t)}j}(2C_j+g_j)\left(T^s_{g_j}-x\right)+\frac{1}{6}v^1_{i^{(t)}j}\left(T^s_{g_j}-x\right)^2\right]dx + \int_{T^e_{gj}}^{T^e_{gj}+e^{max}_{gj}} \frac{1}{C_j}\left[\frac{1}{3}v^2_{i^{(t)}j}(5C_j+g_j)\left(x-T^e_{g_j}\right)-\frac{1}{6}v^2_{i^{(t)}j}\left(x-T^e_{g_j}\right)^2\right]dx$$
$$= \frac{v^1_{i^{(t)}j}\left[a^{max3}_{gj}+3(2C_j+g_j)a^{max2}_{gj}\right]+v^2_{i^{(t)}j}\left[3(5C_j+g_j)e^{max2}_{gj}-e^{max3}_{gj}\right]}{18C_j}$$

References

Alonso-Mora J, Samaranayake S, Wallar A, et al. On-demand high-capacity carpooling via dynamic trip-vehicle assignment. Proceedings of the National Academy of Sciences, 2017, 114(3): 462-467.

Alazzawi, S., Hummel, M., Kordt, P., et al., 2018. Simulating the impact of shared, autonomous vehicles on urban mobility-a case study of Milan. EPiC Ser. Eng. 2, 94–110.

Anderson, P., Daganzo, C.F., 2020. Effect of transit signal priority on bus service reliability[J]. Transp. Res. B Methodol. 132, 2–14.

Antonio, G.P., Maria-Dolores, C., 2022. Multi-agent deep reinforcement learning to manage connected autonomous vehicles at tomorrow's intersections[J]. IEEE Trans. Veh. Technol. 71 (7), 7033–7043.

Chen, Z., He, F., Zhang, L., et al., 2016. Optimal deployment of autonomous vehicle lanes with endogenous market penetration. Transp. Res. Part C: Emerg. Technol. 72, 143–156.

Chen, Z., Li, X., 2021. Designing corridor systems with modular autonomous vehicles enabling station-wise docking: Discrete modeling method[J]. Transp. Res. Part E: Logistics Transp. Rev. 152, 102388.

Cherkesly, M., Desaulniers, G., Laporte, G., 2015. A population-based metaheuristic for the pickup and delivery problem with time windows and LIFO loading. Comput. Oper. Res. 62, 23–35.

Dresner K, Stone P. Multiagent traffic management: A reservation-based intersection control mechanism[C]//Autonomous Agents and Multiagent Systems, International Joint Conference on. IEEE Computer Society, 2004, 3: 530-537.

Cokyasar, T., Larson, J., 2020. Optimal assignment for the single-household shared autonomous vehicle problem. Transp. Res. B Methodol. 141, 98–115.

Farhan, J., Chen, T.D., 2018. Impact of ridesharing on operational efficiency of shared autonomous electric vehicle fleet. Transp. Res. Part C: Emerg. Technol. 93, 310–321.

Ghiasi, A., Hussain, O., Qian, Z.S., et al., 2017. A mixed traffic capacity analysis and lane management model for connected automated vehicles: A Markov chain method. Transp. Res. B Methodol. 106, 266–292.

Gurumurthy, K.M., Kockelman, K.M., 2018. Analyzing the dynamic carpooling potential for shared autonomous vehicle fleets using cellphone data from Orlando, Florida. Comput. Environ. Urban Syst. 71, 177–185.

Hasan, M.H., Van Hentenryck, P., 2021. The benefits of autonomous vehicles for community-based trip sharing[J]. Transp. Res. Part C: Emerg. Technol. 124, 102929.

Levin, M.W., Boyles, S.D., Patel, R., 2016. Paradoxes of reservation-based intersection controls in traffic networks[J]. Transp. Res. A Policy Pract. 90, 14–25.

Levin, M.W., Odell, M., Samarasena, S., Schwartz, A., 2019. A linear program for optimal integration of shared autonomous vehicles with public transit. Transp. Res. Part C: Emerg. Technol. 109, 267–288.

Li, D., Zhu, F., Chen, T., et al., 2023. COOR-PLT: A hierarchical control model for coordinating adaptive platoons of connected and autonomous vehicles at signal-free intersections based on deep reinforcement learning[J]. Transp. Res. Part C: Emerg. Technol. 146, 103933.

Liang, X., de Almeida, C.G.H., An, K., et al., 2020. Automated taxis' dial-a-ride problem with carpooling considering congestion-based dynamic travel times. Transp. Res. Part C: Emerg. Technol. 112, 260–281.

Liang, S., Leng, R., Zhang, H., et al., 2023. Two-stage transit signal priority control method to improve reliability of bus operation considering stochastic process[J]. Transp. Res. Rec. 2677 (2), 1713–1727.

Liu, Z., Song, Z., 2019. Strategic planning of dedicated autonomous vehicle lanes and autonomous vehicle/toll lanes in transportation networks. Transp. Res. Part C: Emerg. Technol. 106, 381–403.

Lokhandwala, M., Cai, H., 2018. Dynamic ride sharing using traditional taxis and shared autonomous taxis: A case study of NYC. Transp. Res. Part C: Emerg. Technol. 97, 45–60.

Mohajerpoor, R., Ramezani, M., 2019. Mixed flow of autonomous and human-driven vehicles: Analytical headway modeling and optimal lane management. Transp. Res. Part C: Emerg. Technol. 109, 194–210.

Pei, M., Lin, P., Du, J., et al., 2021. Vehicle dispatching in modular transit networks: A mixed-integer nonlinear programming model[J]. Transp. Res. Part E: Logistics Transp. Rev. 147, 102240.

Pourgholamali, M., Miralinaghi, M., Seilabi, S.E., et al., 2023. Sustainable deployment of autonomous vehicles dedicated lanes in urban traffic networks[J]. Sustain. Cities Soc. 99, 104969.

Razmi Rad, S., Farah, H., Taale, H., et al., 2020. Design and operation of dedicated lanes for connected and automated vehicles on motorways: A conceptual framework and research agenda. Transp. Res. Part C: Emerg. Technol. 117, 102664.

Rey, D., Levin, M.W., 2019. Blue phase: Optimal network traffic control for legacy and autonomous vehicles. Transp. Res. B Methodol. 130, 105–129.

Ropke, S., Pisinger, D., 2006. An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. Transp. Sci. 40 (4), 455–472.

Scheltes, A., de Almeida, C.G.H., 2017. Exploring the use of automated vehicles as last mile connection of train trips through an agent-based simulation model: An application to Delft, Netherlands[J]. Int. J. Transp. Sci. Technol. 6 (1), 28–41.

Seilabi, S.E., Pourgholamali, M., de Almeida, C.G.H., et al., 2023. Robust design of CAV-dedicated lanes considering CAV demand uncertainty and lane reallocation policy[J]. Transp. Res. Part D: Transp. Environ. 121, 103827.

Seo, T., Asakura, Y., 2022. Multi-objective linear optimization problem for strategic planning of shared autonomous vehicle operation and infrastructure design. IEEE Trans. Intell. Transp. Syst. 23 (4), 3816–3828.

Simonetto, A., Monteil, J., Gambella, C., 2019. Real-time city-scale ridesharing via linear assignment problems. Transp. Res. Part C: Emerg. Technol. 101, 208–232.

Sun, P., Veelenturf, L.P., Hewitt, M., Van Woensel, T., 2018. The time-dependent pickup and delivery problem with time windows. Transp. Res. B Methodol. 116, 1–24.

Sun, P., Veelenturf, L.P., Hewitt, M., et al., 2020. Adaptive large neighborhood search for the time-dependent profitable pickup and delivery problem with time windows. Transp. Res. Part E: Logistics Transp. Rev. 138, 101942.

Tafreshian, A., Masoud, N., 2020. Trip-based graph partitioning in dynamic ridesharing. Transp. Res. Part C: Emerg. Technol. 114, 532–553.

Tian, Q., Lin, Y.H., Wang, D.Z.W., et al., 2022. Planning for modular-vehicle transit service system: Model formulation and solution methods[J]. Transp. Res. Part C: Emerg. Technol. 138, 103627.

Truong, L.T., Currie, G., Wallace, M., et al., 2019. Coordinated transit signal priority model considering stochastic bus arrival time. IEEE Trans. Intell. Transp. Syst. 20 (4), 1269–1277.

Viegas, J., Lu, B., 2004. The intermittent bus lane signals setting within an area[J]. Transp. Res. Part C: Emerg. Technol. 12 (6), 453–469.

Wan, Q.K., Lo, H.K., 2009. Congested multimodal transit network design[J]. Public Transport 1, 233–251.

Wang, J., Lu, L., Peeta, S., et al., 2021. Optimal toll design problems under mixed traffic flow of human-driven vehicles and connected and autonomous vehicles. Transp. Res. Part C: Emerg. Technol. 125.
Wu, Y., Chen, H., Zhu, F., 2019. DCL-AIM: Decentralized coordination learning of autonomous intersection management for connected and automated vehicles[J]. Transp. Res. Part C: Emerg. Technol. 103, 246–260.
Xu, M., Zhai, X., Sun, Z., et al., 2023. Multiagent control approach with multiple traffic signal priority and coordination[J]. J. Transp. Eng. Part A: Syst. 149 (1), 04022124.
Ye, L., Yamamoto, T., 2018. Impact of dedicated lanes for connected and autonomous vehicle on traffic flow throughput[J]. Phys. A: Stat. Mech. App. 512, 588–597.
Zeng, X., Zhang, Y., Jiao, J., et al., 2020. Route-based transit signal priority using connected vehicle technology to promote bus schedule adherence[J]. IEEE Trans. Intell. Transp. Syst. 22 (2), 1174–1184.