Transportation Research Part C xxx (xxxx) xxx

Contents lists available at ScienceDirect

Transportation Research Part C

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

An Autonomous Modular Public Transit service☆

Xi Cheng [a], Yu (Marco) Nie [b], Jane Lin [a,c,*]

[a] Department of Civil, Materials, and Environmental Engineering, University of Illinois Chicago, 842 W. Taylor Street (MC246), Chicago, IL 60607, USA
[b] Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL 60208, USA
[c] Institute for Environmental Science and Policy, University of Illinois Chicago, 842 W. Taylor Street (MC246), Chicago, IL 60607, USA

ARTICLE INFO

Keywords:
Autonomous Modular Vehicle Technology (AMVT)
Autonomous Modular Public Transit (AMPT)
Gridded Fixed-route Transit Network
Pod
Joining/Disjoining of Pods
Agency and passenger costs

ABSTRACT

In this work, we present a proof-of-concept investigation of Autonomous Modular Public Transit (AMPT) at a network scale and compare it with the traditional fixed-route, fixed-vehicle size transit service in terms of total cost, which consists of both agency's capital and operational cost (including energy cost) and passenger time cost. We formulate and solve stylized design models for AMPT on a grid network in a range of demand density scenarios with both homogenous and heterogeneous distributions. The AMPT models explicitly account for pod joining and disjoining (and therefore en-route transfers of passengers) and potential energy savings due to pod train formation (pod platooning), which represent major departures from the traditional transit models in the literature. Numerical results find that AMPT, if designed properly, may save the total cost compared to traditional transit systems thanks to demand responsive pod train capacity, particularly in the low demand scenarios. The cost savings of AMPT are largely attributed to passenger time saving by en-route transfer; the agency cost of AMPT has a mixed picture. The load factor of AMPT generally improves over the traditional transit service. We also show how key parameter values may affect the AMPT costs through sensitivity analysis.

1. Introduction

In the U.S., public transit has long struggled to attract ridership. From 1960 to 2019, the share of commuting trips made by public transit decreased from 12.1 % to around 5.0 % (US Census Bureau, 2019). This led to extremely low utilization of transit buses, whose average load factors ranged between 10.1 % and 12.4 % (National Transit Database, 2015; US Census Bureau, 2019). The COVID pandemic has further exacerbated the situation, pulling even more would-be passengers away from buses. The public transit sector needs to rethink its current service model – often fixed route, schedule, and vehicle size – and look for innovative ones that provide much greater flexibility and efficiency. The rise of demand-responsive transit and micro transit may be viewed as a response to this call for changes. However, these new modes face the scalability challenge and risk worsening traffic congestion by flooding urban streets with a large quantity of small vehicles. The emergence of the so-called Autonomous Modular Vehicle Technology (AMVT) (see, e.g., NEXT.com, CNN.com, 2018), may enable a potential paradigm shift in transit service. AMVT offers both modularity and autonomy. Modularity means vehicle capacity can adapt to passenger demand in real-time. The modular autonomous vehicles, also known as pods and powered by electricity, run on existing road infrastructure either individually or as a connected pod train (by pod joining and

☆ This article belongs to the Virtual Special Issue on "ISTTT25".

* Corresponding author.
E-mail address: janelin@uic.edu (J. Lin).

https://doi.org/10.1016/j.trc.2024.104746
Received 15 January 2024; Received in revised form 15 April 2024; Accepted 30 June 2024

disjoining as seen in Fig. 1). Coupling modularity with autonomy, this Autonomous Modular Public Transit (AMPT) service is poised to offer a promising solution for better utilization in low and high-demand situations.

As an emerging technology, the literature related to AMPT is in its nascent stage. As one of the first papers related to AMPT, Gecchelin and Webb (2019) envisioned a modular dynamic ridesharing (MDRS) system comprised of low capacity (up to 10 people) pods that can operate individually for off-peak low-capacity last mile feeder services or be articulated into a single vehicle (i.e., a pod train) as a quasi-rapid transit system. They claimed that MDRS has the potential to reduce congestion by decreasing the size and number of vehicles and increasing cost-effectiveness and convenience.

In the subsequent literature, most studies have focused on the AMPT service design and operational parameters in a single-line or corridor operation. Some studies investigated the demand responsive design of AMPT. For example, Zhang et al. (2020) proposed an AMPT consisting of a set of trailer modules that can travel locally to serve demand and connect travelers to main modules for long-distance trips. They also compared the modular transit service with the door-to-door shuttle service as the benchmark to showcase the benefits of modular transit. Liu et al. (2021) presented a functional design for flexible-route transit services using self-adaptive capacities of modular autonomous vehicles (MAVs) in response to heterogeneous demand, and demonstrate that the MAV-based service is more effective and attractive than the traditional transit services. Caros and Chow (2021) developed a day-to-day operational policy adjustment to learn the latent demand of a MAV service with an objective of improving service and increasing profits in a mobility service market, which is the first to evaluate a last-mile operating strategy using MAVs allowing stationary transfers.

Some studies focused on varying capacity and dispatching strategy designs. Chen et al. (2020) investigated a time-varying capacity design problem for modular mass transit systems with a focus on the joint design of dispatching headway and vehicle capacity for one-to-one shuttle systems in oversaturated traffic to achieve the optimal tradeoff between general vehicle dispatching cost (mainly comprised of vehicle energy consumption) and customer waiting cost. The problem was formulated in a tactical level continuum approximation (CA) model to illuminate fundamental analytical properties in the optimal design. Chen and Li (2021) designed a transportation corridor system with MAV enabling station-wise docking. Station-wise docking means vehicles can change their formations (or capacity) at any station along the corridor. The problem was solved with a customized branch and bound (B&B) algorithm, which showed a better performance than the commercial solver Gurobi for relatively large instances of real-world applications. In another study, Chen et al. (2021) applied a continuum approximation (CA) modeling approach to the above problem. Compared with the CA method, the B&B method returns solutions with smaller gaps, which can be taken as an exact benchmark algorithm. Dakic et al., (2021) proposed a flexible bus dispatching system with a fleet of fully automated modular bus units, and developed an optimization framework based on the three-dimensional macroscopic fundamental diagram (3D-MFD) to determine the optimal composition of modular bus units and the optimal service frequency. Liu et al. (2020) investigated a minimum fleet size (MFS) problem for a single-line AMPT system. Most recently, Guo et al. (2023) developed a two-phase optimization process for scheduling Modular Autonomous Electric Vehicles in customized bus services. Khan et al. (2023) introduced an AMV bus splitting strategy to improve headway consistency and validated it through macroscopic simulations. Tian et al. (2023) propose a dynamic model that synchronizes scheduling and vehicle formation to meet fluctuating travel demands.

Other studies addressed the in-motion transfer operational strategies by joining and disjoining of modular vehicles. For example, Wu et al. (2021) developed a modular, adaptive, and autonomous transit system (MAATS) and proposed an optimal in-motion transfer strategy to minimize passenger transfers and autonomous modular buses (AMBs) movements. The study also evaluated the benefits of the AMB system from the passengers' perspective and validated the benefit (i.e., shorter travel time, fewer transfers).

Lastly, Lin et al. (2023) discussed the potential of Autonomous Modular Vehicle Technology (AMVT) in transforming public transit and last-mile delivery systems. Our study paper follows the paradigm as envisioned in Lin et al. (2023) and focuses on transit network design for transporting passengers.

In summary, most existing studies subscribe to the notion that AMVT can bring a range of operational benefits to passenger transportation and set out to demonstrate how it may work in an isolated and sometimes highly simplified service system. Studies above typically focus on one corridor or route to dispatch the MAVs. In this work, however, we present a proof-of-concept investigation of AMPT at a network scale and compare it with the traditional fixed-route, fixed-vehicle size transit service in terms of total cost – both capital and operational (including energy cost) on the agency side and time cost on the passenger side – and load factor. We develop stylized design models for AMPT on a grid network in a range of demand density scenarios with both homogenous and heterogeneous spatial distributions. The AMPT models explicitly account for pod joining and disjoining (and therefore enabled en-route transfers of

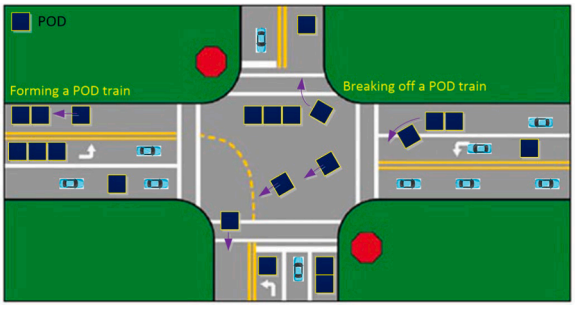


Fig. 1. En-route pod train joining and disjoining.

passengers) and potential energy savings due to pod train formation (platooning), which represent major departures from the traditional transit models in the literature.

The rest of the paper is organized as follows. Section 2 describes the setup of four transit service scenarios we investigate in this paper. Section 3 describes the detailed design models of the four scenarios. Section 4 presents the proposed solution method. Section 5 discusses the results of numerical experiments and sensitivity analysis for model validation and evaluation. Section 6 concludes the study.

2. Problem setting

We consider a fixed-route bus transit system in a square grid network shown in Fig. 2, which is adapted from Chen and Nie (2017). The total area is D by D (sq. km). The street network is depicted by the grey lines in the background, with a spacing of s km. The solid lines in blue represent the transit lines that run in both directions. N is the number of transit lines in each direction. The transit line spacing is D/N. In the rest of the paper the transit network size is denoted by $N \times N$. Each intersection of two perpendicular fixed transit lines defines a designated transit stop. Furthermore, the $N \times N$ network is divided into $N \times N$ equal *square zones* with each transit stop at the centroid of the zone, as shown in Fig. 3.

To simplify the problem formulations to be presented in this paper, we make the following general assumptions in the context of a fixed-line transit service, which apply to both traditional transit and AMPT:

- **Assumption 1**. Passengers walk to the closest stops from their origins and get off at the closest stops to their destinations.
- **Assumption 2**. Passengers travel between these stops as directly as possible with the least possible number of transfers (i.e., there is at most one transfer in a grid transit network).
- **Assumption 3**. Passengers randomly choose the initial direction of travel when their origins and destinations are located at the central square.
- **Assumption 4**. To ensure an apple-to-apple comparison of operating costs, both the baseline traditional transit system and the proposed AMPT system operate with electric autonomous vehicles.
- **Assumption 5**. The road infrastructure for both AMPT and traditional transit service is essentially the same. No additional docking requirements are needed for AMPT.
- **Assumption 6**. All transit vehicles travel at a constant cruising speed in mid-block and constant acceleration and deceleration rates at stops.

Assumption 4 excludes driver cost in the total cost consideration in both services since AMPT pods are driverless. According to the American Public Transportation Association (APTA) 2023 Fact Book, the labor cost accounts for 75 % of the total operating cost of transit bus service (APTA, 2023, pp.3). On the other hand, the cost of a self-driving system is a capital investment and may come down over time as the technology matures. Assumption 4 ensures that the comparison does not favor one or the other. Assumption 5 assumes the same infrastructure cost in both services and therefore infrastructure is not considered in total cost comparison between the two.

Two spatial passenger demand distributions are considered: homogeneous and heterogeneous across the network. Heterogeneous

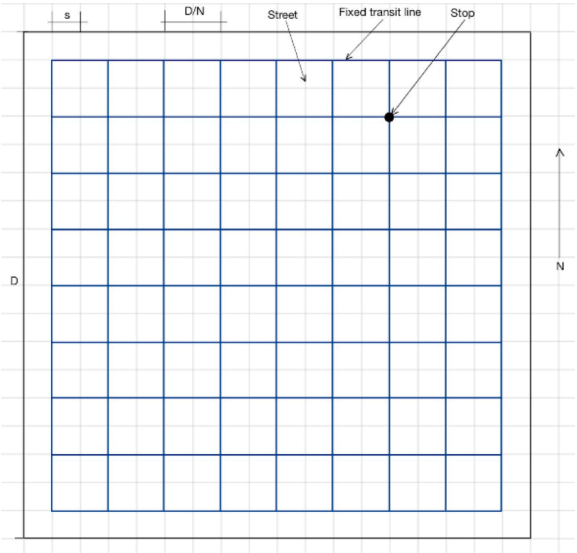


Fig. 2. A gridded transit network. (a 9 × 9 transit network in this example).

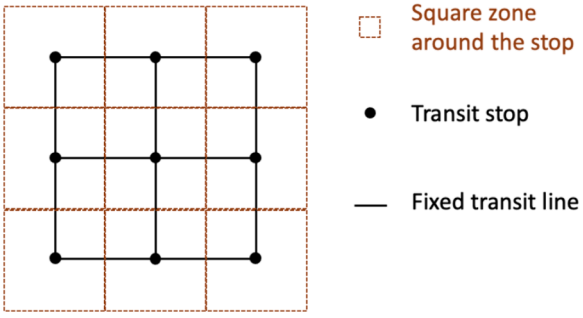


Fig. 3. Illustration of square zones using a 3 × 3 transit network.

demand distribution is assumed to be a continuously differentiable function. The analysis period is one hour, within which the spatial demand distribution remains fixed.

Our analysis covers the following four scenarios, two for each type of demand distribution.

1) **Set 1**: spatially homogeneous demand density:
- **Scenario 1 (P1)**: base scenario – a traditional fixed-route, fixed-vehicle size bus service, modified from Daganzo (2010);
- **Scenario 2 (P2)**: proposed scenario – a fixed-route AMPT service;

2) **Set 2**: spatially heterogeneous demand density:
- **Scenario 3 (P3)**: base scenario – a traditional fixed-route, fixed-vehicle size bus service, modified from Ouyang et al. (2014);
- **Scenario 4 (P4)**: proposed scenario – a fixed-route AMPT service.

In Scenarios 2 and 4, pod joining and disjoining are realized when pods are either dwelling or in motion. We make further assumptions about pod joining/disjoining operations as follows.

- Assumption 7 In theory, a pod (or a subset of pods) can disjoin from its current pod train anywhere on the transit line before turning onto the destination transit line. In this study, for simplification, disjoin is assumed to takes place only at the intersection between the current transit line and the destination transit line for energy saving. This is depicted by the blue pod train in Fig. 4(1b), (2b) and (3b), where the last pod in the blue EB pod train disjoins from it at the intersection and turns to NB, carrying passengers who need to transfer from EB to NB.
- Assumption 8. Similarly, pod joining can take place anywhere on the transit line near an intersection or mid-block, as illustrated in Fig. 4(1c), (2c) and (3c). In this study, it is assumed that, after the disjoining operation described in Assumption 7, the disjoined pod (s) must join and become an integral part of a dispatched pod train on the destination transit line by either speeding up to catch the pod train ahead (Fig. 4(1c) and (2c)) or slowing down to join the incoming pod train (Fig. 4(3c)). As such, the pod capacity can be better utilized without interfering with the pre-scheduled destination transit line operation. Furthermore, the resulting pod train must not exceed the maximum pod train length (L_m) as defined in Assumption 9.
- Assumption 9 A maximum pod train length (L_m) is assumed considering practical mechanical and technological constraints in pod operation.
- Assumption 10. Not all transfer demand needs to or could be satisfied via en-route transfer. In other words, manual transfer – passengers physically get off at a stop and transfer to other bus lines – may take place if needed.
- Assumption 11 There is a small amount of delay in pod joining due to the technical requirements of coupling mechanics. On the other hand, pod disjoining (detachment) is instantaneous (i.e., no delay).

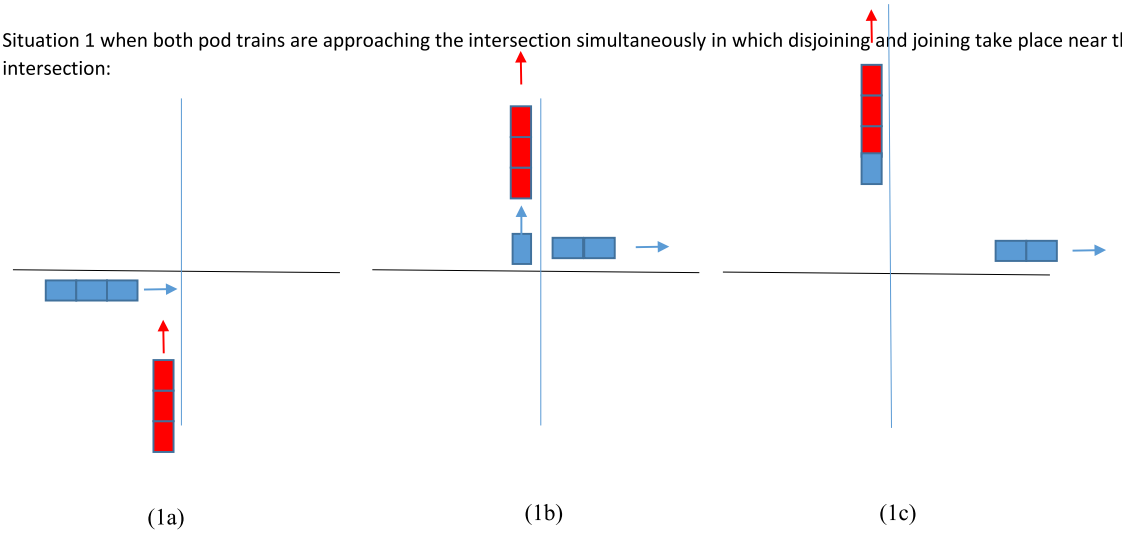
It is worth noting that no joining is allowed between any two dispatched pod trains of the same transit line. Assumption 8 eliminates disjoined pods from operating on their own on the destination transit line. They must become a part of the dispatched pod trains on that transit line.

In this study, the maximum pod train length (L_m) is determined as follows. A current pod prototype (NEXT.com) has a length of 2.67 m, and a typical electric bus length is 12.2 m. Therefore, we assume that a pod train can have up to 5 pods at once, i.e., L_m=5.

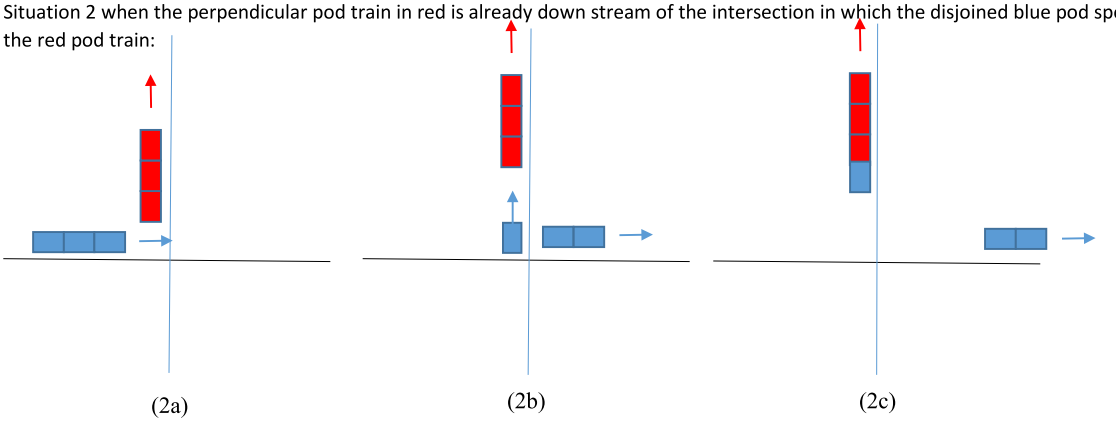
3. Design models

We now present the formulations of P1 and P2, and P3 and P4, between the traditional transit and the proposed AMPT. Most model notations follow Daganzo (2010) for P1 and P2 and Ouyang et al. (2014) for P3 and P4. Table 1 lists all the letter notations used in this paper.

Situation 1 when both pod trains are approaching the intersection simultaneously in which disjoining and joining take place near the intersection:



Situation 2 when the perpendicular pod train in red is already down stream of the intersection in which the disjoined blue pod speeds up to join the red pod train:



Situation 3 when the perpendicular pod train in red is mid-block toward the intersection in which the disjoined blue pod waits to join the red pod train when it crosses the intersection:

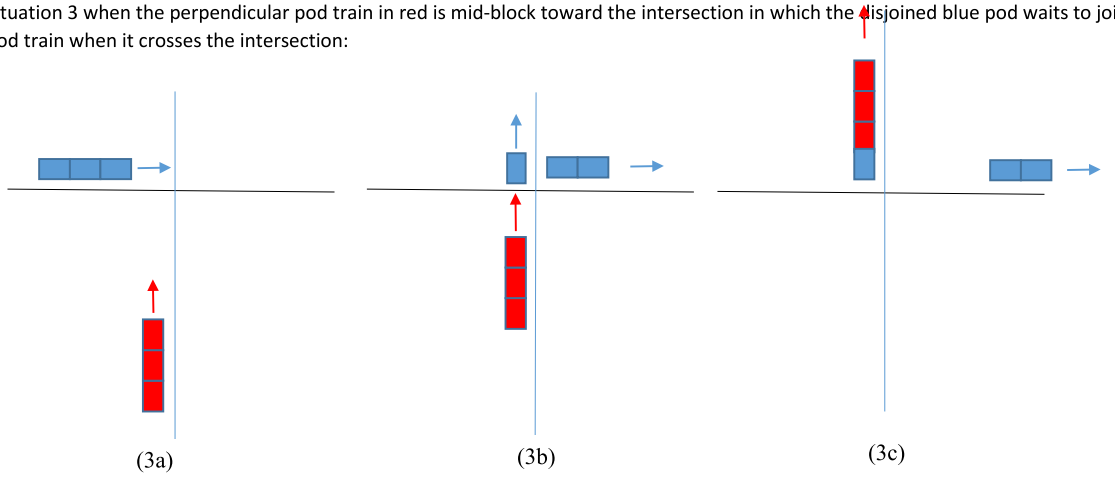


Fig. 4. Illustration of pod train joining and disjoining between two perpendicular directions.

3.1. Spatially homogeneous demand density

3.1.1. Scenario 1 – Traditional transit service (P1)

Following Daganzo (2010), P1 considers costs incurred for both the transit agency and the passengers within one hour of operation.

Table 1
Notations.

Notation	Description
D(km)	Side length of the grid network
s(km)	Distance between two adjacent parallel streets (street spacing)
λ(pax/hr/km^2)	Demand density in number of passenger trips generated per hour per km^2
N	Number of transit lines in each direction (EB, WB, NB, SB) in the grid network
I	Set of vertical transit lines (N-S),$I = \{1, 2, \cdots, N\}, \lvert I\rvert = N$
J	Set of horizontal transit lines (E-W), $J = \{1, 2, \cdots, N\}, \lvert J\rvert = N$
Dir	Set of directions,$Dir = \{NB, SB, WB, EB\}$
l_0(km)	Transit line spacing,$l_0 = D/N$
H(hr)	Transit vehicle headway
μ($/hr)	Value of time
δ(km)	Transfer penalty expressed in terms of the equivalent distance walked
v(km/hr)	Average vehicle operating speed
v_w(km/hr)	Walking speed
τ(sec/stop)	Time lost per stop due to deceleration and acceleration
τ'(sec/pax)	Boarding time per passenger
$\$_C$($/veh • hr)	Capital cost of an electric bus deprecated over an assumed 12-year lifespan
$\$_Q$($/veh • km)	Operating cost per vehicle distance of a bus
$\$_E$($/veh • km)	Energy cost per vehicle distance of a bus
Q(km)	Expected total vehicle distance traveled (VKT) per hour of operation
M	Number of vehicles required for hourly operation
π_Q, π_M	Conversion coefficients of agency costs into equivalent travel time per passenger (see Daganzo (2010))
A(hr/pax)	Average total walking time per passenger
T(hr/pax)	Average total in-vehicle travel time per passenger
W(hr/pax)	Average total waiting time per passenger
v_c(km/hr)	Average vehicle operating speed including all time losses due to acceleration/deceleration at a transit stop and boarding/alighting activities.$v_c = \frac{Q}{M}$
e_T	expected number of transfers per passenger
$\$'_C$($/pod • hr)	Capital cost of a pod deprecated over a 12-year lifespan
$\$'_Q$($/pod • km)	Operating cost per vehicle distance of a pod
$\$'_E$($/pod • km)	Energy cost per vehicle distance of a pod
η_{pl}	Energy saving rate η_{pl} due to platooning of pods (pod train formation), η_{pl} is a function of the number of pods (L) in a pod train p
$t_{ij}^{dir,\mathscr{T}}$	Arrival time of pod train $\mathscr{T}$ at an intersection (i,j) in the direction of dir, where i is the vertical transit line index, and j is the horizontal transit line index,$\forall\, i \in I, j \in J, dir \in Dir, \mathscr{T} \in \left\{1, \cdots, \left\lfloor \frac{1}{H} \right\rfloor\right\}$
τ_f(s/joint)	Time lost due to pod joining in seconds
TJ(joints/hr)	Total number of hourly joints network wide
TPJ(joints/hr)	Total number of potential hourly joints network wide
L(pods)	Number of pods in a pod train
L_m(pods)	Maximum number of pods in a pod train allowed ($L_m = 5$ in this study)
C_a(pax)	Pod capacity (number of seats) – assuming no standing passengers
C_a^b(pax)	Capacity of traditional bus (number of seats) – assuming no standing passengers

The total cost $z(N,H)$ is a function of the number of transit lines $N\,(N>1)$ and dispatching headway H, and consists of these six components (all expressed in time-equivalent cost per passenger, hrs./pax equivalent): [i] agency operating cost of vehicle kilometers of traveled (VKT) and energy use, [ii] agency operating cost of required fleet size taking account time losses due to acceleration/deceleration at a transit stop and boarding/alighting activities, [iii] passenger waiting time, [iv] passenger walking time, [v] in-vehicle travel time, and [vi] transfer time.

$$z(N,H) = \underset{[i]}{\pi_Q Q} + \underset{[ii]}{\pi_M M} + \underset{[iii]}{A} + \underset{[iv]}{T} + \underset{[v]}{W} + \underset{[vi]}{\frac{\delta}{v_w} e_T}$$

In [i] and [ii], $Q = \frac{4ND}{H}$ is the total hourly VKT; $M = \frac{Q}{v_c} = \frac{Q}{v} + \frac{2\tau N^2}{H} + \tau'(1+e_T)\lambda D^2$ is the fleet size (number of vehicles required for hourly operation), where v_c is the average vehicle operating speed after taking account all time losses due to acceleration/deceleration and boarding/alighting activities at a transit stop. Coefficients $\pi_Q = \frac{\$_Q + \$_E}{\lambda D^2 \mu}$ and $\pi_M = \frac{\$_C}{\lambda D^2 \mu}$ convert Q and M into equivalent travel time cost per passenger, respectively (see Daganzo (2010)). A key addition in our formulation is the incorporation of the per km energy cost $\$_E$ in hourly VKT cost. Furthermore, per hour capital cost $\$_C$ of transit bus is calculated over an assumed 12-year life span of a bus.

In [iii], $A = \frac{2}{v_w} \times \frac{1}{2} \frac{D}{N} = \frac{D}{v_w N}$ denotes average total walking time per passenger. It encompasses walking time from both the origin to the nearest starting stop and the nearest alighting stop to the final destination. Average walking time is calculated by the distances of

spatially homogenous distributed passengers to the transit stop at the center of a square zone in the gridded transit network (see Fig. 3).

In [iv], $T = \frac{E}{v_c}$ denotes the average total in-vehicle travel time per passenger, where E is the total in-vehicle travel distance per passenger, $E = d \cdot Prob(Tr = 0) + 2d \cdot Prob(Tr = 1)$. d is the average in-vehicle travel distance of a passenger in one direction, $d = 0.34D$, adopted from Aldaihani et al., (2004). $Prob(Tr = 0)$ is the probability of no transfer in a passenger trip and $Prob(Tr = 1)$ is the probability of having one transfer in a passenger trip – there is at most one transfer in a gridded transit network. They are determined as follows. For a grid transit network of Fig. 3, there are N^2 square zones divided by the fixed transit lines. When the origin and destination are in the same square zone, passengers walk instead of taking the transit. If there is no transfer in a trip, once the origin is given, the destination must be in one of the $2N-2$ square zones. Thus, the probability of no transfer is given by $Prob(Tr = 0) = \frac{2N-2}{N^2-1} = \frac{2}{N+1}$, and $Prob(Tr = 1) = 1 - \frac{2}{N+1} = \frac{N-1}{N+1}$. The total in-vehicle travel distance per passenger is $E = d \cdot Prob(Tr = 0) + 2d \cdot Prob(Tr = 1) = \frac{0.34D \times 2N}{N+1}$.

In [v], $W = \frac{H}{3} + \frac{H}{2}e_T$ denotes the expected passenger delay at stops within a one-hour operation. It consists of waiting time at a starting stop in the first term and delay due to transfer activities in the second term. In the gridded network, each passenger can take either E-W or N-S transit for her trip. At the starting stop, the waiting times of passengers for the E-W and the N-S bus are assumed independently and uniformly distributed in the interval $[0, H]$. Thus, the expected waiting time at the starting stop is $\frac{H}{3}$ (see Murphy (2012)). At a transfer stop, the expected waiting time is $\frac{H}{2}e_T$.

In [vi], transfer penalty per passenger is $\frac{\delta}{v_w}e_T$, where δ is the equivalent walking distance. e_T is the expected number of transfers per passenger defined as follows:

$$e_T = 0 \times Prob(Tr = 0) + 1 \times (1 - Prob(Tr = 0)) = 1 - Prob(Tr = 0) = \frac{N-1}{N+1}$$

The final model formulation for P1 is summarized below. Constraint (1)–(2) guarantees that dispatched buses meet the passenger demand of the network, while Constraint (1)–(3) ensures the feasibility of both variables.

$$(\mathbf{P1}) \min z(N, H) = \pi_Q Q + \pi_M M + A + T + W + \frac{\delta}{v_w}e_T \tag{1-1}$$

s.t.,

$$MC_a^b \geq \lambda D^2 \tag{1-2}$$

$$H > 0, N \in \{2, \cdots, \lfloor D/s \rfloor\} \tag{1-3}$$

where:

$$\pi_Q = \frac{\$_Q + \$_E}{\lambda D^2 \mu} \tag{1-4}$$

$$\pi_M = \frac{\$_C}{\lambda D^2 \mu} \tag{1-5}$$

$$Q = \frac{4ND}{H} \tag{1-6}$$

$$M = \frac{Q}{v} + \frac{2\tau N^2}{H} + \tau'(1 + e_T)\lambda D^2 \tag{1-7}$$

$$W = \frac{H}{3} + \frac{H}{2}e_T \tag{1-8}$$

$$A = \frac{D}{v_w N} \tag{1-9}$$

$$T = \frac{0.68ND}{N+1} \times \frac{M}{Q} \tag{1-10}$$

$$e_T = \frac{N-1}{N+1} \tag{1-11}$$

3.1.2. Scenario 2 – AMPT (P2)

In P2, conventional transit vehicles are replaced with pods and pod trains. There are several key operational distinctions due to this change. First, the initial length of a pod train (L) at dispatching is determined by the minimum of these two: the number of pods to meet the transit line demand, and the maximum pod train length allowed (L_m) to ensure that the dispatched pod train is demand responsive within the technical/operational limit. Furthermore, after each stop a pod train length may change due to joining/disjoining activities to facilitate en-route transfer. The pod train length needs to be updated after each stop. Second, pod joining/disjoining enables

seamless en-route transfers in motion, which incurs no waiting time for the transfer passengers. Lastly, a pod train is regarded as pod platooning in operation. Because of these features, an energy saving rate η_{pl} for a pod train as a function of the pod train length is considered.

There are three basic elements involved in P2: (a) determination of the number of onboard passengers in a pod train upon arrival at a stop, (b) determination of the numbers of en-route and manual transfers and the pod train length upon leaving a stop, and (c) energy saving by pod train formation (pod platooning). They are described as follows.

a. ***Number of onboard passengers upon arrival at a stop***.

From Fig. 2, pod trains run in four directions of the transit network, SB, NB, EB and WB. There are a total of 4 N number of transit lines. In the case of spatially homogeneous demand, it implies that the arrival rate of passengers (λ) at each stop is uniform. So, the total number of passengers arriving and waiting at a stop for all four directions between two successive pod trains is $H\lambda\left(\frac{D}{N}\right)^2$, half of which board the horizontal lines EB and WB and the other half board the vertical lines SB and NB. It is further assumed that for each O-D pair on a directional transit line, the O-D demand is uniform and identical to $\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{1}{N-1}$.

Consider a WB pod train $\mathcal{T}$. Upon arrival at a transit stop k ($k\in\{1,\cdots,N\}$), the number of onboard passengers in $\mathcal{T}$ is denoted by $P_{\mathcal{T}}(k)$, which is divided into three groups as illustrated in Fig. 5(a): (i) those who alight at stop k, denoted by $PD_{\mathcal{T}}(k)$; (ii) those who transfer to other transit lines at stop k, denoted by $PT_{\mathcal{T}}(k)$, including both en-route and manual transfers; and (iii) those who remain in $\mathcal{T}$, denoted by $PS_{\mathcal{T}}(k)$. At stop k, two types of passengers board $\mathcal{T}$: those originating from the square zone served by stop k, denoted by $PB_{\mathcal{T}}(k)$; and passengers transferring from the perpendicular directions, denoted by $PT^{\perp}_{\mathcal{T}}(k)$. The number of onboard passengers in $\mathcal{T}$ upon arrival at stop $k+1$ is then computed in Eq. (2)–(1).

$$P_{\mathcal{T}}(k+1)=P_{\mathcal{T}}(k)-PD_{\mathcal{T}}(k)-PT_{\mathcal{T}}(k)+PB_{\mathcal{T}}(k)+PT^{\perp}_{\mathcal{T}}(k) \tag{2-1}$$

where,

$$PD_{\mathcal{T}}(k)=\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N}\cdot\frac{1}{N-1}+\sum_1^{k-1}\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{N-k}{N}\cdot\frac{1}{N-k}=\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N-1} \tag{2-2}$$

$$PT_{\mathcal{T}}(k)=\sum_1^{k-1}\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{N-k}{N-1}\cdot\frac{1}{N-k}\cdot\frac{N-1}{N}=\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N} \tag{2-3}$$

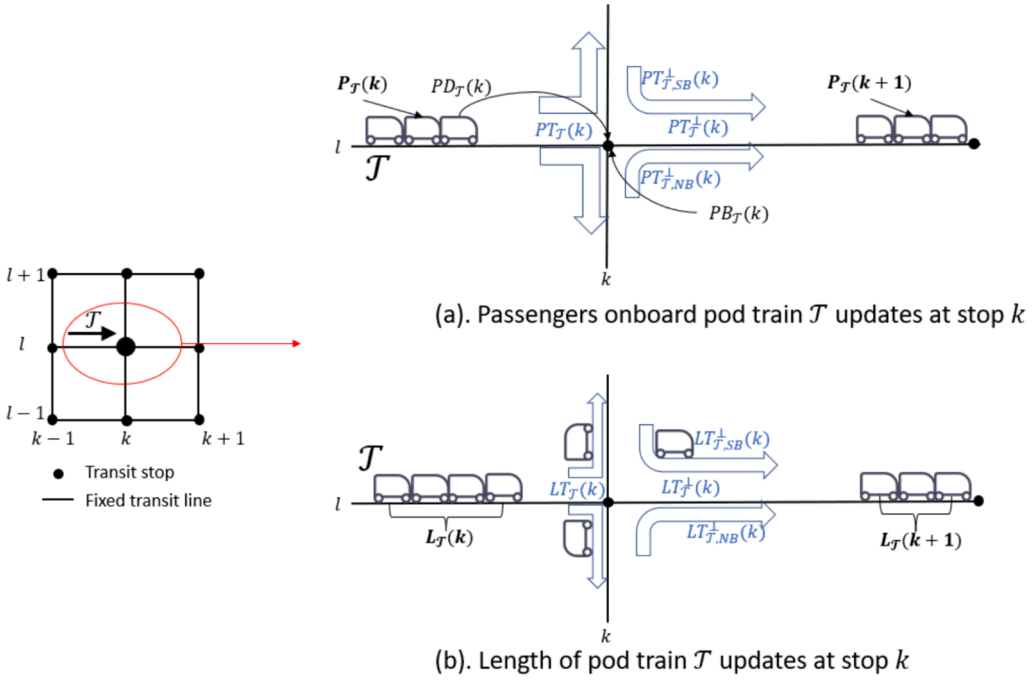


Fig. 5. Pod train $\mathcal{T}$ movement and passenger activities at stop.k

$$PB_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-k}{N-1} \tag{2-4}$$

$$PT^{\perp}_{\mathscr{T},NB}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{l-1}{N} \cdot \frac{N-k}{N-1} \tag{2-5}$$

$$PT^{\perp}_{\mathscr{T},SB}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-l}{N} \cdot \frac{N-k}{N-1} \tag{2-6}$$

$$PT^{\perp}_{\mathscr{T}}(k) = PT^{\perp}_{\mathscr{T},NB}(k) + PT^{\perp}_{\mathscr{T},SB}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-k}{N} \tag{2-7}$$

Derivations of Eqs. (2)–(1) through (2–7) are included in Appendix A. Substituting all the terms in Eq. (2)–(1) with Eqs (2)–(2)-(2–7), the number of onboard passengers arriving at stop k, $P_{\mathscr{T}}(k)$, is defined in Eq. (2)–(8). Derivation of Eq. (2)–(8) is presented in Appendix B.

$$P_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\left(\frac{1}{N}+\frac{1}{N-1}\right)\cdot(N+1-k)\cdot(k-1) \tag{2-8}$$

Proposition 1 In the case of spatially homogeneous demand, upon arrival at stop $\frac{N}{2}+1$, the number of passengers onboard a pod train reaches the maximum, $\frac{1}{8}H\lambda D^2\left(\frac{1}{N}+\frac{1}{N-1}\right)$.

Proof. As derived in Appendix B, $P_{\mathscr{T}}(k)$ is given by

$$P_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\left(\frac{1}{N}+\frac{1}{N-1}\right)\cdot(N+1-k)\cdot(k-1)$$

$$\leq \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\left(\frac{1}{N}+\frac{1}{N-1}\right)\left(\frac{N}{2}\right)^2 (*).$$

If and only if $N+1-k=k-1$ (i.e., $k=\frac{N}{2}+1$), Equality in (*) can be established and $P_{\mathscr{T}}(k)$ reaches the maximal value, $\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\left(\frac{1}{N}+\frac{1}{N-1}\right)\left(\frac{N}{2}\right)^2 = \frac{1}{8}H\lambda D^2\left(\frac{1}{N}+\frac{1}{N-1}\right)$.

Q.E.D.

Algorithm 1 computes the quantities of those onboard passenger compositions at stop k, $\forall k \in \{1,\cdots,N\}$ for spatially homogeneous demand. It is worth noting that the conceptual definition of $P_{\mathscr{T}}(k)$ in Eq. (2–1) also applies in the case of spatially heterogeneous demand.

Algorithm 1 Compute the number of onboard passengers $P_{\mathscr{T}}(k)$

Input: $P_{\mathscr{T}}(1)=0$, $PS_{\mathscr{T}}(1)=0$

for $k=1,2,\cdots,N$ **do**

{

Compute $PB_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{N-k}{N-1}$

Compute $PT^{\perp}_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{N-k}{N}$

Compute $PD_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N-1}$

Compute $PT_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N}$

Update $PS_{\mathscr{T}}(k) = P_{\mathscr{T}}(k) - PD_{\mathscr{T}}(k) - PT_{\mathscr{T}}(k)$

Update $P_{\mathscr{T}}(k+1) = PS_{\mathscr{T}}(k) + PB_{\mathscr{T}}(k) + PT^{\perp}_{\mathscr{T}}(k)$

}

Output: $P_{\mathscr{T}}(k), PS_{\mathscr{T}}(k), PT_{\mathscr{T}}(k), \forall k \in \{1,\cdots,N\}$

b. ***Numbers of en-route and manual transfers and pod train length upon arrival at a stop.***

The initial length of a pod train $\mathscr{T}$ departing the terminal (i.e., pod train length for the pod train $\mathscr{T}$ upon arrival at the first stop), $L_{\mathscr{T}}(1)$, is determined by the maximum number of onboard passengers of the pod train, $\max(P_{\mathscr{T}}(k))$ $(\forall k \in \{1,\cdots,N\})$, along the trip. At the same time, $L_{\mathscr{T}}(1)$ must not exceed the maximum pod train length L_m. Therefore, the initial pod train length for $\mathscr{T}$ at dispatching is expressed as

$$L_{\mathscr{T}}(1) = \min\left\{L_m, \max_{k,\forall k\in\{1,\cdots,N\}}\left\lceil\frac{P_{\mathscr{T}}(k)}{C_a}\right\rceil\right\} \tag{2-9}$$

where C_a is the pod seating capacity in number of passengers.

After initial dispatching, the pod train length changes as a result of en-route transfer activities (pod disjoining/joining). Taking a pod train $\mathscr{T}$ on the EB transit line as an example, Fig. 5(b) depicts how the pod train length varies at stop k, due to one or more of the following activities: (1) disjoined pod(s) to the perpendicular transit lines $LT_{\mathscr{T}}(k)$, and (2) joined pod(s) from the perpendicular transit lines $LT^{\perp}_{\mathscr{T}}(k)$, which are further divided into those pods disjoined from the SB transit line $LT^{\perp}_{\mathscr{T},SB}(k)$ and those from the NB transit line $LT^{\perp}_{\mathscr{T},NB}(k)$. Thus, the pod train length for $\mathscr{T}$ upon arriving at stop $k+1$ is updated as follows:

$$L_{\mathscr{T}}(k+1) = L_{\mathscr{T}}(k) - LT_{\mathscr{T}}(k) + + LT^{\perp}_{\mathscr{T}}(k) \tag{2-10}$$

$LT_{\mathscr{T}}(k)$ is a function of the number of en-route transfers of $\mathscr{T}$ to the perpendicular lines at stop k, denoted by $PET_{\mathscr{T}}(k)$. Algorithm 2 shows the procedure of determining $PET_{\mathscr{T}}(k)$. Note that $LT^{\perp}_{\mathscr{T}}(k)$ is similarly determined for the perpendicular transit line. This is how Algorithm 2 works. When a pod train is approaching a stop, if passengers request a transfer and there are sufficient pods available, the necessary number of pods will disjoin from the main pod train. This disjoint occurs only if there are enough pods remaining to accommodate the passengers still on board. The number of passengers taking en-route transfers is determined by either the total number of passengers transferring or by the pod capacity multiplied by the length of the disjoining pods. Subsequently, the length of the pod train is adjusted to reflect the change, taking into account both the disjoined pods and any new pods joining from perpendicular transit lines. According to Assumption 8, when a disjoint occurs, the separated pods will either accelerate to join a perpendicular pod train or wait for the next opportunity to join, thereby creating a new joint on the transit line. Then the pod train length can be updated according to Eq. (2)–(10).

Recall in Eq. (2)–(3), $PT(k)$ is the total number of transfers at stop k, which is decomposed into two terms: en-route transfer passengers $PET(k)$ and number of manual transfer passengers denoted by $PMT(k)$. Once $PET(k)$ is determined, $PMT(k)$ can be determined by $PT(k) - PET(k)$. The total number of passengers taking manual transfers in one hour is $\frac{4N}{H}\sum_{k=1}^{N} PMT(k)$. And the manual transfer fraction is computed by

$$MTF^{(P2)} = \frac{\frac{4N}{H}\sum_{k=1}^{N} PMT(k)}{\lambda D^2} \tag{2-11}$$

Lastly, a joint (or disjoint) event takes place when there is non-zero en-route transfer occurring at a stop, i.e., $TJ(k) = 1$ when $PET(k) > 0$, which is computed in Algorithm 2. So, the total number of joints is the sum of all $TJ(k)$'s, i.e.,

$$TJ^{(P2)} = \sum_k TJ(k) \tag{2-12}$$

Algorithm 2 Compute the number of en-route transfers $PET(k)$ and manual transfer passengers $PMT(k)$, and update length of pod train $L(k)$

Input: $P(k), PS(k), PT(k), \forall k \in N$, pod capacity C_a

compute $L(1) = \min\left\{L_m, \max_{k,\forall k\in\{1,\cdots,N\}}\left\lceil\frac{P(k)}{C_a}\right\rceil\right\}$
iterate $l^{th}(l \in N)$ WE pod train **do**
  for $k = 1, 2, \cdots, N$ **do**
    if $\left\lceil\frac{PT(k)}{C_a}\right\rceil \le L_l(k) - \left\lceil\frac{PS(k)}{C_a}\right\rceil$ **then**
      Set $PET_l(k) = PT(k)$
      Set $LT_l(k) = \left\lceil\frac{PT(k)}{C_a}\right\rceil$
    else then
      Set $PET_l(k) = C_a \cdot \left(L_{\mathscr{T}}(k) - \left\lceil\frac{PS(k)}{C_a}\right\rceil\right)$
      Set $LT_l(k) = L_l(k) - \left\lceil\frac{PS(k)}{C_a}\right\rceil$
    if $PET_l(k) > 0$ **then**
      Set $TJ(k) = 1$
    else then
      Set $TJ(k) = 0$
  Set $PMT_l(k) = PT(k) - PET_l(k)$
  Compute $LT^{\perp}_{\mathscr{T},NB}(k) = \left\lceil\frac{PT^{\perp}_{\mathscr{T},NB}(k)}{C_a}\right\rceil$
  Compute $LT^{\perp}_{\mathscr{T},SB}(k) = \left\lceil\frac{PT^{\perp}_{\mathscr{T},SB}(k)}{C_a}\right\rceil$
  Update $L_l(k+1) = L_l(k) - LT_l(k) + LT^{\perp}_{l,SB}(k) + LT^{\perp}_{l,NB}(k)$
Output: $PMT(k), LT_l(k), L_l(k), \forall l, k \in \{1, \cdots, N\}$

c. ***Energy saving by pod train formation***

In this study, energy saving rate η_{pl} is assumed a function of the pod train length (L). Since there is no real-world pod train energy saving data, we have adopted the bus platooning results in Sethuraman et al. (2019) as shown in Fig. 6, in which. Specifically, we apply the energy saving rate η_{pl} by number of buses in the platoon with a 1-meter intra-platoon distance (i.e., blue curve) to compute energy saving of a pod train over individual pods.

P2 model formulation

Finally, we have the formulation for P2 as follows.

$$(\mathbf{P2})\min z(N,H) = \pi_Q^{(P2)}Q + \pi_M^{(P2)}M^{(P2)} + W^{(P2)} + A + T^{(P2)} + \frac{\delta}{v_w}\cdot MTF^{(P2)} \text{ s.t., (2–13)}$$

$$L^{(P2)}(k) \le L_m, \forall k \in \{1,\cdots,N\} \tag{2-14}$$

$$N \in \left\{2,\cdots,\left\lfloor\frac{D}{s}\right\rfloor\right\} \tag{2-15}$$

$$H > 0 \tag{2-16}$$

where:

$$M^{(P2)} = \frac{Q}{v} + \frac{2\tau N^2}{H} + \tau'\left(1 + MTF^{(P2)}\right)\lambda D^2 + \tau_f \cdot TJ^{(P2)} \tag{2-17}$$

$$\overline{L}^{(P2)} = \frac{\sum_{k=1}^{N}\frac{D}{N}L^{(P2)}(k)}{D} = \frac{1}{N}\sum_{k=1}^{N}L^{(P2)}(k) \tag{2-18}$$

$$\pi_Q^{(P2)} = \overline{L}^{(P2)}\frac{\$_Q' + \$_E'*\left(1-\eta_{pl}\right)}{\lambda D^2\mu} \tag{2-19}$$

$$\pi_M^{(P2)} = \overline{L}^{(P2)}\frac{\$_C'}{\lambda D^2\mu} \tag{2-20}$$

$$W^{(P2)} = \frac{H}{3} + \frac{H}{2}\hat{A}\cdot MTF^{(P2)} \tag{2-21}$$

$$T^{(P2)} = \frac{0.68ND}{N+1}\times\frac{M^{(P2)}}{Q} \tag{2-22}$$

Compared with P1, we note the key differences in P2. Different from Eq. (1)–(7) in P1, Eq. (2)–(17) substitutes the expected number of transfers e_T with the manual transfer fraction $MTF^{(P2)}$ and includes time loss due to pod train joining actions. Since pod joining/disjoining is the key feature in AMPT, pod train length must be determined in P2. So are energy savings by pod train formation in Eqs (2)–(18) and (2–19) respectively.

3.2. Spatially heterogeneous demand density

3.2.1. Scenario 3 – Traditional transit service (P3)

In P3, demand density varies across the network. P3 is formulated using the continuum approximation (CA) method following Ouyang et al. (2014). The demand density function is assumed to be a continuously differentiable function. Similar to P1, the total cost

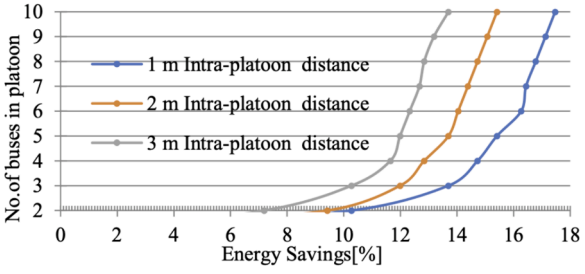


Fig. 6. Impact of Platooning on Energy Consumption (Driving-Cycle with Signal Priority). in Adapted from Fig. 8Sethuraman et al. (2019).

$z(l_0, H)$ is a function of transit line spacing l_0 and headway H, and consists of these seven components: [i] agency cost of delay caused by boarding and alighting, G, [ii] agency cost of total VKT and energy use, V, [iii] agency cost of total vehicular time expended and transit vehicle purchase, TE, [iv] passenger walking time, $A(x,y)$, [v] out-of-vehicle passenger waiting time including at transfers, $W(x,y)$, [vi] in-vehicle travel time, $IVTT(x,y)$, and [vii] extra penalties associated with transfers $D_{transfer}(x,y)$.

$$(\mathbf{P3})\min z(l_0,H) = \int_{y=0}^{D}\int_{x=0}^{D} z(x,y)dxdy = \int_{y=0}^{D}\int_{x=0}^{D}\left\{\frac{1}{\mu}[G+V+TE] + \left[A(x,y)+W(x,y)+IVTT(x,y)+\frac{\delta}{v_w}D_{transfer}(x,y)\right]\right\}dxdy \tag{3-1}$$

s.t.,

$$H > 0 \tag{3-2}$$

$$0 \le l_0 \le D \tag{3-3}$$

where:

$$G = \frac{4c_S}{Hl_0^2} \tag{3-4}$$

$$V = \frac{4(\$_Q + \$_E)}{Hl_0} \tag{3-5}$$

$$TE = \frac{4(\$_C)}{H}\frac{1}{l_0^2}\left(\frac{l_0}{v}+\tau\right) \tag{3-6}$$

$$A(x,y) = \frac{1}{4v_w}\cdot 2l_0\cdot\left(D_{start}(x,y)+D_{end}(x,y)\right) \tag{3-7}$$

$$W(x,y) = \frac{H}{2}\cdot\left(D_{start}(x,y)+D_{end}(x,y)\right) \tag{3-8}$$

$$IVTT(x,y) = \frac{1}{v}\left(F_W(x,y)+F_E(x,y)+F_S(x,y)+F_N(x,y)\right)+\tau\cdot\frac{F_E(x,y)+F_W(x,y)+F_N(x,y)+F_S(x,y)}{l_0} \tag{3-9}$$

According to Ouyang et al., (2014), the demand for bus travel from a zone that encompasses coordinates (x_1,y_1) to another zone that covers coordinates (x_2,y_2) has a density $\delta(x_1,y_1,x_2,y_2)$ that varies smoothly over all its arguments, in an unit of number of passengers per unit time per square area. We denote the expected rate per unit time per area that passengers start and end their trips at a stop located at (x,y) as $D_{start}(x,y)$ and $D_{end}(x,y)$. For a given network size of $D\times D$ (km^2), $D_{start}(x,y)$ and $D_{end}(x,y)$ are computed by

$$D_{start}(x,y) = \int_{\bar{y}=0}^{D}\int_{\bar{x}=0}^{D}\delta(x,y,\bar{x},\bar{y})d\bar{x}d\bar{y} \tag{3-10}$$

$$D_{end}(x,y) = \int_{\bar{y}=0}^{D}\int_{\bar{x}=0}^{D}\delta(\bar{x},\bar{y},x,y)d\bar{x}d\bar{y} \tag{3-11}$$

We denote the approximate expected rate per unit time per area that passengers perform directional transfers at a stop located at (x,y) as $D_{transfer}(x,y)$. $D_{transfer}(x,y)$ is computed by

$$D_{transfer}(x,y) = \frac{1}{2}\int_{\bar{y}=0}^{D}\int_{\bar{x}=0}^{D}\left[\delta(\bar{x},y,x,\bar{y})+\delta(x,\bar{y},\bar{x},y)\right]d\bar{x}d\bar{y} \tag{3-12}$$

Let $F_E(x,y)$, $F_W(x,y)$, $F_S(x,y)$ and $F_N(x,y)$ denote expected fluxes through (x,y) eastbound, westbound, southbound, and northbound. Using the eastbound as an example, $F_E(x,y)$ is calculated as the sum of two terms. The first term, $f_E^1(x,y)$, represents trips that start within a thin horizontal slice through (x,y) and are destined for the east side of the network. The second term, $f_E^2(x,y)$, represents trips that are transferred through (x,y) towards the eastbound. Therefore, $F_E(x,y)$ can be expressed as follows:

$$F_E(x,y) = f_E^1(x,y)+f_E^2(x,y) \tag{3-13}$$

where:

$$f_E^1(x,y) = \frac{1}{2}\int_{x_1=0}^{x}\left[\int_{x_2=x}^{D}\int_{y_2=0}^{D}\delta(x_1,y,x_2,y_2)dy_2dx_2\right]dx_1 \tag{3-14}$$

$$f_E^2(x,y) = \frac{1}{2}\int_{x_2=0}^{D}\left[\int_{x_1=0}^{x}\int_{y_1=0}^{D}\delta(x_1,y_1,x_2,y)dy_1dx_1\right]dx_2 \tag{3-15}$$

Owing to network symmetry, $F_W(x,y)$, $F_S(x,y)$, and $F_N(x,y)$ are derived in a similar fashion to $F_E(x,y)$. Please refer to Ouyang et al., (2014) for details of how Eqs. (3)–(10) through (3–15) are derived.

3.2.2. Scenario 4 – AMPT (P4)

Similarly, the total cost of P4, $z(l_0,H,L)$, is a function of transit line spacing l_0, headway H, and the length of pod train L.

Recall, in P2, we calculate the number of onboard passengers using Eqs (2)–(1) through (2–8). Similarly, we need to determine the number of onboard passengers in the context of a heterogeneous demand distribution for P4. We adopt a similar approach as in P2, with the key distinction that a continuously differentiable demand function is assumed in P4, rather than a uniform distribution in P2. The pod train and passenger activities at a stop depicted in Fig. 5 also apply to P4.

Following P3, the $D \times D$ (km^2) gridded network of Fig. 2 is divided into many sufficiently small square cells as shown in Fig. 7. Each square cell k is defined by its centroid (x_k,y_k). The demand density between an origin cell o and a destination cell d is denoted by $\delta(x_o,y_o,x_d,y_d)$.

We denote (X_i, Y_j) as the coordinates of a stop (i,j), which is shared by vertical transit line i and horizontal line j, respectively. The square zone served by stop (i,j) is denoted by s_{ij}. As seen in Fig. 7, the total number of passengers originated from s_{ij}, denoted by $PB(i,j)$, is the sum of demand in all square cells contained in s_{ij}, which is calculated by

$$PB(i,j) = \int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}} D_{start}(x,y)dxdy$$

Per Assumptions 2 and 3, half of those passengers take the horizontal lines EB/WB, and the other half board the vertical lines NB/SB.

Similar to P2, consider an EB pod train operating on the horizontal line j. The number of onboard passengers in the pod train approaching a transit stop (i,j) is denoted by $P_{EB}(i,j)$. The departing passengers from the pod train at stop (i,j) are divided into two groups as illustrated in Fig. 5(a): those who alight the transit service for good at stop (i,j), denoted by $PD_{EB}(i,j)$; and those who transfer at stop (i,j), denoted by $PT_{EB}(i,j)$, including both en-route and manual transfers. On the other hand, two types of passengers board the pod train at stop (i,j): those originated from s_{ij}, denoted by $PB_{EB}(i,j)$; and those transferred from the perpendicular lines, denoted by $PT_{EB}^{\perp}(i,j)$. Thus, the number of onboard passengers arriving at the next stop $(i+1,j)$ is computed in Eq. (4)–(1):

$$P_{EB}(i+1,j) = P_{EB}(i,j) - PD_{EB}(i,j) - PT_{EB}(i,j) + PB_{EB}(i,j) + PT_{EB}^{\perp}(i,j) \tag{4-1}$$

where,

$$PD_{EB}(i,j) = \int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\left(\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdy + \frac{1}{2}\int_{y=0}^{\frac{Y_{j-1}+Y_j}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdy + \frac{1}{2}\int_{y=\frac{Y_j+Y_{j+1}}{2}}^{D}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdy\right)d\bar{x}d\bar{y} \tag{4-2}$$

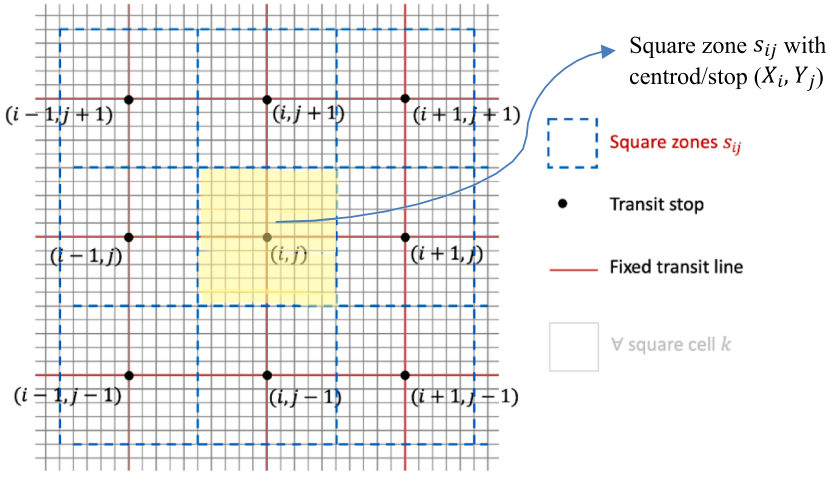


Fig. 7. Definition of square cells and square zones in a gridded square network for P4.

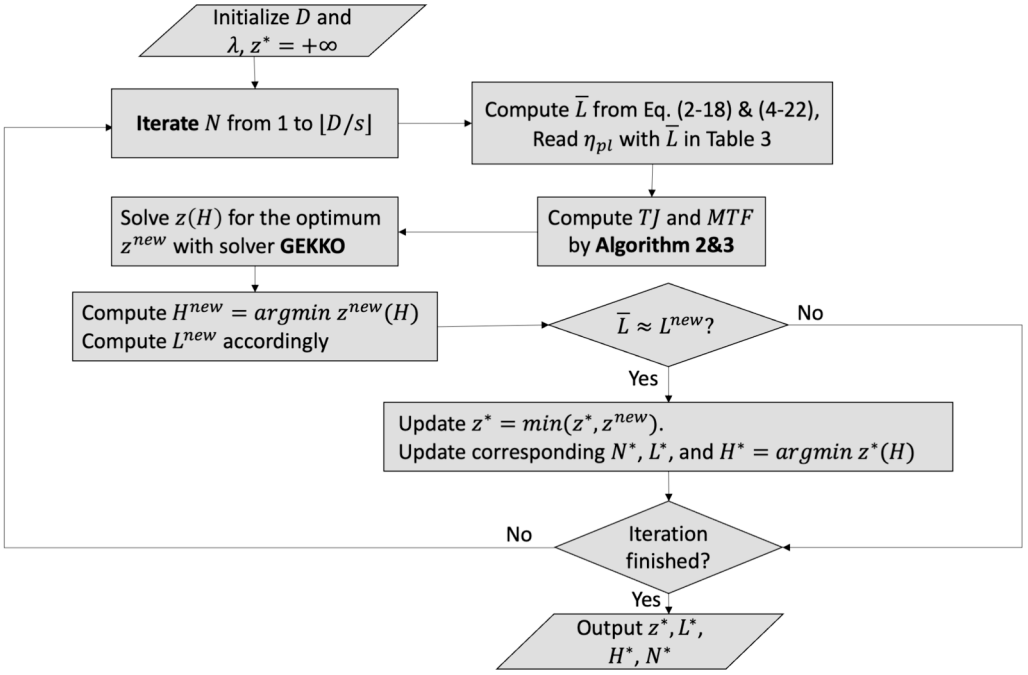


Fig. 8. Algorithm 4 flow chart.

$$PT_{EB}(i,j) = \frac{1}{2}\int_{\bar{y}=0}^{\frac{Y_{j-1}+Y_j}{2}}\int_{\bar{x}=\frac{X_i+X_{i+1}}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y} + \frac{1}{2}\int_{\bar{y}=\frac{Y_j+Y_{j+1}}{2}}^{D}\int_{\bar{x}=\frac{X_i+X_{i+1}}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}} \times \int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y} \tag{4-3}$$

$$PB_{EB}(i,j) = \frac{1}{2}\int_{\bar{y}=0}^{D}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{D}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y} \tag{4-4}$$

$$PT^{\perp}_{EB,NB}(i,j) = \frac{1}{2}\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=0}^{\frac{Y_{j-1}+Y_j}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y} \tag{4-5}$$

$$PT^{\perp}_{EB,SB}(i,j) = \frac{1}{2}\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_j+Y_{j+1}}{2}}^{D}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y} \tag{4-6}$$

$$\begin{aligned} PT^{\perp}_{EB}(i,j) &= PT^{\perp}_{EB,NB}(k) + PT^{\perp}_{EB,SB}(k) \\ &= \frac{1}{2}\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=0}^{\frac{Y_{j-1}+Y_j}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y} + \frac{1}{2}\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_j+Y_{j+1}}{2}}^{D} \\ &\quad \times \int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y} \end{aligned} \tag{4-7}$$

$P_{EB}(i,j)$ is the number of passengers onboard the EB pod train on line j upon arrival at stop (i,j). At the starting stop $(1,j)$ of line j, $P_{EB}(1,j) = 0$. In the subsequent stops, $P_{EB}(i+1,j)$ is updated as follows.

$$
\begin{aligned}
&P_{EB}(i+1,j)-P_{EB}(i,j)=PB_{EB}(i,j)+PT^{\perp}_{EB}(i,j)-PD_{EB}(i,j)-PT_{EB}(i,j)\\
&=\frac{1}{2}\int_{\bar{y}=0}^{D}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{D}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}-\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\left(\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdy\right)d\bar{x}d\bar{y}\\
&-\left(\frac{1}{2}\int_{\bar{y}=0}^{\frac{Y_{j-1}+Y_j}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}+\frac{1}{2}\int_{\bar{y}=\frac{Y_j+Y_{j+1}}{2}}^{D}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}\right)\\
&=\frac{1}{2}\int_{\bar{y}=0}^{D}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{D}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}-\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}\\
&-\frac{1}{2}\left(\int_{\bar{y}=0}^{D}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}-\frac{1}{2}\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}\right)\\
&=\frac{1}{2}\int_{\bar{y}=0}^{D}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{D}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=\frac{X_i+X_{i+1}}{2}}^{\frac{X_i+X_{i+1}}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}-\frac{1}{2}\int_{\bar{y}=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}\\
&-\frac{1}{2}\int_{\bar{y}=0}^{D}\int_{\bar{x}=\frac{X_{i-1}+X_i}{2}}^{\frac{X_i+X_{i+1}}{2}}\int_{y=\frac{Y_{j-1}+Y_j}{2}}^{\frac{Y_j+Y_{j+1}}{2}}\int_{x=0}^{\frac{X_{i-1}+X_i}{2}}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}
\end{aligned}
\tag{4-8}
$$

$PT_{EB}(i,j)$ includes both en-route and manual transfers, denoted by $PET^{(P4)}_{EB}(i,j)$ and $PMT^{(P4)}_{EB}(i,j)$, respectively. Then, the total number of passengers taking manual transfers within one hour on the entire network is summarized as $\frac{1}{H}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(PMT^{(P4)}_{EB}(i,j)+PMT^{(P4)}_{WB}(i,j)+PMT^{(P4)}_{NB}(i,j)+PMT^{(P4)}_{SB}(i,j)\right)$. Thus, the manual transfer fraction in P4 is given by

$$
MTF^{(P4)}=\frac{\frac{1}{H}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(PMT^{(P4)}_{EB}(i,j)+PMT^{(P4)}_{WB}(i,j)+PMT^{(P4)}_{NB}(i,j)+PMT^{(P4)}_{SB}(i,j)\right)}{\int_{\bar{y}=0}^{D}\int_{\bar{x}=0}^{D}\int_{y=0}^{D}\int_{x=0}^{D}\delta(x,y,\bar{x},\bar{y})dxdyd\bar{x}d\bar{y}}
\tag{4-9}
$$

Lastly, a joint (or disjoint) event takes place when there is non-zero en-route transfer occurring at a stop (i,j), i.e., $TJ(i,j)=1$ when $PET(i,j)>0$, which is computed in Algorithm 3. So, the total number of joints is the sum of all $TJ(k)$'s, i.e.,

$$
TJ^{(P4)}=\sum\nolimits_k TJ(i,j)
\tag{4-10}
$$

Algorithm 3 works similarly to algorithm 2 in computing the number of manual transfers $PMT^{(P4)}_{EB}(i,j)$, and updating the length of pod train $L^{(P4)}_j(i)$ after stop (i,j) for P4.

Algorithm 3 Compute the number of manual transfer passengers $PMT^{(P4)}_{EB}(i,j)$, and update length of pod train $L^{(P4)}_j(i)$

Input: $P_{EB}(i,j), PD_{EB}(i,j), PT_{EB}(i,j), \forall k \in N$, pod capacity C_a

iterate $j^{th}(j \in \{1,2,N\})$ EB pod train **do**
  compute $L^{(P4)}_j(1)=\min\left\{L_m, \max_{i,\forall i\in\{1,2,\cdots,N\}}\left\lceil\frac{P_{EB}(i,j)}{C_a}\right\rceil\right\}$
  for $i=1,2,\cdots,N$ **do**
  {
    $PS_{EB}(i,j)=P_{EB}(i,j)-PD_{EB}(i,j)-PT_{EB}(i,j)$
    if $\left\lceil\frac{PT_{EB}(i,j)}{C_a}\right\rceil \le L^{(P4)}_j(i)-\left\lceil\frac{PS_{EB}(i,j)}{C_a}\right\rceil$ **then**
      Set $PET^{(P4)}_{EB}(i,j)=PT_{EB}(i,j)$
      Set $LT^{(P4)}_j(i)=\left\lceil\frac{PT_{EB}(i,j)}{C_a}\right\rceil$
    else then
      Set $PET^{(P4)}_{EB}(i,j)=C_a\cdot\left(L^{(P4)}_j(i)-\left\lceil\frac{PS_{EB}(i,j)}{C_a}\right\rceil\right)$
      Set $LT^{(P4)}_j(i)=L^{(P4)}_j(i)-\left\lceil\frac{PS_{EB}(i,j)}{C_a}\right\rceil$
    if $PET^{(P4)}_{EB}(i,j)>0$ **then**
      Set $TJ(i,j)=1$
    else then
      Set $TJ(i,j)=0$
    Set $PMT^{(P4)}_{EB}(i,j)=PT_{EB}(i,j)-PET^{(P4)}_{EB}(i,j)$
    Compute $LT^{\perp}_{j,NB}(i)=\left\lceil\frac{PT^{\perp}_{EB,NB}(i,j)}{C_a}\right\rceil$

(*continued on next page*)

(*continued*)

Algorithm 3 Compute the number of manual transfer passengers $PMT_{EB}^{(P4)}(i,j)$, and update length of pod train $L_j^{(P4)}(i)$

18: Compute $LT_{j,SB}^{\perp}(i) = \left\lceil \frac{PT_{EB,SB}^{\perp}(i,j)}{C_a} \right\rceil$

19: Update $L_j^{(P4)}(i+1) = L_j^{(P4)}(i) - LT_j^{(P4)}(i) + LT_{j,NB}^{\perp}(i) + LT_{j,SB}^{\perp}(i)$

}20:

Output: $PMT_{EB}^{(P4)}(i,j), L_j^{(P4)}(i), \forall i,j \in \{1,\cdots,N\}$

Given the number of the transit lines and the network size, the coordinates of stop (i,j) are $X_i = \frac{D}{2N} + \frac{D}{N}(i-1)$ and $Y_j = \frac{D}{2N} + \frac{D}{N}(j-1)$. For an EB pod train on transit line j, the right-hand side of Eq. (4)–(8) is a function of i (j is fixed). We denote that by $\mathscr{T}(i)$. That is

$$P_{EB}(i+1,j) - P_{EB}(i,j) = \mathscr{T}(i) \tag{4-11}$$

and,

$$\begin{aligned} P_{EB}(i,j) - P_{EB}(i-1,j) &= \mathscr{T}(i-1) \\ P_{EB}(i-1,j) - P_{EB}(i-2,j) &= \mathscr{T}(i-2) \\ P_{EB}(i-2,j) - P_{EB}(i-3,j) &= \mathscr{T}(i-3) \\ &\vdots \\ P_{EB}(3,j) - P_{EB}(2,j) &= \mathscr{T}(2) \\ P_{EB}(2,j) - P_{EB}(1,j) &= \mathscr{T}(1) \end{aligned} \tag{4-12}$$

Summing up the series of equations in Eqs (4)–(11) and (4–12) above, we have $P_{EB}(i,j) - P_{EB}(1,j)$ on the left-hand side. Since $P_{EB}(1,j) = 0$, it leaves only $P_{EB}(i,j)$. And

$$P_{EB}(i,j) = \sum_{i'=1}^{i-1} \mathscr{T}(i') \tag{4-13}$$

The final model formulation of P4 is presented as follows.

$$\begin{aligned} (\mathbf{P4})\min z(l_0,H,L) &= \int_{y=0}^{D}\int_{x=0}^{D} z(x,y)dxdy \\ &= \int_{y=0}^{D}\int_{x=0}^{D} \left\{ \frac{1}{\mu}\left[G^{(P4)} + V^{(P4)} + TE^{(P4)}\right] + \left[A(x,y) + W(x,y) + IVTT(x,y) + \frac{\delta}{v_w} D_{transfer}(x,y)\cdot MTF^{(P4)}\right] \right\} dxdy \end{aligned} \tag{4-14}$$

s.t.,

$$\overline{L}^{(P4)} \cdot C_a \cdot \frac{TE^{(P4)}}{\$_c} \cdot D^2 \geq \int_{y=0}^{D}\int_{x=0}^{D} D_{start}(x,y) \tag{4-15}$$

$$\overline{L}^{(P4)} \leq L_m \tag{4-16}$$

$$0 \leq l_0 \leq D \tag{4-17}$$

$$H > 0 \tag{4-18}$$

where:

$$G^{(P4)} = \frac{4c_S}{Hl_0^2} \tag{4-19}$$

$$V^{(P4)} = \frac{4\overline{L}^{(P4)}\left(\$'_Q + \$'_E * (1-\eta_{pl})\right)}{Hl_0} \tag{4-20}$$

$$TE^{(P4)} = \frac{4\overline{L}^{(P4)}(\$'_C)}{H}\frac{1}{l_0^2}\left(\frac{l_0}{v} + \tau\cdot MTF^{(P4)} + \tau_f \cdot TJ^{(P4)}\right) \tag{4-21}$$

$$\overline{L}^{(P4)} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} L_j^{(P4)}(i) \tag{4-22}$$

$L_j^{(P4)}(i)$ is computed in Algorithm 3, and

$MTF^{(P4)}$ and $TJ^{(P4)}$ are defined in Eqs (4)–(9) and (4–10), respectively.

Lastly, it is worth noting that the mathematical formulation of P1 and P3, as well as P2 and P4, can be unified under the CA framework. The homogeneous and heterogeneous models consist of the same cost components. If the demand function reduces to a uniform value (homogeneous), the integrals in the CA models (P3 and P4) reduce to those in P1 and P3, respectively. Thus, P1 and P3 (and P2 and P4) are essentially of the same mathematical formulation. They are presented separately for the purpose of demonstrating the methodological continuation between P1 and P2, and between P3 and P4.

4. Solution methods

This section presents the computational implementation of the mathematical models proposed in Section 3. All models are mixed integer nonlinear programming optimization problems. In P2, the decision variable N is an integer; the decision variable H is a continuous variable. The pod train length may vary at stops. Furthermore, the pod train energy saving rate η_{pl} depends on the pod train length. All these increase the complexity in solving P2. P4 faces similar complexity if not more.

Algorithm 4 is devised to solve both P2 and P4, as illustrated by the flow chart in Fig. 8. Note that $\bar{L}$ is computed by Eq. (2–18) for P2 and Eq. (4–15) for P4, and its value is bounded by L_m in the model constraint (2–14) and (4–16), respectively. The value of η_{pl} is taken from Table 2 (Sethuraman et al., 2019) based on the value of $\bar{L}$ in each iteration. The pseudo code of Algorithm 4 is shown below.

P1 and P3 can be solved readily and quickly with an existing solver. In this study, we use the solver GEKKO (Beal et al., 2018) with its default settings as the optimizer to solve P1 and P3.

All computations were performed on a 3.1 GHz Dual-Core Intel Core i5.

Algorithm 4 Solution method for P2 & P4

Input: Network size D, demand density λ in P2, demand density function δ in P4

Initialize $z^* = +\text{Infinity}$
Compute $\bar{L}$ from Eq. (2)–(18) and (4–22)
Read η_{pl} in Table 2 accordingly.
for $N = 1, 2, \cdots, \lfloor D/s \rfloor$ **do**
{
Compute TJ^{P2orP4} in Eq. (2)–(12) or Eq. (4)–(10)
Compute MTF^{P2orP4} by Algorithm 2 or 3
Identify one variable objective function $z(H)$
Identify feasible set of H, $\phi(H) = \{H > 0; \max_{k,\forall k \in \{1,\cdots,N\}} \left\lceil \frac{P(k)}{C_a} \right\rceil \leq \bar{L}, \forall k \in N\}$
Solve the optimization problem for optimal value $z^{new}(H)$ with solver GEKKO:
minimize : $z(H)$
subject to : $\phi(H)$
Compute $H^{new} = argminz^{new}(H)$
Compute L^{new} in Eq. (2)–(18) or Eq. (4)–(22) accordingly
if $|\bar{L} - L^{new}| \leq \frac{1}{2}$ **then**
Update $z^* = min(z^*, z^{new})$.
Update corresponding N^*, L^*, and $H^* = argminz^*(H)$
}
Output: z^* and corresponding L^*, H^*, N^*.

The problems can also be solved in the genetic algorithm (GA) used in Chen and Nie (2017). To evaluate the performance of our proposed Algorithm 4, we compare it with the GA with solving P2A. Fig. 9 plots the running time of the two methods when $D = 20$ km under different demand densities (λ). 100 randomly generated instances were run for each demand scenario, and the average running time was recorded. Fig. 9 demonstrates that Algorithm 4 has much lower computation time than GA, while both provide the same final

Table 2
Corresponding values in Fig. 6 (1 m intra-platoon distance).

No. buses in platooning	Energy savings (%) Driving-Cycle with Signal Priority
1	0
2	10.3
3	13.7
4	14.8
5	15.4
6	16.3
7	16.5
8	16.8
9	17.2
10	17.4

solutions in each scenario.

5. Numerical experiments

With the solution methods presented in Section 4 and the parameter values in Table 3, a number of numerical experiments of P1 – P4 are performed. The results are presented as follows.

In the scenario of heterogeneous demand distributions (i.e., P3 and P4), we consider a mono-centric distribution with the highest demand density at the center of the network, as shown in Fig. 10. The demand density function is assumed to be a continuously differentiable function.

The demand density function in P3 and P4 is adopted from Ouyang et al. (2014) as below.

$$\delta(x_1, y_1, x_2, y_2) = \prod_{i=1}^{2} \left(a_1 + a_2 \sum_{j=1}^{2} exp\left[-\left(a_{3j}x_i - b_{ij}\right)^2 - \left(a_{4j}y_i - \bar{b}_{ij}\right)^2 \right] \right) \tag{5-1}$$

where parameters are defined in Table 4 also adopted from Ouyang et al. (2014). Parameters in Table 4 generates a mono-centric distribution with the highest demand density at the center of the network, as shown in Fig. 10.

5.1. Optimal cost comparisons

Fig. 11 shows that P2 (AMPT) incurs cost saving over P1 across most of the combinations of demand density and network size. In particular, when the demand density is low and the network size is small, P2 is as much as 40 % more cost efficient, which makes intuitive sense since P2 fleet capacity adapts to the demand. Furthermore, Fig. 11 indicates that P2 is generally more cost efficient than the traditional transit service at low demand density, regardless of the network size. Such findings suggest that AMPT be a cost-efficient alternative to the existing transit service in suburban and rural towns and villages where transit demand is low. When the demand density is high and the network is large, AMPT is no longer cost competitive over the traditional transit service considered in the study.

Fig. 12(a) and 12b compares P3 and P4 with varying network size. Fig. 12(a) shows that P4 outperforms P3 in terms of total cost by a margin of 3 % to 12 %. When the network size is smaller, P4 saves more cost from P3 percentage wise. As the region size grows, P4′s saving decreases percentage wise but the absolute saving increases from P3. Fig. 12(b) implies that P4 has considerable savings from P3 in terms of electricity energy cost (around 60 %), due to platooning.

5.2. Cost breakdown

We conducted a thorough analysis of the cost components for P2 in comparison to P1 to identify potential savings. This analysis categorized the costs of P1 and P2 into various categories, including agency capital/operating costs, transfer savings, walking time, and in-vehicle travel time. Our study primarily focused on agency capital/operating costs and transfer savings.

Detailed breakdowns of the cost saving results of Fig. 11 are provided in Figs. 13(a) and (b), highlighting the differences in costs between P1 and P2 in terms of percentage of total difference relative to P1. Fig. 13(a) demonstrates that, at lower demand densities, AMPT's agency costs are more cost-effective than those of the conventional transit service, P1. However, as demand increases, AMPT's agency costs start to exceed those of P1, primarily due to the requirement for a larger fleet of pods. Fig. 13(b) emphasizes the significant cost savings in transfers offered by AMPT compared to traditional transit systems. Notably, as demand increases, transfer savings become increasingly substantial in the overall cost difference between AMPT and P1. Taking both Figs. 13(a) and (b) into account, it

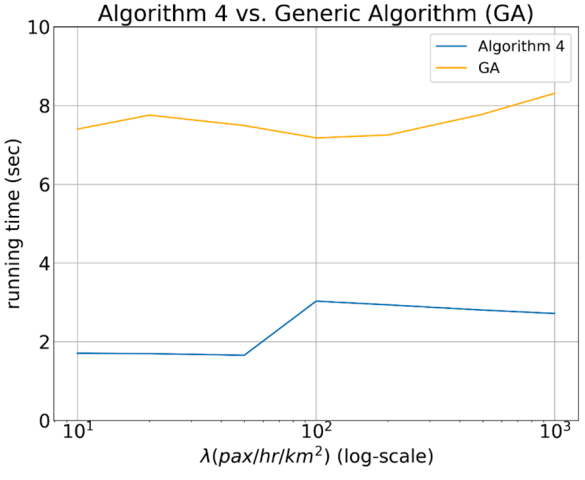


Fig. 9. Running time comparison of Algorithm 4 and GA by solving P2.

Table 3
Fixed values of input parameters.

Notation	Value	Source
λ(pax/hr./km^2)	0 to 1000	
D(km)	2 to 24	
s(km)	0.15	
μ($/hr.)	20	
τ(s)	12	
τ'(s)	1	Chen and Nie (2017)
τ_f(s)	30	
v(km/hr.)	25	
v_w(km/hr.)	2	
δ(km)	0.03	
c_s($/bus-stop)	0.3	Ouyang et al. (2014)
$\$_Q$($/veh • km)	0.342	publicpower.org
$\$_M$($/veh • hr.)	0	–
$\$'_Q$($/pod • km)	0.038	Department of Energy (2021)
$\$'_M$($/pod • hr.)	0	–
$\$_E$($/veh • km)	0.174	publicpower.org
$\$_C$($/veh • hr.)	6.976	Aber (2016)
$\$'_E$($/pod • km)	0.0308	energysage.com
$\$'_C$($/pod • hr.)	0.507	NYC Transit
C_a(pax)	6	next-future-mobility.com
C_a^b(pax)	35	US Department of Transportation (2023)

becomes apparent that while transfer savings play a crucial role in reducing overall costs, the impact of agency costs in AMPT is mixed. It sometimes provides savings and other times leads to higher expenses compared to P1.

5.3. Load factor analysis

In terms of load factor, Figs. 14(a) and 14(b) shows improvement in AMPT (P2 and P4) over the traditional system (P1 and P3). Fig. 14(a) implies that the load factors of P2 outperform those of P1 in all cases. Overall, with an increasing network size, the relative difference of load factors (P2-P1)/P1 increases. Interestingly, however, in the high demand cases (i.e., $\lambda = 500$ and 1000 (pax/hr./sq. km)), the difference in load factors (P2-P1)/P1 increases initially and drops relatively quickly close to zero when the network grows. This finding indicates some trade-off between load factor and network size in AMPT that requires further investigation. Nonetheless, it

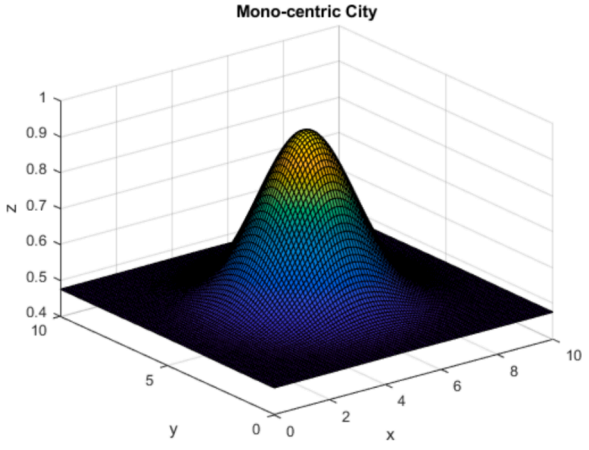


Fig. 10. Mono-centric demand density distribution.

Table 4
Parameter values of the demand density function (Ouyang et al. (2014)).

Parameter	a_1	a_2	a_{31}	a_{32}	a_{41}	a_{42}	b_{11}	b_{12}	b_{21}	b_{22}	$\bar{b}_{11}$	$\bar{b}_{12}$	$\bar{b}_{21}$	$\bar{b}_{22}$
Value	0.0016	0.065	0.5	0	0.5	0	2.5	0	2.5	0	2.5	0	2.5	0

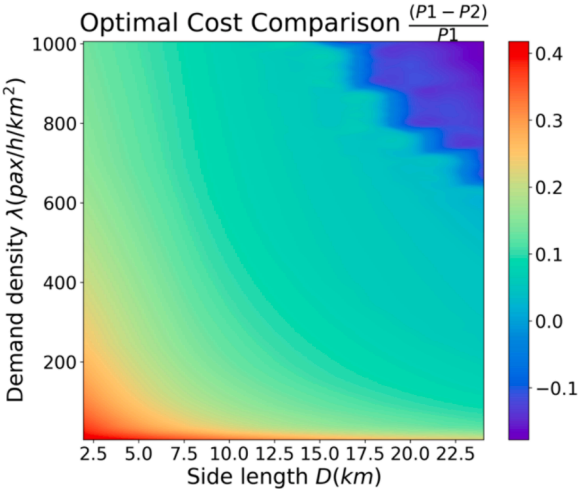


Fig. 11. Optimal cost comparison between P1 and P2.

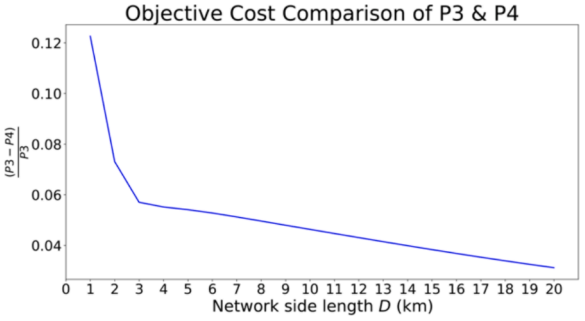


Fig. 12a. Total cost comparison of P3 and P4.

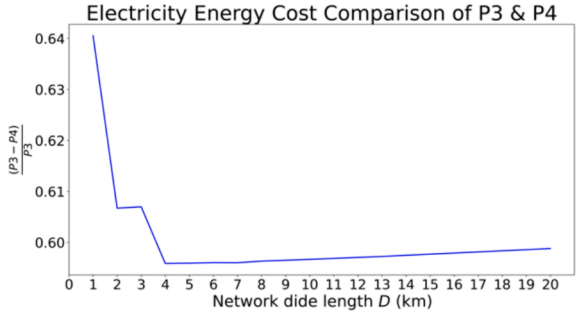


Fig. 12b. Energy cost comparison of P3 & P4.

suggests that AMPT's performance converges to the traditional fixed line transit service in large networks when the demand density is high. In Fig. 14(b), the relative improvement of load factor of P4 from P3 initially decreases rapidly from over 30 % to around 20 % as the network size goes up and later stabilizes at around 20 %. This suggests that the demand-adaptive pod trains maintain high utilization of the pod capacity.

5.4. Sensitivity analysis

The numerical experiment results presented above clearly depend on the parameter values in the models. To provide a fuller evaluation of the models, we conduct a series of sensitivity analysis to explore how some key model parameters may affect the model behavior. As already presented in Section 3, there are numerous model parameters involved in AMPT service design. In the section, we

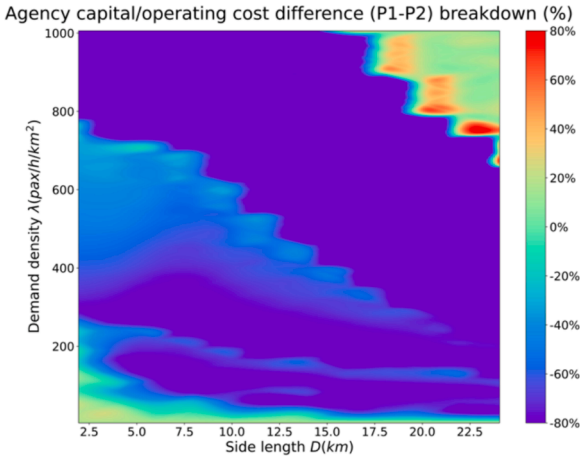


Fig. 13a. Agency cost difference by demand density and network size.

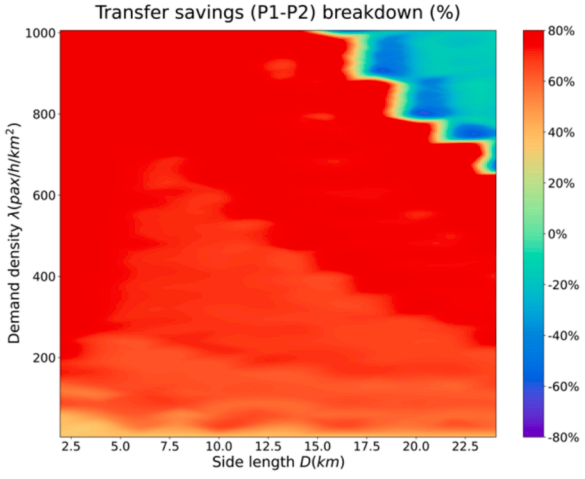


Fig. 13b. Passenger transfer saving by demand density and network size.

focus on these three model parameters, namely the pod joining delay (τ_f), pod energy cost ($\$'_E$), and pod capital cost ($\$'_C$). In the sensitivity analysis, we fix the network size $D = 5$ km and demand density $\lambda = 100$ (pax/hr./km^2), which correspond to the lower left area of Fig. 13(a) where P2A demonstrates noticeable savings over P1.

As expected, if pod joining delay τ_f increases, the cost saving by P2 decreases. Fig. 15 shows the range of decrease in % cost saving when τ_f increases from 0 to 120 s. It is a linear decreasing function from (2–12) and (2–16). In the numerical experiments presented in Section 5, τ_f was set at 30 s; the corresponding cost saving is around 25 %. When τ_f is 120 s, the cost saving is reduced to less than 21 %. A similar trend is observed in pod energy cost $\$'_E$ ($/pod/km) with an almost linear decreasing trend in Fig. 16. In the above numerical analysis, $\$'_E$ was set at $0.031/pod/km, which corresponds to a roughly 24.5 % cost saving by AMPT. When the cost goes up to $0.175/pod/km equivalent to that of traditional bus, the cost saving by AMPT drops to about 18 %. On the other hand, if the pod energy efficiency is improved from the current assumed value, the cost saving by AMPT could be as high as 26 %.

Lastly, the pod capital cost ($\$'_C$) has a non-linear monotonic decreasing effect on the % cost saving by AMPT in Fig. 17. In particular, when its value exceeds $4.5/pod/hr (in the scenario of $D = 5$ km and $\lambda = 100$ pax/hr./km^2), AMPT ceases to offer cost saving over traditional bus service. That value is about 64.5 % of the traditional bus capital cost in the analysis ($6.976/veh/hr).

6. Conclusions

In this study, we formulated and investigated a new AMVT-based fixed-route transit paradigm called AMPT. Because of modularity, AMPT allows varying vehicle capacity in operation and en-route transfers through pod joining and disjoining while in motion. These are unique features of AMPT, departing from the traditional transit service and making the model formulation distinctively more complex. We presented the model formulations for both uniform and heterogeneous demand density distributions and described our

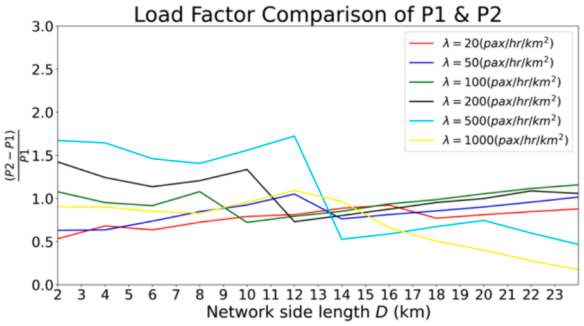


Fig. 14a. Load Factor Improvement (P2&P1).

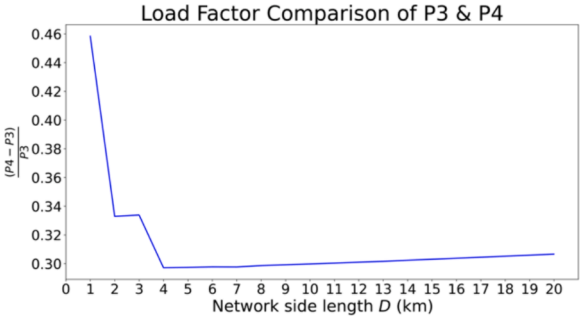


Fig. 14b. Load Factor Comparison (P3&P4).

proposed solution methods. Overall, numerical experiment results show that, if designed properly, AMPT brings a range of operational benefits to a grid transit network compared to the traditional fixed-vehicle size transit system under both spatially homogeneous and heterogeneous demand distribution. Notably, due to its modularity, AMPT brings more significant cost savings and capacity utilization than traditional transit services in low demand density. Most of the cost savings by AMPT are due to passenger time reduction by en-route transfer. Additional cost savings also incur with pod train formation (platooning). But overall the agency cost of AMPT has a mixed picture.

An important limitation of the study is that the problems considered are static, including the passenger demand and the supply of pod trains. Next step will consider dynamic AMPT service. Furthermore, other common network shapes such as zone-based and ring-radial networks for AMPT will be studied. In addition to the fixed-route transit system, AMVT could be viewed as a viable alternative to taxis (or vehicles of transportation network companies), which makes AMVT pods suitable for demand responsive mobility services. This is an on-going investigation, while preliminary analysis of a small network shows that pod train formation and en-route passenger consolidation in a MOD ridesharing service contribute marginally to energy saving (Cheng and Lin, 2024). Further investigation in a large network setting is warranted.

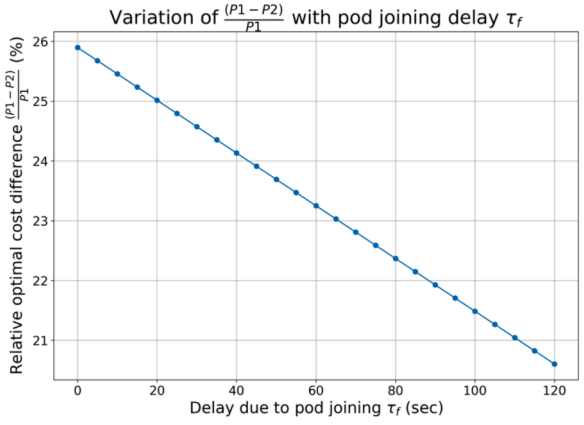


Fig. 15. Effect of pod joining delay τ_f on relative cost saving $\frac{P1-P2}{P1}$ (%).

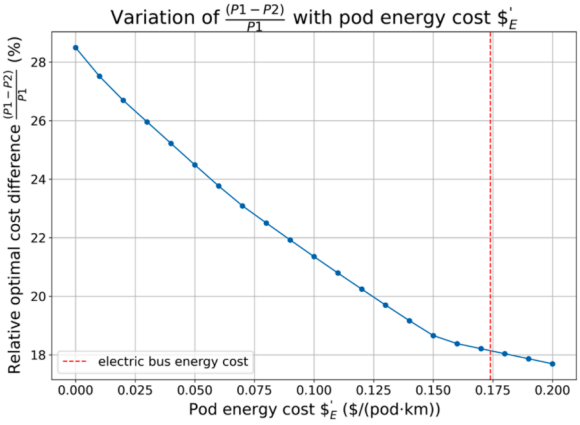


Fig. 16. Effect of pod energy cost $\$'_E$ on relative cost saving $\frac{P1-P2}{P1}$ (%).

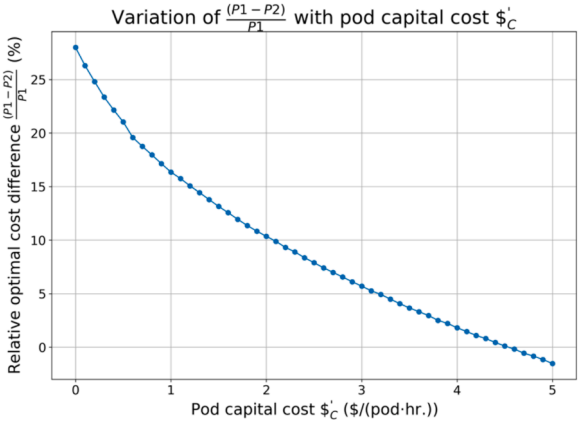


Fig. 17. Effect of pod capital cost $\$'_C$ on relative cost saving $\frac{P1-P2}{P1}$ (%).

CRediT authorship contribution statement

Xi Cheng: Writing – original draft, Formal analysis. **Yu (Marco) Nie:** . **Jane Lin:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data availability

All data used are also described in the paper.

Acknowledgement

This research is funded by the United States National Science Foundation (award # CMMI 2127677).

Appendix A

In the case of spatially homogeneous demand, it implies that the arrival rate of passengers (λ) at each stop is uniform. Within one hour, the number of passengers generated from and to one square in Fig. 4 is $\lambda\left(\frac{D}{N}\right)^2$. So, the number of passengers arriving and waiting at a stop for all four directions between two successive pod trains is $H\lambda\left(\frac{D}{N}\right)^2$, half of which board the horizontal lines EB and WB and the other half board the vertical lines SB and NB.

Consider a WB pod train $\mathscr{T}$. Upon arrival at a transit stop k ($k \in \{1, \cdots, N\}$), $\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2$ passengers originated in the square zone

served by stop k, wait to board the horizontal lines EB and WB. $\frac{k-1}{N-1}$ of these passengers would board EB pod train, while $\frac{N-k}{N-1}$ of them would board WB pod train. Thus, $PB_{\mathscr{T}}(k)$ is given by

$$PB_{\mathscr{T}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-k}{N-1} \tag{A-1}$$

According to Fig. 5(a) and $PB_{\mathscr{T}}(k)$ above, $PT^{\perp}_{\mathscr{T},NB}(k)$ are comprised of passengers onboard NB pod train before vertical stop l who are destinated on horizontal line l after stop k. Thus, $PT^{\perp}_{\mathscr{T},NB}(k)$ reads as

$$PT^{\perp}_{\mathscr{T},NB}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \sum_{l'=1}^{l-1}\frac{N-l'}{N-1}\cdot\frac{1}{N-l'}\cdot\frac{N-k}{N-1} = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-k}{N-1}\cdot\sum_{l'=1}^{l-1}\frac{1}{N-1} = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{l-1}{N}\cdot\frac{N-k}{N-1} \tag{A-2}$$

Similarly, $PT^{\perp}_{\mathscr{T},SB}(k)$ is computed as

$$PT^{\perp}_{\mathscr{T},SB}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \sum_{l'=1}^{N-l}\frac{N-l'}{N-1}\cdot\frac{1}{N-l'}\cdot\frac{N-k}{N-1} = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-k}{N-1}\cdot\sum_{l'=1}^{N-l}\frac{1}{N-1} = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-l}{N}\cdot\frac{N-k}{N-1} \tag{A-3}$$

Thus, $PT^{\perp}_{\mathscr{T}}(k)$, a summation of $PT^{\perp}_{\mathscr{T},NB}(k)$ and $PT^{\perp}_{\mathscr{T},SB}(k)$, is given by

$$PT^{\perp}_{\mathscr{T}}(k) = PT^{\perp}_{\mathscr{T},NB}(k) + PT^{\perp}_{\mathscr{T},SB}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \frac{N-k}{N} \tag{A-4}$$

$PB_{\mathscr{T}}(k)$ and $PT^{\perp}_{\mathscr{T}}(k)$ are passengers who are boarding at stop k for pod train $\mathscr{T}$. Now we derive the number of passengers who are leaving pod train $\mathscr{T}$, which is composed of two groups of passengers, those destinated at stop k, $PD_{\mathscr{T}}(k)$, and those transferring to other transit lines at stop k, $PT_{\mathscr{T}}(k)$.

$PD_{\mathscr{T}}(k)$ is made up of passengers boarding on pod train $\mathscr{T}$ before stop k, $PB_{\mathscr{T}}(k)$, and passengers transferred from vertical lines before stop k, $PT^{\perp}_{\mathscr{T}}(k)$. Thus, $PD_{\mathscr{T}}(k)$ is derived as

$$\begin{aligned}
PD_{\mathscr{T}}(k) &= \sum_{k'=1}^{k-1} PB_{\mathscr{T}}(k')\cdot\frac{1}{N-k'}\cdot\frac{1}{N} + \sum_{k'=1}^{k-1} PT^{\perp}_{\mathscr{T}}(k')\cdot\frac{1}{N-k'} \\
&= \sum_{k'=1}^{k-1}\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{N-k'}{N-1}\cdot\frac{1}{N-k'}\cdot\frac{1}{N} + \sum_{k'=1}^{k-1}\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{N-k'}{N}\cdot\frac{1}{N-k'} \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N}\cdot\frac{1}{N-1} + \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N} \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot(k-1)\cdot\left(\frac{1}{N\cdot(N-1)} + \frac{1}{N}\right) \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot(k-1)\cdot\left(\frac{1}{N-1} - \frac{1}{N} + \frac{1}{N}\right) \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N-1}
\end{aligned} \tag{A-5}$$

According to Assumption 2, passengers take at most one transfer. Thus, passengers transferring to other transit lines at stop k, $PT_{\mathscr{T}}(k)$, are exclusively composed of those who boarded $\mathscr{T}$ before stop k. Thus, $PT_{\mathscr{T}}(k)$ reads

$$PT_{\mathscr{T}}(k) = \sum_{1}^{k-1}\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{N-k}{N-1}\cdot\frac{1}{N-k}\cdot\frac{N-1}{N} = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2\cdot\frac{k-1}{N} \tag{A-6}$$

Therefore, passengers onboard pod train $\mathscr{T}$ is updated as

$$P_{\mathscr{T}}(k+1) = P_{\mathscr{T}}(k) - PD_{\mathscr{T}}(k) - PT_{\mathscr{T}}(k) + PB_{\mathscr{T}}(k) + PT^{\perp}_{\mathscr{T}}(k) \tag{A-7}$$

where $PD_{\mathscr{T}}(k), PT_{\mathscr{T}}(k), PB_{\mathscr{T}}(k), PT^{\perp}_{\mathscr{T}}(k)$ are in Eqs. (A-1)-(A-6).

Appendix B

As defined, $P_{\mathscr{T}}(k)$ is the number of passengers onboard pod train $\mathscr{T}$ upon arrival at stop k. Thus, $P_{\mathscr{T}}(1) = 0$, which is the baseline. In the subsequent stops, $P_{\mathscr{T}}(k)$ is updated as follows.

$$
\begin{aligned}
P_{\mathcal{F}}(k+1) - P_{\mathcal{F}}(k) &= -PD_{\mathcal{F}}(k) - PT_{\mathcal{F}}(k) + PB_{\mathcal{F}}(k) + PT^{\perp}_{\mathcal{F}}(k) \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(-\frac{k-1}{N-1} - \frac{k-1}{N} + \frac{N-k}{N-1} + \frac{N-k}{N} \right) \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{N-2k+1}{N-1} + \frac{N-2k+1}{N} \right) \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{1}{N} + \frac{1}{N-1} \right) \cdot (N-2k+1) \\
&= \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{1}{N} + \frac{1}{N-1} \right) \cdot (N+1) - \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{1}{N} + \frac{1}{N-1} \right) \cdot 2 \cdot k
\end{aligned}
\tag{B-1}
$$

For a given network and headway, we denote $\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{1}{N}+\frac{1}{N-1}\right) \cdot (N+1)$ as constant ($\mathscr{A} > 0$), and $\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{1}{N}+\frac{1}{N-1}\right) \cdot 2$ as $\mathscr{B}$ ($\mathscr{B} > 0$). Then Eq. (B-1) can be rewritten as

$$P_{\mathcal{F}}(k+1) - P_{\mathcal{F}}(k) = \mathscr{A} - \mathscr{B}k$$

and,

$$
\begin{aligned}
P_{\mathcal{F}}(k) - P_{\mathcal{F}}(k-1) &= \mathscr{A} - \mathscr{B}(k-1) \\
P_{\mathcal{F}}(k-1) - P_{\mathcal{F}}(k-2) &= \mathscr{A} - \mathscr{B}(k-2) \\
P_{\mathcal{F}}(k-2) - P_{\mathcal{F}}(k-3) &= \mathscr{A} - \mathscr{B}(k-3) \\
&\vdots \\
&\vdots \\
P_{\mathcal{F}}(3) - P_{\mathcal{F}}(2) &= \mathscr{A} - \mathscr{B}(2) \\
P_{\mathcal{F}}(2) - P_{\mathcal{F}}(1) &= \mathscr{A} - \mathscr{B}(1)
\end{aligned}
\tag{B-3}
$$

We sum up the series of formulas in Eq. (B-3) above. Since $P_{\mathcal{F}}(1) = 0$, left side goes to $P_{\mathcal{F}}(k)$, which reads as

$$
\begin{aligned}
P_{\mathcal{F}}(k) &= \sum_{k'=1}^{k-1} \mathscr{A} - \mathscr{B}k' \\
&= \mathscr{A} \cdot (k-1) - \mathscr{B} \cdot \frac{(1+k-1)\cdot(k-1)}{2} \\
&= \mathscr{A} \cdot (k-1) - \mathscr{B} \cdot \frac{k \cdot (k-1)}{2} \\
&= \left(\mathscr{A} - \frac{\mathscr{B}}{2}k \right) \cdot (k-1) \\
&= \frac{\mathscr{B}}{2}\left(\frac{2\mathscr{A}}{\mathscr{B}} - k \right) \cdot (k-1) \\
&\le \frac{\mathscr{B}}{2}\left(\frac{\frac{2\mathscr{A}}{\mathscr{B}} - 1}{2} \right)^2
\end{aligned}
\tag{B-4}
$$

If and only if $\frac{2\mathscr{A}}{\mathscr{B}} - k = k - 1$ (i.e., $k = \frac{\mathscr{A}}{\mathscr{B}} + \frac{1}{2}$), "=" is satisfied, and $P_{\mathcal{F}}(k)$ reaches the maximal value.

Now, we substitute $\mathscr{A}$ and $\mathscr{B}$ with $\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{1}{N}+\frac{1}{N-1}\right) \cdot (N+1)$ and $\frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \cdot \left(\frac{1}{N}+\frac{1}{N-1}\right) \cdot 2$ respectively. $P_{\mathcal{F}}(k)$ is given by

$$P_{\mathcal{F}}(k) = \frac{1}{2}H\lambda\left(\frac{D}{N}\right)^2 \left(\frac{1}{N}+\frac{1}{N-1}\right) \cdot (N+1-k) \cdot (k-1) \tag{B-5}$$

References

Aber, J. (2016) Electric Bus Analysis for New York City Transit, Columbia University. http://www.columbia.edu/~ja3041/Electric%20Bus%20Analysis%20for%20NYC%20Transit%20by%20J%20Aber%20Columbia%20University%20-%20May%202016.pdf.

Aldaihani, M.M., Quadrifoglio, L., Dessouky, M.M., Hall, R., 2004. Network design for a grid hybrid transit service. Transp. Res. A Policy Pract. 38 (7), 511–530.

American Public Transportation Association (APTA), *2023 Public Transportation Fact Book*, available at https://www.apta.com/research-technical-resources/transit-statistics/public-transportation-fact-book/.

Beal, L.D.R., Hill, D., Martin, R.A., Hedengren, J.D., 2018. GEKKO Optimization Suite. Processes 6 (8), pp. https://doi.org/10.3390/pr6080106.

Caros, N.S., Chow, J.Y.J., 2021. Day-to-day market evaluation of modular autonomous vehicle fleet operations with en-route transfers. Transportmetrica B-Transport Dynamics 9 (1), 109–133.

Chen, Z.W., Li, X.P., 2021. Designing corridor systems with modular autonomous vehicles enabling station-wise docking: discrete modeling method. Transp. Res. E-Log 152.

Chen, Z.W., Li, X.P., Zhou, X.S., 2020. Operational design for shuttle systems with modular vehicles under oversaturated traffic: continuous modeling method. Transp. Res. Part B-Methodol. 132 https://doi.org/10.1016/j.trb.2019.05.018. SI pp. 76–100.

Chen, Z.W., Li, X.P., Qu, X.B., 2021. A continuous model for designing corridor systems with modular autonomous vehicles enabling station-wise docking. Transp. Sci. 1–30.

Chen, P., Nie, Y., 2017. Analysis of an idealized system of demand adaptive paired-line hybrid transit. Transp. Res. Part B: Methodol., Elsevier 102 (C), 38–54.

Cheng, X. and Lin, J. (2024) On-demand Ridesharing with Autonomous Modular Vehicles, presented at *the 103rd Transportation Research Board Annual Meeting*, Washington D.C., January 7-11, 2024.

Daganzo, C.F., 2010. Structure of competitive transit networks. Transp. Res. B Methodol. 44 (4), 434–446.

Dakic, I., Yang, K.D., Menendez, M., Chow, J.Y.J., 2021. On the design of an optimal flexible bus dispatching system with modular bus units: Using the three-dimensional macroscopic fundamental diagram. Transp. Res. Part B-Methodol. 148, 38–59.

CNN.com, "Dubai: A testing ground for future transport", https://www.cnn.com/2017/11/29/middleeast/gallery/dubai-future-transport/index.html, published May 6, 2018.

Department of Energy (2021). Battery-Electric Vehicles Have Lower Scheduled Maintenance Costs than Other Light-Duty Vehicles. FOTW #1190, June 14, 2021. https://www.energy.gov/eere/vehicles/articles/fotw-1190-june-14-2021-battery-electric-vehicles-have-lower-scheduled.

Gecchelin, T., Webb, J., 2019. Modular dynamic ride-sharing transport systems. Economic Anal. Policy 61 (SI), 111–117.

Guo, R., Guan, W., Vallati, M., Zhang, W., 2023. Modular autonomous electric vehicle scheduling for customized on-demand bus services. IEEE Trans. Intell. Transp. Syst. 24 (9), 10055–10066. https://doi.org/10.1109/TITS.2023.3271690.

Khan, Z.S., He, W., Menéndez, M., 2023. Application of modular vehicle technology to mitigate bus bunching. Transp. Res. Part C: Emerg. Technol. 146, 103953 https://doi.org/10.1016/j.trc.2022.103953.

Lin, J., Nie, Y.M., Kawamura, K., 2023. An Autonomous Modular Mobility Paradigm. IEEE Intell. Transp. Syst. Mag. 15 (1), 378–386. https://doi.org/10.1109/MITS.2022.3159484.

Liu, T., Ceder, A., Rau, A., 2020. Using deficit function to determine the minimum fleet size of an autonomous modular public transit system. Transp. Res. Record. No 0361198120945981.

Liu, X.H., Qu, X.B., Ma, X.L., 2021. Improving flex-route transit services with modular autonomous vehicles. Transp. Res. E-Log 149.

Murphy, K.P., 2012. Machine Learning: A Probabilistic Perspective. MIT Press. Exercise 2.17, Expected value of the minimum.

National Transit Database (NTD), 2015 Report Year National Transit summary and Trends: Appendix,https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/2015%20NTST%20Appendix.pdf, last accessed November 30, 2020.

Next-future-mobility.com. https://www.next-future-mobility.com/.

Ouyang, Y., Nourbakhsh, S.M., Cassidy, M.J., 2014. Continuum approximation approach to bus network design under spatially heterogeneous demand. Transp. Res. Part B: Methodol. 68 (ISSN), 3330191–3442615. https://doi.org/10.1016/j.trb.2014.05.018.

Sethuraman G., Liu X., Bachmann, F. R., Xie M., Ongel A., and Busch F., "Effects of Bus Platooning in an Urban Environment," 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 2019, pp. 974-980, doi: 10.1109/ITSC.2019.8917041.

Tian, Q., Lin, Y.H., Wang, D.Z., 2023. Joint scheduling and formation design for modular-vehicle transit service with time-dependent demand. Transp. Res. Part C: Emerg. Technol. 147, 103986 https://doi.org/10.1016/j.trc.2022.103986.

U. S. Department of Transportation (US DOT), 2023. Federal Transit Administration, National Transit Database (Washington, DC: Annual reports), Vehicles, available at. as of Nov 2, 2023. https://www.transit.dot.gov/ntd/ntd-data.

U. S. Census Bureau, 2021. Commuting by Public Transportation in the United States: 2019. Report Number: ACS-48, Issued April 2021, By Michael Burrows, Charlynn Burd, and Brian Mckenzie. https://www.census.gov/content/dam/Census/library/publications/2021/acs/acs-48.pdf.

Wu, J.M., Kulcsar, B., Selpi, Qu., Xb,, 2021. A modular, adaptive, and autonomous transit system (MAATS): A in-motion transfer strategy and performance evaluation in urban grid transit networks. Transp. Res. Part A-Policy and Practice 151, 81–98.

Zhang, Z.H., Tafreshian, A., Masoud, N., 2020. Modular transit: Using autonomy and modularity to improve performance in public transportation. Transp. Res. E-Log 141, 102033.