Cleaner Logistics and Supply Chain 12 (2024) 100167

Contents lists available at ScienceDirect

Cleaner Logistics and Supply Chain

journal homepage: www.journals.elsevier.com/cleaner-logistics-and-supply-chain

Sustainable truck platooning operations in maritime shipping: A data-driven approach

Zhaojing Yang [a], Min Xu [b], Xuecheng Tian [a,*]

[a] Faculty of Business, The Hong Kong Polytechnic University, Hong Kong, China
[b] Department of Industrial & Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Keywords:
Maritime shipping
Low-carbon shipping
Freight transportation
Platooning
Truck operations

ABSTRACT

In liner shipping, stakeholders are increasingly committed to adopting autonomous and environmentally friendly transportation solutions, especially for truck operations managing container transfers. Beyond reducing labor costs, truck platooning technology—which enables autonomous trucks to operate in close formations, thereby significantly decreasing fuel consumption—promises to revolutionize fleets involved in maritime container transport. However, the potential of these benefits hinges on the process of developing and implementing optimization plans that address the specific challenges of container logistics, particularly in integrating truck platooning plans. In response to this need, this study extends the traditional instant-dispatch strategy by proposing a novel, data-driven dispatch strategy. We develop algorithms for both models and conduct extensive experiments focusing on truck operations for sea freight containers. Our findings reveal significant advantages of the data-driven dispatch strategy: it substantially reduces the total costs and fuel consumption associated with truck deliveries compared to the instant-dispatch strategy.

1. Introduction

Intermodal transportation, which involves the transit of sea freight containers to their final destinations using trucks, plays a crucial role in the supply chain management (Li et al., 2023). Numerous countries and regions around the world are actively enhancing their transportation infrastructures, both maritime and terrestrial, to facilitate the efficient flow of containers. However, this rapid expansion of infrastructure—from seaports to inland road networks—poses substantial challenges. These include a significant increase in carbon dioxide (CO_2) emissions and elevated energy consumption (Jiang et al., 2024). In 2018, the transportation sector accounted for 23% of global energy-related CO_2 emissions (IEA, 2020), with the highest dependency on fossil fuels among all sectors (IEA, 2022). In nearly half of all countries, it is the predominant emitter within the economy (UN, 2021). These issues have catalyzed global discussions focusing on balancing the growing demand for transportation with the imperative to reduce CO_2 emissions (Attanasio et al., 2023).

Fortunately, advancements in intelligent transportation systems have enhanced the energy efficiency of transportation networks. A notable innovation is truck platooning, where vehicles travel closely as a convoy, forming a road train. This configuration significantly reduces air drag, leading to aerodynamic efficiencies (Zhang et al., 2020). This technique can yield energy savings of up to 16 % (Bonnet and Fritz, 2000). Additionally, truck platooning improves traffic flow and can increase road capacity by up to 200% (Tsugawa et al., 2016), thereby reducing congestion without the necessity for costly new infrastructure developments (Robinson et al., 2010).

The numerous advantages of platooning underscore the necessity of effective platoon scheduling, particularly when maritime-shipped goods are transferred to land-based transportation networks. Optimal operation and scheduling of platoons not only mitigate greenhouse gas emissions but also enhance road capacity and yield significant cost savings for enterprises, holding substantial practical value. Consequently, developing an optimal transportation scheduling strategy that minimizes costs while efficiently coordinating platoon dispatches for freight transport is essential for linking maritime and terrestrial components of global supply chains. This paper explores the truck platooning scheduling problem with the goal of minimizing total costs incurred during the transportation of the required requests. Utilizing extensive historical data on trucking requests, we aim to devise cost-effective truck platooning dispatch strategies through a data-driven

* Corresponding author.
E-mail addresses: zhaojing.yang@connect.polyu.hk (Z. Yang), min.m.xu@polyu.edu.hk (M. Xu), xuecheng-simon.tian@connect.polyu.hk (X. Tian).

https://doi.org/10.1016/j.clscn.2024.100167
Received 4 June 2024; Received in revised form 21 August 2024; Accepted 21 August 2024
Available online 23 August 2024

mathematical programming model.

2. Literature review

The pre-departure planning of scheduling and routing plays a crucial role in making truck platooning a common practice in the transport sector. To the best of our knowledge, there have been relatively few studies published in this emerging field of platooning research. This paper presents the current state of the literature concerning platoon coordination methods by centrally planning trips in advance. Larson et al. (2016) studied the mixed-integer programming (MIP) algorithm for the combined vehicle routing and platoon scheduling problem, improved the computational efficiency by introducing auxiliary parameters and constraints, and explored the effect of waiting at initial and intermediate points. Larson et al. (2015) formulated the queuing problem as an integer programming for both same and different starting positions, proved its NP-hardness, developed optimal pairing and pivoting heuristics and improved these heuristics with local search for efficiency, but found that large real-world instances remain problematic. Luo et al. (2018) built the mixed-integer linear programming (MILP) efficient clustering heuristic algorithm that integrates routing, scheduling, individual speed selection, and platoon formation/dissolution, which first separates the set of trucks and then routes each group individually. Meisen et al. (2008) presented the mining frequent sub-routes problem for platoon driving, formulated the algorithm to find truck platoon sequential pattern (TPSpan), and solved the problem by pruning parameters to increase the efficiency. Nourmohammadzadeh and Hartmann (2016) formulated a mathematical MIP model considering vehicle deadlines, employing the LINDO solver for small samples and a Genetic algorithm for larger instances, promoting the development of an effective genetic algorithm based on the prior elimination of irrational routes. Sokolov et al. (2017) addressed a combinatorial optimization problem that integrates platoon coordination with vehicle routing to maximize fuel savings by synchronizing routes and departure times, comparing central coordination with an uncoordinated ad hoc approach through simulations, using the MIP formulation by Larson et al. (2016) as a basis, and demonstrating a substantial increase in platooning possibilities through reasonable waiting times at the origin. Zhang et al. (2016) introduced and analyzed a platoon scheduling problem that aims to minimize total delivery costs by considering travel time variance. The model accounts for driving costs, schedule miss penalties, and fuel costs, highlighting the conflicting goals of timely arrival and fuel savings through platooning. Zhang et al. (2017) formulated and analyzed a freight transport platoon coordination and departure time scheduling problem under travel time uncertainty, aiming to minimize expected costs. You et al. (2023) studied the multi-trip container drayage problem with truck platooning, addressing its NP-hardness by proposing a branch-and-price-and-cut algorithm with a route-based set partitioning model and tight linear relaxations to achieve exact solutions. Xu et al. (2022) addressed the truck routing and platooning problem, considering drivers' mandatory breaks, state-and-position-dependent fuel-saving rates, designated intermediate relays, and platoon size limits, aiming to route trucks to their destinations on time with the minimal fuel costs. They used a MILP model and a hybrid algorithm combining partial-MILP and iterated neighborhood search to solve the problem.

Existing studies often assume pre-known information of future demands to create scheduling models, which is unrealistic since we only have access to historical and current requests. This paper seeks to bridge this gap by developing a data-driven optimization approach that leverages historical data to develop more realistic scheduling models for truck platooning.

The main contributions of this paper are as follows. First, we consider two approaches: the traditional approach that dispatches a platoon once requests are received during a decision period, and a data-driven approach that predicts future demands based on the distribution patterns of historical request data and makes the decision by solving a mathematical optimization model. Second, we conduct computational experiments using a truck delivery center as a case study, comparing the total cost and the fuel cost of the two approaches in several cases. These experiments validate the superiority of the data-driven dispatch strategy.

The remainder of this paper is organized as follows. Section 3 provides problem description and formulates the models of the two approaches. Section 4 verifies the proposed method with a concrete example. Section 5 concludes this paper.

3. Problem formulation

Assume that there is one delivery center and a set of U customers, each indexed by u. The delivery center has been making platoon dispatching decisions for T periods, indicating that the historical requests during the past T periods are available. In order to verify the effectiveness of our proposed data-driven methods, we split the original dataset into the training and testing data sets in the ratio of 4:1, indicating that the delivery center designs strategies using the first 80% of data (i.e., the information on the first $0.8T$ periods, including the number of requests generated on each period) and evaluates the performance of the strategies using the last 20% of data. We denote the training data set by T_1, where each period $t \in \{1, 2, \cdots, \mathcal{T}_1\}$, and the testing data set by T_2, where each time period $t' \in \{1, 2, \cdots, \mathcal{T}_2\}$. We adopt the ratio of 4:1 because it is the most widely used data set splitting ratio (Sadeghi et al.,2022). In most cases, other ratios (e.g., 3:1) have little or no effect on model performance (Ferentinos,2018; Tseremoglou et al.,2022).

Now, we use the test set to describe our studied problem. Specifically, at a specific decision period $t' \in \{1, 2, \cdots, \mathcal{T}_2\}$, each customer $u \in \{1, \cdots, U\}$ can place a number of requests, represented by $n^u_{t'}$. In this paper, we assume that a request needs to be satisfied by a truck. The aim is to minimize the total costs, which includes the platoon costs and the delay costs for the planning horizon of the test period. The platoon cost is calculated using the unit cost c_1 for the first truck and the unit cost c_0 of the following trucks. The first truck in the platoon incurs a unit cost c_1 that encompasses the driver's salary, fuel expenses, and the cost associated with overcoming the air resistance. The trucks that hang after the first truck have a unit cost of c_0, which is only the cost of fuel, since there is no driver as well as less air resistance. Then, it can be seen that $c_0 < c_1$.

After collating requests from all customers during a specific decision period, the delivery center must decide whether to dispatch the platoon immediately at the corresponding time period t' (denoted by $x_{t'} = 1$) or to defer the deliveries to subsequent periods (denoted by $x_{t'} = 0$). The unit delay penalty for delaying customer u's request at period t' is represented by $p^u_{t'}$. Consequently, if the delivery center chooses to defer the deliveries to subsequent periods, there will be delay penalties (delay costs), which are equal to the unit delay penalty multiplied by the number of delayed requests, i.e., $\sum_{u=1}^{U} p^u_{t'} n^u_{t'}$. Here, we assume that if a request at a certain period is not immediately met, the delay penalty for that unmet request will only incur at that period; that is to say, if the request is still not met at the next period, the request will not incur delay penalty at the next period.

Now if we choose to send a platoon at period t (i.e., $x_{t'} = 1$), the number of trucks in the platoon, denoted by $D_{t'}$, is the sum of the unsatisfied requests up to period t', denoted by $A_{t'}$, and the new requests generated at period t', $\sum_{u=1}^{U} n^u_{t'}$. Therefore, we have $D_{t'} = A_{t'} + \sum_{u=1}^{U} n^u_{t'}$. Both D_t and A_t are measured by the number of trucks.

3.1. Method I: Instant-dispatch strategy

In this strategy, we assume that as long as there is a request coming during a decision period, the decision maker would dispatch a platoon. The notation is shown in Table 1, and the model, referred to as $[\mathcal{M}_1]$, can be formulated below. Model $[\mathcal{M}_1]$ is followed by Algorithm 1, which solves model $[\mathcal{M}_1]$.

Table 1
Notation for model $[\mathscr{M}_1]$.

Notation	Meaning
$\mathscr{T}_2$	The number of periods in the testing set T_2
t'	The index of the testing period, where $t' \in \{1, 2, .., \mathscr{T}_2\}$
$x_{t'}$	The binary decision variable, denoting whether to dispatch the platoon immediately at period t' or not
c_1	The cost of the first truck in a platoon
c_0	The unit cost of the trucks following the first truck in a platoon
$D_{t'}$	The sum of the requests accumulated before and at period t' after the previous platoon sent
U	The number of customers, indexed by $u \in \{1, 2, \cdots, U\}$
$n_{t'}^u$	A deterministic parameter representing the number of requests by customer u at period t'

$$[\mathscr{M}_1]\min w_1 = \sum_{t'=1}^{\mathscr{T}_2} [x_{t'}(c_1 + c_0(D_{t'} - 1))] \tag{1}$$

subject to

$$x_{t'} \geq \mathbb{I}(D_{t'} \geq 1), \forall t' \in \{1, 2, \cdots, \mathscr{T}_2\} \tag{2}$$

$$D_{t'} = \sum_{u=1}^{U} n_{t'}^u, \forall t' \in \{1, 2, \cdots, \mathscr{T}_2\} \tag{3}$$

$$x_{t'} \in \{0, 1\}, \forall t' \in \{1, 2, \cdots, \mathscr{T}_2\} \tag{4}$$

Objective function (1) aims to minimize the total cost w_1 for the test planning horizon, where $c_1 + c_0(D_{t'} - 1)$ is the cost of dispatching a platoon at the current period t', and $\sum_{t'=1}^{\mathscr{T}_2} [x_{t'}(c_1 + c_0(D_{t'} - 1))]$ is the total cost for all of the testing periods. In Constraints (2), $\mathbb{I}(*)$ is an indicator function. It indicates that the value of the function is 1 if the condition in the parentheses $\mathbb{I}(*)$ is satisfied; so, Constraints (2) indicate that as long as a request is received at period t', then a platoon will be dispatched. For this reason, $D_{t'}$ in Constraints (3) indicates the total number of requests generated at period t'. In $[\mathscr{M}_1]$, since platoons will be dispatched when requests are generated, there will be no delayed cost. Constraints (4) define the domains of decision variables.

Given model $[\mathscr{M}_1]$, we then present Algorithm 1 below to show the solution process of the model.

Algorithm 1. Solution process of model $[\mathscr{M}_1]$

Input: U (the number of customers), T (the number of periods), c_0 (the cost of the first truck), c_1 (the unit cost of the following trucks), n_t^u (the requests from every customer at every period)
Output: w_1 (total cost)
1:Let $w_1 \leftarrow 0$
2:$\mathscr{T}_2 \leftarrow 0.2T$
3:For $t' \leftarrow 1, 2, \cdots, \mathscr{T}_2$ // Enumerate the possible values of dispatch period
4:If $D_{t'} > 0$
5:$x_{t'} \leftarrow 1$
6:$D_{t'} \leftarrow \sum_{u=1}^{U} n_{t'}^u$
7:$w_1 \leftarrow w_1 + x_{t'}(c_1 + c_0(D_{t'} - 1))$
8:Else
9:$w_1 \leftarrow w_1 + 0$
10:End for

The instant-dispatch strategy, while offering immediate response capabilities, presents several significant drawbacks. It leads to higher operational costs due to the increased fuel consumption and vehicle wear and tear, which in turn escalate maintenance expenses. Additionally, more frequent dispatches necessitate more labor hours, increasing driver costs. This strategy also has a negative environmental impact, as inefficient dispatching results in more trips and higher fuel consumption, thereby contributing to a larger carbon footprint and environmental degradation. Scalability is another major issue; as the volume of orders grows, managing instant dispatches becomes increasingly challenging, causing bottlenecks and inefficiencies. This strain on resources and infrastructure further complicates scaling up operations. The strategy's limited flexibility is evident in its inability to adapt to changing conditions such as traffic patterns, weather, or sudden demand spikes, making it inherently reactive rather than proactive. This lack of adaptability hampers the ability to anticipate and mitigate potential issues. Moreover, the instant-dispatch strategy often fails to leverage data analytics, resulting in a lack of insights that could otherwise inform decision making, optimize routes, and predict demand. Consequently, decisions are less informed and more prone to errors, highlighting the need for a more sophisticated, data-driven approach.

3.2. Method II: Data-driven dispatch strategy

From the above section, we know that the training set T_1 and the testing set T_2 are obtained from the same distribution. When making dispatching decisions at one period in the testing set T_2, we do not know the exact number of requests that will be received in the future periods, but we know the number of requests of every period in the training set T_1, which has the same distribution. Therefore, we can use the training set T_1 to develop a data-driven strategy to make a decision on whether to send a platoon or not for each period of the testing set T_2.

For example, there are T data points, where T may equal 20, and the training set T_1 may contain the first 16 data points while the testing set T_2 contains the last 4 data points. Suppose we are at the first period of T_2, which means $t' = 1$, and for every customer u, we receive $n_{t'}^u$ requests. From the decision-making perspective, we do not know the requests that we will receive in the future $T_2 - t'$ periods; therefore, we use T_1 to generate S scenarios of uncertain requests in the future $T_2 - t'$ periods, as an approximation for the future data. Each scenario is represented by s, $s \in \{1, \cdots, S\}$. We choose the simplest method, random sampling, as the method for generating a scenario: we randomly select $T_2 - t'$ periods (with the historical demand information) from the T_1 dataset to form one scenario. This is repeated S times, generating S scenarios, each with U rows representing customers and $T_2 - t'$ columns ($j \in \{1, 2, \cdots, T_2 - t'\}$) representing future $T_2 - t'$ periods. This approach is shown visually in Fig. 1.

Then, we develop an optimization model $[\mathscr{M}_2]$ to obtain the optimal solution $x_{t'}$ using the data of the current period t' and the S scenarios for future periods. The model $[\mathscr{M}_2]$ operates each time when we approach to a new time period to make a decision. When $x_{t'} = 1$, it means we will send a platoon at the current period t'. When $x_{t'} = 0$, we will not send a platoon at t'. The notation is shown in Table 2 and model $[\mathscr{M}_2]$ is constructed below. Model $[\mathscr{M}_2]$ is followed by Algorithm 2, which solves model $[\mathscr{M}_2]$.

$$[\mathscr{M}_2]\min_x x_{t'}(c_1 + c_0(D_{t'} - 1)) + (1 - x_{t'})\sum_{u=1}^{U} p_{t'}^u n_{t'}^u$$

$$+\frac{1}{S}\sum_{s=1}^{S}\sum_{j=1}^{T'}\left[x_j^s\left(c_1 + c_0\left(D_j^s - 1\right)\right) + (1 - x_j^s)\sum_{u=1}^{U} p_j^{u,s} n_j^{u,s}\right] \tag{5}$$

subject to

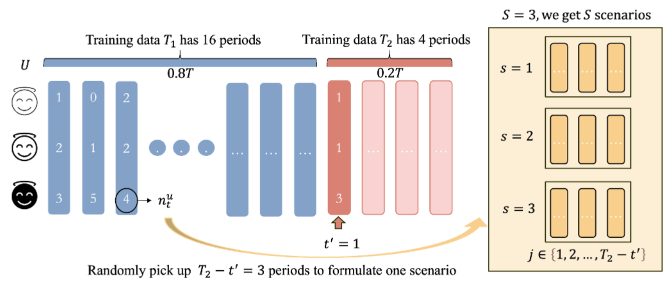

Fig. 1. The data-driven dispatch strategy.

Table 2
Newly introduced notation for model $[\mathscr{M}_2]$.

Notation	Meaning
$p_{t'}^u$	The unit penalty of delaying customer u's requests at period t'
S	The number of scenarios developed
s	The index of scenarios, where $s \in \{1, \cdots, S\}$
T'	The number of future periods, which is $\mathscr{T}_2 - t'$
j	The index of future periods, where $j \in \{1, 2, \cdots, T_2 - t'\}$
x_j^s	The decision variable for the jth future period in scenario s
D_j^s	The sum of the remaining requests at the end of the jth future period in scenario s
$p_j^{u,s}$	The unit delay penalty for customer u's requests for the jth future period in scenario s
$n_j^{u,s}$	The number of requests from customer u at the jth future period in scenario s
A_j^s	The unsatisfied requests up to the jth future period in scenario s

$$D_{t'} = \sum_{u=1}^{U} n_{t'}^u + A_{t'} \tag{6}$$

$$D_j^s = \sum_{u=1}^{U} n_j^{u,s} + A_j^s \forall j \in \{1, \cdots, T'\}, s \in \{1, \cdots, S\} \tag{7}$$

$$A_j^s = A_{j-1}^s + \sum_{u=1}^{U} n_{j-1}^{u,s} - x_{j-1}^s D_{j-1}^s \forall j \in \{2, \cdots, T'\}, s \in \{1, \cdots, S\} \tag{8}$$

$$A_1^s = D_{t'}(1 - x_{t'}) \forall s \in \{1, \cdots, S\} \tag{9}$$

$$x_{T'}^s \geq I\left(D_{T'}^s > 0\right) \forall s \in \{1, \cdots, S\} \tag{10}$$

$$x_{t'} \in \{0, 1\} \tag{11}$$

$$x_j^s \in \{0, 1\} \forall j \in \{1, \cdots, T'\}, s \in \{1, \cdots, S\}. \tag{12}$$

Objective function (5) aims to minimize the total cost, where $x_{t'}(c_1 + c_0(D_{t'} - 1)) + (1 - x_{t'})\sum_{u=1}^{U} p_{t'}^u n_{t'}^u$ is the cost at the current period t' and $\frac{1}{S}\sum_{s=1}^{S}\sum_{j=1}^{T'}\left(x_j^s\left(c_1 + c_0\left(D_j^s - 1\right)\right) + (1 - x_j^s)\sum_{u=1}^{U} p_j^{u,s} n_j^{u,s}\right)$ is the approximated expected cost of dispatching platoons in future T' periods. Constraint (6) indicates that if we choose to send a platoon at period t', the number of trucks in the platoon $D_{t'}$ is the sum of the unsatisfied requests up to period t' (we name it $A_{t'}$, which can be precomputed) and the requests generated at period t', which is $\sum_{u=1}^{U} n_{t'}^u$. Constraints (7) indicate that the number of trucks in the platoon dispatched at the end of the jth future period in scenario s is the sum of the unsatisfied requests up to the jth future period in scenario s (we name it A_j^s) and the requests generated at the jth future period in scenario s from all of the U customers, which is $\sum_{u=1}^{U} n_j^{u,s}$. Constraints (8) indicate that if we have dispatched a platoon at the $(j-1)$th future period in scenario s, A_j^s will be 0; otherwise, A_j^s will be the sum of the unsatisfied requests up to the $(j-1)$th future period in scenario s (we name it A_{j-1}^s) and the requests generated at the $(j-1)$th future period in scenario s from all of the U customers $\sum_{u=1}^{U} n_{j-1}^{u,s}$. Constraints (9) indicate that if we have dispatched a platoon at period t' in scenario s, the unsatisfied requests up to the first next period in scenario s, which is A_1^s, will be 0; otherwise, A_1^s will be $D_{t'}$. Constraints (10) indicate that as long as a request is generated at the last period in scenario s, then a platoon must be dispatched. In practice, we need to linearize Constraints (10) using the Big-M method: $M \bullet D_{T'}^s \geq x_{T'}^s$, where M is a very large number. Constraints (11)–(12) define the domains of the decision variables. In order to better demonstrate the above method as well as to visualize how to solve model $[\mathscr{M}_2]$ to calculate the cost, we outline the detailed pseudo codes in Algorithm 2.

Algorithm 2. Solution process for model $[\mathscr{M}_2]$

Input: All of the parameters in Tables 1 and 2
Output: *platoons* (set of the number of trucks for each platoon sent), dispatch solutions *solutions*, *totalcost*

(*continued on next column*)

(*continued*)

Algorithm 2. Solution process for model $[\mathscr{M}_2]$

Let *totalcost*←0
Let *platoons*←[ ], *solutions*←[]
Let $t' \leftarrow 1; \mathscr{T}_2 \leftarrow 0.2T$
Let $A_{t'} \leftarrow 0$
While $t' < \mathscr{T}_2$
Let $J \leftarrow \mathscr{T}_2 - t'$
Set $n_j^{u,s}, p_j^{u,s}$ by taking J random columns of the training data
Formulate the model $[\mathscr{M}_2]$ with parameters $n_j^{u,s}, p_j^{u,s}$
Solve the model $[\mathscr{M}_2]$ by GUROBI to obtain $x_{t'}, D_{t'}$
If $x_{t'} = 0$
$cost \leftarrow \sum_u p_{t'}^u n_{t'}^u$
$A_{t'} \leftarrow A_{t'} + \sum_{u=1}^{U} n_{t'}^u$
Else
$cost \leftarrow (D_{t'} - 1)c_0 + c_1$
$A_{t'} \leftarrow 0$
End if
totalcost←*totalcost* +*cost*
solutions←*solutions* ∪ $x_{t'}$
platoons←*platoons* ∪ $D_{t'}$
Update the training data by adding $n_{t'}^u$ to the training data
$t' \leftarrow t' + 1$
End while
$t' \leftarrow \mathscr{T}_2$ // Obtain $x_{t'}$ and *cost* for the last period $\mathscr{T}_2$ for testing data
If $\sum_{u=1}^{U} n_{t'}^u > 0$ or $A_{t'} > 0$
$x_{t'} \leftarrow 1; D_{t'} \leftarrow A_{t'} + \sum_{u=1}^{U} n_{t'}^u; cost \leftarrow (D_{t'} - 1)c_0 + c_1$
Else
$x_{t'} \leftarrow 0; D_{t'} \leftarrow 0; cost \leftarrow 0$
End if
totalcost←*totalcost* + *cost*
solutions←*solutions* ∪ $x_{t'}$
platoons←*platoons* ∪ $D_{t'}$

After making the decision for period t', we add the data of $t' = 1$ to the training set T_1 and get the new T_1 set. We move to the next period of T_2, which means $t' = 2$. We start the next round of operation and obtain the optimal solution, and so on.

4. Computational experiments

The experiments are run on a laptop computer equipped with Apple M2 Pro CPU and 16 GB of RAM, and model $[\mathscr{M}_2]$ is solved using the GUROBI solver within Jupyter. We first set the values of parameters for drawing the basic results, and then we conduct computational experiments to examine the impacts of these parameters.

4.1. Experiment and parameter settings

The number of customers is set to $U = 3$. We denote each period by a month and the total number of periods by $T = 120$, so the testing data $\mathscr{T}_2 = 0.2T = 24$, which is 24 months (2 years). When formulating the model $[\mathscr{M}_2]$, we build $S = 3$ scenarios.

The numbers of requests n_t^u and the unit penalties p_t^u are sampled from normal distributions. The mean of the distribution for n_t^u is a random number from the set $\{0, \cdots, 25\}$, while the variance is a number from the set $\{1, \cdots, 5\}$. Similarly, the normal distribution for p_t^u has the mean in the range $\{200, \cdots, 300\}$ and the variance in $\{1, 2, 3\}$.

Next, to set the value of the parameter c_1, we conduct a survey of logistics companies. According to our results, the total cost of one FAW Qingdao 6.8-meter-long truck to deliver goods from Heng Shui to Weifang (420 kms, 20 h) is approximately \$117/workday, including the fuel cost of \$48, the loss charge of \$2, the highway toll fees of \$28, the driver cost of \$36, and the insurance and tax of \$3. Given that each month spans 25 workdays, the value of the parameter is set to $c_1 = 117 \bullet 25 =$ \$2925 per period.

Finally, we set the value of the parameter c_0 by removing the driver cost and lowering the fuel cost with respect to the first truck's cost, c_1.

According to 0, the fuel cost is reduced by approximately 16 %, resulting in only $40 of fuel cost. Adding the loss charge of $2, the highway toll fees of $28, and the insurance and tax for $3, we get $73per day, which corresponds to $c_0 = 73 \bullet 25 = \$1825$ per period.

4.2. Basic results

This subsection analyses the cost saved by following the data-driven dispatch strategy as opposed to the instant-dispatch strategy. We generate and test 30 cases with different distribution parameters for the numbers of requests n_t^u and the unit penalties p_t^u.

Fig. 2 shows the total cost saved by Model $[\mathscr{M}_2]$ (data-driven dispatch strategy) compared to Model $[\mathscr{M}_1]$ (instant-dispatch strategy), calculated as the total cost obtained from Model $[\mathscr{M}_1]$ minus the total cost obtained from Model $[\mathscr{M}_2]$. In half of the cases, including Cases 0, 1, 2, 4, 5, 9, 11, 14, 15, 19, 21, 22, 27, and 29, Model $[\mathscr{M}_2]$ has total costs lower than Model $[\mathscr{M}_1]$. This is the money saved in two years of operation. Compared to the traditional instant-dispatch strategy, data-driven dispatch strategy can help truck delivery centers save a significant amount of costs.

Additionally, in Fig. 3, we have separately illustrated the fuel costs saved by the data-driven dispatch strategy over the instant-dispatch strategy. As shown in Fig. 3, half of the cases, Model $[\mathscr{M}_2]$ has fewer fuel costs than Model $[\mathscr{M}_1]$, indicating that data-driven dispatch strategy is more effective in conserving fuel energy. Furthermore, in all of these cases, there are no cases where Model $[\mathscr{M}_1]$ yields lower total cost and fuel consumption than Model $[\mathscr{M}_2]$.

The credibility of this study lies in the extensive testing of numerous case studies, all of which consistently demonstrated that the data-driven strategy is more cost-efficient than the instant-dispatch strategy.

Based on our results, we recommend that stakeholders consider historical data when making future decisions regarding platooning. By doing so, they can accurately forecast future demand and make informed decisions about platoon size, thereby reducing the operational costs associated with platooning. These measures will also contribute to the preservation of the natural environment and the mitigation of transportation-related issues.

5. Conclusions

We have developed a data-driven framework designed to support the real-time decision-making processes of service providers in the context of truck platooning, with the primary goal of minimizing operational costs. This framework incorporates tailored optimization models and efficient algorithms to ensure the effective scheduling of autonomous trucks, facilitating the sequential fulfillment of multiple delivery requests.

To assess the effectiveness of the proposed decision-making tools, we conducted numerical experiments and case studies. These investigations focus on the influence of delivery request patterns on the cost-saving performance of platoon services. Based on our findings, we provide targeted suggestions for improvement measures and guidelines to relevant stakeholders.

The innovation of this study lies in addressing the future demand uncertainty when making platooning decisions. Specifically, we handle the uncertain demand by using historical demand data. This paper makes contributions to both theoretical knowledge and practical implementations, enhancing innovative industrial mobility and supporting the decarbonization of international trade and commerce. For future truck platoon optimization, historical data can be leveraged to design data-driven models that align closely with actual conditions, thereby benefiting both the company's operations and the environment cleanness.

The methodology presented in the paper has several limitations. First, while the model offers a theoretical solution, its computational complexity may hinder practical applications, particularly for large-scale problems. Second, the paper assumes that the platoon fleet serves three customers in the same end city, meaning that the trucks have identical start and end points. If the end points are in different cities, route planning must be considered when making decisions.

CRediT authorship contribution statement

Zhaojing Yang: Writing – original draft, Visualization, Validation, Software, Methodology, Data curation. **Min Xu:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Xuecheng Tian:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization.

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.

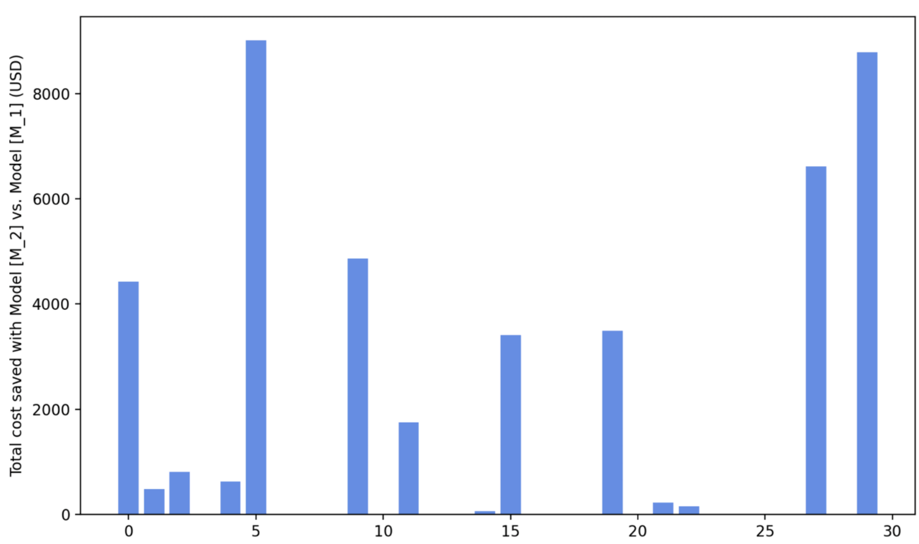

Fig. 2. Total cost saved by Model $[\mathscr{M}_2]$.

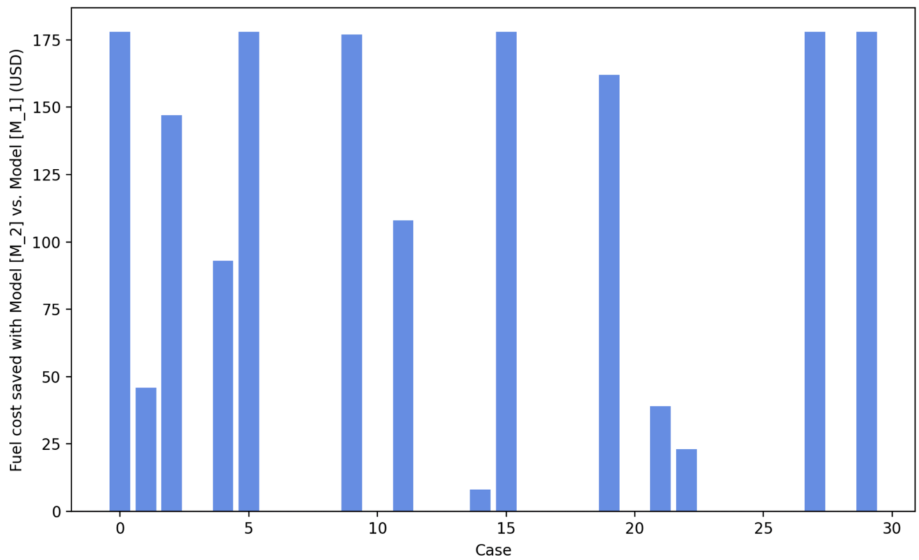

Fig. 3. Fuel cost saved by Model $[\mathscr{M}_2]$.

Data availability

Data will be made available on request.

Acknowledgement

This work is also supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 15222822).

References

Attanasio, G., Battistella, C., Chizzolini, E., 2023. The future of energy management: Results of a Delphi panel applied in the case of ports. J. Clean. Prod. 417, 137947.

Bonnet, C., & Fritz, H., 2000. Fuel consumption reduction experienced by two promote-chauffeur trucks in electronic towbar operation. In Proceedings of the 7th World Congress on Intelligent Systems, Turin.

Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. 145, 311–318.

IEA, 2020. Global energy-related CO2 emissions by sector. Available at: https://www.iea.org/data-and-statistics/charts/global-energy-related-co2-emissions-by-sector/ (accessed 14 August 2024).

IEA, 2022, Transport. Available at: https://www.iea.org/energy-system/transport/ (accessed 14 August 2024).

Jiang, M., Wang, B., Hao, Y., Chen, S., Wen, Y., Yang, Z., 2024. Quantification of CO_2 emissions in transportation: an empirical analysis by modal shift from road to waterway transport in Zhejiang, China. Transp. Policy 145, 177–186.

Larson, J., Munson, T., & Sokolov, V., 2016. Coordinated platoon routing in a metropolitan network. In Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing, Albuquerque.

Li, L., Wang, J., Wang, H., Jin, X., Du, L., 2023. Intermodal transportation hub location optimization with governments subsidies under the Belt and Road Initiative. Ocean Coast. Manag. 231, 106414.

Luo, F., Larson, J., Munson, T., 2018. Coordinated platooning with multiple speeds. Transportation Research Part c: Emerging Technologies 90, 213–225.

Meisen, P., Seidl, T., & Henning, K., 2008. A data-mining technique for the planning and organization of truck platoons. In Proceedings of the International Conference on Heavy Vehicles, Paris.

Nourmohammadzadeh, A., & Hartmann, S., 2016. The fuel-efficient platooning of heavy duty vehicles by mathematical programming and genetic algorithm. In Proceedings of Theory and Practice of Natural Computing: 5th International Conference, Sendai.

Robinson, T., Chan, E., & Coelingh, E., 2010. Operating platoons on public motorways: An introduction to the sartre platooning programme. In Proceedings of the 17th World Congress on Intelligent Transport Systems, Busan.

Sadeghi, S., Akbarpour, A., Abbasianjahromi, H., 2022. Provide a lean and agile strategy for an antifragile sustainable supply chain in the construction industry (residential complex). Clean. Logist. Supply Chain 5, 100079.

Sokolov, V., Larson, J., Munson, T., Auld, J., Karbowski, D., 2017. Maximization of platoon formation through centralized routing and departure time coordination. Transport. Res. Record: J. Transport. Res. Board 2667, 10–16.

Tseremoglou, I., Bombelli, A., Santos, B.F., 2022. A combined forecasting and packing model for air cargo loading: a risk-averse framework. Transport. Res. Record: J. Transport. Res. Board 158, 102579.

Tsugawa, S., Jeschke, S., Shladover, S.E., 2016. A review of truck platooning projects for energy savings. IEEE Trans. Intell. Veh. 1, 68–77.

UN, 2021. UN sustainable transport conference fact sheet: Climate change. Available at: https://www.un.org/sites/un2.un.org/files/media_gstc/FACT_SHEET_Climate_Change.pdf (accessed 14 August 2024).

Xu, M., Yan, X., Yin, Y., 2022. Truck routing and platooning optimization considering drivers' mandatory breaks. Transporta. Res. Part c: Emerg. Technol. 143, 103809.

You, J., Wang, Y., Xue, Z., 2023. An exact algorithm for the multi-trip container drayage problem with truck platooning. Transport. Res. Part e: Logist. Transport. Rev. 175, 103138.

Zhang, L., Chen, F., Ma, X., Pan, X., 2020. Fuel economy in truck platooning: A literature overview and directions for future research. J. Adv. Transp. 2020, 2604012.

Zhang, W., Ma, X., & Jenelius, E., 2016. Planning of heavy-duty vehicle platoon formulation: Basic scheduling problem considering travel time variance. In Proceedings of Transportation Research Board 95th Annual Meeting, Washington, D. C.

Zhang, W., Jenelius, E., Ma, X., 2017. Freight transport platoon coordination and departure time scheduling under travel time uncertainty. Transport. Res. Part e: Logist. Transport. Rev. 98, 1–23.