

Hub-Based Platoon Formation: Optimal Release Policies and Approximate Solutions

Alexander Johansson^{ID}, Ehsan Nekouei^{ID}, *Member, IEEE*, Xiaotong Sun,
Karl Henrik Johansson^{ID}, *Fellow, IEEE*, and Jonas Mårtensson^{ID}, *Member, IEEE*

Abstract—This paper studies the optimal hub-based platoon formation at hubs along a highway under decentralized, distributed, and centralized policies. Hubs are locations along highways where trucks can wait for other trucks to form platoons. A coordinator at each hub decides the departure time of trucks, and the released trucks from the hub will form platoons. The problem is cast as an optimization problem where the objective is to maximize the platooning reward. We first show that the optimal release policy in the decentralized case, where the hubs do not exchange information, is to release all trucks at the hub when the number of trucks exceeds a threshold computed by dynamic programming. We develop efficient approximate release policies for the dependent arrival case using this result. To study the value of information exchange among hubs on platoon formation, we next study the distributed and centralized platoon formation policies which require information exchange among hubs. To this end, we develop receding horizon solutions for the distributed and centralized platoon formation at hubs using the dynamic programming technique. Finally, we perform a simulation study over three hubs in northern Sweden. The profits of the decentralized policies are shown to be approximately 3.5% lower than the distributed policy and 8% lower than the centralized release policy. This observation suggests that decentralized policies are prominent solutions for hub-based platooning as they do not require information exchange among hubs and can achieve a similar performance compared with distributed and centralized policies.

Index Terms—Platoon coordination, optimal control, transport planning, cyber-physical systems, simulation.

Manuscript received 6 April 2023; revised 5 October 2023; accepted 7 November 2023. Date of publication 23 November 2023; date of current version 31 May 2024. The work of Alexander Johansson, Karl Henrik Johansson, and Jonas Mårtensson was supported in part by the Horizon 2020 through the Project ENSEMBLE, in part by the Strategic Vehicle Research and Innovation Programme, in part by the Knut and Alice Wallenberg Foundation by a Wallenberg Scholar Grant, and in part by the Swedish Research Council through the Distinguished Professor under Grant 2017-01078. The work of Ehsan Nekouei was supported by the Start-Up Grant through the City University of Hong Kong under Grant 7200658. The Associate Editor for this article was A.-T. Nguyen. (*Corresponding author: Alexander Johansson.*)

Alexander Johansson is with the FOI Swedish Defence Research Agency, 164 90 Stockholm, Sweden (e-mail: alexander.samimi.johansson@foi.se).

Ehsan Nekouei is with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong (e-mail: enekouei@cityu.edu.hk).

Xiaotong Sun is with the Thrust of Intelligent Transportation, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511458, China, and also with the Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong, SAR, China (e-mail: xtsun@ust.hk).

Karl Henrik Johansson and Jonas Mårtensson are with the Integrated Transport Research Laboratory and the Division of Decision and Control, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, SE-10044 Stockholm, Sweden (e-mail: kallej@kth.se; jonas1@kth.se).

Digital Object Identifier 10.1109/TITS.2023.3332607

1558-0016 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See <https://www.ieee.org/publications/rights/index.html> for more information.

I. INTRODUCTION

A. Motivation

IN THE truck platooning technology, a set of trucks drive with small inter-vehicular distances. Typically, a human driver maneuvers the lead truck in a platoon, while automated driving systems maneuver the follower trucks. The truck platooning technology reduces the operational cost of road transportation by lightening drivers' workload. The fuel consumption level of the follower trucks in a platoon reduces due to the small inter-vehicular distances between trucks. Thus, the truck platooning technology results in significant environmental benefits. For instance, energy savings of approximately 10% for follower trucks have been reported in the literature based on experimental data [1], [2], [3], [4], [5]. Other benefits of truck platooning include increased road capacity and safety and reduced travel time [6], [7], [8]. All these potential benefits motivate the investigations of different platooning architectures and business models, *e.g.*, see, [9], [10], and [11].

B. Related Work

Platoon coordination is a central part of the truck platooning technology. Here, platoon coordination refers to the decision-making of which trucks should form a platoon depending on their travel time schedules and their geographical locations. Trucks can merge into platoons on the road by adjusting their speeds. For example see [12], [13], [14], and [15]. One drawback of on-road platoon formation is their complexity level, as they require trucks to speed up or slow down to merge into platoons on the roads. Such maneuvers might be challenging to perform depending on the surrounding traffic conditions. Another drawback of on-road platoon formation is that the surrounding traffic may be disturbed by the platoon formation or may disturb the platoon formation. An alternative platoon formation approach is hub-based platoon formation. Hubs are locations along highways where trucks can stop and wait for others to form platoons, *e.g.*, freight terminals, gas stations, parking places, tolling stations, and harbors. Under this approach, a platoon coordinator decides on the departure time of trucks from the hubs; the trucks which leave a hub simultaneously form a platoon. Platoon formation at a single hub when the coordinator knows trucks' arrival schedules to the hub was considered in [16], [17], and [18]. In [19] and [20], the authors studied the agent-based platoon formation at multiple hubs in a transportation network,

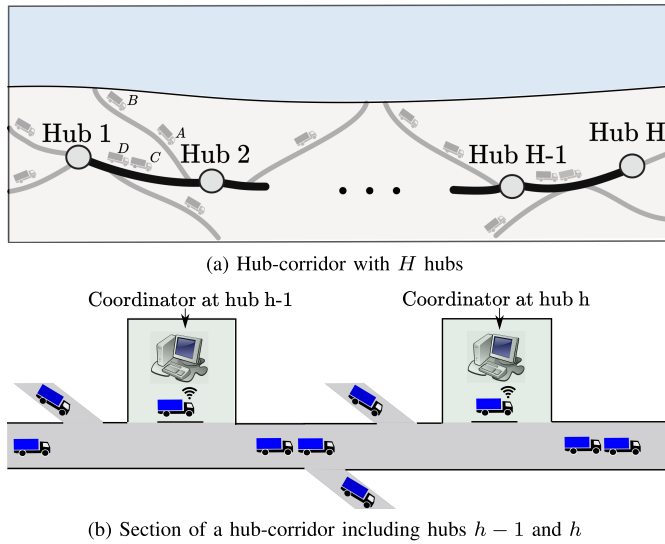


Fig. 1. (a) A hub-corridor with H hubs in a transportation network, and (b) the section of a hub-corridor including hubs $h-1$ and h where coordinators make platoon release decisions.

where trucks individually decide their departure times from hubs. In the agent-based approaches, the trucks are required to share their arrival schedules at hubs with each other. In [21], the authors studied platoon formation at multiple hubs when the decision-making of which platoons should form was distributed on coordinators at the hubs. In the above-mentioned work, the trucks are required to share their arrival schedules with the coordinators at the hubs.

Transportation firms or individual truckers may need to keep their arrival times private for privacy or competitive reasons. For example, the arrival times may be sensitive information for a transportation firm or a trucker to share, as a competitor may use this information to take the firm's customers by offering earlier deliveries. Thus, the arrival schedules may not always be available to the platoon coordinator. Platoon formation at a single hub when the trucks' arrivals are unknown to the hub coordinator in advance was studied in [22] and [23], where the truck arrival process was modeled by a sequence of independent and identically distributed random variables. This paper also considers a platoon formation problem with unknown arrivals but for multiple connected hubs instead of a single hub, as in [22] and [23]. Our paper advances the state-of-the-art by developing efficient decentralized platoon formation policies, where the coordinator at each hub takes into account the downstream hubs' policies when computing its platoon release policy. Thus, the platoon formation problem in this paper becomes a multi-agent decentralized control problem. We finally refer the reader to [24] and [25] for extensive reviews on platoon coordination strategies.

C. Contributions

In this paper, we study the platoon formation in a set of hubs, as shown in Fig. 1a. In our set-up, each hub is equipped with a platoon coordinator, as shown in Fig. 1b, which decides the departure time of trucks from the hubs. The trucks that leave a hub at the same time form a platoon.

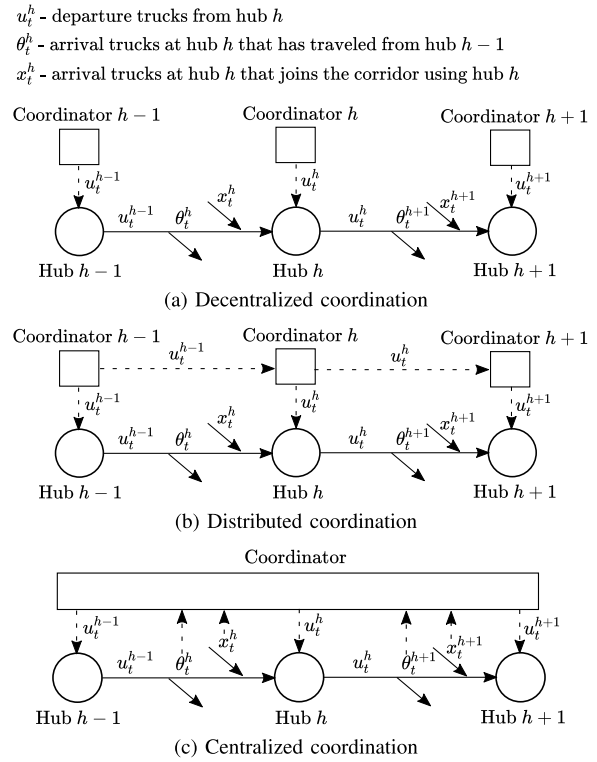


Fig. 2. Coordination under different information structures.

We first develop the optimal release policies in the decentralized case as shown in Fig. 2a, where each coordinator only has access to the history of observed arrivals at its hub, the coordinators do not share any information with each other, and the arrival schedules of trucks are not known *a priori*. We next develop coordination strategies in the distributed and centralized cases. In the distributed case, each coordinator has access to the release decisions of its preceding hub, whereas in the centralized case, the coordinators share their release decisions with each other and the arrival schedules of trucks are known by the coordinators.

The main contributions in this paper are as follows:

- We investigate the structure of the optimal decentralized release policy at a hub when its truck arrivals, *i.e.*, the number of trucks arriving at the hub, are independent over time. Here, we show that a thresholding release policy is optimal. That is, it is optimal for the coordinator to release all the trucks located at the hub when the number of trucks is more than a threshold which is computed by dynamic programming (DP). We use this result to develop approximate thresholding decentralized release policies in cases where truck arrivals are not independent over time by approximating the truck arrivals as independent.
- We study the structure of the optimal decentralized release policy at a hub when the truck arrivals to its preceding hub are independent over time and the preceding hub follows an optimal release policy. We show that the optimal decentralized release policy, in this case, is in the form of a threshold policy which depends on a variable capturing the state of the preceding hub and can be computed using DP. This result is used to develop approximate thresholding decentralized release policies in

cases where truck arrivals at the preceding hub are not independent over time by approximating the truck arrivals at the preceding hub as independent.

- We propose solutions for the distributed and centralized cases, which are receding horizon solutions computed by DP.
- We perform a simulation study for hub-based platoon formation along a highway in Sweden, where the decentralized release policies are compared with the distributed and centralized release policies. The simulation study shows that the performance gap between the decentralized release policies and the distributed and centralized release policies is less than 8%. This observation indicates that the decentralized policies are potential solutions for hub-based platooning since they do not require communication between hubs and achieve a similar performance compared with distributed and centralized policies.

D. Outline

The outline of this paper is as follows. In Section II, the system model of the hub-corridor is formulated, including the arrival process and the platoon release problem at each hub. In Section III, the decentralized platoon release problem at a hub is studied under an independent-arrival assumption. In Section IV, we also study the decentralized platoon release problem at a hub, but under the assumption that the preceding hub has independent arrivals and follows the developed release policy in Section III. In Section V, release policies for distributed and centralized cases are developed. In Section VI, the proposed release policies are evaluated in a simulation study over a hub-corridor along a Swedish highway. Finally, the paper is concluded in Section VII.

II. SYSTEM MODEL

We consider a set of hubs in a transportation network, as shown in Fig. 1a, where trucks can wait for other trucks and form platoons. Examples of hubs include drivers resting areas, gas stations, and harbors. In our set-up, each hub has a platoon coordinator that decides whether the available trucks at the hub will depart as a platoon or wait for more trucks to join. We refer to the set of hubs located along a highway as a hub corridor, which is represented by the thick line in Fig. 1a. The grey lines in Fig. 1a represent roads used by trucks to join or leave the hub-corridor. The hubs are enumerated as 1 to H .

Consider the section of the hub-corridor that includes hubs $h - 1$ and h , as shown in Fig. 1b. The coordinator at hub h decides how many trucks to release from the hub at each time step. The released trucks depart from the hub in the form of a platoon. Let $u_t^h \in \mathbb{Z}_{\geq 0}$ denote the number of trucks released by the coordinator at hub h at time step t . The state of a hub is defined as the number of trucks located at the hub. We use $n_t^h \in \mathbb{Z}_{\geq 0}$ to denote the state of hub h at time step t . The state of each hub changes dynamically based on the truck arrivals and the release decisions, as indicated in Fig. 1b. The state of hub h has the following dynamics

$$n_{t+1}^h = n_t^h - u_t^h + x_{t+1}^h + \theta_{t+1}^h, \quad (1)$$

where $\theta_t^h \in \mathbb{Z}_{\geq 0}$ is the number of trucks that were released as a platoon at hub $h - 1$ and arrive at hub h at time step t , and $x_t^h \in \mathbb{Z}_{\geq 0}$ is the number of trucks that join the hub-corridor via hub h at time step t . The state variable n_t^h is a realization of the random variable denoted N_t^h .

Remark 1: Note that the trucks that arrive at the hub h at time step t can be divided into two groups. The first group is the trucks released as a platoon at hub $h - 1$ and arrive at hub h . We denote the number of trucks in this group by θ_t^h . The second group of trucks is those that enter the hub-corridor via hub h . The number of trucks in this group is denoted by x_t^h . For example, in Fig. 1a, the trucks labeled with A and B join the hub-corridor via the second hub, and the trucks labeled with C and D which arrive at the second hub were released by the first hub as a platoon. In practice, the coordinator at hub h can distinguish between these two groups of trucks by requiring the second group of trucks to send an acknowledgment message to the hub coordinator when they join the corridor. Thus, if the coordinator at hub h does not receive a message from a truck, it knows that the truck was released from hub $h - 1$.

We assume that the truck arrivals are unknown to the coordinators *a priori*. This is realistic if carriers or trucks keep their routes and schedules private or when the travel times are uncertain. To capture the uncertainty in the number of trucks arriving at hubs, the arrival variables x_t^h and θ_t^h are assumed to be realizations of random variables denoted X_t^h and Θ_t^h , respectively. Throughout the paper, we assume that the random variables X_0^h, X_1^h, \dots are independent as these trucks have not been participating in a platoon formation process at a preceding hub, but the independent arrival assumption will in general not hold for the truck arrivals at hub h that have been released at hub $h - 1$. The dependency between the hubs $h - 1$ and h captures that a truck that is released at the hub $h - 1$ may exit the highway before reaching the hub h . We use l^h to denote the likelihood of any truck exiting the highway between the hubs. Moreover, recall that the state variable n_t^h is a realization of the random variable N_t^h , and the release decision u_t^h is a mapping from n_t^h . Thus, the release decision u_t^h can also be viewed as a random variable denoted by U_t^h .

The platoon release problem in the hub-corridor is defined as

$$\max_{\{\mu_0^h, \dots, \mu_T^h\}_h} \mathbb{E} \left[\sum_{h=1}^H \sum_{t=0}^T R^h(N_t^h, U_t^h) \right], \quad (2)$$

where T is the time horizon, and $R^h(n_t^h, u_t^h)$ denotes the reward of the coordinator of hub h for releasing u_t^h trucks when its state is n_t^h . This reward can include the total cost-saving due to a platoon of length u_t^h as well as the waiting cost of $n_t^h - u_t^h$ trucks at the hub. The release decision u_t^h is computed using the release policy μ_t^h based on the available information to the coordinator at hub h , at time step t .

In this paper, we assume that the reward function $R^h(n_t^h, u_t^h)$ is convex in u_t^h and n_t^h . An important example of the reward function is the following convex piecewise-linear form:

$$R^h(n_t^h, u_t^h) = \max \left(0, b^h(u_t^h - 1) \right) - c(n_t^h - u_t^h), \quad (3)$$

where the first term is the platooning profit for releasing u_t^h trucks that form a platoon with $u_t^h - 1$ follower trucks and b^h is the profit per follower truck. The second term is the cost for having $n_t^h - u_t^h$ trucks waiting at the hub, and c is the waiting cost per truck.

This paper focuses on decentralized solutions to the platoon release problem in (2), where the release decisions at each hub only depend on its local state and information. Computing the optimal decentralized release policies in (2) becomes prohibitively difficult as the number of hubs increases. The difficulty in computing an optimal decentralized solution to the problem (2) arises because of the nested information structure of the hubs; each hub has to use its history of observed arrivals to estimate the history of all other hubs' observed arrivals and history of release decisions, and note, the variations of possible histories of observed arrivals and release decisions are numerous even for a few hubs. We refer the reader to [26] and [27] for works on decentralized optimal control under nested information structures. To solve this problem, in this paper, we decompose the optimization problem in (2) into H sub-problems, one for each hub. In the rest of the paper, we will explore different decentralized solutions to the platoon release problem based on our decomposition approach.

Remark 2: During off-peak hours when trucks are unaffected by congestion, the travel times have low variability, and assuming that travel times are deterministic and known a priori is justified. However, travel times may be uncertain during peak hours, and the travel times between hubs can then be modeled as stochastic. Our set-up can be used to study platoon formation under both deterministic and stochastic travel-time scenarios.

Remark 3: The reward function in the convex piecewise-linear form in (3) is accurate if the incremental platooning profit is equal for each follower truck in a platoon and deterministic and constant over the day. Moreover, the cost of waiting at the hub is linear and the same for all trucks. These assumptions may not hold due, for example, to the influence of factors such as traffic conditions and heterogeneous truck properties. However, our reward function in (3) can be applied to a case with heterogeneous trucks where trucks have different platooning benefits. One way is to use the average platooning benefit of the trucks in our problem formulation. The reward function can also be extended to capture the impact of traffic by letting the platooning benefit be a time-varying (deterministic) signal where the benefit is low during peak hours and low during off-peak hours.

III. DECENTRALIZED RELEASE POLICY: SINGLE-HUB APPROACH

In this section, we propose a solution for the decentralized platoon release problem in (2) by decomposing it into H decoupled sub-problems, where each hub maximizes its own reward from platooning and irrespective of the decision-making behavior of other hubs. The decentralized platoon release problem at hub h is

$$\max_{\mu_0^h, \dots, \mu_T^h} \mathbb{E} \left[\sum_{t=0}^T R^h(N_t^h, U_t^h) \right], \quad (4)$$

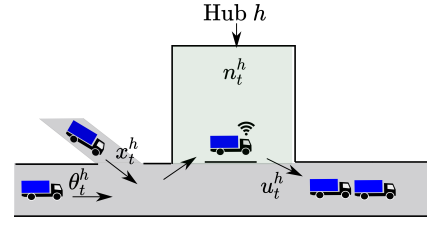


Fig. 3. Single-hub model.

where the objective is to maximize the reward of hub h from platooning. In the next subsection, we study the structure of the optimal release policy at a hub, under the single-hub approach, when its truck arrivals are independent over time. Later, we will use this result to derive an approximate solution for the single-hub approach when truck arrivals are dependent over time.

A. Independent Arrival Case

In this sub-section, we study the structure of the optimal release policy at hub h under the single-hub approach and its arrivals are independent over time, that is, the random variables $X_0^h, \dots, X_T^h, \Theta_0^h, \dots, \Theta_T^h$ are assumed to be independent. Fig. 3 illustrates the single-hub approach. To study the structure of the optimal release policy at hub h , we first derive the Bellman optimality equation associated with the optimal release policy of hub h . Using the Bellman's principle of optimality, the optimal value function associated with the state n_t^h can be expressed as

$$V_t^h(n_t^h) = \max_{u_t^h \in \mathcal{U}_t^h(n_t^h)} R^h(n_t^h, u_t^h) + \mathbb{E}[V_{t+1}^h(n_t^h - u_t^h + X_{t+1}^h + \Theta_{t+1}^h)],$$

where the decision variable u_t^h can take values in the set $\mathcal{U}_t^h(n_t^h) = \{0, \dots, n_t^h\}$ as the maximum number of trucks to release is the number of trucks at the hub. The value function at the terminal time step is $V_T^h(n_T^h) = R^h(n_T^h, n_T^h)$, which is the reward for releasing the remaining trucks. The optimal value function can be computed using standard dynamic programming (DP), and the optimal decision can then be computed using the optimal value function. We omit to present the standard DP algorithm but refer the reader to [28] for a detailed presentation of the standard DP algorithm.

The following theorem gives fundamental results of the structure of the optimal release policy under the single-hub approach when the truck arrivals are independent. The structural results significantly reduce the complexity of computing the optimal policy.

Theorem 1: Consider the decentralized platoon release problem in (4) when the truck arrivals are independent. The following holds:

- The value function $V_t^h(n_t^h)$ is convex in n_t^h .
- The optimal release decision $u_t^{h,*}$ takes values in the set $\{0, n_t^h\}$.

Furthermore, if the reward function has the piecewise-linear form in (3), then there exists a threshold $\rho_t^h \in \mathbb{Z}_{\geq 0}$ such that

$$u_t^{h,*} = \begin{cases} n_t^h & \text{if } n_t^h \geq \rho_t^h \\ 0 & \text{if } n_t^h < \rho_t^h. \end{cases}$$

The threshold ρ_t^h can be computed using dynamic programming.

Proof: See Appendix A. \square

According to Theorem 1, the optimal decision of the coordinator of the hub h is to either to release all its truck or release no trucks. This property is due to the convexity of the reward function, and significantly reduces the search space of optimal release policies, thus leading to a light computational load when solving the Bellman equation by DP. Moreover, a release policy with the structure in Theorem 1 can be used as a low complexity approximate release policy when the truck arrivals are dependent. In such cases, the dependent arrivals can be approximated with an independent arrival process. Then, an approximate release policy can be computed using the structure of Theorem 1 and the approximated arrival process.

B. Dependent Arrival Case

The independent arrival assumption in the previous sub-section is valid if hub h is the only hub in the hub-corridor or if all trucks join the corridor using hub h . The arrivals at hub h will be dependent if, for example, the arriving trucks have participated in a platoon formation process at hub $h-1$. In order to use the solution in Theorem 1 for the dependent arrival case, we approximate the joint distribution of the truck arrival process with the joint distribution of an independent arrival process. That is, we approximate the random variables $\Theta_0^h, \dots, \Theta_T^h$ as independent. An approximate release policy can be computed, as in the previous subsection, using the approximate arrival process with independent arrivals.

Algorithm 1 shows the procedure for computing the single-hub approximate release policy, where for each hub h , one at a time, we first compute the empirical distribution of its arrivals based on arrival data during several episodes (each episode spanning over the time steps 0 to T). To this end, we divide the time steps 0 to T into N intervals denoted $\mathcal{T}_0, \dots, \mathcal{T}_N$. Then, we approximate the distribution of Θ_t^h , for $t \in \mathcal{T}_j$, using the following equation

$$\Pr(\Theta_t^h = \theta) = \frac{\sum_{s \in \mathcal{T}_j} \sum_{e \in \mathcal{E}} \mathbf{1}_{\theta_{s,e}^h = \theta}}{\sum_{s \in \mathcal{T}_j} \sum_{e \in \mathcal{E}} 1}, \quad (5)$$

where the set of episodes is denoted $\mathcal{E} = \{1, \dots, E\}$ and E is the number of episodes, $\theta_{s,e}^h$ is the number of trucks in episode e that have traveled from hub $h-1$ and arrive at hub h at time step s , and $\mathbf{1}_{\theta_{s,e}^h = \theta}$ is the indicator function such that $\mathbf{1}_{\theta_{s,e}^h = \theta} = 0$ if $\theta_{s,e}^h \neq \theta$ and $\mathbf{1}_{\theta_{s,e}^h = \theta} = 1$ if $\theta_{s,e}^h = \theta$. We then use the empirical distributions of $\Theta_0^h, \dots, \Theta_T^h$ to compute an approximate release policy for hub h . Note that the empirical estimator can be applied to either simulated or

real arrival data and recall that the distributions of the random variables X_0^h, \dots, X_T^h are independent as these trucks have not participated in a platoon formation process at a preceding hub.

Algorithm 1 Computation of the Single-Hub Approximate Release Policies

input : Distribution of $\{X_t^h\}_{t,h}$

output: Approximate release policies $\{\mu_t^h\}_{t,h}$

Generate $\{x_{t,e}^1\}_{t,e}$ using the distribution of $\{X_t^1\}_t$

Compute optimal release policy $\{\mu_t^1\}_t$ using DP and Theorem 1

Compute the releases $\{u_{t,e}^1\}_{t,e}$ using $\{\mu_t^1\}_t$ and $\{x_{t,e}^1\}_{t,e}$ for each episode e

for $h = 2, \dots, H$ **do**

 Generate $\{x_{t,e}^h\}_{t,e}$ using the distribution of $\{X_t^h\}_t$

 Generate arrivals $\{\theta_{t,e}^h\}_{t,e}$ using $\{u_{t,e}^{h-1}\}_{t,e}$ for each episode e

 Approximate the distribution of $\{\Theta_t^h\}_t$ using (5)

 Compute approximate release policy $\{\mu_t^h\}_t$ using DP and Theorem 1

 Compute the releases $\{u_{t,e}^h\}_{t,e}$ using $\{\mu_t^h\}_t$, $\{\theta_{t,e}^h\}_{t,e}$ and $\{x_{t,e}^h\}_{t,e}$ for each episode e

end

IV. DECENTRALIZED RELEASE POLICY: TWO-HUB APPROACH

In this section, we propose a two-hub approach for the decentralized platoon release problem in (2) where the release problem is decomposed into H decoupled sub-problems. Under the two-hub approach, each hub finds its optimal release policy based on the release policy of its preceding hub. In the following subsections, we first study the structure of the optimal release policy of a hub when its preceding hub has independent arrivals and the preceding hub follows the release policy in Theorem 1. We will then derive an approximate release policy for a hub when the arrivals to its preceding hub are not independent in time.

A. Independent Arrival Case

In this sub-section, we study the structure of the optimal release policy at hub h when the arrivals to hub $h-1$ are independent in time and hub $h-1$ follows the optimal release policy in Theorem 1. The two-hub model is illustrated in Fig. 4. The random variables $\Theta_0^{h-1}, \dots, \Theta_T^{h-1}$, $X_0^{h-1}, \dots, X_T^{h-1}$ are the independent arrivals at hub $h-1$. The arrivals of trucks at hub h from hub $h-1$, i.e., denoted by the random variables $\Theta_0^h, \dots, \Theta_T^h$, are dependent in time since hub $h-1$ follows the optimal release policy in Theorem 1. The random variables X_0^h, \dots, X_T^h are independent as the trucks that join the hub-corridor using hub h have not participated in a platoon formation process at a preceding hub.

To study the structure of the optimal release policy at hub h , we derive the Bellman optimality equation. Using the

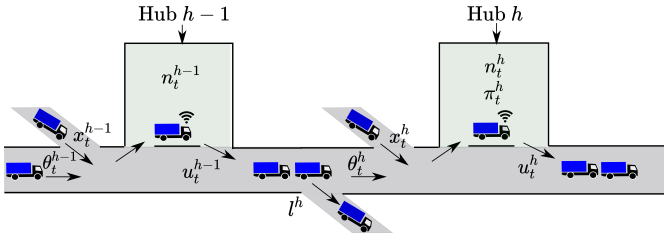


Fig. 4. Two-hub model.

Bellman's principle of optimality, the optimal value function can be expressed as

$$\begin{aligned} V_t^h(n_t^h, \pi_t^h) &= \max_{u_t^h \in \mathcal{U}_t^h(n_t^h)} R^h(n_t^h, u_t^h) \\ &\quad + E[V_{t+1}^h(n_{t+1}^h - u_t^h + X_{t+1}^h + \Theta_{t+1}^h, \pi_{t+1}^h)], \end{aligned} \quad (6)$$

where $\pi_t^h = \{\pi_{t,j}^h\}_{j \in \mathbb{Z}_{\geq 0}}$, and $\pi_{t,i}^h$ is defined as

$$\pi_{t,i}^h = \Pr(D_{t-k^h}^{h-1} = i | \Theta_0^h = \theta_0^h, \dots, \Theta_t^h = \theta_t^h), \quad (7)$$

where $D_{t-k^h}^{h-1} = N_{t-k^h}^{h-1} - U_{t-k^h}^{h-1}$ is the number of trucks at hub $h-1$ at time step $t-k^h$ after releasing, and k^h is the travel time between the hubs $h-1$ and h . Thus, $\pi_{t,i}^h$ is the likelihood that i trucks are at hub $h-1$ after releasing at time step $t-k^h$ given the observed arrivals at hub h . The variable $\pi_{t+1}^h = \{\pi_{t+1,j}^h\}_{j \in \mathbb{Z}_{\geq 0}}$ can be computed recursively as

$$\pi_{t+1,j}^h = \frac{\sum_{i \geq 0} \pi_{t,i}^h p_{t,i,j}^h r_{t,i,j}^h(\theta_{t+1}^h)}{\sum_{i,j \geq 0} \pi_{t,i}^h p_{t,i,j}^h r_{t,i,j}^h(\theta_{t+1}^h)}, \quad (8)$$

where $p_{t,i,j}^h = \Pr(D_{t-k^h+1}^{h-1} = j | D_{t-k^h}^{h-1} = i)$ and $r_{t,i,j}^h(\theta_{t+1}^h) = \Pr(\Theta_{t+1}^h = \theta_{t+1}^h | D_{t-k^h+1}^{h-1} = j, D_{t-k^h}^{h-1} = i)$. The probabilities $p_{t,i,j}^h$ and $r_{t,i,j}^h(\theta_{t+1}^h)$ are determined by the truck arrival distributions, the likelihood of leaving the corridor between the hubs, and the release policy of hub $h-1$, and their precise forms are given in Appendix B. The optimal value function at the terminal time step is $V_T^h(n_T^h, \pi_T^h) = R^h(n_T^h, n_T^h)$.

The Bellman optimality equation in the form of (6) is difficult to solve by DP since π_t^h is a continuous variable. In the rest of this subsection, we will derive the structure of the optimal value function and the optimal release policy that allows us to easily solve the Bellman optimality equation by DP. To this end, let $w_t^h \in \mathbb{Z}_{\geq 0}$ denote the number of time steps since a non-zero truck arrival at hub h from hub $h-1$, which can be computed recursively as

$$w_{t+1}^h = \begin{cases} w_t^h + 1 & \text{if } \theta_{t+1}^h = 0 \\ 0 & \text{if } \theta_{t+1}^h \neq 0. \end{cases}$$

Theorem 2: Consider the decentralized platoon release problem at hub h when the truck arrivals at hub $h-1$ are independent over time and its coordinator follows the release policy in Theorem 1. Then, the following statements hold:

- The optimal value function only depends on n_t^h and w_t^h .
- The value function $V_t^h(n_t^h, w_t^h)$ is convex in n_t^h .

- The optimal release decision $u_t^{h,*}$ takes values in the set $\{0, n_t^h\}$.

Furthermore, if the reward function has the piecewise-linear form in (3), then there exists a threshold $\rho_t^h(w_t^h)$, such that

$$u_t^{h,*} = \begin{cases} n_t^h & \text{if } n_t^h \geq \rho_t^h(w_t^h) \\ 0 & \text{if } n_t^h < \rho_t^h(w_t^h). \end{cases}$$

The threshold $\rho_t^h(w_t^h)$ can be computed using dynamic programming.

Proof: See Appendix C. \square

Theorem 2 implies that the optimal release policy of hub h in the two-hub approach can be computed by solving the Bellman optimality equation using standard DP using w_t^h instead of π_t^h . Note that π_t^h is a continuous random variable which significantly complicates the computation of the optimal policy using DP. However, based on Theorem 2, the optimal release policy can be computed by using w_t^h which is a discrete variable. This significantly facilitates the computation of the optimal policy. Theorem 2 also indicates that similar structural results continue to hold for the optimal release policy as those in Theorem 1 when the arrivals at hub $h-1$ are independent and the coordinator at hub $h-1$ follows a release policy with the structure in Theorem 1. This result is useful as it limits the candidate optimal release policies when seeking an optimal release policy by DP.

B. Dependent Arrival Case

The two-hub solution in the previous subsection is optimal under the assumptions that the arrivals at the preceding hub $h-1$ are independent and its coordinator follows the release policy in Theorem 1. The independent-arrival assumption at hub $h-1$ is valid if hub $h-1$ is the first hub in the corridor. In this subsection, we propose an approximate two-hub solution for hub h by approximating the joint distribution of the arrival process at hub $h-1$ using independent arrivals. That is, we approximate the random variables $\Theta_0^{h-1}, \dots, \Theta_T^{h-1}$ as independent. The approximate distribution allows us to use Theorem 2 to obtain an approximate release solution for the dependent case. The approximate two-hub release policy at hub h is computed using the approximate arrival distribution at hub $h-1$ and DP, as explained in the previous subsection.

Algorithm 2 shows the procedure for computing the two-hub approximate release policy, where for each hub h , we first compute the empirical distribution of the arrivals at hub $h-1$ using equation (5). The empirical distribution of the arrivals is then used as an approximation of the arrival distribution at hub $h-1$. The empirical distribution of $\Theta_0^{h-1}, \dots, \Theta_T^{h-1}$ is used to compute an approximate release policy for hub h .

V. DISTRIBUTED AND CENTRALIZED RELEASE POLICIES

In this section, we propose distributed and centralized release policies, which will be used as benchmarks in Section VI to evaluate the performance of the proposed decentralized release policies and investigate the value of communication among hubs. Under the distributed release

Algorithm 2 Computation of the Two-Hub Approximate Release Policies

input : Distribution of $\{X_t^h\}_{t,h}$
output: Approximate release policies $\{\mu_t^h\}_{t,h}$

Generate $\{x_{t,e}^1\}_{t,e}$ using the distribution of $\{X_t^1\}_t$
 Compute optimal release policy $\{\mu_t^1\}_t$ using DP and Theorem 1
 Generate releases $\{u_{t,e}^1\}_{t,e}$ using $\{\mu_t^1\}_t$ and $\{x_{t,e}^1\}_{t,e}$ for each episode e
 Generate $\{x_{t,e}^2\}_{t,e}$ using the distribution of $\{X_t^2\}_t$
 Compute optimal release policy $\{\mu_t^2\}_t$ using DP and Theorem 2

for $h = 3, \dots, H$ **do**
 Generate $\{x_{t,e}^h\}_{t,e}$ using the distribution of $\{X_t^h\}_t$
 Generate arrivals $\{\theta_{t,e}^{h-1}\}_{t,e}$ using $\{u_{t,e}^{h-2}\}_{t,e}$
 Approximate distribution of $\{\Theta_t^{h-1}\}_t$ using (5)
 Compute approximate release policy $\{\mu_t^h\}_t$ using DP and Theorem 2
 Generate releases $\{u_{t,e}^h\}_{t,e}$ using $\{\mu_t^h\}_t$, $\{\theta_{t,e}^{h-1}\}_{t,e}$ and $\{x_{t,e}^{h-1}\}_{t,e}$ for each episode e
end

policy, each coordinator informs its proceeding coordinator in the hub-corridor about its release decisions. Under the centralized release policy, coordinators share their release decisions and have *a priori* knowledge of their arrivals.

A. Distributed Case

In the distributed case, the coordinator at each hub receives the release decisions of its preceding hub and computes its release decisions using a receding horizon solution with L time steps as the horizon. The distributed release policy requires communication links between hubs which will require installation and maintenance, whereas the decentralized release policies do not require communication among hubs. Let $\mathcal{M}_t^h = \{u_{t-k^h+1}^{h-1}, \dots, u_{t-k^h+L}^{h-1}\}$ denote the release decisions of hub $h-1$ which are known by hub h at time step t . Then, the Bellman optimality equation for hub h can be expressed as

$$V_s^h(n_s^h, \mathcal{M}_t^h) = \max_{u_s^h \in \mathcal{U}_s^h(n_s^h)} R(n_s^h, u_s^h) + E[V_{s+1}^h(n_s^h - u_s^h + X_{s+1}^h + \Theta_{s+1}^h, \mathcal{M}_t^h) | \mathcal{M}_t^h],$$

for $s = t, \dots, t+L-1$, and the value function at the terminal time step of the horizon is $V_{t+L}^h(n_{t+L}^h) = R^h(n_{t+L}^h, n_{t+L}^h)$. The Bellman optimality equation is solved recursively by DP. Then, the computed release decision for time step t is implemented, while the set of releases at the preceding hub and the computed release decisions for time steps $t+1, \dots, t+L$ are updated at the next step.

B. Centralized Case

In the centralized case, the coordinator at each hub computes its release decisions using a receding horizon solution

where the arrivals within the horizon of L time steps are known. The arrivals are known as the coordinators sharing their release decisions in the centralized case, and the trucks inform the coordinators about their arrivals beforehand. The centralized release policy requires knowledge of the schedule and the route of trucks and communication between hubs. However, the decentralized release policies do not rely on any information exchange between hubs. The set of information available to the coordinator of hub h at time step t is denoted as $\mathcal{N}_t^h = \{x_{t+1}^h, \dots, x_{t+L}^h, \theta_{t+1}^h, \dots, \theta_{t+L}^h\}$. The Bellman optimality equation used by the coordinator at hub h to compute the release decision at time step t is expressed as

$$V_s^h(n_s^h, \mathcal{N}_t^h) = \max_{u_s^h \in \mathcal{U}_s^h(n_s^h)} R(n_s^h, u_s^h) + V_{s+1}^h(n_s^h - u_s^h + x_{s+1}^h + \theta_{s+1}^h, \mathcal{N}_t^h),$$

for $s = t, \dots, t+L-1$, and the value function at the terminal time step of the horizon is $V_{t+L}^h(n_{t+L}^h) = R^h(n_{t+L}^h, n_{t+L}^h)$. The Bellman optimality equation is solved recursively by DP, and similar to the distributed case, the release decision for time step t is implemented, while the set of information and the release decisions for time steps $t+1, \dots, t+L$ are updated at the next time step.

VI. SIMULATION STUDY

In this section, we perform a simulation study over a hub-corridor with three hubs in northern Sweden. We first explain the simulation setup, including the hub locations and arrival distributions. Then, we study the performance of the developed coordination solutions under deterministic travel times, which is accurate if the impact from surrounding traffic on the highway is ignored. We also study the impact of uncertainty in travel times between the hubs on the performance of the coordination policies.

A. Setup

Fig. 5 shows a hub corridor between Luleå and Sundsvall in northern Sweden. The hubs are located near the cities Luleå, Skellefteå, Umeå, and the hub-corridor ends near the city of Sundsvall, Sweden. The length of the road segments connecting the hubs are 131 km, 136 km, and 263 km, and we assume that trucks travel with a speed of 80 km/h under free-flow conditions. To obtain realistic arrival distributions for the trucks that join the corridor, we use the real data shown in Fig. 6 which were collected during two working days in 2018–2019 by the Swedish Transport Administration [29]. This figure shows the average hourly truck count on roads that connect to each hub location in Fig. 5 (excluding the roads connecting Luleå, Skellefteå, Umeå, and Sundsvall).

In our simulations, we assume that the number of trucks joining the corridor using hub h at time t , i.e., X_t^h is Poisson distributed with a time-varying mean. More precisely, $X_t^h \sim \text{Poisson}(\lambda_t^h)$, for $t = 0, \dots, T$, where the mean of the process, i.e., λ_t^h , is set according to the data in Fig. 5. The time step length is set to one minute in the simulations. We will study the performance of the proposed decentralized policies under both known travel-times and uncertain travel-times.



Fig. 5. Three-hub-corridor along a highway in Sweden.

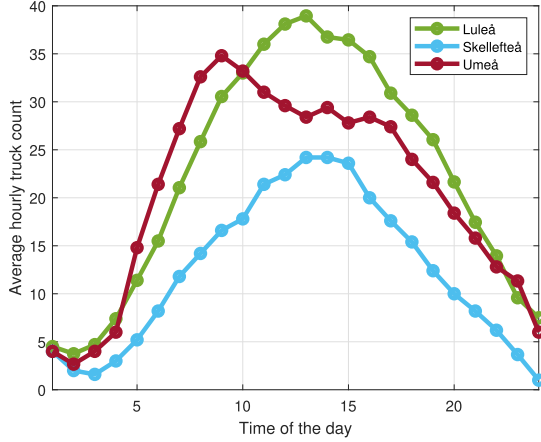
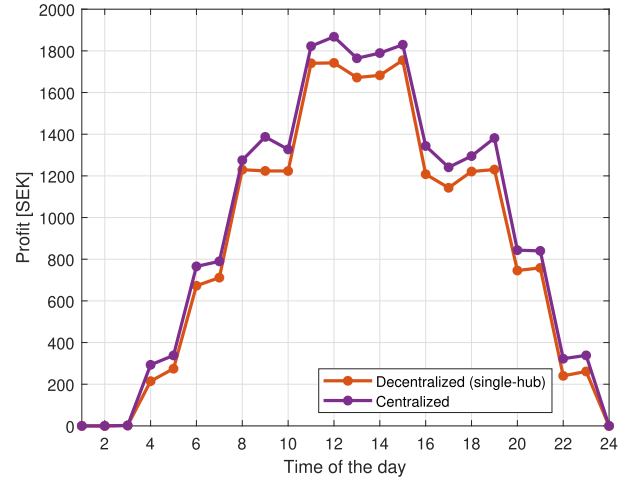


Fig. 6. The average hourly truck count on roads connecting to Luleå, Skellefteå, and Umeå (excluding roads connecting Luleå, Skellefteå, Umeå, and Sundsvall). The data was collected by the Swedish Transport Administration [29].

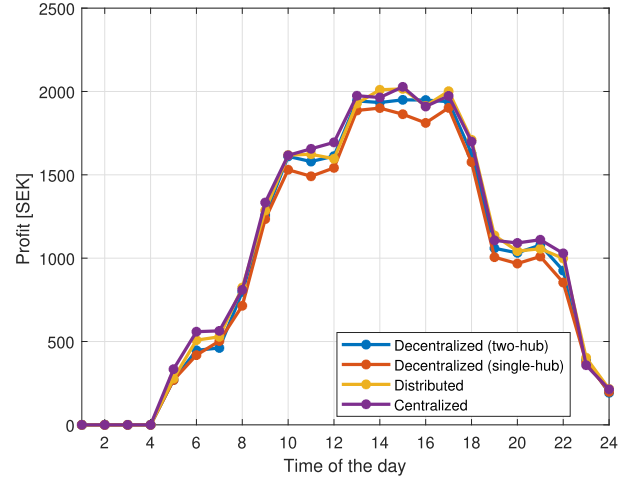
We use the reward function in (3) for our simulations, where the platooning benefit (b^h) is 10% fuel consumption reduction for the follower trucks. The cost of fuel is assumed to be 5 SEK/km in our simulations. The waiting cost per truck in our simulations is equal the hourly driver cost, which is around 200 SEK/h [30], or approximately $c = 3.33$ SEK per time step. The following results are generated by 50 Monte Carlo simulations.

B. Deterministic Travel Times Scenario

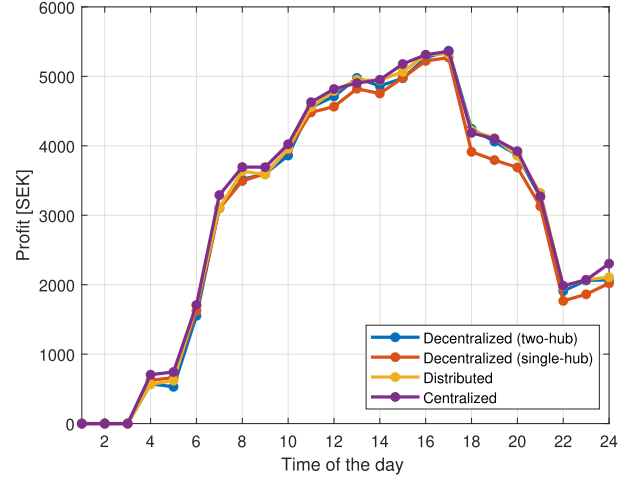
In this subsection, we study the performance of different platoon release policies in the deterministic travel time scenario where the travel times between hubs are known *a priori* and set to the free-flow travel times. Fig. 7 shows the (hourly) average reward in a period of 24 hours under different release policies. According to this figure, the decentralized two-hub solution generally results in a higher average hourly reward than the decentralized single-hub solution. This is because the former policy takes the arrival process and release behavior of its preceding hub into account. Fig. 7 also shows that



(a) First hub



(b) Second hub



(c) Third hub

Fig. 7. Profits as a function of time. In the simulations, the likelihood of any released truck to leave the corridor before the next hub is set to $l^h = 0.5$.

the distributed and centralized release policies outperform the decentralized release policies by a small margin. This indicates that the performance loss due to the decentralized structure of the one-hub and two-hub solutions is relatively small.

Fig. 8 shows the (hourly) average number of released trucks and its standard deviation for each hub. The number

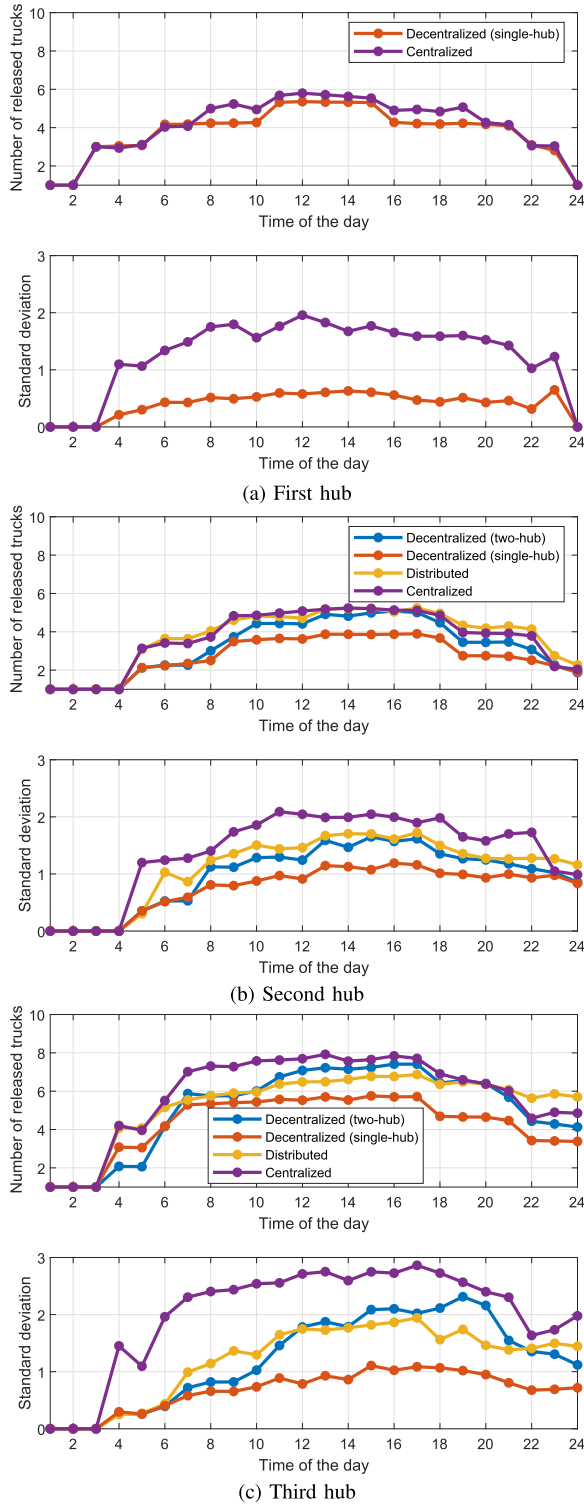


Fig. 8. Platoon length and its standard deviation as a function of time. In the simulations, the likelihood of any released truck to leave the corridor before the next hub is set to $l^h = 0.5$.

of released trucks is referred to as release size in this figure. According to this figure, more trucks are released on average under the decentralized two-hub policy than the decentralized single-hub policy. The standard deviation of the release sizes is also higher for the decentralized two-hub policy. The increased variability in release sizes is due to that the threshold for the

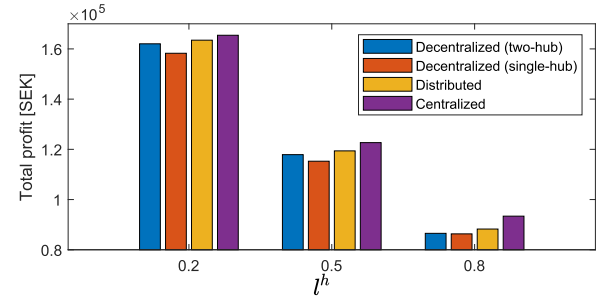


Fig. 9. Total profit as a function of the likelihood of any released truck to leave the corridor before the next hub.

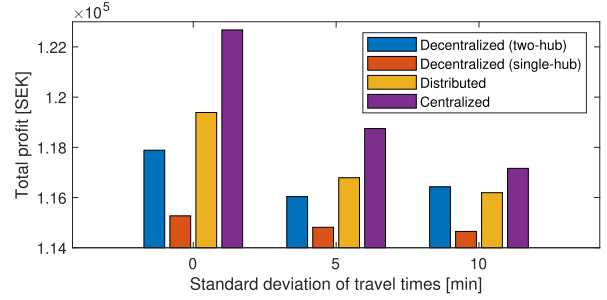


Fig. 10. Total profit as a function of the standard deviation of the travel times. In the simulations, the likelihood of any released truck to leave the corridor before the next hub is set to $l^h = 0.5$.

decentralized two-hub policy depends on the variable w_t^h , i.e., the number of time steps since a non-zero arrival was observed, as shown in Theorem 2. Fig. 8 also shows that the distributed and centralized policies generally have larger release sizes and higher standard deviations than the decentralized policies. This implies that the centralized and distributed policies are more efficient in forming platoons than the decentralized policies.

Fig. 9 shows the average daily reward generated at the hubs as a function of the variable l^h , i.e., the likelihood that a truck leaves the hub-corridor before reaching the next hub in the hub-corridor. Based on this figure, the daily reward decreases as the variable l^h increases. This is because the trucks will visit fewer hubs on average as l^h becomes large, reducing the number of trucks participating in platoons. Moreover, Fig. 9 shows that the difference between the decentralized and distributed policies is small when l^h becomes large. This figure also shows that the centralized release policy achieves higher rewards even when l^h is high. The largest difference between the decentralized and centralized policies is 8% and occurs at $l^h = 0.8$.

C. Uncertain Travel Times Scenario

In this subsection, we will study the impact of travel-time uncertainty on the performance of different release policies. To this end, we model the travel times between the hubs as truncated Gaussian distributed random variables with the free-flow travel times between hubs as the mean of the distribution. The standard deviation of the travel times is varied to study the impact of travel time uncertainty. Fig. 10 shows the average daily reward of the hubs as a function of the standard deviation of the travel-times. According to Fig. 10, the reward slightly decreases with increased uncertainty in

travel times. The performance loss of decentralized policies is less than 2% when the standard deviation of the travel-time is 10 minutes. Also, the performance loss of the distributed and centralized policies is less than 5% when the standard deviation of the travel time is 10 minutes.

VII. CONCLUSION AND FUTURE WORK

This paper studied the platoon release problem in a hub-corridor, in which release times at hubs are coordinated to form platoons. We focused on the decentralized case without communication between coordinators, where the release decisions are based on statistical information of truck arrivals. Under this information structure, we proposed two different decentralized release policies. The first decentralized release policy was shown to be optimal when its arrivals from the preceding hub are independent. The second decentralized release policy is shown to be optimal when the preceding hub has independent arrivals and follows an optimal release policy. In order to evaluate the value of communication among hubs, we also proposed distributed and centralized release policies.

A simulation study of a hub-corridor with three hubs in Sweden was performed to compare the proposed release policies and evaluate the value of sharing information. The simulation study showed that the decentralized release policies achieved profits close to that of the distributed and centralized. More precisely, the profits of the decentralized policy was less than 3.5% and 8% lower than of the distributed and centralized, respectively. This suggests that a viable platooning system can be obtained without sharing truck routes and schedules across the system, which may otherwise be one of the bottlenecks for platoon cooperation among different transportation companies. Moreover, the simulations showed that the likelihood of leaving the system between hubs heavily affects the platoon profits of the proposed release policies. For instance, when the likelihood of leaving the corridor between hubs is high, the benefit of exploiting trucks from the preceding neighboring hub is limited.

For future research, we aim to study platoon release problems in networks with general typologies, expanding the line graph (hub-corridor) considered in this paper. A possibility is to include platoon routing in the decision of the hub coordinators. Another promising future research direction is to use reinforcement learning to seek approximate solutions when the truck arrival process is unknown. Furthermore, another avenue for future work is to investigate if similar properties shown for the platoon release problem in this paper hold when using a more realistic form of the reward function that, for example, captures the influence of traffic and heterogeneous truck properties. Final possible future extensions are to use a traffic simulator, for example, SUMO, to model the travel time uncertainty in the simulation and to include hub and road capacity constraints in the problem formulation.

APPENDIX A PROOF OF THEOREM 1

Before providing the proof, we define

$$G_t^h(n_t^h, u_t^h) = R^h(n_t^h, u_t^h) + E[V_{t+1}^h(n_t^h - u_t^h + X_{t+1}^h + \Theta_{t+1}^h)],$$

and note that

$$V_t^h(n_t^h) = \max_{u_t^h \in \mathcal{U}_t^h(n_t^h)} G_t^h(n_t^h, u_t^h). \quad (9)$$

First, assume that $V_{t+1}^h(n_{t+1}^h)$ is convex in n_{t+1}^h . Then, since convexity is preserved under affine maps and positively weighted summations, we have that $E[V_{t+1}^h(n_t^h - u_t^h + X_{t+1}^h + \Theta_{t+1}^h)]$ is convex in both n_t^h and u_t^h . Therefore, $G_t^h(n_t^h, u_t^h)$ is also convex in both n_t^h and u_t^h . Consider then the relaxed version of the maximization problem in (9) with the decision space $[0, n_t^h]$ instead of $\mathcal{U}_t^h(n_t^h) = \{0, \dots, n_t^h\}$. The relaxed maximization problem is of a convex function over a compact and convex set. Therefore, the maximizer is at one of the boundary points of the relaxed decision space $[0, n_t^h]$, see [31] for proof. Since the boundary points of $[0, n_t^h]$ are elements in $\mathcal{U}_t^h(n_t^h)$, the maximizer of the maximization problem in (9) is also one of these two boundary points. That is, the optimal release decision $u_t^{h,*}$ takes values in the set $\{0, n_t^h\}$. By this fact we have that

$$V_t^h(n_t^h) = \max\{G_t^h(n_t^h, 0), G_t^h(n_t^h, n_t^h)\},$$

which is convex in n_t^h since $G_t^h(n_t^h, 0)$ and $G_t^h(n_t^h, n_t^h)$ are convex functions in n_t^h and the maximizer of two convex functions is convex. Finally, $V_T^h(n_T^h) = R^h(n_T^h, n_T^h)$ is convex by assumption. Thus, $V_t^h(n_t^h)$ is convex for all $t \leq T$ by induction. Conclusions in first part of Theorem 1 follow.

In the second part of Theorem 1, we assume the piecewise-linear form of $R^h(n_t^h, u_t^h)$ in (3). Then we have for $n_t^h \geq 1$ that

$$G_t^h(n_t^h, n_t^h) = n_t^h b^h + F_t,$$

where $F_t = -b^h + E[V_{t+1}^h(X_{t+1}^h + \Theta_{t+1}^h)]$ and

$$G_t^h(n_t^h, 0) = -cn_t^h + E[V_{t+1}^h(n_t^h + X_{t+1}^h + \Theta_{t+1}^h)].$$

Since the optimal release decision $u_t^{h,*}$ takes values in the set $\{0, n_t^h\}$, we have that $u_t^{h,*} = n_t^h$ if and only if

$$\begin{aligned} \Delta(n_t^h) &= G_t^h(n_t^h, n_t^h) - G_t^h(n_t^h, 0) \\ &= n_t^h b^h + F + cn_t^h - E[V_{t+1}^h(n_t^h + X_{t+1}^h + \Theta_{t+1}^h)] \geq 0. \end{aligned} \quad (10)$$

To show that there exists a threshold ρ_t^h such that $u_t^{h,*} = n_t^h$ if and only if $n_t^h \geq \rho_t^h$, we show that the optimal release decision is $u_t^{h,*} = n_t^h \geq 1$ (when the number of trucks at hub h is n_t^h) implies the optimal release decision $u_t^{h,*} = n_t^h + 1$ (when the number of trucks at hub h is $n_t^h + 1$). To show this, it is sufficient to show that $\Delta(n_t^h)$ is non-decreasing in n_t^h . First, assume that there exists a threshold ρ_{t+1}^h such that $u_{t+1}^{h,*} = n_{t+1}^h$ if and only if $n_{t+1}^h \geq \rho_{t+1}^h$. This implies that $V_{t+1}^h(n_{t+1}^h + 1) - V_{t+1}^h(n_{t+1}^h) = b^h$ if $n_{t+1}^h \geq \rho_{t+1}^h$. Since convex functions have increasing differences it follows that

$$V_{t+1}^h(n_{t+1}^h + 1) - V_{t+1}^h(n_{t+1}^h) \leq b^h, \quad (11)$$

for all n_{t+1}^h . Thus, equations (10) and (11) imply

$$\begin{aligned} \Delta(n_t^h + 1) - \Delta(n_t^h) &= b^h + c - E[V_{t+1}^h(n_t^h + 1 + X_{t+1}^h + \Theta_{t+1}^h)] \\ &\quad - E[V_{t+1}^h(n_t^h + X_{t+1}^h + \Theta_{t+1}^h)] \geq b^h + c - b^h > 0. \end{aligned}$$

We regard the threshold $\rho_T^h = 1$ at the terminal time step and the conclusion in the second part of Theorem 1 follows by induction.

APPENDIX B

TRANSITION AND OBSERVATION PROBABILITIES

The optimal release policy at hub $h - 1$ is in the form given in Theorem 1. Consequently, there exist a set $\mathcal{D}_{t-k^h}^{h-1} \subseteq \mathbb{Z}_{\geq 0}$ such that the optimal release decision $u_{t-k^h}^{h-1,*} = n_{t-k^h}^{h-1}$ if $n_{t-k^h}^{h-1} \in \mathcal{D}_{t-k^h}^{h-1}$ and otherwise $u_{t-k^h}^{h-1,*} = 0$. Then, for $i \notin \mathcal{D}_{t-k^h}^{h-1}$ and $j \notin \mathcal{D}_{t-k^h+1}^{h-1}$, we have

$$p_{t,i,j}^h = \begin{cases} \Pr(X_{t-k^h+1}^{h-1} + i = j) & \text{if } j \geq i \wedge j \neq 0 \\ \Pr(X_{t-k^h+1}^{h-1} + i \in \mathcal{D}_{t-k^h+1}^{h-1}) & \text{if } i \neq 0 \wedge j = 0 \\ \Pr(X_{t-k^h+1}^{h-1} \in \mathcal{D}_{t-k^h+1}^{h-1} \vee X_{t-k^h+1}^{h-1} = 0) & \text{if } i = 0 \wedge j = 0 \\ 0 & \text{else,} \end{cases}$$

where the first case is when the number of trucks at hub $h - 1$ after the release decision is non-decreasing due to not releasing, the second case is when the number of trucks at hub $h - 1$ after the release decision becomes zero due to releasing, and the third case is when the number of trucks at hub $h - 1$ after the release decision remains zero due to releasing or no arrivals. Furthermore, for $\theta \neq 0$ and $j = 0$, we have

$$r_{t,i,0}^h(\theta) = \sum_{n \in \mathcal{D}_{t-k^h+1}^{h-1}} \Pr(X_{t-k^h+1}^{h-1} + i = n) \Pr(\Theta_{t+1}^h = \theta | U_{t-k^h+1}^{h-1} = n) / p_{t,i,0}^h,$$

where $\Pr(\Theta_{t+1}^h = \theta | U_{t-k^h+1}^{h-1} = n)$ is determined by l^h , and for $\theta \neq 0$ and $j \neq 0$, we have $r_{t,i,j}^h(\theta) = 0$.

APPENDIX C

PROOF OF THEOREM 2

The optimal release decision at hub $h - 1$, that is $u_t^{h-1,*}$, takes values in the set $\{0, n_t^{h-1}\}$, as shown in Theorem 1. Thus, $r_{t,i,j}^h(\theta_{t+1}^h) = 0$ for $j > 0$ and $\theta_{t+1}^h \neq 0$. This implies $\pi_{t+1,0}^h = 1$ and $\pi_{t+1,j}^h = 0$ for $j > 0$ when $\theta_{t+1}^h \neq 0$. That is, π_t^h is reset to $\pi_t^h = (1, 0, \dots, 0)$ when $\theta_t^h \neq 0$ and the number of time steps since a non-zero arrival was observed (or more precisely w_t^h) is the only variable affecting π_t^h . We can therefore write the bellman equation as

$$V_t(n_t^h, w_t^h) = \max_{u_t^h \in \mathcal{U}_t^h(n_t^h)} R^h(n_t^h, u_t^h) + E[V_{t+1}^h(n_t^h - u_t^h + X_{t+1}^h + \Theta_{t+1}^h, w_{t+1}^h)].$$

The proof of the rest of Theorem 2 is similar to the proof of Theorem 1 and we therefore omit most details but provide the step which is least trivial. If $V_{t+1}^h(n_{t+1}^h, w_{t+1}^h)$ is convex in n_{t+1}^h , then $E[V_{t+1}^h(n_t^h - u_t^h + X_{t+1}^h + \Theta_{t+1}^h, w_{t+1}^h)]$ is convex in n_t^h and u_t^h , and by following similar steps as in Appendix A, this implies that $V_t^h(n_t^h, w_t^h)$ is convex in n_t^h .

Then, $V_t^h(n_t^h, w_t^h)$ is convex for all $t \leq T$ by induction since $V_T^h(n_T^h, w_T^h) = R^h(n_T^h, n_T^h)$ is convex in n_T^h . The remaining part of the proof follows similar steps as the proof of Theorem 1 but including the state w_t^h in the notations.

REFERENCES

- [1] F. Browand, J. McArthur, and C. Radovich, "Fuel saving achieved in the field test of two tandem trucks," Univ. Southern California, Los Angeles, CA, USA, Tech. Rep. UCB-ITS-PRR-2004-20, 2004.
- [2] A. Davila, E. del Pozo, E. Aramburu, and A. Freixas, "Environmental benefits of vehicle platooning," SAE Tech. Paper Ser. 2013-26-0142, Jan. 2013.
- [3] A. Alam, B. Besselink, V. Turri, J. Mårtensson, and K. H. Johansson, "Heavy-duty vehicle platooning for sustainable freight transportation: A cooperative method to enhance safety and efficiency," *IEEE Control Syst. Mag.*, vol. 35, no. 6, pp. 34–56, Dec. 2015.
- [4] S. Tsugawa, S. Jeschke, and S. E. Shladover, "A review of truck platooning projects for energy savings," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 1, pp. 68–77, Mar. 2016.
- [5] R. Bishop, D. Bevely, L. Humphreys, S. Boyd, and D. Murray, "Evaluation and testing of driver-assistive truck platooning: Phase 2 final results," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2615, no. 1, pp. 11–18, Jan. 2017.
- [6] P. A. Ioannou and C. C. Chien, "Autonomous intelligent cruise control," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 657–672, Nov. 1993.
- [7] P. Fernandes and U. Nunes, "Platooning with IVC-enabled autonomous vehicles: Strategies to mitigate communication delays, improve safety and traffic flow," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 91–106, Mar. 2012.
- [8] Y. Jo, J. Kim, C. Oh, I. Kim, and G. Lee, "Benefits of travel time savings by truck platooning in Korean freeway networks," *Transp. Policy*, vol. 83, pp. 37–45, Nov. 2019.
- [9] R. Janssen, H. Zwiijnenberg, I. Blankers, and J. de Kruijff, "Truck platooning," Driving, TNO, Delft, The Netherlands, Tech. Rep. TNO 2014 R11893. [Online]. Available: <https://repository.tno.nl/islandora/object/uuid%3A778397eb-59d3-4d23-9185-511385b91509>
- [10] A. Chottani, G. Hastings, J. Murnane, and F. Neuhaus, *Distraction or Disruption? Autonomous Trucks Gain Ground in U.S. Logistics*. New York, NY, USA: McKinsey Smith Co, 2018. [Online]. Available: <http://https://www.mckinsey.com/industries/travel-transport-and-logistics/our-insights/distraction-or-disruption-autonomous-trucks-gain-ground-in-us-logistics>
- [11] J. Axelsson, T. Bergh, A. Johansson, B. Mårdberg, P. Svenson, and V. Åkesson, "Truck platooning business case analysis," Rise, Stockholm, Sweden, RISE Rep. 2020:07, 2020.
- [12] E. Larsson, G. Sennton, and J. Larson, "The vehicle platooning problem: Computational complexity and heuristics," *Transp. Res. C, Emerg. Technol.*, vol. 60, pp. 258–277, Nov. 2015.
- [13] K.-Y. Liang, J. Mårtensson, and K. H. Johansson, "Heavy-duty vehicle platoon formation for fuel efficiency," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 4, pp. 1051–1061, Apr. 2016.
- [14] S. van de Hoef, K. H. Johansson, and D. V. Dimarogonas, "Fuel-efficient en route formation of truck platoons," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 1, pp. 102–112, Jan. 2018.
- [15] X. Xiong, J. Sha, and L. Jin, "Optimizing coordinated vehicle platooning: An analytical approach based on stochastic dynamic programming," *Transp. Res. B, Methodol.*, vol. 150, pp. 482–502, Aug. 2021.
- [16] W. Zhang, E. Jenelius, and X. Ma, "Freight transport platoon coordination and departure time scheduling under travel time uncertainty," *Transp. Res. E, Logistics Transp. Rev.*, vol. 98, pp. 1–23, Feb. 2017.
- [17] N. Boysen, D. Briskorn, and S. Schwerdfeger, "The identical-path truck platooning problem," *Transp. Res. B, Methodol.*, vol. 109, pp. 26–39, Mar. 2018.
- [18] R. Larsen, J. Rich, and T. K. Rasmussen, "Hub-based truck platooning: Potentials and profitability," *Transp. Res. E, Logistics Transp. Rev.*, vol. 127, pp. 249–264, Jul. 2019.
- [19] A. Johansson, E. Nekouei, K. H. Johansson, and J. Mårtensson, "Strategic hub-based platoon coordination under uncertain travel times," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8277–8287, Jul. 2022.
- [20] T. Bai, A. Johansson, K. H. Johansson, and J. Mårtensson, "Event-triggered distributed model predictive control for platoon coordination at hubs in a transport system," in *Proc. 60th IEEE Conf. Decis. Control (CDC)*, Dec. 2021, pp. 1198–1204.

- [21] A. Johansson, J. Mårtensson, X. Sun, and Y. Yin, "Real-time cross-fleet Pareto-improving truck platoon coordination," in *Proc. IEEE Int. Intell. Transp. Syst. Conf. (ITSC)*, Sep. 2021, pp. 996–1003.
- [22] A. Johansson, V. Turri, E. Nekouei, K. H. Johansson, and J. Mårtensson, "Truck platoon formation at hubs: An optimal release time rule," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 15312–15318, 2020.
- [23] A. Adler, D. Miculescu, and S. Karaman, *Optimal Policies for Platooning and Ride Sharing in Autonomy-Enabled Transportation*. Cham, Switzerland: Springer, 2020, pp. 848–863.
- [24] A. K. Bhoopalam, N. Agatz, and R. Zuidwijk, "Planning of truck platoons: A literature review and directions for future research," *Transp. Res. B, Methodol.*, vol. 107, pp. 212–228, Jan. 2018.
- [25] V. Lesch, M. Breitbach, M. Segata, C. Becker, S. Kounev, and C. Krupitzer, "An overview on approaches for coordination of platoons," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10049–10065, Aug. 2022.
- [26] Y.-C. Ho and K.-C. Chu, "Team decision theory and information structures in optimal control problems-part i," *IEEE Trans. Autom. Control*, vol. AC-17, no. 1, pp. 15–22, Feb. 1972.
- [27] A. Dave and A. A. Malikopoulos, "Decentralized stochastic control in partially nested information structures," *IFAC-PapersOnLine*, vol. 52, no. 20, pp. 97–102, 2019.
- [28] D. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 4. Nashua, NH, USA: Athena Scientific, 2012.
- [29] The Swedish Transport Administration. *Road Flow Map*. Accessed: Jan. 13, 2021. [Online]. Available: <https://vtf.trafikverket.se/SeTrafikinformation>
- [30] *Godstransporter I Sverige En Nulägesanalys Rapport*, Trafikanalys, Stockholm, Sweden, 2016.
- [31] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1970.



Alexander Johansson received the M.Sc. degree in applied mathematics from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2017, and the Ph.D. degree from the Division of Decision and Control Systems, Department of Intelligent Systems, School of Electrical Engineering, KTH Royal Institute of Technology, in 2022. Currently, he is with the FOI Swedish Defence Research Agency. His research interests include optimization, game theory and control for platoon coordination, networked control systems, and robotics.



Ehsan Nekouei (Member, IEEE) is currently an Assistant Professor with the Department of Electrical Engineering, City University of Hong Kong. From 2014 to 2019, he held post-doctoral research positions with the KTH Royal Institute of Technology, Stockholm, Sweden, and The University of Melbourne, Australia. His current research interests include privacy in networked control systems and integrated processing of human decision-making data.



Xiaotong Sun received the Ph.D. degree in civil engineering from the University of Michigan, Ann Arbor, in July 2020. She is currently an Assistant Professor with the Intelligent Transportation Thrust, HKUST (Guangzhou), and an affiliated Assistant Professor with the Civil and Environmental Engineering Department, HKUST. Before joining HKUST, she was a Post-Doctoral Research Fellow with the University of Michigan. Her current research interests lie at the intersection of game theory, optimization, and the management and operations of emerging transportation systems.



Karl Henrik Johansson (Fellow, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree in automatic control from Lund University. He is currently a Swedish Research Council Distinguished Professor in electrical engineering and computer science with the KTH Royal Institute of Technology, Sweden, and the Director of Digital Futures. He has held visiting positions with UC Berkeley, Caltech, NTU, and among other institutions. His research interests are in networked control systems and cyber-physical systems with applications in transportation, energy, and automation networks, areas in which he has coauthored more than 800 journals and conference papers and supervised close to 100 post-doctoral's and Ph.D. students. He is also the President of the European Control Association and a member of the IFAC Council and has served on the IEEE Control Systems Society Board of Governors and the Swedish Scientific Council for Natural Sciences and Engineering Sciences. He has received several best paper awards and other distinctions from IEEE, IFAC, and ACM. He has been awarded Distinguished Professor by Swedish Research Council, Wallenberg Scholar with the Knut and Alice Wallenberg Foundation, Future Research Leader from the Swedish Foundation for Strategic Research, the triennial IFAC Young Author Prize, and IEEE Control Systems Society Distinguished Lecturer. He is a fellow of the Royal Swedish Academy of Engineering Sciences.



Jonas Mårtensson (Member, IEEE) received the M.Sc. degree in vehicle engineering and the Ph.D. degree in automatic control from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2002 and 2007, respectively. He was appointed as a Docent in 2016. He holds a tenure-track position as an Assistant Professor with the Department of Automatic Control, KTH Royal Institute of Technology. He is also engaged as the Program Leader of the Integrated Transport Research Laboratory and the Thematic Leader for the Area Transport in the Information Age within the KTH Transport Platform. His research interests are cooperative and autonomous transport systems, in particular related to heavy-duty vehicle platooning. He is involved in several collaboration projects with Scania CV AB, Södertälje, Sweden, dealing with collaborative adaptive cruise control, look-ahead platooning, route optimization and coordination for platooning, path planning and predictive control of autonomous heavy vehicles, and related topics.