A Decentralized Approach to Autonomous Train Platooning

Anonymous submission

Abstract

This paper presents a decentralized approach for train platooning, combining reinforcement learning and autonomous decision-making. A Double Deep Q-Network (DDQN) model is employed to ensure safe following within a platoon, while novel algorithms are used to manage real-time decisions for virtual coupling and uncoupling. Our method aims to improve network efficiency by maximizing track usage and facilitating coordinated, safe platoon formation, leading to increased rail capacity and flexibility. The proposed methods were evaluated in a custom Gym environment, demonstrating scalability and effectiveness in forming platoons.

Introduction

Efficient train scheduling is crucial for reducing delays, maximizing track use, and meeting growing demand. Decentralized scheduling, where trains rely on local information to coordinate movements, marks a shift from centralized systems, reducing reliance on a single authority. By distributing decision-making across the network, it reduces vulnerability to single-point failures, enhances scalability, and provides greater flexibility to adapt to disruptions or unexpected changes. However, ensuring safety in such a system requires robust coordination mechanisms and clear communication protocols between trains.

Decentralized scheduling and coordination are becoming crucial as demand for rail transportation increases, necessitating higher density and greater flexibility, challenges that centralized systems often struggle to address effectively. This shift is particularly vital in urban metro systems, high-speed rail networks, and freight corridors, where even minor delays can disrupt the entire network. By empowering trains to make localized decisions based on real-time information, decentralized systems enhance flexibility, resilience, and scalability. These attributes make them well-suited for modern rail networks, enabling more efficient use of infrastructure.

This paper introduces novel deep-reinforcement learning-based distributed coordination mechanisms and communication protocols with a particular focus on platooning and virtual coupling. These innovative strategies go beyond traditional systems, such as fixed block and moving block, enabling trains to travel closely together in a coordinated group. This approach improves efficiency, reduces energy consumption, and maximizes track usage (Felez and Vaquero-Serrano 2023).

In this paper, we first start with a literature review that discusses traditional train systems and ongoing efforts to advance them, followed by the research and methods section, and concluding with a discussion of future directions. The research itself is divided into two parts: the first focuses on ensuring safe train following within a platoon, while the second examines decision-making for coupling and uncoupling, as well as handling deviation scenarios.

Literature Review

Traditionally, two main systems have been used to coordinate trains on unidirectional tracks to avoid collisions: fixed block and moving block systems. Figure 1 illustrates these two systems.

In the fixed block system, the track is divided into predefined sections, or ”blocks,” with only one train allowed to occupy a block at any given time. Signals often control these blocks, and a train must wait until the block ahead is cleared before it can proceed. The length of each block is fixed, based on safety considerations, such as braking distances. The advantages of this system are its simplicity, ease of implementation and maintenance, and inherent safety, as only one train can occupy a block at a time. However, it is inflexible because the fixed block length limits track capacity and operational efficiency.

In the moving block system, a safe distance between trains is calculated in real-time based on the speed and braking capacity of each train. Unlike the fixed block system, there are no predefined blocks; instead, the ”block” moves with the lead train. The following train continuously calculates the area of the track that the leading train may occupy. In practice, an absolute braking distance is maintained between the trains, ensuring that the following train can safely stop before reaching the last known position of the leading train. This method reduces headway and increases efficiency, while still maintaining safety.

Virtual Coupling advances the concept of moving blocks allowing trains to operate closer together, almost as if they are physically coupled wagons, but without any actual mechanical connection. Instead of absolute braking distance, it uses relative braking and assumes the train in front is also in motion and won’t stop suddenly (Quaglietta et al. 2022).

The idea is to enable trains to form a platoon where they behave like a single train but retain the flexibility to disconnect or adjust independently. The trains in a virtual coupling system move in a synchronized fashion, with the lead train dictating speed and movement. Figure 1 also illustrates virtual coupling.

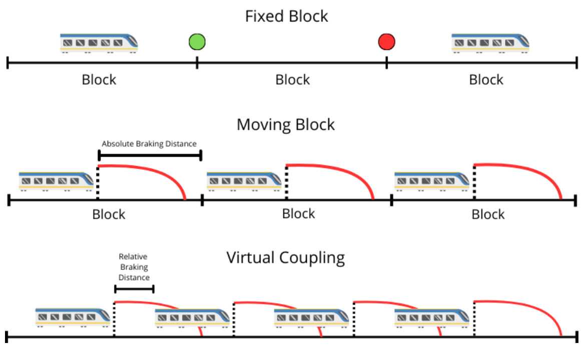

Figure 1: Train coordination systems.

Although real-time, high-precision Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication is needed, this method has some benefits (Goikoetxea 2016). Firstly, trains can couple on the go without interrupting service. This allows flexible routing and better handling of diverse passenger or freight demands. Secondly, trains built by different manufacturers, or designed for different purposes, could interoperate on the same platoon without infrastructure modification since the link is virtual. Thirdly, virtual coupling would allow trains to operate at much shorter intervals, maximizing track usage. It also helps avoid the need for expensive track expansions. Furthermore, platooning enables fuel efficiency and reduces carbon emissions (Tsugawa, Jeschke, and Shladover 2016). Finally, the autonomous capabilities of the trains would potentially reduce the need for physical signaling infrastructure and make the system more adaptive and resilient.

Virtual coupling also has some drawbacks. For trains to operate in close proximity, ultra-reliable, low-latency communication systems are critical to prevent accidents or miscoordination (Flammini et al. 2018). Ensuring the safety of passengers and cargo in virtually coupled trains will require new standards for emergency braking, fault tolerance, and fail-safe protocols. Additionally, While the concept allows trains from different manufacturers to operate together, ensuring they can seamlessly communicate and react to one another will require standardization across the industry.

Extensive research has been conducted in vehicle platooning and virtual coupling, particularly for cars and trucks, focusing on improving traffic flow, fuel efficiency, and safety (Hou et al. 2023). Early concepts for train platooning appeared in (Bock and Bikker 2000), with dynamic coupling and uncoupling proposed in (Konig and Schnieder 2001). Multi-agent systems, where autonomous modular trains operate alongside conventional trains, were introduced in (König, Braun, and Schnieder 2003). Limited advancements followed until the European Shift2Rail initiative (European Research Initiative 2024) revived interest by supporting research into modernized communication technologies.

Communication systems play a crucial role in train coordination, particularly in ensuring safety in modern rail networks. Recent advancements in train-to-train (T2T) and train-to-infrastructure (T2I) communication have significantly improved the ability of rail systems to handle high-density traffic. Reliable technologies for Moving Block Signaling architecture have widely been adopted in systems like the European Rail Traffic Management System (ERTMS) Level 3 (Di Meo et al. 2020) and the China Train Control System Level 3 (Dong et al. 2010). Each train's occupancy zone is continuously updated via real-time bidirectional Train-to-Ground (T2G) communication between the train and wayside infrastructure, such as radio block centers and GSM-R exchange centers. However, virtual coupling requires more advanced technology for wireless train-to-train (T2T) communication. MOVINGRAIL Deliverable D3.3 suggested that 5G cellular networks are well-suited for long-distance inter-platoon communications, while IEEE 802.11 networks can be used for short-range intra-platoon communications, ensuring real-time data exchange for speed and braking coordination (Ge et al. 2024).

Virtual coupling has been analyzed for applications like high-speed rail (Schumann 2016) and metro transit (Luo et al. 2021), leading to the development of various control architectures, primarily centralized and decentralized. Centralized systems optimize train operations globally, coordinating movements across the network. In contrast, decentralized systems enable each train to make decisions autonomously, relying on local environmental knowledge. Platoon formation in decentralized systems is facilitated through communication topologies, such as unidirectional train communication (Felez, Kim, and Borrelli 2019) or leader-issued commands (Zheng et al. 2017).

One of the major challenges in platooning is making effective decisions about coupling and uncoupling. While forming platoons at stop-in stations is an option, it can be time-intensive, prompting proposals for dynamic coupling at operational speeds (Nold and Corman 2021). Significant research has focused on platoon merging for cars on multi-lane highways (Rios-Torres and Malikopoulos 2017), including strategies for integrating a single vehicle into a cooperative platoon (Scholte, Zegelaar, and Nijmeijer 2022). However, relatively little attention has been given to trains. One notable study framed the scheduling of freight trains and their assignment to rail platoons as an optimization problem, specifically addressing corridors equipped with rapid-shunting facilities (Schwerdfeger, Otto, and Boysen 2021).

In recent years, reinforcement learning (RL) has facilitated more autonomous decision-making in platooning, particularly in optimizing speed control coordination. For example, a distributed deep RL model using bidirectional communication was developed to achieve local training while maintaining global consensus (Liu, Ding, and Lv 2020). Additionally, multi-agent reinforcement learning for platooning in highway on-ramp scenarios has been studied, where a centralized training and decentralized execution solution was identified (Chen et al. 2023).

Problem Setup and Proposed Approach

The increasing demand for rail transportation systems highlights the need for more trains to share tracks simultaneously. This demand necessitates train coordination strategies that minimize communication overhead to ensure scalability while maintaining safety and adaptability. Despite the advancements in the field, a gap remains in scalable, fully decentralized approaches, particularly in their application to the railway industry. This paper aims to address this gap by proposing a new decentralized train platooning system where each train operates independently but collaborates locally with nearby trains to form platoons. By relying on unidirectional communication—where each train shares data only with the train directly behind—and bidirectional communication—where each platoon shares data with the platoon in front and behind—the system achieves scalability and efficiency. A custom reinforcement learning framework was implemented to train agents to maintain safe braking distances and dynamically couple or decouple based on shared routes and destinations. This modular approach simplifies coordination and ensures trains can safely adapt to real-world complexities, such as route deviations and varying speeds. The research aims to optimize this coordination system to enhance rail efficiency and maximize track usage.

Our first focus is on situations where trains are coupled. We created a custom OpenAI Gym environment to train a single agent to follow a leader train at a safe braking distance. The idea is that if one agent can learn the optimal policy to maintain a safe distance from the train ahead, then the model can be scaled to multiple trains. Given a unidirectional track and restricted forward movement, this approach is sufficient to prevent collisions within the platoon.

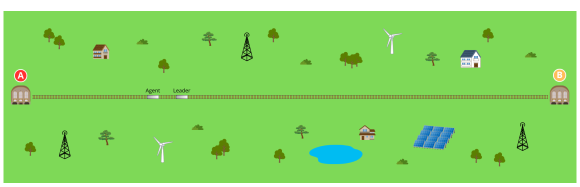

Figure 2: Custom Gym environment.

The agent learns the optimal behavior by interacting with the environment and observing how it responds. It collects feedback in the form of rewards and tries to maximize them. Over time, it learns the optimal action to take for a particular state. Reinforcement learning is suitable in real-world situations where feedback isn't immediately available for every step, such as in this case where communication can be inconsistent or have delays.

The training process adopts a Sim-to-Real approach, where the policy is first learned in a simulated environment and then deployed in real trains. Simulation allows for the creation of a diverse set of scenarios, including edge cases and rare events, enabling thorough policy training under controlled and reproducible conditions. Furthermore, it avoids exposing real-world systems to potentially unsafe or impractical scenarios.

On each episode of the training process, the leader and the following agent start at the same random speed and at a random distance from each other within the communication range. The leader changes speed arbitrarily by repeatedly setting random target speeds and accelerating or decelerating uniformly until achieving them. The episode ends when the leader arrives at Station B, if both trains collide, or if they exceed the reliable communication distance. While the agent is within the communication range, it receives data on the train in front, particularly the speed and position. Then, the state representation becomes the following $\{$distance to the leader, agent's speed, leader's speed$\}$. For each possible state, the agent can take any of the following 3 actions $\{$slow down, maintain speed, speed up$\}$. A speed limit is set to prevent excessive acceleration and collisions.

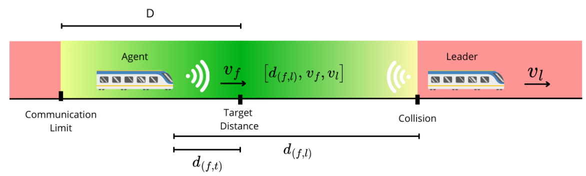

Figure 3: State Representation.

Since the goal is to maintain the same speed for both agents and assuming both have equal and constant deceleration capabilities, the relative braking distance effectively becomes the safety distance denoted as γ. Let s represent the position, v velocity, and a deceleration, with subscripts f for the follower train and l for the leader train. Using the equation $v = v_0 + at$, we set $0 = v_0 + at$ to find the stopping time $t = \frac{v_0}{a}$. Substituting into the equation $x = x_0 + v_0 t - \frac{1}{2}at^2$ for stopping distance gives $x = \frac{v_0^2}{2a}$. Thus, the agent should be positioned at:

$$s_f = s_l + \frac{v_l^2}{2a_l} - \frac{v_f^2}{2a_f} - \gamma \tag{1}$$

Consequently, the relative braking distance is:

$$d = |s_f - s_l| = \left| \frac{v_l^2}{2a_l} - \frac{v_f^2}{2a_f} \right| + \gamma \tag{2}$$

When $a_l = a_f$ and $v_l = v_f$, this simplifies to:

$$d = \gamma \tag{3}$$

The rewards encourage the agent to minimize the distance to the target coupling distance which is the relative braking distance. The agent is given a positive reward between 0 and 1 on each timestamp of an episode. The closer it is to the target distance, the reward is larger as follows in $1 - \frac{d_(f,t)}{D}$. D is the largest possible gap before leaving the communication region. A larger positive reward is also given at the end of the episode to encourage the agent not to collide or get behind ending the episode early.

Algorithm 1: Reward Function

```
d ← |distance_to_leader − target_distance|
reward ← 1 − d/D
if |leader_pos − end_pos| ≤ ϵ then
    if lower_optimal_distance < distance_to_leader < upper_optimal_distance then
        reward ← reward + 100
    else
        reward ← reward + 50
    end if
end if
```

The continuous state space in this problem makes Q-Tables (Watkins and Dayan 1992) impractical, so we implemented a Double Deep Q-Network (DDQN) to approximate Q-values using a neural network. This method also stabilizes learning by reducing the overestimation bias found in traditional DQNs, making it more effective in complex, dynamic train platooning. We trained the DDQN for 5000 episodes with a 0.001 learning rate and an epsilon decay of 0.999, promoting exploration initially and exploitation over time. As shown in 4, total rewards per episode generally increased, with the 200-episode moving average stabilizing around 550. Variability in episode rewards arose due to randomized leader speeds, varied initializations, and occasional extreme starting positions, which sometimes ended episodes prematurely.

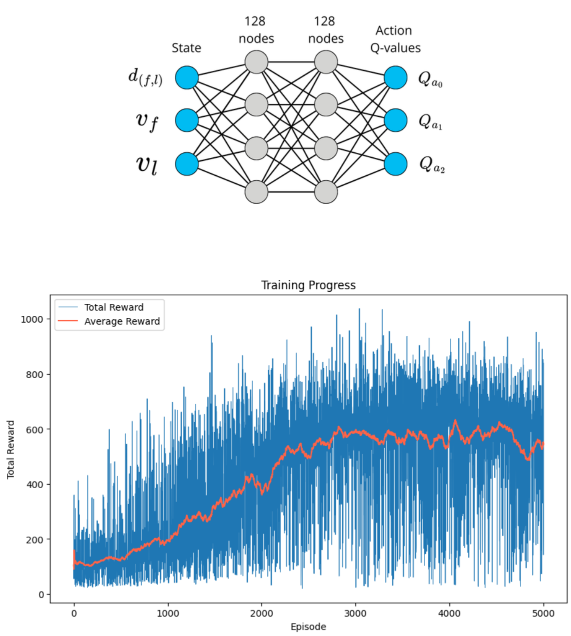

Figure 4: DDQN architecture and total reward plot.

Now that a single agent can follow a leader safely, this approach scales to a multi-agent platoon. In this setup, each train receives data only from the train directly in front and transmits its data to the next train behind it, as shown in Figure 5. The first agent acts as the platoon leader, indirectly controlling the other trains. Communication remains minimal, as each train independently follows the train ahead, requiring no central coordination across the entire platoon.

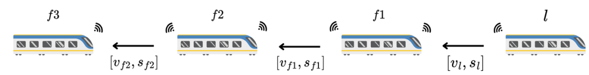

Figure 5: Scaled multi-agent scheme.

During testing of the multi-agent version with five agents, including the platoon leader, we recorded each agent's speed and its distance to the corresponding leader. As shown in Figure 6, the left graph illustrates how all agents adjusted their speeds to align with the train ahead, despite the leader's random speed changes. However, some irregularities can be observed throughout the episode, primarily because the agents were initialized at random distances within the communication range. On the right plot, it is evident that over time, the agents effectively maintained the relative braking distance, indicated by the red horizontal line.

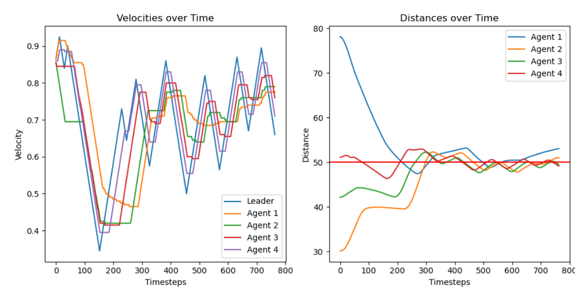

Figure 6: Multi-Agent behavior.

Coupling and Uncoupling

With agents now capable of platooning, the next challenge is to determine the optimal timing and approach for coupling and uncoupling. To explore this, we developed a new environment where a number of agents are initialized on a track, each positioned at a safe distance from the others. Farther down the track, there is a fork where trains can proceed to either Station B or Station C.

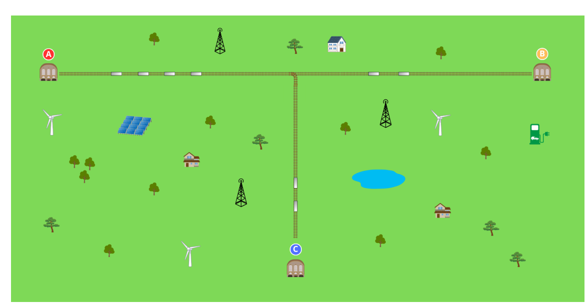

Figure 7: Coupling and uncoupling environment.

Each agent is randomly assigned a destination, and the first agent in the sequence is designated as the leader. The leader sets a target speed—the speed limit—and accelerates until it reaches this goal, while the following agents maintain a safe braking distance, tracking the leader's speed using the model outlined above.

As the platoon approaches a fork in the track, the leader chooses the direction based on its assigned destination. Agents destined for the other station must then separate from the platoon. Communication with the track's switching infrastructure enables a seamless transition. When a platoon member reaches the fork, it checks whether the next train will follow the same path. If so, no change is made. However, if the next agent won't follow the same path, the platoon splits, with the following agent becoming the leader of the new rear platoon. The new leader can adjust speed and slow down if necessary, ensuring a smooth transition through the fork. This approach is flexible and adaptable to various situations, as demonstrated in Figure 8.

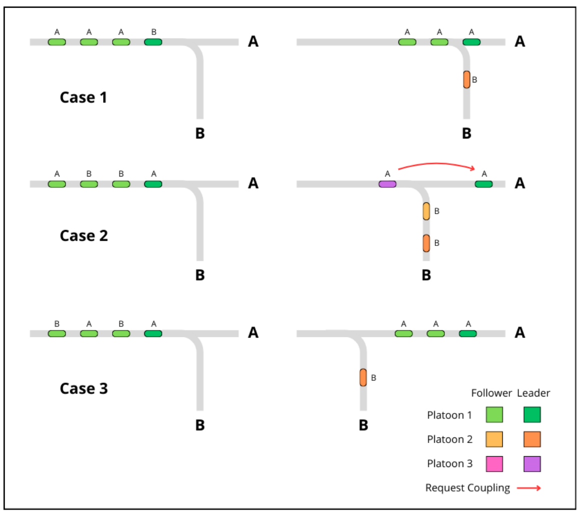

Figure 8: Deviation cases.

Periodically, each leader assesses the position of the platoon ahead and sends a request to join. The leader of the front platoon, based on certain criteria, can decide to wait for the trailing platoon. In a real-world scenario, factors like energy efficiency, the length of the shared route, speed capabilities, and estimated arrival times could influence this decision. For simplicity, in this simulation, leaders allow a joining maneuver if the trailing platoon is within a specified distance and is sharing the same route. In these situations, the front platoon slows its speed until the leader of the rear platoon is within a reliable communication range, allowing coupling to begin. At that point, the rear leader adopts the safe braking distance model trained previously and no longer acts as a leader. Once coupled, the leader is notified and it can speed up to its original target speed. The episode ends when all the platoons arrive at their destination.

Figure 9 shows the velocities of agents during an episode. Initially, all agents are part of a single platoon, with Agents 1, 2, and 4 heading toward Station B, while Agents 3, 5, and 6 are heading to Station C. When Agent 2 reaches the intersection, Agent 3 becomes the leader of a new platoon, as it will not follow the same direction as Agents 1 and 2. When Agent 3 reaches the fork Agent 4 will take a different path, becoming the leader of another new platoon. At this point, Agent 3 forms a single-agent platoon, and three platoons exist in total. However, Agent 4 requests to merge with the platoon of Agents 1 and 2. Upon approval, Agent 1, as the leader, slows down to allow Agent 4 to join. When Agent 4 reaches the fork, Agents 5 and 6, bound for Station C, split and form a new platoon with Agent 5 as the leader. They then request to merge with Agent 3's platoon. As the leader, Agent 3 allows the merge by reducing speed. After both mergers, the agents maintain a safe braking distance and continue at their target speeds as newly formed platoons.

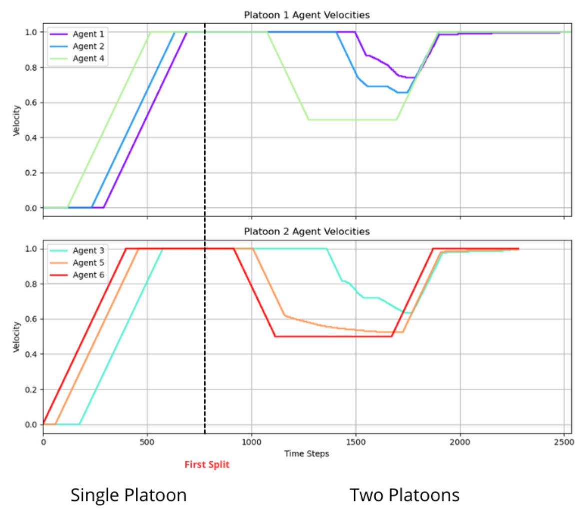

Figure 9: Coupling and uncoupling velocities.

Conclusion and Future Work

This research presents a decentralized train platooning approach using an off-policy RL model that learns from data, enabling each train to autonomously maintain safe distances and dynamically join or leave platoons. The results demonstrate the potential to enhance rail efficiency and increase capacity through autonomous, close-following operations.

Future work will focus on optimizing the decision-making processes for platoon formation, particularly regarding whether the platoon leader should wait for following trains or continue without them. Research will explore strategies to improve coordination when trains are waiting to join a platoon, ensuring smoother merging and reducing delays. Enhancing safety protocols will also be crucial for reliable operations under various conditions. Additionally, SIM-to-REAL techniques could be explored to train the policy in simulation and apply it to real trains with the necessary computing and networking hardware. Lastly, refining the reinforcement learning model will be key to improving scalability and performance at higher speeds.

References

Bock, U.; and Bikker, G. 2000. Design and Development of a Future Freight Train Concept – "Virtually Coupled Train Formations". *IFAC Proceedings Volumes*, 33(9): 395–400. 9th IFAC Symposium on Control in Transportation Systems 2000, Braunschweig, Germany, 13-15 June 2000.

Chen, J.; Zhou, Z.; Duan, Y.; and Yu, B. 2023. Research on Reinforcement-Learning-Based Truck Platooning Control Strategies in Highway On-Ramp Regions. *World Electric Vehicle Journal*, 14(10).

Di Meo, C.; Di Vaio, M.; Flammini, F.; Nardone, R.; Santini, S.; and Vittorini, V. 2020. ERTMS/ETCS Virtual Coupling: Proof of Concept and Numerical Analysis. *IEEE Transactions on Intelligent Transportation Systems*, 21(6): 2545–2556.

Dong, H.; Ning, B.; Cai, B.; and Hou, Z. 2010. Automatic Train Control System Development and Simulation for High-Speed Railways. *IEEE Circuits and Systems Magazine*, 10(2): 6–18.

European Research Initiative. 2024. Shift2Rail. Accessed: 2024-11-12.

Felez, J.; Kim, Y.; and Borrelli, F. 2019. A Model Predictive Control Approach for Virtual Coupling in Railways. *IEEE Transactions on Intelligent Transportation Systems*, 20(7): 2728–2739.

Felez, J.; and Vaquero-Serrano, M. A. 2023. Virtual Coupling in Railways: A Comprehensive Review. *Machines*, 11(5).

Flammini, F.; Marrone, S.; Nardone, R.; Petrillo, A.; Santini, S.; and Vittorini, V. 2018. Towards Railway Virtual Coupling. In *2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)*, 1–6.

Ge, X.; Wu, Q.; Han, Q.-L.; and Zhang, X.-M. 2024. Resilient Virtual Coupling Control of Automatic Train Convoys With Intermittent Communications. *IEEE Transactions on Vehicular Technology*, 73(5): 6183–6195.

Goikoetxea, J. 2016. Roadmap Towards the Wireless Virtual Coupling of Trains. In Mendizabal, J.; Berbineau, M.; Vinel, A.; Pfletschinger, S.; Bonneville, H.; Pirovano, A.; Plass, S.; Scopigno, R.; and Aniss, H., eds., *Communication Technologies for Vehicles*, 3–9. Cham: Springer International Publishing. ISBN 978-3-319-38921-9.

Hou, J.; Chen, G.; Huang, J.; Qiao, Y.; Xiong, L.; Wen, F.; Knoll, A.; and Jiang, C. 2023. Large-Scale Vehicle Platooning: Advances and Challenges in Scheduling and Planning Techniques. *Engineering*, 28: 26–48.

Konig, S.; and Schnieder, E. 2001. Modeling and simulation of an operation concept for future rail traffic. In *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No.01TH8585)*, 808–812.

König, S.; Braun, I.; and Schnieder, E. 2003. Decentralized Management and Operations Control Concept for Railway Freight Transport Services by Multi Agent Systems. *IFAC Proceedings Volumes*, 36(14): 317–322. 10th IFAC Symposium on Control in Transportation Systems 2003, Tokyo, Japan, 4-6 August 2003.

Liu, B.; Ding, Z.; and Lv, C. 2020. Platoon control of connected autonomous vehicles: A distributed reinforcement learning method by consensus. *IFAC-PapersOnLine*, 53(2): 15241–15246. 21st IFAC World Congress.

Luo, X.; Tang, T.; Liu, H.; Zhang, L.; and Li, K. 2021. An Adaptive Model Predictive Control System for Virtual Coupling in Metros. *Actuators*, 10(8).

Nold, M.; and Corman, F. 2021. Dynamic train unit coupling and decoupling at cruising speed: Systematic classification, operational potentials, and research agenda. *Journal of Rail Transport Planning & Management*, 18: 100241.

Quaglietta, E.; Spartalis, P.; Wang, M.; Goverde, R. M.; and van Koningsbruggen, P. 2022. Modelling and analysis of Virtual Coupling with dynamic safety margin considering risk factors in railway operations. *Journal of Rail Transport Planning & Management*, 22: 100313.

Rios-Torres, J.; and Malikopoulos, A. A. 2017. A Survey on the Coordination of Connected and Automated Vehicles at Intersections and Merging at Highway On-Ramps. *IEEE Transactions on Intelligent Transportation Systems*, 18(5): 1066–1077.

Scholte, W.; Zegelaar, P.; and Nijmeijer, H. 2022. A control strategy for merging a single vehicle into a platoon at highway on-ramps. *Transportation Research Part C: Emerging Technologies*, 136: 103511.

Schumann, T. 2016. Increase Of Capacity On The Shinkansen High-speed Line Using Virtual Coupling. In *COMPRAIL 2016*.

Schwerdfeger, S.; Otto, A.; and Boysen, N. 2021. Rail platooning: Scheduling trains along a rail corridor with rapid-shunting facilities. *European Journal of Operational Research*, 294(2): 760–778.

Tsugawa, S.; Jeschke, S.; and Shladover, S. E. 2016. A Review of Truck Platooning Projects for Energy Savings. *IEEE Transactions on Intelligent Vehicles*, 1(1): 68–77.

Watkins, C. J. C. H.; and Dayan, P. 1992. Q-learning. *Machine Learning*, 8(3): 279–292.

Zheng, Y.; Li, S. E.; Li, K.; Borrelli, F.; and Hedrick, J. K. 2017. Distributed Model Predictive Control for Heterogeneous Vehicle Platoons Under Unidirectional Topologies. *IEEE Transactions on Control Systems Technology*, 25(3): 899–910.