PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
الجمهـــوريــــــة الجـزائــريـــــة الديـمـقـراطيـــــة الشـعـبيـــــة
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
وزارة التعليـم العالـــي والـبحـث العلمــي
HIGHER SCHOOL OF COMPUTER SCIENCE AND DIGITAL TECHNOLOGIES-BEJAIA
المـــدرسة العـــليا في علـــوم وتكنــولوجيــات الاعـــلام الالـــي بجـــاية

Dissertation Submitted to the Department Of Computer Science in Partial Fulfillment of the Requirements for Master's Degree in Computer Science

Specialty: Artificial Intelligence and Data Sciences

Submitted By:**Amani Chaimaa Sellam**

Deep Reinforcement Learning for Vehicle Platooning Optimization

Supervised by:

Dr. Nadir Farhi GRETTIA Lab
Mme. Meriem Bouali ESTIN

Members of the Jury:

▪ **Pr .** Hamid KHERBACHI	President	ESTIN
▪ **Dr.** Sid Ali BOUSLA	Examiner	ESTIN
▪ **Dr.** Soraya TOULOUM	Examiner	ESTIN
▪ **Mme.** Ouahiba BOUGLIMINA	Examiner	ESTIN

Academic year: 2023/2024

Dedication

First and foremost, I would like to express my deepest gratitude to my parents, who have been my pillars of support throughout this journey. To my father, who has always been my mirror, reflecting the best version of myself and guiding me with wisdom and patience.

To my mother, who has been my greatest cheerleader, encouraging me at every step and never letting me doubt my abilities.

To my sister Tahani, who never misses a chance to drive me crazy, but in doing so, has always kept me grounded and brought joy to my life.

To my brother Mustafa and my sister Maryam, who might not realize this, but they were my strength and my source of force, always motivating me to push forward.

To my grandma, who prayed for me as long as she was breathing. Her prayers and blessings have been a guiding light in my life, and I will forever be grateful for her love and support.

To my dear friend Cheima, who has always been there to lift me up whenever I felt down. Your unwavering support and belief in me have been instrumental in my achievements, and I am forever thankful for your friendship.

Lastly, I would like to thank all my friends who have stood by me through thick and thin. Your companionship and encouragement have made this journey more enjoyable and less daunting.

This work is as much yours as it is mine. Thank you all.

Abstract

Automated vehicle platooning has emerged as a significant method for improving traffic efficiency, reducing fuel consumption, and enhancing road safety. This thesis investigates the optimization of vehicle platooning using reinforcement learning techniques, specifically focusing on Deep Q-Networks (DQN) integrated with dueling networks and prioritized experience replay (PER). A two-layered approach is employed, where the first layer identifies joinable platoons and evaluates the benefits of joining them, and the second layer uses reinforcement learning to optimize merging, lane-changing, and acceleration strategies. The results of the simulation demonstrate that the proposed method significantly reduces travel time, fuel consumption, and improves the overall safety of the platooning process. Future work includes extending the model to mixed traffic environments with both autonomous and human-driven vehicles, as well as integrating predictive traffic models.

Keywords: Automated vehicle platooning, Deep Q-Networks, Dueling networks, Prioritized experience replay, Reinforcement learning, Traffic efficiency, Fuel consumption, Road safety, Merging strategies, Mixed traffic environments.

Résumé

Le peloton de véhicules automatisés est devenu une méthode significative pour améliorer l'efficacité du trafic, réduire la consommation de carburant et renforcer la sécurité routière. Cette thèse examine l'optimisation du peloton de véhicules en utilisant des techniques d'apprentissage par renforcement, en se concentrant spécifiquement sur les réseaux de neurones profonds (DQN) intégrés avec des réseaux de duel et la reprise d'expérience priorisée (PER). Une approche à deux niveaux est utilisée : le premier niveau identifie les pelotons accessibles et évalue les avantages de les rejoindre, tandis que le second niveau utilise l'apprentissage par renforcement pour optimiser les stratégies de fusion, de changement de voie et d'accélération. Les résultats de la simulation montrent que la méthode proposée réduit de manière significative le temps de trajet, la consommation de carburant et améliore la sécurité générale du processus de peloton. Les travaux futurs incluent l'extension du modèle à des environnements de trafic mixte avec des véhicules autonomes et conduits par des humains, ainsi que l'intégration de modèles prédictifs de trafic.

Mots-clés: Peloton de véhicules automatisés, Réseaux de neurones profonds, Réseaux de duel, Reprise d'expérience priorisée, Apprentissage par renforcement, Efficacité du trafic, Consommation de carburant, Sécurité routière, Stratégies de fusion, Environnements de trafic mixte.

الملخص

ظهرت قيادة القوافل الآلية كطريقة مهمة لتحسين كفاءة المرور وتقليل استهلاك الوقود وتعزيز سلامة الطرق. تبحث هذه الأطروحة في تحسين قيادة القوافل الآلية باستخدام تقنيات التعلم المعزز، مع التركيز بشكل خاص على الشبكات العصبية العميقة المدمجة مع الشبكات المزدوجة وإعادة تشغيل التجارب ذات (DQN) الأولوية (PER). يتم استخدام نهج من طبقتين، حيث تتعرف الطبقة الأولى على القوافل القابلة للانضمام وتقيّم فوائد الانضمام إليها، بينما تستخدم الطبقة الثانية التعلم المعزز لتحسين استراتيجيات الاندماج وتغيير المسار والتسارع. أظهرت نتائج المحاكاة أن الطريقة المقترحة تقلل بشكل كبير من وقت السفر واستهلاك الوقود وتُحسن من أمان عملية قيادة القوافل. تشمل الأعمال المستقبلية توسيع النموذج ليشمل بيئات مرور مختلطة مع وجود كل من المركبات ذاتية القيادة والمركبات التي يقودها البشر، بالإضافة إلى دمج نماذج تنبؤية لحركة المرور.

الكلمات المفتاحية: قيادة القوافل الآلية، الشبكات العصبية العميقة، الشبكات المزدوجة، إعادة تشغيل التجارب ذات الأولوية، التعلم المعزز، كفاءة المرور، استهلاك الوقود، سلامة الطرق، استراتيجيات الاندماج، بيئات المرور المختلطة.

Acronyms

AV Autonomous Vehicle

CAV Connected and Automated Vehicle

DQN Deep Q-Network

LP Linear Programming

MILP Mixed-Integer Linear Programming

MPC Model Predictive Control

PID Proportional-Integral-Derivative

QP Quadratic Programming

RL Reinforcement Learning

V2I Vehicle-to-Infrastructure

V2V Vehicle-to-Vehicle

List of Figures

List of Tables

Contents

Introduction

Automated platooning, a cooperative driving pattern where a group of vehicles moves at a consensual speed while maintaining a small and nearly constant distance between adjacent vehicles, has gained increasing attention due to its potential to improve transportation systems. We explore the optimization of platooning on highways through the application of reinforcement learning techniques.

Automated vehicle platooning control and optimization have great potential to revolutionize optimal control through the application of reinforcement learning techniques. Automated vehicles can learn to negotiate crossings, merge onto highways, navigate dynamic settings, and optimize their platooning behavior to alleviate congestion and enhance traffic efficiency by utilizing deep reinforcement learning algorithms (Peng et al., 2021). Additionally, automated vehicles may coordinate their actions and communicate with one another through the use of reinforcement learning. This allows the vehicles to adjust their speed and make real-time decisions, preventing traffic jams at intersections and ensuring efficient and smooth traffic flow. Vehicles can achieve long-term optimization and short-term constraint satisfaction by integrating robust model-free reinforcement learning with model-based techniques, particularly in modeling techniques for self-driving cars.

Purpose of the Study

This study attempts to maximize the advantages offered by intelligent networked systems, address the increasing demand for transportation by ensuring traffic safety, enhancing road efficiency, and anticipating future changes in the traffic environment through the development of automated driving technology.

The Connected and Automated Vehicle (CAV) system is a novel approach to increasing vehicle speed, decreasing traffic congestion, and lowering the frequency of traffic accidents. It operates by combining computer network communication technology with the current highway traffic system to enable vehicle-to-vehicle and vehicle-road communication. Further, it enables automatic control of inter-vehicle distances and vehicle transverse and longitudinal dynamics.

Firstly, the advancement of automated driving technology removes human error as the main contributor to accidents. Additionally, connected and automated vehicles (CAVs) can anticipate and obtain the driving characteristics of nearby vehicles in real-time, improving their ability to maintain a safe following distance.

Moreover, the intelligent network system allows vehicles to receive real-time traffic information via telematics, optimize route and speed choices, reduce driving time, avoid traffic, and make effective use of the available road resources. The creation of intelligent linked systems has been aided recently by the advancement of 5G communication technologies. However, the practice of combining Internet-connected vehicles with regular cars will continue due to technological constraints, legal restrictions, and ethical considerations.

According to recent studies, the number of self-driving cars on the road is expected to remain below 30% by 2040 [1]. Furthermore, due to human-driven vehicles (HDVs), there is an increase in environmental randomness, vehicle-to-vehicle interference, and the complexity of the decision-making process for self-driving vehicles (SDVs). As a result, researchers are placing greater focus on how to ensure vehicle safety when operating in mixed traffic flows and how this can positively affect traffic flow and road infrastructure.

In situations with mixed traffic flow, platoon driving is a practical way to reduce energy consumption, prevent congestion, and improve traffic efficiency and safety. With the help of an automated driving system and a network of linked vehicles, platoon driving seeks to maximize road capacity while also taking into account the principles of road safety. A leader car and several followers are formed by the leader's driving speed and the distance between the vehicles in the small group of collaborators. The PATH (California Partners for Advanced Transit and Highways) project [2] is an example of a representative project. In addition to implementing the first multilayer system research on fleet cooperative driving technology, this project uses real-vehicle road tests to calibrate a follow-up model that imitates the behavior of intelligent network vehicles.

Technical Overview and Concepts

1.1 Introduction

Platooning, an important innovation in automated vehicle technology, is increasingly seen as a key strategy to improve roadway efficiency [1], minimize environmental impact, and improve traffic safety. This technology allows vehicles to travel in closely-knit groups or "platoons" at highway speeds, which enables several operational efficiencies and safety improvements. The essence of platooning lies in its ability to allow vehicles to communicate and operate in unison, reducing the gaps between them. This synchronization is facilitated by advanced control systems that manage the acceleration and braking of all vehicles in the platoon, effectively minimizing human error and variability in driver reactions.
Moreover, platooning leverages cutting-edge communication technologies that ensure continuous data exchange between vehicles (V2V – vehicle-to-vehicle) and between vehicles and traffic management systems (V2I – vehicle-to-infrastructure).

These technologies are pivotal for the operational integrity and safety of platoons, allowing vehicles to respond collectively and instantaneously to changing road conditions, traffic flows, and potential hazards. The integration of these advanced systems not only helps streamline traffic flow, but also significantly reduces fuel consumption due to reduced aerodynamic drag, thereby lowering emissions. Furthermore, improved traffic management resulting from effective platooning can lead to reduced road congestion, which in turn decreases the overall travel time and increases the predictability of shipping schedules, presenting clear economic benefits.

1.2 Background on Autonomous Vehicles

Autonomous vehicles (AVs) are at the forefront of modern transportation technology, representing a significant shift in how vehicles are operated and integrated into broader traffic systems. The development of AVs is driven by the potential to drastically improve road safety, enhance mobility, and reduce the environmental impact of transportation.

1.2.1 Levels of Autonomy

The progression towards full automation in vehicles is categorized by the Society of Automotive Engineers (SAE) into six distinct levels, from Level 0 (no automation) to Level 5 (full automation) [6]. At Level 0, the vehicle relies entirely on the human driver for control, while Level 5 represents a fully autonomous vehicle capable of handling all driving tasks under any conditions, without the need for human intervention. Intermediate levels involve increasing degrees of automation, where the vehicle can manage certain driving tasks but still requires human oversight, particularly in complex scenarios.

1.2.2 Technologies Enabling Autonomous Driving

AVs rely on a sophisticated array of sensors and communication technologies to perceive their surroundings and make real-time driving decisions. Key technologies include radar, lidar, cameras, and ultrasonic sensors, which provide a comprehensive understanding of the vehicle's environment. These sensors are complemented by advanced software systems that utilize machine learning and artificial intelligence to interpret sensor data, recognize objects, and predict the movements of other road users [6]. Additionally, connectivity features such as Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication enhance the vehicle's ability to respond to real-time traffic conditions and collaborate with other vehicles.

1.2.3 Challenges in Autonomous Driving

Despite the technological advancements, several challenges remain in the widespread adoption of fully autonomous vehicles. These challenges include ensuring the reliability and safety of AVs in complex and dynamic environments, such as urban areas with unpredictable pedestrian behavior or adverse weather conditions that can impair sensor performance. Another significant challenge lies in the ethical and legal implications of autonomous driving, particularly in scenarios where the vehicle must make decisions that could impact human lives [7]. Additionally, there are ongoing concerns about cybersecurity, as the connectivity features that enable AVs to communicate with their environment also make them vulnerable to hacking and data breaches.

1.2.4 The Role of Artificial Intelligence in AVs

Artificial Intelligence (AI) is a cornerstone of autonomous driving technology, enabling vehicles to process vast amounts of data and make decisions in real-time. Machine learning algorithms, particularly deep learning, are used to improve the accuracy of object detection, decision-making, and trajectory planning. These AI-driven systems allow AVs to learn from their experiences, adapt to new driving conditions, and optimize their performance over time [7]. AI also plays a crucial role in developing the vehicle's ability to handle unforeseen situations, which is essential for achieving higher levels of autonomy.

1.2.5 Future Prospects for Autonomous Vehicles

The future of autonomous vehicles is poised to bring about significant changes in personal and public transportation. As AV technology continues to evolve, there is potential for fully autonomous vehicles to become a common sight on roads, leading to safer and more efficient transportation systems. Ongoing research is focused on improving the scalability of AV systems, enhancing their ability to operate in diverse environments, and integrating them into existing transportation infrastructures. The development of Vehicle-to-Everything (V2X) communication will further enhance the capabilities of AVs, allowing them to interact more seamlessly with other vehicles, infrastructure, and even pedestrians [6].

1.3 Problem-Related Concepts

1.3.1 Platoon Optimization

Definition

Optimization in platooning involves fine-tuning operational protocols and algorithms to effectively manage platoon formation, maintenance, and dissolution, ensuring maximum efficiency and safety.

1.3.2 Components of Optimization

- **Platoon Formation** - Algorithms are developed to dynamically form platoons based on real-time traffic data and vehicle capabilities.
- **Platoon Maintenance** - Continuous adjustment of speeds and distances within the platoon to adapt to changing traffic conditions and maintain cohesion.
- **Platoon Dissolution** - Coordinated disbanding of platoons as vehicles approach their destinations to minimize disruptions to overall traffic flow.

1.3.3 Platoon Optimization Motivation

- **Traffic Efficiency** - Optimized platooning maximizes road capacity and reduces congestion, enabling smoother flow of traffic.
- **Safety Improvements** - By reducing human error and enhancing vehicle response times, platooning decreases the likelihood of accidents.

- **Environmental and Economic Benefits** - Reduction in fuel consumption and emissions results in environmental benefits, while operational cost savings from decreased fuel use offer economic advantages.

1.3.4 Methodologies for Optimization

Simulation Tools

SUMO (Simulation of Urban Mobility) is used as a primary tool for simulating various platooning scenarios to assess different strategies under controlled and repeatable conditions.

Application of Reinforcement Learning

Reinforcement learning– particularly Q-learning and deep Q-Learning – is employed to autonomously optimize decision-making processes within platoons, enhancing their ability to respond to dynamic traffic environments effectively.

Communication Protocols

Efficient communication protocols, both vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I), ensure real-time data exchange among vehicles and between vehicles and traffic management systems, crucial for the successful implementation of platooning.

1.4 Solution-Related Concepts

1.4.1 Introduction to Optimization Techniques

Optimization is a fundamental aspect of many problems in engineering, economics, and computer science. The goal of optimization is to find the best solution from a set of feasible solutions, often by maximizing or minimizing a specific objective function. Traditional optimization techniques include linear programming, gradient descent, and genetic algorithms, each with its own strengths and limitations.

Linear programming is widely used for problems that can be expressed as a set of linear equations or inequalities. It is powerful for problems where the relationships between variables are linear and the objective function is well-defined. However, it struggles with non-linear, dynamic, or highly complex environments.

Gradient descent is another popular technique, particularly in the field of machine learning. It iteratively adjusts parameters to minimize a cost function. While gradient descent is effective in finding local minima, it can be trapped in these local minima in non-convex problems and may require careful tuning of hyperparameters like the learning rate.

Genetic algorithms imitate the process of natural selection to evolve solutions over time. They are useful for optimization problems where the search space is large and complex, but they can be computationally expensive and slow to converge to an optimal solution.

1.4.2 Deep Learning and Neural Networks

Deep learning is a subfield of machine learning that has gained prominence due to its ability to model complex patterns in data using artificial neural networks. These networks are inspired by the structure and functioning of the human brain, consisting of layers of neurons that process input data to produce meaningful outputs. The depth of these networks, meaning the number of layers, allows them to learn hierarchical representations of data, which is particularly effective for complex tasks such as pattern recognition and data classification.

Structure of Neural Networks

A neural network is composed of an input layer, one or more hidden layers, and an output layer. Each layer contains a set of neurons, where each neuron is connected to every neuron in the adjacent layers. These connections have associated weights, which are adjusted during the training process to minimize the error between the network's predictions and the actual outcomes.

The basic operation of a neuron involves taking the weighted sum of its inputs, passing this sum through an activation function, and producing an output that is fed to the next layer. The choice of activation function (such as ReLU, sigmoid, or tanh) plays a crucial role in determining the network's ability to capture non-linear relationships in the data.

The hidden layers in a neural network enable it to learn intermediate representations of the data, transforming raw input into a form that is more suitable for the task at hand. As the data passes through each layer, the network progressively extracts higher-level features, making it possible to model complex relationships and patterns.

Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a specific type of neural network particularly well-suited for processing grid-like data structures, such as images. Unlike traditional fully connected networks, CNNs use convolutional layers that apply filters to small regions of the input data, enabling the network to detect local features like edges, textures, and shapes.

A CNN typically consists of several convolutional layers followed by pooling layers, which reduce the dimensionality of the feature maps generated by the convolutional layers. This reduction helps in managing computational complexity and prevents overfitting by focusing on the most relevant features. After the convolutional and pooling layers, fully connected layers are used to integrate the extracted features and make final predictions.

The use of shared weights in convolutional layers (i.e., the same filter applied across different parts of the input) allows CNNs to be more efficient and effective in recognizing patterns regardless of their position in the input, making them the go-to architecture for image and video processing.

Recurrent Neural Networks (RNNs)

Recurrent Neural Networks (RNNs) are designed to handle sequential data, where the order of the data points is significant. Unlike feedforward networks, RNNs have connections that loop back on themselves, allowing them to maintain a memory of previous inputs. This capability makes RNNs particularly effective for tasks where context and temporal dynamics are crucial, such as time series prediction and language modeling.

In an RNN, the output from a given neuron at one time step is fed back into the network as input for the next time step, creating a recursive structure. This allows the network to learn dependencies over time, capturing the sequential nature of the data.

However, RNNs can struggle with learning long-range dependencies due to issues like vanishing and exploding gradients. To address these challenges, variants such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) have been developed. These architectures include mechanisms that control the flow of information, making it easier for the network to retain important information over long sequences.

Training Deep Neural Networks

The training of deep neural networks involves optimizing the weights of the connections between neurons to minimize a loss function, which measures the difference between the predicted output and the true output. This process is typically done using backpropagation, where the error is calculated at the output and propagated back through the network to update the weights.

During training, the network is exposed to a large dataset, and the weights are adjusted iteratively to improve performance. Stochastic gradient descent (SGD) and its variants (such as Adam and RMSprop) are commonly used optimization algorithms to adjust the weights.

One of the key challenges in training deep networks is avoiding overfitting, where the network performs well on training data but poorly on unseen data. Regularization techniques such as dropout, batch normalization, and data augmentation are often employed to mitigate overfitting and improve the network's generalization capabilities.

Challenges and Considerations

Deep learning models, particularly those involving deep neural networks, present several challenges. The training process can be computationally intensive, requiring significant processing power and large datasets. Additionally, the models can be difficult to interpret, often described as "black boxes," which raises concerns about their transparency and trustworthiness.

Moreover, the success of deep learning models heavily depends on the availability of large, labeled datasets. In many domains, obtaining such data can be expensive and time-consuming. Researchers are exploring methods to reduce this dependency through techniques like transfer learning and unsupervised learning, which aim to leverage existing models and unlabeled data to improve learning efficiency.

As the field progresses, addressing these challenges will be critical for the continued development and deployment of deep learning models across various applications.

1.4.3 Introduction to Reinforcement Learning

Reinforcement Learning (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment. The agent is not provided with explicit instructions on how to achieve its goal but instead must discover the best strategies through trial and error, receiving feedback in the form of rewards. Over time, the agent's goal is to maximize the cumulative reward, which requires it to balance exploration (trying new actions to discover their effects) and exploitation (choosing the best-known actions to maximize reward). RL is particularly powerful in dynamic environments where the agent's actions can significantly influence future states and rewards.

1.4.4 Key Concepts

The core components involved in RL include the Agent, Environment, State, Action, and Reward. Each plays a crucial role in the learning process:

- **Agent**: The learner or decision-maker that interacts with the environment.

- **Environment**: The external system with which the agent interacts, often modeled as a Markov Decision Process (MDP).
- **State (s)**: A representation of the current situation in the environment. It encapsulates all the information the agent needs to make a decision.
- **Action (a)**: The set of all possible moves the agent can make. The action taken at any given time influences the next state and the reward received.
- **Reward (r)**: The feedback from the environment, given after the agent takes an action. It indicates how good or bad the action was in terms of achieving the agent's goal.

The process of a RL agent interaction is shown in the following figure:

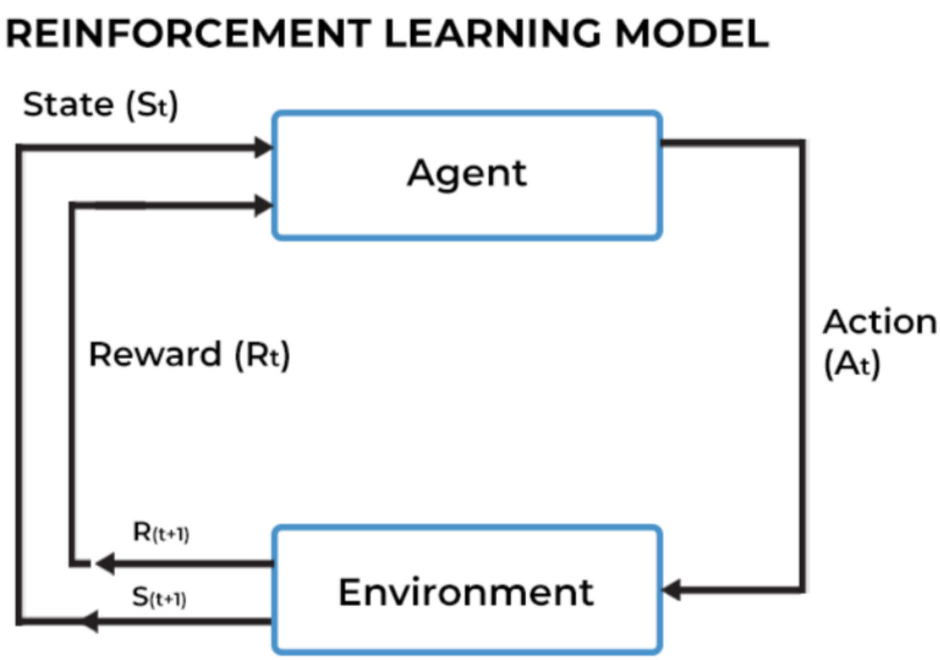


Figure 1.1: How does RL work

1.4.5 Key Enhancements in DQN

DQN introduces several improvements over traditional Q-Learning:

- Experience Replay

- Target Network

1.4.6 Experience Replay

Experience Replay is a fundamental component in Deep Q-Learning that enhances the training process of neural networks by storing the agent's experiences at each timestep, typically in a data structure known as the replay memory buffer. Each experience is stored as a tuple containing the state, action, reward, and subsequent state. During training, batches of experiences are randomly sampled from this buffer. This random sampling helps to break the correlation between consecutive learning samples and ensures that the learning process is stable and efficient. Moreover, using experiences stored over many past episodes allows the network to learn from a more diverse set of states, preventing overfitting to recent experiences and smoothing out learning updates over time.

1.4.7 Target Network

The Target Network in Deep Q-Networks (DQN) is a crucial component designed to stabilize the training process. It involves maintaining two separate neural networks: the primary network and the target network. The primary network continuously updates its weights as it computes the predicted Q-values from the incoming data. In contrast, the target network's weights are updated less frequently, remaining fixed for several training steps. This approach helps to generate more consistent training targets, reducing the risk of rapid fluctuations that can destabilize learning in dynamic environments. By decoupling the target generation process from the learning process, the Q-learning updates become less susceptible to feedback loops that can arise from self-generated data.

1.5 Efficiency of Reinforcement Learning

Reinforcement learning has proven to be highly effective in solving complex optimization problems that are typically challenging for traditional research algorithms. Unlike heuristic or gradient-based optimization methods, RL can operate effectively in environments with high uncertainty and without requiring explicit models or gradient information.

Reinforcement learning algorithms excel by learning optimal strategies directly from raw input data, adapting to changes in the environment, and improving continually through trial and error. This capability allows RL to outperform conventional optimization approaches in tasks that involve sequential decision-making, dynamic environments, and long-term planning, such as robotics, game playing, and autonomous driving. This technique involves storing the agent's experiences and later using a random subset of them to train the network. This helps in reducing correlations in the sequence of observations.

1.5.1 Advantages of Reinforcement Learning in Optimization

Reinforcement Learning (RL) offers a different approach to solving optimization problems, particularly in environments where the optimal solution is not static and must be learned through interaction. Unlike traditional methods that require a fixed model of the environment or a static objective function, RL is model-free and adapts to changes in the environment.

In RL, an agent learns to make decisions that maximize cumulative rewards over time by interacting with its environment. This allows RL to handle optimization in dynamic, uncertain, and complex environments, where the relationships between variables may not be linear or even well-understood.

Adaptive Learning: One of the key advantages of RL is its ability to continuously learn and improve as it interacts with the environment. This adaptive learning process is particularly valuable in real-world scenarios where conditions change over time, making static optimization techniques less effective.

Exploration vs. Exploitation: RL naturally balances exploration (trying new actions to discover better outcomes) and exploitation (choosing the best-known actions to maximize reward). This balance allows RL to avoid local optima that often trap traditional optimization methods.

Scalability: RL algorithms, particularly when combined with deep learning (as in Deep Q-Learning), can scale to handle high-dimensional and complex state spaces that are infeasible for traditional optimization techniques.

Conclusion

This chapter provided a detailed exploration of both the challenges and solutions in the domain of automated vehicular technology, specifically focusing on the concept of platooning and its optimization through advanced machine learning techniques. We began by defining platooning, discussing its significance in enhancing roadway efficiency, reducing emissions, and improving safety. The optimization of platooning was then addressed, highlighting how technologies such as Experience Replay and the Target Network within the Deep Q-Learning framework, as well as the actor-critic architecture, contribute to more stable and effective learning processes. These reinforcement learning strategies are crucial for addressing the dynamic and complex nature of managing platoons, offering robust solutions that outperform traditional optimization algorithms. Overall, the chapter underscores the synergy between sophisticated vehicle coordination challenges and cutting-edge AI techniques, presenting reinforcement learning as a transformative tool for advancing automated transportation systems.

Related Work

2.1 Introduction

The development of connected and automated vehicle (CAV) platooning has garnered significant attention in recent years, driven by the immense potential to not only improve traffic efficiency but also significantly reduce fuel consumption and enhance roadway safety. CAV platooning is fundamentally based on the idea of a group of vehicles traveling in close coordination, where the lead vehicle dictates the pace and trajectory, while the following vehicles mirror these movements. This requires sophisticated control strategies, robust communication systems, and high levels of precision to ensure that the vehicles in the platoon act as a cohesive unit.

The research landscape surrounding platooning is expansive, covering a variety of critical aspects such as vehicle communication, maneuver coordination, and control strategies. Among the most challenging maneuvers are those involving merging and splitting, where vehicles either join or leave the platoon. To manage these maneuvers, two primary control paradigms have been explored: centralized and decentralized systems. Centralized approaches rely on a single controlling entity—usually a lead vehicle or an external server—that makes decisions for the entire platoon. These approaches offer the advantage of optimized decision-making since the central controller has a global view of the platoon's state. However, they are often criticized for their limited scalability and vulnerability to single points of failure. On the other hand, decentralized systems distribute decision-making responsibilities across all vehicles, which enhances the system's robustness and scalability. However, decentralized control can result in suboptimal decisions due to the lack of global state awareness.

Another crucial aspect of vehicle platooning research is trajectory planning. Ensuring that vehicles can merge, split, and execute lane changes safely while optimizing for fuel efficiency and travel time is a complex task. While centralized optimization techniques have historically been favored for trajectory planning due to their ability to globally optimize the platoon's behavior, decentralized methods based on local information are gaining traction for their robustness and scalability.

Machine learning, particularly reinforcement learning (RL), has also recently been explored as a tool for managing platoon operations in dynamic environments. RL enables systems to learn from data, adapt their strategies over time, and make more intelligent decisions without being explicitly programmed for every possible scenario. This shift towards data-driven optimization marks a significant evolution in the field.

This section will review the foundational contributions in CAV platooning, with a focus on centralized and decentralized control protocols, trajectory planning methods, and the application of RL. By exploring these areas in depth, we aim to provide a comprehensive overview of the current solutions and identify the areas where this thesis contributes to advancing the state-of-the-art.

2.2 Overview of Vehicle Platooning Research

Research on Connected and Automated Vehicle (CAV) platooning has evolved significantly over the past decade, primarily due to the potential to drastically improve traffic flow, minimize vehicle emissions, and enhance roadway safety. Platooning involves the coordination of multiple vehicles driving in close proximity, allowing them to take advantage of reduced aerodynamic drag and synchronized acceleration and braking. The benefits are particularly pronounced in terms of fuel efficiency, reduced traffic congestion, and the ability to create safer driving conditions by minimizing human error.

Managing the complex maneuvers that arise during platooning, such as merging into or splitting from the platoon, requires sophisticated control strategies. These strategies are generally divided into two primary categories: centralized and decentralized systems.

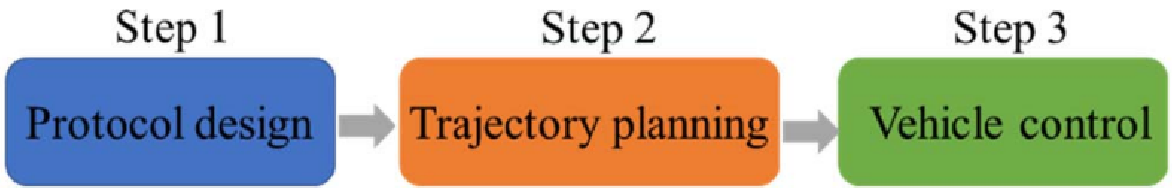


Figure 2.1: The 3-layer framework in literature. This diagram represents the hierarchical structure typically employed in vehicle platooning systems, where the bottom layer handles communication, the middle layer governs vehicle control, and the top layer deals with optimization and decision-making. It highlights how these layers work together to enable a seamless operation of platooning.

2.2.1 Centralized Protocol Design

Centralized protocol design is based on the concept that a single controlling entity, usually the lead vehicle or an external server, manages communication and decision-making for the entire platoon. The main benefit of this approach is that the controller has access to the global state of the platoon, allowing for highly optimized and coordinated decisions. In practice, centralized control involves the lead vehicle gathering real-time data from all other vehicles in the platoon, processing this data, and then sending back optimized commands to ensure that all vehicles operate in unison.

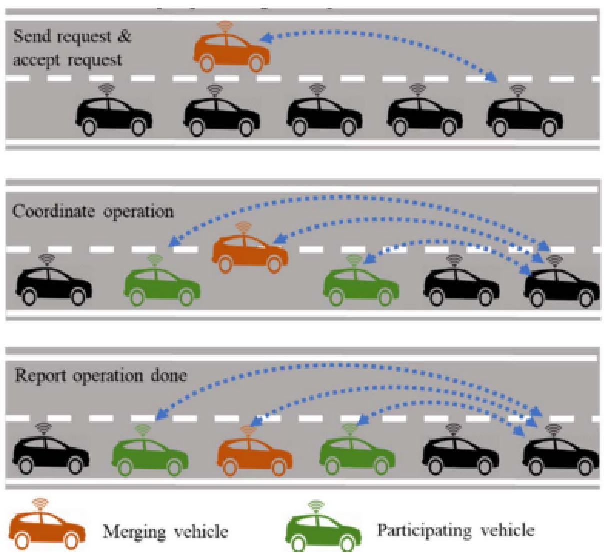
However, the reliance on a single controlling entity introduces potential vulnerabilities. Should the central controller fail, the entire platoon could be compromised. Moreover, as the platoon size grows, the computational and communication demands on the central controller can become overwhelming, limiting the scalability of the approach.

Algorithm:

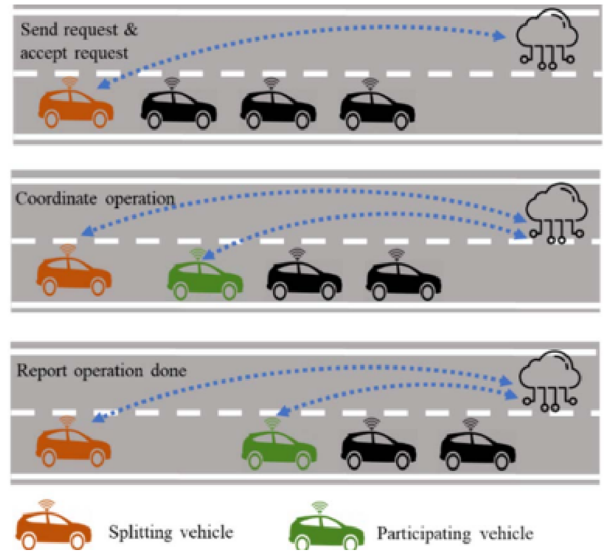
```
Algorithm CentralizedProtocolDesign
Input: Platoon P, CommunicationRange r, ControlCommands C
Output: Coordinated Platoon Behavior

Initialize: Central controller establishes communication links with all
    vehicles
Broadcast initial control commands C to all vehicles
for each time step t do
    Central controller collects data from each vehicle (position, speed, etc.)
    Process collected data to assess platoon state
    Determine new control commands based on the global platoon state
    Broadcast updated control commands to all vehicles
    Monitor execution of commands and adjust if necessary
end for
Finalize the coordinated movement and disband communication
```

This algorithm outlines the steps involved in a centralized protocol design. The central controller first establishes communication with all vehicles in the platoon, collecting real-time data on their position, speed, and other critical parameters. Based on this data, the controller determines the optimal control commands to ensure smooth and efficient operation of the platoon. These commands are then broadcast to all vehicles, and the controller monitors their execution, adjusting them if necessary. Centralized systems excel in their ability to make globally optimal decisions but suffer from scalability issues and vulnerabilities to communication delays or controller failures.



(a) Centralized Merging. This figure depicts how the central controller guides the process of a vehicle merging into the platoon. The vehicle adjusts its speed and position based on commands from the controller, which ensures that the merge occurs smoothly and safely.



(b) Centralized Splitting. This illustration shows how the central controller facilitates a vehicle's exit from the platoon. The controller manages the speed and lane changes necessary for the vehicle to safely leave the platoon while maintaining the integrity of the remaining vehicles.

Figure 2.2: Centralized Merging and Splitting. These figures illustrate the processes of vehicles either merging into or leaving a platoon under a centralized control system. The central controller's role is crucial in coordinating these maneuvers to ensure that the platoon's overall structure remains intact and operationally efficient.

Despite its advantages, the centralized approach is often criticized for being too rigid and lacking scalability. The more vehicles that join the platoon, the more complex the communication and decision-making become. Additionally, any failure in the central controller can cause catastrophic failures across the entire system, as there is no fallback mechanism. This limitation has led to increased interest in decentralized systems.

2.2.2 Decentralized Protocol Design

Decentralized protocol design operates on a fundamentally different principle. Rather than having a single entity responsible for decision-making, control is distributed across all vehicles in the platoon. Each vehicle independently gathers information from its neighbors—typically through sensors or short-range communication—and makes decisions based on this local data. This allows for greater scalability, as the system can accommodate additional vehicles without placing an additional burden on any central controller.

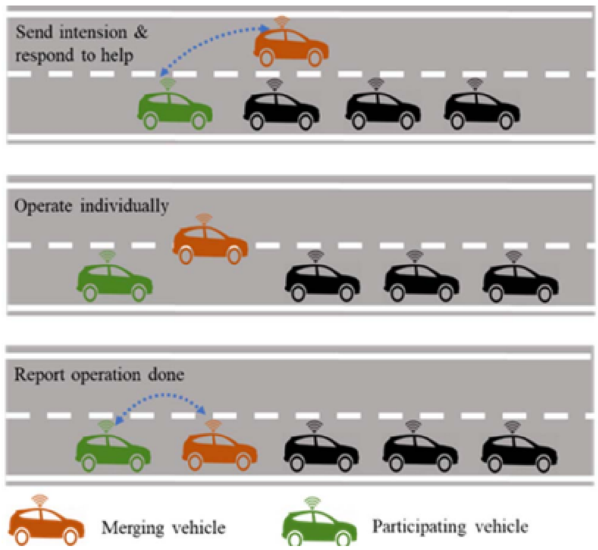
In decentralized systems, each vehicle adjusts its behavior to maintain safe distances and coordinated speeds with neighboring vehicles. These systems tend to be more robust than centralized ones because the failure of a single vehicle does not impact the overall system. However, decentralized systems often sacrifice efficiency for robustness, as vehicles do not have access to the global state of the platoon and must rely on heuristics or local optimization techniques.

Algorithm:

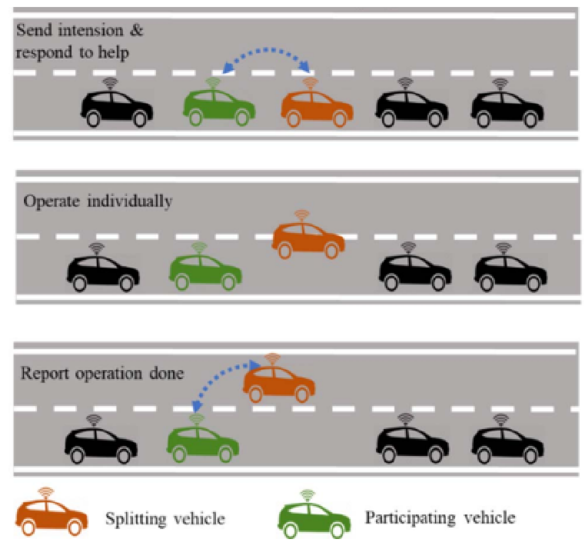
```
Algorithm DecentralizedProtocolDesign
Input: Local vehicle information, Neighboring vehicles' data
Output: Coordinated Local Behavior

Initialize: Each vehicle sets up communication with neighboring vehicles
for each time step t do
    for each vehicle i do
        Receive information from neighboring vehicles (e.g., position, speed)
        Process local and received data to determine vehicle i's state
        Make local decisions (e.g., speed adjustment, lane changes)
        Communicate decisions and state to neighboring vehicles
    end for
end for
Finalize local behavior and adjust as needed based on new information
```

In this decentralized protocol design algorithm, each vehicle collects information from its neighboring vehicles and uses this data to adjust its speed, position, and other parameters in real-time. Unlike centralized systems, decisions are made locally, which enhances the system's robustness. Decentralized systems are highly scalable, and because each vehicle is responsible for its own decisions, they are less prone to catastrophic failure. However, they may struggle to achieve the same level of coordination as centralized systems, particularly during complex maneuvers like merging and splitting.



(a) Decentralized Merging. In this example, a vehicle is attempting to merge into a platoon without the assistance of a central controller. The vehicle relies on data from its neighboring vehicles and makes decisions about speed and lane changes based on local observations.



(b) Decentralized Splitting. Similar to the merging process, decentralized splitting involves a vehicle deciding when and how to leave the platoon based on its local environment. The lack of centralized coordination makes this process more flexible but potentially less efficient.

Figure 2.3: Decentralized Merging and Splitting. These figures illustrate the processes of vehicles either merging into or splitting from a platoon in a decentralized control system. Each vehicle makes its own decisions based on local information, offering a robust yet potentially less efficient alternative to centralized control.

Decentralized systems excel in their ability to scale without imposing excessive communication or computational demands on any single vehicle. However, the lack of centralized coordination can lead to inefficiencies, particularly during complex maneuvers such as merging or splitting. Vehicles must rely on local information to make decisions, which can result in suboptimal performance compared to centralized systems where global information is available.

2.2.3 Comparing Centralized and Decentralized Protocols

Centralized protocols offer significant efficiency in platoon operations due to the coordinated actions among vehicles, making them particularly effective in enhancing traffic mobility for both the platoon and surrounding traffic. This advantage likely explains why many existing studies favor centralized protocols. However, centralized systems depend heavily on advanced Vehicle-to-Vehicle (V2V) and/or Vehicle-to-Infrastructure (V2I) communications, which present challenges in practical implementation, particularly in the early stages of Connected and Automated Vehicle (CAV) development. Additionally, centralized protocols place a substantial computational load on the platoon leader, who is responsible for making decisions for all vehicles. A major drawback of this approach is the reduced system robustness; if the platoon leader encounters a failure—due to hardware malfunctions or cyberattacks—the entire platoon's operation is compromised. On the positive side, only the platoon leader has access to comprehensive information about all vehicles, thereby reducing privacy risks since individual vehicles cannot access data about others.

In contrast, decentralized protocols address many of the limitations associated with centralized systems. In a decentralized setup, each vehicle perceives its environment through sensors and/or short-range communication technologies and adjusts its actions based on the information it gathers. This structure reduces both communication and computational demands, requiring fewer resources in terms of software (such as high-performance computing units) and hardware (like long-range communication devices). Decentralized protocols also offer enhanced system robustness by distributing communication and computation tasks across all vehicles. However, this comes at the cost of operational efficiency, as decentralized systems lack the coordinated vehicle management provided by a central controller, resulting in longer operation times. Moreover, since all vehicles have access to each other's information in a decentralized system, the risk of privacy breaches is higher.

Table 2.1: Management Protocols and Merging/Splitting Positions in Different Studies. This table summarizes various research studies comparing centralized and decentralized platoon management protocols, as well as their approaches to handling vehicle merging and splitting positions in platoons. The columns indicate where the merging or splitting takes place—at the head, middle, or tail of the platoon. The sources cited in this table are used to categorize and compare their approaches and findings, offering a broad perspective on the effectiveness of different management protocols.

Study	**Management Protocol**		**Merging Position**			**Splitting Position**		
	Centralized	**Decentralized**	**Head**	**Middle**	**Tail**	**Head**	**Middle**	**Tail**
[10]	✓		✓		✓		✓	✓
[14]		✓		✓				
[4]		✓		✓				
[9]	✓			✓		✓		
[15]	✓		✓			✓		
[16]		✓				✓	✓	
[2]	✓					✓		
[5]	✓				✓		✓	
[1]	✓					✓	✓	
[8]	✓						✓	
[12]		✓		✓				
[13]		✓	✓			✓		
[11]	✓			✓				
[3]	✓					✓		
[17]		✓	✓	✓		✓	✓	

Table 2.1 provides a detailed comparison of various research studies on platoon management protocols and merging/splitting strategies. The table highlights the diverse approaches taken by different studies, offering insights into the effectiveness of centralized versus decentralized management, as well as the positions where merging and splitting maneuvers are typically performed within a platoon. For instance, studies such as [10] and [15] favor centralized control with merging typically occurring at the head or tail of the platoon, while decentralized systems such as [14] and [4] emphasize flexibility by allowing vehicles to merge in the middle of the platoon. Such comparisons reveal the trade-offs between the two paradigms, with centralized systems often excelling in efficiency and coordination, while decentralized approaches provide greater scalability and robustness.

2.3 Conclusion

In this chapter, we have provided a comprehensive exploration of vehicle platooning strategies, particularly focusing on centralized and decentralized protocol designs, trajectory planning, and their comparative advantages and limitations. Centralized protocols offer optimized coordination by leveraging a global view of the platoon, ensuring efficient management of vehicle behaviors such as merging, splitting, and lane-changing. However, the computational and communication demands placed on a single controller, along with potential single points of failure, limit the scalability and robustness of centralized systems.

On the other hand, decentralized protocols distribute decision-making responsibilities across individual vehicles, enhancing system robustness and scalability. This approach allows vehicles to react dynamically to local conditions without relying on a central controller. While decentralized systems excel in managing large-scale platoons, they may suffer from suboptimal coordination due to the lack of a global perspective. The trade-off between robustness and optimality is a key consideration when choosing between centralized and decentralized designs, and various research works demonstrate these paradigms' performance in different contexts.

The chapter has also discussed the critical role of trajectory planning in ensuring that vehicles can maneuver efficiently and safely within a platoon. Centralized trajectory planning, though computationally intensive, allows for globally optimal decisions, while decentralized methods enable each vehicle to compute its trajectory based on local information, enhancing system flexibility.

Throughout this chapter, we have analyzed multiple studies to highlight how different approaches to vehicle platooning have been developed and implemented. The comparative study in Table 2.1 emphasizes the various strategies researchers have employed in addressing the merging and splitting positions within platoons, illustrating the diversity of approaches in both centralized and decentralized systems.

Looking forward, one of the key challenges in vehicle platooning research will be to balance the strengths of centralized and decentralized systems, potentially developing hybrid approaches that combine the efficiency of centralized coordination with the robustness of decentralized decision-making. Additionally, as autonomous vehicle technology advances, there is a growing need for more sophisticated machine learning techniques, such as reinforcement learning, to dynamically adapt to increasingly complex traffic environments. By addressing these challenges, future research can help pave the way for the large-scale deployment of vehicle platooning systems that improve traffic efficiency, reduce fuel consumption, and enhance road safety on a global scale.

The following chapters will delve deeper into the technical aspects of vehicle control, optimization techniques, and machine learning applications in CAV platooning, building on the foundation established in this chapter to further enhance the understanding and development of these systems.

Conclusion

In this document, we have explored the intricate dynamics of connected and automated vehicle (CAV) platooning, focusing on the critical aspects of protocol design and trajectory planning, as well as the application of reinforcement learning (RL) in optimizing these operations. The discussion has spanned both centralized and decentralized approaches, providing a thorough understanding of their respective strengths, weaknesses, and applicability in various scenarios.

2.4 Summary of Centralized and Decentralized Approaches

The centralized and decentralized approaches to protocol design and trajectory planning represent two fundamental paradigms in the management of vehicle platoons. Centralized systems, where a single entity (usually the platoon leader) orchestrates the actions of all vehicles, offer significant advantages in terms of coordination and optimization. The centralized approach ensures that all vehicles operate in a harmonious and efficient manner, particularly when precise coordination is paramount, such as in urban environments or high-density traffic scenarios.

However, centralized systems are not without their drawbacks. The reliance on a single point of control introduces vulnerability; if the central controller fails, the entire system is at risk. Additionally, as the size of the platoon increases, the computational and communication demands on the central controller escalate, potentially leading to scalability issues. These challenges underscore the need for robust communication networks and sophisticated algorithms to manage the increased complexity.

In contrast, decentralized approaches distribute decision-making across all vehicles in the platoon. This method enhances system robustness, as the failure of a single vehicle does not compromise the entire platoon. Decentralized systems are also more scalable, as each vehicle only needs to communicate with its immediate neighbors, reducing the overall communication burden and allowing the system to adapt more fluidly to changes in platoon composition or traffic conditions.

Despite these advantages, decentralized systems face challenges in achieving the same level of coordination and optimization as centralized systems. Without a global view of the platoon, vehicles must rely on local information and heuristic-based rules, which can result in suboptimal performance. The increased latency in decision-making and the need for complex algorithms to ensure effective coordination further complicate the implementation of decentralized systems.

2.5 The Role of Trajectory Planning in Platooning

Trajectory planning is a pivotal aspect of CAV platooning, ensuring that vehicles can merge, split, and maneuver safely and efficiently within a platoon. The approaches to trajectory planning mirror the division between centralized and decentralized systems. Centralized trajectory planning benefits from a global perspective, allowing for the optimization of trajectories that minimize fuel consumption, reduce travel time, and enhance safety. This optimization is achieved through sophisticated mathematical models and algorithms, such as Linear Programming (LP), Quadratic Programming (QP), and Mixed-Integer Linear Programming (MILP).

However, centralized trajectory planning is heavily dependent on the reliability and speed of communication between the central controller and the vehicles. Any delays or failures in communication can disrupt the execution of planned trajectories, leading to safety risks. Furthermore, the computational demands of centralized planning increase with the size of the platoon, necessitating powerful processors and efficient algorithms to handle the complexity.

Decentralized trajectory planning, on the other hand, allows vehicles to plan their paths based on local information. This approach is more flexible and can adapt quickly to changes in the environment, such as the sudden appearance of obstacles or changes in traffic flow.

Decentralized planning typically employs simpler, rule-based methods, such as linear and nonlinear control laws, to determine vehicle trajectories. While this approach is less optimal than centralized planning, it offers greater robustness and scalability, making it suitable for dynamic and unpredictable environments.

2.6 Reinforcement Learning in Vehicle Platooning

The integration of reinforcement learning (RL) into vehicle platooning represents a significant advancement in the field, offering the potential to overcome some of the limitations of traditional centralized and decentralized approaches. RL algorithms enable vehicles to learn optimal behaviors over time through interaction with the environment, without the need for explicit programming of every possible scenario.

In the context of CAV platooning, RL can be applied to both centralized and decentralized systems. In a centralized system, an RL agent (e.g., the platoon leader) can learn to optimize the overall performance of the platoon by adjusting the trajectories and speeds of all vehicles based on the current state of the system. This learning process is facilitated by techniques such as Q-learning and Deep Q-learning (DQN), which use reward signals to iteratively improve the decision-making policy.

The introduction of techniques like Experience Replay and Target Networks in DQN has significantly enhanced the stability and efficiency of learning in complex, high-dimensional environments. These techniques allow the RL agent to break the correlation between consecutive learning samples and to stabilize the training process by maintaining a separate network for generating target values. The result is a more robust and reliable system that can effectively manage the dynamic and stochastic nature of real-world traffic environments.

In decentralized systems, RL can be employed to allow individual vehicles to learn cooperative behaviors that contribute to the overall efficiency and safety of the platoon. Each vehicle, acting as an independent RL agent, can learn policies that optimize its interactions with neighboring vehicles, leading to emergent behavior that enhances the performance of the entire platoon. This approach is particularly advantageous in scenarios where centralized control is impractical due to communication constraints or the need for rapid adaptation to environmental changes.

However, the application of RL in CAV platooning is not without challenges. One of the primary issues is the need for extensive training data, which can be difficult to obtain in real-world traffic scenarios. Additionally, RL algorithms often require a large number of iterations to converge to an optimal policy, which can be computationally expensive. The exploration-exploitation trade-off is another critical challenge, as RL agents must balance the need to explore new strategies with the goal of exploiting known, successful behaviors.

2.7 Challenges and Future Directions

While significant progress has been made in the development of centralized and decentralized protocols, trajectory planning methods, and the application of RL in CAV platooning, several challenges remain. The scalability of centralized systems continues to be a concern, particularly as platoon sizes increase and traffic environments become more complex. Future research will need to focus on developing more efficient algorithms that can handle larger platoons without compromising performance or safety.

In decentralized systems, the primary challenge lies in achieving a level of coordination and optimization that approaches that of centralized systems. This will likely require advances in communication technologies, such as Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication, as well as the development of more sophisticated decision-making algorithms that can operate effectively with limited information.

The integration of RL into CAV platooning also presents several areas for future research. Improving the sample efficiency of RL algorithms will be crucial for their practical application, as will the development of techniques to reduce the computational demands of training. Additionally, there is a need for more robust exploration strategies that can help RL agents discover optimal policies in complex, high-dimensional environments.

Finally, as CAV platooning technology continues to evolve, it will be important to address the ethical and regulatory challenges associated with its deployment. Ensuring the safety and reliability of platooning systems will be paramount, particularly as they are integrated into mixed traffic environments with both automated and human-driven vehicles. Furthermore, the development of standards and regulations that govern the operation of CAV platoons will be essential to ensure their safe and efficient deployment on public roads.

We have provided a comprehensive exploration of the key concepts and methodologies related to CAV platooning, with a particular focus on protocol design, trajectory planning, and the application of reinforcement learning. Both centralized and decentralized approaches offer distinct advantages and face unique challenges, and the choice between them should be guided by the specific requirements of the application. The integration of RL represents a promising direction for future research, offering the potential to further enhance the efficiency, safety, and adaptability of CAV platooning systems. As research in this field continues, it is likely that new algorithms and technologies will emerge, further pushing the boundaries of what is possible in the realm of connected and automated transportation.

Bibliography

[1] I. J. P. Graells A. S. Ibrahim and C. Canudas de Wit. Platoon-based cooperative merging controller. *IEEE Trans. Intell. Veh.*, 2(2):96–107, 2017.

[2] A. Mahmassani A. Talebpour and F. Farahani. Modeling the network-wide impacts of connected autonomous vehicles. *Transp. Res. Rec.*, 2564(1):144–154, 2016.

[3] M. E. Barrows and R. A. Knapp. Platoon-based traffic signal control and urban corridor management with connected and autonomous vehicles. *Transp. Res. Rec.*, 2674(2):11–22, 2020.

[4] F. A. Bueno and R. De Almeida. Mixed-integer model for platoon formation of autonomous vehicles. *IEEE Trans. Intell. Transp. Syst.*, 19(7):2137–2147, 2018.

[5] Q. Wang C. Xu and B. J. R. A. Mohd. Safety and efficiency evaluation of connected and autonomous vehicles at roundabouts using microscopic traffic simulation. *IEEE Intell. Transp. Syst. Mag.*, 11(3):57–69, 2019.

[6] X. Chen et al. Evaluating decentralized vehicle trajectory planning methods. *Transportation Research Part C*, 103:45–58, 2019.

[7] Y. Chen et al. Future directions in cav platooning: Addressing scalability and ethical challenges. *IEEE Transactions on Intelligent Vehicles*, 7:654–667, 2022.

[8] K. K. Esfahani and A. Talebpour. Investigating the interactions between autonomous vehicles and other road users in urban corridors. *Transp. Res. Rec.*, 2620(1):12–23, 2017.

[9] M. I. Qureshi F. S. Tabar and R. Ghabcheloo. A model predictive control approach to platoon formation of autonomous vehicles at a signalized intersection. *IEEE Trans. Intell. Transp. Syst.*, 20(4):1335–1346, 2019.

[10] R. Hall and C. Chin. Vehicle sorting for platoon formation: Impacts on highway entry and throughput. *Transp. Res. C, Emerg. Technol.*, 13(5):405–420, 2005.

[11] F. Hou and D. Zhao. Intersection management based on connected vehicle technology: A comparison of centralized and decentralized methods. *IEEE Trans. Intell. Veh.*, 2(4):237–244, 2017.

[12] J. Lee and B. Park. Evaluation of cooperative adaptive cruise control system under mixed traffic scenarios. *IEEE Trans. Intell. Transp. Syst.*, 19(2):686–697, 2018.

[13] M. Levin and A. Boyles. A framework for modeling autonomous vehicle control at signalized intersections. *Transp. Res. C, Emerg. Technol.*, 52:37–52, 2015.

[14] M. Levin and M. V. C. Cunningham. Intersection management for autonomous vehicles. *Transp. Res. Part C: Emerg. Technol.*, 65:58–75, 2016.

[15] J. Ploeg S. Gunter and H. Nijmeijer. Cooperative path planning for a group of automated vehicles at intersections. *IFAC-PapersOnLine*, 51(9):44–49, 2018.

[16] J. Bared S. Zhao and A. Stevanovic. Optimization of intersection control for autonomous vehicles in a mixed traffic environment. *Transp. Res. Rec.*, 2672(35):107–117, 2018.

[17] T. Wu Z. Tian and X. Zhang. Cooperative adaptive cruise control in mixed traffic with connected and automated vehicles: A robust approach. *IEEE Trans. Intell. Transp. Syst.*, 21(6):2524–2534, 2020.