Impact of Sensing Errors on Headway Design: From α-Fair Group Safety to Traffic Throughput

Wei Shao, Postdoctoral Researcher, UC Institute of Transportation Studies, University of California, Davis
Zejun Fan, Visiting Student, Royal Melbourne Institute of Technology
Chia-Ju Chen, Ph.D. Student, UCLA Mobility Lab, University of California, Los Angeles
Zhaofeng Zhang, Ph.D. Student, School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe
Jiaqi Ma, Ph.D., Associate Professor, Samueli School of Engineering, Univerity of California, Los Angeles
Junshan Zhang, Ph.D., Professor, Electrical and Computer Engineering Department, University of California, Davis

August 2024

PRE-PRINT
Published in IEEE Transactions on Mobile Computing, vol. , no. 01, pp. 1-15, 5555. DOI: 10.1109/TMC.2024.3389987

Project No.: RIMI-5E

Impact of Sensing Errors on Headway Design: From α-Fair Group Safety to Traffic Throughput

Wei Shao, *Member, IEEE,* Zejun Fan, Chia-Ju Chen, Zhaofeng Zhang, Jiaqi Ma, Junshan Zhang, *Fellow, IEEE,*

***Abstract*—Headway, namely the distance between vehicles, is a key design factor for ensuring the safe operation of autonomous driving systems. There have been studies on headway optimization based on the speeds of leading and trailing vehicles, assuming perfect sensing capabilities. In practical scenarios, however, sensing errors are inevitable, calling for a more robust headway design to mitigate the risk of collision. Undoubtedly, augmenting the safety distance would reduce traffic throughput, highlighting the need for headway design to incorporate both sensing errors and risk tolerance models. In addition, prioritizing group safety over individual safety is often deemed unacceptable because no driver should sacrifice their safety for the safety of others. In this study, we propose a multi-objective optimization framework that examines the impact of sensing errors on both traffic throughput and the fairness of safety among vehicles. The proposed framework provides a solution to determine the Pareto frontier for traffic throughput and vehicle safety. ComDrive, a communication-based autonomous driving simulation platform, is developed to validate the proposed approach. Extensive experiments demonstrate that the proposed approach outperforms existing baselines.**

***Index Terms*—Autonomous Vehicles, Alpha Fairness, Sensing Errors.**

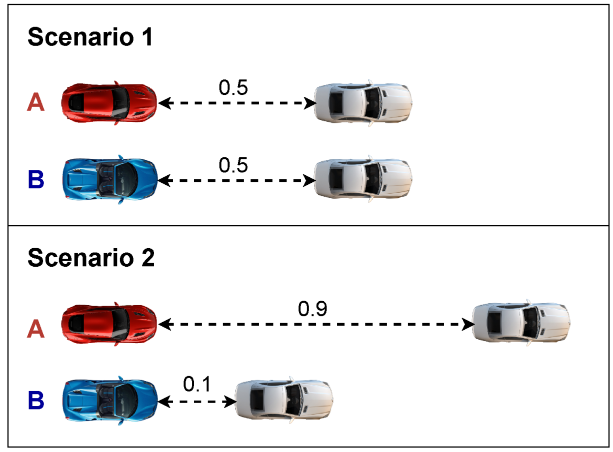


Fig. 1. Illustration of group safety and individual safety. Though the sum of the safety in both scenarios is the same, the risks of each individual vehicle are significantly different.

I. Introduction

AUTONOMOUS driving technology has the potential to revolutionize the transportation industry, profoundly impacting society, the economy, and human daily life [44], [59]. It is envisioned that autonomous driving can significantly enhance road safety by minimizing the number of accidents caused by human errors [21], improving mobility for those unable to drive, optimizing road usage, reducing traffic congestion, and enhancing travel time. Additionally, it presents significant economic benefits, such as cost savings for individuals and businesses. Concerning human daily life, autonomous driving technology can transform how people interact with their vehicles and how cities and suburbs are designed by enabling individuals to use their commute time for other activities, minimizing the need for car ownership, and reducing pollution and traffic in urban areas. However, it also poses significant challenges, including safety issues caused by sensing errors, reduced traffic throughput if not planned well, and fair treatment of all vehicles.

Sensing is a critical component of autonomous driving technology, as it allows the vehicle to detect and respond to its surroundings, including other vehicles, pedestrians, and obstacles. However, sensing errors can occur due to various factors, such as adverse weather conditions, poor lighting, and sensor malfunctions. These errors can result in inaccurate information about the vehicle's surroundings, potentially leading to unsafe driving conditions. Regarding safety in autonomous driving, sensing errors can cause inaccurate estimation of the safety distance (distance headway), potentially increasing the risk of collisions [21], [58]. For instance, imagine a vehicle equipped with a sensor system that can detect the distance between it and the vehicle in front of it. Suppose the sensor system's sensing distance is much greater than the minimum safe distance. In that case, the vehicle may accelerate to catch up to the leading vehicle, potentially causing the distance between the two vehicles shorter than expected. This behavior could increase the risk of a rear-end collision. On the other hand, Underestimating the distance headway can result in an unnecessarily large distance headway, reducing traffic throughput. Determining the appropriate safety distance is a delicate balance that depends on the accuracy and reliability of sensors used to measure the distance, as well as the risk tolerance of the driving model. In practice, perfect sensing is challenging due to errors in onboard sensors such as radar, GPS, and cameras. These errors can result in incorrect calculations of the time to collision (TTC) [30], a critical factor in determining the safety distance. Therefore, addressing sensing errors and developing robust risk tolerance models is critical to ensure safe and efficient autonomous driving.

While much research on studying sensing errors has focused on the influences on the safety issue of a single vehicle named individual safety, there has been limited consideration of group safety and fairness. In traditional safety evaluation methods, group safety is often measured as the sum of the safety indicators for each vehicle, which can be unfair and potentially dangerous for individual vehicles. As illustrated in Figure 1, consider two scenarios involving vehicles A and B. In the first scenario, both vehicles have a time-to-collision (TTC) safety measurement of 0.5. In the second scenario, the TTC measurement for vehicle A is 0.9, while that for vehicle B is 0.1. While the total safety measurement is the same in both scenarios, the risks for each vehicle are significantly different. Vehicle A in the first scenario has more time to avoid a collision, while Vehicle B in the second scenario has little room for error. However, increasing the safety of one vehicle cannot fully compensate for a decrease in another vehicle's safety. Therefore, balancing individual and group safety is a significant challenge for future autonomous driving. Ensuring that no individual vehicle would sacrifice too much of its safety to improve the safety of other vehicles is essential. Fairness, as a crucial concept in modern society, should be integrated into the design and evaluation of autonomous driving systems to ensure that the benefits and risks are distributed equitably among all vehicles on the road.

To address these challenges, we propose FairAV, a multi-objective reinforcement learning-based approach for autonomous driving with stochastic sensing errors and fairness considerations. Compared to existing approaches, FairAV achieves (1) Real-time balance between traffic throughput and safety with sensing errors: Using a multi-objective reinforcement learning approach, FairAV can obtain a Pareto sub-optimal solution in real-time, balancing the two objectives of traffic throughput and group safety with sensing errors. It accomplishes this by setting the traffic throughput and safety as two contradicted objectives and adding a variable-length compensated margin to the sensing headway as the action. (2) Fairness in vehicle safety: We designed a new safety metric called α-fair group safety that guarantees the minimum safety for individual vehicles and strikes a balance between group safety and individual safety. Using our multi-objective optimization framework, we can provide a fairer guarantee of safety among vehicles.

Our main contributions can be summarized as follows:

- To the best of our knowledge, we are the first to propose the concept of safety fairness in autonomous driving. We design a novel fair group safety metric named α-fair group safety to ensure safety fairness among autonomous vehicles. This approach also provides insights into balancing individual and group safety.
- To enhance group safety, we develop a method to counteract the adverse effects of sensing errors on detected headway. Alongside this, we introduce a bi-objective reinforcement learning solution aimed at achieving a balance between group safety and traffic throughput.
- Our extensive experimentation with FairAV, our developed approach, demonstrates its effectiveness in mitigating sensing errors' impact on safety and traffic flow. We also evidence the benefits of incorporating fairness into vehicle safety. Comparatively, our approach shows notable improvements over traditional multi-objective optimization, as evidenced in trials on our developed ComDrive simulation platform.

II. Related Work

In this section, we explore recent advancements in multi-objective optimization and fairness in cooperative driving, specifically within the framework of autonomous vehicle technology.

A. Cooperative Driving

The field of cooperative driving, especially in the realm of autonomous vehicles, is rapidly progressing with a strong focus on improving traffic efficiency and safety.

The initial stages of cooperative driving research delved into the fundamental aspects of vehicle-to-vehicle (V2V) communication. Research by Wang *et al.* [53] demonstrated that leveraging V2V communications for cooperative driving can significantly reduce traffic congestion and enhance both traffic capacity and stability. The concept of vehicle-to-infrastructure (V2I) communication is equally vital in the cooperative driving framework. Djahel *et al.* [14] introduced a model wherein roadside infrastructure, such as traffic light controllers at intersections, transmits traffic signal timing to oncoming vehicles. Utilizing this data, the vehicles collaboratively calculate their optimal speeds and other necessary maneuvers to navigate intersections with minimal delay and, ideally, without the need to stop. Based on the principles of V2V and V2I communication, Wang et al. proposed Connected Automated Vehicles (CAVs), which can achieve optimal traffic efficiency while ensuring safety, energy efficiency, and a comfortable experience for passengers.

Recent studies in cooperative driving have predominantly concentrated on three areas: enhancing cooperative perception, optimizing communication strategies, and refining control tactics. Cooperative perception has been a key focus, employing wireless communication technologies to merge environmental data collected by edge nodes with local sensory inputs. This integration not only boosts the precision of vehicle perception but also minimizes latency and addresses perception blind spots, as explored by Cui *et al.* [9]. Another vital element in this field is the optimization of communication. Addressing network latency fluctuations and the limitations of current mobile network bandwidth is crucial for enhancing the reliability and safety of autonomous vehicle operations, a topic Liu *et al.* [34] have investigated in depth. Lastly, the advancement in control strategies forms a crucial part of cooperative driving. This involves the development of sophisticated approaches like trajectory planning [51], Model Predictive Control (MPC) [41], and the creation of scenario-specific rules to mitigate traffic congestion [10], [55].

These research efforts collectively highlight the transformative potential of cooperative driving in revolutionizing transportation systems. However, they primarily concentrate on assessing the performance metrics of cooperative driving as a collective entity, while often overlooking the aspect of fairness among individual vehicles.

B. Multi-objective Optimization

We have formulated our problem as a bi-objective optimization problem and aim to find a sub-optimal solution that maximizes both objectives while satisfying certain constraints. In multi-objective optimization, three categories of methods can typically be used: enumerative, deterministic, and stochastic [40].

Enumerative methods evaluate all feasible solutions and are impractical for many real-world multi-objective optimization problems due to the vast solution space [8]. Deterministic methods, such as Greedy [38], Hill Climbing [13], and best-first algorithms [36], rely on gradients or derivatives. However, these methods are unsuited for optimizing problems with many design variables and multiple local minima, maxima, and saddle points [8]. Stochastic techniques include meta-heuristic optimization algorithms and iterative approach-based algorithms from a set of randomly generated candidate solutions. Such methods are widely used in multi-objective optimization, particularly in engineering. Although meta-heuristic optimization algorithms can find globally optimal solutions with a well-designed fitness function, their high computational cost and one-time use data are unsuitable for dynamic environments that require real-time decision-making. Furthermore, meta-heuristic approaches rely on a deep understanding of the environment and an explicit physical model, such as the driving behaviors of all vehicles, which is difficult to obtain in a real-world autonomous driving setting.

Recently, multi-objective reinforcement learning (MORL) has become a popular approach for solving multi-objective optimization problems. MORL has many advantages over meta-heuristic optimization methods for real-time and flexible decision-making, which is crucial for autonomous vehicle driving. Reinforcement learning (RL) [27] is a set of algorithms that enable agents to learn how to take action in various situations. In single-objective reinforcement learning, the goal is to find a function that maps states to actions that maximize the reward received over time. In contrast, MORL involves maximizing multiple objectives, each representing performance on a particular goal. MORL addresses sequential decision-making problems that involve balancing competing objectives [25]. Autonomous driving involves making continuous, optimal decisions over time in a complex, high-dimensional environment [31]. This decision-making process can be formalized using classical RL techniques, enabling the autonomous vehicle to learn about and adapt to its environment optimally. However, due to the combinatorial size of this environment, learning and decision-making in autonomous driving present a significant challenge [50]. In recent years, deep reinforcement learning (DRL) has been widely used in autonomous driving, particularly in multi-agent systems. DRL enables agents to generalize their knowledge to new, unseen situations and has led to the development of new algorithms for handling continuous state and action spaces [28]. This has made DRL particularly useful in the context of autonomous driving.

C. Fairness

Recently, there has been a growing interest in fairness in both the design of applications and algorithms. For example, Raman *et al.* studied fairness in ride-sharing platforms and developed methods for reducing income fluctuations and inequality by incorporating fairness constraints into the objective function and redistributing income to drivers [43]. Belahcene *et al.* [3] have also introduced fairness into the joint selection and allocation of items to a population. Regarding algorithms, researchers have used reinforcement learning in pursuit-evasion games to study fairness in multi-agent systems [20], [26], and have incorporated fairness into deep neural networks [1], [15].

The issue of fairness in autonomous driving systems has captured the attention of researchers across various domains [4], [12]. For example, in a dynamic vehicle routing problem with stochastic requests and a given time horizon, Soeffker *et al.* [48] investigated the trade-off between higher acceptance rates and increased fairness. Wang *et al.* [52] developed the first data-driven fairness-aware displacement system to improve overall profit efficiency and fairness for e-taxi fleets. In 2022, Yin *et al.* [61] jointly optimized passenger allocation and vehicle scheduling to maximize transportation network utility, focusing on economic vehicle use and fair passenger allocation. Furthermore, in 2017, Hacker proposed three interventions to promote fairness in the code of autonomous systems, including mandatory "data safe" alternatives, personalized data protection, and procedural rules on algorithmic fairness [23].

Despite these efforts, little attention has been devoted to fairness in the context of driving safety in autonomous driving. Therefore, further exploration is necessary to ensure that safety is equitably distributed among all vehicles on the road.

III. Preliminary

In this section, we introduce the foundational concepts pertinent to our research, including the two objectives of our problem: traffic throughput and safety. Additionally, we will define and discuss the distinctions between group safety and individual safety in the autonomous driving field.

A. Traffic Throughput Φ

Traffic throughput refers to the number of vehicles passing through a given roadway or transportation network within a specific time frame, typically measured in vehicles per hour or day [17]. It is an essential factor in transportation engineering and planning, as it can impact the congestion level on roadways, the travel time for commuters, and the overall efficiency of the transportation network. Maximizing traffic throughput can help reduce traffic congestion and improve the mobility of people and goods while minimizing the number of accidents on the road.

According to [17], the traffic throughput can be defined as:

$$\Phi(T, x) = \frac{m}{T}, \tag{1}$$

where m is the number of vehicles passing the fixed point x during an interval T.

Assume that n is the number of vehicles on the road we design. In our model, the interval T is short. Apparently, $m < n$. Thus, the traffic throughput defined above represents the impact of some vehicles and cannot make a good assessment of all vehicles. Thus, we redefine the traffic flow.

For a circular road, if just one vehicle is on the road, then

$$\Phi = \frac{1}{t} = \frac{v}{s}, \tag{2}$$

where v is the velocity of the vehicle and s is the length of the road.

Given the above definition, for n vehicles,

$$\Phi = \sum_{i}^{N} \frac{v}{s} = \frac{\overline{v}}{s} \times n. \tag{3}$$

Note that n can be calculated as follows:

$$n = \frac{s}{\overline{h}}, \tag{4}$$

where $\overline{h}$ denotes the average distance headway of n vehicles.

Thus, we can transform Φ into this way:

$$\Phi = \frac{\overline{v}}{\overline{h}} \tag{5}$$

Based on Eq. 5, we get the definition of individual traffic throughput:

$$\phi_i = \frac{v_i}{h_i}. \tag{6}$$

Finally, the total traffic throughput can be redefined as:

$$\Phi = \sum_{i=1}^{n} \phi_i = \sum_{i=1}^{n} \frac{v_i}{h_i}, \tag{7}$$

where n is the number of vehicles, v_i is velocity of vehicle i and h_i denotes the headway between vehicles i and $i+1$.

B. Group Safety vs. Individual Safety

To assess the safety of vehicles in traffic, various indicators are used to determine whether a vehicle is operating safely [56]. Generally speaking, individual safety and group safety are the two most popular indicators. In the realm of autonomous driving, the term individual safety refers to protecting the occupants of a single vehicle. In contrast, group safety in autonomous driving pertains to the safety of all road users. The importance of individual and group safety in developing and deploying autonomous driving technology cannot be overstated. Ensuring the safety of all road users is crucial for the widespread adoption of this technology, and careful consideration of individual and group safety concerns is paramount.

1) Individual Safety: Time to collision (TTC) [30] is one of the most popular indicators used to measure a vehicle's individual safety. By estimating the time required for a potential collision, TTC provides a useful metric for evaluating the effectiveness of collision avoidance systems and other safety features. In addition, many modern vehicles now come equipped with collision avoidance systems that use TTC as a critical parameter for assessing the risk of a collision and triggering automatic emergency braking or other safety measures. As such, TTC has become an important metric for evaluating the safety of vehicles and for promoting the development and implementation of advanced collision avoidance technologies. Hence, we use it to measure each vehicle's individual safety ψ_i. Without loss of generality, assume vehicle j is the front vehicle of the subject vehicle i. Then the individual safety of the vehicle i can be defined as:

$$\psi_i = \frac{h_i - \ell_j}{v_i - v_j}, \tag{8}$$

where h_i is the headway of vehicle i, ℓ_j denotes the length of the front vehicle j, and v_i and v_j are velocities of vehicles i and j.

2) Group Safety: In the context of autonomous vehicles, group safety refers to the safety of all individuals sharing the road. In this study, we define group safety as the safety of all autonomous vehicles sharing the same road networks, which is measured as the sum of the individual safety of each vehicle. Specifically, we use the sum of the time-to-collision (TTC) of each vehicle to define the group safety function $\Psi(TTC)$ as follows:

$$\Psi(TTC) = \sum_{i}^{n} \psi_i \tag{9}$$

While this definition can effectively measure the overall safety of vehicles within the same road network, it lacks consideration for fairness among vehicles. This omission can potentially result in harm to specific vehicles in the network. In essence, we cannot disregard individual safety in the pursuit of overall safety. Our objective is to enhance overall safety while ensuring that each individual receives equitable security protection. To address this issue, we introduce a novel metric in the following section.

IV. Problem Formulation

This section presents three proposed solutions to the aforementioned problems. Firstly, we propose an error-aware margin to alleviate the impact of headway sensing errors. Secondly, we introduce an α-fair group safety metric to promote fairness in safety measurements. Lastly, we formulate the problem as a bi-objective optimization problem and present a reinforcement learning-based approach to strike a balance between traffic throughput and group safety.

A. Sensing-error-aware Headway Margin

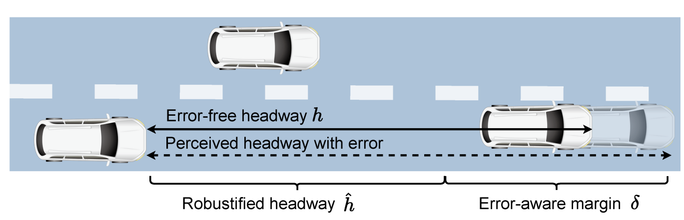


Fig. 2. Illustration of error-free headway and perceptual headway: Error-aware margin denotes an additional margin incurred by sensing errors beyond the error-free headway of the subject vehicle.

As illustrated in Fig. 2, we introduce an error-aware headway margin, denoted as δ, to mitigate the impact of sensing errors on vehicle safety. In specific scenarios where the sensing device's noise persistently underestimates the true headway, the following vehicle may be slowed down. To counteract this, we can employ a negative error-aware margin as a compensatory measure, effectively increasing the perceived headway and allowing the following vehicle to maintain a higher velocity without compromising safety. Determining the optimal δ value depends on various factors, such as the sensor technology employed, environmental conditions, and the characteristics of the headway sensing errors.

Our proposed reinforcement learning-based approach, which is presented later in the paper, is designed to adapt to sensing uncertainties and learn the balance between traffic throughput and group safety in the presence of sensing errors. This learning process intrinsically accounts for the sensing errors when optimizing the error-aware margins.

With the inclusion of the error-aware margin, the corresponding modified individual safety of vehicle i is given by:

$$\hat{\psi}_i = \min\left(\frac{h_i + \delta_i - \ell_j}{v_i - v_j}, \overline{TTC}\right) \tag{10}$$

where $\overline{TTC}$ is the maximum meaningful value for the time-to-collision, and δ_i denotes the margin added to the headway of vehicle i. Other variables are the same as in Eq. 8. We can see that a positive margin is likely to increase individual safety, and a negative one might decrease it.

Additionally, the modified traffic throughput is defined by:

$$\hat{\phi}_i = \sum_{i=1}^{n} \frac{v_i}{h_i + \delta_i}, \tag{11}$$

By incorporating the error-aware margin in the individual safety and traffic throughput equations, we recognize that the specific value of the margin can influence the resulting traffic throughput. This emphasizes the need to strike a balance between safety and efficiency when determining the error-aware margin to avoid promoting overly conservative or aggressive driving behaviors. Furthermore, the inclusion of the error-aware margin may increase the computational complexity of the optimization problem, necessitating more efficient algorithms or approximations to find suitable solutions in a timely manner.

B. α-fair Group Safety

Many autonomous driving systems focus on maximizing group safety. While it is true that in some instances, sacrificing the safety of certain vehicles can result in a substantial enhancement of the overall safety evaluation, this approach may also lead to vehicle collisions, which are unacceptable for safe driving. The current group safety metrics fail to ensure fairness in individual vehicle safety. In response to this gap in considering fairness within group safety assessment, we propose a new metric in this section: α-fair group safety Ψ(fairness). This metric evaluates the group safety of autonomous vehicles that share the same road network while taking fairness into account. Exactly, when the individual security assessments within the group are both higher and more consistent, the resulting α-fair group security metric will be higher as well.

Ensuring the individual safety of each vehicle can be regarded as a single resource allocation problem that has been extensively studied. Based on the work [29], [45], we employ a fairness scalarization function proposed by [29] that tackles fairness-efficiency trade-off to define the fairness of safety in this problem as follows:

$$\Psi = \lambda_f \ell\left(f_\beta(\mathbf{\Psi})\right) + \ell\left(\sum_i \psi_i\right) \tag{12}$$

$$\ell(y) = \text{sign}(y)\log(|y|) \tag{13}$$

$$f_\beta(\mathbf{\Psi}) = \text{sign}(1-\beta) \cdot \left[\sum_{i=1}^{N}\left(\frac{\psi_i}{\sum_j \psi_j}\right)^{1-\beta}\right]^{\frac{1}{\beta}}, \tag{14}$$

where $\beta \in (0,1) \cup (1,\infty)$ determines the type of fairness metric, $\lambda_f \in (0,\infty)$ indicates the emphasis on fairness component, and ψ_i indicates the individual safety of vehicle i.

Although we employ Eq. (12) to guarantee fairness for each vehicle, it does not mean we can sacrifice any individual safety without a limit. Hence, we set a constraint that $\forall \psi_i \in \Psi, \psi_i > 0$. In the real world, $\psi_i > 0$ suggests that the headway of vehicle i is not allowed below the 0 according to Eq. 20. Additionally, Corollary 7 in [29] gives the lower bound of $f_\beta(\mathbf{\Psi})$ as

$$\text{sign}(1-\beta) \cdot \frac{\left(\mu\Gamma^{1-\beta} + 1 - \mu\right)^{\frac{1}{\beta}}}{\left(\mu\Gamma + 1 - \mu\right)^{\frac{1}{\beta}-1}}, \tag{15}$$

where $\Gamma = \frac{\psi_{max}}{\psi_{min}}$, and $0 \leq \mu \leq 1$. We can obtain the final objective of safety is as follows:

$$\Psi = \lambda_f log(f_\beta(\mathbf{\Psi})) + log(\sum_i^N \psi_i). \tag{16}$$

In order to preserve Pareto optimality [6], the necessary and sufficient condition on λ_f such that $\Phi_{\lambda_f}(y) > \Phi_{\lambda_f}(x)$ if y Pareto dominates x is

$$\lambda \leq \left| \frac{\beta}{1-\beta} \right|. \tag{17}$$

This framework adapts the α-fairness concept for safety resource allocation, allowing us to balance individual and group safety objectives. By applying the fairness scalarization function, we model the trade-offs between individual safety and group safety, capturing the fairness-efficiency trade-offs that arise when trying to balance the safety of individual vehicles and the overall group safety.

C. Problem Statement

Given N human-driven or autonomous vehicles, we aim to obtain a sub-optimal set of error-aware margins $\boldsymbol{\delta} = g(\epsilon)$ to simultaneously maximize the group safety Ψ and traffic throughput Φ in a specific road network with uncertain sensing

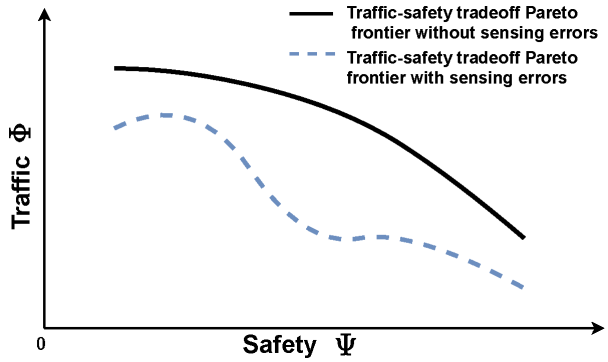


Fig. 3. The solid black line denotes a convex set representing the Pareto front of solutions in an error-free environment. The dashed blue line illustrates the real traffic-safety trade-off with sensing errors.

errors. Hence, our problem can be formalized as a multi-objective optimization problem as follows:

$$\max_{\boldsymbol{\delta}}(\Phi, \Psi). \tag{18}$$

In multi-objective optimization, it is rare to find a feasible solution that simultaneously optimizes all objective functions. Instead, focus is placed on finding Pareto optimal solutions [6], which cannot be improved in any of the objectives without degrading at least one of the other objectives as illustrated in Fig. 3. However, due to sensing errors, it may be difficult to achieve such a Pareto optimal solution in practice. Additionally, the lower safety bound cannot be guaranteed, which can significantly increase collision risks. As a result, we aim to pursue a sub-optimal solution in this study.

V. Methodology

This section provides a detailed exposition of our proposed solution to address the problem at hand. Our approach revolves around the application of multi-agent multi-objective (MAMO) deep reinforcement learning (DRL) to the problem. In addition, we present a novel reward function designed to evaluate traffic flow and safety on a per-vehicle basis, thereby offering a nuanced assessment of system performance. The structure of this section is as follows: First, we delve into the concept of Multi-objective Reinforcement Learning and how we utilize it to transform the problem into a multi-objective Markov decision process (MOMDP). Next, we discuss our implementation of Multi-Agent Reinforcement Learning, explaining our strategy to combat noise in deep Q-networks, particularly within the context of autonomous driving scenarios involving multiple agents. Following this, we outline our unique approach to calculating the reward in a cooperative multi-agent multi-objective system, with particular attention to the individual contribution of each agent. Lastly, we detail the adapted enveloped multi-objective Q-learning algorithm used for the simultaneous maximization of traffic and safety.

A. Problem Representation and Training Process

The multi-objective reinforcement learning (MORL) approach is utilized to address the problem at hand, conceptualizing it as a multi-objective Markov decision process (MOMDP). This is achieved by employing a tuple representation, $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \boldsymbol{r}, \Omega, f_\Omega \rangle$, where $\mathcal{S}$ denotes the state space, $\mathcal{A}$ signifies the action space, $\mathcal{P}(s'|s,a)$ represents the transition distribution, $\boldsymbol{r}(s,a)$ is the vector reward function, Ω is the preference space, and f_Ω is the preference function [60].

The overarching goal of our training procedure is to devise a methodology capable of learning policies that are apt for the MOMDP. The agent, once trained, should be competent to adjust to sub-optimal policies for any arbitrary preference vector $\boldsymbol{\omega}$ within the preference space Ω, with performance assessed during the testing time $\mathcal{T}$. The problem is orchestrated in two distinct phases for training the agent to adapt to new tasks with either predefined or ambiguous preferences. The first phase is the learning phase, where the agent acquaints itself with a suite of optimal policies corresponding to the MOMDP through interactions with the environment and historical trajectories. The second phase, the adaptation phase, employs a sample-efficient learning approach as proposed by [60], adjusting the policies acquired in the learning phase to a novel task with a given or undefined preference. The loss function at each step k during this phase is a composite of two components, L^A, and L^B. The first component, L^A, quantifies the disparity between the predicted value of a state-action pair and its true value, estimated by sampling transitions. The second component, L^B, facilitates the optimization of the loss function, defined as the absolute difference between the predicted and true values of a state-action pair under a given preference. The final loss function is a balanced sum of these components, computed as:

$$(1-\lambda)L^A + \lambda L^B \tag{19}$$

B. Multi-agent reinforcement learning

In this work, we aim to mitigate the effects of noise in deep Q-networks for autonomous driving scenarios involving multiple agents. We propose a multi-objective reinforcement learning (MORL) model to address this multi-objective optimization problem because traditional approaches, such as meta-heuristic approaches, are not well-suited in a dynamic environment such as autonomous driving [5]. First, while MORL requires significant computational resources during the training phase, a well-trained policy can provide a solution instantly since all training processes can be done offline, and the inference time for a deep neural network can be negligible. In contrast, meta-heuristic approaches need to spend a considerable amount of time generating a solution for the current environment since they do not separate the training and inference stages. Second, MORL can provide an adaptive solution to different environments by learning state-action mappings through interaction with the dynamic environment. Conversely, meta-heuristic approaches rely heavily on the current situation and are difficult to adapt to new environments. Additionally, with the increasing development of parallel programming hardware, such as graphical computational units, inferring MORL policies with deep neural networks has become more efficient in recent years. Meanwhile, parallelizing meta-heuristic optimization algorithms is challenging [19].

To solve this problem using MORL, we formulate it as a Markov game in a multi-agent environment [33]. We introduce a variable-length margin to maximize the combined reward of traffic throughput and group safety. All the vehicles involved in the experiment act as agents and take actions (i.e., enlarge margins) to increase the reward.

At each time step, the agents observe the state of the environment and gain experience from each other. The overall reward for the system is taken as the sum of the rewards for traffic throughput and group safety. To implement the MORL approach, we define the state space and action space at each time step t as follows:

State space: The learning agents learn to get experience and update decision policy through observing states of the environment. Based on the optimization objectives, the agents make their decisions using states in the environment. The state space is denoted as $s_t = \{v_i^t, h_i^t, v_{i+1}^t, h_{i-1}^t, v_{i-1}^t\}$, where v_i^t is the velocity of the agent i, and v_{i+1}^t and v_{i-1}^t are the velocities of the front and rear agents of agent i, respectively. Furthermore, h_i^t and h_{i-1}^t represent the headway of agent i and the agent behind i.

Action space: The action space includes the interval of decision (margins) $\mathcal{A} = [\delta_{min}, \delta_{max}]$, where δ_{min} and δ_{max} represent the lower bound and upper bound for margin, respectively.

C. Reward

Our aim is to measure the unique contribution of each agent in a cooperative multi-agent multi-objective system to each system objective. To achieve this, we adapted the difference reward method proposed by Mannion et al. [35]. For each agent, i, the method calculates the difference between the value of objective i evaluated by the entire system and the value of objective i evaluated by the system without agent i.

In our problem, we focus on two primary objectives, namely traffic throughput and group safety. To incentivize agents to contribute towards achieving these objectives, we have defined specific reward structures for each agent.

The reward function of traffic throughput is expressed as:

$$\Phi_i = \Phi - \Phi_{-i}, \tag{20}$$

where Φ_i is the traffic throughput reward of agent i, Φ is traffic throughput of the whole system, and Φ_{-i} denotes the traffic throughput without agent i.

Correspondingly, the reward function of group safety is expressed as:

$$\Psi_i = \Psi - \Psi_{-i}, \tag{21}$$

where Ψ_i is the safety reward of agent i, Ψ is group safety of the whole system, and Ψ_{-i} denotes the group safety without agent i.

To fairly evaluate the contribution of each agent to each objective under different preferences, we need to make use of normalization to uniform the expected returns' scale of different objectives.

Thus, for traffic throughput reward of agent i,

$$r_\Phi = \frac{\Phi_i - \Phi_{imin}}{\Phi_{imax} - \Phi_{imin}} \tag{22}$$

where r_Φ is the normalized traffic throughput reward of agent i, Φ_i is the traffic throughput reward of agent i, Φ_{imax} and Φ_{imin} are the utopia and nadir values for traffic throughout reward of agent i.

For group safety reward of agent i,

$$r_\Psi = \frac{\Psi_i - \Psi_{imin}}{\Psi_{imax} - \Psi_{imin}} \tag{23}$$

where r_Ψ is the normalized group safety reward of agent i, Ψ_i is the group safety reward of agent i, Ψ_{imax} and Ψ_{imax} are the utopia and nadir values for group safety reward of agent i.

Then, the overall reward of the system is expressed as:

$$\boldsymbol{r}(s_t, a_t) = [r_\Phi, r_\Psi] \tag{24}$$

$$f_{\boldsymbol{w}}(\boldsymbol{r}(s_t, a_t)) = \boldsymbol{\omega}^{\boldsymbol{T}} \boldsymbol{r}(s_t, a_t) \tag{25}$$

where $\boldsymbol{w}$ is preference.

D. Algorithm

In this section, we adapted an enveloped multi-objective reinforcement learning architecture [60] to optimize for the simultaneous maximization of traffic and safety. The proposed algorithm is outlined in Algorithm 1.

VI. ComDrive

In this section, we present ComDrive, a new simulation platform for autonomous driving simulation and control that is built upon the reliable driving simulation platform Flow [57]. ComDrive introduces several key innovations and contributions that set it apart from other similar systems:

- Enhanced perception module with support for simulating various sensor types and incorporating sensor noise, allowing for a more realistic representation of real-world driving scenarios.
- A communication module that enables efficient information sharing among agents, providing each agent with a comprehensive understanding of the traffic environment and facilitating more informed decision-making.
- Integration of multiple state-of-the-art controllers that can leverage the proposed margin to handle sensing errors, enhancing the performance of autonomous vehicles in uncertain environments.
- A flexible deep reinforcement learning module that supports various algorithms and can be tailored to the specific requirements and characteristics of different tasks, promoting more effective learning of margins.
- An extensive evaluation module that supports various performance metrics for comprehensive comparison and validation of the proposed system against traditional control algorithms and other deep reinforcement learning approaches.

The ComDrive system architecture, as shown in Fig. 4, comprises five primary modules: perception, communication, evaluation, deep reinforcement learning approaches, and controller. In the following sections, we will introduce the details

Algorithm 1 Traffic-Safety Maximization Optimization via Enveloped Multi-objective Q-learning

Input: the preference $\boldsymbol{\omega}$, the balance weight λ increasing from 0 to 1, an N-dimensional fused noises distribution $\mathcal{E} = \{\varepsilon_1, \varepsilon_2, ..., \varepsilon_N\}$ and controller $\mathcal{C} = \{c_1, c_2, ..., c_N\}$ for each vehicle.

Initialize replay buffer $\mathcal{D}_\tau$, network $\mathbf{Q}_\theta$, and λ = 0
for episode = 1,..., m **do**
  for $t = 0, ..., n$ **do**
    for $i = 0, ...N$ **do**
      Observe state $s_t = \{v_i^t, h_i^t + \varepsilon_i^t, v_{i+1}^t, h_{i-1}^t + \varepsilon_{i-1}^t, v_{i-1}^t\}$
      Sample an margin ϵ-greedy:
      $\delta_t = \begin{cases} \text{random margin in } \mathcal{A} & \text{w.p. } \epsilon \\ max_{\delta \in \mathcal{A}} \boldsymbol{\omega}^T Q(s_t, \delta; \theta) & \text{w.p. } 1-\epsilon \end{cases}$
      Receive a vectorized reward $\boldsymbol{r}_t = [r_\Phi, r_\Psi]$
      Observe $s_{t+1} = \{v_i^{t+1}, \mathcal{C}(s_t), h_i^{t+1} + \varepsilon_i^{t+1}, v_{i+1}^{t+1}, h_{i-1}^{t+1} + \varepsilon_{i-1}^{t+1}, v_{i-1}^{t+1}\}$
      Store transition $(s_t, \delta_t, \boldsymbol{r}_t, s_{t+1})$ in $\mathcal{D}_\tau$
      if update **then**
        Sample N_τ transitions $(s_j, \delta_j, \boldsymbol{r}_j, s_{j+1})$ in $\mathcal{D}_\tau$
        Compute $\mathbf{Q}_j$=
        $\begin{cases} \boldsymbol{r}_j & \text{for } s_{j+1} \\ \boldsymbol{r}_j + \gamma arg \max_{\delta \in \mathcal{A}} \boldsymbol{w}^T \mathbf{Q}(S_{j+1}, \delta; \theta) & 1 \leq j \leq N_\tau \end{cases}$
        update Q_θ
        increase λ along the path p_λ
      end if
    end for
  end for
end for

of each module and highlight the innovations they bring to the field of autonomous driving simulation and control.

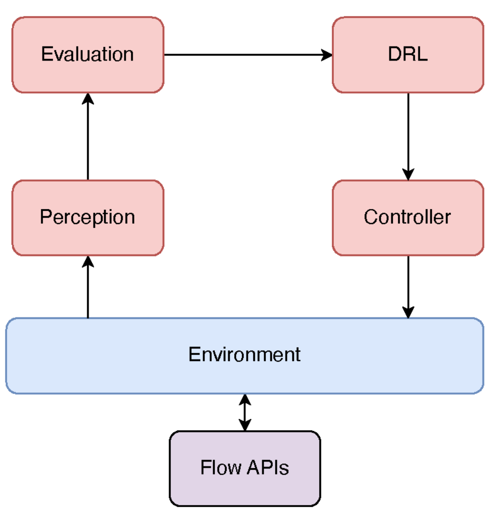


Fig. 4. ComDrive System Architecture

A. Perception

The Perception module plays a crucial role in simulating an array of sensors, such as GPS, radar, and cameras, thereby providing the controllers with the necessary environmental state information. ComDrive offers a highly customizable interface that enables users to incorporate sensors of various types while specifying the sensor name and the data detected by the sensor.

To simulate real-world driving scenarios more accurately, ComDrive allows users to introduce noise into sensor data. This is achieved by defining the noise type when adding a sensor, utilizing a three-dimensional tuple in the format of detect_type, sensor_name, noise_function. Additionally, default values for each sensor noise are established based on cutting-edge research literature [7], [11], [16], [22], [39], [47], [62]. This enhanced perception module, which accommodates various sensor types and incorporates sensor noise, contributes to the realistic representation of autonomous driving simulations.

B. Communication

The Communication module serves as an essential component in facilitating information exchange among agents within the autonomous driving system. This system enables the sharing of vital data, such as velocity, vehicle trajectories, and driving intentions, equipping each agent with a more comprehensive understanding of the traffic environment. Consequently, agents can make more informed decisions regarding their safe and efficient navigation through the environment. The communication process is currently implemented via a centralized server.

C. Controller

This module is responsible for generating control actions based on the state information from the perception module and the margins from the deep reinforcement learning module. ComDrive supports a variety of controllers, including the Intelligent Driver Model (IDM) [54], the Bando-Follow the Leader (Bando-FTL) model [24], and Linear Adaptive Cruise (LAC) [46]. As we continue to evolve ComDrive, we are actively exploring the inclusion of additional autonomous driving models, such as Rajamani model [42], to further enrich our platform's capabilities. These controllers use different strategies to control vehicle acceleration, deceleration, and steering, but they can all use our proposed margin to handle sensing errors. Controllers receive information from the simulation platform, including position, velocity, headway, and other relevant data. Our proposed margin can be applied to any of these controllers to improve their performance in the presence of sensing errors.

D. Deep Reinforcement Learning

The deep reinforcement learning module is responsible for training the margins using a multi-agent multi-objective reinforcement learning algorithm. ComDrive supports multiple deep reinforcement learning algorithms, including DQN, DDPG, and PPO. Users can select the appropriate algorithm based on their task's specific requirements and characteristics. The deep reinforcement learning module receives feedback from the simulation platform, including the environment state and the reward, using this information to train the margin models. Once training is complete, the learned margins can be applied to the controllers, improving their performance in the presence of sensing errors.

E. Evaluation

This module is responsible for evaluating the performance of the ComDrive system. ComDrive supports various evaluation metrics, such as traffic throughput, group safety metrics (e.g., TTC [30]), and individual safety metrics (e.g., headway variance, average velocity). The evaluation module can be used to compare the performance of different configurations (e.g., different sensor types and different controller types) and to validate the effectiveness of the proposed margin model. It can also be used to compare the performance of ComDrive with other approaches, such as traditional control algorithms or other deep reinforcement learning algorithms.

VII. Experimental Studies

The present section is delineated into four primary facets of inquiry. The initial focus is on contrasting the operational efficiency of automobiles that are equipped with and devoid of sensing discrepancies, thereby providing insights into the implications of such errors. Subsequently, the attention is shifted towards a comparative analysis of the proposed framework's performance, both with and without a sensing error-conscious margin. This analysis is conducted in the context of vehicle safety and traffic throughput within the Multi-Objective Reinforcement Learning (MORL) framework. The third element of our study seeks to authenticate the assertion that the fairness safety metric is capable of offering fairness assurances for each vehicle. Concluding our multi-faceted exploration, we assess the efficacy of Deep Reinforcement Learning for the training of margins, comparing it directly with the conventional leading-edge algorithm, the Genetic Algorithm. The overarching aim of these inquiries is to shed light on the potential for improvements in vehicle safety and traffic management.

A. Experiment Setting

Evaluation: Our evaluation focuses on achieving a dual objective: maximizing traffic throughput while ensuring group safety in a specific road network with sensing errors. To assess our model's performance, we employ the metrics of group safety and traffic throughput, quantified using Eq.9 and Eq.7, respectively. Additionally, we delve into individual vehicle performance by considering metrics such as average velocity and average Time-to-Collision (TTC). To categorize encounters as safe or critical [37], we set a threshold of 30 seconds as the upper bound for TTC, indicating absolute safety. To validate the fairness guarantee provided by our proposed fairness metric, we examine the headway distribution. A more concentrated headway distribution and a larger minimum headway value indicate greater safety for each vehicle. Thus, we analyze headway metrics to assess the fairness of individual vehicle safety.

Controller: The Bando-Follow the Leader (Bando-FTL) model, as documented by [24], serves as a foundational framework for simulating Adaptive Cruise Control (ACC) car-following behavior within traffic flow analyses. This model synthesizes the insights of the Bando model [2] and the Follow the Leader model [18], enabling it to dynamically adjust vehicle acceleration or deceleration according to the vehicle's current speed, the optimal speed, and the speed of the preceding vehicle. By accurately modeling and simulating how vehicles interact and follow each other in a traffic stream, the Bando-FTL model contributes to the wider field of car-following models. These models are crucial in traffic engineering and transportation research for their ability to elucidate and forecast traffic behaviors. Consequently, the Bando-FTL controller has been selected for application in our experimental investigations.

Noise: In our experiments, we introduce noises into the measurement of headway for all controllers. These noises are typically modeled as Gaussian random variables in the context of in-vehicle sensing. For example, Li *et al.* [32] modeled headway sensing errors from radar as a Gaussian distribution with a mean of 0.6 and a standard deviation of 0.72. Parker*et al.* [39] modeled headway errors from GPS as a Gaussian distribution with a mean of 0 and a standard deviation of 6. Similarly, camera sensors can generate sensing errors that follow a distribution of $\mathcal{N}(0, 5.587)$ [49]. Based on previous research, we assume these sensing errors follow a Gaussian distribution.

B. The Impact of Sensing Errors

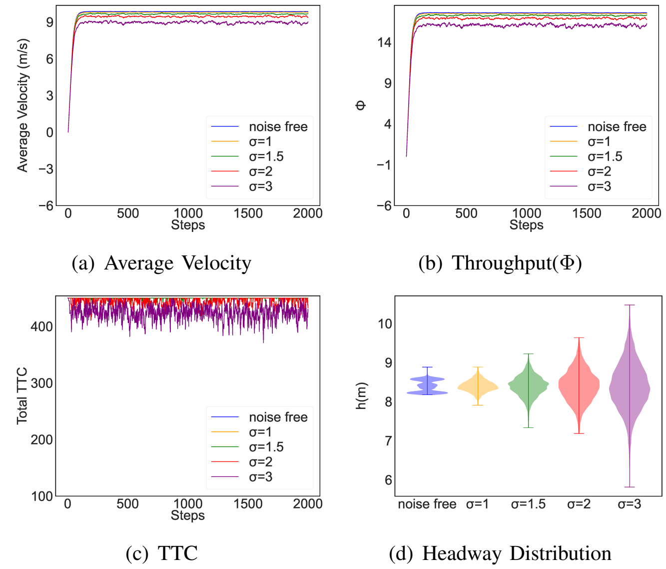
Fig. 5-7 show the impact of sensor errors following different distributions (Gaussian, uniform, and Laplacian) on traffic throughput and safety. In this experiment, 15 vehicles are placed uniformly with equal spacing on a 200m ring road.

Gaussian Distribution: As shown in Fig. 5, as the standard deviation σ increases, the noise has a greater impact on the results, decreasing traffic throughput and safety, both the group and individual safety of vehicles.

Uniform Distribution: Fig. 6 demonstrates that the performance of vehicles in terms of traffic throughput and safety is sensitive to the range of possible values for uniform distribution.

Laplacian Distribution: As depicted in Figure 7, as the scale diversity λ increases for the Laplacian distribution, both traffic throughput and safety decrease.

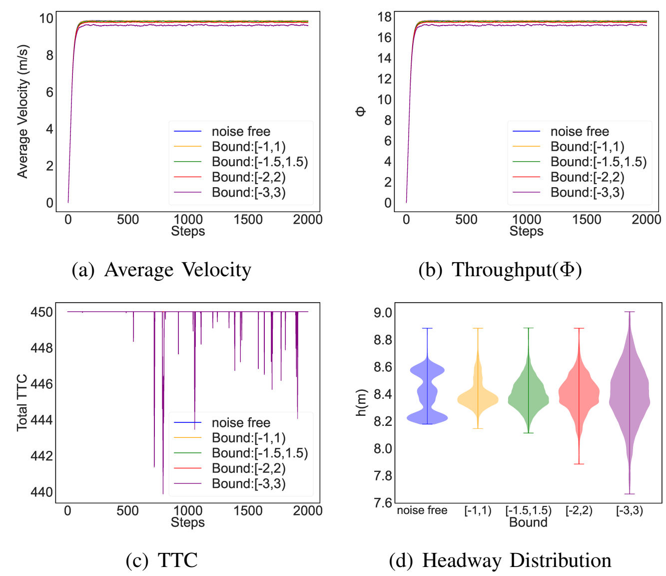
Despite the fact that noise with small parameters (σ for Gaussian distribution, the range of possible values for uniform distribution, and λ for Laplacian distribution) may not significantly impact the overall safety of the vehicles in a group, the minimum value of sensing headway becomes smaller, which increases the risk of collisions and compromises the safety of individual vehicles. It is clear that the presence of sensing errors, regardless of the distribution they follow, leads to a decrease in both traffic throughput and safety for both the group and individual vehicles.



(a) Average Velocity (b) Throughput(Φ)

(c) TTC (d) Headway Distribution

Fig. 5. The impact of Gaussian distribution noise on different metrics(Gaussian noise μ = 0)



(a) Average Velocity (b) Throughput(Φ)

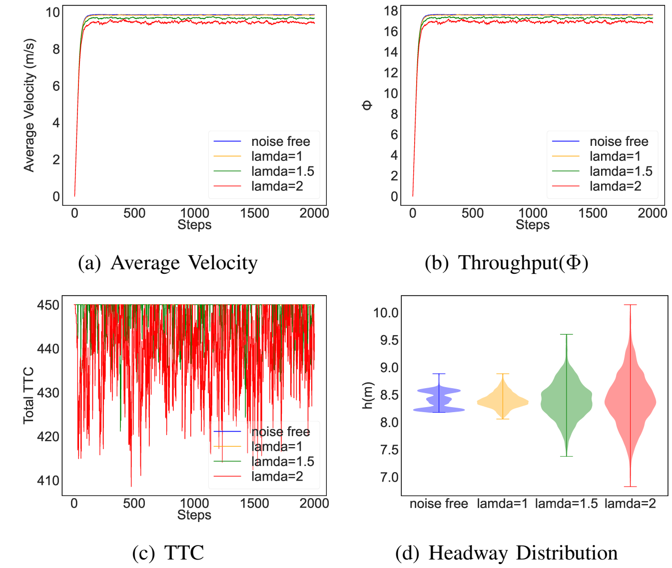
(c) TTC (d) Headway Distribution

Fig. 6. The impact of uniform distribution noise on different metrics

C. The Effectiveness of Headway Margin

This section is divided into two parts to evaluate the effectiveness of our proposed margin on traffic throughput and group safety: training and results.

Training: To validate the effectiveness of our proposed margin on traffic throughput and group safety, we conducted an experiment where we trained the margin using DQN on a 250m ring network with 20 vehicles. We used traffic throughput Φ and $\alpha - fair$ group safety $\Psi(fair)$ as the metrics for training. Since GPS is more widely used than radar and lidar, we consider the sensors to have Gaussian distribution error with a standard deviation of 2.45 in terms of headway, which is consistent with the real GPS distance error reported in [39]. The trained neural network was then used to add margins to sensor data automatically.



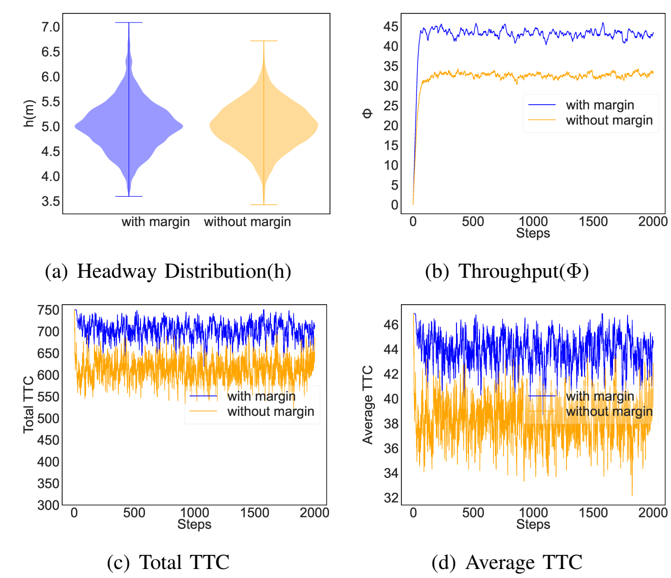
(a) Average Velocity (b) Throughput(Φ)

(c) TTC (d) Headway Distribution

Fig. 7. The impact of Laplace distribution noise on different metrics

Results: To test the generalizability of our proposed margin, we conduct experiments in both pure and mixed road conditions. In pure road condition, all vehicles are equipped with the margin. In mixed road conditions, only half of the vehicles are equipped with the margin, while the other half do not have it.

1) On pure road: To further evaluate the effectiveness of our proposed margin under different levels of noise and the influence of the number of vehicles on the road, we conduct experiments with 25 vehicles on a 250m crowded ring road with two different levels of Gaussian noise in headway (σ=2.45 and σ=1, $\mu = 0$). The experiments are implemented 10 times to obtain reliable results.

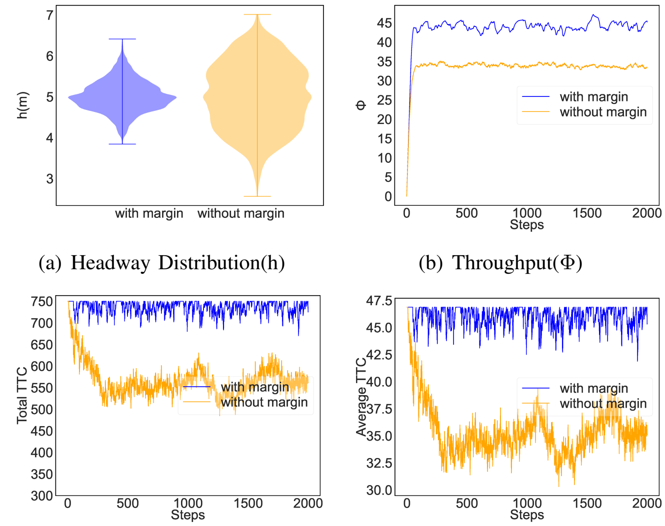


(a) Headway Distribution(h) (b) Throughput(Φ)

(c) Total TTC (d) Average TTC

Fig. 8. Comparison results for controllers with and without margin (σ=2.45)

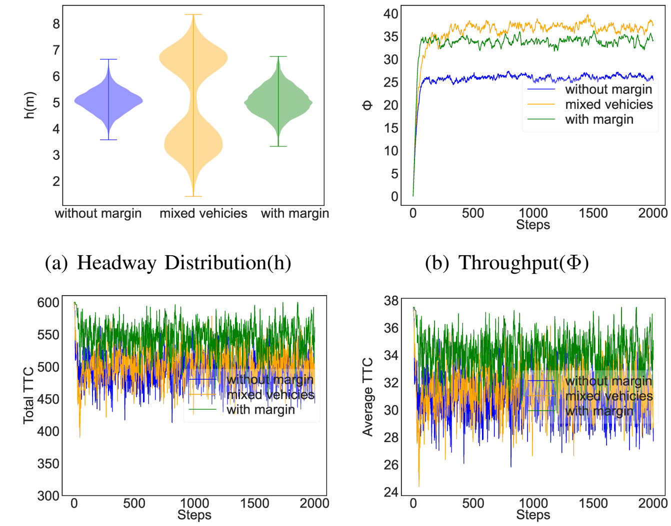
As shown in Fig. 8 and Fig. 9, adding a margin to the sensing headway can improve the performance of autonomous



(a) Headway Distribution(h) (b) Throughput(Φ)

(c) Total TTC (d) Average TTC

Fig. 9. Comparison results for controllers with and without margin (σ=1)



(a) Headway Distribution(h) (b) Throughput(Φ)

(c) Total TTC (d) Average TTC

Fig. 10. Comparison results for mixed road (σ=2.45)

vehicles in terms of traffic throughput and group safety. This is because the margin can alleviate the bias of sensor data, which can cause unnecessary emergency brakes or collisions when the sensing value is lower or larger than the real value, respectively. By adding a positive margin in a proper way, the number of unnecessary emergency brakes can be reduced, and by adding a negative margin in some cases, collisions can be avoided. Our model is able to find the optimal way of adding margins to mitigate data bias and improve vehicle performance.

2) On mixed road: In this section, we investigate the performance of our proposed margin on a mixed road, where half of the vehicles have margins and the other half do not. We set the standard deviation of the Gaussian distribution of sensing errors to 2.45, the number of vehicles to 20, and the road to 250m. The vehicles are randomly placed on the road and the experiments are repeated 10 times to obtain the results.

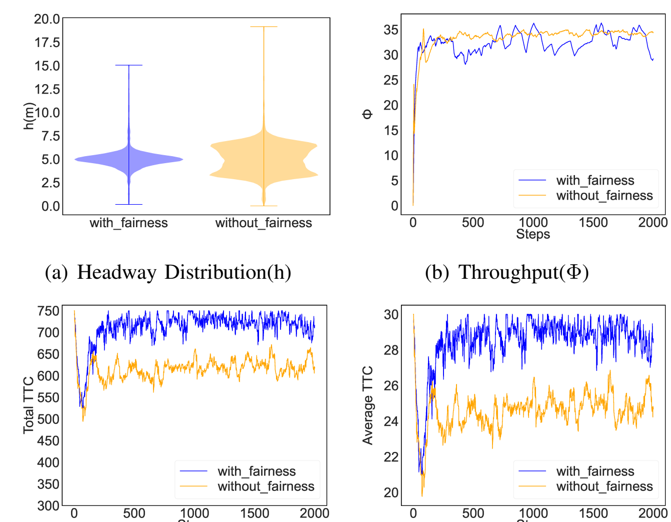
Fig. 10 shows that when our proposed margin model is applied to the mixed road, the performance of the vehicles is better than when no margins are used at all but worse than when margins are applied to all vehicles.

D. The Effectiveness of Fairness Safety Metric

In this section, we compare our framework with and without considering fairness, to validate whether the fairness metric can provide a fairness guarantee for each vehicle. In order to better evaluate the effectiveness of the fairness safety metric, we present the experiment in two parts: training and results.

Training: In the training process, we developed two models for generating sensor-aware margins. The only difference between the two models is the group safety metric that is used. One model uses the TTC-only safety metric $\Psi(TTC)$ while the other model uses the α-fair group safety metric $\Psi(fair)$.

Results: In this experiment, we compare the performance of margin models trained by $\alpha - fair$ group safety with those trained by TTC-only safety. The experiment is conducted on a 250m ring road with 25 vehicles randomly placed. In the first experiment, all vehicles use margin models trained by $\alpha - fair$ group safety, while in the contrast experiment, only vehicles with margin models trained by TTC-only safety are included.



(a) Headway Distribution(h) (b) Throughput(Φ)

(c) Total TTC (d) Average TTC

Fig. 11. Comparison results for controllers with and without fairness (σ=2.45)

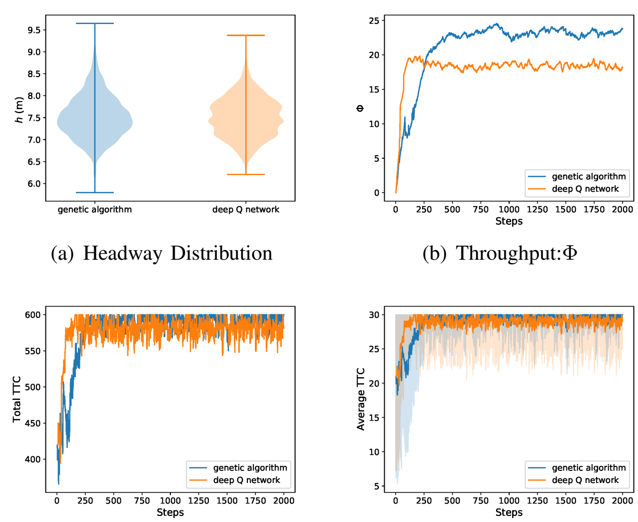
Figure 11 compares the results of our framework with and without fairness considerations. As shown in Figure 11(a), the margin trained with fairness increases the minimum value of sensing headway, ensuring individual safety. However, this may come at the cost of reduced traffic throughput (Φ), as shown in Figure 11(b).

E. Genetic Algorithm and Deep Q-network

In this section, we compare the performance of deep reinforcement learning (DRL) with genetic algorithm in terms of run time, traffic throughput, and TTC-only safety. To do this, we simulate a scenario with 20 vehicles on a 250m ring road with Gaussian headway noise (σ=2.45). We set the lower bound of the margin to -5 and the upper bound of the margin to 5. The fitness function of the genetic algorithm is based on the sum of traffic throughput and TTC. We iterate through different sets of margins, adding them to the sensor data and evaluating their fitness values. The margins with the highest fitness value are chosen as the potential best margins for each iteration.

TABLE I
COMPARISON RESULTS OF TWO ALGORITHMS

algorithm	run time(s)	throughput(1/s)	TTC(s)
genetic algorithm	48098	**21.4**	**580**
DRL	**84**	18.1	576



(a) Headway Distribution

(b) Throughput:Φ

(c) Total TTC

(d) Average TTC

Fig. 12. Results for genetic algorithm and deep Q network methods

According to the results presented in Table I, it can be seen that the genetic algorithm performs better than DRL in terms of traffic throughput and TTC-only safety. However, the run time of the genetic algorithm is significantly longer, making it unacceptable for use in real-time driving situations. Additionally, the minimum value of sensing headway is much lower with the genetic algorithm, which poses a risk to individual safety. Based on these observations, it can be concluded that deep reinforcement learning is a more effective method for training the proposed margin model.

VIII. CONCLUSION

This study has introduced FairAV, a robust multi-objective optimization framework for autonomous vehicles, which considers the challenges posed by sensing errors, the need for traffic throughput, and the imperative of fair safety distribution among vehicles. A novel α-fair group safety metric, the first of its kind, has been proposed, ensuring a minimum safety threshold for individual vehicles, thereby addressing the ethical concerns of any vehicle compromising its safety for the larger group. By employing multi-objective reinforcement learning, FairAV has been shown to strike a real-time balance between traffic throughput and group safety, while concurrently reducing the detrimental effects of sensing errors on detected headway. The efficacy of the approach has been substantiated through comprehensive experiments conducted on the ComDrive autonomous driving simulation platform, demonstrating significant improvements over existing methodologies. This research underscores the critical need for integrating fairness considerations into autonomous driving system design and lays the groundwork for future studies aimed at improving safety without compromising efficiency. Future efforts will explore the application of this approach to more complex traffic scenarios and heterogeneous vehicle types, and investigate the potential of more sophisticated reinforcement learning algorithms and robust optimization methods to further enhance the performance of autonomous driving systems.

ACKNOWLEDGMENT

This work is supported in part by NSF Grants CNS-2203239, CNS-2203412, and CCSS- 2203238.

REFERENCES

[1] Olga Andreeva, Matthew Almeida, Wei Ding, Scott E Crouter, and Ping Chen. Maximizing fairness in deep neural networks via mode connectivity. *IEEE Intelligent Systems*, 2022.

[2] Masako Bando, Katsuya Hasebe, Akihiro Nakayama, Akihiro Shibata, and Yuki Sugiyama. Dynamical model of traffic congestion and numerical simulation. *Physical review E*, 51(2):1035, 1995.

[3] Khaled Belahcene, Vincent Mousseau, and Anaëlle Wilczynski. Combining fairness and optimality when selecting and allocating projects. In *IJCAI*, pages 38–44, 2021.

[4] Jean-François Bonnefon, David Černy, John Danaher, Nathalie Devillier, Veronica Johansson, Tatiana Kovacikova, Marieke Martens, Milos Mladenovic, Paula Palade, Nick Reed, et al. Ethics of connected and automated vehicles: Recommendations on road safety, privacy, fairness, explainability and responsibility. *Ethics and Information Technology*, 2020.

[5] Lorenzo Canese, Gian Carlo Cardarilli, Luca Di Nunzio, Rocco Fazzolari, Daniele Giardino, Marco Re, and Sergio Spanò. Multi-agent reinforcement learning: A review of challenges and applications. *Applied Sciences*, 11(11):4948, 2021.

[6] Yair Censor. Pareto optimality in multiobjective problems. *Applied Mathematics and Optimization*, 4(1):41–59, 1977.

[7] Woo Young Choi, Jin Ho Yang, and Chung Choo Chung. Data-driven object vehicle estimation by radar accuracy modeling with weighted interpolation. *Sensors*, 21(7):2317, 2021.

[8] Carlos A Coello Coello, Gary B Lamont, David A Van Veldhuizen, et al. *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer, 2007.

[9] Guangzhen Cui, Weili Zhang, Yanqiu Xiao, Lei Yao, and Zhanpeng Fang. Cooperative perception technology of autonomous driving in the internet of vehicles environment: A review. *Sensors*, 22(15):5535, 2022.

[10] Gabriel Rodrigues De Campos, Paolo Falcone, and Jonas Sjöberg. Autonomous cooperative driving: A velocity-based negotiation approach for intersection crossing. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pages 1456–1461. IEEE, 2013.

[11] Willard de Jong. Sensor and v2v delay compensation in cooperative adaptive cruise control using an extended kalman filter. 2020.

[12] Celso M de Melo, Stacy Marsella, and Jonathan Gratch. Social decisions and fairness change when people's interests are represented by autonomous agents. *Autonomous Agents and Multi-Agent Systems*, 32(1):163–187, 2018.

[13] Renio Díaz and Alexandro R Suarez. A study of the capacity of the stochastic hill climbing to solve multi-objective problems. In *Proceedings of the Third International Symposium on Adaptive Systems-Evolutionary Computation and Probabilistic Graphical Models, La Habana: Institute of Cybernetics, Mathematics and Physics*, pages 37–40. Citeseer, 2001.

[14] Soufiene Djahel, Nafaa Jabeur, Robert Barrett, and John Murphy. Toward v2i communication technology-based solution for reducing road traffic congestion in smart cities. In *2015 International Symposium on Networks, computers and communications (ISNCC)*, pages 1–6. IEEE, 2015.

[15] Mengnan Du, Fan Yang, Na Zou, and Xia Hu. Fairness in deep learning: A computational perspective. *IEEE Intelligent Systems*, 36(4):25–34, 2020.

[16] Jingliang Duan, Shengbo Eben Li, Yang Guan, Qi Sun, and Bo Cheng. Hierarchical reinforcement learning for self-driving decision-making without reliance on labelled driving data. *IET Intelligent Transport Systems*, 14(5):297–305, 2020.

[17] Lily Elefteriadou et al. *An introduction to traffic flow theory*. Springer, 2014.

[18] Denos C Gazis, Robert Herman, and Richard W Rothery. Nonlinear follow-the-leader models of traffic flow. *Operations research*, 9(4):545–567, 1961.

[19] Hermann Gehring and Jörg Homberger. Parallelization of a two-phase metaheuristic for routing problems with time windows. *Journal of heuristics*, 8(3):251–276, 2002.

[20] Niko A Grupen, Bart Selman, and Daniel D Lee. Cooperative multi-agent fairness and equivariant policies. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 9350–9359, 2022.

[21] Junyao Guo, Unmesh Kurup, and Mohak Shah. Is it safe to drive? an overview of factors, metrics, and datasets for driveability assessment in autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 21(8):3135–3151, 2019.

[22] Qingbei Guo, Yuan Zhang, Jaime Lloret, Burak Kantarci, and Winston KG Seah. A localization method avoiding flip ambiguities for micro-uavs with bounded distance measurement errors. *IEEE Transactions on Mobile Computing*, 18(8):1718–1730, 2018.

[23] Philipp Hacker. Personal data, exploitative contracts, and algorithmic fairness: autonomous vehicles meet the internet of things. *International data privacy law*, 7(4):266–286, 2017.

[24] Amaury Hayat, Xiaoqian Gong, Jonathan Lee, Sydney Truong, Sean McQuade, Nicolas Kardous, Alexander Keimer, Yiling You, Saleh Albeaik, Eugene Vinistky, et al. A holistic approach to the energy-efficient smoothing of traffic via autonomous vehicles. In *Intelligent Control and Smart Energy Management*, pages 285–316. Springer, 2022.

[25] Conor F Hayes, Roxana Rădulescu, Eugenio Bargiacchi, Johan Källström, Matthew Macfarlane, Mathieu Reymond, Timothy Verstraeten, Luisa M Zintgraf, Richard Dazeley, Fredrik Heintz, et al. A practical guide to multi-objective reinforcement learning and planning. *Autonomous Agents and Multi-Agent Systems*, 36(1):1–59, 2022.

[26] Jiechuan Jiang and Zongqing Lu. Learning fairness in multi-agent systems. *Advances in Neural Information Processing Systems*, 32, 2019.

[27] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

[28] B Ravi Kiran, Ibrahim Sobh, Victor Talpaert, Patrick Mannion, Ahmad A Al Sallab, Senthil Yogamani, and Patrick Pérez. Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[29] Tian Lan, David Kao, Mung Chiang, and Ashutosh Sabharwal. *An axiomatic theory of fairness in network resource allocation*. IEEE, 2010.

[30] David N Lee. A theory of visual control of braking based on information about time-to-collision. *Perception*, 5(4):437–459, 1976.

[31] Changjian Li and Krzysztof Czarnecki. Urban driving with multi-objective deep reinforcement learning. *arXiv preprint arXiv:1811.08586*, 2018.

[32] Xin Li, Weiwen Deng, Sumin Zhang, Yaxin Li, Shiping Song, Shanshan Wang, and Guanyu Wang. Research on millimeter wave radar simulation model for intelligent vehicle. *International Journal of Automotive Technology*, 21(2):275–284, 2020.

[33] Michael L Littman. Markov games as a framework for multi-agent reinforcement learning. In *Machine learning proceedings 1994*, pages 157–163. Elsevier, 1994.

[34] Shaoshan Liu, Bo Yu, Jie Tang, Yuhao Zhu, and Xue Liu. Communication challenges in infrastructure-vehicle cooperative autonomous driving: A field deployment perspective. *IEEE Wireless Communications*, 29(4):126–131, 2022.

[35] Patrick Mannion, Sam Devlin, Jim Duggan, and Enda Howley. Reward shaping for knowledge-based multi-objective multi-agent reinforcement learning. *The Knowledge Engineering Review*, 33, 2018.

[36] Radu Marinescu. Best-first vs. depth-first and/or search for multi-objective constraint optimization. In *2010 22nd IEEE international conference on tools with artificial intelligence*, volume 1, pages 439–446. IEEE, 2010.

[37] Michiel M Minderhoud and Piet HL Bovy. Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention*, 33(1):89–97, 2001.

[38] Gerardo Minella, Rubén Ruiz, and Michele Ciavotta. Restarted iterated pareto greedy algorithm for multi-objective flowshop scheduling problems. *Computers & Operations Research*, 38(11):1521–1533, 2011.

[39] Ryan Parker and Shahrokh Valaee. Vehicle localization in vehicular networks. In *IEEE Vehicular Technology Conference*, pages 1–5. IEEE, 2006.

[40] João Luiz Junho Pereira, Guilherme Antônio Oliver, Matheus Brendon Francisco, Sebastião Simões Cunha, and Guilherme Ferreira Gomes. A review of multi-objective optimization: methods and algorithms in mechanical engineering problems. *Archives of Computational Methods in Engineering*, pages 1–24, 2021.

[41] Xiangjun Qian. *Model predictive control for autonomous and cooperative driving*. PhD thesis, Université Paris sciences et lettres, 2016.

[42] Rajesh Rajamani. *Vehicle dynamics and control*. Springer Science & Business Media, 2011.

[43] Naveen Raman, Sanket Shah, and John Dickerson. Data-driven methods for balancing fairness and efficiency in ride-pooling. *arXiv preprint arXiv:2110.03524*, 2021.

[44] Mark Ryan. The future of transportation: ethical, legal, social and economic impacts of self-driving vehicles in the year 2025. *Science and engineering ethics*, 26(3):1185–1208, 2020.

[45] Dingyuan Shi, Yongxin Tong, Zimu Zhou, Bingchen Song, Weifeng Lv, and Qiang Yang. Learning to assign: Towards fair task assignment in large-scale ride hailing. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 3549–3557, 2021.

[46] Mehmet Ali Silgu, İsmet Gökşad Erdağı, Gökhan Göksu, and Hilmi Berk Celikoglu. Combined control of freeway traffic involving cooperative adaptive cruise controlled and human driven vehicles using feedback control through sumo. *IEEE Transactions on Intelligent Transportation Systems*, 23(8):11011–11025, 2021.

[47] Jan Sobotka and Jiri Novak. Digital vehicle radar sensor target simulation. In *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pages 1–5. IEEE, 2020.

[48] Ninja Soeffker, Marlin W Ulmer, and Dirk Mattfeld. On fairness aspects of customer acceptance mechanisms in dynamic vehicle routing. *Proceedings of logistikmanagement*, 2017:17–24, 2017.

[49] Teppei Suzuki and Ikuro Sato. A new way of estimating distances to foregoing vehicles from a monocular camera with error modeling. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 2208–2214. IEEE, 2019.

[50] Victor Talpaert, Ibrahim Sobh, B Ravi Kiran, Patrick Mannion, Senthil Yogamani, Ahmad El-Sallab, and Patrick Perez. Exploring applications of deep reinforcement learning for real-world autonomous driving systems. *arXiv preprint arXiv:1901.01536*, 2019.

[51] Icaro Bezerra Viana, Husain Kanchwala, Kenan Ahiska, and Nabil Aouf. A comparison of trajectory planning and control frameworks for cooperative autonomous driving. *Journal of Dynamic Systems, Measurement, and Control*, 143(7):071002, 2021.

[52] Guang Wang, Shuxin Zhong, Shuai Wang, Fei Miao, Zheng Dong, and Desheng Zhang. Data-driven fairness-aware vehicle displacement for large-scale electric taxi fleets. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, pages 1200–1211. IEEE, 2021.

[53] Nannan Wang, Xi Wang, Paparao Palacharla, and Tadashi Ikeuchi. Cooperative autonomous driving for traffic congestion avoidance through vehicle-to-vehicle communications. In *2017 IEEE Vehicular Networking Conference (VNC)*, pages 327–330. IEEE, 2017.

[54] Yilin Wang and Yiheng Feng. Idm-follower: A model-informed deep learning method for long-sequence car-following trajectory prediction. *arXiv preprint arXiv:2210.10965*, 2022.

[55] Lianzhen Wei, Zirui Li, Jianwei Gong, Cheng Gong, and Jiachen Li. Autonomous driving strategies at intersections: Scenarios, state-of-the-art, and future outlooks. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 44–51. IEEE, 2021.

[56] Bowen Weng, Linda Capito, Ümit Özgüner, and Keith Redmill. A finite-sampling, operational domain specific, and provably unbiased connected and automated vehicle safety metric. *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[57] Cathy Wu, Abdul Rahman Kreidieh, Kanaad Parvate, Eugene Vinitsky, and Alexandre M Bayen. Flow: A modular learning framework for mixed autonomy traffic. *IEEE Transactions on Robotics*, 38(2):1270–1286, 2021.

[58] Runsheng Xu, Yi Guo, Xu Han, Xin Xia, Hao Xiang, and Jiaqi Ma. Opencda: an open cooperative driving automation framework integrated with co-simulation. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 1155–1162. IEEE, 2021.

[59] Runsheng Xu, Hao Xiang, Xu Han, Xin Xia, Zonglin Meng, Chia-Ju Chen, Camila Correa-Jullian, and Jiaqi Ma. The opencda open-source ecosystem for cooperative driving automation research. *IEEE Transactions on Intelligent Vehicles*, 2023.

[60] Runzhe Yang, Xingyuan Sun, and Karthik Narasimhan. A generalized algorithm for multi-objective reinforcement learning and policy adaptation. *Advances in Neural Information Processing Systems*, 32, 2019.

[61] Yonghao Yin, Dewei Li, Zhenyu Han, Xinlei Dong, and Hui Liu. Maximizing network utility while considering proportional fairness for rail transit systems: Jointly optimizing passenger allocation and vehicle schedules. *Transportation Research Part C: Emerging Technologies*, 143:103812, 2022.

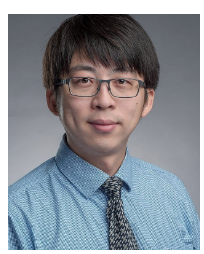
[62] Xiangmo Zhao, Shoucai Jing, Fei Hui, Ronghui Liu, and Asad J Khattak. Dsrc-based rear-end collision warning system–an error-component safety distance model and field test. *Transportation research part C: emerging technologies*, 107:92–104, 2019.

Wei Shao (Member, IEEE) is currently working as a research scientist at the CSIRO Data61, an adjunct fellow at RMIT University, and a visiting researcher at UC Davis. His research interest include AI for cybersecurity, graph neural networks, spatio-temporal data mining, autonomous driving, reinforcement learning, and the Internet of Things.

Zejun Fan is a visiting student. She has a keen interest in graph neural networks, and her current research is focused on the field of causal discovery and autonomous driving.

Chia-Ju Chen received the B.S. in Electrical and Computer Engineering from National Tsing Hua University, Taiwan in 2017 and M.S. from University of Texas at Austin, USA in 2019. She is currently pursuing the Ph.D. degree at UCLA mobility lab. Her research interests include privacy, computer vision and autonomous driving.

Zhaofeng Zhang received his B.Eng. degree in Electrical Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2015. He received his M.S. and Ph.D. degrees in Electrical Engineering from Arizona State University, Tempe, AZ, USA, in 2017 and 2023. Currently, he is a Postdoctoral Researcher at School of Computing and Augmented Intelligence at Arizona State University, Tempe, AZ, USA. His research interests include edge computing, statistical machine learning, deep learning, and optimization.



Jiaqi Ma received the Ph.D. degree in Transportation Engineering from University of Virginia, Virginia, USA, in 2014. He is currently an Associate Professor at the UCLA Samueli School of Engineering and faculty lead of the New Mobility program at UCLA Institute of Transportation Studies. His research interests include intelligent transportation systems, autonomous driving, and cooperative driving automation. He is a member of the TRB Standing Committee on Vehicle-Highway Automation, the TRB Standing Committee on Artificial Intelligence and Advanced Computing Applications, and the American Society of Civil Engineers (ASCE) Connected and Autonomous Vehicles Impacts Committee, and the Co-Chair of the IEEE ITS Society Technical Committee on Smart Mobility and Transportation 5.0. He is Editor in Chief of the IEEE Open Journal of Intelligent Transportation Systems.

Junshan Zhang (Fellow, IEEE) is a professor in the ECE Department at University of California Davis. He received his Ph.D. degree from the School of ECE at Purdue University in Aug. 2000, and was on the faculty of the School of ECEE at Arizona State University from 2000 to 2021. His research interests fall in the general field of information networks and data science, including edge intelligence, reinforcement learning, continual learning, network optimization and control, game theory, with applications in connected and automated vehicles, 5G and beyond, wireless networks, IoT data privacy/security, and smart grid.

Prof. Zhang is a Fellow of the IEEE, and a recipient of the ONR Young Investigator Award in 2005 and the NSF CAREER award in 2003. He received the IEEE Wireless Communication Technical Committee Recognition Award in 2016. His papers have won a few awards, including the Best Student Paper award at WiOPT 2018, the Kenneth C. Sevcik Outstanding Student Paper Award of ACM SIGMETRICS/IFIP Performance 2016, the Best Paper Runner-up Award of IEEE INFOCOM 2009 and IEEE INFOCOM 2014, and the Best Paper Award at IEEE ICC 2008 and ICC 2017. Building on his research findings, he co-founded Smartiply Inc, a Fog Computing startup company delivering boosted network connectivity and embedded artificial intelligence.

Prof. Zhang served as Editor-in-chief for IEEE Transactions on Wireless Communications during 2019–2022, and is a senior editor for IEEE/ACM Transactions on Networking. He was TPC co-chair for a number of major conferences in communication networks, including IEEE INFOCOM 2012 and ACM MOBIHOC 2015. He was the general chair for ACM/IEEE SEC 2017 and WiOPT 2016. He was a Distinguished Lecturer of the IEEE Communications Society.