SYSTEMATIC EVALUATION OF CONTROL STRATEGIES FOR IMPROVING TRAFFIC CONGESTION IN MIXED-AUTONOMY TRAFFIC

Arwa Alanqary
Department of Electrical Engineering and Computer Science
University of California - Berkely, Berkeley, CA, 94720
Email: arwa@berkeley.edu

Jonathan Lee
Department of Electrical Engineering and Computer Science
University of California - Berkely, Berkeley, CA, 94720
Email: jonny5@berkeley.edu

Alexandre Bayen
Department of Electrical Engineering and Computer Science
University of California - Berkely, Berkeley, CA, 94720
Email: bayen@berkeley.edu

Word Count: 4153 words + 4 table(s) $\times$ 250 = 5153 words

Submission Date: August 1, 2022

ABSTRACT

We design a testbed for evaluating and comparing the performance of automated vehicles (AVs) control strategies for reducing traffic congestion and improving the energy efficiency in mixed autonomy traffic. The system includes a microscopic traffic simulator along with specific test scenarios derived from real-world trajectory data representing different traffic conditions. We equip the system with a set of desired KPIs and evaluation metrics that serve as the basis on which we compare and rank the controllers. We present a use case of the system by implementing and comparing a few control strategies ranging from classical to deep reinforcement learning (DRL) based controllers. Our comparison results show that the controllers can achieve an increase of up to 19.15% in the average speed and an improvement of up to 25.39% in the fuel efficiency of the platoon, compared to the baseline of fully human-driven traffic. Through these experiments we illustrate the capabilities of the system in allowing for detailed and fair comparison between control strategies of different nature on a wide range of scenarios with flexible configurations. Finally, the system architecture allows for easy integration of additional control strategies in the future. This aims to promote continues development in the field and to allow input from the broader research community.

Keywords: traffic simulation, mixed-autonomy traffic, automated vehicles, traffic control.

INTRODUCTION

Mixed-autonomy traffic, is a system in which only a fraction of the vehicles are automated (AVs) and they interact with human driven vehicles in the road. The introduction of AVs in such predominantly human drivers traffic has the potential to improve the traffic conditions at both the microscopic and the system level.

Several recent studies have demonstrated, through simulations or field experiments, such positive impact of introducing one or a few AVs on the overall traffic flow. Different metrics have been considered in such studies including improved travel time (*13*), increasing the throughput, improving the system stability (*11*), and reducing fuel consumption (*8*).

There is still a strong emphasis within the intelligent transportation research community to further harness such benefits. This results in a growing number of developed controller that are based on various control strategies and designs to serve the common purpose of improving traffic conditions. The literature is rich in such controllers ranging from classical hand-crafted micro-controllers to modern deep reinforcement learning policies. In this work we highlight the importance of establishing a standard testbeds for common evaluation and comparisons of control strategies of AVs in mixed-autonomy traffic. Systematic evaluation and comparison will not only further understanding of the strengths of existing algorithms, but also reveal their limitations and suggest directions for future research. Further, such benchmarks can accelerate the development of the field by enabling researchers to focus on controller and algorithmic design rather than experimental design for testing and evaluation.

The benefit of standard benchmarks in control and sequential-decision-making research has been realized by multiple research fields. Perhaps the most prominent example comes from the field of reinforcement learning (RL), in which standard benchmarks has accelerated the development of the field and created common grounds for evaluating progress in algorithm design. Examples of RL benchmarks include the Arcade Learning Environment (ALE) (*2*), the benchmarks released with rllab (*4*), and the Multi-Joint dynamics with Contact (MuJoCo) physics engine (*9*).

In contrast, when it comes to the field of mixed autonomy traffic, there's a sever lack of domain specific standardized benchmarks and testbeds that are designed based on realistic traffic scenarios. This makes it difficult to evaluate the practicality of the proposed methods in the literature, compare their performance in a systematic and fair manner, and reveal their limitations. A few studies have attempted to bridge this gap and propose systems and benchmarks for comparison and evaluation in the domain of mixed autonomy control. Ault and Sharon (*1*) developed a toolkit for comparing RL-based controllers for signal control. The work of Vinitsky et. al. (*11*) devised a set of standard benchmarks for RL-based longitudinal controller on problems shockwave minimization, inflow management, efficient merging, and intersection control.

In this work we propose and develop a testbed for the task of smoothing traffic flow through longitudinal control of a AVs in mixed-autonomy traffic driving along a single high-way lane. The testbed consists of a microscopic simulation engine for mixed-autonomy traffic, a set of standard benchmarks inspired by real-world scenarios, and a set of evaluation metrics for comparing and ranking controllers. To the best of our knowledge, we are unaware of a standard set of benchmarks and evaluation tools that are compatible with both RL-based and classical controllers for such task in the domain of mixed-autonomy traffic. By introducing this system, we aim to bridge the evaluation gap and provide a standard way for comparing and evaluating controllers in the literature. Our key contributions in this article:

- Define the task of reducing traffic congestion in mixed-autonomy traffic using AVs.

- Introduce a set of test scenarios for this task based based on real-world trajectory data.
- Develop an architecture for testing and evaluating control algorithms on these scenarios.
- Present a use-case to illustrate the use of the system using multiple control strategies ranging from classical to deep RL controllers.

SETUP

In this section we describe the details of the the mixed-autonomy traffic system and the task of reducing traffic congestion in such system. We then introduce the microscopic simulation componenet of the evaluation testbed.

System Description

The system we consider in this work is a mixed-autonomy platoon consisting of three types of vehicles: (1) a leader vehicle which has a fixed trajectory and takes the leading position in the platoon, (2) autonomous vehicles which are equipped with longitudinal (acceleration based) controllers which we aim to compare, and (3) human driven vehicles whose longitudinal behaviour is governed by a human-driver model and their lateral behaviour is governed by a probabilistic lane changing distribution.

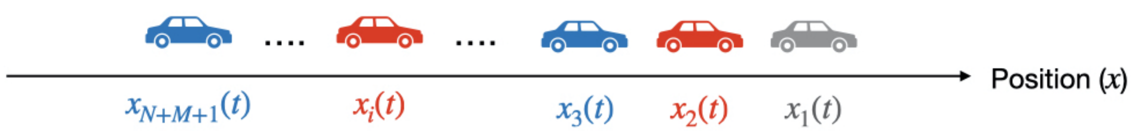


FIGURE 1: Example of a mixed-autonomy platoon consisting of a leader vehicle (gray), automated vehicles (red), and human driven vehicles (blue)

The leader vehicle's trajectory represents the downstream traffic conditions. Congested traffic presents itself in the form of stop-and-go waves in the leader's trajectory. The task is to design a control strategy for all AVs in the system to dampen the backwards propagation of the stop-and-go waves and reduce their overall effect on the platoon. In the next chapter we will introduce specific metrics used to evaluate the performance of the controllers for this congestion reduction task.

Notation

As a common notation, we number that cars in the platoon from front to back with the leader having the index 1. In all scenarios, we have a fixed simulation horizon $T \in R_+$. We denote the time-variable number of vehicles in the platoon $n(k)$ as the total number of vehicles (leader, human-driven, and automated vehicles) at time step k. We note that the number of vehicles can change due to the lane switching disturbance that we will introduce later in the paper. Further, we by $I_a(k)$ the set of indices of the AVs in the platoon at time step k.

At a given time step k, and given a platoon of 1 leader vehicle, N AVs and M human-driven vehicles, we have $n(k) = 1+M+N$ and we denote the state vector $\mathbf{s}(k) = [x_i(k), v_i(k)]_{i=1:n(k)} \in$

$R \times R_+$, where $x_i(k)$ and $v_i(k)$ represent the position and velocity, respectively, of the i^{th} vehicle in the platoon. The position is taken with reference to the initial position of the leader vehicle, and the velocity is restricted to be always positive (i.e. cars can't drive backwards). The action vector is a list of accelerations for each AV in the system and is denoted by $a \in [a_{min}, a_{max}]^N$ where a_{min} and a_{max} represent the minimum and maximum allowable accelerations.

Traffic Simulation

In this work we simulate the performance of mixed-autonomy traffic control policies through the use of microscopic traffic simulators, in which the states and controls of the traffic network are modeled at the level of individual vehicles. Traffic microsimulators have been broadly accepted in the transportation engineering community as a tool to evaluate automation tasks such as adaptive cruise control (ACC) (*6*) and mixed-autonomy traffic flow control (*8*). In the context of microscopic simulation of mixed-autonomy traffic, human-driver models are used to recreate the lateral and longitudinal behavior of human-driven vehicles. In our simulation, the longitudinal behaviour of human-driven cars is governed by the intelligent driver model (IDM) (*10*), a state-of-the-art human-driver model in microscopic traffic simulation.

Since our environment consists of a single lane of a highway, we have a simplified model for lateral human behaviour. It is modeled as a stochastic disturbance in the form of vehicles cutting-in (appearing) into the simulation lane and joining the platoon, or vehicles in the platoon cutting-out (disappearing) from the lane. At every simulation step and for each car in the platoon, this disturbance is modeled as a Bernoulli random variable with parameter taken as a function of the headway and the velocity of this vehicle. In the test scenarios this disturbance can be disabled resulting in a closed system with a fixed number of cars in the platoon.

EVALUATION TESTBED

In this section, we detail the proposed framework for testing and evaluation control strategies for the task of reducing traffic congestion. We start by introducing the test scenarios, followed by the evaluation metrics and KPIs used to compare and rank the controllers.

Scenarios

The test scenarios aim to expose several aspects of the controllers. Namely we aim to test the performance of the controller in cases of (1) congested traffic (2) free flow traffic (3) lane changing. Below we describe each of these scenarios.

Congested Traffic

The congested scenarios highlight cases where stop-and-go waves generate in the platoon starting with the leader and propagate backwards. The purpose of these scenarios is to evaluate the ability of the controller to dampen the waves and reduce the traffic congestion. The scenarios are defined by specifying the leader vehicle's trajectory which experiences stop-and-go waves. We define the leader's trajectory in these scenarios using real-world vehicle trajectory data collected from the I-24 highway during morning peak hours (*7*).

Free-flow Traffic

The free-flow traffic scenarios highlight cases where the trajectory of the leader resembles free flow cases with no congestion. The purpose of these scenarios is to detect any emerging behaviour from the controller that might cause undesired effects that can generate stop-and-go waves or worsen free-flow traffic. Similarly, this benchmark is defined by the leader's trajectory. We extract free flow leader trajectories from the same I-24 dataset.

Lane changing

The lane changing scenarios can be thought of as a variant of the previous two sets of scenarios in which human-driven vehicles can cut-in (appear) into or cut-out (disappear) from the lane. The purpose of this scenario is to evaluate the controllers' safe-handling of lane switching and to expose the consequences of leaving large gaps by the controller. To design these scenarios we enable the lane changing stochastic disturbance in the simulation.

Configuration

We allow some flexibility in the test scenarios' configuration. The user can specify the size and order of the platoon and the penetration rate of AVs. The user can also provide additional test trajectories and evaluate the performance on them. Further the simulation parameters can be easily tenable including the IDM parameters, the size parameter of the lane changing distribution, and the simulation horizon. We note that we provide a default configuration that can be used as a reference for all comparisons.

Tests and Metrics

One of the most important aspects of this evaluation scheme is the design of the metrics and KPIs on which the controllers are compared. The aim is to devise a comprehensive set of metrics to effectively captures the main objective of the control task: reducing traffic congestion. At the same time we aim to detect any side behaviour of the controller that can compromise safety. For this we propose the following metrics:

Average speed.

We use the commonly used metric for evaluating traffic congestion which is the time-average sample-average speed (TASAS) (*3, 11*) defined as

$$TASAS = \frac{\sum_{k=0}^{T}\sum_{i=1}^{n(k)} v_i(k)/n(k)}{T}. \tag{1}$$

Fuel consumption.

Along with increased travel time, fuel consumption wastage is one of the most concerning effect of traffic congestion. It is projected that 2% of the total fuel consumption of vehicles on-highways is wasted due to congestion in 2020, a figure that is project to rise to 4.2% in 2050 (*12*). As such, it is of great interest to measure the potential efficiency improvement provided by the tested control strategies. We use the miles-per-gallon (MPG) metric as a measure of the energy efficiency of the platoon, which is defined as

$$MPG = \frac{\sum_{k=0}^{T} f(T_i,k)x_i(k)}{\sum_{k=0}^{T}\sum_{i=1}^{n(k)} C_i(k)}, \tag{2}$$

where T_i indicates the last time step in which the i^{th} vehicle is in the system, and the function $f(T_i,k) = 1$ if $T_i = k$ and zero otherwise. Here, $C_i(k)$ is the instantaneous fuel consumption of the i^{th} vehicle at time step k. This quantity is computed using the fuel consumption model proposed in (*5*) which is validated using real-world driving data. The MPG represents the total miles traveled by all the vehicles in the platoon divided by the total number of gallons of fuels consumed by those vehicles.

Vehicles crash count.
As a first, and straightforward, measure of safety, we count the number of crashes occurring in the platoon.

Minimum headway.
Another safety metric that we consider in the system is the measure of the minimum space headway achieved by the AVs in the platoon throughout the simulation, defined as

$$\min_{k=0,\ldots,T} \min_{i \in I_a(k)} x_{i-1}(k) - x_i(k). \tag{3}$$

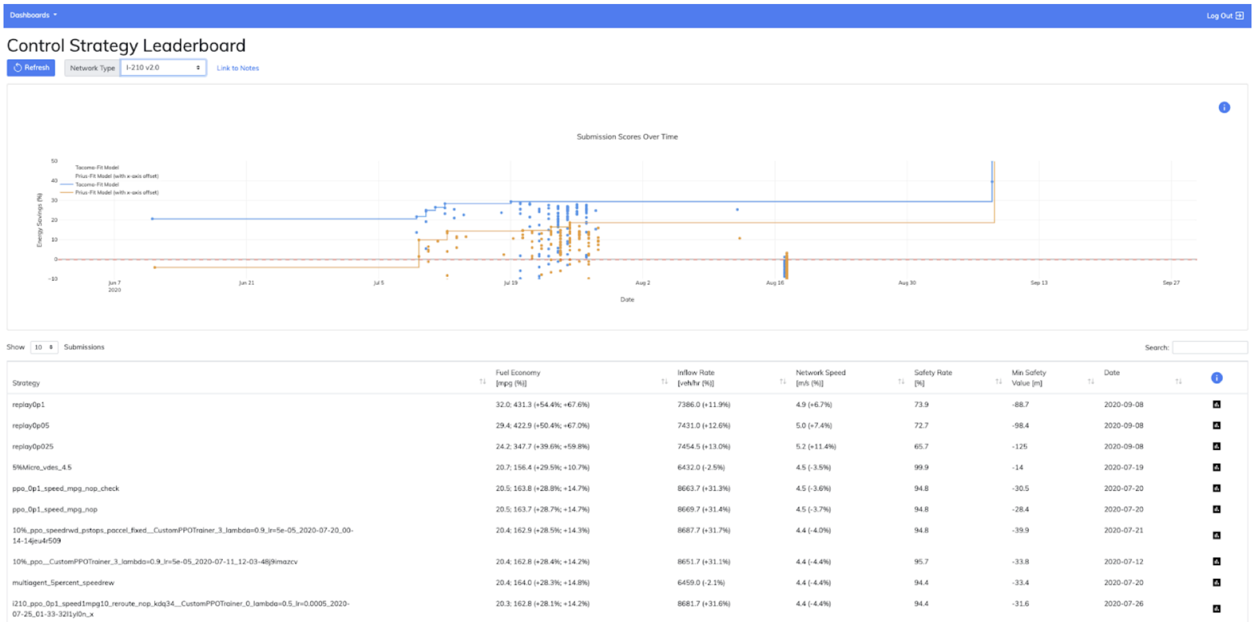
Small space headway might indicate a safety concern, however, we leave the definition of the minimum acceptable headway open for user evaluation as it depends heavily on additional considerations regarding the deployment conditions of the controller.

Controllers
The last component of the evaluation scheme are the algorithms and controller developed for the longitudinal control of the AVs in the platoon. The system architecture is developed to be compatible with and allows easy integration of a wide range of control strategies and algorithms. Any control algorithm can be tested using this proposed evaluation testbed as long as it can be written as a function that takes in the system state vector $s(k) \in R^{n(k)}$ and produces a set of command accelerations $a(k) \in R^N$. For ease of integration we establish a standardizes interface between the simulation and the control algorithm. This architecture is developed with the aim of removing barrier of use and allow for rapid testing and development. This allows us to continuously track progress over time (see Figure 2) and engage effectively with the algorithms developers with minimal overhead from the users' side.

EXPERIMENTS
In this section, we demonstrate a use case of the proposed system using the scenarios described above. We present three different control algorithms for the task of improving traffic congestion and we compare their performance based on the aforementioned evaluation metrics.

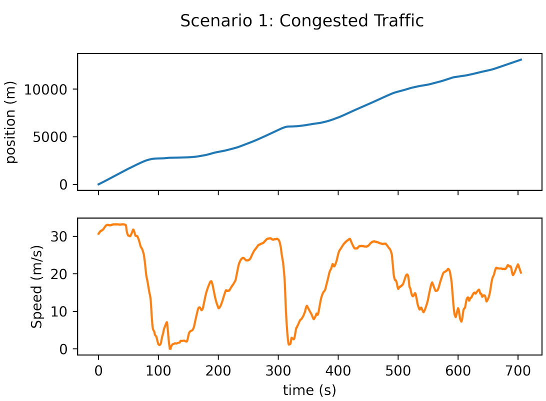
Test Scenarios
We conduct our experiments on a two trajectories, each representing one of the two scenarios: congested and free-flow traffic. The leader's trajectory for these scenarios is shown in Figure 3. For each trajectory we run the experiments with and without lane changing. We fixed the starting platoon for all experiments to be of 200 vehicles. We also use a fix set of initial conditions for all experiments. Two sets of experiments are conducted with AVs penetration rates of 5% and 10% of the initial platoon size. The initial order of the platoon consists of a leader vehicle followed by the AVs equally spaced by human-driven vehicles.

Controllers
We implement three different control strategies for the task of reducing traffic congestion and evaluate their performance using our evaluation system. We use a fully human-driven platoon of the same size as baseline for comparing the metrics (by replacing the AVs with human-driven cars

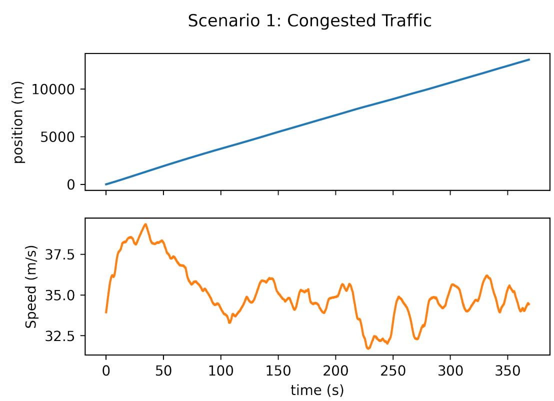


Strategy	Fuel Economy [mpg (%)]	Inflow Rate [veh/hr (%)]	Network Speed [m/s (%)]	Safety Rate [%]	Min Safety Value [m]	Date
replay0p1	32.0; 431.3 (+54.4%; +67.6%)	7386.0 (+11.9%)	4.9 (+6.7%)	73.9	-88.7	2020-09-08
replay0p05	29.4; 422.9 (+50.4%; +67.0%)	7431.0 (+12.6%)	5.0 (+7.4%)	72.7	-98.4	2020-09-08
replay0p025	24.2; 347.7 (+39.6%; +59.8%)	7454.5 (+13.0%)	5.2 (+11.4%)	65.7	-125	2020-09-08
5%Micro_vdes_4.5	20.7; 156.4 (+29.5%; +10.7%)	6432.0 (-2.5%)	4.5 (-3.5%)	99.9	-14	2020-07-19
ppo_0p1_speed_mpg_nop_check	20.5; 163.8 (+28.8%; +14.7%)	8663.7 (+31.3%)	4.5 (-3.6%)	94.8	-30.5	2020-07-20
ppo_0p1_speed_mpg_nop	20.5; 163.7 (+28.7%; +14.7%)	8669.7 (+31.4%)	4.5 (-3.7%)	94.8	-28.4	2020-07-20
10%_ppo_speedrwd_pstops_paccel_fixed__CustomPPOTrainer_3_lambda=0.9_lr=5e-05_2020-07-20_00-14-14jeu4r509	20.4; 162.9 (+28.5%; +14.3%)	8687.7 (+31.7%)	4.4 (-4.0%)	94.8	-39.9	2020-07-21
10%_ppo__CustomPPOTrainer_3_lambda=0.9_lr=5e-05_2020-07-11_12-03-48j9imazcv	20.4; 162.8 (+28.4%; +14.2%)	8651.7 (+31.1%)	4.4 (-4.4%)	95.7	-33.8	2020-07-12
multiagent_5percent_speedrew	20.4; 164.0 (+28.3%; +14.8%)	6459.0 (-2.1%)	4.4 (-4.4%)	94.4	-33.4	2020-07-20
i210_ppo_0p1_speed1mpg10_reroute_nop_kdq34__CustomPPOTrainer_0_lambda=0.5_lr=0.0005_2020-07-25_01-33-32l1yl0n_x	20.3; 162.8 (+28.1%; +14.2%)	8681.7 (+31.6%)	4.4 (-4.4%)	94.4	-31.6	2020-07-26

FIGURE 2: A screenshot from the dashboard of the evaluation testbed illustrating the progress of the controllers in achieving the energy saving metric over time. Each dot represents a new controller submission and the lines (orange and blue) represent the progress of the best achieving controllers on different scenarios.



(a) Leader trajectory for the first scenario



(b) Leader trajectory for the second scenario

FIGURE 3: Experiments test scenarios

in the platoon). Below, we give a short description for each of the controllers. We omit the details of the controllers and their development as it is beyond the scope of this work and are only used as a sample controllers for this experiments.

Microcontroller

This is a hand-designed control strategy. The microcontroller has three major components: (1) computing and trying to follow a desired velocity to ensure a uniform steady-state flow, (2) correction of this desired velocity based on the leader's state to make it locally adaptable, and (3) a headway management component to ensure safe driving at all time. The commanded accelerating of this controller is a combination of the acceleration produced by each of these components.

Deep reinforcement learning (DRL)
This is a control policy trained using a version of the proximal policy optimization (PPO) algorithm. The training was conducted on the I-24 trajectory dataset using a reward function that encompasses the evaluation metrics represented in the system.

Optimized Human-driver Controller (OHDC)
This is an optimization based controller in which we start with a collected of human-driver models and optimize their parameters to achieve better values for the desired evaluation metrics. We also use the I-24 trajectory data and optimize for the parameters that improve the fuel consumption and the average speed for all training trajectories simultaneously.

Results

Scenario 1: Congested Traffic

Without lane changes
In Table 1, we compare the performance of the three controllers against the baseline on the congested traffic scenario with no lane changes. For this experiment, all three control strategies achieve an improvement over the baseline in the average speed and evaluation metrics. With only 5% penetration rate, the controllers can achieve an improvement of up to 12.7% in the average speed, and up to 19% in the fuel consumption, both using the DRL controller. With a penetration rate of 10% the improvement increases to 19.15% for the average speed (using the OHDC), and 25.39% for the fuel consumption (using the Microcontroller). These results are generally consistent with findings of previous studies that suggests improved performance with increased penetration rates of AVs.

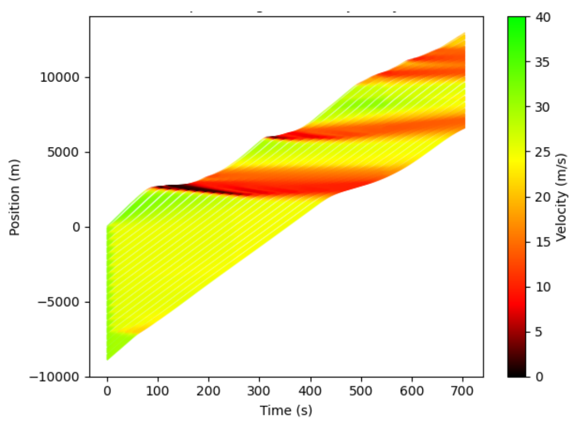
We note from these results that we do not get consistent ranking of the controllers when using different evaluation metrics. That is, the best performing controller in terms of fuel consumption might not be the one performing best in terms of average speed. This highlights the fact that these metrics and not equivalent and highlights the importance of considering them separately.

We also note that all the controllers perform at least as good as the baseline when it comes to the safety metrics. All controllers complete the simulation without causing crashes. Further, the controllers tend to keep a headway distance that is significantly larger than the baseline, except for the the Microcontroller.

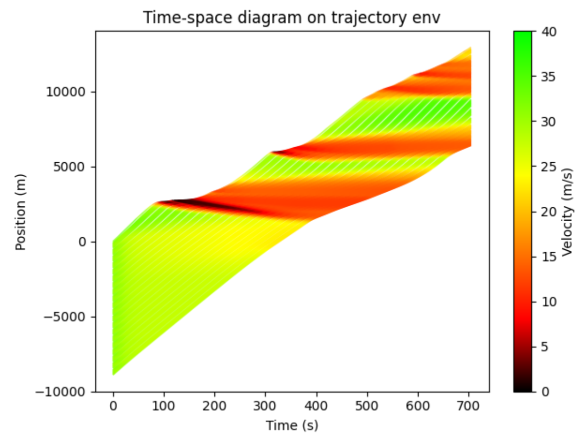
	AS (m/s)		**FC (mpg)**		**CC (#)**		**MH (m)**	
Penetration Rate	5%	10%	5%	10%	5%	10%	5%	10%
Microcontroller	19.00	17.95	44.78	47.26	0	0	2.20	10.32
DRL	19.53	18.98	44.86	46.79	0	0	11.56	22.14
OHDC	18.52	20.65	42.76	41.43	0	0	9.36	9.81
Baseline	17.33		37.69		0		2.50	

TABLE 1: Comparison results for the first scenario **without** lane switching. The evaluation metric in the table are: AS (Average Speed), FC (Fuel Consumption), CC (Crash Count), and MH (Minimum Headway)

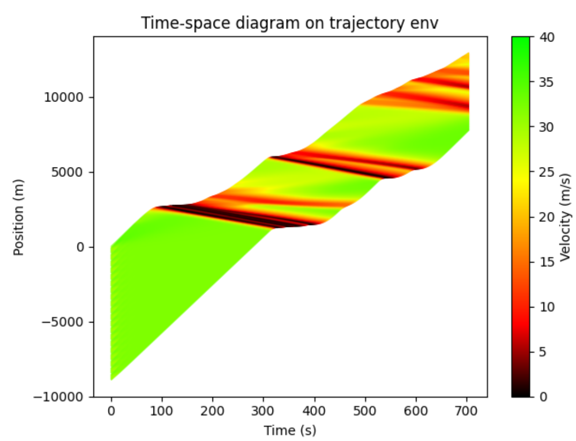
It is worth mentioning that the system provides visualizations of the simulated scenario. As an example, in Figure 4 we include the time-space diagram of the platoon for this scenario and compare between the trajectories resulting from the Microcontroller and the DRL controller. We see from the figure that both controllers effectively smooth the stop-and-go waves. However, we notice some differences in their performance. The micro-controller tends to leave larger gaps and causes some reduction in the speed and the beginning of the trajectory compared to the DRL controller. Though such illustrations are difficult to use for systematic ranking of the controllers it can offer further insights into their performance and behaviour.



(a) Microcontroller



(b) DRL Controller



(c) Baseline

FIGURE 4: Time-space diagram for the congested scenario without lane changes for the platoon with 10% penetration rate.

With lane changes

In table 2, we present the results for the congested traffic scenario with lane changes. The importance of including lane changing behaviour in the simulation is that it indirectly penalizes leaving a large space headway, something a controller can exploit to improve other metrics like the fuel consumption. This is due to the fact the the probability of a lane change increases with the headway gap. As a result, we notice a significant decrease in the performance of all the controllers when it comes to fuel consumption. However, the average speed metric is slightly improved in the case of lane changes. We note that the DRL controller succeeds in maintaining a good performance even in the presence of lane changes, achieving a fuel consumption improvement up to 19.25% and up to 5% increase in the average speed.

	AS (m/s)		**FC (mpg)**		**CC (#)**		**MH (m)**	
Penetration Rate	5%	10%	5%	10%	5%	10%	5%	10%
Microcontroller	18.50	18.56	38.83	39.81	0	0	9.80	14.52
DRL	19.59	19.25	42.03	42.73	0	0	8.33	1.02
OHDC	19.93	23.54	36.22	36.51	0	0	8.67	7.88
Baseline	17.66		35.83		0		2.50	

TABLE 2: Comparison results for the first scenario **with** lane switching. The evaluation metric in the table are: AS (Average Speed), FC (Fuel Consumption), CC (Crash Count), and MH (Minimum Headway)

Scenario 2: Free-flow traffic

Without lane changes

In Table 3, we present the results for the free flow traffic scenario without lane changes. The aim of presenting such results is to observe the performance of the controllers and ensure that they are not worsening the traffic conditions in the free flow scenario. We notice that the controllers managed to improve the fuel consumption over the baseline, though the improvement is far less significant than in the congested traffic scenario. We note however that all the controllers perform slightly worse than the baseline when it comes to the average speed. Such is a behaviour that we need to be mindful of when testing and evaluating controllers. Another observation we make is concerning the minimum headway achieved by the controllers in the free flow conditions. We notice that these are much higher than those for the congested scenario. This is due to the fact that the average speed is also higher in these scenario, which requires a larger space gap to ensure safety. This draws our attention to the potential benefit of revising the minimum headway metric to be the minimum time-headway (define as space headway divided by the vehicle's speed) to give a better understanding of the safety of the controller.

With lane changes

Finally, and for the sake of completion, we present the results of the free flow traffic scenario with lane changes in Table 4. We make similar observations regarding the limited performance in terms of fuel consumption and decreased average speed compared to the baseline. We note that the minimum headway is significantly smaller in the case of lane changes. This might indicate an unsafe condition due to aggressive cut-ins. Thought such behaviour requires additional investigation, and more metrics can be added to the system in the future to indicate such behaviour.

CONCLUSION

In this work we present a simulation testbed for evaluating the performance of control strategies for improving traffic congestion. We present a system consisting of a micro-simulation engine of

	AS (m/s)		**FC (mpg)**		**CC (#)**		**MH (m)**	
Penetration Rate	5%	10%	5%	10%	5%	10%	5%	10%
Microcontroller	33.80	33.56	37.82	38.11	0	0	84.23	86.35
DRL	33.80	33.21	38.19	39.15	0	0	86.03	87.65
OHDC	33.21	32.27	39.12	40.68	0	0	101.83	101.83
Baseline	34.80		35.19		0		43.79	

TABLE 3: Comparison results for the second scenario **without** lane switching. The evaluation metric in the table are: AS (Average Speed), FC (Fuel Consumption), CC (Crash Count), and MH (Minimum Headway)

	AS (m/s)		**FC (mpg)**		**CC (#)**		**MH (m)**	
Penetration Rate	5%	10%	5%	10%	5%	10%	5%	10%
Microcontroller	32.73	32.73	34.31	34.52	0	0	38.35	36.20
DRL	32.71	32.13	34.63	36.09	0	0	42.77	19.27
OHDC	32.47	32.09	34.37	36.40	0	0	37.16	25.97
Baseline	34.42		32.45		0		19.10	

TABLE 4: Comparison results for the second scenario **with** lane switching. The evaluation metric in the table are: AS (Average Speed), FC (Fuel Consumption), CC (Crash Count), and MH (Minimum Headway)

mixed-autonomy traffic, a set of test scenarios, and a set of evaluation metrics to evaluate, compare, and rank different control strategies. We illustrate the ability of our system to integrate a wide range of controllers by implementing and testing three different controllers that are very different in nature. We conducted a few experiments on a variety of scenarios and test cases. Through these experiments we illustrated the use of our system, and demonstrated its features. Furthermore, we identified, through these experiments, a few shortcomings of our proposed evaluation metrics, which opens the door for future improvements of the system by introducing additional metrics to address such deficiencies.

There are multiple interesting avenues for expanding this work in future. One is scaling the simulation engine to accommodate multiple lanes and detailed simulation of the lateral behaviour of both AVs and human-driven cars. Another area of improvement is to enrich the evaluation system with additional scenarios (e.g. merging lanes and bottleneck scenarios) and additional metrics to capture more features of the controller such as passenger's comfort. Finally, we aim to achieve a high level of engagement in the form of both accepting submissions of controllers and allowing community contributions to the development of the system's compoenent.

ACKNOWLEDGEMENTS

The authors would like to thank Nathan Lichtlé, Brent Zhao, and Eric Cheng for their contribution to the back-end and front-end development of the system. The authors would also like to thank Amaury Hayat, Eugene Vinitsky, Adit Shah, and Nathan Lichtlé for their contribution to this work through the development and training of the control strategies used in the experiments.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: A. Alanqary, J. Lee, A. Bayen; data collection: J. Lee; experiments design: A. Alanqary, J. Lee; analysis and interpretation of results: A. Alanqary; draft manuscript preparation: A. Alanqary. A. Bayen. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. James Ault and Guni Sharon. Reinforcement learning benchmarks for traffic signal control. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1)*, 2021.
2. M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling. The arcade learning environment: An evaluation platform for general agents. *Journal of Artificial Intelligence Research*, 47:253–279, jun 2013.
3. Jiaxun Cui, William Macke, Harel Yedidsion, Aastha Goyal, Daniel Urielli, and Peter Stone. Scalable multiagent driving policies for reducing traffic congestion. *arXiv preprint arXiv:2103.00058*, 2021.
4. Yan Duan, Xi Chen, Rein Houthooft, John Schulman, and Pieter Abbeel. Benchmarking deep reinforcement learning for continuous control. In *International conference on machine learning*, pages 1329–1338. PMLR, 2016.
5. Jonathan W Lee, George Gunter, Rabie Ramadan, Sulaiman Almatrudi, Paige Arnold, John Aquino, William Barbour, Rahul Bhadani, Joy Carpio, Fang-Chieh Chou, et al. Integrated framework of vehicle dynamics, instabilities, energy models, and sparse flow smoothing controllers. *arXiv preprint arXiv:2104.11267*, 2021.
6. Vicente Milanés, Steven E Shladover, John Spring, Christopher Nowakowski, Hiroshi Kawazoe, and Masahide Nakamura. Cooperative adaptive cruise control in real traffic situations. *IEEE Transactions on intelligent transportation systems*, 15(1):296–305, 2013.
7. Matthew Nice, Nathan Lichtle, Gracie Gumm, Michael Roman, Eugene Vinitsky, Safwan Elmadani, Matt Bunting, Rahul Bhadani, Kathy Jang, George Gunter, Maya Kumar, Sean McQuade, Chris Denaro, Ryan Delorenzo, Benedetto Piccoli, Daniel Work, Alex Bayen, Jonathan Lee, Jonathan Sprinkle, and Benjamin Seibold. The i-24 trajectory dataset, September 2021.
8. Raphael E Stern, Shumo Cui, Maria Laura Delle Monache, Rahul Bhadani, Matt Bunting, Miles Churchill, Nathaniel Hamilton, Hannah Pohlmann, Fangyu Wu, Benedetto Piccoli, et al. Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. *Transportation Research Part C: Emerging Technologies*, 89:205–221, 2018.
9. Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012.
10. Martin Treiber and Arne Kesting. The intelligent driver model with stochasticity-new insights into traffic flow oscillations. *Transportation research procedia*, 23:174–187, 2017.
11. Eugene Vinitsky, Aboudy Kreidieh, Luc Le Flem, Nishant Kheterpal, Kathy Jang, Cathy Wu, Fangyu Wu, Richard Liaw, Eric Liang, and Alexandre M Bayen. Benchmarks for reinforcement learning in mixed-autonomy traffic. In *Conference on robot learning*, pages 399–409. PMLR, 2018.
12. Zia Wadud, Don MacKenzie, and Paul Leiby. Help or hindrance? the travel, energy and carbon impacts of highly automated vehicles. *Transportation Research Part A: Policy and Practice*, 86:1–18, 2016.
13. Haoran Wei, Xuanzhang Liu, Lena Mashayekhy, and Keith Decker. Mixed-autonomy traffic control with proximal policy optimization. In *2019 IEEE Vehicular Networking Conference (VNC)*, pages 1–8. IEEE, 2019.