Proceedings of the ASME 2023
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC/CIE2023
August 20-23, 2023, Boston, Massachusetts

DETC2023-DRAFT

ROUNDABOUT TRAFFIC: SIMULATION WITH AUTOMATED VEHICLES, AI, 5G, EDGE COMPUTING AND HUMAN IN THE LOOP

Giorgio Previati
Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy

Gianpiero Mastinu
Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy

Elena Campi
Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy

Massimiliano Gobbi
Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy

Lorenzo Uccello
Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy

Álvaro Varela Daniel
Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy

Antonino Albanese
Italtel S.p.A., Via Caldera 21, 20136, Milan, Italy

Alessandro Roccasalva
TECH – CRF S.C.p.A., Strada Torino 50, 10043 Orbassano (TO), Italy

Gabriele Santin
Fondazione Bruno Kessler, Trento, Italy

Massimiliano Luca
Fondazione Bruno Kessler, Trento, Italy and Faculty of Computer Science, Free University of Bolzano

Bruno Lepri
Fondazione Bruno Kessler, Trento, Italy

Laura Ferrarotti
Fondazione Bruno Kessler, Trento, Italy

Nicola di Pietro
Athonet Italy, Via Cà del Luogo 6/8, 36050, Bolzano Vicentino (VI), Italy

ABSTRACT

The aim of the paper is to assess how the traffic of roundabouts could be organized in the future. A mixed traffic is supposed to occur, featuring both fully automated vehicle and vehicles driven by humans. The case study is a part of a comprehensive research funded by European Commission, focusing on how improving the 5G network by Artificial Intelligence (AI) and edge computing.

The traffic flow into a reference roundabout is simulated by SUMO, a Reinforcement Learning (RL) algorithm is derived and drives the automated vehicles into the roundabout. By means of a dynamic driving simulator, a real human drives into the simulated traffic of the roundabout. The human-in-the-loop is

needed to check whether the driver feels comfortable and safe while driving in presence of automated cars.

The preliminary results are quite promising. Drivers seem not worried when they interact with Connected and Automated Vehicles (CAVs) into the roundabout. They seem preferring to share the roundabout with CAVs as they seem feeling the traffic to be more fluent. Such preliminary results deserve an extensive statistical investigation for a final assessment.

Keywords: Roundabout traffic simulations, Reinforcement learning, Driving simulator, Human in the loop

1. INTRODUCTION

Artificial Intelligence has become a major innovative force and it is one of the pillars of the fourth industrial revolution. In fact, European Commission has already highlighted how highly performing, intelligent and secure networks are fundamental for the multi-service Next Generation Internet. The objective of the European project AI@EDGE [1] is to integrate AI-enabled platforms in potentially autonomous decision-making systems or even critical infrastructure. In this context, the project studies, among other practically important applications of AI and 5G, the negotiation of a roundabout by automated vehicles. In the considered scenario, several connected and automated vehicles (CAVs) exchange data related to their trajectories, velocities and accelerations, in order to successfully navigate into a roundabout while interacting with human driven vehicles. A reinforcement learning process is carried out, with the aim of defining a policy able both to drive safely CAVs into the roundabout and to optimize the traffic flow.

Managing road intersections has always been a challenging task as they represent bottlenecks in the road network. Their poor management can produce road congestion and increase travel time, emissions and fuel consumption [2]. In the last decades, roundabouts have drawn particular attention because, thanks to their design, they can be crossed by drivers avoiding the need for a complete stop and allow to reduce the number of conflict points with respect to signalized intersection, as shown in Fig.1, decreasing the severity of crashes and nearly eliminating head-on collisions [3].

In this scenario, connected and automated vehicles can represent an additional improvement because they would allow to increase driving safety, travel efficiency and passenger comfort [5]. For these reasons, it is important to investigate the behaviour of the connected and automated vehicles while approaching the roundabout. Algorithms must be tested through virtual simulations before the implementation on real roads, in order to increase the level of safety associated to the developed solutions.

A roundabout is currently a ‘nightmare’ for engineers dealing with automated driving [6]. Actually, automated cars get often stuck in roundabouts as they are unable to keep the accident risk at the needed minimum level [7, 8]. As a consequence, the roundabout becomes a -or ‘the’- bottleneck for connected and automated cars to be introduced on the market. Despite the many technical challenges related to the negotiation of roundabouts by automated vehicles, several benefits can be envisaged

- Safer traffic (less accidents)
- More efficient traffic (more fluid traffic)
- Less pollution (less vehicles stop and go at the junction)

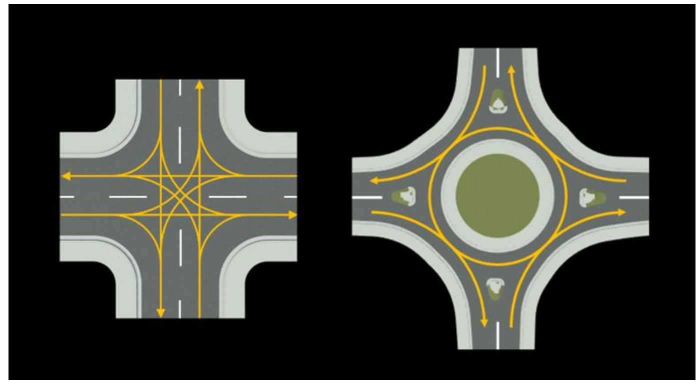

FIGURE 1: CONFLICT POINTS AT FOUR-LEG INTERSECTION AND SIGLE-LANE ROUNDABOUT. CONFLICT POINTS OCCUR WHERE LINES CROSS EACH OTHER OR JOIN [4]

In the literature, the control approaches for manging traffic into a roundabout can be divided, from the network architecture point of view, into 2 main categories [9].

The first one is based on the use of a local system: a local controller gathers all the information from the approaching vehicles in the communication range, makes choices and gives information back to the vehicles; this kind of architecture id referred to as Vehicle-To-Infrastructure (V2I) communication. Such an approach has not been considered in this study because the management of the infrastructure is expensive and complex from the data management point of view.

The second category is based on a multi-agent system: in this case, the decision-making agents are directly the vehicles involved in the intersection area; this approach exploits the so called vehicle-to-vehicle (V2V) communication. Such a direct connection between vehicles suffer of the Non-line-of-sight (NLOS) problem and of the multi hop communication between distant vehicles so does not seem to provide an efficient data exchange for our scenario and is not considered in this research.

Another way to categorize the different algorithms referring to V2X depends on the objective function taken into account [9]. We can have reservation-based algorithms, or trajectory-based algorithms. Both of the two kinds of algorithms are not taken into account in our study due to their complexity and depending on propagation conditions; those conditions affect the control of the real time motion of automated vehicles. This gap is evident, in [10], where an approach is proposed to make automated vehicles navigate the roundabouts. The formation of vehicle platoons is studied (platoon control strategy is a tricky question, addressed e.g. in [11]). Vehicles are informed of the optimal sequence of merging times. The main drawback associated to the use of this approach is that once platoons are formed, the gaps between the vehicles are very small, so any communication delay can produce a detrimental effect and lead to potential collisions.

Due to the complexity of the scenarios and the number of vehicles involved, an innovative approach based on V2N2V (Vehicle-to-Network-to-Vehicle) communication through the MEC/Edge nodes has been approached, leveraging the features of 5G networks to guarantee the quality of service required in this context.

To drive CAVs into the roundabout a multi-agent reinforcement learning algorithms has been used. Here we briefly introduce -for the sake of the reader- the Reinforcement Learning (RL) framework, by summarizing the basic concepts, the relevant literature, and some details on the algorithm employed in our scenario.

RL is a field of Machine Learning inspired by natural learning mechanisms, where animals adjust their actions based on rewards and punishments received from the environment. The classic RL setup is composed of an actor (or agent) that interacts with the surrounding environment, choosing at every instant of time which action to do. The objective of Reinforcement Learning algorithms is to train the agent, for it to automatedly learn the best behavior to maintain in the environment, given specific problem-dependent goals [12]. Based on such idea, an RL system uses a reward function to generate the stimuli from the environment to the agent associated with each state-action couple. The agent is modeled as a policy function, mapping each state perceived by the agent into the space of possible actions that the agent can take. The goals to be achieved by the agent are represented instead by a value function, approximating the long-term gain that an agent is supposed to collect starting from a state (or a state-action couple). Based on the presence (or lack thereof) of a formal model of the environment, RL approaches are divided into model-based and model-free. Model-based RL approaches can investigate prospective scenarios and situations before they are encountered, by using the model to provide a simulated experience, predicting the next state and reward, given a specific state-action couple. On the contrary, model-free methods learn from experience matured by direct interaction with the real environment [13].

The model-based RL approach is known to present difficulties when applied to traffic systems, due to issues in modeling large-dimensional state spaces. Additionally, representing real-world problems as fully observed models is prohibitive in most cases. On the other hand, in recent years the data-driven model-free RL framework has been used in many works to optimize traffic management [14, 15], considering different states perceived by the agent, actions, and reward functions. Examples of possible states are the length of the queue to represent the state of the traffic congestion [16, 17, 18], the position of the vehicles with respect to the stop line, and its relative speed [19, 15]. In the literature, commonly used reward functions are for instance the length of the queue [17, 20], the average delay [14, 21], the cumulative delay [22, 23], the vehicle staying time [15], or sometimes a combination of indices, as in [20], where a weighted sum of the length of the queue, the vehicle delays, the average waiting time, the light switcher state, the number of vehicles and the total travel time of each car is employed. Significant results have been obtained by employing RL techniques in the field of adaptive control of traffic signals to optimize both single intersections [15, 22, 23, 20] and multiple crossroads scenarios [24, 21, 25, 26]. In [16] it is shown that, after training an agent over a single intersection scenario, the obtained insights are beneficial in order to manage multiple intersections. In [18] a Multi-Agent Reinforcement Learning (MARL) framework is proposed to coordinate the potentially conflicting goals of the agents. The roundabout scenario is as well considered in the literature as a benchmark. RL and Deep RL are employed to optimize the capability of artificial agents of executing maneuvers and insertions in a roundabout, training both single automated vehicles and multi-agent systems of CAV in the MARL framework, i.e. [6, 27, 28]. On the algorithmic side, most of the mentioned works use either Q-Learning [29, 30] and Deep Q-Learning [31] algorithms to learn to assess the value function and extrapolate the policy in a greedy or semi-greedy way, or Policy Gradient [32, 33] algorithms to parameterize and learn the policy function as well.

The paper is structured as follows. At first the use case is presented. Then the real scenario is implemented into a virtual environment in which a driving simulator is used. In particular, the SUMO software is introduced together with the RL-learning algorithm which defines the optimal policy to drive CAVs into the roundabout. Finally first investigations on the traffic flow into the roundabout are performed with a human in the loop to assess the driver/passenger comfort during automated driving into a roundabout.

2. AI@EDGE PROJECT: USE CASE ON RUNDABOUT TRAFFIC WITH AUTOMATED VEHICLES

Following the V2N2V approach, specifically developed in the AI@EDGE project, a MEC/Edge node is installed ideally in the centre of the roundabout. It receives data about position, speed and acceleration from all the vehicles involved in the scenario, i.e. by both conventional and automated ones. By means of the AI algorithm, the MEC/Edge node defines the speeds of each automated vehicle, with the aim to obtain a safe traffic flow and a reduction of the air pollution due to vehicles' movement. The MEC/Edge node sends via 5G to the automated vehicles the instructions to accelerate or to brake. We made the hypothesis that the non-automated vehicles, i.e. the ones simply driven by humans, just send data (position, velocity acceleration) to the network but they do not receive any data to control their respective speeds.

In this scenario, vehicles will exchange data among them without relying on a broadcast from an infrastructure authority, basing their decision on a shared policy defined by an AI policy executed at the edge server.

The future non-human driven cars, i.e. CAVs, will share the road with human driven cars. We consider this complicated case, focusing on 20% and 80% market penetration of automated vehicles. Not only vehicles in the roundabout, but also all vehicles approaching the roundabout are connected with the network, i.e. the MEC/Edge node gets the position, velocity and acceleration of each vehicle. A 5G telematic box on board of the car does the job of exchanging data. As addressed above, non-human driven vehicles both receive and transmit data. Conversely, human driven vehicle(s) just provide their data but do not receive any information. The driver cannot be bothered and so he/she drives according to his/her wish.

The performance levels requested to the mobile network by this application must provide low latency and high reliability and represent a core challenge that goes beyond current 5G design. The AI@EDGE project [1] aims to overcome such limitations and to provide data exchange at the required latency. This will allow the execution of the AI policy, able to manage a big number of vehicles.

Such a complex scenario should be extensively tested with human drivers in the loop. Real world tests are clearly prohibitive in terms of availability of a large number of automated vehicles and, most important, for the safety issues related to the presence of human drivers. On the other hand, purely simulated tests in a completely virtual environment cannot represent the human behaviour when human drivers have to interact with automated vehicles.

In this context, the use of a driving simulator seems a viable way to perform such kind of investigations. A dynamic driver simulator can guarantee to test the AI policy in a safe environment, avoiding the risks connected to real traffic and the possibility of congestion and accidents. Secondly, the use of a driving simulator reduces the costs related to the experimental set-up: in fact, it is possible to simulate any number of autonomous vehicles in the scenario and explore different situations simply modifying the parameters involved in the simulations, thus avoiding the costs related to experimental equipment in a real environment. Finally, the use of a simulator ensures the repeatability of the test, as it eliminates the influences due for examples to weather conditions or amount of traffic.

Among the many challenges posed by this use case, multi-sensors data synchronization is certainly one of the biggest, i.e. how to handle such a variety of data sources and how to synchronize them. The accuracy of time synchronization indeed is expected to affect the vehicle control's safety; so handling the sensor data with different frequency and timestamp accuracy remains a very big challenge.

To support such scenario, a real-time data processing coming from various sensors is required. So, powerful computing hardware is essential and, for this reason, different HW acceleration solutions are considered, along with experimentation and comparison of different HW acceleration equipment like e.g Graphic Processor Unit (GPU), Digital Signal Processor (DSP), Field Programmable Gate Arrays (FPGA), Application-Specific Integrated Circuit (ASIC) etc.

In order to integrate so many different technologies and multiple different equipment, together with the innovations brought by AI@EDGE project, strong systems integration capabilities are needed too.

3. HARDWARE: SETUP AT THE DYNAMIC DRIVING SIMULATOR

The test are carried at the DriSMi laboratory of Politecnico di Milano [34]. The laboratory is equipped by a high fidelity last generation cable driven driving simulator. The driving simulator is depicted in Fig.2 and its performance are summarized in Table 1. More details on the driving simulator can be found in [35]. The traffic scenario is generated thanks to a co-simulation between SUMO (Simulator of Urban MObility [36]), which is used as traffic engine, and VI-WorldSim [37], that simulates the motion of the vehicle driven by the human in the loop and allows to have a graphical representation of the traffic scenario. SUMO is in charge of the simulation of all virtual vehicles in the roundabout, comprising virtual human driven vehicles and virtual non-human driven (CAVs) vehicles. SUMO receives the data pertaining to the car driven by the human in the loop (called ego-vehicle) from the driving simulator through VI-WorldSim. The current position of all simulated vehicles is fed to VI-WorldSim which is in charge of all the graphical environment of the driving simulator, of the interface with the human in the loop and of the simulation of the motion of the ego-vehicle accordingly to the request of the real human driver.

FIGURE 2: DRIVING SIMULATOR AT DRISMI LABORATORY OF POLTECNICO DI MILANO

All of the simulations are performed in real time and the corresponding data are stored in a real time database. The human in the loop drives the vehicle from the cockpit of the driving simulator, which moves accordingly to the simulated motion of the ego-vehicle. The cockpit is equipped with a telematic box. The telematic box is connected to the CAN bas of the cockpit, reads the dynamic data of the car, and transmits them to a 5G radio platform. The 5G radio platform transmits the data to and from a edge server and is controlled by a NUC where the advanced services developed in the AI@EDGE project are installed. Finally, the edge server hosts the AI in charge of the policy which is used to control the connected automated vehicles simulated by SUMO.

TABLE 1: DRIVNG SIMULATOR DATA

Physical quantity	Values
Platform size	6m x 6m
Visual system (H)	270°
Visual system (V)	90°
Degrees of freedom	9
Longitudinal acceleration of the base	1.5 g
Lateral acceleration of the base	1.5 g
Vertical acceleration of the cockpit	2.5 g
Longitudinal travel	4.2 m
Lateral travel	4.2 m
Vertical travel	± 298 mm
Yaw angle	± 62°
Roll angle	± 15°
Pitch angle	± 15°

4. 5G AD ROUNDABOUT TRAFFIC

In the proposed scenario, the role of the 5G network is central in allowing an efficient low-latency and reliable data exchange, both uplink, i.e., from the connected vehicles to the edge servers where the AI algorithms run, and downlink, i.e., from the edge servers to the CAVs. A key feature of the experimental setup is the deployment physically very close to the edge servers of a 5G radio access point (a base station) and of a dedicated distributed 5G User Plane Function (UPF). In particular, the UPF is part of the 5G Core Network, and it has the role of forwarding traffic between the base station and the MEC/Edge node. Having the UPF positioned at the edge of the 5G network, very close to the so-called User Equipment (the telematics boxes of the connected vehicles) and the radio access point, enables traffic localization: all the user data is maintained local within the roundabout and its surroundings, and it does not flow through remote locations before reaching its destination, as would happen instead with a centralized non-dedicated UPF. From the point of view of the quality of service, this enables low latencies and supports the efficiency of V2N2V communications; from the network perspective, this also reduces the impact of the considered services on the transport network, preventing relevant amounts of data to be uselessly forwarded throughout it. The UPF of the proposed testbed is fully virtualized and deployed as a virtual machine over a common-off-the-shelf server.

4.1 SUMO

SUMO is a software which allows to simulate several traffic scenarios, such as intersections and roundabouts, in order to understand how the vehicles interact in different situations [36]. It gives the possibility to modify the parameters related to driver behaviour, so it is possible to simulate both human driven vehicles and connected and automated ones in the same traffic scenario. In addition, it is possible to include in a SUMO simulation also externally controlled vehicles (i.e. vehicles driven by a different traffic engine or by a human driver) which can interact with the vehicles simulated by SUMO.

An important hypothesis is that CAVs driving skills can be well simulated by mimicking the human skills. Actually, in our simulations we will start the reinforcement learning algorithms by introducing a IDM (intelligent driver model) [38]. Additionally, the X-by wire controls are ABS (Antilock Braking System), ESP (Electronic Stability Program) and AEB (Automated Emergency Braking) only [39].

4.2 Reinforcement Learning to define roundabout traffic

In this work, we use SUMO as simulation environment and we consider a mixed-drivers scenario: the roundabout is simultaneously occupied by a set of individual drivers controlled by SUMO itself, representing human drivers, and by a group of CAVs, controlled by a policy trained using RL techniques and deployed at the edge. Each AV independently decides at every instant which action to take, specifically the next value of acceleration to actuate and a discrete decision on changing line/maintaining the current one, based on the trained policy. Such decision is based on the agent individual state, built from past acceleration values and past change-line decisions. The policy is trained using Proximal Policy Optimization (PPO) [40], a Policy Gradient RL algorithm that simplifies the Trust Region Policy Optimization (TRPO) algorithm [41]. TRPO, in order to improve training stability, carries on a constrained optimization, enforcing a trust-region constraint expressed by a Kullback–Leibler divergence [45] on the policy update, with the rationale of taking the largest step possible to improve performance, while satisfying a special constraint on how close the probability distributions of old and new policies are. TRPO is quite complicated to implement, and the necessary constrained optimization might be expensive. PPO solves these aspects by substituting the hard constraint of TRPO with a penalty in the objective function.

4.3 First tests with a human in the loop in mixed traffic conditions in a roundabout

We have investigated whether the RL framework has been able to define a behaviour policy for CAVs running into a roundabout which is acceptable and applicable. The driver -which is in the loop- is used as a sensor, i.e. he/she has to check whether it is possible to drive safely and comfortable when CAVs run -together with other human driven cars- into the roundabout. Our target is to understand if the tested policy is somehow appreciated by the drivers and if it is perceived as safe and comfortable or not.

As we mentioned, the goal of our RL solution is to help AV make decisions to minimize the time needed to go through the proposed roundabout. To measure the goodness of our trained policy, we measure the average time needed by a vehicle to go from any entrance to any exit. Given a simulation of 400 seconds with a given percentage of AV and n vehicles $v_1, v_2, \dots, v_n$, we associate to each vehicle a timestamp when it is entering ($t_{v_i}^{in}$) and when it is exiting ($t_{v_i}^{out}$) the roundabout. The average time is then computed as

$$\mu(t) = \frac{\sum_{i=1}^{n}(t_{v_i}^{out} - t_{v_i}^{in})}{n} \quad (1)$$

The results obtained are reported in Table 2:

TABLE 2: AVERAGE TRAVEL TIME AS A FUNCTION OF AV MARKET PENETRATION RATE

AV percentage	$\mu(t)$ 2s]
10	6.33
20	6.27
30	6.18
40	6.11
50	5.88
60	5.26
70	4.73
80	4.72
90	4.69
100	4.66

The tested scenario referred to a three-legged single-lane roundabout. Since the roundabout was small, the manoeuvres inside it were very quick. In order to give the possibility to the testers to have a full overview of the situation, every driver approached the roundabout one time per each leg. He/she waited his/her turn to enter into the circulatory roadway, then he/she took the second exit. This choice allowed the driver into the circulatory roadway to cross another entry flow and to observe the behaviour of the approaching vehicles. The manoeuvres were repeated 2 times, once with 20% of automated vehicles and once with 80% of CAVs. The participants to the test were not previously informed on the presence of automated cars. This was done in order to investigate whether the driver was able to perceive a difference between the two mixed traffic scenarios i.e. when the simulated market penetration of automated cars was 20% or 80% of the total.

FIGURE 3: SCENARIO WITH VEHICLES INSIDE THE ROUNDABOUT

There were 40 vehicles in the traffic scenario, 39 of which were simulated and the other one was driven by the human seated in the cockpit of the driving simulator (Fig.2). The automated vehicles were controlled through the RL behaviour policy, trained by means of reinforcement learning with the objective to minimize the travel time associated to the roundabout. The non-autonomous cars were controlled directly by SUMO, using IDM car-following model, with the aim to reproduce human behaviour as faithfully as possible. The tests were carried on twelve people, whose data regarding age and driving experience are reported in APPENDIX A. At the end of the test, they were informed on the mixed traffic environment they drove in. They were asked to answer some questions. The results obtained are shown in the following tables.

TABLE 3: ANSWERS ASSOCIATED TO TRAFFIC SMOOTHNESS

Regarding traffic smoothness, which of the following statements do you agree with the most?	Number of answers
Traffic in the scenario with 20% of CAVs was definitely smoother than in the scenario with 80% CAVs	0
Traffic in the scenario with 20% of CAVs was partially smoother than in the scenario with 80% CAVs	4
Traffic in the scenario with 20% of CAVs was partially less smooth than in the scenario with 80% CAVs	5
Traffic in the scenario with 20% of CAVs was definitely less smooth than in the scenario with 80% CAVs	1
I did not perceive differences	2

TABLE 4: DATA RELATED TO SAFETY PERCEPTION

Regarding safety perception, which of the following statements do you agree with the most?	Number of answers
In the scenario with 20% of CAVs, I felt definitely safer than in the scenario with 80% CAVs	1
In the scenario with 20% of CAVs, I felt partially safer than in the scenario with 80% CAVs	2
In the scenario with 20% of CAVs, I felt partially less safe than in the scenario with 80% CAVs	3
In the scenario with 20% of CAVs, I felt definitely less safe than in the scenario with 80% CAVs	0
I did not perceive differences	6

TABLE 5: INFORMATION ON THE GLOBAL PREFERENCE OF THE DRIVERS

Globally, which of the 2 scenarios did you prefer?	Number of answers
I definitely preferred the scenario with 20% CAVs with respect to the scenario with 80% CAVs	1
I partially preferred the scenario with 20% CAVs with respect to the scenario with 80% CAVs	2
I partially preferred the scenario with 80% CAVs with respect to the scenario with 20% CAVs	4
I definitely preferred the scenario with 80% CAVs with respect to the scenario with 20% CAVs	2
I cannot say which scenario I preferred	3

As it can be observed in Table 3, there is a slight trend to prefer the scenario where 80% of traffic actors were automated cars. This might mean that an improvement in terms of traffic flow can be perceived by the people in case of increased fraction of CAVs. From the data collected in Table 4, it is possible to see how the feeling of safety is generally not affected by the market penetration rate of automated vehicles. In fact, most of the participants did not feel any difference from this point of view, comparing the situations with the two different percentages of automated cars. This can indicate that the increment in the amount of CAVs does not bother the human driver in his/her usual actions. Finally, the answers to the question reported in Table 5 give encouraging insights on the global interaction of traditional human-driven vehicles with automated ones. In fact, most of the people said to prefer the scenario where 80% of the cars were automated.

Obviously, more extended investigations would be needed, particularly more people should be involved in a statistical analysis able to provide reliable results. The results indicate just a possible trend in the preferences of the participants, but they cannot be considered a final assessment.

The study is intended to proceed further. The test can be improved considering more vehicles running into the roundabout and refining the behaviour policy defined by the RL algorithm.

5. CONCLUSIONS AND FUTURE WORK

The aim of the paper was to assess how the traffic of roundabouts could be organized in the future. Two mixed traffic scenarios were investigated, namely 20% or 80% presence of connected and automated vehicles (CAVs) was considered. The CAVs dynamic behavior into the roundabout was defined after an artificial intelligence (AI) algorithm was used, based on a multi-agent reinforcement learning (Q-learning) scheme. Such a scheme aims to reduce the time vehicles pass through the roundabout. The dynamic behavior policy has been defined by resorting to SUMO, an open software particularly suited to simulate traffic flows driven by AI policies. All of the cars running into the roundabout are assumed to share information on kinematics data (position, velocity and acceleration) via 5G communication. A vehicle-to-network-to-vehicle (V2N2V) innovative approach has been introduced. It is based on a single MEC/Edge node, positioned at the roundabout centre, and leveraging the features of 5G networks to guarantee the quality of service required in this context.

The theoretical simulations show that, beside the latency proper of 5G, the AI algorithm is able to improve the traffic into the roundabout either when the cars are 100% CAVs or 80% or 20%.

By means of a dynamic driving simulator, a real human driver has been put in the loop of the simulated traffic of the roundabout, as mentioned, 20% or 80% presence of connected and automated vehicles (CAVs) was considered. The human-in-the-loop is needed to check whether the driver feels comfortable and safe while driving in presence of automated cars.

Twelve people drove into the driving simulation environment. The preliminary results deserve an extensive statistical investigation for a final assessment. However, some initial results may be presented.

Drivers seem not worried when they interact with CAVs into the roundabout. They seem preferring to share the roundabout with CAVs as they seem feeling the traffic to be more fluent.

In this paper, roundabout traffic is studied utilizing different sensing inputs by simulation. The interesting results might contribute to the development of new intelligent transport system. In order to do that, some aspects could be further investigated in the future.

Thanks to the use of simulations, it is possible to guarantee the repeatability of the setting and to compare the results obtained. The weather may affect drivers' visibility. The influence of this aspect on test results could be investigated in future researches, reproducing different meteorological conditions.

Once the safety and the acceptability of CAVs behavior has been assessed, the policy can be tested on real cars in dedicated proving grounds. This could be a further step to validate the system, still reducing the risks connected to accidents and collisions.

ACKNOWLEDGEMENTS

Dott-ing L. Andreone from Stellantis is acknowledged for having inspired the use case research presented in this paper.

Dr. F. Comolli and Ing. A. Francesconi are gratefully acknowledged for having assisted during tests at the driving simulator at the Politecnico di Milano.

REFERENCES

[1] "The AI@EDGE H2020 Project", last modified December 31, 2021, accessed October 20, 2022, https://aiatedge.eu/.

[2] Filocamo, B., Ruiz, J. A., and Sotelo, M. A., 2019, "Efficient Management of Road Intersections for Automated Vehicles – The FRFP System Applied to the Various Types of Intersections and Roundabouts", *Applied Sciences*, 10(1): 316.

[3] Isebrands, H., 2009, “Crash Analysis of Roundabouts at High-Speed Rural Intersections”, *Transportation research record*, 2096(1): 1-7.
[4] Pratelli, A., Casella, S., Farina, A., and Lupi, M., 2018, “Conventional and unconventional roundabouts: a review of geometric features and capacity models”, *International Journal of Transport Development and Integration*, 2(3): 225-239.
[5] Hang, P., Huang, C., Hu, Z., Xing, Y., & Lv, C., 2021, “Decision making of connected automated vehicles at an unsignalized roundabout considering personalized driving behaviours”, *IEEE Transactions on Vehicular Technology*, 70(5), 4051-4064.
[6] García Cuenca, L., Puertas, E., Fernandez Andrés, J., & Aliane, N., 2019, “Autonomous driving in roundabout maneuvers using reinforcement learning with Q-learning”, *Electronics*, 8(12), 1536.
[7] Garcia Cuenca, L., Sanchez-Soriano, J., Puertas, E., Fernandez Andres, J., & Aliane, N., 2019, “Machine learning techniques for undertaking roundabouts in autonomous driving”, *Sensors*, 19(10), 2386.
[8] Rodrigues, M., McGordon, A., Gest, G., & Marco, J., 2018, “Autonomous navigation in interaction-based environments - A case of non-signalized roundabouts”, *IEEE Transactions on Intelligent Vehicles*, 3(4), 425-438.
[9] Martin-Gasulla, M., & Elefteriadou, L., 2021, “Traffic management with autonomous and connected vehicles at single-lane roundabouts”, *Transportation research part C: emerging technologies*, 125, 102964.
[10] Bakibillah, A. S. M., Kamal, M. A. S., & Tan, C. P., 2019, “The optimal coordination of connected and automated vehicles at roundabouts”, *2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)* (pp. 1392-1397). IEEE.
[11] Negash, N. M., & Yang, J., 2022, “Anticipation-based autonomous platoon control strategy with minimum parameter learning adaptive radial basis function neural network sliding mode control”, *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, *6*(10-06-03-0017), 247-265.
[12] Sutton, R. S., & Barto, A. G., 2018, *Reinforcement learning: An introduction*, MIT press.
[13] Bertsekas, D., & Tsitsiklis, J. N., 1996, *Neuro-dynamic programming*. Athena Scientific.
[14] Van der Pol, E., & Oliehoek, F. A., 2016, “Coordinated deep reinforcement learners for traffic light control”, *Proceedings of learning, inference and control of multi-agent systems (at NIPS 2016)*, *8*, 21-38.
[15] Gao, J., Shen, Y., Liu, J., Ito, M., & Shiratori, N., 2017, “Adaptive traffic signal control: Deep reinforcement learning algorithm with experience replay and target network”, *arXiv preprint arXiv:1705.02755*.
[16] Abdulhai, B., Pringle, R., & Karakoulas, G. J., 2003, “Reinforcement learning for true adaptive traffic signal control”, *Journal of Transportation Engineering*, *129*(3), 278-285.
[17] Li, L., Lv, Y., & Wang, F. Y., 2016, “Traffic signal timing via deep reinforcement learning”, *IEEE/CAA Journal of Automatica Sinica*, *3*(3), 247-254.
[18] El-Tantawy, S., Abdulhai, B., & Abdelgawad, H., 2013, “Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): methodology and large-scale application on downtown Toronto”, *IEEE transactions on Intelligent transportation systems*, *14*(3), 1140-1150.
[19] Liang, X., Du, X., Wang, G., & Han, Z., 2018, “Deep reinforcement learning for traffic light control in vehicular networks”, *arXiv preprint arXiv:1803.11115*.
[20] Wei, H., Zheng, G., Yao, H., & Li, Z., 2018, “Intellilight: A reinforcement learning approach for intelligent traffic light control”, *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2496-2505).
[21] Arel, I., Liu, C., Urbanik, T., & Kohls, A. G., 2010, “Reinforcement learning-based multi-agent system for network traffic signal control”, *IET Intelligent Transport Systems*, *4*(2), 128-135.
[22] Mousavi, S. S., Schukat, M., & Howley, E., 2017, “Traffic light control using deep policy-gradient and value-function-based reinforcement learning”, *IET Intelligent Transport Systems*, *11*(7), 417-423.
[23] Liang, X., Du, X., Wang, G., & Han, Z., 2018, “Deep reinforcement learning for traffic light control in vehicular networks”, *arXiv preprint arXiv:1803.11115*.
[24] Kuyer, L., Whiteson, S., Bakker, B., & Vlassis, N., 2008, “Multiagent reinforcement learning for urban traffic control using coordination graphs”, *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2008, Antwerp, Belgium, September 15-19, 2008, Proceedings, Part I 19* (pp. 656-671). Springer Berlin Heidelberg.
[25] Casas, N., 2017, “Deep deterministic policy gradient for urban traffic light control”, *arXiv preprint arXiv:1703.09035*.
[26] Genders, W., & Razavi, S., 2016, “Using a deep reinforcement learning agent for traffic signal control”, *arXiv preprint arXiv:1611.01142*.
[27] Capasso, A. P., Bacchiani, G., & Molinari, D., 2020, “Intelligent roundabout insertion using deep reinforcement learning”, *arXiv preprint arXiv:2001.00786*.
[28] Konstantinidis, F., Sackmann, M., De Candido, O., Hofmann, U., Thielecke, J., & Utschick, W., 2021, “Parameter Sharing Reinforcement Learning for Modeling Multi-Agent Driving Behavior in Roundabout Scenarios”, *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)* (pp. 1974-1981). IEEE.
[29] Watkins, C. J. C. H., 1989, “Learning from delayed rewards”
[30] Watkins, C. J. C. H., and Dayan, P., 1992, “Q-learning”, *Machine learning*, 8(3-4): 279-292.
[31] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M., 2013,

"Playing atari with deep reinforcement learning", *arXiv preprint arXiv:1312.5602*.
[32] Sutton, R. S., McAllester, D., Singh, S., & Mansour, Y., 1999, "Policy gradient methods for reinforcement learning with function approximation", *Advances in neural information processing systems*, *12*.
[33] Konda, V., & Tsitsiklis, J., 1999, "Actor-critic algorithms", *Advances in neural information processing systems*, *12*.
[34] "DriSMi: Driving Simulator Politecnico di Milano", last modified September 30, 2022, accessed October 20, 2022, https://www.drismi.polimi.it/?lang=it
[35] Previati, G., Mastinu, G., and Gobbi, M., 2022, "Influence of the Inertia Parameters on a Dynamic Driving Simulator Performances", *Society of Allied Weight Engineers 81st Annual Conference, Savannah, Georgia*.
[36] "Simulation of Urban MObility", last modified July 19, 2022, accessed October 20, 2022, https://www.eclipse.org/sumo/
[37] "VI-WorldSim", last modified September 30, 2022, accessed October 20, 2022, https://www.vi-grade.com/en/products/vi-worldsim/
[38] "Car-Following Models", last modified September 30, 2022, accessed October 20, 2022, https://sumo.dlr.de/docs/Definition_of_Vehicles%2C_Vehicle_Types%2C_and_Routes.html#car-following_models
[39] Peng, H., & Chen, X., 2022, "Active safety control of X-by-wire electric vehicles: A survey", *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, *6*(10-06-02-0008).
[40] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O., 2017, "Proximal policy optimization algorithms", *arXiv preprint arXiv:1707.06347*.
[41] Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P., 2015, "Trust region policy optimization", *International conference on machine learning* (pp. 1889-1897). PMLR.
[42] Kullback, S., & Leibler, R. A., 1951, "On information and sufficiency", *The annals of mathematical statistics*, 22(1), 79-86.