

# AVARS - Alleviating Unexpected Urban Road Traffic Congestion using UAVs

Jiaying Guo<sup>1</sup>, Michael R. Jones<sup>2</sup>, Soufiene Djahel<sup>3</sup>, and Shen Wang<sup>1</sup>

**Abstract**—Reducing unexpected urban traffic congestion caused by en-route events (e.g., road closures, car crashes, etc.) often requires fast and accurate reactions to choose the best-fit traffic signals. Traditional traffic light control systems, such as SCATS and SCOOT, are not efficient as their traffic data provided by induction loops has a low update frequency (i.e., longer than 1 minute). Moreover, the traffic light signal plans used by these systems are selected from a limited set of candidate plans pre-programmed prior to unexpected events' occurrence. Recent research demonstrates that camera-based traffic light systems controlled by deep reinforcement learning (DRL) algorithms are more effective in reducing traffic congestion, in which the cameras can provide high-frequency high-resolution traffic data. However, these systems are costly to deploy in big cities due to the excessive potential upgrades required to road infrastructure. In this paper, we argue that Unmanned Aerial Vehicles (UAVs) can play a crucial role in dealing with unexpected traffic congestion because UAVs with onboard cameras can be economically deployed when and where unexpected congestion occurs. Then, we propose a system called "AVARS" that explores the potential of using UAVs to reduce unexpected urban traffic congestion using DRL-based traffic light signal control. This approach is validated on a widely used open-source traffic simulator with practical UAV settings, including its traffic monitoring ranges and battery lifetime. Our simulation results show that AVARS can effectively recover the unexpected traffic congestion in Dublin, Ireland, back to its original uncongested level within the typical battery life duration of a UAV.

**Index Terms**—UAVs, Deep Reinforcement Learning, Traffic Light Control, Unexpected Congestion

## I. INTRODUCTION

Many governments, policy makers, tech giants and SMEs are joining their efforts to support the vision towards achieving Net-Zero transport by 2050 (e.g., in the UK & Ireland). This requires revolutionary solutions, supported by adequate policy changes, to a number of problems facing the wide adoption and/or deployment of cutting-edge technologies such as transport electrification, and the use of drones to offload ground transport, etc. Achieving the above objective also requires innovative solutions to better control traffic congestion, especially after the occurrence of unexpected events (e.g., car crashes and unplanned road work etc.) on the road. Unexpected (aka "non-recurrent") urban road traffic congestion is

caused by such en-route events. It is challenging to accurately predict the traffic impact of such events at any possible time and location due to the lack of such historical data. Therefore, an effective approach to reduce such congestion should react quickly after the event occurrence, rather than predicting the event impact in advance. This fast reaction is effective in reducing additional delays experienced by drivers and the resulting additional emissions.

UAVs are renowned for their rapid deployment and efficiency in supporting emergency management (e.g., life-critical services [1], situations where the communication infrastructure is destroyed, subject to cyber attacks or overloaded [2], [3]). UAVs have also been used for road traffic monitoring on urban and highway roads [4]. However, the use of UAVs for emergency management in the transportation domain (e.g., reducing the impact of non-recurrent congestion), has not been sufficiently studied in the literature, thus we argue in this paper that UAVs have a great potential to be explored in this context. Additionally, although existing adaptive traffic light control systems (e.g., SCATS [5]) work well for the recurrent urban traffic, they cannot be effective for non-recurrent congestion due to their low-frequency and coarse-grained traffic monitoring detectors (i.e., induction loops collect traffic volumes in minutes). Recent DRL-based traffic light signal control can be promising [6]. However, the convergence of the DRL solution remains uncertain when applied in a fast-changing environment. Moreover, the practical deployment may need excessive hardware and software to upgrade existing systems (i.e., install cameras to all major road intersections and deploy DRL-based software on the road side units and central servers). This makes it unnecessary for non-recurrent congestion reduction which only happens very occasionally. To fill the above-mentioned gaps, we propose a system, dubbed AVARS, which is also open-sourced<sup>1</sup> and its main contributions are summarized as follows:

- **AVARS uses UAVs to reduce non-recurrent congestion, rather than just monitoring normal traffic.** To the best of our knowledge, this is the first work that leverages UAVs to reduce non-recurrent congestion in urban traffic, while most others are for traffic monitoring, which we think is not practical as UAVs can only sustain for less than one hour compared to 24/7 deployed cameras in the fixed location. UAVs can be dispatched to congested intersections in a few minutes and provide high-frequency high-resolution data [7], which are well-suited features

<sup>1</sup>Jiaying Guo and Shen Wang are with the School of Computer Science, University College Dublin, Ireland, E-mail: jiaying.guo@ucdconnect.ie, shen.wang@ucd.ie

<sup>2</sup>Michael R. Jones is with the Department of Computing and Mathematics, Manchester Metropolitan University, UK, E-mail: michael.jones@mmu.ac.uk

<sup>3</sup>Soufiene Djahel is with the Department of Computer Science, University of Huddersfield, UK, E-mail: s.djahel@hud.ac.uk

<sup>1</sup><https://github.com/Guojoyy/AVARS>

for recovering from non-recurrent road congestion within a short duration.

- **AVARS is a DRL-based system that contains many practical deployment considerations.** AVARS is carefully designed to achieve the minimum possible upgrade to the existing systems. This includes embedding low-cost wireless communication modules for traffic light controllers only (not vehicles). Thus, it avoids strong assumptions about having full or mix-autonomy traffic with vehicle-to-everything (V2X) communication infrastructures. AVARS also has a practical DRL model (i.e., newly designed state, action, and reward) as it can obtain a more stable convergence under fast-changing road traffic.
- **The effectiveness of AVARS is validated on a widely-used simulator with a realistic urban scenario.** The simulation study shows that AVARS can achieve the largest reduction in travel time, fuel consumption, and CO<sub>2</sub> emission, when unexpected congestion occurs in a city center area of Dublin, Ireland. Compared methods include SCATS [5] (no DRL, no UAV), and IntelliLight [6](DRL but no UAV). Additionally, AVARS can recover non-recurrent congestion back to the original traffic within the 30-minute operation of UAVs, which is well within the normal UAV's battery life. If considering more realistic factors impacting battery duration, only 10-minute traffic signal control by AVARS is sufficient to mitigate congestion significantly.

## II. AVARS - SYSTEM DESCRIPTION

### A. Assumptions

**Communications:** The communications required in our AVARS are only between a UAV and a traffic light controller (i.e., vehicles are not involved) that occur every second (i.e., to be responsive to en-route events). The communication technologies used could be any easy to implement ones such as 4/5G or IEEE 802.11p. In particular, the UAV send the "action" to traffic lights to either switch to the next phase or keep the current phase. Conversely, the traffic light controller can send current traffic light phase back to the UAV as part of current observations (states). This is for the UAV to decide about the recommended action in the next time interval. We assume the communication is entirely reliable without any packet loss and delay.

**Computations:** Most computation loads are in the centralized Traffic Management Center (TMC). Firstly, TMC uses simulators to mimic possible road closures to collect data on the traffic impact. This is necessary for AVARS to select which intersections (i.e. likely the most badly influenced by the events) that UAVs must control after the occurrence of road events. Secondly, the pre-training of DRL models also happens in this simulation. Thirdly, after each assigned UAV completes its task, it brings back to TMC the realistic DRL trajectory data (i.e., state, action, reward, next state...) when UAV is operating AVARS. The trajectory data is also used for pre-training our DRL models. In addition to the computation loads at the TMC

side, UAVs also need to do the following two computation tasks: extract the traffic information from their cameras; and control the traffic light controllers using a pre-trained DRL policy function with the real-time state information being the input.

**UAV Placement:** Upon the receipt of an intersection location (i.e., where traffic events occurred) from the TMC, it is assumed that a UAV can independently transit both safely and efficiently across the environment space. Upon arrival at the intersection location the UAV can maintain both its position and altitude, to deliver a visual representation of the intersection allowing for traffic information extraction.

### B. Design objectives

We design AVARS to meet the following objectives:

- **Be practical:** The system should have a practical deployment plan to meet the limitations of UAVs and DRL approaches. The cameras of UAVs have limited monitoring range. This implicates the maximum length of the road for UAV traffic monitoring should be around 220 meters [7]. Moreover, the limited battery life of a UAV requires that the congestion should be effectively alleviated before the energy runs out (e.g., within 40 minutes since a UAV starts to operate [8]). Additionally, the practical requirement of DRL refers to a stable convergence performance within reasonable amount of iterations. A stable DRL approach can potentially save computation resources and be effective in more general scenarios.
- **Be effective:** The system should be effective to recover the unexpectedly congested traffic within a limited time duration (i.e., UAV battery lifetime). The design of DRL models in terms of state, action, and reward is of vital importance to achieve this objective. The verification of this objective will be done by comparing the traffic evolution when no road closure occurs, as well as the reduction in travel time, fuel consumption, and CO<sub>2</sub> emissions.

### C. System components

We clarify the responsibilities of each different AVARS component as follows:

- **UAVs:** Each UAV is responsible for data collection (i.e., DRL state for decision making and DRL trajectories for pre-training) and sending actions (i.e., switching or not to the next traffic light phase) according to the learned DRL policy function for controlling a single intersection.
- **Traffic lights:** Each traffic light only needs to receive the DRL action from a UAV to set its phase in the next time interval, and reports back to this UAV about the current phase as a part of DRL state.
- **TMC:** Based on numerous simulation results, TMC needs to select a number of key intersections for UAVs to control when unexpected events occur. TMC also needs to train the proposed DRL model using DRL trajectories collected after UAVs complete their AVARS field tasks.

Unlike most of existing works in the literature [9], our AVARS does not rely on any vehicle equipped with advanced technologies such as autonomous or connected vehicles. This saves a large potential cost to upgrade existing infrastructure and removes strong assumptions on collecting vehicle driving information in recent DRL-based traffic light signal control approaches [10].

#### D. System flow

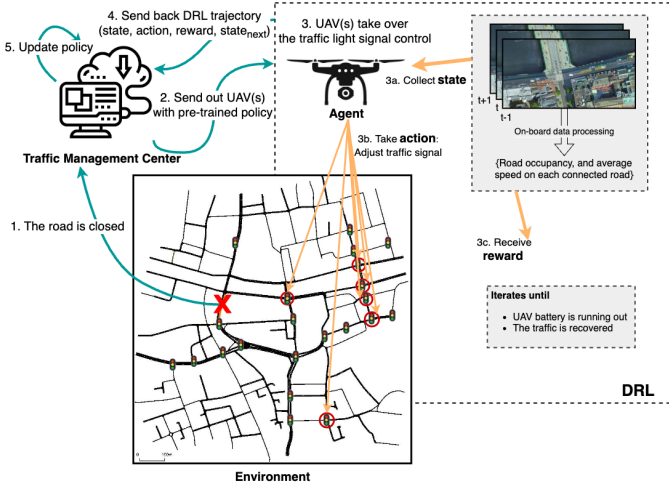


Fig. 1. Illustration of AVARS system flow. The traffic environment is a subset of Dublin city center road network. The highlighted road is the closed road causing unexpected congestion in the surrounding area. UAVs of AVARS control the six red-circled intersections in this scenario, which are influenced most by the road closure.

Suppose a road segment within an urban road network is suddenly closed, due to an incident, leading to queues of vehicles, reduced traffic speed and several affected vehicles needing to reroute to bypass the closed road. This will migrate the traffic demand to several other signalized intersections, which causes unexpected traffic congestion. To efficiently handle such scenarios, AVARS operates as follows (also shown in Fig.1):

- 1) Once a road is closed, the TMC is informed through existing traffic monitoring technologies.
- 2) The TMC suggests several signalized intersections for UAVs to control (six intersections highlighted by red circles in the environment of Fig.1). Each UAV is assigned to a designated intersection. For each selected intersection, all of its directly connected roads should be covered by the monitoring range of the UAV camera.
- 3) When UAVs arrive at their assigned intersections, each UAV starts: a) collecting the traffic information, then converting them to the “DRL state”; b) sending the “DRL action” to adjust the traffic signal according to the pre-trained DRL policy function; and c) calculating the “DRL reward” based on real-time traffic information. These 3 steps iterate every second until the preset operation duration is reached or the UAV enters a low battery state, both triggering a UAV to return to the TMC. Each

UAV returns to TMC with the collected DRL trajectory (state, action, reward, new state) acquired during this operation interval.

- 4) The DRL policy function is updated using newly collected DRL trajectory data on the TMC cloud servers.

#### E. DRL models

- **State:** For a specific intersection that is controlled by a UAV, the state involves the current signal phase of the traffic lights, the occupancy (i.e., the ratio of the total length of vehicles on a specific road to the length of that road) and average vehicle speed on each of the roads directly linked to this intersection. This state information is collected from the UAV’s sensors (e.g., camera) and extracted after onboard data processing [11].
- **Action:** The action that each UAV sends to its controlled traffic light is a binary set, which defines that “1” represents switching to the next phase and “0” means to keep the current phase unchanged for the next time interval.
- **Reward:** Given the DRL state and action at the intersection  $r_j$ , as described in Eq.1, the reward is calculated as the negative value of maximum road occupancy among all incoming roads  $I$  connected to the intersection  $j$ . The reward aims to evacuate traffic on the road that has the highest occupancy as soon as possible.

$$r_j = -\max(O_i), i \in I \quad (1)$$

The road occupancy  $O_i$  for each incoming road is also involved in the UAV state.

To get a stable convergence with a moderate number of iterations, we do not recommend any multi-agent DRL algorithm for AVARS so far [9]. Algorithm 1 shows one completed iteration of AVARS training process to obtain the DRL policy function taken by UAVs to control traffic lights, in which the DRL algorithm can be either Deep Q-learning (DQN) or Proximal Policy Optimization (PPO) [12]. DQN is commonly used in existing traffic light signal control systems [6], [10], having advantages in discrete action space, i.e., a binary set of traffic signal switching in AVARS. The off-policy method is efficient to update the finite Q values on a one-step temporal difference for a higher reward. However, it is also volatile in a fast-changing environment, especially with complex traffic information, which aggravates training convergence. Unlike DQN, PPO can reduce the large variance of policy updates and facilitate training convergence. PPO can also improve sample efficiency by multiple epochs of mini-batch updates.

#### F. Summary

We summarize how the design of AVARS achieves our first objective - *be practical*. Firstly, AVARS avoids large-scale upgrades of hundreds of traffic light controllers by quickly assigning UAVs to very few intersections that are most impacted by the en-route events. Secondly, AVARS uses UAVs’ sensors to obtain real-time traffic information, which is much more flexible than embedding induction loops on the road. Thirdly, AVARS does not consider vehicle-related

---

**Algorithm 1** One DRL training iteration of AVARS

---

**Input:** A set of intersections  $E$  to control

**Output:** Policy function parameter  $\theta$

- 1: Initialize the urban road environment in  $T$  time horizons constrained by UAV battery lifetime
  - 2: Initialize Q-network  $Q(\theta)$  if DQN or policy network  $\pi(\theta)$  if PPO
  - 3: **for** each  $E$  **do in parallel**
  - 4:   UAV arrives at the designated intersection
  - 5:   **for**  $t \leq T$  **do**
  - 6:     Get observation  $s_t$
  - 7:     Execute action  $a_t$  selected through  $Q(s_t, a; \theta)$  or  $\pi(a|s_t; \theta)$ ,  $a \in \{0, 1\}$  and transit to  $s_{t+1}$
  - 8:     Calculate reward  $r_t$  using Eq.1
  - 9:     Collect trajectory  $\tau = (s_t, a_t, r_t, s_{t+1})$  and extend  $B$  with  $\tau$
  - 10:   **end for**
  - 11: **end for**
  - 12: Sample a random batch from  $B$  to update the policy parameter  $\theta$
- 

upgrades (e.g., V2X communication capabilities) and requires minimum upgrades of the traffic lights (e.g., 5G or IEEE 802.11p communication technology) to communicate with UAVs instead of the costly wired cable communications used in today's adaptive traffic light systems. Last but not least, our DRL model is simple enough to calculate. Additionally, the heavy-weighted computation for training DRL only occurs in the centralized TMC, which is practical to implement using existing cloud-based services, rather than distributing these computation loads to multiple roadside units or resource-constrained UAVs.

### III. EVALUATION RESULTS AND ANALYSIS

#### A. Unexpected urban traffic congestion scenarios

The implementation of AVARS is based on FLOW<sup>2</sup>, which is a framework integrating RLlib (DRL library) and SUMO (traffic simulator). The testing map is a subset of Dublin city center road network around the River Liffey, covering approximately 1 square kilometer, as shown in Fig.1. The scenario is extracted from the open data in [13] to simulate the real-world traffic in Dublin city, and the traffic generation lasts 45 minutes with 1168 vehicles in our experiments. We close a road in the center of the testing scenarios for 30 minutes from the 10th minute of the simulation, at which point the vehicles that would have entered the closed road are required to reroute. The radius of the area managed by the TMC at Dublin city center is 5km, given the maximum flying speed of 80 km/h [11], a UAV is capable of direct flight to any intersection within the control region in 3.75 minutes. However, the direct flight may be limited by the urban environment, hence we

<sup>2</sup><https://flow-project.github.io>

assume the start time of UAV control after road closure is 5 minutes.

#### B. Compared scenarios

- **Original:** This scenario contains the regular traffic for a given urban region. It is a reference to verify if the unexpected congestion is recovered. The traffic light signals are static in the scenario, operating on a fixed cycle plan. The duration of a signal plan ranges from 90 seconds to 113 seconds for the given heterogeneous intersections. Green light phases last from 27 seconds to 50 seconds. Every time the green light switches, a 3-second yellow is followed to clear the intersection, and some signal plans combine a 5-second all-red light after the yellow light.
- **Congestion:** This scenario contains unexpected congestion due to one road closure for the same urban region. The traffic light signals are static, the same as the settings in the Original scenario.
- **SCATS [5]:** This scenario contains the simulation of the existing adaptive traffic light controller, SCATS [5], which neither use DRL nor UAVs. The embedded induction loops detect the duration of vehicles passing through the intersection during the green light phase, which can estimate the effective green time. The sum of the ratio of effective green time for each green light phase is the degree of saturation of one signal plan. SCATS provides multiple pre-defined signal plans and selects the signal plan with the minimum degree of saturation [9].
- **IntelliLight [6]:** This scenario contains the DRL-based traffic light signal control model using DQN but not using UAVs. Compared with AVARS, more complicated traffic metrics, such as queue length, number of vehicles, and waiting time of vehicles, are included in state and reward design. Besides, the state is also expanded with vehicles' positions and traffic light phases collected from the image representation. IntelliLight controls the signal switching to reduce queue length and travel delays.

#### C. Result analysis

##### Is DRL training of AVARS more stable to converge?

As shown in Fig.2, after 150 iterations, the training of AVARS(PPO) terminates at a higher average episode reward and smaller standard deviation (approx. -250 and 99) than the initial stage (approx. -766 and 430), which are calculated over 18 parallel episodes for each iteration. It also demonstrates the continuous positive updates of the trained policy. The DQN training of AVARS gets similar results, while the policy update of DQN is much faster, resulting in earlier convergence. Instead, the reward improvement of IntelliLight shown in Fig.3 varies widely across different DRL algorithms. PPO can get a better policy with a higher average episode reward (about -2860) after training compared to the DQN reward (about -4050). However, the standard deviation of either PPO or DQN during training is large, probably due to the complex design of states and rewards to evaluate

traffic. Meanwhile, the complex design of IntelliLight(DQN) results in the training time being approximately three times longer than IntelliLight(PPO). Hence, the benefit of AVARS in terms of states and reward design can also be demonstrated by the above experiments on AVARS and IntelliLight using DQN or PPO. AVARS, with the concise state and reward design, can take advantage of DQN or PPO to get an expected policy to recover unexpectedly congested traffic, which is in line with the second system design objective - *effectiveness*. Additionally, we would simplify the following experiments with only AVARS(PPO) and IntelliLight(DQN) because PPO is more stable than DQN in traffic improvement based on IntelliLight, while DQN is the default DRL algorithm for IntelliLight.

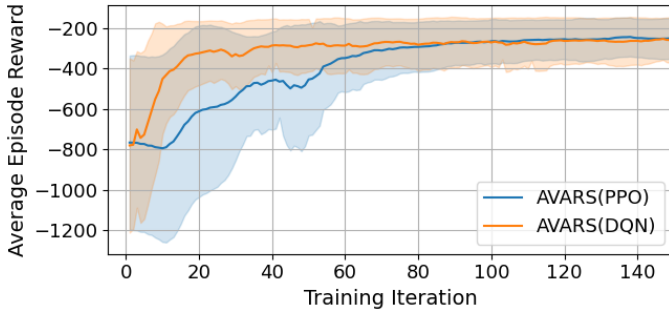


Fig. 2. Evolution of the average episode reward during the training process of AVARS, computed over 18 parallel episodes at each iteration. As for AVARS using PPO, the reward ranges from approx. -766 to -250, and the standard deviation decreases from about 430 to 99.

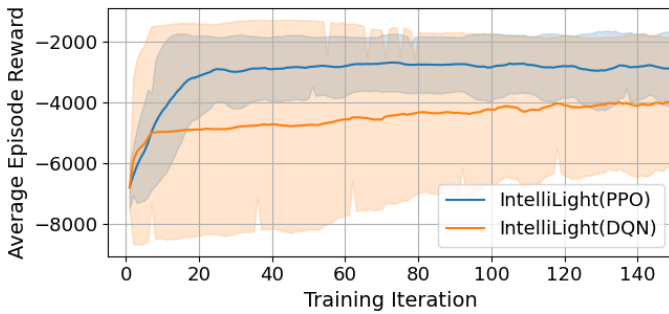


Fig. 3. Evolution of the average episode reward during the training process of IntelliLight, computed over 18 parallel episodes at each iteration. The reward of IntelliLight(DQN) ranges from approx. -6800 to -4050, and the standard deviation decreases from about 3400 to 2200.

**Is AVARS effective?** Vehicles generated for the 45 minutes (i.e., 2700 timesteps), so the number of vehicles running in the simulated region tends to gradually increase during this time interval, as illustrated in the Original scenario of Fig.4. The road closure (the Congestion scenario) has resulted in a sharp increase in running vehicles, up to more than 250. However, after UAVs effectively control the selected intersections for approximately 10 minutes, the number of running vehicles decreases to a level lower than the Original scenario at the same timestep. The largest difference between the Original scenario and our proposed method is 61 vehicles around

the 2500th timestep. SCATS cannot evidently mitigate the congested traffic, and in this scenario, the number of running vehicles shows only minor deviations from the Congestion scenario. The results of IntelliLight are closely aligned to that of the Original scenario. Hence, Fig.4 illustrates that AVARS can alleviate the unexpected congestion.

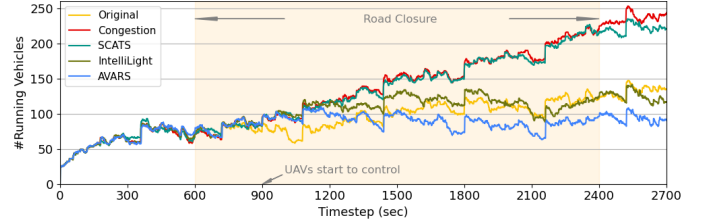


Fig. 4. Evolution of the number of running vehicles in the chosen urban region over the simulated timesteps of an episode under AVARS and other compared scenarios. Road closure starts from the 600th timesteps and lasts 30 mins. UAVs used in AVARS start to control traffic light signals from the 900th timesteps.

Furthermore, we analyze the main traffic statistics including average travel time, fuel consumption, and CO<sub>2</sub> emissions (i.e., we refer to these metrics as “main traffic metrics” in the following descriptions), as shown in Table.I. The results are the average of 10 episodes that terminated after all vehicles completed the trip. Our proposed system AVARS produces the best results. Even compared to the Original scenario, where no road closure is in place, all traffic metrics are significantly reduced. By contrast, the negative impact of one critical road closure is evident, with the value of these main traffic metrics more than doubling up to 126%. Congestion causes massive delays, leading to longer travel times.

All traffic light signal control methods, SCATS, IntelliLight, and AVARS, can alleviate the unexpected traffic congestion. Although SCATS reduces about 30% on these main traffic metrics over the Congestion scenario, the results are still inferior to that of the Original scenario. The results show that SCATS is not efficient in alleviating unexpected traffic congestion. In comparison, both DRL-based methods achieve a substantial improvement in average travel time, fuel consumption, and CO<sub>2</sub> emissions, while our system AVARS achieves the largest reduction, 67%, 72%, and 72%, respectively.

The travel time statistics for all vehicles also demonstrates the benefit of our system for the vehicles influenced by the unexpected congestion, as shown in Table.II. These vehicles have experienced more prolonged travel times due to the congestion. Except for the shortest average travel time, the smallest standard deviation of AVARS demonstrates a general reduction in vehicle congestion delays. More than half of vehicles can finish their trips in about 200 seconds, which is better than the Original scenario. In addition, the longest travel time is decreased from 4599 seconds in the Congestion scenario to 554 seconds. Vehicles can reach their destination earlier, even when travelling longer distances due to rerouting.

**How fast can AVARS alleviate the congestion?** Time periods of 10, 20 and 30 minutes are selected as operational

TABLE I

MAIN TRAFFIC STATISTICS. EXCEPT FOR THE ORIGINAL SCENARIO, THERE IS A ROAD CLOSURE UNDER ALL OTHER SCENARIOS. THE ACHIEVED TRAFFIC IMPROVEMENTS OF SCATS, INTELLILIGHT, AND AVARS ARE BASED ON THE CONGESTION SCENARIO.

	Average travel time (s)	Fuel consumption (l/100km)	CO <sub>2</sub> emissions (g/km)
Original	275.78	20.14	468.62
Congestion	597.31	45.45	1057.20
SCATS	429.61 -28.08%	31.69 -30.28%	737.10 -30.28%
IntelliLight	232.67 -61.05%	14.60 -67.88%	339.55 -67.88%
AVARS	195.42 -67.28%	12.31 -72.92%	286.29 -72.92%

TABLE II

TRAVEL TIME STATISTICS (AVERAGE, 25% PERCENTILE, MEDIAN, 75% PERCENTILE, MINIMUM, MAXIMUM, AND STANDARD DEVIATION) FOR ALL VEHICLES ACROSS ALL COMPARED SCENARIOS.

	Original	Congestion	SCATS	IntelliLight	AVARS
Avg	275.78	597.31	429.61	232.67	195.42
Std	208.83	965.67	614.66	148.08	108.86
Min	2.60	2.50	1.00	2.60	1.00
1st Qtr	127.40	129.52	119.00	119.53	109.55
Med	230.70	215.90	209.00	206.70	193.35
3rd Qtr	387.78	342.48	333.00	304.85	261.00
Max	1365.20	4599.10	3059.00	743.00	553.70

AVARS durations, all shorter than 40 minutes, which represent the typical UAV battery lifetime [8]. Fig.5 suggests that although the longer the UAV operates, the better traffic improvement can be achieved, just 10 to 20 minutes of UAV operation in AVARS should be sufficient to recover the congestion back to its original state.

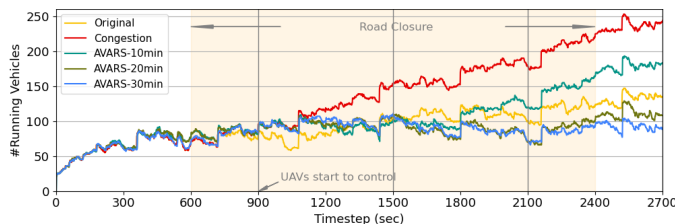


Fig. 5. Evolution of the number of running vehicles in the chosen urban region over the simulated timesteps of an episode under different UAV operation durations of AVARS compared with the Original and Congestion scenarios. UAVs used in AVARS start to control traffic light signals from the 900th timestep and terminate at the 1500th timestep (10 minutes), the 2100th timestep (20 minutes), and the 2700th timesteps (30 minutes), respectively.

## IV. CONCLUSION AND FUTURE WORK

This paper proposes AVARS: a DRL-based traffic light control system using UAVs to reduce unexpected urban road traffic congestion. AVARS takes advantage of UAVs' rapid deployment, and rich traffic monitoring capabilities to timely react to these events. AVARS also supports state-of-the-art DRL methods to deliver an efficient control of traffic light signals, without the need for the costly upgrade of traffic light controllers. We demonstrated that AVARS can effectively alleviate congestion and achieve notable reductions in travel time, fuel consumption, and CO<sub>2</sub> emissions in an urban scenario. Future works may include designing an extended version of AVARS in which multiple UAVs cooperate using multi-agent DRL while addressing its high complexity of cooperation and convergence issues.

## REFERENCES

- [1] J. Bauer, D. Moormann, R. Strametz, and D. A. Groneberg, "Development of unmanned aerial vehicle (UAV) networks delivering early defibrillation for out-of-hospital cardiac arrests (OHCA) in areas lacking timely access to emergency medical services (EMS) in Germany: a comparative economic study," *BMJ open*, vol. 11, no. 1, p. e043791, 2021.
- [2] X. Zhang and L. Duan, "Fast deployment of UAV networks for optimal wireless coverage," *IEEE Transactions on Mobile Computing*, vol. 18, no. 3, pp. 588–601, 2018.
- [3] R. Masroor, M. Naeem, and W. Ejaz, "Efficient deployment of UAVs for disaster management: a multi-criterion optimization approach," *Computer Communications*, vol. 177, pp. 185–194, 2021.
- [4] S. Liu, S. Wang, W. Shi, H. Liu, Z. Li, and T. Mao, "Vehicle tracking by detection in UAV aerial video," *Science China Information Sciences*, vol. 62, no. 2, pp. 1–3, 2019.
- [5] P. Lowrie, "SCATS, sydney co-ordinated adaptive traffic system: A traffic responsive method of controlling urban traffic," 1990.
- [6] H. Wei, G. Zheng, H. Yao, and Z. Li, "Intellilight: A reinforcement learning approach for intelligent traffic light control," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 2496–2505.
- [7] K. Park and R. Ewing, "The usability of Unmanned Aerial Vehicles (UAVs) for pedestrian observation," *Journal of Planning Education and Research*, p. 0739456X18805154, 2018.
- [8] S. R. Hashemi, R. Esmaeeli, H. Aliniagerdroudbari, M. Alhadri, H. Alshammari, A. Mahajan, and S. Farhad, "New intelligent battery management system for drones," in *ASME international mechanical engineering congress and exposition*, vol. 59438. American Society of Mechanical Engineers, 2019, p. V006T06A028.
- [9] H. Wei, G. Zheng, V. Gayah, and Z. Li, "A survey on traffic signal control methods," *arXiv preprint arXiv:1904.08117*, 2019.
- [10] T. Wang, J. Cao, and A. Hussain, "Adaptive traffic signal control for large-scale scenario with cooperative group-based multi-agent reinforcement learning," *Transportation research part C: emerging technologies*, vol. 125, p. 103046, 2021.
- [11] M. A. Khan, W. Ectors, T. Bellemans, D. Janssens, and G. Wets, "Unmanned aerial vehicle-based traffic analysis: A case study for shockwave identification and flow parameters estimation at signalized intersections," *Remote Sensing*, vol. 10, no. 3, p. 458, 2018.
- [12] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [13] M. Guérliau and I. Dusparic, "Quantifying the impact of connected and autonomous vehicles on traffic efficiency and safety in mixed traffic," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020, pp. 1–8.