2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall), 10-13 October 2023, Hong Kong
DOI:10.1109/VTC2023-Fall60731.2023.10333535

Fragility Impact of RL Based Advanced Air Mobility under Gradient Attacks and Packet Drop Constraints

Deepak Kumar Panda[1] and Weisi Guo[1,2]

***Abstract*— The increasing utilization of unmanned aerial vehicles (UAVs) in advanced air mobility (AAM) necessitates highly automated conflict resolution and collision avoidance strategies. Consequently, reinforcement learning (RL) algorithms have gained popularity in addressing conflict resolution strategies among UAVs. However, increasing digitization introduces challenges related to packet drop constraints and various adversarial cyber threats, rendering AAM fragile. Adversaries can introduce perturbations into the system states, reducing the efficacy of learning algorithms. Therefore, it is crucial to systematically investigate the impact of increased digitization, including adversarial cyber-threats and packet drop constraints to study the fragile characteristics of AAM infrastructure. This study examines the performance of artificial intelligence(AI) based path planning and conflict resolution strategies under different adversarial and stochastic packet drop constraints in UAV systems. The fragility analysis focuses on the number of conflicts, collisions and fuel consumption of the UAVs with respect to its mission, considering various adversarial attacks and packet drop constraint scenarios. The safe deep q-networks (DQN) architecture is utilized to navigate the UAVs, mitigating the adversarial threats and is benchmarked with vanilla DQN using the necessary metrics. The findings are a foundation for investigating the necessary modification of learning paradigms to develop antifragile strategies against emerging adversarial threats.**

I. Introduction

The civil aviation infrastructure has experienced a significant annual increase of 1.5% [1], [2]. Furthermore, low-altitude airspace as a part of AAM is witnessing continuous growth in the presence of smaller autonomous aircraft [3]. Free flight for a UAV in AAM offers operational flexibility in airspace, along with the enhancement of safety [4] and fuel efficiency [5], [6]. By deploying sophisticated AI algorithms, automated decentralized conflict resolution mechanisms can alleviate the burden on air traffic controllers to realize the futuristic vision of AAM.

The autonomous AAM necessitates real-time decision-making to resolve conflicts while upholding safety and mission requirements. Deep reinforcement learning (DRL) algorithms offer a promising alternative, as they can handle real-time uncertainty and dynamic interactions among autonomous aircraft during conflict resolution [7], while leveraging real-time navigation information [8]. However, DRL is vulnerable to real-time data integrity issues arising from adversarial disturbances and communication constraints. Attackers, who employ AI techniques [9], increasingly inject adversarial disturbances into decision-making capabilities to maximize impact while minimizing detection. These disturbances can lead to performance degradation in DRL systems, causing suboptimal actions by agents [10]. The review of the various threat models along with its varied impact has been reviewed in Table I. As stated in Table I, various aspects of the UAVs have been affected by the adversarial threats which includes targeted tracking [11], collecting data from ground nodes [12] and offloading policy [13]. Communication constraints also affect the mission requirements of the UAVs [14], particularly in decentralized algorithms where knowledge of nearby UAV positions relies on vehicle-to-vehicle (V-V) communication links [15]. The effect of packet drops on UAV swarms have been studied in [16], [17] where the Bernoulli, Gilbert and extended Gilbert models have been used to model the lossy conditions leading to packet drop constraints.

This paper investigates the vulnerability of the AAM to adversarial attacks and packet drop constraints. The study focuses on the impact of fragility, encompassing various factors such as fuel requirements, conflicts, and collisions with other UAVs. The simulation setup builds upon previous work [24], extending it to include adversarial attacks and communication constraints. The fragility testing involves evaluating the AAM system against emerging adversarial threats, with a particular emphasis on countering the momentum iterative (MIM) attack studied in [24], which targets the communication channel selection of air traffic control (ATC). In this context, a safe RL (Reinforcement Learning) architecture, previously designed to counter the fast gradient sign method (FGSM) [24], is tested for its resilience against the MIM attack. Furthermore, the resilience of the AAM infrastructure is studied concerning packet drop rate models, drawing comparisons with the studies presented in [16] where Bernoulli packet drop rates were considered. The research goes on to benchmark different packet drop rate scenarios across various loss channels incorporating correlated packet drop models, using two and three-state Markov chains [17]. Unlike the approach in [24], which primarily analyzed metrics using mean values, this paper introduces a novel univariate distribution metric. The simulation involves randomly generated targets and centralized control UAVs, and it is compared against vanilla DRL architecture. The contribution of this paper can be summarized as follows:

- Implementation of randomized MIM attack and Gilbert

*This work was not supported by any external funding

[1]The authors are with School of Aerospace, Transport, and Manufacturing, Cranfield University, MK43 0AL Cranfield, U.K Deepak.Panda@cranfield.ac.uk, weisi.guo@cranfield.ac.uk

[2]Weisi Guo is with The Alan Turing Institute, NW1 2DB London, U.K : weisi.guo@cranfield.ac.uk

TABLE I

THE BRIEF REVIEW OF THE ADVERSARIAL ATTACKS ON UAV AND THEIR IMPACT

References	Threat Models	Impact
[11]	Continuous disturbance signal with Gaussian noise	Inability to estimate the flight path of the targeted UAV
[18]	FGSM and BIM	Collision risk and reaching goals
[19]	Pixel level attack and semantic perturbation	Object detection for UAV navigation
[13]	Triggerless backdoor attack on the model parameter	UAV offloading policy
[20]	Adding imperceptible perturbations to the image	Targeted UAV tracking
[21]	Adversarial attack based on forward derivative and optimization	Navigation and control of UAV
[22]	Manipulate and control the input channel of the sensor.	Collision with the obstacles.
[12]	Jamming the input channel	Reduce the signal-to-noise ratio (SNR)
[23]	Time delay attack by delaying the transmission of data packets	Military reconnaissance and strike mission.

and extended Gilbert model on RL based AAM.

- Fragility impact study of AAM with respect to safe DQN under various adversarial attacks and packet drop models.
- Benchmarking the results with respect to vanilla DQN architecture.

II. METHODS

There are two modes of control for UAVs: free flights and regulated flights, as shown in Figure 1. The centralized controller guides UAVs along a predetermined trajectory and is used for infrastructure inspection and monitoring [25], [26]. On the other hand, free-flight UAVs, used for last-mile delivery [27], autonomously detect their trajectory while considering fuel levels and avoiding collisions with other UAVs. The simulation setup is based on [24].

The state variables Φ represent navigational information of the free flight Φ_{free} and the nearest intruders Φ_{int}. The state information is corrupted with packet drops and adversarial disturbance. The effect of state disturbance, is reflected in the performance of the free-flight UAV based on the number of steps it takes to reach the target and the number of conflicts and collisions with the neighboring UAVs. The actions of the free-flight UAV are generated from a DNN architecture as a part of DQN. The second architecture tested here is the safe DQN architecture explained in [24]. The DQN models used in this paper have been trained in [24]. This paper tests the trained model considering disturbed states via several attack and packet drop constraints. The deep RL has been trained to maximize the reward, the minimum number of steps to reach the target and minimize the penalty of conflict and collision. The performance of the DRL architecture has been analyzed concerning the probability distribution of the number of steps to reach the target and the number of conflicts between the free-flight UAV and other UAVs. The following subsection will explain the dynamics of the free flight UAV used in this paper as a part of the environment.

A. Environment

The UAV dynamics are based on [28] and simulated in a 2D environment. In this simulation, (x,y) represents the UAV's position, v represents velocity and ϕ represents the heading angle and their relationship is explained in 1.

$$\begin{aligned} \dot{x} &= v\cos\phi \\ \dot{y} &= v\sin\phi \\ \dot{\phi} &= a \end{aligned} \tag{1}$$

The state space of DRL is further explained in the subsequent subsection.

B. States

As shown in [24], the DRL training reward is optimal where the state space has navigation information from the nearest three UAVs. Each state space consists of the position and speed of free-flying UAVs and information from the three nearest intruders. Hence, the state information is represented as:

$$\Phi = \left\{ \left\{ x_i, y_i, v_{x_i}, v_{y_i} : i \in N(o) \right\}, x_o, y_o, v_{x_o}, v_{y_o}, s, h, \Phi_{\text{goal}} \right\}, \tag{2}$$

where, $N(o)$ represent the set of j free flight UAVs nearest neighbor, with $\|N(o)\| = j$, Φ_{goal} represents the goal coordinates $(x_{\text{goal}}, y_{\text{goal}})$. Here $\{x_i, y_i, v_{x_i}, v_{y_i}\}$ represents position and speed of each intruder. s and h represent the speed and the heading direction of the free flight. Since the number of intruders considered here is 3, the state space dimensionality is 20. The action to be described from the states Φ is described in the next subsection.

C. Action

The action space of the free flight consists of direction or speed control. The action space dictates that the agent can either move in positive or negative velocity, turn left or right, or take no decision. Hence, the action space consists of two parts A_s and A_h, which control the speed and heading direction, respectively, so that the rewards can be maximized. The rewards are explained in the following subsection.

$$\begin{aligned} \mathbf{u} &= A_s \times A_h \\ A_s &= \{+v, 0, -v\} \\ A_h &= \{\phi, 0, -\phi\} \end{aligned} \tag{3}$$

D. Rewards

The goal of free-flight UAVs is to reach the target with a minimum amount of steps and avoid collision while minimizing conflicts. The reward function in achieving this

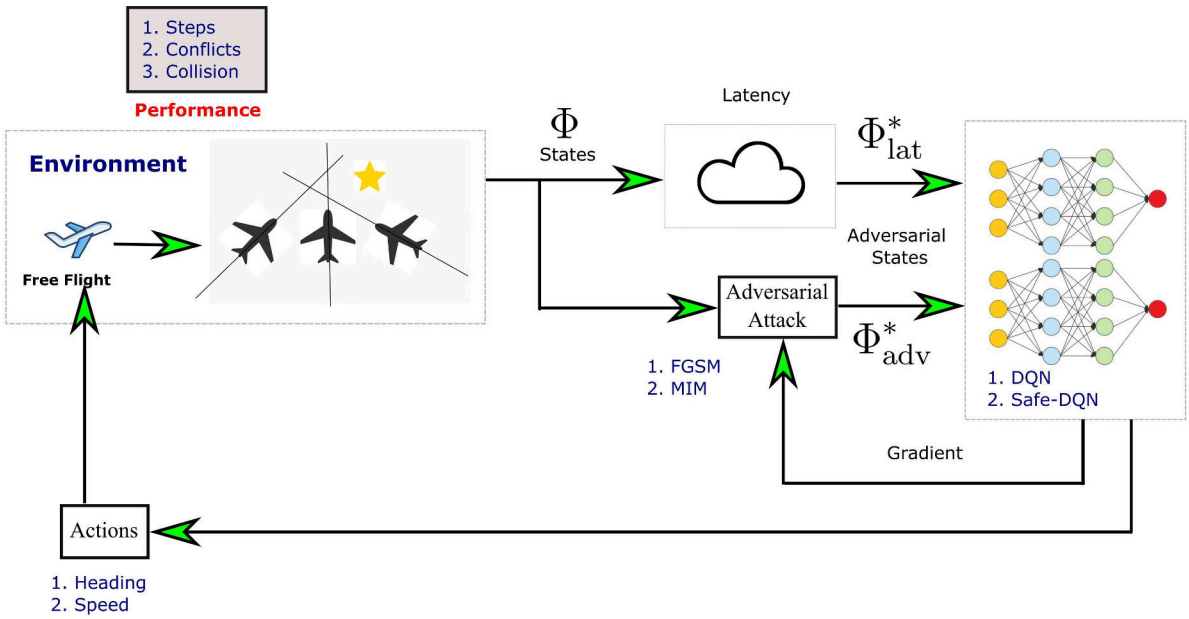

Fig. 1. Schematic of the entire AAM setup with gradient attacks and packet drop rate.

objective has been adapted from [24]. Vanilla DQN, having a single DNN, will consist of a reward function achieving both objectives. However, in safe DQN, separate reward structure for both missions i.e. efficiency and safety, i.e. r^c and r^g has been utilized, which is explained as follows.

$$r^c = \begin{cases} 1 & \text{target achieved,} \\ -0.0001 & \text{otherwise step penalty.} \end{cases} \tag{4}$$

$$r^g = \begin{cases} -1 & \text{if a collision occurs,} \\ -0.5 & \text{if a conflict occurs.} \end{cases} \tag{5}$$

The above reward scheme is used for safety DQN. However, in vanilla DQN all the reward functionalities are shown in a cumulative manner.

$$r = \begin{cases} 1 & \text{target achieved,} \\ -1 & \text{if a collision occurs,} \\ -0.5 & \text{if a conflict occurs,} \\ -0.0001 & \text{otherwise step penalty.} \end{cases} \tag{6}$$

E. Terminal State

The episode is terminated when either of the following occurs:

- When the UAV reaches the goal target while satisfying the necessary safety constraints.
- The UAV, not arriving at the target within 500 steps. which represents the finite battery capacity within the UAV.
- The free-flight UAV collides with another centrally controlled UAV.

III. Deep Reinforcement Learning and State Perturbations

A. DQN and Safe DQN

The task of the RL is to take actions to maximize the reward, as described in the previous section. The optimal policy is obtained from the expected sum of future rewards. For solving the sequential decision-making problem, the optimal value must be estimated. When the agent follows the policy π, the true value of the action u in a state Φ is given as:

$$Q^{\pi}(\Phi, u) = \mathbb{E}_{\Phi \sim E, \mathbf{u} \sim \pi}\left[\sum_{t=0}^{\infty} \gamma^k r_t \,\middle|\, \Phi_0 = \Phi, \mathbf{u}_0 = \mathbf{u}\right], \tag{7}$$

where, $\gamma \in [0,1]$ is the discount factor, r_t is the reward at time t. The optimal policy is defined as $\text{argmax}_{\pi} Q^{\pi}(\Phi, u)$, which is parameterized by the deep neural networks. Here θ represents the parameter of the network. The learning objective is to update the policy parameter θ^Q to minimize the Bellman error represented as:

$$L\left(\theta^Q\right) = \mathbb{E}_{\Phi \sim p^{\beta}, \mathbf{u} \sim \beta, r_t \sim E}\left[\left(Q\left(\Phi_t, \mathbf{u}_t; \theta\right) - y_t\right)\right]. \tag{8}$$

Here, β represents random exploration policy,θ^Q represents the parameters of the Q function, which replaces θ every τ steps respectively. The DQN described above does not incorporate the safety aspects in the value function. The safe DQN focuses on the learning policy which does not violate the safety constraints where two separate Q-networks are proposed for fulfilling efficiency and safety objectives. In safe-DQN, rewards are obtained from goal-oriented and safety-oriented objectives separately. Hence, we can split the

cost function in the Q value as:

$$Q^{\pi}(\Phi,u)=\mathbb{E}_{\Phi\sim E,\mathbf{u}\sim\pi}\left[\sum_{t=0}^{\infty}\gamma^{k}\left[w_{g}C_{\mathrm{g}}+w_{s}C_{\mathrm{s}}\right]\middle|\,\Phi_{0}=\Phi,\mathbf{u}_{0}=\mathbf{u}\right]. \tag{9}$$

Hence, without loss of generality, we can split the final Q value from the above equation (9), as given in [24] as:

$$Q^{\pi}=Q_{c}^{\pi}+Q_{g}^{\pi}. \tag{10}$$

From (10), two deep neural networks are designed to estimate the primary Q-value Q_c^{π} and safety value Q_g^{π}. Hence, the safety Q_g^{π} value is used to validate the safety condition of a certain state-action pair. In safety aware DQN, the action is selected by $u=\mathrm{argmax}_u\left(Q_c^{\pi}(\Phi_t,u)+Q_g^{\pi}(\Phi_t,u)\right)$ if the safety value $Q_g^{\pi}(\Phi_t,u)$ is above a certain threshold value. Once the neural network is formulated, testing the performance with several latencies and adversarial attacks is essential, as explained in the next subsection.

B. Adversarial Attacks on RL

Adversarial examples tend to mislead the output of machine learning models [29]. In RL, adversarial examples generally tend to perturb the input state values with minimum detectability, causing the RL agent to take actions which will mislead the agent away from the required goal or compromise their safety [10]. Fast Gradient Sign Method (FGSM) is a one-time attack popular for its low complexity and is the common adversarial crafting algorithm in DRL. The adversarial states $\Phi_{\mathrm{F}}^{\mathrm{adv}}$ after the FGSM attack is given as:

$$\Phi_{\mathrm{F}}^{\mathrm{adv}}=\Phi+\varepsilon\mathrm{sign}\left(\nabla_{\Phi}L(\Phi,\mathbf{u};\theta)\right), \tag{11}$$

where, $J(\Phi,\mathbf{u};\theta)$ represents the cross entropy loss between all the actions $\mathbf{u}$ and the distribution of the actions based on the Q values. $\nabla_{\Phi}L(\Phi,\mathbf{u};\theta)$ represents the gradient of the loss function with respect to the observation Φ. Perturbation frequency is crucial, as frequent perturbations significantly reduce detectability. Therefore, it is imperative to analyze the frequency of FGSM attacks, aiming for maximum damage inflicted while minimizing the effort.

In equation (11), the attack perturbation operates along a fixed gradient direction across all observations throughout the iterations. However, momentum is incorporated in the iterative attack, which changes the gradient direction [30]. The momentum-based iterative attack has also been studied concerning ATC jamming attack scenarios [31]. The attack can influence the iterative step size based on the difference in the gradients of consecutive observations. The dynamic momentum iterative method (MIM) is described as:

$$\begin{aligned}\alpha_n&=\frac{\left|2\nabla_{\Phi_n}L-\nabla_{\Phi_{n-1}}L\right|}{\left\|2\nabla_{\Phi_n}L-\nabla_{\Phi_{n-1}}L\right\|_1},\\ g_{n+1}&=\mu\cdot g_n+\frac{\nabla_{\Phi_n}*L}{\left\|\nabla_{\Phi_n}*L\right\|_1},\\ \Phi_{n+1}^{*}&=\mathrm{Clip}_{\Phi,\varepsilon}\left\{\Phi_n^{*}+\alpha_n\cdot\mathrm{sign}\left(g_{n+1}\right)\right\}.\end{aligned} \tag{12}$$

As we observe in 12, the step size α_n is dependent on the gradient $\nabla_{\Phi}L(\Phi,\mathbf{u};\theta)$ with respect to current and previous observation Φ_n and Φ_{n-1} respectively.

C. Packet Drop Constraints

The trained DRL is tested with respect to packet drop constraints in state observations. The loss in information, i.e. free flying UAV and intruder navigation data, is considered as packet drop rates while modelling packet drop constraints. Two cases are considered as they require separate communication infrastructure [32]. Let us consider the packet drop rate for the free-flying UAV to be p_{own} and the nearest intruders to be p_{int}. Hence, we can write the state space [16]:

$$\begin{aligned}\Phi_{\mathrm{era}}&=\left\{\Phi_{\mathrm{own}},\Phi_{\mathrm{int}},x_{\mathrm{goal}},y_{\mathrm{goal}}\right\},\\ \Phi_{\mathrm{own}}(t)&=p_{\mathrm{own}}\Phi_{\mathrm{own}}(t)+(1-p_{\mathrm{own}})\Phi_{\mathrm{own}}(t-1),\\ \Phi_{\mathrm{int}}(t)&=p_{\mathrm{int}}\Phi_{\mathrm{int}}(t)+(1-p_{\mathrm{int}})\Phi_{\mathrm{int}}(t-1).\end{aligned} \tag{13}$$

Here $\{p_{\mathrm{own}},p_{\mathrm{int}}\}$ are Bernoulli random variables. Nevertheless, the current Bernoulli approach fails to account for correlated packet drops, which depend on the status of communication channels, denoted as either "good" or "bad" In the former scenario, only a few packets are lost; in the latter, most packets are lost [17], [33]. In the conducted simulation, equal probabilities are assigned to the "good" and "bad" states. However, in the extended Gilbert model, an additional state is introduced alongside the "good" and "bad" states, with equal probabilities for packet transmission and loss.

IV. Results and Discussion

A. Implementation Details

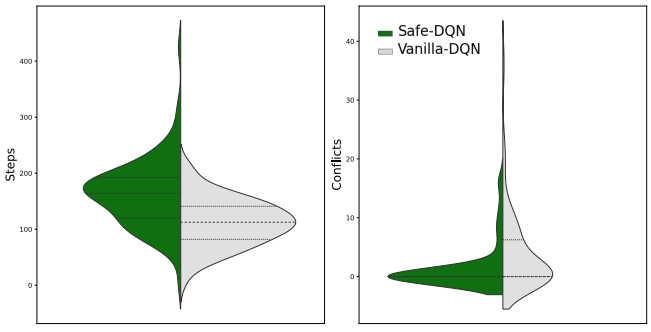

Fig. 2. Univariate plot of the number of steps and conflicts for DQN and Safe-DQN without any state perturbation.

We use the simulation setup described in [24] utilizing the trained vanilla and safe DQN to test the system under various attacks and packet drop constraint scenarios. The system runs for 500 episodes, involving ten centrally controlled UAVs placed randomly at the start and target positions. The maximum number of steps allowed is 500, representing the overall fuel capacity for the journey. Since intruder UAVs and initial positions are randomly generated, analysis has been performed using various fragility metrics of univariate

distribution. Unlike [24], these metrics are not averaged; the univariate distribution provides insight into their values across all randomized scenarios in each episode.

B. RL Performance without State Perturbation

Firstly, we evaluate the performance of the two DQN architectures without state perturbations, as depicted in Figure 2. We find that the safe DQN requires more steps on average to reach the target compared to the vanilla DQN, as it adopts a conservative approach to handle potential state perturbations. Although the average number of conflicts is similar between the two architectures, the vanilla DQN exhibits a higher distribution spread. Therefore, the safe DQN's cautious nature reduces the occurrence of separation loss cases, ensuring enhanced safety at the expense of slightly higher fuel consumption. Next, we will explore the implications when the states are subjected to adversarial perturbations.

C. RL Performance with Adversarial State Perturbation

In [24], the performance of the conflict resolution algorithm under an FGSM attack was examined by varying the perturbation magnitude for the safe DQN architecture. Disturbances were randomly injected at different intervals, and the corresponding performance metrics were measured. It was observed that conflicts increased with higher frequencies and magnitudes of disturbance. However, Figure 3 illustrates that the distribution spread widens with lower attack frequencies for the MIM attack. A noticeable increase was not observed for perturbation intensity of 0.03 with the MIM attack, but the spread of the number of steps was higher for perturbation intensity of 0.05 and 0.07. In contrast, the number of steps did not show significant changes with increasing perturbation intensity in the case of an FGSM attack. Nonetheless, the distribution of the number of steps remains similar for both the momentum iterative and FGSM attacks for higher attack frequencies. Therefore, we can conclude that the iterative attack tends to be more impactful for lower probability attacks and increasing values of perturbation intensity.

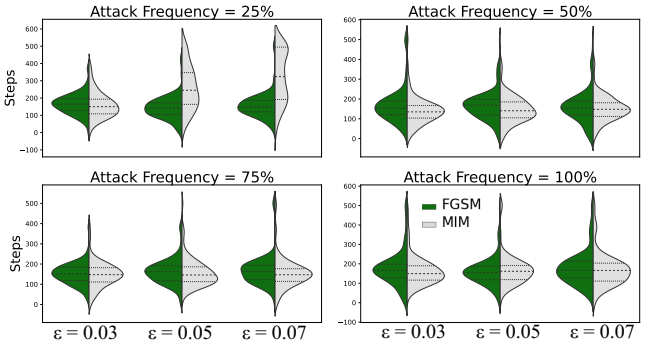

Fig. 3. Univariate plot of the number of steps of free-flying UAV with safe DQN under FGSM and MIM attack.

As we observe in Figure 4, with a lower attack frequency, conflicts are slightly increased, accompanied by a wider

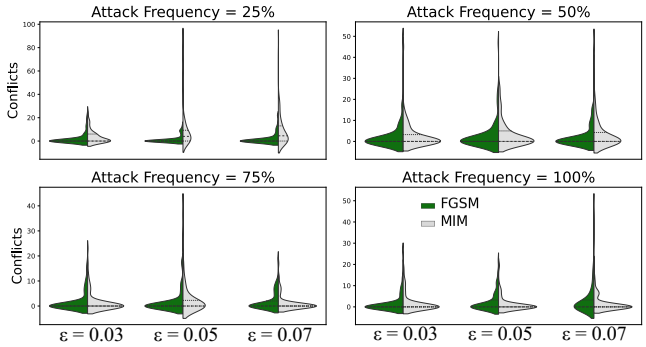

Fig. 4. Univariate plot of the number of conflicts of free-flying UAV with safe DQN under FGSM and MIM attack.

distributional spread as perturbation intensity increases. The conflicts also increase as the attack frequency decreases for MIM attacks. In the case of FGSM attacks, the spread expands as perturbation intensity increases, indicating an increase in the number of conflicts.

Similar to conflicts and the number of steps, it is observed in Figure 5 that, the number of collisions is higher for MIM attacks as compared to FGSM. The number of collisions is way higher for lower randomized attack frequencies, and it increases progressively with an increase in perturbation magnitude. However, we observe a sharp decrease in the number of collisions when the attack frequency is more than 50%. Hence, we can infer that safe RL, designed to mitigate FGSM attacks in [24], is antifragile with respect to MIM attacks when the attack frequency is more than 50%, while it is highly fragile at lower attack frequencies below 50%.

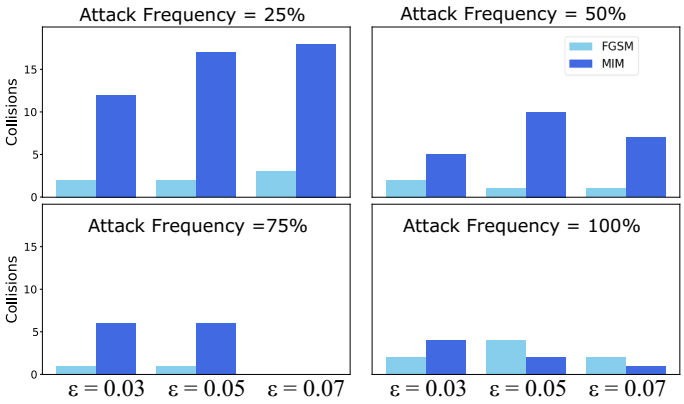

Fig. 5. Number of collisions of the own UAV with other UAVs in case of FGSM and MIM attacks.

D. RL performance with Packet Drop Constraints in Observation States

In the context of packet drops that adhere to the Bernoulli distribution, the drop rate of 50% has been considered [16]. When examining correlated packet drops using the Gilbert-Elliot models, the probability in both good and bad channel states has been considered to be 0.8, resulting in an average packet drop probability close to 0.5, albeit with correlated packet drops. Figure 6 illustrates that both vanilla

and safe DQN approaches exhibit slightly higher numbers of steps and conflicts when correlated packet drop models are employed compared to independent packet drops from the Bernoulli distribution. However, it is essential to note that no significant distinction is observed when there are packet drops of navigation information of the controlled UAV and those of the nearest UAVs.

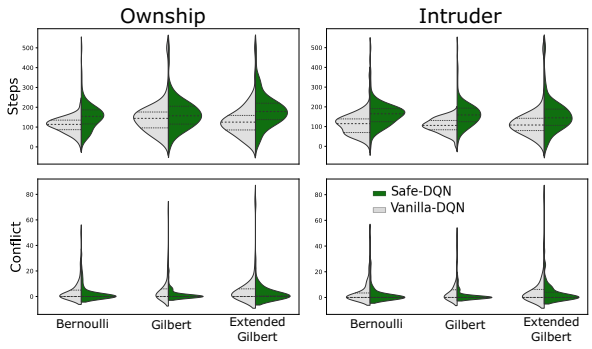

Fig. 6. Univariate plot of the number of conflicts and mission steps of free-flying UAV for different correlated packet drop model of ownship and nearest UAV with respect to safe and vanilla DQN.

However, concerning collisions, the safe DQN approach demonstrates a notable reduction in collision occurrences and proves to be more resilient to varying packet drop rates when compared to the vanilla DQN, as evident from the findings presented in Figure 7. It is also observed that the collision rate is higher when correlated packet drop rates are applied to the ownship navigation information, whereas uncorrelated packet drop rates on navigation information of intruder UAVs result in increased collision rates. These collision rate patterns align with those seen in the 50% attack frequency discussed in the previous subsection. Consequently, it can be inferred that future efforts to develop robust reinforcement learning (RL) strategies tailored explicitly for correlated packet drop rates could exhibit antifragile behaviour against the AAM with an attack frequency of around 50%.

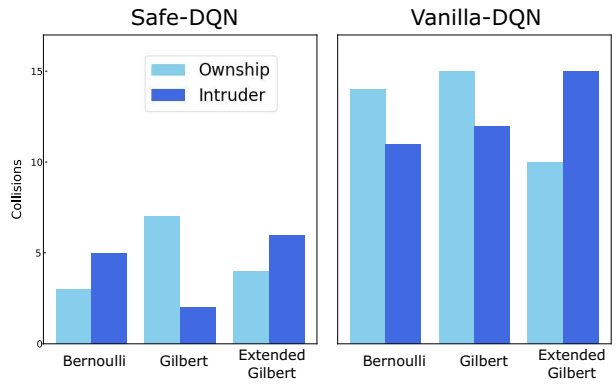

Fig. 7. Plot of the number of collisions of the free flying UAV with other UAVs for different packet drop model of ownship and nearest UAV with respect to safe and vanilla DQN.

V. CONCLUSIONS

In this paper, we have observed the fragility of RL-based free-flying UAV in AAM with respect to adversarial attacks and various uncorrelated and correlated packet drops. The trained RL agent is tested while considering the random number of intruder UAVs and the target position of the free-flying UAV in 500 episodes. The fragility analysis has been conducted with respect to the distributions of the number of steps required to reach the target and the number of conflicts and collisions with the intruder UAVs as a univariate plot. The main takeaways from this research are as follows:

- Safe DQN navigation algorithm, effective for FGSM attacks as studied in [24], is antifragile against MIM attacks at higher frequencies, but fragile against the same at lower frequencies with an increase in perturbation intensity.
- The fuel required for the completion of the mission and the number of conflicts is higher when it experiences correlated packet drops as compared to uncorrelated ones.
- The robust RL strategies designed to mitigate the correlated packet drops can be antifragile against adversarial attacks at around 50%.

Future work will focus on designing collision free navigation trajectories of the free-flying UAVs which will be antifragile against emerging adversarial attacks and packet drop constraints.

REFERENCES

[1] Federal Aviation Administration. *FAA Aerospace Forecast: Fiscal Years 2020–2040;* 2020.

[2] Giulio Rigoni et al. “Delivery with UAVs: a simulated dataset via ATS”. In: *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*. IEEE. 2022, pp. 1–6.

[3] Karthik Balakrishnan et al. “Blueprint for the Sky: The roadmap for the safe integration of autonomous aircraft”. In: *Airbus UTM, San Francisco, CA* (2018).

[4] Jacco M Hoekstra, Ronald NHW van Gent, and Rob CJ Ruigrok. “Designing for safety: the ‘free flight’air traffic management concept”. In: *Reliability Engineering & System Safety* 75.2 (2002), pp. 215–232.

[5] Jimmy Krozel and Mark Peters. “Conflict detection and resolution for free flight”. In: *Air Traffic Control Quarterly* 5.3 (1997), pp. 181–212.

[6] Mario Valenti Clari, Rob Ruigrok, and Jacco Hoekstra. “Cost-benefit study of free flight with airborne separation assurance”. In: *AIAA Guidance, Navigation, and Control Conference and Exhibit*. 2001, p. 4361.

[7] Marc Brittain and Peng Wei. “Autonomous aircraft sequencing and separation with hierarchical deep reinforcement learning”. In: *Learning-Based Decision Making for Safe and Scalable Autonomous Separation Assurance* (2021), p. 79.

[8] Zhuang Wang et al. “Review of deep reinforcement learning approaches for conflict resolution in air traffic control”. In: *Aerospace* 9.6 (2022), p. 294.

[9] Ministry of Defence. *Defence Artificial Intelligence Strategy*. Tech. rep. 2022.

[10] Anay Pattanaik et al. “Robust deep reinforcement learning with adversarial attacks”. In: *arXiv preprint arXiv:1712.03632* (2017).

[11] Qie Hu, Young Hwan Chang, and Claire J Tomlin. "Secure estimation for unmanned aerial vehicles against adversarial cyber attacks". In: *arXiv preprint arXiv:1606.04176* (2016).

[12] Xueyuan Wang and M Cenk Gursoy. "Resilient uav path planning for data collection under adversarial attacks". In: *ICC 2022-IEEE International Conference on Communications*. IEEE. 2022, pp. 625–630.

[13] Shafkat Islam et al. "A Triggerless Backdoor Attack and Defense Mechanism for Intelligent Task Offloading in Multi-UAV Systems". In: *IEEE Internet of Things Journal* 10.7 (2022), pp. 5719–5732.

[14] Cong Pu and Pingping Zhu. "Mitigating Routing Misbehavior in the Internet of Drones Environment". In: *2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring)*. IEEE. 2022, pp. 1–6.

[15] General Aviation Manufacturers Association. *Vehicle-to-Vehicle Datalink Communications: Enabling Highly Automated Aircraft and High-Density Operations in the National Airspace*. Tech. rep.

[16] Scott James, Robert Raheb, and Allison Hudak. "Impact of packet loss to the motion of autonomous UAV swarms". In: *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*. IEEE. 2020, pp. 1–9.

[17] Luca Rosario Buonocore et al. "Effects of packet losses on formation control of unmanned aerial vehicles". In: *IFAC Proceedings Volumes* 47.3 (2014), pp. 1234–1240.

[18] Thomas Hickling, Nabil Aouf, and Phillippa Spencer. "Robust adversarial attacks detection based on explainable deep reinforcement learning for uav guidance and planning". In: *arXiv preprint arXiv:2206.02670* (2022).

[19] Abdullah Hamdi, Matthias Müller, and Bernard Ghanem. "SADA: semantic adversarial diagnostic attacks for autonomous applications". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 07. 2020, pp. 10901–10908.

[20] Changhong Fu et al. "Ad 2 attack: Adaptive adversarial attack on real-time uav tracking". In: *2022 International Conference on Robotics and Automation (ICRA)*. IEEE. 2022, pp. 5893–5899.

[21] Jiwei Tian et al. "Adversarial attacks and defenses for deep-learning-based unmanned aerial vehicles". In: *IEEE Internet of Things Journal* 9.22 (2021), pp. 22399–22409.

[22] Drew Davidson et al. "Controlling {UAVs} with sensor input spoofing attacks". In: *10th USENIX workshop on offensive technologies (WOOT 16)*. 2016.

[23] Wenbin Zhai et al. "ETD: An Efficient Time Delay Attack Detection Framework for UAV Networks". In: *IEEE Transactions on Information Forensics and Security* (2023).

[24] Lei Wang et al. "Explainable and Safe Reinforcement Learning for Autonomous Air Mobility". In: *arXiv preprint arXiv:2211.13474* (2022).

[25] Mouna Elloumi et al. "Monitoring road traffic with a UAV-based system". In: *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE. 2018, pp. 1–6.

[26] Koppány Máthé and Lucian Buşoniu. "Vision and control for UAVs: A survey of general methods and of inexpensive platforms for infrastructure inspection". In: *Sensors* 15.7 (2015), pp. 14887–14916.

[27] Clément Lemardelé et al. "Potentialities of drones and ground autonomous delivery devices for last-mile logistics". In: *Transportation Research Part E: Logistics and Transportation Review* 149 (2021), p. 102325.

[28] Xuxi Yang and Peng Wei. "Autonomous on-demand free flight operations in urban air mobility using Monte Carlo tree search". In: *International Conference on Research in Air Transportation (ICRAT), Barcelona, Spain*. Vol. 8. 2018.

[29] Christian Szegedy et al. "Intriguing properties of neural networks". In: *arXiv preprint arXiv:1312.6199* (2013).

[30] Yinpeng Dong et al. "Boosting adversarial attacks with momentum". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 9185–9193.

[31] Mingqian Liu et al. "Adversarial attack and defense on deep learning for air transportation communication jamming". In: *IEEE Transactions on Intelligent Transportation Systems* (2023).

[32] Kenneth Bandelier et al. "White Paper-Use Cases for Vehicle-to-Vehicle (V2V) Communications for Unmanned Aircraft Systems". In: *Use Cases for Vehicle-to-Vehicle (V2V) Communications for Unmanned Aircraft Systems* (2023), pp. 1–24.

[33] Kelvin K Lee and Samuel T Chanson. "Packet loss probability for real-time wireless communications". In: *IEEE Transactions on Vehicular Technology* 51.6 (2002), pp. 1569–1575.

Fragility impact of RL based advanced air mobility under gradient attacks and packet drop constraints

Panda, Deepak Kumar

2023-12-11

Panda DK, Guo W. (2023) Fragility impact of RL based advanced air mobility under gradient attacks and packet drop constraints. In: 2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall), 10-13 October 2023, Hong Kong
https://doi.org/10.1109/VTC2023-Fall60731.2023.10333535