

5G-based Ground Risk Mitigation for UAVs: A Deep Reinforcement Learning Approach

Mohammed Lahouari Harchaoui*, Sihem Ouahouah[§], Oussama Bekkouché[§], Miloud Bagaa[¶], and Derouiche Abir*

*Ecole Nationale Supérieure d'Informatique (ESI ex.INI), Algiers, Algeria

{jm_harchaoui, a_derouiche}@esi.dz

[§]Aalto University, Espoo, Finland

{firstname.lastname}@aalto.fi

[¶]Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada

miloud.bagaa@uqtr.ca

Abstract—The emergence of the Beyond Visual Line of Sight (BVLOS) operations for Unmanned Aerial Vehicles (UAVs) unlocked a wide range of new applications across various domains, such as urban transportation, package delivery, and aerial surveillance. However, due to the possibility of losing control and collisions, BVLOS operations present several risks to people on the ground. Therefore, it is crucial to minimize safety risks by flying UAVs along paths that traverse less populated areas. Nevertheless, implementing such a solution requires access to real-time data on the population density distribution across UAV operational areas. Consequently, in this paper, we harness the network exposure capabilities of 5G mobile networks, proposing a framework that integrates the UAV Traffic Management (UTM) system with the 5G Core (5GC). The proposed framework can collect real-time information about the density of mobile users in Areas of Interest (AoI), leveraging this data to estimate ground risks and subsequently devise optimized flight paths. Moreover, we propose a Deep Reinforcement Learning (DRL) solution to compute optimized flight paths. The simulation results show the efficiency of our proposed solution to achieve the designed goals in terms of reducing the experienced ground risk and total flight distance.

Index Terms—UAVs, 5G, Risk mitigation, service-based architecture, flight planning, deep reinforcement learning.

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) have a promising role in the emerging use cases of Urban Air Mobility (UAM), facilitating a wide range of applications, such as urban transportation, package delivery, and aerial surveillance [1]. However, realizing this potential requires transitioning from Visual Line-of-Sight (VLOS) to Beyond Visual Line-of-Sight (BVLOS) operations, which encounter several challenges, including safety concerns, and limitations in communication range. Fortunately, 5G networks offer solutions to many of these challenges by providing Ultra Reliable Low Latency Communication (URLLC) to UAVs. Additionally, 5G networks encompass ulterior value-added services such as identification, localization, and demo-graphical data that can be explored. Additionally, progress has been noted in safety regulations for BVLOS operations [2]. Nevertheless, it is emphasized that, despite these advancements, safety concerns remain, particularly in densely populated areas. Indeed, collisions with static or mobile obstacles may pose risks to

people and property on the ground, potentially resulting in fatalities. To address these risks, authorities enforce different risk assessment methodologies [2], [3]. These methodologies assess BVLOS operations for validation, focusing on reducing both air and ground risks by thoroughly analyzing the flight region and creating optimized flight plans.

Various models and algorithms have been developed to automate and enhance the process of risk assessment for BVLOS operations. These models typically involve calculating risks and generating risk maps in the flight region, followed by the application of flight planning algorithms to find the best paths based on this information [4]–[7]. Different approaches consider various types of risks, including the risk of collision with manned aircraft and high buildings (i.e., air risk), as well as the risk of falling on people and properties on the ground (i.e., ground risk).

In this work, we focus on mitigating ground risks, especially the risk of falling on people. This risk is strongly correlated with population density [8], whereby several methods have attempted to estimate using static data from census and registers. However, these datasets often fail to capture the spatiotemporal features necessary for accurate risk assessment, as they provide information primarily about “night-time population” rather than actual changing population presence throughout the day. In recent years, the proliferation of novel digital data sources, such as mobile phones and social media, has revolutionized population dynamics tracking. Mobile phone data, in particular, offers precise digital footprints of individuals in space and time, providing comprehensive tracking over long periods. While not all mobile network generations allow for real-time data collection, 5G enables this capability by providing a mechanism to authorized third-party applications for collecting information about the number of mobile users in Area-of-Interests (AoIs). Our paper introduces a novel risk-aware path planning method that utilizes 5G user mobility data for population density estimation, enabling the calculation of ground risks and the generation of flight paths that mitigate these risks using Deep Reinforcement Learning (DRL). Indeed, through continuous learning from experiences by interacting with the environment, the use of DRL exhibits

a robust adaptation to environmental changes, regardless of their occurrence frequency.

The remainder of this paper is organized as follows. Section II presents the related works. Section III discusses our proposed framework. Our problem formulation and DRL solution are detailed in Section IV. Section V presents the performance evaluation and results analysis of the proposed solution. Finally, Section VI concludes the paper.

II. RELATED WORK

Various research works were conducted alongside the emergence of UAV-based applications to assess risks related to their deployment, particularly regarding the presence of potential impacts of UAVs fatalities on individuals and ground vehicles. The correlation between population density and the fatality risk of UAVs has been investigated and concluded by authors in [8]. Further features are determined and modeled by adopting different mathematical models. The loss of control, collision with people or objects on the ground, and fatality rates in case of collision have been considered and estimated based on a stochastic model using conditional probability [9]. Whereas, a two-dimensional risk map model that uses location, population density, no-fly zones, and obstacles, quantifying the risk to the population on the ground has been suggested in [10]. An enhanced precise probable risk estimation method has been introduced in [11], which analyzes possible fatality positions during the descent process of UAVs.

Authors in [12] have proposed another stochastic-based alternative method for evaluating the likelihood of all events leading to accidents that may cause damage to road traffic. While authors in [13] investigate the severity of UAV fatality impacts by generating probabilistic risk maps using the Monte Carlo simulator under limited environmental uncertainty. A practical study conducted in [14] aimed to model third-party risk in a metropolitan environment by analyzing existing frameworks. The results of similar studies can facilitate the generation of risk-aware maps, thereby enabling various tasks to benefit from them to enhance the implementation of UAV-based applications. One such task is the generation of UAV path planning. In this regard, several works have been conducted using a variety of models and optimization objectives.

Traditional methods such as Dijkstra's algorithm [15] and heuristic algorithms like A* [16] have been employed, along with swarm-based heuristic methods [17]. These methods aim to minimize travel distance and cost or maximize flight duration while considering obstacle avoidance. However, they often overlook risks underneath the UAV's flight path. In response, risk-based path planning has emerged, leveraging risk maps for path planning [4], [5]. Various algorithms have been developed to generate risk maps and address risk-based path-planning problems. The RiskA* algorithm was later introduced to minimize path risk [6]. Genetic algorithms and Dijkstra's methods are also commonly employed and serve as benchmarks for comparison with A*-based approaches [7].

Current risk models regarding ground impacts often overlook the dynamics of population density in metropolitan

areas, thus undermining the accuracy of assessing population distribution, a crucial factor in risk analysis. Regrettably, the aforementioned mathematical adopted models did not overcome this issue.

III. INTEGRATED FRAMEWORK FOR POPULATION DENSITY-AWARE FLIGHT PLANNING

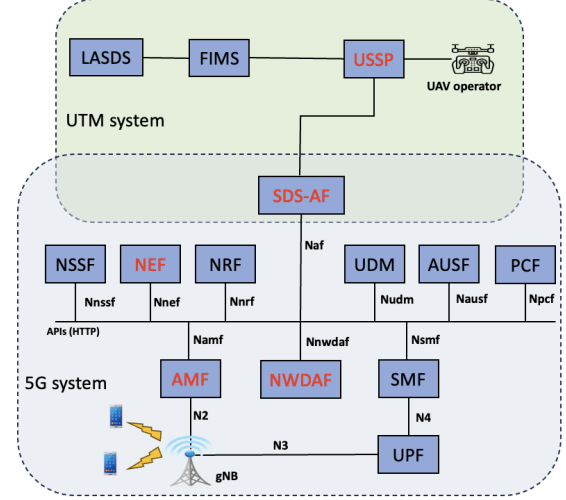


Fig. 1: Integrated 5G-UTM systems

In this section, we discuss the practical realization of our proposal. Fig.1 illustrates the integration of a UTM system with a standardized 5G system [18]. Although UTM systems vary in implementation, they share common concepts of operations and fundamental components [2], [3]. Here, we focus solely on the functionalities of the 5G and UTM system components involved in the population density-aware flight planning process, which are highlighted in red in Fig.1. For more detailed information on other components, interested readers are referred to [2], [18].

- **UAS Service Provider (USSP):** Supports the UAV operator by providing services for flight planning. This includes, for example, generating flight paths that meet the operational requirements of local authorities and enable safe and efficient use of airspace. In our proposal, the USSP is responsible for leveraging the population density information to devise flight plans that minimize ground risk.
- **UAS Supplemental Data Service Application Function (SDS-AF):** supports the USSP by providing data collected from various sources; this includes, for example, terrain and obstacle data, weather data, and population density data. In our proposed framework, the SDS-AF is deployed as a third-party Application Function (AF) in the 5G system [18]. The authorized third-party AFs can collect real-time network data. Thus, the SDS-AF is responsible for collecting data on the number of mobile users in the UAV's operational area and sharing this information with the USSP to build the population density map.

- **Network Exposure Function (NEF)**: plays the role of a proxy that securely exposes network capabilities data to third-party AFs. The NEF is responsible for authenticating and authorizing the SDS-AF, and exposing the mobile user density information from the network to the SDS-AF.
- **Network Data Analytics Function (NWDAF)**: performs the actual collection of network data and analytics then exposes the results to third-party AFs via the NEF [19]. Based on the request from the SDS-AF, the NEF requests the mobile user density information from the NWDAF and forwards the received data to the SDS-AF.
- **Access and Mobility Management Function (AMF)**: possesses the raw user mobility data, which is collected by the NWDAF for deriving mobile user density information [19].

IV. RL-BASED MOBILE NETWORK RISK ASSESSMENT AND MITIGATION

A. System model

1) *UAV navigation model*: We consider an environment, whereby a UAV of interest flies at a fixed speed $\mathcal{V}$ over a fixed altitude; from a starting location (x_s, y_s) till arriving at a target destination (x_d, y_d) . For the sake of simplicity in what follows, we denote the UAV of interest as UAV. Moreover, we assume that the time horizon is slotted to discrete time slots $t = 0, 1, 2, 3, \dots$ of equal length Δt . At the beginning of each time slot t , the UAV in location (x_t, y_t) adjusts its heading direction by taking a rotation angle $a_t \in \{0 \dots 360\}$. By the end of the time slot t , the UAV travels a distance $\Delta d = \Delta t \times \mathcal{V}$ in the adjusted direction and moves to a new location denoted (x_{t+1}, y_{t+1}) . Indeed, the adopted UAV's kinematic model is defined as follows:

$$\begin{cases} x_t = \mathcal{V} \times \cos(a_t) \\ y_t = \mathcal{V} \times \sin(a_t), \end{cases} \quad (1)$$

, whereby $a_t \in \{0 \dots 360\}$. These steps are repeated until the UAV reaches (x_d, y_d) . We notice here that once the distance d_t between the current location (x_t, y_t) and the destination (x_d, y_d) is smaller or equal to a small threshold D_{thr} , we consider that the UAV reaches its target destination. Moreover, we suppose that UAV's flight starts at a specific timestamp that represents the time of day, denoted T_s , and we define the function $\Theta(t)$ that maps discrete time slots t to timestamps of the day as follows:

$$\Theta(t) = \begin{cases} T_s & \text{If } t = 0 \\ \Theta(t-1) + \Delta t & \text{Otherwise} \end{cases} \quad (2)$$

Formally, the first time slot at which the UAV adjusts its direction occurs at timestamp T_s , the second time slot occurs at $T_s + \Delta t$, and so on.

2) *Ground risk assessment model*: The area in which the UAV operates is modeled as a grid of cells denoted $\mathcal{C}$, whereby each cell has time-varying risk levels that depend on the ground population density in that cell. In this work, we adopt the risk assessment model presented in [5]. The proposed

risk model offers a mathematical framework that accounts for multiple risk types for UAVs. Nevertheless, in our formulation, we consider only the risk model of ground impact on people. The risk level experienced by the UAV for people on the ground at a particular time slot t is given by:

$$K_t = \lambda \times \mathcal{M}_{c,t} \times \mathcal{F}_c, \quad (3)$$

, whereby λ is a factor representing the probability of the UAV falling, c is the cell containing the UAV's coordinates (x_t, y_t) , and $\mathcal{M}_{c,t}$ is the number of people at risk in cell c at time slot t expressed by:

$$\mathcal{M}_{c,t} = \mathcal{A}_c \times \mathcal{P}_{c,\Theta(t)}, \quad (4)$$

, with $\mathcal{A}_c$ denoting the area of cell c and $\mathcal{P}_{c,\Theta(t)}$ denoting the population density in cell c at timestamp $\Theta(t)$. The population density is retrieved from the 5G mobile network as discussed in section III, by specifying the Area-of-Interest (AoI) c and the time of the day $\Theta(t)$. Nevertheless, because of the unknown mobility patterns of mobile users, the population density information returned by the NWDAF is stochastic and can vary considerably over time. For instance, some areas might be crowded during the day and vacant during the evening. This implies that the number of people at risk $\mathcal{M}_{c,t}$, and thus K_t , are also stochastic random variables with unknown distribution.

Finally, $\mathcal{F}_c$ denotes the fatality factor in cell c , which depends on the UAV's kinetic (i.e., UAV's acceleration and weight) and the nature of the terrain in cell c . The fatality factor $\mathcal{F}_c$ is expressed by:

$$\mathcal{F}_c = \frac{1}{1 + \left(\sqrt{\frac{I}{I_0} \left(\frac{I_0}{E_{\text{imp}}} \right)} \right)^{\frac{1}{4 \times Sh_c}}} \quad (5)$$

, whereby Sh_c is the sheltering coefficient of cell c that considers the characteristics of the terrain, such as buildings and trees, softening the ground impact to people from falling UAVs. The possible values of Sh_c are presented in TABLE I. Furthermore, I is the impact energy that causes 50% of fatalities when $Sh_c = 0.5$, I_0 is the impact energy that causes fatalities when Sh_c approaches 0, and $E_{\text{imp}} = \frac{m \times V_f^2}{2}$ is the impact of the kinetic energy of the falling UAV, with m denoting the UAV's mass and V_f denoting the fall speed.

TABLE I: The values of the sheltering coefficient

Sh_c	Terrain characteristics
0+	No obstacles
0.25	Sparse trees
0.50	Trees and low buildings
0.75	High buildings
1	Industrial buildings

B. Problem formulation

Our goal is to steer the UAV from its starting location (x_s, y_s) to its destination (x_d, y_d) by controlling its heading direction. The resulting flight path should minimize both, the experienced ground risk, and the traveled distance. This problem involves making a sequence of control decisions over time, where each decision affects future states and decisions.

As a result, we first model our problem using the Markov Decision Process (MDP) framework which is inherently suited for such sequential decision-making problems. In MDPs, the decision maker referred to as “Agent” chooses an action denoted a_t based on the current state of its environment denoted s_t according to a policy, which is a strategy that specifies the best action to take in each state to maximize future rewards. The environment then responds to the agent’s action by transitioning to a new state s_{t+1} , according to a probability distribution known as the transition probabilities. Upon transitioning to the new state, the agent receives a reward (or penalty), denoted as r_{t+1} which quantifies the benefit (or cost) of the chosen action. Through this ongoing process of action selection, state transition, and reward collection, the MDP framework allows the agent to learn the optimal policy that maximizes the cumulative reward over time. An MDP can therefore be defined by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$, whereby $\mathcal{S}$ is the states space, $\mathcal{A}$ is the actions space, $\mathcal{P}$ is the transitions probability function, and $\mathcal{R}$ is the reward function. Hereinafter, we define the components of our MDP formulation:

1) *State space*: The state observed by the agent at the time step t is represented by the vector $s_t \in \mathcal{S}$ defined as follows:

$$s_t = \langle x_t, y_t, x_d, y_d, d_t, \psi_t, \Theta(t) \rangle \quad (6)$$

, whereby (x_t, y_t) are the coordinates of UAV at time step t , (x_d, y_d) are the coordinates of the target destination, d_t is the distance between the (x_t, y_t) of the UAV and (x_d, y_d) , ψ_t is the clockwise angle calculated from (x_t, y_t) to (x_d, y_d) , indicating how far is the direction of the agent in comparing to the targeted destination, and $\Theta(t)$ is the time of the day at time step t . ψ_t is calculated with the following formula:

$$\psi_t = (\text{degrees}(\tan^{-1}(x_d - x_t, y_d - y_t)) + 360) \bmod 360 \quad (7)$$

2) *Actions space*: At each time step t , the agent adjusts the direction of the UAV by a clockwise angle $a_t \in \mathcal{A}$ with $\mathcal{A} = \{0 \dots 360\}$. As discussed in section IV-A1, the UAV will then travel for a distance Δt in the adjusted direction.

3) *Reward function*: The agent seeks to minimize the total traveled distance and the experienced ground risk. Therefore, the reward function when transiting to a state s_t can be expressed by:

$$r_t = \begin{cases} P_{out} & \text{If } (x_t, y_t) \notin (X \times Y) \\ R_{arrv} & \text{If } d_t \leq D_{thr} \\ -w_d \times d_t - w_r \times K_t & \text{If } d_t > D_{thr} \end{cases} \quad (8)$$

Formally, the reward function (8) is crafted to encourage the agent to keep the UAV inside the operations grid and progress toward the target efficiently. The agent is penalized with a high penalty P_{out} when the UAV moves a new location (x_t, y_t) which is out of the considered grid, where $X \times Y$ is the set of all possible locations inside the grid. The agent is rewarded with a high reward of R_{arrv} when the UAV arrives at its target location (i.e., the distance to the target location d_{t+1}

is smaller or equal to a threshold D_{thr}). Finally, the agent is penalized for all transitions to states s_t in which the distance d_t to the target location is higher than D_{thr} . The penalty is expressed as a weighted sum inversely proportional to the distance between the current location and target destination d_t and the experienced ground risk K_t in location (x_t, y_t) , urging the agent to follow the shortest and safest possible flight path to the destination. It is worth noting that both d_t and K_t are normalized to the same range $[0, 1]$, achieving a trade-off between minimized risk and distance, where $w_r + w_d = 1$.

4) *Transition probabilities*: In an MDP, the environment’s *dynamics* describing the state transitions and obtained rewards are given by a function $p : \mathcal{S} \times \mathcal{R} \times \mathcal{S} \times \mathcal{A} \rightarrow [0, 1]$ defined as follows:

$$p(s', r | s, a) \doteq Pr\{s_{t+1} = s', r_{t+1} = r | s_t = s, a_t = a\}, \quad (9)$$

, whereby $p(s', r | s, a)$ defines the probability of transitioning to state s' and receiving a reward r when the agent is in state s and selects action a . However, since the returned reward in equation (8) depends on the random risk K_t which had an unknown distribution, it is not possible to model the transition probabilities, and thus it is not possible to solve the MDP using classical dynamic programming methods. As a result, we propose DRL solution that can learn sub-optimal policies for model-free MDPs even with the unknown dynamics.

5) *Deep Reinforcement Learning Solution*: The agent’s objective is to determine the optimal policy π^* that maximizes the expected cumulative reward over time. Initially, we consider the state value function, $V(s)$, which represents the expected cumulative reward when starting from state s and following a given policy π . Formally, the state value function is defined as:

$$V_\pi(s) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s \right] \quad (10)$$

, whereby R_{t+k+1} represents the reward obtained at time step $t + k + 1$, and γ is the discount factor determining the present value of future rewards.

The optimal policy π^* is the one that maximizes the state value function for all states:

$$V_\pi(s) \leq V_{\pi^*}(s), \quad \forall s \in \mathcal{S}, \forall \pi, \quad (11)$$

This implies that the optimal policy leads to the highest expected cumulative reward starting from any state s . On the other hand, the Q-value, $Q_\pi(s, a)$, represents the expected cumulative reward starting from state s , taking action a , and then following a given policy π thereafter. Formally, the Q-value is defined as:

$$Q_\pi(s, a) = \mathbb{E}_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \mid S_t = s, A_t = a \right]. \quad (12)$$

The optimal policy π^* is the one that maximizes the Q-value for all state-action pairs:

$$Q_\pi(s, a) \leq Q_{\pi^*}(s, a), \quad \forall s \in \mathcal{S}, \forall a \in \mathcal{A}, \forall \pi. \quad (13)$$

The relationship between the state-value function and the action-value function is expressed by:

$$V_\pi(s) = \max_a Q_\pi(s, a), \quad (14)$$

, whereby the maximum expected cumulative reward when starting from state s is achieved by selecting the action that maximizes the Q-value for that state.

Because of the high dimension of our state space and the unknown environment dynamics, we propose a DRL solution that relies on neural networks to learn and approximate the Q-values of actions directly from the raw state inputs. Moreover, we adopt the value-based DRL algorithm, known as Dueling Deep Q Networks (DDQN), a sophisticated variation of the traditional Deep Q-Networks DQN algorithm. DDQN uses neural networks to approximate the Q-values, this network takes the state as input and predicts the Q-values for each action.

$$Q(s, a; \theta) \approx Q^*(s, a) \quad (15)$$

, Whereby θ represents the parameters of the neural network, and $Q^*(s, a)$ is the true optimal Q-value function.

DDQN decomposes the Q-value function into two separate components: The value function $V(s)$, and the advantage function $A(s, a)$. The value function represents the value of being in a particular state regardless of the action taken, while the advantage function evaluates the advantage of taking a specific action over others in the same state. The Q-value function is then reconstructed using these two components:

$$Q(s, a; \theta) = V(s; \theta_v) + (A(s, a; \theta_a) - \frac{1}{|\mathcal{A}|} \sum_{a'} A(s, a'; \theta_a)) \quad (16)$$

, whereby $|\mathcal{A}|$ represents the number of possible actions in state s . The term $A(s, a; \theta_a) - \frac{1}{|\mathcal{A}|} \sum_{a'} A(s, a'; \theta_a)$ helps the agent prioritize actions based on their relative advantages rather than their absolute values, leading to more efficient and effective learning.

The network is trained to approximate both the value and advantage functions simultaneously by minimizing the loss function, typically defined as the mean squared error between the predicted Q-values and the target Q-values:

$$L(\theta, \alpha, \beta) = \mathbb{E}[(Q(s, a; \theta, \alpha, \beta) - y)^2], \quad (17)$$

where y is the target Q-value obtained using the Bellman optimality equation [20].

V. PERFORMANCE EVALUATION

This section evaluates the ability of our solution to achieve its designed goals in terms of minimizing the ground risk for UAVs and the total traveled distance. For this purpose, we compare the gains of our DDQN-based solution with two baseline solutions: *i) Shortest path*: steers the UAV by following the shortest path from the starting location to the target destination; *ii) Safest path*: steers the UAV by following the trajectory that has minimal ground risks. We consider an area of $500 \times 500 \text{ m}^2$ devised to a grid of 100 cells (i.e., the size of each cell is 2500 m^2). The speed of UAV is set to 5 m/s , and Δt is set to 2 s . To mimic the data collected by the NWDAF, we use the dataset published in [21] for the Helsinki capital region. The data of the categories of terrains in each cell was retrieved from the public maps “OpenStreetMap”.

Fig.2 depicts the cumulative rewards achieved by the DDQN algorithm during the training process, we notice that cumulative rewards converge after around 1000 training episodes,

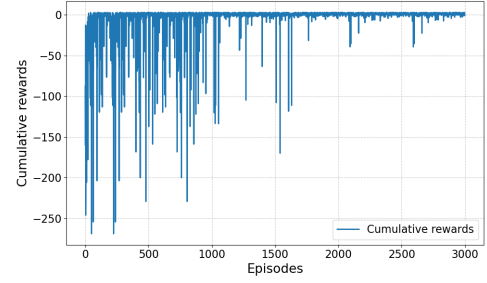


Fig. 2: The achieved cumulative rewards during the training reward process

indicating that the estimated Q-values approximate their true value. This later can be used to derive the optimal policy.

Fig.3 depicts the path plots of a test case for the baseline solutions and our DDQN-based solution, whilst TABLE II presents the risk and distance of the selected paths. We notice that the DDQN solution achieves a tradeoff between the traveled distance and the experienced ground risk, compared to the shortest path solution that selects a path with a shorter distance at the cost of higher ground risk, and the safest path solution that selects a path with a smaller ground risk at the cost longer travel distance.

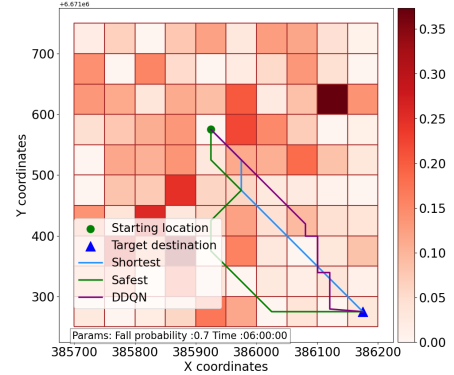


Fig. 3: Paths generated by the different solutions

To further validate our claims, we evaluate the performance of all solutions while varying the ground risk in each cell. For this purpose, we increase the risk expressed in equation(3) by increasing the UAV fall probability λ . We generate 100 different missions with different starting locations and destinations. Then, we devise the UAV flight path for each mission using the three solutions while considering different values for the UAV fall probability. Finally, we calculate the experienced risk and traveled distance for each devised path.

Fig.4 and Fig.5, respectively, present the average traveled distance and the average experienced risk across the 100 scenarios for all solutions when considering different fall probabilities. We notice that the increased UAV fall probability (i.e., increased risk in cells) causes our DDQN-based solution to select paths with longer travel distances and higher risks. The shortest path solution selects the same path for each of the 100 missions regardless of the UAV fall probability and thus its stable achieved travel distance and linearly increased risk

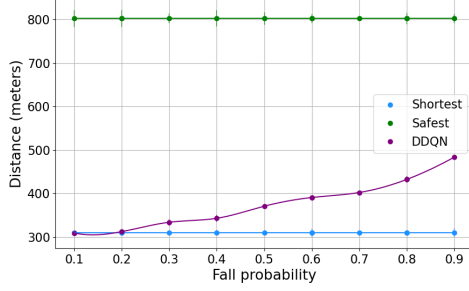


Fig. 4: Simulation results of distance for baseline algorithms and the proposed solution

(the risk of the same paths increases by a λ factor). The safest path solution also selects the same path for each of the 100 missions regardless of the UAV fall probability, this is because when λ increases, the risk increases equally for all cells. Thus, the safest path is always the same. However, the DDQN-based solution always outperforms the safest path solution in terms of travel distance. It also outperforms the shortest-path solution in terms of experienced risk. As a result, our solution achieves a better tradeoff between risk and travel distance.

TABLE II: Achieved results for the different solutions.

	Shortest	Safest	DDQN
Distance	252.4362	388.9550	265.0718
Risk	38.1537	23.4606	28.8061

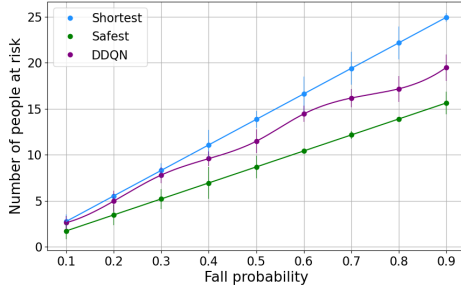


Fig. 5: Simulation results of risk for baseline algorithms and the proposed solution

VI. CONCLUSION

This paper presented a novel framework for mitigating ground risks in UAV operations, particularly under Beyond Visual Line of Sight (BVLOS) conditions, by leveraging the advanced capabilities of 5G networks. Our approach integrated a UTM system with the 5G core to utilize real-time mobile user density data for dynamic risk assessment and flight path optimization. The DRL-based solution proved pivotal in incorporating real-time data and continuously adjusting to changing conditions, computing paths that not only minimize ground risks but also reduce total flight distance, and balancing safety with operational efficiency.

Future work will focus on refining the proposed solution, by exploring the integration of additional real-time data sources

and risk factors. Furthermore, we will expand our models to include multi-UAV scenarios. Ultimately, this research not only advances UAV operational safety but also contributes to the broader implementation of autonomous systems in complex urban landscapes.

REFERENCES

- [1] S. Mishra and P. Palanisamy, "Autonomous advanced aerial mobility—an end-to-end autonomy framework for uavs and beyond," *IEEE Access*, vol. 11, pp. 136 318–136 349, 2023.
- [2] American Federal Aviation Administration, "Unmanned Aircraft System (UAS) Traffic Management (UTM), UTM ConOps v2.0," March 2020.
- [3] SESAR JOINT UNDERTAKING, "U-space concept of operations (conops), fourth edition," July 2023.
- [4] L. De Filippis, G. Guglieri, and F. Quagliotti, "A minimum risk approach for path planning of uavs," *Journal of Intelligent & Robotic Systems*, vol. 61, pp. 203–219, 2011.
- [5] X. Hu, B. Pang, F. Dai, and K. H. Low, "Risk assessment model for uav cost-effective path planning in urban environments," *IEEE Access*, vol. 8, pp. 150 162–150 173, 2020.
- [6] S. Primates, G. Guglieri, and A. Rizzo, "A risk-aware path planning strategy for uavs in urban environments," *Journal of Intelligent & Robotic Systems*, vol. 95, pp. 629–643, 2019.
- [7] J. d. Silva Arantes, M. d. Silva Arantes, C. F. Motta Toledo, O. T. Júnior, and B. C. Williams, "Heuristic and genetic algorithm approaches for uav path planning under critical situation," *International Journal on Artificial Intelligence Tools*, vol. 26, no. 01, p. 1760008, 2017.
- [8] K. Dalamagkidis, K. P. Valavanis, and L. A. Pieg, "Evaluating the risk of unmanned aircraft ground impacts," in *2008 16th mediterranean conference on control and automation*. IEEE, 2008, pp. 709–716.
- [9] S. Bertrand, N. Raballand, F. Viguier, and F. Muller, "Ground risk assessment for long-range inspection missions of railways by uavs," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2017, pp. 1343–1351.
- [10] S. Primates, A. Rizzo, and A. la Cour-Harbo, "Ground risk map for unmanned aircraft in urban environments," *Journal of Intelligent & Robotic Systems*, vol. 97, pp. 489–509, 2020.
- [11] Y. Liu, X. Zhang, Z. Wang, Z. Gao, and C. Liu, "Ground risk assessment of uav operations based on horizontal distance estimation under uncertain conditions," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–12, 2021.
- [12] S. Bertrand, N. Raballand, and F. Viguier, "Evaluating ground risk for road networks induced by uav operations," in *2018 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2018, pp. 168–176.
- [13] B. Levasseur, S. Bertrand, N. Raballand, F. Viguier, and G. Goussu, "Accurate ground impact footprints and probabilistic maps for risk analysis of uav missions," in *2019 IEEE Aerospace Conference*. IEEE, 2019, pp. 1–10.
- [14] S. H. Kim, "Third-party risk analysis of small unmanned aircraft systems operations," *Journal of Aerospace Information Systems*, vol. 17, no. 1, p. 24, 2020.
- [15] E. W. Dijkstra, "A note on two problems in connexion with graphs," in *Edsger Wybe Dijkstra: His Life, Work, and Legacy*, 2022, pp. 287–290.
- [16] M. G. Bell, "Hyperstar: A multi-path astar algorithm for risk averse vehicle navigation," *Transportation Research Part B: Methodological*, vol. 43, no. 1, pp. 97–107, 2009.
- [17] Y. Wu, K. H. Low, B. Pang, and Q. Tan, "Swarm-based 4d path planning for drone operations in urban environments," *IEEE transactions on vehicular technology*, vol. 70, no. 8, pp. 7464–7479, 2021.
- [18] 3GPP TS 23.501, "System architecture for the 5G System (5GS)," *3GPP Technical Specification*, 2024.
- [19] 3GPP TS 23.288, "Architecture enhancements for 5G System (5GS) to support network data analytics services," *3GPP Technical Specification*, 2024.
- [20] S. Richard and B. Andrew, *Reinforcement Learning: An Introduction, Second Edition*. Massachusetts, USA: The MIT Press, 2018.
- [21] C. Bergroth, O. Järvi, H. Tenkanen, M. Manninen, and T. Toivonen, "A 24-hour population distribution dataset based on mobile phone data from helsinki metropolitan area, finland," *Scientific data*, vol. 9, no. 1, p. 39, 2022.