Learning-Based Strategies for UAV Swarm Target Defense: Imitation and Reinforcement Learning Perspective Inspired by the Hawk–Pigeon Game

Hugo Trebert Eliot Morin Micha Iakovlev

December 7, 2025

1 Introduction

As outlined in our first report, the Hawk–Pigeon framework provides a rigorous foundation for studying pursuit–evasion interactions between defending and attacking UAVs. This problem has gained increasing relevance in the current geopolitical context, where unmanned aerial systems play a growing role in surveillance and defense scenarios. The rapid development and deployment of UAV swarms motivate the study of defensive strategies capable of protecting critical targets.
In this report, we refine our methodology by implementing the natural pursuit laws described in the literature.

2 Methods

2.1 Simulation Environment

We developed a 3D multi-agent simulation environment designed to reproduce pursuit evasion interactions between hawks and pigeons. Each agent is modelled as a point-mass UAV with a 6-DOF state:

$$[x,\, y,\, z,\, V,\, \mu,\, \phi]^T,$$

where (x, y, z) is the 3D position, V the airspeed, and (μ, ϕ) the flight-path and heading angles.

The corresponding kinematic equations for the position components are given by

$$\dot{x} = V \cos\mu \cos\phi, \qquad \dot{y} = V \cos\mu \sin\phi, \qquad \dot{z} = V \sin\mu,$$

which describe how the UAV's velocity vector, defined by its airspeed and orientation angles, determines its instantaneous motion in 3D space.

The evolution of airspeed, flight-path angle and heading follows the standard point-mass dynamics(Some terms are explained later):

$$\dot{V} = g\,(n_x - \sin\mu), \qquad \dot{\mu} = \frac{g}{V}\,(n_f \cos\gamma - \cos\mu), \qquad \dot{\phi} = \frac{g\,n_f \sin\gamma}{V\cos\mu}.$$

Hawks and pigeons share the same physical constraints; the only advantage given to hawks is their higher nominal speed. The victory conditions are asymmetric: pigeons win if any of them reaches the target, whereas hawks must capture all pigeons.

Each agent outputs a desired acceleration that reflects its behavioural strategy (detailed later). This abstract command cannot be applied directly to a flying vehicle, so the *Converter* module transforms it into physically meaningful inputs: the longitudinal load factor n_x, the normal load factor n_f, and the bank angle γ.

The mapping between analytical accelerations $(u_x,\, u_y,\, u_z)$ and UAV control inputs $(n_x,\, n_f,\, \gamma)$ is obtained through the converter:

$$n_x = \Gamma_1, \qquad \gamma = \tan^{-1}\!\left(\frac{\Gamma_3}{\Gamma_2}\right), \qquad n_f = \frac{\Gamma_2}{\cos\gamma},$$

where

$$\Gamma = M^{-1}\begin{bmatrix} u_x/g \\ u_y/g \\ u_z\cos\mu/g \end{bmatrix} + \begin{bmatrix} \sin\mu \\ \cos\mu \\ 0 \end{bmatrix}.$$

These quantities determine how the UAV pitches, pulls, and rolls. The *Dynamics* module then updates the state using a fourth-order Runge–Kutta (RK4) integrator, ensuring stable flight evolution throughout the simulation.

The underlying kinematics follow the standard 3D point-mass aircraft model, where the velocity components are reconstructed from (V, μ, ϕ) and the angular rates depend on n_x, n_f, and γ. This formulation allows both hawks and pigeons to generate realistic trajectories consistent with the physics of small aerial vehicles.

Simulation logic is handled by a central *Game* class. Hawks can detect pigeons within a 1 km sensing radius and must commit to a chosen target until it is captured. A capture event occurs when the hawk–pigeon distance falls below 10 m, while a pigeon reaching the target within the same threshold ends the mission in favour of the attackers. The simulation is updated with a 0.1 s timestep and terminates when all pigeons are captured, one pigeon reaches the target, or a maximum duration of 30 s is reached. Initial positions, target location, and number of agents are configurable, allowing full multi-trajectory rollouts under consistent physical and interaction constraints.

2.2 Analytical Baseline and Trajectories

The analytical baseline defines how hawks and pigeons behave in the pursuit evasion game. Each agent follows a deterministic control law that translates its behavioural strategy into a desired acceleration vector. These rules generate the reference trajectories used to analyse the dynamics and assess the effectiveness of the interaction model.

Hawk Behaviour (Defender)

Hawks aim to intercept pigeons before they reach the protected target. Each hawk evaluates all pigeons within a sensing radius of approximately one kilometre and ranks them according to three intuitive criteria:

- **Proximity:** priority is given to the pigeon that is physically closest, as it is typically the quickest to reach.
- **Isolation:** a pigeon drifting away from the flock is easier to capture and therefore strategically valuable.
- **Local density:** selecting a pigeon located in a dense region of the flock can disrupt the attackers more effectively.

For each criterion, the hawk identifies one candidate pigeon. It then selects the overall most advantageous target and pursues it using a geometric interception strategy: rather than steering toward the pigeon's current position, the hawk directs its motion toward a predicted interception point computed from the relative geometry and velocities of both agents. This mechanism produces smooth pursuit trajectories and realistic interception attempts.The hawk's acceleration command is defined as:

$$u_{\text{hawk}} = K_{PN}\,(\omega \times v_{\text{hawk}}) \; - \; K_{PP}\,(\beta \times v_{\text{hawk}}).$$

The term $K_{PN}(\omega \times v_{\text{hawk}})$ corresponds to Proportional Navigation and steers the hawk toward the predicted future position of the pigeon based on the rotation of the line of sight ω. The term $-K_{PP}(\beta \times v_{\text{hawk}})$ corresponds to Proportional Pursuit and realigns the hawk's heading toward the target by reducing the deviation angle β.

Pigeon Behaviour (Attacker)

Pigeons attempt to reach the protected target while avoiding capture. Their control law combines three complementary behaviors:

- **Target following:** a forward acceleration directs the pigeon toward the target, defining its primary trajectory.
- **Escape:** when a hawk enters a predefined threat radius, the pigeon generates a repulsive acceleration that pushes it away from the defender.

- **Collision avoidance:** pigeons repel each other when too close, preventing unrealistic flock collapse and maintaining a coherent formation.

These components result in adaptive motion patterns: The pigeon's control law is the weighted sum of target attraction, predator avoidance and collision avoidance: The pigeon's control law is the weighted sum of the three behaviours:

$$a_{\text{pigeon}} = k_t \, a_{\text{target}} + k_e \, a_{\text{escape}} + k_c \, a_{\text{avoid}}.$$

$$a_{\text{target}} = \frac{p_T - p}{\|p_T - p\|}, \qquad a_{\text{escape}} = \sum_{j \in E_2} k_j^{(2)} \frac{p - p_j}{\|p - p_j\|}, \qquad a_{\text{avoid}} = \sum_{m \in E_3} k_m^{(3)} \frac{p - p_m}{\|p - p_m\|}.$$

(Here, p is the current pigeon position, p_T the target position, p_j and p_m the positions of other pigeons, E_2 the set of pigeons within the hawk's sensing range, and E_3 the set of neighbouring pigeon).Pigeons progress toward their objective, deviate when threatened, and maintain safe spacing within the group. In addition, we introduce adjustable gains for each behavioral component, allowing us to tune their relative importance and explore whether specific weight combinations can yield improved overall performance.

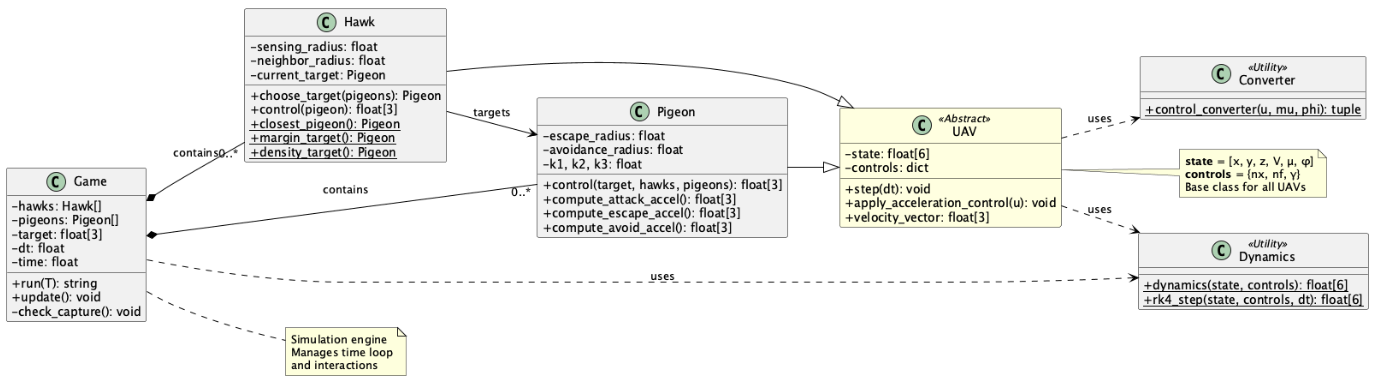

Figure 1: ULM Diagram of your main classes

3 Results

We tested the analytical model in both simple (1 hawk vs. 1 pigeon) and multi-agent configurations (e.g., 2 hawks vs. 2 pigeons), and in all cases the simulation behaved as expected. Depending on the initial geometry, either the hawk successfully intercepted the pigeon or the pigeon reached the target first, demonstrating that the pursuit, escape, and avoidance rules are correctly implemented. The logged 3D trajectories confirm coherent motion for both defenders and attackers, and provide a clear basis for later quantitative evaluation and learning-based experiments.

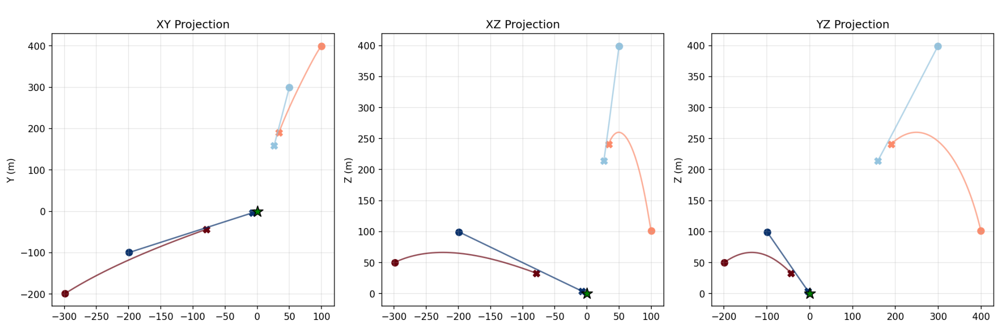

Figure 2: 2D trajectories on a pigeon win simulation

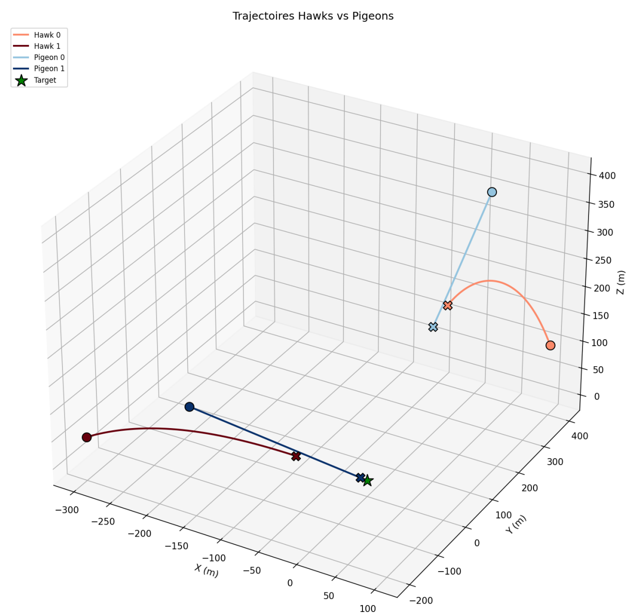

Figure 3: 3D trajectories on a pigeon win simulation

4 Conclusion

In this second report, we implemented a full 3D multi-agent simulation environment and reproduced the analytical pursuit–evasion behaviours that define the Hawk Pigeon interaction. The resulting trajectories validate the model in both simple and multi-agent settings, where interception and escape outcomes naturally emerge from the control laws. This implementation now serves as a reliable framework for generating structured datasets of agent states and expert actions.

These datasets constitute the basis for the next stage of the project: training learning-based controllers that can approximate or potentially improve upon the analytical strategies. Future work will therefore explore imitation learning and reinforcement learning methods to assess whether data driven policies can achieve more robust or efficient defensive behaviours.

Beyond behavioural improvements, learning-based controllers may also reduce runtime computational cost. Once trained, a neural policy can reproduce the analytical guidance laws through a single forward pass, enabling faster decision updates and potentially lighter onboard hardware.

Acknowledgments

Eliot worked on implemented UAV and the hawks class and wrote the report.

Hugo worked on the dynamics/ converter and implemented the pigeon class.

Micha worked on the file game and the visualisation.

References

[1] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[2] W. Ruan, Y. Sun, Y. Deng, and H. Duan, "Hawk–pigeon game tactics for unmanned aerial vehicle swarm target defense," *IEEE Transactions on Cybernetics*, vol. 51, no. 9, pp. 4423–4436, 2021.