Optimal swarm control in a threatening environment

Stanislas de Charentenay[12], Alexandre Reiffers-Masson[1], Gilles Coppin[1], Guillaume Lesueur[2], Jean-Yves Schneider[2]

[1] IIMT Atlantique, LAB-STICC laboratory, Brest, France
stanislas.mareschal-de-charentenay@imt-atlantique.fr
[2] Thales Optronique SAS, Elancourt, France

Mots-clés : *UCAV swarm, Optimal control, Markov decision process, Path planning.*

1 Introduction

Robot swarms have gained a lot of popularity in recent years, especially on the topic of optimal trajectory control. Here, we focus on swarms of unmanned combat aerial vehicles (UCAV) that evolve in a hostile environment. In this case, the swarm members have a high risk of being destroyed during their mission. We want to find a control that avoids the complete destruction of the swarm with the highest probability and allow at least one member of the swarm to reach a target safely. Although the risk of destruction of a swarm member depends strongly on its position in the environment, we know that the presence of other members affect this destruction probability by acting as shield or distracting the attacker for example. In the state of the art, although path planning in dangerous environments has been widely studied for a single UAV and swarms, the swarm effect is often neglected and the swarm is only exploited for target allocation purposes.

In this paper, we aim to exploit this swarm effect to find a control that minimizes the probability for the whole swarm to be destroyed before reaching a target zone. For this we designed two metrics to evaluate a given control and achieved an simple optimization algorithm in the case of a predetermined control for a path planning problem.

2 Model description

We control a swarm of n UCAVs in a 3 dimensional space and aim to reach a target zone $\mathcal{Z} \subset \mathbb{R}^3$. The members of the swarm risk destruction throughout their mission. The studied scenario is modeled by a discrete time Markov decision process.

In our model, each member of the swarm has a position and speed state both in $\mathbb{R}^3$ and a destruction state in $\{0, 1\}$. We add an additional reaching state in $\{0, 1\}$ for the whole swarm that keeps track of whether a member the swarm already reached the target. We concatenate all these states to an overall swarm state $S(t)$ at time t. The actions taken by the swarm correspond to a bounded acceleration in $\mathbb{R}^3$ for each member.

Regarding the evolution of the states, The position and speed state are deterministic, as the position and speed are the discrete derivative of the speed and acceleration respectively. For the evolution of the destruction state, we know that it will stay we assume that we know individual destruction rate functions that represent the probability for one member of the swarm to be destroyed. This probability is dependent on the position of the studied drone, but also on the positions and destruction state of other members to take the swarm effect into account. From these functions, we can get the probabilities $q(d \mid s)$ of destruction state d when in overall state s. The next reaching state of the swarm can be determined from the next destruction and position state, by checking if any members of the swarm are undetected and inside the target zone.

3 Control evaluation

The control of the swarm is modeled by a policy π. To evaluate the policies, we design a reward system based on a cost function that associate to swarm state s the cost $c(s) = -1$ if the reaching state is 0 and $c(s) = 0$ else. As such we can evaluate the policy with the value function $V_\pi(s) = \mathbb{E}_\pi\Big[\sum_{t=0}^{\infty} \gamma^t c(S(t)) \mid S(0) = s\Big]$, with γ the discount factor. The value function can be interpreted as the average discounted time left before reaching the target zone safely, and can be calculated using Bellman equations.

As the state space has a very large number of dimensions, finding an optimal closed loop control will be very hard as the computation cost of classical optimization algorithms will be too high. Instead, it is easier to look at a path planning problem with an open loop control. In this case, the policy is deterministic which implies that the position and speed of the swarm so we can deduce $s(d,t)$ the overall swarm state at time t knowing that we are in detection state d and reaching state 0. As such we get the recursive expression based on Bellman equation :

$$V_\pi(X(t)) = -1 + \sum_{d' \in \{0,1\}^n} q(d' \mid s(d,t)) V(d', t+1)$$

The metric used to evaluate policy π would be $V_\pi(s_0)$ with s_0 the initial state of the swarm. The calculation complexity of this metric is $O(T^n)$ with T the number of time steps in the scenario.

By calculating the instants t_i^* where the planned trajectory of UCAV i enters the target zone, we can get a different expression of the metric by calculating the probability to be in different states at each t_i^*. This expression can be calculated in $O(T2^n n!)$, which is faster for smaller swarms.

Furthermore, using a slightly different cost function, we can calculate the probability to avoid complete destruction of the swarm with similar expressions. This probability makes for a useful additional metric.

4 Path planning optimization

To optimize our path planning control, we used a gradient descent algorithm to maximize the value of our different metrics. We experimented our implementation on a scenarios where the members can act as a shield for others, meaning that the destruction rate is lower when a member is close to its neighbors. Although the resulting trajectories were satisfying for a swarm of 2 or 3 UCAVs, we are still struggling to optimize trajectories for larger swarms. Future work has to be done on clever approximations to reduce the dimension of our state space, such as discretization or aggregation techniques.

Références

[1] Qie, Han, et al. Joint Optimization of Multi-UAV Target Assignment and Path Planning Based on Multi-Agent Reinforcement Learning. *IEEE Access*, 2019, vol. 7, p. 146264-146272

[2] Eun, Yeon-Ju and Bang, Hyochoong. Cooperative control of multiple UCAVs for suppression of enemy air defense. *AIAA 3rd" Unmanned Unlimited" Technical Conference, Workshop and Exhibit*, 2004. p. 6529.

[3] Borges, Nícolas Pereira, Cinara Guellner Ghedini, and Carlos Henrique Costa Ribeiro. SUN-Stealth UAV Networks : A cooperative local model to mitigate UAVs exposure to threats. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019.